

Received January 8, 2020, accepted January 19, 2020, date of publication January 23, 2020, date of current version January 31, 2020. Digital Object Identifier 10.1109/ACCESS.2020.2969055

A Unified Framework and Method for **EEG-Based Early Epileptic Seizure Detection and Epilepsy Diagnosis**

ZIXU CHEN[®], **GUOLIANG LU[®]**, **(Member, IEEE)**, **ZHAOHONG XIE[®]**, **AND WEI SHANG[®]** 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 study was supported by the Shandong Provincial Natural Science Foundation, China (ZR2019MEE063) and the Fundamental Research Funds of Shandong University (2018JC010).

ABSTRACT Electroencephalogram (EEG) contains important physiological information that can reflect the activity of human brain, making it useful for epileptic seizure detection and epilepsy diagnosis. However visual inspection of large amounts of EEG by human expert is time-consuming, and meanwhile there are often inconsistences in judgement between physicians. In this paper, we develop a unified framework for early epileptic seizure detection and epilepsy diagnosis, which includes two phases. In the first phase, the signal intensity is first calculated for each data point of the given EEG, enabling the well-known autoregressive moving average (ARMA) model to characterize the dynamic behavior of the EEG time series. The residual error between the predicted value of learned ARMA model and the actually observed value is used as the anomaly score to support a null hypothesis testing for making epileptic seizure decision. The epileptic seizure detection phase can provide a quick detection for anomaly EEG patterns, but the resulting suspicious segment may include epilepsy or other disordering EEG activities thus required to be identified. Therefore, in the second phase, we use pattern recognition technique to classify the suspicious EEG segments. In particular, we propose a new and practical classifier based on a pairwise of one-class SVMs for epilepsy diagnosis. The proposed classifier requires normal and epilepsy data for training, but can recognize normal, epilepsy and even other disorders that would not be trained in the training samples. This point is practical and meaningful in real clinic scenarios as the input EEG may include other brain disordering diseases besides of epilepsy. We conducted experiments on the publicly-available Bern-Barcelona and CHB-MIT EEG database, respectively, to validate the effectiveness of the proposed framework, and our method achieved classification accuracy of 93% and 94% on them. Comprehensive experimental results, outperforming the *state-of-the-arts*, suggest its great potentials in real applications.

INDEX TERMS Seizure detection, epilepsy diagnosis, change detection, one-class SVM, EEG diagnosis.

I. INTRODUCTION

Epilepsy is one of the most common neurological disorders. Approximately 1% of the world's population has epilepsy, and up to 5% of people may have at least one seizure during their lifetime [1]. This serious disorder is associated with recurrent, unprovoked epileptic seizures resulting from a sudden disturbance of brain function which is characterized by abnormal firing of cortical neurons recruiting neighboring

The associate editor coordinating the review of this manuscript and approving it for publication was Victor Hugo Albuquerque¹⁰.

cells into a critical mass [2]. The seizure occurs at random to impair the normal function of the brain, most of the patients also suffer from numerous other unforeseeable side effects, such as memory loss, depression and other psychological disorders [3], [4]. Therefore it is important to detect and identify the epilepsy at an early stage, so as to help user take appropriate actions/health-care in advance to avoid accidental consequences and ensure the patient's health. As a result, automated seizure detection and epilepsy diagnosis from electroencephalogram (EEG) signals has become an active research topic in past decades [5]-[8].

Epileptic seizure can be considered as disordered brain activities that differ from those that are usual under normal EEG status [9]. In this sense, seizure detection can be accomplished as a novelty detection. Lots of methods have been proposed based on statistical analysis. A well established mechanism is described as follows: a regressive model (e.g., linear model [10], [11], logistic model [12], gaussian model [13], [14]), is first used to characterize the dynamic behavior of EEG signals; temporal anomalies are then calculated to quantify the possibilities that indicates a change happening; hypothesis testing is finally adopted for decision making.

Meanwhile, epilepsy diagnosis has been also attracted much attention in the area of pattern recognition [15]-[17]. It normally relies on feature extraction and pattern classification techniques. Typical features include amplitude, amplitude average, duration, half-wave duration, sharpness ratios, slope attributes of half-waves and so on. However, due to the fact of that EEG may perform in an arbitrary way, the information we obtained from these features is very limited. To overcome this shortcoming, high-level features have been proposed based on time-frequency analysis, non-linear analysis and chaos theory [18]-[20], typically including power spectrum density, entropy and sample entropy, etc. By using these features, kinds of pattern recognition methods can be used as a classifier e.g., support vector machine (SVM), artificial neural network (ANN), or other alternatives such as extreme learning machine (ELM) [21], and so on.

However, as a preliminary assumption, to apply the seizure detection and epilepsy diagnosis separately on the continuously monitored EEG signal may lead to several limitations,

- a) Seizure detection relying on auto-regression analysis, although implemented very fast, lacks of post-processing mechanism to identify the epilepsy because other disorders (and even a normal physiological movement) would be also detected as a seizure [22]. Therefore there exist many false alarms in those detected results for real EEG signals.
- b) Epilepsy diagnosis heavily relies on feature extraction and pattern classification that has been discussed in the above. Accurate and effective feature extraction guarantees reliable diagnosis but needs large time to compute, making the common use of sliding-window strategy limited for continuous EEG diagnosis in practice [23].

A unified and more practical framework that combines seizure detection and epilepsy diagnosis together, is thus more accepted in recent studies [24], [25]. This paradigm overcomes important limitations of existing methods regarding the epileptic seizure detection and epilepsy diagnosis as two separate problems. It detects suspicious epilepsy segments in the seizure detection phase, and then performs the diagnosis on detected segments to identify the epilepsy. It can provide more accurate and reliable diagnosis results with a high computational efficiency. This paper is therefore focused on the epileptic seizure detection and epilepsy diagnosis with a unified framework.

In the rest of this paper, Section II provides an overview of the proposed framework; Section III describes the proposed method; Experimental validations are given in Section IV; Section V shows some discussions; Conclusion is finally made in Section VI.

II. OVERVIEW OF THE PROPOSED FRAMEWORK

As shown in Fig. 1, the proposed framework mainly includes three steps, described as below.

- 1) *EEG signal preprocessing*: the mean filter, as a commonly method for signal smoothing, is first used here to reduce the noise in raw EEG data.
- 2) Epileptic seizure detection: the autoregressive moving average model (ARMA) [26] is used to learn the regularity of past EEG data, and the residual errors are calculated as temporal anomalies that can characterize the dynamics of EEG signals. As such, the deviation of a new EEG input from the observed data can be quantified. Finally, a common null hypothesis testing is performed to produce the seizure detection result.
- 3) Epilepsy diagnosis: the detected suspicious EEG segments may include epilepsy, normal status and even other unknown disordering brain activities. To consider this issue, we propose a new and practical classifier based on a pairwise of one-class SVMs for epilepsy diagnosis. The proposed classifier requires normal and epilepsy data for training, but recognizes normal, epilepsy and even other disorders that were not included in training. This point is practical and very meaningful in real clinic scenarios as the input EEG signal may include other brain disordering diseases besides of epilepsy. The proposed classifier can recognize these disordering patterns without a prior training for them, offering users a more accurate monitoring for the given EEG inputs.

III. METHODOLOGY

This section describes the detail of main technologies used in the framework.

A. SIGNAL PREPROCESSING

As a pre-processing, we accomplish this task by applying a smoothing filter on the raw collected EEG signal, i.e., $x_t \leftarrow x(t) * G$ where x(t) is the observed value of EEG signal at time *t* and *G* is a filter (a *mean* filter with a length of 1×30 in experiments), * is the convolution operation.

B. EPILEPTIC SEIZURE DETECTION

In order to detect the EEG status change that are from normal to abnormal due to the presence of epileptic seizure, two steps are performed as follows: (1) employ a mathematical model to characterize the dynamic behavior of EEG signal, and (2) perform a null hypothesis testing for decision making.

1) DATA MODELING OF EEG SIGNALS

It is a common practice to employ an appropriate model that can extract dynamic characteristics of a given EEG



FIGURE 1. Overview of the proposed framework.

signal [27]. It is however not a trivial task due to the high non-stationarity of EEG signals. Signal intensity, as a basic EEG waveform that can reflect the cortical electrical activity, has been demonstrated promising performance on EEG signal processing [28]. Given a data stream of EEG signal up to time N, i.e., $X = \{x_1, \ldots, x_t, \ldots, x_N\}$, the signal intensity is calculated for each data point by,

$$v_t = \sqrt{\frac{1}{L} \sum_{t-L+1}^{t} x_i^2}$$
 (1)

where v_t is the computed signal intensity at time t, and *L* is the length of local window for calculation (in this paper we set L = 5 empirically). As such, the original EEG data stream is now represented by a sequence of corresponding intensities i.e., $X : V = \{v_1, ..., v_t, ..., v_N\}$. This representation enables to extract linearity characteristics of EEG signals and the ARMA model is employed to achieve this end, described as,

$$\widetilde{v}_t = m_1 v_{t-1} + m_2 v_{t-2} + \ldots + m_p v_{t-p} + \varepsilon_t, \qquad (2)$$

where \tilde{v}_t represents the predicted value at time t, and $m_i(i = 1, 2, 3, ..., p)$ is the i-th ARMA coefficient, ε_t is an independent and identically distributed (i.i.d) white noise.

There are two issues in the use of this model for EEG signal analysis: (1) determination of the order p of model, and (2) estimation of coefficients i.e., $\{m_1, m_2, \ldots, m_p\}$.

- (1) The order of the ARMA term is usually chosen by using the Akaike Information Criteria (AIC) [29]. In this paper, a prior order range $\{p_1^*, \ldots, p_k^*\}$ from 1 to 10 is firstly confirmed empirically in order to reduce the computation burden in the search of optimal *p*.
- (2) Based on {p₁^{*},..., p_k^{*}}, the least square method is used (given in Eq. 3) to estimate the corresponding model coefficients at each order, i.e.,

$$\{m_1, m_2, \dots, m_p\} \leftarrow \min \sum_{t=p+1}^N [\nu_t - \widetilde{\nu}_t]^2.$$
 (3)

(3) The optimal order p, together with corresponding model coefficients i.e., $\{m_1, \ldots, m_p\}$ can be confirmed

with Akaike Information Criteria (AIC) by finding the minimum *AIC* value given as below,

$$AIC(p) = Nln \frac{\sum \sigma_i^2}{N} + 2p \tag{4}$$

where N is observation number and σ is the deviation between predicted data and real data.

The procedure of the above algorithm is provided in **Algorithm I**. Based on a prior estimation, the model construction process of an example EEG *Ind* 0005 from Bern-Barcelona database [30] is shown in Fig. 2. The optimal p is confirmed as 6 and its corresponding coefficients i.e., $\{m_1, \ldots, m_p\}$, can be also confirmed. We accomplish this estimation by an off-line calculation process and use the estimated p directly in the execution of the proposed framework.

Based on the optimal value of p together with corresponding coefficients $\{m_1, \ldots, m_p\}$, the ARMA model can be constructed to describe the dynamic characteristics of the given EEG signal X.

2) DECISION MAKING

Once the ARMA model is constructed, we can use it to quantify the deviation between the predicted value and the



FIGURE 2. The prior estimation of the optimal p.

real value, and then a null hypothesis can be tested for decision making.

1) Anomaly score calculation: we first compute the temporal series s_t which can reflect the possibility that a change occurs. The temporal series s_t is calculated as the deviation between the predicted data and the real data, for the computed signal intensity i.e. $V = \{v_1, \ldots, v_t, \ldots, v_N\}$ where v_t is the real value of the sequence at time *t*. We denote the predicted value according to ARMA model at time *t* using \tilde{v}_t . The s_t can be thus calculated as the mean value of the residual error between the actual value and the predicted value in a local extent (set as five neighboring data points),

$$s_t = \frac{\sum_{i=t-2}^{t+2} |v_i - \widetilde{v}_i|}{5}.$$
 (5)

2) Null hypothesis testing: on the basic of s(t), we can employ a certain test to produce the detection result. Those methods can divided into *real-time* detection methods and *retrospective* detection methods. The retrospective methods, e.g., cumulative sum (CUSUM) test, F_c metric, generalized likelihood ratio test (GLRT), and Friedman test, can produce accurate change detection results but it needs a larger observation delay. As the seizure detection we consider in this paper is expected to detect the seizure at an early stage, we used the real-time detection method. With an assumption of Gaussian distribution [31], [32], the common 3σ criterion is used here to test a null hypothesis in order to realize the real-time detection.

$$H_0 :| s_t - \mu_{t-1} | \le 3\sigma_{t-1}; \text{ No change detected}$$

$$H_1 :| s_t - \mu_{t-1} | > 3\sigma_{t-1}; \text{ Change detected} \quad (6)$$

where μ_{t-1} is the mean value and σ_{t-1} is the standard deviation of the set of $\{s_1, s_2, \ldots, s_{t-1}\}$. H_0 means no change occurs at time t, and H_1 indicates that a change occurs.

A EEG status change implies an occurring of seizure or other disorder activities that are different with the normal state. The seizure detection phase can provide a quick detection for anomaly EEG patterns, but the resulting suspicious segment may include epilepsy or other disordering EEG activities thus required to be identified. To achieve this end, as shown in Fig. 3, a suspicious segment X_c which begins at the detected time c and lasts for a fixed length of l', i.e., $X_c = \{x_c, x_{c+1}, \ldots, x_{c+l'}\}$, is formed for further analysis to identify the epilepsy.¹

C. EPILEPSY DIAGNOSIS

Once an EEG status change has been detected, an automatic analysis of the detected EEG suspicious segment is performed to identify the epilepsy based on pattern recognition techniques, which includes two steps: (1) feature extraction, and (2) EEG classification.

1) FEATURE EXTRACTION

The raw data of suspicious EEG segment is data redundant, it is therefore necessary to extract explanatory parameters from the raw EEG data. As a promising tool to analyze the non-stationary signal, EMD, which is proposed in [33], has been reported good results in EEG signal processing [34]. As an adaptive signal decomposition method, EMD can achieve a high temporal resolution and high frequency resolution simultaneously, which is a great improvement than the classic Short Time Fourier Transform (STFT) [35]. And the decomposed results depend only on the signal itself. In comparison with wavelet transform (WT), EMD is also advantageous because it does not require the selection of basis functions [36]. Conventionally, EMD decomposes a given signal X_c into intrinsic oscillatory components, called intrinsic mode functions which is represented by $imf_i(t)$ using a sifting process, and a residual left after the sifting process, i.e. $r_n(t)$. The formulation of EMD is given in Eq.(7),

$$X_c = \sum_{i=1}^n imf_i(t) + r_n(t)$$
(7)

However, the $imf_i(t)$ s obtained from EMD are always too large and complex as the feature vectors. To deal with this shortcoming, SVD is thus used so as to obtain the stable and simple features as used in [37]. Further more, we give an analysis of the energy from each *imf*, and Fig. 4 shows that the first five *imf* s contain most of the energy of the given signal and were used to extract singular values in our method.

¹The size of window was set as 10240 empirically in the experiment.



FIGURE 3. The suspicious segment confirmation.



FIGURE 4. Selection of informative *imf* s.

The generated singular value vector i.e. $[\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5]$ indicates magnitudes of component signals, and each component signal represents a different spectrum component of the original signal, which supports an informative and sufficient representation for different EEG patterns [38].

Based on EMD and SVD, for a suspicious EEG segment X_c , we can use the corresponding singular value vectors, i.e. $[\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5]$ as features that will be fed to diagnosis.

2) EPILEPSY RECOGNITION

The one-class SVM can be used to diagnose the suspicious EEG segment for epilepsy identification. The one-class SVM is a method to deal with these highly nonlinear classification problems [39]. As a binary classifier, its training can be implemented only using the data of one class. Based on this,

most methods which using one-class SVM are trained by normal data [40], and the classification is normal or abnormal, where the output of '1' indicates normal and '-1' indicates abnormal.

Recall that, there would exist many other brain disordering diseases in real clinic scenarios [41], [42]. It is however impossible to collect all kinds of data to train a classifier considering the specificity and diversity of different disordering diseases in EEG signals. To overcome this shortcoming, in this paper we propose a new classifier for epilepsy recognition. The proposed classifier is composed of a pairwise one-class SVMs, as depicted in Fig. 5, the training and testing procedure of the proposed classifier is given as follows:

- In the training phase, the first one-class SVM in the proposed pairwise classifier is trained with epilepsy EEG samples, and the second SVM is trained with normal EEG samples.
- In the testing phase, we feed the suspicious EEG segments to the proposed pairwise classifier, the outputs of the first trained one-class SVM is given as: 1 indicates an epilepsy status and -1 indicates *non*-epilepsy; the outputs of the second trained one-class SVM is given as: 1 indicates normal and -1 indicates abnormal.
- The decision is made by a combination of the outputs of the pairwise one-class SVMs that is given in Tab. 1. The diagnosis result for the suspicious EEG segment is finally confirmed as: {1, -1} indicates epilepsy EEG status, {-1, 1} indicates normal, {-1, -1} indicates other unknown brain disordering activities that are not included in the training samples and {1, 1} indicates a false recognition, where the first value indicates the output of the first SVM classifier and the second value is the second SVM classifier output.

Based on the new and practical classifier, the detected EEG suspicious segment can be identified whether an epilepsy occurs or not. If an epilepsy is identified, the system will reports an alarm to the user; otherwise it continues the seizure detection for the new EEG comes.



FIGURE 5. The procedures of using the pairwise SVMs classifier for epilepsy identification.

TABLE 1. Combination of two SVMs outputs for diagnosis.

Model\Result	×	Other disorder	Normal	Epilepsy
The 1st SVM	+1	-1	-1	+1
The 2nd SVM	+1	-1	+1	-1

IV. EXPERIMENT

A. VALIDATION ON BERN-BARCELONA EEG DATABASE1) MATERIAL AND EXPERIMENT IMPLEMENTATION

The framework is validated on two experiment setups. Both the EEG data are taken from the publicly available Bern-Barcelona EEG database [30]. These data were obtained from 5 patients with epilepsy including focal and non-focal channels, which were recorded at 1024Hz. Total 3750 pairs of simultaneously recorded signals x and y are randomly selected from the pool of all signals measured at focal and non-focal EEG channels respectively, and divide the recordings into time windows of 10 seconds resulting in 10240 samples totally. We only use x-signals of this database, where we downsampled these signals with a down-sampling rate of 1:40 to reduce the computation burden in the experiment.

We carried out the experiments in Matlab R2018a without using any acceleration programing. The computation environment is: CPU2.30 GHz and RAM 8.00 GB.

2) RESULT ON SEIZURE DETECTION

In this experiment, we firstly randomly select 50 focal and 50 non-focal signal from the described database, and concatenated each pair of a non-focal signal record and a focal signal record to generate new EEG data streams such that each generated data stream contains at least one EEG status change. Three indicators i.e., *precision*, *recall*, F_{score} are employed to evaluate the performance of seizure detection, which are calculated respectively as follows:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F_score = \frac{2 \times precision \times recall}{precision + recall}$$

where *TP* is classified as true positive patterns, *FP* and *FN* are classified as false positive patterns and negative patterns.

We perform the seizure detection (described in Section III-B) on the testing data. In order to provide a quantitative evaluation of the proposed method, in the experiment we perform the method until a true EEG status change is detected. Fig. 6 shows two examples of the seizure detection. Fig. 6 (a) shows a successful detection where we can see that the computed anomaly scores keep a stable trend under EEG non-focal status while changes greatly that exceeds the control limit once the EEG goes into the focal status. The employed null hypothesis testing can detect this change successfully. While, as for the second example shown in Fig. 6 (b), a change that has been detected before the true change was served as a false alarm. The main possible reason is the second testing data has a relatively large data fluctuations compared with the first testing data.

A comprehensive detection result is given: the *precision* of 0.72, *recall* of 1.00 and *F_score* of 0.84 have been obtained by our method. Here, it is worth to note that, the indicator of *recall* achieves satisfactory performance i.e., 1.00 which implies that all true changes have been successfully detected. While, the indicator of *precision* only achieves 0.72, which means that there are some false alarms. The result is reasonable considering the complexity and high noise of the raw EEG signal. The detected change point will be served as the starting time of suspicious segment which will be fed to the epilepsy diagnosis phase for identification.

3) RESULT ON EPILEPSY DIAGNOSIS

In this experiment, 100 focal and 100 non-focal signals are randomly selected as testing data. And 50 focal data are used for the first one-class SVM classifier training and 50 non-focal data for the second one-class SVM classifier training, the residual 50 focal data and 50 non-focal data are used for testing.

We have used the RBF kernel function of SVM. Additionally it is common to confirm the γ and μ in the use of one-class SVM for classification. Specifically, the γ plays an important role in the nonlinear mapping of the input vector from the input space to the high-dimensional space, and the μ controls the proportion of the kernel empirical risk of the confidence interval in the classifier. In other words, the first parameter controls the width of the distribution, and the second one represents the estimation of spurious data in the normal state registry. Accurate epilepsy diagnosis has been achieved using the following parameters: $\gamma = 0.1768$, $\mu = 0.0034$ for the first one-class SVM classifier, and $\gamma = 0.0078$, $\mu = 0.0034$ for the second one.

The result of the epilepsy diagnosis is given in Fig. 7. There are 47 normal segments and 46 epilepsy segments have been correctly classified as shown in the shaded part in the confusion matrix, and one epilepsy segment has been wrongly classified as normal, because there are only two labels in this EEG database. The whole number of this confusion matrix are not 100. Except the data shown in the matrix, one segment



(b) Detection result of test signal Ind0012.

FIGURE 6. Detection results of two EEG examples.

has been classified by our method as ' \times ' and five segments have been classified as 'Other disorder'. Several common indicators are used to evaluate the proposed framework. They are,

$$Accuracy = \frac{TP + TN}{Total \ samples} (\%)$$

$$Sensitivity = \frac{TP}{TP + FN} (\%)$$

$$Specificity = \frac{TN}{TN + FP} (\%)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

$$K = \frac{P_o - P_e}{1 - P_e}$$

where, *TP*, *FP* and *FN* have been defined in the above. And *TN* is classified as true negative patterns. Calculations of p_o and p_e can be referred in [43].



FIGURE 7. Confusion matrix of epilepsy diagnosis based on Bern-Barcelona database.

Thus the classification *Accuracy* of our method is 93%, *Sensitivity* is 97.8%, *Specificity* is 100%, *MCC* is 0.979 and *K* is 0.869 on the average. All these results indicate that the classified results have a good consistence with the true labels. In Fig. 8, we show the testing EEG data that have been recognized as 'normal', 'epilepsy' and 'Other disorder', where spectrograms of each of them is also provided based





FIGURE 8. The epilepsy diagnosis results of 3 EEG examples. In each of them, from top to bottom are the original EEG and the spectrum respectively.

on short time frequency transform (STFT). It is obvious to distinguish the normal EEG from the epilepsy EEG and other disorder EEG, because the normal EEG is more smooth and more regularly visually. The difference between epilepsy and other disorder in EEG is the characteristic waveform, i.e. there appears many spinous slow composite waves in epilepsy segment which is a typical symbol of epilepsy; while, there appears many spike waves in other disorder segment which is rarely appears in normal and epilepsy segments. From the spectrograms of each segment, we can see that most powers of normal EEG mainly exist in a low frequency band (5Hz-20Hz), the epilepsy and other disorder have less power in low frequency band and according to the occurrence of the spike waves in other disorder EEG; meanwhile there appears fluctuations of high frequency band power in other disorder EEG data, which is a difference from the epilepsy EEG.

The diagnosis results have been evaluated by neurological experts. The possible reasons why the testing data have been recognized as 'Other disorder' include,

- *Outside interference*: patients may have a passive frightened or the hardware of device damaged;
- *Other unknown diseases*: patients may have some other unknown diseases, such as the flu, even a sneezing can lead to abnormal EEG status, or the intravenous injection causing pathological involuntary convulsions.

The above analysis also demonstrated the proposed framework can provide a more precise evaluation of EEG status.

B. VALIDATION ON CHB-MIT EEG DATABASE

1) MATERIAL AND EXPERIMENT IMPLEMENTATION

The framework is validated on two experiment setups. Both the EEG data are taken from the publicly available CHB-MIT EEG database [44]. These data were collected from



(b) Detection result of test signal chb03_01.



22 subjects (5 males, ages 3 to 22; and 17 females, ages 1.5 to 19) at the Children's Hospital Boston, consisting of EEG recordings from pediatric subjects with intractable seizures. Total 129 files that include seizures are included in this collection, with a sampling rate at 256Hz and a resolution at 16-bit. Most files contain 23 channel signals, we only use the signal from channel FP1-F7. And the computing environment is the same as described before.

2) RESULT ON SEIZURE DETECTION

In this experiment, we randomly select 50 labeled seizures from this database, the signals that 10 seconds before and after the onset time of each seizure were constructed as a new EEG data stream, such that each constructed data stream contains at least one EEG status change. In Fig. 9, we also give two examples. Fig. 9 (a) shows a successful detection



FIGURE 10. Confusion matrix of epilepsy diagnosis based on CHB-MIT database.

that the calculated anomaly score fluctuate a lot once the EEG appears the seizure waveform. And the null hypothesis testing can detect it successfully. In Fig. 9 (b) we show another example that a change has been detected as a false alarm.



(c) The detected segment as other disorder.

FIGURE 11. The epilepsy diagnosis results of 3 EEG examples. In each of them, from top to bottom are the original EEG and the spectrum respectively.

It can be clearly seen that the false alarm is caused by a local fluctuation. The detection result of this database is given: the precision of 0.84, recall of 1.00, and F_score of 0.91 have been obtained by our method.

3) RESULT ON EPILEPSY DIAGNOSIS

In this experiment, 100 seizure and 100 non-seizure signals are randomly selected as testing data, each signal lasting 10 seconds. And 50 seizure data are used for the first one-class SVM classifier training and 50 non-seizure data for the second one-class SVM classifier training, the residual 50 seizure data and 50 non-seizure data are used for testing.

We also used the RBF kernel function of SVM, and accurate epilepsy diagnosis has been achieved using the following parameters: $\gamma = 0.0884$, $\mu = 0.0034$ for the first one-class SVM classifier, and $\gamma = 0.0625$, $\mu = 0.0034$ for the second one.

The result of the epilepsy diagnosis is given in Fig. 10. There are 48 normal segments and 46 epilepsy segments that have been classified correctly, and one normal segment has been wrongly classified. Additionally, two segments have been classified by our method as '×' and three segments have been classified as 'Other disorder'. We also give the spectrogram of three failed recognition examples in Fig. 11. We can also see that from time-domain, it is obvious to distinguish the normal EEG from the epilepsy EEG and other disorder EEG, the difference between epilepsy EEG and other disorder EEG is the occurrence of the spinous slow composite wave. And in time-frequency-domain, the energy distributions of three EEG status are the same as discussed above. On the overall, the classification is achieved as: accuracy is 94%, sensitivity is 100%, specificity is 97.9%, *MCC* is 0.979 and *K* is 0.887 on the average.

C. COMPARISON SUMMARIZATION

The proposed approach is compared with state-of-the-art works. Table 2 and Table 3 summarizes the comparison results based on the Bern-Barcelona and CHB-MIT

Ref.&Year	Description	Accuracy	Sensitivity	Specificity
[45], 2016	Entropy + LSSVM with Linear kernel	76%	65%	64%
	Entropy + LSSVM with Polynomial kernel	82%	68%	67%
	Entropy + LSSVM with Polynomial RBF kernel	86%	72%	71%
[46], 2016	DWT + SVM with RBF kernel	83.07%	83.05%	83.09%
[47], 2018	ApEn, SampEn, FuzzyEn + NNge classifier	99%	99%	99%
Ours		93%	97.8%	100%

TABLE 2. The classification results in Bern-Barcelona database.

TABLE 3. The classification results in CHB-MIT database.

Ref.&Year	Description	Accuracy	Sensitivity	Specificity
[48], 2015	PS-PCA + LDA and NBC	94.69%	89.10%	94.80%
[49], 2015	FuzzyEn + SVM	98.31%	98.27%	98.36%
[50], 2017	PCA-CSP + ADCD		96%	
[51], 2017	LBP + KNN	99.7%	99.2%	99.8%
Ours		94%	100%	97.9%

database respectively. Specifically, in [45], 50 normal and 50 focal EEG signals were used to validate the algorithm, they transform the input EEG signal to a TF plane using the Stockwell Transform (S-Transform) and compute the entropy measure in the time-frequency domain, finally LSSVM with different kernel functions (Linear, Polynomial, RBF) together with the entropy measure are used to make decision. In [46], they propose a framework using discrete wavelet transform (DWT) and SVM (RBF) for epilepsy diagnosis. In [47], entropies such as approximate entropy (ApEn), sample entropy (SampEn), fuzzy entropy (FuzzyEn) are extracted from EEG data and fed into several classifiers, finally the NNge gave the best performance with the classification accuracy of 99%. In [48], a new multi-channel EEG seizure detection method is presented based on the dynamics of the trajectories in phase space, then the features were extracted based on the Poincarle section together with PCA method, finally a two-layer schem that comprising LDA and NBC was employed as the classifier to make the decision. In [49], they first calculated the fuzzy entropy of EEG signals from different states, then a feature selection method has been used, and finally based on the optimal features, the support vector machine (SVM) was employed to make classifications. In [50], they presented a framework that employs principle component analysis and common spatial patterns to enhance the EEG signals and uses the extracted discriminative feature as an input for adaptive distance-based change point detector to make the final decision. And in [51], a novel framework was proposed, the morphological features were extracted based on the local binary pattern operator, and K-nearest neighbor classifier was used for classification. Through the comparison, we can see that our method doesn't achieve the best classification accuracy. However it is more robust while **TABLE 4.** Computational complexity of the framework in different phases where n_1 and n_2 denote the number of data points in one segment for epileptic seizure detection and epilepsy diagnosis respectively, p is the length of the preliminarily set order range, I is the number of segments for epileptic seizure detection.

Process	Computational complexity
Signal preprocessing	$O(n_1 * l)$
Epileptic seizure detection	$O(n_1 * l * \frac{p^2}{2} + n_1 * l + p * l)$
Epilepsy diagnosis	$O(n_2^3 * log(n_2) + n_2^3)$

the diagnosis made by it relies on the pairwise SVMs, and the diagnosis result include normal, epilepsy and even other disorders that would not be trained in the training samples. This point can help the user diagnose epilepsy and also other physical assessment of the patient, which is practical and very meaningful in realistic clinical scenarios.

V. DISCUSSION ON COMPUTATIONAL COMPLEXITY

It is noted that the computational efficiency is also an important issue in the design of a framework for epilepsy diagnosis. In Tab. 4, we show the computational complexity of proposed framework in different phases. It can be seen that the complexity of the phase epileptic seizure detection is much lower, which indicates it can give us a very quick response when the EEG signal is arriving continuously. The phase of epilepsy diagnosis has a higher computational complexity. It is tolerant because accurate diagnosis is more required than the computation speed. Another reason is that, this phase can also be executed with an off-line manner.

Feature extraction is of great importance in the context of EEG signal processing. We compared the employed feature

TABLE 5. Computation time of feature extraction.

Ref.&Year	Feature	Time (Seconds)
[47], 2018	ApEn	0.054
	SampEn	0.053
	RE	0.003
	FuzzyEn	1.14
[45], 2016	TF entropy	0.4027
Ours		0.317

extraction method with some others reported in references. The computation time of our method was calculated on a data segment with 2560 samples. As shown in Tab. 5, our method takes the least time i.e., 0.317*s*, implying it is fast enough to realize the real-time epilepsy diagnosis.

VI. CONCLUSION

In this paper, we concentrate on the problem of automatic seizure detection and epilepsy diagnosis from EEG signals. We have proposed a new and practical unified framework to achieve this end. A null hypothesis testing is used to detect the seizure in continuous monitoring of EEG signals. As such, suspicious segments can be identified, which is fed to intelligent diagnosis using a novel classifier based on a pairwise one-class SVMs. Experiments were conducted on the public Bern-Barcelona EEG database and CHB-MIT EEG database to investigate the performance of seizure detection and epilepsy diagnosis respectively. Comparison with recently-released results demonstrated that the proposed method can achieve a high performance in terms of classification accuracy, sensitivity and specificity. It is more robust and can give the classification including normal, epilepsy, other disorders that are not included in the training samples. In the future work, we will extract and use more features for EEG signal representation to further improve the performance on the proposed framework.

REFERENCES

- M. Peker, B. Sen, and D. Delen, "A novel method for automated diagnosis of epilepsy using complex-valued classifiers," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 1, pp. 108–118, Jan. 2016.
- [2] R. Wang, G. Q. Zeng, X. Liu, R. Z. Tong, D. Zhou, and Z. Hong, "Evaluation of serum matrix metalloproteinase-3 as a biomarker for diagnosis of epilepsy," *J. Neurolog. Sci.*, vol. 367, pp. 291–297, Aug. 2016.
- [3] R. H. Uk, H. J. Pyo, H. Su-Hyun, C. Eun Ju, S. J. Hyun, L. Sang-Ahm, and K. Joong Koo, "Seizure frequencies and number of anti-epileptic drugs as risk factors for sudden unexpected death in epilepsy," *J. Korean Med. Sci.*, vol. 30, no. 6, pp. 788–792, 2015.
- [4] L. M. Willems, P. S. Reif, S. Knake, H. M. Hamer, C. Willems, G. Krämer, F. Rosenow, and A. Strzelczyk, "Noncompliance of patients with driving restrictions due to uncontrolled epilepsy," *Epilepsy Behav.*, vol. 91, pp. 86–89, Feb. 2019.
- [5] S. R. Mathieson, N. J. Stevenson, E. Low, W. P. Marnane, J. M. Rennie, A. Temko, G. Lightbody, and G. B. Boylan, "Validation of an automated seizure detection algorithm for term neonates," *Clin. Neurophysiol.*, vol. 127, no. 1, pp. 156–168, Jan. 2016.

- [7] J. Birjandtalab, M. B. Pouyan, and M. Nourani, "Unsupervised EEG analysis for automated epileptic seizure detection," in *Proc. 1st Int. Workshop Pattern Recognit.*, Jul. 2016, Art. no. 100110M.
- [8] A. Bhardwaj, A. Tiwari, R. Krishna, and V. Varma, "A novel genetic programming approach for epileptic seizure detection," *Comput. Methods Programs Biomed.*, vol. 124, pp. 2–18, Feb. 2016.
- [9] Y. Kumar, M. L. Dewal, and R. S. Anand, "Epileptic seizures detection in eeg using DWT-based APEN and artificial neural network," *Signal Image Video Process.*, vol. 8, no. 7, pp. 1323–1334, 2014.
- [10] J. M. Antelis, L. Montesano, A. Ramos-Murguialday, N. Birbaumer, and J. Minguez, "On the usage of linear regression models to reconstruct limb kinematics from low frequency EEG signals," *PLoS ONE*, vol. 8, no. 4, Apr. 2013, Art. no. e61976.
- [11] S.-F. Liang, C.-E. Kuo, Y.-H. Hu, Y.-H. Pan, and Y.-H. Wang, "Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 6, pp. 1649–1657, Jun. 2012.
- [12] K. Sun-Hee, F. Christos, and Y. Hyung-Jeong, "Coercively adjusted auto regression model for forecasting in epilepsy eeg," *Comput. Math. Methods Med.*, vol. 2013, no. 2, 2013, Art. no. 545613.
- [13] Y. Li, W. Cui, M. Luo, K. Li, and L. Wang, "Epileptic seizure detection based on time-frequency images of EEG signals using Gaussian mixture model and gray level co-occurrence matrix features," *Int. J. Neur. Syst.*, vol. 28, no. 07, Sep. 2018, Art. no. 1850003.
- [14] A. Quintero-Rincón, M. Pereyra, C. D'Giano, H. Batatia, and M. Risk, "A new algorithm for epilepsy seizure onset detection and spread estimation from EEG signals," *J. Phys., Conf. Ser.*, vol. 705, Apr. 2016, Art. no. 012032.
- [15] H. Göksu, "EEG based epileptiform pattern recognition inside and outside the seizure states," *Biomed. Signal Process. Control*, vol. 43, pp. 204–215, May 2018.
- [16] 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.
- [17] H. Ke, D. Chen, X. Li, Y. Tang, T. Shah, and R. Ranjan, "Towards brain big data classification: Epileptic EEG identification with a lightweight VGGNet on global MIC," *IEEE Access*, vol. 6, pp. 14722–14733, 2018.
- [18] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, Mar. 2015.
- [19] J. Martinez-del-Rincon, M. J. Santofimia, X. Del Toro, J. Barba, F. Romero, P. Navas, and J. C. Lopez, "Non-linear classifiers applied to EEG analysis for epilepsy seizure detection," *Expert Syst. Appl.*, vol. 86, pp. 99–112, Nov. 2017.
- [20] L. Wang, W. Xue, Y. Li, M. Luo, J. Huang, W. Cui, and C. Huang, "Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis," *Entropy*, vol. 19, no. 6, p. 222, May 2017.
- [21] Q. Liu, X. Zhao, Z. Hou, and H. Liu, "Epileptic seizure detection based on the kernel extreme learning machine," *Technol. Health Care*, vol. 25, pp. 399–409, Jul. 2017.
- [22] S. Saxena, S. P. Singh, and K. Makhija, "Seizure disorders," Tech. Rep., Oct. 2015, pp. 1–14, doi: 10.1007/978-1-4939-0779-3_71-1.
- [23] C. Hao, Z. Chen, and Z. Zhao, "Analysis and prediction of epilepsy based on visibility graph," in *Proc. 3rd Int. Conf. Inf. Sci. Control Eng. (ICISCE)*, Jul. 2016, pp. 1271–1274.
- [24] J. Gubbi, S. Kusmakar, A. S. Rao, B. Yan, T. Obrien, and M. Palaniswami, "Automatic detection and classification of convulsive psychogenic nonepileptic seizures using a wearable device," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 4, pp. 1061–1072, Jul. 2016.
- [25] S. Xie and S. Krishnan, "Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis," *Med. Biol. Eng. Comput.*, vol. 51, nos. 1–2, pp. 49–60, 2013.
- [26] G. E. P. Box and G. C. Tiao, "Intervention analysis with applications to economic and environmental problems," *J. Amer. Stat. Assoc.*, vol. 70, no. 349, pp. 70–79, Mar. 1975.
- [27] T. Zhang, W. Chen, and M. Li, "AR based quadratic feature extraction in the VMD domain for the automated seizure detection of EEG using random forest classifier," *Biomed. Signal Process. Control*, vol. 31, pp. 550–559, Jan. 2017.

- [28] N. Loveless and A. Sanford, "The impact of warning signal intensity on reaction time and components of the contingent negative variation," *Biol. Psychol.*, vol. 2, no. 3, pp. 217–226, Jan. 1975.
- [29] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, Dec. 1974.
- [30] 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.
- [31] Y. Bazi, L. Bruzzone, and F. Melgani, "An approach to unsupervised change detection in multitemporal SAR images based on the generalized Gaussian distribution," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.* (IGARSS), Dec. 2004, pp. 1402–1405.
- [32] M. Bosc, F. Heitz, J.-P. Armspach, I. Namer, D. Gounot, and L. Rumbach, "Automatic change detection in multimodal serial MRI: Application to multiple sclerosis lesion evolution," *NeuroImage*, vol. 20, no. 2, pp. 643–656, Oct. 2003.
- [33] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, 1998.
- [34] J. Zhong, S. Qin, and C. Peng, "Study on separation for the frequency bands of eeg signal and frequency band relative intensity analysis based upon EMD," in *Proc. WSEAS Int. Conf. Signal Process.*, 2008, pp. 151–155.
- [35] X. Zhao, B. Ye, and T. Chen, "Study on measure rule of time-frequency concentration of short time Fourier transform," J. Vib. Meas. Diagnosis, vol. 37, no. 5, pp. 948–956, 2017.
- [36] J. Zhang, J. Wei, Y. Hui, S. Wei, L. Shi, and J. Ping, "Epileptic electroencephalogram recognition based on discrete s transform and permutation entropy," J. Biomed. Eng., vol. 34, no. 5, pp. 681–687, 2017.
- [37] P. Baranyi, Y. Yam, A. Varkonyi-Koczy, R. Patton, P. Michelberger, and M. Sugiyama, "SVD-based complexity reduction to TS fuzzy models," *IEEE Trans. Ind. Electron.*, vol. 49, no. 2, pp. 433–443, Apr. 2002.
- [38] H. Hassanpour, M. Mesbah, and B. Boashash, "Time-frequency feature extraction of newborn eeg seizure using svd-based techniques," *EURASIP* J. Adv. Signal Process., vol. 2004, no. 16, pp. 1–11, 2004.
- [39] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Comput.*, vol. 13, no. 7, pp. 1443–1471, Jul. 2001.
- [40] 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.
- [41] H.-C. Lin, S. Xirasagar, H.-C. Lee, C.-C. Huang, and C.-H. Chen, "Association of Alzhemier's disease with hepatitis C among patients with bipolar disorder," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0179312.
- [42] T. Pincus, L. F. Callahan, L. A. Bradley, W. K. Vaughn, and F. Wolfe, "Elevated MMPI scores for hypochondriasis, depression, and hysteria in patients with rheumatoid arthritis reflect disease rather than psychological status," *Arthritis Rheumatism*, vol. 29, no. 12, pp. 1456–1466, Dec. 1986.
- [43] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, Apr. 1960.
- [44] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 2009.
- [45] P. T. Krishnan and P. Balasubramanian, "Automated EEG seizure detection based on S-transform," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCIC)*, Dec. 2016, pp. 1–5.
- [46] D. Chen, S. Wan, and F. S. Bao, "Epileptic focus localization using discrete wavelet transform based on interictal intracranial EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 5, pp. 413–425, May 2017.
- [47] N. Arunkumar, K. Ram Kumar, and V. Venkataraman, "Entropy features for focal EEG and non focal EEG," *J. Comput. Sci.*, vol. 27, pp. 440–444, Jul. 2018.
- [48] M. Zabihi, S. Kiranyaz, A. B. Rad, A. K. Katsaggelos, M. Gabbouj, and T. Ince, "Analysis of high-dimensional phase space via Poincaré section for patient-specific seizure detection," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 3, pp. 386–398, Mar. 2016.

- [49] 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.
- [50] S. Khanmohammadi and C.-A. Chou, "Adaptive seizure onset detection framework using a hybrid PCA–CSP Approach," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 1, pp. 154–160, Jan. 2018.
- [51] P. P. M. Shanir, K. A. Khan, Y. U. Khan, O. Farooq, and H. Adeli, "Automatic seizure detection based on morphological features using onedimensional local binary pattern on long-term EEG," *Clin EEG Neurosci*, vol. 49, no. 5, pp. 351–362, Sep. 2018.



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



GUOLIANG LU (Member, IEEE) 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.



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 of the Second Hospital of Shandong University. His research interests include electroencephalogram, polysomnography, and their applications in neurology and sleep medicine.



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.

• • •