ARTICLE IN PRESS
JID: NEUCOM [m5G;December 20, 2019;1:27]

Neurocomputing xxx (xxxx) xxx

Contents lists available at ScienceDirect

Neurocomputing

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

Brief papers

A novel automatic classification detection for epileptic seizure based on dictionary learning and sparse representation☆

Hong Peng[a], Cancheng Li[a], Jinlong Chao[a], Tao Wang[a], Chengjian Zhao[a], Xiaoning Huo[b], Bin Hu[a,*]

[a] Gansu Provincial Key Laboratory of Wearable Computing, School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China
[b] The Third People's Hospital of Lanzhou, Lanzhou 730050, China

ARTICLE INFO

Article history:
Received 3 July 2019
Revised 2 December 2019
Accepted 3 December 2019
Available online xxx

Communicated by Prof. FangXiang Wu

Keywords:
Epileptic seizure
Electroencephalogram (EEG)
Dictionary learning
Detection
Classification

ABSTRACT

Electroencephalogram (EEG) signals play an important role in the epilepsy detection. In the past decades, the automatic detection system of epilepsy has emerged and performed well. In this paper, a novel sparse representation-based epileptic seizure classification based on the dictionary learning with homotopy (DLWH) algorithm is proposed. The performance of the proposed method evaluates on two public EEG databases provided by the Bonn University and Childrens Hospital Boston-Massachusetts Institute of Technology (CHB-MIT), various classification cases that include health and seizure; non-seizure and seizure; inter ictal (seizure-free interval) and ictal (seizure). The results show that the DLWH only completes the test with 19.671 s compared with the traditional sparse representation methods with high degree of automation, which are better than those obtained using the well-known dictionary learning method. Besides, two publicly available benchmark databases recognition rates are as high as 100%, 99.5%, with average of 99.5% and 95.06%,% respectively, and the results show that the epileptic detection system based on the dictionary learning has a high application value.

1. Introduction

Epilepsy is the second most common serious nervous system disease after the stroke [1]. It is a chronic nervous system disease characterized by the recurrent and unprovoked seizures [2]. About 65 million people worldwide have been affected from the epilepsy. Electroencephalogram (EEG) is an important clinical tool for monitoring, diagnosing and managing the epilepsy-related neurological diseases [3]. Doctors and neurologists analyzed these signals to assess states of the brain, which may be a time consuming process. EEG based on the computer-aided techniques was found very effective in diagnosing the epilepsy compared with other methods [4–6], such as the electrocardiogram (ECOG). Besides, EEG was a convenient and safe technique for monitoring the brain activity [7–9].

Almost one third of epilepsy patients are currently untreated [10,11]. Despite the choice of diet, medication and surgical treatment, they suffered from sudden and unpredictable seizures that had a great impact on their daily lives and temporarily impair their perception, speech, motor control, memory and consciousness [12]. Many new therapies have been researched, and one of the most promising therapies was implantable devices that provided the direct electrical stimulation to affected areas of the brain [13]. These processes relied heavily on the robust algorithms for the epileptic detection to perform effectively. Because epileptic seizures cannot be predicted in a short period of time, it is necessary to continuously re-record EEG to detect the epilepsy. However, it is tedious, time-consuming and costly to analyze the long visual EEG records to find traces of the epilepsy [14]. Therefore, the automatic detection of the epilepsy has been being the goal of many researchers. With the development of technology, the digital EEG data can be fed into the automatic seizure detection systems, allowing doctors to treat more patients in a given time, due to the fact that the automation greatly reduces the time spent examining EEG data [15–17].

☆ This work was supported in part by the National Natural Science Foundation of China [Grant Nos. 61632014, 61210010, 61402211], the National Basic Research Program of China (973 Program) [No. 2014CB744600], the International Cooperation Project of Ministry of Science and Technology [No.2013DFA11140], the Program of Beijing Municipal Science & Technology Commission [No.Z171100000117005], the Project of National Defense Science and Technology Innovation Zone, Talent innovation and entrepreneurship project of Lanzhou (2015-RC-60). Kindly note that H. Peng and C. C. Li are jointly of the first authorship.

* Corresponding author.
E-mail addresses: pengh@lzu.edu.cn (H. Peng), licch17@lzu.edu.cn (C. Li), chaojl15@lzu.edu.cn (J. Chao), wangtao2018@lzu.edu.cn (T. Wang), zhaochj18@lzu.edu.cn (C. Zhao), lzsyhxn@163.com (X. Huo), bh@lzu.edu.cn (B. Hu).

https://doi.org/10.1016/j.neucom.2019.12.010
0925-2312/

In recent years, automatic classification models with epileptic seizures have made a lot of progress in applications of pattern recognition (PR). In addition, there are many classification methods for EEG signals, of which the multi-layer perception neural network (MLPNN) and the support vector machine (SVM) are two widely used classification models. Most automatic epilepsy detection systems are based on the feature extraction and various classification models. In [18], Polat et al. presented a hybrid system with two stages: the feature extraction using (fast fourier transformation) FFT and the decision tree classifier with 98.68% and 98.72% accuracy, respectively. Fu et al. presented a new technique for the seizure classification of EEG signals using the Hilbert–Huang transform (HHT) and SVM in [19]. Further, the spectral entropies and energy features of frequency-bands of the rhythms using hilbert marginal spectrum (HMS) analysis were extracted and combined with the SVM. The entropy was the measure method of uncertainty. Wang et al. applied the Wavelet transform and Shannon entropy to the feature extraction method and fed into k nearest neighbor (KNN) classifier with the best classification accuracy about 100% in [20]. Furthermore, Nicolaou et al. used the Permutation Entropy (PE) as a feature extraction approach in [21]. The low computational complexity of PE constituted as an advantageous feature part of a system for the real-time automatic seizure detection. In [22], Kumar et al. used the discrete wavelet transform (DWT) analysis and the approximate entropy (ApEn) of EEG signals. Abundant and significant results are achieved on the wavelet transform. In [23,24], both Subasi et al. and Guo et al. use the discrete wavelet transform model and made a series of progress. Moreover, Farrikh et al. presented (discrete wavelet transform) DWT and adaptive hybrid feature selection within bagging with multi-layer perception (MLP) detection for epileptic seizure in [25].

In this work, we propose a method based on the dictionary learning and sparse representation to classify seizures for the first time. The presented clinical EEG databases provided by the Bonn University [26] and Children's Hospital Boston-Massachusetts Institute of Technology (CHB-MIT) [27–29], which are also applied to this paper. The results show that the proposed scheme possess both high classification rate and high recognition speed in the seizure identification. In addition, it shows that the automatic detection and diagnosis can be realized in a kind of real-time epilepsy support system. So far, no related studies on the classification scheme has been conducted on the epilepsy based on the dictionary learning with homotopy (DLWH). The remainder of this paper is organized as follows. Related researches on the classification of seizures are provided in Section 2 In Section 3, technical components which include the preprocessing, the linear sparse representation model, the sparse representation by minimization, the dictionary learning with homotopy algorithm and classification approach are introduced. Furthermore, a brief description on the experimental procedures and results are provided in Section 4. There are discussions and conclusions at the end of this paper.

2. Related work

Epileptic seizure classification and detection have been studied for several years. Specifically, beyli presented a combination of three different EEG signals (sets A, D and E) with wavelet coefficients, and the overall classification accuracy rate was 93.17% [30]. Altunay et al. used the linear predictive error energy method for five sets, with their reported acuuracy up to 93.6% [31]. A clustering technique based on the least squares support vector machine (CT-LS-SVM) was presented in [32], with the classification accuracy of EEG 94.18%. In [33], a cross-correlation SVM classifier for EEG classification was presented by Chandaka. In [34], Least-Square Support Vector Machine (LS-SVM) was applied to classify two sets of EEG signals (sets A and E), and the classification accuracy rate reached 99.56%. Furthermore, in [35], the classification results and statistical parameter value indicated that the PNN classifier could distinguish the difference of EEG signals, meanwhile, the total classification accuracy was as high as 97.63%. Liang et al. employed the ELM algorithm that was extract nonlinear features, i,e., approximate entropy, Hurst exponent and scaling index, for the classification of interictal epilepsy and ictal EEG signals [36]. Acharya et al. extracted four entropy features to train seven classifiers, and the fuzzy classifier had been proven that three categories could be distinguished with the overall classification accuracy rate being 98.1% [37]. Li et al. employed the online optimization algorithm for dictionary learning to obtain a new dictionary, meanwhile, they also introduced the elastic net to make the system robust for the test samples [38]. In addition, the EEG database provided by the CHB-MIT is applied in many studies. Ahmad et al proposed a method that did not require requires any preprocessing of the data, which was unsupervised and fully automated [39]. In [40], principal component (PCA) and common spatial patterns (CSP) was applied to enhance the EEG signals, then, the extracted discriminant features was used as input to the distance based adaptive change point detector to identify seizures.

3. Mathematica methods

We propose a dictionary learning framework in this paper. The outline of the system is shown in Fig. 1. It illustrates several parts of this algorithm, which includes the preprocessing, DLWH and classification. Then, we will introduce each part detailedly in the ensuing sections.

3.1. Preprocessing

Before using DLWH classification method to identify the epileptic EEG, it is necessary to preprocess the EEG signals, which makes the characteristics of the seizure period more differentiable. Note that we use 0.1–49 Hz bandpass filter in this paper. Besides, it is conducted to maintain the phenomena, such as sharp waves, spikes and characteristic waves, that often occurs during seizures, so that the difference between EEG signals in the attack period and intermittent EEG signals in the episodes is more obvious. Then, DLWH based on the sparse representation classification method is used to perform pattern recognition on the two types of preprocessed EEG data. As a continuation, a sparse representation matrix is constructed, which uses an EEG training set, and the details are shown in the following subsections and Fig. 1. In order to find coefficients of the vector x, we use ℓ_1 regularization in the test signal y. Thus, we learn to get the dictionary D, which is the sparse representation algorithm in the brain dictionary in seizures and intermittent brain dictionaries, with residuals associated with seizures and non-seizures sub-dictionary being calculated, respectively. Actually, the test signal is classified as two categories, i.e., seizures and non-seizures, by comparing with the reconstructed residuals. Besides, the final classification results are satisfied with the diagnose in real time by the clinician.

In this section, we will introduce the sparse representation model of test signals and training signals. In the mean time, we will depict the usage of the model. Specifically, Fig. 2 shows a dictionary describing the design process of the linear sparse representation model. Firstly, we get the test signal y, and use the same process to get the column of dictionary D. Therefore, the test signal is converted to the vector $y \in \mathbb{R}^m$, and the dimension is consistent with that of the dictionary D. Beyond that, the test signal y can be represented by a linear sparse representation of the column of dictionary D.

Suppose that there is a group of linearly irrelevant base $[\alpha_1, \alpha_2, \ldots, \alpha_n] \in D$, an n-dimensional signal $y \in \mathbb{R}^n$ can be linearly

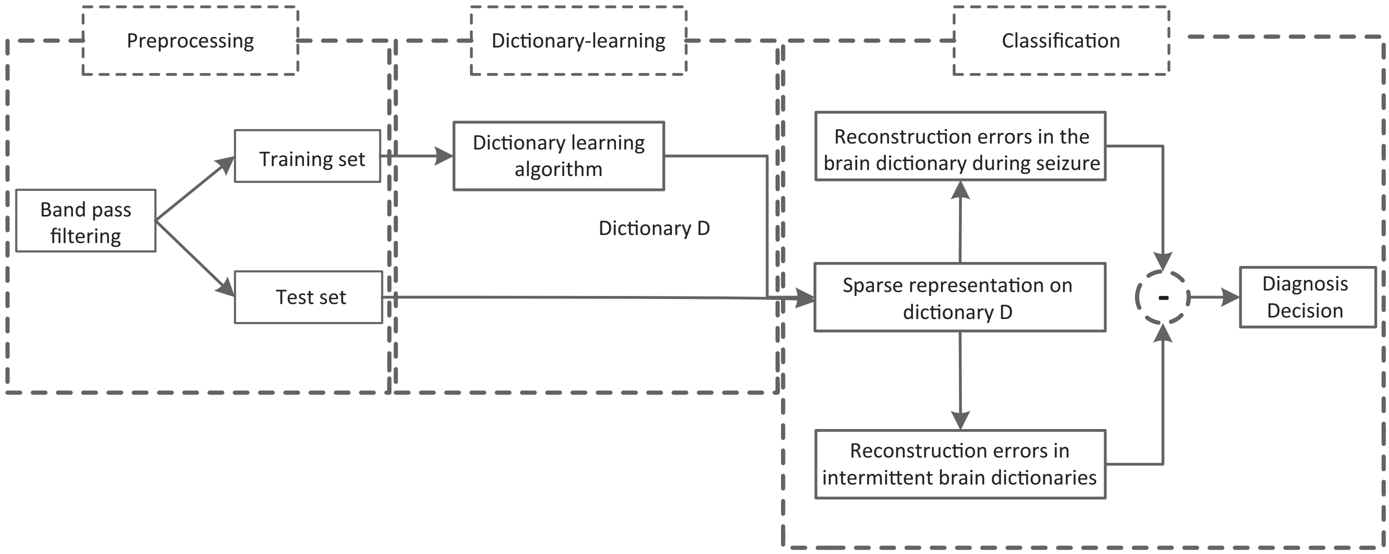

Fig. 1. Frames include three step: pro-processing, constructing dictionary D with dictionary learning algorithm, and diagnosis decision.

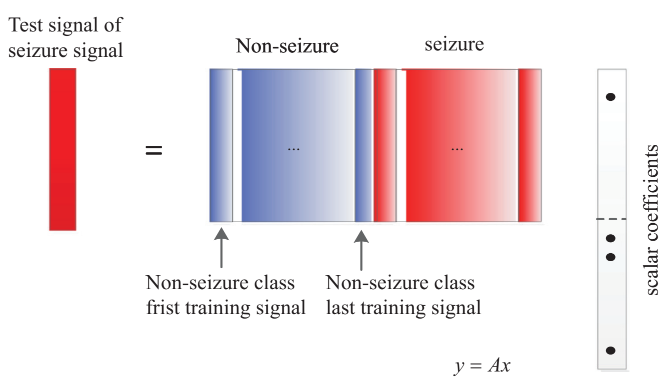

Fig. 2. The design on a dictionary and linear sparse representation model.

represented as

$$y = \sum_{i=1}^{m} x_i\alpha_i = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_i\alpha_m. \tag{1}$$

Eq. (1) can be reconstructed as

$$y = Dx. \tag{2}$$

where $x = [x_1, x_2, \ldots, x_m]^{\mathrm{T}}$ represents the expansion coefficient of the signal y on the basis, and $D \in R^{n \times m}$ is the dictionary matrix. In this research, we assume that a noise-free model is the sparse equation expression.

- If $n > m$, D is an incomplete dictionary, some vector dictionaries in space can not be effectively represented.
- If $n = m$, D is a complete dictionary.
- If $n < m$, D is an incomplete dictionary. In addition, when the system of linear equations is undetermined, the expansion coefficient vector $\mathbf{x}$ has infinite solutions.

If some expansion coefficients are large and others are zero or small, the signal can be sparsely represented. Mallat et al. presented a theory in [41], which illustrates a matching tracking calculation of the super complete dictionary decomposition. The sparse representation of the signal uses base functions as few as possible to represent the signal in a given overcomplete dictionary, which is to simplify and effectively express the original signal [42].

3.2. Sparse representation by minimization

In recent years, Wright et al. presented a classification algorithm based on the sparse representation classification (SRC) and apply it to the face recognition [43]. In the SRC algorithm, test samples are sparsely represented in training dictionaries composed of different types of training samples. Then, the reconstruction errors of different types of training samples are calculated, and compared with complete classification and recognition according to the minimum reconstruction errors. This algorithm breaks the traditional feature extraction and pattern recognition.

The sparsity of coefficients can be defined by ℓ_1-norm, and the less non-zero elements in the coefficient vector, the greater the sparsity of signals. After introducing constraints, the signal y can be expressed as follows:

$$\arg\min\|x\|_0 \qquad \text{subject to } y = Dx \tag{3}$$

Although the ℓ_0-norm plays an important role in the representation of signal sparsity, its solution process is quite complicated. However, in recent researches on the compressed sensing, ℓ_1 minimization technology is applied to the signal reconstruction. Furthermore, many researchers have proved that the ℓ_1-norm optimization problem has the same solution as ℓ_0-norm optimization under the constrained isometric conditions [44]. Thereby the formula (3) can be equivalent to the ℓ_1-norm optimization, which is expressed as

$$\arg\min\|x\|_1 \qquad \text{subject to } y = Dx \tag{4}$$

Specifically, Fig. 3 shows the minimization process of ℓ_1-norm and ℓ_2-norm by two-dimensional examples, which illustrates the reason that ℓ_1 minimization is able to find sparse solutions. In the norm theory, ℓ_1-norm and ℓ_2-norm can be expressed as vectors on the surface of circles and diamonds in, respectively. The grey dotted line corresponds to an isobar of constant model loss as achieved at each combination of ℓ_1 and ℓ_2.

A training dictionary $X = [X_1, X_2, \ldots, X_K]$ is constructed by a K-class training sample set, where $X_i = \left[X_i^1, X_i^2, \ldots, X_i^{m_i}\right] \in \mathbb{R}^{n \times m_i}$ represents the training sample set of class i with $i = 1, 2, \ldots, K$. Since this study only considers two conditions, i.e, epilepsy and non-epilepsy, the value of K is set as 2. For the test sample $y \in \mathbb{R}^n$, the sparse representation using the training dictionary is given as follows:

$$y = X\alpha \tag{5}$$

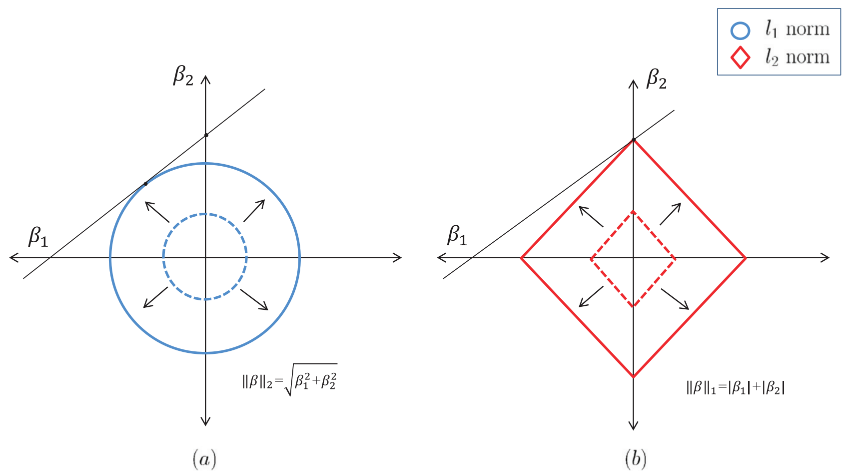

Fig. 3. The minimization process of ℓ_1-norm and ℓ_2-norm.

where $\alpha = \left[\alpha_1^T, \alpha_2^T, \ldots, \alpha_K^T\right]^T$ and α_j represents the sparse representation vectors of test sample y on the jth class training sample set. If the test sample y belongs to class i, the training sample set of class i can better represent the same data y, while the coefficients in other vectors $\boldsymbol{\alpha_j}(j \neq i)$ are zero or smaller. As shown in Eq. (6), the solution of the sparse coefficient is obtained through the minimum $\ell1$-norm optimization problem.

$$\hat{a}_i = \arg\min \|X\alpha - y\|_1 \qquad \text{subject to } \|X\alpha - y\|_2 \le \varepsilon \tag{6}$$

After obtaining the sparse vector $\hat{a}_i$, the test samples is reconstructed by using various training samples X_i and corresponding sparse vector $\hat{a}_i$, and the corresponding reconstruction errors of all types are calculated. Then the test sample y and the most accurate reconstruction training sample set belong to the same category, that is

$$\text{identity}(y) = \arg\min_i \left\| y - X_i \hat{a}_i \right\|_2 \tag{7}$$

After solving the $\ell1$ minimization problem, the non-zero element x must be correspond to class column i. Because the EEG signal is geatly noisy and non-stationary, non-zero elements may exist in the projection coefficient vectors of other classes due to the presence of noise. In order to use the sparse representation result coefficient vector x. The non-zero vector caused by noise will not have a great impact on the reconstruction of the test samples. Thereby, y belongs to the same category as the training samples that can accurately reconstruct it, i.e., the category to which the training samples corresponding to the minimum reconstruction error belongs is the category of the test samples y, and the classification rules are given:

$$class(y) = \arg\min_i r_i(y) \tag{8}$$

Besides, we conclude the SRC algorithm shown in Table 1.

Table 1
Pseudo-code implementation of the SRC method.

Sparse Representation Classification Algorithm (SRC)
Input: Training signal $A \in \mathbb{R}^{m\times 2N_t}$ for I classes, a test signal $y \in \mathbb{R}^{m\times 1}$
Output: The class label y for test data
Step 1: Normalize the columns of A and y
Step 2: Code y over x via ℓ_1-minimization;
solve the convex optimization problem: according to Eq. (4), where the constant ε is to account for the dense small noise in y or to balance the coding error of y and the sparsity of α
Step 3: Calculate the residuals for class i according to Eq. (7), where α_i is the coding coefficient vector associated with class i
Step 4: Output the identity of y; calculate the class (y) according to Eq. (8)
End

3.3. Dictionary learning with homotopy (DLWH)

Dictionary learning is proposed by Mairal et al. [45], which learns from the original training signals, and aims to construct a new dictionary for the test signal. Given a training set $X = [x_1, x_2, \ldots, x_n]$, the goal of the dictionary learning algorithm is to maximize the empirical loss function $f_n(D)$ that is given as follows:

$$f_n(D) = \frac{1}{n}\sum_{i=1}^{n} l(x_i, D) \tag{9}$$

where $l(x_i, D)$ is a loss function, and whether the dictionary could represent the signal commendably is determined by its value, n represents the dimension of the training signals. Every column of the dictionary is a basic vector, which is named atom. In general, the specific atom in dictionary D is an important factor in determining whether a test signal can be well represented. Here, we define the loss function as

$$\ell(x_i, D) = \alpha_i = \arg\min_{\alpha_i} \frac{1}{2}\|x_i - D\alpha_i\|_2^2 + \lambda\|\alpha_i\|_1 \tag{10}$$

where λ denotes the regularization parameters, which improves the generalization ability of this model. Meanwhile, α_i represents the k-dimensional coefficient vector. Besides, to solve this problem, a $\ell2$-norm (less than or equal to 1) is applied to constrain each column of D. According to [46], we define C as the convex set of matrix:

$$C = \left\{ D \in \mathbb{R}^{m\times k}\ s.\ t. \forall j = 1, \ldots, k \left\| d_j \right\|_2 \le 1 \right\} \tag{11}$$

Updating the loss function to Eq. (10), we can get

$$f_n(D, \alpha) = \arg\min_D \in C, \alpha_i \in \mathbb{R}^{k\times n} \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2}\|x_i - D\alpha_i\|_2^2 + \lambda\|\alpha_i\|_1\right] \tag{12}$$

Although $f_n(D, \alpha_i)$ is not a convex optimization function in the case where both the dictionary D and the coefficient vector α are changed, and when one of the dictionary D and the coefficient α is fixed, then $f_n(D, \alpha_i)$ is convex optimization with respect to the remaining one. To deal with this situation, one of the two variables of the dictionary D and the coefficient vector are kept fixed, and the empirical loss function is minimized by the remaining one iteration.

As one of the greedy algorithms, the homotopy method has been applied by many scholars in the field of sparse signal recovery [47]. If a signal x is k-sparse, this algorithm can find solution to Eq. (4) in k iterations. At the same time, Eq. (4) can be changed as an unconstrained optimization problem in (10). The objective function in (10) undergoes a constraint from ℓ_2 to ℓ_1. Therefore, when λ decreases to 0, the solution set x_λ^* of the large λ and $\boldsymbol{\alpha}$ converges to the solution of (4). Furthermore, the solution path is a piecewise constant that is a function of λ. It only changes at the critical value of λ.

Making the objective function of (10) is $F(\alpha)$, then

$$\partial F(\alpha) = -D^T(y - D\alpha) + \lambda \partial \|\alpha\|_1 \tag{13}$$

Since the $\|\alpha\|_1$ of Eq. (13) is not globally differentiable. Therefore, we need to redefine $\|\alpha\|_1$ of the sub-differential as follows:

$$u(\alpha) = \left\{ u \in R^N \left| \begin{array}{l} u_i = \mathrm{sgn}(\alpha_i), \ \ \alpha_{\mathbf{i}} \neq 0 \\ u_i \in [-1, 1], \ \ \alpha_i = 0 \end{array} \right. \right\} \tag{14}$$

The homotopy algorithm through the iterative process of k step with an initial value $\alpha_{(0)} = 0$. During each iteration, and with a certain nonzero λ, furthermore, Eq. (15) can be acquired through $\partial F(\alpha) = 0$:

$$c(\alpha) = D^T y - D^T D\alpha = \lambda u(\alpha) \tag{15}$$

According to Eq. (14), we maintain the sparse support set in the kth iteration.

$$I \in \left\{ i : \left| c^i_{(k)}(\alpha) \right| = \lambda \right\} \tag{16}$$

Then, we calculate the update direction and gradually adjust the non-zero coefficients of $\alpha_{(k)}$. In addition, mathematical derivation and algorithm information can be found in [47,48]. We use the efficient block-coordinate descent algorithm proposed in [49], which does not require any learning rate tuning. For each x_i each column, i.e., d_j (block), of the dictionary D is sequentially updated by the block coordinate decent algorithm [49]:

$$d_i = \frac{1}{A_{j,j}}\left(B_j - DA_j\right) + d_j \tag{17}$$

$$A_i = A_{i-1} + \alpha_i \alpha_i^{\mathrm{T}} \tag{18}$$

$$B_i = B_{i-1} + x_i \alpha_i^{\mathrm{T}} \tag{19}$$

Dictionary learning algorithm is of great significance for the development of the real-time system. Sparse coding with a fixed dictionary is a linear least squares problem. Furthermore, we introduce ℓ_1 regularization that can be solved by the LARS algorithm, which is another derivative of LASSO [49,50]. In the decade research, Marial et al. have employed block coordinates to fall and restart [49]. This method does not require learning rate adjustment, which can avoid the difficulty of setting optimization constraints. Finally, learning dictionary D is obtained by alternating above mentioned sparse optimization and dictionary update steps for each x_i until D is convergent. In this work, we set the number of iteration t as 100 to find the converged dictionary. Besides, we conclude the dictionary learning with homotopy algorithm shown in Table 2.

3.4. Classification

In Fig. 1, we can understand that the various aspects of the classification scheme. After the EEG data is preprocessed by filtering, etc, the learning dictionary is trained by randomly selected EEG training samples, and the dictionary training method is an online learning method. We introduce the ℓ_1 regularization penalty factor constraint, which can affect the sparsity of sparse coding and the generalization ability of model coding.

Table 2
Pseudo-code implementation of the DLWH method.

Dictionary learning with Homotopy
Input: training feature $x \in R^{m \times n}$
Output: learned dictionary D
I: Initialtize the dictionary $D_0 \in C$, N is the number iterations, A_0=0, B_0=0
II: For $t = 1 \ldots N$ do
  select x_t from the original training samples and updata the sparse coefficient,by solving equation.
  compute the α using Homotopy algorithm
  $\alpha = \arg\min_\alpha \frac{1}{2}\|y - D\alpha\|_2^2 + \lambda\|\alpha\|_1$
III: Updata A and B
  $A_i = A_{i-1} + \alpha_i\alpha_i^{\mathrm{T}}$
  $B_i = B_{i-1} + x_i\alpha_i^{\mathrm{T}}$
IV: Updata D using block coordinate decent algorithm using as follows:
  $\min_D \frac{1}{2}\|x_i - D\alpha_i\|_2^2 s.t. \left\|d_j\right\|_2 = 1, j = 1, 2, \ldots, k$
 End for
End while

The test sample y is sparsely coded in the learning dictionary, and the calculation of the sparse coefficient vector is realized by the Lars algorithm [51]:

$$\hat{\beta} = \arg\min_{\beta \in R^k} \left\{ \frac{1}{2}\|y - D\beta\|_2^2 + \lambda_1\|\beta\|_1 \right\} \tag{20}$$

of which λ signifies a constant, $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_i]$, $\hat{\beta}_i$ is the representation coefficient of the test sample on the ith sub-dictionary. Next, we need to calculate the reconstruction error that the test sampley represents on each sub-dictionary. In addition, we can determine the category to which the test sample belongs by reconstructing the error, and the test sample belongs to the category with the smallest reconstruction error.

$$r_i(y) = \left\| y - D_i \hat{\beta}_i \right\|_2^2 \tag{21}$$

where $\hat{\beta}_i$ is the representation coefficient of the test sample on the ith sub-dictionary, and D_i is the dictionary of the ith class. Finally, the label which the test sample belongs to is discriminated based on the reconstruction error.

4. Experimental procedure and dataset

To evaluate the proposed method, we conduct seizure detection experiments on two publicly available EEG databases provided by the Bonn University (2001) [26] and the open-source EEG database from CHB-MIT(http://phys ionet.org/cgi-bin/atm/ATM)

The first EEG database provided by the Bonn University consists of five groups (represented as Z, O, N, F, and S), each contains 100 single-channel EEG fragments with a duration of 23.6 s and a sampling rate of 173.6 Hz. After visual examination of artifacts, such as muscle activity or eye movement, EEG fragments are selected and cut out from the successive multichannel EEG records. The dimension of the original data signal is 4096. The group Z and group O consist of segments from the surface EEG records taken from five healthy volunteers using the standard 10–20 electrode replacement protocol. The volunteers relax their eyes (Z) and close eyes (O) under the wakefulness. N, F and S are derived from the preoperative diagnosis of EEG files. Fragments from epileptogenic areas are recorded in group F, and those from contralateral hemispheric hippocampal structures are recorded in the group N. Set N and F contain only activities measured during epileptic intervals, and set S contains the epileptic activity. The dataset Z includes signals from the normal people, and S contains epileptic seizures.

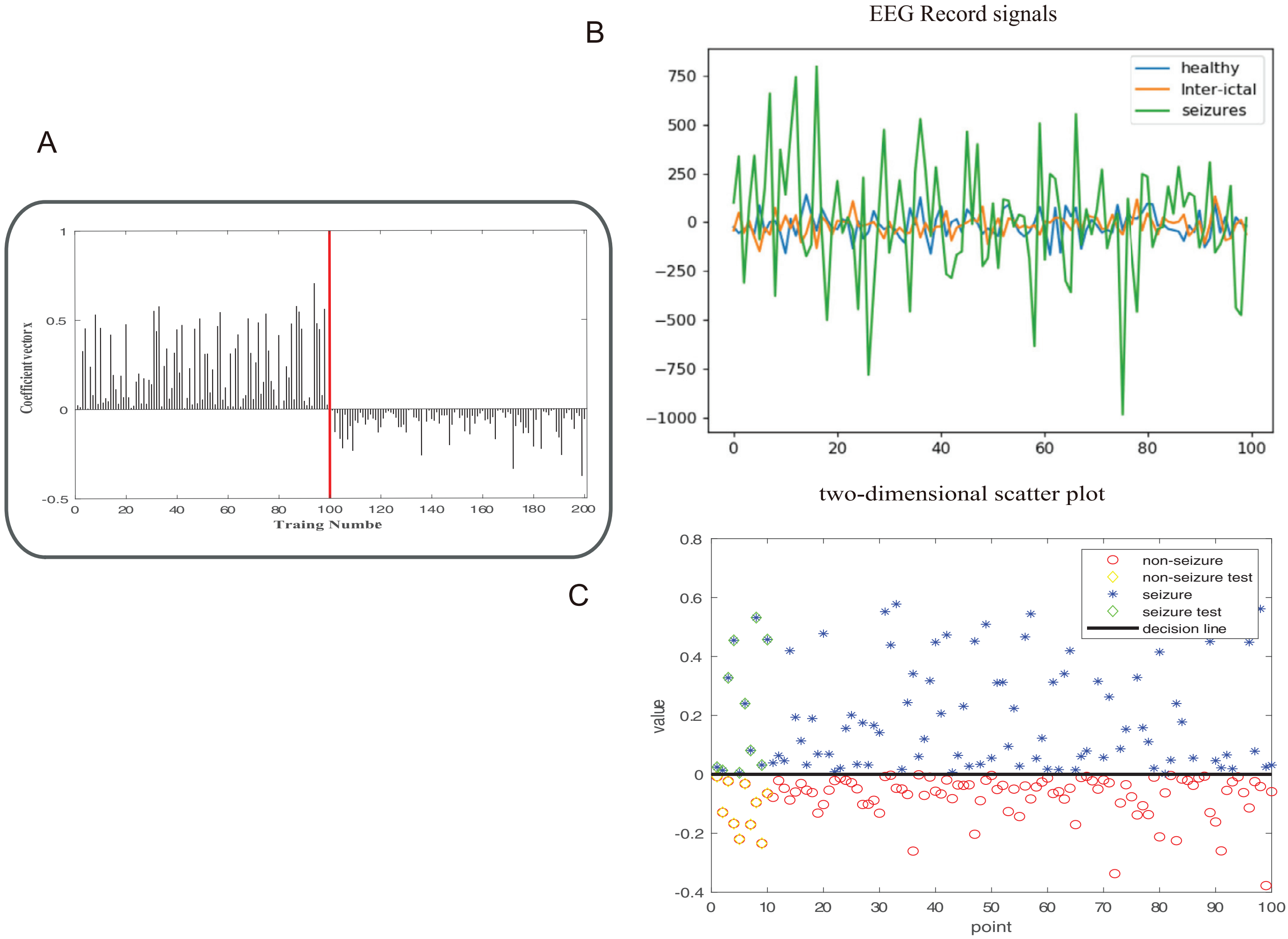

Fig. 4. Signal sparse representation and scatter plot.

EEG signals with the same 128 channel amplifier are recorded. The data is digitalized at 12 bits resolution and 173.6 samples per second. The bandpass filter is set to 0.5-3-60 Hz. The total number of EEG signals is 300 (100 normal signals, 100 intermittent and 100 seizures). Each dataset has 4096 sampling points. Moreover, in order to facilitate the requirements of following classification table, we rename the groups Z, O, N, F and S to the sets A, B, C, D and E.

Another EEG database was provided by the CHB-MIT [27–29] is publicly free and available on www.physionet.org. 23 Children with epilepsy have been recorded by placing 23 electrodes on the scalp of each subject. The scalp EEG dataset has been sampled at 256 Hz. The study included 17 females that ranged in age from 1.5 to 19 years and five males that ranged in age from 3 to 22 years. The age and gender information of a child is unknown. Subjects 1 and 21 were from the same female patient with 1.5 year apart and are considered as two extra patients in this paper. The start and end times of seizures are clearly labeled according to expert judgment, and the number and duration of seizures vary from subject to subject.

5. Result

In the evaluation of epileptic detection systems, the assessment is usually quantified by sensitivity, specificity, recognition accuracy and AUC which are defined as follows:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{22}$$

$$\text{Specificity} = \frac{TN}{TP+FN} \tag{23}$$

$$\text{Recognition Accuracy} = \frac{TN+TP}{TP+FN+TN+FP} \tag{24}$$

$$\text{AUC} = \frac{\sum I(P_M, P_N)}{M*N} \tag{25}$$

• Sensitivity (Sen): the number of true positives/the total number of ictal EEG epochs labeled by the EEG specialists. True positive represents the ictal EEG identified by algorithm and experts.

• Specificity (Spe): the number of true negatives/the total number of interictal EEG epochs labeled by the EEG specialists. True negative represents the interictal EEG identified by algorithm and experts.

• Recognition accuracy (Acc): the number of correctly identified epochs/the total number of epochs.

• Are under curve (AUC): the predicted probability of the positive epochs is greater than the predicted probability of a negative epochs. M represents a positive epochs and N represents a negative epochs.

In this section, we provide the classification results with the Bonn university and CHB-MIT available databases as described in

Table 3
The Bonn dataset: demographic prediction performance comparison by three evaluation metrics. Sens%, Spec% and Acc% represent the sensitivity, specificity and recognition accuracy, respectively.

Classification		Dictionary size											
		173		256		512		1024		2048		4096	
		DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC
E–A	Sens%	93	93	93	86	94	92	95	97	98	98	100	100
	Spec%	83	59	88	47	95	77	100	97	99	99	100	100
	Acc%	88	76	90.5	66.5	94.5	84.5	97.5	97	99.5	98.5	100	100
	AUC	0.9473	0.9056	0.9644	0.7813	0.987	0.947	0.9945	0.9904	0.9994	0.9999	1	1
E–B	Sens%	94	92	96	93	100	93	100	99	100	99	100	100
	Spec%	87	65	90	51	95	72	95	87	97	94	95	90
	Acc%	90.5	78.5	93	72	97.5	82.5	97.5	93	98.5	96.5	97.5	95
	AUC	0.9744	0.9249	0.9905	0.8714	0.9973	0.9574	0.9994	0.9965	0.9996	0.9989	1	1
E–C	Sens%	87	76	89	73	93	73	95	83	97	93	98	96
	Spec%	98	86	98	79	100	99	100	100	100	100	100	100
	Acc%	92.5	81	93.5	76	96.5	86	97.5	91.5	97.5	96.5	99	98
	AUC	0.9773	0.9234	0.9872	0.8345	0.9964	0.9771	0.9993	0.9965	0.9999	0.9998	0.9998	0.9996
E–D	Sens%	88	79	88	69	95	73	95	84	97	92	99	97
	Spec%	95	91	100	85	100	100	100	99	99	100	100	100
	Acc%	91.5	85	94	77	97.5	86.5	97.5	91.5	98	96	99.5	98.5
	AUC	0.9757	0.8569	0.9919	0.8569	0.9985	0.9862	0.9993	0.9954	0.9993	0.9994	0.9996	0.9997
Average	Sens%	**90.5**	85	**91.5**	80.25	**95.5**	82.75	**96.25**	90.75	**98**	95.5	**99.25**	98.25
	Spec%	**90.75**	75.25	**94**	65.5	**97.5**	87	**98.75**	95.75	**98.75**	98.25	**98.75**	97.5
	Acc%	**90.625**	80.125	**92.75**	72.875	**96.5**	84.875	**97.5**	93.25	**98.375**	96.75	**99**	97.875
	AUC	**0.9739**	0.9178	**0.9820**	0.836	**0.9948**	0.9669	**0.9981**	0.9947	**0.9996**	0.9994	**0.9999**	0.9999

Table 4
The CHB-MIT database, the results of different subjects recognition accuracy.

Subject		Dictionary size											
		173		256		512		1024		2048		4096	
		DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC
1	Acc%	84.5	86	88.5	84	91.5	72.5	94.5	96	80	98.5	94.5	85.5
2	Acc%	95.5	93	95.5	93.5	97.5	84.5	97	87	97.5	92.5	98	94
3	Acc%	88.5	88.5	87	89	92	73.5	94	72.5	95.5	72.5	94	86.5
4	Acc%	92	91	93	91.5	94	70	98.5	83	100	81.5	99.5	82.5
5	Acc%	76	83	73.5	80	83.5	70	92	77.5	96	85	95.5	93
6	Acc%	83	83.5	89.5	82.5	92	84.5	95.5	89	96	89.5	96	93.5
7	Acc%	91.91	91.5	94.5	93	96	81	98.5	82	99.5	84.5	98	91.5
8	Acc%	68.5	71.5	78.5	75.5	90.5	74	97.5	75	99.5	82	98.5	91.5
9	Acc%	96.5	71.5	97.5	94.5	98	86	96.5	92.5	99	97	99.5	95.5
10	Acc%	78	77.5	81.5	83.5	84	78	92.5	81	97	76	97	82
11	Acc%	83.5	81.5	83.5	81	92.5	96	95	68.5	93.5	62.5	91	57.5
12	Acc%	72	69.5	79	74.5	89.5	61.5	91	62.5	91	70	93.5	54.5
13	Acc%	74.5	75.5	76.5	75	86.5	50	87	54.5	89	49	89	46.5
14	Acc%	81.5	81.5	86	87	89.5	56	89	60.5	90	62	92.5	60
15	Acc%	87	89	91	88.5	91	56	93	52	92.5	60.5	94.5	54
16	Acc%	79	79	83	74.5	86	55	86	47.5	87.5	56.5	91	61
17	Acc%	76.5	79	77.5	73.5	82.5	65	82	61	91	62	95	58
18	Acc%	90	91	93.5	84	94.5	68	92.5	62.5	95.5	59	98	65.5
19	Acc%	86.5	88	88	81.5	91.5	68	92	62	92	64.5	96	69
20	Acc%	81	76	82.5	79.5	90	55	90.5	64	91.5	57	91	56
21	Acc%	86.5	84.5	86	85	87.5	61.5	93.5	59.5	92	61.5	89	54
22	Acc%	88.5	86.5	89	84	92	66	92.5	65.5	92	64.5	94.5	70
23	Acc%	88.5	89	94.5	90.5	98	77.5	98	82	99	91	99	89.5
24	Acc%	83.5	79	85	77.5	90	72	93	77	95.5	81	97	93.5
Average	Acc%	**83.75**	83.75	**82.79**	83.46	**90.83**	70.06	**92.98**	70.35	**91.61**	72.56	**95.06**	74.35

Section 4, which compare the results with SRC. The sensitivity, specificity, recognition accuracy, and AUC are calculated from the results of each fold. Then, the mean values of these statistical measures over the 10-folds are treated as the actual estimates of classification performance.

All the experiments are executed in Matlab 2018a environment running in Inter core processor with 2.7 GHz. For the Boon EEG dataset, as depicted in Fig. 4(A), the green, yellow, and blue lines represent the original EEG signals of seizures, the inter-ictal, and the health, respectively. We take a segment of each signal, a total of 100 points, and the time point and amplitude as their respective abscissa and ordinate. Furthermore, we can clearly know that the amplitude of the health and inter-ictal are relatively stable, and the spike is not easy to occur. However, the seizures signal is prone to the spike phenomenon. Fig. 4(B) shows 180 trail signals of the epileptic and the non-epileptic.

We draw one EEG signal from sets A, B, C, D and E, respectively. For each EEG subset (A, B, C, D or E). In addition, we consider the following four binary classification problems that use different combinations of these data sets in this paper.

1. Classification of sets A and E: only A and E segments are used and they are classified into two classes: normal or seizure.

2. Classification of sets B and E: only B and E segments are used and they are classified into two classes: normal or seizure.

Table 5
The CHB-MIT database, the results of different subjects specificity.

Subject		Dictionary size											
		173		256		512		1024		2048		4096	
		DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC
1	Spec%	80	76	85	74	89	80	95	67	94	79	92	84
2	Spec%	95	93	96	95	98	83	96	83	95	89	96	93
3	Spec%	92	93	93	96	95	80	97	71	98	70	95	86
4	Spec%	89	88	94	90	93	55	97	75	100	69	99	71
5	Spec%	66	77	69	66	79	67	91	79	94	82	95	91
6	Spec%	80	76	85	74	88	78	93	82	94	87	93	92
7	Spec%	87	89	95	93	98	78	99	76	99	80	96	90
8	Spec%	75	78	86	85	94	80	100	82	100	83	99	92
9	Spec%	96	99	98	98	98	86	95	90	99	95	99	98
10	Spec%	80	80	82	91	91	84	91	83	97	73	98	81
11	Spec%	83	78	82	76	92	64	96	62	93	57	88	44
12	Spec%	70	71	78	71	88	60	92	68	91	75	90	57
13	Spec%	74	73	71	75	89	48	86	50	87	46	86	44
14	Spec%	79	83	83	88	91	59	88	58	89	60	93	60
15	Spec%	84	91	89	85	92	63	93	53	93	62	94	64
16	Spec%	76	79	85	79	86	53	89	53	89	57	92	69
17	Spec%	78	85	75	73	84	62	84	56	90	52	95	44
18	Spec%	91	88	94	75	95	59	94	51	94	45	97	70
19	Spec%	83	81	83	73	91	62	91	54	90	55	94	61
20	Spec%	77	66	76	71	88	46	88	62	91	54	90	56
21	Spec%	86	84	86	87	85	64	95	68	91	63	92	49
22	Spec%	88	86	90	86	93	74	92	71	91	69	93	77
23	Spec%	95	90	96	93	99	67	98	76	99	89	99	83
24	Spec%	89	91	89	93	99	83	97	81	98	79	99	92
Average	Spec%	**83.04**	83.13	**85.83**	82.79	**91.46**	68.13	**93.21**	69.79	**94**	69.58	**94.33**	72.83

Table 6
CHB-MIT database, the results of different subjects sensitivity.

Subject		Dictionary size											
		173		256		512		1024		2048		4096	
		DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC
1	Sens%	89	96	92	94	94	65	94	73	98	81	97	87
2	Sens%	96	93	95	92	97	86	98	91	100	96	100	95
3	Sens%	85	84	81	82	89	67	91	74	93	75	93	87
4	Sens%	95	94	92	93	95	85	100	91	100	94	100	94
5	Sens%	86	89	78	94	88	73	93	76	98	88	96	95
6	Sens%	86	91	94	91	96	91	98	96	98	92	99	95
7	Sens%	91	94	94	93	94	84	98	88	100	89	100	93
8	Sens%	62	65	71	66	87	68	95	68	99	81	98	91
9	Sens%	97	92	97	91	98	86	98	95	99	99	100	93
10	Sens%	76	75	81	76	77	72	94	9	97	79	96	83
11	Sens%	84	85	83	86	93	74	94	75	94	68	94	71
12	Sens%	74	68	80	78	91	63	90	57	91	65	87	52
13	Sens%	75	78	82	75	84	52	88	59	91	52	92	49
14	Sens%	84	80	89	86	88	53	90	63	91	64	92	60
15	Sens%	90	87	93	92	90	49	93	51	92	59	95	44
16	Sens%	82	79	81	70	86	57	83	42	86	56	90	53
17	Sens%	75	73	80	74	81	68	80	66	92	72	95	72
18	Sens%	89	94	93	93	94	77	91	74	97	73	99	61
19	Sens%	90	95	93	93	92	74	93	70	94	74	98	77
20	Sens%	95	86	89	88	92	64	93	66	92	61	92	56
21	Sens%	87	85	86	83	90	59	92	51	93	60	86	59
22	Sens%	89	87	88	82	91	58	93	60	93	60	96	63
23	Sens%	82	88	93	88	97	88	98	88	99	93	99	96
24	Sens%	78	67	81	62	81	61	89	73	93	83	95	95
Average	Sens%	**84.88**	84.38	**86.92**	84.25	**90.21**	69.75	**92.75**	71.92	**95**	75.58	**95.38**	75.88

3. Classification of sets C and E: only C and E segments are used and they are classified into two classes: normal or seizure.

4. Classification of sets D and E: only D and E segments are used and they are classified into two classes: normal or seizure.

Besides, we performed seizure detection on the CHB-MIT EEG database by recognition rate, sensitivity, and specificity, the AUC evaluation indicators judges the performance of the proposed algorithm, and the result has been shown in Tables 4–7, respectively.

There have been many studies on the identification of epileptic EEG signals for this database in recent years. The utilization of the same EEG dataset is necessary for a more precise performance comparison between the proposed method and other existing methods.

In Table 3, we see that the sensitivity, specificity and recognition accuracy of DLWH are higher than SRC under different dictionary size, which is consistent with our previous predictions. When the dictionary size is taken to 4096, we can clearly see that the average sensitivity, specificity and recognition rate are both over 99%, which are significantly higher than SRC index. When the dictionary size is taken to 173, the average sensitivity, specificity and

Table 7
CHB-MIT database, the results of different subjects AUC.

Subject		Dictionary size											
		173		256		512		1024		2048		4096	
		DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC	DLWH	SRC
1	AUC	0.9276	0.9584	0.9688	0.9423	0.9815	0.819	0.9939	0.7805	0.9966	0.8815	0.9946	0.9151
2	AUC	0.9919	0.9861	0.9949	0.9769	0.9948	0.9274	0.9979	0.9474	0.9989	0.964	0.9998	0.9819
3	AUC	0.9494	0.9474	0.9613	0.9571	0.9816	0.7953	0.9902	0.7697	0.9953	0.804	0.9959	0.9498
4	AUC	0.9726	0.9713	0.9749	0.9751	0.994	0.8508	1	0.9177	1	0.9338	1	0.9321
5	AUC	0.8446	0.9011	0.852	0.9028	0.9445	0.7781	0.9903	0.8514	0.9937	0.9204	0.996	0.9755
6	AUC	0.9191	0.9252	0.9724	0.9351	0.9878	0.9078	0.9969	0.9615	0.9959	0.9734	0.9973	0.9786
7	AUC	0.974	0.977	0.9823	0.9776	0.9969	0.8963	0.9998	0.911	1	0.9502	1	0.9629
8	AUC	0.8111	0.7968	0.8935	0.8616	0.9763	0.786	0.9946	0.8061	0.9995	0.8668	0.9992	0.9484
9	AUC	0.9915	0.9915	0.9962	0.9777	0.998	0.9061	0.9988	0.9725	0.9999	0.9959	0.9999	0.9819
10	AUC	0.8806	0.8776	0.9137	0.9322	0.9209	0.875	0.9802	0.8835	0.9984	0.8266	0.9983	0.9032
11	AUC	0.9094	0.9002	0.9035	0.8948	0.9462	0.7829	0.953	0.8078	0.9488	0.739	0.9235	0.7061
12	AUC	0.8295	0.7831	0.8662	0.7972	0.936	0.6768	0.9246	0.717	0.92	0.7645	0.9365	0.5878
13	AUC	0.7669	0.8092	0.8152	0.8042	0.8958	0.5267	0.8682	0.5842	0.8758	0.5103	0.8795	0.4847
14	AUC	0.8714	0.8583	0.9145	0.8902	0.8973	0.6287	0.8872	0.6637	0.9041	0.6801	0.9457	0.6395
15	AUC	0.9194	0.9322	0.949	0.9327	0.9322	0.5613	0.9411	0.5675	0.9606	0.6104	0.971	0.5365
16	AUC	0.8496	0.8433	0.8694	0.8181	0.8791	0.5341	0.8699	0.511	0.8818	0.5882	0.9	0.6116
17	AUC	0.8371	0.8801	0.8477	0.8071	0.8816	0.7455	0.8816	0.7076	0.8924	0.92	0.996	0.6619
18	AUC	0.9372	0.9457	0.9496	0.9311	0.9512	0.7948	0.9546	0.7494	0.9695	0.7192	0.9819	0.6588
19	AUC	0.9246	0.9527	0.9328	0.9247	0.9429	0.762	0.9497	0.7499	0.9724	0.7778	0.9718	0.7925
20	AUC	0.8608	0.8659	0.909	0.8917	0.8987	0.6116	0.9176	0.6847	0.9039	0.6476	0.8998	0.6493
21	AUC	0.8734	0.9092	0.8844	0.9112	0.9227	0.6748	0.9423	0.6432	0.9174	0.6529	0.9169	0.5926
22	AUC	0.9242	0.9289	0.9378	0.9208	0.948	0.7506	0.933	0.7489	0.9392	0.7362	0.9771	0.7885
23	AUC	0.9693	0.9597	0.9851	0.9653	0.9947	0.9106	0.9962	0.9322	0.9966	0.9634	0.9988	0.9706
24	AUC	0.9146	0.8944	0.9378	0.9013	0.9868	0.8375	0.9926	0.8196	0.9979	0.8554	0.9981	0.9744
Average	AUC	**0.9021**	0.9081	**0.9289**	0.9095	**0.9496**	0.7641	**0.9569**	0.7787	**0.9619**	0.795	**0.9699**	0.7993

recognition rate of DLWH are 90.5%, 90.75%, 90.625%, respectively. Nevertheless the SRC gets a bad effect. At the same time, in the evaluation metric of AUC, we can clearly note that when the dictionray size is small (i.e. 173 and 256), the AUC of DLWH is more obvious than that of SRC. When the Dictionary size is 4096, the average AUC is close to 0.99985. As we all know, when the value of AUC is 1, it is a perfect classifier.

Tables 4 –7 represent the recognition accuracy, specificity, sensitivity, AUC of the CHB-MIT 24 subjects and the statistical average of all the results. Then, we could clearly see that the sensitivity, specificity, and recognition accuracy of DLWH are higher than SRC in different dictionary size. When the dictionary size is set to 4096, we can clearly see the average specificity, sensitivity, recognition accuracy and AUC are 94.33%, 95.38%, 95.06%, 0.9699%, respectively, and the difference between DLWH and SRC. Besides, when the dictionary size is 173, the average specificity, sensitivity, recognition rate and AUC are 83.04%, 84.88%, 83.75% and 0.9021% respectively, showing excellent performance. Later, we will discuss the differential interpretation of the results of the EEG datasets at the Bonn University and CHB-MIT.

We explain the dictionary learning scheme in Section 3.3, dictionary size, i.e., the number of columns in the dictionary, which depends on the number of classes and the feature dimension.

Fig. 5 shows the classification results by the different dictionary size, respectively. The subgraph Fig. 5(A)–(D) represents four classification (E–A, E–B, E–C and E–D), and Fig. 5(E) is the statistical average of the four classification cases of Bonn University of EEG dataset. Fig. 5(F) is the result of statistical average of CHB-MIT EEG dataset. Here, we compare classification accuracy for the different dictionary size in SRC and DLWH, respectively. We note that too small number of atoms in the dictionary show poor classification accuracy. Because if the number of atoms in dictionary is too small, new test feature cannot be represented will with similar features in the dictionary. Therefore, it resulted in low discrimination power. Fig. 5 from 2056 to 4096 shows good classification performance for both approach (SRC and DLWH). On the other hand, computation time is also increased with large dictionary size.

6. Discussion

In this research, the sparse representation with the dictionary learning is proposed for the seizure detection, and the optimization algorithm is employed to train the train sets to acquire a better dictionary, which can represent the test sample well. Besides, we improve the generalization ability of the model and the sparsity of sparse representation by introducing ℓ_1 regularization. In the preprocessing part, the bandpass filter is applied to both seizure part and non-seizure background, which makes the data cleaner.

The training signals of all classes can be directly used as the dictionary for the sparse representation. Furthermore, the dictionary learning has been proven to be essential for producing good results, which can process one element of the train signal. In this paper, the residuals for each category are calculated and compared by the dictionary learning and sparse coding which test signals.

The dictionary learning method can achieve faster performance than the traditional algorithm. To compare with the cost of our improved sparse representation method, the time taken in the detection processing is calculated. The average difference in execution time is negligible with the same standard desktop computer with 2.7 GHz core and 8 GB RAM running Matlab 2018a. We compare the running time with the SRC algorithm. In consequence, it can be found that the execution time of DLWH algorithm is 19.671 s, which is less than that of SRC algorithm, i.e., 465.8 s in the training state. Furthermore, there are obvious differences in their running time, which makes the system possibly adapt to develop the online seizure detection system or the portable detection device. Apart from that, in the future work, we will focus on the ability of DLWH for the seizure detection, which satisfies the clinic requirement.

The Bonn University EEG database used in this research has been applied in many researches for evaluating the seizure detection algorithm. It is more desirable to compare the proposed method with other existing work. The performance comparisons between the normal (Set A) and the ictal (Set E) EEG signals with correct classification rates are listed in Table 8. The classification of the normal (Set A) and ictal (Set E) EEG is easier than that of the interictal (Set D) and ictal (Set E) EEG. Hence, the reported

Please cite this article as: H. Peng, C. Li and J. Chao et al., A novel automatic classification detection for epileptic seizure based on dictionary learning and sparse representation, Neurocomputing, https://doi.org/10.1016/j.neucom.2019.12.010

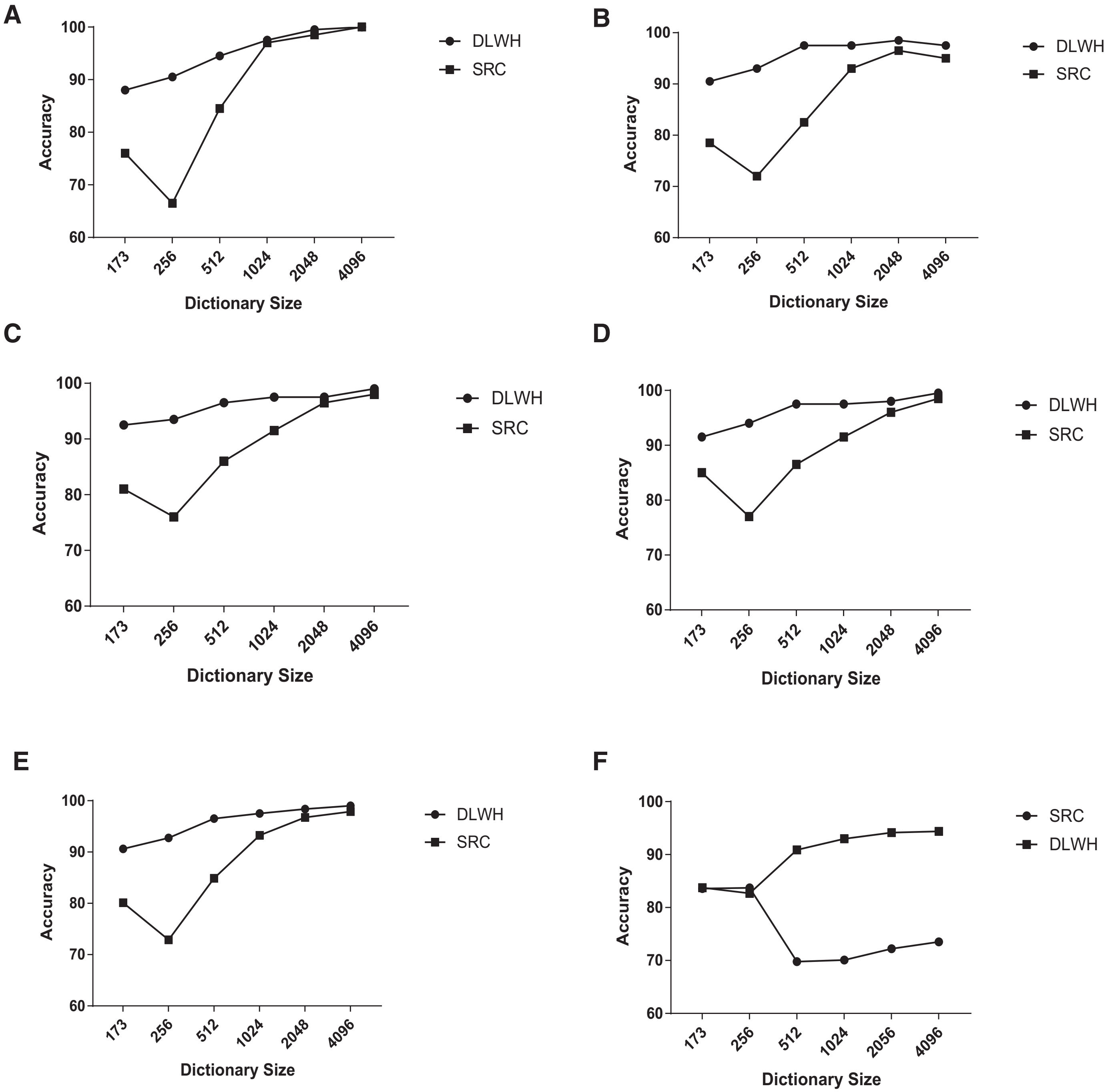

Fig. 5. Comparison of classification accuracy for different dictionary size. The subgraph Fig. 5(A)–(D) represents four classification (E–A, E–B, E–C and E–D) of the Bonn university dataset, Fig. 5(E) represents the average result of the bonn dataset, Fig. 5(F) represents the average result of the CHB-MIT database.

results plotted in Table 8 are all quite high. However, the proposed method has achieved much high accuracy, sensitivity and specificity. Polat et al. employed the fast fourier transformation for the feature extraction and decision tree classification with set A and set E, which produced accuracy rates, sensitivity and specificity of 98.72%, 99.40%, and 99.31%, respectively. Fu et al. presented the HMS and SVM methods, which were used to extract features for the classification of the epileptic EEG signals, yielding accuracy of 99.85%. For set B and set E, the weighted visibility graph combined with SVM is presented, with the 97.25% recognition accuracy satisfactory. For both set C and set E, Siuly et al. presented a clustering technique-based least square support vector machine (CT-LS-SVM), which yielded the recognition rate, sensitivity and specificity of 98.5%, 99.3%, and 97.7%, respectively. For class D and class E, alzami et al. employed the DWT and adaptive hybrid feature selection within bagging with MLP approch, which achieved good results. In this work, the proposed detection system have generated 99.5%, 99%, 100% results in accuracy, sensitivity, and specificity, respectively. Apart from that, we just list some of them, and a detailed description on the other seizure technique is shown in Table 8. For three classes of seizure detection, a number of researchers have implemented different methods, which are listed in Table 6. In particular, Tawfik et al. extracted the valid features by the weighted permutation entropy, and as the input to the SVM classifier, producing a recognition rate of 97.5% [66]. Meanwhile, we also compare and summarize the research

Please cite this article as: H. Peng, C. Li and J. Chao et al., A novel automatic classification detection for epileptic seizure based on dictionary learning and sparse representation, Neurocomputing, https://doi.org/10.1016/j.neucom.2019.12.010

Table 8
Summaries on existing approaches to epilepsy detection using features extracted from EEG signals using the same dataset used in this study.

Classes	Author	Features extracted method	Accuracy (%)	Sensitivity	Specificity
A–E	Nigam [52]	Nonlinear preprocessing filter, diagnostic artificial neural network	99.6%	–	–
	Srinivasan [3] 0	Time-frequency domain feature, recurrent neural network	97.2%	–	–
	Polat [18]	FFT-decision tree classifier	98.72%	99.4%	99.31%
	Subasi [23]	Discrete wavelet transform, mixture of expert model	95%	95%	94%
	Guo [24]	Discrete wavelet transform-relative wavelet energy, MLP	95.2%	98.17%	92.12%
	Wang [20]	Wavelet transform and Shannon entropy, KNN	99.45%	–	–
	Nicolaou [21]	Permutation entropy, SVM	93.55%	–	–
	Fu [19]	Time-frequency image using HHT, SVM	99.13%	–	–
	Fu [53]	HMS analysis, SVM	99.85%	–	–
	Our	**DLWH**	**100%**	**100%**	**100%**
B–E	Supriya [54]	Weighted visibility graph and SVM	97.25%	–	–
	Our	**DLWH**	**97.5%**	**100%**	**95%**
C–E	Samiee [55]	Rational discrete short-time fourier transform and MLP	98.5%	99.3%	97.7%
	Siuly [56]	Clustering technique-based LS-SVM	98.5%	99.3%	97.7%
	Our	**DLWH**	**99%**	**98%**	**100%**
D–E	alzami [35]	DWT and adaptive hybrid Feature selection	97.33%	97.76%	97.49%
	Siuly [58]	Least Square Support Vector Machine (LS-SVM)	94%	88%	100%
	Samiee [55]	Rational Discrete Short Time Fourier Transform	94.90%	95.6%	94.1%
	Kaya [57]	One-dimensional local pattern (1D-LBP)	95.50%	96%	95%
	Siuly [56]	Clustering technique-based least square support vector machine	93.60%	89.4%	97.80%
	Kumar [22]	Discrete wavelet transform analysis and approximate entropy	93%	94%	92%
	Our	**DLWH**	**99.5%**	**99%**	**100%**

Table 9
Performance comparison between the proposed method and other works for classifying three-class problem of the normal (Set A), Interictal (Set D) and ictal (Set E) EEG.

Author	Features extracted method	Accuracy (%)
Chua et al. [59]	Higher order spectra and power spectral density	93.11%
Ubeyli et al. [30]	Wavelet transform+mixture of expert model	93.17%
Orhan et al. [60]	Discrete wavelet transform	96.7%
Acharya et al. [61]	Recurrence quantification analysis parameters+ SVM	95.6%
Song et al. [62]	Sample entropy+extreme learning machine	95.67%
Acharya et al. [63]	Entropies+Higher order spectra+Higuchi fractal dimension+Hurst exponent]+fuzzy	99.7%
Martis et al. [64]	Intrinsic time-scale decomposition+[energy+fractal dimension+sample entropy]	95.67%
Murugavel and Ramakrishnan [65]	Wavelet transform based statistical features	96%
Kaya et al. [57]	1D-LBP+BayesNet	95.67%
Tawfik et al. [66]	Weighted permutation entropy+SVM	97.5%
Our research	**DLWH**	**97.6%**

methods and experimental results with other researches, which are provided in Table 9. Apart from that, Many researches have used the CHB-MIT EEG database for evaluating seizure detection algorithms. Fergus et al. proposed a system for detecting epilepsy with linear discriminant analysis (LDA), which achieved 86.26%, 87.58%, and 86.93% results in accuracy, sensitivity, and specificity, respectively [67]. In particular, Ahammad et al. proposed a method based on wavelet features and certain features of statistical features without wavelet decomposition, which was classified for normal and epileptic EEG signals using the linear classifier, the overall accuracy rate reached 84.2% [69]. Furthermore, a new multi-channel EEG seizure detection method was presented based on the dynamics of the trajectories in phase spacee by Zabihi [70]. Besides, in [71], a new classification approach called collective network of binary classifier (CNBC) was presented, which achieved good performance. Many researchers use different algorithms for seizure detection, which is listed in Table 10.

In Section 5, two EEG databases effectively evaluate our proposed algorithm. The highest recognition rate and the average recognition rate of DLWH in the Boon university EEG database are 100% and 99.99%,% respectively. Meanwhile, the high and average recognition rates of another database CHB-MIT are 99.5% and 95.06%, respectively. Since the Boon university database is directly obtained from the intracranial brain, so the intracranial EEG has the advantages of constant potential, high signal-to-noise ratio and small artifacts, which makes the EEG signal quality outstanding. The CHB-MIT signal is collected from the scalp electrode, which contains more noise and may result in reduced feature quality. Due to the spatial averaging effect of the dura and skull [72], intracranial EEG also includes features not observed in the scalp EEG, and has a high signal-to-noise ratio and fewer artifacts. Besides, the CHB-MIT is collected at the scalp electrode, which contains more noise interference, furthermore, the focal area recorded by

Please cite this article as: H. Peng, C. Li and J. Chao et al., A novel automatic classification detection for epileptic seizure based on dictionary learning and sparse representation, Neurocomputing, https://doi.org/10.1016/j.neucom.2019.12.010

Table 10
Performance comparison of the CHB-MIT database seizure detection research and classification results.

Author	Classifier	Subjects	Sens (%)	Spec (%)	Acc (%)
Fergus et al. [67]	LDA	23	88	88	–
Gill et al. [68]	GMM	12	86.26	87.58	86.93
Ahammad et al. [69]	Linear classifier	24	–	98.5	84.2
Zabihi et al. [70]	LDA and Navie Bayesian	24	88.27	93.21	93.11
Kiranyaz et al. [71]	CNBC	24	93	89	–
Khan et al. [73]	LDA	5	83.60	100	91.80
Our	DLWH	24	95.38	94.33	95.06

the brain electrode contains more redundant information, which may result in the extraction of low quality features.

By contrast, our research based on the sparse representation classification of EEG signals gets the excellent performance, and the ℓ_1-minimization for sparse coding takes less time than other existing methods. Therefore, it is effective to avoid the problem of the feature selection and improve the operation speed. Besides, it will be expected to be applied to the clinic.

In future work, we will consider the post-process procedure to the detection system, such as wavelet filtering, the differential filtering, Kalman filter and kernel trick mentioned [38]. For example, Khan et al. introduced the collar technology to compensate for the missed seizures and make seizures more accurate [73].

7. Conclusions

In this paper, we have proposed a classification method for epilepsy EEG signals with a dictionary learning based on the sparse representation, rather than using the original training sample directly as a dictionary. First, we have employed the EEG training sample set as the basis of sparse representation, and the dictionary to obtain the sparse representation coefficient on the EEG by homotopy algorithm. In this system, a test sample training set by minimizing the ℓ_1-norm has been proposed. Finally, the EEG sample has been tested with a sparsely coded subject learning dictionary. Then, the reconstruction errors of the seizures and non-epileptic EEG training samples on the test samples are calculated to determine the label of the test samples, which are used for making the decision. Therefore, the classification method have avoids some problems, such as the feature selection, and the information carried by the EEG signal that is completely retained. Hence, its fast speed makes sense for the real-time seizure detection. At the same time, we have created the speed of operation, sensitivity and specificity, as well as recognition accuracy improved to meet requirements of automatic seizures detection. Lastly, the proposed seizure detection system demonstrates the practical applications of real-time function.

Declaration of Competing Interest Statement

None

CRediT authorship contribution statement

Hong Peng: Conceptualization, Writing - review & editing, Writing - original draft. **Cancheng Li:** Conceptualization, Writing - review & editing, Writing - original draft, Data curation, Formal analysis. **Jinlong Chao:** Data curation, Formal analysis, Writing - original draft. **Tao Wang:** Writing - review & editing, Writing - original draft. **Chengjian Zhao:** Writing - review & editing, Writing - original draft. **Xiaoning Huo:** Writing - review & editing, Writing - original draft. **Bin Hu:** Conceptualization, Writing - review & editing, Writing - original draft.

Acknowledgments

The author would like to express sincere appreciation to Huiyan Lu at Lanzhou University for her valuable comments and the linguistic modification.

References

[1] J. Jin, R. Chen, X. Zheng, Post-epilepsy stroke: a review, Expert Rev. Neurother. 16 (3) (2016) 341–349.
[2] U.R. Acharya, S.V. Sree, G. Swapna, R.J. Martis, J.S. Suri, Automated EEG analysis of epilepsy: a review, Knowl. Based Syst. 45 (2013) 147–165.
[3] A. Saastamoinen, T. PietilA, A. Varri, M. Lehtokangas, J. Saarinen, Waveform detection with RBF network-application to automated EEG analysis, Neurocomputing 20 (1998) 1–13.
[4] P.D. Watson, K.M. Horecka, R. Ratnam, N.J. Cohen, A phase-locked loop epilepsy network emulator, Neurocomputing 173 (2016) 1245–1249.
[5] E. Pippa, E.I. Zacharaki, I. Mporas, V. Tsirka, M.P. Richardson, M. Koutroumanidis, V. Megalooikonomou, Improving classification of epileptic and non-epileptic EEG events by feature selection, Neurocomputing 171 (2016) 576–585.
[6] M. Parvez, M. Paul, Epileptic seizure detection by analyzing EEG signals using different transformation techniques, Neurocomputing 145 (2014) 190–200.
[7] K. Zeng, J. Yan, Y. Wang, A. Sik, G. Ouyang, X. Li, Automatic detection of absence seizures with compressive sensing EEG, Neurocomputing 171 (2016) 497–502.
[8] B. Hu, X.W. Li, S.T. Sun, M. Ratcliffe, Attention recognition in EEG-based affective learning research using CFS+KNN algorithm, IEEE/ACM Trans. Comput. Biol. Bioinf. 15 (1) (2016) 38–45.
[9] X. Li, D.W. Song, P. Zhang, Y.Z. Zhang, Y.H. Hou, B. Hu, Exploring EEG feature in cross-subject emotiion recognition, Front. Neurosci. 12 (162) (2018).
[10] E.H. Reynolds, The prevention of chronic epilepsy, Epilepsia 29 (S1) (1998) S25–S28.
[11] Z. Zheng, G.Y. Liu, Z.J. Yao, W.H. Zheng, Y.W. Xie, T. Hu, Y. Zhao, Y. Yu, Y. Zou, J. Shi, J. Yang, T.C. Wang, J. Zhang, H. B, Changes in dynamics within and between resting-state subnetworks in juvenile myoclonic epilepsy occur at multiple frequency bands, Front. Neurol. 9 (2018).
[12] R. Fisher, W. Boas, W. Blume, C. Elger, P. Genton, P. Lee, J. Engle, Epileptic seizures and epilepsy: definitions proposed by the international league against epilepsy (ILAE) and the international bureau for epilepsy (IBE), Epilepsia 46 (4) (2005) 470–472.
[13] R.A. Ramadan, A.V. Vasilakos, Brain computer interface: control signals review, Neurocomputing 223 (2017) 26–44.
[14] N. Kannathal, L.C. Min, U.R. Acharya, P.K. Sadasivan, Entropies for detection of epilepsy in EEG, Comput. Methods Progr. Biomed. 80 (2005) 187–194.
[15] J. Gotman, Automatic recognition of epileptic seizures in the EEG. electroencephalogr, Clin. Neurophysiol. 83 (1982) 271–280.
[16] L. Guo, R. Daniel, J. Dorado, C.R. Munteanu, A. Pazos, Automatic feature extraction using genetic programming: an application to epileptic EEG classification, Expert Syst. Appl. 38 (8) (2011) 10425–10436.
[17] L. Guo, D. Rivero, J. Dorado, J.R. Rabunal, A. Pazos, Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks, J. Neurosci. Methods 191 (1) (2010) 101–109.
[18] K. Polat, S. Gnes, Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast fourier transform, Appl. Math Comput. 187 (2) (2007) 1017–1026.
[19] K. Fu, J. Qu, Y. Chai, Y. Dong, Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM, Biomed. Signal Process. Control 13 (2014) 15–22.
[20] D. Wang, D. Miao, C. Xie, Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection, Expert Syst. Appl. 38 (11) (2011) 14314–14320.
[21] N. Nicolaou, J. Georgio, Detection of epileptic electroencephalogram based on permutation entropy and support vector machines, Expert Syst. Appl. 39 (1) (2012) 202–209.
[22] Y. Kumar, M. Dewal, R. Anand, Epileptic seizures detection in EEG using DWT-based apen and artificial neural network, Signal Image Video Process. 8 (7) (2014) 1323–1334.

[23] A. Subasi, EEG Signal classification using wavelet feature extraction and a mixture of expert model, Expert Syst. Appl. 32 (4) (2007) 1084–1093.
[24] L. Guo, D. Rivero, J. Seoane, A. Pazos, Classification of EEG signals using relative wavelet energy and artificial neural networks, in: Proceedings of the First ACM/SIGEVO, Summit on Genetic and Evolutionary Computation, 2010, pp. 1101–1109.
[25] A. Farrikh, J. Tang, Z. Yu, S. Wu, C. Chen, J. You, J. Zhang, Adaptive hybrid feature selection-based classifier ensemble for epileptic seizure classification, IEEE Access 3 (2018) 29132–29145.
[26] R. Andrzejak, K. Lehnertz, F. Mormann, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state, Phys. Rev. E 64 (6) (2001) 061907.
[27] A. Shoeb, H. Edwards, J. Connolly, B. Bourgeois, S.T. Treves, J. Guttag, Patient-specific seizure onset detection, Epilepsy Behavior 5 (4) (2004) 483–498.
[28] A.H. Shoeb, Application of machine learning to epileptic seizure onset detection and treatment, Massachusetts Institute of Technology, Cambridge, MA, USA, 2009 ph. d. dissertation.
[29] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, H.E. Stanley, Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000). e215–e20
[30] E. Beyli, Wavelet/mixture of experts network structure for EEG signals classification, Expert Syst. Appl. 34 (3) (2008) 1954–1962.
[31] S. Altunay, Z. Telatar, O. Erogul, Epileptic EEG detection using the linear prediction error energy, Expert Syst. Appl. 37 (8) (2010) 5661–5665.
[32] Siuly, Y. Li, P. Wen, Clustering technique-based least square support vector machine for EEG signal classification, Comput. Method Progr. Biomed. 104 (3) (2011) 358–372.
[33] S. Chandaka, A. Chatterjee, S. Munshi, Cross-correlation aided support vector machine classifier for classification of EEG signals, Expert Syst. Appl. 36 (2) (2009) 1329–1336.
[34] E.D. beyli, Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals, Expert Syst. Appl. 37 (1) (2010) 233–239.
[35] E. Beyli, Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines, Comput. Biol. Med. 38 (1) (2008) 14–22.
[36] N. Liang, P. Saratchandran, G. Huang, N. Sundararajan, Classification of mental tasks from EEG signals using extreme learning mechine, Int. J. Neural Syst 16 (1) (2006) 29–38.
[37] U.R. Acharya, F. Molinari, S.V. Sree, S. Chattopadhyay, Automated diagnosis of epileptic EEG using entropies, Biomed. Signal Process. Control 7 (4) (2012) 401–408.
[38] J. Li, W. Zhou, S. Yuan, Y. Zhang, C. Li, Q. Wu, An improved sparse representation over learned dictionary method for seizure detection, Int. J. Neural Syst. 26 (1) (2015) 1550035.
[39] M.Z. Ahmad, A.M. Kamboh, S. Saleem, A.A. Khan, Mallats scattering transform based anomaly sensing for detection of seizures in scalp EEG, IEEE Access 5 (2017) 16919–16929.
[40] S. Khanmohammadi, C. Chou, Adaptive seizure onset detection framework using a hybrid PCA-CSP approach, IEEE J. Biomed. Health Inf. 22 (1) (2018) 154–160.
[41] S. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Trans. Signal Process 41 (12) (1993) 3397–3415.
[42] M. Aharon, M. Elad, A. Bruckstein, K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation, IEEE Trans. Signal Process 54 (11) (2006) 4311–4322.
[43] J. Wright, A. Yang, A. Ganesh, S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. 31 (2) (2008) 210–222.
[44] E. Candes, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies? IEEE Trans. Inf. Theory 52 (12) (2006) 5406–5425.
[45] J. Mairal, F. Bach, J. Ponce, G. Sapiro, Z. Andrew, Supervised dictionary learning, Adv. Neural Inf. Process. Syst. 21 (2008) 1033–1040.
[46] J. Mairal, F. Bach, J. Ponce, G. Sapiro, Online learning for matrix factorization and sparse coding, J. Mach. Learn. Res. 11 (2010) 19–60.
[47] M.S. Asif, J. Romberg, L1 homotopy:, A matlab toolbox for homotopy algorithms in l1 norm minimization problems (2013). 12–1
[48] D. Donoho, Y. Tsaig, Fast solution of l1-norm minimization problems when the solution may be sparse, IEEE Trans. Inf. Theory 54 (11) (2008) 4789–4812.
[49] J. Mairal, F. Bach, J. Ponce, G. Sapiro, Online learning for matrix factorization and sparse coding, J. Mach. Learn. Res 11 (2010) 19–60.
[50] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. 11 (7) (1989) 674–693.
[51] H. Lee, A. Battle, R. Raina, A.Y. Ng, Efficient sparse coding algorithms, in: Proceedings of Neural Information Processing Systems, 2007, pp. 801–808.
[52] Y. Song, P. Lio, A new approach for epileptic seizure detection: sample entropy based feature extraction and extreme learning machine, J. Biomed. Sci. Eng. 3 (6) (2010) 556–567.
[53] K. Fu, J. Qu, Y. Chai, T. Zou, Hilbert marginal spectrum analysis for automatic seizure detection in EEG signals, Biomed. Signal Process. Control 18 (2015) 179–185.
[54] S. Supriya, S. Siuly, H. Wang, J. Cao, Y. Zhang, Weighted visibility graph with complex network features in the detection of epilepsy, IEEE Access 4 (2016) 6554–6566.
[55] K. Samiee, P. Kovacs, M. Gabbouj, Epileptic seizure classification of EEG time-series using rational discrete short-time fourier transform, IEEE Trans. Biomed. Eng. 62 (2) (2014) 541–552.
[56] Siuly, Y. Li, P. Wen, Clustering technique-based least square support vector machine for EEG signal classification, Comput. Methods Progr. Biomed. 104 (3) (2011) 358–372.
[57] Y. Kaya, M. Uyar, R. Tekin, S. Yildirim, 1D-local binary pattern based feature extraction for classification of epileptic EEG signals, Appl. Math. Comput. 243 (2014) 209–219.
[58] N. Siuly, Y. Li, P. Wen, EEG Signal classification based on simple random sampling technique with least square support vector machine, Int. J. Biomed. Eng. Technol. 7 (4) (2011) 1752–6426.
[59] K. Chua, V. Chandran, R. Acharya, C.M. Lim, Automatic identification of epilepsy by HOS and power spectrum parameters using EEG signals: A comparative study, in: Proceedings of the Thirtieth Annual International Engineering in Medicine Biology Society Conference, 2011, pp. 3824–C3827.
[60] U. Orhan, M. Hekim, M. Ozer, EEG Signals classification using the k-means clustering and a multilayer perceptron neural network model, Expert sys. appl. 38 (10) (2011) 13475–13481.
[61] U.R. Acharya, S.V. Sree, S. Chattopadhyay, W. Yu, P. Ang, Application of recurrence quantification analysis for the automated identification of epileptic EEG signals, Int. J. Neural Syst. 21 (3) (2011) 199–211.
[62] Y. Song, J. Crowcroft, J. Zhang, Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine, J. Neurosci. Methods 210 (2) (2012) 132–146.
[63] U. Acharya, S. Sree, P. Ang, Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals, Int. J. Neural Syst. 22 (2) (2012) 125002.
[64] R. Martis, U. Acharya, J. Tan, A. Petznick, L. Tong, C. Chua, E. Ng, Application of intrinsic time-scale decomposition (ITD) to EEG signals for automated seizure prediction, Int. J. Neural Syst. 23 (5) (2013) 1350023.
[65] A. Murugavel, S. Ramakrishnan, An optimized extreme learning machine for epileptic seizure detection, Int. J. Comput. Sci. 41 (4) (2014) 212–221.
[66] N.S. Tawfik, S. Youssef, M. Kholief, A hybrid automated detection of epileptic seizures in EEG records, Comput. Electr. Eng. 53 (2016) 177–190.
[67] D. Fergus, A.J. Hussain, D. Al-Jumeily, An advanced machine learning approach to generalised epileptic seizure detection, Intell. Comput. Bioinf., LNCS (2014) 112–118.
[68] A.F. Gill, S.A. Fatima, M.U. Akram, S.G. Khawaja, S.E. Awan, Analysis of EEG signals for detection of epileptic seizure using hybrid feature set, in: Theory and Applications of Applied Electromagnetics, Springer, 2015, pp. 49–57.
[69] N. Ahammad, T. Fathima, P. Joseph, Detection of epileptic seizure event and oneset using EEG, BioMed Res. Int. 2014 (2014) 1–7.
[70] M. Zabihi, S. Kiranyaz, A.B. Rad, A.K. Katsaggelos, M. Gabbouj, T. Ince, Analysis of high-dimensional phase space via poincare section for patient-specific seizure detection, IEEE trans, Neural Syst. Rehabil. Eng. 24 (3) (2016) 386–398.
[71] S. Kiranyaz, T. Ince, M. Gabbouj, D. Ince, Automated patient-specific classification of long-term electroencephalography, J. Biomed. Inf. 49 (2014) 16–31.
[72] A. Shoeb, J. Guttag, Application of machine learning to epileptic seizure onset detection, in: Proceedings of the ICML, 2010.
[73] Y.U. Khan, R. N, O. Farooq, Automated seizure detection in scalp EEG using multiple wavelet scales, in: Proceedings of the International Conference on Signal Processing, Computing and Control, 2012, pp. 1–5.

Hong Peng received the Ph.D. degree from Lanzhou University, Lanzhou, China. From 2010 to 2011, he was as a Visiting Scholar with the Institute of Computer System, ETH Zurich, Switzerland. He is currently an Associate Professor with the School of Information Science and Engineering, Lanzhou University. He is in charge of three projects from National Natural Science Foundation of China, Central College Foundation Project of Lanzhou University and Youth CrossProject of Lanzhou University. He has authored or co-authored over 30 papers in peer-reviewed journals, conferences, and book chapters. He is involved in the work of biosensors, biological signal processing, and emotional characteristics analysis. His research areas include bioinformation processing and ubiquitous affective computing.

Cancheng Li was born in Guangdong, China. He obtained a bachelors degree in 2017 and is currently pursuing a masters degree at Lanzhou University (Universal Perception and Intelligent Systems Laboratory), China. He is also a Guest Student with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. He has been a reviewer of the IEEE ACCESS journal since 2018. His research interests are bioelectrical signal processing and pattern recognition.

Jinlong Chao received the B.E. degree from Gansu agricultural university, Gansu, China, in 2015. He is currently pursuing the Ph.D. degree in computer science and technology at School of Information Science and Engineering, Lanzhou University, Lanzhou, China. His main research interests include EEG, fNIRS and bioelectrical signal processing.

Tao Wang received the bachelors degree in control engineering in 2018 from Zhengzhou University, China. He is currently pursuing a masters degree in information science with the School of Information Science and Engineering, Lanzhou University, China. He is also a Guest Student with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. His research focuses on affective computing, brain–computer interface, machine learning, and pattern recognition.

Chengjian Zhao received the B.E. degree from Chongqing University, Chongqing, China, in 2016. He is currently pursuing the M.E. degree in Communication and information system at School of Information Science and Engineering, Lanzhou University, Lanzhou, China. His main research interests include pattern recognition, neural networks, intelligent information processing, artificial intelligence and bioelectrical signal processing.

Xiaoning Huo is a chief physician and vice-president at the third people's hospital of Lanzhou City. She received a master degree from Gansu University of Chinese Medicine in 2008. Her research interests focus on drug treatment and psychotherapy of mental illness.

Bin Hu is currently a Professor and the Dean of the School of Information Science and Engineering, Lanzhou University, Lanzhou, China, an Adjunct Professor with Tsinghua University, Beijing, China, and a Guest Professor with ETH Zrich, Zrich, Switzerland. He has authored or co-authored over 200 papers in peer-reviewed journals, conferences, and book chapters, including Science (Supplementary), the Journal of Alzheimers Disease, the IEEE Transactions, the IEEE Intelligent Systems, AAAI, BIBM, EMBS, CIKM, and ACM SIGIR. He is an IET Fellow. He is the Co-Chair of the IEEE SMC TC on Cognitive Computing, a Member at Large of the ACM China, and the Vice President of the International Society for Social Neuroscience (China Committee). His work has been funded as a PI by the Ministry of Science and Technology, the National Science Foundation China, the European Framework Programme 7, EPSRC, and HEFCE, U.K. He has served as a Guest Editor of Science for the special issue on Advances in Computational Psychophysiology and an Associate Editor of the IEEE Transactions on Affective Computing, the IEEE Transactions on Computational Social Systems, Brain Informatics, IET Communications, Cluster Computing, Wireless Communications and Mobile Computing, and Security and Communication Networks.

Please cite this article as: H. Peng, C. Li and J. Chao et al., A novel automatic classification detection for epileptic seizure based on dictionary learning and sparse representation, Neurocomputing, https://doi.org/10.1016/j.neucom.2019.12.010