

Research Article

Automatic Epileptic Seizure Detection Using PSO-Based Feature Selection and Multilevel Spectral Analysis for EEG Signals

Qi Sun ¹, Yuanjian Liu ¹, Shuangde Li,¹ and Chaodong Wang^{2,3}

¹College of Electronic and Optical Engineering & College of Flexible Electronics (Future Technology), Nanjing University of Posts and Telecommunications, Nanjing, China

²Department of Neurology & Neurobiology, Xuanwu Hospital of Capital Medical University, Beijing, China

³National Clinical Research Center for Geriatric Diseases, Beijing, China

Correspondence should be addressed to Yuanjian Liu; liuyj@njupt.edu.cn

Received 30 June 2022; Revised 14 September 2022; Accepted 22 September 2022; Published 28 October 2022

Academic Editor: Kathiravan Srinivasan

Copyright © 2022 Qi Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Automatic epileptic seizure detection technologies for clinical diagnosis mainly rely on electroencephalogram (EEG) recordings, which are immensely useful tools for epileptic location and identification. Currently, traditional seizure detection methods based only on single-view features have great limitations for the typical dynamic and nonlinear EEG signals. An objective of this paper is to investigate the effect of multiview feature selection and multilevel spectral analysis methods on the identification of the EEG signals for seizure detection. Here, multiview features are extracted from time domain, frequency domain, and information theory to collect adequate information of EEG signals. And a feature selection algorithm based on particle swarm optimization (PSO) is proposed for automatic seizure detection. Moreover, due to the different frequency components of the EEG signals, they are divided into four kinds of brain waves for multilevel spectral analysis. The effect of these four rhythm waves on seizure detection is compared. Three well-known classifiers are employed to classify EEG signals concerning seizure or nonseizure events. The result shows that the average accuracy, specificity, and sensitivity of classification with the CHB-MIT database are 98.14%, 98.64%, and 96.79%, respectively. The application of the PSO-based feature selection method for automatic seizure detection improves accuracy by 5.99% with the SVM classifier. Compared with the state-of-the-art methods, the proposed method has superior competence with high performance for automatic seizure detection. It is further shown that the feature selection method is an indispensable step in seizure detection. With PSO-based feature selection and multilevel spectral analysis, the θ wave in the frequency range of 4-7 Hz shows better performance in the identification of EEG signals and is more suitable for the proposed method. The PSO-based feature selection algorithm for automatic seizure detection can be a useful assistant tool for clinical diagnosis.

1. Introduction

Epilepsy is a common neurological disease caused by abnormal brain discharge that affects about 1% of the world population. About 35% of patients with epilepsy can be controlled with medication. The other 8-10% can be treated with surgery. And there is no feasibility treatment strategy for the remaining patients [1]. The symptoms of epileptic seizures accompanied by sudden and periodic brain dysfunction

reflect the clinical state of excessive abnormal brain neurons. Electroencephalogram (EEG) can record this process and detect electrical activity in the brain [2, 3]. The difference between the EEG recordings of normal people and those with epilepsy can be found by reading the EEG recordings. This difference mainly reflects the various modes of EEG in the brain, including rhythmic waves of various frequencies, multiple spikes, and complexes of spike and sharp waves. Reading short-term EEG recordings is easy for experts, but long-

term EEG recordings will be a time-consuming task. Automatic seizure detection technologies can effectively help clinical experts diagnose epilepsy disease.

Automatic seizure detection methods have attracted increasing attention from researchers in recent years. Early automatic seizure detection algorithms were mainly based on empirical criteria [4, 5]. These methods may be insensitive to the identification of seizures. Subsequently, some seizure detection methods based on single-view features emerged successively and used frequency-based feature extraction methods to train classifiers [6–9]. Due to the nonlinear characteristic of EEG signals, frequency-based analysis methods had significant limitations and could not effectively differentiate seizure and nonseizure sequences. Tzallas et al. used the time-frequency (TF) analysis method to obtain energy by calculating the probability density function (PDF) of different TF distributions and used the energy as a characteristic to classify the EEG signals. The classification results verified that the TF analysis method was effective for automatic seizure detection [10]. Zhang et al. constructed an EEG montage and a multiple feature extractor for the time series and different frequency components of the signals [11]. Ech-Choudany et al. designed a TF analysis framework based on dissimilarity for seizure detection [12]. The experimental results showed that it achieved high accuracy for various classification problems (including 2, 3, and 5 classification). The TF method provided joint information of multiview features from time and frequency domain, achieving better performance than the early methods that only considered single-view features.

Several feature extraction methods have been proposed for EEG analysis. They are divided into four domains: time-domain analysis, frequency-domain analysis, TF analysis, and nonlinear methods. In [13], the singular value decomposition and the successive spike interval analysis were used to obtain the low and high frequency features of the signal, respectively. The extracted features were then fed into separate artificial feedforward neural networks for classification and each of the networks included two hidden layers. In [14], 21 features were extracted in each 2 s EEG segment, including time domain, frequency domain, and nonlinear features based on entropy. These features were then inputted into linear discriminant analysis (LDA) for classification. Thomas et al. extracted 55 features from each 8 s EEG segment and used the Gaussian mixture model to detect seizure [15]. Ahmed et al. and Temko et al. used the same set of features and fed these features into the support vector machine (SVM) classifier with radial basis function (RBF) and Gaussian dynamic time warping [16, 17]. Zwanenburg et al. obtained 5 features and used SVM to detect whether epilepsy occurred in a newborn [18]. Logesparan et al. summarized 97 shortlisted publications for comparison and obtained 65 features to evaluate the discriminative performance of these features [19]. When the number of extracted features is large, the dimension of the feature vector will increase, resulting in high computational cost for the identification of seizures. The feature selection method can reduce the dimension of features and improve the accuracy of classification.

The purpose of feature selection is to select a subset of relevant features. Existing work has rarely considered the effect of feature selection on performance classification for seizure detection. Some features extracted from the EEG signals are redundant or irrelevant. The feature selection method can remove redundant information and prevent the loss of useful information. The advantages of feature selection methods include the following: (1) a simplified and easier model for researchers or users to interpret [20]; (2) shorter training time [21]; (3) reduction in dimensionality [22]; (4) improvement of the compatibility of learning models for classification data [23]; (5) encode the inherent symmetries in the input space. Feature selection includes two aspects: search methods for new feature subsets and evaluation metric that scores each subset. According to the different evaluation metric, the feature selection algorithms can be divided into three categories: wrapper, filter, and embedded methods [24]. Wrapper methods use a predictive model to score a subset of features. Each new subset is used to train a model and then tested on the validation set. The feature subset is scored by calculating the number of errors (that is, the error rate of the model) on the validation set. Such methods tend to find the optimal feature set for a particular type of model. Filter-based methods employ proxy metrics instead of scoring error rates based on subsets of features. It is also used in the preprocessing step of wrapping methods, so that wrapping methods can still be used when the problem is too complex. Embedded methods try to combine the advantages of both previous methods. It takes advantage of its own variable selection process and performs feature selection and classification simultaneously. The computational complexity of this method tends to fall between the wrapper and filter methods. In this paper, the wrapper method is used to find the best feature set for seizure detection.

Particle swarm optimization (PSO) is a population-based search algorithm first proposed by Kennedy and Eberhart [25, 26]. It does not use evolutionary operators for individuals but regards each of them as a particle (point) without volume in the D-dimensional search space. A particle flies at a certain speed in this space and adjusts its position and velocity dynamically according to its own experience and the experience of its companions [27]. The PSO algorithm is powerful and easy to implement with high computational efficiency. Due to these advantages of the PSO algorithm, it has evolved into a simple and effective optimization algorithm [28, 29]. In this paper, the PSO-based feature selection algorithm is used to optimize the extracted features. The influence of the feature vectors before and after optimization on the classification of EEG signals is explored.

Several classification strategies are used in seizure detection. SVM classifiers are perhaps the most popular classifiers and are widely used in automatic seizure detection [30–33]. The SVM classifiers find a hyperplane in the sample space based on the training set and separate the samples into different classes. Finding the partitioning hyperplane with the maximum margin yields robust classification results and produces the best generalization to unseen samples. Random forest (RF) classifiers are a kind of ensemble learner which

consists of multiple independent decision tree learners [34, 35]. It collects training samples randomly, so that the samples show diversity, which can effectively increase the generalization of the final ensemble. Based on some distance measurements, k -nearest neighbor (k -NN) classifiers find the k samples that are closest to a given sample in the training set and then make predictions based on the information of these k neighbors [36]. In the feature selection algorithm, the corresponding classifier is used to train the model and obtain the best subset of features of the patient. The test set is predicted to evaluate the performance of the classifier.

Many of feature extraction methods have been proposed for seizure detection. However, feature fusion and selection will still affect the final performance of identification. In [19], the authors focused on selecting an optimal set of features by comparing the performance of single feature for classification. In [37], the authors provided nine statistical features in the time domain to solve the problem of unbalanced EEG data sets. And it has been pointed out in [27] that optimized features improve the performance remarkably. Hassan et al. extracted three features by using empirical mode decomposition and selected the most significant feature for the classification [38]. But the features they used for automatic feature selection were from single domain, which will not obtain adequate information from the EEG signals. In this paper, fusional features from multiview are provided to obtain adequate information on the EEG signals. Furthermore, a PSO-based feature selection algorithm is proposed for automatic seizure detection. The algorithm outputs an optimal set of features to improve the performance of identification of seizure detection through three common classifiers.

In addition, epileptic seizures may cause discrimination in certain frequency bands. EEG signals from epileptic patients are nonstationary with time-varying frequency components. Depending on the multilevel spectrum analysis of the EEG signal, the signal is divided into four frequency bands: δ (0.4-4 Hz), θ (4-8 Hz), α (8-12 Hz), and β (12-30 Hz). And the correlation of these bands is analyzed for seizure detection. In this paper, a PSO-based feature selection method is used for automatic seizure detection to classify EEG signals. The framework of the proposed method is depicted in Figure 1. The innovation of this work is to employ PSO-based feature selection to find optimal feature set for multiview features and improve the accuracy of classification for automatic seizure detection with EEG signals. In addition, multilevel spectral analysis method is used to find which frequency band is better for seizure detection. The proposed method can optimize feature vector and improve the performance of identification of the EEG signals. The main contributions of this paper are provided as follows:

- (1) The multiview features are extracted to provide joint information of signals, achieving better performance than methods that only consider single-view features. Therefore, to obtain adequate information of the EEG signals, we extracted multiview features from the time domain, frequency domain, and infor-

mation theory. In addition, a PSO-based feature selection method is proposed for automatic seizure detection to improve the performance of classification

- (2) Due to the different frequency components of the EEG signals, they are divided into four kinds of brain waves. The multilevel spectral analysis method is used to improve the performance of classification
- (3) Three well-known classifiers are employed to classify EEG signals concerning seizure or nonseizure events with optimal feature subset

The content of this paper is organized as follows. In Section 2, the database used to perform simulation experiments, and the preprocessing method is introduced. And the feature extraction and classification of EEG signals are carried out. The experimental results and the analysis are presented in Section 3. Section 4 discusses comparisons with state-of-the-art seizure detection methods with the same database. Finally, the conclusions are described in Section 5.

2. Methods and Materials

2.1. CHB-MIT EEG Database. The public CHB-MIT database provided by Boston Children's Hospital is used for the proposed method. The database contains long-term EEG recordings from 23 neonatal epilepsy patients with intractable epilepsy [39-41]. Data acquisition is performed by placing electrodes on the scalp of patients. The placement of electrodes on the scalp follows the international standard 10-20 system (see Figure 2). The database collects records of 23 patients between the ages of 1.5 and 22. And the detailed description is shown in Table 1. There are 24 cases in the database, each of them contains 9 to 42 consecutive files from a single subject. In 24 cases, a total of 664 edf files and 198 seizure events are included. All EEG signals are sampled at 256 Hz with a 16-bit resolution. Most of the files contain 23 channels (24 or 26 in some cases). At the same time, due to the continuity of the electrode montages, we cannot read the data of some channels in chb4, chb6, chb12, and chb16. Therefore, the data for these four patients are removed.

2.2. Preprocessing of EEG Signals

2.2.1. Segmentation. In order to train the classification model, 4352 segments have been selected from 20 patients. Taking into account the long-term characteristic of EEG signals, a sliding window of 8 s with 50% overlap is used to better achieve seizure detection. In this paper, the data of the ictal and interictal periods are segmented by a window of 8 s with overlap. The segment obtained from the ictal periods of the signals is labeled positive samples. The segment obtained from the interictal periods of the signals is labeled negative samples. Since the total duration of the interictal periods is much longer than that of the ictal periods. In order to avoid the impact on the classification caused by the additional bias towards a certain type of

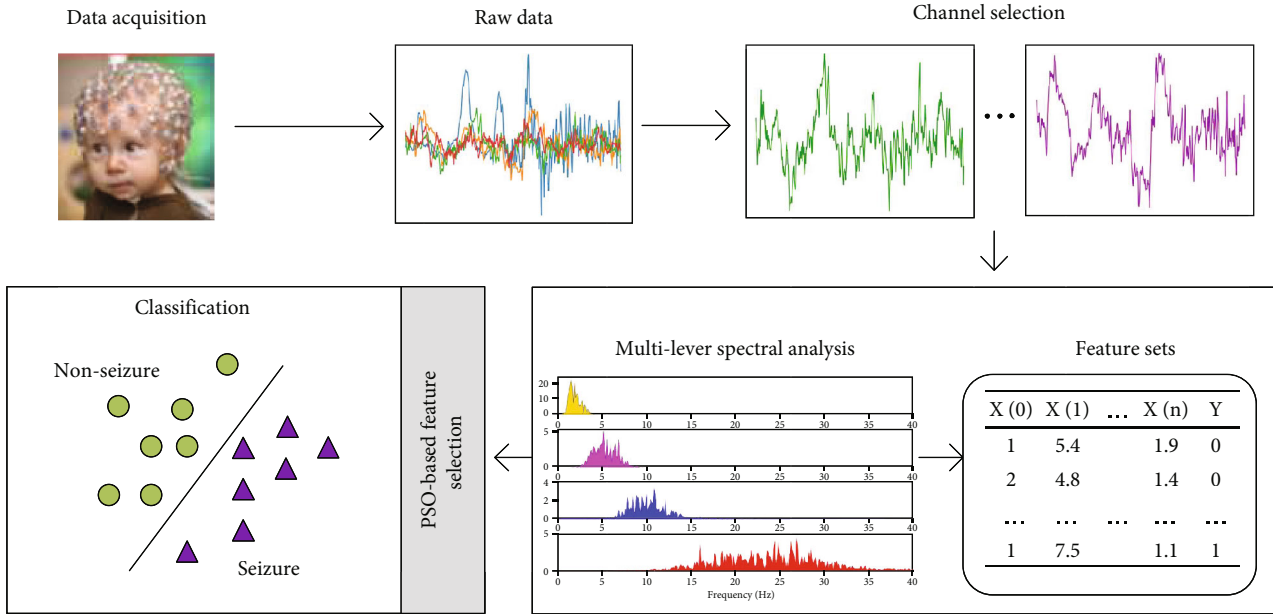


FIGURE 1: The framework of proposed automatic seizure detection method.

samples in the data division process, we ensure consistency of the data distribution during the division process by randomly sampling the negative samples to keep the number of positive and negative samples consistent. Specifically, all seizures of each patient will be used to train the model. And the nonseizure segments are selected randomly to keep the number of seizure segments and nonseizure segments equal.

2.2.2. Filtering. Before implementing the proposed seizure detection method, it is critical to perform signal filtering to remove unwanted components noise from the original EEG signals [42]. For the CHB-MIT database, most epileptiform discharges are below 32 Hz. So, the frequency band commonly used for epilepsy diagnosis is 0.01-32 Hz [43]. Therefore, a fourth-order Butterworth filter is used to filter the EEG signals and preserve the EEG data in that specific frequency band.

2.2.3. Channel Selection. Due to the multichannel characteristic of EEG signals, it is necessary to perform channel selection [44–49]. Most of the cases in CHB-MIT collect 23 channel information. But in some cases, there are 24 or 26. If all channels are used for classification, features cannot be extracted uniformly and will lead to high computational time. Classification performance and computational time must be balanced by selecting channels related to epileptic events in EEG signals for further analysis.

In [46], the authors presented a machine learning based seizure prediction method for channel selection. They found that adaptively selected three to six channels were good enough for the EEG seizure prediction task. In this paper, five channels are selected for analysis of EEG signals. The channel selection method is described below.

The channel with the least standard deviation (SD) is selected firstly. When the spike-rate increases, the amplitude

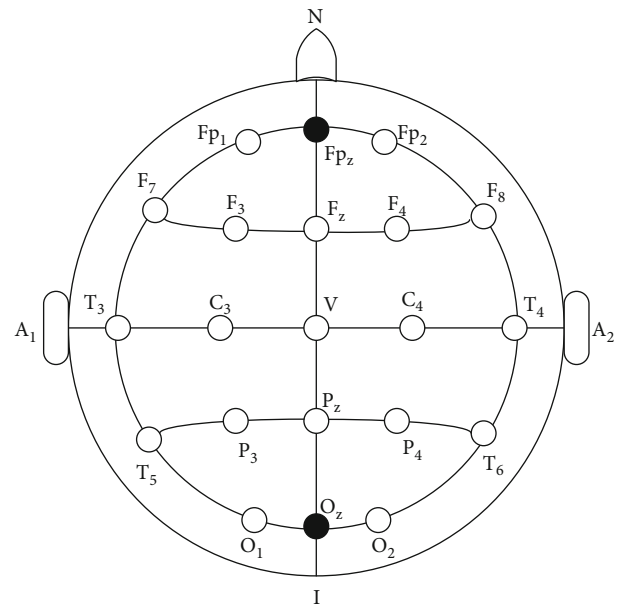


FIGURE 2: The international standard 10-20 system of EEG.

also becomes very high in the seizure segments. Due to unwanted artifacts or muscular movement, it can increase the spike-rate and amplitude in nonseizure segments. Those channels with artifacts will have comparatively high SD, making it hard to distinguish seizure events in long-term EEG recording. Then, the remaining of the four channels are selected according to mutual information (MI). MI is a quantitative measure and is used to find the interdependency or similarity of two random variables. By computing the MI between the first selected channel and the rest of the channels, the four channels that have higher MI with the selected one will be obtained. Specifically, the definition

TABLE 1: The information of the CHB-MIT database.

Patients	Gender-age	Number of seizures ($T_{min} - T_{max}$)(s)	Total seizure time (s)	Total seizure-free time (s)
1	F-11	7(28-102)	449	23475
2	M-11	3(10-83)	175	7983
3	F-14	7(48-70)	409	24791
5	F-7	5(97-121)	563	17437
7	F-14.5	3(87-144)	328	32208
8	M-3.5	5(135-265)	924	17076
9	F-10	4(63-80)	280	34218
10	M-3	7(36-90)	454	50008
11	F-12	3(23-753)	809	9249
13	F-3	12(18-71)	547	28253
14	F-9	8(15-42)	117	25023
15	M-16	20(31-205)	2012	48420
17	F-12	3(89-116)	296	10528
18	F-18	6(31-69)	323	19951
19	F-19	3(78-82)	239	10307
20	F-6	8(30-50)	302	19732
21	F-13	4(13-82)	203	13587
22	F-9	3(59-75)	207	10593
23	F-6	7(21-114)	431	31823
24	NR-NR	16(17-71)	527	42673

of MI of two random variables C_1 and C_2 is introduced as [50]

$$MI(C_1; C_2) = H(C_1) + H(C_2) - H(C_1, C_2), \quad (1)$$

where $H(C_1)$ and $H(C_2)$ are the entropy values of the random variables C_1 and C_2 , respectively. $H(C_1, C_2)$ is the joint entropy of C_1 and C_2 .

2.3. Feature Extraction. Feature extraction methods are highly useful in the automatic detection of epilepsy. Multi-view fusion features will provide adequate information for EEG signals [51]. In this paper, time-domain, frequency-domain, and information theory features are extracted. Time-domain features are obtained from the original signals or from the first and second derivatives of the signals. Frequency-domain features are calculated by the power spectral density of each segment. On the basis of the random characteristic of signals, entropy in information theory is introduced to represent the characteristic information of signals.

2.3.1. Time-Domain Features. Time-domain features are obtained from the original signals or from the first and second derivatives of the signals. Amplitude and interval analysis are applied to extract statistical information in the time domain. For each segment S of signals, maximum, minimum, mean, line length, variance, number of maxima and minima, and root mean square amplitude are computed to obtain the dimensional features according to the recordings of segment S . In addition, the variances of the first difference signal and second difference signal are also considered.

Due to the time-varying nature of the EEG signals, some nondimensional features are used, concerning wave form factor, peak factor, pulse factor, margin factor, skewness, and kurtosis. Furthermore, the number zero crossings of the original signal and its first and second derivatives are also introduced. Nonlinear energy is used to predict seizures in adult epileptic patients and is calculated for each segment shown as [52]

$$NEO = v(t)^2 - v(t-1) * v(t+1). \quad (2)$$

The Hjorth parameters are based on simple statistical calculations on the EEG [53], including activity, mobility, and complexity. Autoregressive modelling (AR) methods are used to analyze the EEG signal [54]. In this paper, the errors of AR modeling with order 1-9 are used. Totally, 31 features are used in the time domain.

2.3.2. Frequency-Domain Features. In frequency domain analysis, the Fourier transform is used to convert the time domain into the frequency domain to obtain features from another perspective. Frequency-domain features are obtained by calculating the power spectral density of each segment, which can reflect the magnitude of signal components at different frequencies. Through the Fourier transform, the peak frequency, median frequency, center frequency, total power, frequency variance, and root mean square frequency are selected. Spectral edge frequencies (SEFs) are calculated as frequencies below which 80%, 90%, and 95% of the total spectral power reside. To obtain a detailed and focal analysis of the frequency, the segment

is decomposed through 8 levels of decomposition with the Daubechies 4 wavelet.

2.3.3. Information Theory. Based on the random characteristic of the signals, the entropy in information theory is introduced to represent the characteristic information of the signals. Entropy is regarded as a numerical measure method of stochastic signal and is widely used to analyze bioelectrical signals, such as EEG signals. In information theory, the entropy of signals is obtained as a feature which is depicted below.

The spectral entropy SE describes the relationship between the power spectrum and the entropy rate. And the calculation formulas is shown as follows:

$$SE = -\sum_i P_i \log(P_i), \quad (3)$$

where P_i is the power spectrum of the i^{th} grouped EEG segment.

The singular value entropy SVE can reflect the distribution obtained from the decomposition and transformation of the signal, which can represent the energy distribution of singular values.

$$SVE = \sum_i \text{svdVal}_i \log(\text{svdVal}_i), \quad (4)$$

where svdVal_i is the i^{th} value of the singular value decomposition of the segment.

The sample entropy SampEn measures the complexity of the time sequence by measuring the probability of generating new patterns in the signals. The large probability of new patterns will yield a high complexity of the sequence [55].

$$\text{SampEn}(m, r, N) = -\ln \left[\frac{A^m(r)}{B^m(r)} \right], \quad (5)$$

where m is the dimension of segmented signals, r is similarity threshold, $A^m(r)$ is the probability of matching $m+1$ points for two grouped sequences with similarity tolerance r , and $B^m(r)$ is the probability of matching m points for two grouped sequences with similarity tolerance of r .

The wavelet entropy WE can reflect the degree of disorder of EEG signals and extract rhythm from nonstationary EEG signals. Through the wavelet energy, the relative energy relationship between the EEG signals is also obtained. Kumar et al. obtained a classification accuracy of 99.75% for normal and interictal EEG signals and 96.30% for interictal and interictal EEG signals by combining wavelet entropy features with a recurrent neural network for classification [56].

$$WE = -\sum_i E_i \log(E_i), \quad (6)$$

where E_i is the relative wavelet energy of the i^{th} subsignal.

Shannon entropy HE characterizes the amount of information contained in the EEG signals. And the calculation of this feature is mainly towards continuous time-varying signals.

$$HE = -\sum_i^N p(x_i) \log_2(p(x_i)), \quad (7)$$

where $p(x_i)$ is the probability of EEG segment x_i . N represents the total number of records of the segment.

In this paper, 46 features of EEG signals are extracted, including 31 time-domain features, 10 frequency-domain features, and 5 features in information theory for each divided segment during the feature extraction process of EEG signals. Specific features are listed in Table 2. The source code in MATLAB to calculate features is uploaded at website: <https://github.com/chloeqisun/multiViewFS/tree/master>.

2.4. Classification. The features previously extracted are inputted into the SVM to train the data and obtain the classifier model. SVMs mainly rely on two assumptions. One is transforming the data into a high-dimensional space implemented by a linear discriminant function, which makes complex classification problems simple. The other is that SVMs use training patterns near the decision surface. It assumes that they can provide more useful information for classification. The classification methods used in this paper mainly consist of two steps: training and testing. The selected kernel of the SVM is the RBF. The number of features obtained from multiview domain is large, and it is linearly inseparable. So, linear kernel is not available. In addition, the training data is further split for validation to avoid overfitting. The formula for RBF kernel is described as follows:

$$K(x_i, x_j) = \exp \left(-\frac{\|x_i - x_j\|^2}{2\sigma^2} \right), \sigma > 0, \quad (8)$$

where σ is the width of the kernel. $K(x_i, x_j)$ is the kernel function, which is based on the dot production of two samples x_i and x_j .

RF is a kind of ensemble machine learning algorithm whose base learner is a decision tree. It is based on bootstrap sampling. In detail, if the feature dimension of each sample is M , the algorithm specifies a constant m , which is much smaller than M , and randomly selects m subsets from M features. Then it selects the optimal subset from these m subsets by pruning the tree. The randomness here makes the algorithm not easy to fall into overfitting and has good antinoise ability. Each decision tree is a classifier. For an input sample, N decision trees will produce N results. RF integrates the classification results of all decision trees through voting, and the class with the most votes will be regarded as the final output. RF is simple and easy to implement with low computational overhead. Compared with a single decision tree, it has higher accuracy and works efficiently on large data

TABLE 2: The list of features.

Type	Feature name
Time domain (31)	Maximum, minimum, mean, variance, line length, skewness, kurtosis, nonlinear energy, root mean square amplitude, the variances of the first and second difference signal, Hjorth parameters (activity, mobility, and complexity), zero crossing (the number of zero crossings of original signal and its first and second derivatives), number of maxima and minima, wave form factor, peak factor, pulse factor, margin factor, errors of AR modeling with order 1-9
Frequency domain (10)	Total power, peak frequency, median frequency, center frequency, frequency variance, root mean square frequency, wavelet energy, spectral edge frequencies (3)
Information theory (5)	Spectrum entropy, singular value entropy, sample entropy, wavelet entropy, energy entropy

sets, which means it can handle data with high-dimensional features without dimensionality reduction.

k -NN is a supervised learning classification algorithm. Given a test sample, the algorithm finds the k closest training samples in the training set based on some distance metric and then makes a prediction based on the information of these k neighbors. Usually, the algorithm employs the voting method to give classification results according to the class of k neighbors.

2.5. PSO-Based Feature Selection Algorithm for Seizure Detection. The phenomenon of feature redundancy occurs after the calculation of feature extraction with a high dimension. For this reason, a PSO-based feature selection algorithm is designed for automatic seizure detection. The algorithm first trains the original feature sets and selects the subset with the highest classification accuracy. Then, the obtained feature vector is inputted into the classifier for classification.

The PSO is initialized as a bunch of random particles (random solutions). Then the optimal solution is found through iteration. In each iteration, each particle i updates itself by tracking two extreme values ($p_{\text{best}}, g_{\text{best}}$). After finding these two optimal values, the particle updates its velocity v_i and position x_i by the following formula:

$$v_{i+1} = w * v_i + c_1 * r_1 * (p_{\text{best}} - x_i) + c_2 * r_2 * (g_{\text{best}} - x_i), \quad (9)$$

$$x_{i+1} = x_i + v_{i+1}, \quad (10)$$

where w is the inertia weight, r_1 and r_2 are random numbers in the range $[0, 1]$. Whereas c_1 and c_2 are learning factors.

In the beginning, 5-fold cross-validation has been used, and (5-1) folds of training data are further split for validation. The trained model is then tested with the fold that has been left aside for testing to find the optimal subset of features. And the ratio of training set, validation set, and test set is 50%, 30%, and 20%, respectively.

The PSO-based feature selection algorithm partitioned the patient data set into training, validation, and test sets, represented by S_{train} , S_{valid} , and S_{test} , respectively. N is the number of particles and max_iter is the maximum iteration of the PSO. For each particle, the algorithm trains the model with current position and computes fitness according to the

following formula:

$$\text{fitness} = \alpha * (1 - \text{Acc}_{\text{avg}}) + \beta * \frac{\text{num_feat}}{\text{max_feat}}, \quad (11)$$

where α and β are weighted parameters. And Acc_{avg} is the accuracy of the classifier computed on S_{train} with 10-fold cross-validation. The original feature size and the number of features selected by the proposed algorithm are represented by max_feat and num_feat , respectively.

In Algorithm 1, the local optimal value and the global optimal value of each particle are updated according to Equation (10). Then, the algorithm updates the speed and position and turns to the next iteration. When the maximum iteration is reached, the algorithm stops and an optimal feature set is obtained. Since 5-fold cross-validation has been used, five feature subsets are selected. In this paper, the feature subset with the best test accuracy Acc_{test} will be selected. Finally, the classifier uses this optimal set of features to evaluate the classification performance of the model.

3. Experimental Result

The proposed PSO-based feature selection method for seizure detection is implemented using MATLAB R2020a and PYTHON 3.8 on a DELL Inspiron 14, Intel Core i5 8th, and RAM 8G. Due to the powerful ability of computation, the signal processing is implemented in MATLAB to obtain feature vectors. Python is used to train a model for the identification of seizures.

The performance evaluation of the method is based on some evaluation metrics, including accuracy (ACC), precision (Pre), sensitivity (SEN), F1-measure (F1), specificity (SPE), negative predictive value (NPV), and area under curve (AUC). Accuracy is one of the most widely used performance metrics in the literature and is defined as the proportion of correctly classified samples in the total samples. Precision is the proportion of positive samples that are correctly predicted. Sensitivity, also known as the true positive rate, is the proportion of seizure samples that are correctly classified as seizure. The higher the sensitivity, the lower the probability of missed detection. F1-measure comprehensively considers precision and sensitivity. Specificity, or true negative rate, refers to the proportion of nonseizure samples that are correctly classified as nonseizure. NPV is a measure of the completeness of the result. The AUC is defined as the area under the receiver operating characteristic (ROC)

Parameters:
N: population size;
patientNum: the number of patients;
K: *K*-fold cross-validation;
max_iter: max iteration of the PSO;
for *p* in *patientNum*: **do**
 for *k* in *K*: **do**
 Data partitioning according to feature vectors of *p*:
 S_{train}, *S_{valid}*, and *S_{test}* represent the training, validation, and test sets, respectively;
 Initialization population: Initialize position and velocity of each particle within permissible range;
 while *t* < *max_iter* **do**
 for *i* in *N*: **do**
 Conduct 10-fold cross-validation on *S_{train}*, and calculate average accuracy *Acc_{avg}*;
 Evaluate the classification accuracy *Acc_{class}* on *S_{valid}*;
 Compute fitness according to (11);
 Update *pbest* and *gbest* optimum of *i*;
 Update the velocity and position of the particle *i*;
 Observe the *Acc_{class}*, when the iteration achieves the best validation accuracy, the training will stop;
 Retrain and build the classifier on *S_{train}* + *S_{valid}* based on the selected feature subset;
 Measure test accuracy *Acc_{test}* on the test set *S_{test}* via the trained classifier;
 Select the feature subset *gbest* with best test accuracy *Acc_{test}*;
 Output: Optimal feature set;

ALGORITHM 1: The proposed PSO-based feature selection algorithm for seizure detection

curve. The ROC curve is based on a series of different cut-off values or thresholds and uses the Sen as the ordinate and the $(1 - \text{Spe})$ as the abscissa. The AUC can measure the performance of the classifier. These evaluation metrics are defined as follows:

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} * 100\%, \quad (12)$$

$$\text{Pre} = \frac{\text{TP}}{\text{TP} + \text{FP}} * 100\%, \quad (13)$$

$$\text{Spe} = \frac{\text{TN}}{\text{TN} + \text{FP}} * 100\%, \quad (14)$$

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}} * 100\%, \quad (15)$$

$$F1 = \frac{2 * \text{Sen} * \text{Pre}}{\text{Sen} + \text{Pre}} * 100\%, \quad (16)$$

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}} * 100\%, \quad (17)$$

where TP (true positive) is the number of seizure samples that are predicted as seizure, TN (true negative) is the number of nonseizure samples that are predicted as nonseizure, FP (false positive) is the number of nonseizure samples that are predicted as seizure, and FN (false negative) is the number of seizure samples that are predicted as nonseizure.

The proposed PSO-based feature selection algorithm for automatic seizure detection can effectively identify seizure and nonseizure from the EEG signals. In the data preprocessing stage, the EEG signals are processed with filters to reserve only the frequency range commonly used for the diagnosis of epilepsy. Additionally, frequency bands com-

monly used for the diagnosis of epilepsy are classified into four different types of brain waves. The discrepancy of these four types of frequency band on the recognition of seizures from EEG signals is compared. At the same time, the effect of feature selection on seizure recognition is also studied.

3.1. Comparisons of the PSO-Based Feature Selection. To investigate the effect of proposed PSO-based feature selection on the classification of seizure detection, we study the performance of seizure detection before and after feature selection. Feature extraction methods obtain features from time domain, frequency domain, and information theory for EEG signals. For each channel, a total of 46 features are extracted. In this paper, five channels are selected for each patient, so the number of features for each patient is 230. Using the multiview feature vectors to classify the EEG signals of each patient can collect adequate information. However, the large number of features may lead to high computational cost. And redundant information sometimes occurs, reducing the performance of classification. Therefore, the PSO-based feature selection algorithm is used to select the best subset of features for a certain patient, and the influence of the classifier on the classification results before and after feature selection is compared. At the same time, three classifiers are also applied to analyze the performance of classification. The results are recorded in Tables 3 and 4. Table 3 shows the classification results obtained by the classifier without feature selection. It can be seen from the table that the accuracy of most patients is above 90% by using multiview feature extraction method. The extracted feature space really achieves better results. But due to the heterogeneity of the EEG signals of the patients, the results among these patients show huge differences. RF obtains better results among these three classifiers

TABLE 3: The classification results without feature selection.

PID	SVM						RF						k-NN								
	Acc	Pre	Sen	F1	Spe	NPV	AUC	Acc	Pre	Sen	F1	Spe	NPV	AUC	Acc	Pre	Sen	F1	Spe	NPV	AUC
1	99.00	97.89	100	98.89	98.26	100	94.55	100	100	100	100	100	100	100	95.01	94.22	95.99	95.06	93.99	95.98	97.4
2	90.75	97.5	83.71	88.83	97.5	87.69	98.21	94.83	96.36	96	95.58	94.29	94.29	100	94.92	94.92	95.56	94.79	95.28	95.56	98.7
3	90.72	88.61	93.44	90.91	87.98	93.27	92.98	96.72	96.14	97.61	96.77	95.17	98.08	99.42	89.05	86.05	93.16	89.38	84.47	92.92	95.15
5	97.74	96.56	99.33	97.91	96.01	99.1	99.83	99.25	98.64	100	99.31	98.36	100	99.97	93.21	93.69	93.4	93.28	94.35	92.24	98.85
7	93.44	95.65	91.01	93.11	96.17	91.32	95.54	93.51	94.99	92.51	93.55	94.79	92.38	99.37	90.19	90.44	89.9	90	91.27	89.12	96.72
8	85.62	90.74	78.69	84.11	92.57	81.52	94.49	95.06	96.19	93.38	94.69	96.54	94.27	98.43	87.64	87.48	88.06	87.75	87.28	87.6	93.4
9	96.8	95.71	97.78	96.54	95.26	98.82	99.49	96.8	97.5	96.9	97.06	96.36	96.9	99.87	94.4	93.85	95.13	94.28	93.6	95.6	97.41
10	98.52	99.05	98.57	98.77	99	97.33	99.95	100	100	100	100	100	100	100	95.56	98.08	93.11	95.5	97.8	93.28	99.16
11	95.19	96.82	93.59	95.15	96.94	93.48	99.01	97.97	99.43	96.26	97.79	99.51	97.8	99.42	95.7	95.66	95.93	95.69	95.31	95.99	97.55
13	92.31	95.24	90.91	93.02	94.12	88.89	97.33	96.38	97.89	94.99	96.31	97.99	95.14	98.11	79.49	75	83.33	78.95	76.19	84.21	81.35
14	84.1	84.57	88.5	83.95	87.14	84.5	88.33	92.18	93.81	90.48	91.32	93.81	92.22	96.81	77.69	81.9	76.48	77.14	81.71	77.4	80.1
15	83.99	88.87	77.99	82.97	90.2	80.32	92.36	97.54	97.45	97.71	97.57	97.44	97.6	99.34	85.38	85.13	86.21	85.4	85.22	85.83	90.82
17	90.58	85.97	97.24	91.06	84.13	96.18	94.17	93.41	91.33	96.47	93.33	92.36	95.38	99.88	81.19	82.54	80.49	80.85	81.74	79.75	90.48
18	88.67	89.63	85.02	86.92	92.79	86.97	94.79	89.41	88.44	91.98	89.55	87.91	91.24	98.51	86.55	87.46	84.94	85.94	88.71	83.98	93.92
19	93.59	96.11	91.06	93.34	96.64	91.34	91.27	97.27	98.33	96.79	97.53	98	96	98.75	93.64	93.5	92.42	92.8	95.49	92.9	97.6
20	85.13	88.92	79.21	83.64	89.84	81.46	90.82	91.7	92.16	92	91.53	91.84	91.68	98.23	84.33	83.4	82.6	82.64	84.56	82.19	86.78
21	89.93	87.32	95.38	90.2	88.44	92.5	99.17	97.71	100	95.28	97.49	100	95.96	99.75	84.31	88.22	82.86	84.16	90.1	80.93	90.6
22	92.57	90.53	95.5	92.63	89.74	95.96	97.27	95.73	96.24	96.67	96.21	95.14	95.56	99.05	95.67	93.46	98.18	95.59	93.86	97.78	96.66
23	95.8	93.3	98.36	95.67	93.89	97.03	99.5	97.92	97.89	98.18	97.94	98.1	97.89	99.56	91.62	88.82	95.77	92.08	87.12	95.58	96.1
24	89.55	96.16	83.85	89.38	95.83	84.01	90.37	92.41	96.41	88.32	91.97	95.83	89.79	97.69	88.63	91.92	85.1	88.12	92.44	86.08	94.52
Average	91.7	92.76	90.96	91.35	93.12	91.08	95.47	95.79	96.46	95.58	95.78	96.17	95.61	99.11	89.21	89.29	89.43	88.97	89.52	89.25	93.66

TABLE 4: The classification results with PSO-based feature selection.

PID	SVM						RF						k-NN								
	Acc	Pre	Sen	F1	Spe	NPV	AUC	Acc	Pre	Sen	F1	Spe	NPV	AUC	Acc	Pre	Sen	F1	Spe	NPV	AUC
1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	97.56	100	95	97.44	100	95.45	100
2	100	100	100	100	100	100	100	97.42	95.68	100	97.71	94.17	100	100	100	100	100	100	100	100	100
3	91.67	100	83.33	90.91	100	85.71	100	97.29	95	100	97.44	94.44	100	100	97.3	95	100	97.44	97.44	100	98.83
5	100	100	100	100	100	100	100	99.24	99.13	99.13	99.13	99.33	99.33	99.97	98.87	99.2	98.67	98.91	99.31	98.33	99.66
7	96.67	100	91.67	95.65	100	94.74	98.61	94.11	93.41	94.65	94.01	93.13	94.38	99.13	93.4	93.3	93.17	93.04	93.64	94.03	93.52
8	95	90.91	100	94.24	90	100	95	96.63	96.08	98	97.02	95.87	97.37	99.31	91.01	92.12	89.32	90.59	92.39	90.13	96.05
9	100	100	100	100	100	100	100	98.4	97.5	100	98.67	96.36	100	99.94	96.8	95.09	98	96.38	95.57	98.75	97.86
10	99.01	98.82	98.95	98.85	99.2	99.09	100	100	100	100	100	100	100	100	99.51	99.13	100	99.56	98.95	100	99.47
11	98.73	100	97.56	98.77	100	97.44	100	98.48	100	96.98	98.47	100	96.94	99.61	97.97	99.43	96.46	97.92	99.53	96.46	99.42
13	92.31	95	90.48	92.68	94.44	89.47	95.24	97.43	100	93.75	96.77	100	95.83	98.91	79.48	73.91	89.47	80.95	70	87.5	82.63
14	92.31	87.5	100	93.33	83.33	100	100	96.67	95.56	100	97.5	92	100	98.57	82.95	92.67	75.87	82.68	93.81	75.95	96.87
15	97.32	100	93.83	96.82	100	95.5	99.88	97.44	97.01	97.84	97.41	97.07	97.87	99.08	96.05	95.81	96.38	96.08	95.67	96.39	98
17	96.3	100	92.31	96	100	93.33	100	93.44	93.79	92.61	93.18	94.41	92.6	100	94.92	93.99	95.6	94.66	94.71	95.54	97.16
18	92.86	80	100	88.89	90	100	97.5	89.43	90.69	88.85	89.16	91.24	88.78	99.22	87.24	89.78	85.82	86.74	91.17	85.56	96.09
19	100	100	100	100	100	100	100	97.27	98.46	95.56	96.7	98	97.33	93.75	97.27	96.18	98	97.05	96.67	98.33	97.67
20	96.42	92.31	100	96	93.75	100	98.44	94.17	95.79	94.13	94.93	93.74	92.2	97.35	85.13	85.21	85.05	84.84	84.79	85.69	90.57
21	100	100	100	100	100	100	100	97.78	100	95.83	97.8	100	95.32	100	88.89	100	81.82	90	100	77.78	98.7
22	100	100	100	100	100	100	100	94.74	90.91	100	95.24	88.89	100	100	95.73	94.57	97.78	96.1	92	97.78	95.29
23	100	100	100	100	100	100	100	97.37	97.84	96.84	97.32	97.95	96.89	100	97.4	98.22	96.51	97.31	97.57	96.9	98.53
24	95.24	100	90.48	95	100	91.3	97.27	91.01	95.3	86.33	90.26	94.91	88.58	97.24	91.94	96.73	87.09	91.54	97.21	87.9	92.06
Average	97.19	97.23	96.93	96.86	97.54	97.33	99.1	96.41	96.61	96.53	96.44	96.08	96.67	99.1	93.47	94.52	93	93.46	94.52	92.92	96.42

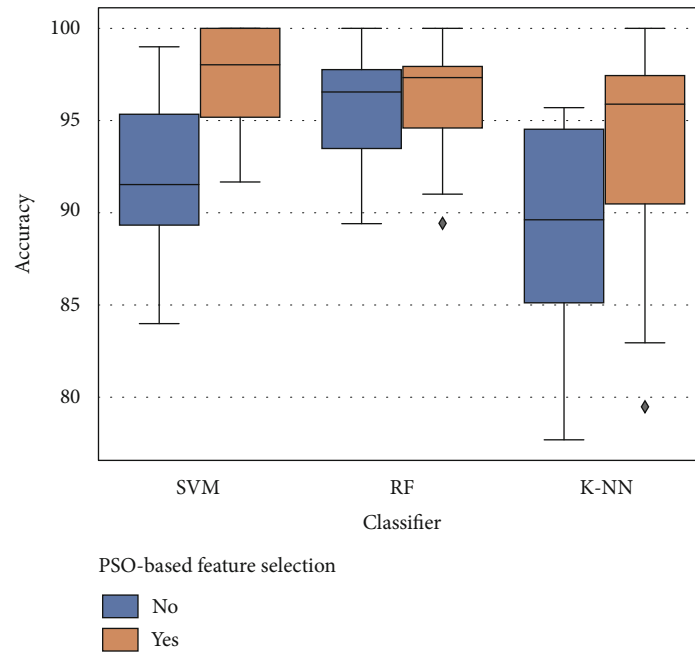


FIGURE 3: The accuracy of feature selection and PSO-based feature selection.

according to the AUC that can reflect the performance of the classifiers. Here, RF is an ensemble algorithm that builds multiple decision trees for classification and selects part of the features to build a decision tree rather than all the features. Table 4 shows the results after using PSO-based feature selection. The average results of the PSO-based feature selection are better than the feature extraction method without feature selection. As shown in Figure 3, after using the PSO-based feature selection algorithm, classification accuracy has been improved significantly. SVM, RF, and k -NN with PSO-based feature selection improve accuracy by 5.99%, 0.65%, and 4.78%, respectively. It can be seen from the figure that the PSO-based feature selection algorithm achieves a better improvement with SVM and k -NN. When the RF classifier is training the model, it always randomly selects a part of the features for training, not all the features, which can be regarded as a kind of feature selection method. Therefore, the classification results do not improve significantly. But after using the PSO-based feature selection algorithm, the results of accuracy, specificity, and sensitivity are also improved significantly. SVM obtains the best result among these three classifiers. The disadvantage of the kernel function RBF of SVM is that it is easy to overfit. So, the training data is further split for validation to avoid overfitting. The proposed algorithm achieves high sensitivity and satisfactory specificity in all 20 patients. Except patients 3, 10, 11, and 14, the classification accuracy of the other patients is above 95% with SVM classifier, which also verifies the effectiveness of the algorithm for the seizure detection of EEG signals.

3.2. Comparisons of the Number of Features. To analyze the number of features after selection based on the PSO algorithm, we recorded the optimal feature subset for each patient. The PSO-based feature selection algorithm pro-

posed in this paper can reduce the dimension of features, and thus reduces the computational time. The dimension of the original feature vector is $46 \times 5 = 230$. After the feature selection algorithm is executed, the average number of features of SVM, RF, and k -NN are 61, 103, and 98, respectively. Figure 4 shows the number of features selected for each patient. As can be seen from the figure, the number of features is reduced with feature selection. The number of features selected by the SVM classifier is lower than that of the RF and k -NN classifier, and the classification accuracy achieved by the SVM classifier is also better than those. Therefore, the SVM classifier shows better performance for seizure detection with the proposed method. Feature extraction or deep processing occupies the main time for automatic seizure detection. In this paper, the run time for feature extraction is provided. For each 8 s segment, Table 5 shows the impact of the selected features on the computational time. The run time of computing the feature vector is 14.71 s for the seizure detection method without feature selection. When the algorithm finds the optimal features, the computational time is reduced observably. The run time is 3.93 s for SVM, which is less than RF and k -NN. This can be seen that the PSO-based feature selection algorithm cannot only reduce the computational time of feature vectors but also improve the accuracy of classification.

3.3. Comparisons of the Four Brain Waves. To detect the detailed information of EEG signals, we used multilevel spectral analysis method to study the effect of the different types of brain waves on the identification of seizure. In the data preprocessing stage, filters are used to obtain the four brain waves in different frequency ranges, including δ , θ , α , and β waves. The δ wave is a kind of brain wave that occurs during deep sleep. The θ wave is usually found in a subconscious state and appears in childhood. And some adults who

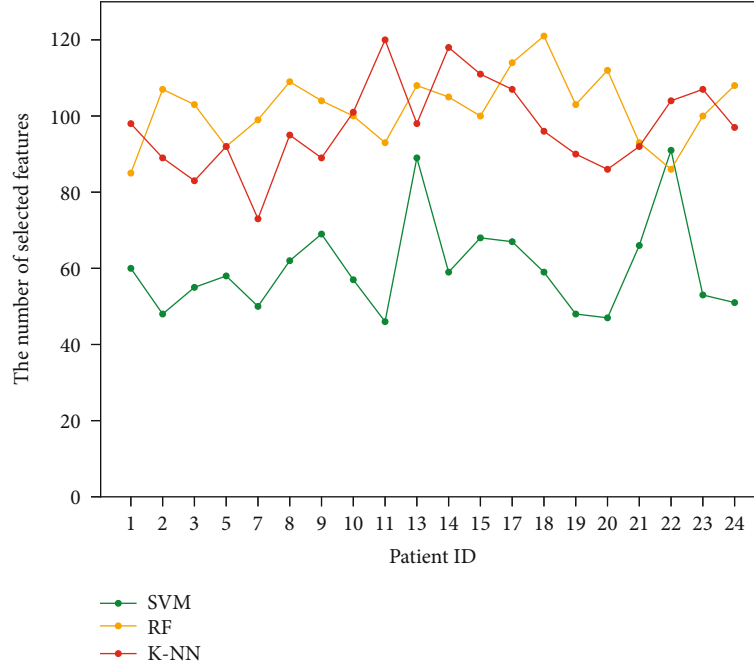


FIGURE 4: The number of features selected by three classifiers.

TABLE 5: The run time of feature extraction.

Classifiers	Without feature selection		PSO-based feature selection	
	Feature numbers	Time (s)	Feature numbers	Time (s)
SVM	230	14.71	61	3.93
RF	230	14.71	103	6.63
k-NN	230	14.71	98	6.31

TABLE 6: The classification results of four brain waves.

Brain waves	Acc	Pre	Sen	F1	Spe	NPV	AUC
δ	95.35	97.13	93.56	95.12	97.02	93.03	98.28
θ	98.14	98.74	96.79	97.69	98.64	96.85	99.28
α	94.6	98.26	90.78	94.19	98.2	91.83	97.57
β	94.66	94.75	93.78	94.03	95.64	94.31	98.02
0.01 – 32Hz	97.19	97.23	96.93	96.86	97.54	97.33	99.1

experience emotional stress or patients with brain disease also have such a wave. The α wave reflects the rhythm of neurons in the brain during wakefulness and rest. The β wave occurs in a state of awake consciousness when people experience stress or anxiety. According to the four brain waves, the EEG data is first filtered through a bandpass filter. The features of the filtered signals are extracted and used to train and classify these four brain waves with the proposed PSO-based feature selection algorithm. The experimental results with the SVM classifier are shown in Table 6. In terms of accuracy, sensitivity, and specificity, the classification performance of θ wave is better than that of other brain

waves and the frequency range of 0.01-32 Hz. The AUCs of these different frequency components are above 97%. Therefore, the SVM classifier can be a useful tool for identifying the seizures from these waves. The θ wave achieves the best result with accuracy of 98.14%. The experimental results show that the θ wave in the frequency range of 4-7 Hz is more suitable for the proposed seizure detection method.

4. Discussions

To further verify the generalizability and effectiveness of the proposed method, the comparison between the proposed seizure detection method and the state-of-the-art methods in the literature is given in Table 7. For feasibility purposes, the same database is used for all experiments. Regarding the CHB-MIT database, EEG signals are scalp data with artifacts that have implications for seizure detection. To verify the effectiveness of the proposed automatic seizure detection method on long-term multichannel seizure EEG signals. The CHB-MIT database is used in the experiment, and the proposed automatic seizure detection method is employed to automatically identify seizure and nonseizure events for EEG signals. Table 7 presents the results of the patient-specific performance evaluation and outlines the state-of-the-art automatic seizure detection method for the CHB-MIT database, where NR indicates that the value is not recorded. As can be seen from the results, the average sensitivity, specificity, and accuracy of the proposed method are better than those of the most recent research, achieving promising performance in this benchmark database.

Due to the different settings of the experimental environment, such as the number of patients and the number of channels, it is difficult to compare directly. Taking into account the feasibility of the comparison, all evaluation

TABLE 7: Comparisons with state-of-the-art seizure detection methods in the CHB-MIT database.

Authors	Methodology	No. of patients-channels	Sen(%)-Spe(%)-Acc(%)
Kiranyaz et al. [57]	Time domain, frequency domain, TF, nonlinear features	21-18	89.0-94.7-NR
Zabihi et al. [58]	Seven features from intersection sequence, LDA	23-23	89.1-94.8-94.6
Li et al. [48]	CE-stSENet	21-5	92.41-96.05-95.96
Chandel et al. [36]	Orthonormal triadic wavelet-based features, k -NN	18-22	98.36-99.62-99.45
Tian et al. [59]	Multiview features, CNN	24-23	96.66-99.14-98.33
Peng et al. [49]	Stein kernel-based SR	20-5	97.85-98.57-98.21
Chen et al. [30]	Discrete wavelet transform, seven wavelet and two statistical features, SVM	18-22	91.71-92.89-92.3
Jiang et al. [31]	Symplectic geometry eigenvalues, SVM	22-5	97.17-99.72-99.62
Hassan et al. [38]	EMD, multilayer perceptron neural network (MLPNN) classifier	23-22	NR-NR-99.57
This work	46 features from time and frequency domain and information theory, PSO-based feature selection, SVM	20-5	96.79-98.64-98.14

results are captured from the same benchmark database. Kiranyaz et al. used multidimensional PSO to evolve a collective network. They achieved a sensitivity above 89% and a specificity above 93% [57]. Zabihi et al. extracted seven features from the intersection sequence by reconstructing the signal trajectories and obtained a sensitivity of 88.27% with a 25% training set [58]. Li et al. proposed a novel framework named CE-stSENet in which the sensitivity, specificity, and accuracy were 92.41%, 96.05%, and 95.96%, respectively [48]. Chandel et al. applied triadic wavelet decomposition for offset and onset detection to obtain feature sets which yielded higher classification performance [36]. Tian et al. first constructed multiview features and then used CNN to learn features from them. They obtained higher accuracy compared to single view and common feature extraction methods [59]. Peng et al. used the Stein kernel-based sparse representation (SR) method to construct EEG signals in symmetric positive definite matrices, and the average accuracy reached 98.21%. And it could also be applied in real-time detection [49]. Chen et al. developed a framework to search for the optimal setting of the discrete wavelet transform and acquired promising performance [30]. Jiang et al. collected features from the decomposition of the symplectic geometry, and the average accuracy of the experimental result was 99.62% [31]. Based on the optimization of the features in our work, the sensitivity, specificity, and accuracy of the experimental evaluation of the data set on average are 98.21%, 98.57%, and 97.85%, respectively. They verified that the method can effectively classify long-term EEG signals. Hassan et al. extracted three features via empirical mode decomposition (EMD) to improve the performance of identification [38].

Existing work rarely considers the temporal and spatial information of the signal and the effect of feature selection for seizure detection. The advantages of this work are: (1) adequate features are extracted from multi domain and can improve the performance of classification for automatic sei-

zure detection; (2) a PSO-based feature selection algorithm is proposed. The results show that it cannot only reduce the computational time but also improve the accuracy of classification; (3) we distinguish different types of brain waves through multilevel spectrum analysis. And the results have proven that θ wave is suitable for the proposed method. Furthermore, optimal features selected by PSO-based feature selection methods can improve the accuracy of classification, so the feature selection method plays a crucial role in seizure detection. And the limitations of this work are:

- (1) Finding a larger EEG dataset that contains more patients. Since our model used to learn various patterns of epileptic seizures is patient-specific, it cannot generalize to the patterns across different patients. Transfer learning is an effective way to train cross-patient model on bigger dataset
- (2) Artifact noises that existed in the original signals prevent us from extracting effective features. In the future work, denoising will be conducted before feature selection in the future work

5. Conclusions

With an increasing requirement for clinical applications, automatic seizure detection methods have been a frontier study in the assistant diagnosis of nervous diseases. To improve the classification efficiency, the channel selection method is used to reduce computational time for the multi-channel EEG. The features of the EEG signals are extracted from time domain, frequency domain, and information theory to obtain adequate information about the EEG signals. Then, an automatic seizure detection method via PSO-based feature selection is presented. Furthermore, three well-known classifiers are used to find seizure or nonseizure events in EEG signals. The experimental results demonstrate

that the seizure detection method via PSO-based feature selection improved accuracy, specificity, and sensitivity by 5.99%, 4.75%, and 6.56% with the SVM classifier, respectively. At the same time, considering the different frequency components of the EEG signals, the signals are divided into four brain waves to test and analyze the effect of different components on seizure detection. The results show that the rhythm signal of θ is more suitable for the proposed automatic seizure detection method.

Data Availability

The public CHB-MIT database provided by the Boston Children's Hospital is used in this paper. It can be found in the website: <https://physionet.org/physiobank/database/chbmit/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Natural Science Fund under Grant no. 61871232 and by the Research Innovation Program for College Graduates of Jiangsu Province in 2021 under Grant no. KYCX21_0718.

References

- [1] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Automatic seizure detection based on time-frequency analysis and artificial neural networks," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 080510, 13 pages, 2007.
- [2] W. Liang, L. Cheng, and M. Tang, "Identity recognition using biological electroencephalogram sensors," *Journal of Sensors*, vol. 2016, Article ID 1831742, 9 pages, 2016.
- [3] M. Antonijevic, M. Zivkovic, S. Arsic, and A. Jevremovic, "Using AI-based classification techniques to process EEG data collected during the visual short-term memory assessment," *Journal of Sensors*, vol. 2020, Article ID 8767865, 12 pages, 2020.
- [4] A. Liu, J. S. Hahn, G. P. Heldt, and R. W. Coen, "Detection of neonatal seizures through computerized EEG analysis," *Electroencephalography and Clinical Neurophysiology*, vol. 82, no. 1, pp. 30–37, 1992.
- [5] J. Gotman, "Automatic detection of seizures and spikes," *Journal of Clinical Neurophysiology*, vol. 16, no. 2, pp. 130–140, 1999.
- [6] A. M. Murro, D. W. King, J. R. Smith, B. B. Gallagher, H. F. Flanigin, and K. Meador, "Computerized seizure detection of complex partial seizures," *Electroencephalography and Clinical Neurophysiology*, vol. 79, no. 4, pp. 330–333, 1991.
- [7] A. J. Gabor, R. R. Leach, and F. U. Dowla, "Automated seizure detection using a self-organizing neural network," *Electroencephalography and Clinical Neurophysiology*, vol. 99, no. 3, pp. 257–266, 1996.
- [8] J. Gotman, D. Flanagan, J. Zhang, and B. Rosenblatt, "Automatic seizure detection in the newborn: methods and initial evaluation," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 356–362, 1997.
- [9] A. J. Gabor, "Seizure detection using a self-organizing neural network: validation and comparison with other detection strategies," *Electroencephalography and Clinical Neurophysiology*, vol. 107, no. 1, pp. 27–32, 1998.
- [10] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time–frequency analysis," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 703–710, 2009.
- [11] Z. Zhang, Y. Ren, N. Sabor et al., "DWT-net: seizure detection system with structured EEG montage and multiple feature extractor in convolution neural network," *Journal of Sensors*, vol. 2020, Article ID 3083910, 13 pages, 2020.
- [12] Y. Ech-Choudany, D. Scida, M. Assarar, J. Landré, B. Bellach, and F. Morain-Nicolier, "Dissimilarity-based time-frequency distributions as features for epileptic EEG signal classification," *Biomedical Signal Processing and Control*, vol. 64, article 102268, 2021.
- [13] H. Hassanpour, M. Mesbah, and B. Boashash, "Time-frequency based newborn EEG seizure detection using low and high frequency signatures," *Physiological Measurement*, vol. 25, no. 4, pp. 935–944, 2004.
- [14] B. R. Greene, S. Faul, W. P. Marnane, G. Lightbody, I. Korotchikova, and G. B. Boylan, "A comparison of quantitative EEG features for neonatal seizure detection," *Clinical Neurophysiology*, vol. 119, no. 6, pp. 1248–1261, 2008.
- [15] E. M. Thomas, A. Temko, G. Lightbody, W. P. Marnane, and G. B. Boylan, "Gaussian mixture models for classification of neonatal seizures using EEG," *Physiological Measurement*, vol. 31, no. 7, pp. 1047–1064, 2010.
- [16] R. Ahmed, A. Temko, W. P. Marnane, G. Boylan, and G. Lightbody, "Exploring temporal information in neonatal seizures using a dynamic time warping based SVM kernel," *Computers in Biology and Medicine*, vol. 82, pp. 100–110, 2017.
- [17] A. Temko, E. Thomas, W. Marnane, G. Lightbody, and G. Boylan, "EEG-based neonatal seizure detection with support vector machines," *Clinical Neurophysiology*, vol. 122, no. 3, pp. 464–473, 2011.
- [18] A. Zwanenburg, P. Andriessen, R. K. Jellema et al., "Using trend templates in a neonatal seizure algorithm improves detection of short seizures in a foetal ovine model," *Physiological Measurement*, vol. 36, no. 3, pp. 369–384, 2015.
- [19] L. Logesparan, A. J. Casson, and E. Rodriguez-Villegas, "Optimal features for online seizure detection," *Medical & Biological Engineering & Computing*, vol. 50, no. 7, pp. 659–669, 2012.
- [20] T. Hastie, R. Tibshirani, and J. Friedman, *An Introduction to Statistical Learning*, 2009.
- [21] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: a review," *Data classification: Algorithms and Applications.*, vol. 2014, pp. 37–64, 2014.
- [22] D. Yarotsky, "Universal approximations of invariant maps by neural networks," *Constructive Approximation*, vol. 55, no. 1, pp. 407–474, 2022.
- [23] A. Kratsios and C. Hyndman, "Neu: a meta-algorithm for universal uap-invariant feature representation," *Journal of Machine Learning Research*, vol. 22, p. 92, 2021.
- [24] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.
- [25] R. Eberhart and J. Kennedy, "Particle swarm optimization, proceedings of the IEEE international conference on," *Neural Networks*, vol. 4, pp. 1942–1948, 1995.

- [26] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory, MHS'95," in *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43, Nagoya, Japan, 1995.
- [27] C. Huang and J. Dun, "A distributed PSO-SVM hybrid system with feature selection and parameter optimization," *Applied Soft Computing*, vol. 8, no. 4, pp. 1381–1391, 2008.
- [28] T. Cuong-Le, T. Nghia-Nguyen, S. Khatir, P. Trong-Nguyen, S. Mirjalili, and K. D. Nguyen, "An efficient approach for damage identification based on improved machine learning using PSO-SVM," *Engineering with Computers*, vol. 38, pp. 3069–3084, 2022.
- [29] X. F. Song, Y. Zhang, D. W. Gong, and X. Z. Gao, "A fast hybrid feature selection based on correlation-guided clustering and particle swarm optimization for high-dimensional data," *IEEE Transactions on Cybernetics*, vol. 52, no. 9, pp. 9573–9586, 2022.
- [30] D. Chen, S. Wan, J. Xiang, and F. S. Bao, "A high-performance seizure detection algorithm based on discrete wavelet transform (DWT) and EEG," *PLoS One*, vol. 12, no. 3, pp. 1–21, 2017.
- [31] Y. Jiang, W. Chen, and M. Li, "Symplectic geometry decomposition-based features for automatic epileptic seizure detection," *Computers in Biology and Medicine*, vol. 116, article 103549, 2020.
- [32] M. Bandarabadi, C. A. Teixeira, J. Rasekhi, and A. Dourado, "Epileptic seizure prediction using relative spectral power features," *Clinical Neurophysiology*, vol. 126, no. 2, pp. 237–248, 2015.
- [33] V. Sridevi, M. R. Reddy, K. Srinivasan, K. Radhakrishnan, C. Rathore, and D. S. Nayak, "Improved patient-independent system for detection of electrical onset of seizures," *Journal of Clinical Neurophysiology*, vol. 36, no. 1, pp. 14–24, 2019.
- [34] M. B. Qureshi, M. Afzaal, M. S. Qureshi, and M. Fayaz, "Machine learning-based EEG signals classification model for epileptic seizure detection," *Multimedia Tools and Applications*, vol. 80, no. 12, pp. 17849–17877, 2021.
- [35] A. H. Ansari, P. J. Cheria, A. Caicedo, G. Naulaers, M. De Vos, and S. Van Huffel, "Neonatal seizure detection using deep convolutional neural networks," *International journal of neural systems*, vol. 29, no. 4, pp. 1–20, 2019.
- [36] G. Chandel, P. Upadhyaya, O. Farooq, and Y. U. Khan, "Detection of seizure event and its onset/offset using orthonormal _triadic_ wavelet based features," *IRBM*, vol. 40, no. 2, pp. 103–112, 2019.
- [37] M. K. Siddiqui, X. Huang, R. Morales-Menendez, N. Hussain, and K. Khatoun, "Machine learning based novel cost-sensitive seizure detection classifier for imbalanced EEG data sets," *International Journal on Interactive Design and Manufacturing*, vol. 14, no. 4, pp. 1491–1509, 2020.
- [38] K. M. Hassan, M. R. Islam, T. T. Nguyen, and M. K. Molla, "Epileptic seizure detection in EEG using mutual information-based best individual feature selection," *Expert Systems with Applications*, vol. 193, pp. 1–11, 2022.
- [39] P. T. PhysioBank, "PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [40] *EEG Time Series Data (Children's Hospital Boston and the Massachusetts Institute of Technology)* website, <https://physionet.org/physiobank/database/chbmit/>. Accessed Sept 2017.
- [41] J. Prasanna, M. S. P. Subathra, M. A. Mohammed, R. Damaševičius, N. J. Sairamya, and S. T. George, "Automated epileptic seizure detection in pediatric subjects of CHB-MIT EEG database—a survey," *Journal of Personalized Medicine*, vol. 11, no. 10, p. 1028, 2021.
- [42] Z. A. Alyasseri, A. T. Khader, M. A. Al-Betar et al., "Multi-objective flower pollination algorithm: a new technique for EEG signal denoising," *Neural Computing and Applications*, vol. 2022, pp. 1–20, 2022.
- [43] L. S. Vidyaratne and K. M. Iftekharuddin, "Real-time epileptic seizure detection using EEG," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 2146–2156, 2017.
- [44] Z. A. Alyasseri, O. A. Alomari, J. P. Papa et al., "EEG channel selection based user identification via improved flower pollination algorithm," *Sensors*, vol. 22, no. 6, pp. 1–16, 2022.
- [45] Z. Abdi Alkareem Alyasseri, O. A. Alomari, M. A. Al-Betar et al., "EEG channel selection using multiobjective cuckoo search for person identification as protection system in healthcare applications," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5974634, 18 pages, 2022.
- [46] N. F. Chang, T. C. Chen, C. Y. Chiang, and L. G. Chen, "Channel selection for epilepsy seizure prediction method based on machine learning," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5162–5165, California USA, 2012.
- [47] A. Bhattacharyya and R. B. Pachori, "A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2003–2015, 2017.
- [48] Y. Li, Y. Liu, W. G. Cui, Y. Z. Guo, H. Huang, and Z. Y. Hu, "Epileptic seizure detection in EEG signals using a unified temporal-spectral squeeze-and-excitation network," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 4, pp. 782–794, 2020.
- [49] H. Peng, C. Lei, S. Zheng et al., "Automatic epileptic seizure detection via stein kernel-based sparse representation," *Computers in Biology and Medicine*, vol. 132, article 104338, 2021.
- [50] M. S. Roulston, "Estimating the errors on measured entropy and mutual information," *Physica D: Nonlinear Phenomena*, vol. 125, no. 3–4, pp. 285–294, 1999.
- [51] Q. U. Mastoi, T. Y. Wah, M. A. Mohammed et al., "Novel DERMA fusion technique for ECG heartbeat classification," *Life*, vol. 12, no. 6, pp. 1–13, 2022.
- [52] M. D'Alessandro, R. Esteller, G. Vachtsevanos, A. Hinson, J. Echaz, and B. Litt, "Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 5, pp. 603–615, 2003.
- [53] N. Päivinen, S. Lammi, A. Pitkänen, J. Nissinen, M. Penttonen, and T. Grönfors, "Epileptic seizure detection: a nonlinear viewpoint," *Computer Methods And Programs In Biomedicine*, vol. 79, no. 2, pp. 151–159, 2005.
- [54] D. P. Burke, S. P. Kelly, P. De Chazal, R. B. Reilly, and C. Finucane, "A parametric feature extraction and classification strategy for brain-computer interfacing," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 1, pp. 12–17, 2005.

- [55] R. Alcaraz and J. J. Rieta, "A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms," *Biomedical Signal Processing and Control*, vol. 5, no. 1, pp. 1–14, 2010.
- [56] S. P. Kumar, N. Sriraam, and P. G. Benakop, *Automated Detection of Epileptic Seizures Using Wavelet Entropy Feature with Recurrent Neural Network Classifier*, TENCON 2008-2008 IEEE Region 10 Conference, Hyderabad, India, 2008.
- [57] S. Kiranyaz, T. Ince, M. Zabihi, and D. Ince, "Automated patient-specific classification of long-term electroencephalography," *Journal of Biomedical Informatics*, vol. 49, pp. 16–31, 2014.
- [58] 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 Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 3, pp. 386–398, 2016.
- [59] X. Tian, Z. Deng, W. Ying et al., "Deep multi-view feature learning for EEG-based epileptic seizure detection," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 10, pp. 1962–1972, 2019.