

## Research Article

# A Pervasive Approach to EEG-Based Depression Detection

**Hanshu Cai,<sup>1</sup> Jiashuo Han,<sup>1</sup> Yunfei Chen,<sup>1</sup> Xiaocong Sha,<sup>1</sup> Ziyang Wang,<sup>1</sup> Bin Hu<sup>1,2,3</sup> ,<sup>1,2,3</sup> Jing Yang,<sup>4</sup> Lei Feng,<sup>5</sup> Zhijie Ding,<sup>6</sup> Yiqiang Chen,<sup>7</sup> and Jürg Gutknecht<sup>8</sup>**

<sup>1</sup>Gansu Provincial Key Laboratory of Wearable Computing, School of Information Science and Engineering, Lanzhou University, Lanzhou, China

<sup>2</sup>CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

<sup>3</sup>Beijing Institute for Brain Disorders, Capital Medical University, Beijing, China

<sup>4</sup>Department of Child Psychology, Lanzhou University Second Hospital, Lanzhou, China

<sup>5</sup>Beijing Anding Hospital, Capital Medical University, Beijing, China

<sup>6</sup>The Third People's Hospital of Tianshui City, Tianshui, China

<sup>7</sup>Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

<sup>8</sup>Computer Systems Institute, ETH Zürich, Zürich, Switzerland

Correspondence should be addressed to Bin Hu; [bh@lzu.edu.cn](mailto:bh@lzu.edu.cn)

Received 31 March 2017; Revised 17 November 2017; Accepted 4 January 2018; Published 6 February 2018

Academic Editor: Haiying Wang

Copyright © 2018 Hanshu Cai 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.

Nowadays, depression is the world's major health concern and economic burden worldwide. However, due to the limitations of current methods for depression diagnosis, a pervasive and objective approach is essential. In the present study, a psychophysiological database, containing 213 (92 depressed patients and 121 normal controls) subjects, was constructed. The electroencephalogram (EEG) signals of all participants under resting state and sound stimulation were collected using a pervasive prefrontal-lobe three-electrode EEG system at Fp1, Fp2, and Fpz electrode sites. After denoising using the Finite Impulse Response filter combining the Kalman derivation formula, Discrete Wavelet Transformation, and an Adaptive Predictor Filter, a total of 270 linear and nonlinear features were extracted. Then, the minimal-redundancy-maximal-relevance feature selection technique reduced the dimensionality of the feature space. Four classification methods (Support Vector Machine, K-Nearest Neighbor, Classification Trees, and Artificial Neural Network) distinguished the depressed participants from normal controls. The classifiers' performances were evaluated using 10-fold cross-validation. The results showed that K-Nearest Neighbor (KNN) had the highest accuracy of 79.27%. The result also suggested that the absolute power of the theta wave might be a valid characteristic for discriminating depression. This study proves the feasibility of a pervasive three-electrode EEG acquisition system for depression diagnosis.

## 1. Introduction

Depression is a common mood disorder, which might cause persistent feeling of sadness, loss of interest, and impairment of memory and concentration. Depressed patients normally experience cognitive impairment and suffer long and severe emotional depression. In severe cases, some patients will experience paranoia and illusion [1]. According to the World Health Organization statistics, >300 million individuals suffer from depression worldwide; approximately 800,000

people die due to it every year [2]. Thus, depression is predicted to become the second most common disease after heart disease by the year 2020 [3]. Hence, the diagnosis of depression in the early curable stages is critical and might save the life of a patient [4].

Presently, the study on the human cerebral is currently under intensive focus in order to understand the mechanism underlying persistent negative emotion and depression. Therefore, the most commonly used diagnosis of depression is a scale-based interview conducted by a psychologist or

psychiatrist. The current international standard mostly used is “In Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)” (DSM-IV) [5], and the clinical test, Mini-Mental State Examination (MMSE), is commonly applied [6]. Other conventional psychometric questionnaires, such as Beck depression inventory (BDI) [7] and Hamilton Depression Rating Scale (HDRS) [8], are also used as screening tools rather than as the instrument for the diagnosis of depression.

The current methods of depression detection are human-intensive, and the results are dependent on the doctor's experience. Furthermore, depressed individuals are less likely to seek help due to fear of stigma and the nature of the disorder. As a result, a large number of depressed patients, not diagnosed accurately, do not receive optimal treatment and adequate recovery period. Therefore, finding convenient and effective methods for the detection of depression is an emerging topic for research. With the latest advances in the sensor and mobile technology, the exploration using physiological data for the diagnosis of mental disorder opens a new avenue for an objective and accurate tool for depression detection. Among all kinds of physiological data, electroencephalogram (EEG) reflects emotional human brain activity in real time [9].

The EEG signal is a recording of the spontaneous, rhythmic, electrical activity of brain neurons from the scalp surface. Since the earliest discovery from the rabbit and monkey brain and the first recording of the human EEG signal by German psychiatrist Hans Berger in 1926, studies on the analytical method of EEG and the interpretation of the association between the brain function and mental disorders have been continued for over a century [10]. Neuroscience, psychology, and cognitive science research showed that a majority of the psychological activities and cognitive behavior could be indicated by EEG [11–13]. The EEG signal is closely related to the brain activities and emotional states, and it could reflect the emotional transformation in real time. Cole and Ray [14] found that the EEG signal collected from the parietal lobe of brain is associated with the cognitive tasks and emotional states. Klimesch et al. found that the alpha waves with low frequency could reflect some of the features of attention, such as vigilance and expectations [15]. Srinivasan et al. demonstrated that the frequency domain features of EEG could be used to predict the level of attention [16]. Therefore, the EEG signal is critical for understanding the processing of human brain information and emotional state transformation.

The studies on EEG could be used to understand the mechanism underlying brain activity, human cognitive process, and diagnosis of brain disease, as well as the field of the Brain Computer Interface (BCI), which has attracted much attention in recent years [17]. Compared to Computed Tomography (CT) and functional Magnetic Resonance Imaging (fMRI), EEG has a higher time resolution, a lower maintenance cost, and a simpler operation method. Thus, as an objective physiological method to obtain data, EEG was proposed as a nonintrusive approach to study cognitive behavior [18–20] and other illness symptoms, such as insomnia [21–23], epilepsy [24–26], and sleep disorder [27]. EEG has also been used in the diagnosis of mental disorders,

such as anxiety [28–30], psychosis [31–34], and depression [35–38]. In addition, depression as a mental disorder with clinical manifestations such as significant depression and slow thinking is always accompanied by abnormal brain activity and obvious emotional alternation. Therefore, as a method tracking the brain functions, EEG can detect these abnormal activities.

The frequency of the EEG signal can be divided into 5 wave-bands: delta wave (<4 Hz), which normally appears in an adult's slow-wave sleep; theta wave (4–8 Hz), which is usually found when someone is sleepy; alpha wave (8–14 Hz), which is normally detected when someone is relaxed; beta wave (14–30 Hz), which commonly appears when someone is actively thinking; and gamma wave (30–50 Hz), which could appear during meditation. The EEG signals undergo changes in the amplitude as well as frequency, while different mental tasks are performed [39–42].

Presently, for research purposes, the most commonly used are 128-electrode and 256-electrode EEG systems [43, 44], which are specifically designed for research purposes. The operation of the instruments was not only difficult to initiate but also it required technicians to apply conductive gel to each electrode on the participant's head before each use. The preparation process alone takes 30 minutes on average. In addition, these EEG systems are expensive. Overall, these systems are not practical for pervasive depression detection.

In the present study, the pervasive three-electrode EEG acquisition system, developed independently by the Ubiquitous Awareness and Intelligent Solutions Lab (UAIS) of Lanzhou University [45], was employed to construct a database containing both depressed patients and normal controls. Thus, the use of the latest data processing technique and machine learning to explore a pervasive EEG-based depression detection system has been the focus of investigation. In order to support this research:

- (1) A pervasive three-electrode EEG acquisition system has been introduced (Section 2.1).
- (2) A psychophysiological experiment has been conducted, in which EEG of 213 participants has been recorded. These physiological data provided a comprehensive database for further analysis, construction, and evaluation of a pervasive EEG-based depression detection system (Sections 2.2 and 2.3).
- (3) Several EEG preprocessing steps and methods were applied on the raw EEG data (Section 3.1).
- (4) 270 features were identified and extracted from the recorded database. By employing a feature selection technique, an optimum feature matrix was constructed for the depression classification process (Section 3.2).
- (5) Four classification algorithms, including *K*-Nearest Neighbor (KNN), Support Vector Machine (SVM), Classification Tree (CT), and Artificial Neural Network (ANN), have been evaluated and compared, using a 10-fold cross-validation (Section 4).

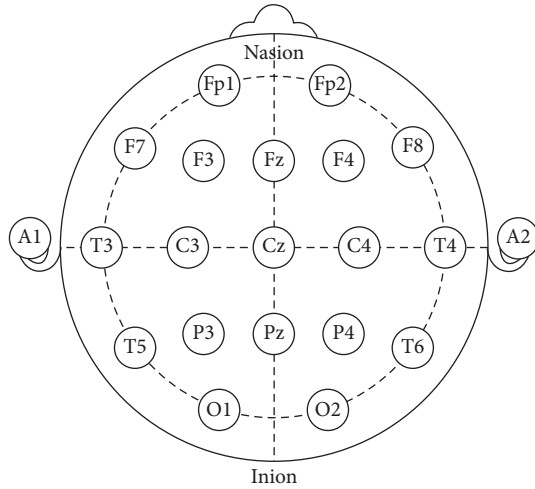


FIGURE 1: The international 10-20 system.

## 2. Pervasive Three-Electrode EEG Database Construction

**2.1. Pervasive Three-Electrode EEG Acquisition System.** The 10-20 system, proposed by Jasper in 1958, defined the name of the electrode and later became the international standard EEG placement system [46]. With the development of sensor technology, the electrode became smaller than that in previous systems and the electrodes recorded a detailed EEG. In 1985, Chatrian et al. added extra electrodes in intermediate sites halfway between those of the existing 10-20 system, thereby expanding it to a 64-electrode system [47]. Due to the complexity of the full-brain 128-electrode and 256-electrode systems, the investigators restricted themselves from mobile and pervasive application. Thus, with the development of universal and pervasive electronic technology, the 8-electrode and 16-electrode systems with small volume were also developed gradually.

As shown in Figure 1, F represents the frontal lobe, T represents the temporal lobe, C represents the center, P represents the parietal lobe, and O represents the occipital lobe. EEG reacts to the biological activity of the brain tissue, thereby indicating the functional status of the brain [48]. The EEG signal collected from the different locations of the scalp reflects a variety of information. For example, EEG from the frontal lobe reflects human memory, computational power, attention, and responsiveness; EEG from the parietal lobe is associated with somatic responses; EEG from the occipital lobe can be used as a reference for visual reactions; EEG from temporal lobe is related to auditory reactions. Therefore, for different research direction and purpose, the appropriate EEG collection location is essential.

Prefrontal cortex is the center of consciousness; thus, the better the control of the forehead cortex, the better the emotional control. Jasper studied the resting-state EEG of severe depression patients showing that when the body suffered from severe depression, the activity of the cerebral cortex was altered [49]. Nauta emphasized that the prefrontal cortex played a major role in different aspects of emotional processes

[50]. Rolls put forward the importance of prefrontal cortex for emotional and motivational processes [51]. Harmon-Jones suggested that the specific forms of anger, or anger elicited in particular contexts, are associated with left-sided prefrontal activation [52]. In conclusion, the above studies have shown that the electrode sites located in the prefrontal cortex are associated with emotional process and psychiatric disorders. Therefore, Fp1, Fp2, and Fpz are the ideal choices of scalp position in the current experiment. The hair in the frontal lobe is absent and contact dry electrode should be sufficient without the need for applying conductive gel. The pervasive three-electrode EEG acquisition system (Figure 2), developed by UAIS from Lanzhou University [53], runs on rechargeable battery and transmits all the EEG data through Bluetooth 2.0 wirelessly. The system is extremely small in size and can be easily placed on the location. The sampling frequency is 250 Hz and according to the EGI engineers, all electrodes have an impedance of  $<50\text{ k}\Omega$ . Since the frequency of EEG is 0.5–50 Hz, the passband of the EEG acquisition is 0.5–50 Hz.

**2.2. Experiment Method.** Compared to the normal controls, depressed patients responded differently to outside stimulus [54, 55]. The feedback of the depressed patients to the positive and negative stimuli weakened. As the positive stimulus feedback weakened further, the overall performance was negative emotions and reflected as such in the emotional response of the different subsystems. In summary, no significant difference was observed in the positive stimulus between normal controls and depressed patients, and depressed patients would produce more negative emotions under negative stimulus as compared to normal controls. Beck's cognitive behavioral model of depression postulated that the depressed patients are likely to support a negative view of themselves, the world, and even the future. In order to maintain this negative self-view, they even resist the environmental feedback that is inconsistent with the view [56]. Epstein et al. suggested that, in comparison to normal controls, depressed patients responded with less bilateral ventral striatal activation to positive stimuli, which leads to the decreased interest in performance of activities [57]. Bylsma et al. proved that depressed patients exhibit less reactivity to all stimuli and events, irrespective of positive or negative nature [58].

Therefore, recording and analysis of the EEG signal in different stimuli may help in the identification of patients with depression. This study was designed to record the participants' EEG in four different cases: in resting state, under negative stimulus, under neutral stimulus, and under positive stimulus. The source of stimulus is soundtracks from the International Affective Digitized Sounds (IADS-2) [59], which is a standardized database of 167 naturally occurring sounds, widely used in the study of emotions.

The experiment was performed in a quiet room. Firstly, the experiment objective and procedures were described to the participants. Then, the pervasive three-electrode EEG acquisition system was placed on the participants' foreheads and checked for reception. After one minute of relaxation, the experiment begins again. At first stage, 90 s of resting-state EEG was recorded. Then, the participants were asked to

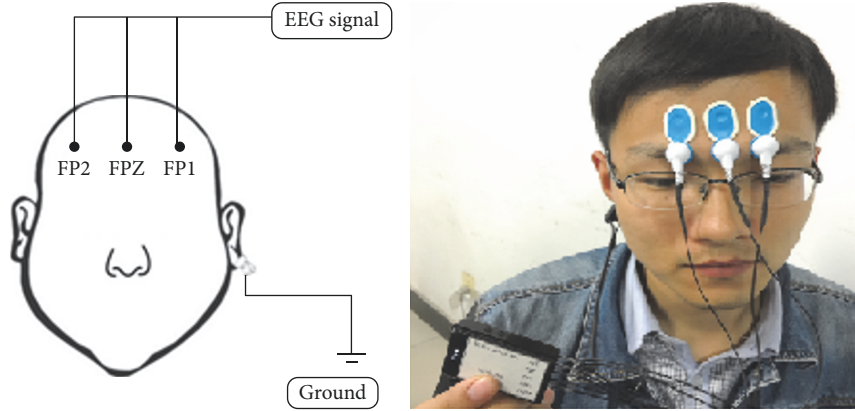


FIGURE 2: Pervasive three-electrode EEG acquisition system using Fp1, Fp2, and Fpz positions.

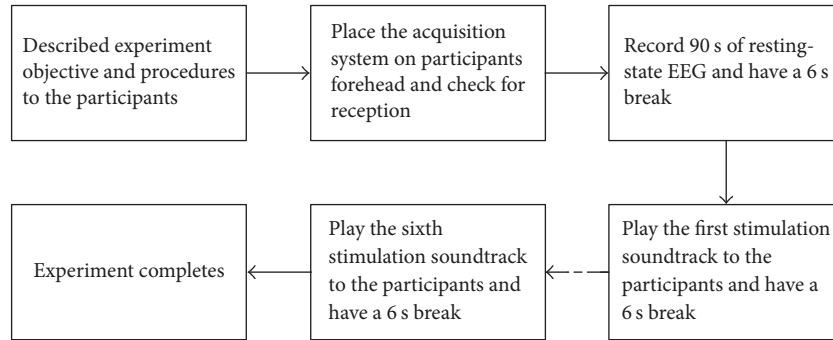


FIGURE 3: Process of EEG acquisition.

TABLE 1: Audio stimulation profile.

Number	Name	Property
(1)	Cattle	Neutral
(2)	Painting	Neutral
(3)	Babies cry	Negative
(4)	Dentist drill	Negative
(5)	Baby	Positive
(6)	Crowd	Positive

remain seated with eyes closed with as little body movements as possible, followed by another minute of rest. In the second stage, stimulation soundtracks will be played to participants. Each soundtrack was 6 s long with a 6 s break between each soundtrack. The process would continue until the experiment is completed. The process of EEG acquisition is shown in Figure 3.

A total of 6 stimulation soundtracks (according to IADS-2) existed, including 2 neutral stimulation soundtracks, 2 negative stimulation soundtracks, and 2 positive stimulation soundtracks. Table 1 describes each audio stimulation.

**2.3. Psychophysiological Database.** Of the total 250 participants, 213 (92 depressed patients and 121 normal controls) completed the experiment, successfully. The raw EEG data

from all the electrodes were recorded. Depressed participants were selected by professional psychiatrists using MMSE [6], which is a 30-point questionnaire used by the psychiatrist during a face-to-face interview to assess the degree of cognitive dysfunction in patients with diffuse brain disorders. In addition, all participants are asked to fill the following scales for cross-referencing:

- (A) The Patient Health Questionnaire (PHQ-9) [60] is a 9-question-based multipurpose instrument for screening, diagnosing, monitoring, and measuring the severity of depression. We chose this questionnaire in order to find the relevance between the EEG characteristic and the severity of depression.
- (B) Life Event Scale (LES) [61] contains 48 questions including events of family, work, and social support. The influence of each event is evaluated for severity, duration, and frequency. We chose this questionnaire for cross-referencing purposes.
- (C) Pittsburgh Sleep Quality Index (PSQI) [62] contains 19 self-reported items, creating 7 components to diagnose sleep disorders. We chose this index to explore the direct link between sleep qualities with depression in EEG.
- (D) Generalized Anxiety Disorder Scale-7 (GAD-7) [63] contains only 7 self-report questions for screening



and measuring the severity of generalized anxiety disorder. We chose this questionnaire for cross-referencing between depression and anxiety.

### 3. Data Processing

In this study, all preprocessing, and data analyses have been implemented using MATLAB software (version R2014a).

**3.1. Preprocessing.** EEG is a noninvasive method of capturing the physiological signal of brainwave activity. However, EEG data recorded are normally mixed with interferences from surrounding environment, such as close-by power line. Furthermore, other physiological signals, including electrocardiogram (ECG), electrooculogram (EOG), and electromyograph (EMG), could also be detected and recorded by EEG sensors [55]. To ensure an accurate result in the feature selection and classification, all the raw data should be denoised first.

ECG is a smooth signal among the physiological electrical signals, with a large amplitude. As the heart is located distally from the head, the ECG signal will be greatly attenuated when spread to the scalp. EMG is produced by muscle contraction, with an amplitude of 10  $\mu$ V to 15 mV. The frequency of EMG is concentrated primarily in the high band > 100 Hz. Power-line interference focuses on fixed operation frequency. In order to remove these interference signals, we followed the results of several investigators. Yang proposed a cascade of three adaptive filters based on the least mean squares (LMS) algorithm and verified that the proposed filter reduced the interference in EEG signals [64]. Tong et al. validated the use of independent component analysis (ICA) for an efficient suppression of the interference of ECG from EEG [65]. The National Institute of Mental Health announced that using an adaptive filter to estimate the contaminants can subtract them from the EEG data [66].

No overlap occurred between the frequency of EEG signal and power-line interferences, EMG and ECG; thus, Finite Impulse Response (FIR) filter based on the Blackman time window was used to remove these interference signals. The adequate linearity of the FIR filter is widely used in modern electronic communication. It can guarantee any amplitude frequency characteristics simultaneously, with strict linear phase-frequency characteristics. In addition, the unit sampling response is finite, which stabilized the filter. In order to reduce the energy leakage of the spectrum, the signal can be truncated by different interception functions. This truncation function is known as the window function. The time domain representation of the Blackman time window is

$$w(n) = \left[ 0.42 - 0.5 \cos\left(\frac{2m}{N-1}\right) + 0.08 \cos\left(\frac{4m}{N-1}\right) \right] \cdot R_N(n), \quad (1)$$

where  $R_N(n)$  is the rectangular window function and  $N$  is the length of truncated data.

The resulting EEG signal is retained only between frequencies in the range of 0.5–50 Hz. However, the frequency of EOG overlaps within this range. Although all participants

were asked to remain seated with eyes closed, their EOG was recorded inevitably while using the prefrontal-lobe EEG sites, such as Fp1, Fp2, and Fpz. A general model for EOG contamination can be described by

$$y(n) = x(n) + F(r), \quad (2)$$

where  $y(n)$  and  $x(n)$  are the samples of the recorded (including noise) and true EEG, respectively,  $r$  represents the source EOG, and  $F$  is an unknown transfer function.

Kalman filter is an optimal recursive data processing algorithm, which has been widely utilized in several applications, such as industrial control systems, radar target tracking, communications and signal processing, aeroengine diagnosis, and intelligent robots. Kalman filter is based on the previous estimated value and the observed value of the current time to estimate the current value of the stated variable. Thus, the frequency of the EOG artifact would not exceed 15 Hz, and the approximate EOG signal and the amplitude of the brain in the low frequency band are small. As a result, the Kalman derivation formula combines the Discrete Wavelet Transformation (DWT) and an Adaptive Predictor Filter (APF) to estimate the pure EOG artifact.

The denoising model proposed in the present study involves the following steps: (1) signal decomposition, (2) ocular artifacts (OA) zones detection, (3) signal prediction, and (4) signal reconstruction. Herein, DWT was used to decompose the EEG signals and detect the OZ zones. The frequency range of the EEG signal was 0–64 Hz, while the OA occurred in 0–16 Hz. The multiscale DWT decomposition was used to extract the low frequency components and non-stationary time series, which were then divided into several approximate stationary time series. Thus, the conventional forecasting methods, such as Kalman filter, can predict the shape of the true wave of decomposition signals accurately. Subsequently, the Adaptive Auto Regressive (AAR) models and an Adaptive Predictor Filter (APF) were applied to improve the prediction. The APF uses an adaptive filter to estimate the future values of signals based on their past values. Finally, the EOG artifacts were removed from the raw EEG signal, and the data were ready for further processing.

**3.2. Features Matrix Construction.** The features matrix consists of  $n$  rows and  $m$  columns, where  $n$  represents the number of EEG data and  $m$  represents the number of features extracted from each EEG. The present study constructed the training effective features matrix using three steps as follows:

- (1) Identify and extract all the efficient features for each set of EEG data, such that each row represents a feature vector.
- (2) Each row of the features matrix is selected by feature selection; that is, the most suitable feature is selected from all the extracted features to form a final feature vector.
- (3) Each row of the feature vectors is tagged by depression or nondepression.

**3.2.1. Feature Extraction.** The EEG signal presents weak, nonlinear, and time-sensitive characteristic, which exhibits typically complex dynamics. The feature of EEG will change with the emotional state transformation. The analysis of

EEG data displayed different linear features such as peak, variance, and skewness that were used in recent literature [67–70]. Efforts have been made in determining nonlinear parameters such as Correlation Dimension for pathological signals, which are shown as useful indicators of pathologies [71]. In order to obtain the feature matrix, we must first perform the feature extraction of the pretreated EEG. The EEG features are mainly divided into Time Domain Features and Frequency Domain Features. Owing to the nonlinearity and randomness of the EEG signal, this study extracts the nonlinear features such as the Correlation Dimension and Shannon Entropy in addition to the above EEG features. Finally, the following features were selected for extraction:

(1) *Time Domain Features.* Time domain constitutes the most intuitive EEG features. The EEG signals are collected at a certain time and frequency. The artifacts are directly removed from the time domain EEG signal, and useful information was extracted as a time domain feature that can be used for continuous prolonged EEG detection. The time domain features extracted in this study include peak, variance, skewness, kurtosis, and Hjorth parameter. Hjorth parameters are indicators of statistical properties used in signal processing in the time domain introduced by Hjorth in 1970 [72]; the parameters include activity, mobility, and complexity. Among them, the activity parameters represent the signal power and the variance of time function. The mobility parameters represent the mean frequency or the proportion of standard deviation of the power spectrum. The complexity parameters represent the change in frequency. These parameters are usually used to analyze the EEG signals for feature extraction.

(2) *Frequency Domain Features.* Frequency domain is a tool for characterizing and classifying the EEG signals. Herein, the frequency domain features are relative centroid frequency, absolute centroid frequency relative power, and absolute power.

(3) *Nonlinear Features.* The EEG signals are nonstationary and random; they also include some of the characteristics of the nonlinear dynamics system. With increasing number of studies on the EEG signals, the nonlinearity has been under intensive focus worldwide. Therefore, processing and analyzing the EEG signal based on the nonlinear dynamics theory become a new research direction. The nonlinear features extracted in this study include C0-complexity, Kolmogorov Entropy, Shannon Entropy, Correlation Dimension, and Power-Spectral Entropy.

(A) The C0-complexity was proposed by Shen et al. [73] to resolve the issue of over-coarse graining preprocessing in Lempel-Ziv complexity (LZC) [74]. The core of the algorithm is to decompose the sequence into regular and irregular components, and the C0-complexity defines the proportion of irregularities in the sequence. The greater the proportion of its share, the closer the time domain signal to the random sequence, and, thus, the greater the complexity. The

doctrine presumes that a signal can be divided into regular part and stochastic components. If  $A_0$  is a measurement of the signal and  $A_1$  is the measurement corresponding to the stochastic part, C0-complexity is defined as the ratio of  $A_1$  and  $A_0$ . Supposedly, the EEG signal to be analyzed is  $\{x(n), n = 0, 1, \dots, N-1\}$  with a length of  $N$  samples; then the C0-complexity can be calculated with the power spectra as follows:

The fast Fourier transform (FFT) of the signal is as follows:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j(2k\pi n/N)} \quad k = 0, 1, \dots, N-1. \quad (3)$$

The mean amplitude of the power spectrum  $X(k)$  is as follows:

$$M = \frac{1}{N} \sum_{k=0}^{N-1} |X(k)|^2. \quad (4)$$

$X(k)$  less than  $M$  are replaced by 0 to obtain a new spectrum series  $Y(k)$ :

$$Y(k) = \begin{cases} X(k) & |X(k)|^2 > M \\ 0 & |X(k)|^2 \leq M. \end{cases} \quad (5)$$

The inverse FFT (IFFT) of  $Y(k)$  is as follows:

$$y(n) = \sum_{k=0}^{N-1} Y(k) e^{j(2k\pi n/N)} \quad n = 0, 1, \dots, N-1. \quad (6)$$

The power of stochastic part  $A_1$  is extracted, and the C0-complexity was estimated:

$$\begin{aligned} A_1 &= \sum_{n=0}^{N-1} |x(n) - y(n)|^2 \\ A_0 &= \sum_{n=0}^{N-1} |x(n)|^2. \end{aligned} \quad (7)$$

$$C0 = \frac{A_1}{A_0}$$

(B) Kolmogorov Entropy was used to measure the rate of loss of information per unit of time. Positive and finite entropy represents that the time series and the dynamic underlying phenomenon are chaotic. Zero entropy indicates a regular phenomenon in the space phase. Infinite entropy refers to a stochastic and non-deterministic phenomenon. Kolmogorov Entropy is defined as the average rate of loss of information as follows:

$$KE = -\lim_{\tau \rightarrow 0} \lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n\tau} \sum_{i_0 \dots i_{n-1}} P_{i_0 \dots i_{n-1}} \ln P_{i_0 \dots i_{n-1}}. \quad (8)$$

- (C) Shannon Entropy was introduced by Shannon in 1948 in an article entitled “A Mathematical Theory of Communication” [75]. The size of the information of a message is directly related to its uncertainty. The amount of information is equal to the amount of uncertainty. Shannon Entropy is a measure of uncertainty of a random variable and a random signal. The larger the entropy, the greater the uncertainty and randomness. In the present study, the entropy used to process EEG can be viewed as a measure of the order in the signal, which measures the skewness and uncertainty [76]. In the case of random variables with known probability distribution, the entropy is defined by

$$H(X) = - \sum_{x \in \chi} p(x) \log p(x), \quad (9)$$

where  $X$  is a random variable with probability distribution  $p(x)$  and alphabet set  $\chi$  [77].

- (D) Correlation Dimension indicates the dynamic features of the EEG signal. The greater the Correlation Dimension number, the complicated the EEG time series. The Correlation Dimension is a fractal dimension, often computed from the time series illustration. It is a simplified phase space diagram constructed from a single data vector. The fundamental Correlation Dimension algorithm was introduced by Grassberger and Procaccia in 1983 [5] and can be expressed as below:

$$CD = \lim_{r \rightarrow 0} \left( \frac{\ln C(r)}{\ln r} \right), \quad (10)$$

where  $C(r)$  is the correlation integral and  $r$  is the radial distance around each reference point.

- (E) Power-Spectral Entropy is a sequence of power density with the frequency distribution obtained by Fourier transform. The calculated entropy of the power spectrum (referred to as Power-Spectral Entropy) can be implemented easily. The Power-Spectral Entropy is used to analyze the timing signals in EEG data. The entropy can be used as a physical indicator to estimate the quality and intensity of brain activity. The larger the entropy, the more active the brain.

All linear and nonlinear features (Table 2) were extracted from alpha wave, beta wave, delta wave, theta wave, gamma wave, and full-band EEG of each electrode (Fp1, Fp2, and Fpz). Therefore, a total of 270 features (15 basic features  $\times$  6 frequencies  $\times$  3 electrodes) were extracted. All the involved linear and nonlinear features are common information about EEG.

**3.2.2. Feature Selection.** Feature Selection is used to select a relevant subset of all available features, which not only yields a small dimensionality of the classification problem but also

TABLE 2: Features used in the feature extraction process.

Name	Property
Centroid frequency	Linear features
Relative centroid frequency	
Absolute centroid frequency	
Relative power	
Absolute power	
Peak	
Variance	
Skewness	
Kurtosis	
Hjorth	
Power-Spectrum Entropy	Nonlinear features
Shannon Entropy	
Correlation Dimension	
C0-complexity	
Kolmogorov Entropy	

reduces the noise (irrelevant features). We further deduced the types of features suitable for suppressing the EEG signal recognition by inspecting the features selected by the applied algorithm.

The feature evaluation function focuses on the relation between the features and the target class, which tends to involve redundant features, influencing the learning accuracy and results. In order to achieve these results, we applied the minimal-redundancy-maximal-relevance (MRMR) technique to perform the feature selection. The MRMR feature selection criterion was proposed by Peng et al. [78] in order to resolve the issue by evaluating both feature redundancy and relevance simultaneously; in particular, max-relevance, denoted as  $\max D(S, c)$ , refers to maximizing the relevance of a feature subset  $S$  to the class label  $c$ . In [1], the relevance of a feature subset is defined as

$$\max D(S, c) = \frac{1}{|S|} \sum_{f_i \in S} \Phi(f_i, c), \quad (11)$$

where  $\Phi(f_i, c)$  denotes the relevance of a feature  $f_i$  to  $c$ .  $\Phi$  could be estimated using any correlation measures.

Feature redundancy is defined based on the pairwise feature dependence. If two relevant features highly depend on each other, the class-discrimination power would not change dramatically if one of the features was removed. Min-redundancy,  $\min R(S)$ , is used to select a feature subset of mutually exclusively features. The redundancy of a feature subset is defined as

$$\min R(S) = \frac{1}{|S|^2} \sum_{f_i, f_j \in S} \Phi(f_i, f_j). \quad (12)$$

MRMR is defined as the simple operator maximizing  $D$  and minimizing  $R$  consecutively. In [1], the incremental search method was used to find the near-optimal features. The

feature subset  $S_{m-1}$  of  $m-1$  selected feature is utilized to select the  $m$ -order feature that optimizes the following criterion:

$$\max_{f_j \notin S_{m-1}} \left[ \Phi(f_i, c) - \frac{1}{m-1} \sum_{f_i \in S_{m-1}} \Phi(f_i, f_j) \right]. \quad (13)$$

**3.2.3. Effective Tagging.** Each feature vector (each row of the feature matrix) has to be marked with a specific emotional tag. In this study, we divided the experimental population into two categories: depressed patients and normal controls. All eigenvectors are tagged as depressed and nondepressed.

## 4. Classification

SVM, KNN, and CT are the widely used classification algorithms in the majority of the EEG-related studies. In the present study, we evaluated the performance of these classifiers (SVM, KNN, and CT) plus the Artificial Neural Network (ANN) classifier in the depression detection process. All classifications and 10-fold cross-validations have been implemented using the MATLAB software (version R2014a).

### 4.1. Classification Techniques

**4.1.1. SVM.** SVM, proposed by Cortes and Vapnik [79] in 1995, is a supervised learning model and regression method. It exhibits several unique advantages in resolving the issue of small sample data, nonlinear data, and high-dimensional pattern recognition [80]. SVM builds a hyperplane or an infinite-dimensional space for classification and regression. The kernel function allows SVM to deal with the nonlinear classification problem by attempting to cluster a feature space based on the known labels, with maximum possible distance between the clusters' borders [79]. In addition, SVM has been widely used in many fields such as text classification [81], image classification [82], biological sequence analysis, biological data mining [83], and handwriting character recognition [84]. In recent years, SVM has also been applied in the field of depression discrimination [85–87]. In the present study, Gaussian Kernel functions have been implemented and evaluated in SVM classification.

**4.1.2. KNN.** KNN algorithm is a nonparametric supervised machine learning method for classification and regression. It was introduced by Dasarathy [88] in 1991 based on instant or lazy learnings. The classifier based on KNN does not require a training phase, and its computational complexity is proportional to the number of documents in the training set. Taken together, if the number of documents in the training set is  $N$ , then the time complexity of the KNN classifier is  $O(n)$ . KNN categorizes the feature spaces into binary or multiclass clusters by employing a training dataset to further classify the data points according to the closest data points to  $K$  in the training dataset. KNN has been used in medical informatics, such as the detection of epilepsy [89], stress [90], and depression [85, 91].

**4.1.3. CT.** CT, also known as decision tree, is a tree structure-based supervised classification model [92], defined by separating and partitioning a feature space, using multiple rules and defining a local model, into which the feature spaces can be categorized as binary or multiclass clusters. Each of the internal nodes represents a property, each edge represents a result, and each leaf represents a class label. Compared to the other classification algorithms, the decision tree is the fastest classification. CT has been used in classifying Alzheimer's disease [93] as well as depression [94].

**4.1.4. ANN.** ANN is a classification method that mimics the structure and function of the biological neural network and consists of an information processing network with wide parallel interconnection of simple units. This network exhibits learning and memory ability, knowledge generalization, and input information feature extraction ability similar to that of the human brain [95]. Neural networks have been used to resolve a variety of difficult tasks using common rule-based programming, such as computer vision [91], speech recognition [96], and mental disorders [97, 98]. ANN is the only unsupervised machine learning classifier used in the present study.

**4.2. Classification Result.** 10-fold cross-validation results of the most optimal performance feature combination sets of each classifier and their accuracy in the detection of depression are shown below: results of resting-state data, neutral audio stimulation data, positive audio stimulation data, and negative audio stimulation data are summarized Tables 3, 4, 5, and 6, respectively.

For resting-state EEG data, KNN achieved the best accuracy of 76.83% using feature combination of absolute power of gamma wave on Fp1, absolute power of theta wave on Fp2, absolute power of beta wave on Fp2, and absolute center frequency of beta wave on Fp2 (Table 3).

For EEG data of participants under neutral audio stimulation, KNN achieved the best accuracy of 74.39% using the feature combination of absolute power of theta on Fp1, center frequency of full-band EEG on Fp2, and peak of full-band EEG on Fp2 (Table 4).

For EEG data of participants under positive audio stimulation, KNN achieved the best accuracy of 79.27% using the feature combination of absolute power of theta wave on Fp1 and absolute power of beta wave on Fp1 (Table 5).

For EEG data of participants under negative audio stimulation, KNN achieved the best accuracy of 77.44% using feature combination of absolute power of theta wave on Fp1, correlation dimension of full-band EEG on Fp1, absolute center frequency of theta wave on Fp2, and absolute power of gamma wave on Fp2 (Table 6).

The results showed that, among all the four classifiers of SVM, KNN, CT, and ANN, KNN performed the best with an average classification accuracy of 76.98% (Figure 4). The absolute power of theta wave appeared in all the best performance feature combination, thereby indicating a potential connection between theta wave and depression. The absolute power of theta wave might be a valid characteristic for pervasive depression discrimination.



TABLE 3: Classification results in resting state data.

Classifier	Feature sets	Accuracy
SVM	Absolute power of theta wave (Fp1), relative power of theta wave (Fp1), relative power of alpha wave (Fp1), absolute center frequency of gamma wave (Fp1)	72.56%
KNN	Absolute power of gamma wave (Fp1), absolute power of theta wave (Fp2), absolute power of beta wave (Fp2), absolute center frequency of beta wave (Fp2)	76.83%
CT	Peak (Fp1), Power-Spectral Entropy of full band EEG (Fp1), absolute power of beta wave (Fp2)	68.29%
ANN	Absolute center frequency of beta wave (Fp1), absolute power of gamma wave (Fp1), Kurtosis of full band EEG (Fp1), absolute power of alpha wave (Fp2), relative center frequency of beta wave (Fp2)	72.56%

TABLE 4: Classification results in neutral audio stimulation data.

Classifier	Feature sets	Accuracy
SVM	Absolute center frequency of theta wave (Fp1), center frequency of full band EEG (Fp2), peak of full band EEG (Fp2)	70.12%
KNN	Absolute power of theta (Fp1), center frequency of full band EEG (Fp2), peak of full band EEG (Fp2)	74.39%
CT	Absolute center frequency of alpha wave (Fp1), Power-Spectral Entropy of alpha wave (Fp1)	67.70%
ANN	Relative center frequency of theta wave (Fp2), Hjorth of full band EEG (Fp2)	73.78%

TABLE 5: Classification results in positive audio stimulation data.

Classifier	Feature sets	Accuracy
SVM	Absolute power of theta wave (Fp1), Kurtosis of full band EEG (Fp1)	68.29%
KNN	Absolute power of theta wave (Fp1), absolute power of beta wave (Fp1)	79.27%
CT	Absolute power of gamma wave (Fp1), absolute power of gamma wave (Fp2)	60.37%
ANN	Absolute power of theta wave (Fp1), power spectral entropy of gamma wave (Fp1), C0-complexity of full band EEG (Fp1), correlation dimension of full band EEG (Fp1), power spectral entropy of theta wave (Fp2), correlation dimension of full band EEG (Fp2)	74.39%

TABLE 6: Classification results in negative audio stimulation data.

Classifier	Feature sets	Accuracy
SVM	Hjorth of full band EEG (Fp2), correlation dimension of full band EEG (Fp2)	67.07%
KNN	Absolute power of theta wave (Fp1), correlation dimension of full band EEG (Fp1), absolute center frequency of theta wave (Fp2), absolute power of gamma wave (Fp2)	77.44%
CT	Absolute power of theta wave (Fp1), power spectral entropy of full band EEG (Fp1), relative power of beta wave (Fp2), peak of full band EEG (Fp2)	71.34%
ANN	Absolute power of theta wave (Fp1), absolute power of beta wave (Fp1), center frequency of full band EEG (Fp1)	71.34%

## 5. Conclusion and Future Work

Depression is a major health concern in millions of individuals. Thus, diagnosing depression in the early curable stages is critical for the treatment in order to save the life of a patient. However, current methods of depression detection are human-intensive, and their results are dependent on the experience of the doctor. Therefore, a pervasive and objective method of diagnosing or even screening would be useful. The present study explores a novel method of depression detection using pervasive prefrontal-lobe three-electrode EEG system, which chooses Fp1, Fp2, and Fpz for electrode sites, according to the international 10-20 system.

Several widely employed psychological scales were used to select the optimal experimental candidates, which encompassed 213 participants (92 depressed patients and 121 normal

controls). Their EEG data of resting state, as well as under sound stimulation, were recorded. The soundtracks were selected from the IADS-2 database, comprising positive, neutral, and negative stimuli.

The FIR filter combining the Kalman derivation formula, DWT, and an APF were applied on the raw EEG data to remove the interference from environment, ECG, EMG, and EOG. Subsequently, 270 linear and nonlinear features were extracted from the preprocessed EEG. Then, the MRMR technique was applied to perform the feature selection. Four classification algorithms, KNN, SVM, CT, and ANN, have been evaluated and compared, using a 10-fold cross-validation. The results exhibited KNN as the best performance classification method in all datasets, with the highest accuracy of 79.27%. The results also demonstrated the feature "absolute power of theta wave" in all the best

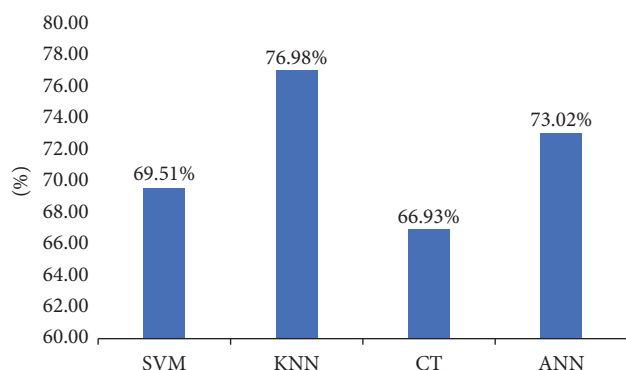


FIGURE 4: Average accuracy of classification on selected features.

performance features of the four datasets, thereby suggesting a robust connection between the power of theta wave and depression; this could be used as a valid characteristic feature in the detection of depression.

The current study postulated that a novel and pervasive system for screening depression is feasible. With a carefully designed model, the pervasive system could reach accuracy similar to the current scale-based screening method; for instance, the accuracy of BDI is reported to be 79–86% in different studies [99–101].

EEG and depression have been under intensive focus of research. In comparison to the study by Knott et al., who collected EEG recordings from 21 scalp sites and conducted univariate analyses for group comparisons and correctly classified 91.3% of the patients and controls [102], the current pervasive three-electrode EEG acquisition system can be measured easily and is rather suitable for the personal use of patients. Moreover, we used a less number of electrodes, thereby reducing the amount of data considerably. Healey and Picard attempted to extract feature patterns from physiological signals for emotional recognition, with an accuracy of 61.8–78.4% [103]. Kim et al. developed a short-term monitoring emotion recognition system based on multiuser physiological signals to extract features and identify three emotions using SVM. The final recognition accuracy was 75% [104]. Compared to these studies, our result has a higher accuracy with faster data processing efficiency. Henriques collected the resting-state EEG (with closed eyes) from 5 depressed patients and 13 normal individuals. The results showed that the activity in the left hemisphere in depressed patients is significantly lower than that in a normal person [105]. Omelchenko and Zaika collected EEG data from 53 depressed patients and 86 normal individuals and demonstrated that patients with depression have a higher delta and theta energy than normal, but a lower alpha and beta energy [106]. Fingelkurts et al. researched the shock components of resting-state EEG from 12 depressed patients and 10 normal persons and found that the brain activates were affected by depression throughout the cerebral cortex [107]. Compared to these studies, the current experiment presented more reliable data to ensure the reliability of the experimental results. Taken together, our data model based on feature extraction and feature selection reduced the amount of data

to be processed with a faster data processing efficiency. In addition, the pervasive three-electrode EEG acquisition system was easier to measure, as well as ensuring the accuracy of data.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Basic Research Program of China (973 Program) (no. 2014CB744600), the National Natural Science Foundation of China (Grants nos. 61210010 and 61632014), and the Program of International S&T Cooperation of MOST (no. 2013DFA11140).

## References

- [1] M. E. P. Seligman, *Helplessness: on Depression, Development, and Death*, WH Freeman/Times Books/Henry Holt & Co, 1975.
- [2] “World Health Organization. Fact sheet on depression [EB/OL]. [2017-02],” <http://www.who.int/mediacentre/factsheets/fs369/en/>.
- [3] World Health Organization, *The World Health Report 2001: Mental health: new understanding, new hope*, World Health Organization, 2001.
- [4] World Health Organization, *Sixty-Fifth World Health Assembly (2012)* [EB/OL], 2012, <http://www.who.int/mediacentre/events/2012/wha65/en/>.
- [5] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, vol. 1, American Psychiatric Association, Arlington, Va, USA, 4th edition, 2000.
- [6] V. C. Pangman, J. Sloan, and L. Guse, “An examination of psychometric properties of the mini-mental state examination and the standardized mini-mental state examination: implications for clinical practice,” *Applied Nursing Research*, vol. 13, no. 4, pp. 209–213, 2000.
- [7] A. Beck T, R. Steer A, and K. Brown G, *Beck depression inventory*, 1996.
- [8] M. Hamilton, “A rating scale for depression,” *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 23, pp. 56–62, 1960.
- [9] B. Lindsley D, *Emotions and the electroencephalogram*, 1950.
- [10] B. Wilder, “EEG in clinical practice,” *Electroencephalography and Clinical Neurophysiology*, vol. 56, no. 5, p. 536, 1983.
- [11] K. G. Jordan, “Continuous EEG and evoked potential monitoring in the neuroscience intensive care unit,” *Journal of Clinical Neurophysiology*, vol. 10, no. 4, pp. 445–475, 1993.
- [12] J. C. Woestenburger, M. N. Verbaten, and J. L. Slangen, “The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain,” *Biological Psychology*, vol. 16, no. 1-2, pp. 127–147, 1983.
- [13] W. Klimesch, “EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis,” *Brain Research Reviews*, vol. 29, no. 2-3, pp. 169–195, 1999.
- [14] H. W. Cole and W. J. Ray, “EEG correlates of emotional tasks related to attentional demands,” *International Journal of Psychophysiology*, vol. 3, no. 1, pp. 33–41, 1985.

- [15] W. Klimesch, M. Doppelmayr, H. Russegger, T. Pachinger, and J. Schwaiger, "Induced alpha band power changes in the human EEG and attention," *Neuroscience Letters*, vol. 244, no. 2, pp. 73–76, 1998.
- [16] R. Srinivasan, S. Thorpe, S. Deng, T. Lappas, and M. D'Zmura, "Decoding attentional orientation from eeg spectra," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 5610, no. 1, pp. 176–183, 2009.
- [17] E. A. Curran and M. J. Stokes, "Learning to control brain activity: A review of the production and control of EEG components for driving brain-computer interface (BCI) systems," *Brain and Cognition*, vol. 51, no. 3, pp. 326–336, 2003.
- [18] A. S. Gevins, G. M. Zeitlin, C. D. Yingling et al., "EEG patterns during 'cognitive' tasks. I. Methodology and analysis of complex behaviors," *Electroencephalography and Clinical Neurophysiology*, vol. 47, no. 6, pp. 693–703, 1979.
- [19] H. Haken, *Principles of brain functioning: a synergetic approach to brain activity, behavior and cognition*, Springer Science & Business Media, 2013.
- [20] F. Fan, Y. Li, Y. Qiu, and Y. Zh, "Use of ANN and Complexity Measures in Cognitive EEG Discrimination," in *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 4638–4641, Shanghai, China, 2005.
- [21] H. Merica, R. Blois, and J.-M. Gaillard, "Spectral characteristics of sleep EEG in chronic insomnia," *European Journal of Neuroscience*, vol. 10, no. 5, pp. 1826–1834, 1998.
- [22] R. R. Rosa and M. H. Bonnet, "Reported chronic insomnia is independent of poor sleep as measured by electroencephalography," *Psychosomatic Medicine*, vol. 62, no. 4, pp. 474–482, 2000.
- [23] H. Merica and J.-M. Gaillard, "The EEG of the sleep onset period in insomnia: A discriminant analysis," *Physiology & Behavior*, vol. 52, no. 2, pp. 199–204, 1992.
- [24] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 2, pp. 205–211, 2007.
- [25] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients," *Physica D: Nonlinear Phenomena*, vol. 144, no. 3, pp. 358–369, 2000.
- [26] K. Lehnertz, G. R. J. Arnhold, and T. Kreuz, "Nonlinear EEG Analysis in Epilepsy: Its Possible Use for Interictal Focus Localization, Seizure Anticipation, and Prevention," *Journal of Clinical Neurophysiology*, vol. 18, no. 3, pp. 209–222.
- [27] R. Williams L, I. Karacan, and C. J. Hirsch, *Electroencephalography (EEG) of human sleep: clinical applications*, John Wiley & Sons, 1974.
- [28] D. K. Hannesdóttir, J. Doxie, M. A. Bell, T. H. Ollendick, and C. D. Wolfe, "A longitudinal study of emotion regulation and anxiety in middle childhood: Associations with frontal EEG asymmetry in early childhood," *Developmental Psychobiology*, vol. 52, no. 2, pp. 197–204, 2010.
- [29] O. Siciliani, M. Schiavon, and M. Tansella, "Anxiety and EEG alpha activity in neurotic patients," *Acta Psychiatrica Scandinavica*, vol. 52, no. 2, pp. 116–131, 1975.
- [30] J. Avram, F. R. Baltes, M. Miclea, and A. C. Miu, "Frontal EEG activation asymmetry reflects cognitive biases in anxiety: Evidence from an emotional face Stroop task," *Applied Psychophysiology and Biofeedback*, vol. 35, no. 4, pp. 285–292, 2010.
- [31] B. A. Clementz, S. R. Sponheim, W. G. Iacono, and M. Beiser, "Resting EEG in first-episode schizophrenia patients, bipolar psychosis patients, and their first-degree relatives," *Psychophysiology*, vol. 31, no. 5, pp. 486–494, 1994.
- [32] R. Diaz-guerrero, J. S. Gottlieb, and J. R. Knott, "The sleep of patients with manic-depressive psychosis, depressive type; an electroencephalographic study," *Psychosomatic Medicine*, vol. 8, no. 6, pp. 399–404, 1946.
- [33] P. A. Davis and H. Davis, "The electroencephalograms of psychotic patients," *The American Journal of Psychiatry*, vol. 95, no. 5, pp. 1007–1025, 1939.
- [34] S. R. Sponheim, B. A. Clementz, W. G. Iacono, and M. Beiser, "Clinical and biological concomitants of resting state EEG power abnormalities in schizophrenia," *Biological Psychiatry*, vol. 48, no. 11, pp. 1088–1097, 2000.
- [35] J. Bureš, O. Burešová, and J. Křivánek, *The mechanism and applications of Leao's spreading depression of electroencephalographic activity*, Academic Press, 1974.
- [36] G. E. Bruder, R. Fong, C. E. Tenke et al., "Regional brain asymmetries in major depression with or without an anxiety disorder: A quantitative electroencephalographic study," *Biological Psychiatry*, vol. 41, no. 9, pp. 939–948, 1997.
- [37] S. A. Reid, L. M. Duke, and J. J. B. Allen, "Resting frontal electroencephalographic asymmetry in depression: Inconsistencies suggest the need to identify mediating factors," *Psychophysiology*, vol. 35, no. 4, pp. 389–404, 1998.
- [38] R. Thibodeau, R. S. Jorgensen, and S. Kim, "Depression, anxiety, and resting frontal EEG asymmetry: a meta-analytic review," *Journal of Abnormal Psychology*, vol. 115, no. 4, pp. 715–729, 2006.
- [39] N. Liang, P. Saratchandran, G. Huang, and N. Sundararajan, "Classification of mental tasks from EEG signals using extreme learning machine," *International Journal of Neural Systems*, vol. 16, no. 1, pp. 29–38, 2006.
- [40] C. W. Anderson, E. A. Stolz, and S. Shamsunder, "Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 3, pp. 277–286, 1998.
- [41] J. D. R. Millán, J. Mouriño, M. Franzé et al., "A local neural classifier for the recognition of EEG patterns associated to mental tasks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 13, no. 3, pp. 678–686, 2002.
- [42] T. Harmony, T. Fernández, J. Silva et al., "EEG delta activity: An indicator of attention to internal processing during performance of mental tasks," *International Journal of Psychophysiology*, vol. 24, no. 1-2, pp. 161–171, 1996.
- [43] "Electrical Geodesics, Inc," <https://www.egi.com/>.
- [44] "Brain Products GmbH," <http://www.brainproducts.com/>.
- [45] "UAIS Lab of Lanzhou University," <http://www.theuais.org/>.
- [46] H. H. Jasper, "The ten twenty electrode system of the international federation," *Electroencephalography and Clinical Neurophysiology*, vol. 10, pp. 371–375, 1958.
- [47] G. E. Chatrian, E. Lettich, and P. L. Nelson, "Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities," *American Journal of EEG Technology*, vol. 25, no. 2, pp. 83–92, 1985.
- [48] S. Sanei and J. A. Chambers, *EEG signal processing*, John Wiley & Sons, 2013.
- [49] H. Jasper, "The ten twenty electrode system of the international federation," *Electroencephalography and Clinical Neurophysiology*, vol. 10, pp. 371–375, 1958.



- [50] W. J. H. Nauta, "The problem of the frontal lobe: A reinterpretation," *Journal of Psychiatric Research*, vol. 8, no. 3-4, pp. 167-187, 1971.
- [51] E. T. Rolls, "On the brain and emotion," *Behavioral and Brain Sciences*, vol. 23, no. 2, pp. 219-228, 2000.
- [52] E. Harmon-Jones, "Contributions from research on anger and cognitive dissonance to understanding the motivational functions of asymmetrical frontal brain activity," *Biological Psychology*, vol. 67, no. 1-2, pp. 51-76, 2004.
- [53] Lanzhou University, "A portable EEG acquisition system," China: CN201520628152.6, 2016-02-17.
- [54] J.-M. Azorin, P. Benhaim, T. Hasbroucq, and C.-A. Possamai, "Stimulus preprocessing and response selection in depression: A reaction time study," *Acta Psychologica*, vol. 89, no. 2, pp. 95-100, 1995.
- [55] W. O. Tatum, B. A. Dworetzky, and D. L. Schomer, "Artifact and recording concepts in EEG," *Journal of Clinical Neurophysiology*, vol. 28, no. 3, pp. 252-263, 2011.
- [56] A. T. Beck, *Depression: Clinical, experimental, and theoretical aspects*, University of Pennsylvania Press, 1967.
- [57] J. Epstein, H. Pan, J. H. Kocsis et al., "Lack of ventral striatal response to positive stimuli in depressed versus normal subjects," *The American Journal of Psychiatry*, vol. 163, no. 10, pp. 1784-1790, 2006.
- [58] L. M. Bylsma, B. H. Morris, and J. Rottenberg, "A meta-analysis of emotional reactivity in major depressive disorder," *Clinical Psychology Review*, vol. 28, no. 4, pp. 676-691, 2008.
- [59] E.-M. Yong, W.-Q. Qian, and K.-F. He, "Penetration ability analysis for glide reentry trajectory based on radar tracking," *Yuhang Xuebao/Journal of Astronautics*, vol. 33, no. 10, pp. 1370-1376, 2012.
- [60] W. Ende, Z. Feng, X. Yanghui, T. Xinxin, and Z. Dan, "Generating Method of Adaptive Linear Signal Based on Kalman Prediction," *Modern Defense Technology*, vol. 40, no. 2, pp. 138-142, 2012.
- [61] Z. Peng, "Aero engine Fault Diagnostics Based on Kalman Filter," *Nanjing University of Aeronautics and Astronautics*, 2008.
- [62] Z. Xiao, W. Ming, and L. Xiaoming, "pplication of extended Kalman filter algorithm in intelligent robot soccer competition," *Electrical and Mechanical Engineering*, vol. 29, pp. 334-338, 2012.
- [63] H. Singh and J. Singh, "A review on electrooculography," *International Journal of Advanced Engineering Technology*, vol. 3, pp. 115-122.
- [64] L. Yang, *Study of Ocular Artifacts Removal Based on Wavelet Transform and Kalman filter*, Lanzhou University, 2016.
- [65] S. Tong, A. Bezerianos, J. Paul, Y. Zhu, and N. Thakor, "Removal of ECG interference from the EEG recordings in small animals using independent component analysis," *Journal of Neuroscience Methods*, vol. 108, no. 1, pp. 11-17, 2001.
- [66] A. Gevins S, W. Du, and H. Leong, *Adaptive interference canceler for EEG movement and eye artifacts*, U.S. Patent 5,513,649[P]. 1996-5-7.
- [67] J. P. M. Pijn, *Quantitative evaluation of EEG signals in epilepsy: nonlinear association time delays and nonlinear dynamics [Ph.D. thesis]*, University of Amsterdam, 1990.
- [68] J. P. M. Pijn, D. N. Veils, M. J. Van Der Heyden, J. DeGoede, C. W. M. Van Veelen, and F. H. Lopes Da Silva, "Nonlinear dynamics of epileptic seizures on basis of intracranial EEG recordings," *Brain Topography*, vol. 9, no. 4, pp. 249-270, 1997.
- [69] S. A. R. B. Rombouts, R. W. M. Keunen, and C. J. Stam, "Investigation of nonlinear structure in multichannel EEG," *Physics Letters A*, vol. 202, no. 5-6, pp. 352-358, 1995.
- [70] C. J. Stam, T. C. A. M. Van Woerkom, and R. W. M. Keunen, "Non-linear analysis of the electroencephalogram in Creutzfeldt-Jakob disease," *Biological Cybernetics*, vol. 77, no. 4, pp. 247-256, 1997.
- [71] U. R. Acharya, O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan, "Non-linear analysis of EEG signals at various sleep stages," *Computer Methods and Programs in Biomedicine*, vol. 80, no. 1, pp. 37-45, 2005.
- [72] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306-310, 1970.
- [73] E. H. Shen, Z. J. Cai, and F. J. Gu, "Mathematical foundation of a new complexity measure," *Applied Mathematics and Mechanics. Yingyong Shuxue He Lixue*, vol. 26, no. 9, pp. 1083-1090, 2005.
- [74] A. Lempel and J. Ziv, "On the Complexity of Finite Sequences," *IEEE Transactions on Information Theory*, vol. 22, no. 1, pp. 75-81, 1976.
- [75] C. E. Shannon, "A mathematical theory of communication," *Bell Labs Technical Journal*, vol. 27, pp. 379-423, 623-656, 1948.
- [76] J. Bruhn, L. E. Lehmann, H. Röpcke, T. W. Bouillon, and A. Hoeft, "Shannon entropy applied to the measurement of the electroencephalographic effects of desflurane," *Anesthesiology*, vol. 95, no. 1, pp. 30-35, 2001.
- [77] C. E. Shannon, "A mathematical theory of communication," *Acm Sigmoble Mobile Computing & Communications Review*, vol. 5, no. 1, pp. 3-55, 1948.
- [78] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226-1238, 2005.
- [79] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.
- [80] M. Hearst, S. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and Their Applications*, vol. 13, no. 4, pp. 18-28, 1998.
- [81] T. Joachims, "Text categorization with support vector machines: learning with many relevant features," in *Machine Learning: ECML-98*, vol. 1398, pp. 137-142, Springer, Berlin, Germany, 1998.
- [82] B. Gaonkar and C. Davatzikos, "Analytic estimation of statistical significance maps for support vector machine based multivariate image analysis and classification," *NeuroImage*, vol. 78, pp. 270-283, 2013.
- [83] R. Cuignnet, C. Rosso, M. Chupin et al., "Spatial regularization of SVM for the detection of diffusion alterations associated with stroke outcome," *Medical Image Analysis*, vol. 15, no. 5, pp. 729-737, 2011.
- [84] C. Bahlmann, B. Haasdonk, and H. Burkhardt, "Online handwriting recognition with support vector machines - A kernel approach," in *Proceedings of the 8th International Workshop on Frontiers in Handwriting Recognition, IWFHR 2002*, pp. 49-54, Canada, August 2002.
- [85] O. Faust, P. C. A. Ang, S. D. Puthankattil, and P. K. Joseph, "Depression diagnosis support system based on eeg signal entropies," *Journal of Mechanics in Medicine and Biology*, vol. 14, no. 3, Article ID 1450035, 2014.



- [86] B. Hosseinifard, M. H. Moradi, and R. Rostami, "Classifying depression patients and normal subjects using machine learning techniques," in *Proceedings of the 2011 19th Iranian Conference on Electrical Engineering, ICEE 2011*, Iran, May 2011.
- [87] I. Kalatzis, N. Piliouras, E. Ventouras, C. Papageorgiou, A. Rabavilas, and D. Cavouras, "Comparative evaluation of probabilistic neural network versus support vector machines classifiers in discriminating ERP signals of depressive patients from healthy controls," in *Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis, 2003. ISPA 2003.*, pp. 981–985, Rome, Italy.
- [88] B. V. Dasarathy, *Nearest neighbor (NN) norms: NN pattern classification techniques*, 1991.
- [89] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K. H. Ng, and J. S. Suri, "Automated diagnosis of epileptic EEG using entropies," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 401–408, 2012.
- [90] J.-S. Wang, C.-W. Lin, and Y.-T. C. Yang, "A k-nearest-neighbor classifier with heart rate variability feature-based transformation algorithm for driving stress recognition," *Neurocomputing*, vol. 116, pp. 136–143, 2013.
- [91] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23–38, 1998.
- [92] M. Hansen, R. Dubayah, and R. Defries, "Classification trees: An alternative to traditional land cover classifiers," *International Journal of Remote Sensing*, vol. 17, no. 5, pp. 1075–1081, 1996.
- [93] C. Lehmann, T. Koenig, V. Jelic et al., "Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG)," *Journal of Neuroscience Methods*, vol. 161, no. 2, pp. 342–350, 2007.
- [94] U. R. Acharya, V. K. Sudarshan, H. Adeli et al., "A novel depression diagnosis index using nonlinear features in EEG signals," *European Neurology*, vol. 74, no. 1-2, pp. 79–83, 2016.
- [95] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *Journal of Biomedical Informatics*, vol. 35, no. 5-6, pp. 352–359, 2002.
- [96] G. Hinton, L. Deng, D. Yu et al., "Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.
- [97] S. D. Puthankattil and P. K. Joseph, "Classification of eeg signals in normal and depression conditions by ann using rwe and signal entropy," *Journal of Mechanics in Medicine and Biology*, vol. 12, no. 4, Article ID 1240019, 2012.
- [98] Y. Li and F. Fan, "Classification of Schizophrenia and depression by EEG with ANNs," in *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 2679–2682, Engineering in Medicine and Biology Society, Shanghai, China, January 2006.
- [99] A. Fischer, M. Fischer, R. A. Nicholls et al., "Diagnostic accuracy for major depression in multiple sclerosis using self-report questionnaires," *Brain and Behavior*, vol. 5, no. 9, Article ID e00365, 2015.
- [100] I. M. Cameron, A. Cardy, J. R. Crawford et al., "Measuring depression severity in general practice: Discriminatory performance of the PHQ-9, HADS-D, and BDI-II," *British Journal of General Practice*, vol. 61, no. 588, pp. e419–e426, 2011.
- [101] C. D. Silberman, J. Laks, C. F. Capitão, C. S. Rodrigues, I. Moreira, and E. Engelhardt, "Recognizing depression in patients with Parkinson's disease: Accuracy and specificity of two depression rating scale," *Arquivos de Neuro-Psiquiatria*, vol. 64, no. 2 B, pp. 407–411, 2006.
- [102] V. Knott, C. Mahoney, S. Kennedy, and K. Evans, "EEG power, frequency, asymmetry and coherence in male depression," *Psychiatry Research: Neuroimaging*, vol. 106, no. 2, pp. 123–140, 2001.
- [103] J. Healey and R. Picard, "Digital processing of affective signals," in *Proceedings of the 1998 23rd IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 1998*, pp. 3749–3752, USA, May 1998.
- [104] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Medical & Biological Engineering & Computing*, vol. 42, no. 3, pp. 419–427, 2004.
- [105] J. B. Henriques and R. J. Davidson, "Left Frontal Hypoactivation in Depression," *Journal of Abnormal Psychology*, vol. 100, no. 4, pp. 535–545, 1991.
- [106] V. P. Omelchenko and V. G. Zaika, "Changes in the EEG rhythms in endogenous depressive disorders and the effect of pharmacotherapy," *Fiziologiya cheloveka*, vol. 28, no. 3, pp. 30–36, 2002.
- [107] A. A. Fingelkurts, A. A. Fingelkurts, H. Ryttsälä, K. Suominen, E. Isometsä, and S. Kähkönen, "Composition of brain oscillations in ongoing EEG during major depression disorder," *Neuroscience Research*, vol. 56, no. 2, pp. 133–144, 2006.

