Received October 8, 2020, accepted November 2, 2020, date of publication November 17, 2020, date of current version January 4, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3038486

An Efficient Method to Detect Sleep Hypopnea- Apnea Events Based on EEG Signals

YAO WANG[1,3], SIYU JI[2], TIANSHUN YANG[2], XIAOHONG WANG[1], HUIQUAN WANG[1,3], AND XIAOYUN ZHAO[4]

[1]School of Life Sciences, Tiangong University, Tianjin 300387, China
[2]School of Electronics and Information Engineering, Tiangong University, Tianjin 300387, China
[3]School of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China
[4]Tianjin Chest Hospital, Tianjin 300000, China

Corresponding author: Xiaoyun Zhao (zxydoctor@163.com)

This work was supported in part by the National Key Research and Development Program of China (No. 2019YF0119400), National Natural Science Foundation of China (No. 61701342, 81901789), the Tianjin Natural Science Foundation (No. 19JCQNJC13100), the Tianjin Science and technology plan project (No. 18ZXRHSY00200, 20ZXGBSY00070), the Science & Technology Development Fund of Tianjin Education Commission for Higher Education (No. 2018KJ212).

ABSTRACT Hypopnea refers to the state in which insufficient alveolar ventilation during night sleep decreases the respiratory airflow by more than 50% of the airflow. However, sleep apnea is a more serious respiratory event, such as complete cessation of respiratory airflow for 10 seconds. The occurrence of hypopnea is a precursor to the occurrence of apnea events and the two are closely connected. In this paper, we propose a method based on the combination of discrete wavelet transform and approximate entropy of EEG signals to detect sleep apnea and hypopnea events. For this purpose, first, data preprocessing is performed on the EEG record data set obtained from Tianjin Chest Hospital, and then infinite impulse response (IIR) Butterworth bandpass filter is used to decompose the data into delta, theta, alpha, beta and gamma. Second, descriptive features are extracted based on sub-bands discrete wavelet transform such as the approximate entropy of high-frequency coefficients. Third, the features are filtered based on Support Vector Machine (SVM) recursive elimination. Finally, several machine learning algorithms including SVM, K-Nearest Neighbor (KNN) and Random forest (RF) are employed to identify the occurrence of sleep hypopnea-apnea events. The highest accuracy rate reached 94.33%, the sensitivity reached 93.10%, and the specificity reached 95.07%. The obtained results validate that the proposed method is an effective and practical diagnostic method to detect the occurrence of hypopnea-apnea events.

INDEX TERMS Sleep apnea hypopnea syndrome, discrete wavelet transform, approximate entropy, machine learning classification model, support vector machine recursive elimination.

I. INTRODUCTION

Sleep disturbances can lead to insufficient sleep at night and mental fatigue during the day, which can have a significant negative impact on our lives. The two main challenging problems in sleep analysis are sleep stage scoring and apnea-hypopnea detection [1]. Hypopnea refers to a 30% drop in oronasal airflow and a 4% drop in blood oxygen saturation for more than 10 seconds, or 50% drop in nose and mouth airflow and 3% drop in blood oxygen saturation for more than 10 s. However, sleep apnea syndrome (SAS) is a sleep disorder in which breathing stops during nocturnal sleep, each time the airflow is stopped for more than 10 s (including 10 s) or the average number of hypoventilations per hour (breathing disorder index) exceeds 5 times. The three types of sleep apnea are: Obstructive Sleep Apnea (OSA), Central Sleep Apnea (CSA), and Mixed Sleep Apnea (MSA). The OSA is one of the most common, frequent and serious sleep disorders. It causes complete obstruction of the upper airway and relaxation of the throat muscles, thereby obstructing the flow of breathing during sleep [12]. The CSA refers to the process in which the brain center stops sending signals to the muscles that control breathing, causing the interruption of respiratory airflow. The MSA is obstructive apnea in the first half and central apnea in the second half or vice versa. The most common type of apnea in clinical practice

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

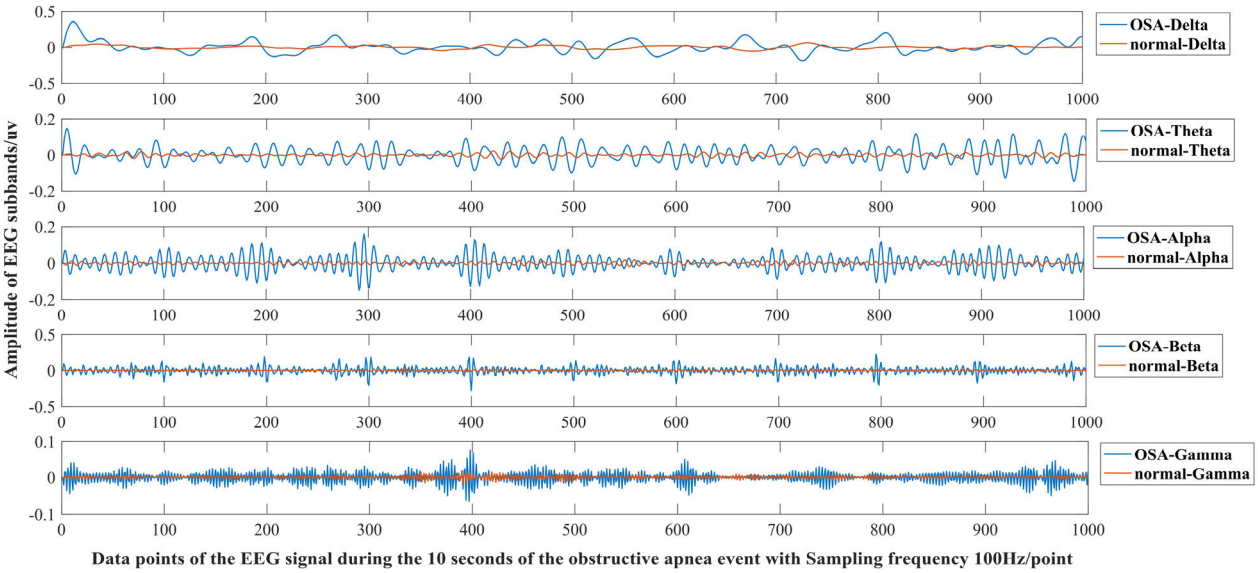

FIGURE 1. Five different changes of brain waves in the event of OSA.

is obstructive apnea. In this paper, the three types of apneas are allocated according to a reasonable proportion proposed by the medical experts of respiratory diseases from Tianjin Chest Hospital, the OSA is 60%, and the MSA and the CSA are 30% and 10%, respectively. Since the characteristics of hypopnea-apnea syndrome are different from other diseases, the symptoms first appear during nighttime sleep, and also promote the development of cardiopulmonary failure [2]–[4]. Therefore, it is particularly important to detect the presence of hypopnea-apnea syndrome in time.

The electroencephalogram (EEG) signals are the spontaneous potential activities generated by brain nerve activity and always present in the central nervous system. These signals are rich in brain activity information and are an important means of brain research, physiological research, and clinical brain disease diagnosis [5]. Moreover, recording the EEG signals is a non-invasive method to measure the brain waves. It is not only related to processing the functional information and physiological mechanisms, but also can realize the memory of multiple processes from simple state to complex state. The EEG signal contains sleep-related information and can provide relevant information for sleep disordered breathing. Therefore, the EEG signal feature extraction and intelligent methods for the corresponding waveband can be used to treat sleep disordered breathing [6], [7]. The EEG signals are widely used in sleep-related problems, and the advantage is that only separated EEG signals are used and no other auxiliary physiological signals are included [8]–[11]. When a hypopnea or apnea event occurs, the EEG signal will show different characteristics in different frequency bands without exception. The EEG signal can be divided into five frequency sub-bands, that is, delta (0-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-32 Hz), and gamma (32-40 Hz). Obviously, when a hypopnea or apnea event occurs, all five-segmented band waves have changes. Take the frequently occurring obstructive apnea as an example, the specific changes that can be observed from the preprocessed data time-domain waveforms are shown in Figure 1.

In this article, the changes of the five different sub-band waves corresponding to the EEG signals are studied and Butterworth filter method is used to extract the sub-bands. The obtained different sub-bands are further subjected to discrete wavelet transform and db3 mother wavelet and four-layer decomposition are selected as the parameters of the transformation. To study the detection of sleep apnea-hypopnea events, we extract features from the first layer of detail coefficient, the second layer of detail coefficient, the third layer of detail coefficient and the fourth layer of detail coefficient respectively. At the same time, the proposed method in this article finds the approximate entropy of the detail coefficient as characteristics of EEG signals during the apnea-hypopnea events. Then a machine learning classification model is established and results are obtained on the test set as the final evaluation index of the feature. Our model achieves the best result with accuracy 94.33%, sensitivity 93.10%, and specificity 95.07% and provides a detection method that uses the EEG signals to automatically detect the hypopnea-apnea events that occur during nighttime sleep.

II. RELATED WORK

Detecting sleep apnea-hypopnea events is a time-consuming and laborious process. Many researchers have applied different physiological signals to the detection of sleep apnea events such as ECG, SPO2, airflow signal and nasal pressure signal. Based on single-lead ECG signal, a non-parametric kernel density-based approach was used for OSA detection and the results obtained were mean accuracy of 82.07%, with mean sensitivity of 83.23% and mean specificity of 80.24% [13]. A method for obstructive sleep apnea severity prediction based on single channel ECG signal was proposed, the accu-

racy of 79.45% for OSA severity classification with sensitivity, specificity, and F-score was achieved [14]. Based on the pulse oximetry signal at night, 16 features (time and frequency statistics, frequency spectrum and nonlinear features) were extracted and genetic algorithm (GA) method was applied in the feature selection stage. Finally, they achieved 87.5% accuracy (90.6% sensitivity and 81.3% specificity) in the test set [15]. In addition, ECG signals and peripheral blood oxygen signals were combined to detect sleep apnea-hypopnea events and the achieved sensitivity, specificity, and accuracy all were around 82% [16]. In Payongkit Lakhan *et al.*, seventeen features (average of maximum and minimum amplitudes from all airflow signal samples etc.) have been extracted from airflow signal and then fed into Deep Neural Networks to binary classify and the cutoff indices at AHI = 5, 15 and 30. The result obtained were accuracy of 83.46%, 85.39% and 92.69% in each cutoff, respectively [17]. Automatic real-time detection of sleep apnea and hypopnea events based on nasal pressure signals has been studied and a good performance has achieved regardless of the severity of AHI, the sensitivity is 86.4%, and the positive predictive value for apnea and hypopnea is 84.5% [18]. The air flow, thoracic and abdominal respiratory movement data have been partitioned and organized into Reasoning Units (RU), which are used in signal segmentation and the static features namely skewness, kurtosis, mean, geomean, variance, and standard deviation are generated from the wavelet packet coefficients, which aid the classification of sleep apnea from normal patients using Partially Connected Cooperative Parallel Particle Swarm Optimization-Support Vector Machine algorithm producing an accuracy of 83.66% [21]. The classification of sleep apnea using cross wavelet transform of airflow and thoracic effort signals in combination of higher order statistics extractions obtained from the Kernel based non-linear Principal Component Analysis (KPCA) produced a sleep apnea classification system with an accuracy value of 85% [22]. Since, EEG signals contain a lot of important information, which reflects activities and abnormal breathing, in this paper, a feature extraction based on EEG signals is proposed to detect sleep apnea-hypopnea events and achieve more accurate results.

III. MATERIALS AND METHODS

A. SUBJECTS

In order to verify the method proposed in this study, the EEG signals from the PSG record of 30 apnea patients with differences in AHI collected in Tianjin Chest Hospital were taken as the research subjects, and the EEG data of the two channels C3-A2 and C4-A1 of the database was used. The sampling frequency was 100 Hz per second and contained the truth labels of apnea-hypopnea events. The time of each apnea-hypopnea event was at least 10s. The detailed information of 30 subjects are as follows: the number of patients with severe apnea (AHI>30) accounted for 19 people, of which 15 were males between 34 and 73 years old and 4 were females between 65 and 78 years old; the number of patients with moderate apnea (15<AHI<30) accounted for 7 people, five males aged between 37 and 55 years old and two females between 54 and 66 years old; the number of patients with mild apnea (5<AHI<15) accounted for 4 people, three males aged between 41 and 70 years old and one female was 54 years old.

TABLE 1. SA dataset specifications (Apnea constructed according to the ratio of OSA: MSA: CSA = 6:3:1).

events	Training set	Test set	Total
Normal	609	203	812
Apnea	305	102	407
Hypopnea	304	101	405
Total	1218	406	1624

Based on the existing subject data, we extracted from 30 subjects all existing epochs of apnea-hypopnea events. Considering that the data we used did not include normal people's data, we extracted epochs without apnea-hypopneas in 30 subjects as normal epochs. Considering sample balance and independence from subjects, we constructed the training and the test sets according to the ratio of 75:25 and ensured that 25% of the test data has never been exposed to the training model. The detailed data set information is shown in Table 1.

The polysomnography data used in this study is the previous archived data that erased the patient's personal biological information and will not cause any harm to the patient's interests. After reviewed by the ethics committee of the author's unit, it met the conditions of ethics exemption, and the patient did not need to sign an informed consent form.

B. PROPOSED METHOD

The detailed steps of the proposed method are shown in Figure 2. Firstly, the original EEG signal data is preprocessed, and decomposed using different Butterworth filters to obtain Delta, Theta, Alpha, Beta, and Gamma waves. Then, discrete wavelet transform is performed on the five seeded EEG signals. Finally, the approximate entropy of the high-frequency coefficients of different layers after each wavelet transform is taken as the feature point of the final classifier.

C. RAW SIGNAL PREPROCESSING AND BAND-LIMITED SIGNAL EXTRACTION

During nighttime sleep, the EEG signals will change significantly as people enter different sleep stages. In order to eliminate the obvious and fluctuating changes between the different EEG data frames, the DC offset of an EEG frame is eliminated by subtracting the mean value of that frame from each sample value. After removing DC, frame amplitude normalization is required to remove the undesirable fluctuation in amplitude occurring in different EEG frames belonging to the same class. For this purpose, after mean value subtraction, normalization of sample values is accomplished with reference to the maximum and the minimize sample values of that frame, which also called min-max normalization, that is, each value of the input EEG data frame minus the minimum

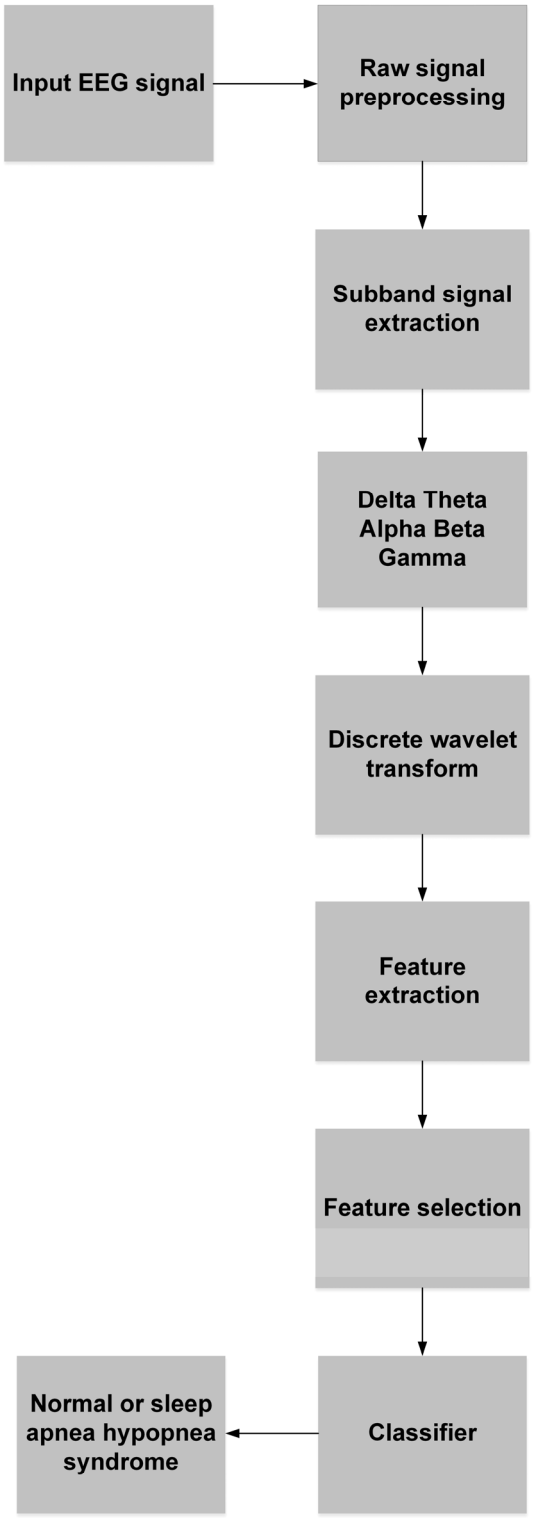

FIGURE 2. Block diagram representing the major steps involved in the proposed method.

value of the frame and the obtained value is divided by the difference between the maximum and minimum values of the frame.

The obtained preprocessed EEG data can be divided into: Delta (0.25-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (12-16 Hz) and Gamma (16-40 Hz). In the proposed method, five band-pass filters are used to extract the band limited EEG signals, which are expected to preserve the local information better than the full band signal.

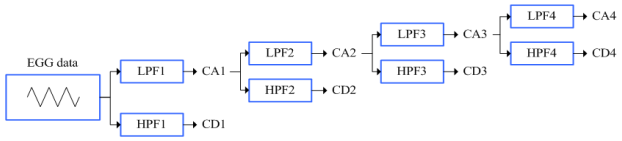

FIGURE 3. Decomposition step diagram of discrete wavelet transform.

D. FEATURE EXTRACTION

The proposed method uses the discrete wavelet transform to perform time-frequency analysis on five different band-limited signals obtained in the previous step. The discrete wavelet transform of a numerical sequence x(t) in any space $L^2(R)$ can be defined as:

$$W_x(j,k) = \int_R x(t)\overline{\psi_{j,k}(t)}\,dt \tag{1}$$

The wavelet prototype expression is defined as:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t}{2^j} - k\right) \tag{2}$$

where ψ is a wavelet prototype. The wavelets are localized in frequency (scale parameter j) and time (shift parameter k) using the preprocessed signal through two complementary filters (i.e. high and low frequency filters). The outputs from the low and high pass filters are referred to as approximation (A1) and detail (D1) coefficients of the first level. The A1 and D1 signals represent the approximate and the detail values of the decomposition signal obtained for the first time, respectively. The approximate value is a coefficient generated by a large scaling factor and represents the low frequency component of the signal. The detail value is a coefficient generated by a small scaling factor and represents the high frequency component of the signal. In the next decomposition, the low-frequency component obtained last time is always decomposed in the same step as the previous decomposition step. By analogy, in this study, the db3 mother wavelet and four-layer decomposition are used on the five different sub-band waves. The detailed decomposition process is shown in Figure 3.

The five sub-band waves corresponding to 609 normal epochs and 609 apnea-hypopnea epochs in the training set perform db3 and four-layer decomposition discrete wavelet transform using C3 channel. Next, 65 absolute values of high frequency coefficients of the second layer of the five-segmented wave of with and without apnea-hypopnea events are randomly selected, which from all training samples take the average value of the second layer of high frequency coefficient. It is found that although the detail coefficient of the gamma wave is not clearly distinguishable from the

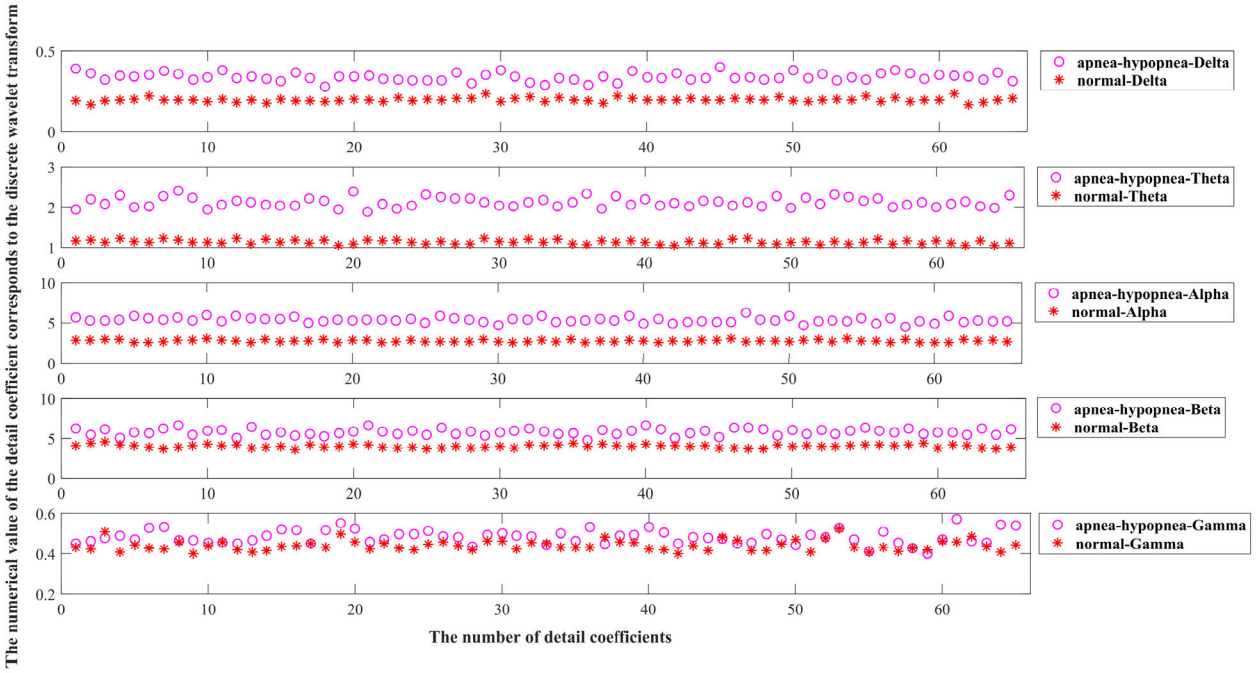

FIGURE 4. The second layer of absolute value of high frequency coefficients of the different five-segment belt EEG signals corresponding to normal and apnea-hypopnea events (C3 channel).

scatter plot, similar changes also occurred in the third and fourth layers. The detailed changes are shown in Figure 4.

Therefore, the proposed method in this article further extracts the features of absolute values of high frequency coefficients to represent the occurrence of apnea-hypopnea events. The ApEn was developed by Pincus as a measure of regularity to quantify the levels of complexity within a time series [26]. It uses a non-negative number to represent the complexity of a time series and reflects the possibility of new information in the time series. The algorithm is defined as follows:

1. Sequence {u(i)} into m-dimensional vector x(i):

$$x(i) = \{u(i), u(i+1), \ldots u(i+m-1)\},$$
$$i = 1,2,\ldots N-m+1$$

2. Calculate the distance between the vector x(i) of each i value and the rest of the vectors x(j), ($i = 0,1,2,\ldots N-m+1, j \neq i$), distance is defined as:

$$d_{ij} = \max |u(i+k) - u(j+k)|, \quad k = 0, 1, \ldots m-1$$

3. For each value of i, the number of $d_{ij} < r$, (given threshold $r > 0$) and the ratio of this number to the total number of vectors $N-m+1$ are recorded as:

$$C_i^m(r) = \{\text{Number of } d_{ij} < r\}/(N-m+1),$$
$$i = 1 \sim N-m+1$$

In the sense of similar tolerance r,
$C_i^m(r)$ reflects the probability that the m-dimensional patterns in the sequence are similar to each other.

4. Take the natural logarithm of $C_i^m(r)$, then average all i, and record it as $\phi^m(r)$:

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r)$$

$\phi^m(r)$ indicates the average closeness of some vectors.

5. For m+1, repeat steps 1-4 to find $C_i^{m+1}(r)$ and $\phi^{m+1}(r)$.
6. The approximate entropy is defined as:

$$ApEn(m, r, N) = \lim_{N\to\infty} [\phi^m(r) - \phi^{m+1}(r)]$$

Usually, N is a finite value, so the approximate entropy is expressed by the following formula:

$$ApEn(m, r, N) = \phi^m(r) - \phi^{m+1}(r)$$

This article uses the approximate entropy of absolute value of high frequency coefficient of a certain layer as a feature point of apnea-hypopnea events. The length of the comparison vector m is selected as 2, the measure of similarity r selects 0.2 times the standard deviation of the original sequence, and N represents the dimension of the input sequence{u(i)}. Considering the effect of the inconsistency of the points of different layers on the approximate entropy result, the average value of entropy is obtained and the data of

TABLE 2. TP TN FP FN parameter meaning.

Detection	Apnea-hypopnea	Normal
Apnea-hypopnea	True Positive (TP)	False Negative (FN)
Normal	False Positive (FP)	True Negative (TN)

TABLE 3. Recursive elimination feature filtering results.

channel	The fourth layer high frequency coefficient approximate entropy	The third layer high frequency coefficient approximate entropy	The second layer high frequency coefficient approximate entropy
C3-A2	Theta Beta Alpha	Gamma Beta Alpha Theta Delta	Gamma Beta Theta
C4-A1	Theta Gamma Beta Alpha Delta	Beta Delta Theta	Beta Gamma Theta
C3 and C4	C4 all sub-band signals and C3-Theta C3-Beta C3-Alpha	C3-Beta C3-Theta C3-Gamma C4-Alpha C3-Alpha C4-Beta	All sub-band signals

different channels are compared with different layers to find the best result.

E. FEATURE SELECTION

The SVM-RFE is a sequence backward selection algorithm based on the maximum separation principle of SVM. The algorithm uses all the features the first time and rejects consecutively the less relevant features by sorting the absolute values of the entries w_i [27]. The rule for removing the features is to first pass the SVM model training samples, then sort the score of each feature and remove the feature with the lowest score. Next, the remaining features are used to train the model again for the next iteration. Finally, the optimal feature list is output in descending order of score. The ranking score of feature A is defined as:

$$C_i = w_i^2 \tag{3}$$

The working algorithm of SVM-RFE is as follows:

1. Input: training samples $\{(x_i,y_i)\}_{i=1}^{N}$, $y_i \in \{+1, -1\}$.
2. Output: feature ranking set R.
3. Initialize the original feature set S = {1,2,…,D}, and feature ranking set $R = [\ ]$.
4. Repeat the following process until the original feature set S = [] and obtain the training samples with candidate feature sets.
5. Use formula:

$$\min \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{N}\alpha_i$$

   train the SVM classifier to get w.
6. Use the formula $C_i = w_i^2$, k = 1, 2, …, |S| to calculate the ranking criteria score.
7. Find the feature with the lowest ranking score:

$$P = \text{argmin}(\min_k C_k)$$

8. Update the feature set R = [P,R].
9. Eliminate the feature with the lowest ranking criterion, S = S(1:P − 1, P + 1 : length(s)).

In order to ensure that the best detection result is found when the feature difference is the most obvious, this study uses the SVM-RFE feature screening method to screen the features extracted from the C3, C4 and C3 and C4 channels. The results of recursive elimination show that different features extracted for different channels contribute differently to the classification. In order to obtain satisfactory results, the features that have small effect or have side effects should be deleted and certain features within the allowable range of the error should be filtered. However, it also appears that each feature extracted has a large contribution rate to the classification result. The detailed results of recursive elimination are shown in Table 3.

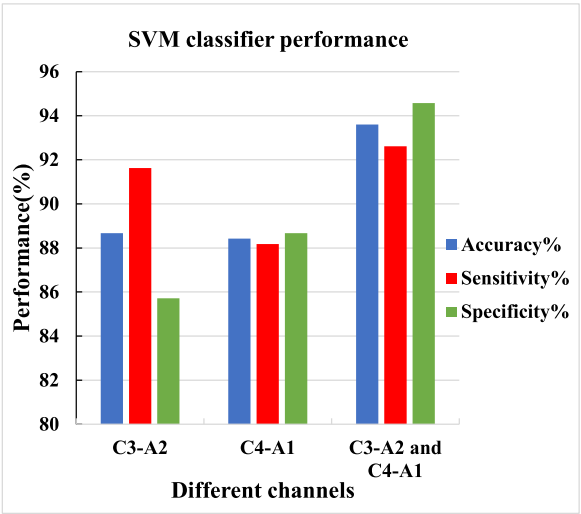

FIGURE 5. The best results obtained for each channel use SVM classifier.

TABLE 4. The statistical results of the classification results corresponding to the high-frequency coefficients of different layers of each channel (Me is Mean; Std is Standard deviation; Iqr is interquartile range).

channel	Me			Std			Iqr		
	Sens.	Spe.	Acc	Sens.	Spe.	Acc	Sens.	Spe.	Acc
C3	83.58	84.89	84.23	11.86	7.66	3.96	16.26	11.45	5.73
C4	81.44	87.03	84.23	7.47	1.50	4.45	11.08	2.22	6.65
C3+C4	88.01	90.64	89.32	6.33	7.69	6.97	8.86	10.35	9.24

IV. RESULTS AND DISCUSSION

A. RESULTS

In order to verify the effectiveness of the model, in addition to the accuracy rate, the indicators such as sensitivity, specificity, and recall rate calculated by the TP, TN, FP and FN parameters are introduced. The parameters are described in Table 2 and the calculation formulas are defined as:

$$\text{accuracy} = \frac{TR + TN}{TP + TN + FP + FN} \tag{4}$$

$$\text{sensitivity} = \frac{TP}{TP + FN} \tag{5}$$

$$\text{specificity} = \frac{TN}{TN + FP} \tag{6}$$

TABLE 5. Results of different classification models with different channels (Ens is Ensembles for Boosting).

Classifier	C3-A2 Sensitivity,Specificity,Accuracy %			C4-A1 Sensitivity,Specificity,Accuracy %			C3-A2 and C4-A1 Sensitivity,Specificity,Accuracy %		
KNN	91.63	85.22	88.42	87.19	88.67	87.93	95.07	92.61	93.84
Ens	92.12	85.22	88.67	94.58	81.77	88.17	90.15	98.03	94.08
Random forest	92.61	84.73	88.67	93.60	83.25	88.42	94.09	94.58	94.33
Mean	92.12	85.06	88.59	91.79	84.56	88.17	93.10	95.07	94.08
Standard deviation	0.49	0.28	0.14	4.01	3.63	0.25	2.60	2.74	0.25
Interquartile range	0.74	0.37	0.19	5.54	5.18	0.37	3.69	4.07	0.37

In the process of building the model, 75% of the training data is used to train the model and the test result of 25% of the test data is used as the final model evaluation indicator. All the classification work was done in MATLAB (R2016 a). According to the feature sequence corresponding to the feature ranking list obtained through the recursive elimination as the input to the classification model (the SVM model first evaluates the classification results by default), different channels corresponding to different layers are discussed and the numbers and the ratios of all training and test sets are strictly in accordance with the structure of Table 1.

In order to eliminate the dimensional influence between indicators, before entering the classification model, both the training and the test data are separately normalized in the same manner. The purpose of this is to avoid the test set being affected by the training set and making the model's predictive ability inaccurate. The normalization approach is also min-max normalization, which is same as the original signal preprocessing method. The best results obtained for each channel are shown in Figure 5 and the average value and the standard error are shown in Table 4.

Different sub-bands corresponding to different channels have different capabilities in detecting the apnea-hypopnea events. When the C3-A2 and the C4-A1 channels are used to detect the events at the same time, the approximate entropy of the second layer of high-frequency coefficients achieves the best results. In addition, when single channel C3-A2 or C4-A1 acts alone, the approximate entropy of the high-frequency coefficients of the second layer achieves the best effect. However, in distinguishing between normal and apnea-hypopnea events, the performance of each channel is different. Such as there is a gap between the accuracy of normal and apnea events that leads to a large range of statistical indicators. Besides, only using SVM algorithm has limitations. Therefore, other classification models in machine learning are selected for verification. The features of each channel that achieves the best results are considered. The detailed results are shown in Table 5.

Analysis of the above results can draw conclusions, that is, after being screened by the SVM-RFE algorithm, when the two channels C3-A2 and C4-A1 work together, the approximate entropy of the corresponding second-layer high-frequency coefficients has the best distinguishing effect. The parameters selected by the KNN model are K is five, distance is euclidean, rule is nearest. The parameters selected by the Ens model are number of tree is 70, integration method is AdaBoostM1. The parameters selected by the RF model are Treebagger way and number of tree is ten. In order to verify the influence of the training set and test set obtained by different division ratios on the model evaluation index and the reliability of the proposed algorithm. The approximate entropy characteristics of the second-layer high-frequency coefficients corresponding to the two channels of C3-A2 and C4-A1 are used as the research object. Select 5:5 and 8:2 respectively which belong to the classic division ratio of the hold-out method divide training set and test set. At the same time, follow the OSA: MSA: CSA=6:3:1 ratio combination to construct the apnea hypopnea data set. The detailed results and model parameters of three models namely RF, KNN and Ens are shown in Table 6 and Table 7.

TABLE 6. Results of different classification models with different channels (the ratio is 5:5, Ens is Ensembles for Boosting, KNN: 4, euclidean, nearest; RF: TreeBagger, nTree=30; Ens: AdaBoostM1, 100, tree).

Classifier	Sensitivity %	Specificity %	Accuracy %
KNN	91.63	97.29	94.46
Random forest	93.60	92.86	93.23
Ensembles for Boosting	90.89	97.78	94.34
Mean	92.04	95.98	94.01
Standard deviation	1.40	2.71	0.68
Interquartile range	2.03	3.69	0.92

TABLE 7. Results of different classification models with different channels (the ratio is 8:2, Ens is Ensembles for Boosting, KNN: 5, euclidean, nearest; RF: TreeBagger, nTree=10; Ens: AdaBoostM1, 120, tree).

Classifier	Sensitivity %	Specificity %	Accuracy %
KNN	91.45	99.38	94.69
Random forest	92.31	99.38	95.20
Ensembles for Boosting	91.45	98.77	94.44
Mean	91.74	99.17	94.77
Standard deviation	0.50	0.35	0.39
Interquartile range	0.65	0.46	0.57

The above is the judgment of sleep apnea hypopnea syndrome, which put together for testing without specific separation of patients of different degrees. Next, we test the

TABLE 8. Classification results and statistical indicators of different classifiers of sleep apnea hypopnea syndrome in mild patients. (KNN:5,euclidean, nearest; RF:nTree=40, Treebagger; Ens:LogitBoost, 350, tree).

Classifier	Sensitivity %	Specificity %	Accuracy %
KNN	86.54	100.00	93.13
Random forest	88.46	96.00	92.15
Ensembles for Boosting	94.23	90.00	92.16
Mean	89.74	95.33	92.48
Standard deviation	4.00	5.03	0.56
Interquartile range	5.80	7.50	0.74

TABLE 9. Classification results and statistical indicators of different classifiers of sleep apnea hypopnea syndrome in moderate patients. (KNN:5, euclidean, nearest; RF:nTree=20, Treebagger; Ens:Bag, 85, tree).

Classifier	Sensitivity %	Specificity %	Accuracy %
KNN	92.86	89.60	91.23
Random forest	94.44	88.80	91.63
Ensembles for Bagging	96.03	86.40	91.24
Mean	94.44	88.26	91.37
Standard deviation	1.59	1.67	0.23
Interquartile range	2.38	2.40	0.30

TABLE 10. Classification results and statistical indicators of different classifiers of sleep apnea hypopnea syndrome in severe patients. (KNN:4, cityblock, nearest; RF:nTree=23, Treebagger; Ens:LogitBoost, 400, tree).

Classifier	Sensitivity %	Specificity %	Accuracy %
KNN	94.03	96.50	95.26
Random forest	93.52	98.00	95.76
Ensembles for Boosting	93.53	97.00	95.24
Mean	93.69	97.16	95.42
Standard deviation	0.29	0.76	0.29
Interquartile range	0.38	1.13	0.39

proposed method of detecting sleep apnea-hypopnea syndrome in this article for events of different degrees and tests from mild, moderate, and severe patients. We selected four patients with mild sleep apnea, including three males and one female with AHI indexes of 9.7, 12.4, 8.2, and 7.1, respectively. In the original data of these four patients, 150 normal breathing events, 75 hypopnea events, 50 obstructive apneas, 12 central apneas and 8 mixed apneas were selected as the training sample dataset and 50 normal events, 23 hypopnea events, 25 obstructive apneas, and 1 central apnea as the test sample dataset.

According to the results obtained above, it is concluded that when the C3-A2 and the C4-A1 channels work together, the corresponding second-layer high-frequency coefficient discrimination results are the best. Therefore, we discussed the discrimination result of the approximate entropy of the corresponding second-layer high-frequency coefficients when the two channels work together. The detailed test results and the statistical indicators are shown in Table 8.

We selected seven patients with moderate sleep apnea and also discussed the detection results of different classifiers corresponding to the approximate entropy of the second-layer high-frequency coefficients of C3-A2 and C4-A1 channels. The seven patients include five males and two females with AHI indexes of 18.3, 19.8, 20.9, 26.7, 21.7, 27.3, and 16.7 respectively. In the original data of these seven patients, 375 normal breathing events, 175 hypopnea events, 100 obstructive apneas, 80 central apneas and 20 mixed apneas were selected as the training sample dataset and 125 normal events, 74 hypopnea events, 30 obstructive apneas, 2 central apnea and 17 mixed apneas as the test sample dataset. The detailed classification results and the statistical indicators are shown in Table 9.

Similarly, the data of 8 severe patients were selected, including five males and three females with AHI indexes of 49.3, 68.9, 66.6, 45.3, 50.9, 68.9, 39.1 and 59.3 respectively. In the original data of these eight patients, 600 normal breathing events, 200 hypopnea events, 200 obstructive apneas, 100 central apneas and 100 mixed apneas were selected as the training sample dataset and 200 normal events, 50 hypopnea events, 50 obstructive apneas, 50 central apnea and 50 mixed apneas as the test sample dataset. The approximate entropy of the high-frequency coefficients of the second layer of the two channels C3-A2 and C4-A1 is discussed and the detailed classification results and the statistical indicators are shown in Table 10.

B. CONCLUSION

Many studies have applied EEG signals to detect sleep apnea. After denoising and wavelet smoothing preprocessing of the EEG signal that is undergoing sleep apnea, each EEG signal frame of 30 seconds is divided into 10 seconds sub-frames and delta band power ratio is extracted as a feature from these sub-frames. These extracted features are used as input to classifiers and accuracies of 84.07% and 84.83% are achieved when KNN and SVM are used, respectively [19]. The entropy is extracted from the multi-band EEG signal as a feature and the Geometric Separability Index (GSI) method is used to score and analyze the features. Based on the KNN algorithm for sleep apnea classification, the accuracy is 87.64% [20]. Discrete wavelet transform and Hilbert transform are combined to extract features of EEG signals that occur in sleep apnea, the selected 30 features were used in the automatic classification of normal breathing and obstructive (OSA) and central (CSA) apnea by a feedforward neural network, The classifier returned the accuracy of 73.9% for the training and 77.3% for the testing set [23]. The given EEG signal test frame is divided into overlapping subframes. Each sub-frame is preprocessed and spectral characteristics are extracted from each pre-processed sub-frame. The statistical features are extracted from the temporal pattern of Beta band energy and used in K nearest neighborhood classifier. The results obtained are accuracy of 82.28%, specificity of 77.72% and sensitivity of 90.58% [24]. Based on the Empirical Mode Decomposition of wavelet reconstructed delta wave of EEG

signal, begins with wavelet transforming an EEG frame and reconstructing the low frequency delta wave from the approximate coefficients, EMD is performed on the reconstructed delta wave to generate intrinsic mode functions. The mean rate of variation and the variance in the first five IMFs of the reconstructed delta wave are extracted as features from each frame. The SVM classifier returned accuracy, sensitivity and specificity rate of 80.43%, 85.59% and 77.87%, respectively [25].

In this paper, we perform discrete wavelet transform based on different sub-band signals of preprocessed EEG signals that are occurring apnea-hypopnea events. The obtained approximate entropy of the high-frequency coefficients of different layers is used as a feature to detect normal breathing and sleep apnea-hypopnea events. The highest classification results are accuracy, sensitivity and specificity rate of 94.33%, 93.10% and 95.07%, respectively, which are higher results compared with the methods used in the previous studies. In addition, the method proposed in this paper has certain clinical significance in the diagnosis of apnea-hypopnea patients.

REFERENCES

[1] T. Schluter and S. Conrad, "An approach for automatic sleep stage scoring and apnea-hypopnea detection," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2010, pp. 1–5.

[2] N. Carroll and M. A. Branthwaite, "Control of nocturnal hypoventilation by nasal intermittent positive pressure ventilation," *Thorax*, vol. 43, no. 5, pp. 349–353, 1988.

[3] C. Guilleminault, G. Kurland, R. Winkle, and L. E. Miles, "Severe kyphoscoliosis, breathing, and sleep," *Chest*, vol. 79, no. 6, pp. 626–630, 1981.

[4] J. Catterall, N. Douglas, C. Calverley, C. Shapiro, V. Brezinova, and H. Brash, "Transient hypoxemia during sleep in chronic obstructive pulmonary disease is not a sleep apnea syndrome," am rev respir dis, vol. 128, no. 1, pp. 24–29, 1983.

[5] T. H. Bullock and R. B. Cowles, "How do brains work?" *Science*, vol. 115, no. 2994, pp. 541–543, 1982.

[6] H. Abdullah, N. C. Maddage, and I. D. Cosic Cvetkovic, "Cross-correlation of eeg frequency bands and heart rate variability for sleep apnoea classification," *Med. Biol. Eng. Comput.*, vol. 48, no. 12, pp. 1261–1269, 2010.

[7] M. Cabrero-Canosa, "An intelligent system for the detection and interpretation of sleep apneas," *Expert Syst Appl.*, vol. 24, no. 4, pp. 335–349, 2003.

[8] J. Zhou, X.-M. Wu, and W.-J. Zeng, "Automatic detection of sleep apnea based on EEG detrended fluctuation analysis and support vector machine," *J. Clin. Monitor. comput.*, vol. 29, no. 6, pp. 767–772, 2015.

[9] S. Taran, V. Bajaj, and D. Sharma, "Robust Hermite decomposition algorithm for classification of sleep apnea eeg signals," *Electron. Letters.*, vol. 53, no. 17, pp. 1182–1184, 2017.

[10] R. Lin, R. G. Lee, C. L. Tseng, H. K. Zhou, C. F. Chao, and J. A. Jiang, "A new approach for identifying sleep apnea syndrome using wavelet transform and neural networks," *Biomed. Eng.*, vol. 18, no. 3, pp. 138–143, 2006.

[11] R. Lin, M.-F. Yeh, R.-G. Lee, and C.-L. Tseng, "Sleep apnea syndrome recognition using the GreyART network," in *Proc. Int. Conf. Electric Inf. Control Eng.*, Apr. 2011, pp. 1–8.

[12] V. Vimala, K. Ramar, and M. Ettappan, "An intelligent sleep apnea classification system based on eeg signals," *J. Med. Syst.*, vol. 43, no. 2, p. 36, Feb. 2019.

[13] L. Chen, X. Zhang, and H. Wang, "An obstructive sleep apnea detection approach using kernel density classification based on single-lead electrocardiogram," *J. Med. Syst.*, vol. 39, no. 5, pp. 1–11, 2015.

[14] N. Banluesombatkul, T. Rakthanmanon, and T. Wilaiprasitporn, "Single channel ECG for obstructive sleep apnea severity detection using a deep learning approach," in *Proc. IEEE Region Conf.*, Jeju, South Korea, Oct. 2018, pp. 2011–2016, doi: 10.1109/TENCON.2018.8650429.

[15] D. álvarez, R. Hornero, J. V. Marcos, and F. del Campo, "Feature selection from nocturnal oximetry using genetic algorithms to assist in obstructive sleep apnoea diagnosis," *Med. Eng. Phys.*, vol. 34, no. 8, pp. 1049–1057, 2012.

[16] B. Xie, "Real-time sleep apnea detection by classifier combination," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 469–477, May 2012.

[17] P. Lakhan, A. Ditthapron, N. Banluesombatkul, and T. Wilaiprasitporn, "Deep neural networks with weighted averaged overnight airflow features for sleep apnea-hypopnea severity classification," in *Proc. IEEE Region Conf.*, Jeju, South Korea, Oct. 2018, pp. 0441–0445, doi: 10.1109/TENCON.2018.8650491.

[18] L. Hyoki, P. Jonguk, and L. Hojoong, "New rule-based algorithm for real-time detecting sleep apnea and hypopnea events using a nasal pressure signal," *J. Med. Syst.*, vol. 40, no. 12, p. 282, 2016

[19] C. Shahnaz, A. T. Minhaz, and S. T. Ahamed, "Sub-frame based apnea detection exploiting delta band power ratio extracted from EEG signals," in *Proc. IEEE Region Conf. (TENCON)*, Nov. 2016, pp. 190–193.

[20] S. Saha, A. Bhattacharjee, M. A. A. Ansary, and S. A. Fattah, "An approach for automatic sleep apnea detection based on entropy of multi-band EEG signal," in *Proc. IEEE Region Conf. (TENCON)*, Nov. 2016, pp. 420–423.

[21] Y. Maali and A. Al-Jumaily, "A novel partially connected cooperative parallel PSO-SVM algorithm: Study based on sleep apnea detection," in *Proc. IEEE Congr. Evol. Comput.*, Brisbane, QLD, Australia, Jun. 2012, pp. 1–8, doi: 10.1109/CEC.2012.6256138.

[22] B. L. Koley and D. Dey, "Classification of sleep apnea using cross wavelet transform," in *Proc. IEEE 1st Int. Conf. Condition Assessment Techn. Electr. Syst. (CATCON)*, Dec. 2013, pp. 275–280.

[23] M. A. Prucnal and A. G. Polak, "Analysis of features extracted from eeg epochs by discrete wavelet decomposition and Hilbert transform for sleep apnea detection," in *Proc. IEEE Eng. Med. Biol. Soc. Conf.*, Jul. 2018, pp. 287–290.

[24] F. Ahmed, P. Paromita, A. Bhattacharjee, S. Saha, S. Azad, and S. A. Fattah, "Detection of sleep apnea using sub-frame based temporal variation of energy in beta band in EEG," in *Proc. IEEE Int. WIE Conf. Electr. Comput. Eng. (WIECON-ECE)*, Dec. 2016, pp. 258–261.

[25] C. Shahnaz and A. T. Minhaz, "Sleep apnea frame detection based on empirical mode decomposition of delta wave extracted from wavelet of EEG signals," in *Proc. IEEE Int. WIE Conf. Electr. Comput. Eng. (WIECON-ECE)*, Dec. 2016, pp. 233–236.

[26] S. M. Pincus and W. M. Huang, "Approximate entropy: Statistical properties and applications," *Commun. Stats*, vol. 21, no. 11, pp. 3061–3077, 1992.

[27] A. R. Hidalgo-Muñoz, M. M. López, A. Galvao-Carmona, A. T. Pereira, I. M. Santos, M. Vázquez-Marrufo, and A. M. Tomé, "EEG study on affective valence elicited by novel and familiar pictures using ERD/ERS and SVM-RFE," *Med. Biol. Eng. Comput.*, vol. 52, no. 2, pp. 149–158, 2014.

YAO WANG was born in 1988. She received the bachelor's degree in biomedical engineering from the Department of Biological Sciences and Engineering, South China University of Technology, in 2011, and the Ph.D. degree in biomedical engineering from the Department of Medicine, Tsinghua University, in 2016. Since 2016, she has been with Tiangong University, as a Lecturer. Her main research interests include auditory signal processing and detection, brain–computer interface research based on auditory signals, and auditory and cognitive engineering research.

SIYU JI was born in 1997. She received the bachelor's degree in communication engineering from the Heilongjiang University of Science and Technology, in 2019. She is currently pursuing the master's degree in electronic and communication engineering with Tiangong University. Her current research interests include EEG signal processing and sleep apnea detection.

TIANSHUN YANG was born in 1996. He received the bachelor's degree in electronic information and engineering from the Anhui University of Technology, in 2019. He is currently pursuing the master's degree in information and communication engineering with Tiangong University. His current research interest includes automatic apnea detection.

XIAOHONG WANG was born in 1994. She received the bachelor's degree in communication engineering from the Shijiazhuang College, in 2018. She is currently pursuing the master's degree with the Tianjin University of Technology. Her current research interest includes automatic apnea detection.

HUIQUAN WANG was born in 1985. He received the bachelor's and Ph.D. degrees in biomedical engineering from the School of Precision Instruments and Optoelectronics Engineering, Tianjin University, the bachelor's degree in international finance from Nankai University, and the Ph.D. degree in instrument science and technology from Tianjin University. From 2011 to 2013, he was a Public Visiting Scholars study with Johns Hopkins University. He is currently the Director of the Department of Biomedical Engineering and the Master Instructor of Tiangong University. His main research interests include wearable medical testing equipment and intervention methods, near-infrared spectroscopy and big data mining algorithms, and multi-modal imaging technology.

XIAOYUN ZHAO was born in 1979. He received the bachelor's and master's degrees from the Medical College, Nankai University, Tianjin, China, in 2004, and the M.D. degree in respiratory medicine from Tianjin Medical University, Tianjin, China, in 2016. From 2004 to 2009, he was a Resident Physician with the Tianjin Chest Hospital. Since 2009, he has been an Attending Physician with the Department of Respiratory, Critical Care and Sleep Medicine (PCCSM), Tianjin Chest Hospital, and he has been the Vice Director and Chief Doctor of the RICU and Sleep Center, since 2013. He is the author of more than 40 articles and ten inventions. His research interests include central regulation mechanism of breathing, intermittent swing of pleural pressure during obstruction of airway, and comprehensive complications of sleep breathing disorders. He is also a Scientist and Tutor of graduate school in biomedical engineering from Department of Tianjin University, Tiangong University, and Tianjin Medical University, Tianjin, China. He is a member of American Thorax Society (ATS), China Sleep Research Association (CSRA), and several other medical and bioengineering associations.