Research Article

Detection and Quantification of Resting Tremor in Parkinson's Disease Using Long-Term Acceleration Data

Han Yuan [1], **Sen Liu** [1], **Jiali Liu,**[2,3] **Hai Lin,**[2,3] **Cuiwei Yang** [1,4] **Xiaodong Cai** [2,3] **Lepeng Zeng,**[5] **and Siman Li**[5]

[1]*Center for Biomedical Engineering, School of Information Science and Technology, Fudan University, Shanghai 200433, China*
[2]*Department of Neurosurgery, Shenzhen Second People's Hospital, The First Affiliated Hospital of Shenzhen University, Shenzhen 518035, China*
[3]*Shenzhen University School of Medicine, Shenzhen 518035, China*
[4]*Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention of Shanghai, Shanghai Engineering Research Center of Assistive Devices, Shanghai 200093, China*
[5]*Active Medical Devices Dept.BU, Lifetech Scientific (Shenzhen) Co., Ltd., Shenzhen 518040, China*

Correspondence should be addressed to Cuiwei Yang; yangcw@fudan.edu.cn and Xiaodong Cai; 13632660199@139.com

Long-term monitoring of resting tremor is key to assess the status of patients suffering from Parkinson's disease (PD), which is of vital importance for reasonable medication. The detection and quantification of resting tremor in reported works rely heavily on specified movements and are not appropriate for long-term monitoring in real-life condition. The purpose of this study is to develop a detection model for long-term monitoring of resting tremor and explore an effective indicator for tremor quantification. This study included long-term acceleration data from PD patients and proposed a resting tremor detection model based on machine learning classifiers and Synthetic Minority Oversampling Technique (SMOTE). Four machine learning classifiers, K-Nearest Neighbor (KNN), Random Forest (RF), Adaptive Boosting (AdaBoost), and Support Vector Machine (SVM), were compared. Furthermore, an indicator called tremor timing ratio (TTR) was defined and calculated for tremor quantification. The detection model with RF classifier achieved the highest overall accuracy of 94.81%. The sample entropy of the acceleration signal was proved most influential in the classification by exploring the feature importance. Through the Kruskal-Wallis test and the Mann-Whitney U test, the TTR had a strong correlation with the subscore of resting tremor in Unified Parkinson Disease Rating Scale (UPDRS). Such two-step evaluation process for resting tremor can detect the tremor effectively and is expected to be applied in long-term monitoring of PD patients in daily life to realize a more comprehensive assessment of PD.

1. Introduction

Parkinson's disease (PD) is a progressive, adult-onset neurological disease and gets more common with age, affecting more than doubled number of sufferers over the past generation [1]. Therapy for Parkinson's disease primarily focuses on ameliorating the symptoms with medication [2]. Detection and quantification of these symptoms are of vital importance for reasonable medication. Resting tremor is one of the most typical presenting signs of PD [3] and about 75%

of patients have symptoms of resting tremor [4]. It is an involuntary shakiness most noticeable in the hands, presents at rest, and disappears with intentional movement. The most popular tool to assess the severity of rest tremor in clinic is rating scale and the Unified Parkinson Disease Rating Scale (UPDRS) is widely accepted [5]. However, the motor examination in UPDRS is based on subjective observation of the clinicians trying to capture a snapshot of the outpatients. It is acknowledged that there is variance among the judgements from different clinicians, because of short time

observation and the diversity in experience. Thus, reliable, objective, and accessible diagnosing methods for resting tremor are urgently expected.

Related works have reported various computer-assisted techniques in the diagnosis and assessment of PD. As the mature voice acquisition equipment facilitates the construction of voice-based data, most study on PD focuses on dealing with speech processing [6]. Besides, approaches based on handwritten spiral images drawn by PD patients are also hotspots in recent research [7, 8]. Both voice-based method and drawing images require that the patients have no cognitive disorder and fail to work in case of the low compliance of patients. Local field potentials (LFP) were used to detect resting tremor in recent works [9, 10]. LFP requires the electrodes to be implanted to the patients during the deep brain stimulation (DBS) and directly deals with brain signal. However, it is difficult to popularize the LFP recordings for the surgical DBS treatment is invasive and costly. Other external body noninvasive signals such as electromyography (EMG) [11] and electroencephalography (EEG) [12] also have been applied to study resting tremor. Both EMG and EEG need fixed electrodes which are cumbersome to patients and not appropriate to long-term monitoring in daily life. In recent years, Leap Motion Controller (LMC), an interactive and noncontract device mainly used for hand gestures and finger position detection, has been introduced to the quantification of hand tremor [13]. Nevertheless, all the methods mentioned above are unsuited for long-term monitoring.

Smartphone and smart watch with motion sensors appeared in recent studies on monitoring the symptoms of PD [14]. A few researchers have employed inertial sensors [15, 16] to assess the symptoms of resting tremor. Lacy et al. [17] presented noninvasive electromagnetic sensors to conduct the finger tapping test in PD patients. However, most of the reported studies designed experiments in which the patients were required to do the assigned tasks to induce tremor and record signal for just a few seconds or a few minutes. Although these tasks make the resting tremor more distinctive and make the assessment more efficiently, the assigned tasks are different from the activities of daily living. In the patients' real life, tremor does not occur all the time and the amplitude of the tremor is always changing with time. It is challenging to collect the dataset in real-life condition and to distinguish the resting tremor from the activities of daily living.

In addition, multiple methods such as joint time-frequency analysis, statistical analysis, and machine learning have been employed to study the tremor so far. Spectral analysis and time-frequency analysis were well utilized in several previous studies [18, 19]. Salarian et al. [20] used frequency and amplitude of the signals to detect and quantify resting tremor through fixed thresholding. Manzanera et al. [21] compared several parametric and non-parametric spectral estimation methods for tremor detection but failed to detect the resting tremor accurately. Moreover, wavelet analysis such as discrete wavelet transform [22] and wavelet coherence analysis [23] provided an approach to analyze tremor signals and a framework called WAKE was proposed by using wavelet decomposition coupled with adaptive Kalman filtering to extract hand tremor [24]. Besides, statistical analyses such as correlation analysis [25] and coherence analysis [26] were also helpful methods for exploring the features corresponding to tremor signals. Deep learning techniques were well applied in the diagnosis of diseases. Zeng et al. proposed a SDPSO-SVM model for diagnosis of Alzheimer's disease [27] and promoted the diagnosis algorithm with deep belief network-based multi-task learning [28]. With advances of deep learning technique, a few researchers attempted to employ neural networks to diagnose or quantify tremor [29, 30]. Yohanandan et al. [31] compared the performance of linear regression models with machine learning classifiers and achieved tremor severity ratings on the Bain-Findley tremor rating scale. Most of the reported works so far focused on detecting and quantifying resting tremor based on fragments of signals. However, a long-term continuous automatic monitoring system for PD in the home-based environment is of great significance, which can help the assessment of the tremor severity.

Wearable devices are suitable for tracking the resting tremor in daily life and enable the automatic detection and quantification of resting tremor, making drug adjustment for patients more reasonable and timelier. In this work, a two-step evaluation process for resting tremor of PD patients based on long-term acceleration data is proposed. Consecutive long-term data of PD patients was collected in home settings and provides statistical information such as tremor prevalence. The tremor detection model with good generalization performance combines Synthetic Minority Oversampling Technique (SMOTE) with machine learning classifiers, providing a judgement on the status of the hand tremor. A new indicator called tremor timing ratio (TTR) is extracted for the tremor quantification based on the consecutive tremor detection and proved to be highly correlated with tremor severity.

2. Materials and Methods

2.1. Data Acquisition. To acquire the long-term acceleration data, a total of 20 patients suffering from PD were invited to participate in the experiment. Before data acquisition, explicit informed consent was obtained. The experiment was carried out in accordance with The Code of Ethics of the World Medical Association with prior approval of the Ethics Committee of Shenzhen Second People's Hospital. The patients were asked to attach the wearable devices with the triaxial accelerometer to both their wrists (see Figure 1). The wearable devices acted like watches and did not interference with the activities of subjects at all. The patients wore the wrist accelerometers consecutively and were allowed to move freely following their own will, same as their daily activities. The acceleration data was sampled at a frequency of 100 Hz and transmitted via Bluetooth wirelessly. The data receiver was exactly fixed in the room where the patients were. In the meantime, the whole process was video filmed using cameras. Several neurologists with rich clinical experience were invited to evaluate the tremor of patients
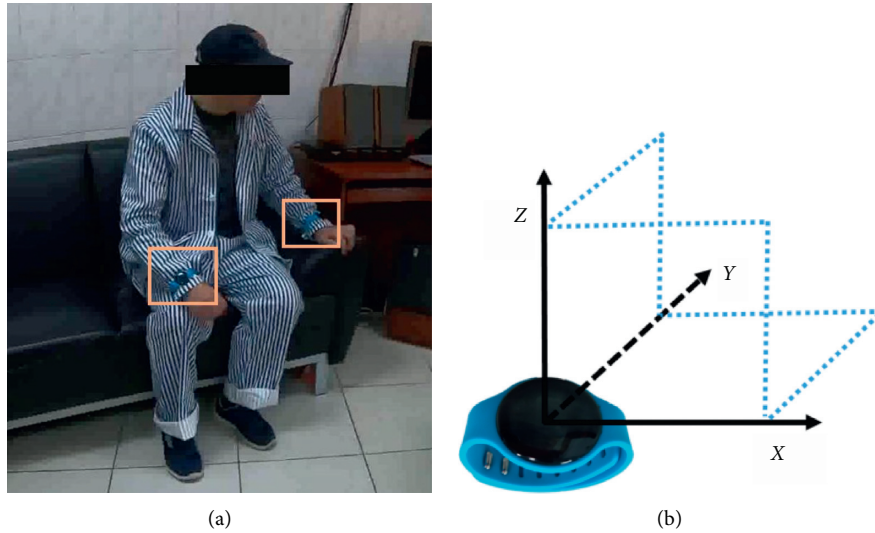
(a) (b)

FIGURE 1: Experimental environment of the data acquisition. The wearable devices with triaxial accelerometer used in this study were placed on both wrists of the patient (a). The wearable device was shown in (b).

according to UPDRS (see Table 1). Because the symptom of the resting tremor might change over time, the patients' scores were repeatedly examined in different periods of the experiment to be applied as the gold standard. The hand status of the patients was further annotated in detail to facilitate the data labelling with assistance of the synchronous video and the acceleration data was then labelled as tremor and nontremor accordingly. Since not all the patients suffered from resting tremor and the resting tremor disappeared after some tremor-dominant patients took the medicine, the mean tremor prevalence was 16.9%. Detailed information of patients and experiment duration is shown in Table 2.

2.2. Framework of Tremor Detection.

The aim of tremor detection was to track the hand status of the patients and to detect whether the resting tremor occurs according to the long-term acceleration data. To develop the resting tremor detection model, all the acceleration data collected was involved. The framework of the detection algorithm is shown in Figure 2, including three key parts: preprocessing, feature extraction, and classification. To meet the requirements of real-time and detection resolution, a 5-second time window was adopted, sliding on the acceleration signals with an incremental step of 1 second. The acceleration data was detected in each time window. Because the distance between the patients and the data receiver was changing as the patients moved and might be too far away from the data receiver, a small amount of data was missing. The sliding window firstly checked whether the acceleration data was missing according to timestamps. If there was no data missing in current time window, the tremor detection would proceed. Otherwise, the data in current window would be skipped and the detection would slide to the next window.

2.3. Preprocessing.

Before feature extraction, preprocessing of raw acceleration data was indispensable. The acceleration signal in the given time window was firstly bandpass filtered with zero-phase digital filter using Kaiser window, respectively. The offset of the triaxial acceleration signals was dependent on the position of the accelerometer and changed with the movements of the hands. Thus, components in extremely low frequency were to be eliminated. Besides, both intentional movement and involuntary resting tremor produced acceleration signals of relatively low frequency and were expected to be preserved after filtering. In this case the passband of the filter was limited to the range from 0.65 to 12.5 Hz. Since the frequency of the resting tremor mainly ranges from 4 to 7 Hz [32], the waveform of resting tremor signal was well preserved after filtering. The comparison of triaxial acceleration signal before filtering and after filtering is shown in Figure 3.

2.4. Feature Extraction.

After preprocessing, features which characterized the resting tremor were extracted. Related works on the characteristics of tremor have been reported, in which amplitude and frequency parameters were proved to be significantly different between the PD patients and healthy controls [33]. A number of characteristics in previous works offer multiple choices for feature extraction. Considering that excess features contain redundant information which may result in bad performance and slow down the detection speed of the model, five features in time domain and frequency domain were finally calculated and were presented in Table 3. Details and definition of the selected features were then clarified below.

In time domain, the feature characterizing the amplitude of signal was extracted. Root-mean-square values of the triaxial acceleration signals were firstly calculated as the total acceleration. The mean amplitude (mAmp) of the

TABLE 1: The evaluation of tremor at rest in head, upper, and lower extremities in motor examinations of UPDRS.

Score	Description
0	Absent.
1	Slight and infrequently present.
2	Mild in amplitude and persistent. Or moderate in amplitude, but only intermittently present.
3	Moderate in amplitude and present most of the time.
4	Marked in amplitude and present most of the time.

TABLE 2: The statistical review of the dataset.

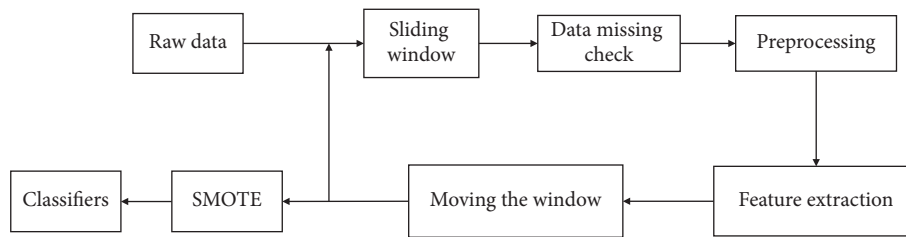| Characteristics	Observations
Number of patients	20
Male/female	15/5
Age (years)	60.7 ± 8.6
Disease duration (years)	5.9 ± 1.5
UPDRS-III before medication	52.7 ± 17.7
UPDRS-III after medication	24.9 ± 14.9
Hoehn and Yahr stage	2.9 ± 0.2
Recording duration (hours)	2.35 ± 1
Tremor prevalence	16.9%

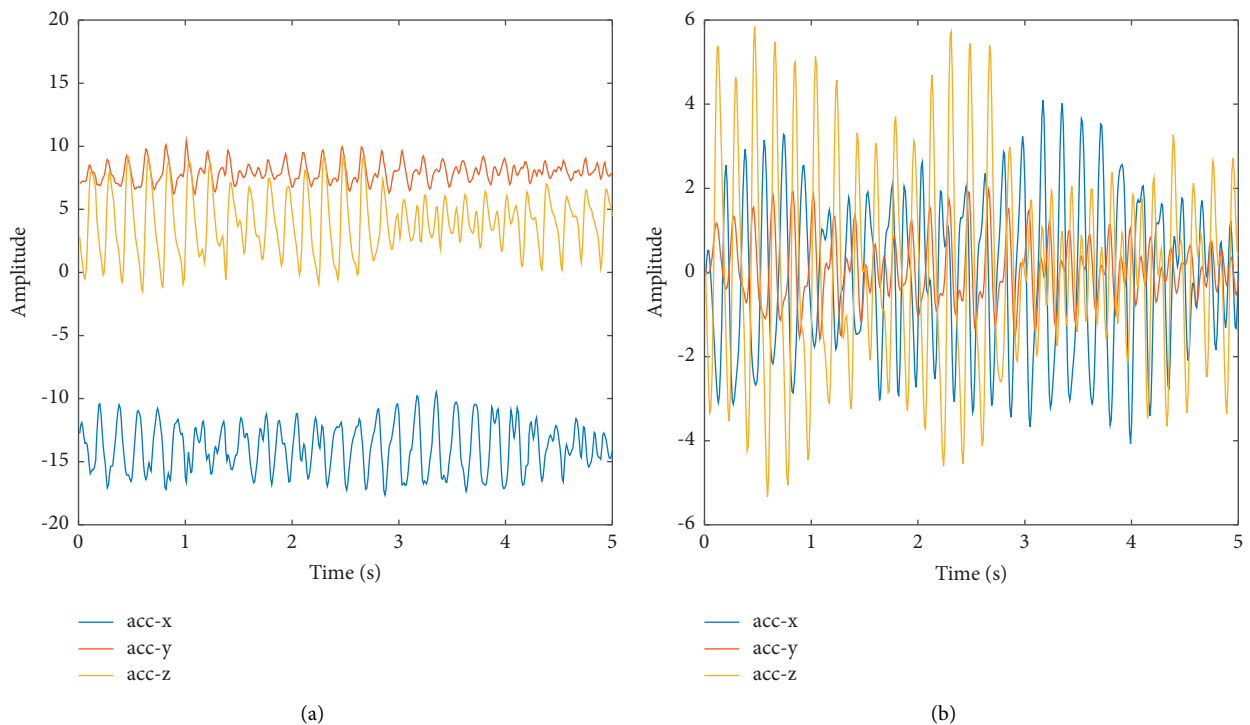FIGURE 2: The framework of the proposed tremor detection model.

(a)

(b)

FIGURE 3: Triaxial acceleration signal before filtering (a) and after filtering (b).

Feature	Symbol	Physical meanings
Mean amplitude	mAmp	Average amplitude of acceleration signal.
Sample entropy	SampEn	Complexity of acceleration signal, quantifying the regularity of acceleration signal.
Dominant frequency	DF	The frequency with highest amplitude in PSD.
Energy	E	The energy of acceleration signal in 5 seconds.
Spectrum concentration	SE	The characteristic of spectrum distribution, describing how frequency component concentrated around dominant frequency.

acceleration signal was computed based on the total acceleration. To calculate the mean amplitude of the acceleration signal, the envelope of the total acceleration was extracted. The mean amplitude of the signal in a time window is defined as the average of the difference between the upper envelope and the lower envelope, given by

$$mAmp = \frac{1}{N} \sum_{n=1}^{N} \left(env_{upper}(n) - env_{lower}(n) \right), \quad (1)$$

where $env_{upper}(n)$ is the upper envelope and $env_{lower}(n)$ is the lower envelope. An instance of the envelope extraction from the acceleration signal is illustrated in Figure 4.

Sample entropy (SampEn) was first introduced by Richman and Moorman [34]. It can be used to characterize the complexity of time series by measuring the probability of a new pattern being generated in the signal. Sample entropy of hand tremor was proved significantly different between the PD patients and healthy controls [35]. In this study, the sample entropy of the total acceleration signal was computed to distinguish the resting tremor status. For the total acceleration signal in a window $\{x(i)|1 < i < N\}$ of N data points, the SampEn is calculated by the following steps:

(1) The acceleration signal is first reconstructed to the delay vector:

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

where $i \leq n - N - m + 1$.

(2) The distance between $X_m(i)$ and $X_m(j)$ is defined as

$$d\left(X_m(i), X_m(j)\right) = \max_{0 \leq k \leq m-1} |x(i+k) - x(j+k)|. \quad (3)$$

(3) A_i is the number of the vector $X_m(i)$ such that $d\left(X_m(i), X_m(j)\right) \leq r$

$$A_i^m(r) = \frac{1}{N-m-1} A_i,$$
$$\quad (4)$$
$$A_m(r) = \sum_{i=1}^{N-m} \frac{A_i^m(r)}{(N-m)}.$$

(4) The SampEn is finally obtained by

$$SampEn(m, r, N) = \log A^m(r) - \log A^{m+1}(r). \quad (5)$$

Fast Fourier Transform (FFT) was applied to the triaxial acceleration signals, respectively, to further obtain the
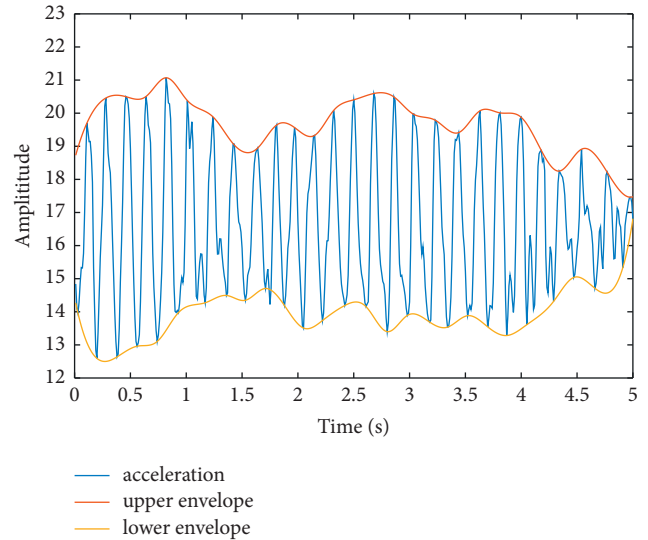

FIGURE 4: Envelope extraction from the total acceleration signal (in blue). The upper envelope (in red) and lower envelope (in yellow) were extracted to obtain the average amplitude of the acceleration signal.

features in frequency domain. The spectrums of the triaxial acceleration signals are obtained by

$$F_x(k) = \sum_{n=1}^{N} acc_x(n) e^{-2\pi jkn/N},$$
$$F_y(k) = \sum_{n=1}^{N} acc_y(n) e^{-2\pi jkn/N}, \quad (6)$$
$$F_z(k) = \sum_{n=1}^{N} acc_z(n) e^{-2\pi jkn/N},$$

where $acc_x(n)$, $acc_y(n)$, and $acc_z(n)$ are the triaxial acceleration signals.

The dominant frequency (DF) was estimated by searching the maximum value of the peaks in the frequency spectrums. Comparing the amplitude of the peaks in triaxial frequency spectrums, the frequency with highest amplitude was defined as the DF. The DF of the triaxial acceleration signals is estimated by

$$F\left(\frac{DF}{f_s}\right) = \max\left\{ F_x(k) \ F_y(k) \ F_z(k) \right\}, \quad (7)$$
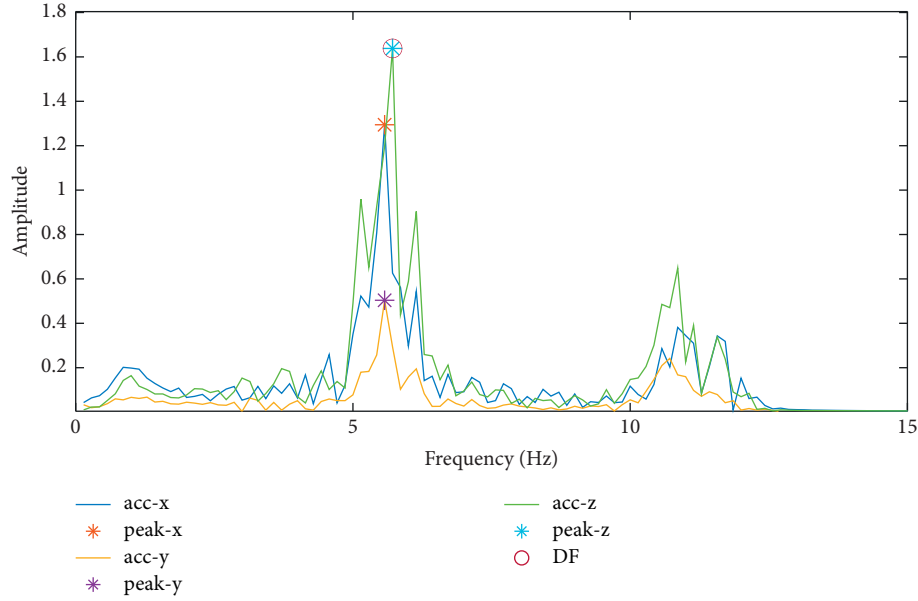
FIGURE 5: DF estimation from the acceleration signals. The triaxial acceleration data is first transformed into frequency domain. The highest peak in each axis is marked by asterisk. The frequency corresponding to the peak with highest amplitude (red circle) is the dominant frequency.

where f_s is the sampling rate. The DF estimation from a window of acceleration signal labelled as resting tremor is depicted in Figure 5.

The power spectral density (PSD) of triaxial acceleration signals was estimated. The energy (E) of the signal was calculated as the integral sum of the triaxial spectrums, written as

$$E = \int P_x(f)\,\mathrm{d}f + \int P_y(f)\,\mathrm{d}f + \int P_z(f)\,\mathrm{d}f, \qquad (8)$$

where $P_x(f)$, $P_y(f)$, and $P_z(f)$ are the PSD of the triaxial acceleration signals.

The spectrum distribution characterized the frequency components of the signal. The power around the DF to the total power described how the spectrum concentrated around the DF. Since the resting tremor was most obvious when the patients were at rest, the spectrum was supposed to be highly concentrated around the DF when resting tremor was present. In the frequency spectrum, the amplitude of the highest peak corresponding to the DF of resting tremor was much higher than other peaks. The ratio of the power in the range of 0.4 Hz around the DF to the total power was named as spectrum concentration (SC) and can be calculated as follows:

$$E_{\mathrm{DF}} = \int_{\mathrm{DF}-0.4}^{\mathrm{DF}+0.4} P_x(f)\,\mathrm{d}f + \int_{\mathrm{DF}-0.4}^{\mathrm{DF}+0.4} P_y(f)\,\mathrm{d}f$$

$$+ \int_{\mathrm{DF}-0.4}^{\mathrm{DF}+0.4} P_z(f)\,\mathrm{d}f, \qquad (9)$$

$$\mathrm{SC} = \frac{E_{\mathrm{DF}}}{E}.$$

It was summarized that each window of 5 s acceleration signal was characterized by five features. The dataset consisted of a total number of 225,066 samples of five features.

Since not all the patients suffered from resting tremor and the tremor may disappear after some tremor-dominant patients took the medicine, the samples of resting tremor only accounted for 16.9%.

2.5. Classification.
Before classification, the SMOTE algorithm was employed to balance the dataset. As the tremor prevalence was only 16.9%, the unbalanced dataset was challenging for the classification. Undersampling and oversampling are two strategies of resampling to solve the unbalanced dataset problem. Random undersampling may reduce the necessary information contained in the dataset and the random oversampling adopts the strategy of simply copying samples to increase the minority class of samples, easy to cause the model overfitting. Therefore, SMOTE, which was an algorithm improved based on the random oversampling algorithm, was employed to deal with the unbalanced dataset [36]. The basic idea of SMOTE is to analyze and simulate the samples of minority class and add the artificial simulated new samples to the dataset. After oversampling with SMOTE algorithm, the proportion of resting tremor samples increased to 50%.

To construct the detection model, four classifiers of machine learning having good classification performance [37, 38] and meeting the real-time requirements were opted to achieve the tremor detection, including K-Nearest Neighbor (KNN), Random Forest (RF), Adaptive Boosting (AdaBoost), and Support Vector Machine (SVM).

K-Nearest Neighbor was proposed by Cover and Hart [39], which was commonly used as a supervised learning method. KNN classifier finds the k training samples closest to the test sample in the training set according to the given distance metric and predicts the results based on the

information of the K-Nearest Neighbors. The selection of the value of k has a great influence on KNN learning model. Therefore, the grid-search method was utilized to find the optimal parameters. Due to the large quantity of the samples, only 10% of the samples were used to search for the optimal parameters.

Random Forest classifier combines a number of decision trees with randomly selected features into a forest. Each decision tree determines the class of the sample independently and the forest chooses the majority of the determination from the decision trees [40]. The RF classifier operates efficiently on large dataset and reserves high accuracy despite the large scale of data missing. Because of the random sampling, the training model has small variance and strong generalization capability.

Adaptive Boosting classifier was first proposed by Schapire [41], which is an ensemble aggregation classifier as well as Random Forest. The core idea is to train different weak classifiers for the same training set and then combine these weak classifiers to make it a stronger one. The adaptation is embodied in the strategy that the weights of the samples misclassified by the previous weak classifiers will be strengthened. Then the updated samples are to be used to train the next weak classifier again. In this work, the Ada-Boost classifier was implemented with the decision tree as the weak learner.

Support Vector Machine model finds hyperplanes which maximize the distance between the closest samples in different classes, hence the name Maximum Margin classifier [42]. SVM addresses the classification problem with small quantity of samples, nonlinear and high dimension data. Similarly, the grid-search method was applied to explore the optimal parameters for the SVM classifier.

2.6. Method of Tremor Quantification. To further evaluate the severity of tremor status, 8 patients with valid tremor data over a long period of time were selected in this experiment. The recorded resting tremor scores ranged from 0 to 3. A total of 33 cases with assigned resting tremor scores were collected, consisting of 9 cases scored as 1, 18 cases scored as 2, and 6 cases scored as 3. Each case was a 15-minute continuous recording of acceleration signal. The detail is shown in Table 4.

By exploiting the time-duration characteristics of different tremor scores, a new indicator was designed and calculated to quantify tremor severity. Tremor timing ratio (TTR) was defined as the ratio of the detected tremor samples (tremor duration) to all the samples (the total duration of tremor samples and nontremor samples) in consecutive fifteen minutes, which can be calculated as follows:

$$\text{TTR} = \frac{\text{samples of tremor}}{\text{samples of tremor} + \text{samples of nontremor}}. \quad (10)$$

The TTR was then used to perform statistical analysis with the rating scores for resting tremor. The Mann-Whitney U test and the Kruskal-Wallis test were adopted.

TABLE 4: Information of the patients for tremor quantification.

Age (years)	59.4 ± 7.0
Gender (F/M)	3/5
Tremor	
Score	1/2/3
NS1	5/5/3
NS2	9/18/6

NS1 = numbers of patients with corresponding tremor scores.
NS2 = numbers of data segments with corresponding tremor scores.

The Mann-Whitney U test or the Wilcoxon rank sum test is a nonparametric test for two populations when samples are independent [43]. The Kruskal-Wallis test is a nonparametric version of classical one-way Analysis of Variance and an extension of the Wilcoxon rank sum test to more than two groups [44]. The Kruskal-Wallis test is valid for data that has two or more groups. It compares the medians of the groups of data to determine whether the samples come from the same population or the populations having the same distribution.

3. Results

3.1. Performance Metrics. In this work, multiple machine learning methods were selected for resting tremor detection. To better assess the performance of the selected methods, the fivefold cross-validation was introduced to the assessment. The dataset was divided into fivefold randomly. In each iteration, onefold of the dataset was kept as the holdout set for testing while the remaining folds were used as the training set. The average accuracy of fivefold cross-validation was defined as the overall accuracy. The performance metrics that were employed to compare those methods include sensitivity, specificity, overall accuracy, and F1 score, which can be expressed as follows:

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%,$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\%,$$

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \times 100\%, \quad (11)$$

$$F1 = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FN} + \text{FP}},$$

where TP, TN, FP, and FN represent the number of true positive, true negative, false positive, and false negative, respectively.

As a measure of consistency, Kappa coefficient was also included as the performance metrics to measure the effectiveness of classification. For the classification problem, the Kappa coefficient reflects a more authentic evaluation than accuracy because the Kappa coefficient makes punishment for the unbalanced dataset. In this tremor detection model, the Kappa coefficient can be calculated as follows:

$$\text{Kappa} = \frac{\text{accuracy} - p_e}{1 - p_e},$$

$$p_e = \frac{(\text{TN} + \text{FN}) \times (\text{TN} + \text{FP}) + (\text{FN} + \text{TP}) \times (\text{FP} + \text{TP})}{(\text{TP} + \text{FN} + \text{TN} + \text{FP})^2}. \tag{12}$$

3.2. Result of Tremor Detection. In the classification of tremor status and nontremor status, a total of 373,894 samples were used in the fivefold cross-validation. The testing results for the selected four classifiers were listed in Table 5. The RF classifier outperformed other classifiers in all aspects and achieved an overall accuracy of 94.81%. The Kappa coefficient of the RF classifier reached 0.90 which implied that the classification results were almost perfectly consistent with the actual labels. The goal of the tremor detection was to distinguish the resting tremor from other hand status of patients. The high sensitivity of RF classifier means the model is able to recognize the tremor status effectively, which is also conductive to the early screens for tremor.

An instance for tremor detection model applied in one of the patients is shown in Figure 6. The model combining with RF classifier was tested on a segment of 60 s acceleration data from both hands. Compared with the gold standard, the error of the detection results was within 2 seconds, which demonstrates the good capability of the model to detect the resting tremor.

3.3. Result of Tremor Quantification. As shown in Table 6, the returned value of p for the Kruskal-Wallis test was 0.0034, which indicated that the Kruskal-Wallis test rejected the null hypothesis that all three data samples came from the same distribution at a 1% significance level. The boxplot (Figure 7) visually presents the summary statistics for each score. The values of p for the Mann-Whitney U test on score 1 and score 2, score 1 and score 3, and score 2 and score 3 were 0.022, 0.0048, and 0.026, respectively, which further verified that the TTR of each score were significantly different from each other. The test results demonstrate that the TTR shows a strong correlation with the subscore of resting tremor in UPDRS and has a prospect of applying in the quantification of resting tremor.

4. Discussion

4.1. Main Findings. This work aims to establish a two-step evaluation process for resting tremor of PD patients using long-term acceleration data. Previous work on tremor detection and quantification is mainly based on short segments of signal and heavily relied on manual signal segmentation [37]. The data cannot represent the activities of daily living and the methods are not suitable for continuous tremor detection. In this work, the long-term monitoring of PD patients mainly lies in two aspects. Firstly, the detection model distinguishes the tremor status from other daily movements. Secondly, based on the consecutive acceleration

recordings, TTR is extracted as the tremor quantification indicator. Statistical information on acceleration data over a continuous period is used to quantify the patient's changing status under continuous tracking. As the results show, the detection model identifies the resting tremor effectively with an accuracy of 94.81%. In comparison, Zhang et al. [45] developed both generic model and person-specific model using SVM classifier and achieved best accuracy of 80.7% and 85.9%. The detection model in this work got a satisfactory result compared to previous achievements. In addition, by testing the algorithm on a Dell XPS 8930 featuring an Intel(R) Core™ i7-8700 CPU @ 3.20 GHz with 8.00 GB RAM, an averaged runtime of the feature extraction is 0.068 seconds per data segment of 5 seconds. The runtime for model training of RF classifier and classification is 20.12 s for all the data. In terms of quantification, TTR assesses the tremor burden by the frequency of resting tremor occurrence, and it is shown to be significantly different between tremor scores. Thus, the proposed two-step evaluation process could detect the resting tremor timely and provide an automatic long-term observation of patients, offering a powerful support to the diagnosis and treatment of PD for clinicians cannot stay with the patient all the time. It is feasible to monitor the resting tremor of PD in daily life, to track the transition of the patients' status, and to realize the automatic quantification of resting tremor.

4.2. Influence of Preprocessing. In data preprocessing, the acceleration signal in the given time window was firstly bandpass filtered. The goal of the bandpass filtering was to remove noise such as gravity artifacts in low frequency. To verify the influence of preprocessing on the overall performance, the detection models were retrained using signals without preprocessing and the result is shown in Table 7. In the absence of preprocessing, the result of classification was significantly decreased compared to that of acceleration signals with preprocessing and the accuracy of RF classifier was 89.38% and reduced by 5.43%. The preprocessing is indispensable to perform before feature extraction.

4.3. Predictor Importance for Features. Among all the classifiers, RF classifier outperformed the others and showed balanced performance for different metrics. Therefore, the predictor importance of the features engaged in classification was estimated through RF classifier. The feature with high importance estimated through RF classifier contains the vital information which differentiates the resting tremor from other daily hand movements of PD patients. The estimated predictor importance was measured by permutation. The out-of-bag permuted predictor importance estimates of the five features in RF classifier are shown in Figure 8. The SampEn gains the largest value of predictor importance followed by the energy of the signal, which means the SampEn is most influential in identifying tremor. The result hints that the complexity of acceleration data measured by SampEn is sensitive to the characteristic of

TABLE 5: The classification results of the four machine learning methods.

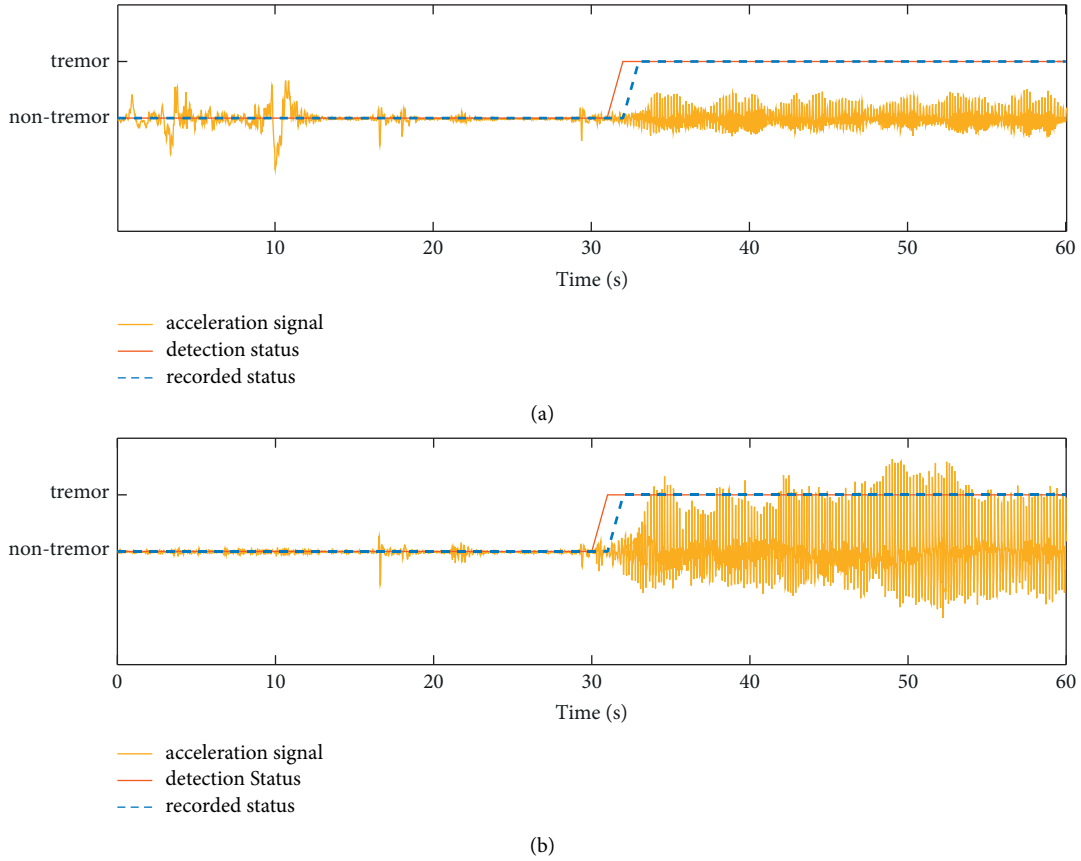| Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)	F1 score (%)	Kappa coefficient
KNN	93.16	93.33	93.24	93.24	0.87
RF	95.18	94.44	94.81	94.83	0.90
AdaBoost	93.09	92.17	92.63	92.67	0.85
SVM	93.59	86.38	89.98	90.33	0.80

(a)

(b)

FIGURE 6: The proposed tremor detection model applied in one of the patients. The detection results of left hand (a) and right hand (b) are depicted.

TABLE 6: The difference of the TTR from three tremor statuses.

	Score 1	Score 2	Score 3	
Score 1	0.46 (0.22)	—	—	
Score 2	0.022*	0.67 (0.14)	—	0.0034**
Score 3	0.0048*	0.026*	0.85 (0.15)	

Values are shown as the median (standard deviation). * p value of Mann-Whitney U test. ** p value of Kruskal-Wallis test.

FIGURE 7: The boxplot of the TTR from different scores.

resting tremor in PD patients. Other features also contributed to the tremor detection algorithm but were less influential compared to SampEn. The difference in feature importance reflects the ability of different features to discriminate between resting tremor and other daily activities. The information contained in the other features is less able to discriminate between tremor and other activities than SampEn and E.
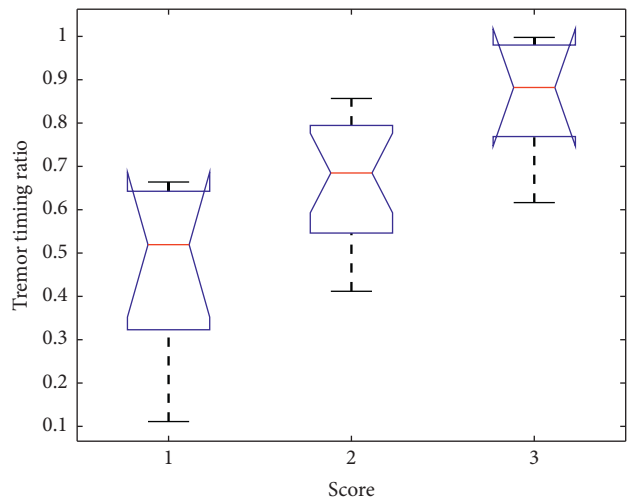
TABLE 7: The classification results of the four classifiers based on data without preprocessing.

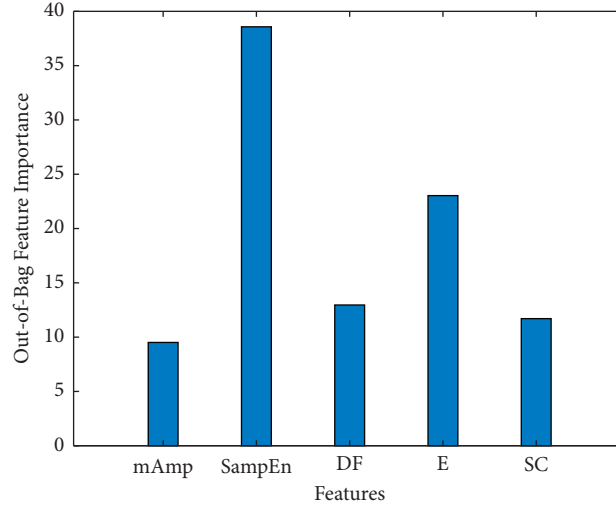| Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)	$F1$ score (%)	Kappa coefficient
KNN	88.53	89.04	88.78	88.75	0.78
RF	89.79	88.96	89.38	89.42	0.79
AdaBoost	88.29	84.51	86.40	86.65	0.73
SVM	72.10	90.72	81.41	79.50	0.63

FIGURE 8: The out-of-bag permuted predictor importance estimates of the features.

TABLE 8: The result for tremor detection under interpatient paradigm.

Patient	Samples		Accuracy (%)	Patient	Samples		Accuracy (%)
	Tremor	Nontremor			Tremor	Nontremor	
1	1883	9866	91.80	11	0	2899	98.72
2	6828	8129	53.16	12	21	5723	96.95
3	14	8763	88.20	13	0	5605	99.23
4	14763	3767	75.28	14	369	6050	95.17
5	2388	5703	74.70	15	9	10059	87.12
6	0	15686	99.63	16	2	15580	99.65
7	0	12814	99.84	17	0	6556	96.39
8	5989	7621	90.74	18	0	19252	98.54
9	179	6811	95.36	19	0	12943	99.53
10	791	5716	87.50	20	2317	3455	84.74

4.4. Tremor Detection under Interpatient Paradigm. In previous studies, two popular evaluation paradigms for analysis of signals in clinical medicine and the biological sciences are intrapatient paradigm and interpatient paradigm. In the intrapatient paradigm, the data in training set and testing set may come from the same patient. In the interpatient paradigm, the data from the same patient never appears in the training set and testing set at the same time. But in practice, the tremor detection model might come across unknown individuals, which requires the model with excellent generalization performance. To test the generalization performance of the resting tremor detection model, the acceleration data collected from 20 patients was again divided into training set and testing set according to different patients and the result is listed in Table 8. The detection model with RF classifier performs well for most of the patients. The weighted average accuracy according to the number of the samples for each patient was calculated as the average accuracy and the model still obtained a high average accuracy of 89.70%, slightly lower than the result of intrapatient paradigm, which proves the good generalization ability of the algorithm.

4.5. Limitation. The main limitation of this work is the data size. Though the number of the subjects is limited, the acceleration data was collected for long enough duration and samples of each subject are adequate. It is really a challenge to collect the long-term acceleration data of PD patients in real-life condition. The dataset needs to be extended to further verify the findings. Secondly, only three kinds of

resting tremor severity were quantified, lacking the score 4. It is because there were very few patients scored as 4 for resting tremor when they were on medication, and the patients diagnosed as 4 were rather severe and unable to participate in the experiment. Therefore, the difference of TTR between the cases scored as 4 and other scores was not validated. Further research is expected by collecting more cases scored as 4.

5. Conclusions

This work developed a two-step evaluation process for resting tremor of PD patients. The proposed tremor detection model in this work can distinguish the resting tremor effectively with good generalization performance. TTR for tremor quantification was proposed and showed significant difference in different scores, which was proved valid in quantifying the severity of the resting tremor. It is notable that the tremor detection model was based on the long-term acceleration data collected in a real-life condition, which made the model feasible to practical application. The proposed automatic two-step method is expected to be applied for real-time monitoring of patients and help to guide the medication for PD in daily life.

Data Availability

The processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authors' Contributions

Han Yuan and Sen Liu have contributed equally to this work.

Acknowledgments

References

[1] GBD 2016 Parkinson's Disease Collaborators, "Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 17, no. 11, pp. 939–953, 2018.

[2] G. K. Wenning, I. Litvan, and E. Tolosa, "Milestones in atypical and secondary Parkinsonisms," *Movement Disorders*, vol. 26, no. 6, pp. 1083–1095, 2011.

[3] C. Váradi, "Clinical features of Parkinson's disease: the evolution of critical symptoms," *Biology*, vol. 9, no. 5, pp. 103–115, 2020.

[4] H. Zach, M. Dirkx, B. R. Bloem, and R. C. Helmich, "The clinical evaluation of Parkinson's tremor," *Journal of Parkinson's Disease*, vol. 5, no. 3, pp. 471–474, 2015.

[5] J. Jankovic, "Parkinson's disease: clinical features and diagnosis," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008.

[6] B. Karan, S. S. Sahu, and K. Mahto, "Parkinson disease prediction using intrinsic mode function based features from speech signal," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 1, pp. 249–264, 2020.

[7] L. C. F. Ribeiro, L. C. S. Afonso, and J. P. Papa, "Bag of samplings for computer-assisted Parkinson's disease diagnosis based on recurrent neural networks," *Computers in Biology and Medicine*, vol. 115, Article ID 103477, 2019.

[8] S. Xu and Z. Pan, "A novel ensemble of random forest for assisting diagnosis of Parkinson's disease on small handwritten dynamics dataset," *International Journal of Medical Informatics*, vol. 144, Article ID 104283, 2020.

[9] C. Camara, P. Isasi, K. Warwick et al., "Resting tremor classification and detection in Parkinson's disease patients," *Biomedical Signal Processing and Control*, vol. 16, pp. 88–97, 2015.

[10] E. Bakstein, J. Burgess, K. Warwick, V. Ruiz, T. Aziz, and J. Stein, "Parkinsonian tremor identification with multiple local field potential feature classification," *Journal of Neuroscience Methods*, vol. 209, no. 2, pp. 320–330, 2012.

[11] A. G. Rabelo, L. P. Neves, A. P. S. Paixão et al., "Objective assessment of bradykinesia estimated from the wrist extension in older adults and patients with Parkinson's disease," *Annals of Biomedical Engineering*, vol. 45, no. 11, pp. 2614–2625, 2017.

[12] O. Sushkova, Y. Obukhov, I. Kershner, A. Karabanov, and A. Gabova, "Classification of early stage Parkinson's disease in EEG and tremor time-frequency features space," *Parkinsonism & Related Disorders*, vol. 22, p. e164, 2016.

[13] A. B. Oktay and A. Kocer, "Differential diagnosis of Parkinson and essential tremor with convolutional LSTM networks," *Biomedical Signal Processing and Control*, vol. 56, Article ID 101683, 2020.

[14] S. Arora, V. Venkataraman, A. Zhan et al., "Detecting and monitoring the symptoms of Parkinson's disease using smartphones: a pilot study," *Parkinsonism & Related Disorders*, vol. 21, no. 6, pp. 650–653, 2015.

[15] H. B. Kim, W. W. Lee, A. Kim et al., "Wrist sensor-based tremor severity quantification in Parkinson's disease using convolutional neural network," *Computers in Biology and Medicine*, vol. 95, pp. 140–146, 2018.

[16] R. López-Blanco, M. A. Velasco, A. Méndez-Guerrero et al., "Essential tremor quantification based on the combined use of a smartphone and a smartwatch: the NetMD study," *Journal of Neuroscience Methods*, vol. 303, pp. 95–102, 2018.

[17] S. E. Lacy, S. L. Smith, and M. A. Lones, "Using echo state networks for classification: a case study in Parkinson's disease diagnosis," *Artificial Intelligence in Medicine*, vol. 86, pp. 53–59, 2018.

[18] M. Lang, S. Hirche, F. M. J. Pfister et al., "A multi-layer Gaussian process for motor symptom estimation in people with Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 11, pp. 3038–3049, 2019.

[19] L. A. Sanchez-Perez, L. P. Sanchez-Fernandez, A. Shaout, J. M. Martinez-Hernandez, and M. J. Alvarez-Noriega, "Rest tremor quantification based on fuzzy inference systems and wearable sensors," *International Journal of Medical Informatics*, vol. 114, pp. 6–17, 2018.

[20] A. Salarian, H. Russmann, C. Wider, P. R. Burkhard, F. J. G. Vingerhoets, and K. Aminian, "Quantification of tremor and bradykinesia in Parkinson's disease using a novel ambulatory monitoring system," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 2, pp. 313–322, 2007.

[21] O. M. Manzanera, J. W. Elting, J. H. van der Hoeven, and N. M. Maurits, "Tremor detection using parametric and non-parametric spectral estimation methods: a comparison with clinical assessment," *PLoS One*, vol. 11, no. 6, Article ID 0156822, 2016.

[22] Y. Pang, J. Christenson, F. Jiang et al., "Automatic detection and quantification of hand movements toward development of an objective assessment of tremor and bradykinesia in Parkinson's disease," *Journal of Neuroscience Methods*, vol. 333, Article ID 108576, 2020.

[23] G. Kramer, A. M. M. Van der Stouwe, N. M. Maurits, M. A. J. Tijssen, and J. W. J. Elting, "Wavelet coherence analysis: a new approach to distinguish organic and functional tremor types," *Clinical Neurophysiology*, vol. 129, no. 1, pp. 13–20, 2018.

[24] S. Shahtalebi, S. F. Atashzar, R. V. Patel, and A. Mohammadi, "WAKE: wavelet decomposition coupled with adaptive Kalman filtering for pathological tremor extraction," *Biomedical Signal Processing and Control*, vol. 48, pp. 179–188, 2019.

[25] R. López-Blanco, M. A. Velasco, A. Méndez-Guerrero et al., "Smartwatch for the analysis of rest tremor in patients with Parkinson's disease," *Journal of the Neurological Sciences*, vol. 401, pp. 37–42, 2019.

[26] F. Vial, P. Kassavetis, S. Merchant, D. Haubenberger, and M. Hallett, "How to do an electrophysiological study of tremor," *Clinical Neurophysiology Practice*, vol. 4, pp. 134–142, 2019.

[27] N. Zeng, H. Qiu, Z. Wang, W. Liu, H. Zhang, and Y. Li, "A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 320, pp. 195–202, 2018.

[28] N. Zeng, H. Li, and Y. Peng, "A new deep belief network-based multi-task learning for diagnosis of Alzheimer's disease," *Neural Computing and Applications*, 2021.

[29] C. R. Pereira, D. R. Pereira, G. H. Rosa et al., "Handwritten dynamics assessment through convolutional neural networks: an application to Parkinson's disease identification," *Artificial Intelligence in Medicine*, vol. 87, pp. 67–77, 2018.

[30] A. Papadopoulos, K. Kyritsis, L. Klingelhoefer, S. Bostanjopoulou, K. R. Chaudhuri, and A. Delopoulos, "Detecting parkinsonian tremor from IMU data collected in-the-wild using deep multiple-instance learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2559–2569, 2020.

[31] S. A. C. Yohanandan, M. Jones, R. Peppard, J. L. Tan, H. J. McDermott, and T. Perera, "Evaluating machine learning algorithms estimating tremor severity ratings on the bain–findley scale," *Measurement Science and Technology*, vol. 27, no. 12, Article ID 125702, 2016.

[32] J. V. D. Wardt, A. M. M. van der Stouwe, M. Dirkx et al., "Systematic clinical approach for diagnosing upper limb tremor," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 91, no. 8, pp. 822–830, 2020.

[33] G. Rigas, A. T. Tzallas, M. G. Tsipouras et al., "Assessment of tremor activity in the Parkinson's disease using a set of wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 478–487, 2012.

[34] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology: Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[35] A. Y. Meigal, S. M. Rissanen, M. P. Tarvainen et al., "Linear and nonlinear tremor acceleration characteristics in patients with Parkinson's disease," *Physiological Measurement*, vol. 33, no. 3, pp. 395–412, 2012.

[36] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[37] H. Jeon, W. Lee, H. Park et al., "High-accuracy automatic classification of Parkinsonian tremor severity using machine learning method," *Physiological Measurement*, vol. 38, no. 11, pp. 1980–1999, 2017.

[38] S. Williams, S. D. Relton, H. Fang et al., "Supervised classification of bradykinesia in Parkinson's disease from smartphone videos," *Artificial Intelligence in Medicine*, vol. 110, Article ID 101966, 2020.

[39] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[40] L. Breiman, "Random forest," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[41] R. E. Schapire, "A brief introduction to boosting," in *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, vol. 2, pp. 1401–1406, San Francisco, CA, USA, July 1999.

[42] H. Yu, "Support vector machine," *Encyclopedia of Database Systems*, Springer, Boston, MA, USA, 2009.

[43] F. Wilcoxon, "Individual comparisons by ranking methods," *Breakthroughs in Statistics*, Springer, New York, NY, USA, 1992.

[44] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American Statistical Association*, vol. 47, no. 260, pp. 583–621, 1952.

[45] A. Zhang, F. De la Torre, and J. Hodgins, "Comparing laboratory and in-the-wild data for continuous Parkinson's disease tremor detection," in *Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, Canada, July 2020.