



Assessment of response to medication in individuals with Parkinson's disease



Murtadha D. Hssayeni^a, Michelle A. Burack^b, Joohi Jimenez-Shahed^c, Behnaz Ghoraani^{a,*}

^a Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL 33431, USA

^b Department of Neurology, University of Rochester Medical Center, Rochester, NY 14642, USA

^c Department of Neurology, Baylor College of Medicine, Houston, TX, USA

ARTICLE INFO

Article history:

Received 15 October 2018

Revised 7 February 2019

Accepted 3 March 2019

Keywords:

Parkinson's disease

Wearable data analysis

Feature extraction and classification

Support vector machine

ABSTRACT

Background and Objective: Motor fluctuations between akinetic (medication OFF) and mobile phases (medication ON) states are one of the most prevalent complications of patients with Parkinson's disease (PD). There is a need for a technology-based system to provide reliable information about the duration in different medication phases that can be used by the treating physician to successfully adjust therapy.

Methods: Two KinetiSense motion sensors were mounted on the most affected wrist and ankle of 19 PD subjects (age: 42–77, 14 males) and collected movement signals as the participants performed seven daily living activities in their medication OFF and ON phases. A feature selection and a classification algorithm based on support vector machine with fuzzy labeling was developed to detect medication ON/OFF states using gyroscope signals. The algorithm was trained using approximately 15% of the data from four activities and tested on the remaining data.

Results: The algorithm was able to detect medication ON and OFF states with 90.5% accuracy, 94.2% sensitivity, and 85.4% specificity. It performed equally well for all the activities with an average accuracy of 91.3% for the activities that were used in the training phase and 88.4% for the new activities.

Conclusions: The developed sensor-based algorithm could provide objective and accurate assessment of medication states that can lead to successful adjustment of the therapy resulting in considerably improved care delivery and quality of life of PD patients.

© 2019 IPEM. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Parkinson's disease (PD) is one of the most common chronic progressive neurological disorders, affecting an estimated seven to 10 million people worldwide and resulting in over \$20 billion in direct and indirect costs a year in America, which is predicted to double by 2040. [1]. PD leads to devastating chronic complications including observable motor symptoms such as tremor, bradykinesia/akinesia (reduced speed and quantity of spontaneous movement), and gait/balance impairment leading to falls. Levodopa is the most common medication used to improve motor impairments in subjects with PD. Unfortunately, prolonged treatment with Levodopa causes troubling motor fluctuations [2] occurring in 50% of patients within 3–5 years of diagnosis and 80% of patients after 10 years [3]. These complications result in frequent fluctuations of “response to treatment intervention” between “ON” state with maximum benefit from Levodopa and “OFF” state with least

benefit from Levodopa, and are a major focus of PD management [4]. The current protocol to address these motor fluctuations involves adjusting therapy, e.g., medication frequency and dosage or Deep Brain Stimulation (DBS) parameters, according to durations in medication ON and OFF states that are obtained from patient's self-reports. Patient self-reports require extensive patient education. Even then, self-reports can be unreliable and suffer from recall bias, resulting in frequent clinical examinations to adjust therapy [5]. Unfortunately, these frequent clinical examinations may not be practical, especially in rural areas where specialty care is less accessible.

Technology-based assessments of patients' response to therapy, using wearable sensors, hold great promise in providing objective measures during subjects' daily, free-living condition that detect PD patients from healthy subjects [6], quantify PD progression [7,8], or can be used by the treating physician to adjust therapy [9]. In this work, our objective is to develop a sensor-based assessment system that can detect patients' responses to treatment interventions (i.e., medication ON and OFF states) that are obtained from patients' self-reports. The development of a sensor-based assessment system that can passively detect clinically important

* Corresponding author.

E-mail addresses: bghoraani@ieee.org, bghoraani@fau.edu (B. Ghoraani).

information on patients' duration in medication ON and OFF states, from natural environment of individuals, will play a significant role in yielding improved PD therapy adjustment strategies to reduce complicated motor fluctuations and associated healthcare costs without the need for less reliable patient self-reports or frequent clinical examinations.

The wide availability of wearable inertial sensors, combined with machine learning algorithms, has led to the development of algorithms for detection of medication ON and OFF states. However, the existing approaches do not fully address the need for clinically actionable information as required for the therapy adjustments. First, some approaches obtained accurate results if the detection was performed for some specific activity, such as walking [10,11] or non-walking [12]. This is problematic as those methods can identify medication states only during those specific activities and are unable to provide a continuous monitoring of the subjects as needed for detection of clinically important information about duration in different medication states. Second, some approaches trade accuracy for continuous monitoring [13–15], or have to use five to seven sensors at different parts of the body in order to provide acceptable accuracy. As a result, the condition under which the subjects need to use the device is very impractical with too many sensors to wear. Third, these systems are based on a “one-size-fits-all” approach and do not consider inter-subject variability. For example, one subject's ON state could be another subject's OFF state depending on the disease stage and each individual's variability with respect to somatotype (body parts are impaired), phenomenology (different motor impairments are present), and severity (e.g. one subject's best movement speed may correspond to a less advanced subject's worst movement speed). However, the underlying algorithms of these systems are modeled using data from a group of well-characterized subjects and are used on new subjects [16]. Hence, they have limited generalizability and still require additional clinical examinations be performed to customize their report for a new subject.

In this work, we developed and validated through experiments an algorithm to automatically detect medication ON and OFF states. Our approach is novel in that it is customized to each subject rather than a “one-size-fits-all” approach, and can continuously detect and report medication ON and OFF states as subjects perform different daily routine activities. To attain our objectives, we tested the hypothesis that the medication ON and OFF patterns in two wearable sensor data can be modeled and identified for every individual, when compared to the gold-standard medication ON and OFF states reported by a neurologist. The population of interest in this study is PD patients with motor fluctuation who go through follow-up visits as frequently as every six months. There will be multiple adjustments in the medications, which are heavily based on the subjects' self-report. Hence, patients would benefit most from an automated and user-friendly system that is trained in their first visit, and then readily used to detect their response to medication (durations in ON/OFF medication states) on a continuous basis and provide these objective measures that can be used by the treating physician to adjust therapy.

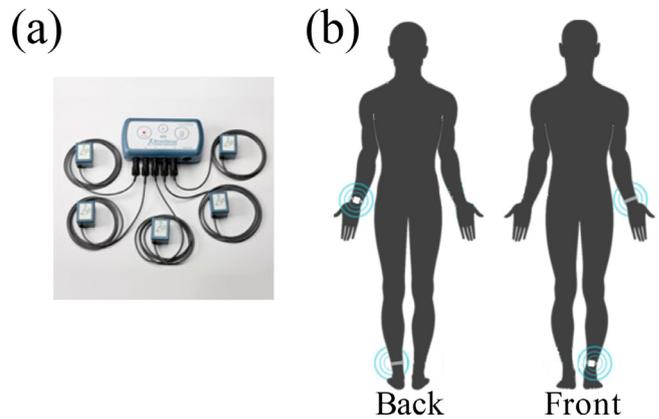


Fig. 1. (a) KinetiSense motion sensor unit. (b) The locations of the wearable sensors.

2. Methods

We developed a new approach to detect medication ON/OFF states of PD subjects using data from two triaxial gyroscope sensors as patients performed a variety of daily-life activities. The approach consists of a feature extraction and selection stage, which represents each segmented window as a set of features, and a classifier based on a support vector machine (SVM) with fuzzy labeling. The details of the data collection and processing, as well as feature extraction and training the classifier, are explained in this section.

2.1. Dataset

A total of 19 PD subjects were selected in this study (see Table 1 for subject demographics). The recording system (Fig. 1(a)) was based on a KinetiSense motion sensor unit (Great Lakes NeuroTechnologies Inc., Cleveland, OH) consisting of a triaxial accelerometer and triaxial gyroscope with 128 Hz sampling rate. Two of these units were mounted on the back and front of the most affected wrist and ankle, respectively, as shown in Fig. 1(b). The study was approved by the institutional review board, and all patients provided written informed consent. The dataset was acquired according to a protocol described elsewhere [17,18]. All the participants were asked to stop their medication the night before the scheduled examination and started the experiment in their medication OFF state. First, they were asked to perform the following seven daily living activities: resting, walking, drinking, dressing, hair brushing, unpacking groceries, and cutting food. These activities were selected as they are routinely referred to during the clinical assessments of Unified Parkinson's Disease Rating Scale (UPDRS) [19] and Unified Dyskinesia Rating Scale (UDysRS) [19], and performing them in a clinical setting is practical. Next, the subjects took their normal anti-parkinsonian medications and performed the same daily activities again. Later, when the subjects were in their ON state, confirmed by a neurologist's direct

Table 1

The subject demographics. AC-UPDRS-III stands for average change between the ON- and OFF-state UPDRS-III, LEDD for Levodopa equivalent daily dose and mAIMS for modified abnormal involuntary movement scale. Values are presented as n, mean \pm STD and/or [range].

Number of subjects	19	OFF UPDRS-III	31 \pm 13 [12–60]
Age [yr–yr]	[42–77]	ON UPDRS-III	14 \pm 8 [4–35]
Sex (M, F)	14, 5	AC-UPDRS-III	14.4 \pm 7.9
Disease duration (yr)	9.2 \pm 3.8	OFF tremor (hand, foot)	0.68 \pm 1.11, 0.53 \pm 0.84
LEDD (mg)	1282.5 \pm 459.8	ON tremor (hand, foot)	0.05 \pm 0.23, 0.00 \pm 0.00
Bradykinesia	16.4 \pm 7	Dyskinesia (mAIMS)	5.2 \pm 3.6

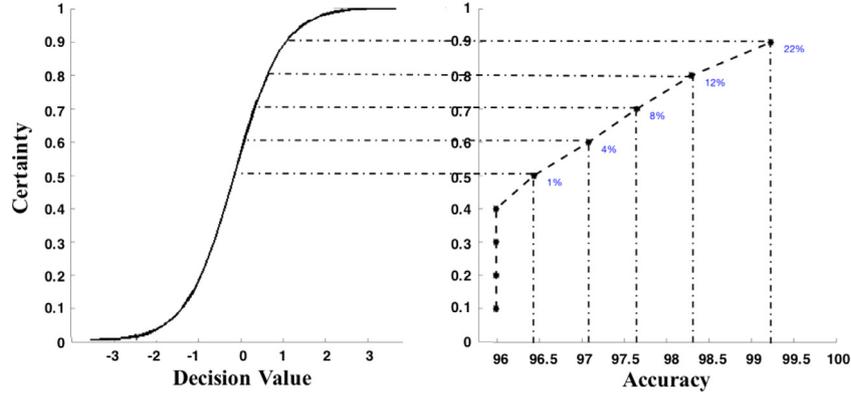


Fig. 2. The plot on the left side is a sample of Platt sigmoid and the plot on the right side is the training accuracy corresponding to different threshold values on the classification certainty. The percentages represent the inconclusive data for nine different threshold values from 50% to 90%.

observation, they were asked to cycle Concurrently, the clinical examinations were performed by a neurologist to measure and record patients' UPDRS part III (UPDRS-III) as well as tremor and bradykinesia sub-scores, and the modified Abnormal Involuntary Movement Scale (mAIMS) to record motor complications related to dyskinesia (abnormal involuntary movements during ON states) sub-score. The UPDRS assessment is composed of four subscales: Parts I and II consider non-motor symptoms and motor experiences of daily living, respectively. UPDRS-III measures the severity of motor complications and is performed by clinical examinations, and part IV measures motor fluctuation and complications. Tremor sub-score for the more affected hand/foot was reported with the following ratings: 0 = absent. 1 = possible rest tremor. 2 = slight and infrequently present. 3 = mild in amplitude and persistent, or moderate in amplitude, but only intermittently present. 4 = moderate in amplitude and present most of the time. 5 = marked in amplitude and present most of the time. Bradykinesia sub-score has a range between 0 and 44 with 0 being the lack of Bradykinesia and 44 referring to severe Bradykinesia. The dyskinesia sub-score has a range of 0–28 with 0 being the lack of random, involuntary movements in the subject's extremities and trunk.

2.2. Data preprocessing

The KinetiSense motion sensors used in the collection provided two motion data types: gyroscope and accelerometer. We used the gyroscope data as is less affected by vibration and mechanical noise than an accelerometer and is expected to better reflect limb rotations that occur during tremor and other PD motor symptoms due to the absence of gravity [17,20,21]. The gyroscope signals were preprocessed by applying a bandpass FIR filter with a pass frequency between (0.5–15 Hz) on the three axes of the recorded signal from each sensor to eliminate low and high-frequency noises. The details of the filter are provided in Supplement Fig. 1 of the Supplement materials. Next, the filtered data was segmented into 5-s-long signals with 4-s overlaps between each segment as shown to be a suitable signal duration to detect bradykinesia and tremor [22]. A few examples of the filtered signal recorded during the medication OFF and ON states are shown in Supplement Fig. 2 of the Supplement materials.

2.3. Selecting training and testing data

In a real-life scenario, the developed algorithm will be trained during a patient's first visit. Then, it will be used to detect the response to medication (ON/OFF medication states) on a continuous basis and report objective information about the duration in ON and OFF states to the treating neurologist for remote medication

adjustments. We selected the training data such that it is feasible to collect during a routine clinical assessment, and does not cause additional burden on the subjects or their treating physicians. Hence, the training data was selected to be short in duration and also from the activities that can be performed in an office setting (i.e., ambulation, drinking, arm resting, and dressing) and are normally performed as part of the routine assessments in Parkinson's disease. The average duration of the ON and OFF training and testing data and the number of analyzed segments are reported in Table 2. Note that there is no overlap between the training and testing data.

2.4. Feature extraction

The total number of extracted features for every window, before any feature selection, was 69 (see Table 3). Twenty-two of these features were extracted from each X, Y, and Z signals (resulting in a total of 66 for all the three axes) and three features were jointly extracted from the three axes' signals. These extracted features were concatenated into one feature vector to represent the data. An alternative approach would be extracting the features from the magnitude signal; however, this method is unable to retain the directional information of the subject's movement. Table 3 lists the extracted features from a segmented window, $\{W_i\}_{i=1:N_{Tr}}$ where N_{Tr} is the number of windows in a subject's training data. The signal power of A-B Hz band (Features 1–3) was calculated as the summation of the signal powers of the A–B Hz band in the Fourier Domain. The spectral entropy (Feature 5) of the spectrum $\{W_i(f)\}_{i=1:N_{Tr}}$ represented the complexity in each window. The dominant frequency feature (Feature 7) was the dominant harmony of the power spectral density as PD subjects in the OFF state present lower dominant frequency with lower power than in their ON state [24]. Average jerk (Feature 10) was the mean of the second derivative of angular velocity to represent the rate of change. Temporal Shannon entropy (Feature 20) represented the complexity and randomness in the signal as defined below:

$$\{H_{W_i}\}_{i=1:N_{Tr}} = - \sum_{b \in B_i} (p(b) * \log_2(p(b))) \quad (1)$$

where B_i is a histogram bin set in the range $(-400, 400)$ with 200 steps for W_i , which are found to be effective in this work, and $p(b)$ is the bin b probability in B_i . The range of the bin set comes from the 75th and 25th percentile of the angular velocity reading. Gini Index (Feature 21) with a value between 0 and 1 was extracted to measure the moving complexity and was calculated as follows:

$$\{G_{W_i}\}_{i=1:N_{Tr}} = 1 - \sum_{b \in B_i} p(b)^2 \quad (2)$$

Table 2

The average duration of the data and number of segments that were used for training and testing the algorithm. Values are presented as mean \pm SD.

	Training		Testing	
	OFF	ON	OFF	ON
Minutes	4.08 \pm 2.61	3.80 \pm 2.68	37.00 \pm 47.79	15.77 \pm 13.30
Segments	223 \pm 149	209 \pm 156	2203 \pm 2880	902 \pm 824
Activities	Part of walking, drinking, resting, and dressing	Part of walking, drinking, resting, and dressing	Remaining of walking, drinking, resting, and dressing + hair brushing, unpacking groceries, and cutting food	Remaining of walking, drinking, resting, and dressing + hair brushing, unpacking groceries, and cutting food

Table 3

List of extracted features.

Feature #	Feature description
1	Signal power of 1–4 Hz band
2	Signal power of 4–6 Hz band [23]
3	Signal power of 0.5–15 Hz band
4	Percentage of the powers for frequencies > 4 Hz [12]
5	Spectral entropy (SH) [21]
6	Peak in the power spectral density [17,24]
7	Dominant frequency [24]
8	Second peak in the power spectral density
9	Secondary frequency associated with the second peak
10	Average jerk
11	Standard deviation (σ) [21]
12	Peak-to-peak in the recorded signal
13	Mean value for each window (μ_{W_i}) [25]
14	Number of autocorrelation peaks
15	Sum of autocorrelation peaks
16	Lag of the first autocorrelation peak [26]
17	First autocorrelation peak (excluding the peak at the origin)
18	Skewness (γ_1) [25]
19	Kurtosis (k) [25]
20	Temporal Shannon entropy [21]
21	Gini index
22	Sample entropy (SE) [27]
23	Cross-correlation ($\rho_{W_i^X W_i^Y}$) between X and Y axes
24	Cross-correlation ($\rho_{W_i^X W_i^Z}$) between X and Z axes
25	Cross-correlation ($\rho_{W_i^Y W_i^Z}$) between Y and Z axes

Sample entropy (Feature 22) was extracted to represent dyskinetic movements in a given window (W_i). Dyskinetic movements are known as abnormal, involuntary movements that occur randomly during medication ON states, and by presenting them in terms of a sample entropy feature, we could improve the classifier's ability to discriminate between the ON states (with dyskinetic movements) and the OFF states with no dyskinetic movements. Sample entropy was calculated by first segmenting W_i into vectors of length m and $m + 1$ defined as $X_m(\tau) = \{W_i(\tau), W_i(\tau + 1), W_i(\tau + 2), \dots, W_i(\tau + m - 1)\}$ where τ is the lag between the segments. Second, the Chebyshev distance was calculated between all the vectors of length $m + 1$ ($d1$) and length m ($d2$). If any of these distances was less than a tolerance r , then their corresponding counter $M1$ and $M2$ was increased by one, respectively, as shown below:

$$M1 = \sum_{j=1}^{N_s-m-1} \sum_{k=\{j+1\}|d1}^{N_s-m-1} 1 \quad (3)$$

$$M2 = \sum_{j=1}^{N_s-m-1} \sum_{k=\{j+1\}|d2}^{N_s-m-1} 1 \quad (4)$$

where $d1 = d[X_{m+1}(j), X_{m+1}(k)] < r$, $d2 = d[X_m(j), X_m(k)] < r$, and $d[., .]$ is the Chebyshev distance. In this work, m and r were set to two and 20% of the standard deviation of the gyroscope signal, respectively. These parameters were selected according to Ref. [27], where the authors show that the calculated sample entropy with those parameters was able to detect dyskinetic movements in PD subjects. Finally, sample entropy was computed as follows:

$$\{SE_{W_i}\}_{i=1:N_W} = -\log \frac{M1}{M2} \quad (5)$$

Cross-correlation feature ($\rho_{W_i^A1 W_i^A2}$) (Features 23–25) was computed to quantify symmetry between the A1 and A2 axes of the recorded gyroscope signal. Since PD subjects tend to bend their limbs in the ON state, these features were designed to use symmetry between different axes to differentiate between the ON and OFF states.

2.4.1. Selection of significant features

For every training and testing set of data, we applied a statistical analysis approach to the features from the training data to select the features that provide high discrimination between the ON and OFF medication states. First, the normality of distribution of every feature was assessed using the Anderson–Darling test. Next, the statistical significance of the normally distributed features was tested using unpaired t -test [28], and for the non-normally distributed features the Wilcoxon rank sum was used. Features with 5% significance level (i.e., p -value < 0.05) between the ON versus OFF medication states were selected. The selected features were used in the training and testing of the classifier as explained in the next section.

2.5. SVM classification with fuzzy labeling

Our prior analysis showed that the support vector machine (SVM) was suitable for modeling ON versus OFF medication states and outperformed the well-known clustering classification methods [29]. Hence, in this work, the well-known SVM approach was employed. The algorithm consists of a training and a testing stage, where the former adjusts the SVM parameters based on the training data and the latter tests the trained SVM model on the testing data.

2.5.1. Training

We used the recursive feature elimination (REF) method to jointly select features and train SVM hyperparameters [30]. REF selects features via a greedy backward selection: it trains the SVM hyperparameters, and obtains a feature weight vector based on minimizing generalization bounds of the leave-one-out cross-validation, using a gradient descent approach. It then removes the feature with the smallest returned absolute-value weight. This process is repeated until a maximum classification accuracy is achieved. We used the Feature Selection Library MATLAB Toolbox to perform the above feature selection process and train the SVM hyperparameters (i.e., linear or RBF kernel, cost parameter ($c \in 2^{\{-2, \dots, 2\}}$), and gamma parameter ($\gamma \in 2^{\{-4, \dots, 4\}}$)). This analysis ensures that a nonlinear SVM classifier will be selected in cases that the feature data is non-linear.

Classification certainty (P): It is known that SVM does not provide any posterior probability that can represent the model classification certainty. Platt [31] proposed a parametric form of sigmoid to provide a measure of posterior probability for SVM that measures the confidence of the trained SVM model in detecting the medication state of each window. This approach uses a parametric

model according to Eq. (6) to fit posterior $P(\mathbf{F} = 1|D)$ where \mathbf{F} indicates feature vectors, D represents the SVM decision values, and 1 refers to the class labels being medication “ON state.” The parameters A and B are fit using maximum likelihood estimation from a union of n -fold sets of D values from the training data. Each of these sets were obtained after training an SVM model on $(n - 1)$ fold of the training data and validating them on the remaining fold to find D values for each feature vector in that fold. The union of all D values as a result of the cross-validation on the training data were used as the training set of the sigmoid. Two scalar parameters (A and B) were estimated by minimizing a negative log likelihood based on the ground truth labels and the training decision values, $D_{1 \times N_{Tr}}$ [31]. We determined these values, $D_{1 \times N_{Tr}}$, from the training data based on 4-fold cross-validation of the four activities in the training data. The estimated parameters (A and B) were then used in Eqs. (6) and (7) to provide a measure of the classification certainty $0 \leq P_n \leq 1$ of a classification decision value M_n .

$$\{P_n\}_{\{n=1:N_{Te}|M_n>0\}} = \frac{1}{1 + e^{A \times M_n + B}} \quad (6)$$

$$\{P_n\}_{\{n=1:N_{Te}|M_n<0\}} = \frac{e^{A \times M_n + B}}{1 + e^{A \times M_n + B}} \quad (7)$$

where a higher value of P_n indicates a higher certainty in the classification decision value. Hence, the decision values with the classification certainty of less than a threshold Th_{Tr} can be reported inconclusive and censored from the medication detection report. Fig. 2 displays one example of the change in the percentage of inconclusive data with respect to thresholds 50% to 90% on the classification certainty. We selected threshold Th_{Tr} for every training data such that there was at most 1% rejection rate in the decision report. The threshold Th_{Tr} along with the estimated parameters A and B were then used in the testing phase.

2.5.2. Testing

The selected features (from the above process) were extracted from the testing data, and passed through the trained SVM model to obtain the decision values ($D_{1 \times N_{Te}}$). Given that the medication state does not normally change so quickly, rapid fluctuating outliers in the predicted decision values were adjusted accordingly. A fuzzy labeling approach was used to adjust the classification label of every second data based on a group of its contiguous feature vectors. This fuzzy labeling process was performed using two averaging filters. The first averaging filter was applied to the decision values ($D_{1 \times N_{Te}}$). The filter had a width of five steps and four-step overlap to reduce the effect of outliers in the contiguous decision values that share part of the signal between each other. The output of this filter was then further processed using a sliding 40-second averaging filter. After filtering, the averaged decision values ($M_{1 \times N_{Te}}$) were used to predict the medication state of every window. The medication state was predicted as medication OFF if $M_n > 0$, and otherwise, it was predicted as medication ON. The trained A and B parameters from the training phase were employed

to estimate the classification certainty for every second n . Eq. (6) was used for medication OFF detection cases and Eq. (7) was used for cases with medication ON detection. The trained threshold Th_{Tr} was used to select and report the decision values with a high classification certainty value, and the remaining data with a certainty value of less than Th_{Tr} was reported as inconclusive.

3. Results

The developed medication state detection algorithm was trained and evaluated using the dataset in Section 2.1. The feature selection and SVM training process resulted in different SVM hyper-parameters and features for different subjects. The linear kernel was selected for five individuals and RBF kernel for the rest of them. The testing accuracy, sensitivity, and specificity results of all the subjects are reported in Table 4. On average, there was 2.7% of inconclusive data with an average classification accuracy of 90.5%, sensitivity of 94.2%, and specificity of 85.4%.

A sample of the medication ON/OFF report that was predicted by the algorithm is shown in Fig. 3. The tall boxes indicate where a medication ON state was predicated, and the shorter boxes show a medication OFF prediction. The red tracing provides the certainty measure of the SVM algorithm with respect to the predicted medication state. A classification certainty of 1 indicates that the SVM classifier predicts with 100% certainty. As indicated by the red tracing, the algorithm detected the medication states with a high certainty for most of the times with the highest number of false detection occurring during the transition from the OFF to ON state.

We further investigated the importance of developing an individualized approach for medication state detection using a short duration of the subject’s data. For this purpose, we trained and tested the developed feature extraction and classification algorithm using a leave-one-subject-out cross-validation. At each trail, we trained a classifier using all the data from all the patients except one, and then tested the classifier on the data of the remaining subject. The average cross-validation accuracy was $84\% \pm 12\%$ with $83\% \pm 19\%$ sensitivity and $87\% \pm 15\%$ specificity. As can be observed, the performance was significantly reduced compared to the individualized approach proposed in this work, demonstrating the importance of a personalized approach for characterizing PD subjects.

3.1. Classification performance and activity type

A medication state monitoring algorithm that can provide accurate and objective information about the duration in ON and OFF states has to continuously detect medication ON/OFF states as the subjects perform a variety of daily routine activities. Here the developed algorithm was trained using part of data from four activities (i.e., walking, resting, drinking, and dressing) and tested on a total of seven activities consisting of four training activities (with no overlap between the training and testing data) and three ad-

Table 4
Testing results of ON/OFF medication state detection.

Performance	Subjects									
	1	2	3	4	5	6	7	8	9	10
Accuracy (%)	99.8	99.5	96.5	95.7	95.4	94.1	93.5	93.4	92.8	92.5
Sensitivity (%)	100	100	99.5	99.8	99.2	99.9	98.4	97.3	95.2	96.0
Specificity (%)	98.7	98.1	88.8	87.6	86.3	85.4	85.5	84.2	85.8	84.9
Performance	Subjects									
	11	12	13	14	15	16	17	1	19	Ave.
Accuracy (%)	90.7	89.3	87.4	86.5	85.6	85.2	84.7	79.8	77.3	90.5
Sensitivity (%)	92.1	91.0	88.2	92.3	86.4	88.9	88.3	89.4	88.5	94.2
Specificity (%)	86.3	85.9	86.3	79.1	84.9	82.7	81.9	76.5	74.2	85.4

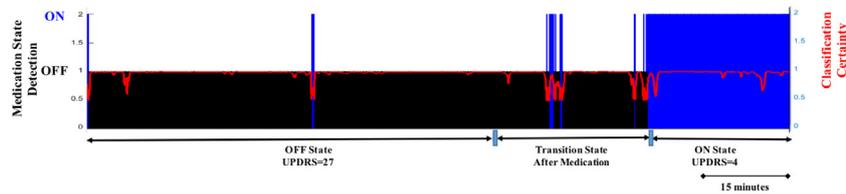


Fig. 3. The classification results (1 medication OFF, 2 medication ON) with the certainty (the continuous line between 0 and 1) for subject 14.

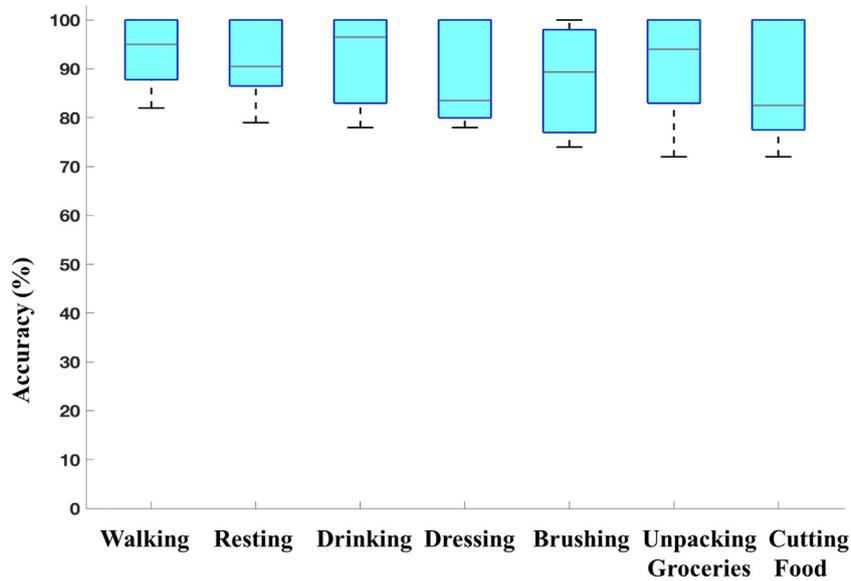


Fig. 4. Testing accuracy for different activities. The central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the minimum and maximum data points.

ditional activities (i.e., brushing, unpacking groceries, and cutting food). To evaluate the robustness of the developed algorithm to the activities that were not included in the classifier training phase, we obtained the average testing accuracy of the algorithm for separate activities (see Fig. 4). The average accuracy for the group of activities in the training phase was 91.3%, and for the group of the new activities, it was 88.4%. We performed a paired *t*-test between the average accuracy of the two activity groups as well as the accuracy of all the 21 possible pairs of activities (e.g., walking and brushing, resting and brushing, etc.). These analyses did not identify any statistically significant difference between classification performance of any of the activities indicating that the algorithm was robust to the new activities that were not used in the training phase.

3.2. Sensor locations and tremor

A fewer number of wearable sensors in a wearable-based medication state monitoring system can improve its user-acceptability and practicality. Hence, we investigated if there was a relationship between the number and location of the sensors and a disease symptom (here tremor) that can help to minimize the number of sensors while achieving the highest classification accuracy. For this purpose, we repeated the classification training and testing phase as described in Section 2.5, using only wrist sensor and only ankle sensor, and calculated the testing accuracy for every subject. Next, the subjects were divided into three symptom groups as diagnosed by the neurologist: without tremor (12 subjects), with dominant tremor on the wrist (4 subjects), and with dominant tremor on the ankle (3 subjects). We compared the classification accuracy (rounded to the nearest one) of three sensor combinations (both wrist and ankle sensors, only wrist sensor, and only ankle sensor) for every symptom group, and counted the number of sub-

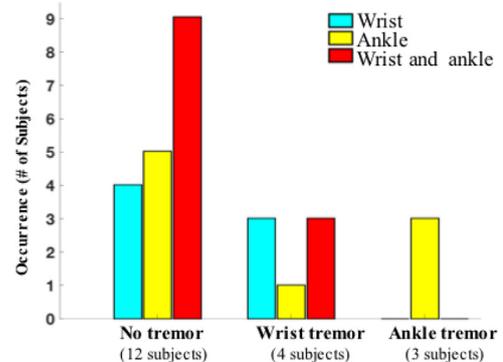


Fig. 5. Occurrence of the sensor combinations with the highest testing accuracy for three symptom groups of no tremor, wrist tremor, and ankle tremor.

jects with the highest accuracy when using a sensor combination. If one subject had the highest classification accuracy using more than one of the sensor combinations, we counted that subject towards all of those sensor combinations. For example, considering the no-tremor group with 12 subjects, our finding was as follows: nine subjects had their highest accuracy using both wrist and ankle sensors; however, one of these subjects resulted in an equally-well accuracy when using only wrist sensor, and two of them achieved an equally-well accuracy when using only the ankle sensor. Three subjects achieved the best classification accuracy only when the wrist sensor was used, and three subjects resulted in their highest classification accuracy only using ankle sensor. These results were counted as 4, 5, and 9 for the best performance using only wrist, only ankle, and both wrist and ankle, respectively. The same process was repeated for the other two symptom groups, and the

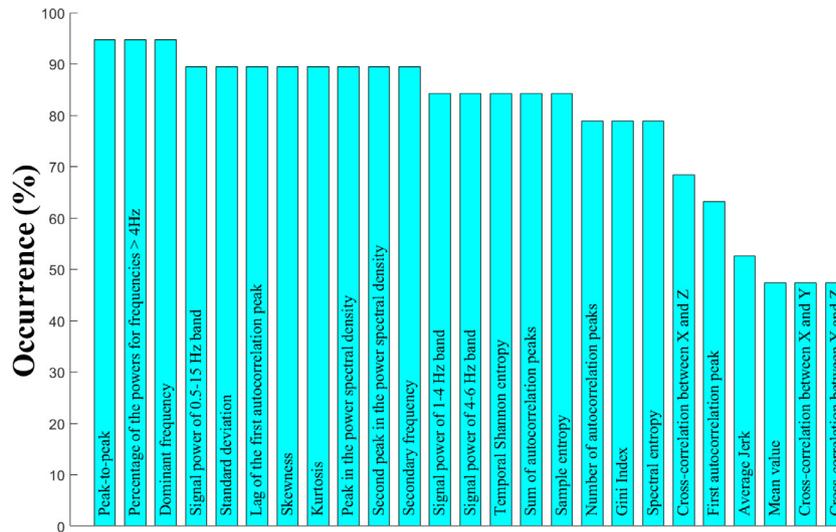
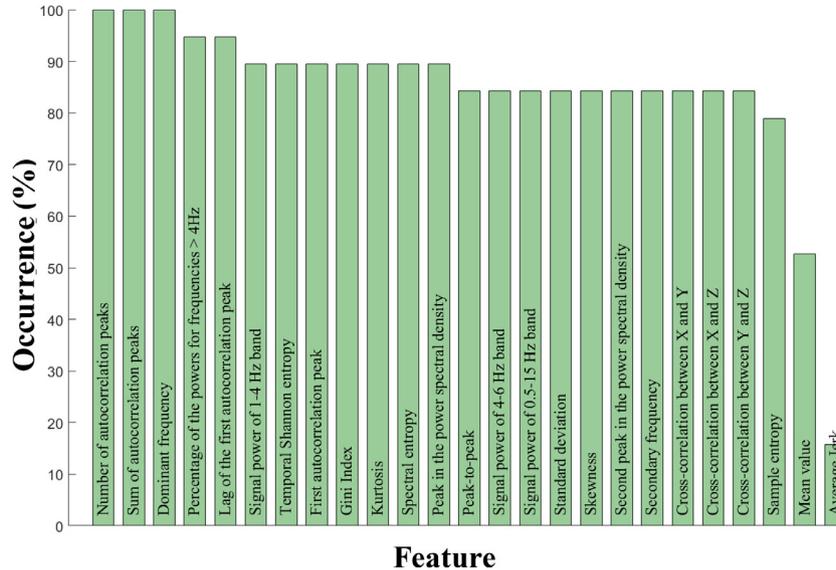
(A) Wrist Sensor**(B) Ankle Sensor**

Fig. 6. Occurrence rate of the selected features from wrist or ankle sensor. The features' names are as defined in Table 3.

results are reported in Fig. 5. One interesting observation from Fig. 5 is that for all the subjects with ankle tremor, using only the ankle sensor was sufficient to achieve the highest classification performance.

3.3. Feature importance to medication state detection

The occurrence rate of the selected features among all the subjects in our study is shown in Fig. 6. An occurrence rate of 100% means that that specific feature provide a high medication ON/OFF discrimination and was selected for all the subjects. We also investigated the relationship between the selected features and tremor and bradykinesia symptoms, and dyskinesia motor difficulty. For every extracted feature type (listed in Table 3), first, we estimated the importance of the feature to the medication state detection with respect to the presence of tremor, bradykinesia, and dyskinesia. This importance was calculated as the difference between the average tremor, bradykinesia, and dyskinesia severity (i.e., clinically-confirmed sub-scores) of two groups of subjects: the subjects for whom the feature was selected and the subjects

for whom the feature was not selected. Next, we displayed the estimated feature importance of all the features for tremor, bradykinesia, and dyskinesia in pairwise plots (see Fig. 7). We repeated this process for the wrist and ankle sensors separately. Finally, we identified the wrist and ankle features that were exclusively important in the presence of severe tremor, bradykinesia, or dyskinesia. For the tremor important features, we determined the features that were common between the two blue shaded areas (i.e., features that were selected in the case of severe tremor and moderate or less bradykinesia and dyskinesia) in Fig. 7(A). For the important bradykinesia features, we identified the features that were common between the two green shaded areas, and for the important dyskinesia features, the ones that were common between the two red shaded areas were selected. For example, in the case of the important tremor features, the features underlying in the first blue area of Fig. 7(A) were feature numbers 1, 4, 5, 9, 15, 17, 18, 20, 22, and 23, and the ones in the second blue area were feature numbers 4, 5, 9, 15, 17, 20, 22, and 23. Hence, the features that were exclusively important to tremor were feature numbers 4, 9, 15, 17, 20, 22, and 23 as they were associated with

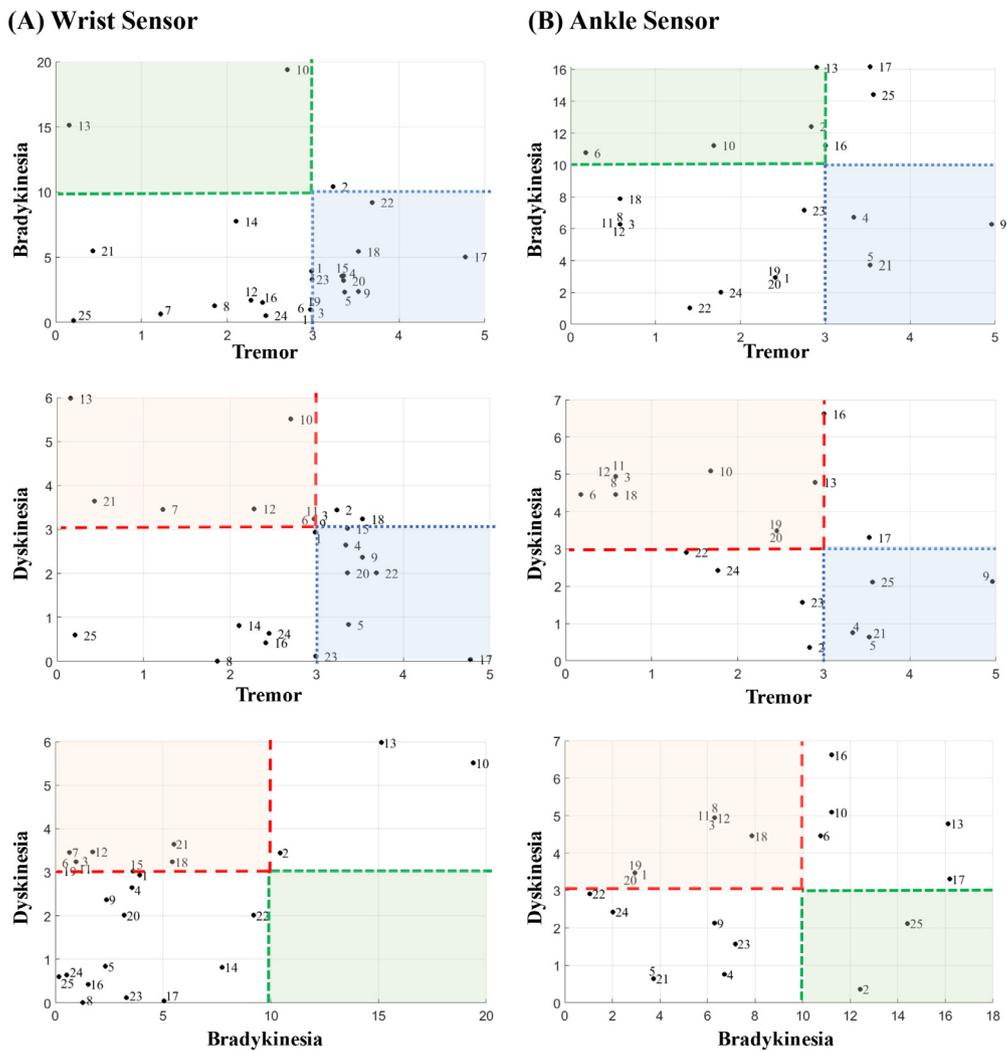


Fig. 7. Importance of the features to medication state detection with respect to the severity of tremor, bradykinesia, and dyskinesia. The features' numbers are as defined in Table 3.

Table 5
Important features to medication state detection with respect to tremor, bradykinesia, and dyskinesia.

	Wrist sensor	Ankle sensor
Tremor	#4: Power frequencies > 4 Hz #9: Secondary frequency #15: Sum of autocorrelation peaks #17: First autocorrelation peak #20: Temporal Shannon entropy #22: Sample entropy #23: Cross-correlation: X and Y axes	#4: Powers frequencies > 4 Hz #5: Spectral entropy #9: Secondary frequency #21: Gini index
Bradykinesia		#2: 4–6 Hz power #3: 0.5–15 Hz power #8: Second peak of power spectral density #11: Standard deviation #12: Signal peak-to-peak #18: Skewness #19: Kurtosis #20: Temporal Shannon entropy
Dyskinesia	#7: Dominant frequency #11: Standard deviation #12: Signal peak-to-peak #21: Gini index	

severe tremor and moderate or less bradykinesia and dyskinesia. This process was repeated for the ankle sensor. Table 5 lists the wrist and ankle features that were exclusively important to the medication ON and OFF detection in the presence of severe tremor, bradykinesia, and dyskinesia.

4. Discussions

We developed a technology-based assessment of PD subjects' medication ON and OFF states, using data from two wearable sensors, to provide objective measures, instead of patient diary or

self-report, that can be used by the treating physician to adjust therapy [9]. We developed a feature extraction and selection algorithm based on SVM classifier with fuzzy labeling. Using concurrently recorded sensor data and gold-standard medication ON and OFF states from 19 PD subjects, we tested our hypothesis that the medication ON and OFF patterns in the data from two wearable sensors can be modeled and identified for every individual. Several interesting observations were made via our analysis: First, our analysis supports our hypothesis by demonstrating that, using the data from two sensors, the developed algorithm was able to predict correct medication states with a high average accuracy of 90.5%, sensitivity of 94.2%, and specificity of 85.4% (Table 4). Moreover, the algorithm performs equally-well for all the seven activities that were considered in this study (Fig. 5). Since we trained the classifier using only part of the data from four of the activities and tested for all the data, including three new activities (Table 2), this observation indicates that the algorithm successfully models the medication state patterns, and can potentially be used to detect medication states during subjects' daily routine activities.

Another interesting observation is that our analysis indicates that the ON/OFF-discriminant features are different between subjects supporting the need for an individualized classifier training approach versus a "one-size-fits-all" approach. As shown in Fig. 6, there is a difference in the set of features that are selected for every subject and from each sensor. Most of the features were selected for more than 70% of the subjects, but some of them (six of the wrist features and two of the ankle features) were selected for less than 70% of the subjects. We also investigated the features that were important to the medication state detection in the presence of severe tremor, bradykinesia, or dyskinesia (see Table 5). We found that the signal power of frequencies higher than 4 Hz and secondary frequency features from both wrist and ankle were important to medication ON/OFF detection in severe tremor cases with faster rhythmic movements of hand or leg during medication OFF states. Similarly, the auto and cross-correlation features in the presence of severe tremor became more evident and were important to differentiating between medication OFF from ON state. The temporal Shannon entropy, sample entropy, spectral entropy, and Gini index features were also important in the presence of severe tremor. These features represent randomness, so their values decrease in the presence of highly structured signals with severe tremor and are important to medication ON/OFF detection. The only feature that was exclusively important in the presence of severe bradykinesia was signal power of 2–4 Hz from the ankle sensor. Slowness of the movement during medication OFF state results in a decrease in the value of this feature, which makes the feature important to detection of medication OFF state with slow movements from medication ON states. The features that were important in the presence of severe dyskinesia included dominant frequency, standard deviation, signal peak-to-peak, and Gini index, Shannon entropy. These features increase with involuntary, random movements of the subjects during ON state and make the medication ON/OFF detection more successful. Other important features in the presence of severe dyskinesia were kurtosis and skewness features. The kurtosis feature decreases and the absolute value of skewness increases with random movements during ON states, which makes these features important to the medication state detection.

We also investigated the possibility of using only one sensor for detecting the medication ON/OFF states. For this purpose, we determined the sensor combination that achieved the highest classification accuracy for every subject and provided its distribution for the presence of tremor in the subject's wrist or ankle (Fig. 6). Our analysis indicates that different sensor numbers and placements provide different performances. It is interesting that using both wrist and ankle sensors does not always result in the best perfor-

mance. All the three individuals with the most dominant tremor on their leg had the best classification accuracy using only a single ankle sensor (see Fig. 6) and three (out of four) individuals with most dominant tremor on their hand had the best classification accuracy using only a wrist sensor. This observation suggests the possibility of a *strategy in selecting the number and placement of sensors*, which is using only an ankle sensor for individuals with the most dominant tremor on the leg and using an ankle and wrist sensor combination for the other subjects. The reliability of such as sensor placement strategy has to be further investigated in future work with a higher number of subjects.

4.1. Comparison to other studies

A summary of some of the previous studies for the classification of PD medication states is listed in Table 6. There are two main types of medication detection approached in the literature. One approach provided medication states only during some specific activities, such as walking or non-walking, [10–12]; however, this approach is unable to provide a continuous monitoring of the subjects as needed for detection of the duration of different medication states. The other approach, similar to our algorithm, makes a medication state detection independent of the activity type [13,14,32], and hence is able to provide duration in different medication states. Our developed algorithm provided the highest sensitivity (94%) and specificity (85%) compared to the state-of-the-art using only two sensors. The second-highest performance was achieved by the work reported by Chelaru et al. [14] with a sensitivity of 90% and specificity of 76%, using five sensors, for 13 subjects with an average change between the ON- and OFF-state UPDRS-III (AC-UPDRS-III) ≥ 16 . In our work, the average classification performance for the subjects who had an AC-UPDRS-III ≥ 16 (eight out of 19 subjects) was 98% sensitivity and 90% specificity, although significantly higher than the performance reported by Salarian [14], it is worth mentioning that there is high variability between the different studies in terms of the number of sensors (range of one to seven), data duration (range of one hour to one week), and subject numbers (range of 12–19), which makes it difficult to perform a fair comparison between the methods.

4.2. Clinical implications and study limitations

The developed algorithm has the potential to establish a wearable-based medication state monitoring system to provide objective measures, from the PD subjects in free-living condition, that can be used by the treating physician to adjust therapy. We used an individualized approach to train a customized classifier for every subject. Similar to the other individualized approaches [10–13], our work uses labeled training data from the same subject, assuming that some gold-standard medication-state data is available for a subject. The objective of this work was to develop an individualized system that can be trained using the data collected during the patients' first clinical visit and then used to automatically monitor the patients between their visits and detect response to medication (durations in ON/OFF medication states) on a continuous basis. Given that the data is available from their first clinical visit, we used about 15% of such data in training the algorithm to make sure that the algorithm is trained for the given patient's disease characteristics. Hence, we designed the classifier training phase based on a short duration of training data from the activities (i.e., ambulation, drinking, resting, and dressing) that can be easily performed in any clinical office. Hence, once the algorithm is trained, it can be readily used as a passive system to monitor medication fluctuations without relying on patient or physician engagement. The algorithm was tested using sensor data as the subjects performed several routine activities in a clinical

Table 6
Comparison of medication ON/OFF state detection methods.

References	# Sensors	# Subjects	Data duration for each subject	Classification method	Activity-independent model (yes, no)	Individualized-model (yes, no, partial)	Results
Keijsers et al. [12]	One tri-axial accelerometer	23	3 h	Linear discriminant and ANN	No (excludes walking)	Partial (Tremor, and non-tremor)	Sens.: 97% Spec.: 97%
Perez-Lopez et al. [10]	One tri-axial accelerometer	7	6 h	Linear discriminant	No (walking and non-walking)	Partial (Threshold on Bradykinesia)	Sens.: 99.9% Spec.: 99.9%
Rodriguez-Molinero et al. [11]	One tri-axial accelerometer	35	1.4–5.5 h	Linear discriminant	No (only walking)	Yes	Sens.: 96% Spec.: 94%
Salarian et al. [14]	Five gyroscopes and accelerometers	13 AC-UPDRS-III ≥ 16	3–6 h	Logistic regression	Yes	No	Sens.: 90% Spec.: 76%
Fisher et al. [32]	Two tri-axial accelerometers	32	4 h (in lab) 1 week (in home)	ANN	Yes	No	Sens.: 55% Spec.: 83%
Hoff et al. [13]	Seven uni-axial accelerometers	15	24 h	Linear discriminant	Yes	Yes	Sens.: 60%–71% Spec.: 66%–76%
Proposed	Two tri-axial gyroscopes	19	1 h	SVM with fuzzy classification	Yes	Yes	Acc.: 91% , Sens.: 94% , Spec.: 85%

lab setting. In our future work, we will investigate the algorithm's performance as the subjects perform a variety of activities in their homes. We will also investigate the sensitivity of the algorithm to the variations in the sensor placements when the subjects wear their sensors at home. Our analysis on selecting the number of sensors and their placement was based on only the tremor symptom. The reliability of this selection to optimize the number of required sensors will be studied using more subjects in our future work.

5. Conclusion

In this paper, we developed a sensor-based assessment system that can continuously detect patients' medication ON and OFF states as the patients perform different types of daily routine activities. A series of novel signal features and feature selection approaches were proposed and integrated into an SVM classifier with fuzzy labeling. The developed algorithm was assessed using data from 19 PD subjects, and the results supported our hypothesis that it was possible to model and detect the subjects' medication ON and OFF state patterns from the data from two wearable sensors as the subjects perform different activities. The proposed algorithm used data from only two sensors on wrist and ankle, and it was able to detect the response to medication during subjects' daily routine activities with an average 90.5% accuracy 94.2% sensitivity, and 85.4% specificity. The developed algorithm would enable a sensor-based assessment system that can provide objective measures about the duration in medication ON and OFF states that are obtained from patients' self-reports. Such a system could provide accurate and timely therapy adjustments in place of less reliable patient self-reports, and as a result, considerably improve both the care delivery and quality of life for the millions of patients afflicted by this debilitating neurodegenerative disorder.

Ethical approval

The study was approved by the institutional review boards of University of Rochester and Great Lakes NeuroTechnologies in accordance with the Helsinki Declaration (2008). All subjects gave written informed consent prior to participation

Conflict of interest

None.

Acknowledgment

The dataset was supported by [Small Business Innovation Research](#) grant offered by [National Institutes of Health](#) (NIH) to Cleveland Medical Devices (1R43NS071882-01A1; T. Mera, PI) and the [National Institute on Aging](#) of NIH grant to Great Lakes NeuroTechnologies Inc. (5R44AG044293).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.medengphy.2019.03.002](https://doi.org/10.1016/j.medengphy.2019.03.002).

References

- [1] National Institute of Neurological Disorders and Stroke. Parkinson's disease: Challenges, progress, and promise. NIH Publication; 2018. No 15-5595
- [2] Dewey RB. Management of motor complications in Parkinson's disease. *Neurology* 2004;62:S3–7.
- [3] Davie C. A review of Parkinson's disease. *Br Med Bul.* 2008;86(1):109–27.
- [4] Jankovic J. Motor fluctuations and dyskinesias in Parkinson's disease: clinical manifestations. *Movem Disord* 2005;20(S11):S11–16.
- [5] Maetzler W, Klucken J, Malcolm H. A clinical view on the development of technology-based tools in managing Parkinson's disease. *Movem Disord* 2016;31(9):1263–71.
- [6] Del Din S, Godfrey A, Galna B, Lord S, Rochester L. Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length. *J. Neuroeng. Rehabil.* 2016;13(1):46.
- [7] Wu Y, Krishnan S. Statistical analysis of gait rhythm in patients with Parkinson's disease. *IEEE Trans Neural Syst Rehabil Eng* 2010;18(2):150–8.
- [8] Wu Y, Chen P, Luo X, Wu M, Liao L, Yang S, et al. Measuring signal fluctuations in gait rhythm time series of patients with Parkinson's disease using entropy parameters. *Biomed Signal Process Control* 2017;31:265–71.
- [9] Rovini E, Maremmani C, Cavallo F. How wearable sensors can support Parkinson's disease diagnosis and treatment: a systematic review. *Front Neurosci* 2017;11:1–41.
- [10] Pérez-López C, Samà A, Rodríguez-Martín D, Català A, Cabestany J, de Mingo E, et al. Monitoring motor fluctuations in Parkinson's disease using a waist-worn inertial sensor. *Adv Comput Intell* 2015:461–74.
- [11] Rodríguez-Molinero A, Samà A, Pérez-Martínez DA, López CP, Romagosa J, Bayés À, et al. Validation of a portable device for mapping motor and gait disturbances in Parkinson's disease. *JMIR mHealth uHealth* 2015;3(1):e9.
- [12] Keijsers NL, Horstink MW, Gielen SC. Ambulatory motor assessment in Parkinson's disease. *Movem Disord* 2006;21(1):34–44.
- [13] Hoff J, Van Der Meer V, Van Hilten J. Accuracy of objective ambulatory accelerometry in detecting motor complications in patients with Parkinson disease. *Clin Neuropharmacol* 2004;27(2):53–7.
- [14] Salarian A. Ambulatory monitoring of motor functions in patients with Parkinson's disease using kinematic sensors, Ph.D. thesis. École Polytechnique Fédérale de Lausanne; 2006.
- [15] Fisher JM, Hammerla NY, Rochester L, Andras P, Walker RW. Body-worn sensors in Parkinson's disease: evaluating their acceptability to patients. *Telemed e-Health* 2016;22(1):63–9.
- [16] Kubota K, Chen J, Little M. Machine learning for large-scale wearable sensor data in Parkinson's disease: concepts, promises, pitfalls, and futures. *Movem Disord* 2016;31(9):1314–26.

- [17] Mera TO, Burack MA, Giuffrida JP. Objective motion sensor assessment highly correlated with scores of global Levodopa-induced dyskinesia in Parkinson's disease. *J Parkins Dis* 2013;3(3):399.
- [18] Pulliam CL, Heldman DA, Brokaw EB, Mera TO, Mari ZK, Burack MA. Continuous assessment of Levodopa response in Parkinson's disease using wearable motion sensors. *IEEE Trans Biomed Eng* 2018;65(1):159–64.
- [19] Goetz CG, Nutt JG, Stebbins GT. The unified dyskinesia rating scale: presentation and clinimetric profile. *Movem Disord* 2008;23(16):2398–403.
- [20] Fabbrini G, Brotchie JM, Grandas F, Nomoto M, Goetz CG. Levodopa-induced dyskinesias. *Movem Disord* 2007;22(10):1379–89.
- [21] Tsipouras M, Tzallas A, Rigas G, Tsouli S, Fotiadis D, Konitsiotis S. An automated methodology for Levodopa-induced dyskinesia: assessment based on gyroscope and accelerometer signals. *AI Med* 2012;55(2):127–35.
- [22] Patel S, Lorincz K, Hughes R, Huggins N, Growdon J, Standaert D, et al. Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors. *IEEE Trans Inf Technol Biomed* 2009;13(6):864–73.
- [23] Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases.. *J Neurol Neurosurg Psychiat* 1992;55(3):181–4.
- [24] Weiss A, Sharifi S, Plotnik M, van Vugt JP, Giladi N, Hausdorff JM. Toward automated, at-home assessment of mobility among patients with Parkinson disease, using a body-worn accelerometer. *Neurorehabil Neural Repair* 2011;25(9):810–18.
- [25] Altun K, Barshan B, Tunçel O. Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Patt Recogn* 2010;43(10):3605–20.
- [26] Roy SH, Cole BT, Gilmore LD, Luca CJ, Thomas CA, Saint-Hilaire MM, et al. High-resolution tracking of motor disorders in Parkinson's disease during unconstrained activity. *Movem Disord* 2013;28(8):1080–7.
- [27] Chelaru MI, Duval C, Jog M. Levodopa-induced Dyskinesias detection based on the complexity of involuntary movements. *J Neurosci Methods* 2010;186:81–9.
- [28] Student. The probable error of a mean. *Biometrika* 1908:1–25.
- [29] Hssayeni MD. Automatic assessment of medication states of patients with Parkinson's disease using wearable sensors, Master's thesis. Rochester, NY: Rochester Institute of Technology; 2017.
- [30] Weston J, Mukherjee S, Chapelle O, Pontil M, Poggio T, Vapnik V. Feature selection for SVMs. *Adv Neural Inf Process Syst* 2001;31(9):668–74.
- [31] Platt JC. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv Large Marg Classif* 1999;10(3):61–74.
- [32] Fisher JM, Hammerla NY, Ploetz T, Andras P, Rochester L, Walker RW. Unsupervised home monitoring of Parkinson's disease motor symptoms using body-worn accelerometers. *Parkinson Relat Disord* 2016;33:44–50.