Parkinsonism and Related Disorders 63 (2019) 111–116

Contents lists available at ScienceDirect

Parkinsonism and Related Disorders

journal homepage: www.elsevier.com/locate/parkreldis

Upper limb motor pre-clinical assessment in Parkinson's disease using machine learning

Filippo Cavallo[a,*], Alessandra Moschetti[a], Dario Esposito[c], Carlo Maremmani[b], Erika Rovini[a]

[a] The BioRobotics Institute, Scuola Superiore Sant'Anna, Viale Rinaldo Piaggio, 34, 56025, Pontedera, Italy
[b] U.O. Neurologia, Ospedale delle Apuane (AUSL Toscana Nord Ovest), Viale Mattei, 21, 54100, Massa, Italy
[c] Co-Robotics srl, Via Franchi, 39, 56033, Capannoli (PI), Italy

ARTICLE INFO

Keywords:
Decision support systems
Idiopathic hyposmia
Inertial wearable sensors
Machine learning
Motor pre-clinical symptoms
Prodromal phase
Parkinson's disease

ABSTRACT

Introduction: Parkinson's disease (PD) is a common neurodegenerative disorder characterized by disabling motor and non-motor symptoms. For example, idiopathic hyposmia (IH), which is a reduced olfactory sensitivity, is typical in > 95% of PD patients and is a preclinical marker for the pathology.
Methods: In this work, a wearable inertial device, named SensHand V1, was used to acquire motion data from the upper limbs during the performance of six tasks selected by MDS-UPDRS III. Three groups of people were enrolled, including 30 healthy subjects, 30 IH people, and 30 PD patients. Forty-eight parameters per side were computed by spatiotemporal and frequency data analysis. A feature array was selected as the most significant to discriminate among the different classes both in two-group and three-group classification. Multiple analyses were performed comparing three supervised learning algorithms, Support Vector Machine (SVM), Random Forest (RF), and Naïve Bayes, on three different datasets.
Results: Excellent results were obtained for healthy *vs.* patients classification (F-Measure 0.95 for RF and 0.97 for SVM), and good results were achieved by including subjects with hyposmia as a separate group (0.79 accuracy, 0.80 precision with RF) within a three-group classification. Overall, RF classifiers were the best approach for this application.
Conclusion: The system is suitable to support an objective PD diagnosis. Further, combining motion analysis with a validated olfactory screening test, a two-step non-invasive, low-cost procedure can be defined to appropriately analyze people at risk for PD development, helping clinicians to identify also subtle changes in motor performance that characterize PD onset.

1. Introduction

1.1. Clinical background

Millions of people worldwide are affected by Parkinson's Disease (PD) [1], a neurodegenerative pathology caused by a significant loss of dopamine in the forebrain, characterized by both cardinal motor symptoms [2] and non-motor manifestations (NMMs) [3]. Today, PD diagnosis is typically made by analyzing motor symptoms based on diagnostic criteria [4]. Clinical scales, such as the Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [5] and the Hoehn & Yahr (HY) [6], are semi-quantitative tools used by neurologists to assess PD patients by assigning a score according to the disease severity. The evaluation can be subjective and affected by variability. Therefore, an objective tool that can help neurologists to quantitatively identify small changes in motion performance is necessary to have an unbiased assessment of the disease.

Since motor symptoms appear when several neurological areas are already damaged, recently, interest has grown toward new diagnostic criteria focusing also on NMMs, which are involved in the neuropathological process and can anticipate the onset of motor symptoms by 5–7 years [7].

Idiopathic hyposmia (IH), a reduced olfactory sensitivity, is a common NMM in 95% of PD patients [8]. The risk of developing PD is about 10–12% greater in healthy adults with IH compared to those without IH [9]. However, IH is not sufficient to identify PD onset, since it has low specificity for PD development. Currently, PD diagnosis is confirmed by imaging techniques, which can reveal dopamine transmitters in the brain (i.e., SPECT-DaTSCAN) [10], or can investigate the PD pathophysiology (e.g., NMR-DTI) [11]. However, these methods are

* Corresponding author.
E-mail addresses: filippo.cavallo@santannapisa.it (F. Cavallo), alessandra.moschetti@santannapisa.it (A. Moschetti), dario.esposito@corobotics.eu (D. Esposito), carlo.maremmani@uslnordovest.toscana.it (C. Maremmani), erika.rovini@santannapisa.it (E. Rovini).

https://doi.org/10.1016/j.parkreldis.2019.02.028
Received 24 September 2018; Received in revised form 19 February 2019; Accepted 19 February 2019

invasive and expensive.

Thanks to their accuracy, unobtrusiveness, low cost, and ease of use, wearable sensors can represent an interesting solution for objective and quantitative evaluation of the motor performance [12]. Moreover, novel machine learning (ML) algorithms can enable mining of data acquired by wearable sensors, providing a useful tool for supporting clinicians in PD diagnosis since the beginning of the pathology [13,14].

In particular, hypothesizing IH as a preclinical biomarker for PD, it would be possible to combine the olfactory screening with a non-invasive motion analysis using wearable sensors to identify the early PD onset [15].

1.2. Related works

ML techniques recently have been applied in studies for PD classification, particularly for speech analysis, comparing different ML algorithms such as Naïve Bayes (NB), J48, Random Forest (RF), and Support Vector Machine (SVM) [16,17]. Generally, a two-group classification was implemented (PD patients vs. healthy subjects of control (HC)), with valuable results obtained. In addition, a multiclass classification work was proposed (i.e., identification of HC, and early, intermediate, and advanced patients) with 93% overall classification accuracy [18].

Some researchers also applied ML to classify PD patients and HC investigating motion abilities. Although most works focused on lower limbs [13], interest in upper limb motion is growing (Table 1). Commonly, the works focused on a single task [19,20], or a single symptom (i.e., tremor) [21,22], while three works involved more exercises [23–25]. Accelerometers and gyroscopes were the most used technologies. Different typologies of classifiers were implemented in these works, without consensus on the most suitable approach for PD assessment. Limited datasets were typically analyzed; thus, the applicability of ML must be demonstrated further.

To the best of our knowledge, none of the studies applies these algorithms to a wide set of features extracted by several exercises, and none consider multiclass classification that includes people at risk for developing PD (i.e., IH subjects). Currently, only imaging diagnostic techniques demonstrated that IH is a preclinical marker of this pathology [8].

In this context, starting from the results obtained by the same authors on lower limb motion [26], this work aims to investigate differences in classes of people (i.e., healthy subjects, IH people, and PD patients), analyze their motor performance in the upper limbs, which are measured using wearable inertial sensors, and compare three different supervised ML approaches. In particular, in this work, the authors propose to:

i) Analyze a comprehensive experimental protocol for a complete motor evaluation of upper limbs in PD using six exercises taken from MDS-UPDRS III and, therefore, increasing the number of exercises with respect to previous works.
ii) Evaluate a wide set of features extracted from the kinematic analysis that enable evaluation of spatial, temporal, and frequency parameters.
iii) Investigate the most suitable ML approach for motor assessment of upper limb performance in PD by using three supervised classifiers (i.e., SVM, RF, and NB).

The final idea is to provide the neurologist with a decision support system, based on the analysis of motor data acquired by inertial sensors during the performance of a structured protocol, applying machine learning techniques aiming to help him in objective clinical diagnosis of PD patients, since the early stage.

2. Materials and methods

2.1. Participants

Three age-matched groups composed of 30 HC (25 males, 5 females, mean age ± standard deviation 65.2 ± 2.5 y), 30 subjects with idiopathic hyposmia (IH) (21 males, 9 females, 66.0 ± 3.2 y), and 30 patients with Parkinson's Disease (PD) (25 males, 5 females, 67.9 ± 8.8 y) were involved in this study. All patients were clinically assessed, and measurements were performed in a clinically defined ON-state. The PD patients were mild to mid (mean MDS-UPDRS III ± SD score: 14.7 ± 8.6; mean HY ± SD score: 1.9 ± 0.8), with lateralization prevalence of the disease almost equally distributed: 9 right, 10 left, and 11 bilateral. Impairments or diseases, other than PD, that could affect the performance of daily activities represented exclusion criteria. IH subjects were recruited through the IPMP-MS Project that provided screening for IH using the IOIT olfactory test [9]. All the subjects signed written informed consent, and the study procedure was approved by the local Ethical Committee (Azienda Sanitaria Locale, Massa, Italy, n°1148/12.10.10) according to the most recent Declaration of Helsinki.

2.2. Instrumentation

A novel wearable device based on inertial measurement units (IMUs) was developed to objectively analyze the upper limb motor performance of the subjects. The patented device, named SensHand V1 (Supplementary Fig. 1), is low cost, low power, non-invasive, small in size, lightweight, wireless, and easy to use [27]. Supplied by a rechargeable LiPo battery, it allows the collection of data with 100 Hz sampling frequency. The device consists of four IMUs: a coordinator on the wrist, and three units on the distal phalanxes of thumb, index, and middle fingers. Each unit is an IMU integrated into the iNEMO-M1 board based on microelectromechanical systems (MEMS) sensors (3-axis gyroscope L3G4200D and 6-axis geomagnetic module LSM303DLHC) and dedicated ARM-based 32-bit microcontroller STM32F103RE (STMicroelectronics, Italy). Data transmission and synchronization between the units is implemented through the CAN-bus standard. The data are sent through Bluetooth to a control station for offline processing. Movements between sensors and anatomical segments are avoided thanks to Velcro straps.

2.3. Experimental protocol

The subjects' motor performance were analyzed by defining an experimental protocol composed of six exercises (forearm pronation/supination (PSUP), hand opening/closing (OPCL), thumb-forefinger tapping (THFF), postural tremor (POST), resting tremor (HRST), and arms swing during the gait (GTAH)) that followed the tasks described in the MDS-UPDRS III.

Before starting the acquisitions, the clinical staff showed the users the correct execution of each exercise, and a preliminary training phase was performed by the subjects to test all the required tasks. During the entire protocol, except for GTAH that implies a standing position, the subjects assumed a comfortable sitting posture, holding right angles at the hip and at the knee. At the beginning of each exercise, the subjects were asked to maintain a specific fixed position for 3 s to acquire a static baseline for each trial. The detailed description of the exercises can be found in Supplementary Table 1.

Every subject was consecutively examined two times, for both the right hand (R_HAND) and left hand (L_HAND). For comparison between groups, the mean value of the two repeated measures was used.

2.4. Pre-processing

Inertial data were acquired and offline processed via Matlab®R2016b (The MathWorks, Inc., Natick, MA, USA).

Table 1
Related works about the combined use of wearable sensors and machine learning algorithms for classification of PD patients and HC, for supporting PD diagnosis. (ACC = Accelerometer, GYR = Gyroscope).

Ref	Technology	Experimental Design	Subjects	Feature extracted	Classifiers	Classifiers Performance
[22]	Smart phone in a custom made glove case (fs = 20 Hz)	Resting task; Postural task (each 30 s)	23 PD, 2 De-Novo PD, 20 HC	Power Spectral Density (PSD)	Bag DT	82% sensitivity, 90% specificity, 0.94 AUC for PD/HC classification
[21]	ACC on wrists, ankles, chest, waist (fs = 62.5 Hz)	Resting tasks (rest in bed and chair, standing up with the hands resting); Postural task; Kinetic tasks (finger-to-nose, finger to finger, walking, picking up/holding objects)	18 PD (10 tremor, 8 other PD symptoms), 5 HC	Dominant frequency (freq) and its amplitude, power spectrum, spectrum entropy, energy	HMM (leave one patient out)	81% accuracy for posture/action detection; 87% accuracy for tremor severity classification; > 95% specificity for tremor/other symptoms discrimination
[20]	ACC, 2 touch sensors on index finger and thumb (fs = 1/0.1 ms)	FT (60 s)	16 PD, 32 HC	Mean and SD of: opening and closing velocity, FT amplitude, total distance of FT, freq, FT interval, finger movement interval, finger contact interval	PCA, logistic regression analysis	Mean of opening velocity and FT total distance decreased, while SD of FT interval increased in accordance with deterioration of the UPDRS FT score. Results reported only in box-plots
[23]	ACC on forearms, arms, thighs, shanks (fs = 100 Hz)	finger-to-nose, FT, hands opening/closing, pronosupination, heel tapping, quiet sitting (30 s, 7 trials)	12 PD (HY 2/3)	Range of amplitude (each channel), RMS value of each ACC signal, peak of the normalized cross-correlation function, time lag to such peak value, dominant freq, freq. ratio, signal entropy	SVM (exponential, rbf, polynomial kernels)	SVM with polynomial kernel had the best performance; 5 s window length optimal to achieve minimum estimation error; Estimation error values of 3.4% for tremor, 2.2% for bradykinesia, 3.2% for dyskinesia
[24]	ACC, wrist-watches on wrists (fs = 20 Hz)	FT, hands opening/closing, pronosupination	12 PD, 12 HC	Range, SD, entropy, time and max freq	Sensitivity, Specificity	83.3% sensitivity, 75% specificity for SD for PD/HC classification
[25]	ACC, GYR on index finger (fs = 128 Hz)	FT as pad-pad and tip-knuckle, pronosupination (each 15 s)	11 PD, 35 HC	Angular velocities	Adaptive filtering algorithms: Ordinary Least Squares, Least Mean Square (LMS), Recursive LS (RLS), Kalman Filter (KF)	AUC 0.781, p = 0.026 for KF in pad-pad FT; AUC 0.828, p = 0.009 for LMS in tip-knuckle FT; AUC 0.869, p = 0.036 for RLS in pronosupination
[19]	GYR on fingertip of index finger	FT (15 s)	10 PD, 10 HC	Cross-sectional areas (CWT analysis)	Quadratic and nearest mean scaled classifiers	94.4% accuracy for quadratic classifier for PD/HC classification
This work	ACC, GYR on wrists, thumb, index and middle fingers	FT, hand opening/closing, pronation/supination, resting tremor, postural tremor, arms swing	30 HC, 30 IH, 30 PD	48 features per limb, including among others: velocity, amplitude, frequency, energy, variations	Optimized SVM, RF, NB	PD/HC F-Measure: 0.947 RF, 0.967 SVM. Best performance for three class classification with RF: 0.803 precision/class, 0.789 recall/class, 0.796 F-measure/class.

Accelerometer and gyroscope data were filtered with a fourth-order low-pass digital Butterworth filter. A 5 Hz cut-off frequency was applied for repetitive exercises (i.e., PSUP, OPCL, THFF, and GTAH) to eliminate high-frequency noise and tremor frequency bands [12,21,26]. Angular rates were integrated using the trapezoidal rule, with sub-intervals of integration equal to 100 ms, to calculate the movement amplitudes. Moreover, a linear drift correction based on the Zero Velocity Update method (ZUPT) was applied step by step to avoid cumulative effects in amplitude calculation [26]. Differently, for HRST and POST, where tremor was investigated, 15 Hz and 20 Hz as cut-off frequency were respectively applied, and an additional fourth-order high-pass digital Butterworth filter was applied with a 0.5 Hz cut-off frequency to avoid the continuous component of the signal due to the static position maintained during these exercises, focusing on the components related to potential tremor. Fast Fourier Transform was applied for frequency analysis implementing the *fft* function provided by Matlab (FFTW library, http://www.fftw.org).

The detailed set of measured parameters and their acronyms for each exercise are reported in Supplementary Table 1 where is also reported the exact sensor that provides useful information for each exercise.

2.5. Feature selection and classification

Combining the three groups of people involved in the study, three datasets were composed, considering IH as an independent class, as healthies, or not considered at all:

- **$2C_{60}$**: 30 HC vs. 30 PD.
- **$2C_{IH}$**: 15 randomly selected IH subjects, considered as healthy persons, and 15 randomly selected HC, obtaining, therefore, a group composed of 30 HC (15 HC plus 15 IH) vs. 30 PD. Testing on this dataset allowed classification of healthy people vs. patients without information about the possible olfactory disorder.
- **$3C_{90}$**: 30 HC vs. 30 IH vs. 30 PD, considering the three groups separately.

Since each extracted parameter resulted in a no parametric distribution according to Kolmogorov-Smirnov test, Kruskal-Wallis test in the three-class case (i.e., dataset $3C_{90}$), and Wilcoxon rank sum test in the two-class case (i.e., datasets $2C_{60}$ and $2C_{IH}$) were performed to evaluate the statistical significance for each feature in distinguishing among the groups. Features with a p-value < 0.05 at least for one side were considered statistically significant and were included in the following analysis. The Spearman's correlation coefficients of the obtained datasets were evaluated to remove highly correlated (rho > 0.85) features. In particular, to create the final dataset, only one feature was kept of the ones with a high correlation coefficient.

Three supervised learning classifiers, namely SVM, RF, and NB, were implemented using functions of Matlab®R2016b to distinguish among the different groups of people, within the aforementioned datasets ($2C_{60}$, $2C_{IH}$, and $3C_{90}$). In particular, regarding the SVM, a third-order polynomial kernel was set, and the hyper-parameters were automatically optimized thanks to the dedicated Matlab function. The analogous optimization process was implemented for NB. Also for the RF, a Matlab function was used, but, in this case, the number of trees was varied using a base two exponential rate with the exponent from 1 to 12 [28]. According to Ref. [28], the Area Under the Curve (AUC) was evaluated. A tradeoff between the AUC and the processing time to obtain it was used as a parameter to choose the optimal number of trees. A ten-fold cross-validation method was applied.

The performances of the classifiers were evaluated in terms of sensitivity or recall, specificity, precision, accuracy, and F-measure calculated as in Ref. [26]. Supplementary Fig. 2 summarizes the methodological approach implemented in this work.

3. Results

This section reports the results obtained from the motor performance assessment using SVM, RF, and NB on three different datasets ($2C_{60}$, $2C_{IH}$, $3C_{90}$).

3.1. Feature selection

Forty-eight features were extracted from upper limbs, both for right and left sides (see Supplementary Table 2 for numerical results, and Supplementary Fig. 3 for graphical results of the trend among the three groups of some of the most significant features). Among them, 39 features per limb produced statistically significant results to differentiate between groups in the $3C_{90}$ dataset, while 35 parameters produced statistically significant results for $2C_{IH}$ and 36 for $2C_{60}$. Thirty-three features were common to all the datasets. According to Spearman's correlation coefficients, highly correlated parameters were removed and final datasets were reduced to 31 features per limb for $3C_{90}$, 27 features for $2C_{IH}$, and 28 features for $2C_{60}$ as highlighted in grey in Supplementary Table 2.

3.2. Classification results

The classification results are reported in Table 2, considering both right and left hand separately, and the hands together.

Focusing on both hands, among the three classifiers, the best results were achieved with RF and SVM when considering the $2C_{60}$ dataset (accuracy and F-measure both equal to 0.95 for RF, and to 0.97 for

Table 2
Comparative results from SVM, RF and NB classifiers in terms of Precision, Recall, Specificity, Accuracy and F-measure for the three different datasets ($3C_{90}$, $2C_{IH}$, $2C_{60}$) calculated on Right Hand, Left Hand and both Hands.

	Precision			Recall			Specificity			Accuracy			F-Measure		
	SVM	RF	NB	SVM	RF	NB	SVM	RF	NB	SVM	RF	NB	SVM	RF	NB
$3C_{90}$															
R_Hand	0.701	0.753	0.722	0.667	0.744	0.711	0.833	0.872	0.856	0.667	0.744	0.711	0.683	0.748	0.716
L_Hand	0.781	0.817	0.687	0.778	0.800	0.689	0.889	0.900	0.844	0.778	0.800	0.689	0.780	0.808	0.688
Hands	0.716	0.803	0.716	0.711	0.789	0.700	0.856	0.894	0.850	0.711	0.789	0.700	0.714	0.796	0.708
$2C_{IH}$															
R_Hand	0.962	0.962	0.893	0.833	0.833	0.833	0.967	0.967	0.900	0.900	0.900	0.867	0.893	0.893	0.862
L_Hand	0.931	0.889	0.897	0.900	0.800	0.867	0.933	0.900	0.900	0.917	0.850	0.883	0.915	0.842	0.881
Hands	0.893	0.963	0.963	0.833	0.867	0.867	0.900	0.967	0.967	0.867	0.917	0.917	0.862	0.912	0.912
$2C_{60}$															
R_Hand	0.967	1.000	0.929	0.967	0.867	0.867	0.967	1.000	0.933	0.967	0.933	0.900	0.967	0.929	0.897
L_Hand	0.967	1.000	0.964	0.967	0.933	0.900	0.967	1.000	0.967	0.967	0.967	0.933	0.967	0.966	0.931
Hands	0.967	1.000	0.933	0.967	0.900	0.933	0.967	1.000	0.933	0.967	0.950	0.933	0.967	0.947	0.933

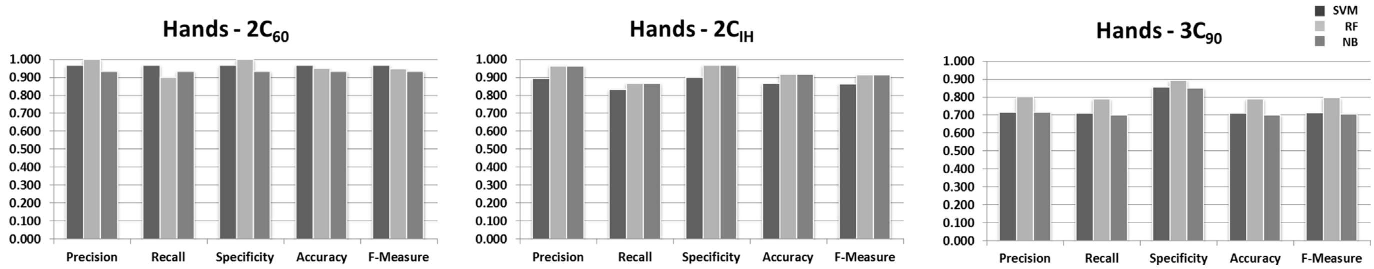
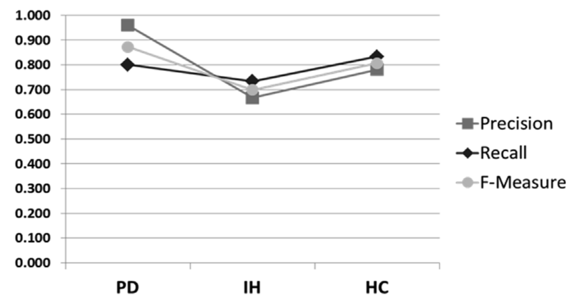


Fig. 1. Comparative classification results for SVM, RF, and NB classifiers, considering both hands in the three datasets $2C_{60}$, $2C_{IH}$, and $3C_{90}$.

(A)

		Predicted Condition		
		PD	IH	HC
True Condition	PD	24	6	0
	IH	1	22	7
	HC	0	5	25



(B)

Fig. 2. Results for $3C_{90}$ dataset for HANDS with RF classifier. (A) Confusion matrix: correct predictions are reported in dark grey, while incorrect predictions are reported in light grey. (B) Obtained values for: Precision/class, Recall/class, and F-Measure/class.

SVM) and with RF and NB for $2C_{IH}$ dataset (accuracy and F-measure equal to 0.92 and 0.91, respectively), while RF gave better performances for the multiclass classification (accuracy equal to 0.79 and F-measure equal to 0.80). Thus, RF appeared as the best classifier (Fig. 1). The results per class are reported for the $3C_{90}$ dataset for both hands in Fig. 2, where PD is the best-identified class, and IH is the worst one, as evidenced also by the confusion matrix. Actually, IH were misclassified both as patients and healthy subjects, while any PD is misclassified as HC and vice versa.

4. Discussion

This work aimed to provide an objective and quantitative assessment of upper limb motion for identifying different classes of people (i.e., HC, IH, and PD). The protocol applied was based on six MDS-UPDRS III tasks. Inertial data collected were processed through ad hoc algorithms for feature extraction. Statistical tests enabled the selection of a feature array as input for ML algorithms. Multiple comparisons were implemented and analyzed in this work, using three different supervised classifiers (i.e., SVM, RF, and NB) that were tested on three datasets (i.e., $2C_{60}$, $2C_{IH}$, and $3C_{90}$), for two-group or three-group classification, based on data from a single limb or both sides (i.e., R_HAND, L_HAND, HANDS).

Since the selected feature array is composed of parameters derived from all six exercises, the use of a comprehensive protocol that does not focus only on a single task is endorsed because it enables a complete analysis of the motor status of subjects, including both motor tasks and tremor analysis. According to literature, the selected feature array should be composed of parameters that are clinically significant and easily understandable for neurologists (e.g., number of movements, amplitudes, and velocities) to significantly enhance support to clinical practice [29,30]. Performances calculated from both hands or a single limb are comparable. Since the PD onset is typically asymmetrical, clinically there is no reason to choose one particular side instead of the other one. Actually, the analysis of both hands is recommended to identify PD even when motor symptoms are unilateral.

The use of ML approaches is appropriate to discriminate among different classes of people. Thus, concerning both hands, the best results were obtained for two-class classification ($2C_{60}$) between HC and PD with RF (0.90 recall, 1.00 specificity) and SVM (0.97 recall and specificity). These results are similar [23] or better than those found for PD/HC discrimination in Refs. [19,22,25] as reported in Table 1. However, no conclusive deduction can be made concerning the most suitable classifier to use since each work used different approaches (e.g., Bag DT, SVM, nearest mean classifiers) and methodologies (single or multi-tasks and different extracted parameters). Nevertheless, the excellent results obtained are promising to sustain the application of the proposed system in clinical practice for supporting clinicians in quantitative assessment and PD diagnosis.

The involvement of a third class of people, such as the IH, allowed the study of the combined datasets in which they could be treated as HC ($2C_{IH}$) or as a separate group ($3C_{90}$). The first group simulates the typical situation where, not knowing whether people are affected by IH, they are labelled as healthy persons, which is clinically reasonable since the possible pathology is latent in them and it will be developed only in 10–12% of IH within 5 years [9]. The results for $2C_{IH}$ achieved the best accuracy (0.92) using RF, which is slightly lower than those obtained in the $2C_{60}$ dataset (i.e., RF 0.95 accuracy) because few IH can have motor performance that is more similar to PD than it is to HC. Finally, the last dataset, $3C_{90}$, allows testing of a multi-group classification, distinguishing between the three classes with an average 0.80 precision and 0.79 recall adopting RF. These results are not directly comparable with any other work in literature, since, to our knowledge, no studies have included motor assessment of IH subjects; nor have they applied ML for multi-class classification based on upper limb motion evaluation. The main difficulties for classification improvement are related to the IH category that shows intermediate motor performance between the HC and PD groups but can be misclassified both as HC and PD, which is clinically justifiable because only a reduced part of the IH will actually develop the disease.

Despite the good results, some limitations are disclosed in this study. First, PD severity or the level of idiopathic hyposmia were not considered, and neither was the correlation with the clinical scale. Thus, it could be interesting, for example, to evaluate whether the IH misclassified as PD have severe olfactory impairments, as well as whether the PD misclassified as IH are mild PD. Second, even if the dataset employed is large enough, these cannot be definitive results because wider samples are required to compute normative data to ensure clinical validation. Thus, the selected feature array should be confirmed with further investigations.

5. Conclusion

In conclusion, the results obtained for two-class analysis are high (0.95 accuracy), and, to the best of our knowledge, this is the first work where motor assessment of IH subjects, as people at risk for developing PD, is evaluated. The system was able to recognize IH as a separate group in a three-class classification, even if it had some difficulty in distinguishing between IH and HC. Although improvements could be applied in future research (e.g., enlarged sample size, IH follow-up, correlation to clinical scale), the good results obtained confirm and show improved results compared to the previous study by the same authors concerning lower limb assessment [26], promoting the idea that quantitative motion evaluation can be a valuable support for neurologists in objective PD diagnosis. Further, by combining motion analysis with a validated olfactory screening test, a two-step, non-invasive, low-cost procedure can be defined to appropriately analyze people at risk for PD development and help clinicians to identify subtle changes in motor performance that characterized PD onset.

Conflicts of interest

The authors declare that they have no conflict of interest.

Authors' contribution

ER was involved in data collection, features extraction, data interpretation, and paper writing. AM was involved in data analysis, data interpretation, and paper writing. DE was responsible for technical development of the used wearable device. CM was the clinical supervisor, involved in protocol writing, experimental session, and interpretation of data. FC was the scientific supervisor, involved in concept design, introduction, and discussion. All authors revised the manuscript and gave final approval of the version to be submitted.

Acknowledgements

This work was financially supported by DAPHNE project (Regione Toscana PAR FAS 2007–2013, Bando FAS SALUTE 2014, CUP J52I16000170002).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.parkreldis.2019.02.028.

References

[1] E.R. Dorsey, B.R. Bloem, The Parkinson pandemic - a call to action, JAMA Neurol 75 (2018) 9–10, https://doi.org/10.1001/jamaneurol.2017.3299.

[2] S. Fahn, Clinical aspects of Parkinson disease, in: R. Nass, S. Przedborski (Eds.), Park. Dis. Mol. Ther. Insights from Model Syst, first ed., Elsevier Inc., 2008, pp. 3–48.

[3] A.H. Schapira, R.K. Chauduri, P. Jenner, Non-motor features of Parkinson disease, Nat. Rev. Neurosci. 18 (2017) 435–450.

[4] D.J. Gelb, E. Oliver, S. Gilman, Criteria for the diagnosis of Parkinson's disease, Arch. Neurol. 56 (1999) 33–39.

[5] C.G. Goetz, B.C. Tilley, S.R. Shaftman, G.T. Stebbins, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, M. Stern, R. Dodel, B. Dubois, R.G. Holloway, J. Jankovic, J. Kulisevsky, A.E. Lang, A.J. Lees, S. Leurgans, P.A. LeWitt, D. Nyenhuis, C.W. Olanow, O. Rascol, A. Schrag, J.A. Teresi, J.J. van Hilten, N. LaPelle, Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results, Mov. Disord. 23 (2008) 2129–2170, https://doi.org/10.1002/mds.22340.

[6] M.M. Hoehn, M.D. Yahr, Parkinsonism: onset, progression, and mortality, Neurology 17 (1967) 427–442, https://doi.org/10.1212/WNL.17.5.427.

[7] J.-A. Palma, H. Kaufmann, Autonomic disorders predicting Parkinson disease, Park. Relat. Disord. 20 (2014) S94–S98, https://doi.org/10.1016/S1353-8020(13)70024-5.

[8] J. Haugen, M.L.T.M. Muller, V. Kotagal, R.L. Albin, R.A. Koeppe, P.J.H. Scott, K.A. Frey, N.I. Bohnen, Prevalence of impaired odor identification in Parkinson disease with imaging evidence of nigrostriatal denervation, J. Neural. Transm. 123 (2016) 421–424, https://doi.org/10.1007/s00702-016-1524-7.Prevalence.

[9] C. Maremmani, G. Rossi, N. Tambasco, B. Fattori, A. Pieroni, S. Ramat, A. Napolitano, P. Vanni, P. Serra, P. Piersanti, M. Zanetti, M. Coltelli, M. Orsini, R. Marconi, C. Purcaro, A. Rossi, P. Calabresi, G. Meco, The validity and reliability of the Italian Olfactory Identification Test (IOIT) in healthy subjects and in Parkinson's disease patients, Park. Relat. Disord. 18 (2012) 788–793, https://doi.org/10.1016/j.parkreldis.2012.03.021.

[10] A. Antonini, R. Benti, R. De Notaris, S. Tesei, A. Zecchinelli, G. Sacilotto, N. Meucci, M. Canesi, C. Mariani, G. Pezzoli, others, 123I-Ioflupane/SPECT binding to striatal dopamine transporter (DAT) uptake in patients with Parkinson's disease, multiple system atrophy, and progressive supranuclear palsy, Neurol. Sci. 24 (2003) 149–150.

[11] C. Atkinson-Clement, S. Pinto, A. Eusebio, O. Coulon, Diffusion tensor imaging in Parkinson's disease: review and meta-analysis, NeuroImage Clin 16 (2017) 98–110.

[12] E. Rovini, D. Esposito, C. Maremmani, P. Bongioanni, F. Cavallo, Using wearable sensor systems for objective assessment of Parkinson's disease, 20th IMEKO TC4 Int. Symp. 18th Int. Work. ADC Model. Test, Benevento, Italy, 2014, pp. 862–867.

[13] E. Rovini, C. Maremmani, F. Cavallo, How wearable sensors can support Parkinson's disease diagnosis and treatment: a systematic review, Front. Neurosci. 11 (2017) 555, https://doi.org/10.3389/fnins.2017.00555.

[14] K. Yang, W.-X. Xiong, F.-T. Liu, Y.-M. Sun, S. Luo, Z.-T. Ding, J.-J. Wu, J. Wang, Objective and quantitative assessment of motor function in Parkinson's disease—from the perspective of practical applications, Ann. Transl. Med. 4 (2016), https://doi.org/10.21037/atm.2016.03.09 90–90.

[15] C. Maremmani, F. Cavallo, C. Purcaro, G. Rossi, S. Salvadori, E. Rovini, D. Esposito, A. Pieroni, S. Ramat, P. Vanni, B. Fattori, G. Meco, Combining olfactory test and motion analysis sensors in Parkinson's disease preclinical diagnosis: a pilot study, Acta Neurol. Scand. 137 (2018) 204–211, https://doi.org/10.1111/ane.12862.

[16] P. Durga, V.S. Jebakumari, D. Shanthi, Diagnosis and classification of Parkinsons disease using data mining techniques, ISSNOnline), Int. J. Adv. Res. Trends Eng. Technol. 3 (2016) 2394–3777 www.ijartet.com.

[17] J. Sujatha, S.P. Rajagopalan, Performance evaluation of machine learning algorithms in the classification of Parkinson disease using voice attributes, Int. J. Appl. Eng. Res. 12 (2017) 10669–10675.

[18] E. Benmalek, J. Elmhamdi, A. Jilbab, Multiclass classification of Parkinson's disease using different classifiers and LLBFS feature selection algorithm, Int. J. Speech Technol. 20 (2017) 179–184.

[19] M. Djurić-Jovičić, V.N. Bobić, M. Ječmenica-Lukić, I. Petrović, S. Radovanović, N.S. Jovicic, V.S. Kostić, M.B. Popović, Implementation of a continuous wavelet transformation in repetitive finger tapping analysis for patients with PD, 22nd Telecommun. Forum, Belgrade, Serbia, 2014, pp. 541–544.

[20] M. Yokoe, R. Okuno, T. Hamasaki, Y. Kurachi, K. Akazawa, S. Sakoda, Opening velocity, a novel parameter, for finger tapping test in patients with Parkinson's disease, Park. Relat. Disord. 15 (2009) 440–444, https://doi.org/10.1016/j.parkreldis.2008.11.003.

[21] G. Rigas, a. T. Tzallas, M.G. Tsipouras, P. Bougia, E.E. Tripoliti, D. Baga, D.I. Fotiadis, S.G. Tsouli, S. Konitsiotis, Assessment of tremor activity in the Parkinson's Disease using a set of wearable sensors, IEEE Trans. Inf. Technol. Biomed. 16 (2012) 478–487, https://doi.org/10.1109/TITB.2011.2182616.

[22] N. Kostikis, M. Arnaoutoglou, C. Kotsavasiloglou, A Smartphone-based tool for assessing Parkinsonian hand tremor, J. Biomed. Heal. Inf. 19 (2015) 1835–1842.

[23] S. Patel, K. Lorincz, R. Hughes, N. Huggins, J. Growdon, D. Standaert, M. Akay, J. Dy, M. Welsh, P. Bonato, S. Member, Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors, IEEE Trans. Inf. Technol. Biomed. 13 (2009) 864–873, https://doi.org/10.1109/titb.2009.2033471.

[24] X. Jia, N. Duroseau, V. Chan, C. Ciraco, R. Wang, S.M. Nia, K. Ho, J.P. Govindavari, F. Delgosha, T. Chan, K.M. Pergament, B. Krishnamachari, A. Farajidavar, Objective quantification of upper extremity motor functions in unified Parkinson's disease rating scale test, 36th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE, Chicago, IL, USA, 2014, pp. 5345–5348.

[25] J.D. Hoffman, J. McNames, Objective measure of upper extremity motor impairment in Parkinson's disease with inertial sensors, 33rd Ann. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE, Boston, MA, USA, 2011, pp. 4378–4381, , https://doi.org/10.1109/IEMBS.2011.6091086.

[26] E. Rovini, C. Maremmani, A. Moschetti, D. Esposito, F. Cavallo, Comparative motor pre-clinical assessment in Parkinson's disease using supervised machine learning approaches, Ann. Biomed. Eng. (2018) 1–12.

[27] F. Cavallo, C. Maremmani, D. Esposito, E. Rovini, P. Dario, M.C. Carrozza, Method and Related Apparatus for Monitoring Biomechanical Performances of Human Limbs, (2014) WO2014108883 A1. PCT/IB2014/058265 of 14/01/2014.

[28] T.M. Oshiro, P.S. Perez, J.A. Baranauskas, How many trees in a random forest? Int. Work. Mach. Learn. Data Min. Pattern Recognit. 2012, pp. 154–168.

[29] R. Matias, V. Paixão, R. Bouça, J.J. Ferreira, A perspective on wearable sensor measurements and data science for Parkinson's disease, Front. Neurol. 8 (2017) 1–7, https://doi.org/10.3389/fneur.2017.00677.

[30] M. Suzuki, H. Mitoma, M. Yoneyama, Quantitative analysis of motor status in Parkinson's disease using wearable devices: from methodological considerations to problems in clinical applications, Parkinson's Dis. 2017 (2017), https://doi.org/10.1155/2017/6139716.