Home Monitoring of Sleep Disturbances in Parkinson’s Disease: A Wearable Solution

Home Monitoring of Sleep Disturbances in Parkinson's Disease: A Wearable Solution

1st Irene Rechichi
Politecnico di Torino
Turin, Italy
irene.rechichi@polito.it

2nd Luca Di Gangi
Politecnico di Torino
Turin, Italy
luca.digangi7@gmail.com

3rd Maurizio Zibetti
Università di Torino
Turin, Italy
maurizio.zibetti@unito.it

4th Gabriella Olmo
Politecnico di Torino
Turin, Italy
gabriella.olmo@polito.it

***Abstract*—Sleep Disorders are the most common and disabling non-motor manifestations of Parkinson's Disease (PD), significantly impairing the quality of life. Monitoring sleep disturbances in PD is a complex task, given the lack of objective metrics and the infrequent neurological assessments. This study proposes a framework for the detection of PD sleep patterns from data collected from 40 subjects (12 PD) through a wearable inertial measurement unit (IMU) during sleep, as well as the automatic assessment of sleep quality. Several features describing overnight motility are proposed and employed in Machine Learning (ML) models to carry out the classification. The best model achieved a 96.2% Accuracy and 93.4% F-1 score in detecting PD subjects from controls, in a Leave-One-Subject-Out cross-validation approach. Sleep quality was assessed with an average accuracy of 79.7% ± 4.4 across the three tested classifiers, and 75% ± 5.25 F-1 score. This suggests the feasibility of characterising overnight motility in PD and effectively monitoring the symptoms' progression through lightweight technology, in a pervasive, e-Health scenario.**

***Index Terms*—Parkinson's Disease, Sleep Disorders, Machine Learning, Telehealth, Wearables**

I. Introduction

The growing incidence of neurodegenerative disorders is gaining an impressive social and healthcare role worldwide, due to their challenging management and the increasing numbers in the aging population. Parkinson's Disease (PD) is the second-most common chronic neurodegenerative disorder, and the fastest growing of the category [1], with a prevalence of up to 10 million people worldwide; this number is projected to double by 2040 [2], as the global population ages.

Nowadays, symptomatic treatment protocols aim at mitigating the motor manifestations, with the support of physical rehabilitation. However, identifying an optimal approach to treatment is challenging, given the unpredictable nature of the disease, as well as the diversity in its symptomatic manifestations, which usually worsen as the disease progresses. PD cardinal features include motor symptoms, such as tremor, rigidity, and bradikynesia [3]. However, non-motor symptoms (NMS) are also highly preponderant in people with PD, and dominate the clinical scenario in advanced PD [4]; they are often poorly recognised and not treated adequately. They entail a broad spectrum, including autonomic disorders, cognitive impairment, behavioural symptoms (such as depression), speech alterations and sleep dysfunctions.

These latter are acknowledged as the most common NMS, with a prevalence up to 90% [5]. Among these, nocturnal hypokinesia – i.e., reduced motility during the night – and morning akinesia affect up to 50% of the subjects already in the earliest stages of the disease [6]. Sleep disorders in PD have also been acknowledged among the most disabling symptoms, with significant effects on the quality of life (QoL) [4].

Nowadays, a critical issue in the management of PD lies in the infrequent neurological assessments, usually scheduled on a 6- to 12-month basis, leading to potential loss of information about daily and monthly fluctuations. Besides, clinical evaluation scales mainly rely on subjective, self-reported metrics [7].

Additionally, diagnosing and monitoring sleep disorders is a lengthy and challenging task. The gold standard is polysomnography (PSG), a diagnostic test performed in a sleep laboratory, which consists in recording biosignals during sleep through numerous electrodes. Though very precise, this technique presents with some disadvantages (i.e., unfamiliar sleeping environment, cumbersome instrumentation), that may significantly hinder diagnosis and follow-up [8].

This introduces the need for effective monitoring of the symptoms' fluctuations by means of objective parameters, possibly through lightweight, pervasive and continuous solutions [9]. Sleep actigraphy represents a possible, minimally invasive solution [10]; it encompasses the use of an inertial device worn on the wrist, that records movements overnight, and enables objective metrics to evaluate sleep.

This work proposes a framework for the automatic detection of PD sleep patterns, as well as classification of good or bad sleep quality (SQ), based on objective and reproducible metrics. The study exploits motility data collected through a wearable, inertial device during sleep on healthy and PD subjects, and Machine Learning (ML) methods to carry out the classification task.

II. Materials

A. Study Participants

The study involved both PD subjects and healthy controls, to assess the ability of the proposed system to characterise motility during sleep, and support follow-up procedures in PD.

Eligible subjects were recruited at the Parkinson Unit (Dept. Neurology, AOU Città Della Salute e della Scienza, Turin), and at the patients' association *Associazione Amici Parkinsoniani ONLUS*, in Turin, Italy. Inclusion criteria for patients were

TABLE I: Demographic characteristics of the population in the study. HC: healthy controls, PD: subjects with Parkinson's Disease.

	Sample	Age	sPSQI	SLEEPS
HC	28 (10 females)	38 ± 10.7 years	6.55 ± 1.27	2.55 ± 1.02
PD	12 (5 females)	68 ± 4.1 years	9.42 ± 4.06	3.68 ± 2.21

defined together with an expert neurologist, specialised in PD and sleep disorders (M.Z.), and included: combined diagnosis of PD and sleep-related disorders, both motor (e.g., nocturnal akinesia, bradykinesia) and non-motor (e.g., excessive daytime sleepiness, low level of activity). Controls were selected on a voluntary basis, among patients' spouses, family members, and respondents to the University's data collection campaign; inclusion criteria required absence of (or familiarity for) Parkinsonisms and other neurodegenerative diseases, absence of diagnosed sleep disorders. Data were collected remotely and data processing was carried out offline; all recordings were conducted at home, in unsupervised settings.

As commonly employed in the clinical practice, participants were asked to fill in two validated questionnaires (further detailed in Section III-A), to investigate on sleep quality and circadian health, and correlate self-reported items with objective metrics. For future analysis, vocal recordings were collected through a smartphone, following the method implemented in a previous work [11].

Data collection was carried out in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the A.O.U. Città della Salute e della Scienza di Torino (Approval Number: 00384/2020); written informed consent for observational study was obtained from all participants. A total of 40 subjects (12 PD) enrolled in the study and took part in the data collection; demographic data are summarised in Table I.

B. Experimental Protocol: Data Acquisition, Sensor Placement

The experimental protocol involved remote collection of motion data during sleep, to characterise sleep motility in Parkinson's Disease. To ensure a successful data acquisition, given the unsupervised approach of the study, all participants were instructed in the use of the device (activation, positioning) and the sleep questionnaires. A user guide, including the complete experimental protocol and information about the device, was also provided to each participant.

Data were collected during the night through a lightweight, commercially-available inertial measurement unit (IMU) (Shimmer3, ShimmerSensing). The device includes the following triaxial sensors: accelerometer, gyroscope, magnetometer, which record variations in linear acceleration, angular velocity, and magnetic field, respectively, along the three axes in space (x, y, z). The sampling frequency was 128 Hz, and the average duration of recordings in this population was of 7.13 ± 0.07 hours. The IMU was worn on the chest through an elastic strap (Figure 1), in order to properly collect motion data from the trunk and characterise whole-body movements [12], which fail to be described in traditional wrist actigraphy.

III. Methods

A. Assessment of Sleep Quality and Sleep Metrics

Two clinically validated questionnaires were employed in the study in order to retrieve self-reported metrics for sleep quality and information on circadian habits.

First, the shortened Pittsburgh Sleep Quality Index (sPSQI) [13] was administered to all study participants, as ground truth for the assessment of overall sleep quality. This survey consists in a shortened (13-item survey), self-reported version of the PSQI survey, commonly employed in the clinical practice to assess sleep quality. The selected 13 questions evaluate sleep health over the following five axes: Sleep Latency, Duration, Efficiency, Disturbances and Daytime Dysfunction, each yielding a relative score. The obtained global score – i.e., the sum of all relative scores – is used to differentiate between *good* and *bad* sleep quality, and is defined in the range (0, 15), with 15 indicating the negative extreme. In the population under study a global score of 7.59 ± 5.13 was obtained; values for the two separate groups are reported in Table I. Following the approach adopted in [11], and after statistical evaluations on the current population, a value of 6.0 was chosen as cut-off between good and bad sleep quality, thus yielding 22 *bad sleepers* and 18 *good sleepers*, overall.

Finally, the SLEEPS [11], a validated, 21-item survey was administered to investigate circadian health. The survey covers the following four areas: (i) General Health, (ii) Work/Study Habits, (iii) Leisure Time Habits, (iv) Sleep Habits, providing a numerical score for each, as well as an overall score (Table I), to be later employed as features in the Machine Learning pipeline (cf. Section III-B2).

B. Data Analysis

Data processing, feature extraction the subsequent analysis were carried out through custom-written MATLAB® (R2022b) and Python code.

1) Data Pre-Processing: Inertial IMU data were pre-processed prior to feature extraction; for the purpose of this work, only data from the triaxial accelerometer and gyroscope were included in the analysis, as they provide enough information regarding body position, axial movements, and overall motility.

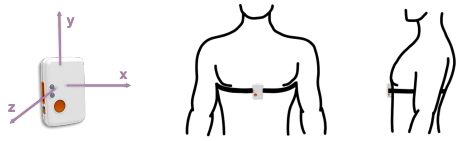

Fig. 1: Sensor placement adopted in the study, along with proper axes of reference.

Raw accelerometry data (acquired as m/s^{-2}) were pre-processed by means of a moving-average filter, with a 1-second sliding window; this method acts as a finite-impulse response (FIR) lowpass filter, to cut out high-frequency noise and lower the effect of abrupt movements. Likewise, data acquired through the gyroscope (deg/s) were pre-processed with a FIR lowpass filter (cutoff frequency: 35 Hz, order: 21), to ensure a reliable detection of turns and velocity of rolling-over.

2) Feature Extraction: As previously introduced, various features were extracted from the available experimental data, with the aim of characterising axial movements and motility in bed.

Given the fact that, up to date, no standardised set of features to characterise Parkinson's Disease sleep patterns is available in the literature, a combination of clinical (polysomnography-related) parameters and motility features was employed in this study. These latter included both novel features, and features previously proposed in night-accelerometry or actigraphy studies [8], [12].

Features related to motion in bed included the reclining angle, the time spent in each sleeping position, the number of turns in bed, and the velocity of turns, and are described in the following paragraphs.

First, sleeping position was assessed in 30-second epochs (following PSG scoring standards), through the inspection of accelerometry data, by taking into account the tilt angle (θ, i.e., the reclining angle in bed) and the mean acceleration in the three directions.

Subsequently, turning events (i.e., turning in bed) were identified. Any detected change in sleeping position is first marked as *turn-candidate*. The event is then labelled as a true turning event if the same sleeping position is maintained for 2 minutes before and after its occurrence, therefore ensuring accurate detection of axial rolling-over. Finally, the number of events (N_{turns}), and the time interval between events (R_{int}) were extracted.

In addition, from the inspection of the angular velocity along the longitudinal axis (y, gyroscope), information regarding the velocity of turns was retrieved. A preliminary peak detection was carried out, by selecting all peaks with amplitude above 85% of the standard deviation of the signal; this threshold was heuristically selected. Then, for each turning event, a 50-second window was chosen as search range for the actual angular velocity peak corresponding to the selected rolling-over (Figure 2). Finally, peak height (velocity of turning, ω-turns) and peak width (duration of turns, T_{turns}) were computed.

To characterise overall night motility, the Activity Index (AI) [14], accounting for variance across the three accelerometry axes (x, y, z), was computed (Equation III-B2). Given that the employed IMU includes a wide-range accelerometer, systematic noise removal was performed prior to the AI computation. The systematic noise (σ_{sys}) was evaluated as the triaxial variance measured when the sensor is in a still, horizontal position, and is averaged across 30-second epochs; for the employed IMU, σ_{sys} measured 27.5 ms^{-2}.

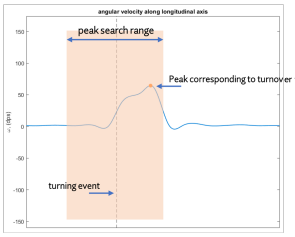

Fig. 2: Angular velocity in the longitudinal axis (y). Peak detection method for computing the velocity and duration of turnover events.

$$AI = \sqrt{\frac{1}{3}[(\sigma_x^2 - \sigma_{sys}) + (\sigma_y^2 - \sigma_{sys}) + (\sigma_z^2 - \sigma_{sys})]}, \quad (1)$$
$$\sigma_{sys} = \sigma_x + \sigma_y + \sigma_z$$

In addition, a novel metric derived from the AI was proposed and employed in this study, as a descriptor of motility trends. This metric is the Average Motility (AM) and it is computed as the moving average of the AI in a 2-minute window, a time-frame considered appropriate for movements during the night, and later mapped to a continuous value in the range [0, 1]. Values of 1 (or close to) represent high overnight activity; its trend is displayed in Figure 3.

Table II displays the set of computed features; for each *Motility* parameter, various statistics (mean, standard deviation, maximum value, minimum value, 25^{th} and 75^{th} percentiles, kurtosis, skewness) were computed and employed as separate features.

3) Feature Inspection and Selection: Given that the set of engineered features included both novel and non-novel features, and that early fusion was performed among features of different categories (Table II), statistical tests were performed on the features to test for their relevance. All statistical analyses were carried out through the open-source tool Jamovi [17]. First, feature normality was assessed through the Shapiro-Wilk test; then, distribution testing was carried out by means of the Student's t test and the non-parametric Mann-Whitney U test for normally- and non-normally distributed features, respectively.

First, the two following configurations were tested: (i) HC vs PD, (ii) Good vs Bad SQ. Second, with the aim of investigating a possible influence of the disease on overall SQ, two additional

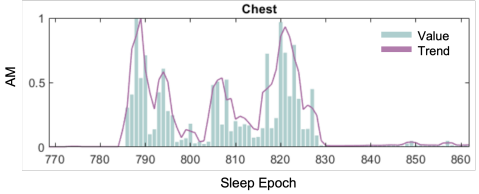

Fig. 3: Portion of the Average Motility variable (trunk movements) for a PD patient.

TABLE II: Features employed in the study, along with their category and proper reference. ⋄: adapted from cited study; ⋆: first proposed in this study.

Feature	Description	Reference
Clinical		
Sleep Onset Latency (SOL)	The amount of time required to fall asleep (min)	various
Wake After Sleep Onset (WASO)	The amount of time the subject is awake during the night (min)	various
Total Sleep Time (TST)	Total hours of sleep	various
Time in bed (TIB)	Lights-off to lights-on interval (h)	various
Sleep Efficiency (SE)	The percentage of time spent asleep while in bed (%)	various
SLEEPS score	Perceived sleep health and quality	[11]
Motility		
Tilt Angle (θ)	Reclining angle in bed	⋆
Sleeping position	Minutes spent in each sleeping position (*supine*, *prone*, *left-side*, *right-side*)	⋆
Number of turns (N_{turns})	Number of turns in bed	⋄ [15]
Rotation interval (R_{int})	Interval between turning events (min)	⋄ [15]
Rotation velocity (ω-turns)	Velocity of turning in bed (deg/s)	⋄ [16]
Rotation acceleration (α-turns)	Acceleration of turning in bed (deg/s^2)	⋄ [16]
Turning duration (T_{turns})	Duration of each turn (s)	⋄ [16]
Stand/Sit Duration (SSD)	Total time spent standing or sitting during the night (min)	⋆
Activity Index (AI)	Level of activity during the night. Range: [0, 1]	[15]
Average Motility (AM)	Overnight motility trend.	⋄ [14]

configurations were tested: (iii) HC with good SQ vs HC with bad SQ (HC_{good} vs HC_{bad}), (iv) HC with bad SQ vs PD with bad SQ (HC_{bad} vs PD_{bad}). No comparison between PD with good SQ and PD with bad SQ was carried out, due to the scarce numerosity of the former group.

For the sake of completeness, correlation between the sleep features and the sPSQI score – i.e., the clinical indicator of SQ – was also tested through Spearman's Correlation; likewise, correlation to the presence of PD was investigated.

Feature selection was carried out by means of the ReliefF algorithm [18], in the configurations (i) and (ii) (the two explored classifications); this step allowed for the reduction of the dataset to be employed in the classification step. The top-K features relevant for the task were kept; the parameter K was heuristically chosen by identifying the elbow on the feature importance scores yielded by the algorithm. For classification, only the selected features were employed in the ML models. Prior to the feature selection process, z-score normalisation was adopted to prevent outliers from affecting the subsequent analysis.

C. Classification

Supervised ML models were employed to automatically discriminate between: (i) HC and PD (to test the capability of the system to detect PD from night-motility data), and (ii) Good vs Bad SQ.

TABLE III: Summary of the employed classifiers and the searched hyperparameters, (parameter and range).

Model	Searched Hyperparameters
SVM	*Kernel function*: linear, polynomial, radial basis, sigmoid *Penalty (C)*: [0.1, 1, 10, 100, 1000] γ: [1, 0.1, 0.001, 0.0001]
KNN	*Minkowski Distance order (p):* [1, 2, 3, 4, 5] *Number of neighbours (K)*: [3, 5, 7] *Weights (W)*: uniform, distance-based
XGBoost	*Number of trees*: [25, 50, 100] *Depth*: [3, 5, 7] *Learning rate*: [0.001, 0.01, 0.1]

Three different models were tested in both configurations; namely, a Support Vector Machine (SVM), a K-Nearest Neighbour (KNN), and eXtreme Gradient Boosting (XGBoost) – this latter being an ensemble method based on decision trees.

Hyperparameters were optimised with a Grid Search approach (50 iterations, Table III), to increase model robustness; within the optimisation process, the F-1 score was elected as metric for model comparison, due to the imbalanced class cardinality in the adopted dataset.

Moreover, given the scarce numerosity of the dataset, a Leave-One-Subject-Out cross-validation (LOSO-CV) approach was adopted for model evaluation and performance comparison across models, to allow for better generalisation capability of the tested classifiers and to limit the effect of overfitting.

Within this framework, at each iteration, one subject is held-out for testing, and the remaining $(N-1)$ subjects are used in the training process. The procedure is repeated for a total of N iterations (where N equals the number of subjects in the dataset). Hence, a total of $N-1$ models are trained. Finally, model performance metrics, namely Accuracy, Recall, and F-1 score (harmonic mean of Precision and Recall) were used to evaluate results.

The complete analysis pipeline is displayed in Figure 4.

IV. Results

A. Statistical Analysis

Several of the variables proposed to describe overnight motility, and employed in this study, proved their discriminative power in characterising PD sleep patterns, as well as sleep quality (cf. Section III-B3, configurations $(i), (ii)$).

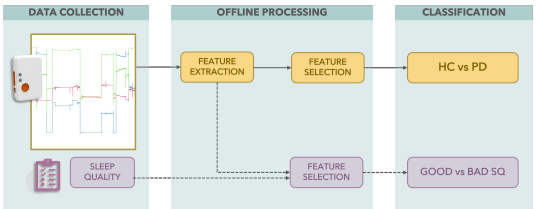

Fig. 4: Overall analysis pipeline of the proposed framework.

Table IV displays the tests results. As appreciable, when testing the HC vs PD configuration, statistical significance was observed in various features characterising overnight body position (θ_{mean}, θ_{p75}), whole-body movements (ω-turns$_{p25}$, ω-turns$_{skew}$), and overall motility (AI$_{skew}$, AM$_{mean}$). Moderate correlation, in terms of Spearman's ρ, with the presence of the disease was also observed for the listed features, with θ_{mean} featuring a value of 0.54; a negative correlation, as expected, was observed with features describing the velocity of turns in bed.

Regarding the Good vs Bad SQ configuration, statistical significance was observed in the variables describing the duration of turns (T$_{turns,std}$, T$_{turns,skew}$), and overnight motility (AI$_{p25}$, AM$_{mean}$); these latter also showed a moderate negative correlation with the sPSQI score, with values of -0.46 and -0.44, respectively.

Finally, to carry out a proper population inspection, and to counteract the possible effect of bias, the initial population was stratified according to sleep quality and the presence of PD, and the other two configurations ((iii) and (iv) in Section III-B3) were investigated.

In particular, in the configuration HC$_{good}$ vs HC$_{bad}$, the features T$_{turns,std}$ and ω-turns$_{p25}$ (turn velocity) resulted significant (p<0.005 and p<0.05, respectively), with Spearman's ρ values (with sPSQI) of 0.45 and -0.64, respectively. Finally, when testing HC$_{bad}$ vs PD$_{bad}$, body position (θ_{mean}), velocity of turns (ω-turns$_{p75}$, ω-turns$_{mean}$) and overall motility (AI$_{mean}$, AM$_{mean}$) resulted significant (p<0.001). These also presented with moderate-to-high correlation with sPSQI, with ρ of 0.64 for θ_{mean}, and -0.51 for AM$_{mean}$. As for velocity of turns, Spearman's ρ was of -0.41 and -0.35 (ω-turns$_{p75}$ and ω-turns$_{mean}$, respectively).

These results, prior to the feature selection process, suggest that whole-body motility is a good descriptor of nocturnal patterns, both for sleep quality and characterisation of sleep in PD.

TABLE IV: Independent Sample statistics of the features employed in the classification, along with their correlation with the target (PD or sPSQI). Mark $*$: high statistical significance.

HC vs PD		
Feature	**Independent Sample Test**	**Correlation** (ρ)
θ_{mean}	<0.005*	0.54
θ_{p75}	<0.001*	0.45
R$_{int,p25}$	<0.05	-0.31
ω-turns$_{p25}$	<0.001*	-0.41
ω-turns$_{skew}$	<0.005*	-0.39
AI$_{skew}$	<0.001*	0.31
AM$_{mean}$	<0.001*	0.32
Sleep Quality		
Feature	**Independent Sample Test**	**Correlation** (ρ)
ω-turns$_{p75}$	<0.05	0.32
T$_{turns,td}$	<0.005*	0.49
T$_{turns,skew}$	<0.001*	0.55
AI$_{p25}$	<0.05	-0.46
AM$_{mean}$	<0.05	-0.44

TABLE V: Classification task: Healthy Controls vs PD Subjects. Performance metrics of the optimised classifiers, employing a LOSO-CV.

	SVM	**KNN**	**XGBoost**
Accuracy	96.2 %	90.4 %	80.7 %
Recall	95.0 %	72.3 %	85.2 %
F-1	93.4 %	80.6 %	70.0 %

TABLE VI: Classification task: Good vs Bad Sleep Quality. Performance metrics of the optimised classifiers, employing a LOSO-CV.

	SVM	**KNN**	**XGBoost**
Accuracy	78.1 %	75.2 %	85.7 %
Recall	74.0 %	73.3 %	78.6 %
F-1	72.0 %	70.8 %	82.5 %

B. Feature Selection

Feature selection was carried out by means of the ReliefF algorithm, in order to find the subset of most discriminative features for each classification task – i.e., HC vs PD and Good vs Bad SQ. The features selected for the classification are shown in Table IV; after inspecting the scores provided by the ReliefF algorithm and identifying the *elbow*, a subset of $K = 7$ and $K = 5$ was chosen for the two tasks. As appreciable, for both classification tasks, the most relevant features were those related to trunk or whole-body movements. This suggests that the sensor placement proposed in this experimental study is suitable to describe overnight motility and sleep quality in PD.

C. Classification Performance

Table V displays the results obtained by the tested classifiers in the configuration HC vs PD; performance metrics are obtained by means of the LOSO-CV and displayed as Accuracy, Recall and F-1 score. As appreciable, all classifiers feature moderately high accuracy, with an average of 89.1% $\pm$ 6.39. The same trend is observed for the F-1 score (81.3% $\pm$ 9.57), suggesting good classification performance. The optimised SVM (C=1, kernel: linear) emerged as best model, with overall accuracy above 96%, accuracy of 95%, as well as an F-1 score of 93.4%. Notably, given the unbalanced cardinality of the dataset (higher number of HC), a high F-1 score ensures model robustness against both false positives and false negatives, thus implying the good predictive power of the employed features in detecting PD from overnight motility patterns.

Likewise, performance metrics for the Good vs Bad SQ classification task are shown in Table VI. Though featuring a slight decrease in performance, the employed models yielded an overall accuracy of 79.67% $\pm$ 4.43, which is indicative of fairly good classification performance. The best scores were attained through a XGBoost classifier (N$_{trees}$=50, depth=5, learning rate=0.1), featuring an accuracy of 85.7%, 78.6% recall, and F-1 score of 82.5%. These results are strongly suggestive of good classification performance, and reflect the good discriminative power of the features employed in the task.

V. Conclusions and Future Work

This study investigated the capability of a lightweight system of monitoring sleep disorders in PD, thus exploring the feasibility of e-Health applications in the scenario of disease management and follow-up. The results attained by the explored ML algorithms proved the efficiency of the objective parameters in the automatic detection of PD subjects, as well as good or bad sleep quality, based on overnight motion patterns recorded through wearable sensors.

Indeed, the proposed system proved effective in detecting PD subjects from motility parameters, with the best ML model attaining an accuracy and F-1 score of 96.2% and 93.4%, respectively. This also reflects the importance of whole-body motion features, rather than relying on traditional wrist-based assessment. To the best of the Authors' knowledge, this is the first study that tackles automatic PD detection from overnight motility patterns through ML.

Finally, automatic classification of sleep quality has also been addressed, based on the same metrics; the best model (XGBoost) achieved an overall accuracy of 85.7%, and a F-1 score of 82.5%. The performance compares well with the literature, though a direct comparison is impractical, as most studies are based on wrist actigraphy and healthy cohorts. Virtually, the results are in line with [19], which reports an average accuracy and F-1 score of 87.9% and 81.9%, respectively. However, the cited study only includes data from young, healthy subjects, and employs deep learning models – therefore, more complex than the one proposed in this work. Comparably, in [20], the Authors exploit ML models though and obtaining poorer results than this study, through a 8-fold CV (accuracy: 72.5%, F-1 N/A), on a young, healthy cohort.

Future work will address the limitations of this study, such as including a larger PD population and stratification according to different disease progression levels, to improve the generalisation capability of the tested models and enhance performance. Future investigations will also include the analysis of muscle activations during sleep and vocal recordings [11], to provide a multi-modal approach for the evaluation of motor and NMS in PD, allowing for personalised treatment. To conclude, this study, though preliminary, provided a framework for minimally invasive sleep studies in unsupervised settings, supporting pervasive, continuous monitoring solutions, thus counteracting the limitations of infrequently scheduled outpatient assessments and possibly improving the QoL of people with PD.

Acknowledgment

The Authors thank the volunteers who took part in the study, and the *Associazione Amici Parkinsoniani Piemonte Onlus* (Turin, Italy) for the fruitful cooperation. This publication is part of NODES, which has received funding from the Italian Ministry for Universities and Research – M4C2 1.5 of PNRR, funded by the European Union NextGenerationEU (grant agreement No. ECS00000036).

References

[1] E. Dorsey, T. Sherer, M. S. Okun, and B. R. Bloem, "The emerging evidence of the parkinson pandemic," *Journal of Parkinson's disease*, vol. 8, no. s1, pp. S3–S8, 2018.

[2] D. K. Simon, C. M. Tanner, and P. Brundin, "Parkinson disease epidemiology, pathology, genetics, and pathophysiology," *Clinics in geriatric medicine*, vol. 36, no. 1, pp. 1–12.

[3] S. Sveinbjornsdottir, "The clinical symptoms of parkinson's disease," *J. Neurochem.*, vol. 139, pp. 318–324, 2016.

[4] K. R. Chaudhuri, D. G. Healy, and A. H. Schapira, "Non-motor symptoms of parkinson's disease: diagnosis and management," *The Lancet Neurology*, vol. 5, no. 3, pp. 235–245, 2006.

[5] A. Stefani and B. Högl, "Sleep in parkinson's disease," *Neuropsychopharmacology*, vol. 45, no. 1, pp. 121–128, 2020.

[6] N. Tambasco, S. Simoni, P. Nigro, F. P. Paoletti, E. Marsili, and P. Calabresi, "Morning akinesia in parkinson's disease: challenges and solutions," *Journal of Parkinsonism and Restless Legs Syndrome*, pp. 57–63, 2016.

[7] C. G. Goetz, B. C. Tilley, S. R. Shaftman, Stebbins, *et al.*, "Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results: MDS-UPDRS: Clinimetric assessment," *Mov. Disord.*, vol. 23, no. 15, pp. 2129–2170, 2008.

[8] C. J. Madrid-Navarro, F. J. Puertas Cuesta, F. Escamilla-Sevilla, M. Campos, F. Ruiz Abellán, M. A. Rol, and J. A. Madrid, "Validation of a device for the ambulatory monitoring of sleep patterns: a pilot study on parkinson's disease," *Frontiers in neurology*, vol. 10, p. 356, 2019.

[9] C. Godinho, J. Domingos, G. Cunha, A. T. Santos, R. M. Fernandes, D. Abreu, N. Gonçalves, H. Matthews, T. Isaacs, J. Duffen, *et al.*, "A systematic review of the characteristics and validity of monitoring technologies to assess parkinson's disease," *Journal of neuroengineering and rehabilitation*, vol. 13, no. 1, pp. 1–10, 2016.

[10] J. E. Maglione, L. Liu, A. B. Neikrug, T. Poon, L. Natarajan, J. Calderon, J. A. Avanzino, J. Corey-Bloom, B. W. Palmer, J. S. Loredo, *et al.*, "Actigraphy for the assessment of sleep measures in parkinson's disease," *Sleep*, vol. 36, no. 8, pp. 1209–1217, 2013.

[11] F. Amato, I. Rechichi, L. Borzì, and G. Olmo, "Sleep quality through vocal analysis: a telemedicine application," in *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pp. 706–711, IEEE, 2022.

[12] J. Sringean, P. Taechalertpaisarn, C. Thanawattano, and R. Bhidayasiri, "How well do parkinson's disease patients turn in bed? quantitative analysis of nocturnal hypokinesia using multisite wearable inertial sensors," *Parkinsonism & Related Disorders*, vol. 23, pp. 10–16, 2016.

[13] O. A. Famodu, M. L. Barr, I. Holásková, W. Zhou, J. S. Morrell, S. E. Colby, and M. D. Olfert, "Shortening of the pittsburgh sleep quality index survey using factor analysis," *Sleep disorders*, vol. 2018, 2018.

[14] J. Bai, C. Di, L. Xiao, K. R. Evenson, A. Z. LaCroix, C. M. Crainiceanu, and D. M. Buchner, "An activity index for raw accelerometry data and its comparison with other activity metrics," *PloS one*, vol. 11, no. 8, p. e0160644, 2016.

[15] M. Louter, W. Maetzler, J. Prinzen, *et al.*, "Accelerometer-based quantitative analysis of axial nocturnal movements differentiates patients with parkinson's disease, but not high-risk individuals, from controls," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 86, no. 1, pp. 32–37, 2015.

[16] J. Sringean, P. Taechalertpaisarn, C. Thanawattano, and R. Bhidayasiri, "How well do parkinson's disease patients turn in bed? quantitative analysis of nocturnal hypokinesia using multisite wearable inertial sensors," *Parkinsonism & Related Disorders*, vol. 23, pp. 10–16, 2016.

[17] "Jamovi - open statistical software for the desktop and cloud." https://www.jamovi.org. Accessed: 2022-12-19.

[18] R. J. Urbanowicz, M. Meeker, W. La Cava, R. S. Olson, and J. H. Moore, "Relief-based feature selection: Introduction and review," *Journal of biomedical informatics*, vol. 85, pp. 189–203, 2018.

[19] A. Sathyanarayana, S. Joty, L. Fernandez-Luque, F. Ofli, J. Srivastava, A. Elmagarmid, T. Arora, S. Taheri, *et al.*, "Sleep quality prediction from wearable data using deep learning," *JMIR mHealth and uHealth*, vol. 4, no. 4, p. e6562, 2016.

[20] O. V. Bitkina, J. Park, and J. Kim, "Modeling sleep quality depending on objective actigraphic indicators based on machine learning methods," *International Journal of Environmental Research and Public Health*, vol. 19, no. 16, p. 9890, 2022.