MDPI

Article

Simplification of Mobility Tests and Data Processing to Increase Applicability of Wearable Sensors as Diagnostic Tools for Parkinson's Disease

Rana M. Khalil [1], Lisa M. Shulman [2], Ann L. Gruber-Baldini [3], Sunita Shakya [3], Rebecca Fenderson [2,†], Maxwell Van Hoven [2,‡], Jeffrey M. Hausdorff [4,5,6,7,8], Rainer von Coelln [2,*] and Michael P. Cummings [1,*]

1 Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD 20742, USA; rmkhalil@umd.edu
2 Department of Neurology, University of Maryland School of Medicine, Baltimore, MD 21201, USA; lshulman@som.umaryland.edu (L.M.S.); rfenders12@gmail.com (R.F.); mvanhoven2@gmail.com (M.V.H.)
3 Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD 21201, USA; abaldin@som.umaryland.edu (A.L.G.-B.); sshakya@som.umaryland.edu (S.S.)
4 Center for the Study of Movement, Cognition, and Mobility, Neurological Institute, Tel Aviv Medical Center, Tel Aviv 6492416, Israel; jeffh@tlvmc.gov.il
5 Department of Physical Therapy, Faculty of Medicine & Health Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
6 Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 6997801, Israel
7 Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL 60612, USA
8 Department of Orthopedic Surgery, Rush University Medical Center, Chicago, IL 60612, USA
* Correspondence: rvoncoelln@som.umaryland.edu (R.v.C.); mcummin1@umd.edu (M.P.C.)
† Current address: Department of Emergency Medicine, Cooper Medical School, Rowan University, Camden, NJ 08103, USA.
‡ Current address: Department of Family Medicine, The Warren Alpert Medical School, Brown University, Pawtucket, RI 02912, USA.

Abstract: Quantitative mobility analysis using wearable sensors, while promising as a diagnostic tool for Parkinson's disease (PD), is not commonly applied in clinical settings. Major obstacles include uncertainty regarding the best protocol for instrumented mobility testing and subsequent data processing, as well as the added workload and complexity of this multi-step process. To simplify sensor-based mobility testing in diagnosing PD, we analyzed data from 262 PD participants and 50 controls performing several motor tasks wearing a sensor on their lower back containing a triaxial accelerometer and a triaxial gyroscope. Using ensembles of heterogeneous machine learning models incorporating a range of classifiers trained on a set of sensor features, we show that our models effectively differentiate between participants with PD and controls, both for mixed-stage PD (92.6% accuracy) and a group selected for mild PD only (89.4% accuracy). Omitting algorithmic segmentation of complex mobility tasks decreased the diagnostic accuracy of our models, as did the inclusion of kinesiological features. Feature importance analysis revealed that Timed Up and Go (TUG) tasks to contribute the highest-yield predictive features, with only minor decreases in accuracy for models based on cognitive TUG as a single mobility task. Our machine learning approach facilitates major simplification of instrumented mobility testing without compromising predictive performance.

Keywords: Parkinson's disease; wearable sensors; machine learning; mobility testing; diagnosis; movement disorders; movement analysis

1. Introduction

Parkinson's disease (PD) is the most common neurodegenerative movement disorder and affects an estimated 7–10 million people worldwide [1]. It is characterized by a range of motor symptoms, including tremors [2,3], bradykinesia [4,5], rigidity [4,6], a slow shuffling gait with small steps [7–9], freezing of gait [10,11], and imbalance [7,12]. Among all these

motor symptoms, gait impairment and imbalance most substantially impact the quality of life of PD patients, ultimately resulting in disability and loss of independence [13,14].

Currently, PD diagnosis is based on clinical assessment, including history and neurological examination, which is inherently subjective in nature and depends on the level of experience of the healthcare provider. Consequently, initial clinical diagnosis by movement disorders experts has limited accuracy (79.6%, increasing to 83.9% at follow-up) and limited overall validity [15]. The need for objective, low-cost, reliable, and valid biomarkers that can accurately identify patients with PD was identified as one of the top priorities in PD research by a National Institutes of Health/National Institute of Neurological Disorders and Stroke (NIH/NINDS) panel [16].

Wearable sensors containing accelerometers and gyroscopes show great promise in objectively detecting the presence of key PD motor symptoms in a convenient and accessible manner [17–21]. The widespread availability of wearable sensors makes it possible to quantitatively assess mobility without costly equipment or a dedicated gait laboratory [22]. Used in instrumented mobility testing, these sensors provide highly granular data on gait, balance, and other motor functions to identify features and patterns indicative of PD [23,24]. Ultimately, the expectation is for wearable sensors to support earlier detection and treatment of PD, leading to better outcomes for patients [18,25]. However, critical challenges have yet to be overcome in order to turn sensors from research instruments into clinical tools widely used to support the diagnosis of PD in clinical practice [26,27].

Studies on the use of wearable sensors to diagnose PD vary widely in terms of which and how many sensors to use and where to place them [28–35]. There is also a large variation in which and how many instrumented mobility tasks participants are required to perform [29,36–38]. Moreover, the methods employed to process sensor data and derive relevant features are complex and vastly different across studies [29,39–42].

We conducted a search of the literature from 2017 to 2024 to identify previous studies that used wearable sensors for PD assessment. This involved reviewing three prior review articles [33–35] and performing a search of the PubMed database with the keywords "(wearable OR IMU) AND Parkinson", which yielded 72 relevant articles. Our survey revealed that the number of sensors used in these studies ranged from 1 to 18, and the number of tasks ranged from 1 to 14, occurring in various combinations (Figure 1). We categorized these studies into four groups based on their objectives. The first group included studies focusing on PD diagnosis. There were 21 studies in this group, using between 1 and 18 sensors and from 1 to 12 tasks [43–63]. Among these, six studies used a single sensor placed on the lower trunk [48,56,59], foot [49], wrist [54], or torso [60]. Additionally, 11 studies used data from a single mobility task—either simple gait tasks [43,46,51,53,57,59,61–63] or more complex tasks [52,56]—with only two of these studies also using a single sensor [56,59]. The second group, comprising 24 studies, focused on quantifying/monitoring the severity of PD symptoms [30,64–86], with 4 using a single sensor [64–67] and 7 using a single mobility task [66,68–73]. Only one study in this group used both a single sensor and a single mobility task [66]. The third group, consisting of 15 studies, aimed to detect freezing of gait in PD [87–101], with 8 using a single sensor [87–92,98,99] and 2 using a single mobility task [93,94]. The last group included 12 studies that performed quantitative characterization of gait or identified imbalance in PD [32,102–112], with 3 using a single sensor [102–104] and 5 using a single mobility task [32,105–107,112]. None of the studies from the third or fourth groups used both a single sensor and a single mobility task. The detailed results of the surveyed studies are presented in Supplementary Table S1.

The reviewed papers exhibit a variety of approaches using wearable sensors for PD assessment, focusing on different tasks and parameters. Some studies emphasize the importance of comprehensive data collection to train symptom detection models [85], objectively measure gait parameters [43,46,48,50–52,69,86,96,102,103,106,112], monitor daily life activities [70,96], or detect early stages of the disease [32,48,57,62,100]. The choice of the number of sensors and their placement is often driven by the specific objectives of each study. For instance, capturing hand motor tests [44,64,78], assessing balance impairments [32,47,66,86,108], and measuring dyski-

nesia severity [77] require specific sensor setups. Sensors are commonly placed on the wrist for tremor and bradykinesia monitoring [54,64,68,74,76,78,80,83,99]; on the foot or shoe to measure gait parameters and detect freezing gait and gait phases [43,49,53,69,88,97,105,106]; on the waist, hips, torso, or lower back for general gait and balance analysis [48,50,59,60,90,91,98,104]; and on the lower limbs to measure several key aspects of PD [61,82,94,101,111,112]. Some studies integrate multiple sensor locations to capture a broader range of symptoms, such as axial rigidity and postural instability, providing comprehensive data [30,45,55,62,63,70,72,73,75,79,96,110]. Overall, the reviewed studies aim to balance detailed data capture with practical application, ensuring accurate and relevant information for their specific goals.
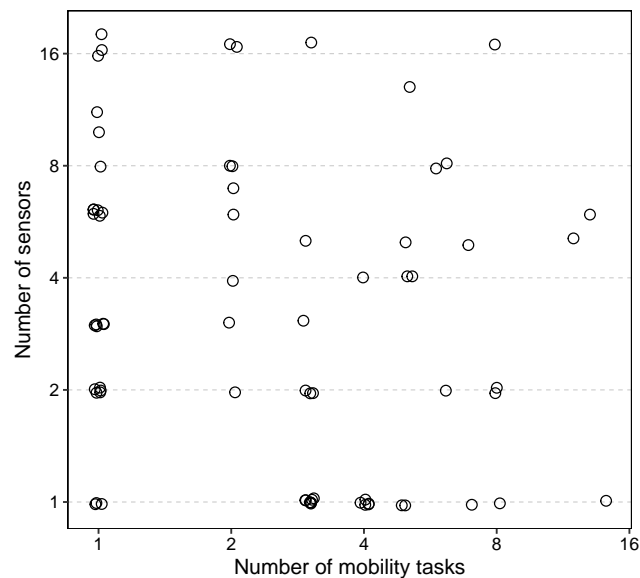

Figure 1. Number of mobility tasks and sensors used in prior studies. The axes are in $\log_2$ scale.

Increasing the number of tasks (abscissa of Figure 1) is associated with more time required from both the patient and clinician, while an increasing number of sensors (ordinate) indicates greater complexity in terms of maintaining, storing, and placing sensors, as well as processing the data. Therefore, the goal of this project is rigorous simplification of the number of sensors, mobility testing, and data processing protocols based on a systematic and data-driven analysis of comprehensive sensor-based mobility data.

Machine learning offers a powerful and flexible analytical approach to accomplish this goal in a large-scale, high-dimensional, and complex dataset collected by wearable sensors [31,113–116]. Furthermore, machine learning can identify and capture hidden patterns and relationships that would be hard to detect using standard observation and statistical methods [117].

Although previous studies have successfully used machine learning and wearable sensors for PD diagnosis [45,49,50,52,53,55,57,59–63], wearable sensors are still not commonly applied in clinical settings due to the complexity and variability of proposed protocols for mobility testing and data processing. Our study objective is to employ machine learning techniques to identify strategies that streamline and simplify the process, reducing the time and effort required by both participants and clinicians while maintaining the predictive performance of models to diagnose (or support the diagnosis of) PD based on instrumented mobility testing. This will ease the application of wearable sensors as a diagnostic tool in clinical practice.

As an initial step toward the simplification of mobility testing, we used a single sensor positioned on the lower backs of participants. Then, we show that sensor data collected from participants performing a single mobility task in a conventional clinic hallway can effectively differentiate individuals with PD from controls. Moreover, we demonstrate that manual segmentation and complex calculations of kinesiological features, which

have been prevalent in prior studies [30,36,38,39,118,119], may be unnecessary. Instead, algorithmic segmentation, followed by the calculation of a small set of sensor-derived features from a single sensor and a single mobility task, offers substantial simplification of both mobility testing and data processing without significantly compromising predictive accuracy. Feature importance analysis revealed that timed up and go (TUG) tasks, which include walking, turning, and sit-to-stand and stand-to-sit transitions, provided more valuable information to our models than simpler mobility tasks such as gait and sway assessments. These results suggest that substantial simplification in mobility testing with a wearable sensor and subsequent data analysis is possible without loss of diagnostic accuracy for PD, facilitating the integration of wearable sensors into clinical practice.

2. Materials and Methods

2.1. Participants

Our study on sensor-based mobility analysis in parkinsonism ("Towards Next-Generation Phenotyping in Parkinson Disease: Quantitative Analysis of Gait and Balance Using a Portable Biosensor Device") at the University of Maryland Movement and Memory Disorders Center (UM-MMDC) enrolled a total of 368 participants between October 2015 and March 2020, including 50 control participants and 318 participants with parkinsonism. Of those, 293 were diagnosed with idiopathic PD by a UM-MMDC movement disorder specialist, applying the UK Parkinson's Disease Society Brain Bank Clinical Diagnostic Criteria [120]. All participants with parkinsonism were also enrolled in the Health Outcomes Measurement (HOME) study at UM-MMDC. The HOME study is a naturalistic cohort study that collects patient- and clinician-reported data during routine office visits. In addition to HOME study participation, all 318 participants with parkinsonism plus 50 non-parkinsonian age-matched control participants were separately enrolled in the mobility analysis in the parkinsonism study. Patients were excluded from the mobility analysis study if they were unable to stand or walk without assistance (Hoehn & Yahr stage 5); presented with other conditions unrelated to parkinsonism that, in the judgment of the treating physician, significantly interfere with gait and balance (e.g., advanced lower extremity arthritis or severe lower extremity neuropathy); or failed to provide signed consent. A total of 24 of the study participants had previously undergone deep brain stimulation (DBS) surgery (20 to the subthalamic nucleus (STN), 3 to the globus pallidus internus (GPi), and 1 to the thalamic ventral intermediate (Vim) nucleus). All participants signed informed consent. A person accompanying a patient (spouse/significant other, sibling, or friend) was approached for enrollment as a control participant if they were of similar age to the patient and did not carry a diagnosis of parkinsonism or another condition significantly affecting their gait and balance. All study procedures for both the HOME study and the mobility analysis study were approved by the University of Maryland Institutional Review Board.

Of 318 participants with parkinsonism, 293 were diagnosed with idiopathic PD. The remaining 25 participants were diagnosed with other forms of parkinsonism. Participants with non-PD forms of parkinsonism were excluded from data analysis for this study. Of 293 participants with a diagnosis of idiopathic PD, one was unable to complete the study procedures and was excluded. Of the remaining 292 PD participants, 30 were excluded from our analysis due to missing data (2 due to technical problems with the wearable sensor and 28 due to a lack of important clinical data). None of the 50 control participants were excluded.

2.2. Clinical Evaluations

Participants were enrolled during routine outpatient visits at the UM-MMDC. All PD participants underwent full clinical assessment according to the HOME study protocol, as previously described [13,121–123]. In brief, collected general data included basic demographics, Cumulative Illness Rating Scale-Geriatrics (CIRS-G) [124], the Older Americans Resource and Services (OARS) Disability Subscale [125] (activities of daily living (ADLs) and instrumental activities of daily living (IADLs)), a questionnaire on the use of assistive

devices and falls (ADF), PROMIS Profile 29 [126], and the Montreal Cognitive Assessment (MoCA) [127]. PD-specific data included PD duration (time since symptom onset and time since diagnosis), PD severity (Total and Motor (Part III) Unified Parkinson's Disease Rating Scale (UPDRS) [128], and modified H&Y staging [14,129]). The HOME study is a naturalistic study, and patients are enrolled during routine clinic visits and on their regular medications. For 87% of assessments performed for this study, the patients rated themselves as being in a medication ON state. Patients who had previously undergone DBS surgery were examined in the stimulation ON state, both for clinical assessments and for the instrumented mobility testing.

2.3. Assessment of Mobility

Mobility testing with a wearable sensor was performed using a previously published protocol [38,130,131]. We adopted the mobility testing protocol from the Memory and Aging Project and the Minority Aging Research Study at the RUSH Alzheimer's Disease Center at Rush University [132] based on the previous successful use of mobility data from this study in parkinsonism. In brief, participants performed mobility tasks with a single wearable sensor strapped to their lower back. For this study, the following data from five of those mobility tasks were analyzed: 1. a 32-foot walk, where participants walk eight feet back and forth twice without stopping, including three turns; 2. standing with eyes open, where participants stand with their feet shoulder-width apart for 20 second with their eyes open; 3. standing with eyes closed, where participants stand with their feet shoulder-width apart for 20 s with eyes closed; 4. two trials of the timed up and go test (TUG), where participants stand up from a seated position in a chair, walk eight feet to a line on the floor, turn, walk back to the chair, make a second turn, and sit down again; and 5. two trials of cognitive TUG (cogTUG), where participants perform the TUG test while counting backwards from 100 in steps of 3.

2.4. Device and Data Collection

Preceding the mobility testing procedure, a small ($106.6 \times 58 \times 11.5$ mm) lightweight (55 g) sensor device (Dynaport MT, McRoberts B.V., The Hague, The Netherlands, technical specifications: https://www.mcroberts.nl/products/movetest/, accessed on 17 June 2024) was strapped to the participant's lower back with a neoprene belt around the waist. This device contains a triaxial accelerometer (range: ± 8 g; resolution: 1 mg) and a triaxial gyroscope (range: ± 2000 dps; resolution: 70 dps) with 100 Hz sampling of three-dimensional acceleration (vertical (az), mediolateral (ay), and anteroposterior (ax)) and rotation (yaw (gz), pitch (gy), and roll (gx)) of the lower trunk. Recording by the sensor was controlled (started and stopped) remotely via wireless Bluetooth signals from a notebook computer using a handheld controller. To facilitate post hoc segmentation of individual mobility tasks, digital markers were placed to label the beginning and end of each mobility task. At the end of each test session, the sensor was removed from the participant and connected to an encrypted notebook computer for data transfer. Raw data files were then transferred to a secure data server.

2.5. Data Extraction

A custom-made, runtime-based Matlab graphic user interface (GUI) was used to view the accelerometry data. The 32-foot walk, TUG, and cogTUG were identified by the typical accelerometry pattern of each of these tasks and by the digital markers recorded at the time of mobility testing. The two 20 s segments corresponding to standing with eyes open and closed, respectively, were identified solely based on the digital markers recorded at the time of mobility testing. The raw accelerometry data for each individual mobility task were saved as separate files for subsequent data processing.

Figure 2 presents an overview of the machine learning pipeline implemented in this study. The following sections describe the steps of the workflow in detail.
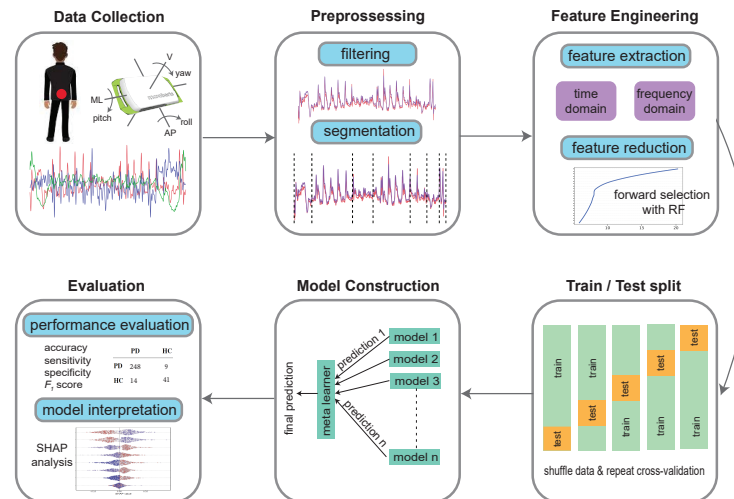
Figure 2. Machine learning pipeline. The process started by collecting data from a wearable sensor on three linear and three angular axes. Filtration was then applied to reduce noise in signals using a Butterworth filter. The composite tasks were further processed by dividing them into segments corresponding to the different subtasks. Using the six signals recorded for each participant and task, a large set of time-domain and frequency-domain features was calculated. This set was reduced using forward selection based on feature importance calculated with random forests. The data were then shuffled multiple times, and a cross-validation framework was applied in each iteration to generate train and test sets. A super-learner model was built using a wide array of base models, and the final prediction was calculated by assigning a weight to the prediction of each base learner. Finally, the performance of the model was evaluated using accuracy, sensitivity, specificity, and F_1 score measures. Model interpretation was performed using SHapley Additive exPlanation (SHAP) analysis (see Supplementary Methods) to quantify the contributions of features to the final model prediction. AP: anteroposterior direction; V: vertical direction; ML: mediolateral direction.

2.6. Segmentation

Some motor tasks performed in this study are composed of several subtasks. For instance, the TUG and cogTUG tasks consist of the following six components: sit to stand, two turns, two walks, and stand to sit. The 32-foot walk task has seven parts representing the four walks and three turns. We developed a segmentation procedure extending a previously described algorithm [133] to include more tasks and subtasks. Our algorithm starts by excluding the stationary segments at the start and end of the signal. These stationary segments result if the recording starts before the participant starts the task or the recording stops after the participant terminates the task. The largest constant parts (with small slope change) common among the six channels and found at the start or the end of the signal were defined as stationary segments and were trimmed. Then, the turns of a task were detected by calculating the angular position from the angular velocity ib the anteroposterior axis (gx) using the trapezoidal integration method [134]. The model for turning fits a line with three horizontal segments and two segments with constant slopes. The transition points from zero to constant and from constant to zero slopes were used to mark the starts and ends of the turns (Supplementary Figure S1a). This method of determining the onset and offset time points of 180° turns has been previously validated [135]. The same integration method was applied to the acceleration data in the vertical direction (az) to determine the sit-to-stand and stand-to-sit parts of the TUG and cogTUG tasks. Specifically, the start and the end of the peak of the fitted line represent the sit-to-stand part, and the start and the end of the valley represents the stand-to-sit part (Supplementary Figure S1b). Finally, our algorithm marked the segments not mapped as turns in the 32-foot walk task and those between the sit-to-stand part and the first turn or between two turns in the TUG and cogTUG tasks as walking segments. Supplementary Figures S2 and S3 show the different components extracted for a participant performing the

TUG and 32-foot walk tasks, respectively. Once we applied our segmentation procedure, the time-domain and frequency-domain features described in the next section were calculated separately for each subtask.

2.7. Feature Engineering

For each of the six channels recorded for each participant and mobility test, we started by applying a low-pass, 20 Hz, zero-lag, 4th-order Butterworth filter [136] to reduce noise in signals. Then, we derived the following sets of features from each signal: statistical measures consisting of 23 frequency-domain features (Supplementary Table S2); 44 time-domain features (Supplementary Table S3); the top-ten power amplitudes from Discrete Fourier Transform (DFT) and Lomb–Scargle Periodogram (LSP) [137,138]; and three time-domain features calculated across the signals of each pair of the six channels (Supplementary Table S4). These sensor-derived features were combined from several previous studies [29,42,139–141]. Kinesiological features were calculated/extracted using a second custom-made, runtime-based, mobility task-specific Matlab graphic user interface (GUI) as previously described [40,41,142]. Additionally, we included features not derived from the sensor data, including demographics and other variables (e.g., age, race, gender, height, OARS, and ADF scores) as predictive variables. Each participant was represented using a total of 26,515 features calculated from the six channels of the segmented subtasks of the 32-foot walk, standing-with-closed-eyes task, standing-with-open-eyes task, the segmented tasks of the TUG task (from the first trial and the second trial and the mean of corresponding features from both trials), and the segmented subtasks of the cogTUG task (from the first trial and the second trial and the mean of corresponding features from both trials). Preprocessing steps were applied to the data before running machine learning models, including removal of variables with constant or infinite values, imputation of missing values using the median of observed values, and normalization of each variable by subtracting its mean and dividing it by its standard deviation.

2.8. Feature Selection

To reduce the number of features, we applied forward feature selection. The goal was to select an optimal subset of features that results in a simpler model with similar or better performance. We started by ranking all features by their importance using a random forest model after randomly undersampling the larger group. Then, the elbow point of the feature importance plot was defined as the point that is farthest away from a straight line drawn from the first to the last point of the plot. This elbow point was used as a cut-off threshold to select the set of candidate features. If the size of this set was n_{thr}, then multiple models were trained using the top one, top two, or . . . , top n_{thr} important features. The best model was selected as the one that minimizes the total Akaike information criterion (AIC) [143], and the corresponding set of features was used in subsequent analyses. This variable reduction strategy results in fewer features compared to the use of principal component analysis, as reported in similar studies [38,144,145], and retains the original features.

2.9. Machine Learning Model

Prior to model training, we split the data into five groups. We used stratified sampling to preserve the class frequencies within each group. Then, a nested-loop framework was applied, where in each iteration of the inner loop, a five-fold cross-validation was implemented by choosing four groups for training and one for testing of the model. This ensured that each participant was included exactly once in the test data. The benefit of this technique is that all observations were used for both training and testing. In the outer loop, we created five replicates by randomly shuffling the whole dataset using different random number generator seeds to create different combinations of participants within each training and testing group. The final class of each participant was defined as the majority vote of the predicted classes from the five outer loops.

Prior to constructing each classifier, a balanced sampling process was applied to the training sets, where the mean sample sizes of the two classes were calculated and used to undersample the majority class and/or oversample the minority class. This mean-based sampling approach provides a balance between simplicity and computational efficiency. An ensemble method combining multiple machine learning algorithms was then applied to the training sets. This approach is known to achieve performance equal to or better than any of the base learners. The user of a super learner (SL) [146] was proposed as a stacked generalization using an ensemble of methods that generates an optimal model by creating a weighted linear convex combination of a set of base algorithms. The super-learner algorithm compares the performance of the user-chosen base algorithms by applying *k*-fold cross-validation to determine the weighted contribution of each base algorithm to the ensemble. We used the R interface H2O with five-fold cross-validation and a generalized linear model as a meta learner. For base models, we included a generalized linear model (GLM) [147], distributed random forests (DRFs) [148], gradient boosting machines (GBMs) [149], extreme gradient boosting (XGBoost) [150], and neural networks [151]. Hyperparameters of the base algorithms were selected using grid search criteria. The final ensemble incorporated 327 base models, including 11 GLM models, 100 DRF models, 100 GBM models, 16 XGBoost models, and 100 neural network models (see Supplementary Files). The class prediction thresholds were selected based on the highest F_1 score.

Given the predicted class of each participant, the performance of the classification model was assessed using overall accuracy, area under the receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, and F_1 score.

2.10. Group Feature Importance

We employed the concept of group feature importance to assess the information of various mobility tasks in predicting the outcome of interest. Group feature importance refers to the assessment of the collective information of a set of features within a specific category. All features derived from each mobility task were grouped together to form five separate groups corresponding to the five mobility tasks. We also created an additional group by combining features not derived from the sensor data, including demographics and other variables (e.g., age, race, gender, height, OARS, and ADF scores). To measure group importance, we adapted the permutation importance method for feature groups within the context of random forests, building upon a previously outlined approach [152]. For each feature group, the grouped variables were simultaneously permuted, and the resultant mean decrease in accuracy prior to and post permutation was computed.

2.11. Statistical Analysis

Data analysis was performed with R version 4.2.3 (2023-03-15) and RStudio. A two-tailed Student's *t*-test was applied to compare continuous variables, and Pearson's χ^2 test was used to compare categorical variables between two groups. Results are presented as mean $\pm$ s.d. for continuous variables. Actual *p* values for the statistical tests are reported.

For each performance metric of the models, we display both the computed values and their respective 95% two-sided confidence intervals (CIs). These intervals were derived via balanced bootstrap resampling with 10K replicates and were adjusted using bias correction and acceleration to better manage bias and skewness in the bootstrap distribution. The bootstrap replicates were generated using the boot function in R package boot version 1.3–30 [153,154], and the boot.ci function from the same package was used to generate the CIs.

3. Results

We collected wearable sensor data from 262 participants with PD (87% were in a medication ON state) and 50 age-matched controls. We partitioned PD participants by their modified Hoehn & Yahr (H&Y) stage, defining mild PD as H&Y $\leq$ 2, moderate PD as H&Y 2.5 and 3, and severe PD as H&Y 4 (note that H&Y 5 patients were excluded

from this study). Although our analysis encompasses four classifiers to distinguish control participants from all PD, mild PD, moderate PD, and severe PD groups, respectively, here, we focus on the results of our models for all PD participants and mild PD participants only. The all-PD classifier best mimics a routine clinical situation with patients in various stages of PD. The mild-PD classifier focuses on the group of PD participants that is most similar to control participants, making these two groups the most challenging to differentiate from one another. Results of the models for moderate- and severe-PD participants are detailed in the Supplementary Results (Section S2.1). Table 1 provides the characteristics of the study cohort.

After the participants performed the five motor tasks (Section 2.3) wearing a Dynaport MT sensor strapped to their lower back, we followed the experimental design described in Figure 2 to predict the final class of each participant.

Table 1. Characteristics of the study cohort.

Feature	Controls ($n = 50$)	All PD ($n = 262$)	Mild PD ($n = 185$)
Age (years, $\bar{x} \pm SD$)	64.1 ± 9.8	66.9 ± 9.3	65.4 ± 9.1
Gender (%male)	38.0	62.0	58.4
Height (cm, $\bar{x} \pm SD$)	168.1 ± 10.9	172.2 ± 10.4	172.0 ± 10.3
UPDRS (total, $\bar{x} \pm SD$)	–	35.5 ± 17.1	29.4 ± 13.4
UPDRS (motor-part III, $\bar{x} \pm SD$)	–	22.2 ± 11.8	18.5 ± 9.7
Disease duration (years, $\bar{x} \pm SD$)	–	7.8 ± 6.5	6.6 ± 5.6
H&Y ($\bar{x} \pm SD$)	–	2.2 ± 0.62	1.9 ± 0.26
stage 1 (n)	–	12	12
stage 1.5 (n)	–	4	4
stage 2 (n)	–	169	169
stage 2.5 (n)	–	35	–
stage 3 (n)	–	25	–
stage 4 (n)	–	17	–

3.1. Classification of PD versus Control

By applying the forward feature selection approach, the total number of features was reduced from 25,705 to 9 for the all-PD classification and 14 for the mild-PD classification (Supplementary Figures S4–S6). Eight of the selected features for the all-PD classification were derived from one or both trials of the cogTUG task; three features were from the second cogTUG trial (cogTUG2), one feature was from the first cogTUG trial (cogTUG1), and four features were the mean of the corresponding features from the two cogTUG trials (mean cogTUG). One feature was derived from the open-eyes standing task. For the mild-PD classification, 13 of the 14 selected features were derived from the cogTUG task; 2 features were from cogTUG2, 1 feature was from cogTUG1, and 10 were from mean cogTUG. One feature was from the eyes-closed standing task. The selected cogTUG features for both classifiers were derived from the turn, walk, and stand-to-sit components of the task, representing acceleration and rotational measures along the vertical, mediolateral, and anteroposterior axes. None of the selected features for both classifiers were from the 32-foot walk or TUG tasks.

Among the 327 base models included in the ensemble classifier, a gradient boosting machine model made the greatest weighted contribution to the final models (Supplementary Figures S7 and S8). The models trained on the selected features achieved an overall accuracy of 92.6% (confidence interval [CI] 88.8%, 94.9%) in distinguishing all-PD from control participants and 89.4% (CI 84.3%, 92.3%) in distinguishing those with mild PD from control participants. Confusion matrices (prediction summaries) and performance metrics for the classifiers are presented in Tables 2 and 3, respectively.

Table 2. Confusion matrices of the classifiers using controls and participants with Parkinson's disease (PD). Rows represent actual classes, and columns represent predictions.

	(a) Controls vs. All PD			(b) Controls vs. Mild PD (H&Y $\leq$ 2)	
	Controls	PD		Controls	PD
Controls	41	9	Controls	36	14
PD	14	248	PD	11	174

Table 3. Classification results of models in distinguishing controls from all Parkinson's disease (PD) and mild-PD (H&Y $\leq$ 2) participants. CI: 95% confidence interval.

Controls vs. PD	All PD	Mild PD
Number of PD participants	262	185
Accuracy (CI) [%]	92.6 (88.8, 94.9)	89.4 (84.3, 92.3)
AUC-ROC (CI)	0.88 (0.83, 0.94)	0.83 (0.77, 0.90)
Sensitivity (CI)	0.95 (0.91, 0.97)	0.94 (0.90, 0.97)
Specificity (CI)	0.82 (0.69, 0.91)	0.72 (0.58, 0.83)
F_1 score (CI)	0.96 (0.93, 0.97)	0.93 (0.90, 0.96)

To further elucidate which mobility tasks were most relevant for the performance of our models, we conducted a group feature importance analysis (see Section 2.10). Figure 3 illustrates the group feature importance of the five mobility tasks and the non-sensor (demographic and clinical) features. The highest-yield group was the one with cogTUG-derived features for both models distinguishing controls from all-PD and mild-PD participants. A substantial decline of at least 64% in importance was observed when comparing the cogTUG group with the remaining groups for either model, underscoring the remarkable discriminatory power inherent in the cogTUG features. More detailed individual feature-level importance scores and Shapley additive explanation (SHAP) values were also calculated (Supplementary Figures S4 and S6, Supplementary Methods, Supplementary Results (Section S2.2)).

To identify potentially confounding factors that may impact our models, we conducted a comparison of demographic and clinical characteristics between correctly classified and misclassified PD and control participants. A two-tailed Student's *t*-test was applied to compare continuous variables, and Pearson's χ^2 test was used to compare categorical variables between separate groups. The comparison results are summarized in Table 4. For both classifiers (all PD and mild PD), false positives (FPs, i.e., controls falsely classified as PD) were significantly older compared to the correctly classified controls (true negatives, TNs). False negatives (FNs, i.e., PD participants classified as controls) were younger compared to the correctly classified PD participants (true positives, TPs), but this difference was not statistically significant. These results suggest that older controls have movement characteristics that confuse the classifier, but there was no significant difference in age between the correctly and falsely classified PD participants. Also, the mean Unified Parkinson's Disease Rating Scale part III (UPDRS_PIII) motor score of the FNs (falsely classified PD participants) was significantly lower than that of TPs (correctly classified PD participants). In contrast to the UPDRS_PIII scores, misclassified PD participants were evenly distributed among the different H&Y stages. Similarly, there was no significant difference in the medication state, sex, MoCA scores, or CIRS-G scores between the correctly and falsely classified PD participants, suggesting that these factors do not have a confounding impact on model performance.
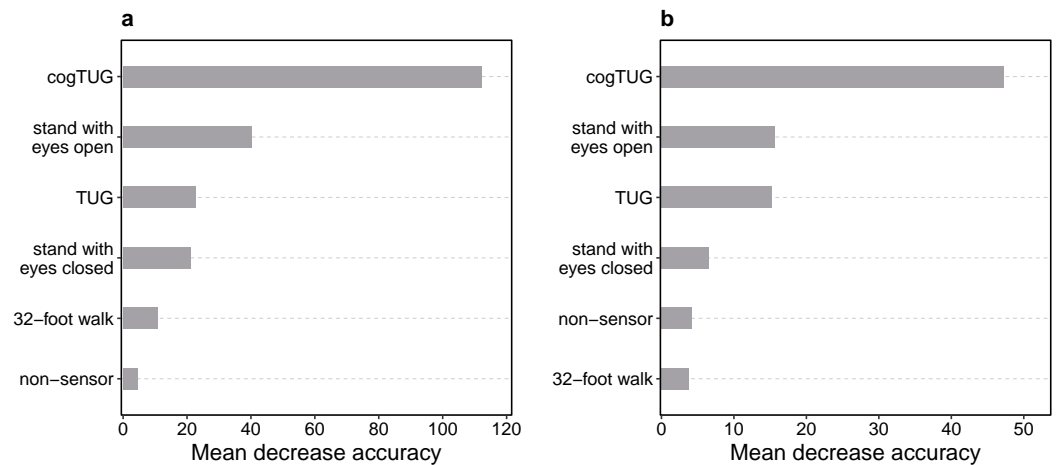
Figure 3. Group feature importance of models distinguishing controls from (**a**) all Parkinson's disease (PD) and (**b**) mild-PD (H&Y $\leq$ 2) participants. Groups are ranked by their importance calculated by permuting features within each group simultaneously and reporting the mean decrease in accuracy between the original and permuted data. Under the null hypothesis that there is no association between the group of predictor variables and the model prediction, permutation should have no or little impact on predictive performance. More detailed individual feature-level importance scores and Shapley additive explanation (SHAP) values are illustrated in Supplementary Figures S4 and S6.

Table 4. Clinical characteristics of correctly classified and misclassified controls and Parkinson's disease (PD) participants.

		Controls			PD		
		False Positive	True Negative	*p*	False Negative	True Positive	*p*
All PD							
	UPDRS_PIII	–	–	–	15.8 ± 8.3	22.5 ± 11.8	0.006
	MoCA	–	–	–	28.6 ± 2.5	27.3 ± 3.0	0.075
	CIRS-G	–	–	–	4.7 ± 4.3	4.9 ± 3.4	0.88
	Age	70.9 ± 7.3	62.6 ± 7.3	0.006	65.9 ± 6.8	66.9 ± 9.3	0.30
	Sex (% male)	11.1	36.0	0.10	50.0	63.3	0.33
	Medication state (% ON)	–	–	–	85.7	70.6	0.44
	H&Y (n)			–			0.47
	1				1	11	
	1.5				1	3	
	2				10	159	
	2.5				1	34	
	3				1	24	
	4				0	17	
Mild PD							
	UPDRS_PIII	–	–	–	13.7 ± 7.4	18.7 ± 9.6	0.03
	MoCA	–	–	–	28.1 ± 1.9	27.5 ± 2.8	0.36
	CIRS-G	–	–	–	4.9 ± 4.4	4.4 ± 3.1	0.72
	Age	71.5 ± 6.3	61.3 ± 6.3	4.0×10^{-5}	65.2 ± 9.1	69.2 ± 7.5	0.12
	Sex (% male)	42.9	36.1	0.71	45.5	60.9	0.32
	Medication state (% ON)	–	–	–	88.2	71.3	0.36
	H&Y (n)			–			0.38
	1				1	11	
	1.5				1	3	
	2				9	160	

3.2. Strategies for Simplifying Mobility Testing and Its Associated Workload

One of the major obstacles preventing widespread use of wearable sensors for PD diagnosis is the increased burden in terms of time and effort in performing sensor-based mobility testing and subsequent data processing and analysis. To overcome these obstacles,

we first focused on possible simplification of the data analysis process. As detailed above, our data processing pipeline includes an algorithmic segmentation approach (Section 2.6) to calculate a set of sensor-derived frequency-domain and time-domain features from the resulting task segments, referred to as "standard models" below. To evaluate the effectiveness of the automated segmentation method and whether segmentation is necessary at all, we constructed alternative models using features derived from data of the unsegmented tasks. These models decreased the accuracy in distinguishing PD from control participants by 3.2% and 5.1% for all-PD and mild-PD classifications, respectively, compared to our standard models (Table 5, second column vs. first column).

We also developed another set of models incorporating calculated kinesiological features (see Section 2.7), in addition to the time- and frequency-domain features from the algorithmic segmentation tasks. These models also reduced the accuracy by 4.8% and 0.50% compared to our standard models (Table 5, third column vs. first column). These findings show that algorithmic task segmentation increases diagnostic accuracy, but complex calculation of kinesiological features does not. These time-consuming and labor-intensive additional steps required to calculate the kinesiological features can therefore be eliminated without compromising the predictive performance of our diagnostic models.

Table 5. Accuracy (%) of super-learner models using unsegmented tasks and a combination of segmented tasks and calculated kinesiological features to distinguish controls from PD participants. The numbers in parentheses represent the difference in accuracy between each model and the corresponding standard model, which used data from segmented tasks without incorporating kinesiological features (same row).

	With Segmented Tasks and No Kinesiological Features	With Unsegmented Tasks and No Kinesiological Features	With Segmented Tasks and Kinesiological Features
All PD vs. controls	92.6	89.4 (−3.2)	87.8 (−4.8)
Mild PD (H&Y $\leq$ 2) vs. controls	89.4	84.3 (−5.1)	88.9 (−0.50)

As a second step towards simplification of the overall workflow, we focused on the tasks of mobility testing itself, by assessing the predictive performance achievable with fewer mobility tasks. Given the outsized importance of cogTUG features in the context of our models based on data from all available mobility tasks (Figure 3), we examined the performance of two drastically simplified models, using TUG- or cogTUG-derived features only, respectively, in addition to the non-sensor (demographic and clinical) features. We applied the feature selection approach to these single-task models (see Supplementary Table S5 for the selected features). For the all-PD classifier, the TUG-only model and cogTUG-only model decreased the accuracy by 2.5% and 3.2%, respectively (Table 6, columns 2 and 3, row 1). For participants with mild PD, applying the cogTUG-only model decreased the accuracy by 2.4%, whereas the accuracy dropped more substantially (by 11.7%) using the TUG-only model compared to the standard model based on data from all tasks (Table 6, columns 2 and 3, row 2). These results suggest that sensor data from a single appropriately chosen mobility task may be sufficient to achieve satisfactory performance in differentiating PD and control participants. Specifically, a dual task paradigm (cognitive plus mobility task) like the cogTUG task achieves superior diagnostic performance to a mobility task alone (like the TUG task) in identifying participants in the early stages of PD.

Our observation that data from additional mobility tasks only offer a limited increase in diagnostic accuracy led us to consider whether further radical simplification would be possible without compromising the performance of our models. Therefore, we tested whether the total duration of TUG or cogTUG tasks as a single feature may be sufficient to build classifiers with predictive performance similar to that of our more comprehensive models described above. Even though we found a significant time difference between PD and control participants for both TUG and cogTUG tasks (see Supplementary Results, Section S2.3), the corresponding logistic regression models using total duration as a single

predictor variable decreased the accuracy in distinguishing PD from control participants by at least 8.6% compared to our standard models (Table 6, columns 4 and 5 compared to column 1). Based on these results, we conclude that total TUG/cogTUG duration alone is not sufficient to capitalize on the discriminatory potential of these complex tasks.

Table 6. Accuracy (%) of alternative models. TUG-only and cogTUG-only models used features derived from the first trial, the second trial, and the mean of corresponding features from both trials. TUG-duration and cogTUG-duration models used the durations of TUG and cogTUG tasks. The numbers in parentheses represent the difference in accuracy between the respective model and the corresponding all-task (standard) model (same row).

	All Tasks	TUG-Only	cogTUG-Only	TUG-Duration	cogTUG-Duration
All PD vs. controls	92.6	90.1 (-2.5)	89.4 (-3.2)	83.3 (-9.3)	84.0 (-8.6)
Mild PD (H&Y $\leq$ 2) vs. controls	89.4	77.9 (-11.7)	87.2 (-2.4)	78.7 (-10.7)	77.0 (-12.4)

4. Discussion

Our study demonstrates the effectiveness of analyzing wearable sensor data to distinguish PD participants from controls using the proposed machine learning pipeline for five mobility tasks. By leveraging a large dataset comprising participants in different stages of PD and controls performing five mobility tasks, along with a carefully selected subset of sensor-derived features and suitable machine learning models, we achieved an accuracy of 92.6% in identifying participants across different PD stages, surpassing the accuracy of 81% previously reported for clinical diagnosis by movement disorder experts [15]. We focused on our modeling results for all-PD participants and for mild-PD participants specifically. The all-PD classifier represents a typical clinical scenario with patients at various stages, while the mild-PD classifier targets the group most similar to control participants, where differentiation is most challenging [155] and where wearable sensors could be most clinically relevant.

Previously published studies have reported accuracies ranging from 73.5% to 96% in discriminating between PD and control participants [29,36,37,39,45,52,55,59,61,63,115, 119,156]. Some studies have utilized the AUC-ROC as a performance measure, with values ranging from 0.86 to 0.99 [50,113,157], compared to an AUC of 0.86 for our model. Additional studies have reported sensitivity and specificity metrics, with sensitives ranging from 40% to 96.2% and specificity ranging from 82% to 96.9% [28,49,60,158], as compared to 95% sensitivity and 82% specificity in our case. Major differences between studies in terms of the number of participants, PD severity, tasks conducted, types and quantity of sensors employed, and reported performance metrics likely account for these differences in diagnostic performance and make direct comparison very challenging.

Analysis of group and individual feature importance revealed that the most important mobility tasks for classifying PD participants are the TUG and cogTUG tasks (Figure 3). Both tasks involve various everyday activities, including walking, turning, and sit-to-stand and stand-to-sit transitions. Our results demonstrate that more complex mobility tasks with multiple mobility subtasks provide diagnostically superior information compared to simpler motor tasks. Models based on the cogTUG task as a classic cognitive–motor dual task, in particular, showed superior performance in differentiating between participants with mild PD and controls (see Table 6). Simpler tasks, such as standing with eyes open, standing with eyes closed, and the 32-foot walk were much less important. Previous studies have reported a strong correlation between these simpler motor assessments and PD [30,159–162]. In the context of our study and with the included mobility tasks and features, we demonstrate that simpler motor tasks are not as informative as more complex ones and, thus, can be omitted from the battery of our mobility tests with very minor loss of predictive performance. This is consistent with the results of a previous study that utilized multiple sensors and a higher number of sensor-derived features [163]. However, our results show that further simplification is possible using a single sensor and a reduced set of features, even with a single mobility task.

In our machine learning framework, we implemented automatic segmentation of the composite tasks into their subtask components, coupled with the calculation of many sensor-derived features combined from prior studies. Segmentation of complex tasks into subtasks (e.g., sit to stand, walk, and turn for the TUG and cogTUG tasks) increased the diagnostic accuracy of our models compared to models based on unsegmented data (see Table 5). This automatic segmentation, as opposed to the use of manual annotations in other studies [39,118], not only expedites the processing but also minimizes subjective biases as a source of errors. Additionally, we employed ensembles comprising disparate machine learning classifiers. This approach distinguishes our study from others that primarily depend on individual models [30,32,39,119,156,158], thereby improving predictive accuracy and robustness. Weighted ensembles have demonstrated performance equal to or better than that of any constituent individual learner [146,164], as they leverage the diverse properties and strengths of various algorithms. Although we observed a modest performance difference of approximately 1% between the ensembles and the models with the highest weight in each ensemble, the process of finding the optimal model requires searching through multiple models and their respective hyperparameters.

Many kinesiological variables can be calculated based on signal processing and analysis of accelerometric and gyroscopic data from a wearable sensor using a developed software. This provides quantitative movement characterization in kinesiological terms, as demonstrated in previous studies [38,43,119,165]. However, these variables require additional manual input for the developed software, and our analysis demonstrates that these kinesiological features are neither necessary nor more helpful than analysis based on sensor-derived features alone in achieving better model performance (see Table 5). This observation suggests that properties of movement beyond those corresponding to conventional kinesiological variables were captured in our feature set derived from data from a wearable sensor and that those properties were instrumental in elevating the predictive performance of our models. Eliminating the additional data processing steps required for the generation of kinesiological variables significantly reduces the workload associated with sensor-based mobility testing and facilitates the use of quantitative mobility testing in research (e.g., testing large numbers of participants) and clinical practice.

Finally, we attempted radical simplification of our models by examining the time to complete either the TUG or cogTUG task as the only feature in our models. Total task duration was significantly shorter for controls compared to PD participants (see Supplementary Results, Section S2.3). However, using elapsed time alone as the basis for discriminating PD was markedly less accurate than models including additional sensor-derived features (see Table 6). Another study also showed moderate accuracy (AUC = 0.699) when discriminating between PD patients and controls using the total duration of a TUG test performed at a self-selected speed [166]. The decrease in model performance demonstrates a trade-off between simplicity of testing and data analysis on the one hand and predictive accuracy (and, hence, the value of mobility testing) on the other hand. It is worth noting that if a model based on total duration alone performed equally as well as our multi-feature-based models, one would conclude that timing the TUG (or cogTUG) task with a stopwatch is just as good as using a state-of-the-art wearable sensor. Our results show that is not, in fact, the case. Rather, our models clearly show the added benefit of using a highly granular mobility analysis provided by a wearable accelerometer/gyroscope.

We observed that increasing the diversity of sensor-derived features improves diagnostic accuracy. Our use of forward feature selection greatly reduced the number of features included in the model. Consequently, it is not the number of features per se but the nature of the features that determines the accuracy. Including multiple tasks gives us some additional informative features, as does segmenting tasks into subtasks. On the whole, the latter appears to be the best approach, as it requires reduced time and effort for both patients and clinicians, as the additional effort is handled algorithmically by a computer.

One of the strengths of our study is the large sample size and the wide spectrum of PD severity of our PD participants. We included 262 participants with PD and 50 control partic-

ipants, compared to an *n* of 13 to 20 in many other previously published studies [28,32,167]. Our PD participants differed widely in terms of disease severity, ranging from mild (H&Y $\leq$ 2) to moderate (H&Y 2.5 and 3) and severe (H&Y 4). Therefore, participants in this study represent the full spectrum of ambulatory PD patients encountered in clinical practice, compared to previous studies that focused on a more limited range of disease severity [28,32,167]. Another strength is the simplicity of instrumented mobility testing using a single wearable sensor. This is in contrast to the multiple sensors employed in many other studies [30,32,36,116] and facilitates clinical use and scalability by streamlining data collection and analysis.

Although our proposed set of features might not be easily interpretable by clinicians, it shows potential for distinguishing between PD and control participants. Given our primary aim of achieving the highest possible diagnostic accuracy and real-world applicability, it is essential to include features that contribute to this goal, regardless of their interpretability. Studies that emphasize interpretability might consider using features that are more clinically explainable and may apply prior correlation analysis, as feature importance measures can show bias towards correlated predictor variables [168,169].

In addition to typical motor symptoms, PD is characterized by non-motor features including cognitive impairment and autonomic dysfunction (e.g., dizziness due to a drop in blood pressure in an upright position). Such non-motor features were not assessed in control participants in our study and, therefore, could not be included in our machine learning models, which is one of the limitations of our study. Incorporating such non-motor manifestations of PD might lead to even more accurate models to predict (or identify) PD participants and will be part of our future work. Additionally, while the H&Y scale is commonly used to assess overall PD severity, it is limited by its inter-rater variability due to its reliance on clinical judgment. This subjectivity introduces potential inconsistencies in the partitioning of PD participants, raising the question of how appropriate the H&Y scale is as frame of reference for objective assessment of PD severity. Information regarding the levodopa equivalent daily dose (LEDD) of each PD participant is not available in the PD clinical database (HOME study) we used for this study. However, ON/OFF medication status, which was available, was not found to be an important feature in any of our machine learning analyses, which is evidence that LEDD and ON/OFF status are not consequential in discrimination of PD patients and controls.

Objective and granular measures of disease manifestation as provided by a wearable sensor are promising candidates as superior markers of disease severity and disease progression. However, longitudinal studies are necessary to test this hypothesis. The cross-sectional nature of the data presented here is another limitation of our study. Longitudinal analysis of a subgroup of our study cohort is currently ongoing. Another aspect worth exploring in more detail in future studies is the potential impact of confounding factors on classification performance. Our analysis of misclassified participants shows that certain variables that are likely to be associated with the level of medical comorbidity, such as age, could influence the sensor data and classification outcomes. Older controls were falsely classified as PD, and PD participants with relatively minor motor symptoms (low UPDRS_PIII motor score) were falsely classified as controls. The differentiation between only mildly affected PD patients and older (non-PD) adults is also among the biggest challenges in clinical routine, which is why, for clinicians, the use of a wearable sensor as a supportive biomarker is relevant and attractive in such situations. That was also the rationale for us focusing on models for mild PD versus controls, in addition to all PD versus controls, in this project. Future studies with an even larger sample size for those particular groups of participants may facilitate the development of more comprehensive models that incorporate movement characteristics to overcome these limitations of our models and to further enhance the accuracy and reliability of sensor-based PD classification/identification.

5. Conclusions

In conclusion, our study demonstrates that diagnostic accuracy for PD exceeding that typically achieved by movement disorders experts in a routine clinical setting is possible with few or even just one complex mobility task instrumented with a single wearable sensor when appropriately analyzed using machine learning. Our results provide a framework for such simplification and the ultimate harmonization of existing approaches to facilitate the use of wearable sensors and machine learning-based PD diagnosis in clinical settings.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/s24154983/s1, Figure S1: (a) An example signal for the angular velocity around the roll axis (gx, black line) and the trapezoidal integration of the signal (red line) for the TUG task. Segments (1) and (2) represent the two turns. (b) The signal for the vertical acceleration (az, black line) and the trapezoidal integration of the signal (red line). Segments (1) and (2) represent the sit-to-stand and stand-to-sit parts. Figure S2: Example results of our segmentation approach for the TUG task applied to all six channels. The components are (1) sit-to-stand, (2) first walk, (3) first turn, (4) second walk, (5) second turn, and (6) stand-to-sit. Figure S3: Example results of our segmentation approach for the 32-foot walk task applied to all six channels. The components are (1) first walk, (2) first turn, (3) second walk, (4) second turn, (5) third walk, (6) third turn, and (7) fourth walk. Figure S4: Feature importance scores (left) and SHAP values (right) of the features included in the all PD vs controls classification model. The first column conveys the task and signal channel. Features are ranked by their importance defined as the mean decrease in area under the curve (AUC) between the original and permuted models. SHAP values were calculated using the kernel SHAP method (see Supplementary Methods). Points in the SHAP plot represent participants from the test sets and SHAP values indicate the impact on model output. The color gradient represents the variable values normalized based on percentile ranks. Figure S5: Correlation matrix of the features included in the all PD vs. controls classification mode. Figure S6: Feature importance scores (left) and SHAP values (right) of the features included in the mild PD (H&Y values $\leq$ 2) vs. controls classification model. The first column conveys the task and signal channel. Features are ranked by their importance defined as the mean decrease in area under the curve (AUC) between the original and permuted models. SHAP values were calculated using the kernel SHAP method (see Supplementary Methods). Points in the SHAP plot represent participants from the test sets and SHAP values indicate the impact on model output. The color gradient represents the variable values normalized based on percentile ranks. Figure S7: Base model coefficients in final super learner classifier using all PD participants and controls. Models with zero coefficients are not shown. Figure S8: Base model coefficients in final super learner classifier using mild PD (H&Y $\leq$ 2) participants and controls. Models with zero coefficients are not shown. Figure S9: Feature importance scores (left) and SHAP values (right) of the features included in the moderate PD (H&Y = 2.5, 3) vs. controls classification model. The first column conveys the task and signal channel. Features are ranked by their importance defined as the mean decrease in AUC between the original and permuted models. SHAP values were calculated using the kernel SHAP method (see Supplementary Methods). Points in the SHAP plot represent participants from the test sets and SHAP values indicate the impact on model output. The color gradient represents the variable values normalized based on percentile ranks. Figure S10: Feature importance scores (left) and SHAP values (right) of the features included in the severe PD (H&Y = 4) vs. controls classification model. The first column conveys the task and signal channel. Features are ranked by their importance as the mean decrease in AUC between the original and permuted models. SHAP values were calculated using the kernel SHAP method (see Supplementary Methods). Points in the SHAP plot represent participants from the test sets and SHAP values indicate the impact on model output. The color gradient represents the variable values normalized based on percentile ranks. Figure S11: Base model coefficients in final super learner classifier using moderate PD (H&Y 2.5 and 3) and controls. Models with zero coefficients are not shown. Figure S12: Base model coefficients in final super learner classifier using severe PD (H&Y 4) and controls. Models with zero coefficients are not shown. Figure S13: Group feature importance of models distinguishing (a) moderate (H&Y = 2.5, 3) Parkinson's disease (PD) and (b) severe PD (H&Y = 4) participants from controls. Groups are ranked by their importance calculated by permuting features within each group simultaneously and reporting the mean decrease in accuracy between the original and permuted models. Under the null hypothesis that there is no association of the group of predictor variables

and the model prediction, permutation should have no or little impact on predictive performance. Figure S14: Comparison of the total duration of (a) TUG and (b) cogTUG tasks between different groups of PD participants and controls. Different PD groups shows significantly longer durations when compared to the controls using two-tailed Student's t-test ($p < 0.0014$). Table S1: Number of sensors and mobility tasks used in a survey of literature from 2017 to 2024. Table S2: Frequency domain features. The power spectrum is obtained using the Discrete Fourier Transform (DFT) of the signal with a fast algorithm, the Fast Fourier Transform (FFT). Table S3: Time domain features. Table S4: Cross-time domain features. Table S5: Features selected by the feature selection approach for TUG-only and cogTUG-only models using features from the first trial, the second trial, and the mean of the corresponding features from both trials. Table S6: Confusion matrix of the classifier using controls and participants with moderate (H&Y 2.5, 3) and severe (H&Y 4) Parkinson's disease (PD). Rows represent actual class and columns represent predictions. Table S7: Classification results of models distinguishing moderate (H&Y = 2.5, 3) and severe (H&Y = 4) Parkinson's disease (PD) participants from controls. Table S8: Clinical characteristics of correctly classified and misclassified controls and moderate and severe Parkinson's disease (PD) participants. Table S9: Accuracy (%) of super learner models using the unsegmented tasks and combination of segmented tasks and calculated kinesiological features to distinguish controls from PD participants. The numbers in parentheses represent the difference in accuracy between each model and the corresponding standard model, which used data from segmented tasks without incorporating kinesiological features (same row). Table S10: Accuracy (%) of alternative models. TUG-only and cogTUG-only models used features derived from the first trial, the second trial, and the mean of corresponding features from both trials. TUG-duration and cogTUG-duration models used the duration of TUG and cogTUG tasks. The numbers in parentheses represent the difference in accuracy between the respective model and the corresponding all-tasks model (same row). Table S11: Comparing the duration of TUG sub-tasks between different groups of PD participants and controls (ctrl). Table S12: Comparing the duration of cogTUG sub-tasks between different groups of PD participants and controls (ctrl). References [170–190] are cited in the Supplementary Materials.

References

1. Pringsheim, T.; Jette, N.; Frolkis, A.; Steeves, T.D. The prevalence of Parkinson's disease: A systematic review and meta-analysis. *Mov. Disord.* **2014**, *29*, 1583–1590. [CrossRef]
2. Geraghty, J.J.; Jankovic, J.; Zetusky, W.J. Association between essential tremor and Parkinson's disease. *Ann. Neurol.* **1985**, *17*, 329–333. [CrossRef] [PubMed]
3. Tan, E.K.; Lee, S.S.; Lum, S.Y. Evidence of increased odds of essential tremor in Parkinson's disease. *Mov. Disord.* **2008**, *23*, 993–997. [CrossRef]
4. Berardelli, A.; Sabra, A.F.; Hallett, M. Physiological mechanisms of rigidity in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **1983**, *46*, 45. [CrossRef]
5. Bologna, M.; Leodori, G.; Stirpe, P.; Paparella, G.; Colella, D.; Belvisi, D.; Fasano, A.; Fabbrini, G.; Berardelli, A. Bradykinesia in early and advanced Parkinson's disease. *J. Neurol. Sci.* **2016**, *369*, 286–291. [CrossRef]
6. Cantello, R.; Gianelli, M.; Bettucci, D.; Civardi, C.; De Angelis, M.S.; Mutani, R. Parkinson's disease rigidity: Magnetic motor evoked potentials in a small hand muscle. *Neurology* **1991**, *41*, 1449–1456. [CrossRef]
7. Pongmala, C.; Roytman, S.; van Emde Boas, M.; Dickinson, O.; Kanel, P.; Bohnen, N.I. Composite measures of motor performance and self-efficacy are better determinants of postural instability and gait difficulties than individual clinical measures in Parkinson's disease. *Park. Relat. Disord.* **2023**, *107*, 105251. [CrossRef]
8. Morris, M.E.; Iansek, R.; Matyas, T.A.; Summers, J.J. The pathogenesis of gait hypokinesia in Parkinson's disease. *Brain* **1994**, *117*, 1169–1181. [CrossRef] [PubMed]
9. Sofuwa, O.; Nieuwboer, A.; Desloovere, K.; Willems, A.M.; Chavret, F.; Jonkers, I. Quantitative gait analysis in Parkinson's disease: Comparison with a healthy control group. *Arch. Phys. Med. Rehabil.* **2005**, *86*, 1007–1013. [CrossRef]
10. Snijders, A.H.; Leunissen, I.; Bakker, M.; Overeem, S.; Helmich, R.C.; Bloem, B.R.; Toni, I. Gait-related cerebral alterations in patients with Parkinson's disease with freezing of gait. *Brain* **2010**, *134*, 59–72. [CrossRef]
11. Zach, H.; Janssen, A.M.; Snijders, A.H.; Delval, A.; Ferraye, M.U.; Auff, E.; Weerdesteyn, V.; Bloem, B.R.; Nonnekes, J. Identifying freezing of gait in Parkinson's disease during freezing provoking tasks using waist-mounted accelerometry. *Park. Relat. Disord.* **2015**, *21*, 1362–1366. [CrossRef] [PubMed]
12. Gao, X.; Chen, H.; Schwarzschild, M.A.; Logroscino, G.; Ascherio, A. Perceived imbalance and risk of Parkinson's disease. *Mov. Disord.* **2008**, *23*, 613–616. [CrossRef] [PubMed]
13. Shulman, L.M.; Gruber-Baldini, A.L.; Anderson, K.E.; Vaughan, C.G.; Reich, S.G.; Fishman, P.S.; Weiner, W.J. The evolution of disability in Parkinson disease. *Mov. Disord.* **2008**, *23*, 790–796. [CrossRef] [PubMed]
14. Hoehn, M.M.; Yahr, M.D. Parkinsonism: Onset, progression, and mortality. *Neurology* **1967**, *17*, 427–442. [CrossRef] [PubMed]
15. Rizzo, G.; Copetti, M.; Arcuti, S.; Martino, D.; Fontana, A.; Logroscino, G. Accuracy of clinical diagnosis of Parkinson disease: A systematic review and meta-analysis. *Neurology* **2016**, *86*, 566–576. [CrossRef] [PubMed]
16. Sieber, B.A.; Landis, S.; Koroshetz, W.; Bateman, R.; Siderowf, A.; Galpern, W.R.; Dunlop, J.; Finkbeiner, S.; Sutherland, M.; Wang, H.; et al. Prioritized research recommendations from the National Institute of Neurological Disorders and Stroke Parkinson's Disease 2014 Conference. *Ann. Neurol.* **2014**, *76*, 469–472. [CrossRef] [PubMed]
17. Maetzler, W.; Domingos, J.; Srulijes, K.; Ferreira, J.J.; Bloem, B.R. Quantitative wearable sensors for objective assessment of Parkinson's disease. *Mov. Disord.* **2013**, *28*, 1628–1637. [CrossRef] [PubMed]
18. Rovini, E.; Maremmani, C.; Cavallo, F. How wearable sensors can support Parkinson's disease diagnosis and treatment: A systematic review. *Front. Neurosci.* **2017**, *11*, 555. [CrossRef] [PubMed]
19. Rodríguez-Molinero, A.; Monsonís, A.S.; Pérez, C.; Rodríguez-Martín, D.; Alcaine, S.; Mestre, B.; Quispe, P.; Giuliani, B.; Vainstein, G.; Browne, P.; et al. Analysis of correlation between an accelerometer-based algorithm for detecting parkinsonian gait and UPDRS subscales. *Front. Neurol.* **2017**, *8*, 431. [CrossRef]
20. Sejdić, E.; Lowry, K.A.; Bellanca, J.; Redfern, M.S.; Brach, J.S. A comprehensive assessment of gait accelerometry signals in time, frequency and time-frequency domains. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 603–612. [CrossRef]
21. Din, S.D.; Godfrey, A.; Rochester, L. Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and Parkinson's disease: Toward clinical and at home use. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 838–847. [CrossRef] [PubMed]
22. Hubble, R.P.; Naughton, G.A.; Silburn, P.A.; Cole, M.H. Wearable sensor use for assessing standing balance and walking stability in people with Parkinson's disease: A systematic review. *PLoS ONE* **2015**, *10*, e0123705. [CrossRef]
23. Lu, R.; Xu, Y.; Li, X.; Fan, Y.; Zeng, W.; Tan, Y.; Ren, K.; Chen, W.; Cao, X. Evaluation of wearable sensor devices in Parkinson's disease: A review of current status and future prospects. *Park. Dis.* **2020**, *2020*, 1–8. [CrossRef]

24. Moreau, C.; Rouaud, T.; Grabli, D.; Benatru, I.; Remy, P.; Marques, A.R.; Drapier, S.; Mariani, L.L.; Roze, E.; Devos, D.; et al. Overview on wearable sensors for the management of Parkinson's disease. *Npj Park. Dis.* **2023**, *9*, 1–16. [CrossRef]

25. Channa, A.; Popescu, N.; Ciobanu, V. Wearable solutions for patients with Parkinson's disease and neurocognitive disorder: A systematic review. *Sensors* **2020**, *20*, 2713. [CrossRef] [PubMed]

26. Espay, A.J.; Bonato, P.; Nahab, F.B.; Maetzler, W.; Dean, J.M.; Klucken, J.; Eskofier, B.M.; Merola, A.; Horak, F.; Lang, A.E.; et al. Technology in Parkinson's disease: Challenges and opportunities. *Mov. Disord.* **2016**, *31*, 1272–1282. [CrossRef] [PubMed]

27. Kubota, K.J.; Chen, J.A.; Little, M.A. Machine learning for large-scale wearable sensor data in Parkinson's disease: Concepts, promises, pitfalls, and futures. *Mov. Disord.* **2016**, *31*, 1314–1326. [CrossRef]

28. Arora, S.; Venkataraman, V.; Zhan, A.; Donohue, S.; Biglan, K.M.; Dorsey, E.R.; Little, M.A. Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Park. Relat. Disord.* **2015**, *21*, 650–653. [CrossRef] [PubMed]

29. Oung, Q.; Hariharan, M.; Lee, H.; Basah, S.; Sarillee, M.; Lee, C. Wearable Multimodal Sensors for Evaluation of Patients with Parkinson Disease. In Proceedings of the 2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 27–29 November 2015; pp. 269–274. [CrossRef]

30. Sotirakis, C.; Su, Z.; Brzezicki, M.A.; Conway, N.; Tarassenko, L.; FitzGerald, J.J.; Antoniades, C.A. Identification of motor progression in Parkinson's disease using wearable sensors and machine learning. *npj Park. Dis.* **2023**, *9*, 1–8. [CrossRef]

31. Cole, B.; Roy, S.; Luca, C.D.; Nawab, S. Dynamical learning and tracking of tremor and dyskinesia from wearable sensors. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 982–991. [CrossRef]

32. Castelli Gattinara Di Zubiena, F.; Menna, G.; Mileti, I.; Zampogna, A.; Asci, F.; Paoloni, M.; Suppa, A.; Del Prete, Z.; Palermo, E. Machine Learning and Wearable Sensors for the Early Detection of Balance Disorders in Parkinson's Disease. *Sensors* **2022**, *22*, 9903. [CrossRef] [PubMed]

33. Albán-Cadena, A.C.; Villalba-Meneses, F.; Pila-Varela, K.O.; Moreno-Calvo, A.; Villalba-Meneses, C.P.; Almeida-Galárraga, D.A. Wearable sensors in the diagnosis and study of Parkinson's disease symptoms: A systematic review. *J. Med. Eng. Technol.* **2021**, *45*, 532–545. [CrossRef]

34. Mughal, H.; Javed, A.R.; Rizwan, M.; Almadhor, A.S.; Kryvinska, N. Parkinson's Disease Management via Wearable Sensors: A Systematic Review. *IEEE Access* **2022**, *10*, 35219–35237. [CrossRef]

35. Kour, N.; Gupta, S.; Arora, S. Sensor technology with gait as a diagnostic tool for assessment of Parkinson's disease: A survey. *Multimed. Tools Appl.* **2022**, *82*, 10211–10247. [CrossRef]

36. Kuhner, A.; Schubert, T.; Cenciarini, M.; Wiesmeier, I.K.; Coenen, V.A.; Burgard, W.; Weiller, C.; Maurer, C. Correlations between motor symptoms across different motor tasks, quantified via random forest feature classification in Parkinson's disease. *Front. Neurol.* **2017**, *8*, 607. [CrossRef] [PubMed]

37. Abujrida, H.; Agu, E.; Pahlavan, K. Smartphone-Based Gait Assessment to Infer Parkinson's Disease Severity using Crowdsourced Data. In Proceedings of the 2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT), Bethesda, MD, USA, 6–8 November 2017; pp. 208–211. [CrossRef]

38. Buchman, A.S.; Leurgans, S.E.; Weiss, A.; VanderHorst, V.; Mirelman, A.; Dawe, R.; Barnes, L.L.; Wilson, R.S.; Hausdorff, J.M.; Bennett, D.A. Associations between quantitative mobility measures derived from components of conventional mobility testing and parkinsonian gait in older adults. *PLoS ONE* **2014**, *9*, e86262. [CrossRef] [PubMed]

39. Palmerini, L.; Mellone, S.; Avanzolini, G.; Valzania, F.; Chiari, L. Quantification of motor impairment in Parkinson's disease using an instrumented timed up and go test. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2013**, *21*, 664–673. [CrossRef] [PubMed]

40. Weiss, A.; Herman, T.; Plotnik, M.; Brozgol, M.; Maidan, I.; Giladi, N.; Gurevich, T.; Hausdorff, J.M. Can an accelerometer enhance the utility of the Timed Up & Go Test when evaluating patients with Parkinson's disease? *Med. Eng. Phys.* **2010**, *32*, 119–125. [CrossRef] [PubMed]

41. Weiss, A.; Mirelman, A.; Buchman, A.S.; Bennett, D.A.; Hausdorff, J.M. Using a body-fixed sensor to identify subclinical gait difficulties in older adults with IADL disability: Maximizing the output of the Timed Up and Go. *PLoS ONE* **2013**, *8*, e68885. [CrossRef]

42. Arora, S.; Venkataraman, V.; Donohue, S.; Biglan, K.M.; Dorsey, E.R.; Little, M.A. High Accuracy Discrimination of Parkinson's Disease Participants from Healthy Controls using Smartphones. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 3641–3644. [CrossRef]

43. Schlachetzki, J.C.; Barth, J.; Marxreiter, F.; Gossler, J.; Kohl, Z.; Reinfelder, S.; Gassner, H.; Aminian, K.; Eskofier, B.M.; Winkler, J.; et al. Wearable sensors objectively measure gait parameters in Parkinson's disease. *PLoS ONE* **2017**, *12*, e0183989. [CrossRef]

44. Krupicka, R.; Viteckova, S.; Cejka, V.; Klempir, O.; Szabo, Z.; Ruzicka, E. BradykAn: A Motion Capture System for Objectification of Hand Motor Tests in Parkinson DISEASE. In Proceedings of the 2017 E-Health and Bioengineering Conference (EHB), Sinaia, Romania, 22–24 June 2017; pp. 446–449. [CrossRef]

45. Abtahi, M.; Bahram Borgheai, S.; Jafari, R.; Constant, N.; Diouf, R.; Shahriari, Y.; Mankodiya, K. Merging fNIRS-EEG Brain Monitoring and Body Motion Capture to Distinguish Parkinson's Disease. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 1246–1253. [CrossRef] [PubMed]

46. Castaño-Pino, Y.J.; González, M.C.; Quintana-Peña, V.; Valderrama, J.; Muñoz, B.; Orozco, J.; Navarro, A. Automatic Gait Phases Detection in Parkinson Disease: A Comparative Study. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 798–802. [CrossRef]

47. Phan, D.; Horne, M.; Pathirana, P.N.; Farzanehfar, P. Measurement of Axial Rigidity and Postural Instability Using Wearable Sensors. *Sensors* **2018**, *18*, 495. [CrossRef] [PubMed]
48. Del Din, S.; Elshehabi, M.; Galna, B.; Hobert, M.A.; Warmerdam, E.; Suenkel, U.; Brockmann, K.; Metzger, F.; Hansen, C.; Berg, D.; et al. Gait analysis with wearables predicts conversion to Parkinson disease. *Ann. Neurol.* **2019**, *86*, 357–367. [CrossRef] [PubMed]
49. Rovini, E.; Maremmani, C.; Moschetti, A.; Esposito, D.; Cavallo, F. Comparative Motor Pre-clinical Assessment in Parkinson's Disease Using Supervised Machine Learning Approaches. *Ann. Biomed. Eng.* **2018**, *46*, 2057–2068. [CrossRef] [PubMed]
50. Rehman, R.Z.U.; Del Din, S.; Shi, J.Q.; Galna, B.; Lord, S.; Yarnall, A.J.; Guan, Y.; Rochester, L. Comparison of Walking Protocols and Gait Assessment Systems for Machine Learning-Based Classification of Parkinson's Disease. *Sensors* **2019**, *19*, 5363. [CrossRef]
51. Buckley, C.; Galna, B.; Rochester, L.; Mazzà, C. Quantification of upper body movements during gait in older adults and in those with Parkinson's disease: Impact of acceleration realignment methodologies. *Gait Posture* **2017**, *52*, 265–271. [CrossRef]
52. Cai, G.; Shi, W.; Wang, Y.; Weng, H.; Chen, L.; Yu, J.; Chen, Z.; Lin, F.; Ren, K.; Zeng, Y.; et al. Specific Distribution of Digital Gait Biomarkers in Parkinson's Disease Using Body-Worn Sensors and Machine Learning. *J. Gerontol. Biol. Sci. Med. Sci.* **2023**, *78*, 1348–1354. [CrossRef] [PubMed]
53. Jovanovic, L.; Damaševičius, R.; Matić, R.; Kabiljo, M.; Simic, V.; Kunjadic, G.; Antonijevic, M.; Zivkovic, M.; Bacanin, N. Detecting Parkinson's disease from shoe-mounted accelerometer sensors using Convolutional Neural Networks optimized with modified metaheuristics. *PeerJ Comput. Sci.* **2024**, *10*, e2031. [CrossRef] [PubMed]
54. Battista, L.; Romaniello, A. A New Wrist-Worn Tool Supporting the Diagnosis of Parkinsonian Motor Syndromes. *Sensors* **2024**, *24*, 1965. [CrossRef]
55. Yue, P.; Li, Z.; Zhou, M.; Wang, X.; Yang, P. Wearable-Sensor-Based Weakly Supervised Parkinson's Disease Assessment with Data Augmentation. *Sensors* **2024**, *24*, 1196. [CrossRef]
56. Bailo, G.; Saibene, F.L.; Bandini, V.; Arcuri, P.; Salvatore, A.; Meloni, M.; Castagna, A.; Navarro, J.; Lencioni, T.; Ferrarin, M.; et al. Characterization of Walking in Mild Parkinson's Disease: Reliability, Validity and Discriminant Ability of the Six-Minute Walk Test Instrumented with a Single Inertial Sensor. *Sensors* **2024**, *24*, 662. [CrossRef]
57. Nair, P.; Baghini, M.S.; Pendharkar, G.; Chung, H. Detecting early-stage Parkinson's disease from gait data. *Proc. Inst. Mech. Eng. H* **2023**, *237*, 1287–1296. [CrossRef]
58. Keloth, S.M.; Viswanathan, R.; Jelfs, B.; Arjunan, S.; Raghav, S.; Kumar, D. Which Gait Parameters and Walking Patterns Show the Significant Differences Between Parkinson's Disease and Healthy Participants? *Biosensors* **2019**, *9*, 59. [CrossRef]
59. Trabassi, D.; Serrao, M.; Varrecchia, T.; Ranavolo, A.; Coppola, G.; De Icco, R.; Tassorelli, C.; Castiglia, S.F. Machine Learning Approach to Support the Detection of Parkinson's Disease in IMU-Based Gait Analysis. *Sensors* **2022**, *22*, 3700. [CrossRef]
60. Marin, F.; Warmerdam, E.; Marin, Z.; Ben Mansour, K.; Maetzler, W.; Hansen, C. Scoring the Sit-to-Stand Performance of Parkinson's Patients with a Single Wearable Sensor. *Sensors* **2022**, *22*, 8340. [CrossRef]
61. Gourrame, K.; Griškevičius, J.; Haritopoulos, M.; Lukšys, D.; Jatužis, D.; Kaladytė-Lokominienė, R.; Bunevičiūtė, R.; Mickutė, G. Parkinson's disease classification with CWNN: Using wavelet transformations and IMU data fusion for improved accuracy. *Technol. Health Care* **2023**, *31*, 2447–2455. [CrossRef] [PubMed]
62. Meng, L.; Pang, J.; Yang, Y.; Chen, L.; Xu, R.; Ming, D. Inertial-Based Gait Metrics During Turning Improve the Detection of Early-Stage Parkinson's Disease Patients. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2023**, *31*, 1472–1482. [CrossRef] [PubMed]
63. Caramia, C.; Torricelli, D.; Schmid, M.; Muñoz-Gonzalez, A.; Gonzalez-Vargas, J.; Grandas, F.; Pons, J.L. IMU-Based Classification of Parkinson's Disease From Gait: A Sensitivity Analysis on Sensor Location and Feature Selection. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 1765–1774. [CrossRef]
64. Mahadevan, N.; Demanuele, C.; Zhang, H.; Volfson, D.; Ho, B.; Erb, M.; Patel, S. Development of digital biomarkers for resting tremor and bradykinesia using a wrist-worn wearable device. *npj Digit. Med.* **2020**, *3*, 5. [CrossRef]
65. Lin, Z.; Dai, H.; Xiong, Y.; Xia, X.; Horng, S.J. Quantification Assessment of Bradykinesia in Parkinson's DISEASE Based on a Wearable Device. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Republic of Korea, 11–15 July 2017; pp. 803–806. [CrossRef]
66. Borzì, L.; Mazzetta, I.; Zampogna, A.; Suppa, A.; Irrera, F.; Olmo, G. Predicting Axial Impairment in Parkinson's Disease through a Single Inertial Sensor. *Sensors* **2022**, *22*, 412. [CrossRef]
67. Ymeri, G.; Salvi, D.; Olsson, C.M.; Wassenburg, M.V.; Tsanas, A.; Svenningsson, P. Quantifying Parkinson's disease severity using mobile wearable devices and machine learning: The ParkApp pilot study protocol. *BMJ Open* **2023**, *13*, e077766. [CrossRef] [PubMed]
68. Bobić, V.; Djurić-Jovičić, M.; Dragašević, N.; Popović, M.B.; Kostić, V.S.; Kvaščev, G. An Expert System for Quantification of Bradykinesia Based on Wearable Inertial Sensors. *Sensors* **2019**, *19*, 2644. [CrossRef]
69. Han, Y.; Liu, X.; Zhang, N.; Zhang, X.; Zhang, B.; Wang, S.; Liu, T.; Yi, J. Automatic Assessments of Parkinsonian Gait with Wearable Sensors for Human Assistive Systems. *Sensors* **2023**, *23*, 2104. [CrossRef]
70. Heijmans, M.; Habets, J.; Herff, C.; Aarts, J.; Stevens, A.; Kuijf, M.; Kubben, P. Monitoring Parkinson's disease symptoms during daily life: A feasibility study. *npj Park. Dis.* **2019**, *5*, 21. [CrossRef] [PubMed]
71. Wu, Z.; Jiang, X.; Zhong, M.; Shen, B.; Zhu, J.; Pan, Y.; Dong, J.; Xu, P.; Zhang, W.; Zhang, l. Wearable Sensors Measure Ankle Joint Changes of Patients with Parkinson's Disease before and after Acute Levodopa Challenge. *Park. Dis.* **2020**, *2020*, 2976535. [CrossRef]

72. Wu, X.; Ma, L.; Wei, P.; Shan, Y.; Chan, P.; Wang, K.; Zhao, G. Wearable sensor devices can automatically identify the ON-OFF status of patients with Parkinson's disease through an interpretable machine learning model. *Front. Neurol.* **2024**, *15*, 1387477. [CrossRef]

73. Pugh, R.J.; Higgins, R.D.; Min, H.; Wutzke, C.J.; Guccione, A.A. Turns while walking among individuals with Parkinson's disease following overground locomotor training: A pilot study. *Clin. Biomech.* **2024**, *114*, 106234. [CrossRef]

74. Delrobaei, M.; Memar, S.; Pieterman, M.; Stratton, T.W.; McIsaac, K.; Jog, M. Towards remote monitoring of Parkinson's disease tremor using wearable motion capture systems. *J. Neurol. Sci.* **2018**, *384*, 38–45. [CrossRef] [PubMed]

75. Memar, S.; Delrobaei, M.; Pieterman, M.; McIsaac, K.; Jog, M. Quantification of whole-body bradykinesia in Parkinson's disease participants using multiple inertial sensors. *J. Neurol. Sci.* **2018**, *387*, 157–165. [CrossRef]

76. di Biase, L.; Summa, S.; Tosi, J.; Taffoni, F.; Marano, M.; Cascio Rizzo, A.; Vecchio, F.; Formica, D.; Di Lazzaro, V.; Di Pino, G.; et al. Quantitative Analysis of Bradykinesia and Rigidity in Parkinson's Disease. *Front. Neurol.* **2018**, *9*, 121. [CrossRef]

77. Delrobaei, M.; Baktash, N.; Gilmore, G.; McIsaac, K.; Jog, M. Using Wearable Technology to Generate Objective Parkinson's Disease Dyskinesia Severity Score: Possibilities for Home Monitoring. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 1853–1863. [CrossRef] [PubMed]

78. Singh, M.; Prakash, P.; Kaur, R.; Sowers, R.; Brašić, J.R.; Hernandez, M.E. A Deep Learning Approach for Automatic and Objective Grading of the Motor Impairment Severity in Parkinson's Disease for Use in Tele-Assessments. *Sensors* **2023**, *23*, 9004. [CrossRef] [PubMed]

79. Prakash, P.; Kaur, R.; Levy, J.; Sowers, R.; Brašić, J.; Hernandez, M.E. A Deep Learning Approach for Grading of Motor Impairment Severity in Parkinson's Disease. In Proceedings of the 2023 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Sydney, Australia, 24–27 July 2023; pp. 1–4. [CrossRef]

80. Bremm, R.P.; Pavelka, L.; Garcia, M.M.; Mombaerts, L.; Krüger, R.; Hertel, F. Sensor-Based Quantification of MDS-UPDRS III Subitems in Parkinson's Disease Using Machine Learning. *Sensors* **2024**, *24*, 2195. [CrossRef] [PubMed]

81. Huo, W.; Angeles, P.; Tai, Y.; Pavese, N.; Wilson, S.; Hu, M.; Vaidyanathan, R. A Heterogeneous Sensing Suite for Multisymptom Quantification of Parkinson's Disease. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 1397–1406. [CrossRef] [PubMed]

82. Yan, F.; Gong, J.; Zhang, Q.; He, H. Learning Motion Primitives for the Quantification and Diagnosis of Mobility Deficits. *IEEE Trans. Biomed. Eng.* **2024**, 1–10. [CrossRef] [PubMed]

83. Hssayeni, M.D.; Jimenez-Shahed, J.; Burack, M.A.; Ghoraani, B. Wearable Sensors for Estimation of Parkinsonian Tremor Severity during Free Body Movements. *Sensors* **2019**, *19*, 4215. [CrossRef] [PubMed]

84. Zajki-Zechmeister, T.; Kögl, M.; Kalsberger, K.; Franthal, S.; Homayoon, N.; Katschnig-Winter, P.; Wenzel, K.; Zajki-Zechmeister, L.; Schwingenschuh, P. Quantification of tremor severity with a mobile tremor pen. *Heliyon* **2020**, *6*, e04702. [CrossRef] [PubMed]

85. Lonini, L.; Dai, A.; Shawen, N.; Simuni, T.; Poon, C.; Shimanovich, L.; Daeschler, M.; Ghaffari, R.; Rogers, J.; Jayaraman, A. Wearable sensors for Parkinson's disease: Which data are worth collecting for training symptom detection models. *npj Digit. Med.* **2018**, *1*, 64. [CrossRef] [PubMed]

86. Tsakanikas, V.; Ntanis, A.; Rigas, G.; Androutsos, C.; Boucharas, D.; Tachos, N.; Skaramagkas, V.; Chatzaki, C.; Kefalopoulou, Z.; Tsiknakis, M.; et al. Evaluating Gait Impairment in Parkinson's Disease from Instrumented Insole and IMU Sensor Data. *Sensors* **2023**, *23*, 3902. [CrossRef]

87. Chomiak, T.; Xian, W.; Pei, Z.; Hu, B. A novel single-sensor-based method for the detection of gait-cycle breakdown and freezing of gait in Parkinson's disease. *J. Neural Transm.* **2019**, *126*, 1029–1036. [CrossRef]

88. Punin, C.; Barzallo, B.; Clotet, R.; Bermeo, A.; Bravo, M.; Bermeo, J.P.; Llumiguano, C. A Non-Invasive Medical Device for Parkinson's Patients with Episodes of Freezing of Gait. *Sensors* **2019**, *19*, 737. [CrossRef]

89. Li, B.; Zhang, Y.; Tang, L.; Gao, C.; Gu, D. Automatic Detection System for Freezing of Gait in Parkinson's Disease Based on the Clustering Algorithm. In Proceedings of the 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Xi'an, China, 25–27 May 2018; pp. 1640–1644.

90. Camps, J.; Samà, A.; Martín, M.; Rodríguez-Martín, D.; Pérez-López, C.; Moreno Arostegui, J.M.; Cabestany, J.; Català, A.; Alcaine, S.; Mestre, B.; et al. Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn inertial measurement unit. *Knowl.-Based Syst.* **2018**, *139*, 119–131. [CrossRef]

91. Samà, A.; Rodríguez-Martín, D.; Pérez-López, C.; Català, A.; Alcaine, S.; Mestre, B.; Prats, A.; Crespo, M.C.; Bayés, À. Determining the optimal features in freezing of gait detection through a single waist accelerometer in home environments. *Pattern Recognit. Lett.* **2018**, *105*, 135–143. [CrossRef]

92. Capecci, M.; Pepa, L.; Verdini, F.; Ceravolo, M.G. A smartphone-based architecture to detect and quantify freezing of gait in Parkinson's disease. *Gait Posture* **2016**, *50*, 28–33. [CrossRef] [PubMed]

93. Reches, T.; Dagan, M.; Herman, T.; Gazit, E.; Gouskova, N.A.; Giladi, N.; Manor, B.; Hausdorff, J.M. Using Wearable Sensors and Machine Learning to Automatically Detect Freezing of Gait during a FOG-Provoking Test. *Sensors* **2020**, *20*, 4474. [CrossRef] [PubMed]

94. Pham, T.T.; Moore, S.T.; Lewis, S.J.G.; Nguyen, D.N.; Dutkiewicz, E.; Fuglevand, A.J.; McEwan, A.L.; Leong, P.H. Freezing of Gait Detection in Parkinson's Disease: A Subject-Independent Detector Using Anomaly Scores. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2719–2728. [CrossRef] [PubMed]

95. Masiala, S.; Huijbers, W.; Atzmueller, M. Feature-set-engineering for detecting freezing of gait in Parkinson's disease using deep recurrent neural networks. *arXiv* **2019**, arXiv:1909.03428.

96. Mancini, M.; Shah, V.; Stuart, S.; Curtze, C.; Horak, F.; Safarpour, D.; Nutt, J. Measuring freezing of gait during daily-life: An open-source, wearable sensors approach. *J. Neuroeng. Rehabil.* **2021**, *18*, 1. [CrossRef] [PubMed]
97. Marcante, A.; Di Marco, R.; Gentile, G.; Pellicano, C.; Assogna, F.; Pontieri, F.E.; Spalletta, G.; Macchiusi, L.; Gatsios, D.; Giannakis, A.; et al. Foot Pressure Wearable Sensors for Freezing of Gait Detection in Parkinson's Disease. *Sensors* **2021**, *21*, 128. [CrossRef]
98. Sigcha, L.; Costa, N.; Pavón, I.; Costa, S.; Arezes, P.; López, J.M.; De Arcas, G. Deep Learning Approaches for Detecting Freezing of Gait in Parkinson's Disease Patients through On-Body Acceleration Sensors. *Sensors* **2020**, *20*, 1895. [CrossRef]
99. Bikias, T.; Iakovakis, D.; Hadjidimitriou, S.; Charisis, V.; Hadjileontiadis, L.J. DeepFoG: An IMU-Based Detection of Freezing of Gait Episodes in Parkinson's Disease Patients via Deep Learning. *Front. Robot. AI* **2021**, *8*, 537384. [CrossRef]
100. Pardoel, S.; Shalin, G.; Nantel, J.; Lemaire, E.D.; Kofman, J. Early Detection of Freezing of Gait during Walking Using Inertial Measurement Unit and Plantar Pressure Distribution Data. *Sensors* **2021**, *21*, 2246. [CrossRef]
101. Shi, B.; Tay, A.; Au, W.L.; Tan, D.M.L.; Chia, N.S.Y.; Yen, S.C. Detection of Freezing of Gait Using Convolutional Neural Networks and Data From Lower Limb Motion Sensors. *IEEE Trans. Biomed. Eng.* **2022**, *69*, 2256–2267. [CrossRef] [PubMed]
102. Rennie, L.; Löfgren, N.; Moe-Nilssen, R.; Opheim, A.; Dietrichs, E.; Franzén, E. The reliability of gait variability measures for individuals with Parkinson's disease and healthy older adults—The effect of gait speed. *Gait Posture* **2018**, *62*, 505–509. [CrossRef] [PubMed]
103. Myers, P.S.; McNeely, M.E.; Pickett, K.A.; Duncan, R.P.; Earhart, G.M. Effects of exercise on gait and motor imagery in people with Parkinson disease and freezing of gait. *Park. Relat. Disord.* **2018**, *53*, 89–95. [CrossRef] [PubMed]
104. Zadka, A.; Rabin, N.; Gazit, E.; Mirelman, A.; Nieuwboer, A.; Rochester, L.; Del Din, S.; Pelosin, E.; Avanzino, L.; Bloem, B.R.; et al. A wearable sensor and machine learning estimate step length in older adults and patients with neurological disorders. *npj Digit. Med.* **2024**, *7*, 142. [CrossRef] [PubMed]
105. Haji Ghassemi, N.; Hannink, J.; Roth, N.; Gaßner, H.; Marxreiter, F.; Klucken, J.; Eskofier, B.M. Turning Analysis during Standardized Test Using On-Shoe Wearable Sensors in Parkinson's Disease. *Sensors* **2019**, *19*, 3103. [CrossRef] [PubMed]
106. Nguyen, A.; Roth, N.; Haji Ghassemi, N.; Hannink, J.; Seel, T.; Klucken, J.; Gaßner, H.; Eskofier, B. Development and clinical validation of inertial sensor-based gait-clustering methods in Parkinson's disease. *J. Neuroeng. Rehabil.* **2019**, *16*, 77. [CrossRef] [PubMed]
107. Liuzzi, P.; Carpinella, I.; Anastasi, D.; Gervasoni, E.; Lencioni, T.; Bertoni, R.; Carrozza, M.C.; Cattaneo, D.; Ferrarin, M.; Mannini, A. Machine learning based estimation of dynamic balance and gait adaptability in persons with neurological diseases using inertial sensors. *Sci. Rep.* **2023**, *13*, 8640. [CrossRef]
108. Stack, E.; Agarwal, V.; King, R.; Burnett, M.; Tahavori, F.; Janko, B.; Harwin, W.; Ashburn, A.; Kunkel, D. Identifying balance impairments in people with Parkinson's disease using video and wearable sensors. *Gait Posture* **2018**, *62*, 321–326. [CrossRef]
109. Jehu, D.; Nantel, J. Fallers with Parkinson's disease exhibit restrictive trunk control during walking. *Gait Posture* **2018**, *65*, 246–250. [CrossRef]
110. Zhang, X.; Jin, Y.; Wang, M.; Ji, C.; Chen, Z.; Fan, W.; Rainer, T.; Guan, Q.; Li, Q. The impact of anxiety on gait impairments in Parkinson's disease: Insights from sensor-based gait analysis. *J. Neuroeng. Rehabil.* **2024**, *21*, 68. [CrossRef] [PubMed]
111. Romijnders, R.; Warmerdam, E.; Hansen, C.; Welzel, J.; Schmidt, G.; Maetzler, W. Validation of IMU-based gait event detection during curved walking and turning in older adults and Parkinson's Disease patients. *J. Neuroeng. Rehabil.* **2021**, *18*, 28. [CrossRef] [PubMed]
112. Lukšys, D.; Griškevičius, J. Application of nonlinear analysis for the assessment of gait in patients with Parkinson's disease. *Technol. Health Care* **2022**, *30*, 201–208. [CrossRef] [PubMed]
113. Zhang, H.; Deng, K.; Li, H.; Albin, R.L.; Guan, Y. Deep learning identifies digital biomarkers for self-reported Parkinson's disease. *Patterns* **2020**, *1*, 100042. [CrossRef] [PubMed]
114. Aich, S.; Pradhan, P.M.; Park, J.; Sethi, N.; Vathsa, V.S.S.; Kim, H.C. A validation study of freezing of gait (FoG) detection and machine-learning-based FoG prediction using estimated gait characteristics with a wearable accelerometer. *Sensors* **2018**, *18*, 3287. [CrossRef] [PubMed]
115. Klucken, J.; Barth, J.; Kugler, P.; Schlachetzki, J.; Henze, T.; Marxreiter, F.; Kohl, Z.; Steidl, R.; Hornegger, J.; Eskofier, B.; et al. Unbiased and mobile gait analysis detects motor impairment in Parkinson's disease. *PLOS ONE* **2013**, *8*, e56956. [CrossRef] [PubMed]
116. Patel, S.; Lorincz, K.; Hughes, R.; Huggins, N.; Growdon, J.; Standaert, D.; Akay, M.; Dy, J.; Welsh, M.; Bonato, P. Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 864–873. [CrossRef] [PubMed]
117. Bzdok, D.; Altman, N.; Krzywinski, M. Statistics versus machine learning. *Nat. Methods* **2018**, *15*, 233–234. [CrossRef]
118. Negrini, S.; Serpelloni, M.; Amici, C.; Gobbo, M.; Silvestro, C.; Buraschi, R.; Borboni, A.; Crovato, D.; Lopomo, N.F. Use of Wearable Inertial Sensor in the Assessment of Timed-Up-and-Go Test: Influence of Device Placement on Temporal Variable Estimation. In Proceedings of the International Conference on Wireless Mobile Communication and Healthcare, Vienna, Austria, 14–15 November 2017; pp. 310–317.
119. Palmerini, L.; Rocchi, L.; Mellone, S.; Valzania, F.; Chiari, L. Feature selection for accelerometer-based posture analysis in Parkinson's disease. *IEEE Trans. Inf. Technol. Biomed.* **2011**, *15*, 481–490. [CrossRef]
120. Gibb, W.R.; Lees, A.J. The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **1988**, *51*, 745–752. [CrossRef]

121. Keshtkarjahromi, M.; Abraham, D.S.; Gruber-Baldini, A.L.; Schrader, K.; Reich, S.G.; Savitt, J.M.; Coelln, R.V.; Shulman, L.M. Confirming Parkinson disease diagnosis: Patterns of diagnostic changes by movement disorder specialists. *Park. Dis.* **2022**, *2022*, 5535826. [CrossRef] [PubMed]

122. Ghourchian, S.; Gruber-Baldini, A.L.; Shakya, S.; Herndon, J.; Reich, S.G.; von Coelln, R.; Savitt, J.M.; Shulman, L.M. Weight loss and weight gain in Parkinson disease. *Park. Relat. Disord.* **2021**, *83*, 31–36. [CrossRef]

123. Shulman, L.M.; Gruber-Baldini, A.L.; Anderson, K.E.; Fishman, P.S.; Reich, S.G.; Weiner, W.J. The clinically important difference on the unified Parkinson's disease rating scale. *Arch. Neurol.* **2010**, *67*, 64–70. [CrossRef]

124. Miller, M.D.; Paradis, C.F.; Houck, P.R.; Mazumdar, S.; Stack, J.A.; Rifai, A.; Mulsant, B.; Reynolds, C.F. Rating chronic medical illness burden in geropsychiatric practice and research: Application of the cumulative illness rating scale. *Psychiatry Res.* **1992**, *41*, 237–248. [CrossRef]

125. Fillenbaum, G.G.; Smyer, M.A. The development, validity, and reliability of the OARS multidimensional functional assessment questionnaire. *J. Gerontol.* **1981**, *36*, 428–434. [CrossRef] [PubMed]

126. Cella, D.; Choi, S.W.; Condon, D.M.; Schalet, B.; Hays, R.D.; Rothrock, N.E.; Yount, S.; Cook, K.F.; Gershon, R.C.; Amtmann, D.; et al. PROMIS® adult health profiles: Efficient short-form measures of seven health domains. *Value Health* **2019**, *22*, 537–544. [CrossRef]

127. Nasreddine, Z.S.; Phillips, N.A.; Bédirian, V.; Charbonneau, S.; Whitehead, V.; Collin, I.; Cummings, J.L.; Chertkow, H. The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **2005**, *53*, 695–699. [CrossRef] [PubMed]

128. Goetz, C.G. The unified Parkinson's disease rating scale (UPDRS): Status and recommendations. *Mov. Disord.* **2003**, *18*, 738–750. [CrossRef]

129. Goetz, C.G.; Poewe, W.; Rascol, O.; Sampaio, C.; Stebbins, G.T.; Counsell, C.; Giladi, N.; Holloway, R.G.; Moore, C.G.; Wenning, G.K.; et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: Status and recommendations. *Mov. Disord.* **2004**, *19*, 1020–1028. [CrossRef]

130. von Coelln, R.; Dawe, R.J.; Leurgans, S.E.; Curran, T.A.; Truty, T.; Yu, L.; Barnes, L.L.; Shulman, J.M.; Shulman, L.M.; Bennett, D.A.; et al. Quantitative mobility metrics from a wearable sensor predict incident parkinsonism in older adults. *Park. Relat. Disord.* **2019**, *65*, 190–196. [CrossRef] [PubMed]

131. Dawe, R.J.; Leurgans, S.E.; Yang, J.; Bennett, J.M.; Hausdorff, J.M.; Lim, A.S.; Gaiteri, C.; Bennett, D.A.; Buchman, A.S. Association between quantitative gait and balance measures and total daily physical activity in community-dwelling older adults. *J. Gerontol. Biol. Sci. Med. Sci.* **2018**, *73*, 636–642. [CrossRef] [PubMed]

132. Bennett, D.A.; Buchman, A.S.; Boyle, P.A.; Barnes, L.L.; Wilson, R.S.; Schneider, J.A. Religious orders study and Rush memory and aging project. *J. Alzheimer'S Dis.* **2018**, *64*, S161–S189. [CrossRef]

133. Koop, M.M.; Ozinga, S.J.; Rosenfeldt, A.B.; Alberts, J.L. Quantifying turning behavior and gait in Parkinson's disease using mobile technology. *IBRO Rep.* **2018**, *5*, 10–16. [CrossRef]

134. Burden, R.L.; Faires, J. *Numerical Analysis*, 9th ed.; Brooks/Cole: Pacific Grove, CA, USA, 2015.

135. Salarian, A.; Zampieri, C.; Horak, F.B.; Carlson-Kuhta, P.; Nutt, J.G.; Aminian, K. Analyzing 180 degrees turns using an inertial system reveals early signs of progression of Parkinson's disease. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009; pp. 224–227.

136. Butterworth, S. On the theory of filter amplifiers. *Exp. Wirel. Wirel. Eng.* **1930**, *7*, 536–541.

137. Lomb, N.R. Least-squares frequency analysis of unequally spaced data. *Astrophys. Space Sci.* **1976**, *39*, 447–462. [CrossRef]

138. Scargle, J.D. Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J.* **1982**, *263*, 835–853. [CrossRef]

139. Aich, S.; Youn, J.; Chakraborty, S.; Pradhan, P.M.; Park, J.H.; Park, S.; Park, J. A supervised machine learning approach to detect the on/off state in Parkinson's disease using wearable based gait signals. *Diagnostics* **2020**, *10*, 421. [CrossRef]

140. Dehkordi, M.B.; Zaraki, A.; Setchi, R. Feature extraction and feature selection in smartphone-based activity recognition. *Procedia Comput. Sci.* **2020**, *176*, 2655–2664. [CrossRef]

141. Mehta, A.; Vaddadi, S.K.; Sharma, V.; Kala, P. A Phase-Wise Analysis of Machine Learning Based Human Activity Recognition using Inertial Sensors. In Proceedings of the 2020 IEEE 17th India Council International Conference (INDICON), New Delhi, India, 10–13 December 2020; pp. 1–7. [CrossRef]

142. Weiss, A.; Herman, T.; Plotnik, M.; Brozgol, M.; Giladi, N.; Hausdorff, J.M. An instrumented timed up and go: The added value of an accelerometer for identifying fall risk in idiopathic fallers. *Physiol. Meas.* **2011**, *32*, 2003. [CrossRef]

143. Akaike, H., Information Theory and an Extension of the Maximum Likelihood Principle. In *Selected Papers of Hirotugu Akaike*; Springer: New York, NY, USA, 1998; pp. 199–213. [CrossRef]

144. Buchman, A.S.; Dawe, R.J.; Leurgans, S.E.; Curran, T.A.; Truty, T.; Yu, L.; Barnes, L.L.; Hausdorff, J.M.; Bennett, D.A. Different combinations of mobility metrics derived from a wearable sensor are associated with distinct health outcomes in older adults. *J. Gerontol. Biol. Sci. Med. Sci.* **2020**, *75*, 1176. [CrossRef]

145. Poole, V.N.; Dawe, R.J.; Lamar, M.; Esterman, M.; Barnes, L.; Leurgans, S.E.; Bennett, D.A.; Hausdorff, J.M.; Buchman, A.S. Dividing attention during the Timed Up and Go enhances associations of several subtask performances with MCI and cognition. *PLoS ONE* **2022**, *17*, e0269398. [CrossRef]

146. der Laan, M.J.; Polley, E.C.; Hubbard, A.E. Super learner. *Stat. Appl. Genet. Mol. Biol.* **2007**, *6*, 25. [CrossRef] [PubMed]

147. Nelder, J.A.; Wedderburn, R.W.M. Generalized Linear Models. *J. R. Stat. Soc. Ser. A* **1972**, *135*, 370–384. [CrossRef]

148. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

149. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]

150. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]

151. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]

152. Gregorutti, B.; Michel, B.; Saint-Pierre, P. Grouped variable importance with random forests and application to multiple functional data analysis. *Comput. Stat. Data Anal.* **2015**, *90*, 15–35. [CrossRef]

153. Canty, A.; Ripley, B.D. boot: Bootstrap R (S-Plus) Functions. R Package Version 1.3-30. 2024. Available online: https://cran.r-project.org/web/packages/boot/index.html (accessed on 17 June 2024).

154. Davison, A.C.; Hinkley, D.V. *Bootstrap Methods and Their Applications*; Cambridge University Press: Cambridge, UK, 1997.

155. Adler, C.H.; Beach, T.G.; Hentz, J.G.; Shill, H.A.; Caviness, J.N.; Driver-Dunckley, E.; Sabbagh, M.N.; Sue, L.I.; Jacobson, S.A.; Belden, C.M.; et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease: Clinicopathologic study. *Neurology* **2014**, *83*, 406–412. [CrossRef]

156. Das, R. A comparison of multiple classification methods for diagnosis of Parkinson disease. *Expert Syst. Appl.* **2010**, *37*, 1568–1572. [CrossRef]

157. Muniz, A.M.S.; Liu, H.; Lyons, K.E.; Pahwa, R.; Liu, W.; Nobre, F.F.; Nadal, J. Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the effect of subthalamic stimulation in Parkinson disease on ground reaction force during gait. *J. Biomech.* **2010**, *43*, 720–726. [CrossRef] [PubMed]

158. Tien, I.; Glaser, S.D.; Aminoff, M.J. Characterization of gait abnormalities in Parkinson's disease using a wireless inertial sensor system. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2010**, *2010*, 3353–3356. [CrossRef] [PubMed]

159. Ferreira, M.I.A.S.N.; Barbieri, F.A.; Moreno, V.C.; Penedo, T.; Tavares, J.M.R.S. Machine learning models for Parkinson's disease detection and stage classification based on spatial-temporal gait parameters. *Gait Posture* **2022**, *98*, 49–55. [CrossRef] [PubMed]

160. Rehman, R.Z.U.; Buckley, C.; Micó-Amigo, M.E.; Kirk, C.; Dunne-Willows, M.; Mazzà, C.; Shi, J.Q.; Alcock, L.; Rochester, L.; Din, S.D. Accelerometry-based digital gait characteristics for classification of Parkinson's disease: What counts? *IEEE Open J. Eng. Med. Biol.* **2020**, *1*, 65–73. [CrossRef]

161. Moon, S.; Song, H.J.; Sharma, V.D.; Lyons, K.E.; Pahwa, R.; Akinwuntan, A.E.; Devos, H. Classification of Parkinson's disease and essential tremor based on balance and gait characteristics from wearable motion sensors via machine learning techniques: A data-driven approach. *J. Neuroeng. Rehabil.* **2020**, *17*, 1–8. [CrossRef] [PubMed]

162. Kleanthous, N.; Hussain, A.J.; Khan, W.; Liatsis, P. A new machine learning based approach to predict freezing of gait. *Pattern Recognit. Lett.* **2020**, *140*, 119–126. [CrossRef]

163. Mirelman, A.; Frank, M.B.O.; Melamed, M.; Granovsky, L.; Nieuwboer, A.; Rochester, L.; Din, S.D.; Avanzino, L.; Pelosin, E.; Bloem, B.R.; et al. Detecting sensitive mobility features for Parkinson's disease stages via machine learning. *Mov. Disord.* **2021**, *36*, 2144–2155. [CrossRef] [PubMed]

164. Kabir, M.F.; Ludwig, S. Enhancing the performance of classification using super learning. *Data-Enabled Discov. Appl.* **2019**, *3*, 5. [CrossRef]

165. Demonceau, M.; Donneau, A.F.; Croisier, J.L.; Skawiniak, E.; Boutaayamou, M.; Maquet, D.; Garraux, G. Contribution of a trunk accelerometer system to the characterization of gait in patients with mild-to-moderate Parkinson's disease. *IEEE J. Biomed. Health Inform.* **2015**, *19*, 1803–1808. [CrossRef]

166. Van Uem, J.M.T.; Walgaard, S.; Ainsworth, E.; Hasmann, S.E.; Heger, T.; Nussbaum, S.; Hobert, M.A.; Micó-Amigo, E.M.; Van Lummel, R.C.; Berg, D.; et al. Quantitative Timed-Up-and-Go Parameters in Relation to Cognitive Parameters and Health-Related Quality of Life in Mild-to-Moderate Parkinson's Disease. *PLoS ONE* **2016**, *11*, e0151997. [CrossRef]

167. Shawen, N.; O'Brien, M.; Venkatesan, S.; Lonini, L.; Simuni, T.; Hamilton, J.; Ghaffari, R.; Rogers, J.; Jayaraman, A. Role of data measurement characteristics in the accurate detection of Parkinson's disease symptoms using wearable sensors. *J. Neuroeng. Rehabil.* **2020**, *17*, 52. [CrossRef]

168. Strobl, C.; Boulesteix, A.L.; Kneib, T.; Augustin, T.; Zeileis, A. Conditional variable importance for random forests. *BMC Bioinform.* **2008**, *9*, 307. [CrossRef]

169. Gregorutti, B.; Michel, B.; Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **2017**, *27*, 659–678. [CrossRef]

170. Cooley, J.W.; Tukey, J.W. An algorithm for the machine calculation of complex fourier series. *Math. Comput.* **1965**, 19, 297–301. [CrossRef]

171. Phinyomark, A.; Phukpattaranont, P.; Limsakul, C. Feature reduction and selection for EMG signal classification. *Expert Syst. Appl.* **2012**, *39*, 7420–7431. [CrossRef]

172. Altin, C.; Er, O. Comparison of different time and frequency domain feature extraction methods on elbow gesture's EMG. *Eur. J. Interdiscip. Stud.* **2016**, *5*, 35. [CrossRef]

173. Sinderby, C.; Lindstrom, L.; Grassino, A.E. Automatic assessment of electromyogram quality. *J. Appl. Physiol.* **1995**, *79*, 1803–1815. [CrossRef]

174. Pepa, L.; Ciabattoni, L.; Verdini, F.; Capecci, M.; Ceravolo, M. Smartphone based fuzzy logic freezing of gait detection in Parkinson's disease. In Proceedings of the 2014 IEEE/ASME 10th International Conference on Mechatronic and Embedded Systems and Applications (MESA), Ancona, Italy, 10–12 September 2014; pp. 1–6. [CrossRef]

175. Bao, L.; Intille, S.S. Activity recognition from user-annotated acceleration data. In Proceedings of the International Conference on Pervasive Computing, Vienna, Austria, 21–23 April 2004; pp. 1–17. [CrossRef]

176. Phinyomark, A.; Phukpattaranont, P.; Limsakul, C. Fractal analysis features for weak and single-channel upper-limb EMG signals. *Expert Syst. Appl.* **2012**, *39*, 11156–11163. [CrossRef]

177. Hasni, H.; Yahya, N.; Asirvadam, V.S.; Jatoi, M.A. Analysis of electromyogram (EMG) for detection of neuromuscular disorders. In Proceedings of the 2018 International Conference on Intelligent and Advanced System (ICIAS), Kuala Lumpur, Malaysia, 13–14 August 2018; pp. 1–6. [CrossRef]

178. Sukumar, N.; Taran, S.; Bajaj, V. Physical actions classification of surface EMG signals using VMD. In Proceedings of the 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 3–5 April 2018; pp. 0705–0709. [CrossRef]

179. Kaiser, J. On a simple algorithm to calculate the 'energy' of a signal. In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Albuquerque, NM, USA, 3–6 April 1990; Volume 1, pp. 381–384. [CrossRef]

180. Penzel, T.; Kantelhardt, J.; Grote, L.; Peter, J.; Bunde, A. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. *IEEE Trans. Biomed. Eng.* **2003**, *50*, 1143–1151. [CrossRef]

181. Higuchi, T. Approach to an irregular time series on the basis of the fractal theory. *Phys. D Nonlinear Phenom.* **1988**, *31*, 277–283. [CrossRef]

182. Katz, M.J. Fractals and the analysis of waveforms. *Comput. Biol. Med.* **1988**, *18*, 145–156. [CrossRef] [PubMed]

183. Gneiting, T.; Ševčíková, H.; Percival, D.B. Estimators of fractal dimension: Assessing the roughness of time series and spatial data. *Stat. Sci.* **2012**, *27*, 247–277. [CrossRef]

184. Quiroz, J.C.; Banerjee, A.; Dascalu, S.M.; Lau, S.L. Feature selection for activity recognition from smartphone accelerometer data. *Intell. Autom. Soft Comput.* **2017**, 1–9. [CrossRef]

185. Ayman, A.; Attalah, O.; Shaban, H. An efficient human activity recognition framework based on wearable IMU wrist sensors. In Proceedings of the 2019 IEEE International Conference on Imaging Systems and Techniques (IST), Abu Dhabi, United Arab Emirates, 9–10 December 2019; pp. 1–5. [CrossRef]

186. Batool, M.; Jalal, A.; Kim, K. Sensors technologies for human activity analysis based on SVM optimized by PSO algorithm. In Proceedings of the 2019 International Conference on Applied and Engineering Mathematics (ICAEM), Taxila, Pakistan, 27–29 August 2019; pp. 145–150. [CrossRef]

187. Castiglia, S.F.; Trabassi, D.; Conte, C.; Ranavolo, A.; Coppola, G.; Sebastianelli, G.; Abagnale, C.; Barone, F.; Bighiani, F.; De Icco, R.; et al. Multiscale entropy algorithms to analyze complexity and variability of trunk accelerations time series in subjects with Parkinson's disease. *Sensors* **2023**, *23*, 4983. [CrossRef] [PubMed]

188. Venables, W.; Ripley, B. *Modern Applied Statistics with S*; Springer: New York, NY, USA, 2002.

189. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley: New York, NY, USA, 1990.

190. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA, 4–9 December 2017; pp. 4768–4777.