




IMU-Based Classification of Parkinson's Disease From Gait: A Sensitivity Analysis on Sensor Location and Feature Selection

Carlotta Caramia , Diego Torricelli, *Member, IEEE*, Maurizio Schmid , *Member, IEEE*,
Adriana Muñoz-Gonzalez, Jose Gonzalez-Vargas , Francisco Grandas,
and Jose L. Pons, *Senior Member, IEEE*

Abstract—Inertial measurement units (IMUs) have a long-lasting popularity in a variety of industrial applications from navigation systems to guidance and robotics. Their use in clinical practice is now becoming more common, thanks to miniaturization and the ability to integrate on-board computational and decision-support features. IMU-based gait analysis is a paradigm of this evolving process, and in this study its use for the assessment of Parkinson's disease (PD) is comprehensively analyzed. Data coming from 25 individuals with different levels of PD symptoms severity and an equal number of age-matched healthy individuals were included into a set of 6 different machine learning (ML) techniques, processing 18 different configurations of gait parameters taken from 8 IMU sensors. Classification accuracy was calculated for each configuration and ML technique, adding two meta-classifiers based on the results obtained from all individual techniques through majority of voting, with two different weighting schemes. Average classification accuracy ranged between 63% and 80% among classifiers and increased up to 96% for one meta-classifier configuration. Configurations based on a statistical preselection process showed the highest average classification accuracy. When reducing the number of sensors, features based on the joint range of motion were more accurate than those based on spatio-temporal parameters. In particular, best results were obtained with the knee range of motion, calculated with four IMUs, placed bilaterally. The obtained findings provide data-driven evidence on which combination of sensor configurations and classification methods to be used during IMU-based gait analysis to grade the severity level of PD.

Manuscript received December 31, 2017; revised April 10, 2018, July 17, 2018, and August 1, 2018; accepted August 5, 2018. Date of publication August 12, 2018; date of current version October 15, 2018. This work was supported in part by the Spanish Ministry of Economy, Industry and Competitiveness under Grant PI 17/02007 and in part by the H2020 Project EURO-BENCH “European Robotic framework for bipedal locomotion benchmarking,” g.a. n. 779963. (*Corresponding author: Maurizio Schmid.*)

C. Caramia and M. Schmid are with the Department of Engineering, Roma Tre University, Rome 00154, Italy (e-mail: carlotta.caramia@uniroma3.it; maurizio.schmid@uniroma3.it).

D. Torricelli, J. Gonzalez-Vargas, and J. L. Pons are with the Cajal Institute, Spanish National Research Council, Madrid 28006, Spain (e-mail: diego.torricelli@csic.es; jose.gonzalez@ottobock.de; jose.pons@csic.es).

A. Muñoz-Gonzalez and F. Grandas are with the Movement Disorders Unit, Universidad Complutense Madrid, Madrid 28040, Spain (e-mail: adri.munoz86@gmail.com; francisco.grandas@salud.madrid.org).

Digital Object Identifier 10.1109/JBHI.2018.2865218

Index Terms—Machine learning, wearable sensors, gait analysis, body sensor networks, feature extraction, Parkinson's disease.

I. INTRODUCTION

PARKINSON'S disease (PD) is a long-term neurodegenerative disorder of the central nervous system that causes motor—tremor at rest, rigidity, bradykinesia and postural instability—and non-motor manifestations [1]. The disease progresses over time and the symptoms usually grow both in severity and in quantity [2], increasing the chance of severe complications, causing a general worsening of the patient's quality of life. The number of affected individuals is growing [3] and, thus, the associated healthcare costs [4].

In the clinical practice, the diagnosis of PD is entirely based on neurological examinations, focused on the observation of motor signs, and on the completion of rating scales, such as the Unified Parkinson's Disorders Rating Scale (UPDRS), or the Hoehn and Yahr (H&Y) [6]. Possible limitations associated with this process are the difficulty to capture reliable fluctuations over time caused by medical response, or to monitor changes that, by definition, appear over longer periods of time than those needed for a clinical examination [7], [8]. This may lead to omit some important details, with possible late diagnosis—mainly in the early stages of the disease.

The extraction of accurate and unbiased data based on quantitative measurements is critical for the understanding of the aforementioned pathophysiological phenomenon [9]. The ability to provide clinicians with quantitative measures of motor performance represents an added value to the analysis of this disease, and the association of these measures with those already used in the clinical practice may lead to a more accurate assessment [7]. The ideal technological platforms must be simple to use, applicable for daily clinical practice, unobtrusive, and possibly low-cost. Moreover, the provided information should be reliable, accurate and easy to interpret [2].

In this scenario, Inertial Measurement Units (IMUs) almost entirely fit the previous requirements. IMUs have been initially introduced in the scientific community of aerospace engineering, as means to estimate attitude of flying objects [10], and to help guidance in navigation systems [11]. They were then

integrated into a variety of commercial vehicles, and their use is now being even more popular in unmanned systems, while miniaturization made it possible to open new application fields in consumer electronics, such as fitness tracking systems, gaming, and sports technology.

In clinical practice, IMU technology has been borrowed from the previously cited application fields to replace motion capture systems. IMU-based gait analysis is a useful tool for the assessment [12], the screening [13] and the diagnosis of PD [14], due to its ability to capture a wide spectrum of gait characteristics induced by motor impairments [15]. Based on instrumented motion analysis, PD gait has been associated with: increased double support time [16]; reduced step/stride length and gait speed [17], [18], [19]; decreased gait symmetry and regularity [14]; reduced range of rotation of thigh, knee [16], [20], trunk [17] and foot [18]; reduced knee symmetry [20].

Based on modifications to walking patterns, machine learning (ML) techniques were introduced to identify gait patterns of people with PD and discriminate them from those of healthy subjects: load sensors [21], also in combination with spatio-temporal kinematic features [22]; motion capture systems and force plates [23], [24], [25]; sensorized walkways [26] were used to this end. Focusing on IMU-based measures, Tien *et al.* [27], identified PD walking patterns by selecting features related to the range of motion of the foot, and using a support vector machine (SVM); Barth *et al.* [28] used statistical features, together with some step features, using linear discriminant analysis (LDA). Klucken *et al.* [15] included step, sequence and frequency-dependent features, combining information coming from different tasks in addition to straight walking, as inputs to an AdaBoost classifier. The aforementioned works used one IMU placed on each foot. While simplifying the experimental setup, this approach excludes those motor symptoms present in the other parts of the body, e.g., proximal lower limb joints, trunk, or arms, which are commonly considered in the clinical evaluation of PD.

In the direction of including motor symptoms of the whole body, Cuzzolin *et al.* [29] used Hidden Markov Models to represent gait sequences from lumbar IMUs, in association to a k-Nearest Neighbor (k-NN) classifier. Arora *et al.* [30] used only statistical features extracted from a smartphone accelerometer during postural sway and gait tests for the classification process through a random forest.

The aim of this work is to identify which of the gait features considered in this heterogeneous scenario are able to better distinguish the presence of PD from age-matched healthy subjects. We propose a concurrent analysis of the effects of IMU sensors location on the extraction of these features, and compared the classification ability of each configuration by applying most of the ML techniques available in the literature. The ultimate goal is to verify whether it is possible to integrate these techniques directly into a cyber-physical system based on IMU technology to monitor the progress of the disease. This approach may be used as a complementary tool to the current observational methods of gait analysis in PD.

II. MATERIALS AND METHODS

A. Participants and Procedure

27 patients with idiopathic PD at different stages of the H&Y motor scale (stage I, II and III, age 43-88, 13 males) and 27 healthy individuals (age range 41-79, 13 males) were recruited for the experiments. The neurologist involved in the experiments selected the patients from a digital archive of the Hospital General Universitario Gregorio Marañón, Movement Disorders Unit, Madrid, Spain. Patients were chosen according to specific inclusion criteria: stage I, II or III of the H&Y motor scale, absence of motor, physical, or mental problems other than those caused by PD, ability to walk independently for at least 30 meters without any aid, BMI < 30. Two patients were discarded because the required conditions of inclusion were not fulfilled; consequently, two age-matched healthy adults were not considered. This process resulted in the inclusion of 8 subjects belonging to stage I, 9 to stage II, and 8 to stage III. Each participant gave written informed consent prior to the experiment, according to the declaration of Helsinki.

The experiments were performed in an internal hallway of the hospital. Participants were asked to walk along a straight path of 15 meters, repeating this task three times. Anthropometric measures of foot, shank, thigh and trunk length were recorded prior to the experiment.

B. Instrumentation and Data Recording

The measuring instrumentation included the Motion Capture System Tech-MCS (Technaid S.L., Spain), consisting of eight IMUs, one portable acquisition hardware (Tech-Hub), and a dedicated software. Each IMU ($36 \times 26 \times 11$ mm, 10 gr) is composed of a triaxial accelerometer (± 4 g), a gyroscope (± 2000 °/s), and a magnetometer (± 8.1 gauss). All the sensors are connected to the Tech-Hub ($156 \times 100 \times 43$ mm, 225 gr), placed at the waist level on the left side, to synchronize and store data (sampling rate 50 samples/s) into a micro-SD card.

The sensors were located in the lower and upper parts of the body (see Fig. 1): one sensor was placed on each foot dorsum, one on each shank, one on each thigh, one on the chest and one in the back side on the lumbar zone through elastic belts.

The Tech-MCS software directly calculates the angles between body segments, taking the lumbar sensor as the body reference system, and stores the data for their post processing.

C. Feature Extraction and Grouping

Two categories of parameters were extracted from raw data: range of motions (RoMs) and spatio-temporal parameters. Each RoM is defined as the difference between the maximum and minimum angle drawn in the sagittal plane between two adjacent articular segments within one gait cycle; RoMs are calculated for ankle, knee, hip and chest.

The spatio-temporal parameters are the step length, step time, stride time and stride speed. Strides were identified based on the heel strike event, calculated as the time when the knee reached its maximum extension after its maximum flexion.

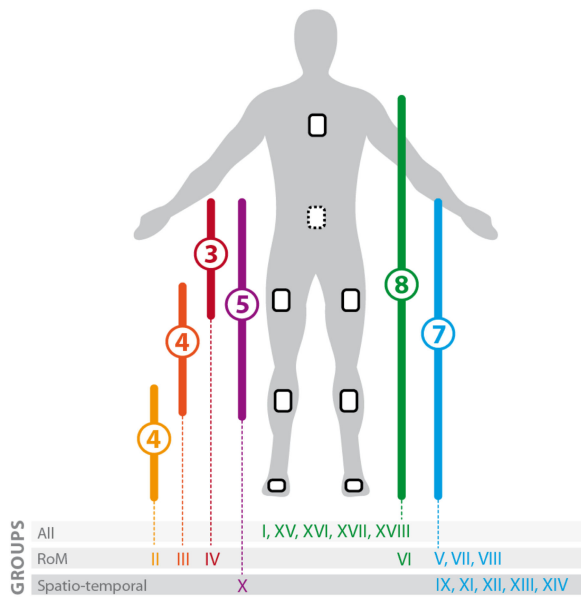


Fig. 1. Sensor locations and grouping definition. Numbers circled in color refer to the number of sensors needed to extract features of that specific group (I–XVIII, see text for details), with lines showing the sensors locations.

Step length, i.e., the distance between the toes of the two feet during the double support phase, was estimated by trigonometry, using the relative orientation angles between the relevant body joints (ankle, knee, hip) and the anthropometric measures of foot, shank and thigh: specifically, in correspondence to each foot initial contact, toe position is estimated by considering relative orientation angles between the relevant body segments; step length was then estimated by calculating the distance between the two toe estimated positions and normalized by the subject's height.

Step time was defined as the difference in time between two consecutive heel strikes of opposite legs.

Stride speed was calculated as the ratio between stride length (i.e., the sum of two consecutive step lengths) and stride time.

Only the second and third walking trials were considered for the data processing. The first trial was used to allow the subject to get used to the experimental setup and the task. The first and last three steps of each trial were excluded to avoid acceleration and deceleration due to start-stop transitions.

For each parameter we calculated the following three features: mean value (μ) across strides; standard deviation (SD) across strides; coefficient of variation (CoV), defined as the ratio of SD to μ . All features were calculated after concatenating both walking trials into one dataset, thus obtaining, for each subject, a single value for each feature. All features, except those relative to the trunk RoM, were calculated for both sides. To represent the symmetry between sides, we calculated two additional features, “ratio” and “asymmetry”, calculated as the ratio and difference between the left and the right-side values. This resulted in a set of 87 features, which were organized into 18 groups according to the following criteria (see Table I):

- Group I included all 87 features.

- Groups II to VIII included features derived exclusively from RoMs. Of these, three groups (II, III, and IV) were joint-specific, two groups (V and VI) considered more than one joint, and the remaining two groups (VII and VIII) specifically included asymmetry and variability features from RoMs.
- Groups IX to XIV included features derived exclusively from spatio-temporal parameters. Of these, one group (XII) included all features, three groups (IX, X, and XI) included features associated with only one spatio-temporal parameter, and the remaining two groups (XIII and XIV) included asymmetry and variability features from all the spatio-temporal parameters.
- Group XV included features resulting statistically different between controls and patients, according to a t-test (p -value < 0.05).
- Groups XVI and XVII included a subset of independent features resulting from a principal component analysis (PCA) [31], using a threshold of 90% on the cumulative variance. In group XVI, PCA was applied to the entire set of 87 features, while in group XVII it was applied to the set of features included in group XV.
- A last group (XVIII), called *random*, is introduced for statistical comparison against the other groups: it included ten vectors, each composed of features randomly chosen from the entire dataset. The number of components corresponded to the number of features of the group that showed the highest value of the averaged classification accuracy. This dataset was used as reference to compare the classification accuracy of each feature group with respect to a baseline random dataset (see Section E. for details).

Figure 1 shows the sensors locations and the number of sensors that are needed for each group.

D. Classification

Classification was then evaluated in two phases: first, the overall samples of healthy and PD groups, regardless of their H&Y assigned stage, were included to test the overall ability to distinguish between controls and patients; then, the procedure was repeated to test the classification accuracy between the population sample of each PD stage (8 patients) against an age-matched sample of healthy individuals.

To test the classification accuracy in the first phase, six different ML techniques were applied to each of the 18 datasets: Naive Bayes (NB) [32], LDA [28], k-NN [29], Decision Tree (DT) [33], and SVM with both linear [34] and non-linear kernel (rbf) bases [27]. For each ML technique, the internal parameter configuration was chosen heuristically based on the one that provided the best results in terms of classification accuracy in the testing process. The classification accuracy was calculated by applying a 5-fold cross validation procedure with 100 repetitions, to obtain reliable performance lowering the bias introduced by data splitting [30], [35].

In the second phase, we selected only one ML technique—the SVM with non-linear kernel basis, which obtained the highest

TABLE I
GROUPS DEFINITION

Group Name	Group	Included parameters	Extracted features	No. of gait features	No. of sensors	Body location
Overall	I	Joint and spatio-temporal parameters	μ , SD, CoV of R&L side, asymmetry and ratio	87	8	R&L ankle, shank, thigh; trunk, chest
Joint-specific RoMs	II	Ankle RoM	μ , SD, CoV of R&L side, asymmetry and ratio	12	4	R&L ankle, shank
	III	Knee RoM		12	4	R&L shank, thigh
	IV	Hip RoM		12	3	R&L thigh; trunk
Multiple joint RoMs	V	RoM: ankle, knee, hip	μ , SD, CoV of R&L side, asymmetry and ratio	36	7	R&L ankle, shank, thigh; trunk
	VI	RoM: ankle, knee, hip, chest		39	8	R&L ankle, shank, thigh; trunk, chest
RoMs asymmetry	VII	RoM: ankle, knee, hip	μ , SD, CoV of asymmetry	9	7	R&L ankle, shank, thigh
RoMs variability	VIII	RoM: ankle, knee, hip	CoV of R&L side	6	7	R&L ankle, shank, thigh
Individual spatio-temporal	IX	Step Length	μ , SD, CoV of R&L side, asymmetry and ratio	12	7	R&L ankle, shank, thigh; trunk
	X	Step Time		12	5	R&L shank, thigh; trunk
	XI	Speed		12	7	R&L ankle, shank, thigh; trunk
Multiple spatio-temporal	XII	Step Length, Step Time, Speed & Stride Time	μ , SD, CoV of R&L side, asymmetry and ratio	48	7	R&L ankle, shank, thigh; trunk
Spatio-temporal asymmetry	XIII	Step Length, Step Time, Speed & Stride Time	μ , SD, CoV of asymmetry	12	7	R&L ankle, shank, thigh; trunk
Spatio-temporal variability	XIV	Step Length, Step Time, Speed & Stride Time	CoV of R&L side	8	7	R&L ankle, shank, thigh; trunk
Statistically Different (St-D)	XV	Joint and spatio-temporal parameters	μ , SD, CoV of R&L side, asymmetry and ratio	43	8	R&L ankle, shank, thigh; trunk, chest
PCA on Overall	XVI	Joint and spatio-temporal parameters	μ , SD, CoV of R&L side, asymmetry and ratio	16	8	R&L ankle, shank, thigh; trunk, chest
PCA on St-D	XVII	Statistically Different parameters	μ , SD, CoV of R&L side, asymmetry and ratio	13	8	R&L ankle, shank, thigh; trunk, chest
Random	XVIII	Joint and spatio-temporal parameters	μ , SD, CoV of R&L side, asymmetry and ratio	13*	8	R&L ankle, shank, thigh; trunk, chest

R&L = right and left.

* The number of features in group XVIII corresponded to that of the group that showed the highest value of the average classification accuracy.

classification accuracy in the first phase—as representative of the overall pool of ML techniques.

We used the Machine Learning toolbox of Matlab R2017a to define and cross-validate all the classifiers.

A seventh classifier, called Majority of votes, was then created as a weighted combination of the six basic classifiers. Two types of this classifier (A and B) were created, based on two different weighting schemes: in type A, the weights associated to each classifier corresponded to the training accuracy values obtained in that group, divided by the sum of all the classifiers' accuracies; in type B, equal weights were associated to all classifiers. In both cases, the subject was finally assigned to the class with the highest weighted sum. In type B, if the number of votes was equal, the subject was assigned to both classes in equal parts. For both conditions, each subject was tested separately with the six trained classifiers. Repeating this process for all the subjects allowed calculating a classification accuracy for each group.

E. Evaluation of Classification Accuracy

The classification accuracy values obtained for each repetition underwent statistical analysis to check for significant differences associated with the feature group choice and ML technique used. To this end, normality of classification accuracy values was checked with the Lilliefors test. If normality was granted,

then a 2-way ANOVA, with feature group and ML technique as factors, was performed (Friedman's test was used if normality was rejected at 0.05 level of significance). Post-hoc analysis (Fisher's Least Significance Difference) was then performed to check for individual pairwise differences among the factor items. In particular, all the individual pairwise combinations between machine learning techniques were checked, while, for the feature groups, pairwise post-hoc analysis was performed on each feature group against the random one, taken as reference.

III. RESULTS

Table II shows the mean values (\pm standard deviation) of step length, step time, stride time and speed values for five datasets, i.e., each H&Y stage-specific PD dataset, the entire sample of patients, and the overall sample of healthy subjects.

Classification accuracy of the different feature groups, on the overall population sample containing all the controls and PD patients, is presented in Table III. These values were obtained averaging the accuracies of the six classifiers (individually shown in Table IV), grouped with respect to the number of sensors necessary to calculate that specific group of features.

The joint-specific hip RoM group (IV) was the only one defined by the use of 3 sensors, and its classification accuracy resulted lower, yet not significantly different, than the random

TABLE II
MEAN VALUE \pm STANDARD DEVIATION OF GAIT FEATURES

	Stage I (no.8)	Age-matched healthy (stage I)	Stage II (no.8)	Age-matched healthy (stage II)	Stage III (no.8)	Age-matched healthy (stage III)	All stages (no.25)	All healthy (no.25)
StepLength *	0.45 \pm 0.05	0.51 \pm 0.06	0.39 \pm 0.07	0.48 \pm 0.07	0.34 \pm 0.07	0.46 \pm 0.07	0.39 \pm 0.08	0.48 \pm 0.07
StepTime[s]	0.57 \pm 0.02	0.52 \pm 0.04	0.54 \pm 0.09	0.55 \pm 0.05	0.53 \pm 0.04	0.54 \pm 0.3	0.55 \pm 0.17	0.54 \pm 0.04
StrideTime[s]	1.15 \pm 0.03	1.04 \pm 0.08	1.08 \pm 0.17	1.09 \pm 0.09	1.06 \pm 0.07	1.09 \pm 0.07	1.10 \pm 0.11	1.07 \pm 0.08
Speed[s⁻¹ *	0.79 \pm 0.09	0.98 \pm 0.13	0.76 \pm 0.21	0.88 \pm 0.18	0.65 \pm 0.14	0.85 \pm 0.17	0.72 \pm 0.17	0.90 \pm 0.16
Hip RoM [°]	44.6 \pm 6.4	49.0 \pm 4.8	41.7 \pm 6.7	45.6 \pm 6.9	34.3 \pm 6.7	43.8 \pm 8.4	39.8 \pm 8.0	45.2 \pm 7.7
Knee RoM [°]	67.1 \pm 6.3	74.0 \pm 3.9	59.6 \pm 7.6	66.3 \pm 6.6	53.1 \pm 8.5	65.9 \pm 6.6	59.8 \pm 9.2	69.2 \pm 6.5
Ankle RoM [°]	37.1 \pm 5.8	33.5 \pm 5.8	28.2 \pm 5.1	32.5 \pm 5.1	25.2 \pm 2	34.2 \pm 6.0	30.1 \pm 7.5	34.1 \pm 6.0

Values with * asterisk are normalized by height [m].

TABLE III
AVERAGE CLASSIFICATION ACCURACY OF THE DIFFERENT FEATURES
GROUPS CLUSTERED BY THE NUMBER OF SENSORS USED. IN EACH
CLUSTER, GROUPS ARE SORTED BY DECREASING VALUE

No. of sensors	Groups	Accuracy across classifiers [%]
3	Joint-specific Hip RoM (IV)	65.96 \pm 2.73
4	Joint-specific Knee RoM (III)* Joint-specific Ankle RoM (II)	76.31 \pm 3.04 65.98 \pm 5.70
5	Individual spatio-temporal Step Time (X)	68.23 \pm 2.77
7	RoMs variability (VIII)* RoMs asymmetry (VII)* Multiple joint RoMs (V) Spatio-temporal variability (XIV) Individual spatio-temporal Step Length (IX) Multiple spatio-temporal (XII) Individual spatio-temporal Speed (XI) Spatio-temporal asymmetry (XIII)	79.62 \pm 2.40 77.82 \pm 3.85 76.75 \pm 2.78 74.14 \pm 2.48 74.05 \pm 3.74 71.02 \pm 3.19 70.88 \pm 4.32 62.99 \pm 3.53
8	PCA on St-D (XVII)* Statistically different (XV)* Multiple joint RoMs with chest (VI) Overall (I) PCA on Overall (XVI)	79.96 \pm 4.55 77.38 \pm 4.25 76.46 \pm 2.60 75.15 \pm 4.55 64.27 \pm 2.33
8	Random (XVIII)	69.94 \pm 3.26

* The asterisk denotes groups with classification accuracy significantly higher than the random group.
The dotted line represents the rank position of the random group if it was in that specific sensors cluster.

group, taken as reference (XVIII). For the joint-specific groups, 4 sensors allowed to assemble the ankle RoM (II) and the knee RoM (III). The latter resulted significantly more accurate than the random group, while the ankle showed lower accuracies. Individual spatio-temporal Step Time (X), which needs 5 sensors, showed lower accuracy than XVIII. Eight groups needed 7 sensors: RoMs variability (VIII) showed the best accuracy and a significant increase with respect to reference, while spatio-temporal asymmetry (XIII) gave the worst classification performance, and significantly worse than XVIII. Also, RoMs asymmetry (VII)

showed a significantly higher accuracy than reference. The application of PCA on St-D (XVII) led to the use of 8 sensors and showed classification accuracy significantly higher than the reference group. This group resulted composed of the following 13 features: knee RoM SD, and ankle RoM μ and CoV, both sides; step length CoVs and speed μ , taken bilaterally; the ratio of the knee RoM CoV, and that of the ankle RoM μ ; hip RoM asymmetry CoV. PCA on Overall (XVI) returned, conversely, the worst accuracy among all in the eight-sensor cluster. In this cluster, PCA on St-D and the St-D (XV) groups showed significantly higher accuracies than that in the reference. Among the multiple joint RoMs groups, the one including the chest sensor (VI) showed a classification accuracy similar to that obtained without its use (V). Multiple joint RoMs revealed higher estimation accuracy than multiple spatio-temporal. Similar results were obtained comparing the mean values across all groups with joint RoMs (II-VIII) with those related to spatio-temporal parameters (IX-XIV): 74.1% to 70.2%, respectively.

Regarding the comparison between the basic classifiers, the mean accuracy values obtained with each classifier, averaged among all groups, were the following: SVM rbf gave the best average accuracy (75.6%), then k-NN (73.0%), NB (72.7%), LDA (72.5%), SVM linear (72.0%), and DT (68.9%). SVM rbf and k-NN were significantly more accurate than the others. SVM rbf accuracy resulted also significantly higher than NB. The Majority of votes approach produced an average classification accuracy up to 85.8% for type A, and to 83.9% for type B. The obtained classification performances (see last column of Table IV) were always higher than those of the basic classifiers, with an average increase in the range 5–20%. In particular, Majority of votes type A, applied on the feature group VI, achieved a 96% accuracy.

Table V presents individual accuracies obtained with each H&Y stage-specific subset, using SVM rbf. In the stage I, accuracies over 80% were obtained with groups I, V, VI, VII, IX, XV and XVII. In particular, the last two groups resulted in accuracies higher than 90%. Each single feature belonging to group XVII—the group with the highest accuracy value—was then used to train the SVM rbf classifier. The Ankle μ returned value of accuracy of 67.1%, the Ratio Hip SD 71.3%, the Stride Time μ 81.3%, the Asymmetry Hip μ 74.9%, the Asymmetry Step Length SD 75.4% and the Asymmetry Step Time μ 71.8%.

TABLE IV
ACCURACY FOR EACH GROUP AND FOR EACH CLASSIFIER [%]

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
LDA	75.26	67.22	77.34	64.28	76.20	75.94	77.36	77.34	73.70	65.72	73.34	72.48	59.86	75.82	77.82	64.78	79.62	71.10±4.32
NB	77.02	59.74	76.76	63.64	75.38	75.14	76.60	78.12	76.82	68.42	72.46	74.14	65.26	74.68	79.52	65.00	78.02	71.29±5.15
k-NN	80.14	69.60	78.20	69.04	78.14	78.80	79.94	80.26	73.44	68.52	73.10	73.06	62.04	75.74	79.20	67.24	84.62	73.01±3.50
SVM	72.30	72.18	70.70	63.66	78.48	77.60	80.98	78.82	72.72	64.30	70.52	70.40	58.90	73.16	76.76	61.00	81.14	68.46±7.01
SVM rbf	78.52	68.98	79.36	69.70	80.08	79.08	81.08	79.06	79.28	71.86	73.50	70.88	66.72	75.92	81.62	65.58	83.22	71.73±4.21
DT	67.68	58.14	75.52	65.46	72.22	72.20	70.94	84.10	68.34	69.54	62.34	65.14	66.06	69.54	69.38	62.02	73.12	64.02±6.23
Majority of votes A	92.00	80.00	84.00	72.00	94.00	96.00	90.00	86.00	86.00	76.00	78.00	84.00	88.00	82.00	92.00	88.00	90.00	84.00±6.53
Majority of votes B	90.00	79.00	84.00	73.00	92.00	91.00	90.00	85.00	83.00	75.00	78.00	83.00	83.00	81.00	87.00	85.00	88.00	82.00±5.01

TABLE V
ACCURACY FOR EACH PARAMETERS GROUP AND FOR EACH POPULATION GROUP WITH SVM RBF [%]

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII
H&Y I	89.25	69.38	74.38	69.75	80.06	84.00	84.31	69.25	85.31	77.94	65.13	78.69	62.88	65.88	93.88	66.94	94.50
H&Y II	63.25	70.88	83.69	50.50	69.13	66.31	81.44	80.00	62.00	71.50	56.25	68.63	47.07	76.88	86.00	64.13	87.75
H&Y III	92.38	75.31	95.63	93.19	97.44	96.88	73.56	93.75	82.63	83.19	80.38	84.44	67.13	93.63	94.94	94.25	93.63

For the pair healthy-PD stage II, accuracies higher than 80% were obtained with groups III, VII, VIII, and again XV and XVII. Stage III showed accuracies higher than 90% for most of the parameter groups; IX-XII were lower, while II, VII, XIII showed very low accuracies. Subjects belonging to stage II and III are older than in stage I (respectively, 64.7 ± 6.9 , 73.2 ± 11.5 , 74.6 ± 4.7).

IV. DISCUSSION

A. Type of Classifiers

No single classifier analyzed in this study provided a clear predominance in terms of classification performance among all feature groups.

This evidence is in line with the postulation that no algorithm can be uniformly accurate over all datasets, due to the cost function and sample data characteristics [36]. For this reason, we explored the possibility to leverage on the results from the individual classifiers, to construct two meta-classifiers based on the Majority of votes approach. These provided significantly higher performance, overcoming the weaknesses of each classifier, through the use of appropriate weights. In particular, between the two types, the weighting scheme based on the training accuracies (type A) provided the highest performance for all the feature groups.

The observed performance increase coming from the use of multiple classifiers uncovers an additional evidence: the misclassified samples obtained from the individual classifiers are not necessarily the same, and this leaves the door open to possible increases in accuracy if more intelligent classification schemes are present. To our knowledge, this is the first time

that the Majority of votes is introduced for a feature-based instrumented assessment of PD, and the obtained results seem to confirm the promising direction shown by this approach.

B. Sensors Location

The number and location of sensors has a direct effect on the parameters to be extracted. The joint-specific knee RoM group, obtained with only 4 sensors, showed an accuracy comparable with (and sometimes higher than) the Overall group, which required 8 sensors. This is particularly relevant for clinical applications, in which a setup composed of half of the sensors can considerably lower the costs and time of the clinical evaluation.

The features extracted from the chest sensor, when included in the RoMs group, did not significantly improve the classification performance (see accuracy for groups V and VI). This may be due to the combination of two factors: i) the low incremental influence of variations to the trunk movement characteristics during gait other than those already captured by the other sensors; ii) the noisy nature of the data coming from the chest sensor, probably due to its fastening, which varied across individuals given the subjects' anatomy and sex. Moreover, from the preliminary feature selection of group XV, chest features were not significantly different between the populations. Thus, even if the features associated with the chest sensor were used to preliminarily define the St-D features needed for the definition of group XV and, subsequently, group XVII (which showed the highest accuracy), the sensors effectively needed to classify with groups XV and XVII required just 7 sensors, thus excluding the chest.

The knee joint appeared to play a major role among the lower limb joints in the assessment of the PD gait, as results on group

III highlighted. This is in line with some studies in the literature, where significant differences in the knee kinematics between healthy and PD subjects were found [16]. Surprisingly, the same did not appear for the ankle, as its classification power (group II) resulted among the lowest ones, while a few studies in the literature provide ground for ankle sensor use [18], [27], [37]. We do not have a clear justification for that, even if we need to consider that, for the features extracted in this study, only ankle plantarflexion and dorsiflexion was considered, thus omitting possible contributions coming from using joint kinematics along the other two planes. Finally, for the hip, this does not play a significant role in the classification ability, confirming previous studies in the literature [8].

The sensor location analysis here provided is motivated by the potential of enabling gait assessment in home setting environment. Lab-based assessment is often affected by alteration of the walking behavior due to formal and non-familiar environment [38], and restricted to a limited duration of data recording. A minimal sensors setup assumes great relevance for home-based assessment, in terms of improved portability, comfort and usability. Meanwhile, these requirements affect the limited number of obtainable gait features, resulting in lower accuracy and reliability of the estimations. Finding the optimal balance between these requirements is one of the great research challenges in wearable systems for healthcare.

C. Type of Parameters and Features

RoM parameters appeared as key contributors for the classification accuracy. This was not unexpected, since joint excursion is directly affected from the presence of PD [39].

Regarding the spatio-temporal parameters, the classification accuracy of this group is led by the Step Length, whose variations as a consequence of the presence of PD are well reported in the literature [18], [19]. A reduced value of step length is characteristic of walking patterns of subjects with PD and thus, an evident relevance of this parameter was expected. Due to the variety of data in the two groups, its classification power should not be taken for granted. Temporal group parameters, on the other hand, determined poor discriminatory power. Not all researches agreed in the definition of this parameter as discriminating [17], [28].

The combination of spatio-temporal and RoMs parameters led to higher accuracies than each of the two groups taken separately. However, this is more marked if this combination is accompanied by a preliminary selection process, which takes into account only those features that are more sensitive to the presence of the disease symptoms: this is what is behind the feature selection process introduced in the Statistically Different group and in the PCA on St-D group. It is no surprise that, if this process is not included, the obtained results coming from the totality of features show limited improvement. The same also applies to the unsupervised feature selection process used in the PCA on Overall group.

The asymmetry and variability of RoMs groups (VII and VIII) provide very good results in terms of classification accuracy. Previous studies showed higher variability and reduced

left-right symmetry in the lower extremity [20] and linked this to a defective load control ability in subjects with PD [40]. At stage I of H&Y scale, the involvement is unilateral, resulting in an asymmetry between the affected and non-affected side. With the progression of the disease, there is a worsening in the symptoms, with bilateral involvement, even if the asymmetry remained. The results obtained with asymmetry and variability of RoMs may be linked with these manifestations, also provided that a non-negligible number of individuals with PD participating in this study was assigned to H&Y I. We do not have an explanation for the rather low accuracy obtained with asymmetry of spatio-temporal parameters. It may be speculated that asymmetry in geometry (i.e., RoMs) is not necessarily mirrored into asymmetry of the spatio-temporal gait characteristics [41], even if decreases in temporal harmony of gait have been linked to the manifestation of PD motor symptoms [42].

D. PD Staging of Patients

In the current study, the PD group included individuals assigned to the three stages of H&Y scale. This inclusion leads to an increase of the inter-subject variability in the entire PD group, and to a reduction in gait differences with healthy subjects, making the classification process harder. That could be the reason of the limited performance of the individual classifiers, when compared with the literature data. As indicated in Table VI, related works are very heterogeneous in terms of kind and level of population included. To our knowledge, none of the works using whole body IMUs considered the three H&Y stages altogether. Despite that, the results obtained by combining the classifiers made it possible to reach values of accuracy that are promising, also given the fact that more than 72% of misclassifications were associated with controls and H&Y stage I subjects.

The additional analyses performed on individual PD groups, relating to the specific staging, showed accuracy values close to those reported in the literature. We did not observe a marked increase of accuracy moving from stage I to stage III. As a matter of fact, a number of groups showed lower classification accuracy at PD stage I as compared with PD stage II. This does not come necessarily unexpected, since each subset was composed of age-matched healthy individuals whose ages, for stage II and III were pretty higher than those for stage I. This evidence might suggest that PD staging from I to II through H&Y does not necessarily reflect into major modifications of gait parameters, other than those associated with age increase. To support this hypothesis, when considering asymmetry (group VII), the classification accuracy decreases with increasing stage, confirming that the unilateral involvement present at stage I is less present in higher stages. When considering variability (groups VIII and XIV), instead, a uniform trend across stages was observed. This may be linked with the ability of variability measures to predict motor severity in PD [43]. Regarding the ability of each feature to classify between healthy and H&Y at stage I, we obtained satisfactory results for the mean value of the stride time, and for some parameters of asymmetry (which is in line with the definition of stage I), even if the overall

TABLE VI
SIZE AND CHARACTERISTICS OF POPULATION SAMPLES IN THE LITERATURE STUDIES, AND ACHIEVED ACCURACY VALUES

Studies	no. of PD patients, age (yrs)	PD stage	no. of healthy subjects, age (yrs)	Acc. (%)	no. & type of ML techniques	type of features	no. & location of sensors
[21] [22]	93, 66.3	stage 2-3 H&Y	73, no info on age	60.2-85.3 52.8-94.4	4, SVM linear 1, LDA	Sp-T	### ###
[23] [24]	12, no info on age	no info on stage	20, no info on age	68.8-90.6 62.5-100	4, NB 4, SVM rbf	Sp-T & RoMs	### ###
[25]	23, 68.5±6.1	stage 2 H&Y	26, age-matched	80.4-92.6	5, RF	Sp-T	###
[26]	40, 59.8±10.6	stage 1.16±0.24 H&Y	40, age-matched	72.0-76.0	1, SVM-rbf	Sp-T	###
[30]*	10, 65.1±9.8	UPDRS _{III} 19.6±6.7 Quest ₃₉ score 18.5±16.9	10, 57.7±14.3	Se. 98.5 Sp. 97.5	3, RF	Statistical	Not provided
[15]*	50, 63.9±10.6	stage 2.1±0.9 H&Y	42, 60.0±11.2	81-82	4, Adaboost	Statistical	2, each foot
[29]*	42, 65.1±9.7	stage 2.2±0.9 H&Y	39, 60.7±11.8				
	156, 67.2±8.0	median stage 1 H&Y UPDRS _{III} 17	424, 51.9±10.0	85.5±4.7	1, k-NN	Gait sequences	1, L4
[27]*	24, no info on age	11 "great disturbance", 13 "no such disturbance"	16, no info on age	Se. 93.3 Sp. 96.8	1, SVM rbf	RoMs	2, each foot
[28]*	28, 63.4±9.3	14 UPDRS _{III} 9.0±3.6	16, 64.9±6.9	Se. 88.0 Sp. 86.0	5, LDA	Sp-T & Statistical	2, each foot
This study*	25, 71.1±8.5	8 stage 1, 9 stage 2, 8 stage 3 H&Y	25, 66.7±9.8	Table III	6 + Majority of Votes classifier		
	8, 64.8±6.9	stage I	8, 65.4±4.7				
	8, 73.2±11.5	stage II	8, 69.9±12.8	Table IV	1, SVM rbf	Sp-T & RoMs	3-8, each foot, each shank, each thigh, waist, chest
	8, 74.6±4.7	stage III	8, 71.38±4.3				

* asterisk denotes studies where IMUs were used.

Acc. reports the accuracy range obtained from spatio-temporal (Sp-T) and ranges of motion (RoMs) data only.

Se. Sensitivity, Sp. Specificity.

no. & type of ML techniques reports the number of ML techniques used and the name of that one with the best accuracy, where RF refers to Random Forest, LDA to Linear Discriminant Analysis, SVM to Support Vector Machine, NB to Naïve Bayes and k-NN to k-Nearest Neighbor).

accuracy of these features, if taken individually, is outperformed by the feature group classifier, thus reinforcing the hypothesis that taking into account multiple features from gait is necessary for better classification ability.

E. Study Limitations

In this study, all the feature groups were defined based on summary parameters extracted from the angle profiles, thus inherently compressing information coming from raw data. While the feature set was very ample, excluding raw data from the feature set may hide possible effects at the profile level.

Moreover, the information associated with upper limb movements was not analyzed. Future studies might include sensors on this location, both on right and left sides, probably providing features which improve the classification performance, especially for patients at the stage I of the H&Y scale with unilateral involvement. The same may appear as the consequence of including motor tasks other than gait.

Referring to the stage-specific classification, we globally confirmed the results obtained in literature, even if some of the gait parameters groups led to lower accuracies than those expected: this might be related to the relatively small sample size for these individual subsets (16 subjects in total).

With a higher sample size, possible confounding factors coming from age discrepancies across different stages could be minimized.

Regarding the sensor placement and location, the positioning of the sensor at the chest level resulted subject-dependent, i.e., associated with anatomy and sex of the subject. This may have led to possible confounding effects in terms of classification accuracy. Even if this could be seen as a limitation of the study, at the same time it reflected the clinical environment conditions. With this study protocol, we intended to evaluate the classification power of IMUs in an ambulatory environment, where no additional requirements in terms of patient set up were considered other than those directly associated with IMU use, in order to verify the effectiveness of such a solution in an uncontrolled ambulatory setting.

V. CONCLUSION

The ability of IMU-based gait analysis to discriminate patients with PD at different severity stages from age-matched healthy individuals has been shown in this study to relevantly depend on the number and location of sensors used to extract the parameters, and only to a lesser extent from the kind of ML technique used. The increase of sensor number in fact does not

determine a direct increase in classification accuracy, unless this is accompanied by a thorough data-driven feature selection process from the variety of gait parameters that can be extracted. Regarding the choice of the ML technique to be used to this end, we were able to confirm that there is no unique solution leading to an overall optimum in accuracy; rather, we demonstrated that, by combining results coming from different ML techniques, a relevant increase in accuracy can be obtained, also in the case of samples including individuals with rather different staging of PD.

While the field of Parkinson's disease encompasses a variety of motor and non-motor symptoms that cannot be directly captured by the sole use of sensors while walking, with the results obtained in this research it will be possible for researchers of the relevant community to select which features, and which sensor placements could be more profitably used to define those gait-related manifestations that are associated with the severity of the pathology.

The availability of wearable sensor networks that are able to capture movement data in unconstrained scenarios has represented a big step towards the decentralization of the monitoring process in a variety of clinical applications. With the integration of computational and classification capabilities, such as those analyzed in this study, these devices will move the clinical process closer to the decentralization of diagnostic decisions as well, including those that, with specific reference to PD, are directly related to the adjustments of drug dosage for the treatment of motor symptoms.

REFERENCES

- [1] J. Jankovic, "Parkinson's disease: Clinical features and diagnosis," *J. Neurol. Neurosurg. Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008.
- [2] R. Bhidayasiri and P. Martinez-Martin, "Clinical assessments in Parkinson's disease: Scales and monitoring," *Int. Rev. Neurobiol.*, vol. 132, pp. 129–182, 2017.
- [3] E. R. Dorsey, B. P. George, B. Leff, and A. W. Willis, "The coming crisis: Obtaining care for the growing burden of neurodegenerative conditions," *Neurology*, vol. 80, pp. 1989–1996, 2013.
- [4] M. E. Micò Amigo *et al.*, "Is the assessment of 5 meters of gait with a single body-fixed-sensor enough to recognize idiopathic Parkinson's disease-associated gait?," *Ann. Biomed. Eng.*, vol. 45, pp. 1266–1278, 2017.
- [5] S. Fahn and R. L. Elton, "Unified Parkinson's disease rating scale," in *Recent Developments in Parkinson's Disease*, vol. 2, S. Fahn, C. D. Marsden, D. Calne, and M. Goldstein, Eds. Florham Park, NJ, USA: McMillan Health Care, 1987, pp. 153–163.
- [6] M. M. Hoehn and M. D. Yahr, "Parkinsonism onset, progression, and mortality," *Neurology*, vol. 17, no. 5, pp. 427–427, 1967.
- [7] A. J. Espay *et al.*, "Technology in Parkinson's disease: Challenges and opportunities," *Movement Disorders*, vol. 31, no. 9, pp. 1272–1282, 2016.
- [8] O. Sofuwa, A. Nieuwboer, K. Desloovere, A. M. Willems, F. Chavret, and I. Jonkers, "Quantitative gait analysis in Parkinson's disease: Comparison with healthy control group," *Arch. Phys. Med. Rehabil.*, vol. 86, no. 5, pp. 1007–1013, 2005.
- [9] C. F. Pasluosta, H. Gassner, J. Winkler, J. Klucken, and B. M. Eskofier, "An emerging era in the management of Parkinson's disease: Wearable technologies and the internet of things," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 6, pp. 1873–1881, Nov. 2015.
- [10] J. Baziw and C. T. Leondes, "In-flight alignment and calibration of inertial measurement units-part I: General formulation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-4, no. 4, pp. 439–449, Jul. 1972.
- [11] S. J. Merhav, "A nongyroscopic inertial measurement unit," *J. Guid. Control Dyn.*, vol. 5, no. 3, pp. 227–235, 1982.
- [12] F. Parisi *et al.*, "Body-sensor-network-based kinematic characterization and comparative outlook of UPDRS scoring in leg agility, sit-to-stand, and gait tasks in Parkinson's disease," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 6, pp. 1777–1793, Nov. 2015.
- [13] S. Patel *et al.*, "Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 864–873, Nov. 2009.
- [14] M. Demonceau *et al.*, "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.*, vol. 19, no. 6, pp. 1803–1808, Nov. 2015.
- [15] J. Klucken *et al.*, "Unbiased and mobile gait analysis detects motor impairment in Parkinson's disease," *PLoS One*, vol. 8, no. 2, pp. 1–9, 2013.
- [16] A. Salarian *et al.*, "Gait assessment in Parkinson's disease: Toward an ambulatory system for long-term monitoring," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 8, pp. 1434–1443, Aug. 2004.
- [17] C. Curtze, J. G. Nutt, P. Carlson-Kuhta, M. Mancini, and F. B. Horak, "Levodopa Is a double-edged sword for balance and gait in people with Parkinson's disease," *Movement Disorders*, vol. 30, no. 10, pp. 1361–1370, 2015.
- [18] H. C. Chang, Y. L. Hsu, S. C. Yang, J. C. Lin, and Z. H. Wu, "A wearable inertial measurement system with complementary filter for gait analysis of patients with stroke or parkinson's disease," *IEEE Access*, vol. 4, pp. 8442–8453, 2016.
- [19] S. Del Din, A. Godfrey, B. Galna, S. Lord, and L. Rochester, "Free-living gait characteristics in ageing and Parkinson's disease: Impact of environment and ambulatory bout length," *J. Neuroeng. Rehabil.*, vol. 13, no. 46, pp. 1–12, 2016.
- [20] N. Toosizadeh, J. Mohler, H. Lei, S. Parvaneh, S. Sherman, and B. Najafi, "Motor performance assessment in Parkinson's disease: Association between objective in-clinic, objective in-home, and subjective/semi-objective measures," *PLoS One*, vol. 10, no. 4, pp. 1–15, 2015.
- [21] D. Chang, M. Alban-Hidalgo, and K. Hsu, "Diagnosing Parkinson's disease from gait," stanford.edu, Dec. 31, 2014. [Online]. Available: <http://cs229.stanford.edu/proj2014/Daryl%20Chang,%20Marco%20Alban-Hidalgo,%20Kevin%20Hsu,%20Diagnosing%20Parkinson's%20from%20Gait.pdf>. Accessed on: Aug. 22, 2018.
- [22] S. V. Perumal and R. Sankar, "Gait monitoring system for patients with Parkinson's disease using wearable sensors," presented at the *IEEE Healthcare Innovation Point-Of-Care Technologies Conf.*, Cancun, Mexico, 2016.
- [23] H. H. Manap, N. Md. Tahir, and R. Abdullah, "Anomalous gait detection using Naïve Bayes classifier," presented at the *IEEE Symp. on Industrial Electronics and Applications*, Bandung, Indonesia, 2012.
- [24] N. M. Tahir and H. H. Manap, "Parkinson disease gait classification based on machine learning approach," *J. Appl. Sci.*, vol. 12, no. 2, pp. 180–185, 2012.
- [25] F. Wahid, R. K. Begg, C. J. Hass, S. Halgamuge, and D. C. Ackland, "Classification of Parkinson's disease gait using spatial-temporal gait features," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 6, pp. 1794–1802, Nov. 2015.
- [26] M. Djurić-Jovičić, M. Belić, I. Stanković, S. Radovanović, and V. S. Kostić, "Selection of gait parameters for differential diagnostics of patients with de novo Parkinson's disease," *Neurol. Res.*, vol. 39, no. 10, pp. 853–861, 2017.
- [27] I. Tien, S. D. Glaser, and M. J. Aminoff, "Characterization of gait abnormalities in Parkinson's disease using a wireless inertial sensor system," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2010, pp. 3353–3356.
- [28] J. Barth *et al.*, "Biometric and mobile gait analysis for early diagnosis and therapy monitoring in Parkinson's disease," in *Proc. 33rd Annu. Int. Conf. IEEE Eng. Med. Biol. Sci.*, 2011, vol. 2011, pp. 868–871.
- [29] F. Cuzzolin *et al.*, "Metric learning for Parkinsonian identification from IMU gait measurements," *Gait Posture*, vol. 54, pp. 127–132, 2017.
- [30] S. Arora, V. Venkataraman, S. Donohue, K. M. Biglan, E. R. Dorsey, and M. A. Little, "High accuracy discrimination of Parkinson's disease participants from healthy controls using smartphones," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2014, pp. 3641–3644.
- [31] L. Rocchi, L. Chiari, and A. Cappello, "Feature selection of stabilometric parameters based on principal component analysis," *Med. Biol. Eng. Comput.*, vol. 42, no. 1, pp. 71–79, 2003.
- [32] W. H. Wu, A. Bui, M. A. Batalin, D. Liu, and W. J. Kaiser, "Incremental diagnosis method for intelligent wearable sensor systems winston," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 5, pp. 553–562, Sep. 2007.

- [33] J. S. Wang, C. W. Lin, Y. T. C. Yang, and Y. J. Ho, "Walking pattern classification and walking distance estimation algorithms using gait phase information," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2884–2892, Oct. 2012.
- [34] J. Kamruzzaman and R. K. Begg, "Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2479–2490, Dec. 2006.
- [35] G. James, D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning, With Application in R*. New York, NY, USA: Springer, 2013.
- [36] D. H. Wolpert, "The lack of a priori distinctions between learning algorithms," *Neural Comput.*, vol. 8, no. 7, pp. 1341–1390, 1996.
- [37] M. Svehlk *et al.*, "Gait analysis in patients with Parkinson's disease off dopaminergic therapy," *Arch. Phys. Med. Rehabil.*, vol. 90, no. 11, pp. 1880–1886, 2009.
- [38] C. Zampieri, A. Salarian, P. Carlson-Kuhta, J. G. Nutt, and F. B. Horak, "Assessing mobility at home in people with early Parkinson's disease using an instrumented Timed Up and Go test," *Parkinsonism Related Disorders*, vol. 17, no. 4, pp. 277–280, 2011.
- [39] M. E. Morris, F. Huxham, J. McGinley, K. Dodd, and R. Iansek, "The biomechanics and motor control of gait in Parkinson disease," *Clin. Biomech.*, vol. 16, no. 6, pp. 459–470, 2001.
- [40] V. Dietz, "Neurophysiology of gait disorders: Present and future applications," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 3, pp. 333–355, 1997.
- [41] G. Vervoort, A. Bengevoord, E. Nackaerts, E. Heremans, W. Vandenberghe, and A. Nieuwboer, "Distal motor deficit contributions to postural instability and gait disorder in Parkinson's disease," *Behav. Brain Res.*, vol. 287, pp. 1–7, 2015.
- [42] M. Iosa *et al.*, "Loss of fractal gait harmony in Parkinson's disease," *Clin. Neurophysiol.*, vol. 127, no. 2, pp. 140–146, 2016.
- [43] S. Lord, K. Baker, A. Nieuwboer, D. Burn, and L. Rochester, "Gait variability in Parkinson's disease: An indicator of nondopaminergic contributors to gait dysfunction?," *J. Neurol.*, vol. 258, pp. 566–572, 2011.