The 7th IEEE International Conference on E-Health and Bioengineering - EHB 2019
Grigore T. Popa University of Medicine and Pharmacy, Iasi, Romania, November 21-23, 2019

Parkinson's Disease Detection from Gait Patterns

Alexandra-Georgiana Andrei[1], Alexandra-Maria Tăuțan[1], Bogdan Ionescu[1],
Research Center CAMPUS, University Politehnica of Bucharest,
Bucharest, Romania, alexandra.tautan@upb.ro

***Abstract*—Parkinson's disease (PD) patients display abnormal gait patterns with impairments and postural instability. In this paper, we propose an automatic system for extracting gait parameters. Various features were extracted from force sensors and analyzed using a threshold-based algorithm and machine learning techniques with the objective to identify the most significant features that would best characterize the presence of the disease. A machine learning algorithm using support vector machine method was developed to identify the presence of the disease. The analyses of the results show that the machine learning algorithm has the best accuracy of 100% in distinguishing between the two groups when looking at features based on stride, swing and stance phases.**

Keywords—Parkinson's disease detection, gait analysis, machine learning.

I. Introduction

Parkinson's disease (PD) is a slowly progressive neurodegenerative disorder that mainly affects people over the age of 60. It is ranked the second most common neurodegenerative disease next to Alzheimer's disease [1]. The main motor symptoms are caused by the deterioration of dopamine-producing neurons. Typically, a patient with PD presents the following symptoms: tremor and bradykinesia (slowness in movement) at an early stage and gait disorders and postural instability with the worsening of the disease.

It is important to determine the tremor and gait disorders because the order in which the symptoms appear differs from one patient to another. Diagnosis of PD can be difficult, especially in the early stages. There is currently no special test or biomarker available for diagnosis. Abnormal gait characteristics that describe PD and were considered in this study are:

(i) *Heel to toe characteristics* – while in normal gait the heel strikes the ground before the toes (heel to toe walking), in PD gait stepping is characterized by flat foot strikes (the entire foot is placed on the ground at the same time) or heel walking (toes touch the ground before the heel) in more advanced stages of the disease [2]. Furthermore, patients with PD have reduced impact at heel strikes. Therefore, the forces exerted on the ground are lower than the forces exerted by a healthy subject;

(ii) *Vertical ground reaction force* (VGFR) – the vertical ground reaction force signal has two peaks in normal gait – the first one when the foot strikes the ground and the second peak corresponding to the push-off force from the ground. PD gait is characterized by reduced forces for heel contact and the push-off phase. The VGRF signal has peaks with reduced height [2];

(iii) *Bradykinesia* – describes the slowness of movement. Patients with PD present a lower movement speed than normal subjects and their gait is characterized by smaller steps [3].

In this paper, we aim to automatically detect PD patients from healthy subjects by using automatic machine learning techniques and analyzing heel to toe characteristics, vertical ground reaction forces and bradykinesia from recorded force signals characterizing gait. The paper is organized as follows. Section II presents the overview of the literature and positions our contribution. Section III describes the proposed approaches. Section IV deals with the analysis of the data and discusses the experimental results. Finally, Section V concludes the paper and discusses future work.

II. Previous Work

Over time, many research studies have developed different systems for PD monitoring using different types of sensors, feature sets and methods of analysis. A large number of techniques have been developed for the automatic detection of PD using different methods based on machine learning algorithms such as neural networks or support-vector machines (SVM). Many techniques present algorithms for a single symptom detection. Given the heterogeneous nature of PD symptoms, this is in most cases not sufficient; few studies focus on detecting multiple motor symptoms.

For instance, Barth et al. [4] studied and compared the gait pattern of normal subjects and PD patients. They used different types of classifiers. Among the classifiers used, LDA (Linear Discriminant Analysis) provided the best classification with a sensitivity of 88% and a specificity of 86%. In the experiments conducted by Salarian et al. [5] the results showed that the stride length, stride velocity and swing time of Parkinsonian patients were lower than those of healthy subjects. Also, they discovered that the stance time in Parkinson's patients was higher in comparison to healthy subjects. Using a force sensor worn by the subjects, Okuno et al. [6] obtained similar results. Further, Tahir and Manap [7] extracted kinetic and kinematic features based on force measurement. Features include the step length, stride time, walking speed, vertical ground reaction forces at heel and toe contact, and the kinematic features as the

978-1-7281-2603-6/19/$31.00 ©2019 IEEE

angle of ankle, knee and hip at heel strike and toe off position. Statistical analysis showed that step length, walking speed and VGRF were among the significant features that would differentiate a PD patient from normal subject. Frenkel-Toledo et al. [8] studied and compared the walking speed and gait variability between PD and normal subjects. The result was that the PD patients had an increased variability of stride time and swing time as compared to normal subjects.

The goal of this paper is to analyze the features extracted from gait measurements which would lead to the detection of PD. When combining multiple features extracted from the gait signals, we obtain better results than those presented in literature.

III. PROPOSED APPROACHES

A. Overview

The input data consists of vertical ground reaction force records of subjects as they walked at their usual, self-selected pace for approximately 2 minutes on level ground. Data processing was carried out in several stages as presented in the diagram from Figure 1. The pre-processing stage consisted of filtering data using a highpass filter. Features were extracted from vertical ground reaction force records. Classification was made in two classes – PD patients and healthy subjects. The last stage consists of the validation of the classification model.

A gait cycle consists of two phases: swing and stance phases. The stance phase represents 60% of the gait cycle and it begins with the initial contact - the heel strike and ends in the toe-off event. The swing phase represents 40% of a gait cycle and it contains the moments when the foot is not in contact with the ground. In this paper, the swing and stance phases were identified. They represent parameters for PD detection. In the following we detail each of the processing steps.

B. Extracting PD characteristics from Gait

1) Pre-processing. In the pre-processing phase, the data, which includes all the three experiments, was filtered using a Chebyshev type II high-pass filter with a cut-off frequency of 0.8 Hz to remove the noise obtained from the changes in orientation of the subject's body and other types of low frequency noise.

2) Feature Extraction. The filtered signal was used to extract a series of gait parameters using peak detection methods and duration detection algorithms. From the peak detection algorithm, various kinetic parameters were extracted. Duration detection algorithms have been developed for extracting temporal and spatial parameters. An overview of the extracted features is available in Table I.

For the determination of several parameters, only the left foot was taken into consideration. An example of VGRF signal is presented in Figure 2. To eliminate the effect of gait initiation, first seconds of VGFR were discarded. In Figure 3, VGFR are plotted against time for the left foot. Points P1-P4 are plotted to mark a gait: P1, P4 – heel strike, P2 – maximal weight acceptance, P3 – toe off. The period between P1 and P4 represents the stride time (double step). Additionally, the time between P1 and P3 represents the duration of the stance phase and the time between P3 and P4 represents the duration of swing phase.

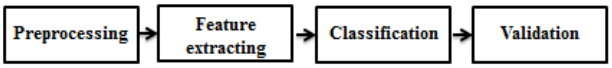

Figure 1. Data processing stages.

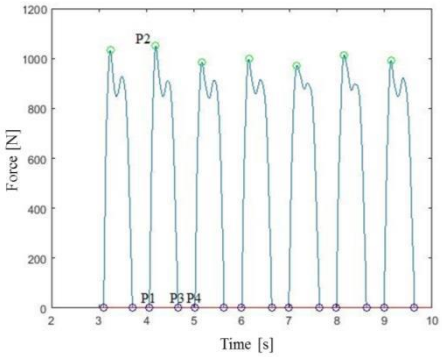

Figure 2. Force reading for a patient with PD. Points P1-P4 mark a gait cycle: P1, P4 – heel strike, P2 – maximal weight acceptance, P3 – toe off.

TABLE I. List of features extracted from the vertical ground reaction force data.

Features	Description
Step length (m)	The distance measured from the heel print of one foot to the heel print of the other foot.
Step time (sec)	The time needed to take a step.
Stride length (m)	The distance between two successive placements of the same foot.
Stride time (s)	Time needed to complete a gait cycle (stride).
Swing phase (%)	The phase during which the foot is not in contact with the ground.
Stance phase (%)	The phase during which the foot remains in contact with the ground.
Cadence (steps/min)	The rate at which a person walks.
Heel force (N)	The force printed on the ground during the heel strike.
Maximal weight acceptance force (N)	The force printed on the ground during the period during early stance phase at which the knee is fully extended and accepting the full weight.
Mid-stance force (N)	The force printed on the ground the time in which the entire foot is in contact with it.
Push off force (N)	The force printed on the ground during the terminal stance phase.
Toe force (N)	The force printed on the ground at the beginning of swing phase.

3) Ground reaction force. By processing the force readings using peak detection and sequence detection methods, all the forces corresponding to each phase of the foot in contact with the ground were extracted. These included heel contact, maximal weight acceptance, mid stance, push off and tow off as depicted in Figure 3. These are further described in Table II. The forces were normalized by the weight of each patient.

C. Classification

1) Threshold-based PD detection. An initial threshold based algorithm was developed for PD detection. Threshold values extracted from literature were applied to the computed features [13] [14]. The threshold values are detailed in Table III.

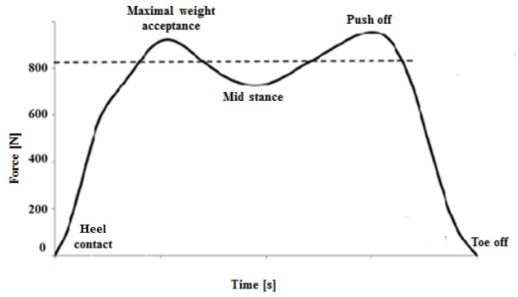

Figure 3. The vertical ground reaction force acting on a healthy patient during the gait cycle. Representation of all the forces corresponding to each phase of the foot in contact with the ground.

TABLE II. Detection methods for all the forces corresponding to each phase of the foot in contact with the ground

Force	Method to identify it
Heel force	Is identified at the beginning of the support phase and represents the start of the heel attack. Its value is found by identifying the force that succeeds the balance phase of the same foot in which the vertical force acting on the ground is 0 (swing phase)
Maximal weight acceptance force	Peak detection method.
Mid-stance force	Peak detection method.
Push off force	Peak detection method.
Toe force	Represents the beginning of the balance phase and is identified by the value preceding the balance phase of the same foot in which the vertical force acting on the ground is zero.

TABLE III. Employed threshold values

Feature	Literature value
Swing phase [%]	40% from a gait cycle
Velocity	1.4 m/s
Stride time	1.09 s
Cadence	90-110 steps/min
Stride distance	150 cm

2) Support Vector Machine (SVM). A support vector machine is a type of machine learning algorithm that performs classification or regression tasks for different data groups. This technique creates a hyperplane that separates data from two different classes.

In order to differentiate normal gait patterns from PD gait patterns, we experimented with different SVM parameters. Both a Linear and a Gaussian kernel were used. A Linear kernel is recommended when the data is linearly separable. The Gaussian kernel is an example of a basic radial functional kernel. The SVM classifier with the Gaussian kernel is simply a weighted linear combination of the kernel function computed between a data point and each of the support vectors.

IV. EXPERIMENTAL RESULTS

A. Database

The dataset used in this study was downloaded from Physionet [9], "Gait in Parkinson's Disease" database [10]. It consists of measures of gait from 93 patients with PD and 73 healthy subjects. The database includes the vertical ground reaction force records of subjects as they walked at their usual, self-selected pace for approximately 2 minutes on level ground. Underneath each foot, there were 8 sensors measuring force. The output of each of these 16 sensors has been digitized and recorded at a sampling rate of 100 Hz. The records also included two signals that reflect the sum of the 8 sensor outputs for each foot [9]. The database comprises 3 different experiments called "Si", "Ju" and "Ga" [8][11][12]. These are summarized in Table IV.

TABLE IV. Database

Database	Healthy Subjects	PD Patients
Si	17	29
Ju	25	29
Ga	18	28
Mean age	66.3 years	66.3 years
Gender	55 % male	63% male

During the experiment, the patients followed the medication prescribed by the doctors. According to the experiment protocol, the patients moved at their normal walking pace for two minutes. Additional information on every patient's speed, height and weight is also included in the dataset.

B. Validation

A k-fold cross-validation method was used, that partitions the data into three and five sets or folds. For each set it trains a model and assesses its performance. Multiple combinations of features, kernel types and number of folds were tested for obtaining the most accurate results.

C. Statistical Analysis

One-way analysis of variance (ANOVA) test was used to determine if there are any significant differences between the mean values of the two groups, healthy subjects and PD patients.

In Figure 4, there is an example of an ANOVA analysis of the step distance for the two groups of subjects. The minimum value is 0.46 m for the healthy subjects and 0.3 m for the PD patients, the mean value is 0.66m for the healthy subjects and 0.54m for the PD patients and the maximum value is 0.8m for the healthy subjects and 0.65m for the PD patients. Significant differences between the mean values of the two classes was observed by obtaining p-values smaller than 0.05. The other parameters were similarly analyzed and all yielded significant differences between the two groups.

D. Classification Results

Using the SVM classifier, high accuracy values were obtained for all three groups considering different features. The results of the classification and the input features are described in Table V. The best result in terms of accuracy is obtained on the "Ju" and "Si" study group using a Gaussian kernel for classification. However, the validation method and the input features were different.

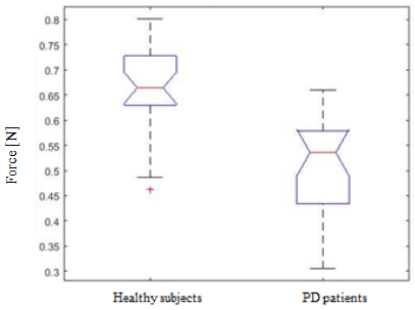

Figure 4. ANOVA representation for step length that provides information about minimum, mean and maximum values of the step length for the two classes of patients in the study group "Ju".

TABLE V. Best accuracy rates obtained for the three test groups.

Group of study	Accuracy rate	Features	Algorithm characteristics
"Ju"	100%	Stride length, stride time.	Gaussian kernel, three folds.
	90.9%	Swing and stance phases, stride length, step length.	Gaussian kernel, five folds.
	88.88%	Stride length, stride time, step length.	Linear kernel, five folds.
"Ga"	88.88%	Swing and stance phases, stride time and length, step time and length, cadence, heel force.	Gaussian kernel, five folds.
"Si"	100%	Swing and stance phases, stride time and length, step time and length, cadence, heel force, maximal weight acceptance.	Gaussian kernel, five folds

TABLE VI. Comparison between the accuracies obtained in this study and the results from other studies.

Group of study	This study: SVM Gaussian kernel	Reference method: Threshold-based	Shyam et. al [15]	Alam et. al [16]
"Ju"	100%	57.4%	92.5%	-
"Ga"	88.88%	73.3%	92.5%	83.1%
"Si"	100%	57.4%	90%	-

The results of the classification obtained with a Gaussian Kernel SVM are compared with the reference threshold method and some of the results available from literature. The comparison is presented in Table VI. The machine learning algorithm significantly outperforms the threshold-based reference method. Higher accuracies are obtained than those found in literature.

V. Conclusions

In this study, a machine learning approach and a threshold-based algorithm were investigated on a database of gait data of healthy subjects and PD patients. By using VGRF, this paper proposed a set of features which can successfully differentiate pathological from healthy gait. The most accurate classifier was found using a Gaussian Kernel SVM, in a three-fold validation for "Ju" group and five-folds for "Si" and "Ga". The best classification was obtained from features based on stride, swing and step phases. The threshold-based PD detection does not show a good performance as the values extracted from the literature vary over a large range. Our results show that these thresholds cannot be used as a safe threshold for disease detection. Taking into consideration the performances of other studies, the machine learning algorithm used in this study showed better results.

For future work, the goal is to create a real-time algorithm based on clinical data. Other classification methods can be used (LDA, neural networks) in order to obtain a higher accuracy.

VI. References

[1] P. Foundation, "Parkinson's Disease Foundation," 11 June 2019. [Online]. Available: https://www.parkinson.org/.

[2] J. R. Hughes, S. Bowes, A. Leeman, C. O'Neill, A. Deshmukh, P. Nicholson, S. Dobbs and R. Dobbs, "Parkinsonian abnormality of foot strike: a phenomenon of ageing and/or one responsive to levodopa therapy?".

[3] M. Morris, R. Iansek, T. Matyas and J. Summers, "Stride length regulation in Parkinson's disease. Normalization strategies and underlying mechanisms", *Brain*, p. 119, 1996.

[4] J. Barth, J. Klucken, P. Kugler, T. Kammerer, R. Steidl, J. Winkler, J. Hornegger and B. Eskofier, "Biometric and mobile gait analysis for early diagnosis and therapy monitoring in Parkinson's disease," in *Conf. Proc IEEE Eng Med Biol Soc*, Boston, MA, USA, 2011.

[5] A. Salarian, H. Russmann, F. Vingerhoets, C. Dehollain, Y. Blanc, P. Burkhard and K. Aminian, "Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 8, pp. 1434-43, 2004.

[6] R. Okuno, S. Fujimoto, J. Akazawa, M. Yokoe, S. Sakoda and K. Akazawa, "Analysis of spatial temporal plantar pressure pattern during gait in Parkinson's disease," in *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vancouver, BC, Canada, 2008.

[7] N. Tahir and H. H. Manap, "Parkinson Disease Gait Classification based on Machine Learning Approach," *Journal of Applied Sciences*, vol. 12, no. 2, pp. 180-185, 2012.

[8] S. Frenkel-Toledo, N. Giladi, C. Peretz, T. Herman, L. Gruendlinger and J. M. Hausdorff, "Effect of gait speed on gait rhytmicity in Parkinson's disease: variability of stride time and swing time respond differently," *Journal of Neuroengineering Rehabilitation*, vol. 2, 2005.

[9] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng and H. Stanley, "PhysioBank, PhysioToolkit and PhysioNet: components of a new research resource for complex physiological signals.," *Circulation*, vol. 101, no. 23, pp. 215-20, 2000.

[10] J. Hausdorff, J. Lowenthal, T. Herman, L. Gruendlinger, C. Peretz and N. Giladi, "Rhythmic auditory stimulation modulates gait variability in Parkinson's disease," *European Journal of Neuroscience*, vol. 26, no. 8, 2007.

[11] G. Yogev, H. Giladi, C. Peretz, S. Springer, E. Simon and J. Hausdorff, "Dual tasking, gait rhytmicity, and Parkinson's disease: which aspects of gait are attention demanding?," *European Journal of Neuroscience*, vol. 22, no. 5, 2005.

[12] D. B. Victor, "Evaluarea mersului. Teste," *Universitatea Valahia, Targoviste*, 2013.

[13] J. Hollman, E. McDade and R. Petersen, "Normative spatiotemporal gait parameters in older adults," *Gait Posture*, vol. 34, no. 1, pp. 111-8, 2011.

[14] S. V. Perumal and R. Sankar, "Gait and tremor assessment for patients with Parkinson's disease using wearable sensors," *ICT Express*, vol. 2, no. 4, pp. 168-174, 2016.

[15] M. Alam, A. Garg, T. Munia, R. Fazel-Rezai and K. Tavakolian, "Vertical ground reaction force marker for Parkinson's disease," *PLoS One*, 2017.