Parkinson's Disease Classification Using Inertial Signals from Smartwatches and Tree-Based Classifiers

Pedro L. S. Lobo[1], Ana Luísa de Bastos Chagas[1], Giordana de Farias F. B. Bucci[1], Rogerio Salvini[1], Fabrizzio Soares[1], Juliana Felix[1,2]

[1]Instituto de Informática, Universidade Federal de Goiás
[2]Escola Politécnica e de Artes, Pontifícia Universidade Católica de Goiás

{pedro_lemes, analuisa23, giordanabucci}@discente.ufg.br
{rogeriosalvini, fabrizzio, julianafelix}@ufg.br

Abstract. *This study evaluates the use of statistical features extracted from inertial signals obtained from smartwatches to assist in distinguishing patients with Parkinson's disease, differential diagnoses, and healthy controls. Using a public dataset comprising data from 469 individuals, tree-based classifiers were applied. The results indicate that simple motion information can support the diagnosis of Parkinson's disease in a non-invasive and accessible manner.*

1. Introduction

Parkinson's disease (PD) is a chronic, progressive neurodegenerative disorder that primarily affects the motor system [Bloem et al., 2021]. It impacts approximately 2–3% of individuals over the age of 65 [Poewe et al., 2017] and is characterized by the gradual loss of dopaminergic neurons in specific brain regions [Zhou et al., 2023]. This neuronal degeneration reduces dopamine synthesis, causing symptoms such as tremors, muscle rigidity, bradykinesia, postural instability, and impaired motor control [Braak and Braak, 2000]. As traditional clinical diagnosis can be expensive and time-consuming, there is increasing interest in developing non-invasive, alternative and low cost approaches to enhance diagnostic precision and improve patient care.

Since motor symptoms appear early [Maetzler and Hausdorff, 2012], several studies have explored the application of artificial intelligence and motion signal analysis to assist in PD diagnosis. Among these approaches, one study employed an IMU-based system to quantify total body tremor, demonstrating a significant correlation with clinical UPDRS (Unified Parkinson's Disease Rating Scale) scores [Delrobaei et al., 2017]. In a recent study, a protocol was conducted using two synchronized smartwatches and a smartphone application to collect inertial signals during motor tasks designed by neurologists, achieving a balanced accuracy of 78.99% in distinguishing PD patients from healthy controls, and 69.18% between PD and other movement disorders [Varghese et al., 2024]. These results demonstrate the potential of wearable devices for detecting motor patterns associated with PD and reinforce the feasibility of non-invasive machine learning-based methods.

This study extends previous research by evaluating statistical features derived from inertial smartwatch signals using tree-based classification methods. It also explores a differential diagnosis versus control scenario, an aspect not previously examined in this dataset, to assess the ability of these features to distinguish among Parkinson's disease

(PD), differential diagnosis (DD), and healthy control (CO) groups. This work is organized as follows. Section 2 describes the methodology, Section 3 presents and discusses the results, and Section 4 provides the conclusions.

2. Materials and Methods

2.1. Dataset

The public dataset *PADS - Parkinson's Disease Smartwatch Dataset*[1] was used, comprising 469 participants, distributed among PD (276), DD (114), and CO (79). The dataset was collected at the University of Münster (Germany) between 2018 and 2021, involving clinical questionnaires and 11 motor tasks monitored by two *Apple Watch Series 4* smartwatches (100 Hz). In this work, only smartwatch data were used.

2.2. Feature Extraction and Selection

Features from the motion signals were extracted from the inertial sensor data (accelerometers and gyroscopes) collected from the left and right wrists during different motor tasks. Each motion segment was represented by the three Cartesian axes X, Y, Z and the vector magnitude M, calculated as the Euclidean norm of these axes. For each axis, five features were extracted: the mean, standard deviation, variance, kurtosis, and skewness. Additionally, the mean Euclidean distance between consecutive samples of the segment was calculated from the magnitude, as shown in Eq. 1.

$$\text{Mean distance: } d_{\text{mean}} = \frac{1}{N-1} \sum_{i=1}^{N-1} \sqrt{(X_{i+1} - X_i)^2 + (Y_{i+1} - Y_i)^2 + (Z_{i+1} - Z_i)^2} \quad (1)$$

Feature extraction considered all combinations of tasks, sensors, wrists, and axes available in each experiment. Each participant performed 11 motor tasks using two wrists and two sensors, with further details about the especific tasks provided in [Varghese et al., 2024]. The complete feature extraction resulted in a feature vector composed of 924 attributes as detailed in Eq. 2,

$$\underbrace{11 \times 2 \times 2 \times 3 \times 5}_{\text{features extracted from the 3 axes (X, Y, Z)}} + \underbrace{11 \times 2 \times 1 \times 2 \times 6}_{\text{features extracted from the magnitude, including the mean distance}} = 924 \quad (2)$$

where the first term corresponds to the five statistics calculated for each of the three axes of each sensor and wrist, and the second term corresponds to the six measures calculated for the magnitude, including the mean distance. In this way, the final feature vector of each participant integrates detailed information about the movements performed under different experimental conditions, allowing a comprehensive analysis of motor patterns and providing a robust basis for applications in classification and machine learning methods.

After extraction, the features were manually selected into different experimental sets, varying according to tasks, sensors, wrists, and measures of interest. This approach allowed testing multiple combinations of attributes to investigate the impact of each subset of features on classifier performance. To address participant laterality, we also performed experiments retaining only the dominant wrist columns.

[1] https://physionet.org/content/parkinsons-disease-smartwatch/1.0.0/

2.3. Classification

Participant classification was performed using four different tree-based classifiers: *Random Forest*, *Extra Trees*, *XGBoost*, and *LightGBM*. Analyses were conducted for different binary classification scenarios, namely $PD \times CO$, $PD \times DD$, and $DD \times CO$. Each scenario was repeatedly evaluated with the different feature sets defined in the previous step, allowing comparison of the impact of manual attribute selection on model performance. Evaluation was performed using stratified *5-fold* cross-validation, ensuring separation between training and test sets in each iteration. Data were standardized using the *z-score* transformation, and the SMOTE (Synthetic Minority Over-sampling Technique) was applied to the training set to address class imbalance. The performance of the classifiers was evaluated in terms of accuracy, balanced accuracy, and weighted F1-score, enabling a comparative assessment of different algorithms and data preprocessing scenarios.

3. Results and Discussion

Table 1 presents the best results obtained for each classification scenario after exploring various feature sets, and Table 2 contains the adopted feature configurations.

Table 1. Classification results with SMOTE for different scenarios.

Scenario	Model	Acc (%)	Bal Acc (%)	F1 (%)
PDxCO	Random Forest	78.31	69.83	78.51
	Extra Trees	**82.25**	70.23	81.16
	XGBoost	78.03	69.56	78.29
	LightGBM	81.13	**74.31**	**81.34**
PDxDD	Random Forest	70.26	60.17	68.46
	Extra Trees	72.31	62.43	70.74
	XGBoost	70.77	63.87	70.34
	LightGBM	**74.62**	**67.37**	**73.86**
DDxCO	Random Forest	75.68	75.41	75.57
	Extra Trees	**82.39**	**82.48**	**82.36**
	XGBoost	77.76	77.93	77.84
	LightGBM	76.73	76.89	76.73

The best balanced accuracy results were achieved by the *LightGBM* model in the PD $\times$ CO scenario (74.31%), and in the PD $\times$ DD scenario (67.37%), while *Extra Trees* obtained the best result in the DD $\times$ CO scenario (82.48%). These values demonstrate competitive performance compared to the original study by Varghese et al. [2024], showing that simple statistical features extracted from inertial signals can provide relevant information for distinguishing subjects with Parkinson's Disease, differential diagnoses, and healthy controls. Furthermore, this study extends the original evaluation by Varghese et al. [2024] including the DD $\times$ CO scenario, allowing a more comprehensive analysis of the discriminative capacity of inertial signals in varied clinical contexts.

4. Conclusion

This work explored simple statistical features extracted from inertial signals obtained from smartwatches, which were able to distinguish individuals with Parkinson's Disease, differential diagnoses, and healthy controls with a balanced accuracy of 74.31%. The results indicate that machine learning-based approaches can achieve performance comparable to more complex previous studies, even when using a reduced set of attributes and

Table 2. Data configurations for different classification scenarios.

Scenario	Configuration	
PDxCO	Axes: Magnitude Sensors: Accelerometer, Gyroscope Features: Mean, Standard Deviation, Variance, Kurtosis, Skewness, Mean Distance Wrist: Dominant	Tasks: StretchHold, HoldWeight, DrinkGlass, CrossArms, TouchNose
PDxDD	Axes: Magnitude Sensors: Gyroscope Features: Mean, Standard Deviation Wrist: Left and right	Tasks: Relaxed1, Relaxed2, RelaxedTask1, RelaxedTask2, StretchHold, HoldWeight, DrinkGlass, CrossArms, TouchNose, Entrainment1, Entrainment2
DDxCO	Axes: X, Y, Z, Magnitude Sensors: Accelerometer, Gyroscope Features: Standard Deviation Wrist: Left and right	Tasks: Relaxed1, Relaxed2, RelaxedTask1, RelaxedTask2, StretchHold, HoldWeight, DrinkGlass, CrossArms, TouchNose, Entrainment1, Entrainment2

without relying on invasive techniques. In addition to validating the diagnostic potential of wearable sensors, this study extends the analysis by including the DD $\times$ CO scenario, reinforcing the applicability of this methodology in differentiating motor disorders. Thus, the findings highlight the potential of low-cost devices and tree-based supervised learning models as auxiliary tools for diagnosing Parkinson's Disease and similar neurodegenerative conditions.

Acknowledgements

The authors would like to acknowledge the support of CNPQ – Conselho Nacional de Desenvolvimento Científico, CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Financing Code #001, and CAPES/PDPG n. 30/2022 – Programa Emergencial de Solidariedade Acadêmica.

References

Bloem, B. R., Okun, M. S., and Klein, C. (2021). Parkinson's disease. *The Lancet*, 397(10291):2284–2303.

Braak, H. and Braak, E. (2000). Pathoanatomy of parkinson's disease. *J Neurol*, 247 Suppl 2:II3–10.

Delrobaei, M., Memar, S., Pieterman, M., Stratton, T., McIsaac, K., and Jog, M. (2017). Towards Remote Monitoring of Parkinson's Disease Tremor using Wearable Motion Capture Systems. *Journal of the Neurological Sciences*, 384.

Maetzler, W. and Hausdorff, J. M. (2012). Motor signs in the prodromal phase of parkinson's disease. *Movement Disorders*, 27(5):627–633.

Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkmann, J., Schrag, A.-E., and Lang, A. E. (2017). Parkinson disease. *Nature Reviews Disease Primers*, 3(1):1–21.

Varghese, J., Brenner, A., Fujarski, M., van Alen, C. M., Plagwitz, L., and Warnecke, T. (2024). Machine Learning in the Parkinson's disease smartwatch (PADS) dataset. *npj Parkinson's Disease*, 10(1):9.

Zhou, Z. D., Yi, L. X., Wang, D. Q., Lim, T. M., and Tan, E. K. (2023). Role of dopamine in the pathophysiology of parkinson's disease. *Translational Neurodegeneration*, 12(1):44.