2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Early Detection of Parkinson's Disease Using Center of Pressure Data and Machine Learning

Rabie Fadil, Asenath Huether, Robert Brunnemer, Andrew P. Blaber, Jau-Shin Lou, Kouhyar Tavakolian

Abstract— **Parkinson's disease (PD) is a progressive neurodegenerative disorder resulting in abnormal body movements. Postural instability is one of the primary motor symptoms of PD and contributes to falls. Measurement of postural sway through center of pressure (COP) data might be an objective indicator of Parkinson's disease. The goal of this work is to use machine learning to evaluate if different features of postural sway can differentiate PD patients from healthy controls. Time domain, frequency domain, time-frequency, and structural features were extracted from COP data collected from 19 PD patients and 13 healthy controls (HC). The calculated parameters were input to various machine-learning models to classify PD and HC. Random Forest outperformed the rest of the classifiers in terms of accuracy, false negative rate, F1-score, and precision. Time domain features had the best performance in differentiating PD from HC compared to other feature groups.**

I. Introduction

Parkinson's disease (PD) is a neurodegenerative disease resulting in tremors, muscle rigidity, bradykinesia, and postural instability. Postural instability can contribute to an increased risk of falls and injuries, and early postural unsteadiness is associated with more rapid disease progression [1]. While evidence of postural instability tends to emerge in moderate to late stages of PD, measures of biomechanical variables can detect differences in postural sway in patients with mild PD motor symptoms compared to healthy age-matched controls [2, 4].

Impairment of postural control is evident in the aging process and is often related to changes in the sensorimotor systems required for postural control [5]. These same systems influence postural control in PD, in addition to age, disease progression, orthostatic hypotension, gait disturbances, dopamine replacement therapy, and muscle loss and weakness [6]. Center of pressure (COP) refers to the point where the pressure of the body over the soles of the feet is concentrated. The deviations in the location of COP from its origin can be used to quantify body sway. Measures of center of pressure (COP) can differentiate between young and older adults in a passive standing position [3, 4]. Various COP measures have also been used in the evaluation of postural instability in PD, although there is significant variability in comparison groups, ON-OFF states, and sensory manipulations [2, 7, 8]. Compared to previous research, our study evaluated PD patients in ON state, during an eyes-open, passive standing position similar to their experience on a daily basis. Due to the subjective nature of traditional motor symptom assessments, we used a machine learning approach to identify COP variables and methods of analysis that were best able to classify PD patients from healthy controls (HC). Global variables in time and frequency domains allow for more conclusive evaluations of the contributing factors to postural unsteadiness. In the same manner, structural analyses were designed to capture dynamic postural changes during standing [9].

II. Materials and Methods

A. Dataset

Participants were required to lie supine on a tilt table for 5-minutes of baseline recording, followed by tilting the table to 70 degrees for 15 minutes to induce an orthostatic challenge. After which, participants were asked to stand upright on a force platform (AMTI's AccuSway Optimized™ multi-axis force platform) with their heads facing forward for 5 minutes. COP data were recorded from 19 patients with PD (age: 65 ± 5 years; height; 172 ±9 cm; weight: 110± 6 kg) and 13 healthy controls (HC; age: 67± 8 years; height: 164±9 cm; weight: 71±9 kg) in the standing position, and at a sampling frequency of 2 kHz. COP data (in mm) for each subject were split into medio-lateral (ML) and anterior-posterior (AP) components. All data were recorded at the Sanford Brain & Spine Center in the Parkinson's Research Laboratory, Fargo, ND, US. We terminated the experiment immediately if the participant showed signs of discomfort, uneasiness, nausea, or upon request. The Sanford Health IRB approved the protocol (IRB #1445) and we obtained written informed consent from all participants.

B. Features Extraction

COP signals (ML and AP) were low-pass filtered at a cutoff frequency of 20 Hz and then resampled to 100 Hz before further processing. The resampled AP and ML time series were referenced to their means. The last 4 min of standing data were used to extract a set of 30 handcrafted features. The extracted features are a combination of time (distance) and frequency domain features, time-frequency domain features, and structural features. This research is not concerned with making an exhaustive list of all of these features, so only the most common and relevant are considered here and discussed. All features used in this work are listed in Table 1.

a) Time Domain Features

Time domain features measure postural sway by estimating a parameter correlated with either the COP's displacement from the stabilogram's central point or its velocity [4, 10]. The mean distance represents the average distance from the mean COP. Based on the total distance traveled, path length

R.F., R.B., and K.T. Authors are with the University of North Dakota Biomedical Engineering Department, Grand Forks, ND, 58202 USA (corresponding author: 701.777.4446; kouhyar@und.edu).
A.H., J-S.L., Authors are with Peltier Parkinson's Disease Research Laboratory, Department of Neurology, Sanford Health, Fargo, ND.
A.P.B., Author is with the Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC, Canada.
J-S.L., Author is with the Department of Neurology, University of North Dakota, School of Medicine and Health Sciences.
The first three authors contributed equally to this study.

quantifies the magnitude of the two-dimensional displacement. The shorter the path length, the better the postural stability. It is considered a reliable outcome measure in a variety of populations and equilibrium conditions. The ellipse area measures 95% of the total area covered in the ML and AP directions. This parameter measures the overall postural performance (a smaller area reflects better performance). Mean sway velocity is calculated by summing the resultant distance between consecutive points on the COP path divided by time. Resultant velocity, for each of the AP and ML components, reflects the efficiency of the postural control system (a smaller velocity reflects better postural control).

b) Frequency Domain Features

Frequency domain parameters of COP describe the preferential involvement of specific neuronal loops in postural control [9]. This is performed by integrating amplitudes within low frequency (0–0.5 Hz) bands, which mostly account for visuo-vestibular regulation, medium frequency (0.5–2 Hz) for cerebellar participation, and high frequency (2–20Hz) for proprioceptive participation [9, 11, 12]. Spectral analyses of COP sway were analyzed using Fast Fourier Transform (FFT) to calculate the power spectral density areas for AP and ML data in these three frequency bands.

c) Structural Features

Structural features of the COP signal function explain the non-linear and complex nature of the postural control system. The stabilogram diffusion analysis (SDA) approach analyzes the dynamic properties of the COP signal [9]. Using SDA, two distinct patterns of postural regulation emerged, reflecting short-term sway movement from equilibrium without recovering (open-loop) and long-term oscillations back and forth around equilibrium (closed-loop). Diffusion coefficients and scaling exponent values extracted from SDA analysis showed to be an objective measurement of postural instability considering the ML and the AP axis as well as the plane of support [13, 14]. Other COP structural features are based on fractal analysis, which provide additional information about the underlying dynamics of postural sway. Fractal dimension methods detect chaos in COP signals, with sensitivity to capture small changes in postural control caused by age, visual disturbances, or neurological pathologies [9]. Sample entropy is another nonlinear dynamic parameter extracted from AP and ML postural data to measure the irregularity of COP signals (the higher the sample entropy value, the more irregular the COP time series). The regularity of postural sway is representative of the efficiency of its control system, with lower sample entropy in neurologically pathological patients compared to healthy controls [9, 13].

a) Time-Frequency Domain Features

Time-frequency analysis quantifies changes in a signal's spectral characteristics over time [15]. Empirical mode decomposition (EMD) of COP data extracts intrinsic mode functions (IMFs). IMFs are the local oscillations that make up the signal, as well as the residual, which represent the local patterns of the COP. The IMFs can be thought of as a set of narrow band non-stationary signals. In EMD, each IMF is a function with symmetric upper and lower envelopes, and the number of extrema and the number of zero crossings must be the same or differ at most by one.

TABLE I. LIST AND DESCRIPTION OF FEATURES EXTRACTED FROM COP DATA AND USED IN MACHINE LEARNING

Time Domain Features
The average distance from the mean COP
The average ML distance from the mean ML displacement
The average AP distance from the mean AP displacement
Root-mean-square distance of the center of pressure data
Root-mean-square distance of the ML data
Root-mean-square distance of the AP data
Average velocity of COP
Average velocity of the COP in the ML direction
Average velocity of the COP in the AP direction
95% confidence ellipse area of the COP data
The first axis of the 95% confidence ellipse
Empirical Mode Decomposition
Minimum of the 6th intrinsic mode function for AP data
Coefficient of variation of the 4th intrinsic mode function for AP data
Minimum of the 1st intrinsic mode function for ML data
Range of the 1st intrinsic mode function for ML data
Stabilogram Diffusion Analysis
Scaling exponents of the short-term region of AP data
Scaling exponents of the short-term region of ML data
Scaling exponents of the short-term region of the COP data
Power spectral features
Total power spectral density area for ML data
Power spectral density area for ML data in the low frequency band
Normalized power spectral density area for ML data in the low frequency band
Normalized power spectral density area for ML data in the high frequency band
Ratio of medium frequency to high frequency power spectral density areas for ML data
Power spectral density area for AP data in the high frequency band
Power spectral density area for AP data in the medium frequency band
Normalized power spectral density area for AP data in the high frequency band
Ratio of medium frequency to high frequency power spectral density areas for AP data
Sample Entropy
Sample entropy of AP data
Sample entropy of ML data
Fractal Dimension
Higuchi fractal dimension of the ML data

Unlike the Fourier transform and wavelet analysis, EMD makes no assumptions about the signal's composition and does not depend on any particular wavelet basis. Statistical features such as mean, minimum, maximum, coefficient of variation, skewness, variance, and kurtosis can be extracted from different IMFs obtained from ML and AP time series and can characterize postural sway [16, 17].

Each of these feature groups has the potential ability to describe different aspects of sway and can differentiate between subject groups. However, the set of parameters that can discriminate between PD patients and healthy older adults using machine learning is yet to be investigated.

C. Classification

In this work, we evaluated the performance of different machine learning models in discriminating between PD patients and healthy older adults. These techniques include Random Forest, Support Vector Machine (SVM), Decision Tree, K-Nearest Neighbor, Neural Network (NN), and Gaussian Naive Bayes. Each model was trained and tested using a stratified 5-fold cross-validation procedure in which the feature set and corresponding groups (PD and HC) were divided into 5 non-overlapping splits. Four of the splits (80 percent of the data) were used for training and the remaining split (20 percent of the data) was used for testing. The process was repeated 5 times, each one using a different split on the data for testing. To assess the performance of the proposed method, accuracy, false negative rate (FNR), F1-score, and precision [18]. In this work, a false-negative response is more important than a false-positive answer because the former might lead to a delayed diagnosis of PD patients.

III. Results and Discussion

The obtained results for the considered classifiers are shown in Table II. On average, RF outperformed the rest of the classifiers while considering all metrics. K-Nearest Neighbor presented the worst performance in terms of accuracy, false negative rate, and F1-score. We believe this is because K-Nearest Neighbor performance declines when the dimension of data is high. Random Forest did well in classifying PD with an F1-score of 0.83, and a precision of 0.87 compared to HC. Moreover, Random Forest has a lower FNR (0.15 vs 0.23) in predicting PD compared to HC (Fig. 1).

Random Forest performed better in all metrics when trained on time domain features compared to the frequency domain, structural and EMD features. It was able to outperform the initial model that was trained with 30 parameters (Fig. 2). EMD features also achieved promising results in terms of accuracy, F1-score, and precision when compared to the frequency domain and structural features. Moreover, EMD features correctly identified 90% (FNR = 0.1) of the Parkinson's patients in the testing dataset (Fig. 3) compared to the rest of the feature categories.

TABLE II. Results comparison of different classifiers

Classifier	Accuracy	FNR	F1-score	Precision
Random Forest	0.81	0.19	0.8	0.86
NN	0.79	0.25	0.75	0.85
Gaussian Naive Bayes	0.78	0.21	0.77	0.80
SVM	0.72	0.23	0.80	0.67
Decision Tree	0.70	0.29	0.69	0.71
K-Nearest Neighbor	0.64	0.39	0.60	0.69

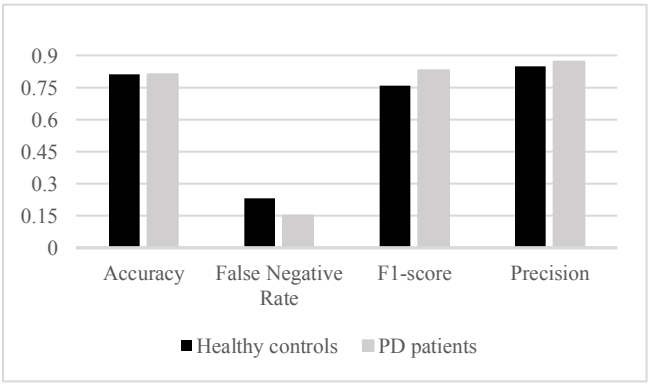

Figure 1. Performance results of Random Forest on the set of 30 features

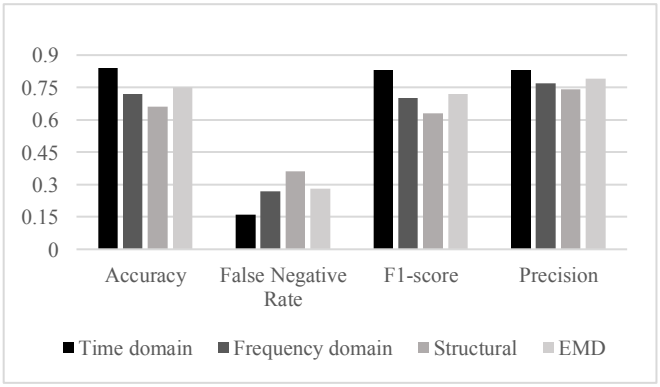

Figure 2. Performance results of Random Forest in differentiating between PD patients and healthy controls when considering different feature groups

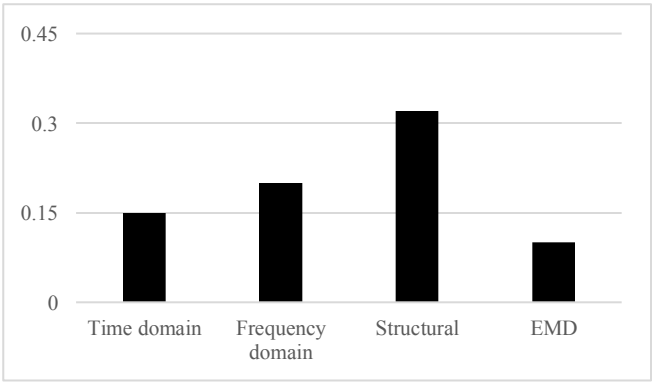

Figure 3. False Negative Rate of Random Forest for detecting PD when considering different feature groups

IV. CONCLUSION

In this study, various machine-learning approaches were investigated along with different feature categories to differentiate PD patients from healthy controls. The most suitable classifier was found by testing Random Forest, Support Vector Machine, Decision Tree, K-Nearest Neighbor, Neural Network, and Gaussian Naive Bayes. The best classification performance was obtained using Random Forest and time domain features, while empirical mode decomposition parameters showed promising results and have the potential to discriminate PD patients from HC. Overall, it is believed that the proposed sway features coupled with a machine learning approach have the potential for clinical application in the early diagnosis and detection of Parkinson's disease.

REFERENCES

[1] Jankovic *et al.*, "Variable Expression of Parkinson's Disease: A base-line analysis of the DAT ATOP cohort," *Neurology*, vol. 40, pp. 1529–1534, 1990.

[2] N. Chastan, B. Debono, D. Maltête, and J. Weber, "Discordance between measured postural instability and absence of clinical symptoms in Parkinson's disease patients in the early stages of the disease," *Mov. Disord.*, vol. 23, no. 3, pp. 366–372, 2008, doi: 10.1002/mds.21840.

[3] B. Tests, "Aging and Postural Control," pp. 1–9, 1990.

[4] T. E. Prieto, J. B. Myklebust, R. G. Hoffmann, E. G. Lovett, and B. M. Myklebust, "Measures of postural steadiness: Differences between healthy young and elderly adults," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 9, pp. 956–966, 1996, doi: 10.1109/10.532130.

[5] R. Johansson and M. Magnusson, "Human PosturalDynamics," *Crit Rev. Biomed. Eng.*, vol. 18, pp. 413–437, 1991.

[6] S. Grill, "Postural instability in Parkinson's disease.," *Md. Med. J.*, vol. 48, no. 4, pp. 179–181, 1999.

[7] M. Mancini, P. Carlson-Kuhta, C. Zampieri, J. G. Nutt, L. Chiari, and F. B. Horak, "Postural sway as a marker of progression in Parkinson's disease: A pilot longitudinal study," *Gait Posture*, vol. 36, no. 3, pp. 471–476, 2012, doi: 10.1016/j.gaitpost.2012.04.010.

[8] J. M. Schmit *et al.*, "Deterministic center of pressure patterns characterize postural instability in Parkinson's disease," *Exp. Brain Res.*, vol. 168, no. 3, pp. 357–367, 2006, doi: 10.1007/s00221-005-0094-y.

[9] T. Paillard and F. Noé, "Techniques and Methods for Testing the Postural Function in Healthy and Pathological Subjects," *Biomed Res. Int.*, vol. 2015, 2015, doi: 10.1155/2015/891390.

[10] Bigelow, Kimberly Edginton. "Identification of key traditional and fractal postural sway parameters to develop a clinical protocol for fall risk assessment in older adults." PhD diss., The Ohio State University, 2008.

[11] Suzuki, Makoto, Hiroyuki Fujisawa, Hiroto Suzuki, Shingo Kawakami, Kenichi Murakami, and Chie Miki. "Frequency analysis of the center of pressure in tandem stance in community-dwelling elderly." Journal of physical therapy science 29, no. 5 (2017): 828-831.

[12] Baratto, Luigi, Pietro G. Morasso, Cristina Re, and Gino Spada. "A new look at posturographic analysis in the clinical context: sway-density versus other parameterization techniques." Motor control 6, no. 3 (2002): 246-270.

[13] Rizzato, Alex, Gerardo Bosco, Michael Benazzato, Antonio Paoli, Giulia Zorzetto, Attilio Carraro, and Giuseppe Marcolin. "Short-term modifications of postural balance control in young healthy subjects after moderate aquatic and land treadmill running." Frontiers in physiology 9 (2018): 1681.

[14] Peterka, Robert J. "Postural control model interpretation of stabilogram diffusion analysis." Biological cybernetics 82, no. 4 (2000): 335-343.

[15] Schumann, Timothy, Mark S. Redfern, Joseph M. Furman, Amro El-Jaroudi, and Luis F. Chaparro. "Time-frequency analysis of postural sway." Journal of biomechanics 28, no. 5 (1995): 603-607.

[16] Pachori, Ram Bilas, D. J. Hewson, Hichem Snoussi, and Jacques Duchêne. "Analysis of center of pressure signals using empirical mode decomposition and Fourier-Bessel expansion." In TENCON 2008-2008 IEEE Region 10 Conference, pp. 1-6. IEEE, 2008.

[17] Chou, Li-Wei, Kang-Ming Chang, Yi-Chun Wei, and Mei-Kuei Lu. "Empirical Mode Decomposition-Derived Entropy Features Are Beneficial to Distinguish Elderly People with a Falling History on a Force Plate Signal." Entropy 23, no. 4 (2021): 472.

[18] Fadil, Rabie, Andie Jackson, Badr Abou El Majd, Hassan El Ghazi, and Naima Kaabouch. "Classification of Microcalcifications in Mammograms using 2D Discrete Wavelet Transform and Random Forest." In 2020 IEEE International Conference on Electro Information Technology (EIT), pp. 353-359. IEEE, 2020.