Using FSR Data for Refined Early Detection of Parkinson’s Disease with Gait Analysis

Rohan Pavuluri

Abstract

Early detection of Parkinson’s disease (PD) is important for improving patient quality of life and slowing disease progression. This research explores the use of force-sensitive resistors (FSRs) in gait analysis to identify early indicators of PD. We use the Gait and Neurological Disorders (GaitND) dataset from PhysioNet. This dataset includes gait data from individuals with Parkinson’s disease, Huntington’s disease, amyotrophic lateral sclerosis, and healthy controls. Analyzed key time-series features such as stride intervals, swing intervals, stance intervals, and double support intervals for the left and right limb. Also we used advanced machine learning techniques like ensemble methods and optimization algorithms to enhance classification accuracy. Using a genetic algorithm we achieved an accuracy of 81.08%. The findings indicate the potential of FSR-based wearable technologies for non-invasive and continuous monitoring of gait patterns. This research also shows the feasibility of using FSRs into other types of diagnostic tools for better PD detection and for early prevention.

Introduction

Background and Context

Parkinson’s disease (PD) is a neurodegenerative disorder characterized by progressive motor and non-motor symptoms and it significantly impacts a patient's quality of life [1]. As the disease advances, it leads to severe motor dysfunction, cognitive decline, and autonomic disturbances [2]. Early detection and diagnosis are important for management and treatment. They can potentially slow disease progression and improve outcomes for the patient. However, the use of traditional diagnostic methods like clinical neurological examinations and imaging techniques only identify PD in later stages when symptoms are pronounced and irreversible neuronal damage has happened. Intervention can be most beneficial at early stages because this diagnostic delay demonstrates the need for innovative approaches to detect the disease in its earliest [3] [4].

Gait, the manner in which a person walks, is a prominent motor symptom affected by PD and can serve as a vital indicator of the disease [5]. Some gait disturbances in PD are bradykinesia (slowness of movement), shuffling steps, and reduced arm swing. These may appear before other more obvious symptoms. Because of this gait analysis is useful for early PD detection. This analysis can provide details into motor function and early biomarkers of the disease.

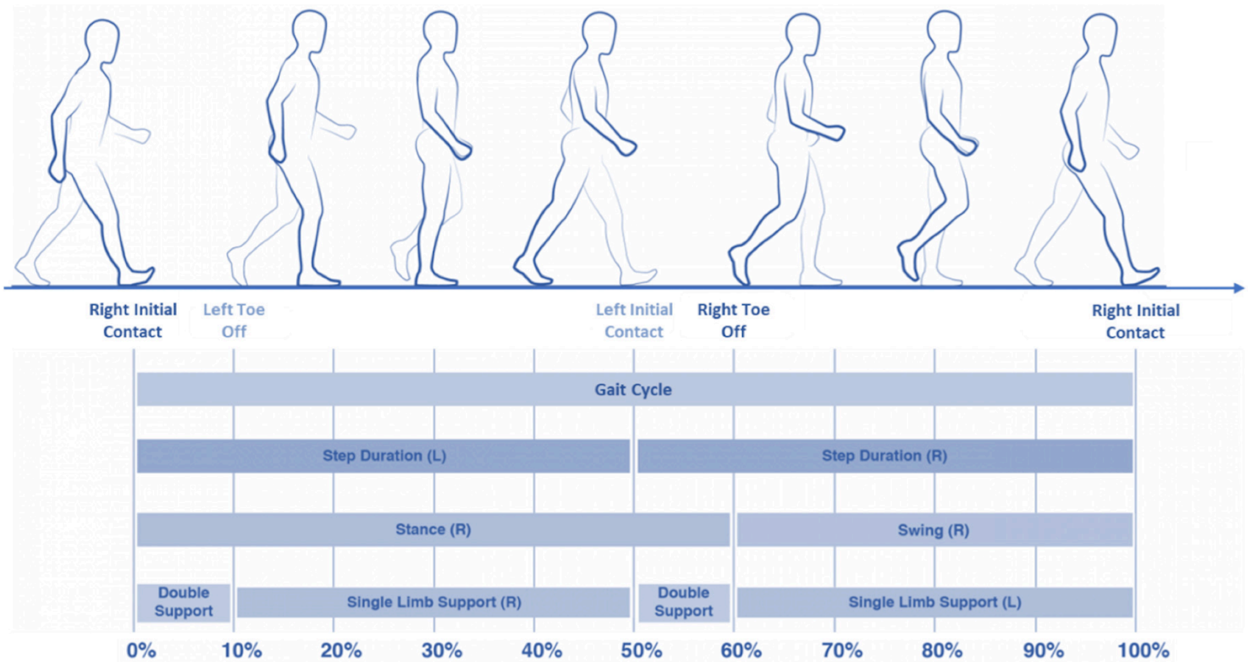

Figure 1: Gait Cycle [6]

Problem Statement and Rationale

Despite advancements in diagnostic technologies, early detection of Parkinson's disease remains a significant challenge. Current methodologies are either resource-intensive or dependent on observable symptomatic progression, delaying critical intervention opportunities. The underutilization of wearable sensors such as FSRs in personalized, predictive diagnostics demonstrates a gap in existing research. Furthermore, some current machine learning models have already been applied to gait data but their predictive accuracy is suboptimal due to limited use of ensemble methods and advanced algorithms for checking accuracy. This study seeks to address these challenges by using FSR-based gait data from the Gait and Neurological Disorders (GaitND) dataset with innovative ensemble techniques to enhance prediction accuracy for early Parkinson's diagnosis.

Significance and Purpose

Recent advancements in wearable technology may be promising for continuous and non-invasive monitoring of gait patterns [7]. FSRs are studied and found to be advantageous for gait analysis [8]. FSRs are thin, flexible sensors that measure the force exerted on them. This makes them ideal for their location into wearable devices such as insoles or smart footwear [9]. Their high sensitivity and ease of use allow for it's use in precise detection of subtle changes in gait dynamics. Subtle changes can be indicative of early PD. The ease of using wearable devices and FSRs in everyday life allows for continuous monitoring, and this offers real-time data on a patient's gait and movement patterns [10]. This real-time data collection is important for noticing the onset of symptoms that may fluctuate or worsen progressively, which could provide a more comprehensive picture of the disease's impact over time.

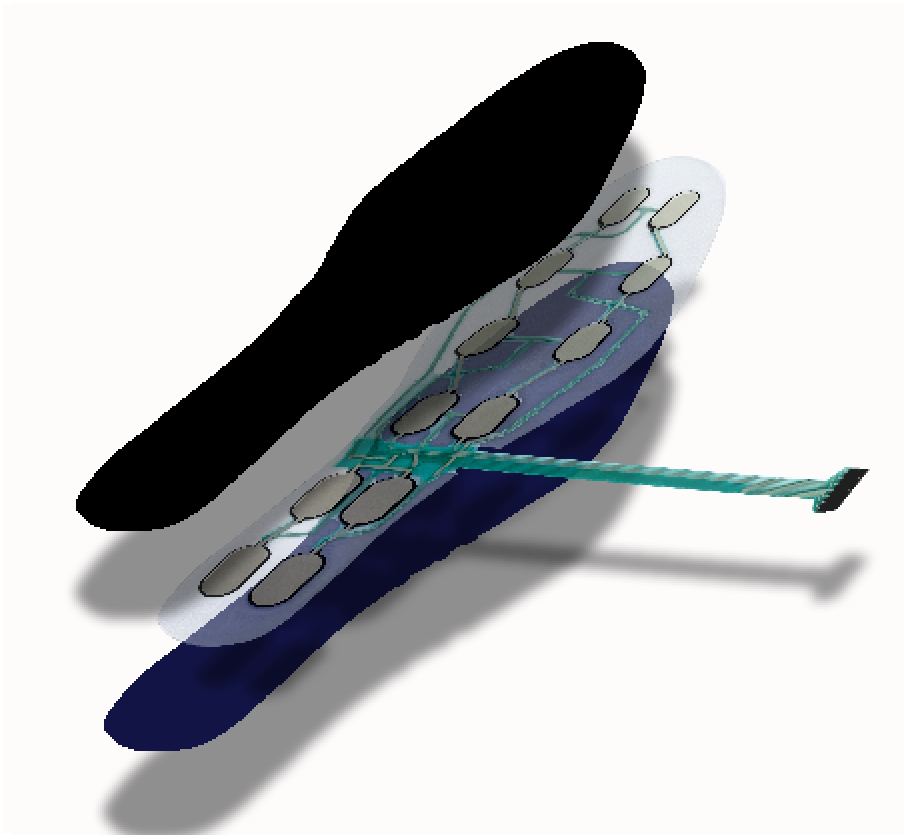

Figure 2: Force-sensitive resistors [11]

The application of Artificial Intelligence (AI) in this context is promising, as it allows for the processing and analysis of data with greater accuracy and efficiency than traditional methods [12]. AI algorithms can be trained to recognize complex patterns and subtle deviations in gait, or in particular features which we present later, that may be indicative of early PD [13]. This capability enhances the early diagnosis and allows clinicians to initiate treatments that can slow the disease and improve daily life for the patient [14]. This paper introduces an altered perspective on current diagnostic practices with the use of potential of wearable technology and AI for early disease detection. I analyzed gait data captured by FSRs and used AI algorithms to detect patterns. In doing so, I aimed to identify early changes in gait that could be indicative of the onset of Parkinson's disease before more severe symptoms develop.

Objectives

The main objectives in this research are:

- Analyze time-series gait data from the GaitND dataset. (Stride, swing, stance, and double support intervals)

- Develop an ensemble machine learning model tested by various classification algorithms to get the highest predictive accuracy.
- Determine the effectiveness of force-sensitive resistors (FSRs) as a tool to continuously monitor and early detect Parkinson's disease.
- Explore the applications of the system in real-world diagnostics.

Scope and Limitations

The scope of this study is limited to gait data from the GaitND dataset. This dataset contains stride intervals, swing intervals, stance intervals, and double support intervals for the left and right sides. The data provides comprehensive gait information but is of a relatively small sample size (only containing 64 records) and this may reduce the generalizability of the findings. There also may be inherent biases in machine learning algorithms that affect the validity of the findings.

Theoretical Framework

This study is based in the theoretical framework of biomechanical analysis and predictive modelling. Our application of machine learning algorithms allows for greater understanding between gait patterns and neurodegeneration. Additionally, our use of an ensemble model allows us to aggregate predictions from multiple models which improves predictive reliability.

Methodology Overview

First, we use the GaitND dataset to extract and preprocess gait features. Features such as Gait Stability Ratio, Right Swing Stance Ratio, and Left Swing Stance Ratio are combined using ensemble techniques, including genetic algorithms and random search to optimize weights and enhance our prediction accuracy.

Methods

Research Design

The study used a cross-sectional observational design, analyzing pre-collected gait data from the GaitND dataset. This approach enabled a comparative analysis of Parkinson's disease patients and healthy controls using machine learning techniques to evaluate predictive accuracy.

Participants & Data Collection

The Gait and Neurological Disorders (GaitND) dataset from PhysioNet is a comprehensive collection of gait data; this data set aims in quantifying gait dynamics in individuals with neurodegenerative diseases [16]. The dataset includes records from 64 subjects, categorized into four groups: Parkinson's disease (n = 15), Huntington's disease (n = 20), amyotrophic lateral sclerosis (ALS) (n = 13), and healthy controls (n = 16). Of which I use the 15 Parkinson's disease data and the 16 healthy controls. The gait data were collected using force-sensitive resistors (FSRs), which provide an output proportional to the force exerted under the foot. The raw signals from these sensors were processed to derive stride-to-stride measures of footfall contact times. The dataset includes four files per record, identified by the subject group (e.g., ”park” for Parkinson's disease) and an arbitrary ID number [16].

These time series have not been filtered, allowing for detailed analysis and potential application of various signal processing techniques.

Additionally, the dataset includes a clinical description file (subject-description.txt) which provides detailed clinical information for each subject. This includes variables such as age, gender, height, weight, walking speed, and disease severity or duration. For Parkinson's disease patients, the Hoehn and Yahr score is provided, indicating the stage of the disease. While control subjects are assigned an arbitrary value of 0 [16].

Variables(Features) and Measurements

Feature extraction plays an important role in using raw gait data to determine meaningful metrics that can be used for predicting Parkinson's disease (PD). Given that gait abnormalities are one of the hallmark symptoms of PD, force-sensitive resistors (FSRs) provide data by capturing the forces exerted by the foot during walking [17]. In this study, we extract features related to temporal, spatial, and kinetic aspects of gait to analyze the distinctive patterns in PD patients compared to healthy controls [6]

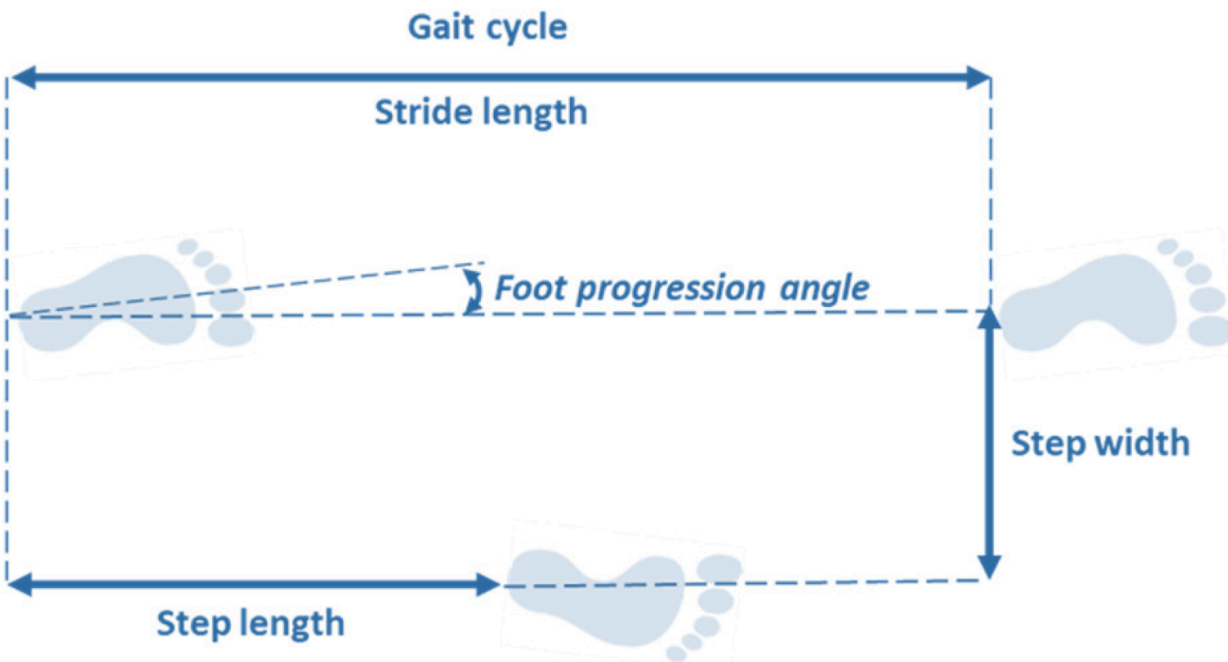

Figure 2: Gait Features [6]

1) Temporal Features

Temporal features provide insights into the timing of different phases of the gait cycle, which are often altered in Parkinson's patients.

- Stride Interval (Left and Right): The time duration between consecutive footfalls of the same foot. This feature helps in identifying irregularities in gait rhythm.

$$\text{Stride Interval} = t_{i+1} - t_i \tag{1}$$

- Swing Interval (Left and Right): The time duration during which the foot is off the ground. Parkinson's patients often exhibit reduced swing times.

$$\text{Swing Interval} = t_{\text{toe-off}} - t_{\text{heel-strike}} \tag{2}$$

where $t_{toe-off}$ is the time when the foot leaves the ground, and $t_{heel-strike}$ is the time when the foot initially contacts the ground.

- Stance Interval (Left and Right): The time during which the foot is in contact with the ground. Prolonged stance times are a common characteristic in PD patients.

$$\text{Stance Interval} = t_{\text{heel-strike}} - t_{\text{toe-off}} \tag{3}$$

2) Spatial Features

Spatial features relate to the distribution and symmetry of forces applied during walking.

- Foot Pressure Distribution: The average pressure distribution across different regions of the foot (e.g., heel, midfoot, forefoot). Altered pressure distributions can be indicative of compensatory mechanisms in PD patients.
- Symmetry Index: A measure of the symmetry between left and right foot pressures. Parkinson's disease often leads to asymmetrical gait patterns.

$$\text{Symmetry Index} = \frac{|P_{\text{left}} - P_{\text{right}}|}{(P_{\text{left}} + P_{\text{right}})} \tag{4}$$

3) Kinetic Features
Kinetic features focus on the forces generated during walking and how they change over time.
   - Peak Pressure (Left and Right): The maximum force exerted by each foot during the gait cycle. Reduced peak pressures can indicate weakened muscular control.
   - Force-Time Integral: The area under the force-time curve represents the total force exerted during the stance phase. This feature helps in understanding the overall force production capability of the individual.
   - Variability of Force: The standard deviation of forces recorded across multiple gait cycles. Increased variability is often observed in Parkinson's patients due to inconsistent gait patterns.

$$\text{Variability of Force} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - \bar{F})^2} \tag{5}$$

Procedure

Force-sensitive resistors (FSRs) were used to measure pressure distribution and kinetic features in the gait data. These sensors detect changes in resistance as a function of applied force which makes them ideal for capturing foot pressure during walking. The FSRs were strategically placed under specific regions of the foot—heel, midfoot, and forefoot—to capture temporal and kinetic data relevant to gait analysis [17].

The raw FSR data was recorded in the form of continuous time series, where the pressure changes were logged at predefined intervals for each footstep. From this data, key features such as foot pressure distribution, peak pressure, and the force-time integral were extracted.

Each FSR output was converted into force measurements using the sensor's calibration curve, ensuring accurate quantification of the pressure exerted during walking.

To ensure consistency and comparability across subjects, the raw FSR outputs were normalized. Given the variability in gait patterns due to differences in subject characteristics such as weight, height, and walking speed, normalization was essential for reducing the effect of these confounding factors.

Following the extraction of features, a feature selection process was carried out to identify the most relevant predictors of Parkinson's disease. Statistical methods, such as correlation analysis, were employed to reduce dimensionality and highlight the features with the highest predictive power [6]. The features that consistently showed strong correlations with PD diagnosis are the Double Support Interval, Swing Interval Variability, and Peak Pressure on the left foot. These features were then used as inputs to the learning models. By analyzing these features, we can understand the underlying biomechanics of Parkinson's gait and improve the accuracy of early diagnosis. These features will then be used in conjunction with machine learning algorithms to develop predictive models aimed at early diagnosis and monitoring of Parkinson's disease.

Data Analysis

In this paper, multiple machine learning techniques were explored to predict PD based on the gait data that has been obtained from force-sensitive resistors. The dataset was split into training and testing data with 80% of the points used as training and the remaining 20% as testing data.

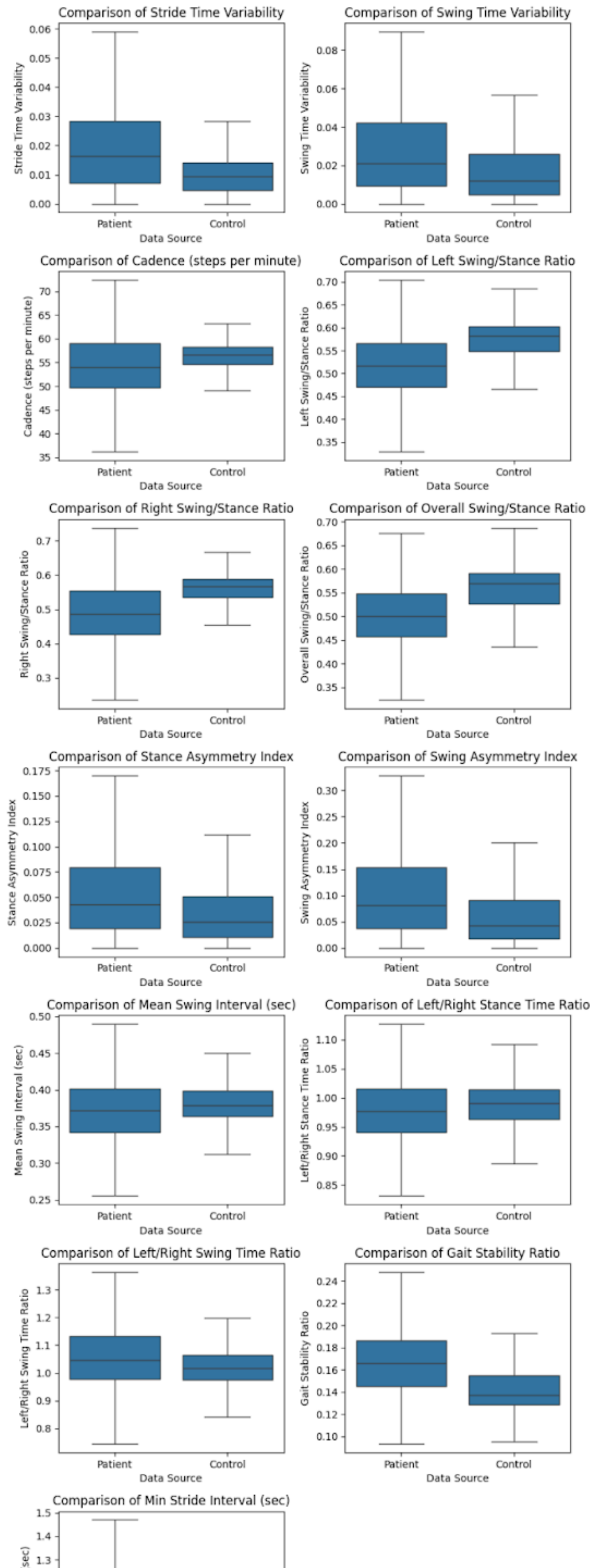
Comparison of Stride Time Variability
Stride Time Variability
Patient
Control
Data Source
Comparison of Swing Time Variability
Swing Time Variability
Patient
Control
Data Source
Comparison of Cadence (steps per minute)
Cadence (steps per minute)
Patient
Control
Data Source
Comparison of Left Swing/Stance Ratio
Left Swing/Stance Ratio
Patient
Control
Data Source
Comparison of Right Swing/Stance Ratio
Right Swing/Stance Ratio
Patient
Control
Data Source
Comparison of Overall Swing/Stance Ratio
Overall Swing/Stance Ratio
Patient
Control
Data Source
Comparison of Stance Asymmetry Index
Stance Asymmetry Index
Patient
Control
Data Source
Comparison of Swing Asymmetry Index
Swing Asymmetry Index
Patient
Control
Data Source
Comparison of Mean Swing Interval (sec)
Mean Swing Interval (sec)
Patient
Control
Data Source
Comparison of Left/Right Stance Time Ratio
Left/Right Stance Time Ratio
Patient
Control
Data Source
Comparison of Left/Right Swing Time Ratio
Left/Right Swing Time Ratio
Patient
Control
Data Source
Comparison of Gait Stability Ratio
Gait Stability Ratio
Patient
Control
Data Source
Comparison of Min Stride Interval (sec)

Figure 3: Side by side boxplots of features relation with patient or control.

1) *Individual Classifier Model Performance:* By quick inspection of the parallel boxplots generated in feature extraction, we note that the most differing 3 features are Gait Stability Ratio, and Left Swing/Stance Ratio, Right Swing/Stance Ratio.

   With this information we quickly test just based off values in each of the respective columns independently of each other and obtain the following raw percentages of accuracy in identification:

Gait Stability Ratio	65.32%
Left Swing/Stance Ratio	67.89%
Right Swing/Stance Ratio	42.21%

Table 1: Individual classifier accuracy on three features

   From this we note that Gait Stability Ratio and Left Swing/Stance Ratio seemed far more promising than Right Swing/Stance Ratio so for further overall predictive power, a brute force manual ensemble approach was applied.

2) *Manual Ensemble:* This technique was very primitive and just relied on an initial guess on weights assuming that Gait Stability and Left Swing/Stance would be better by giving them 0.4 weights each and Right Swing/Stance a 0.2 weight. This weightage determines how much of the final chance that each feature can modify.

Manual Ensemble Accuracy	72.91%

Table 2: Manual ensemble overall accuracy with weights [0.4, 0.4, 0.2]

   Even with initial arbitrary values, a significantly higher accuracy was obtained, leading us to seek a more definitive approach. In the next approach, we iteratively run through all possible weights which is more brute force.

3) *Iterative Brute Force / Dynamic Optimization Ensemble:* In this approach, we quickly note that smaller and more precise intervals take exponentially longer to run, which motivates us to look for more faster optimizations. Regardless this method gave the following data:

Iterative Brute Force / Dynamic Optimization Ensemble Accuracy	73%

Table 3: Iterative brute force or dynamic optimization ensemble accuracy

- Weights: [0.01, 0.495, 0.495]
    - 0.01 was for Gait Stability
    - 0.495's were for Left and Right Swing/Stance.

Surprising to see a similar accuracy regardless of vastly different weights, but only causing us to understand that we may need more features in our analysis. However, with 3 features it already takes an astronomical amount of time to get more precise, so for 13 it would take even longer. To quicken the speed we start using advanced optimization techniques.

4) *Advanced Optimization Techniques:* To further improve performance, optimization methods such as random search, evolutionary algorithms, and Bayesian optimization were used.
    a) Scipy Optimization:
    Using the scipy.optimize library, weights were assigned to all 13 features resulting in a slight improvement to *73.04%.*
    b) Random Search:
    A random search method was employed to explore the solution space, achieving a significant improvement in accuracy:
        - After 200 iterations, the best accuracy was: *78.44%.*
        - The corresponding best weights were: [0.09484909, 0.11365376, 0.00501936, 0.0895144, 0.12566684, 0.16114454, 0.09320969, 0.06588727, 0.01492724, 0.05828005, 0.04906177, 0.11174149, 0.0170445].
    c) Evolutionary Algorithms (Genetic Algorithm):
    Finally, a genetic algorithm was applied using the DEAP library to optimize the weight assignment across all 13 features. This method achieved the best overall accuracy of **81.08%**.
        - The best weights identified were: [0.06394503, 0.19941245, 0, 0.13801787, 0.13534947, 0.16612872, 0, 0.01421639, 0.01052521, 0.14310884, 0.00938601, 0.11991001, 0].

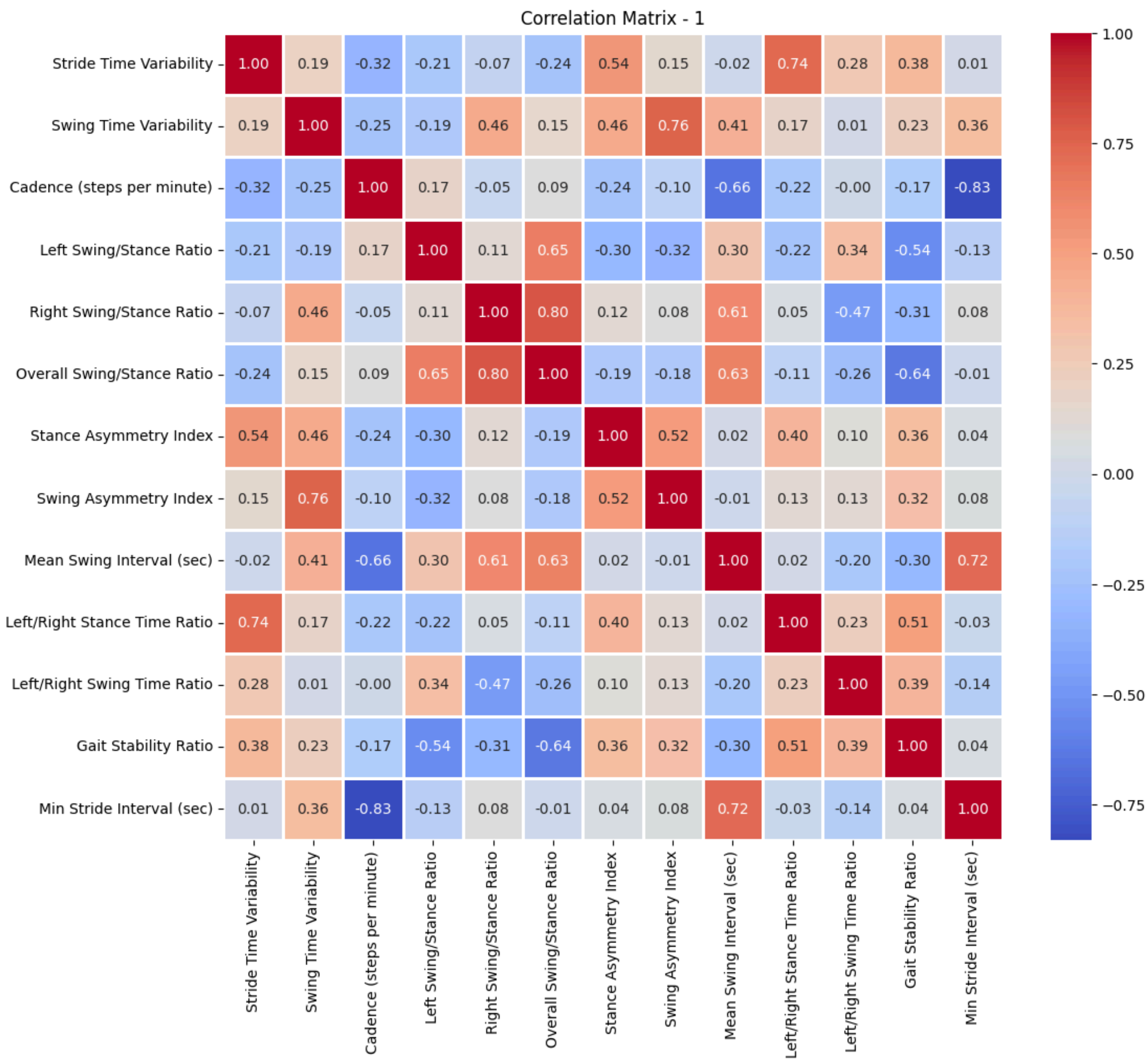

Figure 4: Correlation matrix for 13 different features.

Ethical Considerations

The study followed ethical guidelines, we used publicly available data from PhysioNet. The dataset includes anonymous records, ensuring confidentiality. Ethical concerns related to participant consent and data use were addressed by the original dataset creators. This secondary analysis aligns with the FAIR (Findable, Accessible, Interoperable, and Reusable) principles.

Discussion

Restatement of Key Findings

This study aimed to evaluate the early detection of Parkinson's disease (PD) using the analysis of gait data, force-sensitive resistors (FSRs), and machine learning models. The most notable finding was the significant improvement in prediction accuracy through the use of an ensemble approach. Our genetic algorithm method achieved a peak accuracy of 81.08%. Additionally, gait-specific features such as stride intervals, swing intervals, and stance intervals were identified as key indicators of PD progression, supporting their use in predictive modeling.

Implications and Significance

The findings of this research could result in the development of more effective early intervention strategies and lead to improved diagnostic tools and wearable devices. Such devices then contribute to better management of Parkinson's disease which improves lives of patients. By using advanced technology with traditional clinical practices, this approach aims to revolutionize the early diagnosis and monitoring of PD; this allows timely interventions and improving the quality of life for patients [12]. As such, this research contributes to the field of neurodegenerative disease management and also the potential of using technology with healthcare [15].

Connection to Objectives

The primary objective was to determine if the integration of advanced computational methods and FSR-data could enhance the early detection of PD. This objective was successfully met, as demonstrated by our optimized ensemble model's performance. However, the study also revealed unexpected disparities in individual model accuracies, for example, the low reliability of the Right Swing Stance Ratio model (42.21%).

Limitations

Once again, there were limitations in this study.

1. The GaitND dataset only has 64 records which may not fully represent the diverse characteristics of patients with PD over different demographics and stages of disease.
2. The reliance on such a dataset reduces our generalizability of results.
3. The weighting approaches may introduce biases.
4. The use of FSRs may face practical challenges in terms of development since they are not widely used for long-term patients.

Closing Thought

The use of wearable technology and machine learning gave unprecedented opportunities for advancing early diagnostic tools for neurological disorders. Future research can introduce more accessible and personalized healthcare solutions after addressing our identified limitations. The findings of this study demonstrated the potential of innovative methodologies in enhancing the quality of life for individuals with Parkinson's disease.

References

1. W. Poewe, K. Seppi, C. M. Tanner, G. M. Halliday, P. Brundin, J. Volkmann, A.-E. Schrag, and A. E. Lang, “Parkinson disease,” *Nature reviews Disease primers*, vol. 3, no. 1, pp. 1–21, 2017.
2. L. V. Kalia and A. E. Lang, “Parkinson’s disease,” *The Lancet*, vol. 386, no. 9996, pp. 896–912, 2015.
3. V. Syam, S. Safal, O. Bhutia, A. K. Singh, D. Giri, S. S. Bhandari, and R. Panigrahi, “A non-invasive method for prediction of neurodegenerative diseases using gait signal features,” *Procedia computer science*, vol. 218, pp. 1529–1541, 2023.
4. T. Gasser, “Molecular pathogenesis of parkinson disease: insights from genetic studies,” *Expert reviews in molecular medicine*, vol. 11, p. e22, 2009.
5. A. Mirelman, P. Bonato, R. Camicioli, T. D. Ellis, N. Giladi, J. L. Hamilton, C. J. Hass, J. M. Hausdorff, E. Pelosin, and Q. J. Almeida, “Gait impairments in parkinson’s disease,” *The Lancet Neurology*, vol. 18, no. 7, pp. 697–708, 2019.
6. L. Di Biase, A. Di Santo, M. L. Caminiti, A. De Liso, S. A. Shah, L. Ricci, and V. Di Lazzaro, “Gait analysis in parkinson’s disease: An overview of the most accurate markers for diagnosis and symptoms monitoring,” *Sensors*, vol. 20, no. 12, p. 3529, 2020.
7. S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, “A review of wearable sensors and systems with application in rehabilitation,” *Journal of neuroengineering and rehabilitation*, vol. 9, pp. 1–17, 2012.
8. M. F. Shaikh, Z. Salcic, and K. Wang, “Analysis and selection of the force sensitive resistors for gait characterisation,” in *2015 6th International Conference on Automation, Robotics and Applications (ICARA)*, pp. 370–375, IEEE, 2015.
9. A. Pantelopoulos and N. G. Bourbakis, “A survey on wearable sensor- based systems for health monitoring and prognosis,” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 40, no. 1, pp. 1–12, 2009.
10. U. Verma, P. Tyagi, and M. Kaur, “Artificial intelligence in human activity recognition: a review,” *International Journal of Sensor Networks*, vol. 41, no. 1, pp. 1–22, 2023.

11. *Flexible gait analysis piezoresistive insole force sensitive resistor for smart wear.* FSRTEK. (2021, July 9). https://www.fsrtek.com/flexible-gait-analysis-piezoresistive-insole-force-sensitive-resistor
12. U. K. Patel, A. Anwar, S. Saleem, P. Malik, B. Rasul, K. Patel, R. Yao, A. Seshadri, M. Yousufuddin, and K. Arumaithurai, "Artificial intelligence as an emerging technology in the current care of neurological disorders," *Journal of neurology*, vol. 268, pp. 1623–1642, 2021.
13. J. Figueiredo, C. P. Santos, and J. C. Moreno, "Automatic recognition of gait patterns in human motor disorders using machine learning: A review," *Medical engineering & physics*, vol. 53, pp. 1–12, 2018.
14. B. Y. Kasula, "Ai applications in healthcare a comprehensive review of advancements and challenges," *International Journal of Management Education for Sustainable Development*, vol. 6, no. 6, 2023.
15. Parkinson's Foundation, "Detecting early parkinson's with a wearable movement-tracking device," 2024. Accessed: 2024-10-26.
16. J. M. Hausdorff, A. Lertratanakul, M. E. Cudkowicz, A. L. Peterson, D. Kaliton, and A. L. Goldberger, "Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis," *Journal of applied physiology*, 2000.
17. S. Mahalingam, S. Mahendra, H. Sanjay, and B. V. Hiremath, "Integration of gyroscopes and force sensitive resistors for gait analysis of lower limb injury rehabilitation," in *2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT-2019),* 2019.