Walk this Way: Using Gait Analysis to Detect Parkinson's Disease

Akanksha Cheeti, Meghana Chava, Swagarika Jaharlal Giri

Abstract

Parkinson's disease is a progressive neurological disorder that results in uncontrollable movements due to dysfunction in the brain, characterized by symptoms such as tremors, rigidity, and difficulty with coordination and balance. Gait analysis can be used to quantitatively evaluate and monitor these motor symptoms for diagnosis, treatment planning, and tracking disease progression. In this study, we analyzed three binary classification datasets using frequency domain transformation, coefficient of variability, and statistical feature extraction techniques. We applied machine learning methods such as SVM, logistic regression, decision trees, K-nearest neighbors, and multilayer perceptron to determine the most effective feature transformation method. We further utilized ensemble and attention-based models to evaluate the efficiency of the model to classify Healthy Control vs Person with Parkinson's. Additionally, explored the concept of transfer learning by testing our second and third datasets on models trained on the first dataset to establish relationships between different domains of data.

I. Introduction

Parkinson's disease is a progressive neurodegenerative disorder that affects over 10 million people worldwide. It is characterized by a range of motor and non-motor symptoms, including tremors, rigidity, slowness of movement, and postural instability. One of the most common and significant motor symptoms of Parkinson's disease is gait impairment, that can lead to falls, reduced mobility, and increased risk of disability. Quantitative gait analysis using special walkway systems is a helpful way to measure and track problems with walking in people with Parkinson's disease. This objective approach can be useful for identifying issues early on and monitoring how well treatments are working. Additionally, advanced technologies such as wearable sensors and machine learning algorithms can provide new insights into gait impairment and may lead to the development of more effective interventions for individuals with Parkinson's disease.

The motivation for this work is twofold. Firstly, there are currently no existing VR applications for analyzing gait performance and stability in individuals with Parkinson's disease, and the use of VR is still considered an area of ongoing research and development. By simulating different walking conditions and terrains, VR can provide valuable insights into the impact of Parkinson's disease on gait and inform the development of more effective treatments. Secondly, the diagnosis of Parkinson's disease is primarily based on clinical observations and subjective assessments, which can be unreliable and lead to treatment delays. The motivation for using gait analysis to detect Parkinson's disease is to develop more accurate and objective diagnostic tools. Gait analysis can provide a more quantitative and objective measure of gait impairment, allowing for earlier diagnosis and intervention and ultimately leading to better patient outcomes. Additionally, gait analysis can be used to monitor disease progression and evaluate the effectiveness of treatments over time. Ultimately, the development of a reliable gait analysis tool

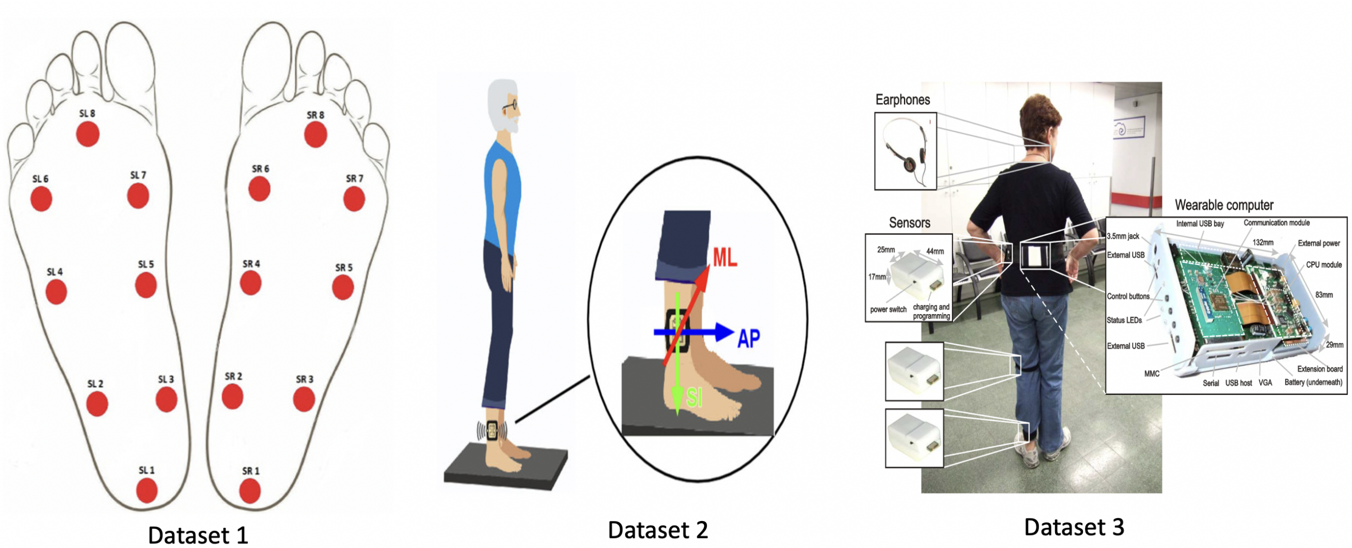

Fig. 1: Datasets

for Parkinson's disease could have significant implications for improving patient care and advancing our understanding of the disease. By combining the potential of VR with gait analysis, this work aims to achieve these goals and improve the lives of those with Parkinson's disease.

To better understand the features that are most effective for gait analysis, we employed a range of machine learning models including decision tree (DT), k-nearest neighbor (KNN), multilayer perceptron (MLP), and support vector machine (SVM). In addition, we utilized advanced techniques such as transfer learning, attention models, and ensemble learning to refine our analysis using multiple datasets. Through this analysis, we were able to identify the key features that best differentiate between healthy individuals and those with Parkinson's disease.

Description	Dataset 1	Dataset 2	Dataset 3
Name	Gait in Parkinson's Disease	Public Data Set of Videos, Inertial Measurement Unit, and Clinical Scales of Freezing of Gait.	Daphnet dataset
Total dataset size	306	71	17
Parkinson Disease	214	42	13
Control	92	29	4
Sensor Placed	Beneath the feet	Ankle	Upper Leg , Ankle and Trunk
Sensor Description	8 sensors underneath each foot that measure the vertical ground reaction force.	Triaxial gyroscope(angular position rate sensor) and a triaxial accelerometer(acceleration sensor) on Ankle.	Triaxial accelerometer(acceleration sensor) on Trunk, Ankle and Upper leg.
Motion	Walking at self-paced speed	Turning-in-place task, alternating 360°turns to their right,then 360°to their left.	Sraight line walking, walking with numerous turns, and finally a more realistic activity of daily living (ADL) task
Time	2 min	2 min	2 min

Fig. 2: Dataset overview

II. Related work

One promising approach for detecting Parkinson's disease is through gait analysis, which involves measuring various aspects of a person's walking pattern. Previous studies like [Sofuwa et al.(2005)] aimed to compare gait characteristics between individuals with Parkinson's disease and a healthy control group using a quantitative gait analysis approach. The study found that individuals with Parkinson's disease exhibited slower gait speed, reduced stride length, increased cadence, and increased double support time compared to the healthy control group. These gait characteristics were associated with disease severity. While, [Bailey et al.(2018)], [McIntosh et al.(1997)] use statistical approaches to test rhythmic auditory stimulation function on PD gait traces and concluded the potential usage of rhythmic auditory stimulation (RAS) as an intervention to improve mobility and reduce fall risk of PD patients. This suggests gait analysis has the potential to revolutionize Parkinson's disease detection and management, improving outcomes and quality of life for millions of people.

III. Dataset

In this work, we worked with three datasets. Figure 1 and Figure 2 gives you an overview of our three Datasets.

Dataset 1: The Gait in Parkinson's disease dataset [Goldberger et al.(2000)] from Physionet will be used as the primary source of data to develop and test the gait analysis algorithm for the early detection of Parkinson's disease.

The dataset contains gait data collected from sensors located under the left and right foot of a person. Each line of data contains 19 columns, including time, vertical ground reaction force (VGRF) on each of the eight sensors located under the left and right foot, and total force under each foot. The X and Y coordinates of each sensor location inside the insole are also provided. These coordinates can be used to calculate a proxy for the location of the center of pressure (COP) under each foot during walking. The dataset can be analyzed to extract meaningful features from the gait data, such as step duration, stride length, walking speed, spectral entropy, and peak frequency, to build machine-learning models for predicting the presence of Parkinson's Disease in new patients based on their gait data.

Dataset 2: A Public Data Set of Videos, Inertial Measurement Unit, and Clinical Scales of Freezing of Gait in Individuals With Parkinson's Disease During a Turning-In-Place Task [Ribeiro De Souza et al.(2022)]. A convenience sample of 35 idiopathic PD patients with FoG(16 females and 19 males) was recruited to participate in this study. The entirety of the dataset is kept in the ASCII format, with each piece of data being contained within a text file named by its corresponding ID and session number. The data within each file consists of a header and 15,360 rows, with each row containing nine columns of information. These columns include 1) frame number, 2) time in seconds, 3) mediolateral accelerometer readings in units of gravitational acceleration, 4) anteroposterior accelerometer readings in units of gravitational acceleration, 5) vertical accelerometer readings in units of gravitational acceleration, 6) mediolateral gyroscope readings in degrees per second, 7) anteroposterior gyroscope readings in degrees per second, 8) vertical gyroscope readings in degrees per second, and 9) a flag indicating the presence or absence of Freezing of Gait (FoG) events, as determined by movement disorder specialists who reviewed video recordings. A value of 0 indicates no FoG event, while a value of 1 indicates the presence of a FoG event.

Dataset 3: The Daphnet dataset[Bachlin et al.(2009)] is a collection of data on Parkinson's disease patients, specifically focusing on Freezing of Gait (FoG), a common symptom of the disease. The dataset includes measurements from inertial measurement units (IMUs) worn by the patients during clinical evaluations, as well as video recordings of the evaluations. The data is stored in ASCII format and includes information such as accelerometer and gyroscope readings, as well as a flag indicating the presence or absence of FoG events, as determined by movement disorder specialists who reviewed the video recordings.

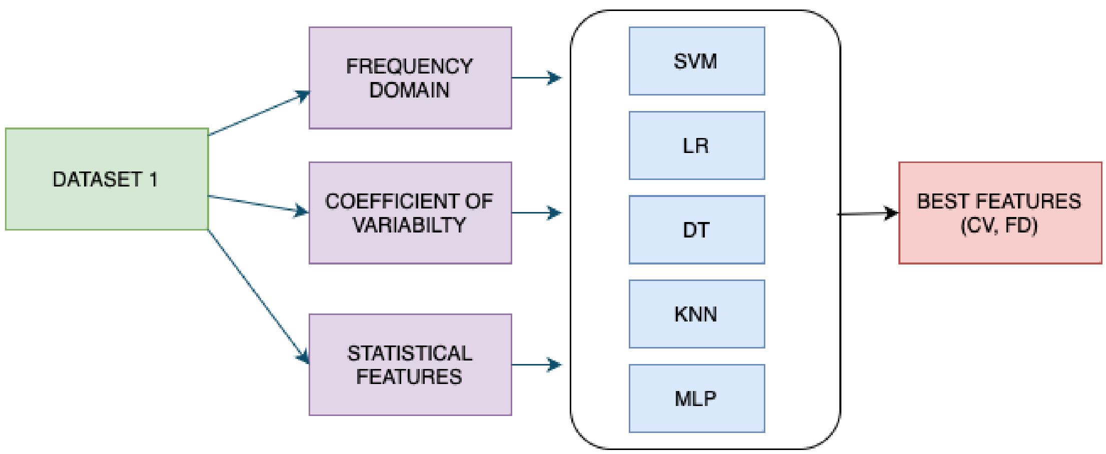

Fig. 3: Feature selection

IV. Methodology

Feature Selection: The dataset used for this project had sensor recordings for each of the patients, stored in separate text files. In order to combine the patient details into a single dataset, we extracted features from the sensor data using methods like frequency domain analysis, coefficient of variability, and statistical analysis.

We tested the performance of these feature extraction methods separately using various machine learning models, including Support Vector Machine, Logistic Regression, Decision Trees, and MultiLayer Perceptron. After analyzing the results, we found that the frequency domain and coefficient of variability features provided the best accuracy. As a result, we decided to use these features for our prediction tasks.

Model Selection for Dataset1: We aim to predict if a person has Parkinson's disease or not. We proposed a high-level architecture shown in Fig 4. It involves using the Frequency domain and Coefficient of variability features. As the features are in different sizes, we further performed PCA Dimensionality reduction to embed the features into one layer and then we merged the datasets.

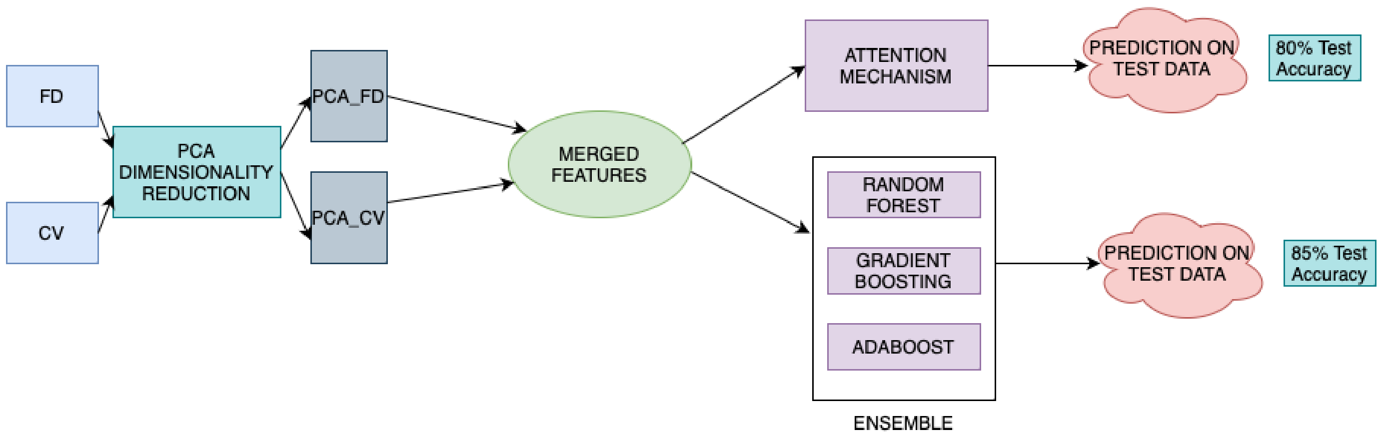

Fig. 4: Proposed Architecture for Dataset 1

We used two mechanisms namely, the attention-based mechanism and the Ensemble technique. The attention-based mechanism helped us focus on important parts of the data and ignores the less important ones. It resulted in better accuracy. In the ensemble model, we combined multiple machine learning models including Random Forest, Gradient Boosting, and AdaBoost, we were able to capture different patterns in data. After evaluating the performance, we got good prediction results on test data for both models.

V. Results

Accuracy	DT	KNN	MLP	Attention	Ensemble
All Sensors	69%	72.2%	77.7%	79.6%	85.18%
Side only	63%	70.3%	75.7%	74.6%	84.88%
Exclude Diagonal	59%	68.51%	73.7%	70.9%	81.18%

Fig. 5: Results on Dataset 1

Evaluation results on Dataset 1:
In this section, we present the prediction accuracies of various machine learning models trained on all the sensors, as well as on subsets of sensors. Specifically, we investigated the impact of using two subsets of sensors: (1) side-only sensors which exclude top and bottom sensors(Excludes 1,8,9,16) and (2) Excluding diagonal sensors: including 2, 3, 6, and 7 sensors on the left and right foot. The results are shown in figure 5.

We observed that the model performed well when using all the sensors, achieving high accuracy in predicting Parkinson's disease. However, when using only the side-only sensors, the accuracy dropped slightly as we ignored four sensors and the model's accuracy was the least when excluding diagonal sensors. Although using all sensors was optimal, the side-only sensors showed reasonable accuracy in predicting Parkinson's disease. In conclusion, our experiments suggest that while using all sensors provides the best accuracy, using only side-only sensors can also be a feasible option to a certain extent.

Evaluation results on Merged Dataset 2 and Dataset 3
We observed a commonality between Dataset 2 and Dataset 3. Our dataset 2 had turn-in-place task data and our dataset 3 had the combination of various walking tasks like walking in a straight line, walking with numerous turns, and walking while doing various day-to-day activities. We observed that both datasets had ankle acceleration sensor readings in common. So we combined both datasets to have a common dataset with walking tasks and turn-in-place tasks.

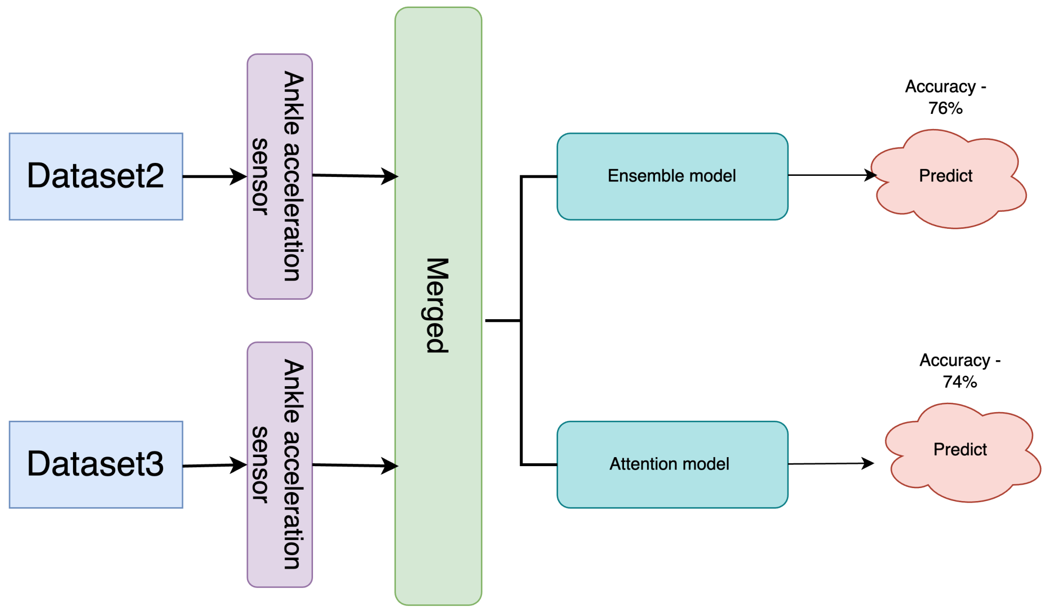

Fig. 6: Results on Merged dataset

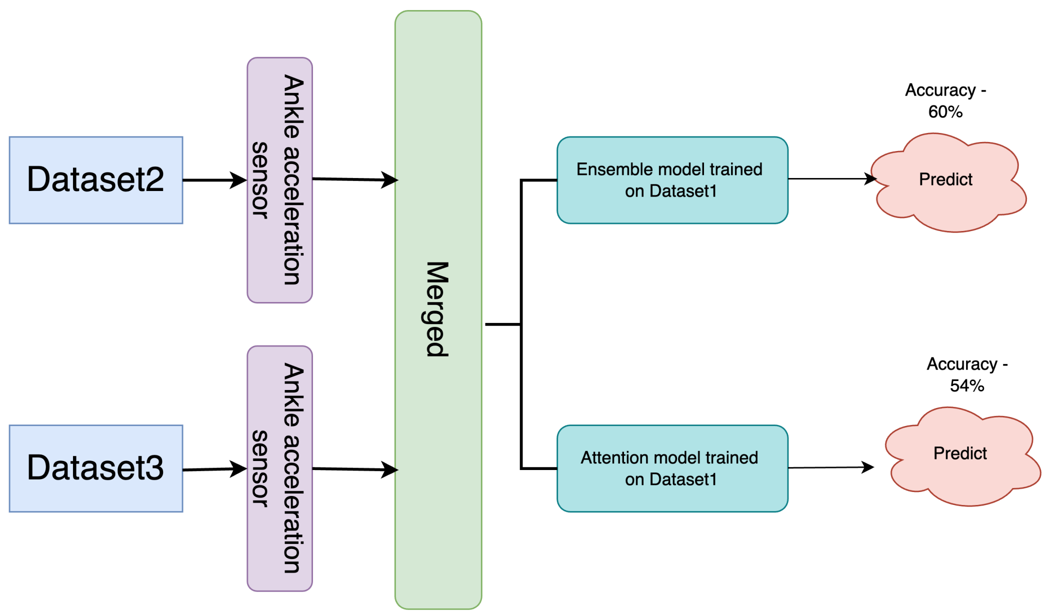

Fig. 7: Transfer learning on Merged dataset

We selected our top two models the Ensemble model and the Attention model for evaluation. We got a result of 74% on the Ensemble model and 72% on the Attention model.

We also wanted to try the idea of transfer learning. Transfer learning is a machine learning technique that involves leveraging knowledge learned from one task or domain to improve performance on another task or domain. In our case, we have trained a model on sensor data collected beneath the foot sensors, and we wanted to test it on data from a triaxial accelerometer for Parkinson's disease. By leveraging knowledge learned from the foot sensor data, we are attempting to improve the performance of the model on the new task. For our merged dataset, we got a result of 60% on the Ensemble model that was originally trained on Dataset 1 and 56% on the Attention model that was originally trained on Dataset 1.
This suggests that there might be some relation between the sensor data collected from beneath the foot and the ankle

acceleration sensor data.

VI. Virtual Reality(VR) Application using Gait analysis data

These personalized VR programs utilize the patient's gait analysis data to create an immersive and adaptive environment that can improve their gait, balance, and overall quality of life. Here are a few applications,

1) Virtual reality gait training: Create immersive simulations for personalized gait training exercises.
2) Virtual reality assessment: Assess balance and movement control using VR simulations.
3) Virtual reality rehabilitation games: Design games to encourage gait improvement with real-time performance feedback.
4) Virtual reality home exercise program: Develop personalized home exercise programs with feedback and adaptation capabilities.
5) Virtual reality visual feedback: Provide 3D visual feedback to correct errors and improve overall gait quality.

References

[Bachlin et al.(2009)] Marc Bachlin, Meir Plotnik, Daniel Roggen, Inbal Maidan, Jeffrey M Hausdorff, Nir Giladi, and Gerhard Troster. 2009. Wearable assistant for Parkinson's disease patients with the freezing of gait symptom. *IEEE Transactions on Information Technology in Biomedicine* 14, 2 (2009), 436–446.

[Bailey et al.(2018)] Christopher A Bailey, Federica Corona, Mauro Murgia, Roberta Pili, Massimiliano Pau, and Julie N Côté. 2018. Electromyographical gait characteristics in Parkinson's disease: effects of combined physical therapy and rhythmic auditory stimulation. *Frontiers in neurology* (2018), 211.

[Goldberger et al.(2000)] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *circulation* 101, 23 (2000), e215–e220.

[McIntosh et al.(1997)] Gerald C McIntosh, Susan H Brown, Ruth R Rice, and Michael H Thaut. 1997. Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* 62, 1 (1997), 22–26.

[Ribeiro De Souza et al.(2022)] Caroline Ribeiro De Souza, Runfeng Miao, Júlia Ávila De Oliveira, Andrea Cristina De Lima-Pardini, Débora Fragoso De Campos, Carla Silva-Batista, Luis Teixeira, Solaiman Shokur, Bouri Mohamed, and Daniel Boari Coelho. 2022. A public data set of videos, inertial measurement unit, and clinical scales of freezing of gait in individuals with parkinson's disease during a turning-in-place task. *Frontiers in Neuroscience* 16 (2022), 832463.

[Sofuwa et al.(2005)] Olumide Sofuwa, Alice Nieuwboer, Kaat Desloovere, Anne-Marie Willems, Fabienne Chavret, and Ilse Jonkers. 2005. Quantitative Gait Analysis in Parkinson's Disease: Comparison With a Healthy Control Group. *Archives of Physical Medicine and Rehabilitation* 86, 5 (2005), 1007–1013. https://doi.org/10.1016/j.apmr.2004.08.012