Machine Learning Models for Parkinson's Disease Gait Assessment and Medication Adherence from Smartphone Sensor Data

By

HAMZA ABUJRIDA

A Dissertation Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering

at

WORCESTER POLYTECHNIC INSTITUTE

December 2022

Professor Emmanuel O. Agu, Advisor

Professor Kaveh Pahlavan, Committee Member

Professor Xinming Huang, Committee Member

Professor Adam C. Lammert, Committee Member

ABSTRACT

Parkinson's Disease (PD) is a neurodegenerative chronic disorder with multiple motor and non-motor symptoms. People afflicted with Parkinson's Disease experience severe problems with performing daily activities including their gait (the way a person walks), which frequently links to a less steady walk, that arises from changes in posture, slowness of movement (bradykinesia), and a shortened stride. This distinctive walk is called 'Parkinsonian gait. When a PD patient develops a Parkinsonian gait, they start to experience festination: progressively shorter but accelerated steps forward, often in a shuffling manner. Other symptoms include slowness of gait, hesitation of starting gait aka Freeze of Gait (FoG), difficulty making turns, and postural instability leading to frequent falls. Some features of Parkinsonian gait are likely to become more pronounced over time, particularly festination, stooped posture, and FoG. PD patients feel unsteady and lose confidence because of the fear of falling. Consequently, their social activities and their quality of life get severely impacted.

As PD has no ultimate cure, physicians aim to delay PD complications, especially those that degrade the patient's quality of life such as motor symptoms and dyskinesia. Patients' lack of adherence to prescribed medication is a major challenge for physicians, especially for patients suffering from chronic conditions. The Centers for Disease Control and Prevention (CDC) estimates that medication non-adherence causes 30 to 50 percent of chronic disease treatment failures and 125,000 deaths per year in the USA [119]. In PD patients particularly, adherence varies between 10% and 67% [120].

Since changes in PD gait can be a good measure for inferring the progression and severity of the disease to inform early intervention, gait has been part of the motor section of the Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS). MDS-UPDRS motor part is mostly assessed by a professional clinician. To reduce the cost of care, remote assessment in patients' homes has recently become an alternative tool for monitoring the progression of Parkinson's disease (PD). Smartphones, particularly, provide an affordable, accessible, and easy-to-use platform for PD gait sensing. Smartphones are ubiquitous, portable, and user-friendly. Equipped with triaxial Accelerometers and gyroscopes in addition to powerful CPUs, smartphones offer the potential for remote gait assessment in the patient's home environment.

In this dissertation, to facilitate an accurate remote Parkinsonian gait assessment, we proposed and rigorously evaluated a novel Deep-Learning (DL) based gait analysis system that assesses the severity of PD gait based on 30-seconds walk given by the patient, facilitating home-based clinical monitoring by remotely assessing the PD patient gait. Specific preprocessing steps were utilized including the calculation of the moving average, subtracting signal mean, and detection of gait strides. These techniques resulted in smoothing the signal, filtering of noise, and cancellation of gravity/breeding effects. These steps facilitate DL automatic feature extraction and eliminated the need for any kind of signal conversion.

The most significant contribution of our work is the proposal of a deep-learning-based system that comprehensively classifies 3 PD symptoms: the severity of FoG, walking imbalance, and shaking/tremors from data gathered in one study. Prior work has trained and tested separate models to analyze each of these PD gait anomalies separately, the model we introduced is a single model that achieved impressive results for all of the PD gait symptoms. This was challenging because the model's parameters had to be jointly tuned in order to establish relationships with different sets of PD symptom labels, all while using the same dataset as an input. To achieve the ultimate results, we investigated four different approaches based on multiple Machine Learning (ML) algorithms. The first approach employs the extraction of hand-crafted features as input to Machine learning algorithms, we conducted supervised classification experiments using 10-fold cross-validation and measured the performance of different models. In the second approach, we encoded the walking signal to an image format using the Gramian Angular Field (GAF) encoder. We employed the concept of Transfer Learning on the top image-based models such as ResNet50, Inception, SqueezeNet, and EfficientNet. The third approach employed variations of the Long Short Term Memory models, we investigated the simple LSTM, CNN-LSTM, and parallel LSTM models. These first three models had limitations, necessitating research and development of a fourth method. The first ML approach, could not achieve an acceptable performance when classifying various walks, mainly because the handcrafted features were not able to linearly or non-linearly discriminate between the different classes. The second and Third approaches suffered data overfitting, because of the models' over-complexity, which could not be justified by our dataset, due to a large number of trainable parameters.

Consequently, a DL multi-layer Conventional Neural Network (CNN) model was introduced, this model operates on 1Dimensional convolution filters to classify 30 seconds of walking data into one of five severity levels. Our DL network was able to Classify the PD Walking-Balance, Shaking/Tremor, and Freeze of Gait (FoG) symptoms, with an accuracy of: 99.1%,98.4%, and 98.2% respectively.

Another important contribution of this work is the model's ability to discriminate between PD patients on- vs off-medication and baseline HC walk. Unlike methods such as Drug-Bottles and urine or blood test that monitor discrete medication-related events, our approach analyzes data corresponding to continuous windows of time, submitted by PD patients every time they walk before/after taking their medication. By training our model on walking segments recorded before and after medication, we were able to present a medication adherence system that operates with an accuracy of 98.2%, which facilitates remote medication adherence. Finally, our DL-based gait analysis system was successfully applied to more than 450 participants from the independent dataset (mPower dataset). This system is proven to be applicable in home environments and capable of providing an accurate PD gait assessment in a telemedicine fashion.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1
INTRODUCTION

Background and Motivation

1.1.1 Parkinson's Disease

Parkinson’s disease is the most common neurodegenerative movement disorder with a worldwide prevalence estimated at 16 million people **[1]**. This number is expected to double by 2050. Symptoms of PD become noticeable when the brain cells that produce dopamine begin to die off. Dopamine, which works as a neurotransmitter, plays an important role in the control and coordination of the human brain. With less dopamine, the human brain loses its ability to control movements, leading to tremor, stiffness, and muscle pain **[2]** as shown in Figure (1.1). PD symptoms start as slight tremors, usually in one of the hands, then gradually increase to include a noticeable tremor, rigidity, akinesia, and postural instability.

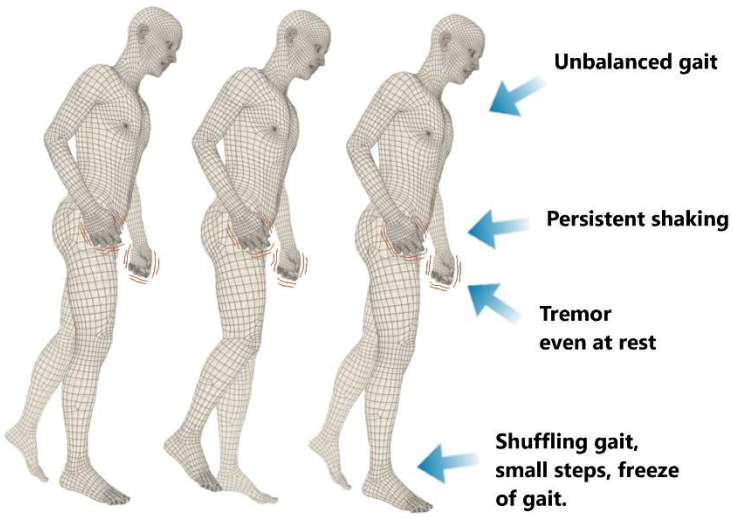


Figure 1.1: Symptoms of Parkinson’s disease

Some features of Parkinsonian gait are likely to become more pronounced over time, particularly festination, stooped posture, and Freeze of Gait (FoG). PD patients feel unsteady and lose confidence because of the fear of falling. Consequently, their social activities and their quality of life get severely impacted.

Today there is no therapeutic treatment for PD. However, there are drugs that mitigate the symptoms. L-dopa drug for instance is commonly used by many PD patients to enhance the brain's supply of dopamine and reduce the severity of PD symptoms when taken regularly. As PD progresses, gait symptoms can be grouped into episodic disturbances of gait initiation including FoG. And continuous disturbances of the step-to-step dynamics that cause shuffled and unbalanced gait. The severity of these symptoms helps physicians decide the overall stage of PD, and hence recommend the right medication dose.

PD care and treatment are costly. In general, hospitalization is the largest component of PD health system costs (69% of total costs) **[3]**. Because the condition of PD patients mostly improves at the early stages of the disease, it is important to detect PD as early as possible and monitor its progression in the early stages. When PD progresses to a severe state, patients lose their ability to walk, talk and can experience depression, fatigue, and memory loss, resulting in a lack of adherence to prescribed medication. Approximately only 67% of PD patients were found to be adherent to their PD medications **[4]** with the level of non-adherence increasing as the daily dosage increases **[5].** Non-adherence to medication raises the mean annual medical cost to $15,826 compared to $9,228 for adherent PD patients (71% increase). Non-adherence also leads to more hospitalization (2.3 vs. 1.8) and office visits (17.0 vs. 15.9) **[6]**, which all increase the burden of PD complications on PD patients.

1.1.2. Problems Addressed by this Dissertation

This dissertation tackled 2 gait-related problems using smartphone sensor data.

- Problem 1: Remote PD assessment (Walking Balance, Shaking/Tremor and FoG) from gait

  Remote assessment of gait in patients' homes has become a valuable tool for monitoring the progression of Parkinson's disease (PD) and to reduce the cost of care resulting from frequent hospital visits and inpatient days. However, these measurements are often not as accurate or reliable as clinical evaluations because it is challenging to objectively distinguish the unique gait characteristics of PD. We explore the inference of patients' stage of PD from their gait using machine learning analyses of data gathered from their smartphone sensors. Specifically, we investigate supervised machine learning (ML)

models to classify the severity of the motor part of the UPDRS (MDS-UPDRS 2.10-2.13). Our goals are to facilitate remote monitoring, infer the patient PD stage based on their gait, and to find out which ML classifier types can discriminate the severity of PD gait anomalies.

- Problem 2: Medication adherence inference from gait.

Patients' lack of adherence to prescribed medication is a major challenge for physicians, especially for patients suffering from chronic conditions. The Centers for Disease Control and Prevention (CDC) estimates that medication non-adherence causes 30 to 50 percent of chronic disease treatment failures and 125,000 deaths per year in the USA [119]. In PD patients particularly, adherence varies between 10% and 67% [120].

The goal of this dissertation is to remotely determine whether PD patients have taken their medication, by analyzing gait data gathered from their smartphone sensors. Using this approach, physicians can track the level of medication adherence of their PD patients

1.1.3 MDS-UPDRS RATING SCALE

The most widely used rating scale for PD is the Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS) **[7**]. The MDS-UPDRS was developed to evaluate various aspects of Parkinson's disease including non-motor and motor experiences of daily living. It includes a motor evaluation and characterizes the extent and burden of disease across various populations. The scale can be used in a clinical setting as well as in research. On this scale, PD anomalies are rated on a scale of zero (normal) to four (severe PD). The MDS-UPDRS features sections that require independent completion by people with Parkinson's and their care-givers, and sections to be completed by the clinician. The main sections are:

- Part 1: non-motor experiences of daily living.
- Part 2: motor experiences of daily living.
- Part 3: motor examination.
- Part 4: motor complications.

In our work, we use the MDS-UPDRS as our primary PD measure to gauge PD gait severities. Table (1.1) shows the gait-related questions utilized in this dissertation.of the motor part of MDS-UPDRS.

Table 1.1: The MDS-UPDRS gait questions utilized

Question	Variable name	Variable details
Over the past week, have you usually had shaking or tremor?	MDS-UPDRS2.10	one of: {'Normal', 'Slight', 'Mild', 'Moderate', 'Severe'} mapping to {0, 1, 2, 3, 4}
Over the past week, have you usually had problems with balance and walking?	MDS-UPDRS2.12	one of: {'Normal', 'Slight', 'Mild', 'Moderate', 'Severe'} mapping to {0, 1, 2, 3, 4}
Over the past week, on your usual day when walking, do you suddenly stop or freeze as if your feet are stuck to the floor?	MDS-UPDRS2.13	one of: {'Normal', 'Slight', 'Mild', 'Moderate', 'Severe'} mapping to {0, 1, 2, 3, 4}

1.1.4 Gait Cycle

The natural human gait cycle starts with the contact of one foot and ends with new contact of the same foot with the ground as shown in Figure (1.2). Each cycle consists of stance and swing phases. The stance phase is when the foot is in contact with the ground. Swing, is as the name implies, refers to the period when the foot is airborne. It can be seen from Figure (1.2) that one gait stride contains two steps. Step length is the distance between the point of initial contact of one foot and the point of initial contact of the opposite foot.

Various gait features can be extracted directly from the gait cycle, including step width, stance time, Step time, and Gait velocity. Gait features can capture the aspects of PD gait, and can highlight the severity of gait anomalies compared to a regular Healthy Control (HC).

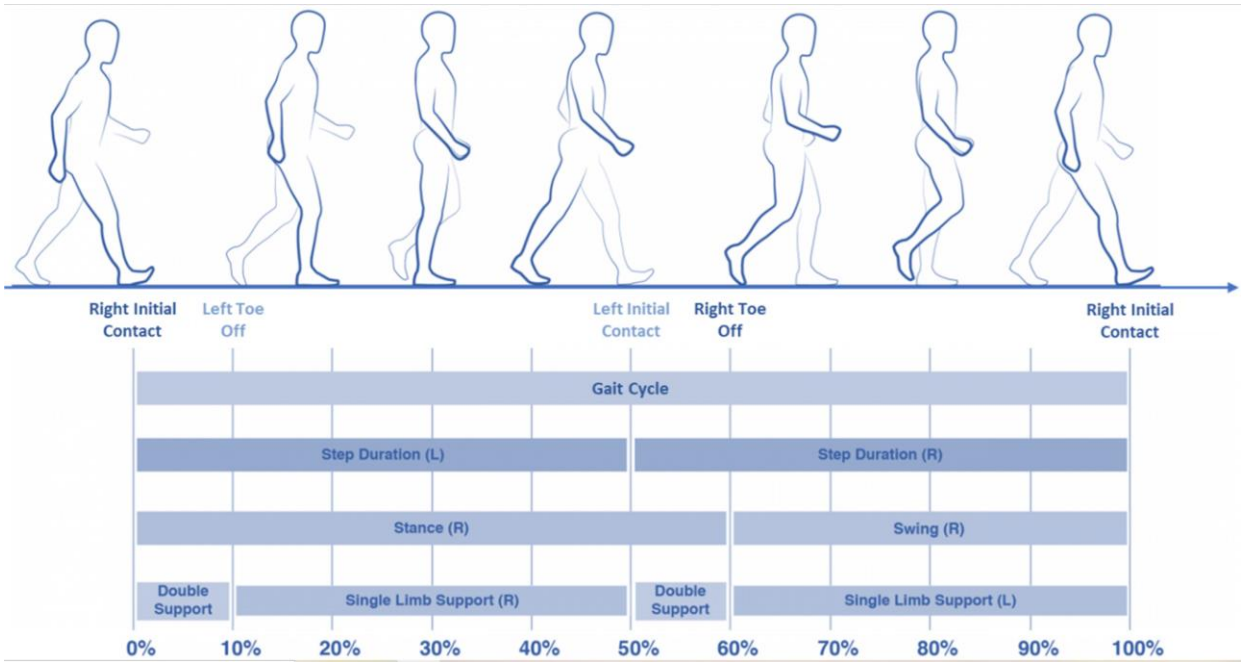


Figure 1.2: Gait Cycle

1.1.5 The Need for Remote Monitoring

The majority of PD patients are cared for by an informal caregiver, usually a family member, the physical, financial, and mental work required by the caregiver is significant [121]. The medical community faces challenges as well, in terms of PD diagnostics and delivering/adjusting the therapy. With the increase in the number of people living with PD [1], existing methods of managing PD for both the caregiver and medical community will not scale up with the challenges, therefore remote monitoring and intervention are required now more than ever.

Remote methods not only monitor patients' gait and medication adherence continuously but also can be used to remind them to take their medications on time. There are multiple methods for effective PD remote monitoring. Sensor-rich wearables and widely owned smartphones, however, stand out and present an opportunity as viable platforms for remote monitoring and PD medication adherence.

1.1.6 Smartphone Sensing

Smartphones are now owned by over 85% of the US population **[8]**. With this high adoption rate, smartphones can be used to sense the behavior and health of the smartphone user. Moreover, smartphones contain powerful processors for analyzing sensor data, making them a unique portable platform for mobile health applications. Researchers can conduct large-scale studies on millions of participants with the use of smartphone sensors to capture the fingerprint of the patient's unique behaviors. Smartphone sensing for healthcare can facilitate remote assessment, follow-up and enable doctors to fine-tune medication to meet each patient's needs. Several smartphone sensors have been demonstrated to be useful in assessing the symptoms of PD, including the use of microphone for voice analysis, accelerometer/gyroscope for gait assessment, and screen tapping for tremor/shaking evaluation **[9-14].** Mobile health applications could play a key role in reducing healthcare costs and the burden of PD on patients, especially those living in remote areas.

Home capture yielded an increased number of participants compared with similar studies that were performed in clinics. However, the home collection method has it is own challenges **[66].** Sources of error include factors such as the variance of different devices/sensors used, the lack of expert proficiency in subjects' self-assessment of gait severities, variability of the environment in which the assessment was performed, and level of subject adherence to the smartphone app instructions, which dramatically affects what subjects record as an observation. These factors ultimately lead to inconsistency in analyzing and classifying each activity and increase the margin of prediction error. PD data collected in the home environment is more realistic. But it is more confounded by noise and is more challenging to analyze than data collected in a lab/clinic or controlled environment. Any algorithm that analyzes a home-gathered dataset has to identify which records contain valid gait data. Another major challenge for home-based datasets is the chance of data mislabeling. Patient or their unofficial caregiver may erroneously label a walking record with a mild severity label instead of moderate or moderate instead of severe.

Goals

In this dissertation, multiple ML models were investigated. The models explore the inference of PD gait by analyzing the patients' Smartphone walk data of 30 seconds into one of five severity levels (Normal, Slight, Mild, Moderate, Severe). Following the traditional classification approach, our model aims to extract discriminative features from the raw accelerometer and gyroscope data without the need for the domain expertise required to extract handcrafted gait features. We aim to use a patient-friendly approach, which allows patients, with the help of their caregiver, to self-record their walk in the convenience of their home. The patient then uploads the walk record for offline gait analysis.

The goal of PD gait analysis is the ability to assess disease severity remotely and help the patient adhere to prescribed medication by differentiating the gaits signals recorded before and after taking medication. By experimenting with different ML algorithms, we aim to identify which machine learning classification model best distinguishes the severity of PD anomalies for motor aspects of the MDS-UPDRS.

Given the gait signal from the inertia accelerometer and gyroscope sensors, and the signal magnitude, the patient walk signal can be expressed as:

$$X = [\alpha x\,(i),\ \alpha y\,(i),\ \alpha z\,(i),\ MagNG\alpha(i),\ \omega x\,(i),\ \omega y\,(i),\ \omega z\,(i),\ MagNG\omega(i)]T$$

where i denotes discrete-time, α indicates acceleration, ω represents rotation and T denotes the Stride length. Our goal is to classify the accelerometer and gyroscope input X into one of five severity levels (Normal, Slight, Mild, Moderate, Severe), we will express those classes as $Y = (y_1, y_2, y_3, y_4, y_5)$, The output of our model can be expressed as $o_{(i)} = P(y_i \mid X)$, or the probability of $(X \in y_i)$. If the gait severity of X is y, then y can be expressed as:

$$y = \arg max_{y(i)} \{o_{(i)} \mid 1 \le i \le 5\} \qquad (1)$$

As can be seen from equation (1) the class with the maximum probability will be the output of the model.

Desirable attributes of our proposed smartphone-based system include:

- **Low cost:** the wide adoption of smartphones minimizes the cost of our approach to patients and their caregivers and eliminates the need for acquiring dedicated hardware such as infrared cameras, motion sensors, or any purpose-built hardware.
- **Continuous Remote Monitoring**: Unlike methods such as smart drug bottles and urine or blood test that monitor discrete medication-related events episodically from drug bottles or blood samples respectively, our approach analyzes data corresponding to continuous windows of time, submitted by PD patients every time they walk before/after taking their medication.
- **Minimal burden while recording patient data in the home setting:** walking is a simple activity that presents a minimal burden to PD patients for data collection.
- **Anomalous gait detection:** our system not only differentiates walks before/after medication but can also discriminate between regular walks for HC and the severity of PD walks. This can help with the early diagnosis of PD gait.

Challenges

1. *Real-world gait data is noisy:* PD data collected in the home environment is more realistic, as people act naturally without any proctor restrictions. However, such data is more noisy and is more challenging to analyze than data collected in a lab/clinic or controlled environment. Noise examples includes Contradictory examples (examples of different class labels for same subject), attribute noise (Erroneous, missing or in-complete values) in addition to any variations of the sensors signals due to the difference in the smartphone models. Self-assessment has its many challenges and might not always generate reliable results. Any algorithm that analyzes a home-gathered dataset has to first identify which records contain valid gait data among several other non-valid records.
2. *Patients may misinterpret instructions:* Another challenge is the chance of misinterpreting the task, while they are asked to walk 30 seconds in a straight line, PD patients may choose not to follow the instructions, which results in an invalid gait record.
3. *Provision of erroneous PD ground truth labels:* Another major challenge for home-based datasets is the possibility of data mislabeling. Patient or their unofficial caregiver may Underestimate, overestimate, or forget to estimate and label the walking record. Therefore,

they can erroneously label a walking record with a mild severity label instead of moderate or moderate instead of severe.

4. *Imbalanced datasets:* Having participants signing up online will often lead to imbalanced dataset classes, as the number of participants signing up with severe and moderate severities is often less than the particpants with mild and slight severities. Imbalanced dataset can lead to bias in the classification. Models that work on such datasets have to be prone to ML bias to perform well.

Organization of Dissertation

The material presented in Chapters 2 to 6 is roughly organized based on the overall approach we are presenting for the PD gait analysis, as briefed in the following sections.

1.4.1 Chapter 2: Related Work

In chapter 2 we discuss the Related Work. Prior work has explored various methods to analyze PD gait. In General, prior methods can be classified into two main categories, wearables, and non-wearables. Wearable sensors utilized in prior work include inertia sensors, portable-wearable sensors, pressure sensors, and electromyography. Non-wearable sensors include floor sensors, image-processing including cameras, and stereoscopic vision. Since our approach utilizes smartphones, we also present the studies that used smartphone technology analyze PD gait .

1. 4.2 Chapter 3: Background on Data Gathering,StudyDataset and Data Preprocessing

Chapter 3 discusses the MPower PD dataset that we used to evaluate our proposed ML/DL approaches, the criteria we used to select participants and the walking records to be included in our evaluation. We also discuss the signal processing techniques that we implemented to pre-process the walking signal and prepare it for both feature extraction and raw DL data analysis. In addition to that, we present the methods we used to segment the walking signal

using time-based and Stride-based segmentation. Lastly, we describe the algorithm utilized to convert the walking signal to an image for image-based DL investigation.

1. 4.3 Chapter 4: Machine Learning Gait analysis

In chapter 4 we explore the inference of patients' stage of PD from their gait using traditional machine learning classification algorithms. Specifically, we investigate supervised machine learning (ML) models to classify the severity of the motor part of the UPDRS (MDS-UPDRS 2.10-2.13). Our goals are to facilitate remote monitoring and to answer the following questions: 1) What is the patient PD stage based on their gait? 2) Which features are best for understanding and classifying PD gait severities? 3) Which ML classifier types best discriminate PD patients from healthy controls (HC)? and 4) Which ML classifier types can discriminate the severity of PD gait anomalies?

From the accelerometer and gyroscope sensor data, statistical, time, wavelet, and frequency domain features were extracted, as well as other lifestyle features derived directly from participants' survey data. We conducted supervised classification experiments using 10-fold cross-validation and measured the model precision, accuracy, and area under the curve (AUC). We found that the best classification model, best feature, highest classification accuracy, and AUC were 1) random forest and entropy rate, 93% and 0.97, respectively, for walking balance (*MDS-UPDRS-2.12*); 2) *bagged trees* and MinMaxDiff, 95% and 0.92, respectively, for shaking/tremor (*MDS-UPDRS-2.10*); 3) *bagged trees* and entropy rate, 98% and 0.98, respectively, for freeze of gait; and 4) random forest and MinMaxDiff, 95% and 0.99, respectively, for distinguishing PD patients from HC.

1. 4.5 Chapter 5: Deep Learning Gait analysis (DeePaGait)

In chapter 5 we introduce DeePaGait, similar to DeePaMed, it is a data-driven neural network model, that focuses on the inference of PD gait by analyzing the patients' Smartphone walk data. DeePaGait consists of a multilayer 1D-CNN, that classifies 30 seconds of data into one of five severity levels (Normal, Slight, Mild, Moderate, Severe). With 1D CNN, DeePaGait follows a similar process to image classification models. DeePaGait extracts discriminative

features from the raw accelerometer and gyroscope data without a need for comprehensive data conversion or domain expertise for gait features extraction. DeePaGait uses a patient-friendly approach, which allows patients, with the help of their caregiver, to self-record, and label their gait data in the convenience of their home. The patient then uploads the walk record for offline DeePaGait analysis. Our DeePaGait DL network was able to Classify the PD Walking-Balance, Shaking/Tremor, and Freeze of Gait (FoG) symptoms, with an accuracy of: 99.1%,98.4%, and 98.2% respectively.

1. 4.4 Chapter 6: Deep learning-Based Medication adherence (DeepaMed)

Chapter 6mainly describes DeePaMed, a data-driven neural networks methodology, that uses the Smartphone's built-in accelerometer and gyroscope sensors to evaluate PD patients' medication adherence and response to medication based on their gait. DeePaMed consists of a multilayer Conventional Neural Network (CNN), that analyzes smartphone gait data. DeePaMed autonomously extracted discriminative features from raw triaxial accelerometer and gyroscope gait data. One-dimensional (1d) filters were used to extract gait features from individual signal components, and multi-dimensional filters captured overall signal variations. As DeePaMed runs on patients' smartphone data, its simplicity and low cost facilitate monitoring of medication adherence of PD patients by having them walk at their homes before and after taking medication. While walking, the smartphone in the PD patients' pockets seamlessly records sensor data for offline DeePaMed analysis.

Our DeePaMed model was able to discriminate PD patients on- vs off-medication and baseline HC walk with an accuracy of **98.2%**. The accuracy of our CNN model surpassed that of traditional Machine Learning methods by over **17%**. We also found that our model performed best with inputs containing a minimum of 10 full gait strides.

The ML approach discussed in Chapter (4) employs the extraction of hand-crafted features as input to Machine learning algorithms, we conducted supervised classification experiments using 10-fold cross-validation and measured the performance of different models. In the DL approach (Chapters 5-6), we encoded the walking signal to an image format using the Gramian

Angular Field (GAF) encoder. We employed the concept of Transfer Learning on the top image-based models such as ResNet50, Inception, SqueezeNet, and EfficientNet. The third approach (Chapters 5-6) employed variations of the Long Short Term Memory models, we investigated the simple LSTM, CNN-LSTM, and parallel LSTM models. These first three models had limitations, necessitating research and development of a fourth method. The first ML approach, could not achieve an acceptable performance when classifying various walks, mainly because the handcrafted features were not able to linearly or non-linearly discriminate between the different classes. The second and Third approaches suffered data overfitting, because of the models' over-complexity, which could not be justified by our dataset, due to a large number of trainable parameters.

Consequently, a DL multi-layer Conventional Neural Network (CNN) model was introduced (Chapters 5-6), this model operates on 1Dimensional convolution filters to classify 30 seconds of walking data into one of five severity levels. Our DL network was able to Classify the PD Walking-Balance, Shaking/Tremor, and Freeze of Gait (FoG) symptoms, with an accuracy of: 99.1%,98.4%, and 98.2% respectively.

Contributions

Our main contributions can be summarized in the following points:

- **Proposed DeePaGait**: a data-driven neural network model, that focuses on the inference of PD gait by analyzing the patients' Smartphone walk data. DeePaGait consists of a multilayer 1D-CNN, that classifies 30 seconds of data into one of five severity levels (Normal, Slight, Mild, Moderate, Severe). Our DeePaGait DL network was able to Classify the PD Walking-Balance, Shaking/Tremor, and Freeze of Gait (FoG) symptoms, with an accuracy of: 99.1%,98.4%, and 98.2% respectively.
- **Proposed a neural networks model that comprehensively classifies 3 PD symptoms: the severity of FoG, walking imbalance, and shaking/tremors from data gathered in the same study:** While prior work has trained and tested separate models to analyze each anomaly of PD gait. DeePaGait achieved impressive results for all 3 PD symptoms. This was challenging because DeePaGait needed to tune the

network parameters to correlate with a different set of labels per PD symptom, all while using the exact dataset as an input.

- **Introduced DeePaMed:** a data-driven neural networks methodology, that uses the Smartphone's built-in accelerometer and gyroscope sensors to evaluate PD patients' medication adherence and response to medication based on their gait. DeePaMed consists of a multilayer Conventional Neural Network (CNN), that analyzes smartphone gait data. Our DeePaMed model was able to discriminate PD patients on- vs off-medication and baseline HC walk with an accuracy of 98.2%. The accuracy of our CNN model surpassed that of traditional Machine Learning methods by over 17%. We also found that our model performed best with inputs containing a minimum of 10 full gait strides.
- **Proposed task-specific preprocessing steps for efficient use of raw sensor data:** Preprocessing included calculation of the moving average, subtracting signal mean, and detection of gait strides. These techniques resulted in smoothing the signal, noise filtering, and cancellation of gravity/breeding effects. These steps facilitate DL automatic feature extraction and eliminated the need for any kind of signal conversion.
- **1D CNN for analyzing multiple aspects of PD gait:** While one-dimension CNN has been proved to be effective on many diverse tasks, to the best of our knowledge, ours is the first work at this scale to analyze multiple aspects of PD gait, using a 1D CNN network.

CHAPTER 2
RELATED WORK

Prior work explored various methods to analyze PD gait. Based on the sensor's placement, gait analysis methods can be classified into two main categories, wearables, and non-wearables. In each category, a variety of sensors and methodologies have been investigated such as smartphones, smartwatches, infrared cameras, and force plates. We will summarize each subcategory in the next sections. For completeness, we reviewed prior work related to medication adherence, walking balance, and the work around ML features, especially the lifestyle features.

2.1 Wearable Gait analysis

Wearable sensors, shown in Figure (2.1), utilized include inertia sensors **[15],** portable-wearable sensors **[16],** pressure sensors **[17],** and electromyography **[18]**. Wearable sensors are relatively cost-friendly. But most of them are not convenient for the patients and require proctoring in a clinical environment. Wearable sensors need to be purchased and worn by the patient which requires a commitment from the patient.

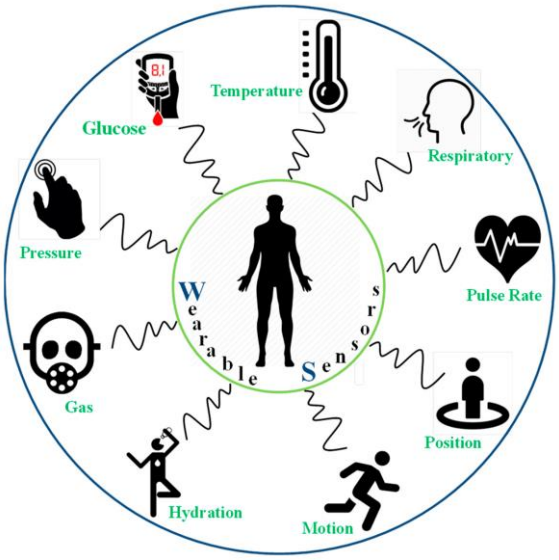


Figure 2.1 Wearable Sensors

2.1.1 Inertia sensors

The inertial sensor is one of the most widely used types of sensors in gait analysis, inertial sensor measures the acceleration and angular velocity of an object along three mutually perpendicular axes [121]. Accelerometers (measures acceleration in 3 dimensions), gyroscopes (measures angular velocity in 3 dimensions), and magnetometers (measures magnetic field) can be the component of the same Inertial Measurement Unit (IMU) as shown in Figure (2.2). Inertial sensors measure velocity, acceleration, orientation, and gravitational forces and can be used to study gait initiation **[19],** assess standing balance **[20],** and quantify bradykinesia **[21]** as shown in Table (2.1).

An inertial sensor can be attached to the feet, legs, or waist **[22]**. Magnetometer included because it can provide information that cannot be determined by both an accelerometer and gyroscope **[23].** The inertial sensor data can be processed on the fly or uploaded for offline processing.

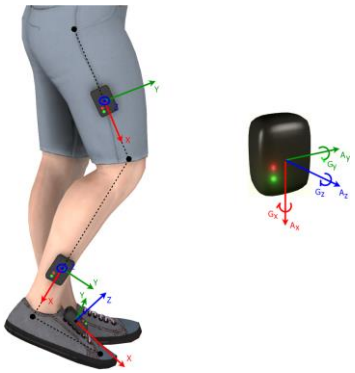


Figure 2.2 : Inertia Sensor

Table 2.1: Studies utilizing nertia sensors for gait analyses

Author	**Year**	**PD Patients /Controls**	**Aspect of PD**	**Device/ Sensor placement**	**Test location**	**Method**	**Metric**
Chen et al [24]	2021	50/50	Detect abnormal PD gait.	Shoe sensor.	Hospital /clinic	1D CNN on raw sensors data.	Accuracy 91.4%.
Mazilu et al [25]	2016	11/0	FoG	wrist sensors	Lab settings	ML and hand-	FoG hit rate of 0.9, and

						crafted features.	specificity between 0.66 and 0.8
Perez et al[26]	2020	2/0	detection of four gait events	foot-mounted	Lab settings	Modified SVM approach	F 1 -scores were 0.987 and 0.95

2.1.2 Goniometer (Angle Measurement)

In orthopedics, a goniometer is used as a device that measures an angle or permits the rotation of an object to a definite position. Goniometers read angle changes based on the change in the internal resistance of the sensor, angle changes has been useful for the determination of gait parameters [122]. The Goniometer sensor is easy to set up and use **[30]**. They are commonly used to study the angles for ankles, knees, hips, and metatarsals as shown in Figure (2.3). Several studies utilized Goniometer for gait analysis , those are listed in Table (2.2) below.

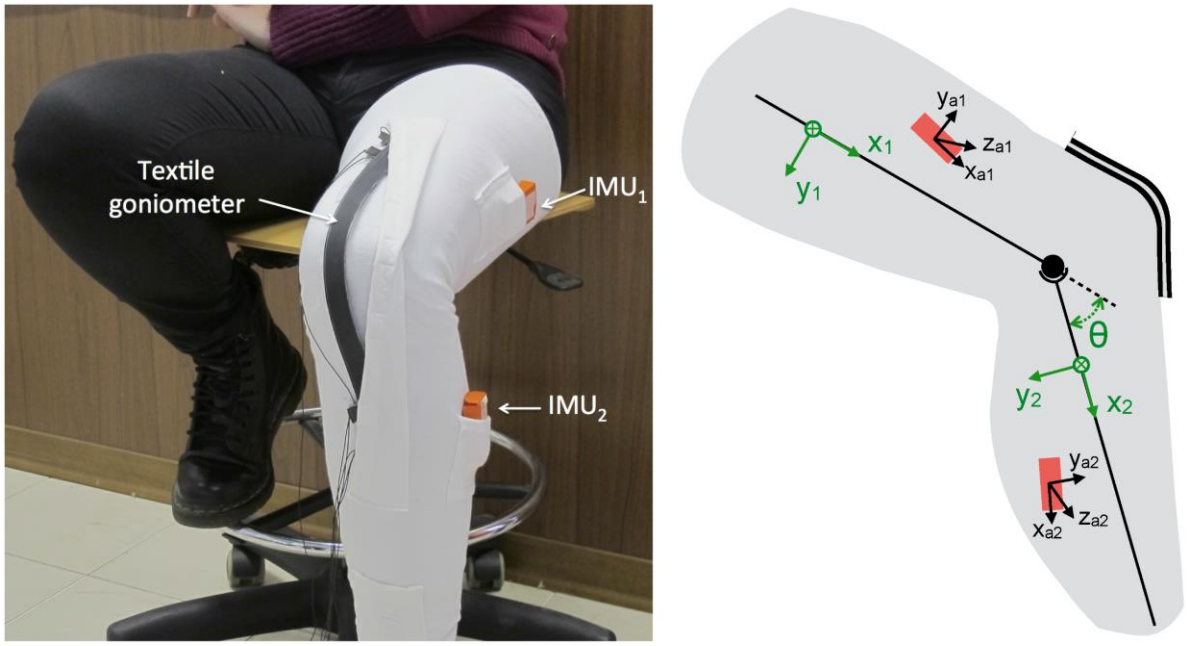


Figure 2.3: Goniometer Sensor

Table 2.2: Studies that utilized Goniometers/angle sensor for gait analyses

Author	**year**	**PD Patients /Controls**	**Aspect of PD**	**Device/ Sensor placement**	**Test location**	**metrics**

[27] Wang et al.	2020	18/25	disease severity of PD	Ankle	Lab settings	angle metrics were significantly smaller in those with PD: mean step angle (F1,48=69.75, P<.001, partial eta-square=0.59), initial step angle (F1,48=15.56, P<.001, partial eta-square=0.25), and last step angle (F1,48=61.99, P<.001, partial eta-square=0.56).
[28] Meg et al.	2005	12/12	PD step length.	Hip, knee, ankle, pelvis	Laboratory	Stride length.
[29] Wegan et al.	2018	15/0	trunk posture angle	Chest	Home	a significant decrease (average −5,4°) in trunk angle from the baseline period to the intervention period

2.1.3 Pressure and Electromyography

Force sensors measure the ground reaction force under the foot as the subject walks over them, converting this force to a current or voltage proportional to the force applied, as shown in Figure (2.4). Sometimes this sensor is installed in the shoes to save cost **[31].** Several previous studies used pressure and force sensors to study stride length variation of PD patients and also to study FoG. **[32].**

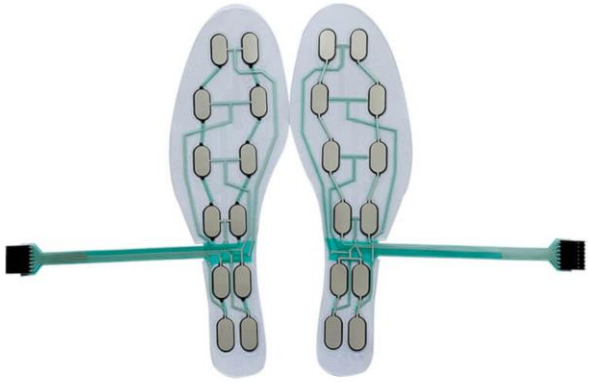


Figure 2.4: Foot Peressure Sensor

Electromyography sensor measures the small electrical signals generated by the muscles when they are in motor activity. Some Electromyography Sensors have a wireless interface to

evaluate gait. **[33].** They can be used to study the PD postural disorders in Parkinson's disease patients, and to explore muscular activity in the Pisa syndrome **[34]** Table (2.3).

Table 2.3: Gait studies using Pressure and Electromyography sensors

Author	**year**	**PD Patients /Controls**	**Aspect of PD**	**Device/ Sensor**	**Test location**	**Metrics**
[35] Marcante et al	2020	20/0	FoG	Shoe pressure sensor	clinical and ecological settings	detected 90% of the FoG episodes. The false positive rate was 6% and the false negative rate was 4%.
[36] Cando et al.	2016	5/1	Detection of FoG	Shoe pressure sensor	Laboratory settings	reduction in the freezing duration is 50.94%, also the time to complete the trial reduces showing an improvement of 34.25%.
[37] Shalin et al.	2021	11/0	LSTM to detect and predict FOG	Shoe pressure sensor	Laboratory settings	The model correctly detected 95% of freeze episodes. 82.1% sensitivity and 89.5% specificity for one-freezer-held-out cross-validation.
[38] Kugler et al.	2013	5/5	recognition of PD using surface electromyography during gait.	Leg muscles.	Laboratory settings	Sensitivity and specificity of 0.90 using leave-one-subject-out cross-validation of SVM classifier.

2.1.4 Impulse Radio and Ultrasound sensors

The Impulse Radio Ultra-WideBand (IR-UWB) sensor idea is based on the concept of sending very short (typically 2-3 ns long) signals, then tracking the body's movement based on the reflected waves from the target body [60]. IR-UWB is used to measure activity for movement disorders **[39]**. This technology is also used for step and gait phase detection **[40].**

Ultrasonic sensors are based on the same idea of sending a sound and estimating the movement based on the received reflected signal[33]. This technology is used for measuring step length and gait phase detection **[41]**. It is also useful in analyzing gait symmetry and coordination **[42]** Table (2.4).

Table 2.4: Gait Analyses studies that utilized IR-UWB and Ultrasonic sensors

Author	year	PD Patients /Controls	Aspect of PD	Device/ Sensor	Test location	Metrics
[43] Blumrosen, G., et al,	2010	0/0 (simulation)	quantifies and analyzes tremor. estimated the tremor frequency	UWB radar detection.	Laboratory settings	The approximated amplitude decreases with frequency.
[42] Ashhar et al	2017	0/5	assess the human gait symmetry	ultrasonic transmitters were Placed just above the ankle joint	Laboratory settings	RMSE in millimeters is between 25.36 and 28.78
[41] Wahab et al.	2011	N/A	Time of Flight (Tof)	Ultrasonic system in an instrumented shoe.	Laboratory settings	Sensor output is proportional to the Tof.

2.2 Non-wearable Gait analysis

Non-wearable sensors include floor sensors **[44],** image-processing including cameras**[45]** , and stereoscopic vision **[46]**. The majority of the Non-wearable analysis technologies are complex and expensive, so they cannot be deployed in the PD patient home environment.

2.2.1 Floor sensors

Floor sensors are the most basic method to collect initial data about a person's gait pattern. Floor sensors are installed in a special floor mat. The mat can detect when a person walks and starts recording force and pressure measurements for processing **[47,48].** Several studies used floor sensors to estimate and characterize gait impairments, including PD gait fibromyalgia **[49-52]**. Those studies are presented in Table (2.5).

Table 2.5: Gait Analyses studies utilizing Floor sensors

Author	year	PD Patients /Controls	Aspect of PD	Device/ Sensor	Test location	Metris
[48] Lee et al.						

[49] Mondal, et al.	2019	70/37	Compare gait stability and gait parameters	Floor sensors	Research center (lab settings).	mean velocity for HC (99.19cm/s) compared to PD patients (73.90cm/s, P value 0.0001). cadence was comparable (103.29 vs. 103.39, P value 0.966).
[50] Yang et al	2008	18/17	relationships between gait and dynamic balance	Floor sensors	Lab settings.	People with early-stage PD exhibited significantly slower walking speed, shorter stride length, and smaller forward MV than the comparison group.
[51] Rehman et al.	2019	93/103	compare the impact of gait assessment systems on the performance of (SVM) and (RF)	Floor sensor and an accelerometer attached at the lower back Axivity	Lab settings.	SVM performed better than RF. during CW Axivity significantly outperformed Floor sensors(AUC: 87.83 ± 7.81% vs. 80.49 ± 9.85%);
[52]	2002	11/11	distinguish gait characteristics of PD patients	Floor sensor	Lab settings.	Mean Normalized Velocity (MNV) was 0.83 for PD at preferred speed and 1.14 at fast speed, the non-impaired was 1.33, and 1.70 respectively. PD patients have lower FAP scores, shorter step lengths, and a long step time.

2.2.2 Camera-based Gait Analyses

Camera-based systems consist of single or multiple cameras that record a person's walk and process the images to obtain information about their gait pattern. This technology is used for several applications including gait recognition **[53]**, and other medical applications like

studying changes in the subject path **[54]**, kinematic **[55],** and PD gait **[56-59]**. Image-processing studies are presented in Table (2.6).

Table 2.6 : Camera-based Gait Analyses studies

Author	**year**	**PD Patients /Controls**	**Aspect of PD**	**Device/ Sensor**	**Test location**	**Metrics**
[56] Tahir et al	2012	12 / 20	Recognize gait pattern of PD.	Infrared camera/ force plate.	Clinic	Both ANN and SVM achieve a high classification accuracy of >95%.
[57] Tucker et al.	2015	7/0	adherence to medication protocols, based on gait variations	line of motion sensing (Infrared Cam)	Lab settings	discriminate on and off medication, with 97% for some and 78% for multiple patients.
[58] Rocha et al.	2014	3/3	discriminating PD vs HC. and PD ON and OFF).	RGB-D camera (Microsoft Kinect)	Clinic (University Hospital)	the variance of center shoulder velocity is the highest discriminative to distinguish between HC vs PD, and ON vs OFF states ($p = 0.004$).
[59]	2021	49/0	Prediction of Parkinsonian Gait	ceiling-mounted camera	Clinic	accuracy of the model that only used gait features was 82.8%, while the model that also used joint trajectories had an accuracy of 94.2%.

2.2.3 IR Thermography

Infrared thermography is a technique used to recognize human gait, by creating an image based on the temperature of the human body surface (the skin). This sensor reading depends on the emissivity of the human skin [**60]**.

2.3 Smartphone Sensing

With the advent of technology, smartphones can now provide an affordable, accessible, and easy-to-use alternative to PD gait sensing. Smartphones are ubiquitous, portable, and user-friendly. Equipped with triaxial accelerometers and gyroscopes and powerful CPUs, smartphones provide a viable alternative for remote gait assessment in the home environment. This smartphone technology has been used to analyze PD gait in several studies **[61],** including our prior work **[62,63].** A comparison of PD smartphone studies for gait analysis is presented in Table (2.7), which includes a description of each study's methodology, outcome measure, number of PD participants, smartphone placement, and the study performance metrics.

Several PD studies enabled participants to perform periodic assessment activities in their homes over extended periods using mobile health apps **[64,65]**. Accelerometer and gyroscope sensors have been demonstrated as useful in assessing gait, tremor, and walking balance **[62].** Those previous studies were able to quantify multiple gait modalities of PD, including walking imbalance and FoG.

Table 2.7: Smartphone Gait studies

Author	**year**	**PD Patients /Controls**	**Aspect of PD**	**Device/ Sensor**	**Test location**	**Limitation/Outcome**
[67] Mazilu, Sinziana, et al	2012	10/0	FoG.	Smartphone and wearable accelerometers	Clinic	Small number of participants. Did not combine gait as well as lifestyle features.
[65] Bot et al	2016	5718/ 1087	Data collection	Smartphone sensors	Remote	Did not combine gait as well as lifestyle features.
[68] Arora et al	2015	10/10	Voice, posture, gait, finger tapping, and response time	Smartphone sensors	Clinic and remote	Small number of participants. Did not combine gait as well as lifestyle features.
[69] Ellis et al	2015	12/12	Gait variability	Smartphone accelerometer, gyroscope, and	Clinic	Small number of participants. Did not combine gait as well as lifestyle features.

				heel-mounted footswitch sensors		
[70] Printy et al	2014	26/0	Bradykinesia.	Smartphone gyroscope, accelerometer, touch screen, microphone, and front- camera	Clinic	Small number of participants. Did not combine gait as well as lifestyle features.
[71] Kim, Hanbyul, et al	2015	15/0	FoG	Smartphone gyroscope, accelerometer	Clinic	Small number of participants. Did not combine gait as well as lifestyle features.
[72] Sharma, Vinod, et al	2014	0/5	Facial tremors, speech, dyskinesia, gait abnormalities	Smartphone/Smart watch accelerometer, front camera, microphone	Remote	Small number of participants. Did not combine gait as well as lifestyle features.
[73] Zhan, Andong, et al	2016	121/105	Voice, balance, gait, dexterity, and reaction time.	Smartphone/acceler ometer, touch screen, mic.	Remote	Did not combine gait as well as lifestyle features.
[74] Lee, Chae Young, et al.	2016	57/87	Bradykinesia	Smartphone/screen, mechanical tapper	Clinic	Small number of participants. Did not combine gait as well as lifestyle features.
[14] Kassavetis, Panagiotis, et al.	2016	14/0	Tremor, bradykinesia	Smartphone/acceler ometer, touch screen,	Clinic	Small number of participants. Did not combine gait as well as lifestyle features.
[63] Abujrida et al.	2019	340/116	Analyzed various aspects of PD gait.	Smartphone Accelerometer and Gyroscope	home	Uneven data bins (classes). Accuracy can be improved with DL methods.
[75] Zhang et al.	2019	247	12 weeks Continuous monitoring, with smartphones, uncontrolled environment.	Smartphone sensors. (CNN with Spectrogram signal conversion).	home or office	Unnecessarily conversion of time series by calculating spectrogram, adding complexity, no controls in the study, limiting the study to binary classification.
[76] Kan et al.	2018	0/10	Smartphone app recording medication intake and ball-game app to measure tremor.	Smartphone Screen/Accelerometer/ gyroscope.	Lab settings	Small # participants who are not PD patients. ML/DL methods can extract informative features from sensor data.

[77] Lakshminaray ana et al.	2017	158/0	Smartphone app recording adherence using questionnaires.	Smartphone app (Smartphone Screen)	home	depending on PD patients' inputs instead of measuring change by smartphone sensors.
Lo et al, **[78]**	2019	237/0	falls, FoG, postural instability, Smartphone placed in Pocket, armband	Smartphone microphone, triaxial accelerometer, and screen.	in clinic and home	Random falls, FoG, and postural instability: AUC = 0.94, 0.95, and 0.9, respectively
Fiems et al, **[79]**	2020	59/0	Sway score to predict future falls. Smartphone Placed at chest harness	triaxial accelerome ter	laboratory setting (clinic)	AUC for ABC: 0.76 (Mini-BESTest): 0.72, MDS-UPDRS: 0.66, and sway: 0.65
Pepa et al [**80**]	2020	44/0	FoG Handcrafted features from the accelerometer, range of values classification. Placed at the right or left side of the hip	triaxial accelerome ter	laboratory and home	FoG Se =0.85 Sp = 0.95 Acc=0.92 AUC = 0.91
Chen et al. **[81]**	2020	37/35	discriminate PD participants from HCs, and estimate the PD disease severity (MDS-UPDRS total scores). rouser pocket/belt pouch	Smartphone microphone, triaxial accelerometer, gyroscope	Home environment	Feature selection and shallow ML. Pearson correlation of 0.72 ($p<0.0001$) and an RMSE of 16.58 between the estimated and the observed MDS-UPDRS total scores
Su et al. **[82]**	2021	52/0	correlation of the calculated gait features to UPDRS.	accelerometer, gyroscope, and compass, Smartphone placed in The front pocket of the pants	Hospital/clinic	Handcrafted features. Participants who walked with greater stride time variability exhibited a greater UPDRS III score.
Borzì et al. **[83]**	2020	42/7	Handcrafted features from the accelerometer, fed to shallow ML. Discrimination PD/HC, Discrimination between postural stability levels	Accelerometer. waist-mounted	Hospital/clinic	(Binary classification) Discrimination PD/HC Accuracy = 100% PD with different levels of postural stability: Accuracy 72%-100%
Memedi, Mevludin, et al **[84]**	2013	95/10	Tremor, bradykinesia	Touch-pad handheld computer	Clinic and Remote	Small number of participants. Did not combine gait as well as lifestyle features.

A good number of studies rely on handcrafted features and traditional ML algorithms. However, handcrafted features need to be extracted by a domain expert, and need to be

evaluated on a large sample of participants to prove the validity of the features. It can also be noted from Table (2.7) that, each study focuses on one aspect of PD gait, i.e., classification/detection of FoG or posture stability, or the differentiation of PD gait from regular HC gait. Almost all studies have a small number of participants except the study by Lo et al, **[78]**, although the model of this study can be improved using DL, to eliminate the dependency on handcrafted features.

2.4 Studies Related to Medication Adherance

One of the applications of our system is the evaluation of PD patients' medication adherence and response to medication based on their gait. As part of our literature review, we analyzed several technological methods that focus on medication adherence.

Prior work has explored various methods for measuring medication non-adherence, these include the reliance on physician's judgment [8,9], patient self-reporting and management [15], tablet count [7], tracking drug levels in urine and blood [11], and using electronic devices [12,13,14,17]. While tracking the drug level in blood and urine is accurate, it requires visiting a lab, which presents a burden for patients. Moreover, as the majority of the medications are taken outside of the clinic, physicians tend to rely more on self-reporting and self-management. As PD progresses, the loss of memory and depression dramatically increases the patients' non-adherence, which further increases as the patients' daily dose increases [10]. Passive, continuous, and unobtrusive methods of monitoring medication adherence would increase patients' compliance and reduce missed episodes. Compared to traditional methods. Electronic measures and AI seem to be the most practical approach to track PD patients' medication intake.

A comparison of PD studies and electronic adherence measurement studies is presented in Table (2.8). A good number of studies **[85-87]** rely on Medication Event Monitoring System (MEMS) drug bottles, there are several commercial products, but the main function is to record the date/time of opening the medication bottle. There are multiple issues with this approach,

First, sometimes the patient may open the medicine bottle but do not consume the medication. Secondly, MEMS approaches do not monitor the patients' symptoms after taking the medication to also monitor the patient's response. Infrared motion sensing presented in **[57]** provides a viable approach to compare the differences in gait before and after taking medication. But due to the complexity of the system, it is not feasible in the patient's home environment. The approach using smartphone sensors presented in **[75]** is suitable for PD medication adherence problems, the processing of the walk time series and calculation of spectrograms from the raw sensor data resulted in unnecessary complexity and reduced the overall accuracy of the model. The Smartwatch approach **[88]** seems to be promising. However, it needs to be validated on a larger dataset.

Table 2.8 : Medication Adherence studies

Author	**year**	**Participants**	**Duration of the study**	**Device/ Sensor**	**Pros**	**Limitation**
Grosset et al **[85]**	2005	68	12 weeks	MEMS drug bottle	Dosage reminders, long-life battery, bottle opening sensing and count.	Patient misuse of the device, patients lost device, no guarantee for taking the dose except opening-closing the bottle, limited # participants.
Grosset et al **[86]**	2009	112	4 weeks			
Leopold et al **[87]**	2004	39	4 weeks			
Kalantarian et al **[88]**	2015	20	Duration of data collection.	Smartwatch detection of taking medication	Automatic detection of medication intake.	The use of hand-crafted features can lead to overfitting. ML/DL methods can improve accuracy and reduce overfitting.
Koesmahargyo et al **[89]**	2020	4,182	varied	Machine learning, smartphone front-facing camera	Dosage reminders, uncontrolled environment.	Accuracy is low. can be improved with DL methods.
Tucker et al **[57]**	2015	7	Duration of data collection session(s).	line of motion sensing (Infrared Cam)	Examine the difference in symptoms between on/off medication states.	Camera angle calibration and sensor constraints make it difficult to self-recording at home. model overfit on individual participants. small #participants.
Zhang et al **[75]**	2019	247	12 weeks	Smartphone sensors.	Continuous monitoring,	Unnecessarily conversion of time

				(CNN with Spectrogram signal conversion).	with smartphones, uncontrolled environment.	series by calculating spectrogram, adding complexity, no controls in the study, limiting the study to binary classification.

2.5 Lifestyle Features for assessing PD Gait

Lifestyle features including age, gender, smoking-history and recent history of exercise, are often ignored when analyzing PD gait. To the best of our knowledge, lifestyle features have not been included when assessing gait severities of PD using engineered sensor features. Van et al [90] found that PD rapidly increased over the age of 60 years, with only 4% of the cases under the age of 50 years. The rate for men (19.0 per 100,000, 95% CI: 16.1, 21.8) was 91% higher than that for women (9.9 per 100,000, 95% CI: 7.6, 12.2). Smoking reduced tremor, rigidity, bradykinesia, and gait disturbance, including frozen gait. These effects lasted for approximately 10-30 min after smoking a cigarette and relieved PD symptoms during the off-medication period [91]. Fertl et al [92] found a significant reduction in physical activity during the course of the disease, but no complete abandonment of sports was observed. Swimming, hiking, and gymnastics were the favored sports. Reuter et al [93] concluded that motor disability in PD patients can be improved by intensive sports activities in the early to medium stages of PD, A comparison of PD lifestyle studies for gait analysis is presented in Table (2.9),

Table 2.9 PD lifestyle studies

Author	**Year**	**Participants**	**Lifestyle Features**	**Duration of the study**	**Approach**	**Outcome**
Joshi et al [123]	2010	487	Age , Gender	18 months	ML Decision tree , Bagging, BF tree, Random Forest, RBF networks , and Neural Networks method	PD classification with accuracy of 99.25%. Stroke, diabetes, genes and age play a major role in the development of PD

KC Paul et al [130]	2019	360	Consumbtion of coffee, tea, alcohol. smoking, physical activity	(2 to 4 examinations; conducted in 2007-2014).	physical examinations to one of the Hoehn & Yahr (H&Y) stages (1-5) based on clinical descriptions of each stage	This population-based study suggests that lifestyle factors influence PD progression and mortality.
I Reuter et al[93]	1999	16	influence of an intensive exercise training on motor disability	20 weeks	UPDRS Evaluations were performed before the start of the study (exam. 1), after 7 wk (exam 2), 14 wk (exam 3), and 20 wk (exam 4/long-term effect).	UPDRS Σ score ($P <$ 0.0001) improved significantly by exercise training. Six weeks after termination of the training program, the majority of the patients had lost only minor components of their regained motor skills.

Prior work utilized different sensors and ML algorithms to solve the problem of PD gait assessment. While many studies utilized the widely adopted Smartphone sensors, very few studies utilized the 1D CNN approach based on gait-cycle segments. For instance [128] utilized Vertical ground reaction force (VGRF) with 1D CNN, was able to achieve 88.7% accuracy, [131] utilized a visioni system with digital camera and 1D-CNN to achieve a highest accuracy of 79.3%. [129] also utilized (VGRF) and 1D CNN and achieved achieved an accuracy of 85.3% in Parkinson's severity prediction.

In this dissertation we propose a deep-learning-based system that comprehensively classifies 3 PD symptoms: the severity of FoG, walking imbalance, and shaking/tremors from data gathered in one study, in addition to the model ability to predict medication adherence. Prior work has trained and tested separate models to analyze each of these PD gait anomalies separately, the model we introduced is a single model that achieved impressive results for all of the PD gait symptoms. This was challenging because the model's parameters had to be jointly tuned in order to establish relationships with different sets of PD symptom labels, all while using the same dataset as an input. To the best of our knowledge, our 1D CNN based

model is the first to to Classify the PD Walking-Balance, Shaking/Tremor, and Freeze of Gait (FoG) symptoms, with an accuracy of: 99.1%,98.4%, and 98.2% respectively. Our model was also able to discriminate PD patients on- vs off-medication and baseline HC walk with an accuracy of 98.2%. The accuracy of our CNN model surpassed that of traditional Machine Learning methods by over 17%., and surpassed the best-published results achieved by prior 1D CNN smartphone models by over 7% [128,129].

CHAPTER 3
BACKGROUND ON DATA GATHERING STUDY, DATASET AND PRE-PROCESSING

3.1 Dataset

Data were acquired from the mPower study **[65]** a clinical observational study on PD conducted entirely through an iPhone (Apple Inc., Cupertino CA, USA) app interface, Figure (3.1) shows the versions of the smartphone used in the study. The mPower study interrogated aspects of movement disorder through surveys and continuous sensor-based recordings from participants with and without Parkinson's disease.

3.1.1 Dataset Overview

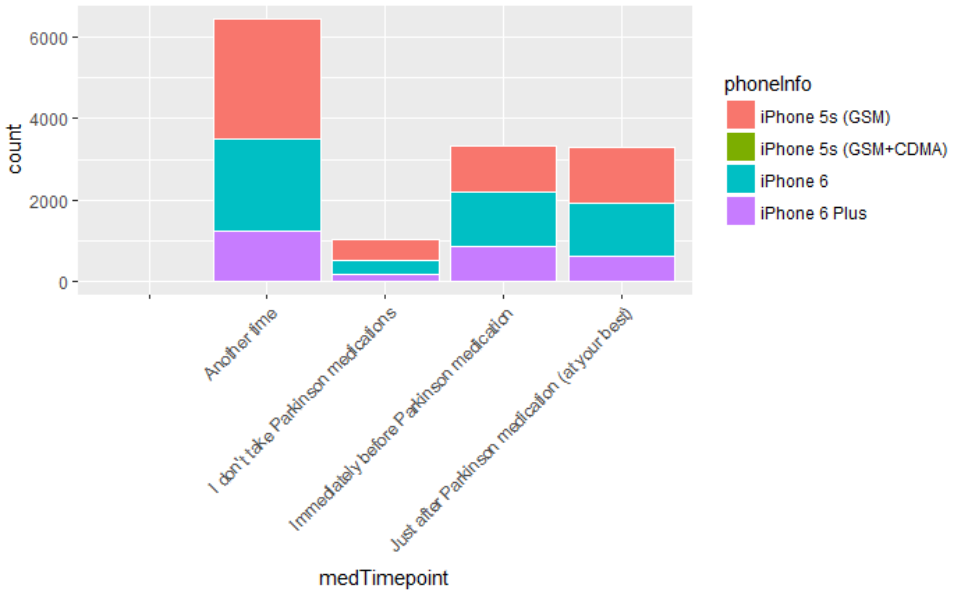
The mPower study had a large enrollment (N= 9520) of participants who opted to share data broadly and contributed at least two measurements. The goal of the study was to help establish baseline variability of real-world activity measurement collected via mobile phones that might ultimately lead to quantification of the ebbs and flows of PD symptoms. The collected mPower activities included 35410 walking, 78887 tapping, and 8569 memory records as shown in Table (3.1). Subjects conducted the PD tests using an iPhone smartphone running the mPower data-gathering application. Participants self-reported PD severities and contributed activities several times during the day, before/after taking medication, and at another time of the day, as shown in Figure (3.1).

Table 3.1: The mPower Dataset

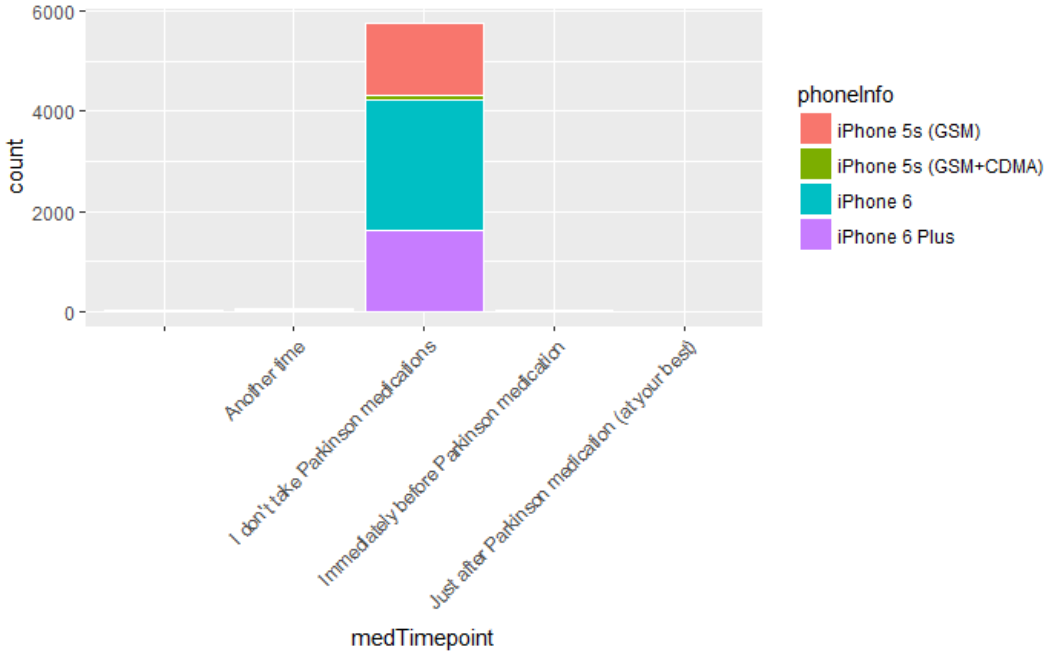
Activity	Frequency	Participants	Records
Demographic Survey	Once	6,805	6,805
Walking	**Daily**	**3,101**	**35,410**
Tapping	Daily	8,003	78,887
Voice	Daily	5,826	65,022
Memory	Daily	968	8,569

Not all participants complied with the application protocol, and therefore their number of recorded activities varied from a few to hundreds of recordings. Approximately 658 PD

patients performed 24001 walks, and 2165 HC performed 10585 walking activities in the first 6 months of the study. Only 815 participants performed at least 3 walking activities.



(3.1a) Time of walking for PD Patients



(3.1b) Time of walking for Healthy Controls

Figure 3.1: Smartphone version and Time of walking

Participants filled out surveys including a subset of the UPDRS Section I (non-motor experiences of daily living) and Section II (motor experiences of daily living). Participants also completed a demographic survey, which included information on their general health history, PD history, and general lifestyle questions.

3.1.2 Walking activity of the dataset

The collected mPower activities included 35410 walking's. But the number of walking records per participant varied from a few to hundreds. 658 PD patients performed 24001 walks, and 2165 HC performed 10585 walks in the first 6 months of the study. Figure (3.2) shows a histogram of the walking activities of all participants.

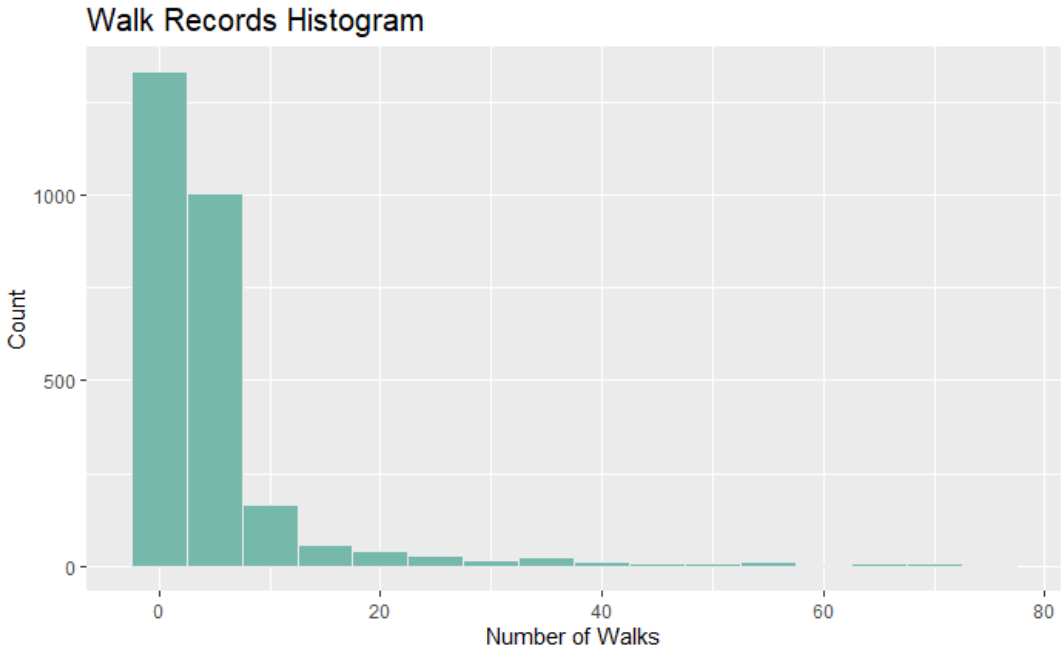


Figure 3.2: Number of Walks per participant

3.1.3 Participants Selection

In our work, we included only PD patients that contributed 3 walks before and 3 after taking medication. Healthy controls were selected if they contributed at least 3 walks in total. We also excluded activity records in which key values of the demographic survey were missing or certain sensor readings were missing. Because the MDS-UPDRS survey data are used for labeling the walk records, we filtered out participants whose survey data were not complete. Any walking record that had missing values such as sensor reading or key demographic values

has been dropped. In a few instances in which lifestyle questions had missing values, they were replaced (inputted) by the mean of the feature. The above subject selection rules yielded a working dataset with 152 PD patients (NPD = 152), and 304 healthy controls (NHC = 304), as shown in Table (3.2). For participants who performed more than 3 walks in each category (before/after medication or at another time), we selected walks performed close to the date on which participants completed the demographic survey to increase the accuracy of labels for each activity. PD symptoms are known to calm down after medication **[94]**. Therefore, we wanted to capture the patient symptoms at peak occurrence, so we mostly analyzed the walks recorded before taking medication, unless both before/after medication walks are required for some experiments.

Participants also filled UPDRS survey describing the severity of gait anomalies, a sample of the UPDRS questions is shown in Table (1.1), and their corresponding results for the selected sample are shown in Figure (3.3).

Table 3.2: Selected Sample

PD Diagnosis	**Classes**	**Age Mean± std**	**Male: Female**	**RACE**	**Participants**	**Records**	**Training Records**	**Testing Records**
PD Patients	Before Medication	63.57± 8.09	89:63	"White/Caucasian" 74.5% "Latino/Hispanic" 7.2% "Mixed" 2.7% "South Asian" 2.7% "Middle Eastern" 2.3% "Black or African" 2.1% "East Asian" 2.1% "Other" 6.1%	152	456	410	46
	After Medication					456	410	46
Healthy Controls	Another Time	40.14± 15.45	213:91	"White/Caucasian" 93.1% "Latino/Hispanic" 1.1% "Mixed" 0.56% "South Asian" 0.56% "Middle Eastern" 0.56% "Black or African" 1.1% "East Asian" 1.1% "Other" 1.7%	304	912	820	92

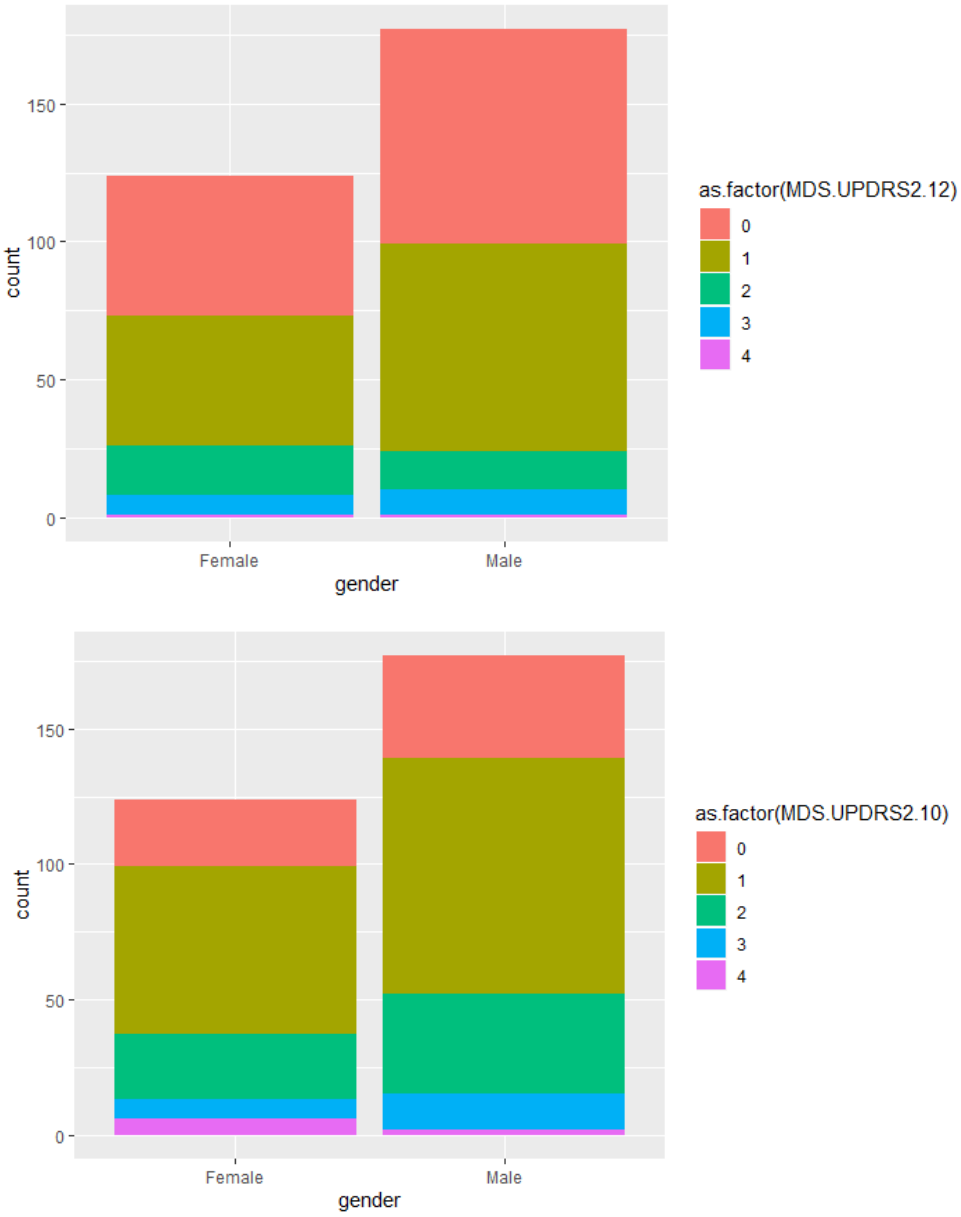


Figure 3.3: UPDRS2.10 (Shaking) and UPDRS2.12 (walking imbalance) severity by gender

3.2 Signal Preprocessing

Signal pre-processing involves several steps. Those will be discussed in the following few sections. Steps include signal filtering and segmentation. Some ML models needed feature extraction, while for DL methods we operated on raw signal components.

3.2.1 Smartphone Gait Signal Capture

To perform the walking activity, participants were asked to walk for 30 seconds in a straight line while placing the smartphone in their pants' front pocket (Figure 3.4), stand for 30 seconds, and subsequently turn and walk back for 30 seconds. The raw data gathered included the smartphone's gyroscope and accelerometer data sampled at 100 Hz, as well as pedometer values and the time of the activity. Our work only analyzes the outbound walking to enable uniform analysis across patients and Healthy Controls (HC) by avoiding FoG and walking imbalance events that usually occur when PD patients attempt to start walking after they turn between the outbound and inbound walks.

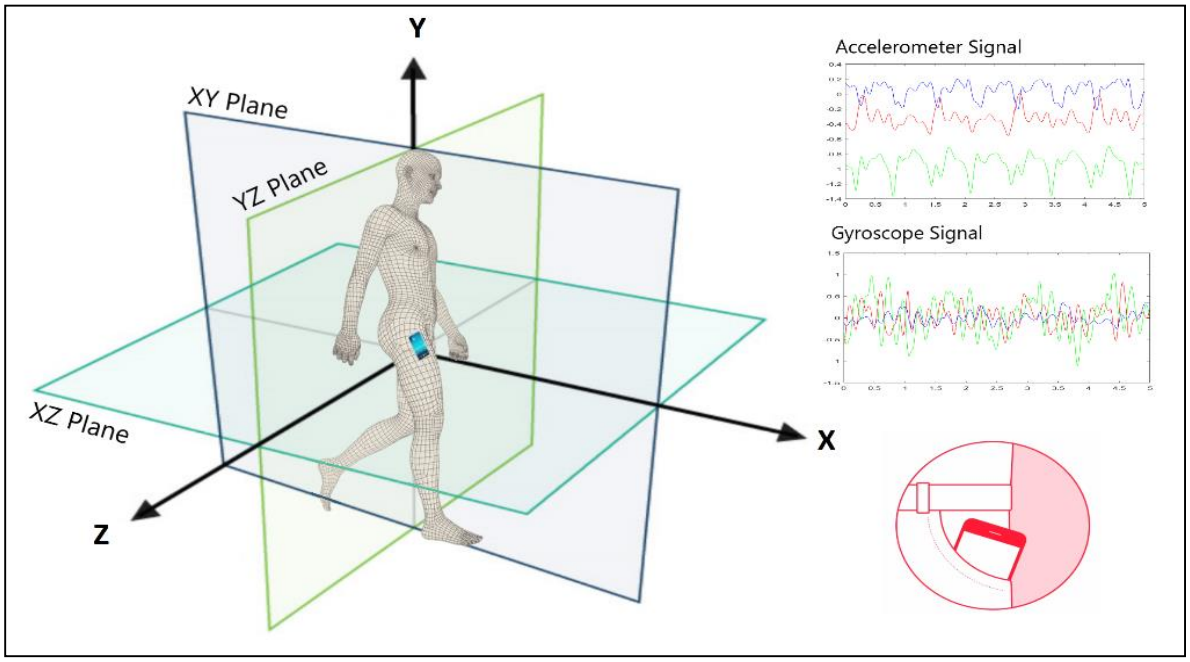


Figure 3.4: Participant Walking with Smartphone in pocket

3.2.2 Signal smoothing and filtering

Signal preparation start by pre-processing the sensor data and reorganizing it into a readable format. We obtained the two main signals from the smartphone sensors, i.e., acceleration and rotation vectors from the accelerometer and gyroscope, respectively:

$\alpha(i) = [\alpha x\ (i),\ \alpha y\ (i),\ \alpha z\ (i)]\ T$ (in m/s2)

$\omega(i) = [\omega x\ (i),\ \omega y\ (i),\ \omega z\ (i)]\ T$ (in deg/s)

where i denotes discrete time, α indicates accelerometer, and ω represents gyroscope.

PD tremor and balance classification have traditionally been captured by sway metrics derived from raw accelerometer values. However, we believe that features derived from the smartphone gyroscope data should supply additional information because it records angular velocity. To facilitate feature extraction, sensors were smoothed by computing the moving average (n=5) and removing sudden changes. The moving average calculation replaces each value in the sequence with the average of several points around it and is given by the following formula:

$$MA_{\alpha x} = \frac{1}{n}\sum_{i=0}^{n-1}\left(\alpha x_{(-i)}\right)$$

3.2.3 Preparing signal components

The values of the three accelerometers and gyroscope axes are used to calculate the signal magnitudes, after which the signal's mean is subtracted to eliminate gravity or any constant factors such as breathing. The resulting formula is given below:

$$MagNG_{\alpha} = \left(\sum_{i=1}^{n}\left(\left\|\alpha_{(i)}\right\|\right) - \overline{Mag_{\alpha}}\right) \quad (1)$$

$$MagNG_{\omega} = \left(\sum_{i=1}^{n}\left(\left\|\omega_{(i)}\right\|\right) - \overline{Mag_{\omega}}\right) \quad (2)$$

where *MagNGα* and *MagNGω* are the vector magnitudes of the acceleration and the rotation rates, respectively; and $\overline{Mag_{\alpha}}$ and $\overline{Mag_{\omega}}$ are the means of the acceleration and rotation rates, respectively.

3.3 Signal Segmentation

Depending on the algorithm used, we proceeded with different types of signal segmentation; Time-Based segmentation for ML and cycle-based segmentation for DL algorithms.

3.3.1 Time-Based Segmentation

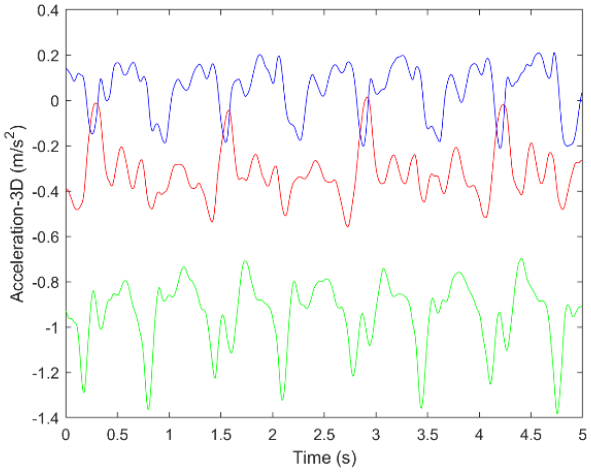
Time-based signal segmentation is used to prepare the signal for ML feature extraction. Signals are first divided into 5-second non-overlapping segments, then features extracted for each of the 5-seconds portions of the signal. To give an example of a feature calculation, the computation of average step time is summarized below.

To compute the average step time metric, we initially found the peaks of the accelerometer axis with the largest magnitude. Only those peaks that are greater than a minimum peak height (MPH) are considered a step. MPH is calculated by the following:

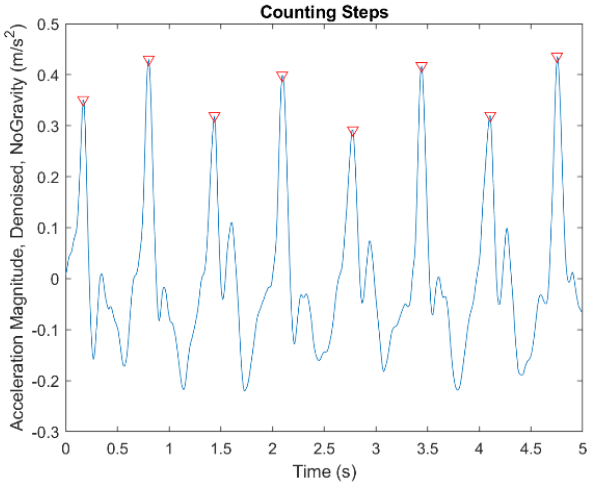
$MPH_\alpha = \overline{Mag_\alpha} + \sigma_\alpha$

where σ_α is the standard deviation of α.

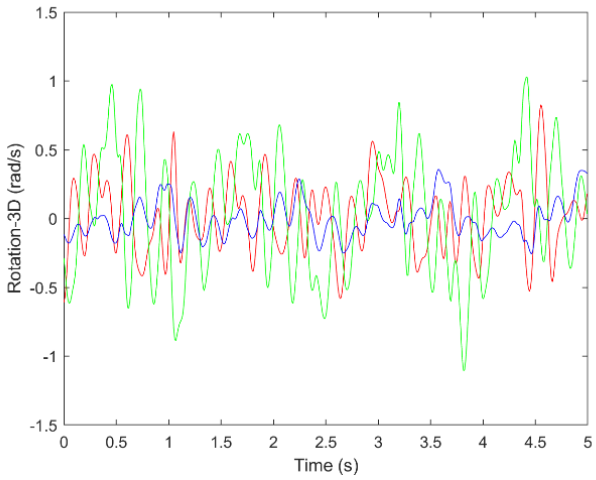
Figure (3.5a) shows a sample of the accelerometer and gyroscope signals on the three axes after preprocessing, smoothing, and removing the gravity component. Figure (3.5b) shows the peaks detected, which are used to estimate the steps for this walking segment. Our methodology does not require the passing of the signal through filters. Figure (3.5c-d) shows the signal components and magnitudes.



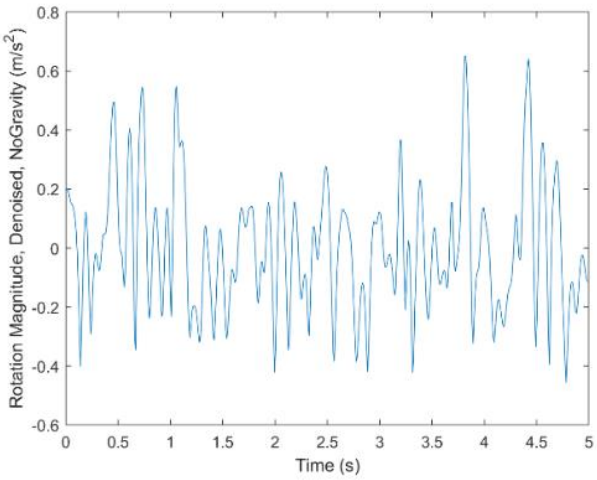
(3.5a) Acceleration 3D signal



(3.5b) Acceleration Magnitude and gait peak detection.



(3.5c) Rotation 3D signal



(3.5d) Rotation Magnitude

Figure 3.5: Acceleration and Rotation signals and gait peak detection

3.3.2 Cycle-Based Segmentation

One of the issues of the mPower dataset is that participants' compliance with the Smartphone app varied from one participant to another. Also, the correct positioning of the Smartphone in the participants' front pockets could not be verified. Consequently, the resulting walking records had gait and non-gait data. Therefore, as our goal was to analyze data for periods when the participants walked, it was essential to identify and extract gait cycles from noisy sensor data. Inspired by Cyclepro **[95]**, we extracted gait segments from the time series and split the accelerometer and gyroscope signals into one-stride segments, according to the following steps; First, the signal magnitude was calculated according to equation (1). Then multiple templates were generated. A template, which represents one walking stride, is selected based on two consecutive local minima's of the signal magnitude. Multiple templates were selected and then the top template was chosen by comparing the template std (statistical standard deviation) with the std of the total walking signal magnitude. Cross-correlation was calculated between the top templates and the *MagNGα* signal to identify the repetitive walking pattern. Cross-correlation was calculated based on equation (3) below:

$$C(i) = \frac{\sum_{j=1}^{T} Mag_t(j) MagNG_\alpha (i+j)}{\sum_{j=1}^{T} Mag_t(j)^2} \quad (3)$$

Where Mag_t is the Template Signal Magnitude.

The output of equation (3) is normalized and used for detecting the walking strides. First, a window of regular cadence size was superimposed and the local maxima's within this window was found. Second, within the window, only the highest maxima were selected and the rest of the maxima's were filtered out. Finally, the final peak was selected using Otsu's method, which finds the similarity of the peaks by minimizing the sum of inner variance within the peaks and separating the invalid peaks.

The resulting maxima's from the final step are used to split Gyroscope and Accelerometer signals into 1-Stride segments. Records of valid gait cycles are then sorted, correctly labeled, and saved for DL processing. The result was a dataset with gait data segmented into single-stride segments, each labeled with the time of the walk, before/after medication for PD patients, and at-another-time for HC.

3.4 Features Extraction

From the two calculated magnitudes MagNGα and MagNGω (referred to as x_i in subsequent sections), the time, frequency, statistical, and wavelet domain features were extracted. After pre-processing the data, gait features were extracted using Matlab (Mathworks) from the accelerometer and gyroscope data gathered during the walking activity. Sway area features were calculated for the gyroscope data. The extracted features were subsequently combined into larger data frames, and multiple datasets were derived and used in PD classification.

3.4.1 Time-Domain Features

Time-Domain features were calculated directly from x_i as shown in Table (3.3). The table shows feature names, descriptions, and definitions.

Table 3.3: Time-Domain features

Time Domain Features and their use cases for Gait analysis			
S.N	*Feature*	*Feature definition*	*Description*

1	*Number of Steps*	*Local Peaks*	*The number of steps taken in a given time interval*
2	*Average Step Time*	$\frac{time}{\#Steps}$	*The average time elapsed for each step*
3	*Average Cadence*	$\frac{\#steps}{time}$	*The ratio of the total number of steps to th total time*
4	*Skewness**	$\frac{\frac{1}{n}\sum(x_i-\mu_x)^3}{\left[\frac{1}{n}\sum(x_i-\mu_x)^2\right]^{3/2}}$	*Asymmetry of the signal distribution*
5	*Coefficient of Variation o Step Time*	$\frac{\sqrt{\frac{1}{n}\sum(interval_i-\mu_{interval})^2}}{\mu_{interval}}$	*The within-subject standard deviation of the stride interval divided by the mean stride interval*
6	*Average Step Length*	$\frac{0.084}{averageStepTime}+1.89$	*The average distance covered by each step*
7	*Gait Velocity*	$\frac{\left(\frac{0.084}{averageStepTime}+1.89\right)}{averageStepTime}$	*The ratio of the total distance covered by the total time*
8	*Minimum and Maximum Difference**	$\max(x_i)-\min(x_i)$	*A global maximum of one step minus a global minimum of one step averaged over all steps of one subject*
9	*Root Mean Square**	$\sqrt{\frac{1}{n}\sum {x_i}^2}$	*Root Mean Square or quadratic mean is a statistical measure*
10	*Entropy Rate**	$-\sum possibility_{unique\ freq} \times log_2(possibility_{unique\ freq})$	*The uncertainty measure of the signal, and the regularity of a signal when it is anticipated that consecutive data points are related*
11	*Sway Area** X.Y, Y.Z, X.Z*	$\pi(AB)$	*Area of an ellipse that encloses the 95 percent confidence interval of all observed gyroscope points in the XY, YZ, and XZ planes. (A and B are the lengths of the semi-major and semi-minor axes of the ellipse)*

3.4.2 Frequency-Domain Features

Frequency domain features were calculated after computing the fast Fourier transform (FFT) and power spectral density (PSD) as shown in Table (3.4). Frequency domain features were subsequently extracted for each walking segment record.

Table 3.4: Frequency domain features

Frequency Domain Features and their use cases for Gait analysis			
12	*Harmonic Ratio**	$\frac{\sum_{i=1,3,5,\dots} V_i}{\sum_{j=2,4,6,\dots} V_j}$	*Harmonic Ratio quantifies the harmonic composition of the accelerations for a given stride via DFT*
13	*Average Power**	$\frac{total\ power\ of\ the\ signal}{bandwidth\ of\ the\ signal}$	*The mean of the total power underneath the curve of the PSD estimate for a signal*
14	*The ratio of Spectral Peak* (with Welch, FFT DCT)*	$\frac{\max(power_{freq})}{mean(power_{freq})}$	*The ratio of the energies of low and high-frequency bands*
15	*Signal Noise Ratio**	$\frac{power_{signal}}{power_{noise}}$	*Power of the whole signal over the power of its computed noise*
16	*The energy in Band 0.5 to 3Hz**	$\int_{0.5}^{3} psd_f\ df$	*The energy in a frequency band describes components of distinct frequencies in the signal, and the frequency range is recommended as 0.5 Hz to 3 Hz*
17	*Windowed Energy in Band 0.5 to 3Hz**	$\int_{0.5}^{3} windowed\ psd_f\ df$	*The energy in the frequency band of 5-second windows with an overlap of 2.5 seconds; windows from the complete signal sequence are averaged*
18	*Peak Frequency**	$\max(power_f)$	*The maximum spectral power*
19	*Spectral Centroid**	$\frac{\sum f \times {power_f}^2}{\sum {power_f}^2}$	*The frequency that divides the spectral power distribution into two equal parts*
20	*Bandwidth**	$\frac{\sum (f - spectralCentroid)^2 \times {power_f}^2}{\sum {power_f}^2}$	*The difference between the uppermost and lowermost frequencies/range of frequencies in the signal (Weighted Average)*

3.4.3 Wavelet-Domain Features

Wavelet domain features were calculated after calculating the Discrete Wavelet Transform (DWT) of the *MagNGx* signal. Showing in Table (3.5).

Table 3.5: Frequency domain features

Wavelet Domain Features and their use cases for Gait analysis

21	*Wavelet Bandwidth**	$\frac{cA' * cA}{(cA' * cA + cD' * cD)}$	*The relative energy contribution in a time-frequency band*
22	*Wavelet Entropy Rate**	$-\sum possibility_{unique\ freq} \times log_2(possibility_{unique\ freq})$	*Wavelet entropy represents the signal disorder in the time-frequency domain*

3.4.1 Statistical Features

Statistical features were calculated directly from x_i as shown in Table (3.6). The table shows feature names, descriptions, and definitions.

Table 3.6: Statistical-domain features

Statistical Features and their use cases for Gait analysis			
23	*Zeroth-Lag Cross-Correlation Coefficient**	$\frac{\sum(x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum(x_i - \mu_x)^2 \sum(y_i - \mu_y)^2}}$	*The agreement or similarity between 2 directional acceleration signals*
24	*Kurtosis**	$\frac{\frac{1}{n}\sum(x_i - \mu_x)^4}{\left[\frac{1}{n}\sum(x_i - \mu_x)^2\right]^2}$	*The extent to which the distribution of signal amplitudes lies predominantly on the left of the mean amplitude*
25	*Standard Deviation**	$\sqrt{\frac{1}{n}\sum(x_i - \mu_x)^2}$	*Measure for signal spreading, defined a the square of standard deviation*

3.4.2 Lifestyle Features

Lifestyle features were extracted directly from the mPower dataset surveys. Those features are shown in Table (3.7).

Table 3.7: Lifestyle features extracted

Lifestyle Features and their use cases for Gait analysis		
26	*GELTQ.1a*	*The number of times the participant performed strenuous exercise for more than 15 minutes over the past week.*
27	*GELTQ.1b*	*The number of times the participant performed moderate exercise for more than 15 minutes over the past week.*
28	*GELTQ.1c*	*The number of times the participant performed minimal effort exercise for more tha 15 minutes over the past week.*
29	*Smoked*	*Ever smoked? (True/false question).*
30	*Age*	*Participant's age (a number in years).*

31	*Years. Smoking*	*Number of years participant has smoked (a number)*
32	*Packs.per.day*	*The number of packets smoked per day.*
32	*Gender*	*Female/male*

3.5 Features Selection

Feature selection is the process of minimizing the number of input variables when developing a predictive model. Having irrelevant features can increase the training time and degrade the models' overall performance. Limiting the model learning to the relevant features only, reduces the computational cost, overfitting, and training time. After feature extraction, we performed ML-based feature selection, due to its advantages as explained in the sections below.

3.5.1 Statistical Feature selection

Statistical-based feature selection involves evaluating the relationship between input variables and the target variable (or label) using statistical measures and selecting those variables that have the strongest relationship with the label. These methods can be fast and effective when the relation between input variables and the label is linear and predictive. However, statically feature selection does not do well when the relation is non-linear. Other feature selection mechanisms will be used in that case.

3.5.2 Machine learning-based Feature selection

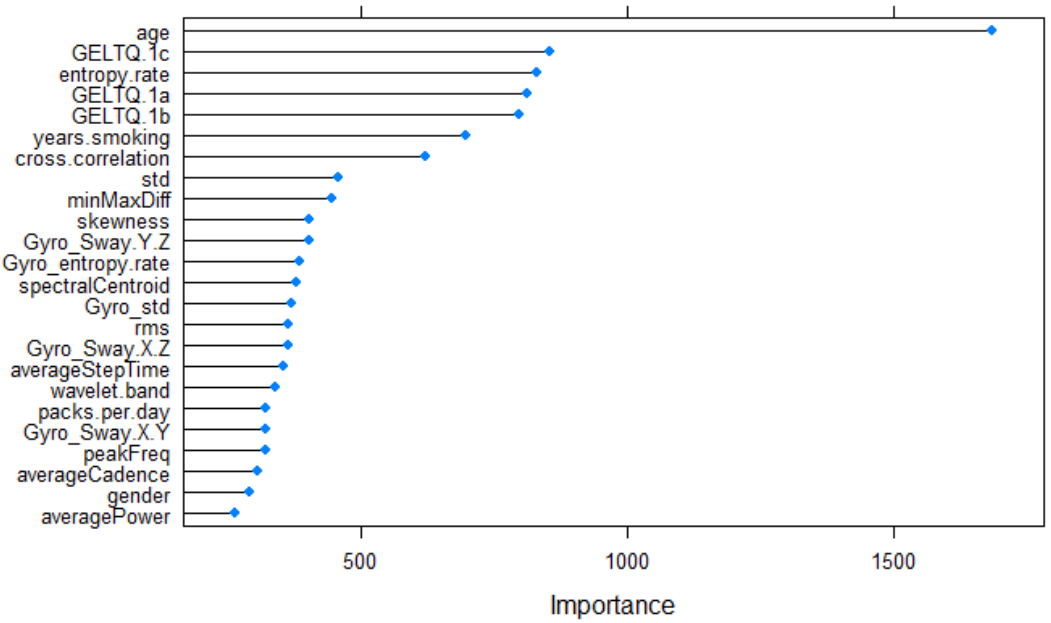
We established which features had statistically significant correlations with the MDS-UPDRS surveys and quantified the level of walking anomalies while patients walked for 30 seconds in a straight line. The walk data were labeled using participant self-assessments of their walk, which were used as labels for ML models. In constructing a decision tree, the importance of each feature is calculated by the decrease in the prediction error (mean squared error) and the increase of information gain, when the decision tree is split by the feature variable.

Figures (3.6-3.8) below show the degree of importance of the selected features with and without consideration of lifestyle features.

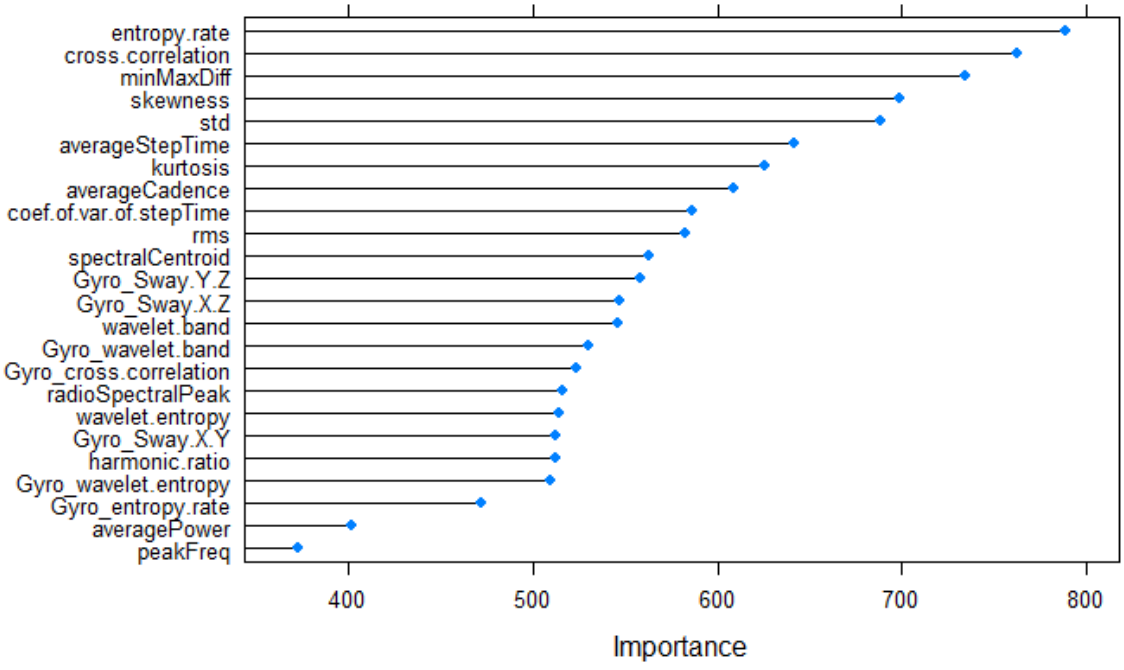
From Figure 3.6, the entropy rate is the most important feature for differentiating the walking balance severity, The mean of entropy rate decreases with the increase in gait severities due to the irregularity of the walking signal associated with PD patients, which was captured by the accelerometer. This finding agrees with the results of our prior work. As a demographic feature, age supersedes any calculated gait feature, which agrees with the findings in prior work **[90]** and **[96]**. However, the effect observed in this study is not the effect of aging on walking balance but is mostly a result of PD complications, as we address further in the discussion section.

Shaking and tremor can be inferred from gait features **[73].** However, most prior studies were based on a limited number of participants, as explained in Table (2.7)**.** Using data from participants in the mPower dataset, the most important features that discriminate the level of shaking/tremor were identified in Figure (3.7). We noticed that lifestyle features are strongly important in classifying shaking/tremor. We also noticed that multiple gyroscope features are highly important in predicting shaking/tremor.

Freeze of gait (FoG) has been studied extensively, and detection of FoG using smart sensors is possible **[67,71].** Previous studies were based on a small number of participants and did not study the effect of lifestyle features and sway area as a gait feature. We were able to discriminate the severity of FoG reported by PD patients. We noticed that both the accelerometer and gyroscope features are highly predictive of FoG, as shown in Figure 3.8.

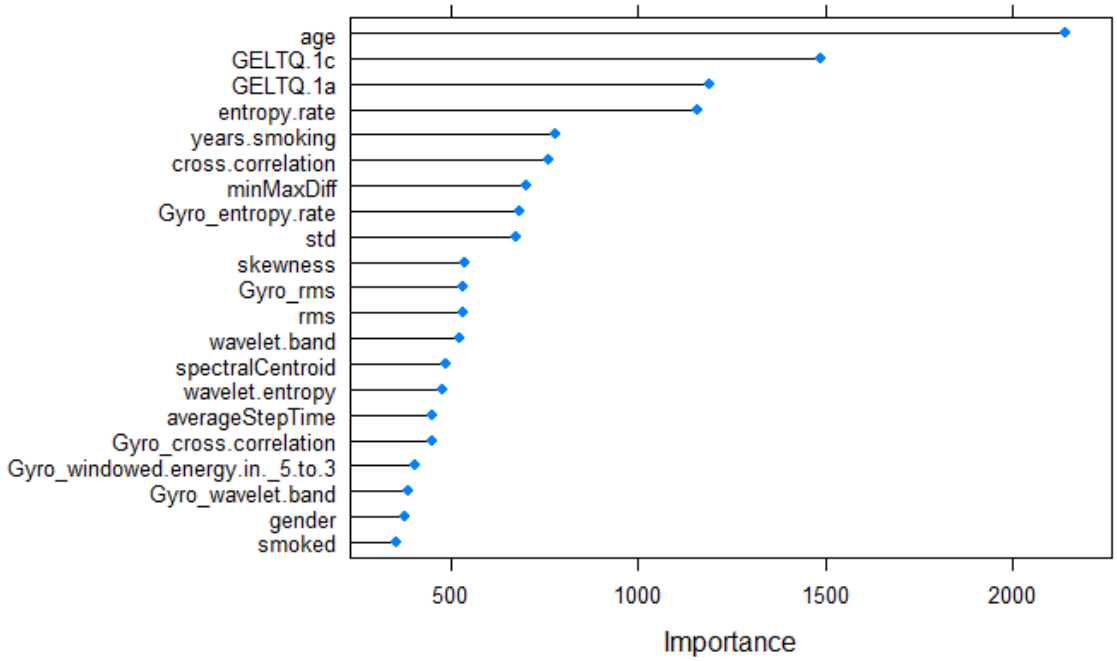


(3.6a) Features by Importance for Walking Balance including lifestyle features

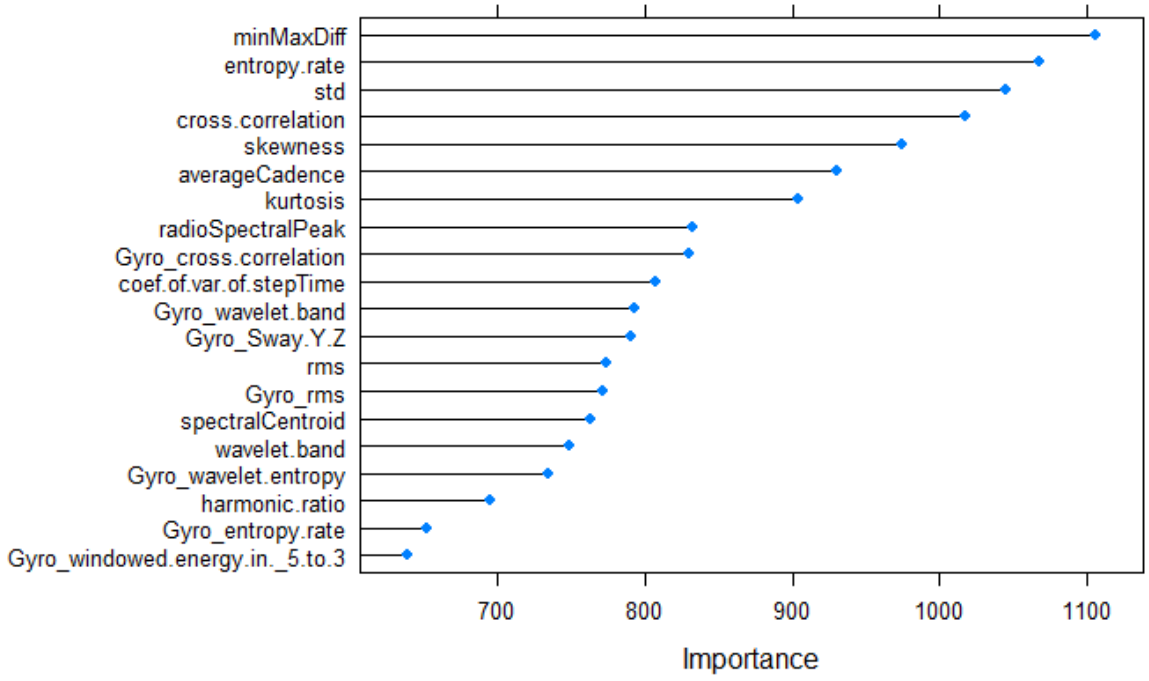


(3.6b) Features Importance for Walking Balance

Figure 3.6: Features Importance – Walking Balance

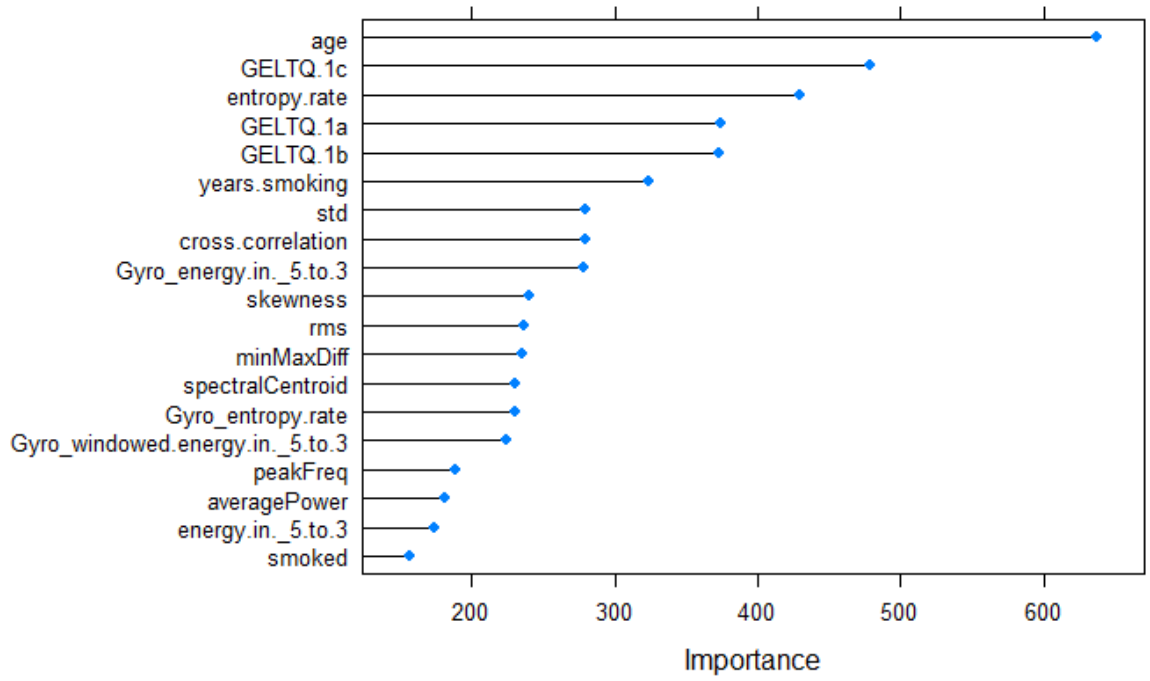


(3.7a) Features by Importance for Shaking/Tremor including Lifestyle

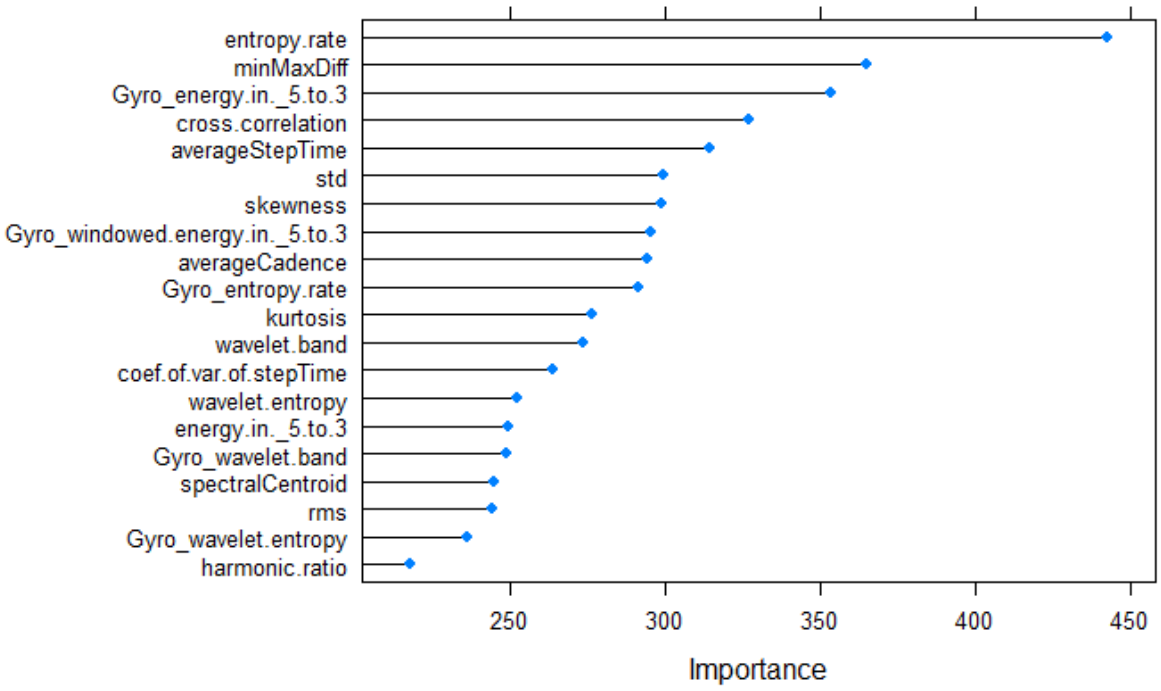


(3.7b) Features Importance for Shaking/Tremor

Figure 3.7: Feature Importance – Shaking/Tremor



(3.8a) Features by Importance for FoG including Lifestyle



(3.8b) Features Importance for FoG

Figure 3.8: Feature Importance – FoG

3.6 Encoding Walking Signals as Images

The main reason behind the boom of DL models is the success of computer vision and image classification applications. To leverage the development brought by computer vision models, we encoded the walking signal into an image format. There are multiple ways we considered to covert time-series to an image including; color-coded spectrogram (CCS), Markov Transition Field (MTF), and Gramian Angular Field (GAF). We choose GAF due to its impressive results for different time-series datasets **[97]**.

Given a Time series $\{x_1, x_2, \ldots, x_n\}$, GAF can be calculated by first normalizing the signal using the equation:

$$\tilde{x}_i = \frac{(x_i - \min(x_i)) + (x_i - \max(x_i))}{(\max(x_i) - \min(x_i))}$$

From the normalized time series and the time stamp values t_i we can calculate the polar representation using the equation:

$$\begin{cases} \phi = \arccos(\tilde{x}_i), -1 \leq \tilde{x}_i \leq 1, \quad \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, \quad t_i \in \mathbb{N} \end{cases}$$

Where the time series consists of N timestamps t_i. The resulting map on the polar coordinates is unique and reversible for each time series. The polar representation also preserves the absolute temporal relations, so that we will not lose these temporal patterns when converting the walking signal. To capture the correlation between different time intervals the gram matrix is calculated according to the following equation:

$$G = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \cdots & \cos(\phi_1 + \phi_n) \\ \vdots & \ddots & \vdots \\ \cos(\phi_n + \phi_1) & \cdots & \cos(\phi_n + \phi_n) \end{bmatrix}$$

Figure (3.9) below shows the GAF conversion of the 10-Strides walking signal and its equivalent polar representation. All the selected participants' walks were encoded to images

using the above GAF encoder and made available for further processing by DL image-classification models.

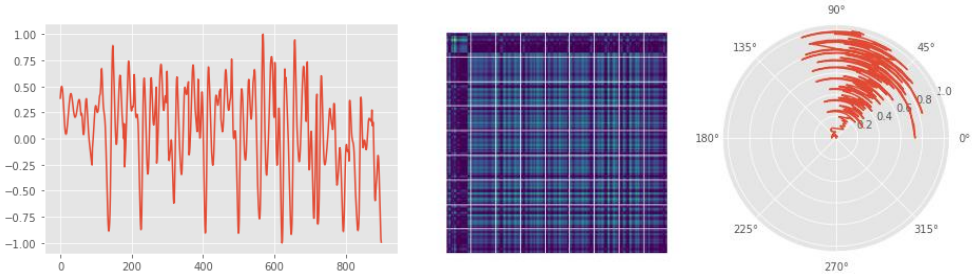


Figure 3.9: GAF for a walking signal and the corresponding polar representation

3.7 Transfer Learning

Transfer Learning (TL) is a machine learning method where a pre-trained baseline model can be reused as the starting point for a second model on a new ML problem Figure (3.10). TL allows repurposing the trained model, which is often trained on a large amount of data, on a new task as an optimization that allows rapid progress when modeling the second task.

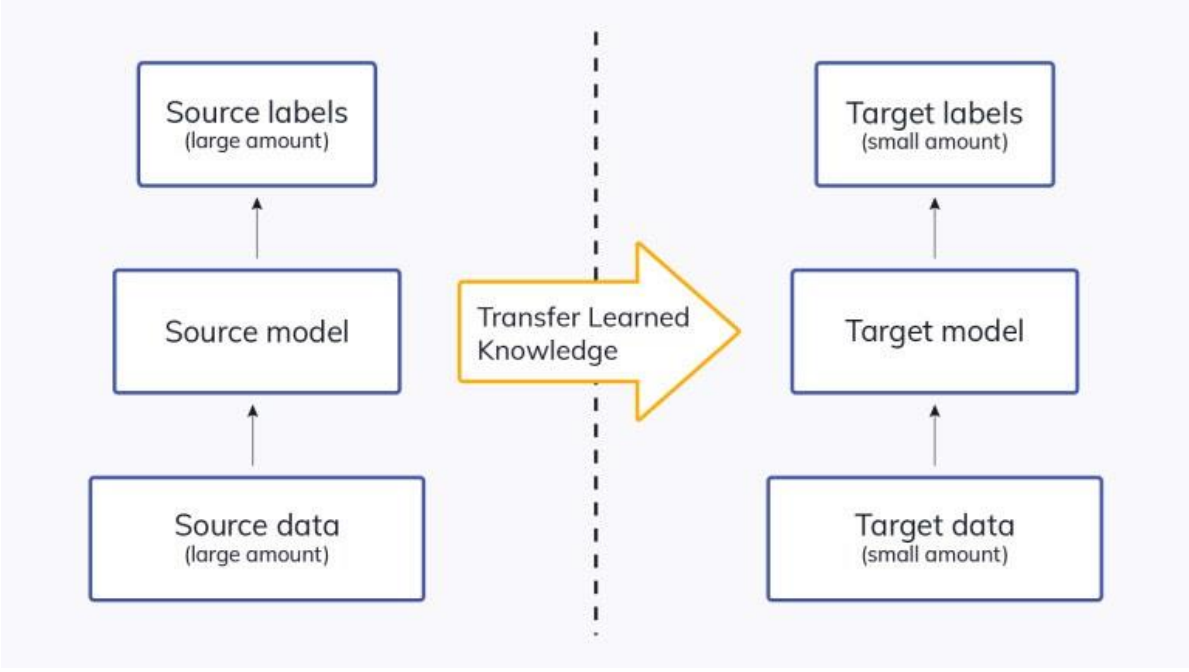


Figure 3.10 Transfer Learning

TL helps with solving ML problem with limited data, whereby applying transfer learning to a new task, one can achieve significantly higher performance than training with only a small amount of data.

In image classifications, researchers prefer to start from a pre-trained model that already knows how to classify objects and has learned general features like edges, shapes in images, than training a model from scratch. ResNet, AlexNet, and Inception are typical examples of models that have the basis of TL, therefore they are used as a baseline models for TL.

CHAPTER 4
MACHINE LEARNING GAIT ANALYSIS

In this chapter, we present the prediction of PD gait aspects based on supervised ML classification using hand-crafted features. The overall methodology that we followed is illustrated in Figure (4.1). From the accelerometer and gyroscope sensor data, statistical, time, wavelet, and frequency domain features were extracted, in addition to other lifestyle features that were derived directly from participants' survey data. Supervised classification experiments were conducted using 10-fold cross-validation and the model precision, accuracy, and area under the curve (AUC) were measured and reported.

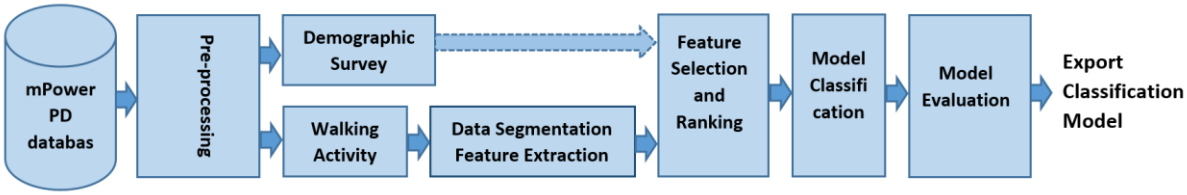


Figure 4.1: Flow Diagram for data collection, feature extraction, and classification

4.1 MODEL EVALUATION

To ensure efficient use of all the data, supervised classification experiments was performed using 10-fold cross-validation and measured the model precision, accuracy, and area under the curve (AUC). Cross-validation was performed by partitioning data into 10 disjoint folds at the population level. For each fold, the model was trained using the out-of-fold observations. Then the model performance was assessed using the in-fold data. The average performance metrics were calculated over all folds. Cross-validation requires multiple fits but gives a good estimate for the predictive accuracy of the final model trained and tested with all the data.

4.1.1 EVALUATION METRICS

• **Accuracy**: The percentage of correctly predicted samples over the total number of samples. Accuracy formulated as

$$= \frac{\text{Correctly Classied Samples}}{\text{Total Testing Samples}}$$

• **Precision**: Precision is the fraction of the correctly predicted samples, to the total positive predicted samples.

Precision measures the model's robustness against false positives.

$$= \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

• **Recall (**Sensitivity**)**: Is the fraction of the correctly predicted positive samples to the total positive samples in the Class. Recall gives an idea of the classification misses.

$$= \frac{\text{True Positive}}{\text{True Positive + False Negative}}$$

Area Under Curve (AUC): is the area under the curve for the plot of sensitivity over 1-specificity (True Negative / [False Positive +True Negative]) across thresholds

4.2 INFERRING MDS-UPDRS

Using only the selected features shown in Figures (3.6-3.8), we investigated multiple ML algorithms families, including Decision Trees (DT), Discriminant Analysis (DA), Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Ensemble Classifiers (EC) including Random Forest (RF) and Logistic Regression (LR) for classifying the walking balance, tremor/shaking and freezing of gait of 152 PD and 304 HC subjects. For walking balance, entropy rate and cross-correlation were found to be the best features for classifying the walking balance severity, with p-values of 0.2792094 and 2.481161e-07, respectively. Note that a lower p-value does not necessarily guarantee better ML performance. [124]

4.2.1 Walking Balance (WB)

Random forest is the best classifier for distinguishing between walking balance severities, with an accuracy of 93%, precision of 92%, and AUC of 0.97. Table (4.1) compares the performance, accuracy, and AUC of each classifier type.

Table 4.1: Walking Balance (WB) Evaluation

4.2 Walking Balance:		***Precision***	***Accuracy***	***AUC***
Accelerometer, Gyroscope Posturography and Lifestyle Features	**Random Forest**	**92%**	**93%**	**0.97**
	Bagged Trees	88%	90%	0.95
	Cubic SVM	72%	81%	0.92
	Weighted KNN	63%	82%	0.86
	Logistic Regression	71%	72%	0.78
	Fine Tree	75%	83%	0.88
	Quadratic Discriminant	71%	71%	0.75

4.2.2 Shaking Tremor (ST)

For shaking/tremor**,** MinMaxRate and EntropyRate were found to be the best features for classifying severities. Bagged trees was the best classifier for distinguishing shaking/tremor severity with an accuracy of 95%, precision of 95%, and AUC of 0.92. as shown in Table (4.2) below.

Table 4.2: Tremor Evaluation

4.3 Shaking Tremor:		***Precision***	***Accuracy***	***AUC***
Accelerometer, Gyroscope Posturography and Lifestyle Features	Random Forest	85%	83%	0.93
	Bagged Trees	**95%**	**95%**	**0.92**
	Cubic SVM	63%	68.8%	0.86
	Weighted KNN	62%	68%	0.77
	Boosted Trees	71%	68%	0.83
	Fine Tree	60%	72%	0.87
	Linear Discriminant	48%	61%	0.74

4.2.3 Freeze of Gait (FoG)

For FoG, features including entropy rate, MinMaxRate, and gyroscope energy successfully discriminated FoG severity. Table (4.3) shows that Bagged Trees was the best classifier for distinguishing FoG severity with an accuracy of 98%, precision of 96%, and AUC of 0.98.

Table 4.3: FoG Evaluation

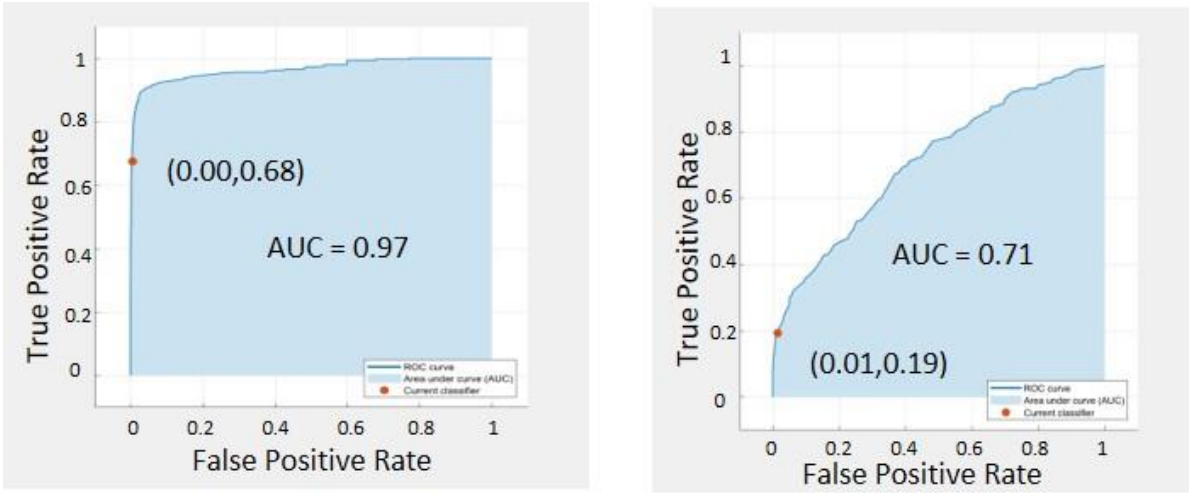
4.4 Freeze of Gait (FoG):		***Precision***	***Accuracy***	***AUC***
Accelerometer, Gyroscope Posturography and Lifestyle Features	Random Forest	92%	96%	0.90
	Bagged Trees	**96%**	**98%**	**0.98**
	Fine Gaussian SVM	92%	93%	0.96
	Weighted KNN	91%	92%	0.95
	Boosted Trees	90%	91%	0.93
	Fine Tree	93%	94%	0.95
	Linear Discriminant	89%	87%	0.71

4.3 Classifying Patients and Controls

One of the goals of our work is to be able to discriminate PD patients from healthy controls (HC) based on gait features. Using our selection of of 152 PD and 304 HC subjects, ML analysis was performed using the subject's response to the question "Have you been diagnosed by a medical professional with Parkinson disease?" on their enrolment questionnaire [**APPENDIX A**] . This question was answered one time when participants filled out a demographic survey to report whether they had ever been diagnosed with PD.

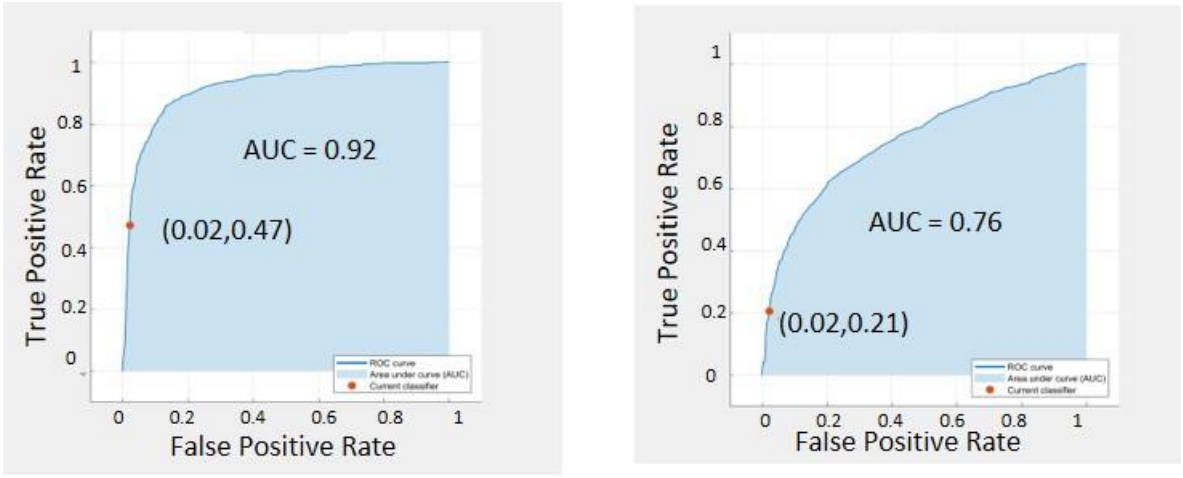
Cross-validation was performed by partitioning data into 10 disjoint folds at the population level. Six ML algorithms were used to classify participant gait features, namely, bagged trees, fine Gaussian SVM, subspace KNN, boosted trees, fine trees, and linear discriminant. Entropy rate and MinMaxDiff were the top features to successfully discriminate HC from PD. Random forest was the best for discriminating between PD and HC, with an accuracy of 95%, precision of 94%, and AUC of 0.99.

To isolate the effect of lifestyle and demographic/age features on model performance, we ran the ML classification algorithms using gait features while excluding lifestyle features, and the results of our classifiers were significantly degraded. For the walking balance, the best classifier results were random forest (Acc 77%, AUC 0.71). Figure (4.2) shows the effect on the total performance of the model and how the result deteriorated. For shaking/tremor, the accuracy of the bagged trees models degraded to Acc 72% and AUC 0.76, as shown in Figure (4.3). For FoG, the accuracy of bagged trees slightly degraded to Acc 91% and AUC 0.92 but was still accurate. It can be noted from Figure (4.4) how AUC slightly deteriorated when lifestyle features were excluded.



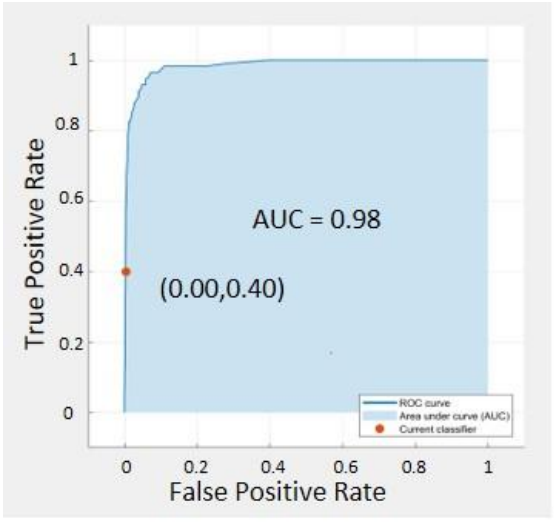
(4.2a) AUC for Walking Balance with lifestyle features. (4.2b) AUC for Walking Balance without lifestyle features.

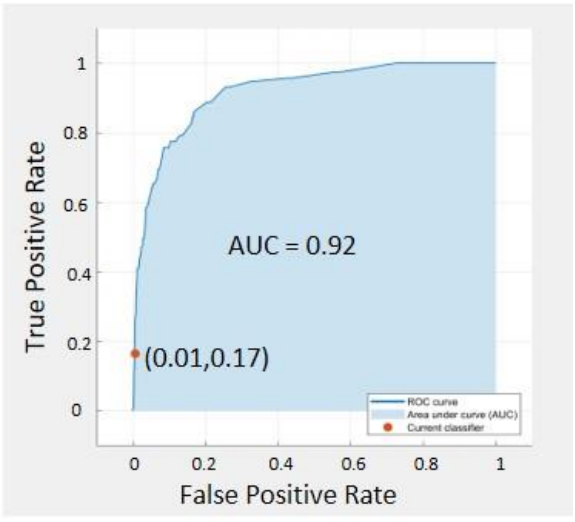
Figure 4.2: Walking Balance AUC



(4.3a) AUC for Shaking/Tremor with lifestyle features. (4.3b) AUC for Shaking/Tremor without lifestyle features.

Figure 4.3: Shaking/Tremor AUC





(4.4a) AUC for FoG with lifestyle features. (4.4b) AUC for FoG without lifestyle features.

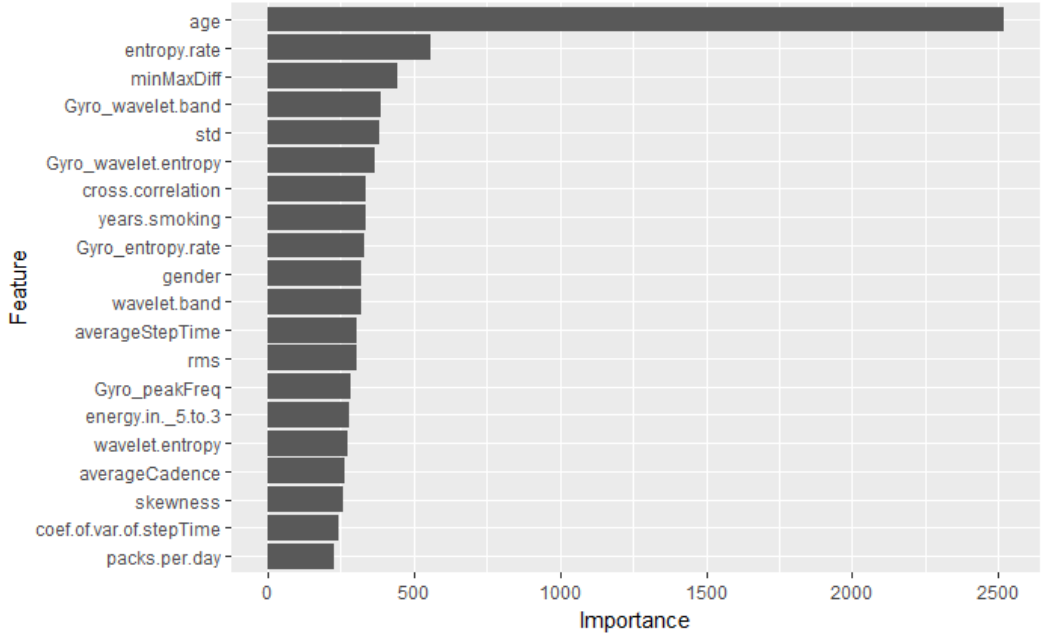
Figure 4.4: FoG AUC

Table 4.4: Comparison of ML algorithms for PD Patients vs. HC Classification

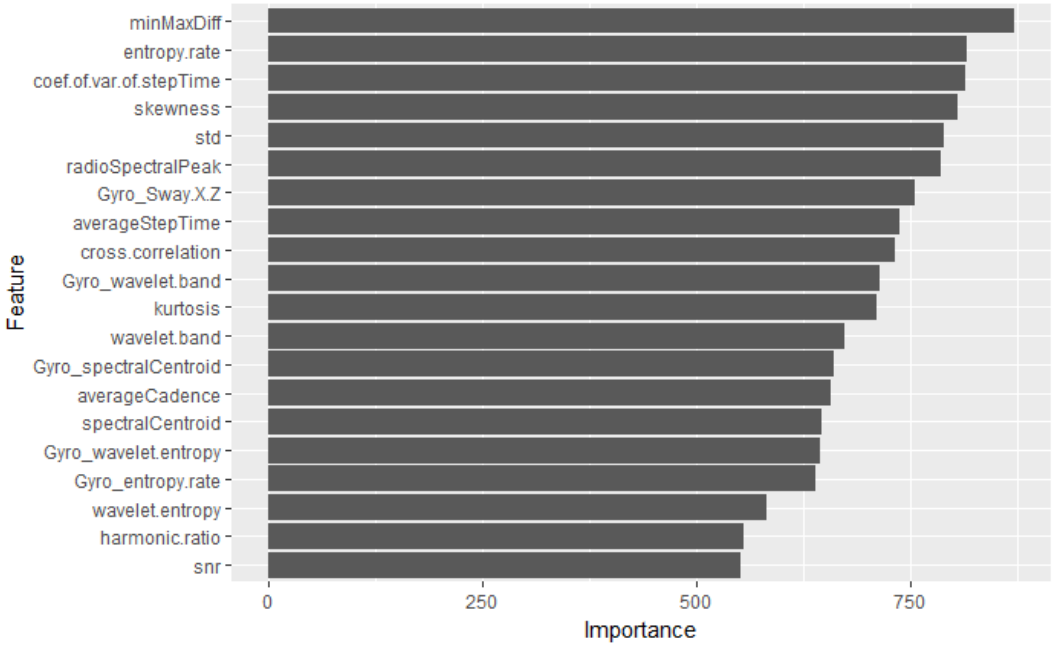
Classifier details		***Precision***	***Accuracy***	***AUC***
Accelerometer, Gyroscope Posturography and Lifestyle Features	**Random Forest**	**94%**	**95%**	**0.99**
	Bagged Trees	92%	93%	0.95
	Fine Gaussian SVM	88%	88%	0.96
	Subspace KNN	91%	90.4%	0.92
	Boosted Trees	84%	90%	0.97
	Fine Tree	90%	91.2%	0.96
	Linear Discriminant	83%	85.5%	0.91

In table (4.4) above, a comparison of the performance, accuracy, and AUC for the classifiers is presented. In this comparison, lifestyle features helped to significantly improve the result of classification. Specifically, the false positive rate is significantly higher (worse AUC curve) if lifestyle features are not included. Removing the lifestyle features led to a degradation of the classification results. Entropy rate and MinMaxDiff remained the top features. However, the accuracy of the random forest model deteriorated to Acc 82% and AUC 0.88. Figure (4.5) below shows the top 20 features selected for discriminating between PD patients and HC with and without

lifestyle features. Note that lifestyle features are less important in differentiating PD patients from HC.



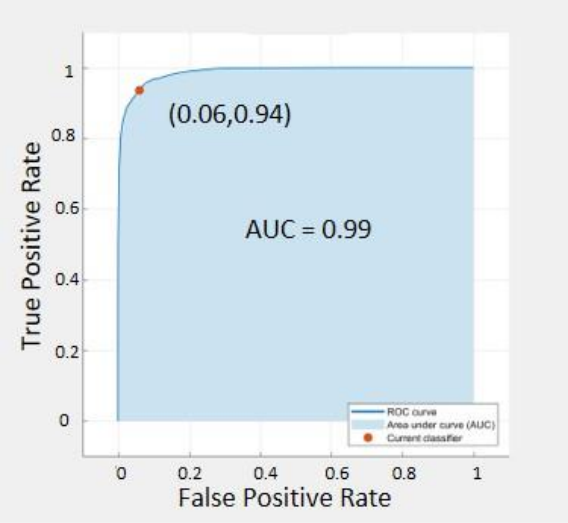
(4.5a) Feature Importance for PD Patients vs HC (including lifestyle)



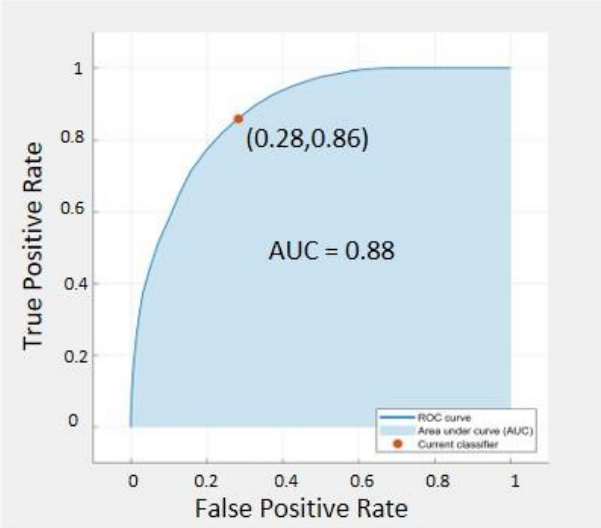
(4.5b) Feature Importance for PD Patients vs HC (no lifestyle)

Figure 4.5: Feature Importance of PD vs HC

The AUC also decreased significantly when lifestyle features were excluded. Figure (4.6) below shows the AUC for HC/PD classification with and without lifestyle features.



(4.6a) AUC for HC/PD Patient Classification with lifestyle features.



(4.6b) AUC for HC/PD Patient Classification

Figure 4.6 : AUC for PD patient vs HC classification with and without lifestyle features

4.4 Improving Model Accuracy with Ensemble methods

We attempted different variations of the ensemble random forest in order to tune its parameters. Random forest is composed of multiple estimators (decision trees) and aggregates their output to return the final ensemble result. If we have a classification problem with a data set in the form of $(X_{1,}Y_1), \ldots\ldots (X_{n,}Y_n)$, where X is a d-dimensional predictor variable and Y is a univariate response, to predict Y with J classes, $Y_i \in \{0, 1, \ldots, \mathrm{J}-1\}$, the target function of interest is $\mathrm{P}[Y = j|X = x]$ ($j = 0, 1, \ldots, \mathrm{J}-1$)

Random forest works by drawing a_n observations $(X'_{1,}Y'_1), \ldots\ldots (X'_{n,}Y'_n)$ at random with replacement from the original data set. These drawn observations are the only observations considered in growing M different randomized trees to obtain different estimates $g_1(\cdot), g_2(\cdot), g_3(\cdot) \ldots g_M(\cdot)$. In `R` random forest implementation, this number of (trees in the

forest) is represented by (ntree). The resulting estimator functions can be written as follows:

$g_n(\cdot) = h_{n,}\big((X'_{1,}Y'_1), \ldots\ldots (X'_{n,}Y'_n)\big)(\cdot)$

where the function $h_n(.)$ defines the estimator as a function of the dataset.

At each cell of each tree, a split is performed based on several variables mtry has been chosen randomly among the overall number of variables (p). The construction of individual trees is stopped when each cell contains less than nodesize points.

For any query point $x \in X_i$, each tree predicts Yi by growing the tree and making the final estimation that only depends on the a_n preselected data points. Because the overall decision is obtained via a majority vote among the classification trees, we can construct an ensemble-based function estimate gens (·) by taking linear combinations of the individual estimates:

$$g_{ens}(\cdot) = \sum_{k=1}^{M} \big(c_k g_{k)}\big)$$

For ensemble bagging and **[98]** original random forests, the linear combination coefficients c_k =1/M are averaging weights, which also result in variance reduction.

Tuning the forest parameters might result in a computational burden, particularly for large datasets with hundreds and thousands of observations and variables. Due to the manageable size of this study dataset, we tuned the following forest parameters with an affordable computational cost:

1- Number of trees to grow (ntree, _Acc)
2- Number of variables randomly sampled at each split (mtry_Acc)
3- Maximum number of terminal nodes (mx_Acc)
4- Minimum size of terminal nodes (nodeSize_Acc).

Figure (4.7) shows different parameters of the random forest ensemble method and their effects on the model performance. When addressing classification problems, it is usually recommended to set nodesize to 1, and mtry to $\sqrt{(p)}$ **[99]**.

Extensive discussion exists in the literature relative to the influence of `mtry` on the overall performance of the model. **[100]** show that different values of `mtry` did not affect the classification rates of their model and that other performance metrics (sensitivity, specificity, kappa, and ROC AUC) were stable under different values of `mtry`. However, **[101]** show that `mtry` had a strong influence on predictor variable importance estimates. Additionally, **[102]** claims that the default value of `mtry` is too small. Therefore, their approach was to make `mtry` as large as possible (limited by available computing resources). We do not fully agree with the last finding, and we noticed that the overall accuracy improved significantly by increasing `mtry`. However, the relationship is not linear because the accuracy was maximized when `mtry` is 60% and 70% of p. By default, the maximum number of leaf nodes is set to the maximum possible. We experimented with limiting this parameter `mx`, which led to a negative impact on the overall accuracy.

It is clear that the forest variance decreases as M grows. Thus, more accurate predictions are likely to be obtained by choosing a large number of trees (`ntree`). The computational cost of increasing a forest increases linearly with M, and thus a good choice results from a trade-off between computational complexity and accuracy. Finally, the default value of the parameter `nodesize` is 1 for classification and 5 for regression. These values are often reported as good choices **[103]**, and even though this selection is not supported by solid theory, our results agree with those findings.

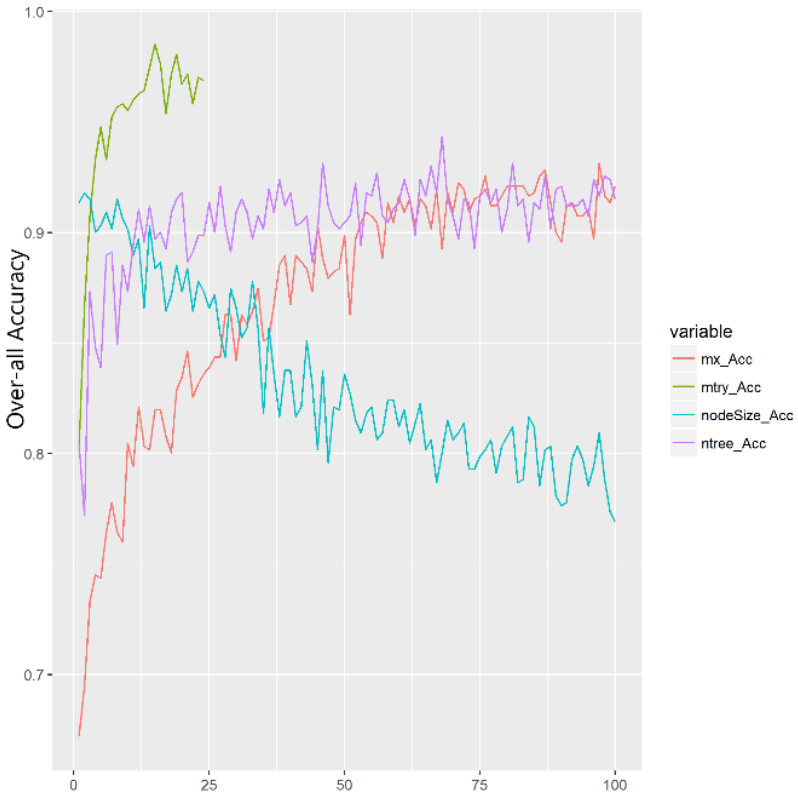


Figure 4.7: Random Forest Parameter Tuning

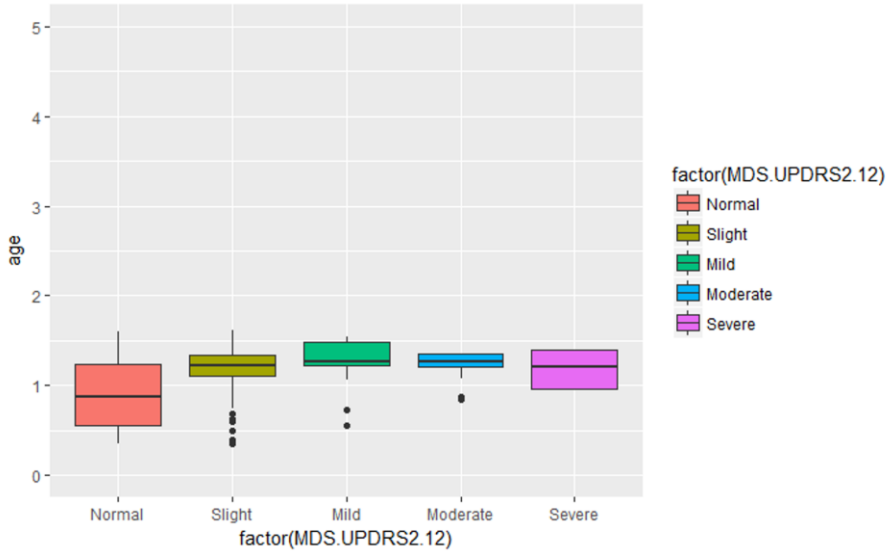
4.5 Top Features

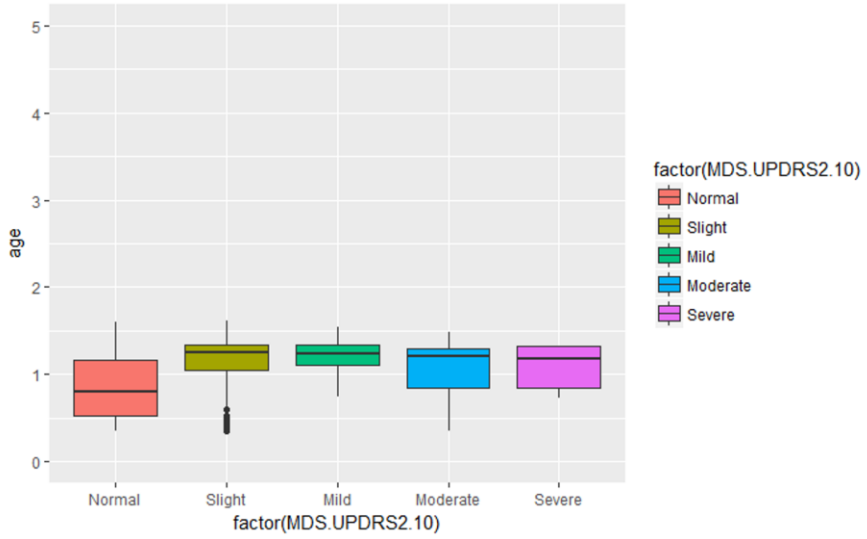
The ensemble model uses decision trees with high variance and low bias as base learners. At each node of a decision tree, the split feature is found based on information gain (I.G.) or the more computationally low-cost Gini impurity reduction method. The information gain due to a feature summed across all the levels of decision trees determines its feature importance. Random forest and bagged trees are composed of multiple decision trees, and thus the importance of a feature is the normalized sum of I.G. delivered by that feature across all trees. The output of these separate trees is aggregated and returned as the final ensemble result.

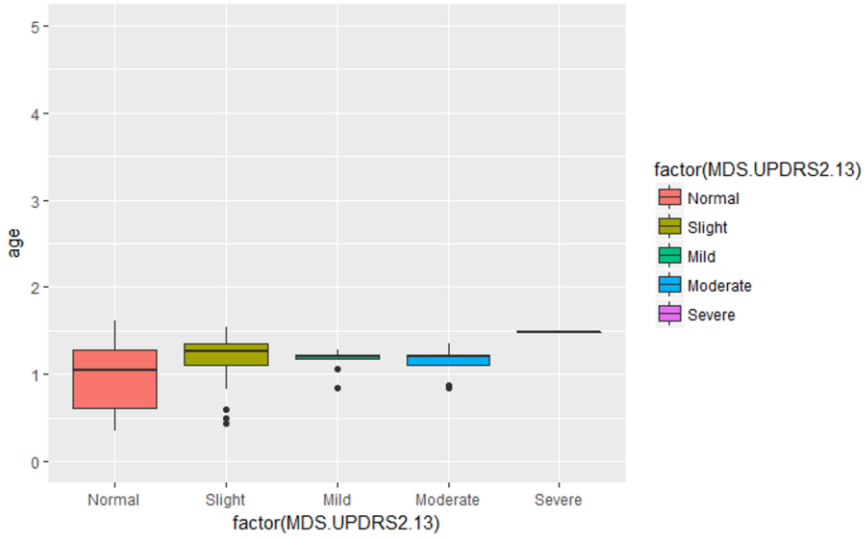
The correlation coefficient is commonly used to evaluate the degree of linear association between two variables. However, it can be shown that a correlation coefficient close to one might also be obtained for a clear curved relationship, depending on the nature of the ML algorithm used, and the selection of features based on correlation can be misleading. We found that the selection of features based on the ML ensemble led to a set of features with high

predictive power when used with nonlinear algorithms. Some of our top features do not have a linear relationship with the response variable. For example, age does not correlate linearly with the label, as shown in Figure 4.8a. Prior studies have found that PD incidence rates for both men and women increased rapidly after the age of 60 years **[90]**. Based on our normalized age feature, elderly people, in general, have higher severities, but as Figure (4.8a) shows, severity does not necessarily increase with age.

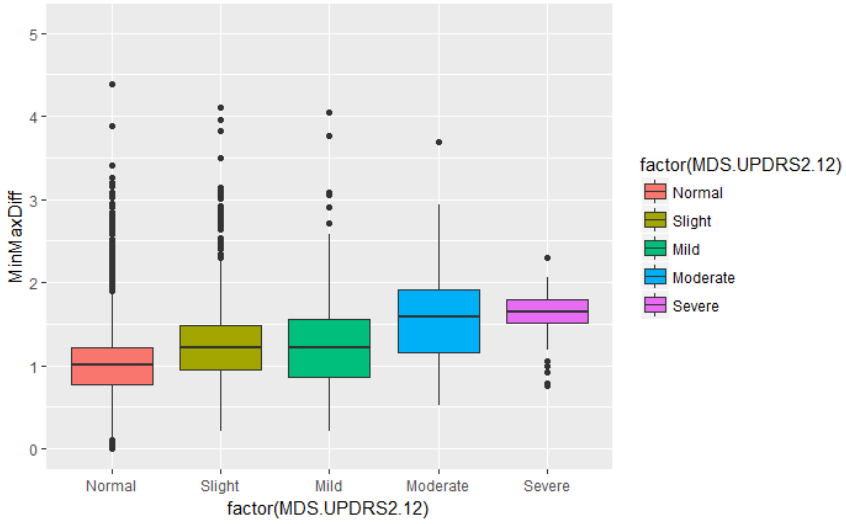
In contrast, top gait features (entropy rate and minMaxDiff) correlate linearly with gait severities. Figure (4.8b) shows that minMaxDiff always increased as gait severities worsened, which occurs due to differences in step swing that are captured by accelerometer peaks. The mean of entropy rate (Figure (4.8c)) decreases with the increase in gait severities due to the irregularity of the walking signal associated with PD patients, which is captured by the accelerometer. Please note that the gait features in Figure (4.8) are normalized at the participant level, whereas age is normalized at the population level.

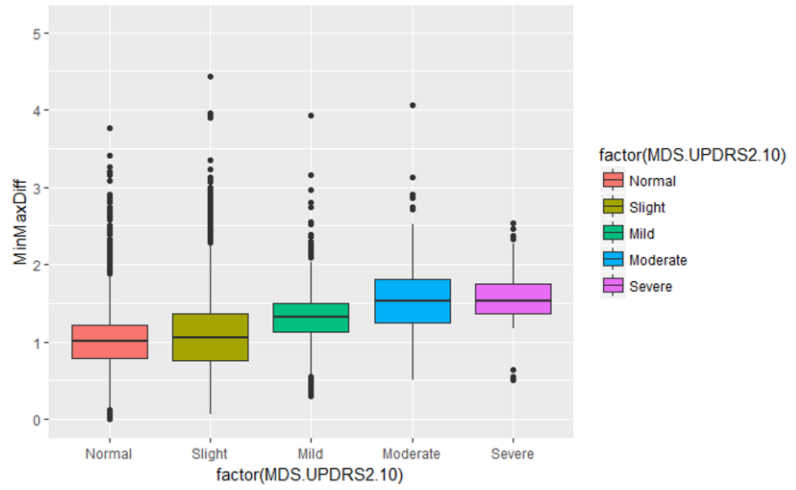


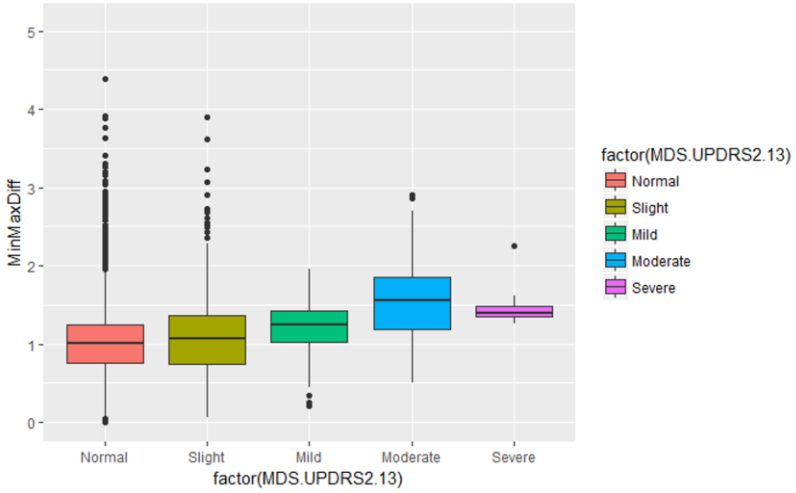




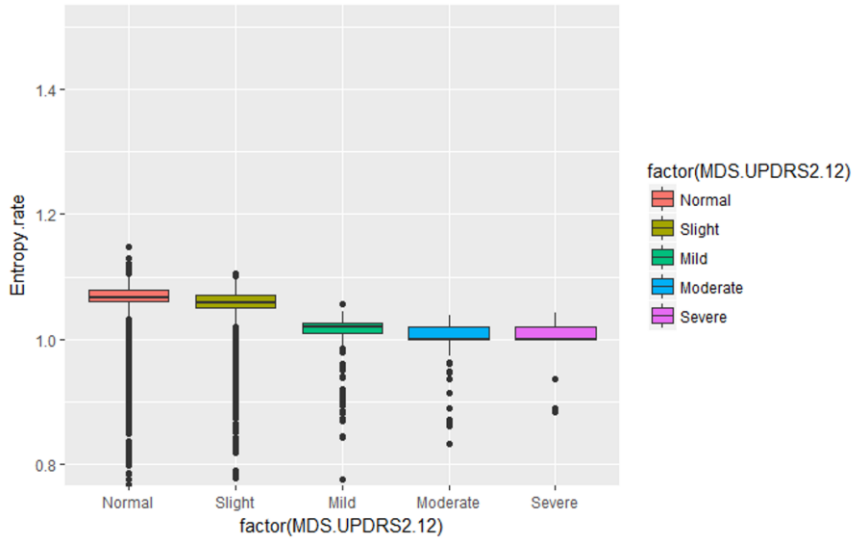
(4.8a) Age vs. Severity of Gait anomalies.

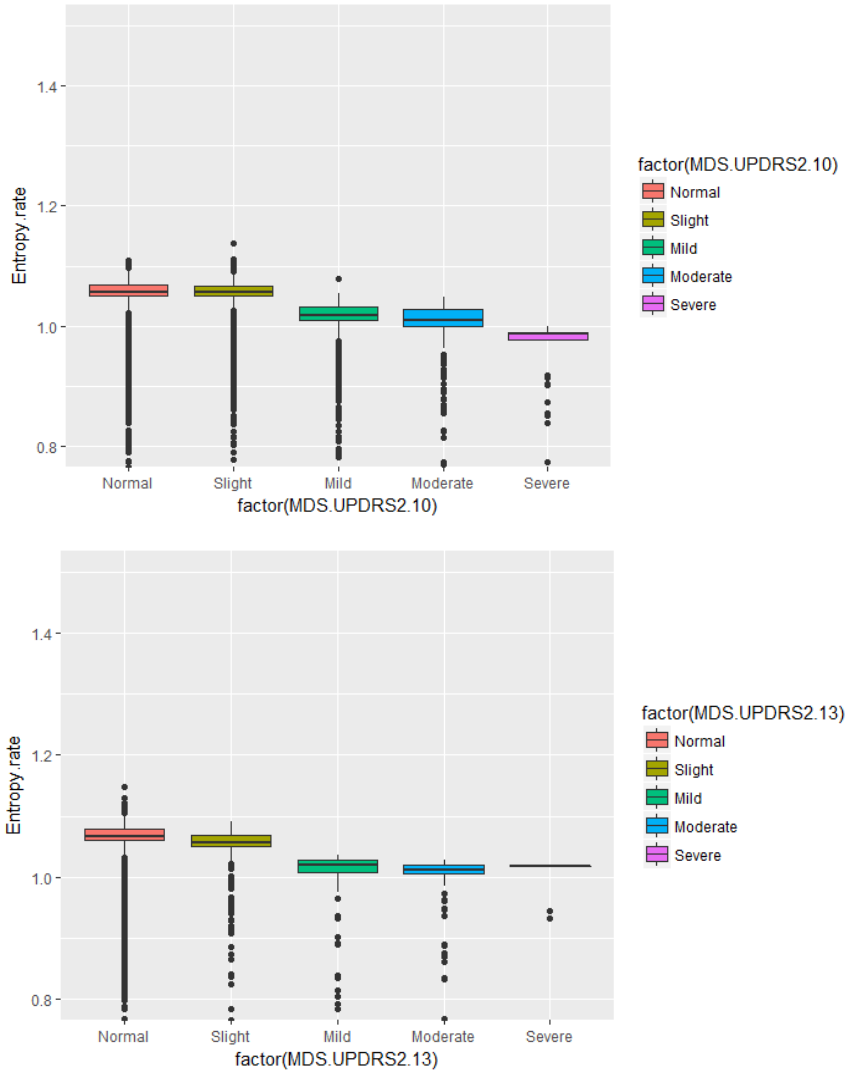






(4.8b) MinMaxDiff vs. Severity of Gait anomalies.





(4.8c) Entropy rate vs. Severity of Gait anomalies.

Figure 4.8: Top Features vs. Severity of Gait anomalies

In our analyses, we found that the gyroscope sway areas contributed significantly to the classification of walking balance, as shown in Figure (3.6a), but it had only minor contributions to shaking/tremor classification and no contribution to FoG classification. The accelerometer sensor and features were more useful in classifying shaking/tremor and FoG. Gyroscope-based analyses of sway can supply a powerful tool for early clinical trials and for monitoring the treatment efficacy for balance disorders in PD patients. Gyroscope sway area calculation can also be used in the online assessment of MDS-UPDRS walking balance (MDS-UPDRS 2.12).

4.6 Chapter Summary

Remote measurement of gait has become an important tool for monitoring the progression of PD. Although measurements reduce hospital visits and offer convenience to both PD patients and the healthcare provider, the validity of these measurements compared with assessments in the clinic continues to be a challenge. In our work, we addressed the unique gait characteristics of PD and inferred the stage of each PD gait modality through machine learning classification of smartphone sensor data collected by a mobile health application. This work contains three main contributions in this regard: (1) Combination of time, frequency, and statistical features with sway area and lifestyle features to remotely infer the level of PD walking modalities for a large set of participants; (2) identification of the most important features that offer deeper ailment understanding and classification of PD gait modalities; and (3) determination of the best ML algorithm for analyzing each gait modality and the one that best discriminates PD patients from HC. Although the classification results were affected by the subjective nature of PD labels assigned by patients based on their responses to the MDS-UPDRS questions, we were able to demonstrate with a relatively large number of participants that remote and automatic PD patient classification based on sensor activity data can supply objective assessments of PD-related gait patterns and severity of gait anomalies, which ultimately has the potential to improve remote healthcare for PD patients.

CHAPTER 5
DEEP LEARNING GAIT ANALYSIS (DEEPAGAIT)

In this chapter, we present the prediction of PD gait aspects based on DL approach. The overall methodology that we followed is illustrated in Figure (5.1). From the accelerometer and gyroscope sensor data we extracted features using DeePaGait; our DL multi-layer Conventional Neural Network (CNN). DeePaGait operates on 1Dimensional convolution filters to classify 30 seconds of walking data into one of five severity levels (Normal, Slight, Mild, Moderate, Severe). We conducted supervised classification experiments and measured the model precision, accuracy, Recall, and F1-score.

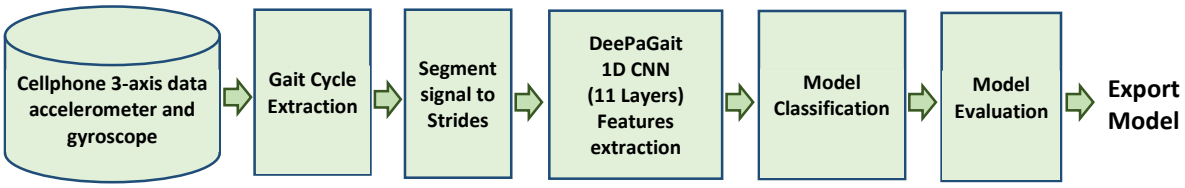


Figure 5.1 : Flow Diagram for data collection, feature extraction, and classification using DeePaGait

5.1 DEEPAGAIT NETWORK ARCHITECTURE

Our DeePaGait network consists of the feature extraction layers followed by fully-connected and classification layers, as shown in Figure (5.2). Feature extraction consists of 3 sets of 1D convolutions layers, the first two sets are followed by pooling layers, to downsize the feature map and reduce model complexity. The last set is followed by a fully-connected (FC), and softmax classification layer. Two dropout layers have been added to reduce the model overfitting and to improve the model's overall performance. Table (5.1) lists all of the DeePaGait layers, the input size of each layer, and the number of trainable parameters. "None" refers to variable batch size. We set an input image patch size of 128 in our classification experiments.

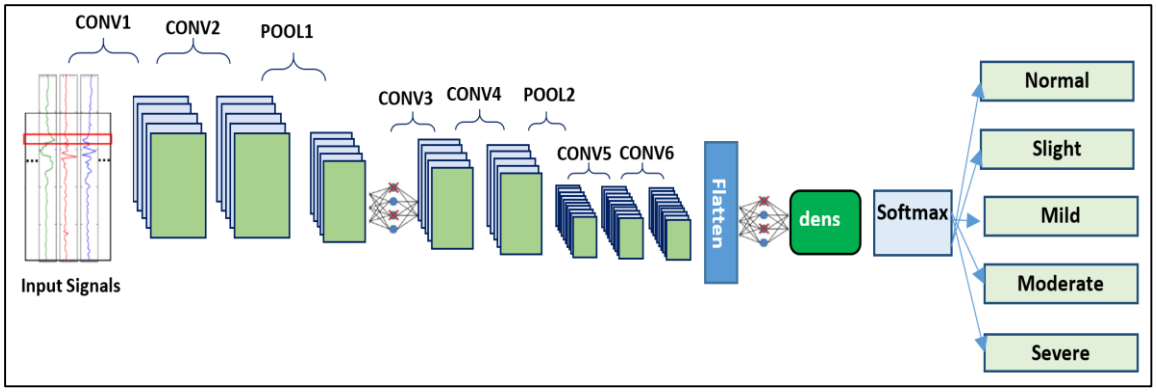


Figure 5.2 : DeePaGait Network Architecture

Table 5.1: DeePaGait CNN structure and parameters

Layer (type)	**Output Shape**	**Param #**
CONV-1 (Conv1D)	(None, 900, 256)	7936
CONV-2 (Conv1D)	(None, 898, 64)	49216
Pool-1 (MaxPooling1D)	(None, 449, 64)	0
dropout_1 (Dropout)	(None, 449, 64)	0
CONV-3 (Conv1D)	(None, 447, 64)	12352
CONV-4 (Conv1D)	(None, 445, 64)	12352
Pool-2 (MaxPooling1D)	(None, 222, 64)	0
CONV-5 (Conv1D)	(None, 220, 64)	12352
CONV-6 (Conv1D)	(None, 218, 64)	12352
Flatten_1 (Flatten)	(None, 13952)	0
dropout_2 (Dropout)	(None, 13952)	0
FC (Dense)	(None, 32)	446496
Softmax Classification (Dense)	(None, 5)	165

Total params: 553,221

Trainable params: 553,221

Non-trainable params: 0

5.2 EVALUATION

In our work [**62,63]**, we were able to discriminate PD patients from HC based on their gait. We also successfully classified gait severity using Traditional ML models based on handcrafted and lifestyle features. By using DL analysis, we extend our work to study the

most challenging gait symptoms of PD, which are: Walking Balance, Shaking, and Freeze of Gait. In the following sections, we present the results of our innovative DeePaGait algorithm, and then we compare DeePaGait performance to several variations of LSTM deployments, in addition to the state-of-the-art pre-trained, CNN image-classification networks. To evaluate the model performance, the following widely-used metrics were adopted:

5.2.1 EVALUATION METRICS

• **Accuracy**: The percentage of correctly predicted samples over the total number of samples. Accuracy formulated as

$$= \frac{\text{Correctly Classied Samples}}{\text{Total Testing Samples}}$$

• **Precision**: Precision is the fraction of the correctly predicted samples, to the total positive predicted samples.

Precision measures the model's robustness against false positives.

$$= \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

• **Recall (**Sensitivity**)**: Is the fraction of the correctly predicted positive samples to the total positive samples in the Class. Recall gives an idea of the classification misses.

$$= \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

F1-score: is the weighted average of Precision and Recall, which takes false positives and false negatives into account

$$= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

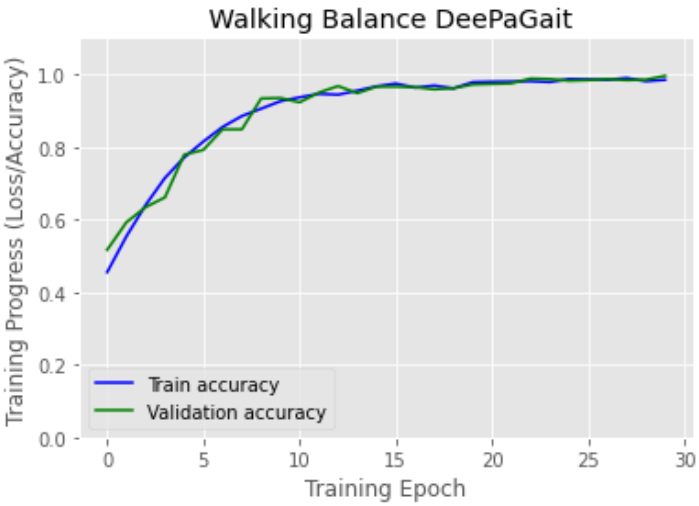
5.3 INFERRING MDS-UPDRS

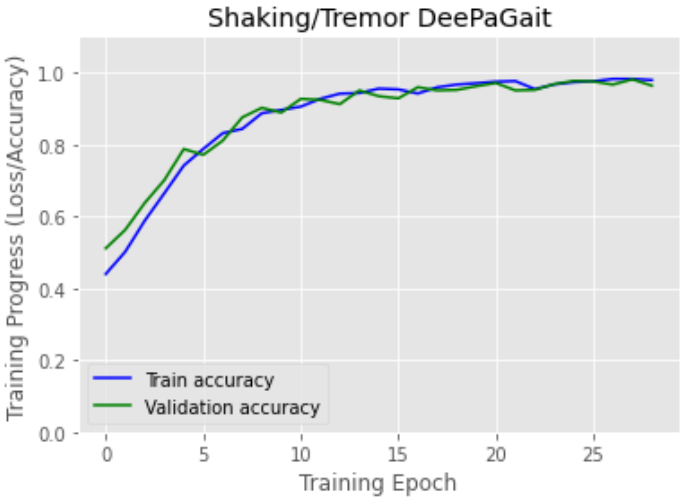
Based on 456 walking records from 152 PD participants, we conducted a supervised classification experiment by training the DeePaGait network on 90% of the data and testing on the remaining 10%. Our results including performance metrics are presented in Table (5.2). It can be noticed that DeePaGait outperformed traditional ML classifiers by at least 6% for the WB problem and 4% in the case of Shaking/Tremor classification.

Table 5.2: Performance metrics of DeePaGait

	Precision	***Recall***	***F1-Score***	***Testing Accuracy***
Walking Balance	***0.991***	***0.992***	***0.993***	***0.991***
Shaking/Tremor	***0.984***	***0.981***	***0.983***	***0.984***
Freeze of Gait	***0.983***	***0.981***	***0.982***	***0.982***

It can be seen from Figure (5.3) that DeePaGait is very adaptive to the gait classification problem compared to traditional ML [63], that is because: First, the mPower gait signals are non-linear and very noisy, which made the classification problem very challenging for ML. DeePaGait could adapt to the gait signals because it uses multiple layers with adaptive non-linear activation functions. Second, the depth of the network adapted to the data volume, and that reduced the effect of overfitting. Last, hand-crafted features are limited in their ability to extract very discriminative features without overfitting the training data.





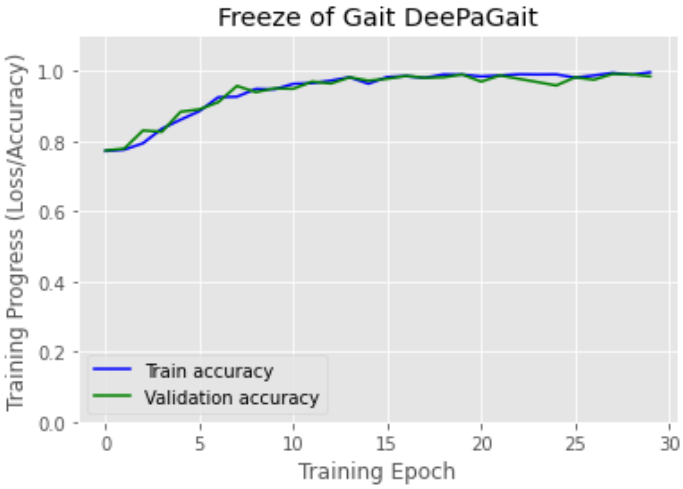


Figure 5.3: DeeePaGait Accuracy per epoch

5.4 Comparison of DeepaGait to Baseline Models

5.4.1 Image-based Models Evaluation

Classifying images on pre-trained DL models has become increasingly common. Those models can help the network learn the general-purpose features based on previous training on huge datasets. Therefore, pre-training saves building a model from scratch and can significantly improve the performance of medical imaging diagnosis **[104 -106]**. Examples of such pre-trained models include ResNet50, Inception, SqueezeNet, and EfficientNet, which we utilized as baselines for comparison with DeepaGait..

ResNet50

ResNet-50 is a 50-layers deep CNN. The pretrained version of the network is trained on more than a million images from the ImageNet database [107]. ResNet50 uses residual blocks and skip connections to combat the problems of vanishing gradient and network degradation. The pre-trained network can take input images of size, 224x224, and can classify the image into 1 of 1000 classes. The network has learned a rich feature representation for a wide range of images, which makes It a perfect candidate for many image classification problems such as the classification of Alzheimer's disease [125] , Breast cancer fiagnosis [126] and human gait identification [127].

Inception

Inception-v3 is a 48-layers deep CNN, that is based on the original work by Szegedy, et.al **[108]**. The main improvement that Inception network brings is the use of inception blocks. These involve convolving the same input with multiple filters and concatenating their results. Inception-v3 was pretrained on million images from the ImageNet database [107] and attained an accuracy greater than 78.1%. The model learned rich feature representation through it is building blocks that included convolutions, average pooling, max pooling, concatenations, dropouts, and fully connected layers. We considered the Inception-V3 network because it was able to classify images of size 299x299 into one of thousand classes.

EfficientNet

EfficientNet is a 237 layers CNN that takes images of the size 224x224. EfficientNet was trained on a million images from the ImageNet database **[107]** and fine-tuned on new datasets. The base EfficientNet-B0 network imploy the inverted bottleneck residual blocks to reduces the number of parameters and matrix multiplications. The pretrained EfficientNet was applied (as a TL) to various datasets, the results show that EfficientNets consistently achieve better accuracy parameters than existing CNN Models **[109].**

2.5.2.3 Transfer Learning (TL) Approach

To compare the performance of our DeePaGait model to the state-of-the-art retrained models, we slightly modified those models and evaluated them using our dataset of GAF images. the overall TL approach is shown in Figure (5.4) and outlined in the steps below:

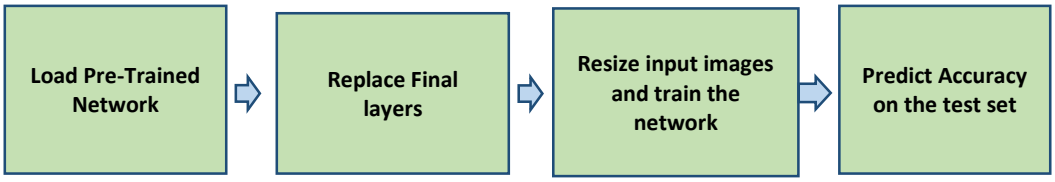


Figure 5.4 : Illustration of the mechanism for TL

Step1: All Accelerometer and Gyroscope signals were encoded to images representation using the GAF method.

Step2: Data were randomly split into 90% training set and 10% as a test set.

Step3: The last CONV, FC, and classification layers were adapted to match our 5 classes' gait severity data.

Step4: The network was trained.

Step5: The network was evaluated on the test set and the performance metrics reported.

Using the method of TL shown in Figure (5.5), we created transferred pre-trained models, trained, and tested ResNet50, Inception, and EfficientNet networks.

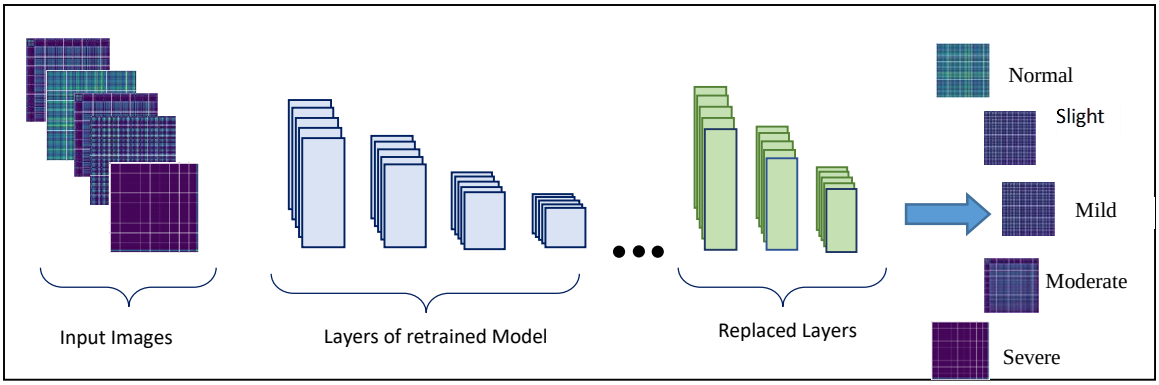


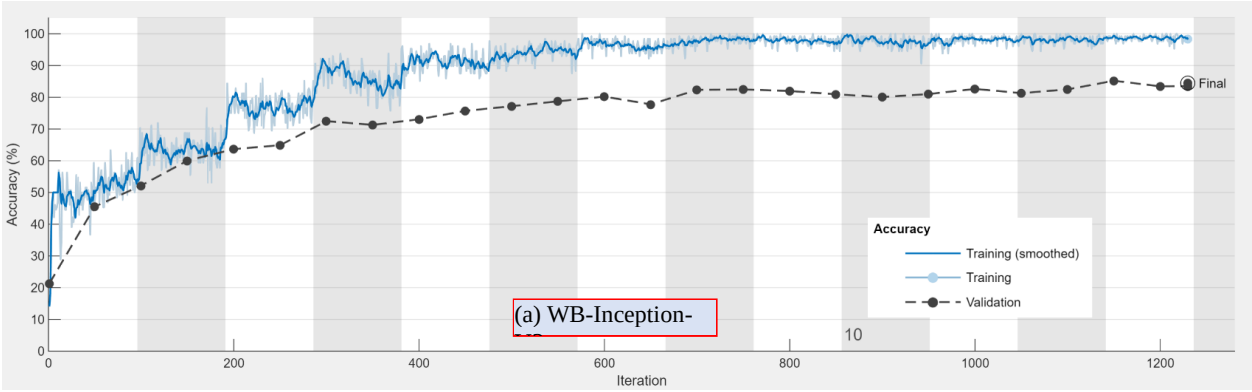
Figure 5.5 : TL using the Pre-trained model

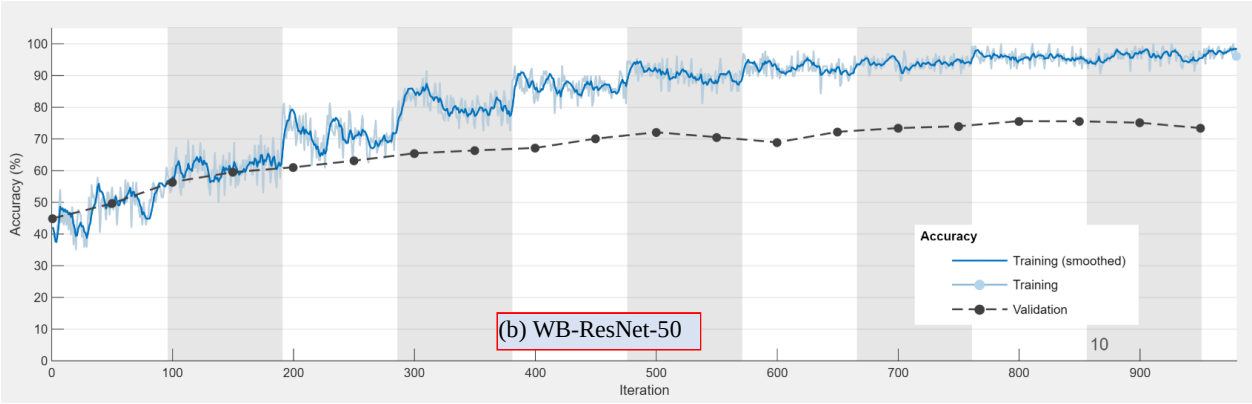
Using the GAF sensor image representation, we converted the gait signals into images. We then investigated some of the top image-classification networks using the method of TL. Details of the results are presented in table (5.3) and figure (5.6) shows that DeePaGait outperformed the ResNet-50, EfficientNet, and Inception-V3 pre-trained image-based models. The table compares the TL model's results to DeePaGait in terms of network depth, the memory size needed for training, and the model complexity represented by the number of trainable parameters. In all of our gait classification experiments, the Inception-V3 model was the best performer, with an accuracy higher than 84% in the WB and ST problems, and an accuracy of 92.5% for the FoG classification.

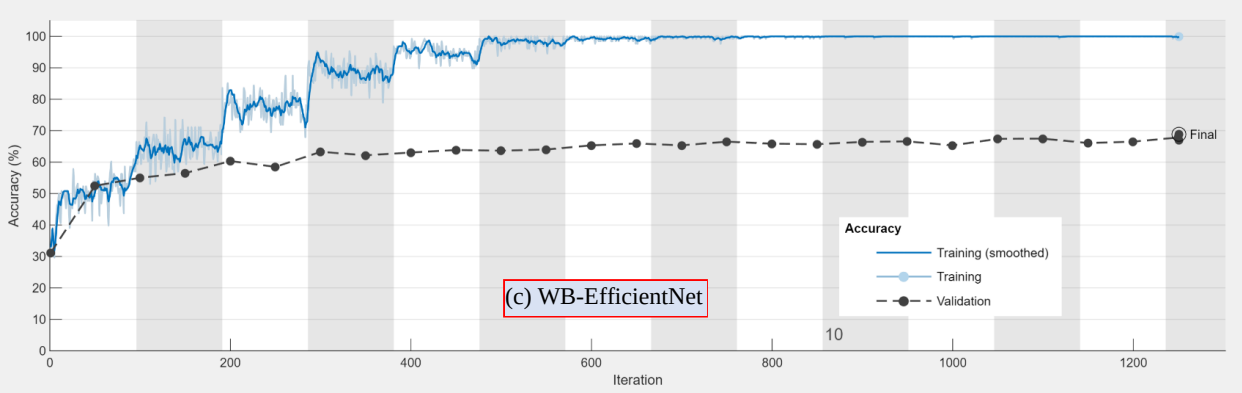
Table 5.3: Comparison of Pretrained TL models and DeePaGait

	Model	***Depth (layers)***	***Parameters (millions)***	***Size (MB)***	***Training Accuracy***	***Testing Accuracy***	***Accuracy difference(pp *)***
Walking Balance	**DeePaGait**	**7**	**0.55**	**0.98**	**0.993**	**0.991**	**0.2**
	ResNet-50	50	25.6	96	0.978	0.746	23.2
	EfficientNet	82	5.3	20	0.994	0.689	30.5
	Inception-V3	48	23.9	89	0.986	0.845	14.1
Shaking /Tremor	**DeePaGait**	**7**	**0.55**	**0.98**	**0.986**	**0.984**	**0.2**
	ResNet-50	50	25.6	96	0.977	0.713	26.4
	EfficientNet	82	5.3	20	0.992	0.651	34.1
	Inception-V3	48	23.9	89	0.993	0.863	13
Freeze of Gait	**DeePaGait**	**7**	**0.55**	**0.98**	**0.987**	**0.982**	**0.5**
	ResNet-50	50	25.6	96	0.993	0.883	11
	EfficientNet	82	5.3	20	0.992	0.848	14.4
	Inception-V3	48	23.9	89	0.993	0.925	6.8

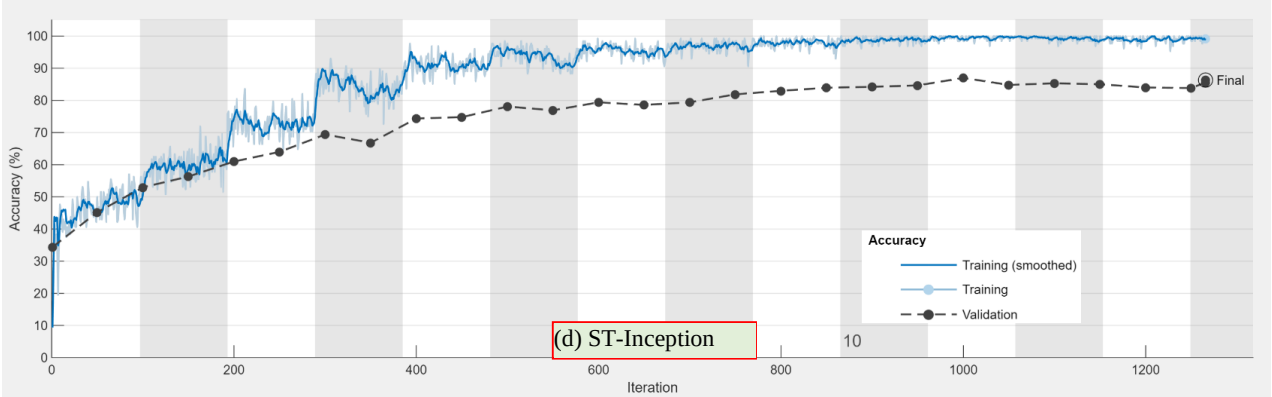
** (percentage points)*

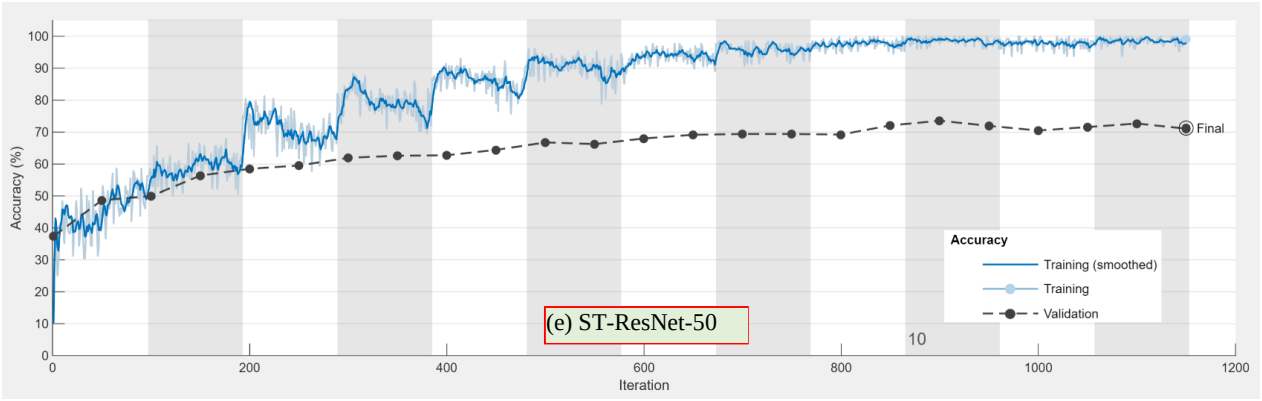


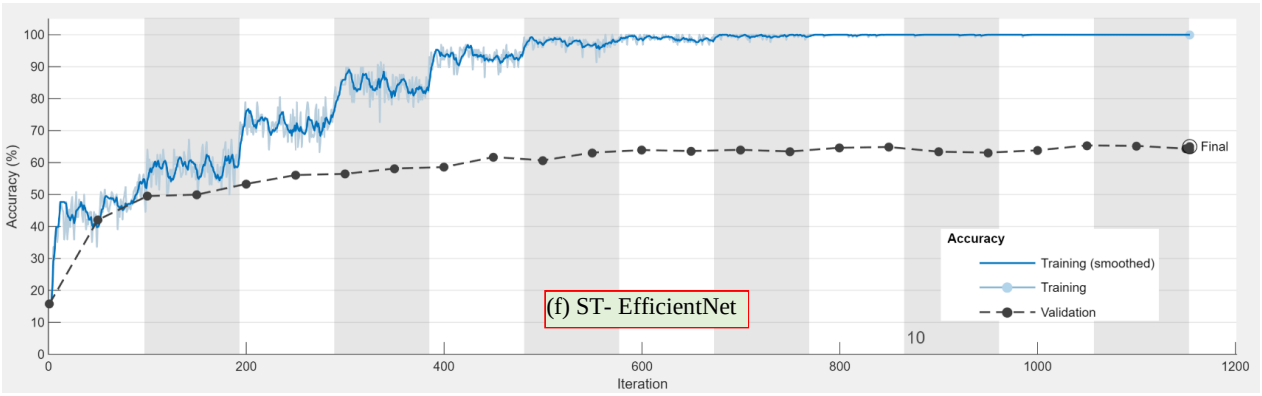




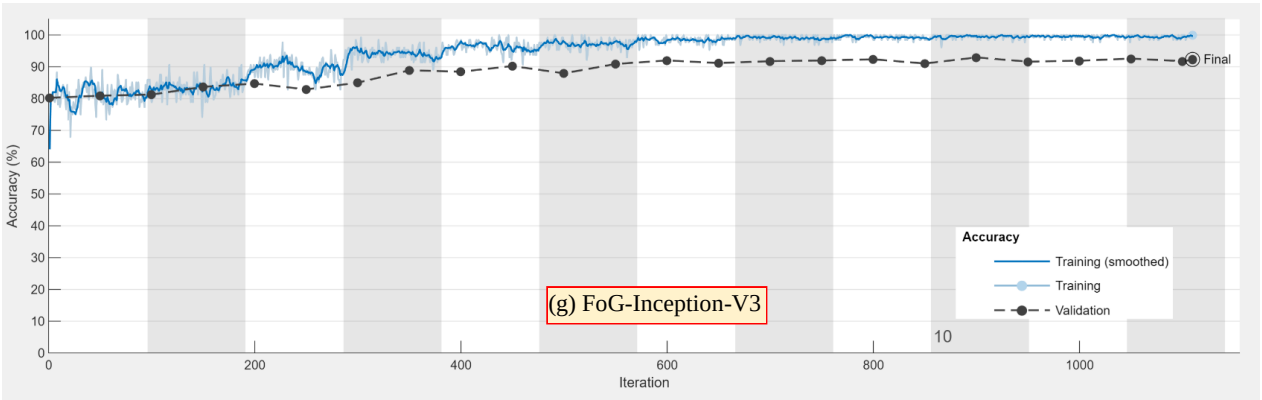
(5.6a) TL Accuracy per epoch for WB

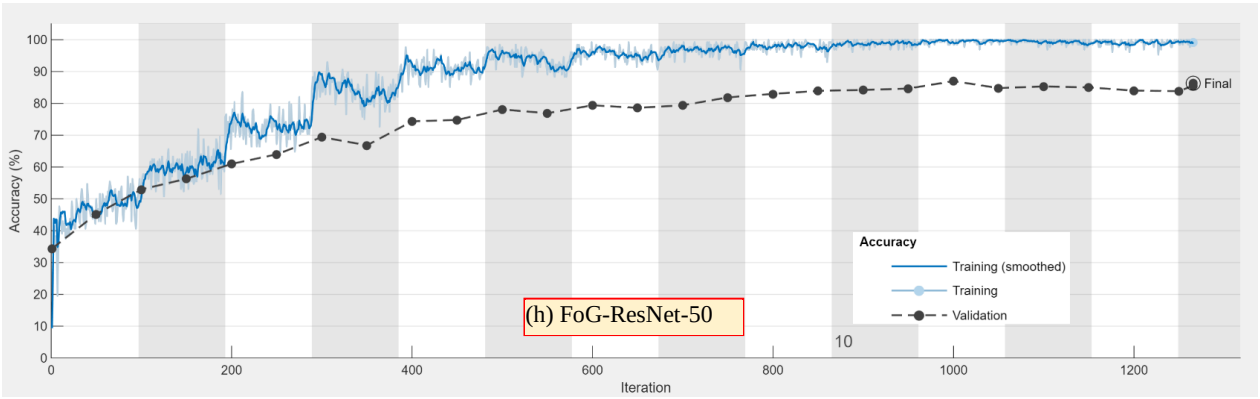


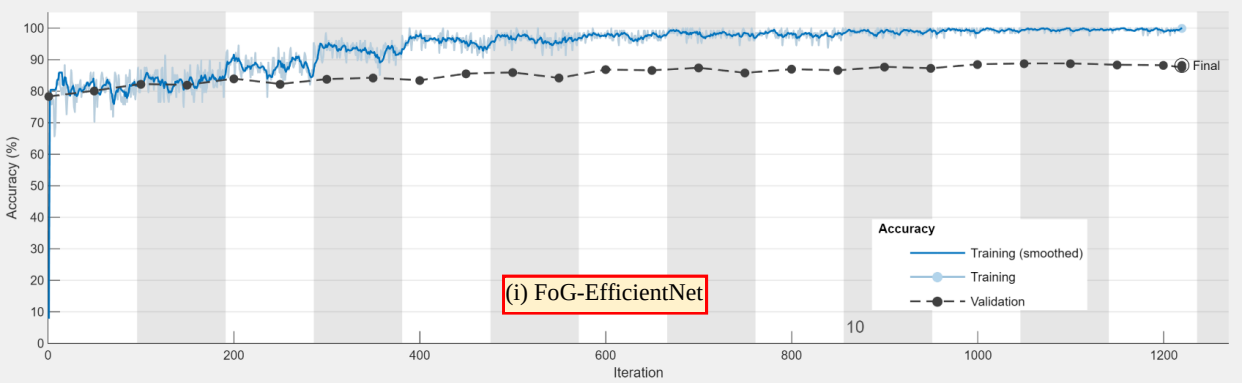




(5.6b) TL Accuracy per epoch for ST







(5.6c) TL Accuracy per epoch for FoG

Figure 5.6 : TL Accuracy per epoch

Table (5.3) compares the model overfitting by presenting both the training and testing accuracy. The last column shows the difference between Training and Testing accuracies in percentage points. We can see that all TL models suffered some level of overfitting, with the severity of overfitting increasing with the increase of the number of trainable parameters. Figure (5.6) shows that overfitting started on all TL models from the 4th epoch, and continued steadily until the end of the training session, for all of the TL models.

5.4.2 LSTM Models Evaluation

The parallel LSTM network has been investigated recently for Human Activity Recognition (HAR) problem and has proven to outperform shallow ML algorithms, Uni-LSTM, and Bi-LSTM networks **[110,111]**. Figure (5.7) shows the parallel LSTM architecture that we used, it consists of 6 parallel LSTM sub-networks, each sub-network operates on one dimension of the 3D Accelerometer and Gyroscope signals. The output of all the sub-networks is combined using a concatenation layer and fed to the FC layer followed by a final softmax layer that generates the overall class prediction.

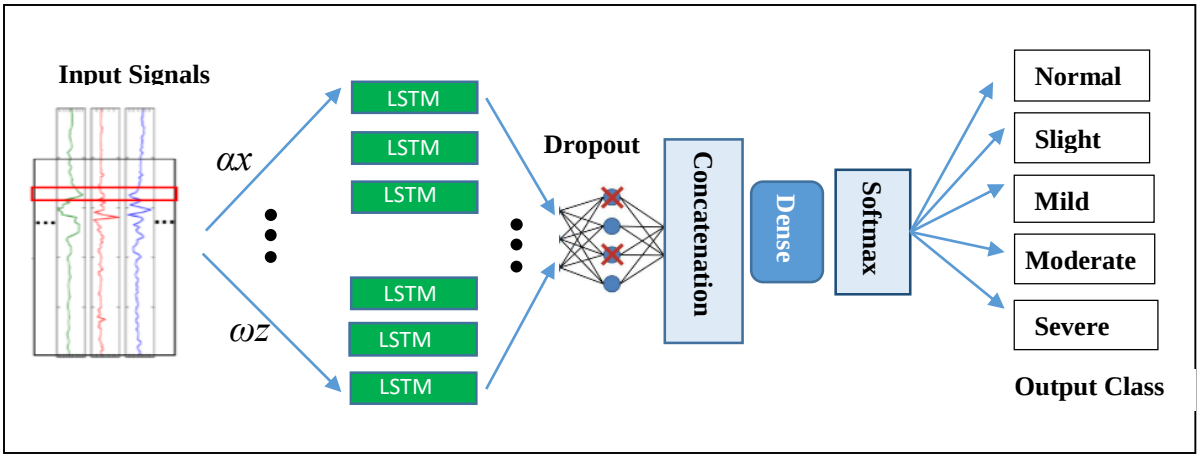


Figure 5.7 : Parallel LSTM Architecture

The deep learning method of the serial CNN-LSTM showed superior performance in the problem of gait authentication and human activity tracking **[112,113]**. Figure (5.8) shows the CNN-LSTM architecture we used to compare to our DeePaGait model. The CNN-LSTM model consists of CNN layers that work as spatial feature extractors. The extracted features are then flattened and fed to the LSTM network that extracts the temporal features just before the final classification is done using Fully-Connected (FC) and a Softmax layer.

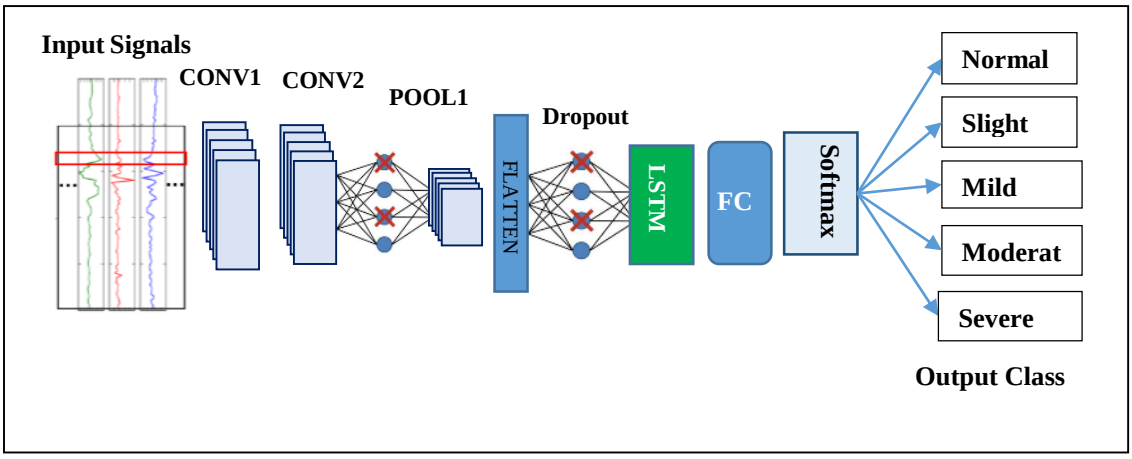


Figure 5.8: CNN-LSTM Architecture used

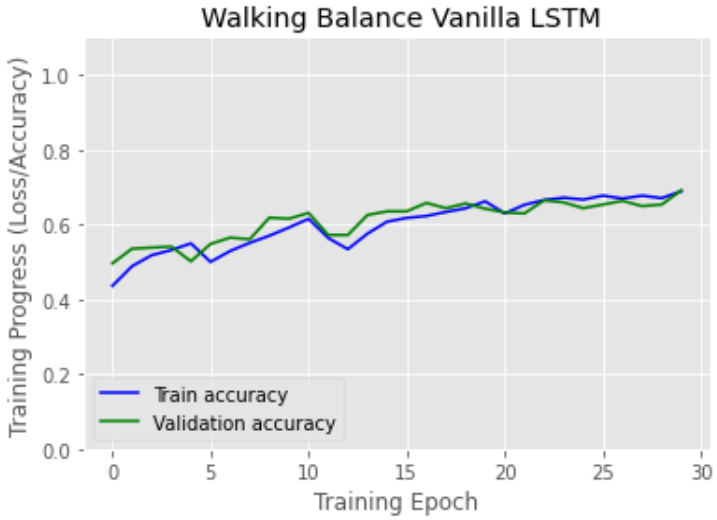
We investigated the performance of the Vanilla LSTM, parallel LSTM, and CNN-LSTM models and compared the performance to our DeePaGait model. Results are shown in Table (5.4). From the table, we can see that the proposed DeePaGait model with the 1D CNN

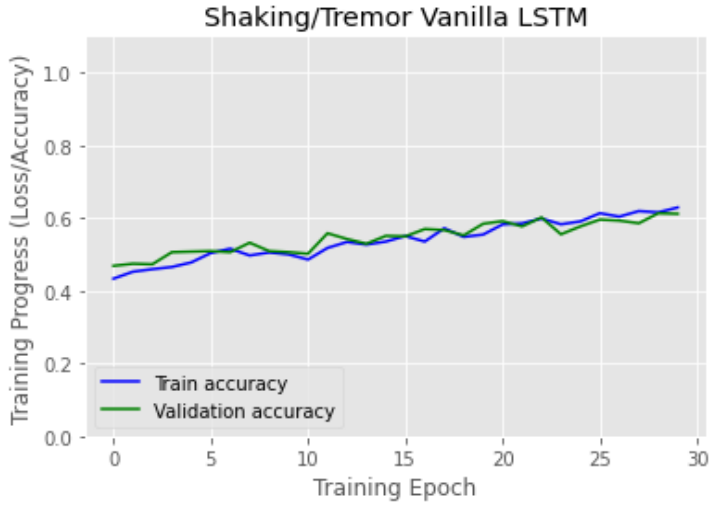
architecture allows the best distinction between PD gait severity levels. The F1-score similarity for different gait impairments indicates the stability and consistency of the DeePaGait algorithm.

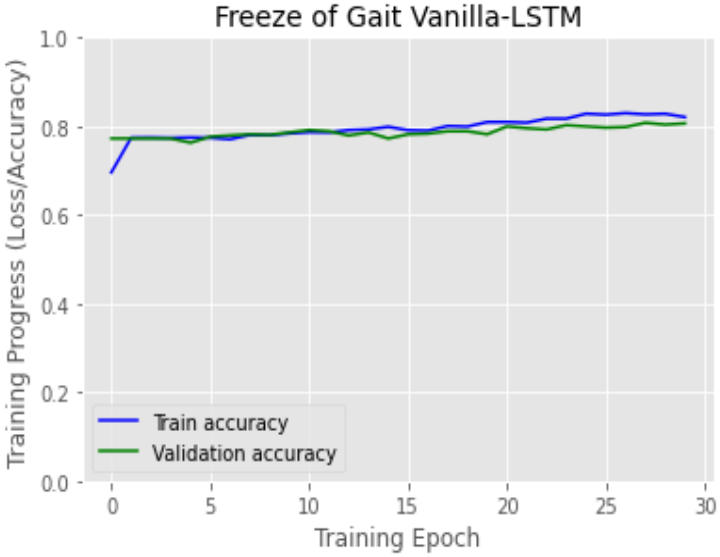
Table 5.4: Comparison of LSTM models to DeePaGait

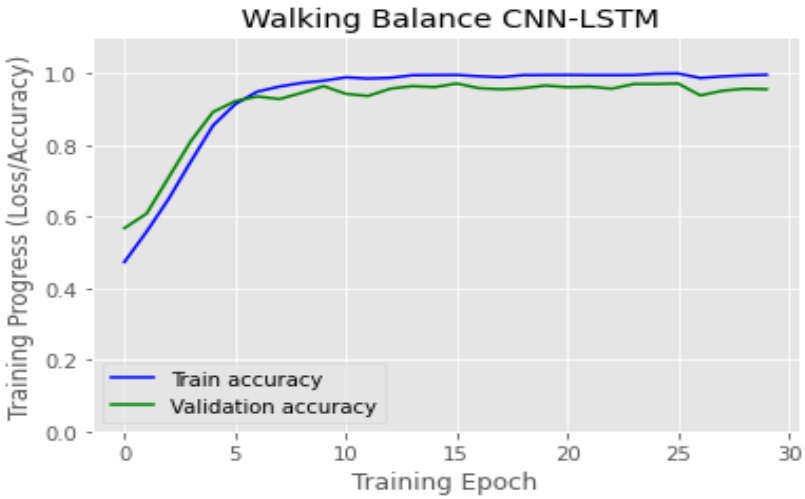
	Gait Impairment	***Precision***	***Recall***	***F1-Score***	***Testing Accuracy***
DeePaGait	***Walking Balance***	0.991	0.992	0.993	**0.991**
	Shaking/Tremor	0.984	0.981	0.983	**0.984**
	Freeze of Gait	0.983	0.981	0.982	**0.982**
LSTM	***Walking Balance***	0.699	0.574	0.628	0.673
	Shaking/Tremor	0.715	0.456	0.553	0.612
	Freeze of Gait	0.827	0.793	0.809	0.819
CNN-LSTM	***Walking Balance***	*0.955*	*0.951*	*0.953*	*0.952*
	Shaking/Tremor	*0.965*	*0.963*	*0.964*	*0.966*
	Freeze of Gait	*0.969*	*0.969*	*0.969*	*0.970*
Parallel LSTM	***Walking Balance***	*0.835*	*0.819*	*0.827*	*0.825*
	Shaking/Tremor	*0.793*	*0.776*	*0.785*	*0.787*
	Freeze of Gait	*0.835*	*0.830*	*0.833*	*0.859*

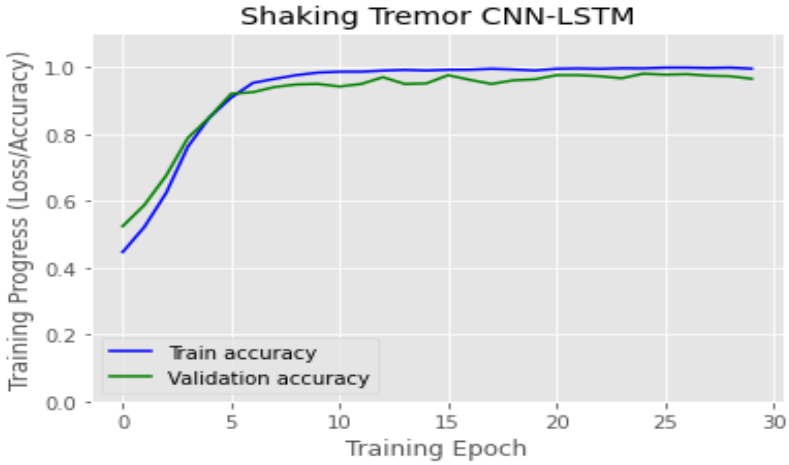
CNN-LSTM model is the best of all other models that we studied as shown in Figure (5.9), we believe CNN-LSTM performed better because it combines the power of CNN spatial features with LSTM temporal features before making the final classification decision. However, the CNN-LSTM model slightly overfits the training data as shown in Table (5.5). By looking at the two other variations of LSTM, we notice that Vanilla LSTM has low performance because the network is missing the required depth, while parallel LSTM learns very slowly and significantly overfits on the training set because of the model over-complexity for the gait analysis problem.

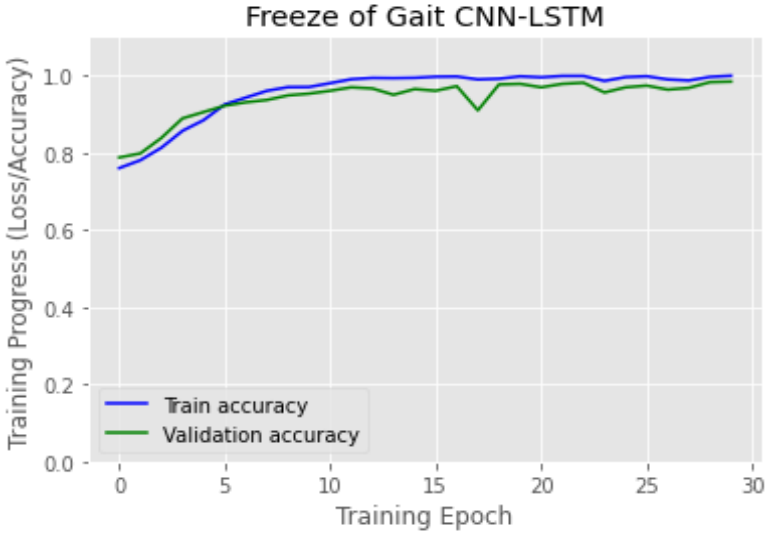

Walking Balance Vanilla LSTM
Training Progress (Loss/Accuracy)
Training Epoch
Train accuracy
Validation accuracy


Shaking/Tremor Vanilla LSTM
Training Progress (Loss/Accuracy)
Training Epoch
Train accuracy
Validation accuracy


Freeze of Gait Vanilla-LSTM
Training Progress (Loss/Accuracy)
Training Epoch
Train accuracy
Validation accuracy


Walking Balance CNN-LSTM
Training Progress (Loss/Accuracy)
Training Epoch
Train accuracy
Validation accuracy
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30


Shaking Tremor CNN-LSTM
Training Progress (Loss/Accuracy)
Training Epoch
Train accuracy
Validation accuracy
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30


Freeze of Gait CNN-LSTM
Training Progress (Loss/Accuracy)
Training Epoch
Train accuracy
Validation accuracy
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
15
20
25
30

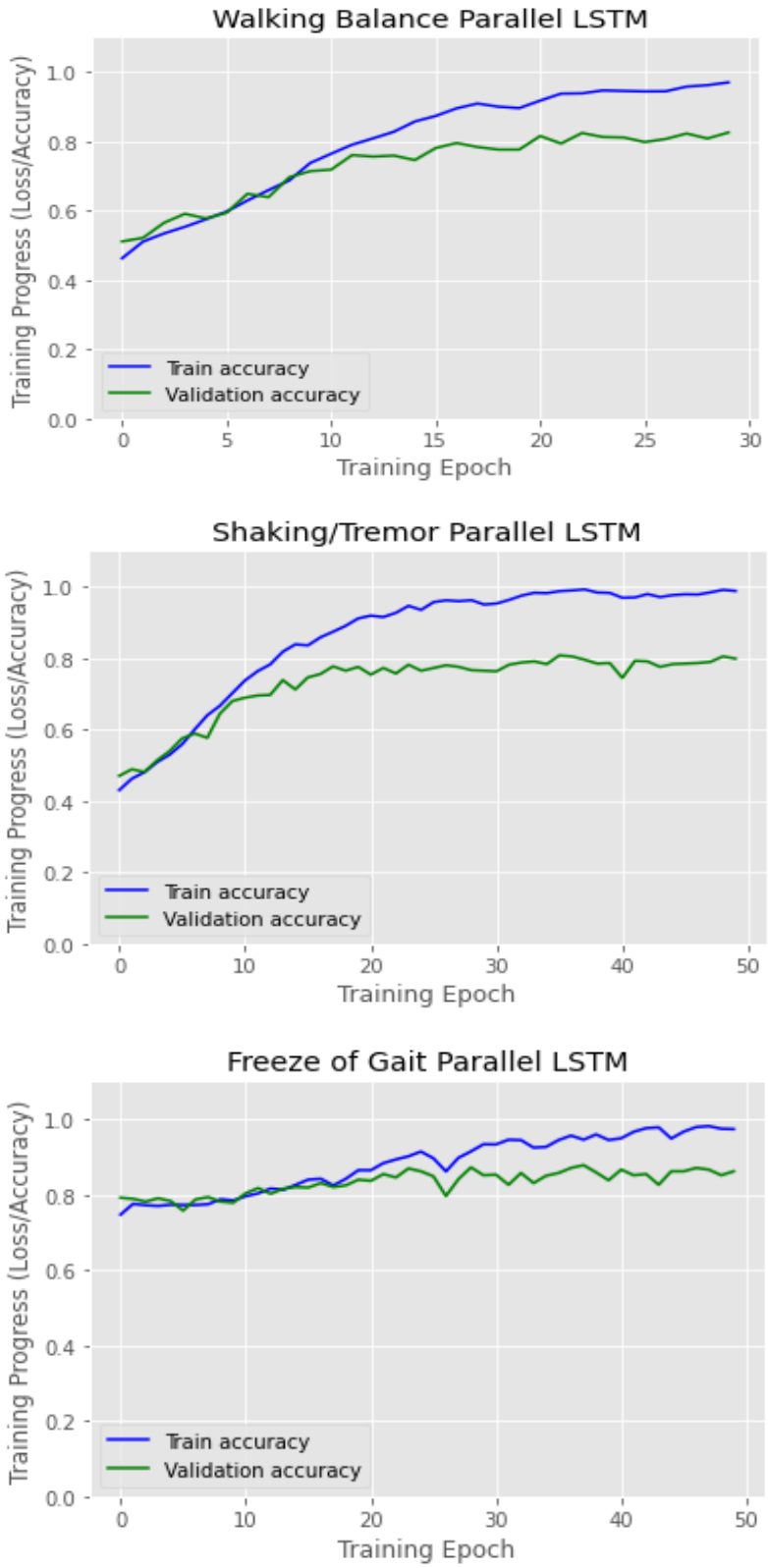


Figure 5.9 : Results for LSTM, Parallel LSTM, and CNN-LSTM models

It is well known that increasing DL model complexity can make the model vulnerable to overfitting, instead of learning discriminative features, the model memorizes the gait impairment pattern of the training set. By comparing to LSTM models in Table (5.5) we can see that DeePaGait can generalize to the testing set, indicating that our DeePaGait was able to learn discriminative features that separate the severities classes of PD gait impairments.

Table 5.5: Comparison of LSTM models and DeePaGait Overfitting

	Model	***Depth (layers)***	***#Param's (millions)***	***size (MB)***	***Training Accuracy***	***Testing Accuracy***	***Difference***
Walking balance	**DeePaGait**	**11**	**0.55**	**0.98**	**0.993**	**0.991**	**0.2**
	LSTM	3	0.08	0.31	0.689	0.673	1.6
	Parallel LSTM	5	38.4	151.4	0.978	0.825	15.3
	CNN-LSTM	9	2.9	11.35	0.991	0.952	3.9
Shaking/ Tremor	**DeePaGait**	**11**	**0.55**	**0.98**	**0.986**	**0.984**	**0.2**
	LSTM	3	0.08	0.31	0.629	0.612	1.7
	Parallel LSTM	5	38.4	151.4	0.991	0.787	20.4
	CNN-LSTM	9	2.9	11.35	0.994	0.966	2.8
Freeze of Gait	**DeePaGait**	**11**	**0.55**	**0.98**	**0.987**	**0.982**	**0.5**
	LSTM	3	0.08	0.31	0.829	0.819	1
	Parallel LSTM	5	38.4	151.4	0.987	0.859	12.8
	CNN-LSTM	9	2.9	11.35	0.988	0.970	1.8

5.5 Chapter Summary

DeePaGait is a data-driven neural network model, that explores the inference of PD gait by analyzing the patients' smartphone walk data. DeePaGait can distinguish the severity of gait aspects based on the smartphone's acceleration and rotation signals. After experimenting DeePaGait on 152 PD patients, we demonstrated that gait severities can be accurately predicted using smartphone sensing of the motor symptoms of PD gait. Our DeePaGait DL network was able to classify the severity of Walking-Balance, Shaking/Tremor, and Freeze of Gait (FoG), with an accuracy of: 99.1%,98.4%, and 98.2% respectively. To the best of our knowledge, the accuracy of our DeePaGait model surpassed the best-published results achieved by prior 1D CNN smartphone models by over 7% [128,129]. Despite the challenges of working with a self-labeled, crowdsourced dataset, we were able to demonstrate that gait classification based on smartphone sensor data is feasible and has potential value as a diagnostic support tool. In future work, we plan to infer the overall UPDRS score based on gait, we would also like to experiment DeePaGait on multiple independent datasets, and multiple age groups. We are also planning to evaluate DeePaGait in a live deployment in the future.

CHAPTER 6
DEEP LEARNING-BASED MEDICATION ADHERENCE (DEE-PA-MED)

In this chapter, we present the prediction of medication adherence based on DL approach and compare it to the tranditional ML approach. Figure (6.1) presents our DL methodology. The first stage is data mining of the mPower gait records and surveys. We processed the mPower data and extracted the relevant survey information and the 3D gait signals. Data were then arranged in a usable format. The two main signals were the smartphone's acceleration (from the accelerometer sensor) and rotation (from the gyroscope sensor). These are given below:

$\alpha(i) = [\alpha x\ (i),\ \alpha y\ (i),\ \alpha z\ (i)]\ T$ (in m/s2)

$\omega(i) = [\omega x\ (i),\ \omega y\ (i),\ \omega z\ (i)]\ T$ (in deg/s)

where i denotes discrete-time, α indicates acceleration, and ω represents rotation. The sensor data corresponding to gait cycles were fed to DeePaMed; a multilayer Conventional Neural Network (CNN), crafted for patches of gait strides. DeePaMed classified 30 seconds of a walk as either PD patient "On" vs. "Off" medication, or if the gait data belongs to an HC. We conducted supervised classification experiments and measured the model precision, accuracy, Recall and F1-score.

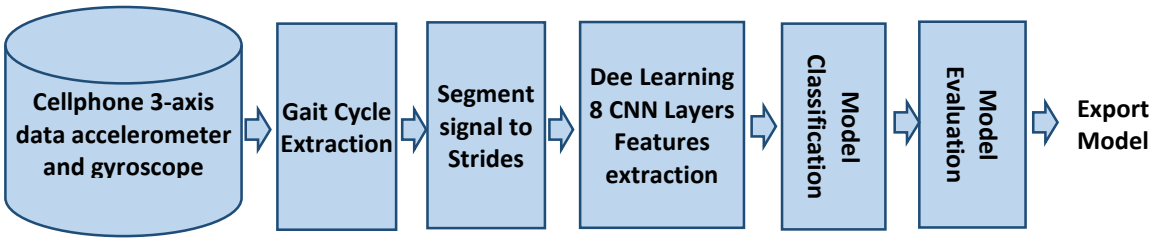


Figure 6.1 Flow Diagram for data collection, feature extraction, and classification using DeePaMed

6.1 DEEPAMED NETWORK ARCHITECTURE

Inspired by prior neural networks-based work by Zou et al **[114]**, we created the CNN architecture shown in Figure (6.2). The network consists of the feature extracting layers followed by fully-connected and classification layers, as shown in Table (6.1). Feature extraction consists of four convolutions and two pooling layers. While Convolution layers generate the features, pooling layers are added to downsize the features map.

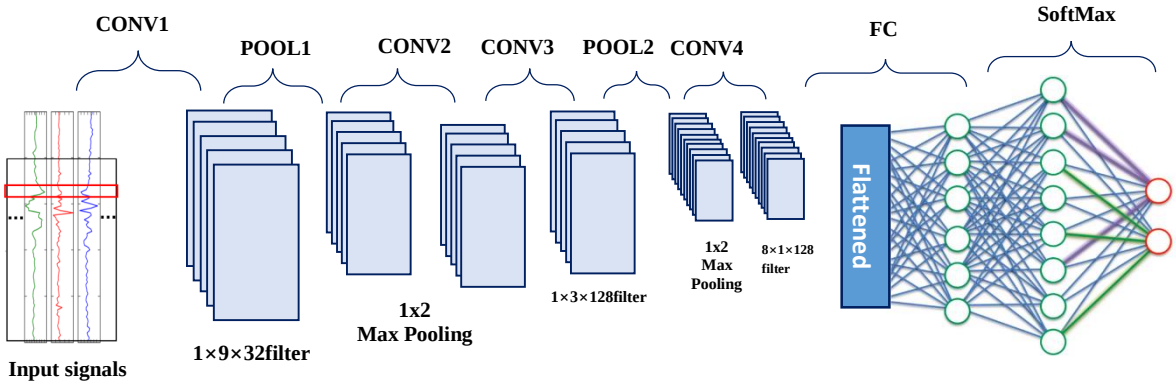


Figure 6.2 : DeePaMed Architecture

The mPower study collected data in an uncontrolled environment. The participant had the freedom of positioning the phone in their pocket, which led to an inconsistent axis orientation from one participant to another. To consider this, axis independent features were extracted individually using one-dimensional filters, enabling the capture of the PD signal variation on that particular axis. The overall signal variation is captured by feeding the accelerometer and gyroscope signal magnitudes to the network. Following this mechanism, we ensured that every spike of PD gait fluctuation was captured and analyzed throughout the 30-sec walk, in chunks of N-Strides patches. Axis interdependent features were extracted at a later stage in the network, specifically at the Conv4 layer, using 8x1 filters. Convolutional and pooling layers are followed by fully connected and classification layers.

Table 6.1: DeePaMed CNN structure

Layer Name	Filter size	Filters	Feature Map
Conv1	1x9	32	8x90x32
Pool1	1x2	N/A	8x45x 32
Conv2	1x3	64	8x45x 64
Conv3	1x3	128	8x45x 128
Pool2	1x2	N/A	8x22x 128
Conv4	8x1	128	1x22x 128

6.2 EVALUATION METRICS

Based on 760 walking records from 456 participants, we conducted a supervised classification experiment by training the ML algorithms and our DeePaMed network on 90% of the data and testing on the remaining 10%. To evaluate modules' performance, the following performance metrics, that are widely used in mobile health applications, have been adopted:

• Accuracy: The percentage of correctly predicted samples. Accuracy formulated is

$$= \frac{\text{Correctly Classified Samples}}{\text{Total Testing Samples}}$$

• Precision: Precision is the fraction of the correctly predicted samples, to the total positive predicted samples.

Precision measures the robustness of the tested module against false positives.

$$= \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

• Recall: Also called the Sensitivity, is the fraction of the correctly predicted positive samples to the total samples in the classification class. Recall gives an idea of the classification misses.

$$= \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

F1-score: is the weighted average of Precision and Recall, it takes false positives and false negatives into account

$$= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

6.3 MODELS EVALUATION

To infer medication adherence, we experimented with solving the classification problem by traditional Machine Learning (ML) algorithms, deep learning DeePaMed architecture, and other DL methods, these methods will be discussed in the following sections.

6.3.1 ML Medication Adherence

Using the same methodology explained in chapter (4), we estimated PD patients' medication adherence, by classifying walks into two categories, Before and After medication. Time and statistical features were calculated directly from the MagNGα and MagNGω. Frequency domain features were calculated from the Fast Fourier Transform (FFT) and Power Spectral Density (PSD). Frequency domain features were subsequently extracted for each walking segment record. Wavelet domain features were extracted from the discrete wavelet transform (DWT) of the MagNGx signal. After the extraction of features, a dataset was prepared and fed to different ML algorithms for medication adherence classification.

Various ML algorithms families were experimented with, including Decision Trees (DT), Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Logistic Regression (LR), Naïve Bayes, and Ensemble Classifiers (EC) including Random Forest (RF) and bagged trees. The results including performance metrics are presented in Table (6.2). Random Forest was the best classifier for distinguishing between medication states, with an accuracy of 83.4%, precision of 83.0%, and F1-Score of 85%.

Table 6.2 : Comparison of ML and DeePaMed algorithms

Medication Inference:	***Precision***	***Accuracy***	***Recall***	***F1-Score***
Random Forest	**83.0%**	**83.4%**	**87.1%**	**85.0%**
Bagged Trees	80.1%	80.2%	78.0%	79.0%
Cubic SVM	74.9%	74.9%	75.1%	75.0%
Weighted KNN	78.5%	78.6%	77.8%	78.1%

Logistic Regressio	55.3%	55.6%	54.9%	55.1%
Fine Tree	74.7%	74.9%	73.8%	74.24%
Naïve Bayes	61.75%	52.9%	60.0%	60.9%

Based on the results seen in Table (6.2), it can be observed that ML could not achieve an acceptable performance when classifying various walks before/after taking medication and at another time (HC), mainly because the handcrafted features were not able to linearly or non-linearly discriminate between the different PD classes.

6.3.2 Image-based Models Evaluation

To evaluate our DeePaMed model, we compared DeePaMed performance to the state-of-the-art pre-trained DL models mentioned above, in terms of model architecture (depth and parameter size), and performance (Training, and Testing Accuracy). We adopted the method of transfer learning (TL) shown in Figure (6.3) and followed the procedure outlined below:

Step1: We converted all the Accelerometer and Gyroscope signals to an image representation using the GAF method.

Step2: Image data were randomly split into 90% training set and 10% as a test set.

Step3: The network architecture was adapted to our medication adherence prediction task by replacing the last CONV, FC, and classification layers to match our 3 classes of data.

Step4: The network was trained.

Step5: The network was evaluated on the test set and the performance metrics reported.

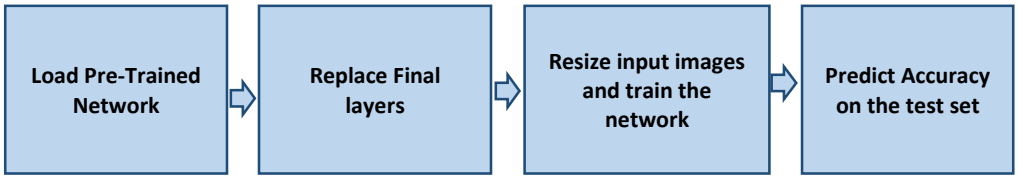


Figure 6.3 : Illustration of the mechanism for TL

Using the method of TL shown in Figure (6.4), we investigated different DL models on the challenging mPower dataset. Detailed results are presented in Table (6.3). The table shows the complexity of the model in terms of parameter size, the memory needed, and the depth of the network. Table (6.3) also shows the training and test accuracy in comparison to our DeePaMed model. The TL models with the best performance were Inseption-V3, and ResNet-

50 which agree with the results of PDMove **[75]**. Inception-V3 and ResNet-50 achieved a testing accuracy of 87% and 81% respectively. However, by examining the training and test accuracies, we can observe that overfitting occurs in all of the models. All the models performed well on the training set, as shown in Figure (6.5). But once validated on the unseen testing set TL models misclassify many of the walk images to the wrong class. This observation will be further discussed in the discussion section.

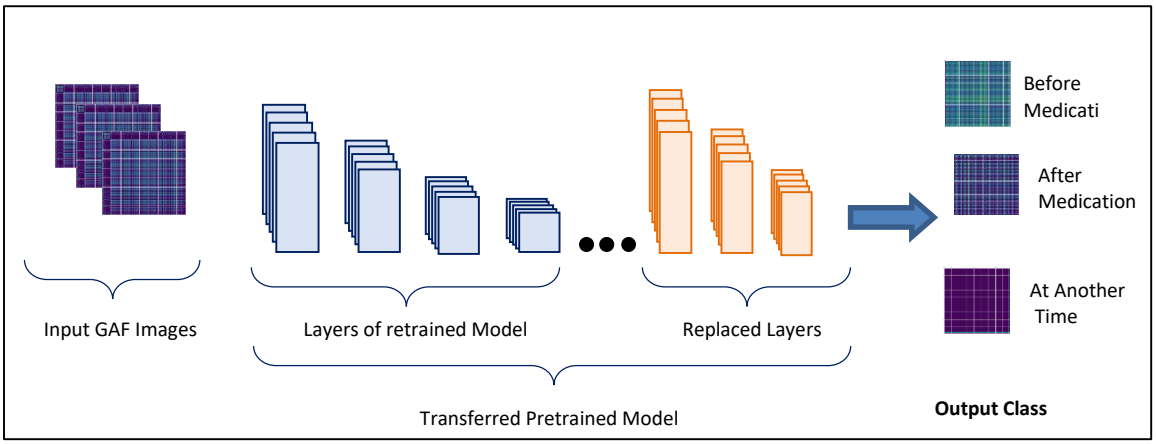


Figure 6.4: TL using the Pre-trained model

Table 6.3: Comparison of Pretrained TL models and DeePaMed

Model	***Depth (layers)***	***Parameters (millions)***	***size (MB)***	***Training Accuracy***	***Testing Accuracy***
DeePaMed	**7**	**0.26**	**0.98**	**99%**	**98%**
ResNet-50	50	25.6	96	99%	81%
EfficientNet	82	5.3	20	99%	73%
Inception-V3	48	23.9	89	99%	86%
ShuffleNet	50	1.4	5.4	99%	70%

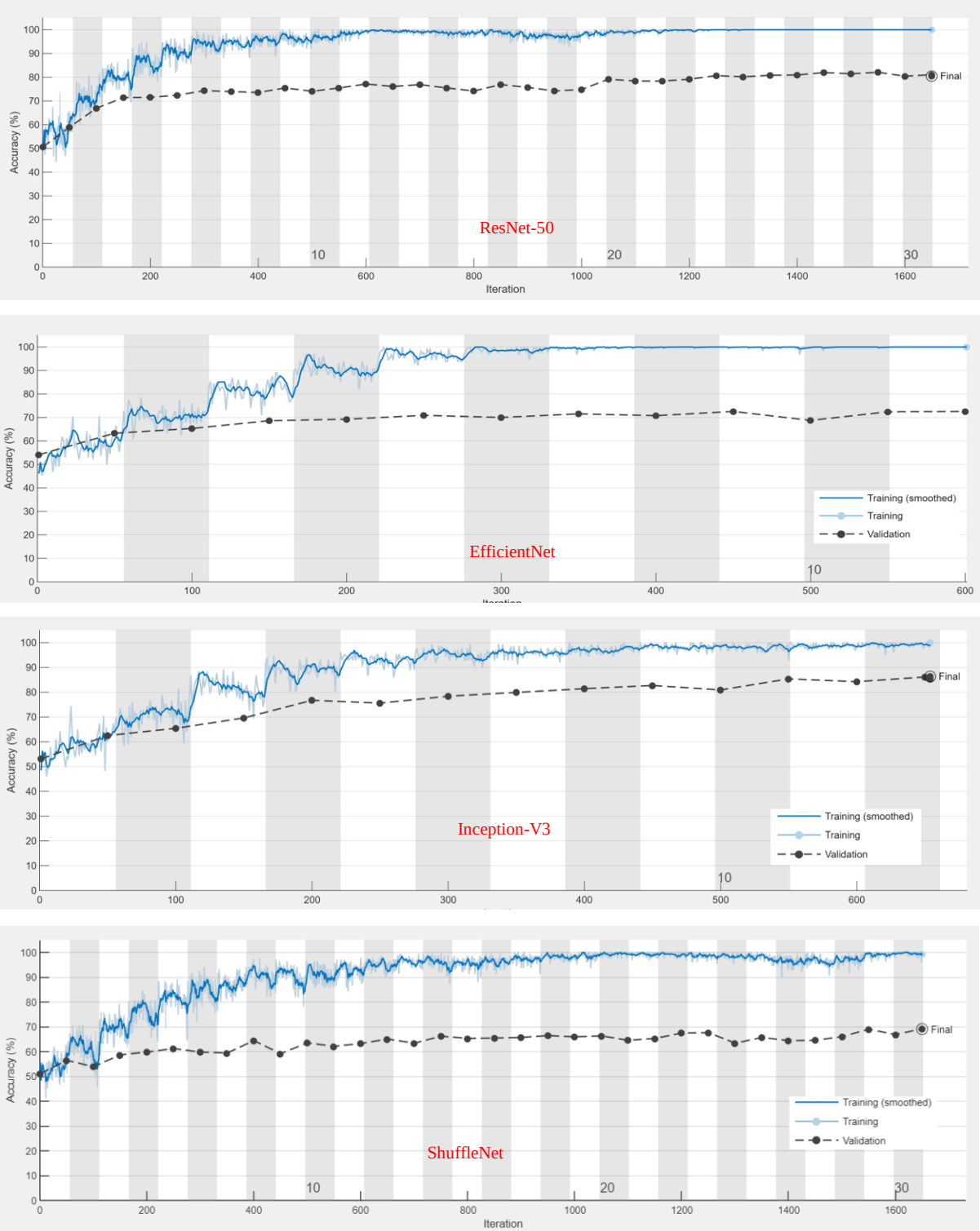


Figure 6.5 : Results of Pretrained TL models

6.3.3 LSTM Models Evaluation

Deep learning methods such as LSTM and variations of CNN-LSTM have achieved promising results in the problem of gait authentication **[115]** and human activity tracking **[116]**. We investigated the performance of the Vanilla LSTM, parallel LSTM, and CNN-LSTM models and compared the performance to our DeePaMed model. Results are shown in Table (6.4).

As shown in Figure (6.6), the CNN-LSTM model uses CNN layers as automatic feature extractors and the LSTM for time series prediction. CNN-LSTM has previously been used for activity recognition and image/video description **[117]**, as well as gait analysis **[113].** The model extract features using CNN Convolution layers. The extracted features are then flattened and provided as input to the LSTM network to extract features before the final classification is done using the softmax layer. The results of CNN-LSTM are the best of all the models that we studied as shown in Figure (6.8), we believe this happens because it combines both the differentiation power of 1D Conv layers and the time series prediction of the LSTM network. However, the CNN-LSTM network also overfits the training data, while the accuracy on the testing set did not exceed 93%.

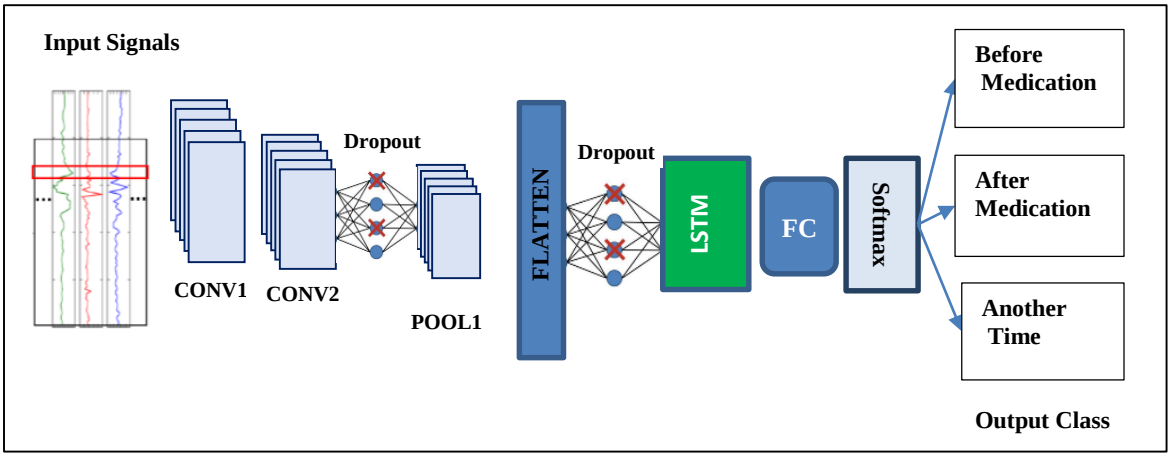


Figure 6.6 : CNN-LSTM Architecture used

A parallel LSTM architecture has been explored for Human Activity Recognition (HAR) and led to better performance than shallow ML algorithms with a performance comparable to CNN models **[110].** The parallel LSTM model we utilized is presented in Figure (6.7). The architecture consists of 6 parallel LSTM nodes, where each part of the 3D Accelerometer and

Gyroscope signals is fed to one LSTM layer. The output of all LSTM layers is combined using a concatenation layer, followed by an FC layer and softmax for the overall prediction calculation. One issue with this network is that it learns slowly, and it is performance is lower than DeePaMed, as shown in Figure (6.8)

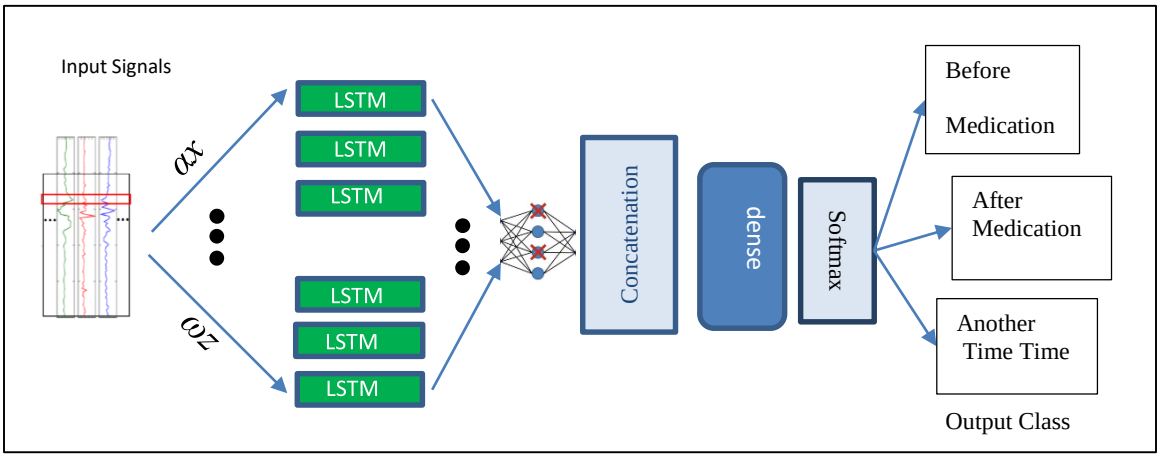


Figure 6.7 : Parallel LSTM Architecture used

Table 6.4: Comparison of LSTM models and DeepaMed

Model	***Depth (layers)***	***Parameters (millions)***	***size (MB)***	***Training Accuracy***	***Testing Accuracy***
DeePaMed	**7**	**0.26**	**0.98**	**99%**	**98%**
LSTM	3	0.08	0.31	61%	56%
Parallel LSTM	5	38.4	151.4	99%	79%
CNN-LSTM	9	2.9	11.35	99%	93%

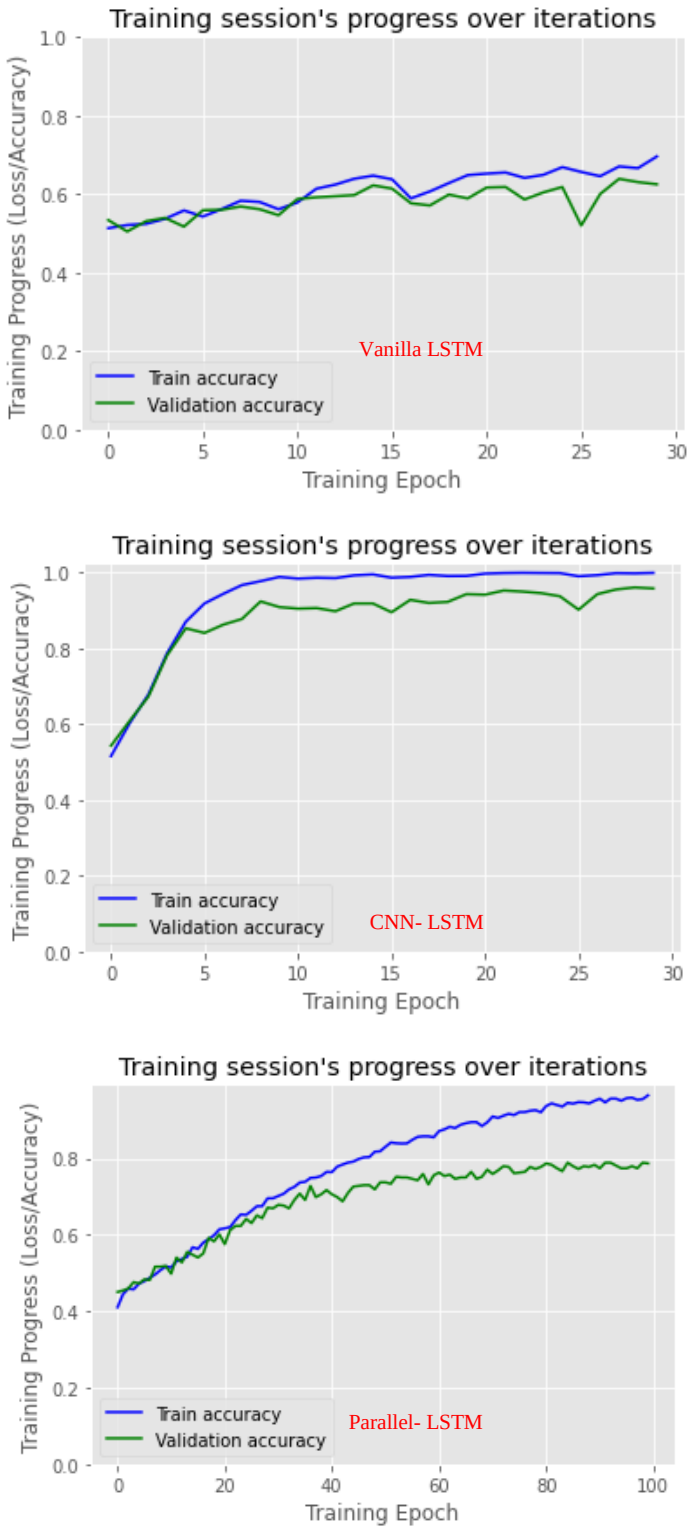


Figure 6.8: Results for LSTM, Parallel LSTM, and CNN-LSTM models.

6.3.4 DeePaMed Evaluation

To overcome the inadequate performance with ML algorithms and due to the demonstrated performance of neural networks on various gait analysis and classification problems, we explored a deep learning approach (DeePaMed) to extract various abstractions of gait features over multiple convolutions and pooling layers.

The classification accuracy with 2 strides segments was 86.6% as presented in Table (6.5). This result exceeds the performance of traditional ML algorithms by a reasonable margin, However, PD step-to-step signal variations due to the effects of tremor, and shuffling made it hard for DeePaMed to distinguish those variations from the signal calm that happens after taking medication, this led us to experiment different segment sizes.

It can be seen from Table (6.5) that we got the best results with 10-Strides segments. We computed the ROC curve for this best-performing model. ROC curve is a good measure of how the model distinguishes between classes, by plotting FPR on the x-axis against TPR on the y-axis. As expected FPR is low for high TPR, particularly for this model, which leads to an area under the ROC curve of 0.97/1, shown in Figure (6.9).

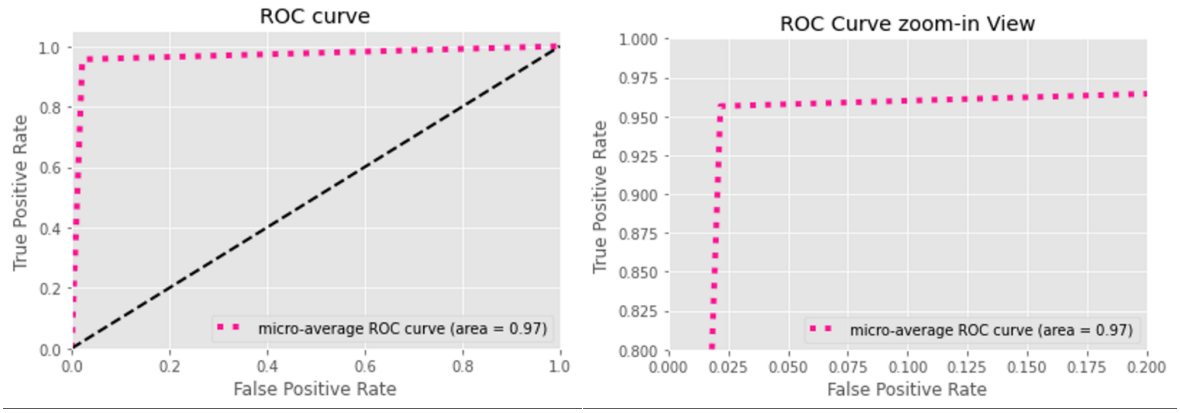


Figure 6.9 : ROC curve for the best performing 10-Strides model

6.4 Improving DeePaGait Performance

6.4.1 Improving Performance by Inputting Multiple Strides

We experimented with varying the input data sizes, consisting of a different number of strides. The results as can be seen in Table (6.5) shows a performance increase as we fed input containing more strides to the network. With more strides fed into the network, DeePaMed could learn stride-to-stride variations better, an attribute that distinguishes PD gait from regular HC walk. Since taking the medication calms the patient and reduced PD gait anomalies, stride-to-stride variations are also reduced, improving our results for discriminating the walk before and after taking medication.

Table 6.5: DeepaMed performance with various numbers of strides

Stride overlaps	***Strides***	***Accuracy***	***Precision***	***Recall***	***F1 Score***
1	2	86.6%	86.3%	85.6%	85.6%
2	3	88.6%	86.6%	88.3%	87.6%
3	4	91.9%	90.4%	91.0%	93.34%
4	5	93.7%	92.4%	93.3%	92.7%
5	6	93.5%	92.0%	92.0%	92.0%
6	7	96.5%	95.6%	95.6%	96.0%
7	8	97.48	96.7%	97.0%	97.0%
8	9	97.1%	96.3%	96.3%	96.3%
9	10	**98.2%**	**97.7%**	**97.7%**	**98.0%**

6.4.2 Improve DeePaMed Performance by Model Tuning

We explored improving performance further by tuning DeePaMed's parameters. Dropout is a technique to reduce overfitting in Neural Networks. Dropout works by randomly dropping out some neurons' output during the training phase, at a rate specified by the dropout rate. The goal is to generalize the model and prevent complex coadaptation. When studying the impact of stride length, we fixed the dropout rate to 50% But we noticed that training accuracy surpassed the testing accuracy regardless of the number of strides used. Here the number of strides were fixed at 10 and varied the dropout rate. Results are presented in Table (6.6) below. We notice

that the difference between Training and Testing accuracy reached the lowest level when using 80% as the training dropout rate. The dropout rate was always kept at 0% (or no dropout) during the testing phase.

Table 6.6: DeePaMed performance with a variable dropout rate

	Dropout keep_probe	***Dropout rate = 1-keep_prob***	***Train Accuracy***	***Test Accuracy***	***Notes***
1	1	0	100%	88.1%	10 strides and 50 epochs
2	0.9	0.1	100%	90.9%	
3	0.8	0.2	100%	95.4%	
4	0.7	0.3	100%	97.7%	
5	0.6	0.4	100%	97.3%	
6	0.5	0.5	100%	98.2%	
7	0.4	0.6	100%	97.9%	
8	0.3	0.7	100%	98.9%	
9	**0.2**	**0.8**	**99.6%**	**99.2%**	
10	0.1	0.9	98.4%	99.1%	

As the dropout rate was increased to 0.8, the testing accuracy got very close to the training accuracy, and overfitting decreased dramatically. Figure (6.10) shows the performance improvement while training the network, the network loss and accuracy were measured for both the training and testing sets, as Figure (6.10) shows the accuracy surpassed 90% after 20 epochs. Also, very small differences can be noticed between the training and testing plots demonstrating that our models did not suffer from overfitting, compared to TL models.

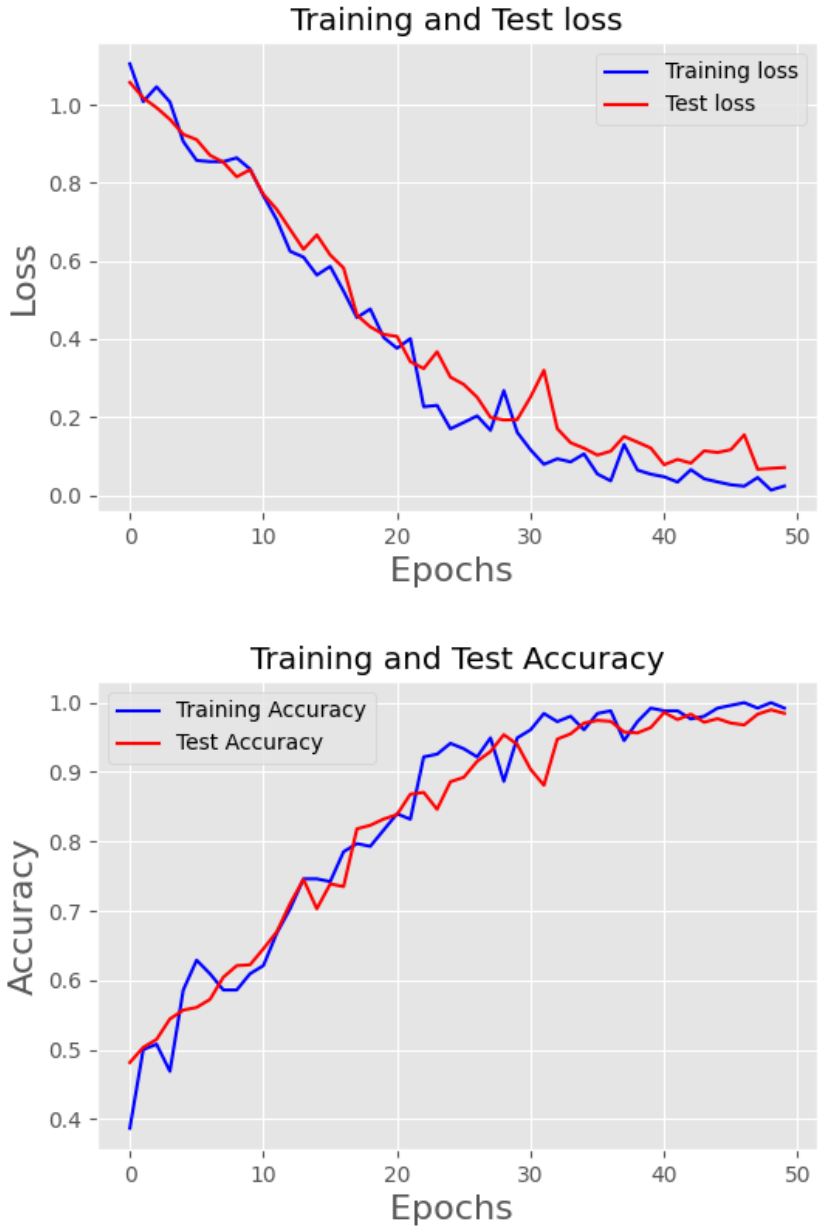


Figure 6.10 : 10-Strides model Performance

6.5 Chapter Summary

PD Patients' lack of adherence to prescribed medication is a major challenge that PD physicians face. The existing methods of remotely measuring medication adherence are inconvenient, not continuous or passive. We introduced DeePaMed, a deep learning smartphone approach that

can differentiate PD walk before and after medication for PD patients. DeePaMed can also distinguish a non-PD walk by HC's based on the smartphone's acceleration and rotation signals. After experimenting DeePaMed on 452 participants, we prove that medication non-adherence can be accurately predicted using smartphone sensing of the motor symptoms of PD gait. Our DeePaMed model was able to discriminate PD patients on- vs off-medication and baseline HC walk with an accuracy of **98.2%**. The accuracy of our CNN model surpassed that of traditional Machine Learning methods by at least **17%**. We also found that the CNN model performed best with inputs containing a minimum of 10 full gait strides and a 0.8 dropout rate. Our findings suggest that remote monitoring of medication adherence using the smartphone is feasible.

CHAPTER 7
CONCLUSION

7.1 DISCUSSION

The majority of prior PD gait studies have been conducted in a clinic/lab-closed environment, under a proctor's supervision. This resulted in datasets with a limited number of participants. The participants' errors were corrected by proctors hid any possible human errors, and hence the performance was unrealistic. The mPower dataset we adopted here is collected by participants in their home environment. While this led to a very noisy dataset, as people act freely and randomly in real life, participants' movements were realistic. Our signal processing and gait extraction technique successfully identified and extracted only valid gait records. Our Signal Processing included smoothing the gait signals, subtracting signal mean, and identifying gait cycles. These steps filtered out all the non-gait data and fed only a valid combination of gait strides to our model, which helped significantly in improving the performance of the overall model.

The performance achieved by our PD models are possibly higher than would be achieved in the real world deployments for the following reasons:

- **Our PD training set was selected to be balanced:** However, in the real world an imbalanced dataset is likely because PD occurs in a relatively small percentage of the population. Specifically, the mPower study had more participants with Slight and Mild severities than Moderate and Severe. Ultimately, imbalance is likely to reduce the performance of our gait analyses models, which were trained on balanced dataset.
- **Our training dataset was split using random splitting, which may have caused data leakage between the training set to the testing set**, which in turn may have led to testing the classifier with data records that are very similar to the records seen during the training process. A more realistic result would have been achieved using subject-level splitting wherein all of each subject's data appears in either the training or test set but not partially in either one. Subject-level data splitting was not possible due to the limited number of subjects who performed all the required walks. Exploring subject-level splitting is an interesting potential future research direction.

- **Limited diversity in our participant sample**; The mPower dataset had 74.5% White/Caucasian participants while all other races combined did not exceed 20%. Moreover, 6% of subjects opted not to disclose their race. Our model might have performed well on our current predominantly white/Caucasian dataset but has not been evaluated on a diverse dataset with respect to race/age-group, which may challenge and possibly reduce the performance of our models.
- **The nominal values generated by on-device sensors of smartphones made by different manufacturers can vary by up to 30%, which could affect the performance of our PD gait analyses models significantly;** A study [135] Investigated accelerometer and gyropscope sensors on 36 devices (including smartphones and smartwatches), that sensor values of smartphones from different manufacturers differed by up to 30%. The paper also found that training models targeted at similar groups of devices outperformed models that generated one model for all phone models. The mPower data was collected using an iphone (v5 and v6 (Apple Inc., Cupertino CA, USA)) smartphone. If our model to be deployed on other devices-models, sensor bias introduced by the differences in their sampling rates could reduce the performance of our gait analyses models.

Some commercial mobile apps compromise security by accessing smartphone sensors. These apps can record the microphone, monitor location, and take photos, all without the user's permission. To address these privacy concerns, the mPower app gives full control of the sensors' recording to the user and limits the gait data collection to 30 seconds before and 30 seconds after taking medication. Furthermore, our signal processing includes smoothing the signal, subtracting the signal mean, and calculating of signal magnitude, all of which will help hide the personal signal variation. Also, our experiments are conducted at the population level using de-identified patient data. The goal is to identify and characterize signal changes that reflect the relatively calm gait that occurs after the medication is taken. In comparison to the segment-level classification that we explore, subject-level classification will require more data per subject. Consequently, an algorithm with more depth will be needed to classify the subject unique walking signal.

As PD progresses to a severe state, patients tend to miss medication doses for various reasons, including depression and memory loss. The level of non-adherence increases as the daily dosage increase [5]. Patients also start developing resistance to taking medication when they notice that medications are not effective. A study found that 20 to 40 out of 100 people with Parkinson's noticed that the drugs are becoming less effective after five years of treatment [132]. By comparing walks before and after medication. DeepaMed can continuously detect and alert both patient and caregiver of nonadherence and medication ineffectiveness, resulting in a reduction of clinic visits and providing more data points for the treatment process.

An overfitted model is a statistical model that contains more parameters than can be justified by the data [133]. When such a model predicts a trend in very noisy data, like in our case the mPower dataset, it tends to overfit the training set, and perform less well on the test set. DeePaMed and DeePaGait models overpass the performance of all the pre-trained TL models and LSTM models, that is because DeePaMed has the right complexity and depth to predict the medication adherence trends in the mPower data.

Due to the complexity of PD, no one treatment that fits all patients. Instead of standardizing the treatment, and to embrace the concept of "Personalized Medicine (PM)", treatment needs to be prescribed based on the susceptibility of specific subtypes of PD to side effects with consideration of lifestyle, genetic framework, personality, and pharmacogenetics. [134]. In PM, Physicians adjust the treatment plan based on occasional hospital visits and input from patients. Alternately, DeepaMed can provide continuous inputs from the patients, so that doctors can monitor and adjust the individual treatment to fit each patient's unique case.

7.2 CONCLUSION

Nearly 10 million people worldwide are living with Parkinson's disease (PD) [118]. The progression of the disease can be inferred from changes in the patients' gait to inform early intervention. Due to the high cost of hospital visits and in-patient days, remote measurement of gait has become an important tool for monitoring the progression of PD. Although measurements reduce hospital visits and offer convenience to both PD patients and the

healthcare provider, the validity of these measurements compared with assessments in the clinic continues to be a challenge.

Another major challenge that PD physicians face is the PD Patients' lack of adherence to prescribed medications. The existing methods of remotely measuring medication adherence are inconvenient, not continuous or passive.

Due to their near-ubiquitous ownership, smartphone sensing is one of the very effective, highly available ways to classify gait. Equipped with triaxial Accelerometers and gyroscopes in addition to powerful CPUs smartphones provided a potential alternative for remote gait assessment in the home environment. This smartphone technology used to analyze PD gait in several studies

In our work, we addressed the unique gait characteristics of PD and inferred the stage of each PD gait modality through deep learning classification of smartphone sensor data collected by a mobile health application. Specifically, we introduced a data-driven neural network model, that explores the inference of PD gait by analyzing the patients' smartphone walk data. Our model can distinguish the severity of gait aspects based on the smartphone's acceleration and rotation signals. Our model can also differentiate PD walk before and after medication for PD patients. DeePaMed can also distinguish a non-PD walk by HC's based on the smartphone's acceleration and rotation signals.

After experimenting on 452 participants, we prove that medication non-adherence can be accurately predicted using smartphone sensing of the motor symptoms of PD gait. Our model was able to discriminate PD patients on- vs off-medication and baseline HC walk with an accuracy of 98.2%. The accuracy of our CNN model surpassed that of traditional Machine Learning methods by at least 17%. We also found that the CNN model performed best with inputs containing a minimum of 10 full gait strides and a 0.8 dropout rate. Our findings suggest that remote monitoring of medication adherence using smartphone is feasible

We Also prove that gait severities can be accurately predicted using smartphone sensing of the motor symptoms of PD gait. Our DL network was able to classify the severity of Walking-Balance, Shaking/Tremor, and Freeze of Gait (FoG), with an accuracy of: 99.1%,98.4%, and 98.2% respectively. To the best of our knowledge, the accuracy of our DeePaGait model surpassed the best-published results achieved by prior 1D CNN smartphone models by over 7%. Despite the challenges of working with a self-labeled, crowdsourced dataset, we were able to demonstrate that gait classification based on smartphone sensor data is feasible and has potential value as a diagnostic support tool.

Although the classification results were affected by the subjective nature of PD labels assigned by patients based on their responses to the MDS-UPDRS questions, we were able to demonstrate with a relatively large number of participants that remote and automatic PD patient classification based on sensor activity data can supply objective assessments of PD-related gait patterns and severity of gait anomalies, which ultimately has the potential to improve remote healthcare for PD patients.

7.3 FUTURE WORK

Planned future work includes the exploration of other signal segmentation strategies and various segmentation window lengths, and techniques such as Bayesian segmentation. Future work also includes the evaluation of DeePaGait model on PD gait datasets that are diverse in terms of participant ages and race, and also a similar number of walks for each severity level. We would also investigate evaluating our models on datasets collected at lab/clinc environment and work with doctors to collect additional data to validate our models using both home-collected and clinic-collected data.

Another area we would like to work on is evaluating our models on data collected using different devices and smartphone-models, to ensure the models robustness to sensors bias and changes in the sensors sampling frequencies.

Model improvement is to be investigated by performing classification using subject level data splitting, in order to evaluate the model using data from unseen subjects during the training stage. Also, there are other associated comorbidities that may affect gait, including dementia and Hip and knee osteoarthritis, and age differences, which may confound gait signals. In future, we will collect data on and investigate the effects of these associated comorbidities and age-groups and explore methods to factor them in.

The exploration of combining walking with other activities such as voice and tapping and memory to predict the PD stage is to be investigated. Since features from those activities can improve the prediction accuracy and facilitate remote follow-up with PD patients. Based on all the activities, the inference of the overall UPDRS score is planned to be investigated. Finally, in future, both DeePaMed/DeePaGait models will be evaluated in a live development.

BIBLIOGRAPHY

[1] Muangpaisan, Weerasak, et al. "A systematic review of the worldwide prevalence and incidence of Parkinson's disease." Journal of the Medical Association of Thailand 94.6 (2011): 749.

[2] Parkinsons' Foundation, "Postural Instability" https://www.parkinson.org-/Understanding-Parkinsons/Symptoms/Movement-Symptoms/Postural-Instability, Accessed on June 2021

[3] Bohingamu Mudiyanselage, Shalika, et al. "Cost of living with Parkinson's disease over 12 months in Australia: A prospective cohort study." Parkinson's Disease 2017 (2017).

[4] Kulkarni, Amit S., et al. "Medication adherence and associated outcomes in medicare health maintenance organization-enrolled older adults with Parkinson's disease." Movement disorders: official journal of the Movement Disorder Society 23.3 (2008): 359-365.

[5] Saini, Sameer D., et al. "Effect of medication dosing frequency on adherence in chronic diseases." Methods (2009).

[6] Davis KL, Edin HM, Allen JK. Prevalence and cost of medication nonadherence in Parkinson's disease: evidence from administrative claims data. Mov Disord. 2010 Mar 15;25(4):474-80. doi: 10.1002/mds.22999. PMID: 20131374.

[7] Goetz, Christopher G., et al. "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results." Movement disorders: official journal of the Movement Disorder Society 23.15 (2008): 2129-2170.

[8] Pew Research Center, "Mobile Fact Sheet" https://www.pewresearch.org/internet/fact-sheet/mobile/ , Accessed on April 2021

[9] Tsanas A et al 2010 Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests IEEE Trans. Biomed. Eng. 57 884–93.

[10] Bayestehtashk A et al 2015 Fully automated assessment of the severity of Parkinson's disease from speech Comput. Speech Lang. 29 172–85

[11] Patel S et al 2009 Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors IEEE Trans. Inf. Technol.Biomed. 13 864–73

[12] Horak F B and Mancini M 2013 Objective biomarkers of balance and gait for Parkinson's disease using body-worn sensors Mov. Disorders 28 (2013) 1544–51

[13] LeMoyne R, Mastroianni T, and Grundfest W 2013 Wireless accelerometer configuration for monitoring Parkinson's disease hand tremor Adv. Parkinson's Dis. 2 62–7

[14] Kassavetis P et al 2016 Developing a tool for remote digital assessment of Parkinson's disease Mov. Disorders Clin. Pract. 3 59–64

[15] Tao, Weijun, et al. "Gait analysis using wearable sensors." Sensors 12.2 (2012): 2255-2283.

[16] de Oliveira Gondim, Ihana Thaís Guerra, et al. "Portable accelerometers for the evaluation of spatio-temporal gait parameters in people with Parkinson's disease: an integrative review." Archives of Gerontology and Geriatrics 90 (2020): 104097.

[17] Marcante, Andrea, et al. "Foot pressure wearable sensors for freezing of gait detection in Parkinson's disease." Sensors 21.1 (2020): 128.

[18] Mazzetta, Ivan, et al. "Wearable sensors system for an improved analysis of freezing of gait in Parkinson's disease using electromyography and inertial signals." Sensors 19.4 (2019): 948.

[19] Bonora, Gianluca et al. "Gait initiation is impaired in subjects with Parkinson's disease in the OFF state: Evidence from the analysis of the anticipatory postural adjustments through wearable inertial sensors." Gait & posture vol. 51 (2017): 218-221. doi:10.1016/j.gaitpost.2016.10.017

[20] Ghislieri, Marco et al. “Wearable Inertial Sensors to Assess Standing Balance: A Systematic Review.” Sensors (Basel, Switzerland) vol. 19,19 4075. 20 Sep. 2019, doi:10.3390/s19194075

[21] Memar, Sara et al. “Quantification of whole-body bradykinesia in Parkinson's disease participants using multiple inertial sensors.” Journal of the neurological sciences vol. 387 (2018): 157-165. doi:10.1016/j.jns.2018.02.001

[22] Muro-De-La-Herran, Alvaro, Begonya Garcia-Zapirain, and Amaia Mendez-Zorrilla. "Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications." Sensors 14.2 (2014): 3362-3394.

[23] Tao, Weijun, et al. "Gait analysis using wearable sensors." Sensors 12.2 (2012): 2255-2283.

[24] Chen, Fangzhe, et al. "Gait Analysis Based Parkinson’s Disease Auxiliary Diagnosis System." Journal of Internet Technology 22.5 (2021): 991-999.

[25] Mazilu, Sinziana, et al. "The role of wrist-mounted inertial sensors in detecting gait freeze episodes in Parkinson’s disease." Pervasive and Mobile Computing 33 (2016): 1-16.

[26] Perez-Ibarra, Juan C., Adriano AG Siqueira, and Hermano I. Krebs. "Identification of gait events in healthy and Parkinson’s disease subjects using inertial sensors: A supervised learning approach." IEEE Sensors Journal 20.24 (2020): 14984-14993.

[27] Wang, Jingying, et al. “Measurement of Step Angle for Quantifying the Gait Impairment of Parkinson's Disease by Wearable Sensors: Controlled Study.” JMIR mHealth and uHealth vol. 8,3 e16650. 20 Mar. 2020, doi:10.2196/16650

[28] Morris, Meg, et al. "Three-dimensional gait biomechanics in Parkinson's disease: evidence for a centrally mediated amplitude regulation disorder." Movement disorders: official journal of the Movement Disorder Society 20.1 (2005): 40-50.

[29] Van Wegen, Erwin EH, et al. "Sensor assisted self-management in Parkinson's disease: A feasibility study of ambulatory posture detection and feedback to treat stooped posture." Parkinsonism & related disorders 46 (2018): S57-S61.

[30] Muro-De-La-Herran, Alvaro, Begonya Garcia-Zapirain, and Amaia Mendez-Zorrilla. "Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications." Sensors 14.2 (2014): 3362-3394.

[31] Bae, Joonbum, and Masayoshi Tomizuka. "A tele-monitoring system for gait rehabilitation with an inertial measurement unit and a shoe-type ground reaction force sensor." Mechatronics 23.6 (2013): 646-651.

[32] Shah, Jesal, et al. "Increased foot strike variability in Parkinson's disease patients with freezing of gait." Parkinsonism & related disorders 53 (2018): 58-63.

[33] Rainoldi, Alberto, G. Melchiorri, and I. Caruso. "A method for positioning electrodes during surface EMG recordings in lower limb muscles." Journal of neuroscience methods 134.1 (2004): 37-43.

[34] Geroin, Christian, et al. "Pisa syndrome in Parkinson's disease: electromyographic quantification of paraspinal and non-paraspinal muscle activity." Functional Neurology 32.3 (2017): 143.

[35] Marcante, Andrea et al. "Foot Pressure Wearable Sensors for Freezing of Gait Detection in Parkinson's Disease." Sensors (Basel, Switzerland) vol. 21,1 128. 28 Dec. 2020, doi:10.3390/s21010128

[36] Cando, O. Alvarado, K. Robles Hidalgo, and B. Chacón Palacios. "A low-cost vibratory stimulus system to mitigate freezing of gait in Parkinson's disease." 2016 IEEE ANDESCON. IEEE, 2016.

[37] Shalin, Gaurav, et al. "Prediction and detection of freezing of gait in Parkinson's disease from plantar pressure data using long short-term memory neural-networks." Journal of neuroengineering and rehabilitation 18.1 (2021): 1-15.

[38] Kugler, Patrick, et al. "Automatic recognition of Parkinson's disease using surface electromyography during standardized gait tests." 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2013.

[39] Yim, Daehyeon, et al. "Quantified activity measurement for medical use in movement disorders through IR-UWB radar sensor." Sensors 19.3 (2019): 688.

[40] Qi, Yongbin, et al. "Using wearable UWB radios to measure foot clearance during walking." 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2013.

[41] Wahab, Yufridin, and Norantanum Abu Bakar. "Gait analysis measurement for sport application based on ultrasonic system." 2011 IEEE 15th international symposium on consumer electronics (ISCE). IEEE, 2011.

[42] Ashhar, Karalikkadan, Cheong Boon Soh, and Keng He Kong. "A wearable ultrasonic sensor network for analysis of bilateral gait symmetry." 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2017.

[43] Blumrosen, G., et al. "Non-contact UWB radar technology to assess tremor." XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010. Springer, Berlin, Heidelberg, 2010.

[44] Chee, Rachel, et al. "Gait freezing in Parkinson's disease and the stride length sequence effect interaction." Brain 132.8 (2009): 2151-2160.

[45] Muñoz Ospina, Beatriz, et al. "Age matters: objective gait assessment in early parkinson's disease using an RGB-D camera." Parkinson's Disease 2019 (2019).

[46] Pachoulakis, Ioannis, and Konstantinos Kourmoulis. "Building a gait analysis framework for Parkinson's disease patients: Motion capture and skeleton 3D representation." 2014 International Conference on Telecommunications and Multimedia (TEMU). IEEE, 2014.

[47] Menz, Hylton B., et al. "Reliability of the GAITRite® walkway system for the quantification of temporo-spatial parameters of gait in young and older people." Gait & posture 20.1 (2004): 20-25.

[48] Middleton, Lee, et al. "A floor sensor system for gait recognition." Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05). IEEE, 2005.

[49] Mondal, Banashree, et al. "Analysis of gait in Parkinson's disease reflecting the effect of l-DOPA." Annals of Movement Disorders 2.1 (2019): 21.

[50] Yang, Yea-Ru, et al. "Relationships between gait and dynamic balance in early Parkinson's disease." Gait & posture 27.4 (2008): 611-615.

[51] Rehman, Rana Zia Ur, et al. "Comparison of walking protocols and gait assessment systems for machine learning-based classification of parkinson's disease." Sensors 19.24 (2019): 5363.

[52] Nelson, Arthur J., et al. "The validity of the GaitRite and the Functional Ambulation Performance scoring system in the analysis of Parkinson gait." NeuroRehabilitation 17.3 (2002): 255-262.

[53] Pratheepan, Yogarajah, Joan V. Condell, and Girijesh Prasad. "The use of dynamic and static characteristics of gait for individual identification." 2009 13th International Machine Vision and Image Processing Conference. IEEE, 2009.

[54] Iwashita, Yumi, Ryo Kurazume, and Koichi Ogawara. "Expanding gait identification methods from straight to curved trajectories." 2013 IEEE Workshop on Applications of Computer Vision (WACV). IEEE, 2013.

[55] Arias-Enriquez, Omar, Mario I. Chacon-Murguia, and Rafael Sandoval-Rodriguez. "Kinematic analysis of gait cycle using a fuzzy system for medical diagnosis." 2012 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS). IEEE, 2012.

[56] Tahir, Nooritawati Md, and Hany Hazfiza Manap. "Parkinson Disease Gait Classification based on Machine Learning Approach." Journal of Applied Sciences 12.2 (2012): 180-185.

[57] Tucker, Conrad S., et al. "Machine learning classification of medication adherence in patients with movement disorders using non-wearable sensors." Computers in biology and medicine 66 (2015): 120-134.

[58] Rocha, Ana Patrícia, et al. "Parkinson's disease assessment based on gait analysis using an innovative RGB-D camera system." 2014 36th annual international conference of the IEEE engineering in medicine and biology society. IEEE, 2014.

[59] Sabo, Andrea, et al. "Prediction of Parkinsonian Gait in Older Adults with Dementia using Joint Trajectories and Gait Features from 2D Video." 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2021.

[60] Xue, Zhaojun, et al. "Infrared gait recognition based on wavelet transform and support vector machine." Pattern recognition 43.8 (2010): 2904-2910.

[61] Abou, Libak, et al. "Gait and Balance Assessments using Smartphone Applications in Parkinson's Disease: A Systematic Review." Journal of medical systems 45.9 (2021): 1-20.

[62] Abujrida, Hamza, Emmanuel Agu, and Kaveh Pahlavan. "Smartphone-based gait assessment to infer Parkinson's disease severity using crowdsourced data." 2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT). IEEE, 2017.

[63] Abujrida, Hamza, Emmanuel Agu, and Kaveh Pahlavan. "Machine learning-based motor assessment of Parkinson's disease using postural sway, gait and lifestyle features on crowdsourced smartphone data." Biomedical Physics & Engineering Express 6.3 (2020): 035005.

[64] Zhan, Andong, et al. "High frequency remote monitoring of Parkinson's disease via smartphone: Platform overview and medication response detection." arXiv preprint arXiv:1601.00960 (2016).

[65] Bot, Brian M., et al. "The mPower study, Parkinson disease mobile data collected using ResearchKit." Scientific data 3 (2016): 160011.

[66] Kubota, Ken J., Jason A. Chen, and Max A. Little. "Machine learning for large-scale wearable sensor data in Parkinson's disease: Concepts, promises, pitfalls, and futures." Movement disorders 31.9 (2016): 1314-1326.

[67] Mazilu, Sinziana, et al. "Online detection of freezing of gait with smartphones and machine learning techniques." Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on. IEEE, 2012.

[68] Arora, Siddharth, et al. "Detecting and monitoring the symptoms of Parkinson's disease using smartphones: a pilot study." Parkinsonism & related disorders 21.6 (2015): 650-653.

[69] Ellis, Robert J., et al. "A validated smartphone-based assessment of gait and gait variability in Parkinson's disease." PLoS one 10.10 (2015): e0141694

[70] Printy, Blake P., et al. "Smartphone application for classification of motor impairment severity in Parkinson's disease." Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE. IEEE, 2014.

[71] Kim, Hanbyul, et al. "Unconstrained detection of freezing of Gait in Parkinson's disease patients using smartphone." Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE. IEEE, 2015.

[72] Sharma, Vinod, et al. "SPARK: personalized parkinson disease interventions through synergy between a smartphone and a smartwatch." International Conference of Design, User Experience, and Usability. Springer, Cham, 2014.

[73] Zhan, Andong, et al. "High frequency remote monitoring of Parkinson's disease via smartphone: Platform overview and medication response detection." arXiv preprint arXiv:1601.00960 (2016).

[74] Lee, Chae Young, et al. "A validation study of a smartphone-based finger tapping application for quantitative assessment of bradykinesia in Parkinson's disease." PloS one 11.7 (2016): e0158852.

[75] Zhang, Hanbin, et al. "Pdmove: Towards passive medication adherence monitoring of parkinson's disease using smartphone-based gait assessment." Proceedings of the ACM on interactive, mobile, wearable, and ubiquitous technologies 3.3 (2019): 1-23.

[76] Kan, Valerii, et al. "STOP: A smartphone-based game for Parkinson's disease medication adherence." Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. 2018.

[77] Lakshminarayana, Rashmi, et al. "Using a smartphone-based self-management platform to support medication adherence and clinical consultation in Parkinson's disease." NPJ Parkinson's disease 3.1 (2017): 1-10.

[78] Lo, Christine, et al. "Predicting motor, cognitive & functional impairment in Parkinson's." Annals of clinical and translational neurology 6.8 (2019): 1498-1509.

[79] Fiems, Connie L., et al. "Does a sway-based mobile application predict future falls in people with Parkinson disease?." Archives of physical medicine and rehabilitation 101.3 (2020): 472-478.

[80] Pepa, Lucia, et al. "A fuzzy logic system for the home assessment of freezing of gait in subjects with Parkinsons disease." Expert Systems with Applications 147 (2020): 113197.

[81] Chén, Oliver Y., et al. "Building a machine-learning framework to remotely assess Parkinson's disease using smartphones." IEEE Transactions on Biomedical Engineering 67.12 (2020): 3491-3500.

[82] Su, Dongning, et al. "Simple smartphone-based assessment of gait characteristics in Parkinson disease: validation study." JMIR mHealth and uHealth 9.2 (2021): e25451.

[83] Borzì, Luigi, et al. "Smartphone-based evaluation of postural stability in Parkinson's disease patients during quiet stance." Electronics 9.6 (2020): 919.

[84] Memedi, Mevludin, et al. "Automatic and objective assessment of alternating tapping performance in Parkinson's disease." Sensors 13.12 (2013): 16965-16984.

[85] Grosset, Katherine A., Ian Bone, and Donald G. Grosset. "Suboptimal medication adherence in Parkinson's disease." Movement disorders: official journal of the Movement Disorder Society 20.11 (2005): 1502-1507.

[86] Grosset, Donald, et al. "Adherence to antiparkinson medication in a multicenter European study." Movement disorders: official journal of the Movement Disorder Society 24.6 (2009): 826-832.

[87] Leopold, Norman A., Marcia Polansky, and Marcia R. Hurka. "Drug adherence in Parkinson's disease." Movement disorders: official journal of the Movement Disorder Society 19.5 (2004): 513-517.

[88] Kalantarian, Haik, Nabil Alshurafa, and Majid Sarrafzadeh. "Detection of gestures associated with medication adherence using smartwatch-based inertial sensors." IEEE Sensors Journal 16.4 (2015): 1054-1061.

[89] Koesmahargyo, Vidya, et al. "Accuracy of machine learning-based prediction of medication adherence in clinical research." Psychiatry Research 294 (2020): 113558.

[90] Van Den Eeden, Stephen K., et al. "Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity." American journal of epidemiology 157.11 (2003): 1015-1022.

[91] Ishikawa, Atsushi, and Tadashi Miyatake. "Effects of smoking in patients with early-onset Parkinson's disease." Journal of the neurological sciences 117.1-2 (1993): 28-32.

[92] Fertl, E., A. Doppelbauer, and E. Auff. "Physical activity and sports in patients suffering from Parkinson's disease in comparison with healthy seniors." Journal of Neural Transmission-Parkinson's Disease and Dementia Section 5.2 (1993): 157-161.

[93] Reuter, Iris, et al. "Therapeutic value of exercise training in Parkinson's disease." Medicine and Science in Sports and Exercise 31.11 (1999): 1544-1549.

[94] Haslinger, B., et al. "Event-related functional magnetic resonance imaging in Parkinson's disease before and after levodopa." Brain 124.3 (2001): 558-607

[95] Ma, Yuchao, et al. "CyclePro: A robust framework for domain-agnostic gait cycle detection." IEEE Sensors Journal 19.10 (2019): 3751-3762.

[96] Mayeux, Richard, et al. "A population-based investigation of Parkinson's disease with and without dementia: relationship to age and gender." Archives of Neurology 49.5 (1992): 492-497.

[97] Wang, Zhiguang, and Tim Oates. "Encoding time series as images for visual inspection and classification using tiled convolutional neural networks." Workshops at the twenty-ninth AAAI conference on artificial intelligence. 2015.

[98] Breiman, Leo. "Random forests." Machine learning 45.1 (2001): 5-32.

[99] Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2:18–22

[100] Cutler, D. Richard, et al. "Random forests for classification in ecology." Ecology 88.11 (2007): 2783-2792.

[101] Strobl, Carolin, et al. "Conditional variable importance for random forests." BMC Bioinformatics 9.1 (2008): 307.

[102] Genuer R, Poggi J-M, Tuleau-Malot C (2010) Variable selection using random forests. Pattern RecognitLett 31:2225–2236

[103] Díaz-Uriarte R, Alvarez de Andrés S (2006) Gene selection and classification of microarray data usingrandom forest. BMC Bioinform 7:1–13

[104] Deepak, S., and P. M. Ameer. "Brain tumor classification using deep CNN features via transfer learning." Computers in biology and medicine 111 (2019): 103345.

[105] Hussein, Sarfaraz, et al. "Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches." IEEE transactions on medical imaging 38.8 (2019): 1777-1787.

[106] Deniz, Erkan, et al. "Transfer learning based histopathologic image classification for breast cancer detection." Health information science and systems 6.1 (2018): 1-7.

[107] Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009.

[108] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[109] Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." International conference on machine learning. PMLR, 2019.

[110] Yu, Tao, et al. "A multi-layer parallel lstm network for human activity recognition with smartphone sensors." 2018 10th International conference on wireless communications and signal processing (WCSP). IEEE, 2018.

[111] Ding, Congzhang, et al. "Continuous Human Activity Recognition through Parallelism LSTM with Multi-Frequency Spectrograms." Remote Sensing 13.21 (2021): 4264.

[112] Xia, Kun, Jianguang Huang, and Hanyu Wang. "LSTM-CNN architecture for human activity recognition." IEEE Access 8 (2020): 56855-56866.

[113] Anwary, Arif Reza, et al. "Insole-based Real-time Gait Analysis: Feature Extraction and Classification." 2021 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL). IEEE, 2021.

[114] Zou, Qin, et al. "Deep learning-based gait recognition using smartphones in the wild." IEEE Transactions on Information Forensics and Security 15 (2020): 3197-3212.

[115] Zeng, Xin, et al. "Gait-based implicit authentication using edge computing and deep learning for mobile devices." Sensors 21.13 (2021): 4592

[116] Sawicki, A., and K. Saeed. "Application of LSTM Networks for Human Gait-Based Identification." International Conference on Dependability and Complex Systems. Springer, Cham, 2021.

[117] Donahue, Jeffrey, et al. "Long-term recurrent convolutional networks for visual recognition and description." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[118] Parkinson's Foundation "Statistics" https://www.parkinson.org/Understanding-Parkinsons/Statistics , Accessed on April 2022

[119] U.S. Foods and Drugs Administration (FDA) "Why You Need to Take Your Medications as Prescribed or Instructed" https://www.fda.gov/drugs/special-features/why-you-need-take-your-medications-prescribed-or-instructed. Accessed on June 2021

[120] Straka, Igor, et al. "Adherence to pharmacotherapy in patients with Parkinson's disease taking three and more daily doses of medication." Frontiers in neurology 10 (2019): 799.

[121] Schrag, Anette, et al. "Caregiver-burden in Parkinson's disease is closely associated with psychiatric symptoms, falls, and disability." Parkinsonism & related disorders 12.1 (2006): 35-41.

[121] Ramachandran, Mahesh, Ashok Veeraraghavan, and Rama Chellappa. "Video stabilization and mosaicing." The Essential Guide to Video Processing (2009): 109-140.

[122] Maranesi, Elvira, et al. "A stereophotogrammetric-based method to assess spatio-temporal gait parameters on healthy and Parkinsonian subjects." 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2015.

[123] Joshi, Sandhya, et al. "Classification of Alzheimer's disease and Parkinson's disease by using machine learning and neural network methods." 2010 Second International Conference on Machine Learning and Computing. IEEE, 2010.

[124] Hubbard, Raymond, and R. Murray Lindsay. "Why p values are not a useful measure of evidence in statistical significance testing." Theory & Psychology 18.1 (2008): 69-88.

[125] Fulton, Lawrence V., et al. "Classification of Alzheimer's disease with and without imagery using gradient boosted machines and ResNet-50." Brain sciences 9.9 (2019): 212.

[126] Al-Haija, Qasem Abu, and Adeola Adebanjo. "Breast cancer diagnosis in histopathological images using ResNet-50 convolutional neural network." 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). IEEE, 2020.

[127] Ni, Zhongfei, and Binke Huang. "Human identification based on natural gait micro-Doppler signatures using deep transfer learning." IET Radar, Sonar & Navigation 14.10 (2020): 1640-1646.

[128] Hoang, Ngoc Son, et al. "Gait classification for Parkinson's Disease using Stacked 2D and 1D Convolutional Neural Network." 2019 International Conference on Advanced Technologies for Communications (ATC). IEEE, 2019.

[129] El Maachi, Imanne, Guillaume-Alexandre Bilodeau, and Wassim Bouachir. "Deep 1D-Convnet for accurate Parkinson disease detection and severity prediction from gait." Expert Systems with Applications 143 (2020): 113075.

[130] Paul, Kimberly C., et al. "The association between lifestyle factors and Parkinson's disease progression and mortality." Movement Disorders 34.1 (2019): 58-66.

[131] Kaur, Rachneet, et al. "A Vision-Based Framework for Predicting Multiple Sclerosis and Parkinson's Disease Gait Dysfunctions-A Deep Learning Approach." IEEE Journal of Biomedical and Health Informatics (2022).

[132] 25. InformedHealth.org, “Parkinson’s disease: Medication for Parkinson's disease” https://www.ncbi.nlm.nih.gov/books/NBK293715/, Accessed on Oct 2022

[133] Everitt, Brian S., and Anders Skrondal. "The Cambridge dictionary of statistics." (2010).

[134] Titova, Nataliya, and K. Ray Chaudhuri. "Personalized medicine in Parkinson's disease: time to be precise." Movement Disorders 32.8 (2017): 1147-1154.

[135] Stisen, Allan, et al. "Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition." Proceedings of the 13th ACM conference on embedded networked sensor systems. 2015.

APPENDICES

APPENDIX A

MPOWER DEMOGRAPHICS SURVEY.

From: The mPower study, Parkinson disease mobile data collected using ResearchKit

Question	Variable name	Variable details
How old are you?	age	integer
What is your sex?	gender	one of {'Female', 'Male', 'Prefer not to answer'}
Which race do you identify with?	race	check all that apply {'Black or African', 'Latino/Hispanic', 'Native American', 'Pacific Islander', 'Middle Eastern', 'Caribbean', 'South Asian', 'East Asian', 'White or Caucasian', 'Mixed'}
What is the highest level of education that you have completed?	education	one of {'2-year college degree', '4-year college degree', 'Doctoral Degree', 'High School Diploma/GED', 'Master's Degree', 'Some college', 'Some graduate school', 'Some high school'}
What is your current employment status?	employment	one of {'A homemaker', 'A student', 'Employment for wages', 'Out of work', 'Retired', 'Self-employed', 'Unable to work'}
What is your current marital status?	maritalStatus	one of {'Divorced', 'Married or domestic partnership', 'Other', 'Separated', 'Single, never married', 'Widowed'}
Are you a spouse, partner or care-partner of someone who has Parkinson disease?	are-caretaker	one of {'true', 'false'}

Question	Variable name	Variable details
Have you ever participated in a research study or clinical trial on Parkinson disease before?	past-participation	one of {‘true’, ‘false’}
How easy is it for you to use your smartphone?	smartphone	one of {‘Difficult’, ‘Easy’, ‘Neither easy nor difficult’, ‘Very Difficult’, ‘Very easy’}
Do you ever use your smartphone to look for health or medical information online?	phone-usage	one of {‘true’, ‘false’, ‘Not sure’}
Do you use the Internet or email at home?	home-usage	one of {‘true’, ‘false’}
Do you ever use the Internet to look for health or medical information online?	medical-usage	one of {‘true’, ‘false’}
Did you happen to do this yesterday, or not?	medical-usage-yesterday	one of {‘true’, ‘false’, ‘don’t know’}
Do you ever use your smartphone to participate in a video call or video chat?	video-usage	one of {‘true’, ‘false’}

Question	Variable name	Variable details
Have you been diagnosed by a medical professional with Parkinson disease?	professional-diagnosis	one of {‘true’, ‘false’}
In what year did your movement symptoms begin?	onset-year	integer input
In what year were you diagnosed with Parkinson disease?	diagnosis-year	integer input
In what year did you begin taking Parkinson disease medication? Type in 0 if you have not started to take Parkinson medication.	medication-start-year	integer input
What kind of health care provider currently cares for your Parkinson disease?	healthcare-provider	one of {‘Don't know’, ‘General Neurologist (non-Parkinson Disease specialist)’, ‘Nurse Practitioner or Physician's Assistant’, ‘Other’, ‘Parkinson Disease/Movement Disorder Specialist’, ‘Primary Care Doctor’}
Have you ever had Deep Brain Stimulation?	deep-brain-stimulation	one of {‘true’, ‘false’}
Have you ever had any surgery for Parkinson	surgery	one of {‘true’, ‘false’}

Question	Variable name	Variable details
disease, other than DBS?		
Have you ever smoked?	smoked	one of {'true', 'false'}
How many years have you smoked?	years-smoking	integer input
On average, how many packs did you smoke each day?	packs-per-day	one of {1, 2, 3, 4, 5}
When is the last time you smoked (put today's date if you are still smoking)?	last-smoked	year last smoked
Has a doctor ever told you that you have any of the following conditions? Please check all that apply.	health-history	Multiple choice from {'Acute Myocardial Infarction/Heart Attack', 'Alzheimer Disease or Alzheimer dementia', 'Atrial Fibrillation', 'Anxiety', 'Cataract', 'Kidney Disease', 'Chronic Obstructive Pulmonary Disease (COPD) or Asthma', 'Heart Failure/Congestive Heart Failure', 'Diabetes or Prediabetes or High Blood Sugar', 'Glaucoma', 'Hip/Pelvic Fracture', 'Ischemic Heart Disease', 'Depression', 'Osteoporosis', 'Rheumatoid Arthritis', 'Dementia', 'Stroke/Transient Ischemic Attack (TIA)', 'Breast Cancer', 'Colorectal Cancer', 'Prostate Cancer', 'Lung Cancer', 'Endometrial/Uterine Cancer', 'Any other kind of cancer OR tumor','Head Injury with Loss of Consciousness/Concussion', 'Urinary Tract infections', 'Obstructive Sleep Apnea', 'Schizophrenia or Bipolar Disorder', 'Peripheral

Question	Variable name	Variable details
		Vascular Disease', 'High Blood Pressure/Hypertension', 'Fainting/Syncope', 'Alcoholism', 'Multiple Sclerosis', 'Impulse control disorder', 'AIDS or HIV', 'Liver Disease', 'Leukemia or Lymphoma', 'Ulcer Disease', 'Connective Tissue Disease', 'Coronary Artery Disease', 'Anemia','Asthma'}

APPENDIX B
(MDS-UPDRS)Part II: Motor Aspects of Experiences of Daily Living (M-EDL)

2.4 EATING TASKS — SCORE

Over the past week, have you usually had troubles handling your food and using eating utensils? For example, do you have trouble handling finger foods or using forks, knives, spoons, chopsticks?

- 0: Normal: Not at all (no problems).
- 1: Slight: I am slow, but I do not need any help handling my food and have not had food spills while eating.
- 2: Mild: I am slow with my eating and have occasional food spills. I may need help with a few tasks such as cutting meat.
- 3: Moderate: I need help with many eating tasks but can manage some alone.
- 4: Severe: I need help for most or all eating tasks.

2.5 DRESSING

Over the past week, have you usually had problems dressing? For example, are you slow or do you need help with buttoning, using zippers, putting on or taking off your clothes or jewelry?

- 0: Normal: Not at all (no problems).
- 1: Slight: I am slow, but I do not need help.
- 2: Mild: I am slow and need help for a few dressing tasks (buttons, bracelets).
- 3: Moderate: I need help for many dressing tasks.
- 4: Severe: I need help for most or all dressing tasks.

2.6 HYGIENE — SCORE

Over the past week, have you usually been slow or do you need help with washing, bathing, shaving, brushing teeth, combing your hair, or with other personal hygiene?

- 0: Normal: Not at all (no problems).
- 1: Slight: I am slow, but I do not need any help.
- 2: Mild: I need someone else to help me with some hygiene tasks.
- 3: Moderate: I need help for many hygiene tasks.
- 4: Severe: I need help for most or all of my hygiene tasks.

2.7 HANDWRITING

Over the past week, have people usually had trouble reading your handwriting?

- 0: Normal: Not at all (no problems).
- 1: Slight: My writing is slow, clumsy or uneven, but all words are clear.
- 2: Mild: Some words are unclear and difficult to read.
- 3: Moderate: Many words are unclear and difficult to read.
- 4: Severe: Most or all words cannot be read.

2.8 DOING HOBBIES AND OTHER ACTIVITIES

Over the past week, have you usually had trouble doing your hobbies or other things that you like to do?

- 0: Normal: Not at all (no problems).
- 1: Slight: I am a bit slow but do these activities easily.
- 2: Mild: I have some difficulty doing these activities.
- 3: Moderate: I have major problems doing these activities, but still do most.
- 4: Severe: I am unable to do most or all of these activities.

1.13 FATIGUE — SCORE

Over the past week, have you usually felt fatigued? This feeling is not part of being sleepy or sad.

- 0: Normal: No fatigue.
- 1: Slight: Fatigue occurs. However it does not cause me troubles doing things or being with people.
- 2: Mild: Fatigue causes me some troubles doing things or being with people.
- 3: Moderate: Fatigue causes me a lot of troubles doing things or being with people. However, it does not stop me from doing anything.
- 4: Severe: Fatigue stops me from doing things or being with people.

Part II: Motor Aspects of Experiences of Daily Living (M-EDL)

2.1 SPEECH

Over the past week, have you had problems with your speech?

- 0: Normal: Not at all (no problems).
- 1: Slight: My speech is soft, slurred or uneven, but it does not cause others to ask me to repeat myself.
- 2: Mild: My speech causes people to ask me to occasionally repeat myself, but not every day.
- 3: Moderate: My speech is unclear enough that others ask me to repeat myself every day even though most of my speech is understood.
- 4: Severe: Most or all of my speech cannot be understood.

2.2 SALIVA AND DROOLING — SCORE

Over the past week, have you usually had too much saliva during when you are awake or when you sleep?

- 0: Normal: Not at all (no problems).
- 1: Slight: I have too much saliva, but do not drool.
- 2: Mild: I have some drooling during sleep, but none when I am awake.
- 3: Moderate: I have some drooling when I am awake, but I usually do not need tissues or a handkerchief.
- 4: Severe: I have so much drooling that I regularly need to use tissues or a handkerchief to protect my clothes.

2.3 CHEWING AND SWALLOWING

Over the past week, have you usually had problems swallowing pills or eating meals? Do you need your pills cut or crushed or your meals to be made soft, chopped, or blended to avoid choking?

- 0: Normal: No problems.
- 1: Slight: I am aware of slowness in my chewing or increased effort at swallowing, but I do not choke or need to have my food specially prepared.
- 2: Mild: I need to have my pills cut or my food specially prepared because of chewing or swallowing problems, but I have not choked over the past week.
- 3: Moderate. I choked at least once in the past week.
- 4: Severe: Because of chewing and swallowing problems, I need a feeding tube.

2.9 TURNING IN BED	SCORE
Over the past week, do you usually have trouble turning over in bed? 0: Normal: Not at all (no problems). 1: Slight: I have a bit of trouble turning, but I do not need any help. 2: Mild I have a lot of trouble turning and need occasional help from someone else. 3: Moderate: To turn over I often need help from someone else. 4: Severe: I am unable to turn over without help from someone else.	☐
2.10 TREMOR	
Over the past week, have you usually had shaking or tremor? 0: Normal: Not at all. I have no shaking or tremor. 1: Slight: Shaking or tremor occurs but does not cause problems with any activities. 2: Mild: Shaking or tremor causes problems with only a few activities. 3: Moderate: Shaking or tremor causes problems with many of my daily activities. 4: Severe: Shaking or tremor causes problems with most or all activities.	☐
2.11 GETTING OUT OF BED, A CAR, OR A DEEP CHAIR	
Over the past week, have you usually had trouble getting out of bed, a car seat, or a deep chair? 0: Normal: Not at all (no problems). 1: Slight: I am slow or awkward, but I usually can do it on my first try. 2: Mild: I need more than one try to get up or need occasional help. 3: Moderate: I sometimes need help to get up, but most times I can still do it on my own. 4: Severe: I need help most or all of the time.	☐

2.12 WALKING AND BALANCE	SCORE
Over the past week, have you usually had problems with balance and walking? 0: Normal: Not at all (no problems). 1: Slight: I am slightly slow or may drag a leg. I never use a walking aid. 2: Mild: I occasionally use a walking aid, but I do not need any help from another person. 3: Moderate: I usually use a walking aid (cane, walker) to walk safely without falling. However, I do not usually need the support of another person. 4: Severe: I usually use the support of another person to walk safely without falling.	☐
2.13 FREEZING	
Over the past week, on your usual day when walking, do you suddenly stop or freeze as if your feet are stuck to the floor? 0: Normal: Not at all (no problems). 1: Slight: I briefly freeze, but I can easily start walking again. I do not need help from someone else or a walking aid (cane or walker) because of freezing. 2: Mild: I freeze and have trouble starting to walk again, but I do not need someone's help or a walking aid (cane or walker) because of freezing. 3: Moderate: When I freeze I have a lot of trouble starting to walk again and, because of freezing, I sometimes need to use a walking aid or need someone else's help. 4: Severe: Because of freezing, most or all of the time, I need to use a walking aid or someone's help.	☐

This completes the questionnaire. We may have asked about problems you do not even have, and may have mentioned problems that you may never develop at all. Not all patients develop all these problems, but because they can occur, it is important to ask all the questions to every patient. Thank you for your time and attention in completing this questionnaire.