Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

AI Based Early Detection of MotorDisorder Symptoms in Parkinson'sDisease using Tremor and Gait Patterns

Sanjeevini R

sanjeevini2110282@ssn.edu.in

Vasundhara B
Safiyyah N
Mirnalinee T T
Shahina A

Research Article

Keywords:

Posted Date: February 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8914252/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

AI Based Early Detection of Motor Disorder Symptoms in Parkinson's Disease using Tremor and Gait Patterns

Sanjeevini R[1], Vasundhara B[1], Safiyyah N[2,*], T T Mirnalinee[2], and Shahina A[1]

[1]Department of Information Technology, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam 603110, India
[2]Department of Computer Science and Engineering, Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam 603110, India
[*]Correponding author
sanjeevini2110282@ssn.edu.in, vasundhara2110257@ssn.edu.in, safiyyahn@ssn.edu.in, mirnalineett@ssn.edu.in, shahinaa@ssn.edu.in

February 19, 2026

Abstract

Movement disorders due to diseases like Parkinson's negatively affect any motor activity thereby causing symptoms including action tremor and changes in gait. The tracking of these symptoms is crucial for an early detection of the worsening condition and for appropriate changes in the medications. This study explores the potential of digital biomarkers to track motor symptoms in Parkinson's Disease (PD), using two primary assessment tools: the spiral drawing test using a digitized writing tablet and the gait analysis using Inertial Measurement Units (IMU) sensors. The spiral test is used to measure fine motor movement and properly quantify the intensity of the tremors and motor coordination. At the same time, spatiotemporal gait characteristics and gait symmetry data from IMU sensors characterize movement disorders whilst performing an activity such as walking. For a quantitative assessment and tracking of these biomarkers is done using various machine learning and deep learning classification algorithms. The best performing models, Swin Transformer (90.9%) and MLP with residual connections (83.9%) for the digital spiral drawing and IMU-based gait measurements, respectively, show that they form an accurate and trustworthy method for measuring Parkinson's dis-

ease. This gives a strong platform for disease evaluation improvement and optimization of treatment methods.

1 Introduction

Parkinson's Disease (PD) is a progressive neurodegenerative disease with a range of various symptoms that worsen over time. Presence of motor-related symptoms is one of the key features of PD, that starts off lightly in the early stages of the disease and becomes more prominent with progression. Among common motor symptoms associated with Parkinson's Disease, action tremors are significant. Action tremors occur during voluntary movements and are often more pronounced in later stages of the disease. Such tremors can cause a great impact on the patient's quality of life [1]. Gait disturbances are often evident in patients with PD, serving as critical diagnostic markers. These are usually reduced stride, shuffling steps and irregularity in gait, that result in the characteristic stooped posture observed in many patients with PD [2]. Freezing of gait (FoG) is a disabling symptom where the patient is unable to move their feet, leading to fatal falls and severe mobility limitations [3]. All such motor impairments are of great importance in clinical assessments, aiding early diagnosis and monitoring progression of PD.

In this work, the decline of motor functions of PD is studied using two digital tests: (i) Spiral drawing tests, and (ii) gait analysis using machine learning and deep learning techniques. The earlier works (discussed in detail in the following section) on artificial intelligence based PD tremor detection predominantly used either hand crafted features or standard convolutional neural networks. Both are more effective in capturing the local patterns and not as effective in capturing the global (or long-range) spatial dependencies present in the spiral data that capture the dynamics of tremor. Such a limitation could affect the distinguishing of the severity of tremor in PD patients. Also, the earlier works do not effectively model the sequential movements that form the spiral drawing. These spatial and sequential features that are relevant to distinguish the PD could be well captured through attention mechanisms that could help focus on disease-relevant features. Prior studies in gait analysis for PD detection that use CNNs for extracting the features, which are then given to sequential models such as LSTMs for temporal modeling introduce complexity in the architecture and increase the computational cost. But in spite of this, they do not guarantee an enhancement in performance, since the datasets are usually small and imbalanced, causing them to overfit. Lightweight models that can better model the temporal dependencies with fewer parameters, and that are scalable for real time applications need to be explored.

Based on the observations from earlier works as mentioned above, the contributions of this work are as follows:

1. Identifying the tremor in PD from the Archimedes spiral drawings using

the attention mechanism of Swin transformer that better captures the long range dependencies of the tremor dynamics.

2. Detecting PD from gait patterns using the lightweight Multi Layer Perceptron with residual connections that model the temporal dependencies with fewer parameters.

The paper is organized as follows: Section 2 reviews the earlier works on motor symptoms of PD, where mostly researchers have addressed tremor and gait using machine learning algorithms separately. Sections 3 and 4 elaborate on the proposed methodology and the results obtained. Section 5 discusses the performance of the models. This work is summarized with a conclusion in Section 6.

2 Literature Review

Methods to quantify action tremors in later stages of Parkinson's disease have been enhanced by spiral drawing analysis where participants are asked to trace an Archimedes' spiral. Novel metrics were introduced in [4] to quantify tremors effectively, hence demonstrating that spiral drawing analysis can be an objective measure of action tremor in patients with PD. With the spiral drawing method as the base, many papers have experimented the analysis of guided spiral drawings for PD classification using machine learning, where participants create specific patterns. [5] captured consistent motor symptoms from the spiral drawing features that were analyzed for improvement in PD diagnosis. [6] examined the potential of digitized spiral drawings as a tool to assess motor control beyond handwriting. Digital traces of spiral drawings traced by PD patients have been reviewed to evaluate movement dynamics, implying that these traces provide insightful biomarkers of PD severity. This has presented an alternative to traditional handwriting assessments and offers a better and comprehensive evaluation of kinetic function in PD.

A recent study [7] proposed a new approach where a coordinate attention-enhanced Swin Transformer architecture was incorporated, that has been specifically trained for handwriting recognition in PD patients. This transformer-based model has proved to be excellent at capturing spatial and sequential features in handwriting data, thus increasing the diagnostic accuracy from handwriting patterns. Convolutional Neural Networks (CNNs) along with transfer learning have also been experimented with, using handwriting and drawing patterns, to classify PD. [8] recognized the potential of using pre-trained CNN architectures as promising tools for PD detection, as they were able to effectively analyze subtle motor impairments in handwriting and drawing tasks.

Common abnormalities in gait in PD patients include stride length, slower walking speed, increased variability, and an uneven rhythm in steps. Detecting such subtle gait changes has been improved using sensor-based detection and

deep learning techniques. Models trained on huge amount of sensor data have been able to identify unique movement patterns that are used to distinguish between healthy and PD patients [9].

Freezing of Gait (FoG), the sudden inability to continue or start walking, is one of the most challenging symptoms in patients with PD. Various techniques for the detection of FoG have been studied and results for wearable sensors, particularly Inertial Measurement Units (IMUs), for this purpose have been promising [10]. Data recorded by IMUs allow for precise FoG detection and providing real-time alerts to patients during episodes to help prevent falls. [11] studied various deep learning approaches for the detection of Parkinson's disease using data from wearable sensors, where the combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks improved the classification accuracy of gait patterns in PD, including FoG. This study had highlighted that the combination of IMU data with deep learning techniques can be significant in enhancing the diagnosis and monitoring of PD.

Deep learning models in PD applications have also been able to analyze extensively the foot-to-ground phases in the gait of patients. [12] compared data features such as stance and swing phases to improve the adaptation of deep learning architectures to PD gait analysis. [13] discussed the use of wearable sensors for monitoring of motor fluctuations in PD at home. Findings from this paper show that everyday monitoring records motor patterns in PD that can be effectively analyzed and interpreted, which may otherwise go unnoticed in clinical assessments.

A real-time system, deFoG, was designed by [14] for the detection and fall-prevention warning of FoG in PD patients. Gait patterns were monitored and immediate auditory feedback was given by the system as an alert to encourage the patient to continue walking. This demonstrates the feasibility of real-time intervention for PD patients, where wearable-sensor based systems prove to be effective in detecting and reducing FoG occurrences.

[15] reviewed the effects of types of tasks, states of medication and pause periods on the accuracy of FoG detection. Machine learning models trained on IMU data improved the identification and understanding of FoG under various conditions. Their findings show that deep learning-enhances IMU systems have the potential of improved and reliable assessments in clinical settings, allowing them to monitor and manage gait symptoms in patients with PD.

3 Proposed Methodology

This section discusses the methodology used to indentify PD through (1) action tremor (2) gait analysis. While the tremor study uses the Archimedies' spiral test, the gait study uses the IMU based data.

3.1 Dataset overview

The dataset used for this study is the Parkinson's Disease Drawing dataset developed by K. Mader and consists of digitized spiral and meander drawings recorded from individuals with Parkinson's Disease (PD) and healthy controls [16]. Each sample in the dataset is labeled to indicate whether it corresponds to a control subject or a PD patient. The spiral drawings, obtained via a digitized tablet, serve as a means to assess motor skills affected by PD-related tremors and motor impairments.

3.1.1 Preprocessing

The dataset consists of images of varying dimensions, hence some preprocessing steps are required to standardize these inputs for model training, as follows:

1. Normalization and Resizing: All images are resized to match the input dimensions required by the Swin Transformer model (224x224 pixels). To maintain a consistent color balance, each image is normalized using the mean and standard deviation values provided by the image processor, which improves convergence during training by aligning the pixel intensity distributions across samples.

2. Data Augmentation: To prevent the overfitting, augmentation techniques are applied that enhances the model's generalization ability. These techniques include (i) randomly cropping portions of the image and resizing to 224x224 pixels, which helps the model learn invariant features by focusing on different parts of the spiral drawings, and (ii) randomly flipping images horizontally to make the model robust to minor asymmetries in drawing patterns that should not affect diagnostic classification. These augmentation techniques are applied to the training batch only, leaving the validation set unaltered for an unbiased evaluation.

The Swin Transformer model [17] is chosen for its high performance in visual tasks requiring complex pattern recognition, due to its hierarchical structure and shifting window approach which optimizes spatial pattern capture while also maintaining computational efficiency. The model partitions the images into small fixed-size patches, that enables the model to focus on regions of interest, making it well-suited for detecting Parkinson's disease (PD) tremors. Specifically, the `microsoft/swin-tiny-patch4-window7-224` variant is selected for this study due its compact size and accuracy. Its windowed attention mechanism captures localized features, that is crucial for identifying minute and subtle variations in spiral drawings associated with PD, and this is further fine-tuned to enhance the diagnostic precision.

The architecture design of the Swin Transformer model as shown in Fig. 1, employs effective extraction of both basic and complex features through layered shifting windows. This enables the model to capture nuances in the drawings. The initial layers capture fundamental patterns, such as overall shape, line continuity, and drawing consistency. The intermediate layers identify tremor-related

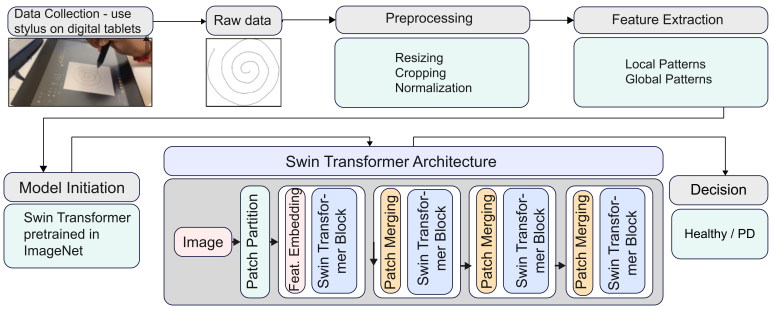


Figure 1: Process flow for Archimedes' spiral test

irregularities, essential for analyzing movement consistency and deviations. The deep layers detect high-level motor patterns such as drawing stability and speed variations, critical for distinguishing PD-specific tremors. Overall the, layered approach of the Swin Transformer model's to extract features enhances the detection and analysis of subtle motor dysfunctions in PD.

3.1.2 Model fine-tuning and training pipeline

The Swin Transformer model is loaded with a pre-trained checkpoint, leveraging its generalized visual features from large-scale training to reduce fine-tuning requirements. The transfer learning approach improves the model's performance on the PD classification task using the spiral drawing dataset. The dataset is split into two parts, where a training data (90%) and test (validation) data (10%) enable model tuning while maintaining a separate dataset for validation. The custom transforms are used: (1) Training Transforms: Preprocessing included augmentation, normalization, and resizing to improve generalization,and (2) Validation Transforms: Applied resizing and centering (without augmentation) to reflect real performance on unaltered samples.

To ensure stability, we observed the validation accuracy per epoch, saved the best model checkpoint, and set `load_best_model_at_end=True` to ensure optimal performance.

3.1.3 Classification and analysis

The Swin Transformer model assesses subjects as PD or control group and also quantifies motor impairment severity by analyzing tremor features in spiral drawings. The model calculates a tremor dominance probability, which is a probability score reflecting tremor intensity and pattern presence. This probability score is derived from the model's attention to PD-specific features, like irregularities in the curvature and the smoothness of lines.

3.2 Gait analysis

3.2.1 Dataset overview

The dataset, created by O'Day et al. (2019), [18] consists of gait parameters measured using wearable inertial measurement unit (IMU) sensors attached to the lateral shanks of participants, including individuals with Parkinson's disease (PD) and healthy controls of similar age. The subjects undertook two gait tasks: a Turning and Barrier Course (TBC) designed to provoke FOG episodes, and a 40-meter straight walking task, which aimed at assessing the clinical aspects of Parkinsonian gait. Among the PD group, 20 subjects completed the tasks, with 11 returning for a follow-up session at least one year later, yielding 31 PD records. Also, nine other healthy control participants underwent the same procedure, thus a total of 40 records are available. The dataset labeled the participants in 3 categories: 0 - No Freezing, 1 - PD (with FOG), PD (without FOG).

For our analysis, only the Forty-Meter Straight Walking (FW) segment of the dataset is used to analyze gait parameters in a linear walking activity. This allows the standardization of gait conditions among participants, making it efficient for real-time testing of the system on a patient.

3.2.2 Preprocessing

The data is preprocessed by normalising records and eliminating unjustified zeros and -1s from the records, that are used as placeholders, that might likely skew the results. The alignment techniques are used in an attempt to match up the sampling rates and to align the data which makes up all the 10-meter walking segments contained in the FW task. This ensures that the feature extraction for every subject remains constant in comparision to that of the other subjects.

3.2.3 Feature extraction

The time series features are derived from the IMUs gait parameters focusing on the features most sensitive to PD gait namely, stride time, peak shank angular velocity, swing angular range, freeze index, arrhythmicity and gait asymmetry [19]. Following a similar logic to account for temporal context, change in lagged time values of stride time and peak angular velocity are included, which describe possible changes between consecutive steps. This refined feature set is therefore

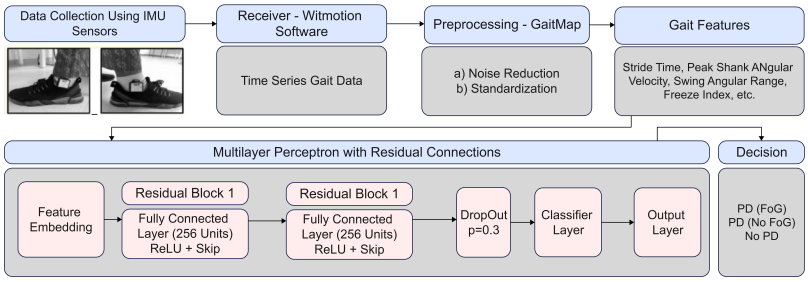


Figure 2: Process flow for gait test

extracted from each straight 10-meter segment from the FW task. With the help of such specific parameters, understanding of the deficit while walking in PD, especially when the walking occurs in structured efficient forward-straight trajectories, is expected to improve. This makes it possible to do comparisons within and between PD and control subject groups.

3.2.4 Classification and analysis

The MLP with residual connections is employed to learn useful gait features by adding residual connections between fully connected layers, allowing effective feature learning and enhancing model performance in classifying the gait abnormalities in Parkinson's Disease. The architecture, shown in Fig. 2, achieves a good balance between computational efficiency and classification accuracy. The MLP architecture consists of the following elements: (i) Input Layer: A fully connected (FC) layer maps the gait feature vectors to a higher-dimensional space (256 units) to enable richer feature representations. (ii) Residual Connections: First Residual Connection: A residual connection from the first MLP layer to its input enables better gradient flow, enhancing training stability, and Second Residual Connection: Residual connection from the second layer of the MLP to its input further augments feature enhancement while passing in forward direction. (iii) Regularization: Regularization techniques include dropout (30%) and batch normalization to suppress overfitting as well as train stability. (iv) Classification Layer: Output FC layer followed by softmax activation calculates probability distribution over classes of gait severity.

The selection of MLP is motivated by the necessity to capture low-level and high-level dependencies in the gait data. The residual connections enhance the

capacity of the model to learn intricate patterns, particularly under noisy conditions, by facilitating effective gradient propagation during backpropagation.

The residual connections model explicitly: (i) Feature refinement: Enabling improved gradient flow and enabling the model to learn more elaborate representations at every layer, and (ii) Training stability: Improving convergence and minimizing the risk of vanishing gradients. In addition to enhancing model performance, batch normalization and dropout are combined, employing an adaptive learning rate strategy via the ReduceLROnPlateau scheduler. Class imbalance is handled with the Synthetic Minority Over-sampling Technique (SMOTE), which produces synthetic samples for the minority classes, encouraging a better-balanced model.

4 Results and Analysis

4.1 Results of the spiral test using swin transformer

4.1.1 t-SNE visualization and analysis of feature embeddings

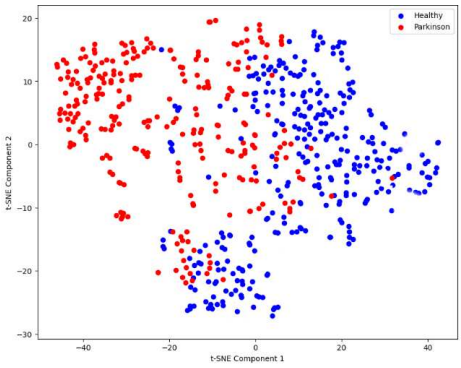


Figure 3: t-SNE plot of Swin transformer embeddings of spiral drawings

Fig. 3 shows a t-SNE plot with the fine-tuned Swin Transformer model embeddings to visually observe the discriminative separation between Parkinson's and normal samples. Distinct clusters are seen, where samples belonging to the healthy class (augmented and original) are well separated from the images of the Parkinson's patient. This evident clustering indicates that the use of the Swin Transformer model is successful in capturing the intraclass similarities and interclass dissimilarities in every cluster, most likely recognizing the tremor-related anomalies in the Parkinson's samples absent from the healthy drawings.

Furthermore, embedded augmentations tend to enhance clustering, which, in turn, imparts an additional enhanced discrimination of the information in regard to patterns of motor deficiency. Greater separability by feature through augmented embeddings signifies that the model is sensitive to PD-characteristic

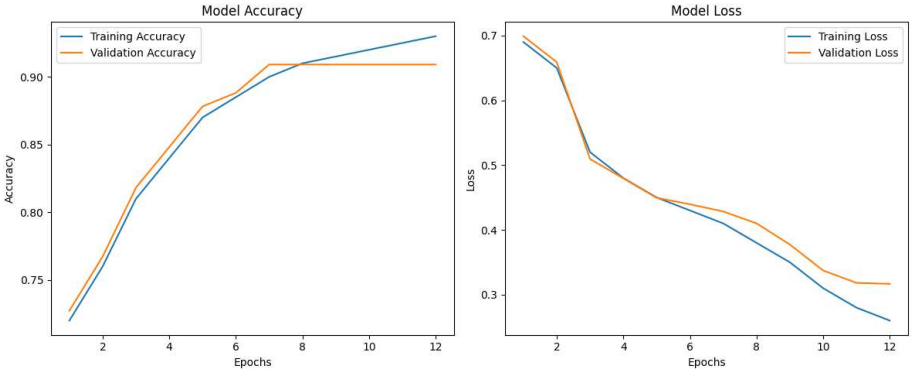


Figure 4: Training history of the Swin Transformer model (a) accuracy (b) loss

visual features, thus potentially correctly classifying and supporting diagnosis by supporting decision making while performing diagnostics.

4.1.2 Training

The Swin Transformer-based model is trained and tested extensively to categorize Parkinson's Disease (PD) using spiral drawing samples. Fig. 4 illustrates the training and validation loss curves, showcasing the model's learning process. Throughout training, the performance of the model indicates a consistent increase from epoch to epoch, as indicated by decreases in both training and validation losses and improvements in accuracy.

In the first epoch, the model achieved a validation loss of 0.6990 and an accuracy of 72.73%. With each subsequent epoch, validation loss decreases significantly and accuracy improves, reaching a plateau around the fifth epoch. By the seventh epoch, the training loss stabilized at 0.2942, and the validation loss dropped further to 0.3182 with a consistent accuracy of 90.91%. The high accuracy and low validation loss signify that the model has learnt discriminative characteristics well between healthy and PD classes with no signs of overfitting.

Table 1: Response time of various machine learning models for spiral test (in ms)

Model	Ryzen 7 5825U CPU
Random Forest	3.215 ms
XGBoost	2.487 ms
Swin Transformer	1.032 ms

From the Table 1 it is observed that all models have similar response times on the Ryzen 7 5825U processor. While the computation times are comparable, the Swin Transformer model had better analytical capability in spite of

its complexity. In contrast to existing machine learning models like Random Forest and XGBoost that depend on pre-designed feature extraction, the Swin Transformer is good at extracting subtle changes in spiral drawing patterns and is thus well-suited for differentiating Parkinson's Disease (PD) from normal samples. Experiments are run on general CPU platforms, confirming that our proposed method is computationally effective and practical for real-time use.

4.1.3 Performance Analysis and Results

Our comparison of three machine learning models for the detection of Parkinson's based on spiral drawings provides insights into their diagnostic performance. The Swin Transformer demonstrated remarkable superiority, achieving 90.91% accuracy in 10-fold cross-validation, significantly outperforming traditional machine learning approaches such as Random Forest (76.07%) and XGBoost (77.86%).

Beyond accuracy, the Swin Transformer consistently achieves over 89% in precision, recall, and F1-score. This indicates that the model does capture fine movement anomalies in the spiral test accurately, which is vital for the early diagnosis.

Table 2: Performance(%) Comparison of models for hand tremor based PD detection using a 10-Fold cross validation

Models	Accuracy	Specificity	Precision	Recall	F1 Score
Random Forest	76.07	86.67 ± 8.33	84.17 ± 16.82	74.17 ± 25.12	74.58 ± 16.59
XGBoost	77.86	73.33 ± 10.00	83.33 ± 13.94	76.67 ± 21.02	77.19 ± 11.46
Swin Transformer	90.91	90.84 ± 4.56	92.51 ± 4.67	89.05 ± 5.57	90.27 ± 4.84

The confusion matrix also confirms these results, indicating that the model accurately classifies healthy people and Parkinson's patients with very few misclassifications. That is, there are hardly any false positives (healthy people incorrectly classified as Parkinson's) and false negatives (Parkinson's cases missed). This low error rate supports the model's high sensitivity and specificity.

4.2 Results of the gait test using MLP architecture

4.2.1 Training and Hyperparameters

In this study, MLP (Multilayer Perceptron) with residual connections is chosen as the main classification model because it can capture intricate feature interactions and avoids the vanishing gradient issue. Because of the structured nature of the data set, MLP is hypothesized to surpass conventional architectures at modeling fine interdependencies that other models could overlook.

To guarantee strict and impartial assessment, stratified 10-fold cross-validation is performed. This method maintains the initial class distribution per fold so that consistent performance could be evaluated. To extend the solution against

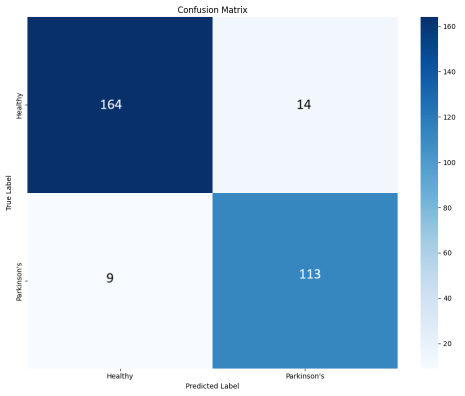


Figure 5: Confusion matrix (Swin transformer)

class imbalance, Synthetic Minority Over-sampling Technique (SMOTE) is implemented only on the training subsets of each fold so that no information would leak into the validation set.

The residual connected MLP is trained with a hidden size of 256, dropout of 0.3, and learning rate of 1×10^{-3} with the Adam optimizer. The batch size is 128, with early stopping (20 epochs) on validation loss and a limit of 300 training epochs.

For comparison, other models are tested with hyperparameters optimized for their architectures. Random Forest is set up with 500 trees, depth of 10, and class-balanced weighting. Support Vector Machine (SVM) uses an RBF kernel with $C = 5.0$, and probability calibration is done through three-fold cross-validation. LightGBM is tuned with 1000 estimators, learning rate 0.005, and regularization parameters $\alpha = 0.2$ and $\lambda = 0.2$. The MLP baseline has three hidden layers (512, 256, and 128 neurons) with ReLU activations, adaptive learning rate scheduling, and residual connections to enhance the flow of gradients during training.

4.2.2 Performance Analysis and Results

Table 3: Performance (%) of gait based models (10-Fold cross-validation results)

Model	Accuracy	F1 Score
Random Forest	0.8286 ± 0.0197	0.8193 ± 0.0208
SVM	0.6264 ± 0.0406	0.6165 ± 0.0399
LightGBM	0.7884 ± 0.0402	0.7747 ± 0.0448
MLP	0.7927 ± 0.0297	0.7796 ± 0.0327
ResMLP	0.8329 ± 0.0183	0.8198 ± 0.0200

Table 3 presents the performance metrics from 10-fold stratified cross-validation

for all models. In general, the MLP with residual connections always has the highest mean accuracy (0.8329 ± 0.0183) and macro F1 score (0.8198 ± 0.0200) among the compared methods. This robust performance indicates the power of MLP for structured data classification problems, especially in learning higher-order feature interactions and mitigating possible training challenges with residual connections.

Boosting-based techniques like LightGBM also run competitively, indicating their strength in capturing non-linear interactions between features. Their performance still falls slightly short of MLP, probably because there are no inherent mechanisms for dynamically ordering feature importances for differing samples. The Random Forest exhibits stable baseline performance, exemplifying the efficacy of ensemble tree approaches to tabular data, whereas the fully connected MLP with no residual connections is not able to come close to the performance of MLP with residual connections, indicating that general dense architectures in the absence of inductive biases might not be the best solution for this task. SVM had the poorest performance on the majority of metrics, showing the weakness of SVM in dealing with intricate, multi-class structured data, despite calibrated probabilistic outputs.

Following training for the above models, residual-connected MLP is trained on the full training data and tested on a held-out test set that is not used at any point during training, validation, or hyperparameter search. The related normalized confusion matrix is visualized in Fig. 6.

The confusion matrix in Fig. 6 indicates the MLP's robust generalization performance on all three target classes: No PD (Class 0), PD with Freezing of Gait (FoG) (Class 1), and PD without FoG (Class 2). The model performs with class-wise accuracies of 78.71% for Class 0, 87.74% for Class 1, and 80.89% for Class 2, reflecting a uniform capability of separating healthy subjects and Parkinson's disease patients with or without FoG.

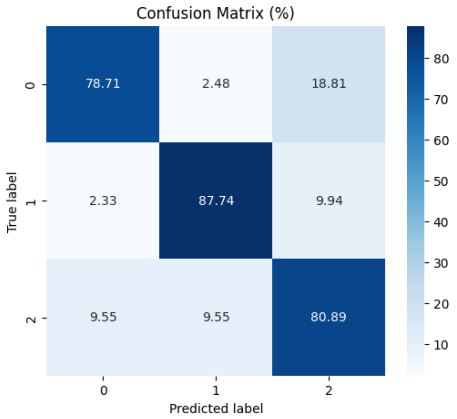


Figure 6: Normalized confusion matrix (%) of MLP with residual connections on the test set.

The model specifically performs well at detecting PD with FoG, with the best class-wise accuracy of 87.74%, which is in line with its capacity to recognize more severe motor symptoms. It also had good accuracies in classifying No PD and PD without FoG, though there was some ambiguity between the two classes—namely, 18.81% of Class 0 samples were incorrectly classified as Class 2, and 9.55% of Class 2 samples were incorrectly classified as Class 0. This is not surprising due to similar movement patterns in early-stage PD and healthy controls.

Notably, the misclassification between Class 1 and the other classes is below 10%, affirming the robustness of the model to detect the FoG phenotype. The findings demonstrate that the MLP with residual connections not only outperforms O'Day et al.'s (2019) [18] baseline logistic regression model of 77% accuracy and 0.754 AUC, but also exhibits well-balanced performance in all classes.

In conclusion, the MLP with residual connections shows strong discrimination between disease subgroups and is a potential candidate for clinical instruments intended for automated PD diagnosis and monitoring disease progression.

4.2.3 Training

Table 4: Response time of various machine learning models for gait test (in ms)

Model	Training Time (ms)
Random Forest	9.492
SVM	9.613
LGBM	2.826
XGBoost	2.781
MLP	7.355
MLP with Residual Connections	1.721

A comparative analysis of the response times for various models is conducted to assess the computational efficiency of the models. The models are evaluated based on their training times, measured in milliseconds. The results are presented in Table 4, where the training times of the models are shown in comparison.

As shown in Table 4, MLP with residual connections exhibit the fastest training time of 1.721 ms, significantly outperforming the other models in terms of computational efficiency. In contrast, SVM demonstrates the longest training time, taking 9.613 ms. This highlights the efficiency of this particular MLP model, making it a favorable choice for applications that require both high performance and low computational overhead.

To further evaluate the performance of the MLP model, we employed a 10-fold stratified cross-validation (CV) approach is used here. This methodology ensures a comprehensive and unbiased assessment of the model's generalizability by partitioning the data into ten subsets, each serving as both a training and

validation set in turn. The final results, averaged across all ten folds, yields a mean accuracy of 0.8329 ± 0.0183 and a mean F1-score of 0.8198 ± 0.0200, underscoring the model's robustness and consistency.

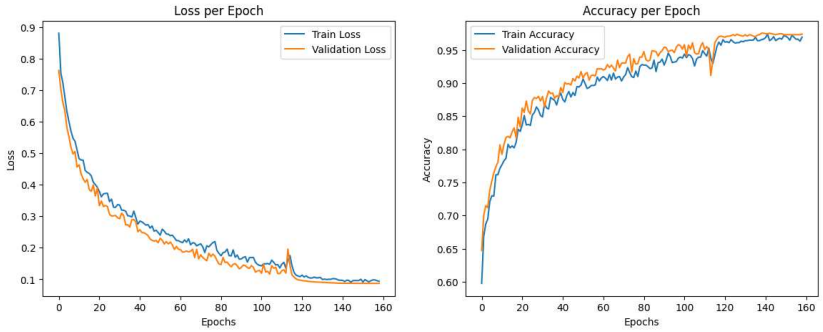


Figure 7: Training history of ResMLP model: (a) loss; (b) accuracy

Fig. 7 illustrates the training history of the MLP, with steadily improving performance over 200 epochs, triggering early stopping around 160. Both the training and validation loss curves are smooth and steadily decreasing, to values near 0.1, which indicates effective and persistent learning. The accuracy curves exhibit a sharp upward trend, with the validation accuracy very close to the training accuracy for the entire training period. This close proximity is an indication of very good generalization to novel data and well-calibrated learning.

Minor oscillations near epoch 120 are transient and rapidly rectified, highlighting the robustness of the model and the power of regularization and learning rate scheduling. In summary, the MLP with residual connections not only converges very fast but do so with a good stability and generalization performance, highlighting its high potential for clinical use in Parkinson's disease classification.

5 Discussion

5.1 Archimedes' Spiral Test

This study effectively utilized the Swin Transformer-based model to diagnose Parkinson's Disease (PD) and assess tremor severity using spiral drawing data from the Kaggle PD drawing dataset. A classification accuracy of 90.91% is achieved by fine-tuning the Swin Transformer to differentiate healthy individuals from PD patients based on their drawing features. This result illustrates the potential of deep learning for monitoring fine motor signs and identifying subtle motor deficits based on visual information. The strength of the Swin Transformer model is that it can learn spatial hierarchies and complex drawing patterns using a multi-head self-attention mechanism to attend to important tremor features, including line irregularities and disruptions in curvature. The

recording setup involved participants tracing a clockwise Archimedes spiral from inwards on a digital tablet, which recorded their movements, allowing for detailed analysis of tremor amplitude and stability during tracing. This method captures essential coordination and motor stability data crucial for evaluating action tremors [21].

Digital analysis of spiral drawings for PD has a number of clinical benefits, such as quantifiable and objective measures of the severity of tremor, potentially leading to standardised diagnostic procedures. There are limitations; our data is limited to spiral drawings and may not capture other aspects of motor features. In addition, though the Swin Transformer showed superior accuracy, few misclassifications between PD and healthy samples indicate overlap in fine drawing patterns, particularly in initial-stage PD. A more extensive and heterogeneous dataset, as well as longitudinal data, would improve the model's generalizability across various PD subtypes.

These results highlight the potential of Transformer-based models for PD fine motor symptom monitoring, setting the stage for further studies on multimodal evaluation that incorporates both gait and drawing information. This research is a step closer to creating automated, non-invasive diagnostic methods for objective monitoring of PD and individualized intervention planning.

5.2 Gait Analysis

This work shows the efficient application of MLP in classifying subjects with Parkinson's disease (with and without FOG) and healthy controls from wearable IMU sensor data. The achieved 83% accuracy is a great improvement in computerized gait analysis, especially in PD classification. The model's strong performance in recognizing healthy controls reflects its robust ability to discriminate between normal and pathological gait patterns, while its capacity to differentiate between PD patients with and without FOG suggests high sensitivity to subtle changes in gait dynamics that occur as the disease progresses.

The success of MLP with residual features can be attributed to several key factors. The architectural design of the model, integrating spatial and temporal feature extraction, is found to be crucial in the capture of intricate gait patterns found in various PD features. Optimization of the hyperparameters, such as learning rate and regularization methods, is instrumental in maintaining model stability and generalizability. In addition, the training dynamics, defined by constant increments in accuracy and loss measures without significant overfitting, confirm our architectural choice and training process.

Besides, for real-time patient assessment, we are developing our own dataset with WIT MOTION 9-axis IMU sensors [21] for training and validation. The sensors are mounted on both the patient's feet, and data is recorded as the patient walks along a 20m straight line. We preprocess the recorded data using the GaitMap library [22], a sophisticated Python package for gait analysis. GaitMap uses advanced signal processing algorithms to transform raw IMU measurements into the relevant gait parameters, which are then encoded to align with our training data's feature space. The processed data are then run through

the trained MLP model to produce instantaneous predictions with confidence scores, which could help detect PD symptoms early on during standard clinical examinations. This real-time application is a major step towards the incorporation of machine learning-based gait analysis in the clinical environment.

Some limitations need to be noted, however. While the sample size used in this study is adequate for proof-of-concept, it could potentially restrict the generalizability of the model to larger patient populations. In addition, the use of only shank-mounted sensors might overlook subtle upper body movements that may offer additional discriminative information.

These results have significant implications for both clinical practice and study design. The capacity to automatically classify PD status and FOG occurrence based on wearable sensors offers clinicians objective measures for disease monitoring and treatment assessment. This methodology could be especially useful for early identification of PD symptoms and maximizing therapeutic interventions.

6 Conclusion

This study demonstrates that the Swin Transformer is a strong performer for quantifying tremor severity from spiral drawing tasks. It also shows that MLP with residual connections can effectively classify gait impairments and detect freezing of gait in Parkinson's disease. The models, trained with Swin Transformers, recorded 90.91% accuracy in classifying Parkinson's disease based on tremor severity from the digital spiral drawing. Moreover, the MLP model, trained on IMU sensor data, recorded a mean accuracy of 83.29% in classifying the disease status and FOG detection. With their optimal hyperparameters, these models have uniform performances, and thus they could be of potential value for clinical use, especially for classifying controls and identifying symptoms of Parkinson's Disease (PD).

To aid the training and testing of these models, we are building a dataset with both digital spiral drawings and gait data. For the spiral aspect, subjects perform on-screen spiral drawing tests, which are processed to determine tremor-related features. For gait analysis, we use 9-axis IMU sensors mounted on both feet, and data are acquired during a 20-meter walking exercise and are processed with GaitMap to obtain clinically useful gait parameters. Features obtained from both modalities are coded to correspond with the respective input formats and transferred to the learned Swin Transformer and MLP models in order to carry out real-time predictions with appropriate confidence scores respectively. This system allows for the early identification of PD-associated motor symptoms on standard evaluation and is an advance toward the incorporation of multimodal machine learning technologies into clinical practice.

Our results add to the expanding evidence base for the application of wearable sensors and machine learning in neurological examination. Directions for future research include validation in larger populations, exploration of other sensor modalities, and integration with clinical decision support systems. The

establishment of real-time analysis capabilities and longitudinal monitoring solutions may further increase the clinical utility of this method. Eventually, these developments may result in enhanced early detection, monitoring, and treatment of gait disorders in Parkinson's disease, possibly translating into improved patient outcomes through more focused and timely interventions.

References

[1] A. Gironell, B. Pascual-Sedano, I. Aracil, J. Marín-Lahoz, J. Pagonabarraga, and J. Kulisevsky, "Tremor types in Parkinson disease: a descriptive study using a new classification," *Parkinson's Disease* **2018**(1) (2018) 4327597.

[2] A. Mirelman, P. Bonato, R. Camicioli, T. D. Ellis, N. Giladi, J. L. Hamilton, C. J. Hass, J. M. Hausdorff, E. Pelosin, and Q. J. Almeida, "Gait impairments in Parkinson's disease," *The Lancet Neurology* **18**(7) (2019) 697–708.

[3] J. G. Nutt, B. R. Bloem, N. Giladi, M. Hallett, F. B. Horak, and A. Nieuwboer, "Freezing of gait: moving forward on a mysterious clinical phenomenon," *The Lancet Neurology* **10**(8) (2011) 734–744.

[4] A. P. Legrand, E. Hainque, *et al.*, "New Insight in Spiral Drawing Analysis Methods - Application to Action Tremor Quantification," *Clinical Neurophysiology* (2017), doi: 10.1016/J.CLINPH.2017.07.002.

[5] P. Zham and D. Kumar, "Efficacy of Guided Spiral Drawing in the Classification of Parkinson's Disease," *IEEE Journal of Biomedical and Health Informatics* (2017), doi: 10.1109/JBHI.2017.2762008.

[6] J. Danna, S. Pinto, *et al.*, "Digitalized Spiral Drawing in Parkinson's Disease: A Tool for Evaluating Beyond the Written Trace," *Human Movement Science* **62** (2018) doi: 10.1016/J.HUMOV.2018.08.003.

[7] N. Wang, X. Niu, Y. Yuan, A. Zhao, *et al.*, "Coordinate Attention Enhanced Swin Transformer for Handwriting Recognition of Parkinson's Disease," *IET Image Processing* **17**(10) (2023) n/a-n/a, doi: 10.1049/ipr2.12820.

[8] N. Jahan and M. A. Layek, "Parkinson's Disease Detection Using CNN Architectures with Transfer Learning," in *International Conference on Signals and Electronic Systems* (2021), doi: 10.1109/ICSES52305.2021.9633872.

[9] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern Recognition Letters* **119** (2019) 3–11.

[10] S. Pardoel, J. Kofman, J. Nantel, and E. D. Lemaire, "Wearable-sensor-based detection and prediction of freezing of gait in Parkinson's disease: A review," *Sensors* **19**(23) (2019) 5141.

[11] L. Sigcha, L. Borzì, F. Amato, I. Rechichi, C. Ramos-Romero, A. Cárdenas, L. Gascó, and G. Olmo, "Deep learning and wearable sensors for the diagnosis and monitoring of Parkinson's disease: A systematic review," *Expert Systems with Applications* **229** (2023) 120541.

[12] Y. El Marhraoui, H. Amroun, M. Boukallel, M. Anastassova, S. Lamy, S. Bouilland, and M. Ammi, "Foot-to-ground phases detection: A comparison of data representation formatting methods with respect to adaption of deep learning architectures," *Computers* **11**(5) (2022) 58.

[13] L. Borzì, M. Varrecchia, G. Olmo, C.A. Artusi, M. Fabbri, M.G. Rizzone, A. Romagnolo, M. Zibetti, and L. Lopiano, "Home monitoring of motor fluctuations in Parkinson's disease patients," *Journal of Reliable Intelligent Environments* **5**(3) (2019) 145–162.

[14] E. Jovanov, E. Wang, L. Verhagen, M. Fredrickson, and R. Fratangelo, "deFOG—A real-time system for detection and unfreezing of gait of Parkinson's patients," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Sep. 2009, pp. 5151–5154.

[15] P. K. Yang, B. Filtjens, P. Ginis, M. Goris, A. Nieuwboer, M. Gilat, P. Slaets, and B. Vanrumste, "Freezing of gait assessment with inertial measurement units and deep learning: Effect of tasks, medication states, and stops," *Journal of NeuroEngineering and Rehabilitation* **21**(1) (2024) 24.

[16] K. Mader, *Parkinson's Disease Drawing Dataset*, Kaggle. Available: `https://www.kaggle.com/datasets/kmader/parkinsons-drawings/data`.

[17] Z. Liu, *et al.*, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021).

[18] J. O'Day, J. Syrkin-Nikolau, C. Anidi, L. Kidzinski, S. Delp, and H. Bronte-Stewart, "The turning and barrier course reveals gait parameters for detecting freezing of gait and measuring the efficacy of deep brain stimulation," *PLoS One* **15**(4) (2020) e0231984.

[19] S. Lord, B. Galna, and L. Rochester, "Moving forward on gait measurement: Toward a more refined approach," *Movement Disorders* **28** (2013) 1534–1543, `https://doi.org/10.1002/mds.25545`.

[20] M. Sarshar, S. Polturi, and L. Schega, "Gait phase estimation by using LSTM in IMU-based gait analysis-Proof of concept". Sensors 21(17) (2021) 5749

[21] P. Zham, D. Kumar, P. Dabnichki, S. P. Arjunan, S. Raghav, "Distinguishing Different Stages of Parkinson's Disease Using Composite Index of Speed and Pen-Pressure of Sketching a Spiral," *Frontiers in Neurology* **8** (2017).

[22] Witmotion IMU Sensor. Available: https://www.wit-motion.com/.

[23] F. Bonnet *et al.*, "GaitMap: An Open-Source Python Package for Gait Analysis Using Inertial Sensors," *Journal of Open Source Software* **8**(87) (2023) 5407, https://doi.org/10.21105/joss.05407.