Journal of Biostatistics and Epidemiology

J Biostat Epidemiol. 2024;10(3): 327-341

Original Article

Analysis of Vertical Ground Reaction Force Data in Predicting Parkinson's Disease

Varun Jain*

Faculty of Health Sciences, McMaster University, Hamilton, Canada.

ARTICLE INFO

Received 19.06.2024
Revised 13.07.2024
Accepted 11.08.2024
Published 09.09.2024

Key words:
Fourier analysis;
Machine learning;
Support vector machine (SVM);
Frequency analysis;
Power spectrum analysis;
Biomedical signals

ABSTRACT

Introduction: Parkinson's disease is a complex, progressive neurodegenerative disorder known to negatively impair patient gait. Therefore, with gait and vertical ground reaction force data, an association can be made between the data and Parkinson's disease.

Methods: Data from 146 participants; 93 with Parkinson's disease and 73 without Parkinson's disease was obtained from a PhysioNet database for use in this article. A Fourier Analysis and several support vector machine learning models were computed in MATLAB to classify whether an individual had Parkinson's disease.

Results: From the Fourier analysis, it was determined that a statistically significant difference was present between the vertical ground reaction force data of individuals with and without Parkinson's disease. Additionally, it was found that a Minimum Classification Error Optimized SVM machine learning model using Bayesian statistics was able to classify individuals with Parkinson's disease using vertical ground reaction force data at an accuracy of 67.1%, and sensitivity of 80.43%.

Conclusion: Therefore, it can be determined that vertical ground reaction force can predict Parkinson's Disease with considerable accuracy which could be improved with an increased number of participants.

Introduction

Parkinson's Disease

Parkinson's disease (PD) is a complex, progressive neurodegenerative disorder known to impact movement in patients.[1] PD patients usually experience a myriad of symp-toms that change over time, including tremors, postural instability, ataxia, disequilibri-um, and bradykinesia.[1-3] PD is one of the most common neurological disorders globally, with an estimated prevalence of 1% of the population over 60 years old, and it is expected to increase as the global population ages.[1] According to the World Health Organization, around 10 million people worldwide are currently living with PD, with significant numbers in both developed and developing countries.[2] In many re-gions, the disease burden is expected to

*.Corresponding Author: Jainv8@mcmaster.ca

Copyright © 2024 Tehran University of Medical Sciences. Published by Tehran University of Medical Sciences.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license (https://creativecommons.org/licenses/by-nc/4.0/). Noncommercial uses of the work are permitted, provided the original work is properly cited.

rise by more than 50% in the next two decades due to demographic shifts.[1, 3] Alongside the characteristic movement impair-ments, PD patients often experience associated non-motor symptoms related to depression, psychosis, behavioural changes, and cognitive impairments.[2-3] One prevalent non-motor cognitive complication is Parkinson's disease dementia; a significant decline in cognitive ability, specifically memory and thinking, in individuals living with PD.[2-3] In fact, the Movement Disorder Society (MDS) estimates that 30% to 40% of PD patients live with Parkinson's disease dementia, and 83% of patients with PD will ultimately de-velop dementia within their lifetime.[2]

PD arises primarily due to the reduction of neurons that produce dopamine (dopaminergic neurons) within the substantia nigra; a basal ganglia structure within the midbrain.[1-2] Dopamine is an important neurotransmitter that is involved in initiating and inhibiting movement; low levels of dopamine can cause abnormal brain activity and ultimately lead to impaired movement.[2] There is no exact cause for the reduced number of dopaminergic neurons, though both genetic and environmental factors have been suggested as potential contributors.[3] Genetic mutations and environmen-tal triggers such as toxin exposure are potentially involved, however, in the majority of PD patients, the cause is unknown.[1]

Gait Disorders in Parkinson's Disease

Gait impairment manifests in the majority of PD patients, eventually progressing to a total loss of mobility.[4] Patients with PD tend to have altered cadence and stride length when walking, in comparison to healthy subjects, resulting in rhythmic inconsistencies.[4] Initially, patients usually present with rigid, slow & small-stepped walking patterns, commonly referred to as freezing & shuffling gait.[4-5] Many patients also experience festinations, which lead to a very inefficient gait rhythm.[4] Due to their altered gait patterns, PD patients tend to have diminished multi-tasking abilities, specifically be-cause their lack of attention and disequilibrium can provoke falls.[4] These disturb-ances can significantly contribute towards disability, cognitive impairments, psychoso-cial distress, and ultimately a lower quality of life, especially as the disease progresses.[4]

Importance of Models

Biomedical models, particularly those utilizing advanced data analysis methods, offer a powerful approach to understanding and predicting the progression of PD. In recent years, the use of biomechanical data, such as vertical ground reaction forces (VGRF), has gained attention for its potential to predict disease severity and progression.[5] VGRF data can provide valuable insights into the patient's gait dynamics, offering a non-invasive and quantitative method for assessing motor function.[5] However, while VGRF data has been utilized in various clinical settings, its full potential in modeling PD remains underexplored.[5]

This study aims to address this gap by using computational modelling that incorporates VGRF data to predict whether an individual has PD. By using mathematics and machine learning techniques, this research seeks to identify key biomarkers in gait that correlate with PD, offering a novel approach

to early detection and monitoring. The innovation of this work lies in the integration of biomechanical data with predictive modeling, potentially leading to more accurate and individualized diagnostic tools for PD, which could significantly improve patient outcomes and management strategies.

Methods

Dataset

The PhysioNet database was used within this analysis to obtain gait measurements using VGRF (Vertical Ground Reaction Force) from healthy individuals and individuals with PD. Data was collected from 93 patients with idiopathic PD, and 73 healthy subjects which are to be used as controls.[6] VGRF was measured using 16 load sensors (8 on each foot) placed at the points seen in Figure 1.[7] These sensors were sampled at 100 Hz and quantified the force applied (Newtons) by the pa-tient while walking.[6] Trials were approximately 2 minutes in length and took place on level ground in order to reduce external influences on gait rhythmicity.[6] With the trials, there were no set requirements re-garding the distance travelled, nor were these parameters provided. This may influence the gait-specific spatiotemporal features that could be ob-tained from the VGRF data.

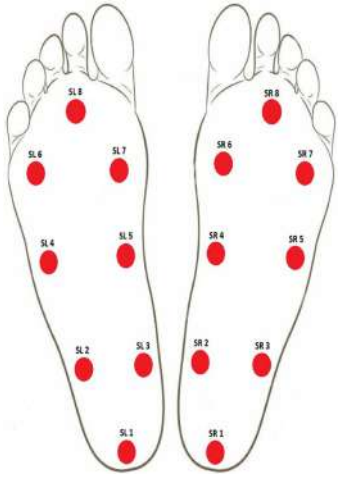

Figure 1. Location of VGRF sensors under each foot (7)

This dataset is valid for helping to treat the disease because it quantifies the difference in rhythmic gait patterns between PD patients and healthy controls. This allows for the study of stride dynamics and variability between groups under usual walking conditions. Additionally, variability in gait patterns can be used to measure disease severity, and help clinicians determine the best route of action when it comes to medications and surgical therapies.

Data was collected from 93 patients with PD, and 73 healthy controls via load sensors which output the VGRF at an instantaneous point in time. The participants performed gait exercises for approximately two minutes. However, upon cleaning the dataset, it was noted that there were encoding issues with two of the participant data files (1 PD patient data file and 1 control data file). Therefore, these files were excluded from the dataset used within this analysis. Thus, the resultant dataset had 92 patients with VGRF data representative of Parkinsonian gait and 72 individuals with data representative of normal gait.

Using this dataset, a Fourier Analysis was performed such that the frequency components within the VGRF signals can be analyzed for both PD patients and control subjects. In addition to this, a Support Vector Machine model was developed and trained such that VGRF data can be used to determine whether an individual has Parkinson's Disease.

VGRF data in itself is low frequency data, with 15 Hz being the maximum frequency used in

research.[8] However, most research studies analyze between 0.5 to 8 Hz as this passband encompasses all of the VGRF data typically seen in healthy and PD gait trials.[9] Based on this information, a passband of 0.5 Hz to 15 Hz was chosen for this analysis. This frequency range removes DC offset (0 Hz), Power Line Frequency (60 Hz), and other high frequency noise. Using a Butterworth Bandpass filter, the VGRF signal was filtered. Using the filtered VGRF data, a Fourier Analysis and a Machine Learning Model were both used to model the data.

Fourier Analysis Model Generation

The Fourier Analysis was used to analyze the frequency components and the power within the VGRF signals for each of the load sensors for both the control subjects and patients with PD. With this analysis, a set of 36 control subjects and 46 PD patients were used.

From literature it has been observed that the rigidity and "freezing" typically seen in Parkinsonian gait results in high frequency components (3 to 8 Hz) within the signal (10). These components were reported to be primarily apparent when participants were standing in place involuntarily, a common symptom seen in PD patients termed freezing of gait (FOG).[10] To identify FOG using the power spectra of VGRF data, the VGRF signal was characterized into two bands: Locomotor band & Freeze band.[10] The locomotor band consists of low frequency signals between 0.5 to 3 Hz and is representative of the frequency seen in normal gait.[10] The freeze band contains high frequency signals between 3 to 8 Hz and is representative of freezing of gait.[10]

The resulting Fourier analysis was performed in MATLAB via a fast fourier transform (fft) to generate the amplitude spectrum of the VGRF signal. The following equation was used with the amplitude spectra to determine the Power Spectral Density estimates of the signal.

$$PSD(\omega) = \frac{\left|\left(Y(\omega)\right)\right|^2}{L}$$

This equation uses the theory behind a periodogram of an array of length L to determine the power spectral estimates. This equation re-sulted in the generation of a power spectral sequence with a frequency range fr om -25 to 25 Hz. The power spectral sequences were aver-aged so that trends and observations could be visually made between the conditions. It is expected that patients with Parkinson's Disease will exhibit a peak in the power spectrum at a higher frequency when com-pared to the healthy subjects, primarily due to the high frequency com-ponents arising from characteristic Parkinsonian gait.

In order to further investigate the differences between the low fre-quency and high frequency bands, the band power in the locomotor (0.5 to 3 Hz) and the freeze band (3 to 8 Hz) was calculated. The band power for both of these bands were calculated for each of the 16 sensors for both the control and PD group. It is expected that patients with Parkin-son's disease would exhibit a larger band power at high frequencies compared to the control group based on their freezing of gait.

Machine Learning Model Generation

Given that specific categories for the classification algorithm were al-ready established (PD or Healthy), a supervised learning model was ap-propriate for this

analysis.

Support Vector Machines

Support Vector Machines (SVMs) are a supervised machine learning al-gorithm mainly used for classification. An SVM essentially represents different classifications in a multidimensional space, where the object is placed within decision boundaries based on how the model is trained. SVMs use an iterative approach so that error is minimized, maximizing accuracy.

For this analysis, an SVM was used to predict whether an individual has PD based on their VGRF data. To improve the accuracy of the model, it was trained and tested on two different datasets. Both da-tasets contain VGRF data for 46 patients with PD, and 36 healthy con-trols.

The data from each participant was iterated through to determine the average swing time for each foot. Swing time is defined as the time in which the analyzed foot is not in contact with the ground during gait tri-als. Swing time is an important spatiotemporal feature within gait studies, which have been previously included as a feature in some machine learning models when predicting the incidence of PD.[7] Within this analysis, the duration of all the swing times for a specific participant for a specific sensor were taken and averaged to determine the average swing time during a trial.

Following the swing time calculation, the low and high frequency band power reported within the VGRF signal was calculated from each individual power spectrum.

Both the average swing time for an individual, along with their low and high frequency band power were used as features within the ML model. The model compared its predictions to the initial classification la-bels during the training and testing phase, and optimized its algorithm to decrease the classification error.

Proposed Models

The training and testing datasets were used to develop and validate three different SVM models, each model was developed using different parameters, and different optimization strategies.

The first SVM model was a Quadratic SVM (Quadratic Kernel) as it resulted in the highest accuracy and minimum classification error com-pared to a Linear SVM (Linear Kernel), Cubic SVM (Cubic Kernel), Fine Gaussian SVM, Medium Gaussian SVM, & Coarse Gaussian SVM (Gaussian / RBF Kernel with coarse distinctions).

The second model used Bayesian optimization to minimize the clas-sification error between predicted and actual results. Specifically, Bayesian optimization is used to set the hyperparameters including the kernel function, scale, and box constraint level. To do this, an expected improvement per second plus acquisition function was used to evalu-ate the ”goodness” of a point based on its classification over the course of 30 iterations. The model was configured based on the hyperparame-ters that result in the least amount of classification error.

The third model uses the framework developed by the second model, except that it included a Principal Component Analysis (PCA).

All three of these models were initially trained on the training dataset and were then validated using the testing dataset. These models output their respective confusion matrices and ROC curves, which were ana-lyzed to determine

their effectiveness.

Fourier Model Analysis

For the Fourier Analysis, an initial Jarque-Bera test was performed in MATLAB which determined that the average power spectrum data for the control and PD subjects did not come from a normal distribution. Table 1 displays the results, with 0 representing normality and 1 representing that the data is not normal.

Given that the power spectral data is not normal, a Mann-Whitney U test was performed to determine whether there were significant differences between the control and PD subjects. Additionally, an averaged power spectral sequence was created for each of the 16 VGRF sensors for both control and PD subjects. For the Mann-Whitney U test, the following assumptions were made:

- The VGRF data is independent and continuous
- The participants used to collect VGRF data from were randomly selected from a stratified population
- The data is not normal

Machine Learning Model Analysis

For the Machine Learning Models, two datasets were created as previously mentioned: Training Dataset (82 individuals, 36 Control and 46 PD) & Testing Dataset (82 in-dividuals, 36 Control and 46 PD). Both datasets consisted of numerous individuals with a stratified division between their respective classes based on the data obtained from the PhysioNet database. The training dataset was used to train the model and allow it to optimize its parameters, while the testing dataset was used to validate the model's efficacy.

Each of the three models were trained and tested separately, outputting a confusion matrix, and an ROC curve. These served as the statistical analysis of the model, and it provided information on the model's ability to differentiate between control and PD subjects. It is important to note that a diagnosis of PD was set to as positive for the confusion matrix and the ROC curve.

The confusion matrix was used to determine the sensitivity and specificity of the model, while the ROC curve was used to evaluate the performance of the classifier.

Results

Fourier Analysis Results

After following the aforementioned Fourier Analysis methodology, the average power spectral sequences were plotted against frequency and it was visually determined that the PD group seems to exhibit increased power at a higher frequency compared to the control

Table 1. Jarque-Bera Test Results for Averaged Power Spectral Data

Group	VGRF 1	VGRF 2	VGRF 3	VGRF 4	VGRF 5	VGRF 6	VGRF 7	VGRF 8
Control Left Foot	1	1	1	1	1	1	1	1
Control Right Foot	1	1	1	1	1	1	1	1
PD Left Foot	1	1	1	1	1	1	1	1
PD Right Foot	1	1	1	1	1	1	1	1

group.

In order to investigate the higher frequency differences further, a bar plot depicting the band power in the locomotor (0.5 to 3 Hz) and the freeze band (3 to 8 Hz) was calculated as seen in Figure 2a & 2b. From these figures, it can be determined that there is a significant amount of variability for the magnitude of the low frequency components between VGRF sensors in control and PD participants.

In order to test whether differences exist between the power spectral data for each sensor, a Mann-Whitney U test was performed. The p-values for each of the comparisons done by the Mann-Whitney U test can be seen in Table 2.

From the Mann-Whitney U test, it can be concluded that there was a statistically significant difference between the populations for most of the sensors. However, it appears that no significant difference was ob-served for VGRF 3 on the Left Foot and VGRF 8 on both feet.

Machine Learning Results

With respect to the SVM models, the confusion matrices and ROC curves for the training and testing phase can be found in Figures 3 to 8. From the ROC curves, it can be determined that the performance of the models in classifying PD using VGRF data was fairly similar. They all had an AUC of approximately 0.71, which indicates that the model is fair in classifying VGRF data, but its performance can be improved. A summary of the SVM model results can be found in Table 3.

From the table, it can be determined that

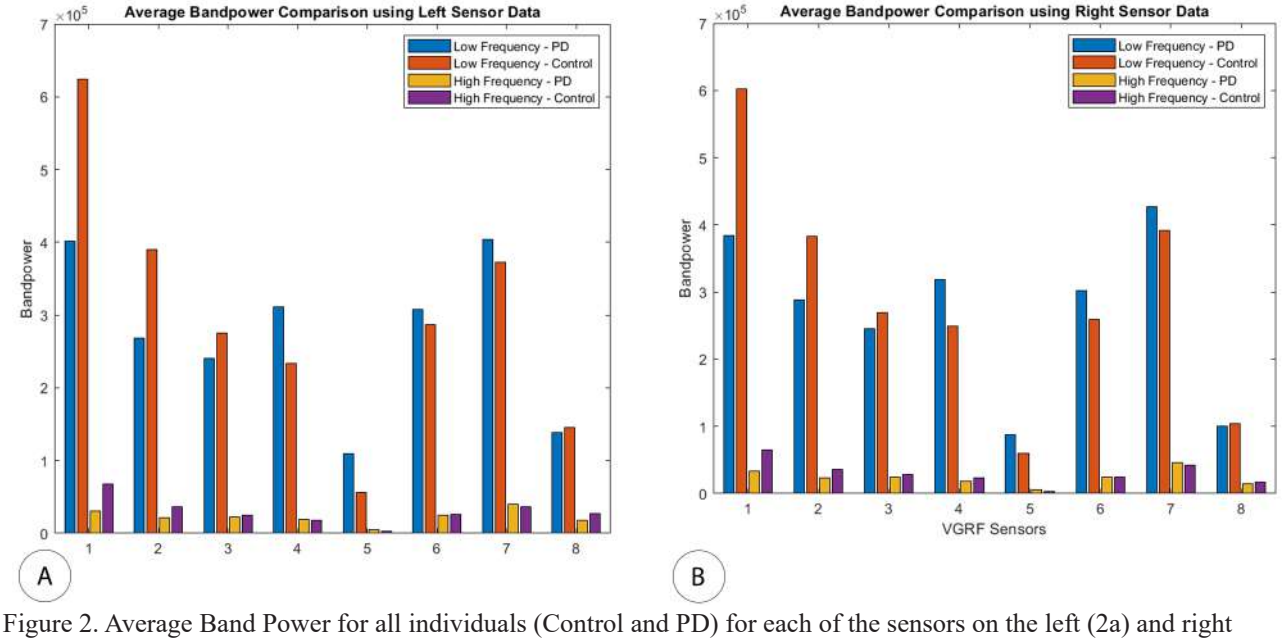

Figure 2. Average Band Power for all individuals (Control and PD) for each of the sensors on the left (2a) and right foot (2b) separated by high and low frequency components.

A) Average Band Power for all individuals for each of the sensors on the left foot separated by high and low frequency components.

B) Average Band Power for all individuals for each of the sensors on the right foot separated by high and low frequency components.

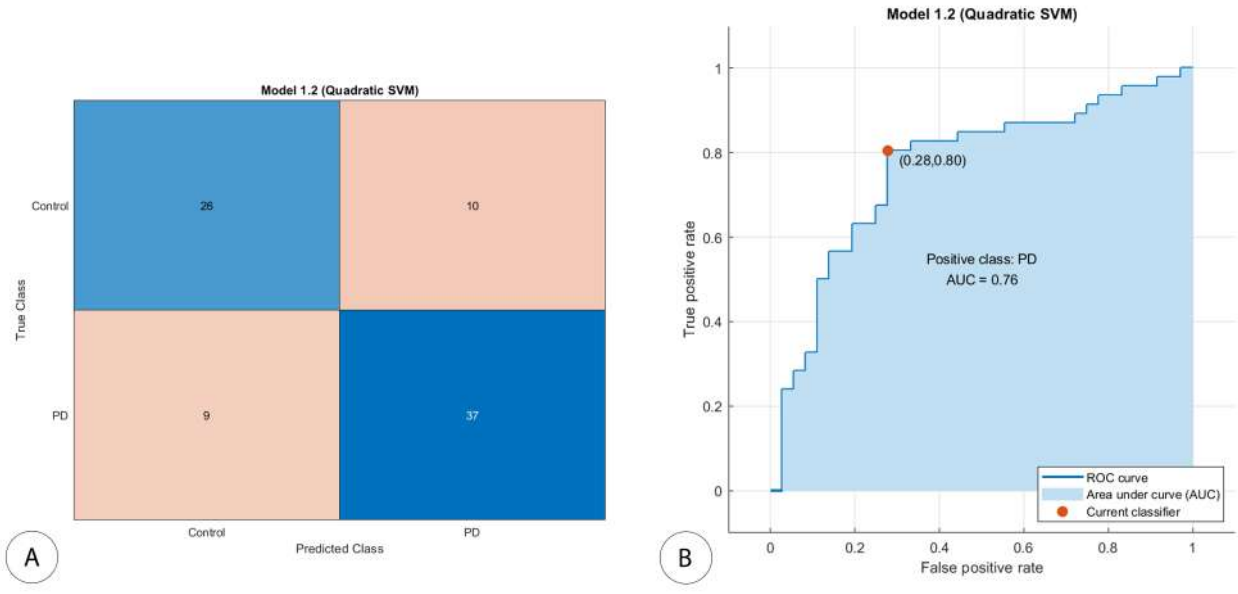

Figure 3. Results of Machine Learning Model 1 against the training dataset, including the Confusion Matrix (3a) and the ROC curve (3b)
A) Confusion Matrix of Machine Learning Model 1 with the Training Dataset
B) ROC Curve including the calculated AUC using Machine Learning Model 1 with the Training Dataset

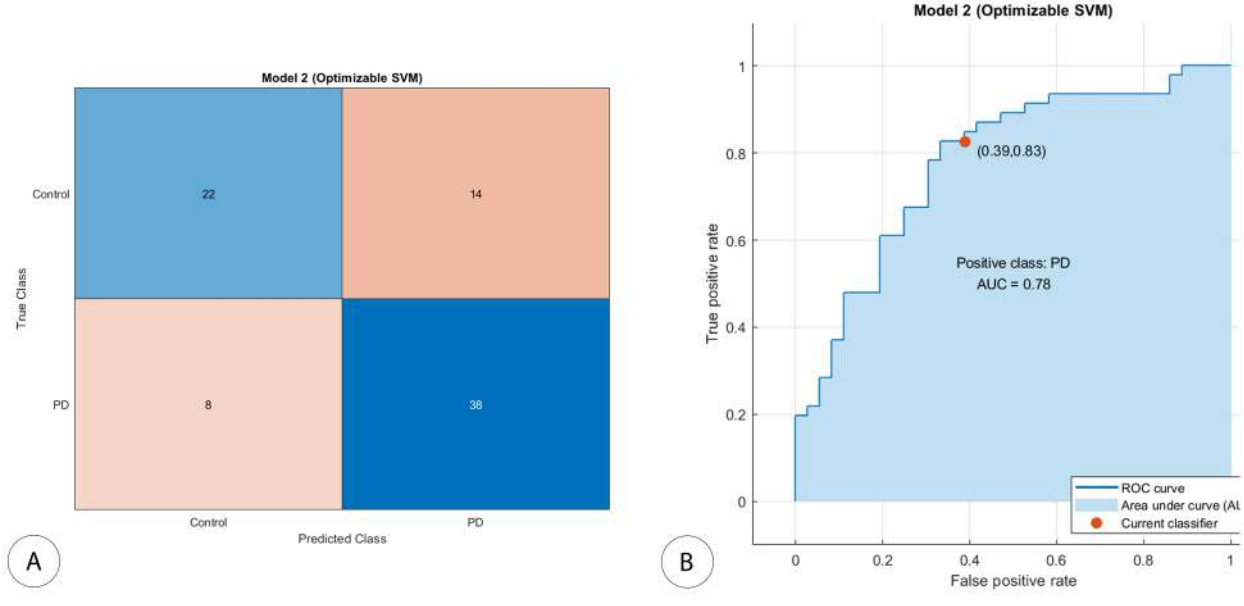

Figure 4. Results of Machine Learning Model 2 against the training dataset, including the Confusion Matrix (4a) and the ROC curve (4b)
A) Confusion Matrix of Machine Learning Model 2 with the Training Dataset
B) ROC Curve including the calculated AUC using Machine Learning Model 2 with the Training Dataset

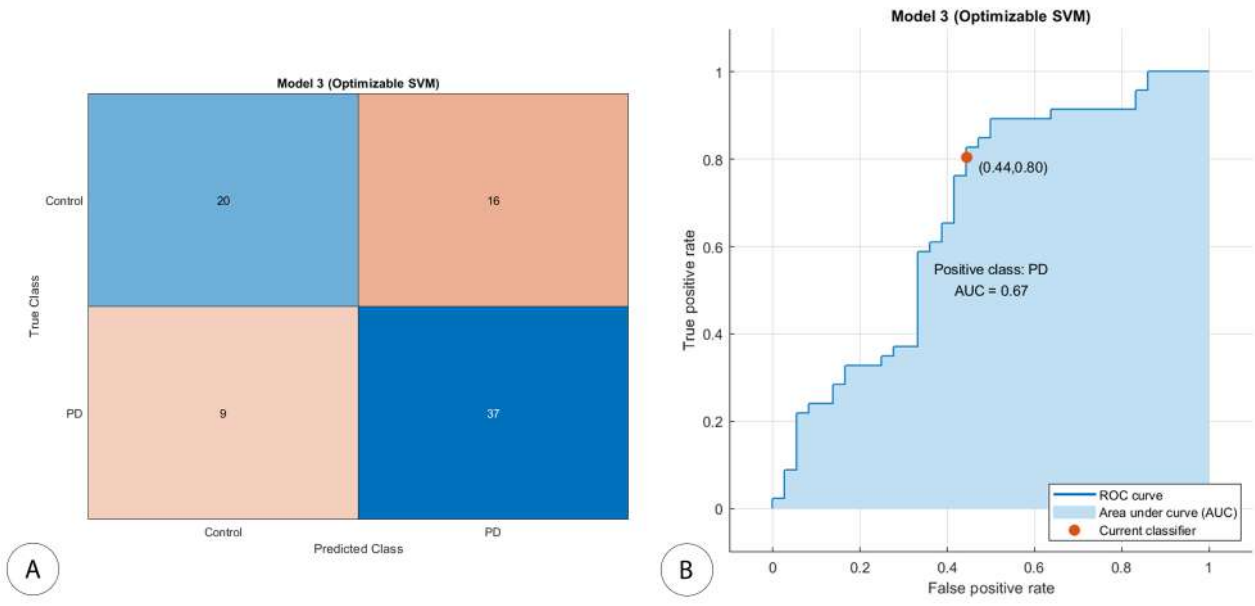

Figure 5. Results of Machine Learning Model 3 against the training dataset, including the Confusion Matrix (5a) and the ROC curve (5b)
A) Confusion Matrix of Machine Learning Model 3 with the Training Dataset
B) ROC Curve including the calculated AUC using Machine Learning Model 3 with the Training Dataset

Table 2. Mann-Whitney U Test Results for Averaged Power Spectral Data

Group	VGRF 1	VGRF 2	VGRF 3	VGRF 4	VGRF 5	VGRF 6	VGRF 7	VGRF 8
Left Foot	6e-43	2e-24	0.7020	2e-23	2e-04	1e-07	5e-04	0.1309
Right Foot	1e-32	2e-24	0.0012	4e-16	1e-08	6e-08	0.0057	0.4567

Table 3. Summary of Machine Learning Model Results

Models	Accuracy		Sensitivity		Specificity		AUC	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
Model 1	76.8%	67.1%	80.43%	55.56%	72.22%	76.09%	0.76	0.71
Model 2	73.2%	67.1%	82.61%	80.43%	61.11%	50.00%	0.78	0.73
Model 3	69.5%	63.4%	80.43%	71.74%	55.56%	52.78%	0.67	0.70

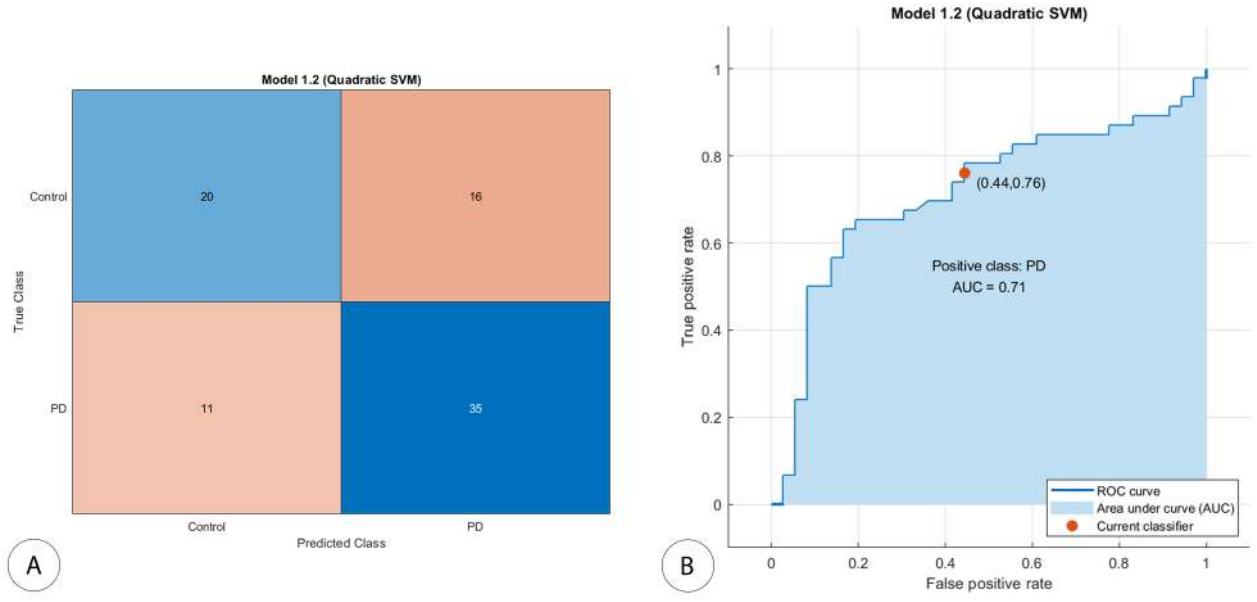

Figure 6. Results of Machine Learning Model 1 against the testing dataset, including the Confusion Matrix (6a) and the ROC curve (6b)
A) Confusion Matrix of Machine Learning Model 1 with the Testing Dataset
B) ROC Curve including the calculated AUC using Machine Learning Model 1 with the Testing Dataset

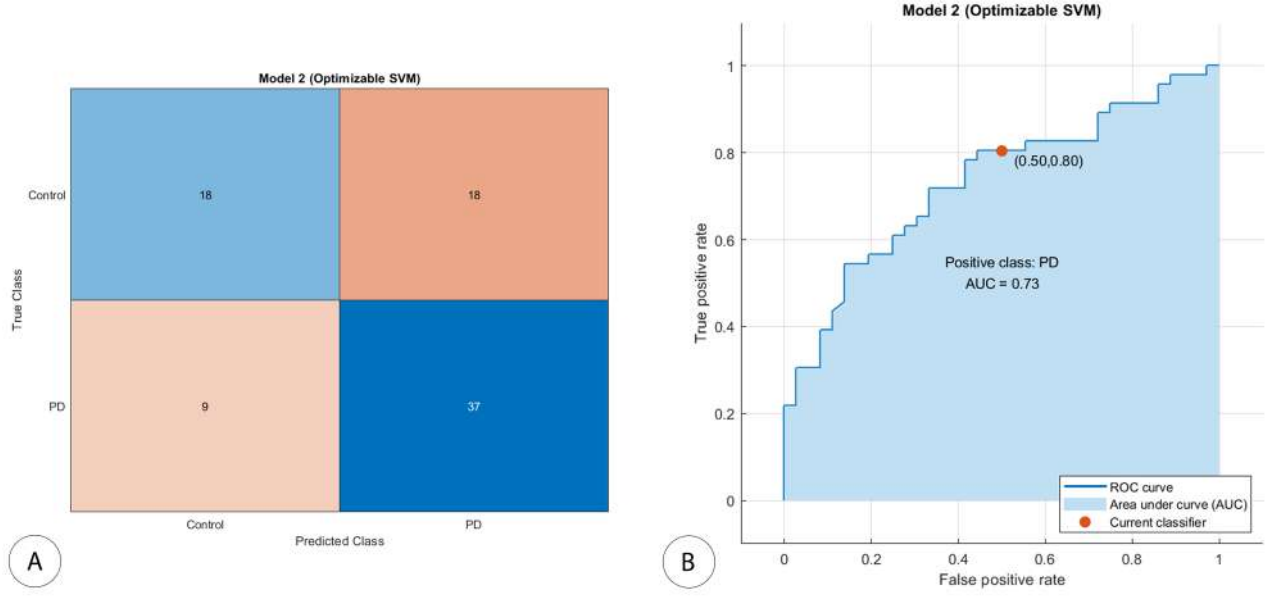

Figure 7. Results of Machine Learning Model 2 against the testing dataset, including the Confusion Matrix (7a) and the ROC curve (7b)
A) Confusion Matrix of Machine Learning Model 2 with the Testing Dataset
B) ROC Curve including the calculated AUC using Machine Learning Model 2 with the Testing Dataset

Model 1 (Quadratic SVM) yields the best accuracy and specificity post-training and post-validation. Additionally, Model 2 (Optimized SVM) yields the best sensitivity post-training and post-validation. Lastly, it can be seen that Model 2 has the best AUC using the training and testing dataset, thereby indicating it has the best performance when classifying PD subjects using VGRF data. However, when determining if the differences in AUC were statistically significant via the Hanley and McNeil approximation, it was determined that none of the models were significantly different from each other.[12] This can be seen in Tables 4-6.

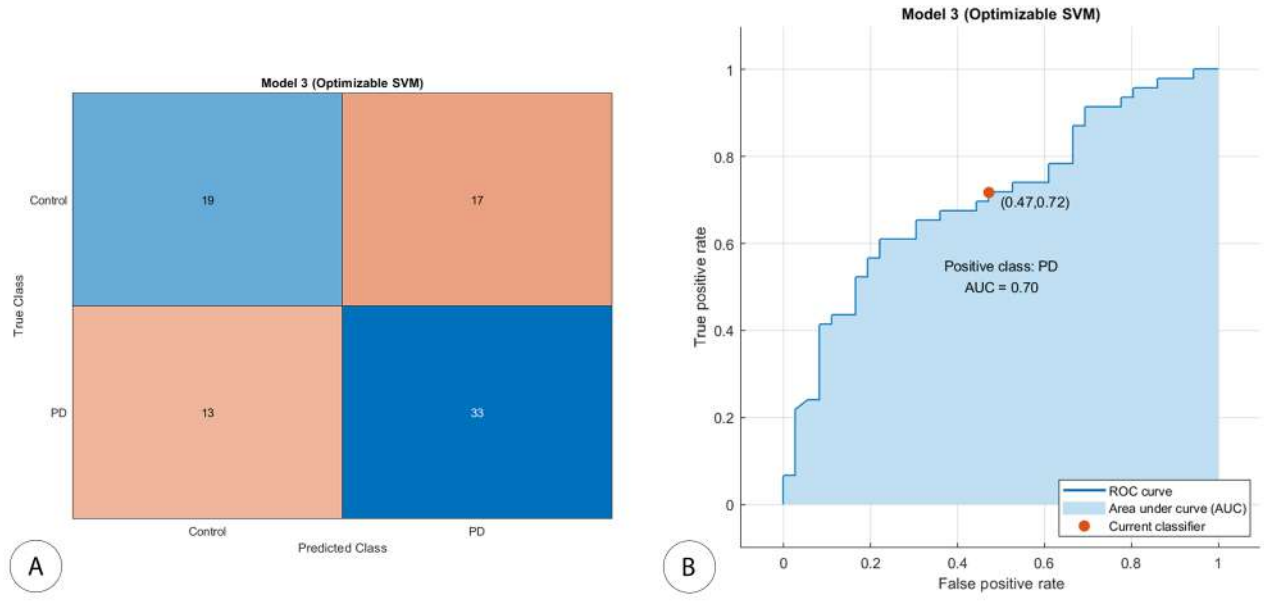

Figure 8. Results of Machine Learning Model 3 against the testing dataset, including the Confusion Matrix (8a) and the ROC curve (8b)
A) Confusion Matrix of Machine Learning Model 3 with the Testing Dataset
B) ROC Curve including the calculated AUC using Machine Learning Model 3 with the Testing Dataset

Table 4. Area Under Curve comparisons between Training and Testing Models

Models	Training	Testing	P Values
Model 1	0.76	0.71	0.258
Model 2	0.78	0.73	0.251
Model 3	0.67	0.70	0.358

Table 5. Area Under Curve comparisons between Training Model Types

Model Comparisons	AUC #1	AUC #2	P Values
Model 1 ↔ Model 2	0.76	0.78	0.391
Model 1 ↔ Model 3	0.76	0.67	0.127
Model 2 ↔ Model 3	0.78	0.67	0.078

Table 6. Area Under Curve comparisons between Testing Model Types

Model Comparisons	AUC #1	AUC #2	P Values
Model 1 ↔ Model 2	0.71	0.73	0.400
Model 1 ↔ Model 3	0.71	0.70	0.450
Model 2 ↔ Model 3	0.73	0.70	0.353

Discussion

This study aimed to explore the potential of using Vertical Ground Reac-tion Force (VGRF) data in both Fourier analysis and machine learning models for predicting Parkinson's disease (PD) based on gait abnormali-ties. The primary findings include:

1) Fourier analysis revealed significant differences in low and high-frequency power between PD and control subjects,
2) machine learning models successfully classified PD and control sub-jects based on gait data, although sensitivity varied across models.

Fourier Analysis Discussion

The Fourier analysis demonstrated significant differences in the low and high-frequency power spectra between PD and control groups, though some inconsistencies in trends were noted. Initially, it was hypothesized that control subjects would exhibit higher low-frequency power than PD subjects, but this was not consistently observed across all sensors. Spe-cifically, while VGRF 1 (heel) in control subjects showed higher low-frequency power, this was not the case for PD subjects, where VGRF 1 showed similar power to VGRF 7 (first metatarsal) and VGRF 4 (lateral arch). This variation could be attributed to the rigidity of PD gait, where pressure distribution is more uniform across the foot, contrary to the more localized pressure in healthy subjects. This aligns with previous studies suggesting that PD patients often demonstrate altered gait pat-terns with reduced variability in foot pressure distribution, potentially due to motor dysfunction and rigidity.[11] A figure of the foot pressure distri-bution in normal, healthy patients can be seen in Figure 9.[11]

Regarding the high-frequency components, we initially hypothesized that PD subjects would show increased high-frequency activity due to freezing of gait, a common feature in PD. However, no clear differences were observed between the two groups. One possible explanation for this discrepancy is experimental limitations, such as participant pauses or rotational movements during the trials, which may have influenced the high-frequency data. This suggests that while high-frequency com-ponents may be indicative of PD, further refinement in experimental de-sign, particularly addressing gait interruptions, is needed for clearer dif-ferentiation. This finding contrasts with some studies that have observed elevated high-frequency components in PD patients during gait analysis.[11] Thus, further research should explore the specific impact of motor fluctuations on high-frequency data.

In terms of statistical analysis, the Mann-Whitney U test confirmed significant differences in band power between control and PD subjects for most sensors, reinforcing

the notion that PD gait disturbances manifest through altered pressure distribution. However, exceptions were noted for VGRF 3 (medial arch) and VGRF 8 (toes), which did not show as much variation. It is hypothesized that these areas may not be as in-fluenced by rigidity as other regions of the foot, suggesting that rigidity impacts pressure distribution more in the rearfoot and midfoot than in the forefoot. Additionally, the use of average power spectral estimates in the Fourier analysis may have introduced errors, and individual spectral estimates may offer more precise insights into group differences. These results highlight the need for careful consideration of methodological choices when analyzing VGRF data.

Machine Learning Discussion

The machine learning models demonstrated fair classification per-formance, with an area under the curve (AUC) of approximately 0.71 for all models. This suggests that the models were able to differentiate be-tween PD and control subjects based on gait data with moderate accu-racy. Notably, there was no significant overfitting, as evidenced by simi-lar performance in both the training and testing phases of the models. However, the sensitivity of Model 1 dropped by about 25% between training and testing, which could be due to suboptimal model parame-ters or the use of an inappropriate kernel for the data. This finding un-derscores the importance of model optimization and the potential benefit

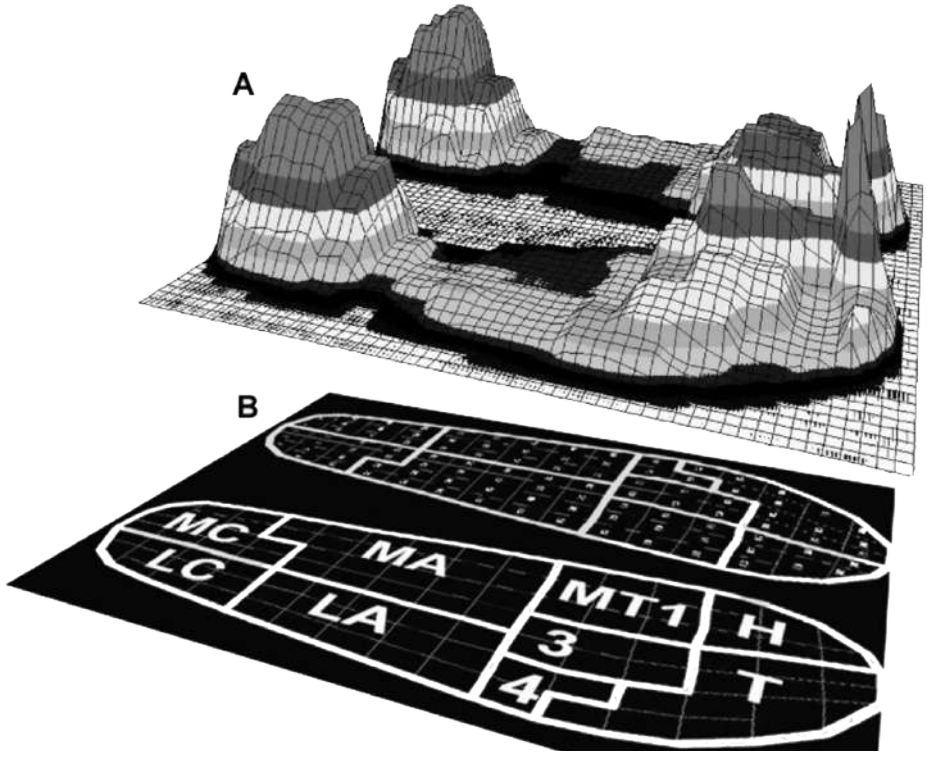

Figure 9. Foot pressure distribution patterns observed during walking. A. 3D distribution of pressure for one healthy subject. B. The nine anatomical areas superimposed on the sole of a foot (MC = medial calcaneus, LC = lateral calcaneus, MA = medial arch, LA = lateral arch, MT1 = first metatarsal, 3 = second and third metatarsal, 4 = fourth and fifth metatarsal, H = hallux, and T = toes)

of exploring alternative kernels to improve sensitivity.

Interestingly, Model 3, which incorporated Principal Component Analysis (PCA) for feature reduction, performed worse than Model 2, which used all features without dimensionality reduction. This aligns with previous research indicating that feature selection or reduction methods like PCA can sometimes result in the loss of important information, thereby impairing model performance.[8] Therefore, it is con-cluded that all features used in Model 2 were relevant for PD classifica-tion and should be considered in future models. However, given that no statistically significant differences were observed between the models, further research must be conducted to determine feature selection and optimization algorithms that can be used to improve the model's perfor-mance in PD classification.

Limitations and Future Directions

Several limitations of this study should be addressed in future re-search. First, the sample size in both the Fourier analysis and machine learning components was limited, which may have restricted the statisti-cal power and generalizability of the findings. A larger cohort, including a broader range of PD stages, would improve the robustness and clinical applicability of the results. Second, the experimental design for the Fourier analysis could be refined, particularly by controlling for potential interruptions in gait such as pauses or rotational movements that might affect high-frequency data. Lastly, the machine learning models would benefit from the inclusion of more gait features, particularly those related to spatial aspects of walking, to further enhance classification accuracy.

Conclusion

In conclusion, this study provides valuable insights into the use of VGRF data for predicting PD and highlights the potential of Fourier analysis and machine learning models as tools for PD de-tection. The significant differences observed in low-frequency components of VGRF data be-tween PD and control subjects emphasize the importance of gait abnormalities as biomarkers for PD. Furthermore, the successful application of machine learning models demonstrates the potential of automated gait analysis for PD classification. However, further refinement in both data collection and model optimization is required to improve the accuracy and clinical utility of these methods. The integration of additional features and the use of larger, more diverse da-tasets will be key to advancing these techniques for early detection and monitoring of Parkin-son's disease.

Acknowledgements

The author gratefully acknowledges Dr. Taylor deVet for her support.

Conflicts of Interest

None Reported. No sources of funding as well.

References

1. "Parkinson's disease -Symptoms and causes," Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/symptomscauses/syc-

20376055 (accessed Oct. 15, 2021).

2. A. Dancis and V. T. Cotter, "Diagnosis and Management of Cognitive Impairment in Parkinson's Disease," The Journal for Nurse Practitioners, vol. 11, no. 3, pp. 307–313, Mar. 2015, doi: 10.1016/j.nurpra.2014.11.023.

3. "Parkinson's Disease," National Institute on Aging. http://www.nia.nih.gov/health/parkinsons-disease (accessed Oct. 15, 2021).

4. G. Ebersbach, C. Moreau, F. Gandor, L. Defebvre, and D. Devos, "Clinical syndromes: Parkinsonian gait," Movement Disorders, vol. 28, no. 11, pp. 1552–1559, 2013, doi: 10.1002/mds.25675.

5. L. di Biase et al., "Gait Analysis in Parkinson's Disease: An Overview of the Most Accurate Markers for Diagnosis and Symptoms Monitoring," Sensors (Basel), vol. 20, no. 12, p. 3529, Jun. 2020, doi: 10.3390/s20123529.

6. J. M. Hausdorff et al., "Dual tasking, gait rhythmicity, and Parkinson's disease: Which aspects of gait are attention demanding?," European Journal of Neuroscience, vol. 22, no. 5, pp. 1248–1256, Sep. 2005, doi: 10.1111/j.1460-9568.2005.04298.x.

7. B. E, B. D, V. K. Elumalai, and U. K, "Data-driven gait analysis for diagnosis and severity rating of Parkinson's disease," Medical Engineering & Physics, vol. 91, pp. 54–64, May 2021, doi: 10.1016/j.medengphy.2021.03.005.

8. "Parkinson Disease Gait Classification based on Machine Learning Approach." https://scialert.net/abstract/?doi=jas.2012.180.185 (accessed Oct. 15, 2021).

9. M. N. Alam, A. Garg, T. T. K. Munia, R. Fazel-Rezai, and K. Tavakolian, "Vertical ground reaction force marker for Parkinson's disease," PLOS ONE, vol. 12, no. 5, p. e0175951, May 2017, doi: 10.1371/journal.pone.0175951.

10. M. B¨achlin, J. Hausdorff, D. Roggen, N. Giladi, M. Plotnik, and G. Tr¨oster, "Online Detection of Freezing of Gait in Parkinson's Disease Patients: A Performance Characterization," BODYNETS 2009 4th International ICST Conference on Body Area Networks, p. 11, Apr. 2009, doi: 10.4108/ICST.BODYNETS2009.5852.

11. M. J. Hessert, M. Vyas, J. Leach, K. Hu, L. A. Lipsitz, and V. Novak, "Foot pressure distribution during walking in young and old adults," BMC Geriatrics, vol. 5, no. 1, p. 8, May 2005, doi: 10.1186/1471-2318-5-8.

12. Hanley JA, McNeil BJ. The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. Radiology, 1982, 143, 29-36.