Predictive Analysis of Parkinson's disease from Brain MRI images and from Gait Sensor Data

Nikhil Sharma
23MCS0020
MTech, CSE, SCOPE
VIT Vellore, TN, India
nikhil.sharma2023a@vitstudent.ac.in

Bhavesh Kumar Modi
23MCS0045
MTech, CSE, SCOPE
VIT Vellore, TN, India
bhavesh.kumar2023a@vitstudent.ac.in

Anil Kumar K
Associate Professor Sr, SCOPE
VIT Vellore, TN, India

Abstract: **Parkinson's Disease (PD) remains a significant challenge in the medical sector due to its progressive and debilitating nature. Early-stage detection with high accuracy continues to elude clinicians. This study aims to leverage advancements in machine learning, particularly deep learning algorithms, to address this challenge. By focusing on common neurodegenerative movement disorders like PD, we seek to identify key symptoms and determine optimal treatment initiation points to mitigate its effects. Our research explores the integration of gait sensor data with other motor, non-motor, and brain image data to enhance diagnostic accuracy. Through a comprehensive analysis of multi-modal data, including gait sensor data alongside motor and non-motor symptoms, we aim to improve the clinical diagnosis of PD. By combining features from different modalities, we anticipate achieving better accuracy in predicting PD compared to single-modality approaches. Furthermore, this study represents a pioneering effort in the realm of computer vision by merging features from disparate subjects and modalities. The implementation of deep learning techniques promises to unlock new insights and potential breakthroughs in PD prediction and management.**

Keywords: ***Parkinson's disease, deep learning, multimodal data, gait sensor, brain imaging, classification, non-invasive diagnosis, Deep neural networks, Feature extraction, Feature fusion.***

1. INTRODUCTION

Parkinson's Disease, a well-known incurable central nervous system disorder has affected nearly 8.5 million individuals around the world. The symptoms of this disease turn up slowly and damage the motor activity system of a human. Some of the early symptoms are rigidity, tremor, slowness of movement, difficulty in walking with some behavioural problems. These symptoms become more visible at the final stage of this disease. Even today we do not know the main cause of Parkinson's disease and researchers are still figuring out whether there is any similarity among PD patients. We have seen that at the advanced phase of this disease patients endure stiff muscles, olfactory loss, rapid eye movement, and sleep disorder which affects their everyday life. Although at an early stage we often fail to notice the mild symptoms of this disease and after some years these symptoms convert into chronological neuro-degenerative disorder. Researchers and doctors find it difficult to diagnose Parkinson's at an early stage due to its variety of disorders. Several months of observations give us a clear picture of those symptoms in an individual. Today's evolution of machine learning and deep learning techniques has created a path to help doctors with the diagnosis of Parkinson's disease. Scientists have started experimenting with this disease and the patient's symptoms characteristic. At the early stage of this experiment, various authors only looked at the direction of the finding similarity of the same disease patients. But later, researchers find that several commodities are affecting this diagnosis process of Parkinson's. Experimenting with only one type of symptom only gives us accuracy for a particular type of data. Still, we have seen that a patient with Parkinson's disease may have several other symptoms which we have ignored during the experiment. So, scientists have come up with a differential identification of PD which will consider several symptoms of this disease while characterizing an individual as a PD or Healthy Control patient. The above consideration is not enough as in the medical sector we need nearly 100% accuracy and the technique mentioned before has a high error rate in the detection of Parkinson's disease.

Scientists have found that the problem in the clinical diagnosis of PD is mostly because clinical identification is based on clinical tests and patients' responses to different medicine with assistance from neuro-images of the brain. The other challenge we have already talked about is early detection. In this scenario, the structure inside the brain changes followed by some subtle symptoms. The brain image of a PD patient shows that the most affected areas are the substantia nigra and basal ganglia. Due to the loss of neurons, the disease shows a dopaminergic effect in a patient which is much before any motor symptoms. The doctor generally tries to identify PD at the initial phase so that neuro-protective medicine can slow down the progressive nature of this disorder and can help in preventing clinical PD symptoms. The use of ML techniques in the identification of PD become more popular in the last ten years and so. This algorithm helps in finding a pattern in clinical data or images and tries to classify PD and HC patients. In this experiment, we try to classify Parkinson's disease patients from gait sensor data and brain MRI images.

We have previously seen that brain image can assist in diagnosing PD patients. On the other hand, Gait analysis studies human movement, body mechanics, and muscle activity. It is typically performed using some sensors attached under the foot and collecting the pressure sensor data for two to three minutes based on that data a neurosurgeon analyses the result

and identify the patient has PD or not depending on his knowledge and visual experience. As we are handling two different sources of data where there is no relationship between any of the records between two datasets as patients from each of the datasets are different from each other. For this scenario, we have planned to use the multi-modal framework to combine the characteristic of each of the datasets altogether to classify PD and Healthy Control people efficiently and more accurately. We will experiment with this concept against each set of datasets where we will collectively observe the improvement of using gait data with brain images, DNA-RNA samples, and clinical data of the PPMI database. We have seen in medical diagnosis doctors used to give his decision based on observation of several medical diagnosis like MRI, CT, SPECT, f-MRI images along with clinical measurement like motor and non-motor symptoms data. But in general, there is always a problem of missing medical images of a patient due to unavailability of MRI, CT machine or any other critical situations. In that case doctors must to rely on only on motor and non-motor symptoms data available at that moment. Here we have described a possibility of classifying PPMI data coming from a patient by combining with Gait sensor data which is of different patients. Idea is to check if there can be a model that can effectively classify a patient data by combining with its missing clinical data.

2. Literature review

Research studies that are based on machine learning techniques have been done until now a can be categorized into three parts. Those are 1. Studies to discriminate PD and HC using machine-learning Techniques; 2. Differential diagnosis; 3. Initial detection of Parkinson's disease. In the following subsection, we have summarized ML techniques used in imaging studies where the researchers have considered functional MRI (fMRI), structural MRI, PET, and SPECT images.

2.1 Studies to discriminate PD and HC using machine-learning Techniques:

Previously we have seen that data coming from a single source has not the ability to describe the important behaviours of PD. That is why researchers started considering combined data of images and clinical data while classifying PD patients while data are measured on different scales. S. Ghosh, P. K. Mandal, S. Dutta Roy, and Prashanth.[16] inferred that combinational analysis of CSF, non-motor, and image data can contribute to the preclinical diagnosis of PD. Whereas Enrico Glaab[5] and his team proved that incorporating blood sample data with PET image data can enhance the diagnosis process of distinguishing PD and healthy control people. Another researcher Thomas J Hirschauer and his team[8] proposed an approach of using neuro-pathological data of the brain with motor and non-motor symptoms data in an enhanced probabilistic neural network. The result came out surprisingly better than any other classical machine learning algorithms with high accuracy than any other previous experiment.

2.2 Studies on different diagnoses of PD using machine-learning Techniques:

We have already discussed that there are many variations of PD. It is crucial to differentiate each of them for separate treatment. In this context, some experiments have been done in recent years. Dopaminergic images, structural MRI, functional MRI, and diffusion tensor images worked very well to separate each category of Parkinson's disease with the help of popular machine learning techniques like SVM, logistic regression, etc. But the idea of combining multi-modal features can be fruitful in the case of distinguishing PD variation. This technique has been followed in some of the recent experiments such as Andrea Cherubini and his group[2] combined DTI image and morphometry which is based on voxel and used in support vector machine to separate patients with PSP and PD. This experiment defined that automated pattern recognition can be beneficial for detecting patients having PSP or PD. Another study by G. Du, M. M. Lewis, S .Kanekar , and his team[4] proved that apparent transverse relaxation rate and DTI image can have the ability to be an important marker for the differential identification of PD disorder.

2.3 Experiments on the initial phase identification of PD using ML Techniques:

As we have discussed in some of the previous sections, it will be beneficial for a patient if we detect Parkinson's disease at an early stage so that many treatments can be possible to delay the progression of this disorder. Some studies have worked towards the same aim of combining image data with clinical data for improving the initial diagnosis of PD. Dan Long, Jinwei Wang, and their team[12] worked with structural data and resting-state functional magnetic resonance image (MRI) data, from image data they collected ALFF (amplitude of low frequency fluctuations), RFCs (regional functional connectivity strength), ReHo (regional homogeneity), characteristic and from the other data they extracted the cerebrospinal fluid (CSF), the gray matter (GM) and the white matter (WM). Using a two-sample t-test they were able to classify early PD patients using SVM with an accuracy close to 87% which is higher than single source image data. Another study by Oliveira etal.[14] experimented with SPECT imaging data where they extracted seven features from each brain hemisphere and applied classical machine learning techniques. The result showed the highest classification accuracy of 97% when they used all the features combinedly other than using those separately which also had better results compared to any other previous studies. Considering the research studies discussed till now, we have identified some limitations and challenges. As we know that clinical diagnosis data is very much prone to error, so judgment based on clinical data for classifying Parkinson's disease patients is not a good idea. So other than the supervised learning method, we can apply unsupervised techniques which will try to find underlying data patterns to label the data more accurately. But they have their technical problem in applying image data as these approaches are not good at extracting features from the image. The second problem is machine learning techniques can identify the data pattern but the whole process is like a black box and it is invisible to researchers and sometimes it is hard to understand. This whole situation is against the rule of evidence-based medicine. So, we need to identify disease-specific features using this machine-learning approach. And the last most important challenge is over-fitting in machine learning. This problem generally happens when the machine learning model when it performs relatively better on the train dataset but fails miserably on a test dataset. In recent studies, it has been proved that too much heterogeneity in Parkinson's disease data causes less generalizability even if we consider only a single source of data. So, we need to improve our data with A large number of data samples and also have to use proper validation methods to help the model generalize the data.

3. Proposed Problem Statement

After a considerable understanding of the above-mentioned limitation, we have come up with the idea of creating a multi-modal framework using two different sources of data for disease classification. To achieve our goal, we have combined Gait analysis data with clinical data and imaging data for better performance improvement in classifying Parkinson's disease. Along with that, we have also used an efficient cross-validation method for eliminating the over-fitting problems in machine learning model. We have tried to reduce the error rate in classification to make this process more reliable in clinical diagnosis. Also, this experiment has tried towards classifying single modal data more accurately using multimodal features. So collectively we can summarize our contribution to this experiment as follows:

- Enhancement of PD diagnosis and early detection using Multi-modal data.
- Combine Gait analysis data features with MRI data.
- Validate and optimize the model to make the process more accurate and reliable.
- Reduce error rate in diagnosis.
- Evaluate the performance of the framework with single modal architecture.

4. Description of the Dataset

In this experiment, we have chosen two datasets to work with. One of them is the Gait dataset from the Physio net database and the other one is clinical and images brain image data from the PPMI database. In the following sub-section, we have discussed about the datasets in a detailed manner.

4.1 Gait Analysis Data:

The Physio net Gait database includes measurement of gait analysis data of 93 individuals with PD with a mean age of 66.3 years of which 63% are male patients, along with 73 healthy control individuals with a mean age of 66.3 years of which 55% are male patients. The database contains the records of force applied to the ground vertically by subjects as they walked in a normal fashion with their usual speed for almost 2 minutes on plain ground. Under the foot of each leg, there are eight sensors to calculate strength as a function of time. The outcome of each of these sixteen sensors has been stored at a 100 Hz sample rate. This record will help us to experiment on the force data and centre of pressure data as a function of time and location. This repository has also contained demographic information like measurement of disorder severity using the Hoehn and Yahr staging and the UPDR (Unified Parkinson's Disease Rating) scale. Before standardizing the data we have converted the time series data of each patients in more readable format. Also, each sensor data is converted two seven statistical measurements, so that there is no loss of information.

4.2 PPMI Dataset:

The second dataset that has been used in this experiment was collected from the PPMI (Parkinson's Progression Markers Initiative) database. This database has records which are also known as the multi-centre and international study that is used to identify PD features for diagnosis and to detect disease progressiveness. This database stores the data in a particular format where the participants are needed to examine 6 clinical assessments and neuro-image data for the two most important areas in the brain. Those clinical assessments are: first, the Movement Disorder Society-sponsored revision of the United Parkinson's Disease Rating Scale (MDS-UPDRS) Part I, which has the records of the non-motor experiences of everyday life; second the MDS-UPDRS Part II, which has the records of the motor experiences of everyday life; third the MDS-UPDRS Part III, which contains records that observe motor function; fourth, the Montreal Cognitive Assessment (MoCA), which contains data of cognitive function; fifth, the Scales for Outcomes in PD – Autonomic (SCOPA-AUT), which contains records of autonomic function; and sixth, the University of Pennsylvania Smell Identification Test (UPSIT), that records data regarding assessment of olfactory function. From the 1141 participants whose data is stored in the PPMI repository, 666 participants which are 58.4% of total participants, have undergone all of the above-mentioned clinical and diagnostic assessments are added to this database. Now from the 666 participants whose data was added to the study, 189 participants which are 28.4% are detected with PD, whereas 62 subjects which are 9.3% of the total participants shows scans without evidence of dopaminergic deficits, and the rest of the 415 participants that is 62.3% are healthy participants. The rate of exclusion of each individual does not vary notably between each class of data. There are a few things to keep in mind:

- The clinical data for everyone is collected over 12 visits along with an initial treatment result.
- Here the clinical motor assessment data is observed at three-month intervals during the first year and from the next year, it was recorded at six-month intervals.
- The behavioural and cognitive diagnoses were conducted at twelve-month intervals for all the individuals.
- VMAT-2 (Australia only) or AT image data was recorded at the time of 12-month, 24- month, and 48-month appointments for PD subjects, and for SWEDD participants it was 24-month visits, and for healthy control subjects, it was recorded only for baseline.
- Each participant undergoes MRI at baseline, and almost 50% of the individual undergoes DTI imaging at baseline, and DAT imaging it was also conducted at the same longitudinal intervals.
- There was a scheduled follow-up for the collection of blood at three-month intervals during the first year which is followed by another six-month interval.
- For each participant, CSF data were collected at 6, and twelve-month visits, followed by twelve-month intervals.
- Everyone undergoes a urine test at 12-month intervals.

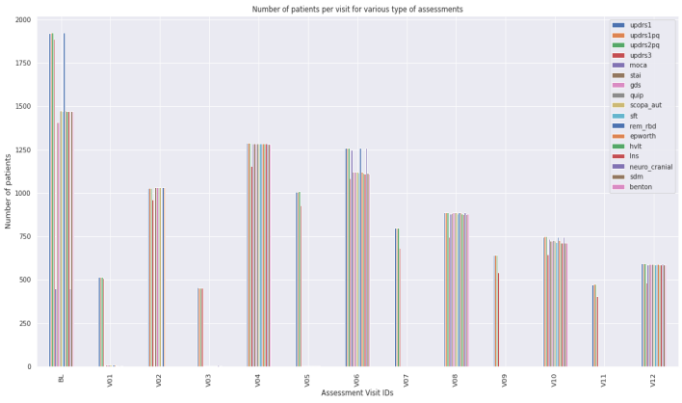


Fig 1: Clinical Diagnosis Distribution over the number of visits

Before starting the analysis, we have seen the distribution of each clinical assessment mentioned above over several visits starting from initial treatment (Baseline) to the last visit for an individual. In total there are total 13 patient visit records in this

database. From the above bar-plot distribution graph we can see that some of the visits do not have all the assessment records. So, we need to consider only those visits that have the above-mentioned clinical assessment record. In this regard, we have selected the initial visit along with visits no: 02, 04, 06, 08,10, and 12 for this analysis. We have identified a total of 476 patients who participated in all the assessments till the last visit.

5. Experiment

Before using those two datasets in a multi-modal framework, we have done some analysis on those datasets separately to understand how much accuracy they can produce if we use those datasets separately.

5.1 Gait Dataset:

Using the Physio net Gait dataset we experimented using a deep learning model for the identification of PD at an initial phase using only motor-based symptoms. We have created a small convolutional neural network model for this analysis which consists of 1D convolutional input layer along with several other layers which are described below:

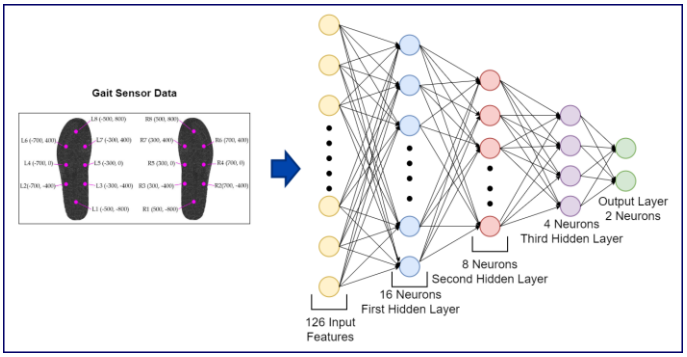
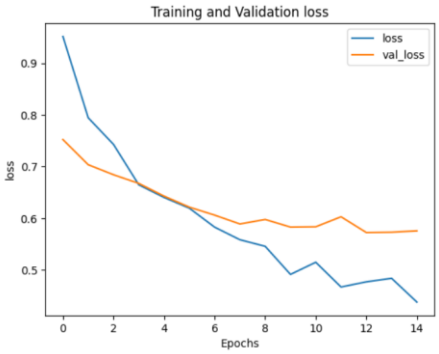
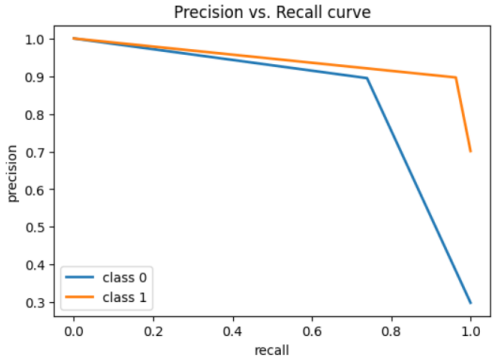


Fig 2: Gait Data Neural Network Model

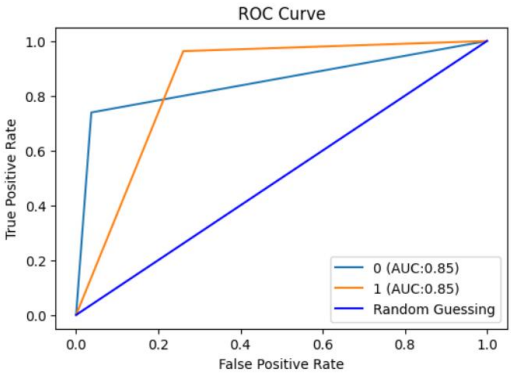
- The Second layer is another 1D convolutional layer with shape 16 and "Relu" activation function.
- The third layer is a dropout layer with a dropout probability of 0.3.
- Next layer is a 1D convolutional layer with shape 8 and "Relu" activation function.
- The fifth layer is a simple is a dropout layer with a dropout probability of 0.3.
- Then the next layer is a fully connected dense layer with shape 4 and activation function selected as "Relu".
- And the output layer is also a fully connected dense layer with output shape 2 and the activation function is "Softmax".



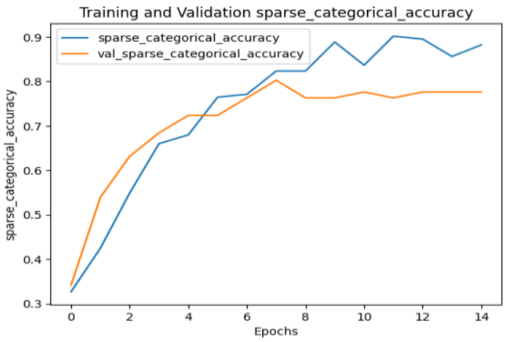
(a) Training and Validation loss



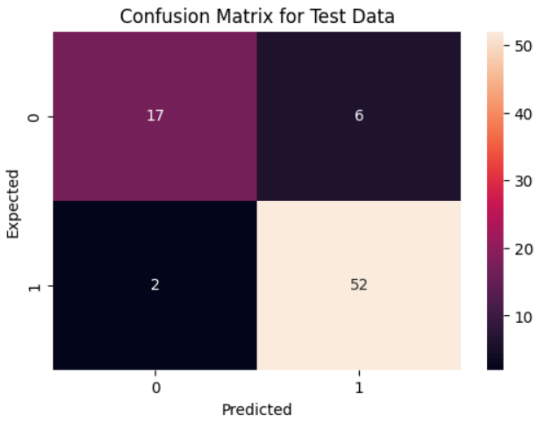
(b) Precision vs. Recall curve



(c) ROC Curve



(d) Training and Validation accuracy



(e) Confusion Matrix on Test Data

Classification Report

	precision	recall	f1-score	support
0	0.89	0.74	0.81	23
1	0.90	0.96	0.93	54
accuracy			0.90	77
macro avg	0.90	0.85	0.87	77
weighted avg	0.90	0.90	0.89	77

Cohen Kappa Score
0.7389830508474576

(f) Classification Report

Fig 3. Result using Gait Model on Gait data:

While compiling the model we have chosen "Sparse Categorical Cross-Entropy" as the loss function with "RMSprop" optimizer and learning rate as 0.0001. We have ran this model for 15 epochs with a batch size of 5 and achieved accuracy of 89.61% along with Cohen Kappa score of 0.7389 and ROC AUC score as 0.8510. From the above confusion matrix of test data, we can see that number of false positive classes while detecting PD patients from that of a healthy individual is much more than the false negative case. For an individual with PD, if the model detects negative, it will be difficult to implement in critical clinical diagnosis. Also, the ROC accuracy has not crossed above 90%, so we need to think about making a more complex model to understand the underlying data pattern.

5.2 PPMI Dataset:

From the PPMI dataset, we will use the potential biomarkers of Parkinson's disease like motor-based symptoms, and non-motor-based symptoms which are recorded after multiple clinical diagnoses of a particular patient. This analysis was done with aim of classifying Parkinson's disease patients into three categories based on the rate of disease progression. We have considered four motor-based symptom data, one neuro-logical assessment data, and twelve non-motor-based symptom data records which in total gives us 1595 across the several visits mentioned previously. Similar to Gait model we have created a small neural network model which is described below:

- The Second layer is another 1D convolutional layer with shape 128 and "Relu" activation function.
- The third layer is a dropout layer with a dropout probability of 0.5.
- Next layer is a 1D convolutional layer with shape 64 and "Relu" activation function.
- The fifth layer is a simple is a dropout layer with a dropout probability of 0.5.
- Then the next layer is a fully connected dense layer with shape 32 and activation function selected as "Relu".
- And the output layer is also a fully connected dense layer with output shape 2 and the activation function is "Softmax".

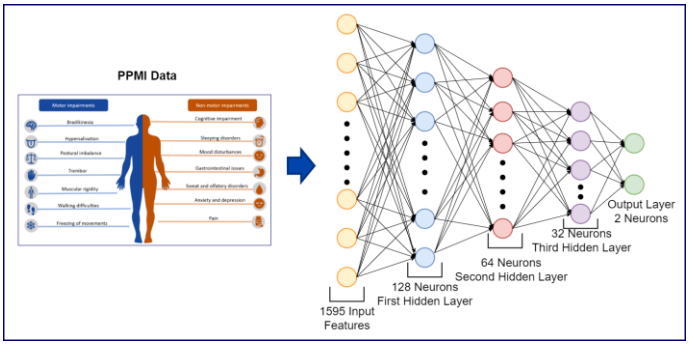
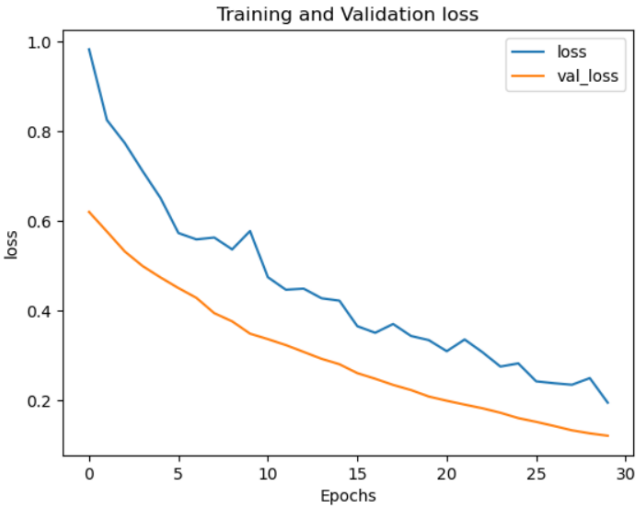
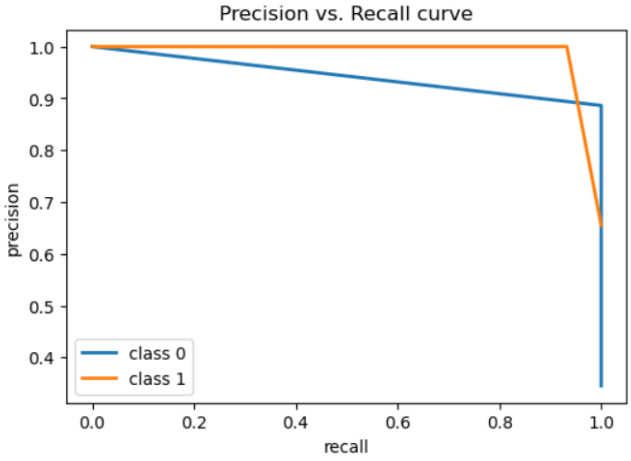


Figure 4: PPMI Data Neural Network Model.

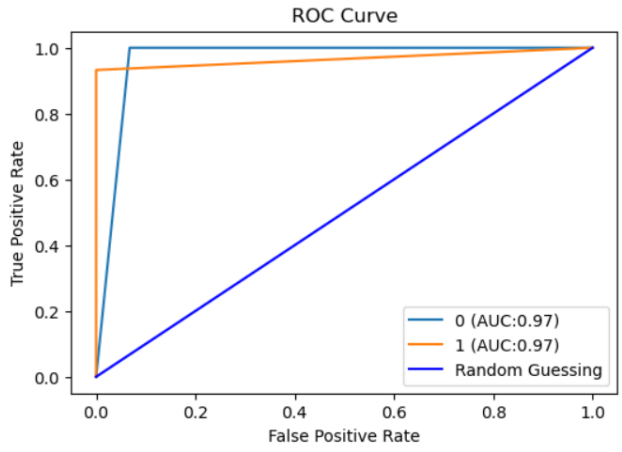
Similar to the Gait model we have chosen "Sparse Categorical Cross-Entropy" as the loss function with "RMSprop" optimizer and learning rate as 0.0001. We have run this model for 15 epochs with a batch size of 5 and it resulted in accuracy of 92.05% along with Cohen Kappa score of 0.8329 and ROC AUC score as 0.9391. From the below results we can observe there is a scope for achieving higher accuracy along with better AUC score. Also we need to reduce miss classified PD data as for small test set it is still more.



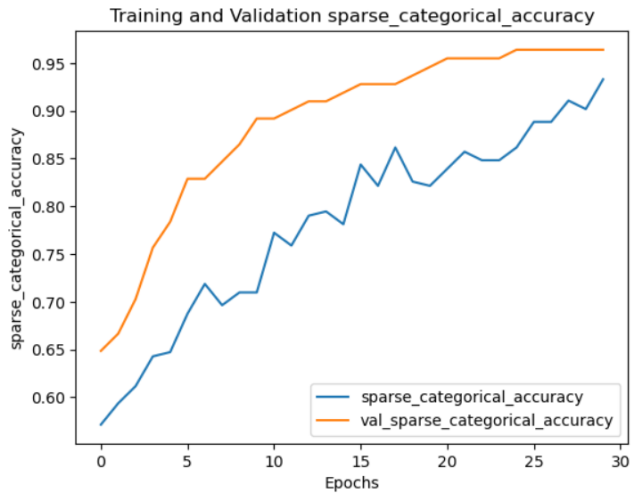
(a) Training and Validation loss



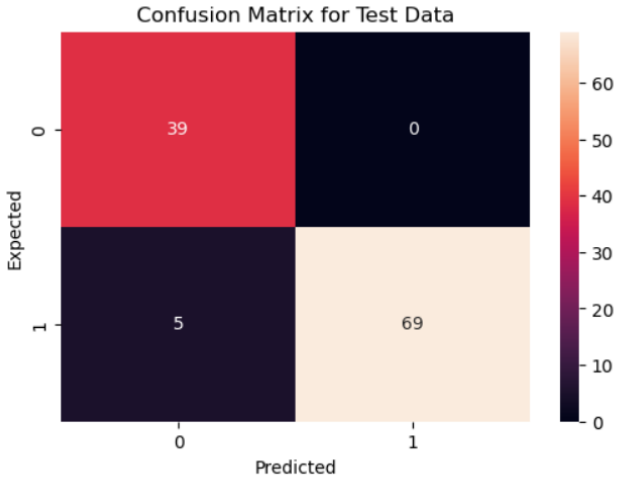
(b) Precision vs. Recall curve



(c) ROC Curve



(d) Training and Validation accuracy



(e) Confusion Matrix on Test Data

Classification Report

	precision	recall	f1-score	support
0	0.89	1.00	0.94	39
1	1.00	0.93	0.97	74
accuracy			0.96	113
macro avg	0.94	0.97	0.95	113
weighted avg	0.96	0.96	0.96	113

Cohen Kappa Score
0.9049941146796704

(f) Classification Report

Fig 5: Result using PPMI Model On PPMI data

5.3 Multi Modal Analysis:

After experiment on individual dataset, we have achieved 89% accuracy on Gait dataset and 91% accuracy on PPMI dataset but we can see that error rate still very high on both the cases. For that we have tried to explore the recent development in multi-modal analysis in machine learning where we can combinedly classify patient data. In that context we have tried to prepare data using both the dataset. In the following section we have described about data generation technique for multi-modal architecture.

5.3.1 Data Generation:

Data Generation: We have divided this part in several steps for better understanding. At first it starts with converting the Gait sensor data in more precise format to remove complexity while merging with PPMI data.

- It starts with applying seven important statistical approach on each column of sensor data like Minimum, Maximum, Mean, Median, Standard Deviation, Skewness, and Kurtosis.
- For each subject the whole time series data boils down to 1 single row with 18x7 = 126 features (each attribute converted to 7 other columns).
- For PPMI dataset, we have considered 7 visits data along with baseline visit for motor and non-motor symptoms data which gives 1595 number of features.
- Gait data is collection of 93 PD patients and 73 HC patients whereas PPMI consist of 294 PD patients and 154 HC patient records.
- Before merging, we have separated both the data based on each class distribution.
- Generated combined dataset by implementing cross product between those two-dataset depending on same class.
- For training 50%, with validation and testing 25% each of those datasets are used.
- In total, after splitting and cross product, it produced data for training 9484, validation 2363 and for testing 2517 rows records with 1721 features.

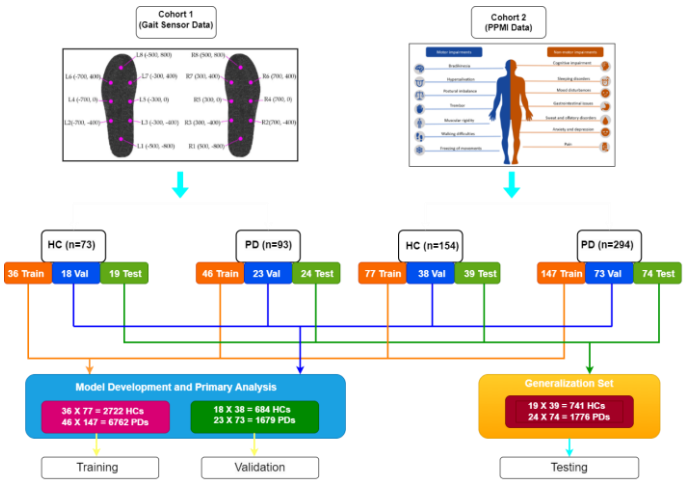


Fig. 6: Data Generation Diagram.

5.3.2 Triplet Loss:

The aim of triplet loss function is to learn from the distributed embeddings representation of data point in a manner in high dimensional vector space such that comparatively similar data points are projected in nearby space on the other hand dissimilar data points are projected far away from each other. The formula is given below:

$$L = max(d(a, p) - d(a,n) + margin, 0)$$

In the above representation it describes how we can calculate triplet loss which in return helps in measuring gradient in network. Here the variable “a” signifies the anchor sample, “p” signifies positive sample and “n” signifies negative sample. As we know that the similarity between a and p should be more than the similarity between a and n.

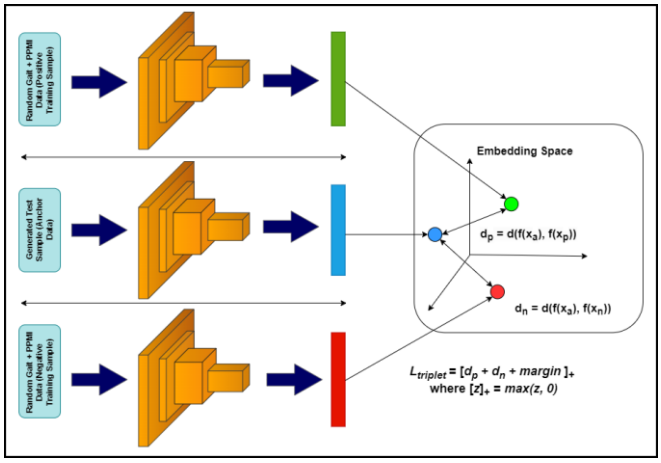


Fig 7: Triplet Loss Diagram.

There is also a hyperparameter called margin that is added to the loss function. This hyperparameter represents how far away the dissimilarities need to be. It assists us to recognise two samples much more efficiently. Eventually we measure the gradients, and these gradients helps us to update the biases and weights of the Siamese network. During training period, we collect an anchor sample along with a positive and negative sample selected randomly to calculate loss and update the gradients of the network.

5.3.3 Training Data Preparation:

To create triplets of we need a positive sample (whose label should be same with anchor sample), an anchor sample and a negative sample (whose label should be same with anchor sample). Both the PPMI and Gait dataset contain two classes. So, for each combined Gait and PPMI data (consider it as anchor sample) from train set, two samples are selected randomly from the same set where one of which is same class as the anchor sample (consider it as positive sample) and another is taken from different class (consider it as negative sample) from training set. Furthermore, for every batch we have selected n number of triplets.

5.3.4 Multi-Modal Architecture:

As we are handling two different types of information, we must think about combining two architectures like Gait and PPMI to create a framework that can work on whole features. This model will have the ability to extract the features from two different networks and concatenate those features and pass it through an aggregate network where the appropriate loss function will predict its label. We have implemented the idea of metric space method where it can identify if a pair of data is similar or not. That is why Siamese metric comes into picture which can calculate difference between two embedding vector and pass it through an energy function and predict that if that pair is similar or not. We have proposed a model that uses two other small architectures as backbone for helping it in creating feature vectors for Siamese network. As we have already discussed about individual Gait and PPMI neural network model, so we are not going for further discussion about those models again in this section. For input to Siamese network, which is basically a framework that uses input coming from more than one modality, we think of extracting second last layer of Gait and PPMI data set as a modified feature vectors for multi-modal input. Those feature vectors will have the characteristic of Gait analysis and PPMI diagnostic values individually which will be concatenated before considering them as an input to Siamese network. Combinedly, the Gait and PPMI features will create 36 feature vectors (4 vector from Gait model and 32 from PPMI model) and we will consider as input to Siamese network. Now the Siamese network will also have its own network architecture, that is discussed below:

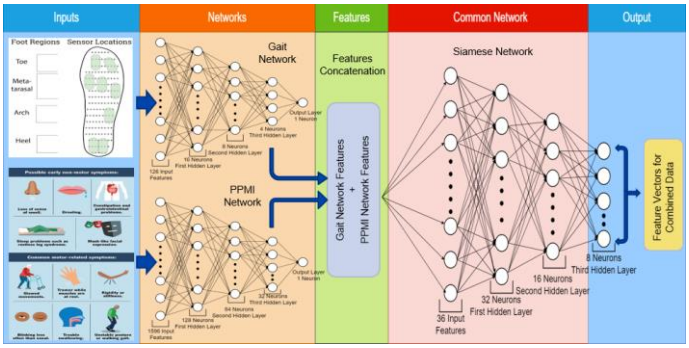


Fig 8: Siamese Network Diagram.

- Input layer with dimension of 36 features which are coming from Gait and PPMI network.
- Next a dense layer with size 32 and activation function 'Relu'. After that there is a dropout layer with rate 0.5 and then,
- Again, a dense layer with kernel size 16 with Relu activation function.
- And it ends with a dropout layer with rate 0.5 followed by a dense layer of kernel size 8 with same activation function that we have used before.

5.3.5 Result:

For this architecture, We have used RMSprop optimizer and set the learning rate at 0.0001 and compiled the model. Also, to validate the model against each set of training dataset we have generated several training sets in a randomized manner. After giving the input, the framework ran for 30 epochs with a batch size of 128 for each of these training set. We have validated our model with several test set data using cosine similarity method. This model has produced positive similarity of 97.45% between anchor and positive sample and negative similarity of 95.06% between anchor and negative test samples. Now to check our experiment in a real-world scenario we have collected all those separated Gait and PPMI data from test set without label. Then to check model performance only for Gait data, we have randomly selected 5 positive (PD patients) and 5 negative (HC patients) PPMI data from train set and combine them individually against a single Gait test sample which will generate 10 (Gait + PPMI) samples for each gait test samples. Now those generated data will act as anchor in triplet model along with the positive and negative samples are taken from training set of combined Parkinson's disease patient data and Healthy Control patient data records respectively. So, there will be 10 triplet model output against 10 set of triplet inputs.

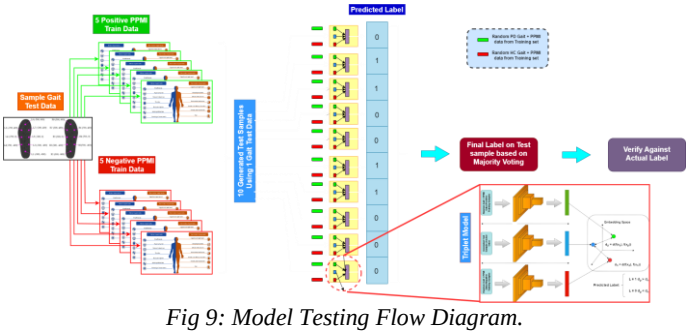
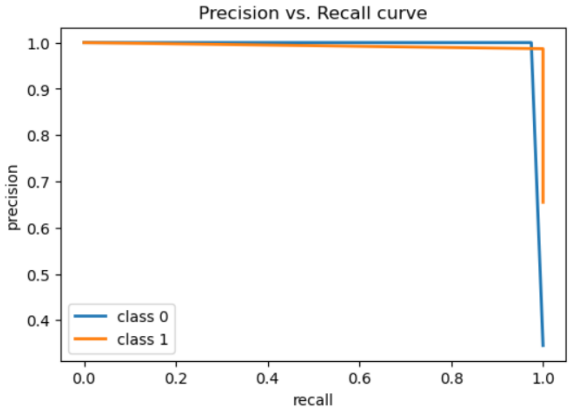
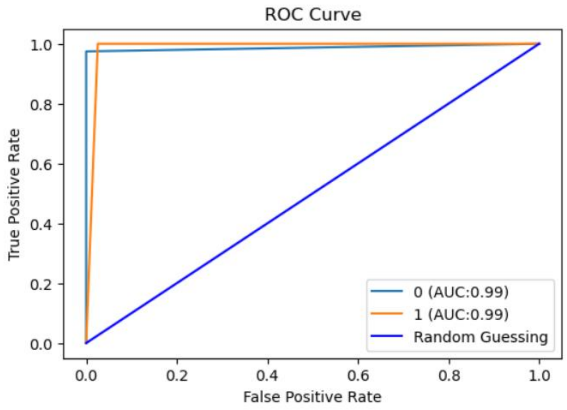


Fig 9: Model Testing Flow Diagram.

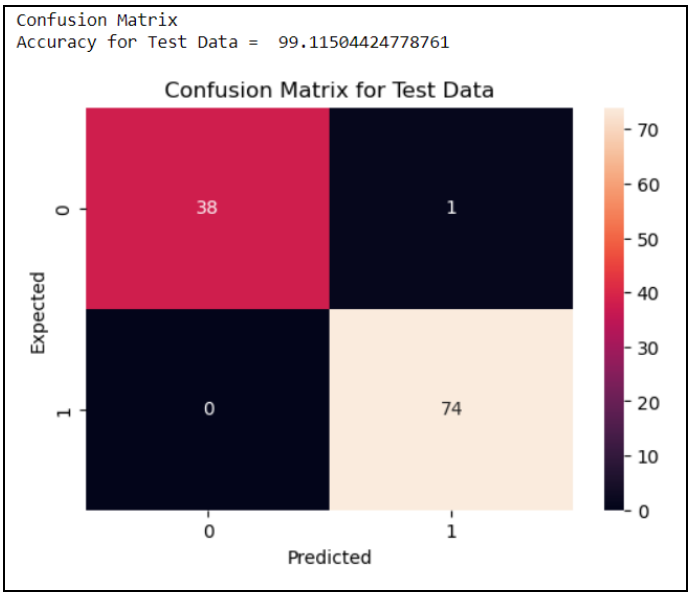
Here the model will calculate distance similar to triplet loss model, but the only difference is that it will generate output as 1 if distance between anchor and positive is less than distance between anchor and negative and 0 if the opposite case happens. So, these 10-triplet model will generate 10 labels against each model, and we have considered majority voting technique to decide the predicted label on that particular Gait test data. We have followed similar technique for testing PPMI test data where a single PPMI test data is combined against 5 positive (PD patients) and 5 negative (HC patients) Gait training set sample and rest of the setup is same as Gait data testing model. And lastly, we have verified the predicted output against the actual label of Gait and PPMI test set data.



(a) Precision vs. Recall curve



(b) ROC Curve

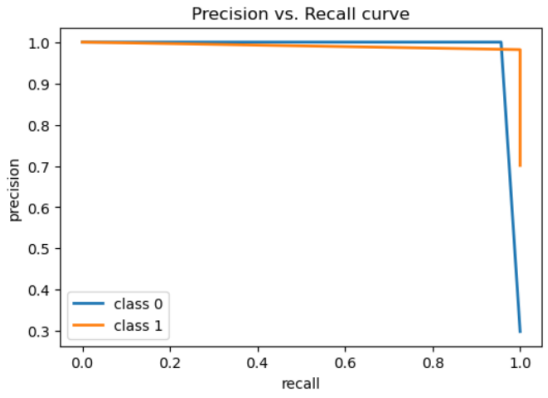


(c) Confusion Matrix

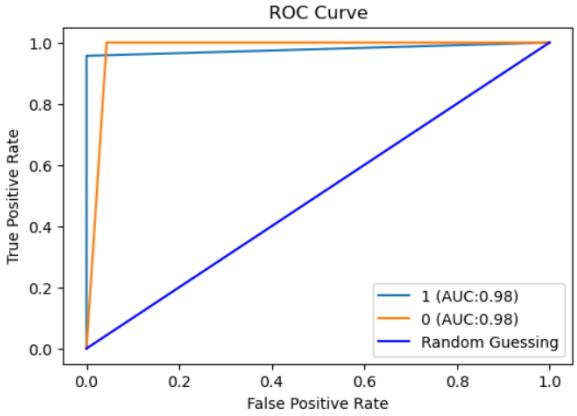
```
Classification Report
              precision    recall  f1-score   support

           0       1.00      0.97      0.99        39
           1       0.99      1.00      0.99        74

    accuracy                           0.99       113
   macro avg       0.99      0.99      0.99       113
weighted avg       0.99      0.99      0.99       113


Cohen Kappa Score
0.9803032944047412
```

(d) Classification report on PPMI test data.

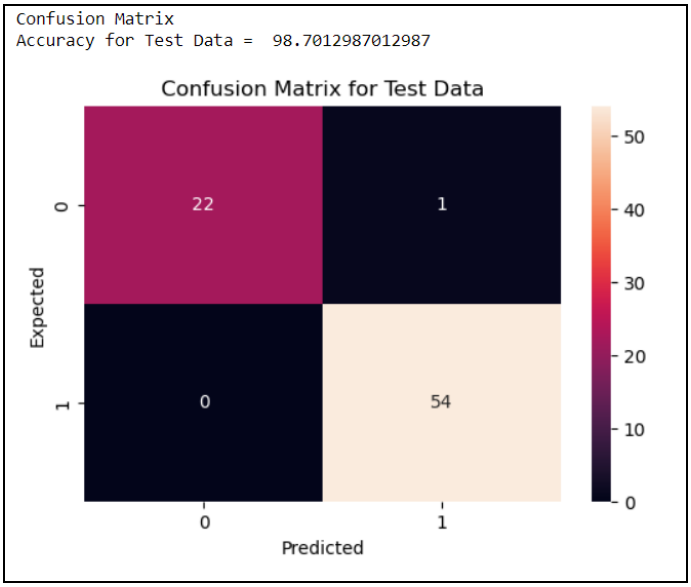
Fig 10: Gait Test Data Results



(a) Precision vs. Recall curve



(b) ROC Curve



(c) Confusion Matrix

```
Classification Report
              precision    recall  f1-score   support

           0       1.00      0.96      0.98        23
           1       0.98      1.00      0.99        54

    accuracy                           0.99        77
   macro avg       0.99      0.98      0.98        77
weighted avg       0.99      0.99      0.99        77


Cohen Kappa Score
0.968609865470852
```

(h) Classification report on PPMI test data.
Fig 11. PPMI Test Data Results

For Gait test samples we have achieved an accuracy of 98.7%, Cohen kappa score of 0.9686 and ROC AUC score as 0.9782 along with sensitivity as 95% and specificity as 94%. On the other side, the model has produced accuracy of 99.23%, Cohen Kappa score of 0.9613 and ROC AUC score as 0.9864 with sensitivity as 98% and specificity as 97.36% for PPMI test set samples. So, if we consider our previous experiment on individual dataset using individual model it is evident that, proposed architecture has able to improve the accuracy along with Cohen Kappa score and ROC AUC value using those multi-modal features.

6. Conclusion

This experiment is an extension to the experiment on achieving higher accuracy and lower error using newly developed feature level fusion technique used aggressively in Computer Vision. We have tried to add more functionality to use features coming from different sub- jects. The unavailability of MRI, SPECT, f-MRI, CT images of patients on many occasions make the situation harder for a doctor to detect Parkinson's disease based on only clinical diagnosis. Our experiment shows that even if patient have only Gait data or only PPMI data, we can still detect if he/she has the disease or not. This will eventually help us to implement this facility in situations where some doctor has to diagnosis a patients based on several modality's data. In future This experiment can be extended towards detecting other serious disease where doctor faces problem of missing data (missing modalities). Also, there is scope of adding more PPMI clinical diagnosis data if a greater number of patient's records are available in PPMI database. We have successfully evaluated the possibility of achieving better accuracy using feature level fusion of several modalities where data comes from different subjects. The attempt to combine the features of different domain comes out to be a unique combination to find pattern in Parkinson's disease patient data. While the aim is for accurate prediction in medical domain it has also eliminate the need for having all kinds of information required for classifying a particular disease of a patient.

7. References

[1] Zhao A, Qi L, Li J, Dong J, and Yu H. A hybrid spatiotemporal model for detection and severity rating of Parkinson's disease from gait data. *Neurocomputing*, 2018.

[2] Cherubini and A. et al. Magnetic resonance support vector machine discriminates be- tween Parkinson's disease and progressive supranuclear palsy. In *International Parkin- son and Movement Disorder Society*, 2014.

[3] Anirudh Reddy Cingireddy, Robin Ghosh, Venkata Kiran Melapu, Sravanthi Joginipelli, and Tor A K wembe. Classification of Parkinson's disease using motor and non-motor biomarkers through machine learning techniques. *International Journal of Quantitative Structure Property Relationships*, 2022.

[4] G. et al. Du. Combined diffusion tensor imaging and apparent transverse relaxation rate differentiate Parkinson disease and atypical parkinsonism. In *American Journal of Neuroradiology*, 2017.

[5] Glaab E, Trezzi JP, Greuel A, Jäger C, Hodak Z, Drzezga A, Timmermann L, Tittgemeyer M, Diederich NJ, and Eggers C. Integrative analysis of blood metabolomics and pet brain neuroimaging data for Parkinson's disease. In *Neurobiology of Disease*, 2019.

[6] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., R. Mark, Stanley, and H. E. (2000). Physio bank, physio toolkit, and physio net: Components of a new research resource for complex physiologic signals. In *Proc. CVPR*, 2000.

[7] MS Hema, R Maheshprabhu, K Sudheer Reddy, M Nageswara Guptha, and V Pandimurugan. Prediction analysis for Parkinson disease using multiple feature se- lection and classification methods. *Multimedia Tools and Applications*, 2023.

[8] H. Hirschauer, T. J. and Adeli, Buford, and J. A. Computer-aided diagnosis of parkin- son's disease using enhanced probabilistic neural network. In *Journal of Medical Systems* 2015.

[9] Tong J, Zhang J, Dong E, and Du S. Severity classification of Parkinson's disease based on permutation-variable importance and persistent entropy. *Multimedia Tools and Applications*, 2021.

[10] Açici K, Erdas CB, As¸urog lu T, Toprak MK, Erdem H, and Ogul H. A random forest method to detect Parkinson's disease via gait analysis. *Engineering Applications of Neural Networks*, 2017.

[11] Charles Leger, Monique Herbert, and Joseph F X DeSouza. Non-motor clinical and biomarker predictors enable high cross-validated accuracy detection of early pd but lesser cross-validated accuracy detection of scans without evidence of dopaminergic deficit. *Frontiers in Neurology*, 2020.

[12] D. et al. Long. Automatic classification of early Parkinson's disease with multi-modal imaging. In *PLOS ONE*, 2012.

[13] Noella N, Gupta D, and Priyadarshini J. Diagnosis of Parkinson's disease using gait dynamics and images. *Proc Int Conf Recent Trends Adv Compute* 2019.

[14] Oliveira, F. P. M., Faria, D. B. and Costa, D. C. and Castelo-Branco, M., Tavares, and
J. M. R. S. Extraction, selection and comparison of features for an effective automated computer-aided diagnosis of Parkinson's disease based on [123i] spect images. In *European Journal of Nuclear Medicine and Molecular Imaging*, 2018.

[15] Ghaderyan P and Fathi G. Inter-limb time-varying singular value: A new gait feature for Parkinson's disease detection and

[16] stage classification. measurement. *Multimedia Tools and Applications*, 2021.

[17] R Prashanth, Sumantra Dutta Roy, Pravat K Mandal, and Shantanu Ghosh. High- accuracy detection of early parkinson's disease through multimodal features and ma- chine learning. *International Journal of Medical Informatics*, 2016.

[18] Veeraragavan S, Gopalai AA, Gouwanda D, and Ahmad SA. Parkinson's disease diagnosis and severity assessment using ground reaction forces and neural networks. *Front Physiol*, 2020. Priya SJ, Rani AJ, and Ubendran N. Improving the prediction accuracy of Parkinson's disease based on pattern techniques. *5th International Conference on Devices, Circuits and Systems*, 2020.

[19] Priya SJ, Rani AJ, and Ubendran N. Improving the prediction accuracy of Parkinson's disease based on pattern techniques. *5th International Conference on Devices, Circuits and Systems*, 2020.

[20] As¸urog˘lu T, Açıcı K, Erdas¸ CB, Toprak MK, Erdem H, and Og˘ul H. Parkinson's disease monitoring from gait analysis via foot-worn sensors. *Biocybern Biomed Eng.*, 2018.

[21] Zeng W, Liu F, Wang Q, Wang Y, Ma L, and Zhang Y. Parkinson's disease classification using gait analysis via deterministic learning. *Neurosci Lett.*, 2016.

[22] Liu X and et al. A dual-branch model for diagnosis of Parkinson's disease based on the independent and joint features of the left and right gait. *Applied Intelligence*, 2021.

[23] Xia Yand Yao Z, Ye Q, and Cheng N. A dual-modal attention-enhanced deep learning network for quantification of Parkinson's disease characteristics. *IEEE Trans Neural Syst rehabil Eng.*, 2020.