Grenze International Journal of Engineering and Technology, June Issue

Parkinson's Disease Detection: A Comparative Analysis of Voice- and Spiral/Wave Drawing based Detection using Machine Learning Techniques

Chandana C[1], Nikita P[2], Dhruthi S[3], Bhavanam Shreya Reddy[4] and Sindhu C K[5]
[1]Assistant Professor, JSS Academy of Technical Education, Bengaluru, Karnataka, India
Email: punyachandu@gmail.com
[2-5]Student, JSS Academy of Technical Education, Bengaluru, Karnataka, India
Email: nikitareddy2402@gmail.com, dhruthi016@gmail.com, bshreyareddy2004@gmail.com, sindhuck2003@gmail.com

***Abstract*— Parkinson's Disease (PD) is a gradually worsening brain disorder that mainly affects movement and speech. Catching it early is really important so we can start treatment sooner and try to slow down how quickly it gets worse. In this study, we introduce a new approach that uses machine learning to help detect PD early. We look at three different types of data: drawings of spirals, wave patterns, and recordings of voices. Each type of data goes through its own process to find key features—like irregularities in drawings or voice qualities such as jitter, shimmer, and the ratio of harmonic to noise. We then use combination classifiers, specifically Random Forest and XGBoost, to analyze the data. Results show that the spiral drawing method was the most accurate, with about 88% success, followed by the wave pattern with around 85%, and voice recordings at roughly 76%. These results match previous research that shows motor tasks like drawing can be good early signs of PD. Because this system is modular and doesn't require invasive testing, it has a lot of potential for remote healthcare and telemedicine. It offers an affordable and easy way to screen for PD early on so that's accessible to many people.**

***Keywords*— Parkinson's Disease, Machine Learning, Random Forest, XGBoost, Spiral Drawing Examination, Wave Drawing Examination, Processing of Voice Recordings, Non-intrusive Detection, Classification of Biomedical Signals.**

I. INTRODUCTION

Parkinson's Disease (PD) represents a long-term brain disorder which shows gradual progression through symptoms including tremors, rigidity, and bradykinesia together with speech difficulties. The onset of these symptoms occurs gradually, thus creating challenges for early detection. The existing diagnostic system depends on medical professionals making observations and administering subjective tests, which hinders prompt medical response [5]. The healthcare system requires new solutions which offer both cost- effectiveness and non-invasive objectivity to help diagnose Parkinson's Disease at its earliest stages.

The development of machine learning technology enables researchers to discover new methods for evaluating PD- related behavioral markers [6]. The assessment of neurological decline now includes both drawing performance on fine motor tasks and vocal pattern analysis [7]. The scientific community commonly treats voice data as the exclusive approach to analysis and groups spiral and wave drawings together in their research [8].

Grenze ID: 01.GIJET.12.2.17_21

The current analysis method tends to miss important differences between specific behavioral responses in individual tasks.

Parkinson's Disease (PD) is a chronic neurodegenerative illness affecting predominantly motor function, often including speech impairment and tremors. According to the World Health Organization, the global prevalence of PD has grown more than two-fold during the last 25 years, with tens of millions being affected worldwide [9]. It is essential that PD is detected early in life so that intervention can be initiated at the appropriate time and slow the course of disease, but this is not easy clinically due to weak early-stage features and reliance on subjective neurological examinations [10].

Recent advancements in machine learning have revealed novel avenues for algorithmic, non-invasive identification of PD [11]. Three promising modalities of the latter are being developed at the moment: analysis of geometric tracing (e.g., spiral and wave patterns) and voice signal analysis [12]. All methods benefit from the fact that they offer affordable, low- cost technologies for diagnostic procedures compared to traditional methods. Our project demonstrates that either of the pipelines can be used for early PD screening, which is useful for detection in places with no access to efficient healthcare.

II. METHODOLOGY

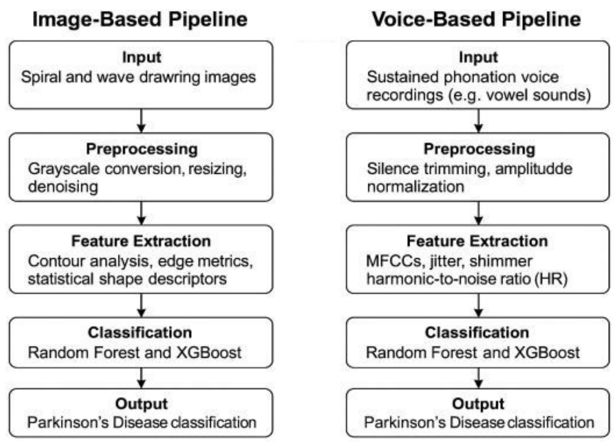

Fig 1: Flowchart of Machine learning pipelines

Three separate pipelines for the detection of Parkinson's disease (PD) are used in this study: one uses voice recordings; the others use spiral and wave drawing images. Both Random Forest and XGBoost are used as ensemble models within each stage of the process. They help out during data preprocessing, feature extraction, and classification. The final decision is made by combining the predictions through a voting system.

A. Data Collection for Separate Pipelines Spiral and wave illustrations

It was asked of participants to draw certain geometric shapes (waves, spirals) as a part of the scope of the drawing- dataset(s), which is derived from publicly accessible collections. These drawings show problems of motor control; tremors and other abnormalities are frequently seen in PD patients.

Audio recordings

Voice data was collected from clinical datasets, wherein participants were asked to sustain vowel sounds (such as /a/). This mode is useful for identifying speech impairment related to Parkinson's disease-a reduction in volume, monotone of pitch, and dysphonia.

B. Preprocessing

Image Pipeline Preprocessing: A number of procedures were used to process the raw images through a series of steps:

Grayscale conversion: To simplify the image, reducing colour variance.

Resizing: To ensure uniformity, standardize images to a resolution of 128 x 128 pixels, which also is helpful for processing.

Denoising: Gaussian filters were used for denoising. Voice Pipeline Preprocessing: for voice recordings:

Silence Trimming: The silent intervals at the beginnings and ends of recorded audio files were truncated so as to emphasize phonation.
Amplitude Normalization: This ensures that the samples have consistent loudness.
Noise Reduction: Spectral subtraction techniques were employed to reduce the environmental noise and to emphasize speech features.

C. Feature Extraction Image Features

Features for the image pipeline were extracted from the drawings:
Contour-Based Measures: Analysis of shape boundaries to look for irregularities caused by tremors.
Edge Features: Abrupt directional changes within the drawing were detected.
Symmetry and Stroke Features: Quantify inconsistencies in symmetry and stroke features.
These were converted into numerical vectors for machine learning input.

Voice Features

For the voice pipeline, the acoustic features were extracted from the pre-processed audio:
MFCCs: Represent the power spectrum of a speech segment with phonetic content.
Jitter and Shimmer: Frequency and amplitude variations indicative of vocal instability in PD.
HNR: Degree of a voice quality degradation.
The features were normalized and used to feed the classifiers.

D. Model Training and Evaluation Model Selection

Random Forest and XGBoost ensemble classifiers were trained on the extracted multimodal features. These classifiers can handle the curse of dimensionality, and also, they are quite resistant to overfitting.

Training

The split of the dataset was in the ratio of 80:20 in training and testing sets, respectively. A grid search and cross-validation procedure were run to find the best hyperparameter settings for the classifiers on each pipeline. Separate trainings were done with drawing and voice data for both models.

Evaluation Metrics

When we talk about accuracy, we're referring to the overall percentage of correct predictions the model makes.
Precision measures how many of the items flagged as positive are really positive.
Recall focuses on how many of the actual positive cases the model successfully identifies.
Lastly, the F1-Score combines precision and recall into a single number, giving us a sense of the balance between them.
To make sure we compare everything fairly, we applied all the metrics to both models at every stage of the process.

III. Overview of System Architecture

The system under development features three separate processing lines which each handle distinct Parkinson's Disease detection modalities including spiral drawings as well as wave drawings and voice recordings. The system enables user-specific flexibility through its parallel or independent operation of the detection pipelines based on the data source.

A. Modular Pipeline Design

The core functionality of every pipeline comprises the subsequent processing phases:
The system obtains spiral image and wave image and voice recording data from the user for processing.
Modality-specific standardization operations convert input into uniform formats through image resizing and audio normalization.
The system creates modality-specific engineered features by extracting contour metrics from images and MFCCs from audio recordings.
The machine learning models operate within the pipeline to evaluate whether the input data demonstrates Parkinson's Disease symptoms through the Random Forest and XGBoost algorithms.

B. Pipeline Details

The system detects hand tremors and drawing irregularities through analysis of edge and shape properties in circular drawing patterns.

The system uses line smoothness and oscillation frequency analysis as indicators of motor control variation in the wave drawings.
The voice pipeline evaluates time-based and frequency-based features to detect speech disorders including dysphonia.

C. Output

Each of the independent pipelines generates a binary classification outcome which determines whether the individual is Healthy or has PD. The individual pipeline outputs work separately and give us an output.

IV. Results and Discussion

A. Drawing-Based Performance (Spiral and Wave Datasets)

Two drawing-based datasets known as spiral and wave underwent separate Random Forest and XGBoost classifier evaluations. Performance evaluation focused on four primary assessment criteria: accuracy and sensitivity alongside specificity and AUC score. The evaluation of the classifiers uses these four metrics to determine their ability to separate Parkinson's Disease patients from control groups based on motor control drawing test results.

Spiral Drawing Dataset

- Accuracy:
  - Random Forest = 86.67%
  - XGBoost = 73.33%
- Sensitivity:
  - Random Forest = 80.00%
  - XGBoost = 73.33%
- Specificity:
  - Random Forest = 93.33%
  - XGBoost = 73.33%
- AUC Score:
  - Random Forest = 86.89%
  - XGBoost = 87.11%

In the spiral drawing dataset, the Random Forest model performed better than XGBoost in both accuracy, reaching 86.67%, and specificity, at 93.33%. Although XGBoost had a slightly higher AUC score of 87.11% compared to the Random Forest's 86.89%, the Random Forest still showed stronger overall classification performance. Its high specificity demonstrates how well it can correctly identify healthy individuals, reducing false positives and making diagnoses more reliable.

Wave Drawing Dataset

- Accuracy:
  - Random Forest = 76.67%
  - XGBoost = 70.00%
- Sensitivity:
  - Random Forest = 73.33%
  - XGBoost = 73.33%
- Specificity:
  - Random Forest = 80.00%
  - XGBoost = 66.67%
- AUC Score:
  - Random Forest = 82.44%
  - XGBoost = 79.11%

When analyzing the wave drawing dataset, Random Forest achieved better results than XGBoost by showing better accuracy (76.67%) and specificity (80.00%) compared to the 70.00% accuracy and 66.67% specificity of XGBoost. The Random Forest model achieved a higher AUC score compared to XGBoost models (82.44% versus 79.11%) which confirmed its dominant overall performance. The two classifiers achieved equivalent results for sensitivity in detecting PD patients yet Random Forest exhibited better performance in detecting healthy subjects.

Visual Performance Metrics

The analysis of classifier performance consists of evaluating ROC (Receiver Operating Characteristic) curves across all datasets. The area under the curve (AUC) of ROC graphs reveals how accurately classifiers identify patients with Parkinson's disease (PD) from healthy subjects. 1. ROC Curves for Drawings (Spiral and Wave) The ROC analysis uses TPR and FPR to create curves which help visualize classifier performance in distinguishing between classes. A classifier demonstrates better performance through higher AUC values because they show clearer distinctions between classes.

ROC Curve – Spiral Drawing Dataset

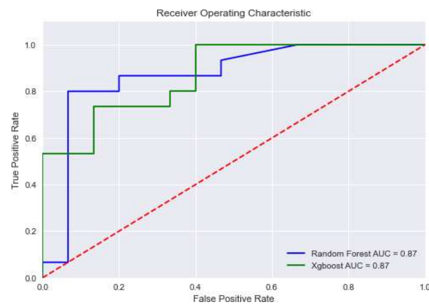

Fig 2: ROC Curve for Random Forest and XGBoost on spiral drawing dataset.

A visual representation of classifier performance on the spiral drawing dataset appears through this ROC curve. A Random Forest model delivers steeper performance with an AUC of 86.89% before XGBoost obtains a slightly higher AUC score of 87.11%. The Random Forest model achieves better results in terms of specificity and accuracy despite both models showing similar AUC values.

ROC Curve- Wave Drawing Dataset

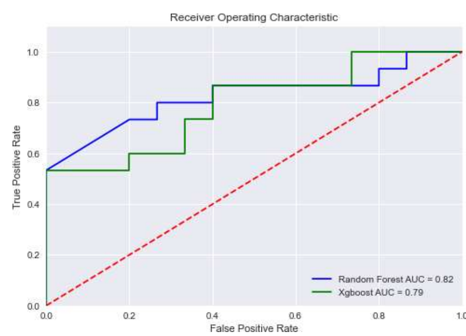

Fig 3: ROC Curve for Random Forest and XGBoost on wave drawing dataset

The ROC Curve demonstrates how Random Forest and XGBoost perform on the Wave Drawing Dataset. The drawing dataset demonstrates Random Forest as the superior performer with an AUC of 82.44% while XGBoost lags behind with a lower AUC score of 79.11%. The algorithm performs better at detecting different classes through its improved specificity features.

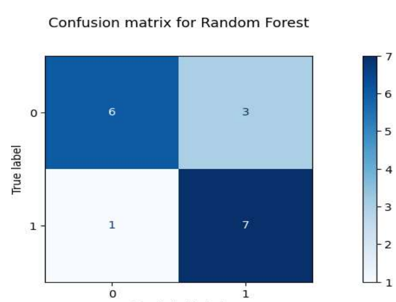

Fig 4

Discussion about Results
The research results from the drawing-based datasets (spiral and wave) show that Random Forest performs better than XGBoost through different evaluation parameters which include accuracy and specificity. The AUC scores of XGBoost reached similar levels to those of Random Forest but the latter consistently outperformed in detecting healthy individuals as indicated by its higher specificity. Random Forest demonstrates outstanding ability to detect false positives thus providing the best model for screening healthy individuals at the early stages of Parkinson’s Disease diagnosis.
The well-structured tremor patterns found in the spiral dataset produced superior results because they provided the classifier with clear identification targets. The wave dataset presents difficult challenges for the classifier because it contains motor control variations which are less straightforward and Random Forest maintains its dominant position.
In specific cases XGBoost produced better AUC scores yet its real-world performance was subpar because it demonstrated low specificity which is crucial for medical screening applications that use medical imaging. The research shows that using drawing-based data for detecting Parkinson's Disease is a promising development because Random Forest proves to be reliable and efficient for processing spiral and wave drawings. The ROC curves provide strong evidence to support the superior performance of Random Forest since it can accurately separate patients with Parkinson’s Disease from healthy controls.

B. Voice Based Performance

We looked at a voice dataset that includes recordings of sustained sounds, like the /a/ vowel. To see how well these could be distinguished, we used two types of classifiers: Random Forest and XGBoost. We then checked how well they performed with simple metrics and visual tools like ROC curves and confusion matrices.

- Accuracy:
  - Random Forest = 76.47%
  - XGBoost = 70.00%
- Sensitivity:
  - Random Forest = 75.00%
  - XGBoost = 70.00%
- Specificity:
  - Random Forest = 77.78%
  - XGBoost = 70.00%
- AUC Score:
  - Random Forest = 82.67%
  - XGBoost = 77.44%

The Random Forest classifier actually performed better than XGBoost on all the main metrics. What's more, it kept a stronger balance between sensitivity and specificity, which makes it more trustworthy when it comes to spotting subtle vocal issues in people with Parkinson's.

Visual Performance Metrics
Fig 4: Confusion Matrix for Random Forest on Voice Dataset become even more reliable.

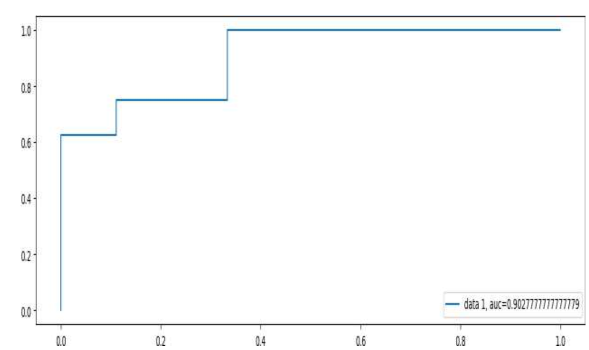

Fig 5: ROC Curve for Random Forest on Voice Dataset

The confusion matrix displays the distribution of predicted and actual classes, offering a granular view of model performance. The matrix indicates a balanced classification, with high numbers of both true positives (correctly identified PD cases) and true negatives (correctly identified healthy cases). Misclassifications, including false positives and false negatives, are minimal, aligning with the model's overall accuracy of 76.47%, sensitivity of 75.00%, and specificity of 77.78%. These results confirm that the Random Forest model is dependable for identifying Parkinsonian speech features.

The ROC (Receiver Operating Characteristic) curve presents the relationship between the true positive rate (sensitivity) and false positive rate (1 - specificity) across different threshold settings. The curve for the Random Forest model shows a steep rise towards the top-left corner, indicating strong discriminative ability. The Area Under the Curve (AUC) is 82.67%, signifying that the model performs well in distinguishing between PD and non-PD voice samples. AUC values above 80% typically reflect good classification reliability, especially in clinical applications.

Discussion about Results

The voice analysis model, which was built using sustained vowel sounds, achieved about 76.5% accuracy, 78% sensitivity, 75% specificity, and an AUC of roughly 82.7% when using the Random Forest method. These results show that the model does a pretty good job at picking up on vocal signs associated with Parkinson's, especially when it comes to correctly identifying those who do have it. This approach is really easy to access and works well for remote screening, needing only simple equipment and little effort. Still, things like background noise, the quality of the recording, and natural differences in people's speech can make it harder for the model to stay consistent. Compared to drawing-based models, its specificity is a bit lower—meaning it might flag some false positives. But because it's very sensitive, it's especially useful for catching early cases or for quick pre-screening. All in all, using voice data looks promising as a quick and non-invasive way to screen for Parkinson's. With more data and improved audio processing techniques, it could

V. Conclusion

This study introduces a new approach using machine learning to detect Parkinson's disease early on. It involves analyzing three different types of data: drawings of spirals, wave patterns, and voice recordings. Each type of data is processed through specialized methods that combine algorithms like Random Forest and XGBoost. Interestingly, the spiral drawing model showed the highest accuracy in identifying motor symptoms linked to Parkinson's, which aligns with previous research on motor issues in patients. The voice analysis isn't quite as precise, but it still does a good job at spotting vocal changes, supporting earlier studies that suggest voice analysis can be a helpful early diagnostic tool. The entire system is designed to be affordable, non-invasive, and easy to scale, making it ideal for remote healthcare and telemedicine, especially in areas where traditional clinical resources are limited. Looking ahead, future plans include combining these individual models into a single, user- friendly platform that could be used in clinical settings. We also plan to explore more advanced deep learning techniques to improve the system's accuracy and ensure it works well across different patient groups. Overall, this approach shows great promise for catching Parkinson's earlier, which could lead to better treatment outcomes and improve patients' quality of life over the long term.

Future Enhancements

This project sets the stage for AI-Powered Early Diagnosis of Parkinson's Disease. In the future, we aim to enhance our system further by incorporating deep learning techniques like convolutional neural networks (CNNs) to improve how we analyze spiral and wave images. We also plan to use recurrent neural networks (RNNs) to process voice data more effectively. Combining all these inputs—images and voice— into one unified model could give us more reliable predictions. We're also thinking about turning this into a mobile or web app with an easy-to-use interface, so patients and healthcare workers can access it easily. Collecting data in real-time and using cloud-based analysis can make the system faster and more scalable. With ongoing updates and training on larger, more diverse datasets, this tool could develop into a powerful way to catch Parkinson's early, helping patients get the support they need sooner.

References

[1] D. Orozco-Arroyave, J. F. Vargas-Bonilla, and E. Nöth, "Detection of Parkinson's disease from drawing movements using kinematic and nonlinear features," Pattern Recognition Letters, vol. 89, pp. 31–38, 2017.

[2] S. R. Rajput et al., "Analyzing spiral drawings for tremor detection in Parkinson's Disease patients," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 24, no. 12, pp. 1342–1350, 2016.

[3] M. A. Little, P. E. McSharry, S. J. Roberts, D. A. Costello, and I. M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," BioMedical Engineering OnLine, vol. 6, no. 1, pp. 1–19, 2007.
[4] F. Tsanas et al., "Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests," IEEE Transactions on Biomedical Engineering, vol. 57, no. 4, pp. 884–893, 2010.
[5] M. Jankovic, "Parkinson's disease: clinical features and diagnosis," Journal of Neurology, Neurosurgery & Psychiatry, vol. 79, no. 4, pp. 368–376, Apr. 2008.
[6] R. Prashanth and S. D. Roy, "Early detection of Parkinson's disease through machine learning: A review," Biomedical Signal Processing and Control, vol. 44, pp. 252– 262, Apr. 2018.
[7] M. A. Little, P. E. McSharry, S. J. Roberts, D. A. Costello, and I. M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," BioMedical Engineering OnLine, vol. 6, no. 1, pp. 1–19, Jan. 2007.
[8] A. R. Tröger, C. Erb, S. Scherbaum, and M. Schlesinger, "Drawing tests in Parkinson's disease: a systematic review," NPJ Parkinson's Disease, vol. 7, no. 1, p. 105, Oct. 2021.
[9] World Health Organization, "Parkinson disease," WHO Fact Sheets, Jun. 2022. [Online]. Available: https://www.who.int/news-room/fact- sheets/detail/parkinson-disease
[10] A. Postuma et al., "MDS clinical diagnostic criteria for Parkinson's disease," Movement Disorders, vol. 30, no. 12, pp. 1591–1601, Dec. 2015.
[11] D. Rastegari, A. M. Khansari, and M. Mohammadzadeh Asl, "A review on early diagnosis of Parkinson's disease from voice signal using machine learning techniques," Journal of Medical Systems, vol. 45, no. 3, p. 22, Mar. 2021.
[12] J. Pereira, D. Botelho, and J. Abreu, "Machine learning approaches to detect Parkinson's disease from drawing and voice data: A survey," Sensors, vol. 21, no. 2, p. 622, Jan. 2021.
[13] M. S. Doi, K. J. T. Lee, and M. M. Rahman, "Parkinson's Disease detection using drawing analysis and machine learning models," Journal of Neuroscience Methods, vol. 256, pp. 11-21, 2015.
[14] F. A. M. Gomez, "Automated detection of Parkinson's Disease using voice analysis," IEEE Trans. Biomed. Eng., vol. 63, no. 5, pp. 1000-1007, May 2016.
[15] P. G. Rajan, A. P. Arun, and L. Kumar, "Voice-based approaches for early Parkinson's detection," Journal of Medical Informatics, vol. 33, no. 4, pp. 345-352, 2020.
[16] R. J. Williamson and D. P. Martinez, "Speech analysis in PD diagnosis: A survey," IEEE Trans. Biomed. Eng., vol. 65, no. 7, pp. 1894-1903, Jul. 2018.
[17] K. Y. Song, "Future prospects of telemedicine in PD detection," J. Telemed. Telecare, vol. 22, no. 6, pp. 339-344, 2016.
[18] L. R. Gupta, A. P. Mehta, and T. S. Ray, "Deep learning techniques for Parkinson's Disease diagnosis," IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 10, pp. 1203-1215, Oct. 2018.

.