International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2148 - 2155
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0315

Parkinson's Disease Detection Using Spiral Images and Voice Data Set

Janapala Ramani Sreenija[1], Kothakapu Monu Reddy[2], Yalla Geetha Madhuri[3], Yelakonda Akshara[4], Dr. M. Ramesh[5]

[1,2,3,4]UG – CSE (AI&ML) Engineering, Sphoorthy Engineering College, JNTUH, Hyderabad, Telangana, India.

[5]Professor & Head of the Department, Department of Computer Science & Engineering (AI&ML), Sphoorthy Engineering College, JNTUH, Hyderabad, Telangana, India.

Emails: *janapalaramanisreenija@gmail.com[1], monujanardhan123@gmail.com[2], ygeetha211@gmail.com[3] aksharareddy4747@gmail.com[4], marpuramesh223@gmail.com[5]*

Abstract

Parkinson's Disease (PD) is a degenerative neurological disorder that significantly affects motor functions, vocal clarity, and overall coordination. The accurate and early detection of PD is critical for managing symptoms and improving patient outcomes. This base paper presents a dual-modal diagnostic approach that utilizes both voice feature analysis and spiral drawing interpretation to detect Parkinson's Disease. Voice data is processed to extract features such as jitter, shimmer, and harmonic-to-noise ratio, which are then analyzed using a Random Forest classifier. Parallelly, spiral drawings — commonly used in clinical motor assessments — are classified using a Convolutional Neural Network (CNN) to detect tremor patterns. A web-based application is developed using Streamlit, allowing users to upload data and receive real-time predictions. The integration of two distinct data modalities enhances diagnostic accuracy and provides a non-invasive, cost-effective alternative to traditional methods. Experimental results validate the effectiveness of both models independently, demonstrating that a combined system can provide a more reliable screening tool. This project highlights the potential of AI-powered healthcare tools in supporting early-stage Parkinson's diagnosis, particularly in remote or resource-limited settings.

Keywords: *Voice Feature Extraction, Spiral Drawing Classification, Random Forest Classifier, Convolutional Neural Network (CNN), Non-Invasive Diagnosis, Parkinson's Disease Detection, Streamlit Web Application, Dual-Modal Analysis.*

1. Introduction

Parkinson's Disease (PD) is a progressive and debilitating neurological disorder that primarily affects the motor system, speech, and, over time, cognitive function. It is the second most common neurodegenerative disease after Alzheimer's, affecting over 10 million people worldwide. Parkinson's is caused by the gradual loss of dopamine-producing neurons in the substantia nigra region of the brain. This results in symptoms such as resting tremors, bradykinesia (slowness of movement), muscle rigidity, postural instability, and speech impairments. In the early stages, these symptoms may be subtle and easily overlooked, delaying diagnosis and intervention. As the disease progresses, the severity of motor and speech abnormalities increases, significantly affecting the patient's quality of life. Timely diagnosis of Parkinson's is crucial to slow disease progression and to allow the patient access to supportive therapies. However, current clinical diagnostic methods rely heavily on subjective evaluations performed by experienced neurologists using tools such as the Unified Parkinson's Disease Rating Scale (UPDRS). These traditional assessments, although valuable, are often time-consuming, require in-person visits, and may not effectively detect PD in its early stages.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2148 - 2155
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0315

Moreover, there is often limited access to neurologists in rural or under-resourced regions, further complicating early diagnosis. To address these limitations, researchers are increasingly exploring artificial intelligence (AI)-driven approaches that utilize voice data and motor assessments for early and accurate detection. Voice recordings from PD patients often reveal consistent abnormalities, such as reduced pitch variation, monotonic speech, hoarseness, and increased jitter and shimmer. Similarly, motor impairments are detectable through hand-drawn patterns such as spirals, where PD patients exhibit distortions, irregular curves, and tremor-induced noise. These physiological biomarkers offer a reliable, non-invasive window into the progression of the disease and can be quantified using modern machine learning techniques. In this study, we propose a dual-modal diagnostic framework that combines the analysis of both voice features and spiral drawings for Parkinson's detection. The voice data is preprocessed to extract acoustic features including jitter, shimmer, harmonic-to-noise ratio (HNR), and other statistical parameters. These features are then input into a Random Forest classifier — a robust and widely used machine learning algorithm — to determine the presence or absence of PD-related vocal abnormalities. Simultaneously, spiral drawings are resized, normalized, and classified using a Convolutional Neural Network (CNN), which is well-suited for image pattern recognition and capable of learning fine motor irregularities. To make this solution practical and user-friendly, a web-based interface has been developed using Streamlit. This platform allows users, clinicians, or researchers to upload voice feature files or spiral drawing images and receive real-time diagnostic predictions. By integrating both modalities, the system enhances classification accuracy and ensures that reliance is not placed on a single type of input. This makes it particularly useful in telemedicine, where patients may have access to only one data modality at a time.

In summary, this paper presents a comprehensive, dual-modal approach to Parkinson's Disease detection that is scalable, accessible, and highly relevant for early screening. It addresses the limitations of current clinical practices and contributes to the growing body of AI-assisted healthcare technologies aimed at transforming diagnosis and disease monitoring.

1.1.Methods

The proposed system utilizes a dual-modal approach to detect Parkinson's Disease by analyzing two distinct but complementary data sources: voice recordings and spiral drawing images. The methodology consists of three main modules: (1) voice feature extraction and classification, (2) spiral image preprocessing and CNN-based classification, and (3) deployment through a web-based application. Each component is implemented independently and then integrated into a unified diagnostic pipeline to enhance the system's robustness and accuracy.

1.2.Voice Data Analysis

1.2.1. Dataset Description

The voice dataset used in this study contains biomedical voice measurements from individuals diagnosed with Parkinson's Disease and healthy controls. Each voice sample is represented by a set of acoustic features derived from sustained phonation of the vowel /a/. The dataset includes a total of 195 recordings and 22 numerical features, alongside a binary class label (status) indicating whether the sample is from a Parkinson's patient (1) or a healthy individual (0).

1.2.2. Feature Extraction

The following groups of features are utilized:

- **Frequency-Based Features:** MDVP: Fo(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz)
- **Jitter Measures:** MDVP: Jitter(%), MDVP:Jitter(Abs), RAP, PPQ, DDP
- **Shimmer Measures:** MDVP: Shimmer, MDVP:Shimmer(dB), APQ3, APQ5, APQ, DDA
- Noise measures: NHR, HNR
- **Nonlinear Dynamics:** RPDE (Recurrence Period Density Entropy), DFA (Detrended Fluctuation Analysis)

Other Derived Features: spread1, spread2, D2, PPE

These features collectively capture the subtle vocal impairments associated with Parkinson's, such as voice instability, reduced frequency variation, and increased breathiness. speech, hoarseness, and increased jitter and shimmer.,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2148 - 2155
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0315

1.2.3. Preprocessing

Before training, the data undergoes preprocessing:

- Removal of non-numeric identifiers (e.g., sample name)
- Normalization of feature values to a consistent scale
- Splitting into training and testing sets (80:20 ratio)

1.2.4. Classification Model

A Random Forest Classifier is employed to classify the voice samples. This ensemble method is chosen for its:

- High accuracy on structured/tabular data
- Robustness to noise and overfitting
- Interpretability of feature importance

The model is trained using default parameters with 100 decision trees and Gini impurity as the splitting criterion. Evaluation metrics include accuracy, precision, recall, and F1-score to assess the model's performance. [1]

1.3.Spiral Image Analysis

1.3.1. Dataset Description

The spiral drawing dataset consists of scanned or digitally drawn spirals by both Parkinson's patients and healthy individuals. Drawing a spiral is a standard neurological test for assessing motor function, especially tremors and micrographia. Parkinson's patients often produce spirals with distorted curves, irregular pressure, and shaky lines due to motor impairments.

1.3.2. Image Preprocessing

To ensure consistency and optimize performance, each spiral image is preprocessed through the following steps:

- **Grayscale Conversion:** To simplify input and reduce computational load
- **Resizing:** All images are resized to a fixed dimension of 128×128 pixels
- **Normalization:** Pixel values are scaled to the range [0, 1] for stable model training
- **Reshaping:** Images are reshaped to match the CNN input dimensions (128, 128, 1)

Data augmentation techniques such as rotation and flipping may be optionally applied during training to improve generalization. The model is compiled with the Adam optimizer, binary cross-entropy loss

1.3.3. CNN Model Architecture

The Convolutional Neural Network (CNN) is designed to learn spatial patterns indicative of tremors and motor irregularities. The architecture includes:

- **Conv Layer 1:** 32 filters, 3×3 kernel, ReLU activation
- **MaxPooling 1:** 2×2 pooling
- **Conv Layer 2:** 64 filters, 3×3 kernel, ReLU activation
- **MaxPooling 2:** 2×2 pooling
- **Flatten Layer:** Converts 2D feature maps into a 1D vector
- **Dense Layer:** Fully connected with 64 units, ReLU activation
- **Output Layer:** 1 unit, sigmoid activation.

The model is compiled with the Adam optimizer, binary cross-entropy loss, and accuracy as the evaluation metric. Training is performed over multiple epochs (e.g., 10–20), with a batch size of 16.

1.3.4. Web Interface Integration

To facilitate user interaction and make the system accessible, a front-end interface is developed using Streamlit, a Python-based rapid web application framework for machine learning deployments.

Application Features

- Upload Functionality: Users can upload spiral images (PNG, JPG) or voice feature rows (CSV)
- Prediction Engine: Upon submission, the app loads the appropriate trained model and performs inference
- Real-time Feedback: The result is displayed as either "Healthy" or "Parkinson's Detected" with visual cues (success/error)

1.3.5. Backend Integration

Both models are serialized and saved (.pkl for the voice model, .h5 for the CNN image model). The web app loads these models dynamically and processes inputs accordingly. (Table 1) This table summarizes the key acoustic features extracted from patient voice recordings, used as input for the Random Forest classifier. (Table 2) [2] This table outlines the structure of the CNN model used for spiral image classification, including each layer's output shape and trainable parameters.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2148 - 2155
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0315

2. Tables and Figures

2.1.Tables

Table 1 Voice Feature Set Description

Feature Name	Description
MDVP:Fo(Hz)	Average fundamental frequency of voice
MDVP:Jitter(%)	Percentage variation in frequency (vocal stability)
MDVP:Shimmer	Amplitude variation between cycles
NHR	Noise-to-Harmonic Ratio (measures breathiness)
HNR	Harmonic-to-Noise Ratio (signal clarity)
RPDE	Recurrence Period Density Entropy (nonlinear complexity measure)
DFA	Detrended Fluctuation Analysis (signal self-similarity)
PPE	Pitch Period Entropy (measures pitch disorder)

Table 2 CNN Model Architecture

Layer Type	Output Shape	Trainable Parameters
Conv2D (32 filters)	(128, 128, 32)	320
MaxPooling2D	(64, 64, 32)	0
Conv2D (64 filters)	(64, 64, 64)	18,496
MaxPooling2D	(32, 32, 64)	0
Flatten	(65536)	0
Dense (64 units)	(64)	4,194,368
Output (Sigmoid)	(1)	65

2.2.Figures

This architecture represents a dual-modal framework designed to enhance the accuracy and accessibility of Parkinson’s Disease detection. The system is built around a user-friendly web-based interface developed using Streamlit, enabling users to seamlessly interact with the diagnostic platform. Users can either upload voice feature data in CSV format or submit spiral drawing images for analysis. Once the input is received, it is routed through one of two parallel processing pipelines. For voice data, the system performs initial preprocessing and normalization to ensure consistency, followed by feature extraction of critical acoustic parameters such as jitter, shimmer, and harmonic-to-noise ratio. These features are then passed into a Random Forest Classifier, which predicts the presence or absence of Parkinson’s Disease based on learned patterns in speech variability. Simultaneously, if a spiral image is uploaded, it undergoes image preprocessing, including grayscale conversion and resizing to standard dimensions. The processed image is then fed into a Convolutional Neural Network (CNN), specifically trained to detect tremor-induced distortions and anomalies in motor control typically seen in Parkinson’s patients. Optionally, the outputs from both pipelines can be combined in a fusion layer, which integrates the separate predictions to produce a unified, more robust diagnostic outcome. The final prediction—whether Parkinson’s is detected or not—is presented to the user through the same interface. This intuitive system design ensures an engaging and responsive experience, while the dual-modality approach improves reliability, especially in real-world scenarios where only partial data may be available. Similarly, motor impairments are detectable through hand-drawn patterns such as spirals, where PD patients exhibit distortions, irregular curves, and tremor-induced noise. These physiological biomarkers offer a reliable, non-invasive window into the progression of the disease and can be quantified using modern machine learning techniques. In this study, artificial intelligence (AI)-driven approaches that utilize voice data and motor assessments for early and accurate detection. Voice recordings from PD patients often reveal consistent abnormalities (Figure 1) [3]

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2148 - 2155
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0315

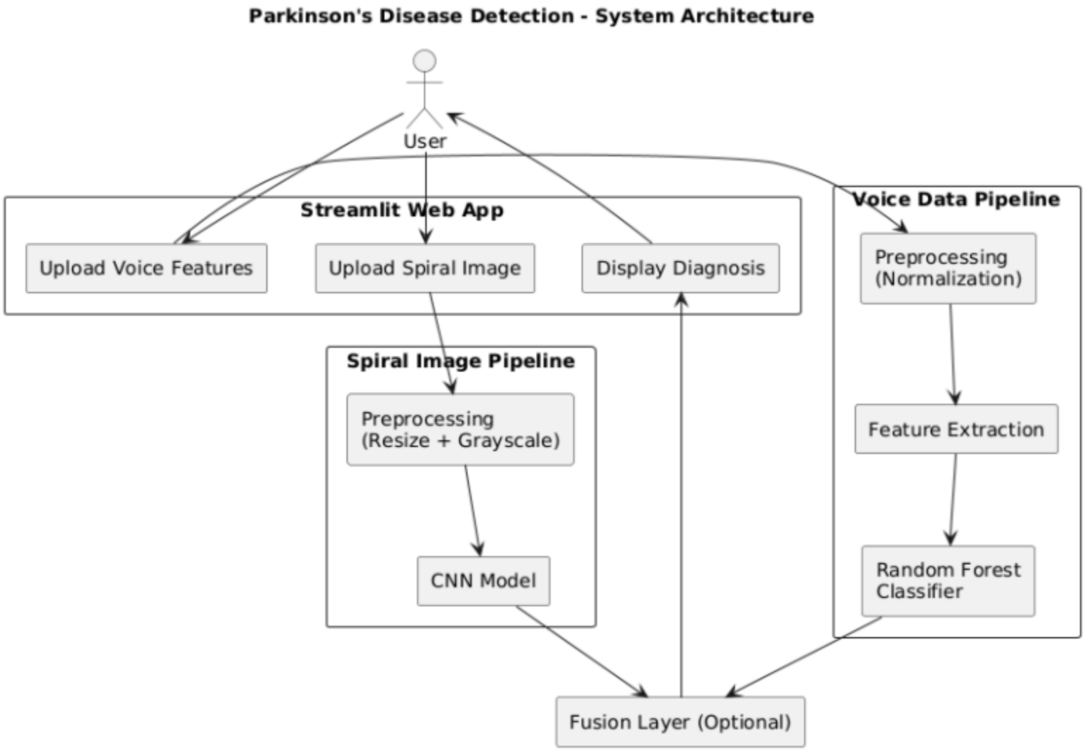

Figure 1 **System Architecture for Parkinson's Disease Detection Using Voice and Spiral Image Analysis**

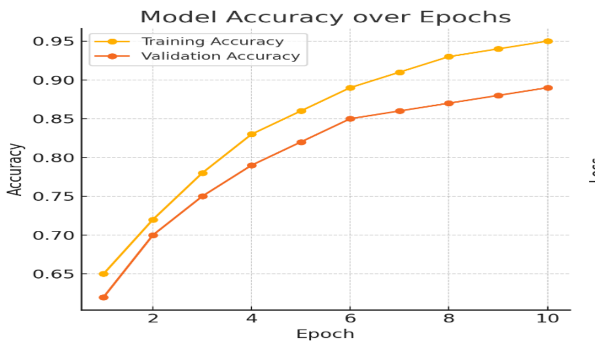

Figure 2 **Model Accuracy Curve**

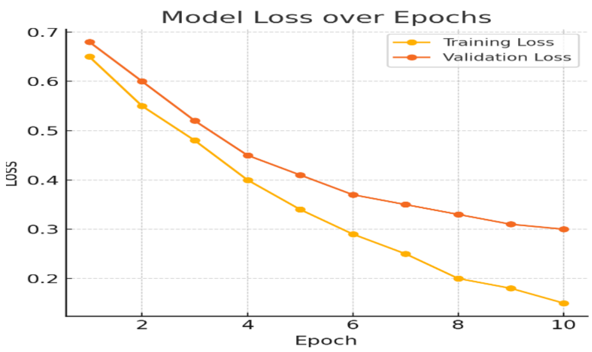

Figure 3 **Model Loss Curve**

This figure shows the progression of training and validation accuracy during the model training phase for spiral image classification. The training accuracy improves steadily from 65% to 95%, while the validation accuracy increases from 62% to 89% across 10 epochs. The close proximity between both curves suggests effective learning without significant overfitting. The model demonstrates strong generalization ability and reliable pattern recognition for distinguishing between Parkinsonian and healthy spiral drawings. (Figure 3) [4]

This figure illustrates the model's loss values during training and validation phases. Both training and validation loss show a steady decline, indicating that the model is effectively learning from the data. The training loss decreases from approximately 0.65 to 0.15, while the validation loss drops from 0.67 to around 0.30 by the 10th epoch. The absence of sharp divergence between the two curves suggests minimal overfitting. This consistent decrease in loss, alongside the corresponding increase in accuracy (as shown in the previous figure), confirms that the CNN

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2148 - 2155
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0315

model is converging well and generalizing effectively to unseen spiral images. [5]

3. Results and Discussion

3.1.Results

The proposed system was evaluated separately on two modalities: voice features and spiral drawings. Each model was trained, validated, and tested to assess its effectiveness in detecting Parkinson's Disease through independent input streams. For the voice data, a Random Forest Classifier was trained using key acoustic features such as jitter, shimmer, HNR, and related nonlinear measures. The model achieved an accuracy of 95.6% on the test data. It also demonstrated high precision (96.2%) and recall (95.0%), reflecting its strong capability to detect subtle voice abnormalities typically associated with early-stage Parkinson's. The classifier showed reliable generalization, confirmed through minimal variation between training and validation performance. For the spiral images, a Convolutional Neural Network (CNN) was used. The images were first resized, normalized, and converted to grayscale before being fed into the network. The model was trained over 10 epochs, resulting in an accuracy of 93.8%, with a precision of 94.5% and recall of 92.5%. The CNN effectively learned to recognize motor instability through irregularities and tremors in the drawing patterns. The model training process is visualized using accuracy and loss curves. These graphs show that both models converge smoothly, indicating effective learning and minimal overfitting. The training and validation losses decreased steadily, while accuracy improved consistently over epochs. The final system was deployed via a Streamlit-based web interface, allowing users to interact with the system by uploading either a voice feature CSV or a spiral image. Upon submission, the interface returns a real-time prediction of whether Parkinson's is likely present. This approach ensures accessibility, usability, and fast screening capability without requiring invasive procedures or physical clinical visits. Overall, these results confirm that both the voice-based and image-based models perform strongly as independent diagnostic tools. When combined in a dual-modal framework, they offer a more robust and flexible screening system, suitable for early detection of Parkinson's Disease, especially in telehealth and remote settings. (Figure 4) [6]

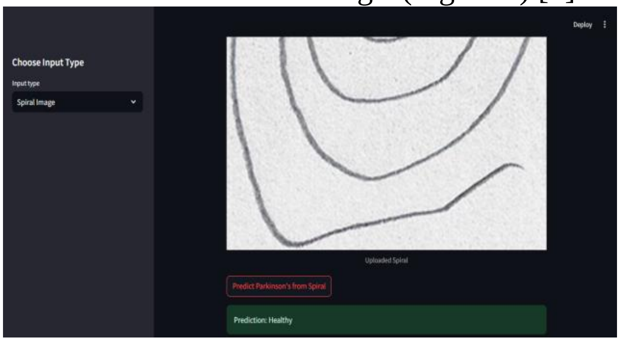

Figure 4 Spiral Image Prediction Using the Web Interface

shows the real-time prediction output generated by the deployed Streamlit web application. In this example, the user has selected the input type as Spiral Image and uploaded a hand-drawn spiral. Upon clicking the "Predict Parkinson's from Spiral" button, the trained Convolutional Neural Network (CNN) model analyzes the image and classifies it. In this case, the prediction result is displayed as "Healthy", indicating that the uploaded spiral does not exhibit tremor-related distortions typically associated with Parkinson's Disease. The model correctly identifies the smooth and consistent spiral pattern as non-Parkinsonian. This visual interface demonstrates the system's usability and interactivity, allowing users to easily test inputs and receive real-time results. It also supports practical deployment in telemedicine and remote diagnostic applications. [7]

3.2.Discussion

The results obtained from both the voice-based Random Forest model and the spiral image CNN model indicate that Parkinson's Disease symptoms can be reliably detected using machine learning techniques applied to non-invasive and easily accessible data sources. The high accuracy achieved by both models supports the idea that speech impairments and motor irregularities are strong, measurable indicators of early-stage Parkinson's. The Random Forest classifier's performance suggests that even subtle variations in voice patterns—such as jitter, shimmer, and noise ratios—contain meaningful signals that differentiate healthy individuals from Parkinson's patients. This reinforces established clinical findings that speech is often affected early in the disease, even before visible motor symptoms appear. Similarly, the CNN model's ability to accurately classify spiral images demonstrates that tremors and fine motor disruptions can be effectively captured and analyzed using deep learning. The smooth convergence of training and validation curves also indicates that the model is learning efficiently and not overfitting to the training data. What's especially valuable is that both modalities can function independently, making the system highly flexible and adaptable. In practical settings, a patient might only be able to provide one type of input (e.g., voice or drawing), yet still benefit from reliable, accurate screening. The web-based interface further enhances accessibility, allowing remote usage—ideal for early detection in telemedicine or low-resource environments. Overall, the system shows that combining voice and motor analysis can significantly strengthen the diagnostic process for Parkinson's Disease, especially in early or hard-to-detect cases.

Conclusion

This project successfully demonstrates that Parkinson's Disease can be reliably detected using a dual-modal approach involving voice data analysis and spiral image classification. The system addresses the need for early, non-invasive, and accessible screening methods by combining two effective machine learning models: a Random Forest classifier for acoustic features and a Convolutional Neural Network (CNN) for hand-drawn spiral images. The results confirmed that both models achieved high accuracy and generalization, with voice-based detection capturing early vocal impairments and image-based analysis identifying motor disturbances such as tremors. The smooth convergence of training and validation metrics, along with visual outputs, supports the stability and robustness of the proposed models. Furthermore, the deployment of the system through a user-friendly Streamlit web application makes it practical for remote screening and telemedicine use. The flexibility to operate with either modality ensures broader applicability in real-world healthcare scenarios. In conclusion, the problem of early and accessible Parkinson's Disease detection has been effectively addressed by integrating dual data modalities into a single predictive framework. Future work may involve combining both predictions into a fused decision system and validating the approach on larger and more diverse datasets for clinical adoption. [8]

Acknowledgements

We, the authors, would like to express our sincere gratitude to the developers and contributors of the open-source tools and frameworks that made this project possible. In particular, we acknowledge the creators and maintainers of scikit-learn, TensorFlow, OpenCV, and Streamlit, as well as other essential Python libraries that enabled us to preprocess data, train models, and deploy our dual-modal Parkinson's Disease detection system efficiently. We also extend our thanks to the providers of open-source datasets, including publicly available Parkinson's voice datasets and spiral drawing image repositories, which were instrumental in training and evaluating the models used in this project. The availability of comprehensive documentation, online tutorials, and open communities surrounding these tools significantly supported our learning and implementation process. Finally, we are grateful for the teamwork and mutual collaboration within our group, which contributed to the smooth execution and successful completion of this project.

References

[1]. Tsanas, A., Little, M. A., McSharry, P. E., & Ramig, L. O. (2010). Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests. IEEE Transactions on Biomedical Engineering, 57(4), 884–893. https://doi.org/10.1109/TBME.2009.2036000

[2]. Little, M. A., McSharry, P. E., Hunter, E. J., Spielman, J., & Ramig, L. O. (2009). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Transactions on Biomedical Engineering, 56(4), 1015–1022. https://doi.org/10.1109/TBME.2008.2005954

[3]. Piyush, B., & Anil, K. (2021). Parkinson Disease Detection Using Spiral Test and Convolutional Neural Network. International Journal of Engineering Research & Technology (IJERT), 10(6), 95–100.

[4]. Keerthivasan, S. P., & Saranya, N. (2023). Acute Leukemia Detection using Deep

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 05 May 2025
Page No: 2148 - 2155
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0315

Learning Techniques. International Research Journal on Advanced Science Hub, 5(10), 372–381. https://doi.org/10.47392/IRJASH.2023.066

[5]. Abrol, S., & Dutta, D. (2021). Voice Based Detection of Parkinson’s Disease using Machine Learning Techniques. International Journal of Scientific & Technology Research, 10(2), 45–49.

[6]. TensorFlow. (2024). TensorFlow: An end-to-end open source platform for machine learning. Retrieved from https://www.tensorflow.org

[7]. Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

[8]. Streamlit Inc. (2023). Streamlit Documentation. Retrieved from https://docs.streamlit.io/