International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 1
Jan-Feb 2026, PP 93-105

AI-Based Real-Time Parkinson's Diseasse Detection and Monitoring System Using Voice Analysis

[1]Kanimozhi S, [2]Hemalatha S, [3]Devasri M, [4]Kanisha S

Department Of Computer Science and Engineering Kongunadu College of Engineering and Technology Tiruchirappalli,India

Abstract - Parkinson's disease (PD) is a progressive neurodegenerative disorder in which speech impairment appears at an early stage, making voice analysis a reliable and non-invasive biomarker for diagnosis and monitoring. However, many existing speech-based systems rely on manually engineered acoustic features and focus primarily on disease detection without providing objective severity assessment or continuous monitoring. This project presents an end-to-end AI-based real-time Parkinson's disease detection and monitoring system using voice analysis. Speech signals are transformed into log-Mel spectrogram representations to preserve spectral and temporal characteristics and are analyzed using a Conformer-based deep learning architecture that integrates convolutional layers with self-attention mechanisms. A multi-task learning strategy is employed to simultaneously identify Parkinson's disease and estimate its severity level. To improve robustness, data augmentation techniques are applied to handle noise and speaker variability. The system further incorporates an AI-driven recommendation module that provides personalized health suggestions based on the predicted results. A secure cloud- connected dashboard enables real-time storage, visualization, and tracking of patient data over time. Additional modules, including speech emotion analysis, smart alerts via email and SMS, automated PDF report generation, mobile application integration, and data encryption, enhance usability, safety, and accessibility. The proposed system supports early detection, objective disease monitoring, and scalable deployment in clinical and remote healthcare environments.

Keywords - Parkinson's Disease, Voice Analysis, Speech Biomarkers, Log-Mel Spectrogram, Conformer Architecture.

Introduction

Parkinson's disease (PD) is a long-term neurodegenerative disorder that gradually affects motor control, speech, and overall quality of life. It is commonly associated with symptoms such as tremor, muscle rigidity, slowed movements, and impaired coordination. In addition to these motor symptoms, changes in speech often appear during the early stages of the disease. These vocal changes include reduced loudness, monotone pitch, imprecise articulation, and irregular speech rhythm, making voice an important indicator for early diagnosis and disease monitoring.

Conventional methods for diagnosing Parkinson's disease rely mainly on clinical observation and neurological examinations conducted by specialists. Although effective, these assessments are subjective, require in-person hospital visits, and are not well suited for continuous monitoring. As a result, early-stage symptoms may remain unnoticed, and disease progression may not be consistently tracked. With the growing number of Parkinson's patients worldwide, there is a strong need for automated, objective, and accessible diagnostic tools.

Recent advances in artificial intelligence (AI) and speech processing have significantly contributed to the development of voice-based systems for Parkinson's disease (PD) detection. Since Parkinson's disease affects motor control, it also impacts speech production, leading to vocal symptoms such as reduced loudness, monotone pitch, breathiness, imprecise articulation, and irregular speech rhythm. These vocal impairments can be detected using computational analysis, making speech a non-invasive, low-cost, and easily accessible biomarker for early diagnosis and monitoring.

Despite these advancements, many existing voice-based Parkinson's detection systems rely heavily on manually extracted acoustic features, such as pitch, jitter, shimmer, formant frequencies, and harmonic-to-noise ratios. These features are typically fed into traditional machine learning models like Support Vector Machines (SVM), Random Forests, or k-Nearest Neighbors (k-NN). While such approaches have shown promising results in controlled settings, they have several limitations.

One major drawback is that handcrafted features require expert knowledge and may fail to represent the full complexity of speech signals. Speech is inherently a dynamic and temporal process, and manually selected features often capture only isolated or static characteristics, ignoring long- range temporal dependencies and subtle variations over time. As a result, traditional models may struggle to generalize well across different speakers, recording conditions, and disease stages.

To address these limitations, this work proposes an AI- based real-time Parkinson's disease detection and monitoring system using voice analysis. The system employs an end-to- end deep learning framework based on a Conformer architecture, which effectively models both local speech characteristics and long-range temporal dependencies. By adopting a multi-task learning strategy, the system simultaneously performs Parkinson's disease classification and severity estimation from speech recordings. Furthermore, the system integrates cloud-based storage, real- time visualization, personalized recommendations, and secure data handling, making it suitable for both clinical use and remote healthcare applications. This approach aims to support early diagnosis, objective disease assessment, and continuous monitoring in a scalable and user-friendly manner.

II. Problem Statement

Parkinson's disease is a progressive neurological disorder that is often diagnosed only after noticeable motor symptoms appear, by which time significant neuronal damage has already occurred. Although speech impairment emerges at an early stage of the disease, it is not routinely or systematically used in clinical practice for early diagnosis or continuous monitoring. Existing diagnostic procedures rely heavily on in-person neurological examinations and subjective clinical rating scales, which limit frequent assessment and make long-term tracking difficult.

Current speech-based Parkinson's disease detection systems commonly depend on manually engineered acoustic features and conventional machine learning models. These approaches require expert knowledge, are sensitive to noise and recording conditions, and often fail to capture subtle temporal variations in speech signals. Moreover, most existing solutions are designed only to classify whether a person has Parkinson's disease, without providing an objective estimation of disease severity or supporting real- time monitoring.

There is a lack of an integrated, secure, and scalable system that can automatically analyse speech, accurately detect Parkinson's disease, estimate its severity, and continuously track patient progress in real-world environments. Addressing this gap requires an end-to-end intelligent solution that eliminates manual feature extraction, supports remote and real-time usage, ensures data security, and provides meaningful feedback for patients and healthcare professionals.

III. Literature Survey

Little et al. (2010) presented one of the earliest studies on Parkinson's disease detection using voice recordings. The authors extracted acoustic features such as jitter, shimmer, and fundamental frequency from sustained vowel phonation and applied classical machine learning classifiers. Their work demonstrated that speech signals can serve as an effective biomarker for distinguishing Parkinson's patients from healthy individuals.[1] Tsana's et al. (2012) proposed a speech-based telemonitoring system for Parkinson's disease assessment. Instead of focusing only on disease detection, the study emphasized estimating clinical severity using regression models mapped to UPDRS scores, highlighting the importance of continuous and remote disease monitoring.[2]

Orozco-Arroyave et al. (2014) analysed articulatory and phonatory speech characteristics in individuals with Parkinson's disease. The study revealed that temporal variations and articulation impairments in speech strongly correlate with disease progression, supporting the use of advanced temporal modelling techniques.[3]

Eskidere et al. (2016) applied feature selection methods combined with Support Vector Machines for Parkinson's disease classification using voice data. Their approach improved classification accuracy by removing redundant acoustic features but relied heavily on manual feature engineering.[4]

Sakar et al. (2018) introduced a comprehensive speech dataset containing multiple types of vocal tasks such as sustained vowels, words, and sentences. The authors evaluated various machine learning classifiers and demonstrated that using diverse speech tasks enhances Parkinson's disease detection performance.[5]

Ma et al. (2020) conducted a clinical study on voice changes in Parkinson's disease and identified early speech impairments such as reduced loudness, monotonic pitch, and articulation errors. Their findings supported speech analysis as a non-invasive tool for early diagnosis.[6]

Gullapalli and Mittal (2022) presented a detailed review of machine learning techniques used for Parkinson's disease detection through speech analysis. The study highlighted limitations of traditional models in capturing long-term temporal dependencies and recommended deep learning- based approaches.[7]

Allayah et al. (2023) developed an automated Parkinson's disease detection system using acoustic features and ensemble learning methods with recursive feature elimination. Although the system achieved improved accuracy, it primarily focused on disease classification without severity estimation.[8]

Momeni et al. (2024) proposed a mobile-based voice analysis system for Parkinson's disease detection. Speech samples recorded using smartphones were analyzed using machine learning algorithms, demonstrating the feasibility of low-cost, remote, and accessible screening solutions.[9]

Recent studies by Xu et al. (2025) explored transformer- based and Conformer-based deep learning models for speech-based Parkinson’s disease assessment. These models utilized spectrogram representations and multi-task learning to jointly perform disease detection and severity estimation, achieving improved robustness and scalability.[10]

IV. Existing System

Current approaches for Parkinson’s disease assessment are largely dependent on traditional clinical evaluation methods and limited automated speech analysis techniques. In medical practice, diagnosis is typically performed through direct neurological examination, where specialists observe motor symptoms and assign scores using clinical rating scales. Although widely accepted, this approach requires physical hospital visits and does not support frequent or continuous monitoring of patients.

Automated Parkinson’s disease detection systems using speech have been explored as an alternative non-invasive solution. Most of these systems follow a conventional pipeline in which predefined acoustic features—such as voice perturbation measures and spectral descriptors—are manually extracted from speech recordings. These features are then used as inputs to classical machine learning classifiers, including support vector machines, decision trees, and ensemble-based models. While such systems demonstrate potential for distinguishing affected individuals from healthy speakers, their effectiveness is strongly influenced by the quality of handcrafted features and controlled recording conditions.

More recent developments have introduced basic deep learning models applied to speech representations. However, these models are often shallow in design and are limited in their ability to represent long-term temporal dependencies and subtle articulatory variations present in Parkinsonian speech. In addition, the majority of existing solutions treat Parkinson’s disease analysis as a binary classification problem, providing no structured estimation of disease severity or progression.

Another limitation of existing systems is the lack of integration with real-time monitoring tools, secure data storage, and patient-oriented feedback mechanisms. Most implementations function as standalone analytical models rather than complete healthcare solutions, reducing their practicality for remote monitoring, telemedicine, and long-term disease management.

V. Methodology

The proposed AI-based Parkinson’s disease detection and monitoring system follows a structured, end-to-end approach that integrates speech processing, deep learning, and secure cloud technologies to support early diagnosis and continuous monitoring. Speech recordings are captured through a web or mobile application using standard microphones, eliminating the need for specialized external hardware. The recorded voice signals undergo preprocessing steps such as noise reduction, silence removal, and normalization to ensure consistent audio quality. To preserve both spectral and temporal characteristics of speech, the processed signals are converted into log-Mel spectrogram representations, which serve as the input to the deep learning model. A Conformer- based neural network is employed to analyze speech patterns by combining convolutional layers for effective local feature extraction with self-attention mechanisms that model long- range temporal dependencies. The system adopts a multi-task learning strategy that simultaneously performs Parkinson's disease classification and estimates disease severity levels from the same speech input, enabling comprehensive assessment from a single recording.

To improve robustness and generalization across diverse speakers and real-world recording environments, data augmentation techniques are applied during the training phase. Based on the model's predictions, an AI-driven recommendation module provides personalized health suggestions, such as advising clinical consultation or speech therapy when necessary. All analysis results are securely stored in a cloud-connected dashboard that enables real-time visualization and long-term tracking of patient data. Additional functionalities—including speech emotion analysis, automated alerts via email or SMS, downloadable PDF report generation, mobile application integration, and strong data encryption—further enhance usability, safety, and accessibility. Through this integrated methodology, the proposed system delivers a scalable, non-invasive, and user- friendly solution for the early detection and continuous monitoring of Parkinson's disease in both clinical and remote healthcare environments.

VI. Proposed System

The proposed system presents an AI-based, voice-driven framework for the early detection and continuous monitoring of Parkinson's disease. It is designed as a two-phase architecture that integrates deep learning, speech signal processing, and cloud-based monitoring to enable accurate diagnosis and real-time assessment. The system leverages speech as a non-invasive biomarker and employs a Conformer-based multi-task learning model to simultaneously perform Parkinson's disease detection and severity estimation. By combining offline model training with real-time voice analysis, personalized recommendations, and secure data management, the proposed system provides a scalable, reliable, and user- friendly solution suitable for both clinical and remote healthcare environments.

PHASE 1: Model training (Back end)

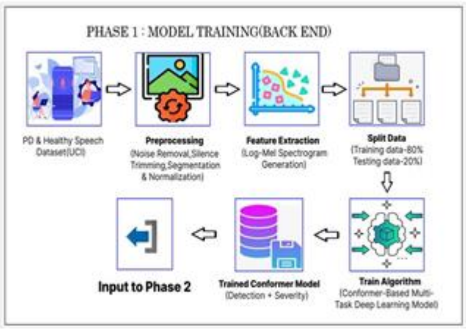

Fig. 1. Model Training and Knowledge Generation

- Speech Dataset Collection: In Phase 1, speech samples from Parkinson's disease patients and healthy individuals are collected using the Parkinson's Speech Dataset available from the UCI Machine Learning Repository. The dataset includes multiple speech tasks such as sustained vowels, spoken numbers, words, and short sentences, enabling comprehensive analysis of vocal characteristics associated with Parkinson's disease.

- Audio Preprocessing: The collected speech signals undergo preprocessing to improve data quality and consistency. This step includes background noise removal, silence trimming, speech segmentation, and signal normalization. Preprocessing minimizes the impact of recording variations and ensures uniform input for feature extraction and model training.
- Feature Extraction Using Log-Mel Spectrograms:
- After preprocessing, the cleaned audio is converted into log- Mel spectrograms, which capture both frequency and temporal speech patterns and are well suited for deep
- learning analysis of Parkinsonian speech. This enhances the model's ability to detect subtle voice impairments.
- Dataset Splitting: The extracted features are divided into training and testing sets, typically using an 80% training and 20% testing ratio. This split enables effective learning while allowing independent evaluation of the model's performance.
- Conformer-Based Multi-Task Model Training: A Conformer-based deep learning architecture is trained using the prepared dataset. The model combines convolutional layers for capturing local speech patterns with self-attention mechanisms for learning long-range temporal dependencies. A multi-task learning strategy is adopted, where one output branch performs Parkinson's disease detection and the other estimates disease severity levels. The trained model is saved and forwarded as learned knowledge for real-time inference in Phase 2.

PHASE 2: Real time detection (Front end)

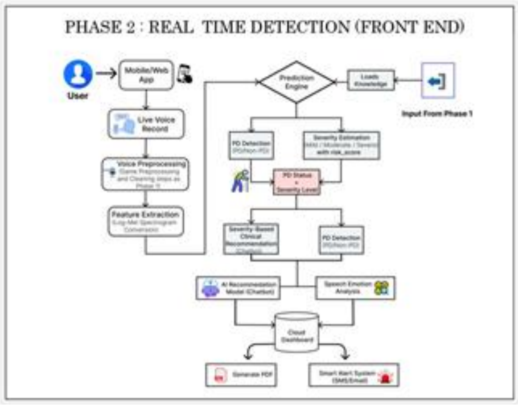

Fig. 2. Real-Time Prediction, Monitoring, and Decision Support

- Phase 2, users interact with the system through a mobile or web application. The application allows users to record live speech samples using built-in microphones, ensuring easy access without requiring specialized equipment.
- Real-Time Voice Preprocessing: The recorded speech samples undergo the same preprocessing steps used in Phase 1, including noise reduction, silence trimming, and normalization. Applying an identical preprocessing pipeline ensures consistency between training and real-time prediction.

- Feature Extraction for Real-Time Input: The preprocessed speech is converted into log-Mel spectrogram features, which are fed into the trained Conformer model. This step ensures that real-time input data matches the feature representation learned during training.
- Prediction Engine: The prediction engine loads the trained Conformer model from Phase 1 and performs multi-task inference. The model simultaneously predicts Parkinson's disease status (PD or Non-PD) and estimates disease severity levels such as mild, moderate, or severe along with an associated risk score.
- AI-Based Recommendation and Emotion Analysis :
- Based on the prediction results, the system generates severity-based clinical recommendations to guide users toward appropriate medical consultation or monitoring. Additionally, a speech emotion analysis module evaluates emotional patterns in the user's voice, providing supplementary insights into mental well-being.
- Cloud Dashboard, Reporting, and Alerts: All prediction outcomes, recommendations, and emotional analysis results are securely stored in a cloud-connected dashboard. The dashboard supports real-time visualization and long-term tracking of patient data. The system also generates downloadable PDF reports and triggers smart alert notifications via email or SMS when critical severity levels are detected. Data encryption techniques ensure the privacy and security of sensitive information.

VII. Result and Discussion

Fig. 3. Overview of the proposed AI-powered Parkinson's Disease voice monitoring system.

The proposed AI-based Parkinson's Disease monitoring system was evaluated using real-time voice samples collected via a web platform. By combining signal processing, machine learning, and deep learning, the system effectively detects early voice abnormalities, estimates risk levels, and enables continuous, non-invasive monitoring through a user-friendly dashboard.

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 1 Issue 2
Nov-Dec 2025, PP 46-62

Fig. 4. Functional modules of the proposed Parkinson's Disease voice analysis system.

Performance Evaluation of the Voice-Based Parkinson's Detection System

The effectiveness and reliability of the proposed system were assessed through multiple test cases using recorded voice samples and real-time user interaction:

Voice Feature Analysis Accuracy: The system successfully extracted clinically relevant voice features such as jitter, shimmer, harmonics-to-noise ratio (HNR), pitch variation, and fundamental frequency. Abnormalities in jitter and shimmer were clearly identified, while stable HNR and pitch values indicated normal vocal stability in low-risk cases.

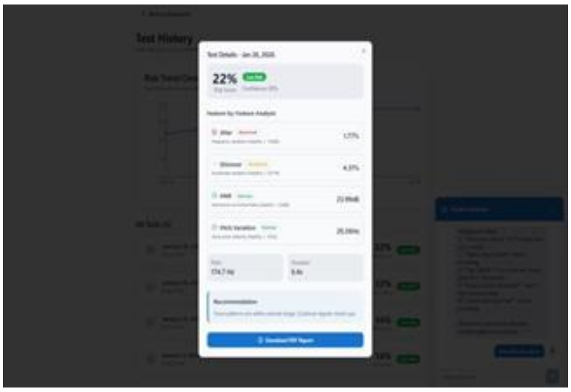

Fig. 5. Extracted voice features used for Parkinson's Disease risk assessment.

Risk Score Estimation: Based on the extracted features, the system generated a quantitative risk score (e.g., 16%–22%) along with qualitative risk categories such as Low, Medium, and High. This dual representation improved interpretability for both patients and clinicians.

International Journal of Statistics and Mathematical Research Trends
E-ISSN: 3108-0871, Volume 2 Issue 1
Jan-Feb 2026, PP 46-54

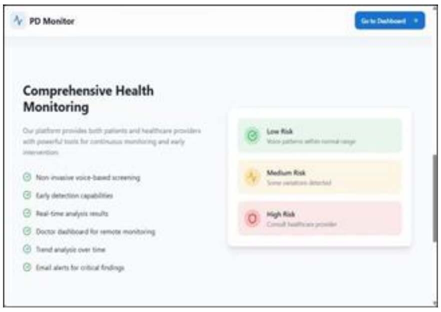

Fig. 6. Risk-level classification generated by the proposed system.

Early Parkinson’s Detection: The AI model accurately distinguished between normal and Parkinsonian
voice patterns, enabling early identification of poten7t)ial symptoms before severe motor signs appear.

Trend Analysis and Progress Monitoring: The dashboard showed risk score trends across multiple tests, helping users and doctors track disease progression or stability over time through clear line graphs.

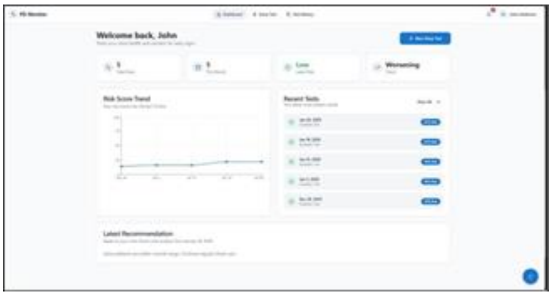

Fig. 7. Risk score trend visualization and recent voice test history.

Real-Time Voice Testing: The system supported short-duration voice recordings (5–7 seconds) and delivered instant analysis results, making it practical for frequent monitoring without clinical visits.

Automated PDF Report Generation: A detailed voice analysis report was automatically generated, summarizing patient information, risk assessment, extracted features, and clinical recommendations. This report can be used for medical review and documentation.

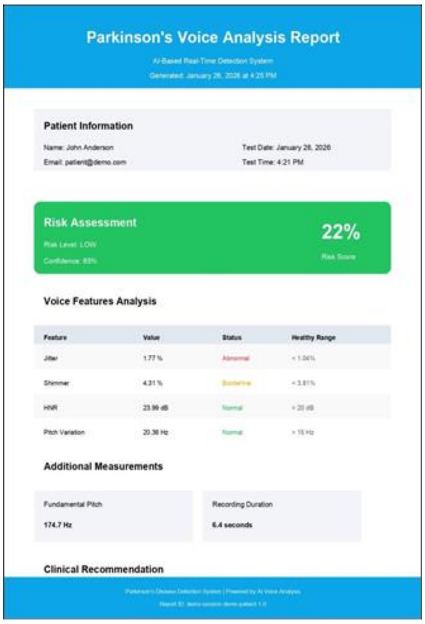

Fig. 8. Automatically generated Parkinson's Disease voice analysis report.

Security, Privacy, and Data Protection

To ensure patient data confidentiality and system reliability, multiple security mechanisms were incorporated:

- Secure User Authentication: User access to the dashboard and test history was protected through authenticated login mechanisms, preventing unauthorized access.
- Encrypted Health Data Storage: Voice samples, analysis results, and personal information were securely stored, ensuring compliance with healthcare data protection standards.
- Privacy-Preserving Voice Analysis: The system analyzed only essential acoustic features rather than storing raw voice content permanently, reducing privacy risks.
- Controlled Report Sharing: Generated PDF reports could be downloaded or shared only by authorized users, maintaining data ownership and confidentiality.

Integration of AI, User Interaction, and Clinical Support

The proposed system successfully combines artificial intelligence with an intuitive user interface to enhance usability and clinical relevance:

- AI-Powered Decision Support: Machine learning models analyzed complex voice patterns and provided clear diagnostic insights, supporting doctors in early decision- making.
- Health Assistant Chatbot: An integrated AI chatbot guided users on proper voice recording techniques, system usage, and general Parkinson's-related queries, improving user engagement.

Fig. 9. AI-based health assistant chatbot for user guidance and support.

Doctor and Patient Dashboards: Separate dashboards enabled patients to track their own health while allowing clinicians to remotely review test results and trends.

Non-Invasive and Cost-Effective Monitoring: Since the system relies only on voice input, it eliminates the need for expensive medical equipment, making Parkinson's screening accessible and affordable.

Clinical Significance and System Evaluation

The experimental results confirm that voice-based analysis is a reliable biomarker for early Parkinson's Disease detection. The proposed system not only identifies abnormal speech patterns but also supports long-term monitoring through trend analysis and risk scoring.

Compared to traditional clinical assessments, the system offers a non-invasive, scalable, and patient-friendly alternative. The integration of AI-driven analytics, real-time feedback, and secure reporting demonstrates strong potential for deployment in telemedicine and preventive healthcare environments.

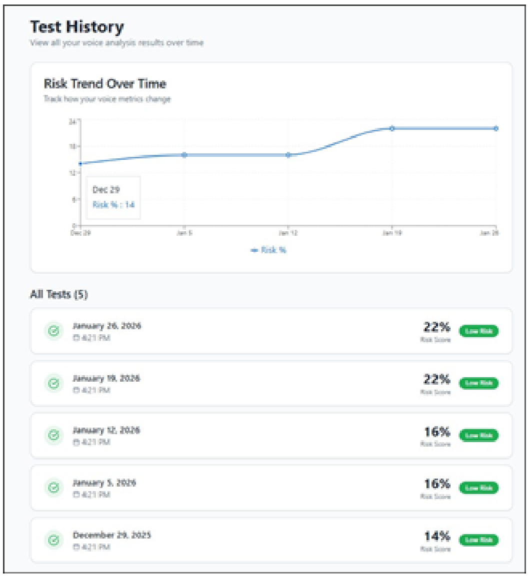

Fig. 10. Risk trend visualization and test history dashboard showing Parkinson's Disease risk scores across multiple voice assessments over time.

The dashboard provides a clear visualization of risk score trends across repeated voice tests, enabling clinicians and users to monitor disease progression or stability over time. Line graphs and test-wise risk summaries support continuous evaluation and early clinical decision-making.

VIII. Conclusion

This project presented an AI-based real-time Parkinson's disease detection and monitoring system using voice analysis as a non-invasive biomarker. By leveraging deep learning with a Conformer-based multi-task architecture, the system successfully performs both Parkinson's disease classification and severity estimation from speech signals. The integration of robust preprocessing, log-Mel spectrogram representation, and intelligent prediction enables accurate identification of abnormal vocal patterns associated with Parkinson's disease. In addition to detection, the system provides continuous health monitoring through trend analysis, personalized AI-driven recommendations, and secure cloud-based data visualization. The experimental results and system deployment demonstrate that the proposed framework offers a scalable, user-friendly, and clinically meaningful solution suitable for early diagnosis and remote healthcare applications, reducing dependency on hardware-intensive clinical assessments. It also supports remote access for patients and clinicians, making it ideal for telemedicine and follow-up care.

Future Scope

The proposed system can be further enhanced in several directions to improve accuracy, usability, and clinical adoption. Future work may include expanding the model to support multi-language and accent-independent voice analysis, enabling broader global accessibility. Incorporating additional biomarkers such as handwriting dynamics, facial expressions, or gait patterns could strengthen disease severity assessment through multimodal analysis. The integration of federated learning can improve model performance while preserving patient privacy by allowing decentralized training across healthcare institutions. Furthermore, real-time clinician feedback loops and integration with electronic health record (EHR) systems can support personalized treatment planning. Optimizing the system for mobile and wearable platforms will also enable continuous, passive monitoring, making the solution more suitable for long-term telemedicine and preventive healthcare deployments.

References

1. A. Allayah, S. Bourennane, and L. Khelifi, "Parkinson's disease diagnosis using ensemble learning and speech features," Applied Soft Computing, vol. 123, pp. 108902, 2022.
2. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," IEEE Transactions on Biomedical Engineering, vol. 56, no. 4, pp. 1015–1022, Apr. 2009.
3. A. Ma, A. Lau, and S. Thyagarajan, "Voice changes in Parkinson's disease: What are they telling us?" Journal of Clinical Neuroscience, vol. 72, pp. 1–7, 2020.
4. B. E. Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgen, S. Delil, H. Apaydin, and O. Kursun, "Collection and analysis of a Parkinson speech dataset

with multiple types of sound recordings," IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 4,
pp. 828–834, July 2013.

5. J. R. Orozco-Arroyave, F. Hönig, J. D. Arias-Londoño, J. F. Vargas- Bonilla, S. Skodda, J. Rusz, and E. Nöth, "Automatic detection of Parkinson's disease in running speech spoken in three different languages," Journal of the Acoustical Society of America, vol. 139, no. 1, pp. 481–500, Jan. 2016.
6. N. Momeni, H. Karimi, and S. Samavi, "Smartphone-based Parkinson's disease detection using voice analysis," IEEE Access, vol. 12, pp. 23145–23156, 2024.
7. O. Eskidere, F. Karakurt, and F. H. Ertaş, "Support vector machine based classification of Parkinson's disease using voice data," International Journal of Advanced Computer Science and Applications, vol. 7, no. 10, pp. 122–126, 2016.
8. S. Gullapalli and V. Mittal, "Machine learning approaches for Parkinson's disease detection using speech signals: A review," Biomedical Signal Processing and Control, vol. 68, pp. 102646, 2021.
9. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests," IEEE Transactions on Biomedical Engineering, vol. 57, no. 4, pp. 884–893, Apr. 2010.
10. Y. Xu, H. Zhang, J. Qin, and Z. Wu, "Transformer-based deep learning models for speech-based Parkinson's disease detection and severity estimation," IEEE Journal of Biomedical and Health Informatics, vol. 29, no. 2, pp. 845–856, Feb. 2025.