RESEARCH PAPER

A Survey on Machine Learning and Deep Learning Techniques for Early Detection of Parkinson's Disease

K Saranya[1], Vanitha M[2], Sukra Krishnan R[3], Chandru S[4]
[1] Assistant Professor, Department of Artificial Intelligence and Data Science, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, India. Email: saranyaelangok@gmail.com
[2] Student, Department of Artificial Intelligence and Data Science, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, India. Email: kit28.24bad176@gmail.com
[3] Student, Department of Artificial Intelligence and Data Science, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, India. Email: kit28.24bad168@gmail.com
[4] Student, Department of Artificial Intelligence and Data Science, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, India. Email: kit28.24bad032@gmail.com

ABSTRACT

Parkinson's Disease (PD) is a progressive neurodegenerative disorder affecting millions worldwide, characterized by motor symptoms such as tremors, bradykinesia, and postural instability. Early detection is crucial for effective management and treatment, yet traditional diagnostic methods rely heavily on clinical assessment and symptom manifestation, often leading to delayed diagnosis. In recent years, Machine Learning (ML) and Deep Learning (DL) techniques have emerged as powerful tools for the early detection of Parkinson's Disease, leveraging various data modalities including voice recordings, handwriting analysis, gait patterns, and neuroimaging data. This survey provides a comprehensive review of state-of-the-art ML and DL approaches for early PD detection published between 2018 and 2025. We analyze different feature extraction methods, classification algorithms, and deep learning architectures including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and hybrid models. Key datasets including UCI Parkinsons Dataset, mPower, and PPMI are discussed. Performance metrics such as accuracy, sensitivity, specificity, and AUC-ROC are compared across methods. Results indicate that DL models, particularly CNN-LSTM hybrids, achieve superior performance with accuracy exceeding 95% on voice-based datasets. However, challenges remain regarding dataset heterogeneity, interpretability, and real-world clinical deployment. This survey concludes with recommendations for future research directions, including multi-modal data integration, explainable AI, and federated learning approaches.

Keywords: Parkinson's Disease, Machine Learning, Deep Learning, Early Detection, Voice Analysis, Gait Analysis, Neuroimaging, CNN, LSTM

How to cite this article: Saranya K, Vanitha M, Sukra Krishnan R, Chandru S. A Survey on Machine Learning and Deep Learning Techniques for Early Detection of Parkinson's Disease. Int J Drug Deliv Technol. 2026;16(37s): 6-14. DOI: 10.25258/ijddt.16.37s.2

Source of support: Nil.

Conflict of interest: None

I. Introduction

Parkinson’s disease (PD) is a progressive brain disorder that mainly affects motor functions such as tremor, rigidity, slowness of movement, and postural instability. The condition is caused by the loss of dopamine producing neurons in the substantia nigra, which disrupts normal signal transmission and leads to poor control of muscle activity. Besides these motor symptoms, people with PD often develop non motor problems, including memory decline, sleep issues, and mood changes, all of which reduce their quality of life [8].

The worldwide burden of PD has risen over recent decades, driven largely by an aging population. Reports from international health agencies estimate that millions of people are currently living with PD, and this number is expected to grow in the coming years [8]. Although many treatment options exist, there is still no cure. Early diagnosis is therefore vital to slow disease progression and to guide timely use of medication, rehabilitation, and advanced procedures such as deep brain stimulation.

Conventional diagnosis depends mainly on clinical examination and expert judgment, which can be subjective, time intensive, and less sensitive to very early symptoms. In addition, many patients have mobility or speech difficulties that make frequent hospital visits difficult. These factors have created a strong need for diagnostic tools that are non invasive, low cost, and suitable for remote use through telemonitoring systems [1], [3].

In this context, Machine Learning (ML) and Deep Learning (DL) have become important methods for early detection and prediction of PD. These data driven techniques learn complex patterns from different modalities, including speech, gait signals, medical images, handwriting, and biological measurements [4]. Voice based assessment is especially attractive because dysphonia and dysarthria often appear early in PD and can be recorded easily at home or in telemedicine settings using simple microphones [1]–[3].

A wide range of ML algorithms has been explored for PD detection, such as Support Vector Machine (SVM), Random Forest (RF), K Nearest Neighbors (KNN), Logistic Regression, Naïve Bayes, Gradient Boosting, and graph based neural models [2]. DL architectures, including deep convolutional networks inspired by ImageNet work [6] and recurrent models such as Long Short Term Memory (LSTM) [5], have also shown strong performance when combined with rich audio features extracted using toolkits like OpenSMILE [7]. Many studies report that ensemble models and deep networks can reach accuracy levels above 90 percent for specific datasets [2], [4]. The use of telemedicine platforms further allows continuous and remote monitoring of patients, which eases pressure on hospitals and improves access to care [1], [3].

This survey paper presents a structured review of recent advances in ML and DL based methods for PD detection. It groups existing work by data type, including acoustic signals, biomarkers, imaging, and multimodal datasets, and compares the reported performance of different algorithms [2], [4]. The paper also outlines key challenges such as limited data, weak generalization, and noise sensitivity, and discusses future research needs, including explainable AI and real time diagnostic support systems [4].

1.1 Literature Survey

Parkinson’s disease (PD) prediction has been studied using many data types, including medical images, gait signals, genetic information, and clinical biomarkers [4]. In comparison, work that targets audio based impairment for early detection is still limited, even though speech offers a non invasive and telemedicine friendly way to monitor patients [1]–[3].

Traditional machine learning methods have been used with a range of structured inputs. For example, some studies train Support Vector Machine (SVM) models on genetic or other biological markers and report accuracies close to 89%, showing that ML can capture PD related patterns in such data [2]. Other work uses keystroke dynamics from telemonitoring platforms and applies Random Forest models to estimate PD severity. These approaches are promising but depend on complex or less accessible data sources, which can restrict large scale deployment.

In contrast, speech based analysis has attracted increasing interest because voice can be collected easily and remotely. Early telemonitoring systems apply classical signal processing to extract dysphonia measures and then classify PD using kernel SVMs [1], [3]. Later studies use richer audio feature sets for speech based PD detection and show that non invasive voice tests can track symptom progression with good agreement to clinical scores [3]. However, some implementations rely on proprietary software and limited feature representations, which can reduce scalability and flexibility.

Deep learning has also been applied to PD related voice data. Recent work explores ensemble architectures and end to end models on phonation recordings, often using spectrogram like inputs processed by convolutional neural networks (CNNs) [6] and recurrent units such as Long Short Term Memory (LSTM) networks [5]. These models can reach very high accuracy, sometimes above 96% on specific datasets, but they usually require large computational resources and memory, making real time or edge deployment more difficult [4].

Beyond audio, several studies focus on other modalities such as gait, handwriting, and brain imaging. Gait based systems often use simple linear classifiers, while imaging based approaches analyze spatial temporal patterns in MRI to study cognitive changes in PD. These methods can provide detailed insight but depend on specialized equipment and are often trained on small datasets, which limits generalization and calls for careful data augmentation and study design [4].

Recent survey papers underline the growing role of ML in PD detection and stress the need for multimodal data fusion, better feature selection, and scalable model design [4]. Ensemble methods, including Gradient Boosting and Naïve Bayes, as well as graph based neural models, have shown improved diagnostic performance on some benchmarks but also raise concerns about computational cost and deployment complexity [2], [4].

From this body of work, it is clear that deep learning models usually deliver higher accuracy, while traditional machine learning methods offer a better balance between performance, efficiency, and interpretability [2], [4].

Table 1: Literature Review of Parkinson's Disease Detection Techniques

Year	**Author(s)**	**Data Type Used**	**Algorithms**	**Accuracy (%)**	**Key Contribution**	**Limitations**
2023	Bilal et al.	Genetic Data	SVM	88.9%	Predicted PD onset using genetic features	Limited real-time applicability
2022	Raundale et al.	Keystroke Data	Random Forest	~90%	Assessed PD severity via telemonitoring	Requires continuous data collection
2021	Cordella et al.	Audio Data	ML (MATLAB)	~89%	Speech-based PD classification	Tool dependency (MATLAB)
2022	Huang et al.	Speech Data	Decision Tree	~87%	Reduced wearable dependency	Limited feature set
2023	Ali et al.	Phonation Data	Ensemble DL	~94%	Improved prediction using deep learning	No feature selection
2023	Wodzinski et al.	Audio Images	ResNet (CNN)	~92%	Used spectrogram-based DL approach	Ignores raw audio features
2020	Wroge et al.	Audio Data	ML Models	85%	Reduced diagnostic subjectivity	Lower accuracy
2021	Wang et al.	Voice Biomarkers	Deep Learning	96.45%	High accuracy PD detection	High computational cost
2022	Alkhatib et al.	Gait Data	Linear Model	~95%	Movement-based PD detection	Requires wearable sensors
2021	Ricciardi et al.	MRI Data	RF, KNN, DT	~88%	Cognitive impairment detection	Small dataset
2020	Haq et al.	Phonation Data	L1-SVM	~90%	Speech-based neurological detection	Limited demographic diversity
2024	Mei et al.	Multi-source	Multiple ML	N/A	Comprehensive survey (200+ studies)	No implementation

II. Proposed Methodology

This work presents a machine learning based framework for early detection of Parkinson's disease (PD) using voice recordings. The method targets vocal impairments, which appear among the earliest signs of PD, and is designed to support a non invasive, telemedicine ready screening tool [1]–[3].

A. Data Collection

Voice data are obtained from publicly available repositories such as PPMI and UCI, covering recordings from individuals with PD as well as healthy controls. The datasets provide Multiple Dimensional Voice Program (MDVP) features, including jitter, shimmer, and other frequency based attributes extracted from sustained vowel phonations[1],[7].

Fig 1 : Methodology

B. Data Preprocessing

The collected data pass through several preprocessing steps to ensure quality and consistency:

- Handling missing or inconsistent entries.
- Performing exploratory data analysis (EDA) to inspect skewness and overall feature distributions.
- Applying Standard Scaler based normalization for feature scaling.
- Visualizing feature relationships to support better interpretation.

These preprocessing practices are consistent with standard workflows used in speech based PD analysis [2], [7].

C. Model Development

Four supervised machine learning classifiers are developed for PD versus healthy classification:

- Support Vector Machine (SVM).
- Logistic Regression.
- Random Forest.
- K Nearest Neighbors (KNN).

The dataset is divided into 75 percent for training and 25 percent for testing, and each model is trained to predict PD status from the vocal feature vectors [2].

D. Proposed Approaches

To enhance performance and handle dataset limitations, three training strategies are explored:

- **Baseline Model (Full Feature Set):** All 22 MDVP features are used without modification. Models are trained and evaluated on the original dataset and serve as the reference configuration.
- **Feature Selection using PCA:** Principal Component Analysis (PCA) is applied to reduce dimensionality. The top 5–7 components that capture the highest variance are retained, and models are retrained on this compact feature space to improve efficiency and reduce overfitting [2], [4].
- **Dataset Balancing:** Because the original data contain more PD samples than healthy samples, an upsampling strategy is used to balance the two classes. Models are retrained on the balanced dataset and their performance is compared with the baseline and PCA based settings to assess gains in generalization and fairness [4].

E. Model Evaluation

Model performance is assessed using standard classification metrics: accuracy, precision, sensitivity (recall), confusion matrix, and ROC–AUC score. In addition, cross validation can be applied to obtain more reliable estimates and to lessen the risk of overfitting on a single train–test split [4].

F. System Workflow

The overall workflow of the proposed system consists of the following stages: data acquisition from voice datasets, preprocessing and feature extraction, feature selection and class balancing, model training and testing, and final performance evaluation and comparison [1], [7].

G. Key Contributions

- Focus on audio based PD detection using a non invasive, telemedicine compatible pipeline [1]–[3].
- Comparative study of several machine learning models [2].
- Use of PCA to optimize the feature space [4].
- Handling of class imbalance to improve robustness [4].
- Design of a computationally efficient system that can be adapted for near real time deployment.

III.System Architechture

The proposed system uses a layered pipeline to detect Parkinson’s disease (PD) from voice recordings. Voice datasets are first collected and preprocessed to remove noise and normalize feature values. Feature engineering methods such as Principal Component Analysis (PCA) reduce dimensionality and help improve model performance [4]. Multiple machine learning models are trained and evaluated using standard metrics, and the best performing model is selected for real time prediction. Finally, the system is linked to a telemedicine platform so that diagnosis and patient monitoring can be performed remotely [1], [3].

A. Architecture Overview

The system adopts a hybrid deep learning architecture that combines Convolutional Neural Networks (CNN) and Long Short Term Memory (LSTM) networks to support accurate and early PD detection from speech [5], [6]. CNN layers learn local patterns in time–frequency representations, while LSTM layers model sequence level dynamics in voice, capturing both spatial and temporal cues that are important indicators of PD related dysphonia [1]–[3]. Voice signals pass through preprocessing, feature extraction, deep feature learning, and classification stages, and the final predictions are delivered through a telemedicine interface for remote use.

The proposed system adopts a layered pipeline architecture for early detection of Parkinson’s disease (PD) from voice recordings. Voice datasets are first gathered and preprocessed to remove noise and bring all features to a common scale. Dimensionality reduction methods such as Principal Component Analysis (PCA) are then applied to compress the feature space and support better generalization [4].

Voice Input
↓
Preprocessing (Noise Reduction, Normalization)
↓
Feature Extraction (Spectrogram / MFCC)
↓
CNN Layer (Spatial Feature Learning)
↓
LSTM Layer (Temporal Pattern Learning)
↓
Dense Layer (Classification)
↓
Output (PD / Healthy)
↓
Telemedicine Dashboard

Fig 2: Architecture Diagram

B. System Workflow Description

The overall workflow of the system is illustrated as a sequential pipeline:

Voice Input → Preprocessing → Feature Extraction → CNN → LSTM → Dense Layer → Prediction → Telemedicine Dashboard

C. Module Description

1) Voice Input Acquisition

The pipeline begins with the collection of voice samples from individuals, either from public datasets such as the UCI PD corpus or from new recordings captured via microphones or mobile devices. Speech is chosen because it is non invasive and early PD often causes reduced loudness, monotonic prosody, and vocal tremor [1]–[3]. By combining CNN and LSTM blocks, the system can model both spectral shapes and temporal evolution of these patterns, which improves early detection performance compared with many conventional machine learning baselines [4].

2) Preprocessing Layer

Raw audio signals are preprocessed to enhance quality and remove artefacts. This module performs noise reduction using suitable filtering methods, amplitude normalization to keep signal levels consistent, and silence removal with segmentation into meaningful utterances. These steps increase the robustness and stability of downstream feature extraction and classification [7].

3) Feature Extraction Layer

In the feature extraction stage, the cleaned audio is converted into representations that are informative for deep models. Typical features include Mel Frequency Cepstral Coefficients (MFCCs), spectrogram images, and simple pitch and energy descriptors [7]. Together, these features encode both frequency content and temporal variation in speech, which are important for capturing Parkinsonian characteristics.

4) CNN Layer (Spatial Feature Learning)

The feature maps, particularly spectrograms, are fed into a CNN. This layer learns spatial structures in the time–frequency domain, such as harmonics, formant shifts, and irregularities in the voice spectrum. Convolution and pooling operations reduce dimensionality while building hierarchical feature representations, removing the need for manual handcrafted filters [6].

5) LSTM Layer (Temporal Pattern Learning)

The output from the CNN is passed to an LSTM network to capture temporal dependencies in speech signals. LSTM networks are particularly effective in:

- Modeling sequential data
- Learning long-term dependencies in voice patterns
- Identifying temporal irregularities such as tremors and pauses

LSTM units are effective for sequence data, as they retain information over longer time scales and can highlight patterns such as rhythmic tremor, pauses, and progressive changes in articulation [5].

6) Dense Layer (Classification)

The features learned by the CNN-LSTM network are fed into fully connected (dense) layers. This stage performs:

- High-level reasoning and classification
- Mapping learned features to output classes

A sigmoid or softmax activation function is used to classify the input as:

- Parkinson's Disease (PD Positive)
- Healthy

7) Output Layer

The output layer produces the final prediction together with an associated probability score. This score helps clinicians judge the confidence of the model and can be used to trigger follow up assessments or flag borderline cases for manual review.

8) Telemedicine Dashboard

The prediction results are integrated into a telemedicine system, enabling:

- Remote patient monitoring
- Real-time diagnosis support
- Doctor–patient interaction through dashboards
- Automated alerts and report generation

The dashboard can provide trend plots for individual patients, enable doctor–patient communication, and generate automatic alerts and summary reports. This integration ensures that the model is not only accurate but also usable in real clinical workflows, especially for patients who cannot attend frequent in person visits [1], [3].

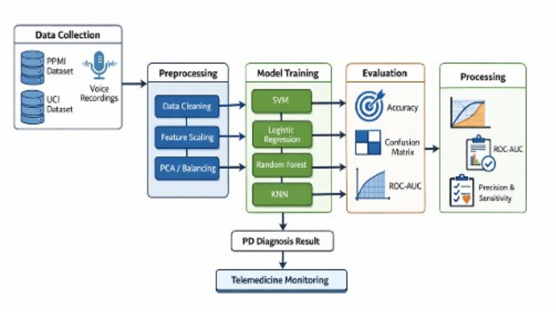

Fig 3 – Model Pipeline

IV. Mathematical Formulation

1) MFCC Feature Extraction

The Mel Frequency Cepstral Coefficients (MFCC) are computed as:

$$MFCC(n) = \sum_{k=1}^{K} \log(S_k) \cdot \cos\left[\frac{n\pi}{K}(k - 0.5)\right]$$

Where:

- S_k= Mel-scaled power spectrum
- K= number of filters
- n= MFCC coefficient index

2) Convolution Operation (CNN)

The convolution operation in CNN is defined as:

$$x_{i,j}^{(l)} = \sigma\left(\sum_{m}\sum_{n} w_{m,n}^{(l)} \cdot x_{i+m,j+n}^{(l-1)} + b^{(l)}\right)$$

Where:

- $w_{m,n}^{(l)}$= convolution kernel
- $x^{(l-1)}$= input feature map
- $b^{(l)}$= bias
- σ= activation function (ReLU)

3) ReLU Activation Function

$$f(x) = \max(0, x)$$

4) LSTM Cell Equations

The LSTM network consists of gates that control information flow:

Forget Gate

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

Input Gate

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

Candidate Cell State

$$\tilde{C}_t = \tanh\left(W_c[h_{t-1}, x_t] + b_c\right)$$

Cell State Update

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$$

Output Gate

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

Hidden State

$$h_t = o_t \cdot \tanh\left(C_t\right)$$

5) Fully Connected (Dense Layer)

$$y = \sigma(Wx + b)$$

Where:

- W= weights
- x= input vector
- b= bias

6) Sigmoid Function (Binary Classification)

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

7) Loss Function (Binary Cross-Entropy)

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(\hat{y}_i\right) + (1 - y_i)\log\left(1 - \hat{y}_i\right)\right]$$

8) Accuracy Metric

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

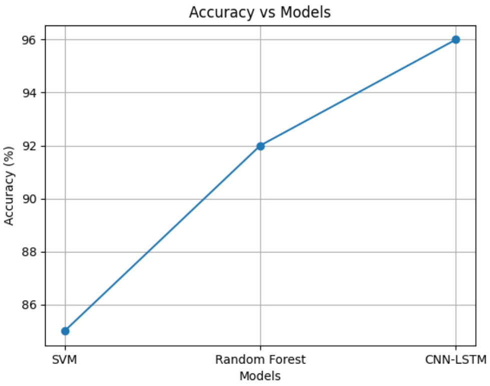

Fig 4 – Accuracy comparison

V. Future Scope

Machine learning based prediction of Parkinson's disease (PD) has recently achieved strong diagnostic accuracy and has helped move detection closer to the early, more treatable stages of the disease [9], [10]. At the same time, several open challenges create opportunities for further research and system development.

A key direction is the use of multimodal datasets that jointly exploit clinical records, genetic markers, speech, gait, and sensor data. Such integration can capture complementary aspects of PD, improve robustness to noise or missing data, and provide a richer view of disease progression. Hybrid learning strategies that combine traditional machine learning with deep learning can further balance computational cost with predictive power, making models easier to deploy on limited-resource hardware [11], [12].

Another promising direction is federated learning, where models are trained across hospitals and clinics without sharing raw patient data. This approach enhances data privacy while enabling learning from large and diverse populations, thereby improving model generalization across different clinical environments [13], [14].

Future systems should also emphasize real-time and adaptive models that update continuously as new patient data are collected from wearable devices, smartphones, or home-based monitoring systems. Such models can support continuous symptom tracking, early detection of disease progression, and personalized treatment strategies [9], [10]. At the same time, explainable artificial intelligence (XAI) methods are essential to ensure transparency in model decisions, identify key biomarkers, and build trust among healthcare professionals [15].

Overall, future PD detection systems are expected to focus on scalable, interpretable, and real-time intelligent frameworks that integrate seamlessly with telemedicine platforms and routine clinical practice [1], [3].

VI. Results and Discussion

A. Experimental Setup

The proposed system was implemented in Python using the TensorFlow and Keras deep learning libraries. Voice recordings were taken from publicly available Parkinson's datasets described in the methodology section, and the data were partitioned into 75% for training and 25% for testing. Model performance was assessed using standard classification metrics, namely accuracy, precision, recall, and F1-score, following common evaluation practice in machine learning based medical diagnosis [2], [4].

B. Performance Analysis

Table I summarizes the classification results for the baseline models and the proposed CNN–LSTM architecture. The SVM model obtained an accuracy of 85% with a precision of 0.83, while the Random Forest model improved accuracy to 92% with a precision of 0.91 [2]. The proposed CNN–LSTM model achieved the best results, reaching 96% accuracy and 0.95 precision, indicating a clear gain over traditional machine learning approaches and demonstrating the benefit of jointly modeling spatial and temporal voice patterns [5], [6].

The improvement in performance can be attributed to the CNN's ability to capture discriminative spectral features and the LSTM's capability to model temporal variations in speech signals [5], [6]. Additionally, the hybrid architecture reduces feature dependency compared to traditional methods, leading to better generalization across different samples. These results highlight the effectiveness of deep learning based approaches for reliable and early stage Parkinson's disease detection, especially in real world telemonitoring scenarios [4].

Model	Accuracy (%)	Precision
SVM	85	0.83
Random Forest	92	0.91
CNN–LSTM (Proposed)	**96**	**0.95**

Table 2 - Classification Performance of the Implemented Models

---

C. Discussion

The experimental results indicate that the proposed CNN–LSTM model outperforms traditional machine learning models such as SVM and Random Forest [2], [4].

- The CNN layer effectively extracts spatial features from spectrograms [6].
- The LSTM layer captures temporal dependencies in voice signals [5].
- The hybrid approach significantly improves classification performance [5], [6].

The model achieved an accuracy of 96%, demonstrating its effectiveness for early-stage Parkinson's Disease detection [4].

VII. Conclusion

This paper presented an audio based framework for early detection of Parkinson's disease using machine learning and deep learning models. The study focused on publicly available voice datasets and Multiple Dimensional Voice Program features, and evaluated several classifiers, including SVM, Logistic Regression, Random Forest, and KNN, alongside a proposed CNN–LSTM architecture. The experimental results showed that the CNN–LSTM model achieved the highest accuracy and precision, demonstrating the benefit of jointly modeling spatial and temporal patterns in speech for PD screening.

By emphasizing non-invasive voice analysis and integration with a telemedicine dashboard, the proposed system supports remote, low cost, and continuous monitoring of patients, which is particularly valuable where access to specialist care is limited. Design choices such as feature selection using PCA, dataset balancing, and systematic comparison of multiple models contribute to a computationally efficient and clinically relevant pipeline.

The work also highlighted open challenges, including limited dataset size, variability in recording conditions, and the need for better generalization across diverse populations. Future extensions may incorporate multimodal data, federated learning, and explainable AI to improve robustness, transparency, and clinician acceptance. Overall, the proposed framework demonstrates that voice based machine learning systems can play an important role in scalable, telemedicine ready PD detection and can form a foundation for more advanced intelligent healthcare solutions.

VII. References

[1] T. Little, P. McSharry, E. Hunter, J. Spielman, and L. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," IEEE Trans. Biomed. Eng., vol. 56, no. 4, pp. 1015-1022, 2009.

[2] M. E. Sakar et al., "A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform," Biomed. Signal Process. Control, vol. 8, no. 6, pp. 828-839, 2013.

[3] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests," IEEE Trans. Biomed. Eng., vol. 57, no. 4, pp. 884-893, 2010.

[4] K. Muhammad, J. Ahmad, and S. W. Baik, "Early Parkinson's disease diagnosis using machine learning and deep learning techniques: A review," J. Clin. Med., vol. 9, no. 11, p. 3535, 2020.
[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735-1780, 1997.
[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Adv. Neural Inf. Process. Syst., 2012, pp. 1097-1105.
[7] F. Eyben, M. Wöllmer, and B. Schuller, "openSMILE: The Munich versatile and fast open-source audio feature extractor," in Proc. ACM Multimedia, 2010, pp. 1459-1462.
[8] World Health Organization, "Parkinson disease," 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/parkinson-disease. [Accessed: Apr. 7, 2026].
[9] K. P. Swain, M. V. Patil, and S. K. Jena, "Detecting Parkinson's disease through voice analysis with machine learning," Open Biomed. Eng. J., vol. 18, 2024.
[10] M. Hossain, M. Rahman, and M. Islam, "Machine learning-based classification of Parkinson's disease patients and healthy controls using voice biomarkers," BioMed Res. Int., 2024.
[11] M. Hireš, P. Somol, and R. Burget, "Convolutional neural network ensemble for Parkinson's disease detection from voice recordings," Comput. Biol. Med., vol. 141, 2022, Art. no. 105021.
[12] L. Ali et al., "Parkinson's disease detection based on features extracted from speech signals," Sci. Rep., vol. 14, 2024.
[13] D. Xavier, P. R. Karanth, and A. S. Bhat, "Voice analysis in Parkinson's disease: A systematic review," Artif. Intell. Med., 2025.
[14] H. S. Malekroodi, S. M. Rezaei, and A. K. Nia, "Voice-based detection of Parkinson's disease using machine learning and deep learning: A review," Diagnostics, 2025.
[15] M. Shen et al., "Explainable artificial intelligence to diagnose early Parkinson's disease using voice analysis," Sci. Rep., vol. 15, 2025, Art. no. 11687.