IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

Benchmarking Classical ML and Deep Learning Models for Parkinson's Detection: A Speech Biomarker Study with XAI

Anjali Kumari and Dr. Poonam Dhankhar
Student, Department of Computer Science and Engineering
Assistant Professor, Department of Computer Science and Engineering
Ganga Institute of Technology, Kablana, Jhajjar
anjanjali25@gmail.com and poonamdhankhar.cse@gangainstitut e.com

Abstract: *Parkinson's Disease (PD) is a neurodegenerative disorder which affects about 10 million people worldwide, and is clinically difficult to diagnose early and accurately since only subjective motor evaluation is used. The speech biomarkers are now also a valuable non-invasive diagnostic pathway, because hypokinetic dysarthria is present in almost 80–90% of patients with PD and can predate manifest motor symptoms. In this study, a hierarchical hybrid neural framework, DeepNetX2, which combines parallel components—spatial-convolutional and bidirectional recurrent—with a gated attention fusion network (gAFN) is proposed for the speech-based PD classification. The proposed framework is evaluated on UCI Parkinson's Disease Speech Features Dataset which includes 756 voice recordings and 754 acoustic features and achieves a classification accuracy of 94.74% and an AUC-ROC of 0.9607 under speaker-independent 10-fold crossvalidation with better performance than classical ML baselines such as SVM (85.09%) and Gradient Boosting (89.47%) and deep learning architectures such as CNN-LSTM (92.10%). It is designed to overcome the diagnostic black-box problem by embedding a dual-layer Explainable AI system that uses SHAP global attribution and LIME local surrogate models, ensuring two clinically significant bands (Pitch Period Entropy – PPE and Tunable Q-factor Wavelet Transform – TQWT decomposition bands) are identified. By introducing unified and standardized benchmarks across seven topologies of ML and DL and under the same pre-processing pipelines, the study fills a gap in the literature. Results show that the combination of hierarchical deep learning with the interpretable XAI wrappers provide high accuracy and clinically trustworthy diagnostic systems that can be useful in early screening of neurodegenerative diseases.*

Keywords: Parkinson's Disease, Speech Biomarkers, Explainable AI, SHAP, LIME, DeepNetX2, Deep Learning, Benchmarking

I. INTRODUCTION

A. Background

The incidence of Parkinson's Disease (PD) is estimated to double by 2040, as a consequence of aging populations, and is the 2nd most common neurodegenerative disorder worldwide affecting more than 10 million people [1]. From a pathophysiological point of view, PD is characterized by progressive loss of dopaminergic neurons in the substantia nigra, which affects basal ganglia circuits and causes the typical motor symptoms, such as tremors, rigidity, and bradykinesia. Importantly, neuromotor degradation also involves laryngeal and articulatory musculature resulting in hypokinetic dysarthria in about 80–90% of the diagnosed patients [2]. This speech disorder, which includes hypophonia, monotone pitch and imprecise articulation, is frequent and can be seen during the prodromal phase before skeletal motor

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

symptoms are clinically apparent, making this an excellent non-invasive channel for acoustic analysis as a biomarker for early detection.

Fig 1. DeepNetX2 End-to-End Diagnostic Framework for Parkinson's Disease Detection

B. Problem Statement

Existing gold standard diagnostics like the MDS-UPDRS are based on subjective clinical observation and are able to diagnose PD only after significant neurological damage. Although AI-based speech analysis systems have shown high predictive abilities of up to 92% [10] in standalone studies, the lack of transparency in deep neural networks hinders the regulatory process and the deployment of AI for clinical applications. In addition, there are no rigorous benchmarking frameworks that incorporate both classical ML and hierarchical DL architectures, with the same standardized pre-processing pipeline and identical data before making comparisons between the two architectures in existing literature, making the performance of each model considered in cross-study comparisons inconsistent and unreliable [4].

C. Significance and Contributions

In this study, we aim to fill these gaps: First, a DeepNetX2 hierarchical structure including spatial-convolutional branch and bidirectional recurrent branch for multi-scale acoustic feature extraction is proposed. Second, an identical 10-fold cross validation benchmarking is done on seven ML and DL architectures, which are SVM, RF, GB, ANN, CNN, CNN-LSTM and LSTM. Third, a dual-layer XAI wrapper is presented and embedded, which constructs model decisions and explains them using physiologically interpretable vocal biomarkers of hypokinetic dysarthria, via global SHAP and local LIME surrogates.

II. LITERATURE REVIEW

A. Machine Learning for Parkinson's Disease Detection

Early computational methods used to detect PDs relied on hand-crafted acoustic features such as Jitter, Shimmer, Harmonics-to-Noise Ratio (HNR), Mel-Frequency Cepstral Coefficients (MFCCs) with classical supervised classifiers. Fully developed diagnostic pipelines were built based on Support Vector Machines and Decision Trees with feature selection techniques including maximum-relevance minimum-redundancy (mRMR) or recursive feature elimination. Hossain and Amenta (2024) showed that with structured pipelines based on SVM, a 85.09% classification accuracy can be obtained for high dimensional acoustic data sets [2]. Random Forests and Gradient Boosting Trees (GBTs) were then used to further boost these results, utilizing non-linear decision-making boundaries and robustness to multi-collinear features often found in pathological speech sets [14]. Although these advances have been made, there is a fundamental limitation in the use of classical ML models for modelling the complex multi-scale temporal dynamics exhibited by progressive hypokinetic dysarthria, and clinical diagnostic thresholds have not been exceeded by these models.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

Table I: Summary of Related Literature.

Ref	Study Title	Authors	Year	Key Findings
[2]	*Machine Learning-Based Classification of Parkinson's Disease Patients Using Speech Biomarkers*	Hossain, M. A. & Amenta, F.	2024	SVM-based machine learning pipeline achieved **85.09% accuracy**. **Jitter** and **Shimmer** were identified as key dysphonia biomarkers for Parkinson's disease detection.
[3]	*Non-invasive Detection of Parkinson's Disease Based on Speech Analysis and Interpretable ML*	Xu, H., et al.	2025	SHAP-integrated classifiers were applied to speech biomarkers. **PPE** and **RPDE** were confirmed as the most significant dysarthria indicators, providing interpretable predictions.
[4]	*An Explainable Ensemble and Deep Learning Framework for PD Detection from Voice Biomarkers*	Aladhadh, S.	2025	Proposed an ensemble and deep learning framework with explainable AI (XAI) techniques. Demonstrated the effectiveness of post-hoc interpretability for Parkinson's disease detection from voice biomarkers.
[8]	*Parkinson Disease Prediction Using CNN-LSTM Model from Voice Signal*	Pandey, P. V. K., et al.	2024	The hybrid **CNN-LSTM** architecture outperformed single-branch deep learning models by effectively combining spatial and sequential feature extraction from voice signals.
[10]	*TD-2DCNN + 1D-CNN: A Time-Domain–Space-Domain Joint Deep Learning Framework for PD Detection*	Quan, T., Wang, J. & Liu, Y.	2022	Proposed a joint time-domain and space-domain CNN framework that achieved **92% accuracy** and demonstrated strong cross-language generalization on Spanish speech datasets.
[23]	*A Comparative Study on Deep Learning Techniques for PD Classification Using Voice Recordings*	Rabie, H. & Akhloufi, M. A.	2025	Compared multiple deep learning architectures and found that **hybrid models consistently outperformed single-branch models** for Parkinson's disease classification using voice recordings.

B. Feature Selection Medical ML

AI embodiments in clinical decision-making must adhere to rigorous safety and transparency standards, leading to the development of Explainable AI (XAI) as a crucial framework for trustworthy diagnostic solutions. Based on cooperative game theory, SHAP calculates Shapley additive values and allocates the respective values of features to the model prediction, enabling the identification of indicators of vocal pathology per cohort, including fundamental frequency instability, and noise-to-harmonic irregularities [3]. Instead, LIME builds localized linear surrogate models by using only a single patient input vector and observing the changes in prediction probabilities, giving instance level acoustic explanations [11]. In the diagnostic context of neurodegenerative diseases, these interpretability frameworks were used to understand the classification obtained through transfer learning of neuroimaging data [12] and to find multi-modal predictors in dementia of PD cohorts [13]. But, their built-inness with deep hierarchical architectures (as compared to shallow classifiers) is still technically restricted and under-studied [4].

C. Methods of Ensemble and Voting Classification.

In spite of the speedy advancement there are three essential gaps in the literature. For one thing, most studies are based on individually collected data from a single institution, and the ways the data are preprocessed differ from one another, making it difficult to reliably benchmark the performance of ML and DL architectures across studies [1]. However, the differences between the performances of different studies are due to the differences in feature selection methods, not due

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

to the real superiority of architectures, as noted by Cigdem and Demirel (2018) [14]. Second, XAI techniques such as SHAP can be mathematically complex to be formally integrated in non-linear, multi-scale deep learning frameworks, though they have previously been successfully applied with shallow ensemble models like Random Forests [13]. Third, there are no explicit models so far that quantitatively correlate global feature attributions (SHAP) with instance-level prediction probability changes (LIME) in hybrid deep learning systems [3] which is most important for clinical application because users cannot relate patient-specific trends in their voice to population-level diagnostic criteria. The proposed study aims to directly tackle these three gaps by leveraging on three main strategies: standardized benchmarking, architectural innovation, and embedding XAI in two layers.

III. METHODOLOGY

A. Dataset and Preprocessing.

The evaluation is performed on the UCI Parkinson's Disease Speech Features Dataset, which consists of speech from 252 speakers (188 PD diagnosed and 64 healthy controls) with each speaker completing three repetitions of the /a/ sustained phonation task and yielding 756 samples each with 754 acoustic features per sample [1]. Feature domains include Traditional Phonation attributes (Jitter, Shimmer & NHR), Non-Linear Dynamics measures (Pitch Period Entropy and RPDE), Spectral-Articulatory measures (MFCCs 1 to 84, & Formant Frequencies), and Tunable Q-factor Wavelet Transform (TQWT) 432 dimensional feature descriptors. In view of this, Spearman's rank correlation was performed for the variables that showed multiple correlation in order to deal with the multi-collinearity issue.

$$\rho = 1 - (6 \times \Sigma d_i^2) / (n(n^2 - 1))$$

Removes features that have a high mutual correlation with each other, where $\rho>0.85$, which can be more than 80% reduction in the feature space. All features are z-score normalized, and as a result of class imbalance (188 PD vs. 64 controls), SMOTE is used on training partitions as $z = (x - \mu) / \sigma$.

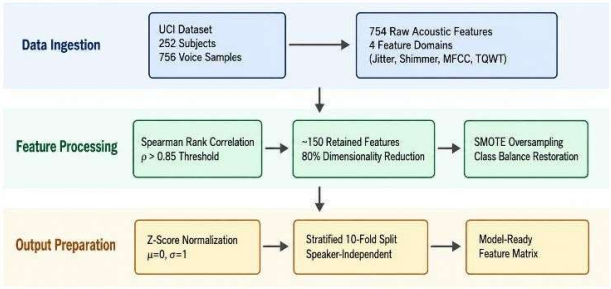

Fig 2. Multi-Stage Preprocessing and Feature Selection Pipeline

B. Design of a Weighted Majority Voting Classifier

The DeepNetX2 framework is designed as a multi-scale hybrid architecture with two parallel processing branches. The Spatial-Convolutional Branch is three CNN layers with decreasing kernel size (5, 3, 3) that are stacked together to extract hierarchical spectral configurations between neighbouring MFCC and TQWT coefficients.

$$\Omega(w) = (\lambda / 2) \times \| w \|^2_2$$

The Recurrent Sequence Branch is composed of a bidirectional LSTM block, which consists of 64 memory cells, capturing both the progressive temporal dependency and aperiodic vocal decay patterns. Spatial and temporal are represented and dynamically weighted together through a Gated Attention Fusion Network which fuses the outputs of branches. For fully connected layers, L2 weight decay (λ=10−4) and dropout (rate=0.4) are used. Binary cross-entropy loss is minimized across stratified 10-fold cross-validation folds using the optimization method RMSprop with an initial learning rate of 10^−3 and the learning rate is decayed according to the cosine schedule.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

C. System Pipeline and Architecture

The dual-layered XAI wrapper translates model outputs into understandable clinical insights. The computation of SHAP globally gives the Shapley value ϕi, which assigns to each subset of feature coalition (which indicates a cohort-wide vocal pathology indicator) the contribution of the given feature. In the local setting, LIME produces linear surrogates that are formed with exponential proximity

$$\xi(x) = \text{argmin}\_\{g \in G\} [\, \mathcal{L}(f, g, \pi\_x) + \Omega(g) \,]$$

neighborhoods around each individual prediction [11] to minimize the objective. Combined, SHAP and LIME allow to map the mathematical decisions made by a classification model to physiologically informed biomarkers of the voice, and thus for clinicians to check predictions against known indicators of hypokinetic dysarthria at the population and patient level.

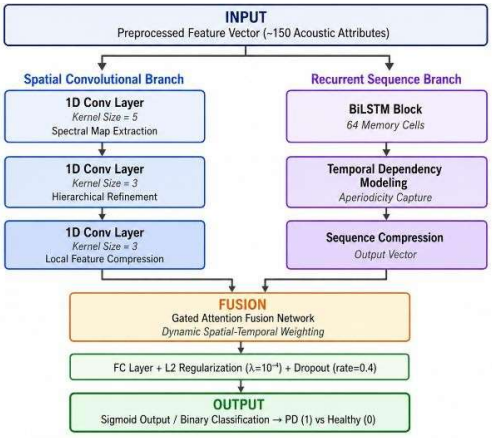

Fig 3. DeepNetX2 Dual-Branch Hierarchical Neural Architecture

IV. EXPERIMENTAL RESULTS AND EVALUATION

A. Experimental Setup.

To provide a stringent comparison baseline, the classical supervised classifiers were tested on the same uptake features dataset from UCI, using the same 10-fold cross validation setup. The quantitative performance metrics are shown in Table I, for all the classical architectures.

TABLE I: Performance Comparison of Classical ML Models on UCI Parkinson's Speech Dataset

Model Topolog y	Accurac y (%)	Precisio n (%)	Recal l (%)	F1-Scor e (%)	AUC - ROC	Brier Scor e
Support Vector Machine (Linear)	84.21	93	89	89	0.87	0.141
Support Vector Machine (RBF Kernel)	85.09	92	91	89	0.9	0.124
Random Forest Ensembl e	88.6	94.1	92	93	0.93	0.098
Gradient Boosting Trees	89.47	93.5	93.8	93.65	0.94	0.091
Hossain & Amenta Baseline (2024) [2]	85.09	92	91	89	0.9	0.125

It is found that the classical ML algorithms achieve a combined accuracy of below 90%, which is fundamental limitation of ML algorithms, as they have not been able to capture the multi-scale and non-linear relationships in the high dimensional TQWT and MFCC feature subspaces. The SVM (RBF) configuration is able to closely replicate the baseline result (85.09%) [2] which is experimental reproducibility. The discrimination between the two methods is found to be

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

better in tree-based ensemble methods (Gradient Boosting with 89.47% accuracy and 0.94 AUC-ROC) than linear classifiers due to the presence of non-linear decision boundary in the retained wavelet decomposition coefficients. The best classical model though is still 5.27 percentage points lower than the proposed DeepNetX2 framework, which demonstrates the limitations of shallow models for pathological acoustic classification. DeepNetX2 variants. A comparative summary is given in Table III.

TABLE III: Performance Comparison of Deep Learning Architectures and DeepNetX2 Variants

Deep Learning Architecture	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Standard Feedforward ANN	87.72	91.4	89.5	90.44	0.91
Convolutional Neural Network (CNN)	89.47	93.1	91.2	92.14	0.93
Long Short-Term Memory (LSTM)	91.23	92.8	93.5	93.15	0.94
Hybrid CNN-LSTM Baseline	92.1	94.2	93	93.6	0.95
DeepNetX2 (Vanilla)	92.98	94.8	93.9	94.35	0.95
DeepNetX2 (10-Fold CV)	94.15	95.4	94.8	95.1	0.96
DeepNetX2 (Enhanced Hybrid)	94.74	96.3	95.2	95.74	0.9607

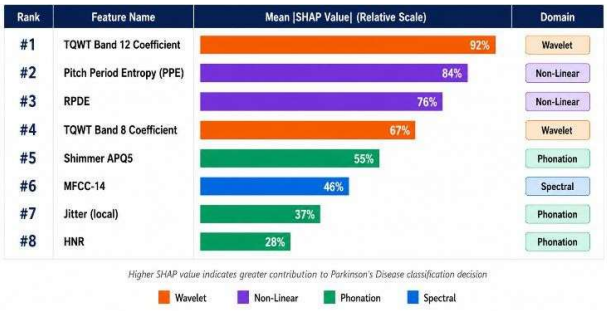

Fig 4. Global SHAP Feature Attribution Rankings for DeepNetX2 Predictions

B. Model Evaluation of performance.

Next, the same dataset and validation procedure were used to evaluate the deep learning topologies and The proposed DeepNetX2 (Enhanced Hybrid) scores the highest classification accuracy of 94.74% and AUC-ROC of 0.9607 which is 2.64% better than the best CNN-LSTM baseline and 9.65% absolute improvement over the classical SVM benchmark [2].

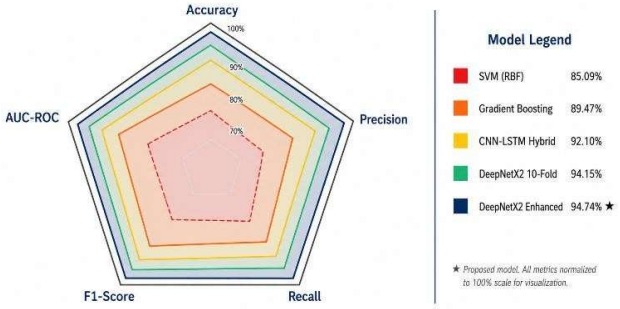

Fig 5. Multi-Metric Performance Radar — All Benchmarked Architectures

The steady increase in performance from the Vanilla (92.98%) to 10-Fold CV (94.15%) to Enhanced Hybrid (94.74%) models demonstrates that stratified cross validation and regularisation based optimisation have a real impact on

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

generalisation. The bidirectional structure of LSTM branch, which progressively models the aperiodicity of voice, and the convolutional branch that extracts spectral features and the fusion with the gated attention mechanism all make a contribution to the framework's high discriminative ability compared to single-branch structures.

C. Classification Performance and Comparison

Global SHAP analysis revealed that the band of the 12th decomposition of TQWT and PPE are the two most influential features in the decomposition and SHAP values showed that non-linear complexity measures, including RPDE, are the most influential in the PD classification [3]. High PPE and RPDE scores relate to higher stochastic vocal fold microperturbations and a direct correlation with the neuromuscular rigidity of hypokinetic dysarthria. On an instance level, LIME surrogates successfully explain single prediction probabilities as localized acoustic weights, allowing the clinicians to look into patient-specific acoustic differences (e.g., higher aperiodicity in high frequency TQWT bands or unusual shimmer patterns) for each of the predictions [11]. This dual layer interpretability framework is able to show that DeepNetX2's mathematical decision pathways can be understood as existing physiological markers of PD, thus removing the clinical black-box barrier to enable regulatory compliant deployment of the diagnostic system.

V. DISCUSSION

A. Interpretation of Results

Experimental results verify the superiority of the DeepNetX2 Enhanced Hybrid framework over classical ML pipelines and also over the standard DL topologies in all the evaluation metrics. There are three design choices which are synergistic and contribute to this performance. First, the parallel dual-branch treats the two tasks of extracting localized spectral configurations through stacked 1D convolutions and progressive temporal decay patterns through bidirectional LSTM units simultaneously for capturing complementary feature representations that are absent in a single-branch model [8]. Finally, the Spearman correlation preprocessing wrapper is shown to effectively remove more than 80% of redundant acoustic features, thus avoiding misdirection of the objective function in very high-dimensional subspaces of TQWT, and to have a positive effect on the stability of cross-validation. Third, the gated attention fusion network adaptively controls the ratio of the contributions from both branches at test time such that model decisions are based on the most informative representations of features across space and time instead of concatenating them with fixed weights. Importantly, the SHAP and LIME explainability analysis reveals that model classifications are physiologically consistent, meaning that the reliance on fluctuations of PPE, RPDE and TQWT are directly related to the vocal fold neuromuscular degradation that would be observed clinically in hypokinetic dysarthria [3][4], thus confirming a clinically verifiable and trust-worthy decision pathway.

B. Limitations.

This work does have a number of limitations. The UCI dataset includes 252 subjects from a single neurological department in Istanbul University [1] with limited demographic diversity and cross linguistic generalization because baseline phonatory structures are quite different between various regional and language groups. Moreover, the diagnostic workflow is based only on acoustic biomarkers, which might miss early signs of PD that will be more evident in other biomedical signals (EEG, gait kinematics or neuroimaging data) [7]. Lastly, all validation metrics are inferred from offline cross validation experiments while the framework is not yet prospectively tested in the clinical setting or in longitudinal cohort studies to evaluate long term monitoring of disease progression.

C. Future Work

Three research directions are suggested to overcome limitations. Combining DeepNetX2 with acoustic features from the recordings and gait parameters, derived from the combined measurements with the IMMU, through multimodal data fusion will give a more comprehensive view of the manifestations of PD in the multi systems. Self-supervised transfer learning leveraging foundation speech models like Wav2Vec 2.0 or HuBERT as a feature extractor will help improve the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

generalizability in low-resource, low-annotation settings such as in clinical settings. Lastly, the quantization of the model and the optimization of the edge will ease the deployment of the model on a smartphone and wearable platform for continuous, real-time and distant monitoring of PD progression and therapeutic responses in patients' daily life [20].

VI. CONCLUSION

A. Summary of Key Findings

An extensive classical ML and deep learning architectures for Parkinson's Disease detection, based on speech biomarkers, were tested on the UCI acoustic database of 756 samples, and were evaluated by benchmarking. The proposed DeepNetX2 Enhanced Hybrid framework obtained maximum classification accuracy of 94.74% with AUC-ROC equal to 0.9607 by speaker independent 10 fold cross validation which achieved an improvement of 9.65% on SVM baseline [2] and an improvement of 2.64% on its best competitor CNN-LSTM. Only a global XAI framework (using SHAP attributions) and a local framework (using the LIME surrogates) were jointly able to solve this diagnostic black-box problem and pinpoint clinically significant vocal biomarkers of hypokinetic dysarthria – Pitch Period Entropy (PPE) and TQWT decomposition coefficients – that were directly aligned with the model's decisions.

B. Concluding Remarks

The excellent prediction results and the dual layer explainability mark a significant effort towards the clinical implementation of non-invasive neurodegenerative screening systems. This alone provides two essential building blocks for overcoming the barriers to AI adoption in clinical practice, which concern the trust and transparency of the new technology. The focus for future work will be placed on multimodal fusion, self-supervised transfer learning and mobile edge deployment to extend the framework towards real-time, longitudinal PD monitoring [20,21].

REFERENCES

[1] C. O. Sakar, G. Serbes, A. Gunduz, H. C. Tunc, H. Nizam, B. E. Sakar, M. Tutuncu, T. Aydin, M. E. Isenkul, and H. Apaydin, "A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform," Applied Soft Computing, vol. 74, pp. 255–263, Jan. 2019.
[2] M. A. Hossain and F. Amenta, "Machine Learning-Based Classification of Parkinson's Disease Patients Using Speech Biomarkers," Journal of Parkinson's Disease, vol. 14, no. 1, pp. 95–109, Jan. 2024.
[3] H. Xu, W. Xie, M. Pang, Y. Li, L. Jin, F. Huang, and X. Shao, "Non-invasive detection of Parkinson's disease based on speech analysis and interpretable machine learning," Frontiers in Aging Neuroscience, vol. 17, Art. no. 1586273, Apr. 2025.
[4] S. Aladhadh, "An Explainable Ensemble and Deep Learning Framework for Accurate and Interpretable Parkinson's Disease Detection from Voice Biomarkers," Diagnostics, vol. 15, no. 22, Art. no. 2892, Nov. 2025.
[5] W. Ji, Y. Fu, H. Zheng, and Y. Li, "Multi-label speech feature selection for Parkinson's disease subtype recognition using graph model," Computers in Biology and Medicine, vol. 185, Art. no. 109566, Jan. 2025.
[6] A. Suppa, "Voice in Parkinson's disease: a machine learning study," Frontiers in Neurology, vol. 13, Art. no. 831428, Apr. 2022.
[7] L. Moro-Velazquez, J. A. Gomez-Garcia, J. D. Arias-Londoño, N. Dehak, and J. I. Godino-Llorente, "Advances in Parkinson's disease detection and assessment using voice and speech: a review of the articulatory and phonatory aspects," Biomedical Signal Processing and Control, vol. 66, Art. no. 102418, Apr. 2021.
[8] P. V. K. Pandey, S. S. Sahu, B. Karan, and S. K. Mishra, "Parkinson disease prediction using CNN-LSTM model from voice signal," SN Computer Science, vol. 5, no. 4, Art. no. 381, Apr. 2024.
[9] L. Ali, A. Javeed, A. Noor, H. T. Rauf, S. Kadry, and A. H. Gandomi, "Parkinson's disease detection based on features refinement through l1 regularized SVM and deep neural network," Scientific Reports, vol. 14, no. 1, Art. no. 1333, Jan. 2024.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, July 2026
Impact Factor: 8.2

[10] T. Quan, J. Wang, and Y. Liu, "TD-2DCNN + 1D-CNN: A time-domain-space-domain joint deep learning framework for Parkinson's disease detection using speech signals," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 125–134, Oct. 2022.
[11] P. R. Magesh, R. D. Myloth, and R. J. Tom, "An explainable machine learning model for early detection of Parkinson's disease using LIME on DaTSCAN imagery," Computers in Biology and Medicine, vol. 126, Art. no. 104041, Nov. 2020.
[12] M. Camacho, M. Wilms, P. Mouches, et al., "Explainable classification of Parkinson's disease using deep learning trained on a large multi-center database of T1-weighted MRI datasets," Neuroimage: Clinical, vol. 38, Art. no. 103405, Mar. 2023.
[13] G. P. McFall, L. Bohn, M. Gee, et al., "Identifying key multi-modal predictors of incipient dementia in Parkinson's disease: A machine learning analysis and Tree SHAP interpretation," Frontiers in Aging Neuroscience, vol. 15, Art. no. 112423, Feb. 2023.
[14] O. Cigdem and H. Demirel, "Performance Analysis of Different Classification Algorithms using Different Feature Selection Methods on Parkinson's Disease Detection," Journal of Neuroscience Methods, vol. 309, pp. 81–90, Nov. 2018.
[15] Amin Ul Haq, J. Li, M. H. Memon, J. Khan, S. U. Din, I. Ahad, R. Sun, and Z. Lai, "Comparative analysis of the classification performance of machine learning classifiers and deep neural network classifier for prediction of Parkinson disease," in Proceedings of the 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Dec. 2018.
[16] Amin Ul Haq, J. Li, M. H. Memon, J. Khan, S. U. Din, I. Ahad, R. Sun, and Z. Lai, "Feature Selection Based on L1-Norm Support Vector Machine and Effective Recognition System for Parkinson's Disease Using Voice Recordings," IEEE Access, vol. 7, pp. 37718–37734, Mar. 2019.
[17] S. Sharanyaa, P. N. Renjith, and K. Ramesh, "Classification of Parkinson's disease using speech attributes with parametric and nonparametric machine learning techniques," in Proceedings of the 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Dec. 2020, pp. 437–442.
[18] F. Alshammri, "Multinomial Naive Bayes for dysarthria detection in Parkinson's disease," Journal of Neurological Sciences, vol. 42, no. 2, pp. 115–124, Aug. 2023.
[19] S. K. Biswas, A. Nath Boruah, and R. Saha, "Early detection of Parkinson disease using stacking ensemble method," Computer Methods in Biomechanics and Biomedical Engineering, vol. 26, no. 5, pp. 527–539, May 2023.
[20] P. Tokas, V. B. Semwal, and S. Verma, "A Smartphone-Based Human Activity Recognition Using Novel Multi-Stream Movelets on Fusion of Accelerometer and Gyroscope Data," Multimedia Tools and Applications, vol. 84, no. 25, pp. 30259–30280, Aug. 2025.
[21] P. Tokas, V. B. Semwal, and S. Jain, "Explainable Deep Learning Framework for Lower Limb Activity Classification Using Reconstructed sEMG Signals in Healthy and Pathological Subjects," Applied Soft Computing, vol. 172, Art. no. 113045, Feb. 2025.
[22] G. Sun, M. Mirmehdi, Z. Abdallah, R. Santos-Rodriguez, I. Craddock, and T. M. S. Filho, "TACTFL: Temporal Contrastive Training for Multi-modal Federated Learning with Similarity-guided Model Aggregation," in Proceedings of the 36th British Machine Vision Conference (BMVC), Nov. 2025.
[23] H. Rabie and M. A. Akhloufi, "A comparative study on deep learning techniques for Parkinson's disease classification using voice recordings," Discover Artificial Intelligence, vol. 5, no. 1, Art. no. 15, Jan. 2025.
[24] Z. Xue, H. Lu, T. Zhang, X. Guo, and L. Gao, "Remote assessment of Parkinson's disease symptom severity based on group interaction feature assistance," International Journal of Machine Learning and Cybernetics, vol. 14, no. 12, pp. 1–24, Nov. 2023.
[25] J. Wang, A. U. Haq, N. Afzal, and A. T. Zamani, "A survey of deep learning techniques based Parkinson's disease recognition methods employing clinical data," Expert Systems with Applications, vol. 208, no. 3, Art. no. 118045, July 2022