A Next-Generation Feature-Selective Deep Hybrid Architecture for Attention-Driven Parkinson's Disease Classification

Raghuvira Pratap Adimulam
Department of Networking and Communications, School of Computing, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu District, Tamil Nadu, India
ra1482@srmist.edu.in

Suresh Annamalai
Department of Networking and Communications, School of Computing, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu District, Tamil Nadu, India
prisu6esh@yahoo.com (corresponding author)

Received: 9 February 2026 | Revised: 21 March 2026 and 3 April 2026 | Accepted: 4 April 2026

DOI: https://doi.org/10.48084/etasr.18067

ABSTRACT

Parkinson's Disease (PD) is among the most widespread neurodegenerative disorders, characterized by cardinal symptoms including bradykinesia, shortened strides, abnormal gait and posture, and other mobility impairments. PD detection primarily relies on clinician observation and evaluation of cardinal motor symptoms. However, conventional diagnostic methods are often constrained by subjectivity, as they depend on interpreting subtle movement alterations, which may lead to misclassification. Meanwhile, non-motor signs of PD may initially be subtle and could be attributed to several other conditions. Thus, these signs are frequently neglected, making the detection of PD at an early stage difficult. To overcome these obstacles and enhance the identification and assessment of PD, artificial intelligence, particularly Machine Learning (ML) and Deep Learning (DL), has been applied. Therefore, this study presents a Robust Parkinson's Disease Classification Framework using Advanced Deep Learning (RPDCF-ADL) model. The aim is to develop a reliable model for the accurate classification of PD patients and healthy control subjects. Initially, raw data undergo preprocessing, including data cleaning and data transformation, to improve data quality and consistency. Following that, a minimum Spearman Maximum Mutual Information (mSMMI) method is utilized to select the most informative features. The selected features are then fed into a hybrid classification model that integrates a Transformer with a Deep Belief Network (T-DBN) to effectively capture complex feature dependencies and non-linear patterns for precise PD classification. Furthermore, the AdamP optimization algorithm is adopted for weight optimization. The experimental validation of the RPDCF-ADL method achieved superior accuracy values of 98.61%, 98.86%, and 97.62% compared with existing models on the PD audio, PD, and HandPD datasets, respectively.

Keywords-Parkinson's Disease (PD); feature selection; medical data; Deep Learning (DL); AdamP optimization; Transformer network

I. INTRODUCTION

Parkinson's Disease (PD), also referred to as neurodegeneration, is a long-term, neurological, and degenerative motor condition characterized by the gradual demise of dopamine-generating brain cells [1]. Dopamine is a natural brain chemical created by nerve cells that acts as a messenger, permitting neurons to communicate with each other. PD results from disrupted neuronal signalling because of dopamine deficiency. In PD, neurons in the substantia nigra, a small region of the brain, are primarily affected [2]. However, PD is linked to a wide range of debilitating symptoms, including both motor and speech issues, which could lead to functional impairment, cognitive decline, and the deterioration of other essential functions. Cognitive decline is especially concerning, as it is associated with a high risk of developing dementia, with around half of PD patients developing progressive dementia within a decade [3]. Cognitive decline often starts early in the course of the disease and is associated with structural brain changes, including gray matter changes in the temporal areas, frontal, hippocampus, and parietal lobes, along with white matter alterations in the cingulate gyrus and corpus callosum [4]. Moreover, differentiating pathological aging from prodromal PD presents a significant diagnostic challenge [5].

Although PD has diverse clinical manifestations, overlapping symptoms often hinder accurate diagnosis and may contribute to misclassification [6]. If PD is detected at an early stage, it is easier to slow its progression. Thus, early diagnosis and accurate detection of this disease are essential for effective treatment and to reduce unnecessary medical testing that increases economic burdens and other risks [7]. Currently, Machine Learning (ML) and Deep Learning (DL) are significantly supporting clinicians in the timely disease identification. The benefit of ML is its capability to examine a vast amount of data and recognize patterns that cannot be detected by human evaluators [8]. In PD diagnosis, ML models could examine patient data such as medical valuations, genetic data, and imaging scans to detect subtle variations that might signify the existence of PD. However, various diseases and disorders have been detected using DL, and the outcomes often surpass those of conventional methods [9]. DL neural networks are especially suitable for extracting higher-level features, which enhance disease classification accuracy because of their remarkable generalization ability [10].

A. Research Contributions

To enhance the precision of automated PD identification, this paper presents a Robust Parkinson's Disease Classification Framework using Advanced Deep Learning (RPDCF-ADL) model. The RPDCF-ADL model aims to classify PD into corresponding classes accurately. The RPDCF-ADL model follows a multi-faceted pipeline encompassing data preprocessing, feature selection, hybrid Transformer-based classification, and weight optimization. The effectiveness of the RPDCF-ADL method is assessed using benchmark PD datasets across diverse evaluation metrics. The core contributions of this manuscript are listed as follows:

- Employs a minimum Spearman Maximum Mutual Information (mSMMI), a hybrid approach combining Maximum Mutual Information (MMI) and minimum Spearman (mS), for feature selection via dimensionality reduction, enhancing discriminative capability and minimizing model complexity.
- Designs a hybrid classification model that integrates a Transformer and a Deep Belief Network (T-DBN) for capturing complex feature dependencies and non-linear representations for precise PD diagnosis.
- Utilizes the AdamP algorithm for weight optimization, improving convergence stability and achieving better model performance.

II. LITERATURE REVIEW ON PARKINSON'S DISEASE DETECTION

This section presents an overview of existing studies on PD detection. Authors in [11] presented an automatic and objective architecture for classifying the intensity of speech disorders in PD patients employing ML and audio features. Authors in [12] introduced an innovative approach to detect Freezing of Gait (FoG) events based on movement signals. The model also combined a bottleneck attention technique into a traditional Bidirectional Long Short-Term Memory (BiLSTM) system. In [13], a multimodal dataset was proposed, which consists of Electromyogram (EMG), Electroencephalogram (EEG), and gait Accelerometer (ACC) signals for detecting FoG events. For a robust FoG detection process, the authors presented an innovative DL-driven method known as Self-FoGNet that incorporates the self-organized operational neuron (Self-ONN) layer for improving feature extraction. Also, the system incorporated feature refinement and effectively utilized the hierarchical feature space across multiple layers, allowing efficient classification. Authors in [14] presented an efficient DL method and speech-based error metrics to classify Parkinsonian and healthy subjects and analyze the disease condition. Utilizing five-fold cross-validation, the Audio Spectrogram Transformer (AST) and Vision Transformer (ViT) models were trained on vocal recordings to yield high-accuracy detection results.

Authors in [15] introduced an advanced Internet of Things (IoT)-assisted approach for medical data examination. An Evolutionary DBN (EDBN) was utilized for collecting and processing patient health data. The composite classifier integrated Convolutional Neural Network (CNN), Swarm-Driven Deep Feature Self-Organizer (SDFSO), and Long Short-Term Memory (LSTM), which were employed for classifying diseases with better precision. Hyperparameter optimization and dropout normalization were utilized to prevent overfitting. Authors in [16] developed a composite approach, which combines manual gait feature extraction with DL-assisted analysis employing a 3D-ResNext technique for enhancing FoG detection. Firstly, both DL and manual features were extracted, and secondly, such patterns were incorporated into an LSTM system for categorizing gait patterns into Pre-FoG, FoG, and Walk. Authors in [17] proposed a new method for detecting PD through analysis of audio signals employing Feature-Based Deep Neural Network (FB-DNN) models. An autoencoder, a particular form of Artificial Neural Network (ANN), was used to efficiently acquire complex patterns in audio data. The DNN was used in the classification process. Authors in [18] presented four DL approaches with a composite method for the early identification of PD. To enhance model efficacy, the Grey Wolf Optimizer (GWO) was employed for fine-tuning the parameters of these techniques.

III. PROPOSED METHODOLOGY

The RPDCF-ADL model aims to precisely classify PD patients and healthy control subjects utilizing biomedical data. The workflow of the RPDCF-ADL approach consists of four major phases: data preprocessing, feature selection, classification, and parameter optimization. Figure 1 illustrates the end-to-end architecture of the RPDCF-ADL model.

A. Data Preprocessing Pipeline

Data preprocessing is vital to ensure consistent inputs for analysis; consequently, data cleaning and data transformation are applied in this research. The PD data are liable to comprise missing values, which, if not addressed, may result in misleading outcomes [19]. An imputation technique, such as median imputation, is used to replace missing values with the median of the corresponding feature. Next, categorical variables are identified during the evaluation of contextual data. One-hot encoding is employed to convert these variables into binary inputs that can be processed by the model. Specifically, the categories are transformed into binary values indicating the presence or absence of PD.

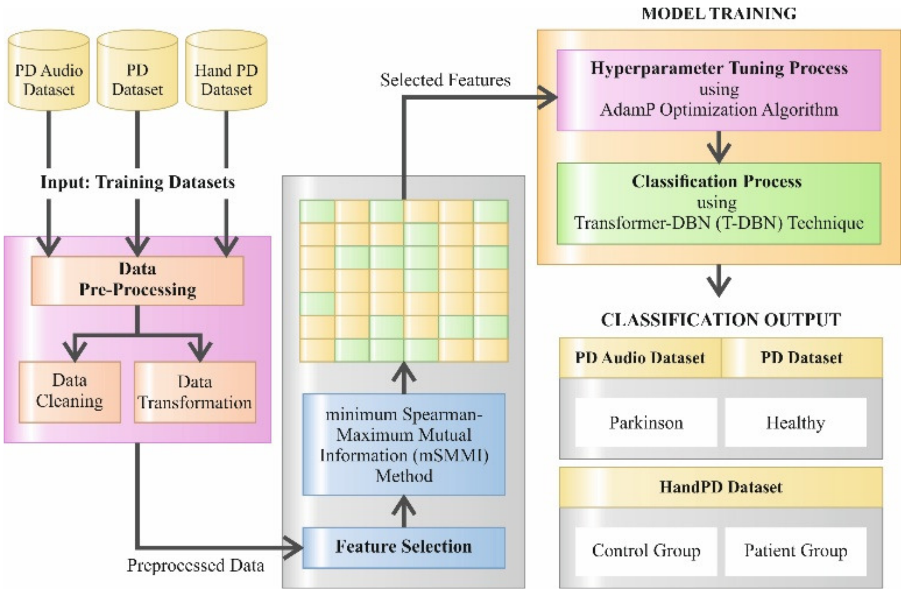

Fig. 1. End-to-end architecture of the RPDCF-ADL model.

B. Feature Selection Using Minimum Spearman Maximum Mutual Information

Feature selection plays a critical role in high-dimensional data analysis by reducing redundancy and improving model performance [20]. A new filter technique called mSMMI is proposed, which combines MMI to measure the relevance between features and class labels and mS to quantify redundancy among features, while adaptively balancing these components using the dual parameters u and v. A preliminary filtering phase reduces the feature dimensionality, in which the mSMMI technique iteratively selects features that exhibit higher relevance to the class label while preserving minimal redundancy with previously selected features.

In order to dynamically adjust the contributions of relevance and redundancy in the selection method, the weighting parameters s and y are computed. These variables are adaptively updated over the iterations to balance the impact of MMI and mS, thus avoiding the dominance of either redundancy or relevance across different datasets. The integrated selection method is defined as follows:

$$mSMMI(p_i) = \begin{cases} s \cdot MMI(F_{object}, L), & i = 0 \\ s \cdot MMI(F_{object}, L) - y \cdot mS_1(p_1, F_{object}), & i = 1 \\ s \cdot MMI(F_{object}, L) - y \cdot mS_2(P, F_{object}), & i > 1 \end{cases} \quad (1)$$

Here, the parameter i specifies the current iteration, P signifies the selected feature subset, and F_{object} denotes the candidate features. The adaptive weighting parameters are described by,

$$s = \cos\left(\frac{i}{I} \cdot \frac{\pi}{2}\right) \quad (2)$$

$$y = \sin\left(\frac{i}{I} \cdot \frac{\pi}{2}\right) \quad (3)$$

Here, the parameter I signifies the maximum number of iterations. During the selection process, the impact of relevance gradually decreases, whereas the importance of redundancy progressively increases, thus allowing the identification of stable feature subsets.

C. Deep Hybrid Classification Architecture

The T-DBN architecture is employed to effectively capture intricate feature dependencies and non-linear patterns for PD classification. Here, tabular biomedical features are treated as a sequence of feature tokens, where each feature represents a token and the sequence length corresponds to the number of selected features after mSMMI-based feature selection. Each feature token is embedded into a high-dimensional space using a learnable linear projection layer. To retain the relative importance and ordering of features, positional encoding is incorporated into the token embeddings. The resulting contextual embeddings are aggregated (via mean pooling) and fed as input to the DBN.

1) Transformer

The Transformer is a revolutionary neural network model based on the self-attention mechanism, designed to overcome the limitations of sequential data processing by enabling parallel computation and handling long-range dependencies [21]. The resulting sequence is then processed by the Transformer encoder to capture inter-feature dependencies and global contextual relationships. The input sequence is first projected into a constant-dimensional space, and then positional encoding enhances the embeddings with sequential information. The encoded data are processed by the Transformer encoder, where the Multi-Head Attention (MHA) mechanism assesses contextual correlations between components through query, key, and value (Q, K, V) relations. The outputs are normalized and passed through Feed-Forward (FF) layers and integrated with residual connections to form the final representation. This model leverages the strength of Transformers to capture sequential patterns effectively.

2) Deep Belief Network

The DBN is a generative probabilistic model that consists of multiple layers of stochastic latent variables. By stacking Restricted Boltzmann Machines (RBMs), a DBN is formed, where layer-wise pre-training is used for unsupervised hierarchical feature learning from input data. By employing greedy layer-wise pre-training followed by supervised fine-tuning, DBNs progressively learn more abstract representations from raw data. In the DBN architecture, the network is constructed by successively stacking RBMs. Finally, the output layer is connected to the last hidden layer to support supervised tasks such as classification. The core of a DBN lies in the RBM, which is defined by the relationships between hidden and visible variables.

$$E(v,h) = -v^T W h - b^T v - c^T h \tag{4}$$

where h refers to the hidden vector; v denotes the visible vector, W indicates the weight matrix that connects v and h, and b and c represent the bias vectors for visible and hidden layers, respectively.

$$P(v,h) = \frac{1}{Z}\exp\big(-E(v,h)\big) \tag{5}$$

$$Z = \sum_{v,h} \exp\big(-E(v,h)\big) \tag{6}$$

Here, Z refers to the partition function, which ensures that probabilities are properly normalized. Using the joint distribution, the marginal probability of the visible vector is given by:

$$P(v) = \frac{1}{Z}\sum_h \exp\big(-E(v,h)\big) \tag{7}$$

This expression measures the likelihood of an observed input v under the RBM by summing over all possible hidden configurations. RBMs enable efficient learning due to the conditional independence between visible and hidden units. The conditional probability of hidden units given the visible vector is defined as:

$$P\big(h_j = 1\big|v\big) = \sigma\big(\textstyle\sum_i v_i W_{ij} + c_j\big) \tag{8}$$

Here, σ is the sigmoid activation function. Similarly, the reconstruction of visible units from hidden units is given by:

$$P\big(v_j = 1\big|h\big) = \sigma\big(\textstyle\sum_j h_j W_{ij} + b_j\big) \tag{9}$$

The contrastive divergence learning algorithm is based on these conditional probabilities and is used to estimate the gradient of the log-likelihood function. Finally, stacking multiple RBMs forms a DBN, where the hidden activations of one layer serve as the visible inputs to the next. The layer-wise generative model is expressed as:

$$P(v, h^1, h^2, \ldots, h^L) = P(v|h^1) \prod_{l=1}^{L-1} P\,(h^l|h^{l+1}) P(h^L) \tag{10}$$

Here, $h^1, h^2, \ldots, h^L$ denote the hidden layers of the DBN, representing increasingly abstract feature representations.

D. AdamP Strategy

The AdamP optimizer is used for updating the network weights to improve convergence and achieve better generalization performance. The AdamP method is based on the concept of Projected Gradient Normalization [22]. AdamP improves upon Adam by projecting the gradient g_t onto the subspace orthogonal to the parameter vector θ_t , thereby focusing on directional updates. The essential mathematical form is given by:

$$g_{perp} = g_t - \frac{g_t * \theta_t}{\|\theta_t\|^2} * \theta_t \tag{11}$$

Here, θ_t denotes the parameter vector, g_t is the raw gradient, and g_{perp} represents the component of the gradient orthogonal to θ_t. The term $\frac{(g_t * \theta_t)}{\|\theta_t\|^2} * \theta_t$ corresponds to the projection of the gradient onto θ_t, capturing the component aligned with the parameter direction. Afterwards, AdamP utilizes g_{perp} to update Adam's first moment m_t and second moment v_t estimates. The parameter update rule is given by:

$$\theta_{t+1} = \theta_t - \eta * \frac{\hat{m}_t}{\sqrt{\hat{v}_t + e}} - \eta * \lambda * \theta_t \tag{12}$$

Using these updates, the adaptive term $\frac{\hat{m}_t}{\sqrt{\hat{v}_t + e}}$, now computed from g_{perp}, improves the update direction of the parameters. The weight decay term $-\eta * \lambda * \theta_t$ controls the parameter magnitude independently. By adapting to normalization layer scales and focusing on direction rather than magnitude, AdamP avoids overly aggressive weight decay. This results in more stable and consistent optimization, ultimately leading to improved performance and better generalization of the final model.

IV. EXPERIMENTAL EVALUATION AND DISCUSSION

In this section, the performance evaluation of the proposed model is presented in detail using three datasets and multiple evaluation measures. This assessment demonstrates the reliability of the proposed model in PD classification. Table I presents the details of the datasets.

TABLE I. DATASET SUMMARY

Class	PD audio	PD	HandPD
Parkinson	147	564	—
Healthy	48	192	—
Control Group	—	—	296
Patient Group	—	—	72
Total	195	756	368

The proposed model is implemented using the Python environment. The parameter settings of the proposed model are as follows: embedding dimension 128, Transformer with 4 layers, 8 heads, feed-forward size 256, and dropout rate 0.2. The DBN uses 3 layers with units [128, 64, 32], followed by a softmax classifier. Training uses AdamP (learning rate = 0.001), batch size 32, 100 epochs, and early stopping with patience of 10.

A. Dataset Description

The performance evaluation of the RPDCF-ADL method is conducted using three datasets: the PD audio dataset [23, 24], the PD dataset [25, 26], and the HandPD dataset [27, 28]. For experimental validation, 70% of the data are used for training and 30% for testing. A brief description of each dataset is provided below:

- PD audio dataset: This dataset contains biomedical voice measurements from 31 individuals, 23 of whom have PD. Each column represents a specific voice measure, and each row corresponds to one of 195 voice recordings from these individuals. The dataset includes a total of 24 features, of which 19 features were selected.
- PD dataset: The data used in this study were collected from 188 individuals with PD aged between 33 and 87 at the Neurology Department of Cerrahpaşa Faculty of Medicine, Istanbul University. The control group consists of 64 healthy individuals aged between 41 and 82. The dataset contains 756 samples with 754 features, of which 325 features were selected.
- HandPD dataset: The dataset consists of handwritten samples from two groups: (i) Patient Group (individuals affected by PD) and (ii) Control Group. The dataset includes 92 individuals, comprising 18 healthy subjects and 74 patients. The dataset contains a total of 166 features, of which 98 features were selected.

B. Performance Evaluation on the Three Datasets

Table II displays the PD identification results of the RPDCF-ADL method during the training phase using three datasets. According to the PD audio dataset, the average values include an accuracy of 98.61%, precision of 99.50%, sensitivity of 98.61%, specificity of 98.61%, and an F1-score of 99.05%.

TABLE II. PD DETECTION RESULT OF THE RPDCF-ADL METHOD WITH TRPH USING THREE DATASETS

Dataset	Training Phase					
	Class label	**Accuracy (%)**	**Precision (%)**	**Sensitivity (%)**	**Specificity (%)**	**F1-score (%)**
PD audio	Parkinson	100.00	99.01	100.00	97.22	99.50
	Healthy	97.22	100.00	97.22	100.00	98.59
	Average	98.61	99.50	98.61	98.61	99.05
PD	Parkinson	99.74	97.44	99.74	93.24	98.57
	Healthy	93.24	99.28	93.24	99.74	96.17
	Average	96.49	98.36	96.49	96.49	97.37
HandPD	Control Group	99.51	98.56	99.51	94.12	99.03
	Patient Group	94.12	97.96	94.12	99.51	96.00
	Average	96.82	98.26	96.82	96.82	97.52

The PD dataset achieves an average accuracy of 96.49%, precision of 98.36%, sensitivity of 96.49%, specificity of 96.49%, and an F1-score of 97.37%. Finally, the HandPD dataset obtains an average accuracy of 96.82%, precision of 98.26%, sensitivity of 96.82%, specificity of 96.82%, and an F1-score of 97.52%.

The comparative analysis of the RPDCF-ADL approach with existing methods is presented in Table III [29-32], where the simulation results demonstrate its superior performance.

TABLE III. COMPARATIVE ANALYSIS OF THE RPDCF-ADL METHOD WITH EXISTING METHODS ON THREE DATASETS

Dataset	**Metric**	**1D-Convnet [31]**	**ELM-BAT [31]**	**DWI [30]**	**Lenet5-AlexNet [31]**	**Inception V3 [31]**	**RPDCF-ADL**
PD audio	Accuracy (%)	94.10	96.74	82.00	95.10	95.00	98.61
	Precision (%)	92.64	86.37	90.53	86.84	94.36	99.50
	Sensitivity (%)	93.06	93.56	86.94	87.95	87.68	98.61
	F1-score (%)	86.04	90.35	86.69	88.48	94.32	99.05
	Processing time (s)	4.74	2.47	5.77	7.62	3.80	1.11
PD	**Metric**	**CFT-ADC [30]**	**Heuron IPD [29]**	**QSM-NMS Composite [29]**	**DenseNet [31]**	**ResNet [31]**	**RPDCF-ADL**
	Accuracy (%)	97.00	90.48	89.42	93.00	96.27	98.86
	Precision (%)	91.46	93.91	95.89	94.07	89.91	99.73
	Sensitivity (%)	92.09	96.40	92.47	93.84	89.39	98.86
	F1-score (%)	89.17	90.17	92.26	93.42	96.71	99.29
	Processing time (s)	5.57	6.22	3.92	6.13	7.58	1.65
HandPD	**Metric**	**FDG-DWI [30]**	**MultiParkNet [32]**	**EfficientNetB0 [32]**	**MobileNetV2 [32]**	**KSVM+FS [32]**	**RPDCF-ADL**
	Accuracy (%)	82.00	96.74	94.95	93.80	95.89	97.62
	Precision (%)	85.35	82.97	94.38	82.37	92.05	99.45
	Sensitivity (%)	86.48	94.16	83.64	86.93	84.13	97.62
	F1-score (%)	86.57	94.20	90.11	83.64	82.41	98.50
	Processing time (s)	5.04	6.33	5.71	6.93	5.76	2.14

For the PD audio dataset, the RPDCF-ADL model achieves a processing time of 1.11 s, significantly lower than those of 1D-Convnet, ELM-BAT, DWI, LeNet-5–AlexNet, and Inception V3, which require 4.74 s, 2.47 s, 5.77 s, 7.62 s, and 3.80 s, respectively.

Similarly, for the PD dataset, the RPDCF-ADL model records the lowest processing time of 1.65 s, whereas CFT-ADC, Heuron IPD, QSM-NMS Composite, DenseNet, and ResNet require 5.57 s, 6.22 s, 3.92 s, 6.13 s, and 7.58 s, respectively.

Finally, for the HandPD dataset, the RPDCF-ADL model achieves a processing time of 2.14 s, whereas FDG-DWI, MultiParkNet, EfficientNetB0, MobileNetV2, and KSVM+FS require higher times of 5.04 s, 6.33 s, 5.71 s, 6.93 s, and 5.76 s, respectively.

V. CONCLUSION

In this article, the Robust Parkinson's Disease Classification Framework using Advanced Deep Learning (RPDCF-ADL) model is presented to precisely classify Parkinson's Disease (PD) patients and healthy control subjects. Primarily, data preprocessing is carried out through data cleaning and data transformation to enhance data quality and consistency. Subsequently, the minimum Spearman Maximum Mutual Information (mSMMI) model identified the most informative features by eliminating redundant features. Besides, the hybrid Transformer–Deep Belief Network (T-DBN) model learns complex feature dependencies and non-linear patterns for accurate PD classification. Moreover, the AdamP optimization algorithm is employed for parameter tuning to enhance overall model performance. The RPDCF-ADL method achieves superior accuracy values of 98.61%, 98.86%, and 97.62% on the PD audio, PD, and HandPD datasets, respectively, compared with existing methods.

DECLARATION OF COMPETING INTERESTS

The authors declare no conflict of interests.

ACKNOWLEDGMENT

Not applicable in this work.

DATA AVAILABILITY

The datasets used to train and test the proposed approach are publicly available at [23-28].

REFERENCES

[1] M. Oumaima, B. Achraf, J. Abdelilah, and H. Ahmed, "Non-invasive detection of Parkinson’s disease using voice analysis and machine learning techniques," *International Journal of Speech Technology*, vol. 29, no. 1, Dec. 2025, Art. no. 3, https://doi.org/10.1007/s10772-025-10231-y.

[2] X. Bi *et al.*, "Diagnosis of Parkinson’s disease by Archimedes spiral dynamic handwriting analysis based on Dual-head Multimodal Networks," *Biomedical Signal Processing and Control*, vol. 112, Feb. 2026, Art. no. 108448, https://doi.org/10.1016/j.bspc.2025.108448.

[3] J. Asma Malica, K. Dinesh, P. Elavarasu, K. Devendran, C. Sagana, and R. Manjuladevi, "Prediction of Parkinson’s Disease Using Ensemble Learning," in *Proceedings of International Conference on Power Engineering and Intelligent Systems*, Srinagar, Uttarakhand, India, 2025, pp. 125–138, https://doi.org/10.1007/978-981-96-9720-5_10.

[4] W.-C. Wang, C.-C. Chiu, and S.-J. Yeh, "Deep learning classification for prodromal Parkinson’s disease based on spatiotemporal features of specific finger movements," *Biomedical Signal Processing and Control*, vol. 113, Mar. 2026, Art. no. 109216, https://doi.org/10.1016/j.bspc.2025.109216.

[5] S. Das, G. Kumar, and R. Sharma, "Prediction of Diabetes, Cardiac and Parkinson’s disease using machine learning and streamlit framework," *Multimedia Tools and Applications*, vol. 84, no. 24, pp. 28853–28872, July 2025, https://doi.org/10.1007/s11042-024-20244-5.

[6] W. Xiong, L. Ma, and H. Li, "Efficient Neural Network Classification of Parkinson’s Disease and Schizophrenia Using Resting-State EEG Data," *Brain Topography*, vol. 38, no. 3, Feb. 2025, Art. no. 32, https://doi.org/10.1007/s10548-025-01102-5.

[7] Y. Zhang *et al.*, "A novel approach for handwriting recognition in Parkinson’s disease by combining flexible sensing with deep learning technologies," *Sensors and Actuators A: Physical*, vol. 385, Apr. 2025, Art. no. 116287, https://doi.org/10.1016/j.sna.2025.116287.

[8] S. Saleh, B. Cherradi, O. El Gannour, S. Hamida, and O. Bouattane, "Predicting patients with Parkinson’s disease using Machine Learning and ensemble voting technique," *Multimedia Tools and Applications*, vol. 83, no. 11, pp. 33207–33234, Mar. 2024, https://doi.org/10.1007/s11042-023-16881-x.

[9] N. Singh and P. Tripathi, "An ensemble technique to predict Parkinson’s disease using machine learning algorithms," *Speech Communication*, vol. 159, Apr. 2024, Art. no. 103067, https://doi.org/10.1016/j.specom.2024.103067.

[10] H. Taherdoost, "Deep Learning and Neural Networks: Decision-Making Implications," *Symmetry*, vol. 15, no. 9, Sept. 2023, Art. no. 1723, https://doi.org/10.3390/sym15091723.

[11] M. Alqam, A. Hossen, S. Al-Abri, R. Al-Hmouz, and M. Deveci, "Classification of the Severity of Speech Disorders in Patients with Parkinson’s Disease Using Machine Learning Techniques," *International Journal of Computational Intelligence Systems*, vol. 19, no. 1, Dec. 2025, Art. no. 31, https://doi.org/10.1007/s44196-025-01109-7.

[12] S. Abbasi and K. Rezaee, "Deep Learning–Based Prediction of Freezing of Gait in Parkinson’s Disease With the Ensemble Channel Selection Approach," *Brain and Behavior*, vol. 15, no. 1, 2025, Art. no. e70206, https://doi.org/10.1002/brb3.70206.

[13] F. R. Sayem *et al.*, "A novel deep learning approach with feature injection technique for automated gait detection in parkinson’s disease using multi-modal data," *International Journal of Data Science and Analytics*, vol. 22, no. 1, Jan. 2026, Art. no. 47, https://doi.org/10.1007/s41060-026-01031-x.

[14] B. Perrone, F. Amato, and G. Olmo, "Voice classification in Parkinson’s disease: A deep learning approach using transformers and error rate metrics," *Biomedical Signal Processing and Control*, vol. 113, Mar. 2026, Art. no. 108954, https://doi.org/10.1016/j.bspc.2025.108954.

[15] S. A. Ansari, M. Luqman, S. Ali, P. Singh, M. Diwakar, and A. Bijalwan, "A Smart IoT-Based System for Early Prediction of Parkinson’s Disease Using Swarm-Assisted Convolutional Neural Networks," *International Journal of Computational Intelligence Systems*, vol. 19, no. 1, Jan. 2026, Art. no. 4, https://doi.org/10.1007/s44196-025-01102-0.

[16] N. Mehra *et al.*, "Hybrid Model for Freezing of Gait Detection in Parkinson’s Disease: Integrating Manual Features and Deep Learning," *International Journal of Computational Intelligence Systems*, vol. 19, no. 1, Jan. 2026, Art. no. 38, https://doi.org/10.1007/s44196-025-01148-0.

[17] P. Valarmathi, Y. Suganya, K. R. Saranya, and S. Shanmuga Priya, "Enhancing parkinson disease detection through feature based deep learning with autoencoders and neural networks," *Scientific Reports*, vol. 15, no. 1, Mar. 2025, Art. no. 8624, https://doi.org/10.1038/s41598-025-88293-w.

[18] B. Majhi *et al.*, "An improved method for diagnosis of Parkinson’s disease using deep learning models enhanced with metaheuristic algorithm," *BMC Medical Imaging*, vol. 24, no. 1, June 2024, Art. no. 156, https://doi.org/10.1186/s12880-024-01335-z.

[19] A. Puviarasu and S. V K, "Enhanced IoT security: privacy-preserving federated learning model for accurate, real-time intrusion detection across devices," *Ain Shams Engineering Journal*, vol. 17, no. 1, Jan. 2026, Art. no. 103866, https://doi.org/10.1016/j.asej.2025.103866.

[20] Y. Shi, Y. Zheng, and X. Bai, "A multiple filter-wrapper feature selection algorithm based on process optimization mechanism for high-dimensional omics data analysis," *Plos One*, vol. 20, no. 12, Dec. 2025, Art. no. e0338051, https://doi.org/10.1371/journal.pone.0338051.

[21] A. E. Oskouei, M. Kaveh, F. Hernando-Gallego, and D. Martín, "Hybrid Time Series Transformer–Deep Belief Network for Robust Anomaly Detection in Mobile Communication Networks," *Symmetry*, vol. 17, no. 11, Oct. 2025, Art. no. 1800, https://doi.org/10.3390/sym17111800.

[22] A. Pasha *et al.*, "Leveraging ANFIS with Adam and PSO optimizers for Parkinson’s disease," *Heliyon*, vol. 10, no. 9, May 2024, Art. no. e30241, https://doi.org/10.1016/j.heliyon.2024.e30241.

[23] M. 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, https://doi.org/10.1109/TBME.2008.2005954.

[24] "Parkinson’s Disease Data Set." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/vikasukani/parkinsons-disease-data-set.

[25] C. O. 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," *Applied Soft Computing*, vol. 74, pp. 255–263, Jan. 2019, https://doi.org/10.1016/j.asoc.2018.10.022.

[26] "Parkinson’s Disease (PD) classification." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/dipayanbiswas/parkinsons-disease-speech-signal-features.

[27] C. R. Pereira *et al.*, "A Step Towards the Automated Diagnosis of Parkinson’s Disease: Analyzing Handwriting Movements," in *2015 IEEE 28th International Symposium on Computer-Based Medical Systems*, Sao Carlos, Brazil, 2015, pp. 171–176, https://doi.org/10.1109/CBMS.2015.34.

[28] "HandPD dataset." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/claytonteybauru/spiral-handpd.

[29] T. Welton *et al.*, "Classification of Parkinson’s disease by deep learning on midbrain MRI," *Frontiers in Aging Neuroscience*, vol. 16, Aug. 2024, Art. no. 1425095, https://doi.org/10.3389/fnagi.2024.1425095.

[30] Y. Chang, J. Liu, S. Sun, T. Chen, and R. Wang, "Deep learning for Parkinson’s disease classification using multimodal and multi-sequences PET/MR images," *EJNMMI Research*, vol. 15, no. 1, May 2025, Art. no. 55, https://doi.org/10.1186/s13550-025-01245-3.

[31] S. Pechetti and B. S. Rao, "Optimized MobileNetV3: a deep learning-based Parkinson’s disease classification using fused images," *PeerJ Computer Science*, vol. 9, Nov. 2023, Art. no. e1702, https://doi.org/10.7717/peerj-cs.1702.

[32] A. Sar, P. S. Puri, H. Naz, S. Aich, T. Choudhury, and L. A. Gabralla, "Multi-modal deep learning framework for early detection of Parkinson’s disease using neurological and physiological data for high-fidelity diagnosis," *Scientific Reports*, vol. 15, no. 1, Oct. 2025, Art. no. 34835, https://doi.org/10.1038/s41598-025-21407-6.