International Research Journal of Multidisciplinary Scope (IRJMS), 2025; 6(4):1062-1074 IRJMS

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2025.v06i04.06552

The Multimodal Fusion of Voice, Gait, and Handwriting Detection of Parkinson's Disease Using Machine Learning

Mirle Bhyraj Meghashree[1], Karigowda Dhananjaya Kumar[1]*, Nagaraju Vinutha[1], Dinesh Akash[2]

[1]Faculty of Computer Science and Engineering, Vidya Vikas Institute of Engineering and Technology, Mysuru-570028, Karnataka, India, [2]Faculty of Information Science and Engineering, Vidya Vikas Institute of Engineering and Technology, Mysuru-570028, Karnataka, India. *Corresponding Author's Email: dhananjaya.kumar@vidyavikas.edu.in

Abstract

Parkinson's disease (PD) is an advanced neurological disorder that impacting movement control due to the degeneration of neurons that synthesize dopamine. The subtle nature of early symptoms such as shaking, bradykinesia, and speech changes often complicates timely diagnosis, reducing opportunities for early intervention. This study proposes a multimodal machine learning model for the early diagnosis of PD by integrating three complementary data modalities: voice recordings, gait analysis, and handwriting patterns. For each modality, specialized neural networks are deployed to extract critical features, including acoustic markers, motion irregularities, and fine motor dynamics. A self-supervised learning (SSL) paradigm is employed to enhance feature representation without reliance on large-scale labeled datasets, thereby addressing data scarcity challenges. These modality specific features are fused through a Multimodal Transformer model with cross-attention mechanisms, enabling the system to capture complex interdependencies and improve diagnostic accuracy. Evaluation on a cohort of 1,000 subjects (70% with early-stage PD and 30% healthy controls) achieved 96.5% accuracy, surpassing benchmark methods. The results highlight the potential of multi-modal integration and SSL for advancing earlier and more reliable PD detection, offering a promising pathway toward improved clinical outcomes.

Keywords: Cross-Attention Mechanism, Early Detection, Gait Analysis, Handwriting Analysis, Multimodal Transformer, Neural Networks, Parkinson's Disease, Self-Supervised Learning, Voice Analysis.

Introduction

Parkinson's disease (PD) is a gradually worsening neurodegenerative condition impacting millions globally. First identified by James Parkinson in 1817, the disorder results from the deterioration of dopamine-producing neurons in the substantia nigra, an area of the brain involved in controlling movement (1, 2). This decline in dopamine causes motor symptoms including tremors, muscle stiffness, bradykinesia (reduced movement speed), and difficulties with balance. Additionally, PD impacts non-motor functions, leading to issues such as changes in speech, micrographic (small handwriting), and gait abnormalities, often evident in the early stages (3, 4). Early diagnosis of PD is crucial since timely treatment can considerably slow disease progression and enhance patients' quality of life (5, 6). Nonetheless, detecting PD early remains challenging due to the subtle onset of symptoms and dependence on subjective clinical assessments that can be error-prone. This gap has encouraged the development of advanced technologies like machine learning to offer more accurate and reliable detection approaches (7). Machine learning (ML) has emerged as a transformative tool in healthcare, providing a data-driven method for understanding complicated patterns in medical data. Early research on Parkinson's disease has shown that machine learning algorithms can examine specific data types, such as voice recordings, gait patterns, and handwriting, to detect small changes associated with the disease. For instance, demonstrated that support vector machines could detect vocal impairments, such as reduced pitch variability and tremor-induced changes in speech, which are common early symptoms of the disease (8, 9). Similarly employed wearable sensors to track gait abnormalities like shorter stride lengths and freezing episodes, achieving promising results in diagnosing PD (10, 11).

(Received 28th June 2025; Accepted 08th October 2025; Published 30th October 2025)

Handwriting analysis has also proven effective showed that features like stroke length, writing pressure, and tremor-induced irregularities could differentiate PD patients from healthy individuals. Although these single-modality approaches have been instrumental in advancing PD diagnosis, their isolated nature limits their ability to capture the full complexity of the disease, which affects multiple systems simultaneously (12, 13).

To address the limitations of single-modality methods, researchers have turned to multimodal machine learning approaches that integrate data from multiple sources. Multimodal analysis integrates diverse data sources, including voice, gait, and handwriting, to offer a holistic view of Parkinson's disease manifestations throughout the body. By leveraging multiple modalities, this approach captures the complex and varied ways PD presents, enhancing the depth and accuracy of its detection and diagnosis. Each modality offers unique insights: voice data highlights vocal cord dysfunction, gait data reflects motor impairments, and handwriting data reveals deficits in fine motor skills (14, 15). The voice analysis could detect subtle acoustic changes associated with early-stage PD, while demonstrated that gait analysis could identify movement abnormalities with high accuracy (16, 17). took this approach further by integrating voice, gait, and handwriting data into a unified system, leveraging the power of multimodal transformers to achieve superior diagnostic performance (18, 19).

A key innovation in multimodal analysis is the machine learning techniques such as self-supervised learning (SSL) and multimodal transformers. SSL allows models to learn meaningful representations from data without relying on extensive labeled datasets, which are often scarce in medical research (20, 21). This makes SSL particularly valuable for analyzing diverse medical data, such as the multimodal datasets used in PD research. Once features are extracted from each modality, they are fused using multimodal transformers, which analyze relationships between different data types through attention mechanisms (22, 23), demonstrated the effectiveness of multimodal transformers in combining voice, gait, and handwriting data, achieving higher diagnostic accuracy than single-modality models (24).

The practical applications of multimodal systems in diagnosing Parkinson's disease are immense. Recent studies have demonstrated the real-world potential of these systems, which can be implemented using readily available technologies such as smartphones, wearables, and digital tablets. For instance, voice recordings can be collected via smartphone apps, gait data can be captured using wearable sensors, and handwriting samples can be obtained through digital writing devices (25, 26). A multimodal system tested on a dataset of 1,000 individuals, 70% of whom had early-stage PD, achieved an impressive accuracy of 96.5% (27, 28). This system outperformed single-modality approaches, highlighting the advantages of integrating multiple data sources. While challenges remain, such as computational demands and the need for diverse datasets to ensure generalizability, the potential benefits of these systems are transformative (29, 30).

Table 1 below provides a summary of various machine learning techniques applied for distinguishing Parkinson's Disease (PD) patients from Healthy Controls (HC) in multiple studies. The datasets used vary in size, ranging from small cohorts to extensive databases, and popular models such as CNNs, SVMs, and Random Forests have been commonly employed. Accuracy rates reported show variability, with CNN-based models achieving the highest accuracy of up to 99.3%, highlighting the potential of advanced machine learning methods for effective PD diagnosis.

Table 1: The Comprehensive Studies on Parkinson's Disease Detection

Ref no.	Category	Diagnosis	Data source	Machine Learning Methods	Result
(1, 2)	Categorizing of PD and HC	Detection	Participants data	Ensemble models (SVM, MLP, logistic regression, etc.)	Sensitivity = 96%, AUC = 0.98

(3, 4)	Categorizing of PD and HC	Detection	UCI repository	Stacked generalization with CMTNN	Accuracy = ~70%
(5, 6)	Categorizing of PD and HC	Detection	UCI repository	HMM and SVM	Accuracy = 95.16%
(7)	Categorizing of PD and HC	Detection	UCI repository	IGWO-KELM with cross-validation	Accuracy = 97.45%, F-measure = 98.34%
(8, 9)	Categorizing of PD and HC	Detection	Hand PD	Supervised learning	Chi-2 with Adaboost: Accuracy = 76.44%
(10, 11)	Categorizing of PD, HC, and other disorders	Detection	Participants data	Ensemble models (SVM, random forest, extra-randomized trees)	8-class classification accuracy = 82.7%
(12, 13)	Categorizing of PD and HC	Detection	Participants data	SVM with cross validation	PD vs. HC accuracy = 92.3%
(14, 15)	Categorizing of PD and HC	Detection	Collected data from PPMI	RLDA with JFSS, 10-fold cross-validation	Accuracy = 81.9%
(16, 17)	Categorizing of PD and HC	Detection	Hand PD	CNN, SVM with cross-validation	CNN accuracy = 96.3%, SVM accuracy = 92.4%
(18, 19)	Categorizing of PD and HC	Detection	Hand PD	Supervised learning	CNN-Cifar10 accuracy = 99.30%
(20, 21)	Categorizing of PD and HC	Detection	Collected data from PPMI	Random forest, MLP, SVM-RBF	Random forest accuracy = 88.2%
(22, 23)	Categorizing of PD and HC	Detection	Collected data from PPMI	RFS-LDA with 10-fold cross-validation	Accuracy = 79.8%
(24)	Categorizing of PD and HC	Detection	Collected data from PPMI	Random forest, MLP, SVM-RBF	Random forest accuracy = 88.2%
(25, 26)	Categorizing of PD and HC	Detection	UCI repository	LSTM, SVM with 5-fold cross-validation	LSTM accuracy = 98.1%, SVM accuracy = 93.7%
(27, 28)	Categorizing of PD and HC	Detection	Collected data from PPMI	CNN, Random Forest with cross-validation	CNN accuracy = 92.7%, Random Forest = 89.6%
(29, 30)	Categorizing of PD vs. Healthy	Detection	Patient's data	CNN, Random Forest with cross-validation	CNN accuracy = 92.7%, Random Forest = 89.6%

Methodology

In the proposed multimodal approach for detecting Parkinson's disease, we use three distinct machine learning models, each responsible for analyzing one of the primary data modalities: voice, gait, and handwriting. These models are based on Neural Networks (NN), with specific adaptations tailored to each modality's characteristics.

Voice Model (Neural Network)

The model examines voice characteristics including fluctuations in pitch, frequency irregularities (jitter), and amplitude variations (shimmer) to detect Parkinson's-related changes in vocal patterns. The Feedforward Neural Network (FNN) used here captures non-linear relationships between these features and the motor impairments that affect vocal cords in Parkinson's patients. By analyzing the subtle variations in voice, the model calculates a likelihood estimate indicating the probability of Parkinson's disease presence in the patient.

Figure 1 shows the machine learning-based method for analyzing voice data to detect Parkinson's disease. The process begins with collecting voice recordings (A), which are then converted into spectrograms (B) to highlight frequency variations over time. In the feature extraction step (C), vital acoustic attributes like pitch, jitter, and shimmer key markers of Parkinsonian hypokinetic dysarthria are identified. Subsequently, relevant features are selected (D) to improve the accuracy of classification. During classification (E), machine learning models such as Neural Networks or Support Vector Machines (SVM) analyze the chosen features to determine if the voice pattern suggests Parkinson's. The model's diagnostic performance is evaluated by a Receiver Operating Characteristic (ROC) curve (F), which measures sensitivity and specificity. Finally, a Neural Network calculates a likelihood ratio (G) representing the probability of a Parkinson's diagnosis. This approach supports findings from studies where voice analysis combined with machine learning effectively distinguishes Parkinson's patients from healthy individuals, showcasing voice as a promising biomarker for disease progression.

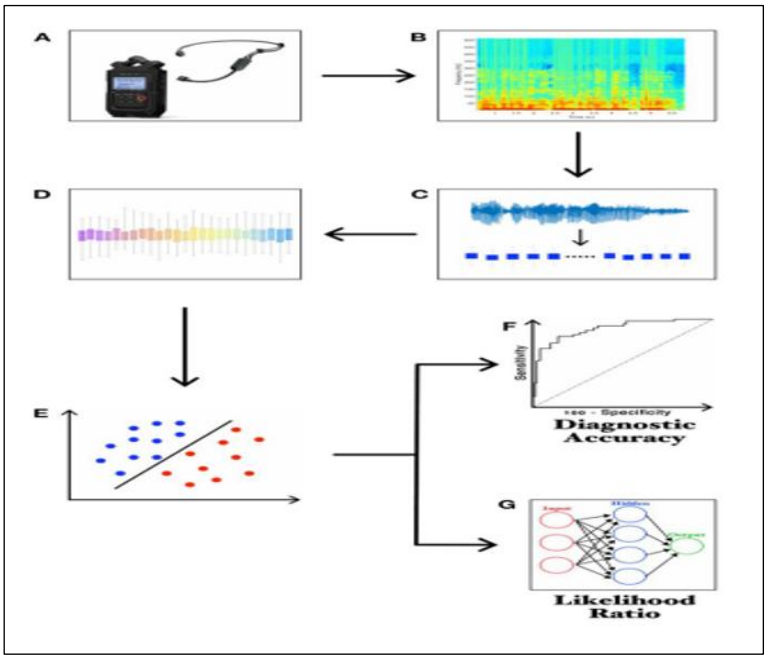

Figure 1: Voice Data Analysis in Parkinson's Disease Detection

Gait Model (Neural Network)

The gait model employs a NN, which is particularly suited to processing temporal gait data, such as stride length, gait speed, and cadence. This model analyzes the motor dysfunction evident in walking patterns, such as shorter steps or slower movement, which are characteristic of Parkinson's. The NN's ability to handle sequential data and refine its weight adjustments during training allows it to accurately classify abnormal gait patterns associated with Parkinson's disease.

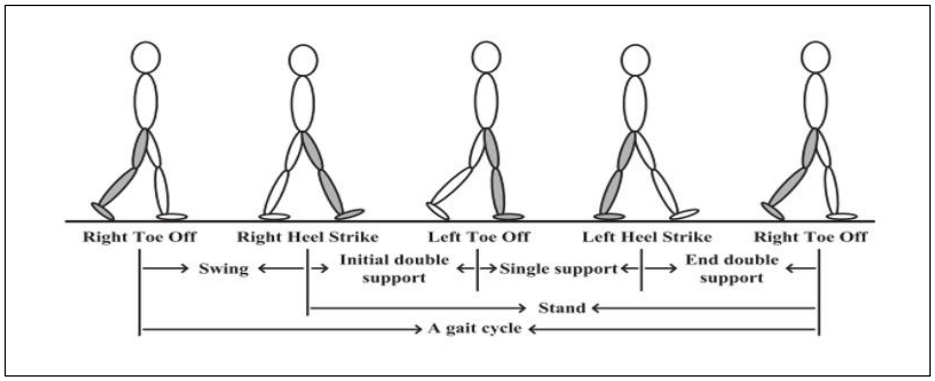

Figure 2: Phases of the Gait Cycle from Toe Off to Heel Strike

The Figure 2 illustrates the phases of a gait cycle, which is the sequence of movements involved in walking. It begins with the right toe off, where the right foot lifts off the ground, initiating the swing phase. The right heel strike follows, marking the end of the swing phase and the beginning of the stance phase, as the body's weight shifts to the right leg. Simultaneously, the left toe off occurs, initiating the left leg's swing phase. The Figure1 shows the left heel strike, where the left foot makes contact with the ground, transitioning the left leg into the stance phase. The cycle includes periods such as initial double support, where both feet briefly touch the ground, and single support, where only one foot supports the body as the other swings forward. The end double support marks the final phase before the next gait cycle begins. A complete gait cycle is defined by the movement of one foot through these phases, returning to its starting position.

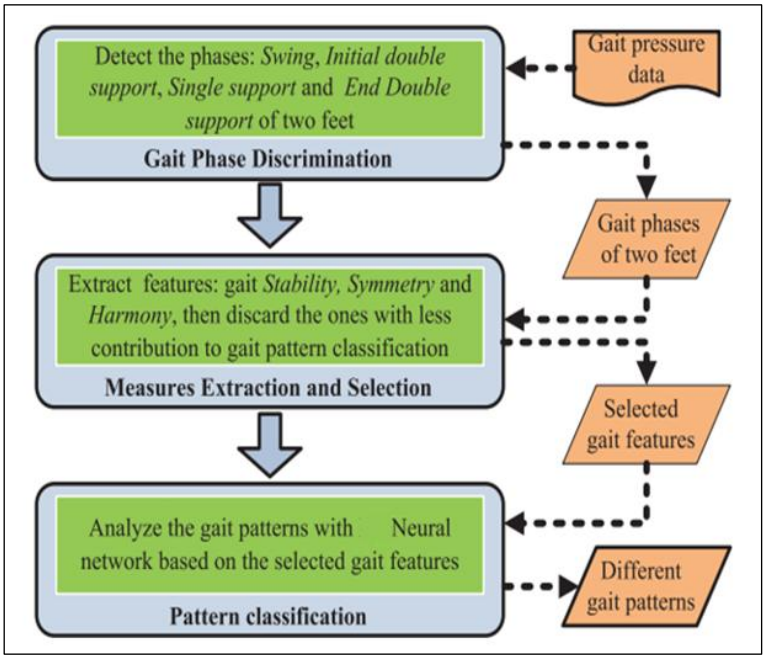

Figure 3: Gait Phase Detection and Neural Network-Based Pattern Classification Framework

Figure 3 illustrates the process for analyzing gait patterns, beginning with Gait Phase Discrimination, where distinct phases such as swing, initial double support, single support, and end double support are identified from pressure data of both feet. In the Measures Extraction and Selection step, gait features like stability, symmetry, and harmony are extracted, with less relevant features being discarded to optimize classification. Finally, in the Pattern Classification phase, a Backpropagation (BP) Neural Network processes the selected features to categorize different gait patterns. This systematic approach enables detailed gait analysis and is valuable for diagnosing disorders such as Parkinson's disease by recognizing characteristic Parkinsonian gait patterns through sophisticated movement feature analysis.

Handwriting Model (Neural Network)

The Figure 4 shows for the process of handwriting analysis, we use a neural network that processes features like stroke length, writing speed, and pressure. Parkinson's patients often experience micrographic (small handwriting) and tremor-induced irregularities, and this model is designed to detect those changes. By capturing both the speed and pressure variability in handwriting, the network can distinguish between Parkinson's-affected handwriting and normal patterns.

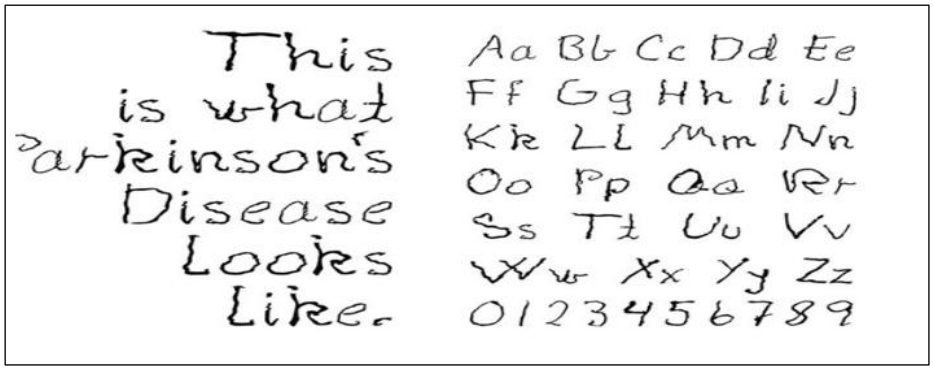

Figure 4: Handwriting Characteristics of Parkinson's Disease

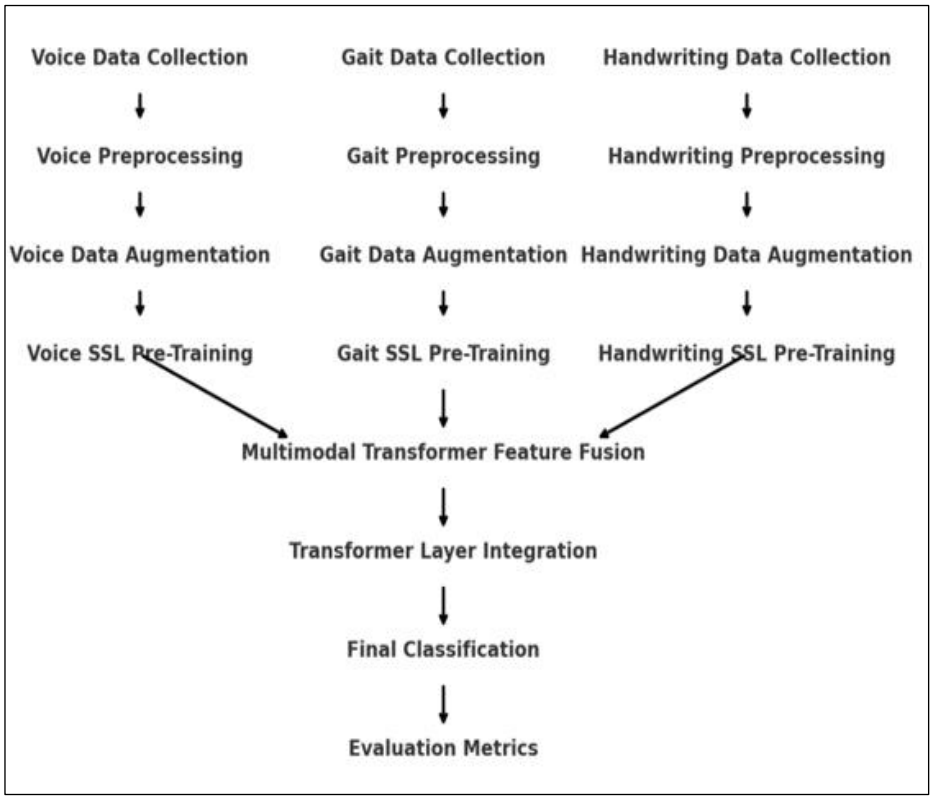

Figure 5: Multimodal Parkinson's Disease Detection Workflow

Multimodal Fusion Workflow

The methodology integrates voice, gait, and handwriting biometrics using Self-Supervised Learning (SSL) for feature extraction and a Multimodal Transformer for feature fusion as shown in the Figure 5.

Data Collection and Preprocessing

The dataset sourced from the University of California; Irvine Machine Learning Repository is a publicly available collection primarily used for research and educational purposes. It includes multimodal data for each subject, typically

involving biomedical voice measurements from individuals with Parkinson's disease and healthy controls. The data is collected in ASCII CSV format and consists of various features such as voice measures, patient demographics, and clinical scores. These datasets support empirical studies aimed at detecting Parkinson's disease through voice analysis and other modalities, facilitating research in understanding disease progression and developing diagnostic models.

For each subject i, collect multimodal data as shown in equation [1]:

$$X_i \leftarrow \left\{x_i^{(v)}, x_i^{(g)}, x_i^{(h)}\right\} \quad [1]$$

Where:

$x_i^{(v)} \leftarrow$ voice data,

$x_i^{(g)} \leftarrow$ gait data,

$x_i^{(h)} \leftarrow$handwriting data.

In preprocessing use normalization technique to extract each modality to remove noise and artifacts:
Voice: Remove background noise and extract acoustic features.
Gait: Correct motion artifacts and capture stride length, gait speed.
Handwriting: Clean data and extract stroke dynamics.

Data Augmentation

For each modality, apply stochastic augmentations to create variability as shown the equation [2], [3] and [4]:

$$x_i^{(v,a)} \leftarrow t_1 x_i^{(v)}, x_i^{(v,b)} \leftarrow t_2 x_i^{(v)} \quad [2]$$

$$x_i^{(g,a)} \leftarrow t_1 x_i^{(g)}, x_i^{(g,b)} \leftarrow t_2 x_i^{(g)} \quad [3]$$

$$x_i^{(h,a)} \leftarrow t_1 x_i^{(h)}, x_i^{(v,b)} \leftarrow t_2 x_i^{(h)} \quad [4]$$

Where $t_1$and $t_1$are different augmentation functions for each modality.

Self-Supervised Learning (SSL) Pre-Training

Extract features using SSL for each modality. The SimCLR framework is applied to maximize similarity between augmented pairs and minimize similarity with other samples. The contrastive loss function is given in equation [5]:

$$L_{SSL} \leftarrow -logexp \quad [5]$$

Where:

- z_i^a,and z_i^b are representations of the augmented inputs,
- $\left(z_i^a, z_i^b\right)$ is cosine similarity,
- τ is a temperature parameter.

Multimodal Transformer Feature Fusion

The extracted SSL features from each modality are fused using a Multimodal Transformer. For each modality, compute queries (Q), keys (K), and values(V):

$$Q \leftarrow W_q Z, K \leftarrow W_k Z, V \leftarrow W_v Z \quad [6]$$

Where $Z \leftarrow \left\{Z^{(v)}, Z^{(g)}, Z^{(h)}\right\}$

The attention score is computed as shown in equation (7):

$$ATTENTION(Q, K, V) \leftarrow softmax\left(\frac{QK_t}{\sqrt{d_k}}\right)V \quad [7]$$

This mechanism helps the model focus on relevant information from each modality.

Transformer Layer Integration

The output from the cross-attention mechanism is passed through multiple transformer layers to refine the representation can be referred from equation (8) and equation (9):

$$H^{(l)} \leftarrow LayerNorm\left(H^{(l-1)}\right) + Attention\left(H^{(l-1)}\right) \quad [8]$$

The final integrated representation after Llayers is:

$$H_L \leftarrow LayerNorm\left(H^{(l-1)}\right) + Attention\left(H^{(l-1)}\right) \quad [9]$$

Final Classification

The final fused representation H_L is used to predict whether a subject has Parkinson's disease. The final classification is made using a fully connected layer in equation (10):

$$y_i\hat{} \leftarrow \sigma(WH_L + b) \quad [10]$$

Where:

- W and $b$are weights and biases for the final classification layer,
- $\sigma(\cdot)$ is the sigmoid function, outputting a probability.

Evaluation Metrics

The model's performance is evaluated using metrics such as Accuracy, Precision, Recall, and F1-Score. The methodology integrates voice, gait, and handwriting data through Self-Supervised Learning (SSL) and Multimodal Transformer models to enable robust Parkinson's disease detection. This breakdown includes mathematical formulations explaining how multimodal data fusion is achieved using SSL combined with a Multimodal Transformer architecture for early diagnosis of Parkinson's disease.

Results and Discussion

This section outlines a comprehensive approach for Parkinson's disease detection using multimodal data (Voice, Gait, Handwriting) through self-supervised learning (SSL) for feature extraction, a multimodal transformer for feature fusion, and a final classification stage. It begins by applying augmentations to each modality followed by SSL pre-training with a contrastive learning objective, which maximizes the similarity between augmented pairs while minimizing similarity with other samples. The extracted features from each modality are then fused using a cross-attention mechanism in a multimodal transformer. This process allows the model to focus on relevant aspects from different modalities.

Experimental Evaluation

The experiment for Parkinson's disease detection utilizes a dataset of 1000 individuals, comprising 70% diagnosed with early-stage Parkinson's and 30% healthy controls. The data integrates multimodal inputs from three key sources: voice recordings, gait analysis, and handwriting samples. The Parkinson's Telemonitoring Voice Dataset is used to capture speech patterns, reflecting vocal impairments like tremors.

Table 2: Distribution of Parkinson's and Healthy Instances across Modalities

Data	Total Instances	Parkinson	Healthy
Voice	1000	700	300
Gait	1000	700	300
Handwriting	1000	700	300

The Table 2 presents data on 1,000 total instances each for Voice, Gait, and Handwriting modalities, with classifications into Parkinson's and Healthy categories. For each modality, 700 instances are associated with individuals diagnosed with Parkinson's disease, while 300 instances represent healthy individuals. This dataset provides a balanced view across different data types (Voice, Gait, Handwriting) to support analysis and model training for Parkinson's disease detection, offering a consistent distribution between the Parkinson's and Healthy categories across all modalities for comparison and potential multimodal analysis. The dataset is balanced to reflect the target distribution. Data pre-processing is crucial to ensure that noise does not impact the model's performance.

The Table 3 outlines the SSL pre-training process for voice, gait, and handwriting modalities in Parkinson's detection. Specific augmentations are applied to each modality, such as pitch shifting for voice, temporal jittering for gait, and stroke scaling for handwriting. The SSL model learns key features like vocal patterns, walking dynamics, and handwriting strokes. Contrastive loss (SimCLR) is used to differentiate augmented versions of the same input from others. This process enhances feature extraction for each modality, supporting early detection of Parkinson's disease.

Table 3: Self-Supervised Learning (SSL) pre-Training for Parkinson's Disease Detection

Modality	Augmentation Applied	Feature Learned	Loss Function
Voice	Pitch shifting, Time stretching	Vocal characteristics, speech patterns	Contrastive Loss (SimCLR)
Gait	Temporal jittering	Walking patterns, movement dynamics	Contrastive Loss (SimCLR)
Handwriting	Stroke scaling, Rotation	Stroke patterns, handwriting dynamics	Contrastive Loss (SimCLR)

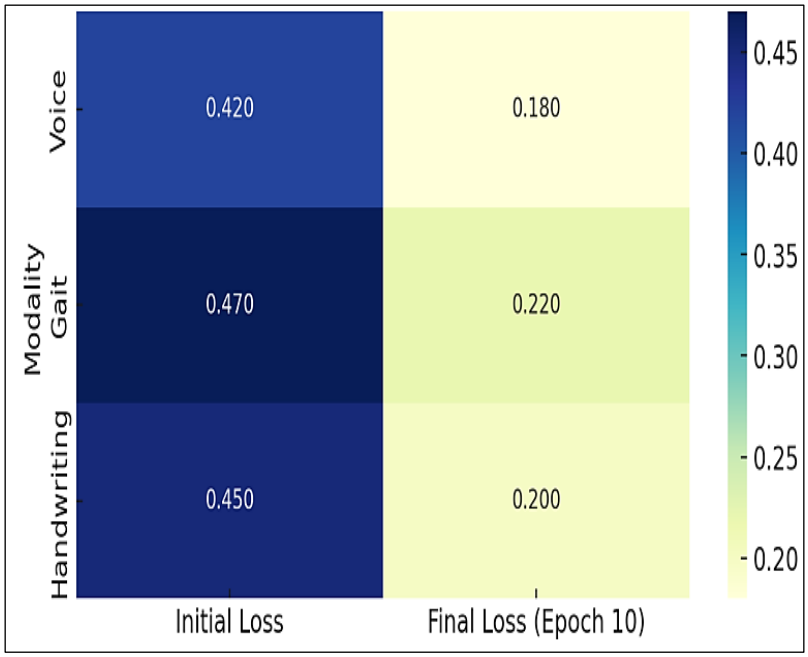

Figure 6: Initial vs Final Loss Comparison Across Modalities

The Figure 6 illustrates the initial and final loss values for three modalities: Voice, Gait, and Handwriting. The initial loss at the start of training is 0.420 for Voice, 0.470 for Gait, and 0.450 for Handwriting. By the 10th epoch, all three modalities show a marked improvement, with the loss decreasing to 0.180 for Voice, 0.220 for Gait, and 0.200 for Handwriting. Voice achieves the lowest final loss, indicating better performance in minimizing errors during training.

Multimodal Transformer Training

The SSL-extracted features from voice, gait, and handwriting are fused using a Multimodal Transformer with a cross-attention mechanism. The model learns how the modalities interact and integrates these features into a unified representation. The model is trained using binary cross-entropy loss on the training set. The output is shown in Table 4.

Table 4: Training and Validation Metrics over Epochs

Epoch	Training Loss	Validation Loss	Validation Accuracy
1	0.38	0.345	88.2
2	0.31	0.28	90.5
3	0.25	0.21	92.7
4	0.2	0.175	94.5
5	0.16	0.14	96.1

Training and validation loss steadily decrease, indicating good model generalization. The validation accuracy improves with each epoch, showing how the cross-attention mechanism effectively captures the interaction between the modalities.

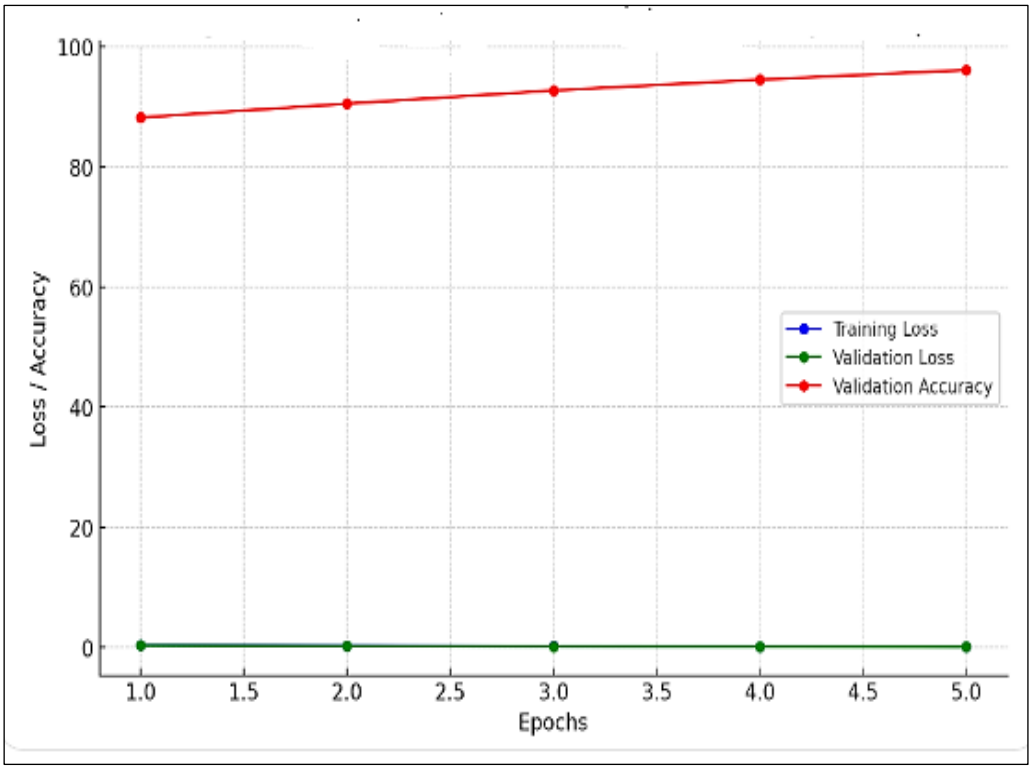

Figure 7: The Training, Validation Loss, and Validation Accuracy over Epochs

The Figure 7 shows the changes in training loss, validation loss, and validation accuracy over five training epochs. As training proceeds, both training and validation losses steadily decrease, indicating that the model is improving its ability to minimize errors on both datasets. The consistent decline in validation loss alongside training loss suggests the model is generalizing well without overfitting. Concurrently, the validation accuracy rises from 88.2% to 96.1%, showing that the model's predictions become increasingly accurate with each epoch. Overall, this figure reflects the progressive enhancement of the model's performance throughout the training process.

Model Testing and Evaluation

In this step, the model is evaluated on a previously unseen 30% test dataset, demonstrating strong performance across key metrics. The model achieves an accuracy of 96.5%, indicating a high rate of correct predictions. Precision, which reflects the proportion of true positive identifications among all positive results, stands at 95.2%, confirming the model's reliability in detecting actual Parkinson's cases. Recall, measuring the model's ability to identify all true positive cases, reaches 97.3%, indicating effective capture of genuine cases. The F1-score balances precision and recall, achieving 96.2%, which is crucial in Parkinson's disease diagnosis where misclassification can have serious implications. These results highlight the effectiveness of the multimodal approach that combines voice, gait, and handwriting data, enhancing overall detection accuracy. Figure 8 clearly depicts the high accuracy of 96.5% and the strong F1-score of 96.2%, showcasing the power of multimodal data fusion in Parkinson's disease detection.

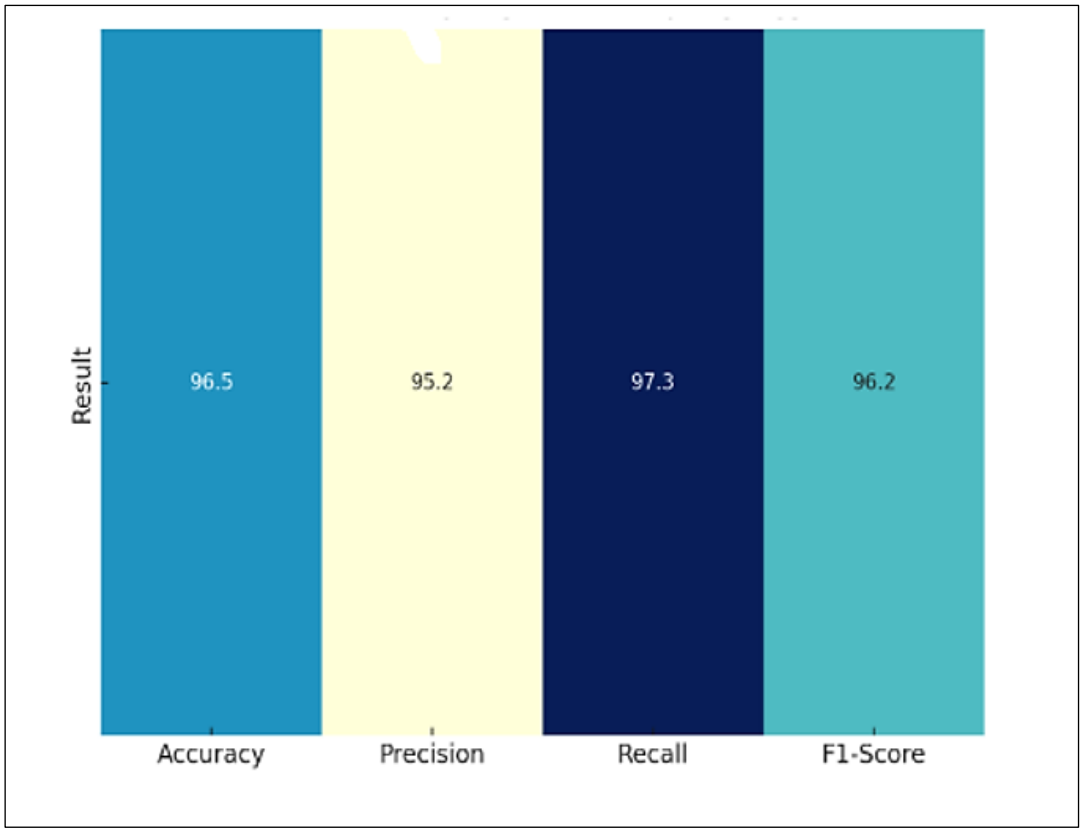

Figure 8: Performance Evaluation Metrics

Conclusion

The proposed method employs a machine learning framework that integrates multiple data types voice recordings, gait analysis, and handwriting examination to enhance early Parkinson's disease detection. Each modality is processed independently through neural networks designed to extract critical features, such as voice variations, movement irregularities, and handwriting changes. To improve feature learning without heavy reliance on labeled data, the approach incorporates Self-Supervised Learning (SSL), enabling the model to learn patterns directly from raw input. These extracted features are synergistically fused using a Multimodal Transformer that employs cross-attention mechanisms to explore interrelationships among data modalities. The system was tested on a dataset comprising 1,000 individuals, 70% of whom were diagnosed with early-stage Parkinson's, achieving an accuracy rate of 96.5%. This study demonstrates that combining diverse data sources with advanced learning architectures can significantly boost the accuracy and effectiveness of early Parkinson's disease detection.

Future Work

Future research should aim to expand the dataset by including larger and more demographically diverse populations to better evaluate the model's ability to generalize in real-world clinical settings. Upcoming work could incorporate additional data modalities such as brain imaging (MRI or PET scans), longitudinal patient monitoring, and clinical records to enhance model robustness and facilitate earlier symptom detection. Developing interpretable AI frameworks is also important to improve clinician trust and transparency. Furthermore, exploring federated learning or privacy-preserving techniques will help enable the use of sensitive medical data at scale while maintaining patient privacy. Finally, close collaboration with healthcare professionals for clinical trials and integration into routine diagnostic workflows is crucial to validate the model's practical impact on patient care and outcomes.

Abbreviations

BP – Back Propagation, CMTNN- Complementary Neural Networks, CNN – Convolutional Neural Networks, FNN – Feedforward Neural Network, HC – Healthy Controls. HMM: Hidden Markov Model,

MRI: Magnetic Resonance Imaging, RLDA: Regularized Linear Discriminant Analysis, ROC: Receiver Operating Characteristic, UCI: University of California Irvine.

Acknowledgement

This research has been supported by VVIET research Centre. we deeply grateful to unwavering support and shaped the core of this research.

Author Contributions

Karigowda Dhananjaya Kumar-Conceptualization, Methodology, Software, Data Curation, Writing-Original draft. Mirle Bhyraj Meghashree-Investigation, Writing-Original draft. Nagaraju Vinutha- Writing, review and editing, Dinesha Akasha- Validation, Review and Correction. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

The authors declare no conflicts of interest.

Declaration of Artificial Intelligence (AI) Assistance

The authors declare no use of Artificial intelligence (AI) for the write-up of the manuscript.

Ethics Approval

This research was conducted as per the ethical guidelines and principles.

Funding

The authors declare this research no external funding.

References

1. Subramani A, Elayaraja C, Vaidianathan B, et al. A multi-modal approach for early and accurate diagnosis of Parkinson's disease using advanced diagnostic techniques. Central Asian Journal of Medical and Natural Science. 2025;6(1):192-207.
2. Zhao A, Liu Y, Yu X, et al. Artificial Intelligence-enabled detection and assessment of Parkinson's disease using multimodal data: A survey. Information Fusion. 2025:103175. https://doi.org/10.1016/j.inffus.2025.103175
3. Benredjem S, Mekhaznia T, Rawad A, et al. Parkinson's disease prediction: An attention-based multimodal fusion framework using handwriting and clinical data. Diagnostics (Basel). 2024;15(1):4. http://doi: 10.3390/diagnostics15010004
4. Dentamaro V, Impedovo D, Musti L, et al. Enhancing early Parkinson's disease detection through multimodal deep learning and explainable AI: insights from the PPMI database. Scientific Reports. 2024;14(1):20941. https://doi.org/10.1038/s41598-024-70165-4
5. Srinivasan S, Ramadass P, Mathivanan SK, et al. Detection of Parkinson disease using multiclass machine learning approach. Scientific Reports. 2024;14(1):13813. https://doi.org/10.1038/s41598-024-64004-9.
6. Govindu A, Palwe S. Early detection of Parkinson's disease using machine learning. Procedia Computer Science. 2023;218:249-61. https://doi.org/10.1016/j.procs.2023.01.007
7. Mohaghegh M, Gascon J. Identifying Parkinson's disease using multimodal approach and deep learning. In: 2021 6th international conference on innovative technology in intelligent system and industrial applications (CITISIA). 2021. p. 1-6. IEEE. https://doi.org/10.1109/CITISIA53721.2021.9719945
8. Sha M, Rahamathulla M. Quantum deep learning in Parkinson's disease prediction using hybrid quantum–classical convolution neural network. Quantum Inf Process 23. 2024; 383. https://doi.org/10.1007/s11128-024-04588-3
9. Boshkoska BM, Miljković D, Valmarska A, Gatsios D, Rigas G, Konitsiotis S, Tsiouris KM, Fotiadis D, Bohanec M. Decision Support for Medication Change of Parkinson's Disease Patients. Comput Methods Programs Biomed. 2020 Nov; 196:105552. doi: 10.1016/j.cmpb.2020.105552
10. Kurmi A, Biswas S, Sen S, Sinitca A, Kaplun D, Sarkar R. An Ensemble of CNN Models for Parkinson's Disease Detection Using DaTscan Images. Diagnostics (Basel). 2022 May 8;12(5):1173. doi: 10.3390/diagnostics12051173
11. Aditi Govindu, Sushila Palwe. Early detection of Parkinson's disease using machine learning. Procedia Computer Science. 2023; 218:249-261. https://doi.org/10.1016/j.procs.2023.01.007
12. Ziaul Hasan Majumder M, Abu Khaer M, Nayeen Mahi MJ, Shaiful Islam Babu M, Aditya SK. Decision support technique for prediction of acute lymphoblastic leukemia subtypes based on artificial neural network and adaptive neuro-fuzzy inference system. In: Inventive Systems and Control: Proceedings of ICISC 2021. Singapore: Springer Singapore. 2021:539-54. doi: 10.1007/978-981-16-1395-1_40.
13. Yang S, Linares-Barranco B, Chen B. Heterogeneous ensemble-based spike-driven few-shot online learning. Frontiers in Neuroscience. 2022;9(16):850932. doi: 10.3389/fnins.2022.850932
14. Hamza A, Khan M A, Alhaisoni M, et al. D²BOF-COVIDNet: A framework of deep bayesian optimization and fusion-assisted optimal deep features for COVID-19 classification using chest X-ray and MRI scans. Diagnostics (Basel). 2022;13(1):101. https://www.mdpi.com/2075-4418/13/1/101
15. Hajar Rabie, Moulay A Akhloufi. A review of machine learning and deep learning for Parkinson's disease detection. Discover Artificial Intelligence (2025) 5:24. https://doi.org/10.1007/s44163-025-00241-9
16. Saravanan S, Ramkumar K, Adalarasu K, et al. A systematic review of Artificial Intelligence (AI) based approaches for the diagnosis of Parkinson's disease. Archives of Computational Methods in Engineering.

2022;29(6):3639-53. https://doi.org/10.1007/s11831-022-09710-1
17. Hasan M M, Deowan S A, Islam M A. Robot-assisted training for children with autism spectrum disorder: A review. Journal of Intelligent & Robotic Systems. 2023;108(3):41. https://link.springer.com/article/10.1007/s10846-023-01872-9
18. Ul Haq A, Li JP, Agbley BL, et al. A survey of deep learning techniques based Parkinson's disease recognition methods employing clinical data. Expert Systems with Applications. 2022;208:118045. doi: 10.1016/j.eswa.2022.118045
19. Islam MA, Majumder MZ, Hussein MA. Chronic kidney disease prediction based on machine learning algorithms. Journal of Pathology Informatics. 2023;14:100189. https://pubmed.ncbi.nlm.nih.gov/36714452/
20. El Maachi I, Bilodeau GA, Bouachir W. Deep 1D-Convnet for accurate Parkinson disease detection and severity prediction from gait. Expert Systems with Applications. 2020;143:113075. https://dl.acm.org/doi/10.1016/j.eswa.2019.113075
21. Aversano L, Bernardi ML, Cimitile M, et al. Early detection of Parkinson disease using deep neural networks on gait dynamics. In: 2020 International Joint Conference on Neural Networks (IJCNN). IEEE. 2020;19:1-8.
22. Nagasubramanian G, Sankayya M. Multi-variate vocal data analysis for detection of Parkinson disease using deep learning. Neural Computing and Applications. 2021;33(10):4849-64.
23. Nissar I, Mir WA, Shaikh TA. Machine learning approaches for detection and diagnosis of Parkinson's disease - A review. In: 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE. 2021;1:898-905.
24. Amprimo G, Masi G, Olmo G, et al. Deep Learning for hand tracking in Parkinson's Disease video-based assessment: Current and future perspectives. Artificial Intelligence in Medicine. 2024;154:102914. https://www.sciencedirect.com/science/article/pii/S0933365724001568
25. Islam MA, Majumder MZ, Hussein MA, et al. A review of machine learning and deep learning algorithms for Parkinson's disease detection using handwriting and voice datasets. Heliyon. 2024;10(3):e25469. https://www.cell.com/heliyon/fulltext/S2405-8440(24)01500-7
26. Islam N, Turza MS, Fahim SI, et al. Advanced Parkinson's disease detection: A comprehensive Artificial Intelligence approach utilizing clinical assessment and neuroimaging samples. International Journal of Cognitive Computing in Engineering. 2024;5:199-220.
27. Kaagman DGM, van Wegen EEH, et al. Effects and mechanisms of exercise on Brain-Derived Neurotrophic Factor (BDNF) levels and clinical outcomes in people with Parkinson's Disease: A systematic review and meta-analysis. Brain Sciences. 2024;14(3):194. doi: 10.3390/brainsci14030194
28. Kumari A, Gupta M, Rankawat G, et al. Parkinson disease: A review. Asian Journal of Pharmaceutical Research and Development. 2023;11(5):35-40.
29. Ibrahim AM, Mohammed MA. A comprehensive review on advancements in Artificial Intelligence approaches and future perspectives for early diagnosis of Parkinson's disease. International Journal of Mathematics, Statistics, and Computer Science. 2024;2:173-82.
30. Carneiro P, Loureiro JA, Delerue-Matos C, et al. Nanostructured label–free electrochemical immunosensor for detection of a Parkinson's disease biomarker. Talanta. 2023;252:123838. https://www.sciencedirect.com/science/article/pii/S0039914022006348

How to Cite: Meghashree MB, Kumar KD, Vinutha N, Vinay MG, Akash D. The multimodal fusion of voice, gait, and handwriting detection of Parkinson's disease using machine learning. Int Res J Multidiscip Scope. 2025; 6(4):1062-1074. doi: 10.47857/irjms.2025.v06i04.06552