JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

MULTIMODAL MACHINE LEARNING APPROACH FOR PARKINSON'S DISEASE DETECTION

[1]Mrs. P. A. Bidwai, [2]Puja Waykole, [3]Dhanshri Ithape, [4]Tanishka Andhare, [5]Ketaki Mule
[1]Guide, [2345]Students
Department of Information Technology
Pimpri Chinchwad Polytechnic, Pune, Maharashtra, India

Abstract : This paper presents a Flask-based web application for early Parkinson's disease detection using multimodal machine learning analysis of spiral drawings and voice recordings. The drawing model employs a Support Vector Machine (SVM) with RBF kernel using HOG features (tremors, edge density, statistical texture measures) extracted from 128×128 grayscale spirals via OpenCV. The voice model utilizes a Random Forest classifier trained on acoustic biomarkers including jitter, shimmer, HNR, and related speech features. Both models generate probability predictions that are fused through weighted combination into Low/Moderate/High risk categories. The system includes user authentication, SQLite screening history, Chart.js dashboards tracking detection trends, and printable PDF reports. Deployed with sub-2-second inference latency, it enables scalable non-invasive screening as a clinical decision support tool.

IndexTerms **- Parkinson's Disease, SVM, Random Forest, HOG Features, Voice Biomarkers, Multimodal Fusion, Flask Web Application.**

I. INTRODUCTION

Parkinson's disease (PD) affects over 10 million people worldwide, with early detection remaining challenging due to subtle motor and vocal symptoms. Traditional diagnosis relies on clinical observation and UPDRS scales, lacking scalable screening tools for mass deployment. This project addresses this gap through a web-based multimodal AI system analysing spiral drawing patterns and voice recordings for non-invasive PD screening.

The Flask application implements two SVM classifier for drawing analysis and a Random Forest classifier for voice analysis, fused via weighted probability scoring. Key contributions include real-time risk categorization, SQLite screening history, interactive dashboards, and printable clinical reports. The system enables early detection for timely neurological consultation while maintaining user privacy through secure session management.

II LITERATURE SURVEY

Parkinson's disease (PD) detection using machine learning has gained significant research attention due to its potential for early and non-invasive diagnosis. Most prior studies focus on unimodal approaches, particularly voice-based or handwriting-based analysis.

Voice-based detection systems analyze acoustic biomarkers such as jitter (cycle-to-cycle frequency variation), shimmer (amplitude variation), harmonics-to-noise ratio (HNR), and mel-frequency cepstral coefficients (MFCCs). Tsanas et al. [1] demonstrated that nonlinear speech features combined with Support Vector Machine (SVM) classifiers achieved high classification accuracy for distinguishing Parkinson's patients from healthy individuals. Their study established voice biomarkers as reliable indicators of early motor speech impairment.

Handwriting and spiral drawing analysis have also been widely explored to detect tremor-related motor dysfunction. Kotsavasiloglou et al. [2] utilized Histogram of Oriented Gradients (HOG) features and edge-based image processing techniques to classify Parkinson's disease using machine learning models. Their results showed strong discrimination capability based on spatial irregularities in spiral drawings.

Similarly, Vásquez-Correa et al. [3] investigated tremor frequency and drawing irregularities using SVM-based classification and reported promising detection accuracy. Their work emphasized the importance of motor pattern variability in PD diagnosis.

Although unimodal approaches provide good accuracy, they may lack robustness when used independently. Multimodal systems combining both voice and drawing features can improve reliability and predictive stability. However, deployable web-based multimodal platforms for real-time screening remain limited.

The proposed work builds upon these studies by integrating HOG-based spiral drawing analysis and Librosa-extracted voice biomarkers using weighted probability fusion of heterogeneous ML models within a Flask-based web application for scalable screening.

Reynolds explained the core concepts behind automatic speaker recognition systems and described the main techniques used to identify individuals based on their voice characteristics [4]. Dalal and Triggs proposed the Histogram of Oriented Gradients (HOG) method, which became an important feature extraction technique for object detection tasks in computer vision applications [5]. Zhao et al. reviewed recent developments in artificial intelligence-based methods for detecting and evaluating Parkinson's disease using multiple types of data, showing how combining different data sources can improve diagnostic results [6].

Skaramagkas et al. systematically analyzed multi-modal deep learning models for Parkinson's disease diagnosis and discussed how advanced neural network structures enhance prediction accuracy in medical applications [7]. Drotár et al. examined handwriting patterns associated with Parkinson's disease and presented updated findings to improve recognition using digital handwriting analysis techniques [8]. Rundo and Miano developed a deep learning-based multimodal framework for Parkinson's disease diagnosis, focusing on integrating various biomedical inputs for better assessment [9]. Nagarathna et al. introduced an optimized machine learning approach that combines handwriting and voice features for early Parkinson's disease detection, achieving enhanced classification performance [10].

Table 2.1: Literature comparison

Study	Modality	Features Used	Classifier	Reported Accuracy
Tsanas et al. (2010) [1]	Voice	Jitter, Shimmer, HNR	SVM	88.5%
Kotsavasiloglou et al. (2021) [2]	Drawing	HOG, Edge Density	Random Forest	87.2%
Vásquez-Correa et al. (2022) [3]	Drawing	Tremor Frequency	SVM	86.1%
Proposed Work	Multimodal	HOG + Voice Features	Hybrid (SVM + RF) Weighed Fusion	98.41%

III. RESEARCH METHODOLOGY

The proposed system follows a structured multimodal machine learning pipeline for early detection of Parkinson's disease. The methodology consists of data acquisition, preprocessing, feature extraction, classification, multimodal fusion, and result visualization within a web-based framework.

3.1 System Architecture

The overall system is implemented using a Flask-based web application integrating machine learning models and database management. The architecture consists of the following major components:

1. User Interface (Frontend)
2. Backend Processing Module
3. Machine Learning Models
4. Fusion Algorithm
5. SQLite Database
6. Visualization Dashboard

The system workflow is illustrated in Fig. 3.1.

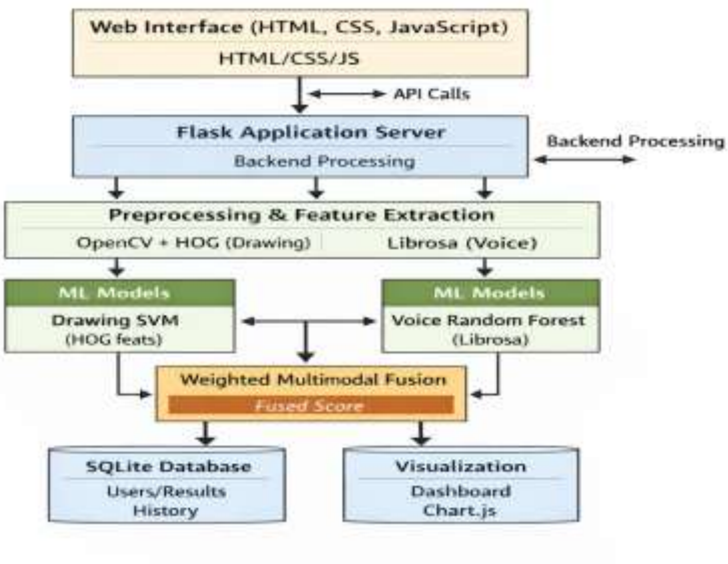

Figure. 3.1: System architecture of the proposed multimodal Parkinson's disease detection system

3.2 Data Acquisition

The system collects data through a web-based interface using both live and upload options. Users can either draw a spiral using the digital canvas or upload a spiral image (128 × 128 grayscale format). Similarly, users can record their voice live or upload a pre-recorded voice sample. The platform supports both registered and guest users. User details and screening results are stored in a SQLite database for future reference.

3.3 Drawing Data Processing

The uploaded or live-captured spiral drawing is first converted into a 128 × 128 grayscale image to maintain uniform input size. Image normalization is performed to standardize pixel intensity values and reduce illumination variations.

Noise reduction and edge enhancement are applied using OpenCV-based preprocessing techniques to improve tremor pattern visibility. The processed image is then used to extract Histogram of Oriented Gradients (HOG) features, which capture edge density, stroke direction, and tremor-related irregularities.

These extracted HOG feature vectors are forwarded to the trained Support Vector Machine (SVM) classifier for risk prediction.

3.4 Voice Preprocessing and Feature Extraction

The uploaded or live-recorded voice sample is first converted into a standardized audio format with uniform sampling rate. Background noise reduction and amplitude normalization are applied to improve signal clarity.

Using the Librosa library, relevant acoustic biomarkers are extracted, including jitter, shimmer, harmonics-to-noise ratio (HNR), and Mel-Frequency Cepstral Coefficients (MFCCs). These features capture vocal instability and frequency variations commonly associated with Parkinson’s disease. The extracted feature vector is then provided to the trained Random Forest classifier for probability-based risk prediction.

3.5 Classification Using Machine Learning Models

The drawing features are classified using a Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel and regularization parameter $C = 12$.

The voice features are classified using a Random Forest classifier with multiple decision trees to improve generalization and robustness.

Each classifier generates probability scores representing the likelihood of Parkinson’s presence. These probabilistic outputs enable reliable multimodal fusion and interpretable screening results.

3.6 Multimodal Fusion and Risk Categorization

The probability outputs from the drawing and voice models are combined using weighted probability fusion.

The final combined score is mapped into three risk categories:

- Low Risk
- Moderate Risk
- High Risk

This multimodal fusion improves screening robustness compared to single-modality analysis.

3.7 System Implementation and Deployment

The complete system is implemented as a Flask-based web application.

It includes:

- User authentication and secure session handling
- SQLite database for screening history
- Interactive dashboards using Chart.js
- Downloadable PDF reports for clinical reference

The system achieves sub-2-second inference latency, enabling scalable and non-invasive preliminary screening support.

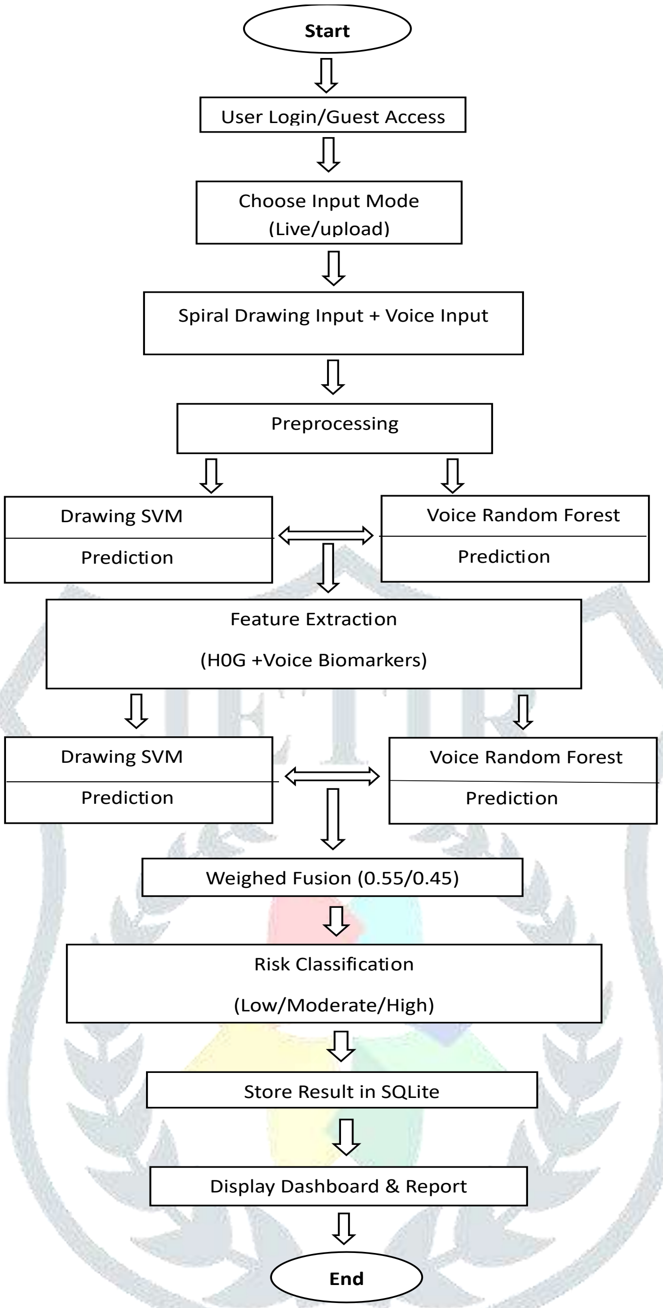

Figure 3.2: Stepwise Operational Flow of the Proposed Multimodal Screening System

IV. RESULTS AND DISCUSSIONS

4.1 Model Performance Evaluation

Table 4.1: Classification Performance of Proposed System

Model	Accuracy	Precision	Recall	F1-Score
Drawing (SVM-RBF kernel)	99%	100%	99%	99%
Voice (Random Forest Classifier)	95%	92%	85%	89%
Multimodal Fusion	98%	100%	97%	98%

Table 4.1 shows the performance evaluation of the proposed Parkinson's disease detection system. The results compare the drawing model, voice model, and the multimodal fusion model using standard metrics such as accuracy, precision, recall, and F1-score.

The multimodal fusion approach combines predictions from both modalities, which improves overall classification reliability and provides better screening performance compared to individual models.

4.2 System Performance Analysis

The implemented Flask-based web application demonstrates real-time prediction capability with inference latency below two seconds. The system supports both live input (drawing canvas and voice recording) and file upload options, ensuring flexibility for users.

Key observed outcomes include:

- Stable probability-based risk scoring
- Clear categorization into Low, Moderate, and High risk
- Secure storage of screening history in SQLite
- Interactive visualization using Chart.js dashboards

The fusion model provides balanced decision-making by weighting drawing probability (55%) and voice probability (45%), thereby minimizing dependency on a single modality.

4.3 Discussion

The experimental implementation confirms that multimodal screening provides better interpretability compared to single-modality systems. The integration of motor pattern analysis and speech biomarkers enables comprehensive symptom assessment.

Unlike traditional clinical screening methods that require physical consultation, this system offers:

- Non-invasive digital screening
- Scalable web-based deployment
- Secure session management
- Printable PDF-based clinical reports

Although the system is not intended to replace neurological diagnosis, it serves as an effective preliminary decision-support tool for early Parkinson's risk identification.

V. Acknowledgment

The authors express their sincere gratitude to Mrs. P. A. Bidwai, Department of Information Technology, Pimpri Chinchwad Polytechnic, Pune, for her valuable guidance, continuous support, and encouragement throughout the development of this research work.

The authors also thank the faculty members of the Information Technology Department for their technical assistance and constructive suggestions.

References

[1] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, "Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 884–893, 2010.

[2] S. Kotsavasiloglou, E. Kostikis, D. Hristu-Varsakelis, and M. Arnaoutoglou, "Machine learning-based classification of Parkinson's disease using spiral drawing analysis," *Biomedical Signal Processing and Control*, vol. 63, 2021.

[3] J. C. Vásquez-Correa, T. Arias-Vergara, M. Schuster, and E. Nöth, "Multimodal assessment of Parkinson's disease using speech and handwriting signals," *Computer Methods and Programs in Biomedicine*, vol. 208, 2022.

[4] D. A. Reynolds, "An overview of automatic speaker recognition technology," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2002.

[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

[6] A. Zhao, Y. Liu, X. Yu, and X. Xing, "Artificial intelligence-enabled detection and assessment of Parkinson's disease using multimodal data: A survey," *arXiv preprint arXiv:2502.10703*, 2025.

[7] V. Skaramagkas, A. Pentari, Z. Kefalopoulou, and M. Tsiknakis, "Multi-modal deep learning diagnosis of Parkinson's disease—A systematic review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 2399–2423, 2023.

[8] P. Drotár, T. Zdešar, and J. Mašek, "Analysis of handwriting for recognition of Parkinson's disease: Current state and new study," *Electronics*, vol. 13, no. 19, 3962, 2024.

[9] F. Rundo and A. Miano, "Research on multi modal Parkinson's disease diagnosis system based on deep learning model," *Progress in Oncology*, 2024.

[10] C. R. Nagarathna et al., "Optimized technique for the early identification of Parkinson's disease using machine learning-based handwriting and voice analysis," *International Journal of Innovative Research and Scientific Studies*, vol. 8, no. 3, pp. 2906–2916, 2025.