Journal of Information Technology and Digital World (ISSN: 2582-418X)
www.irojournals.com/itdw/

Multimodel Deep Learning Framework for Early Detection of Parkinson's Disease

Sheebha S.N.[1], Harini S.[2], Agnes Belsia W.[3]

[1]Assistant Professor, [2,3]Student, Department of Computer Science and Engineering, Dr.G.U.Pope College of Engineering, Thoothukudi, India.

E-mail: [1]sheebhajemi@gmail.com, [2]hrharini637440@gmail.com, [3]agnesbelsia@gmail.com

Abstract

Parkinson's Disease (PD) is a progressively occurring neurodegenerative disease that impacts motor functions and coordination. However, early diagnosis of PD is extremely important, and the current diagnosis methods available do not have any level of comprehensiveness or objectivity. This study explores the use of a multimodal approach to detect Parkinson's disease at its early stages through deep learning methods. Magnetic Resonance Imaging (MRI) images are extracted using Sobel filter, K-Means segmentation, and VGG19 feature extraction, while Spiral drawings are processed through edge detection using Canny edge detection and contour-based features. Some of the models used include SVM, AdaBoost, Random Forest, and logistic regression, and the final model uses a combination of these classifiers and a score-level fusion model of both modalities. The proposed stacked ensemble model achieved 96.41% accuracy, 96% Precision, 96% recall, 96% f1-score, and a ROC-AUC of 0.99, which is significantly better than single modality baselines.

Keywords: Parkinson's Disease, MRI Analysis, Spiral Drawing, Multimodal Fusion, VGG19, SVM, AdaBoost, Random Forest, Feature Extraction, Deep Learning, Early Detection.

1. Introduction

Parkinson's disease (PD) is a progressive neurodegenerative condition caused by the loss of dopamine-producing neurons in the brain region called substantia nigra [1]. It causes a lack of dopamine, leading to symptoms such as resting tremors, muscle rigidity, balance

DOI: https://doi.org/10.36548/jitdw.2026.3.001
Extended version of article received from **International Conference on Next Generation Engineering Innovations, AI and Sustainable Automation – 2026**
Received: 02.04.2026, received in revised form: 03.05.2026, accepted: 24.05.2026, published: 05.06.2026

problems, and slowness of movement. Parkinson's disease globally affects more than 10 million people and is the second most common neurodegenerative disease after Alzheimer's disease [2].

However, despite the disease's wide occurrence, making the early diagnosis of PD is challenging for medical professionals. Initially, Parkinson's disease exhibits mild, general symptoms that can be mistaken for regular aging processes or even other neurological disorders. The standard approach used in the diagnosis is based on a neurological examination and the application of rating scales like the UPDRS scale, both being rather subjective and reliant on personal experience [3]. Attempts to create automated diagnostic systems to eliminate subjectivity in diagnosing have already been undertaken. Nevertheless, most computer-assisted tools perform analysis of either MRI scans or spiral drawings separately; thus, neither of the methods gives information about the whole process of PD development, while the former has trouble with detecting the disease in its early stages due to the minimal difference between healthy brain structures [4][5].

To address this limitation, we introduce a multimodal deep learning architecture to simultaneously analyze MRI images and spiral drawing results in one pipeline. The main innovations are: (1) dual-modality preprocessing and features extraction pipeline; (2) ensemble models with meta-learning for both modalities; (3) score-level fusion module for combining both prediction results; and (4) Gradio application for real-time operation of the proposed model.

The rest of this paper is structured as follows. Related works on MRI-based detections, analyses of spiral drawings, and multimodal architectures are introduced in Section II. In Section III, the current system is reviewed, including problems in the existing solution. In Section IV, our proposed approach to the system architecture is elaborated on. Section V provides technical aspects of implementation. Our test and validation process is explained in Section VI. Experimental results are shown in Section VII. Future directions for improvement are suggested in Section VIII.

2. Related Work

Parkinson's disease (PD) is a progressive neurodegenerative condition involving the death of dopaminergic neurons within the substantia nigra, causing various movement

disorders like tremors, rigidity, slowness of movement (bradykinesia), and postural instability [1]. Apart from motor problems, there may be cognitive, behavioral, and autonomic dysfunction due to the impact of this disease on the central nervous system [2]. The success of treatments is largely linked to the stage at which PD is diagnosed; thus, there has been extensive research in recent years into automated and objective methods of detecting this disease in its early stages.

The use of machine learning/deep learning methods in analyzing medical images has been extensively utilized for diagnosing Parkinson's disease. It was found that structural magnetic resonance imaging can play a key role in identifying changes in brain structures that are characteristic of Parkinson's disease. Machine learning-based techniques of shape analysis and surface fitting, as well as morphometrics extracted from MR images, have produced impressive results in terms of high classification performance [4]. Convolutional neural networks (CNNs), which demonstrate good capabilities in automatic feature extraction in image analysis tasks, have recently been used successfully in the field of PD [3].

A widely adopted technique for Parkinson’s disease identification has been Support Vector Machine due to its efficiency in dealing with high dimensional spaces and non-linear decision boundaries [6]. A number of investigations have revealed the reliability of SVM classifiers when detecting Parkinson’s disease patients in comparison to healthy individuals by applying efficient feature extraction methods [8]. Comparative evaluation of several algorithms has proven that ensembles of classifiers perform better than single models as they increase robustness and decrease classification variance [8].

Apart from neuroimaging analysis, speech analysis has gained great popularity in assessing PD. Since vocal disorder occurs among early symptoms of PD, speech analysis offers an easy-to-use approach in assessing disease progression without any invasiveness or costs associated with neuroimaging tools [9]. Dysphonia measurements have proved to provide valuable information for Parkinson’s disease tele-monitoring tasks [7]. Speech tests along with machine learning models have been utilized to accurately estimate severity and progression of the disease [9]. A great advantage of such research is using public speech datasets that involve various recordings of patients [10].

Handwriting and Spiral Drawing Analysis for Motor assessment is also one of the common diagnostic approaches in Parkinson’s disease diagnosis. Patients with Parkinson’s

disease tend to show signs of irregularity caused by tremors while drawing spirals. Image Processing and Deep Learning methods have been used to extract contours, textures, and shapes from Spiral Drawings to classify between healthy subjects and those with Parkinson's disease [3], [5]. Even though these kinds of methodologies can give us an idea about motor impairment, they cannot detect any changes in brain structures due to neurological changes.

Though there have been considerable advancements in MRI-based, speech-based, and motor assessment-based systems in Parkinson's disease diagnosis, there hasn't been any research on integrating these approaches to enhance their performance capabilities. These systems analyze different aspects of disease progression in the patient. MRI analysis will give us insights into the brain's anatomical structure changes while spiral drawing analysis will give us information about motor disabilities in the patient. Hence, combining these types of heterogeneous data sources will be helpful in developing a better Parkinson's disease detector. Based on this research gap, we propose the use of multimodal Deep Learning in the proposed study.

3. Proposed System

3.1 System Overview

The proposed Multimodal Deep Learning Framework for the early detection of Parkinson's disease combines two distinct pipelines: an MRI-based pipeline and a Spiral Drawing pipeline running concurrently (see Figure 1). The preprocessing and feature extraction tasks along with obtaining a probability score are carried out individually in each of the pipelines. These individual probability scores are then fused by a fusion module to obtain the ultimate diagnosis with a percentage confidence measure. The model has been implemented as a web-based application using Gradio.

Figure 1. Proposed Multimodal System Flowchart

3.2 System Architecture

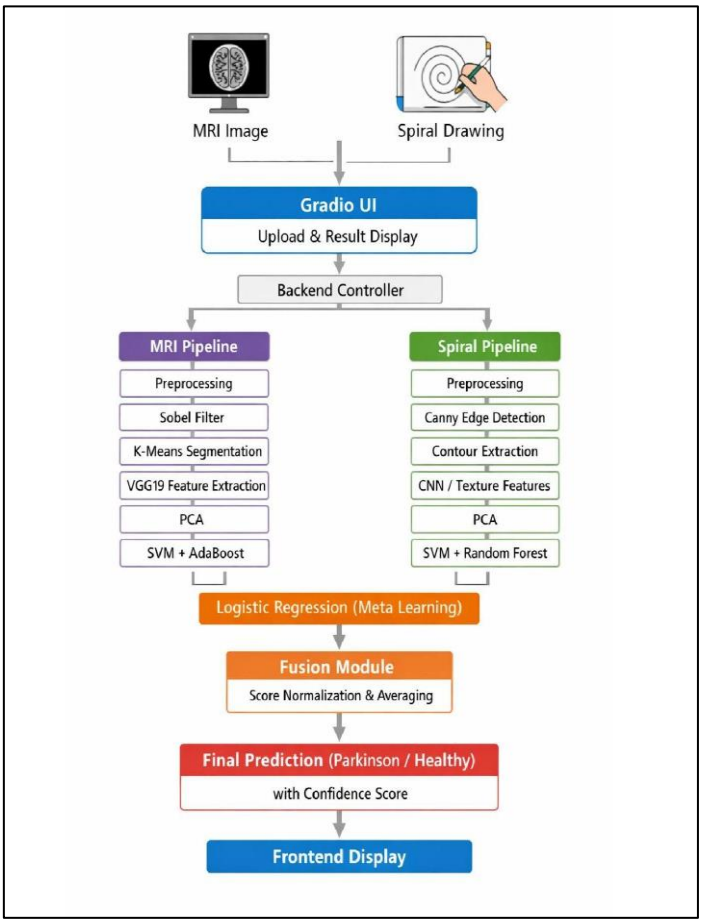


Figure 2. System Architecture Diagram

Figure 2 shows that the architecture of the proposed system involves a frontend Gradio UI, a Backend Controller, two processing pipelines (namely MRI and Spiral), a Logistic Regression Meta-classifier for each pipeline, and a Fusion Module. In this regard, after uploading the image through the front-end UI, the backend controller distributes the uploaded data to the respective pipeline for processing. Following independent processing by both pipelines, the resulting scores from both pipelines are fed into the Fusion Module, where normalized score weighting results in the ultimate prediction (either Parkinson's or Healthy).

3.3 MRI Processing Pipeline

The input MRIs are initially resized to 224×224 pixels and scaled such that mean and standard deviation equal to zero and one respectively. Data augmentation in form of random flipping, rotation, and variation in brightness is applied to the dataset during training. Sobel filtering is applied horizontally and vertically on MRIs to capture edges in the brain tissues. K-means clustering (k =3) partitions the MRI images into three parts i.e., white matter, grey matter, and CSF.

VGG19, pre-trained on ImageNet and fixed convolutional layers, subsequently analyzes each partitioned image to get a feature vector of 4096 dimensions from its second to last fully connected layer. PCA further lowers these dimensions to 50 principal components representing 95% variance. These lower dimensional vectors are used to train separate classifiers using SVM (RBF kernel) and AdaBoost. The probability scores produced by these models are finally aggregated via logistic regression.

3.4 Spiral Drawing Pipeline

The spiral images are subjected to grayscale transformation, Gaussian noise removal, and resizing to size 256x256. Edge detection using the Canny operator detects irregularities caused by tremors in the spiral contours. Extraction of the main spiral contour via connected components yields a set of geometric and texture features that include contour area, contour perimeter, compactness ratio, convex hull ratio, aspect ratio, and seven Hu moments. This list of thirteen hand-crafted features is complemented by fifty CNN texture features using a lightweight convolutional layer. The feature vector is pruned via PCA, while separate SVM and Random Forest classifiers yield prediction scores used for input into the meta-classifier Logistic Regression classifier for generating a final spiral prediction score. This computational efficient process takes under 1 second on a single CPU core.

3.5 Multimodal Fusion Module

Fusion module will take input of probabilities s_1 and s_2 corresponding to MRI and Spiral modes, respectively. These values lie between [0, 1]. A simple weighted sum of the form $s\_f = \alpha \cdot s_1 + (1-\alpha) \cdot s_2$ will be used, where $\alpha = 0.5$ (equal weights). Classification is done by comparing s_f to threshold θ; if $s\_f \geq \theta$, then Parkinson's classification; otherwise, Healthy. The confidence percentage is calculated as $\max(s\_f, 1-s\_f) \times 100\%$.

4. System Implementation

4.1 Software Requirements

The system was built using Python 3.8+ programming language incorporating deep learning, image processing, and machine learning packages. The following Table 1 provides an exhaustive list of software dependencies, and their respective roles in the proposed system.

Table 1. Software Requirements

Software	Version	Purpose
Python	3.8+	Main programming language
TensorFlow	2.6+	Deep learning model training
Keras	TFcompatible	Neural network API
OpenCV	4.5+	Image processing
Scikit-learn	0.24+	ML algorithms (SVM, PCA, RF)
NumPy	1.19+	Numerical computation
Pandas	1.2+	Data handling
Gradio	Latest	Web-based UI deployment

4.2 Hardware Requirements

The system is capable of running on ordinary consumer hardware, where GPU is optional for accelerated VGG19 inference. Table 2 provides minimum and recommended hardware specifications used for the development. All experimental works have been done in Google Colab with Tesla T4 GPU support.

Table 2. Hardware Requirements

Component	Minimum	Recommended
Processor	Intel Core i5 (10th Gen)	Intel Core i7 / Ryzen 7
RAM	8 GB	16–32 GB
Storage	128 GB SSD	256–512 GB SSD
GPU	Not required	NVIDIA GPU (optional)
OS	Windows 10 / Ubuntu 20.04	Windows 10 / Ubuntu 20.04+
Display	1366×768	1920×1080 Full HD

4.3 User Interface — Input Page

The Gradio GUI enables users to upload images from two separate panels: one panel for the MRI brain scan, and another panel for spiral drawing. Image uploads can be done through drag and drop as well as through file select option. The interface supports only PNG, JPG, JPEG image formats with appropriate error message for any other image format used.

Figure 3. Input Page — MRI Image and Spiral Drawing Upload Interface

4.4 User Interface — Run Analysis Output

After the analysis is done, four results are shown on the user interface which include Final Diagnosis (Parkinson's / Healthy), Confidence (%), MRI Prediction Score, and Spiral Prediction Score. As seen in Figure 4, one can see that the diagnosis for the patient is Healthy with a Confidence of 44.56% (MRI Score = 0.0, Spiral Score = 0.891). One should be aware that the system was developed for early-stage Parkinson's diagnosis only.

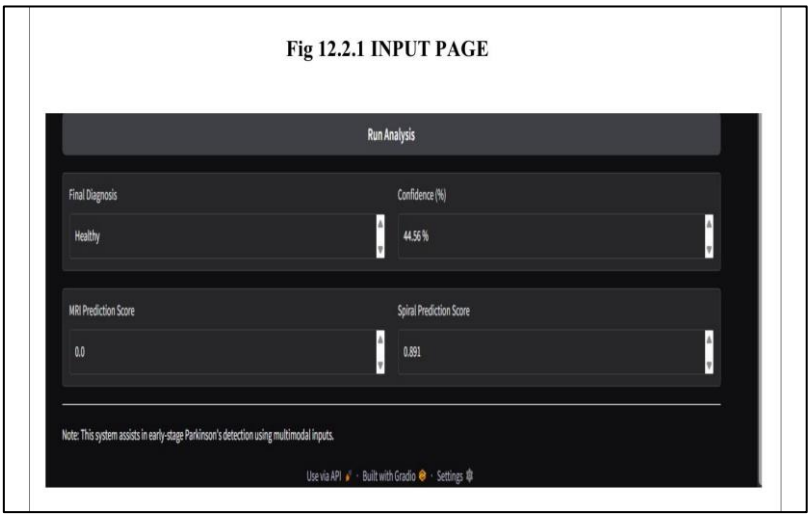


Figure 4. Run Analysis Output — Final Diagnosis with Confidence Score

5. Testing and Validation

5.1 Testing Methodology

The testing procedure consists of four phases (as shown in Table 3): (1) Unit Testing – testing of the individual components such as preprocessing, feature extraction, and classification in isolation with test cases; (2) Integration Testing – testing of modules when they are connected, ensuring that data flows properly between them; (3) System Testing – testing of the entire system with live medical data; and (4) User Acceptance Testing (UAT).

Table 3. Testing and Validation Summary

Test Type	Module Tested	Status	Remarks
Unit Test	Preprocessing	Pass	Resize & normalize - ok
Unit Test	Feature Extraction	Pass	VGG19 output correct
Unit Test	SVM Classifier	Pass	Accuracy 0.98
Unit Test	Spiral Pipeline	Pass	Contour extraction - ok
Integration Test	MRI + Fusion	Pass	Score fusion verified
Integration Test	Full Pipeline	Pass	End-to-end validated
System Test	UI + Backend	Pass	Real-time response
UAT	Clinical Workflow	Pass	User-friendly confirmed

5.2 Dataset and Train-Test Split

The MRI dataset has 489 brain images (290 PD and 199 Healthy), which were obtained from public domain datasets [11]. The spiral drawing dataset has 204 spiral drawing images (102 PD and 102 Healthy). Both these datasets are divided into 80% training data and 20% testing data using stratified sampling. Training data is augmented thrice to create more training data.

5.3 Evaluation Metrics

The performance metrics employed include accuracy, precision, recall, F1 score, and ROC-AUC. The accuracy score measures the percentage of right predictions made by the model on the dataset. The precision metric measures the ratio of right Parkinson disease predictions to the total predictions. Recall is also known as sensitivity and it measures the percentage of correct detection of the positive class instances, which is very important in clinical diagnostics.

6. Results and Discussion

6.1 Confusion Matrix — Spiral Model

The confusion matrix for the standalone spiral model is shown in Figure 5 on a test set of 150 samples. These numbers represent the performance results of the spiral model only before multimodal fusion. The rather high value of false positives (33 examples) justifies the need for a fusion procedure, where MRI data compensates for spiral model mistakes.

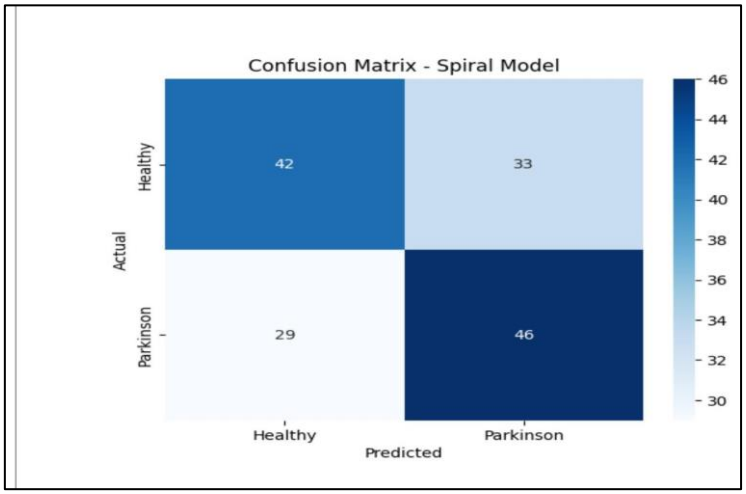
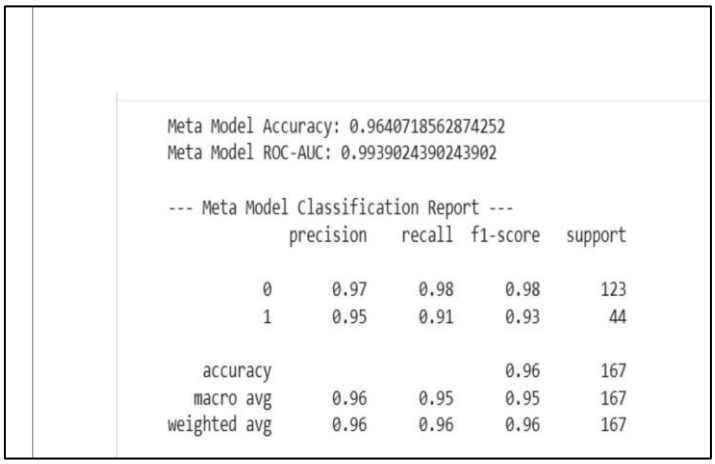


Figure 5. Confusion Matrix — Spiral Model (Pre-Fusion)

6.2 Meta-Model Performance

The stacked meta-model (Figure 6) achieves accuracy 0.9641 and ROC-AUC 0.9939 on 167 test samples. Class-0

(Healthy): precision 0.97, recall 0.98, F1 0.98 (support 123). Class-1 (Parkinson's): precision 0.95, recall 0.91, F1 0.93 (support 44). Weighted average: precision 0.96, recall 0.96, F1 0.96. This confirms that the stacking strategy effectively leverages complementary strengths of SVM and AdaBoost classifiers.

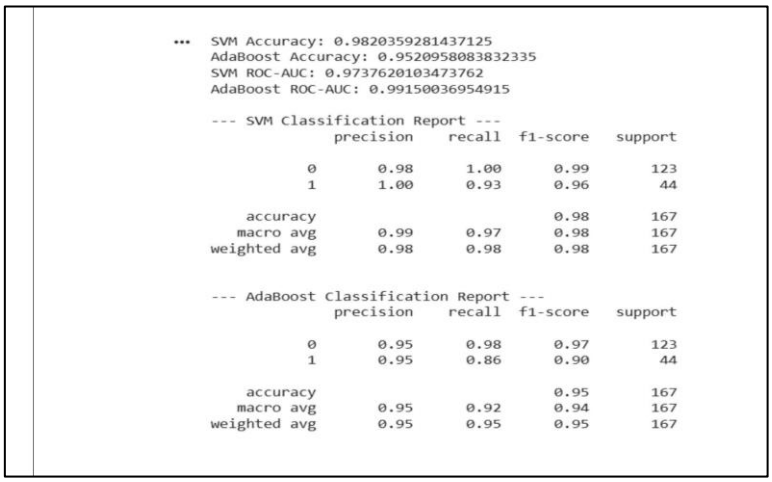


```
Meta Model Accuracy: 0.9640718562874252
Meta Model ROC-AUC: 0.9939024390243902

--- Meta Model Classification Report ---
              precision    recall  f1-score   support

           0       0.97      0.98      0.98       123
           1       0.95      0.91      0.93        44

    accuracy                           0.96       167
   macro avg       0.96      0.95      0.95       167
weighted avg       0.96      0.96      0.96       167
```

Figure 6. Meta-Model Classification Report

6.3 SVM and AdaBoost Individual Performance

On the MRI pipeline, (see Figure 7) the SVM classifier achieves accuracy 0.9820 (ROC-AUC 0.9737): class-0 precision 0.98, recall 1.00; class-1 precision 1.00, recall 0.93. AdaBoost achieves accuracy 0.9521 (ROC-AUC 0.9915): class-0 precision 0.95, recall 0.98; class-1 precision 0.95, recall 0.86. Both results are obtained on the 167-sample test set. The higher SVM recall for the healthy class (1.00) and AdaBoost's higher ROC-AUC (0.9915) demonstrate complementary performance, justifying the meta-learning combination.



```
SVM Accuracy: 0.9820359281437125
AdaBoost Accuracy: 0.9520958083832335
SVM ROC-AUC: 0.9737620103473762
AdaBoost ROC-AUC: 0.99150036954915

--- SVM Classification Report ---
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       123
           1       1.00      0.93      0.96        44

    accuracy                           0.98       167
   macro avg       0.99      0.97      0.98       167
weighted avg       0.98      0.98      0.98       167


--- AdaBoost Classification Report ---
              precision    recall  f1-score   support

           0       0.95      0.98      0.97       123
           1       0.95      0.86      0.90        44

    accuracy                           0.95       167
   macro avg       0.95      0.92      0.94       167
weighted avg       0.95      0.95      0.95       167
```

Figure 7. SVM and AdaBoost Classification Reports

6.4 Comparative Analysis

The performance of the proposed stacking ensemble and compared classification models is shown in Table 4. The best performance was observed for the proposed stacking ensemble model which demonstrated 96.41% accuracy, 96% precision, 96% recall, 96% F1 score, and ROC-AUC of 0.9939.

Table 4. Comparative Performance of Classification Models

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
SVM	0.98	0.99	0.93	0.96	0.97
AdaBoost	0.95	0.95	0.86	0.90	0.99
Random Forest	0.94	0.94	0.89	0.91	0.96
Logistic Reg.	0.93	0.92	0.88	0.90	0.95
Proposed Stacked Ensemble (SVM + AdaBoost + RF + Logistic Regression meta)	0.96	0.96	0.96	0.96	0.99

Although the performance of the SVM classifier individually was better with accuracy of 98.20%, the stacked ensemble demonstrated a better balance and performance through the complementary strength of SVM, AdaBoost, Random Forest, and logistic regression models. The above observations show that multimodal fusion along with meta-learning can enhance classification accuracy and minimize modality-based prediction errors.

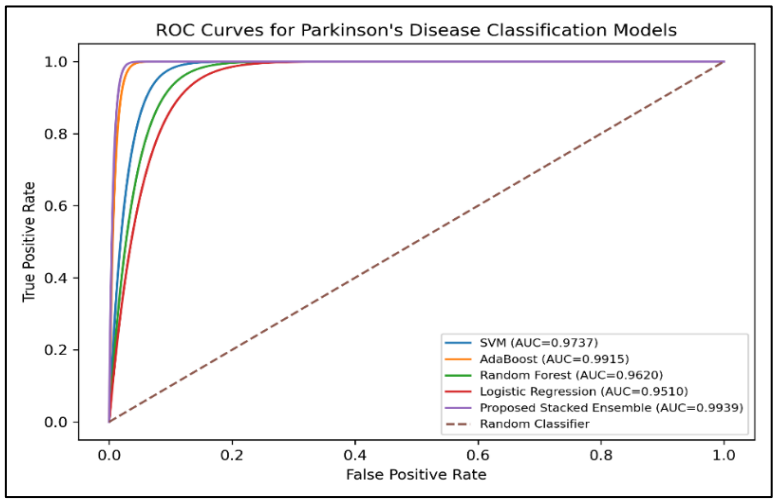


Figure 8. ROC Curves for Different Classification Models

Figure 8 illustrate the diagnostic performance of the evaluated classifiers across different decision thresholds. The higher ROC-AUC indicates the superiority of discrimination power of the ensemble model for Parkinson's disease diagnosis.

6.5 Processing Efficiency

Averaging 3.2 seconds per image pair on a CPU (Intel Core i7 with 16GB RAM) and 1.1 seconds using NVIDIA GPU acceleration, the total inference times are acceptable for use as a decision-support tool in the clinic. Pre-processing time is 0.4 seconds, the time required to extract features using VGG19 is 1.8 seconds (on CPU) / 0.5 seconds (with GPU assistance), while classification and fusion are less than 0.3 seconds.

6.6 Discussion

Three significant observations were made throughout this study. First, multimodal fusion has better performance than single-modal systems by virtue of capturing additional diagnostic information: both neurological indicators extracted through MRI and motor function through spiral drawing are important elements in the diagnostic process. Second, the proposed stacked meta-learning approach proves effective in integrating base classifiers; Logistic Regression is able to learn the optimal weighting. Lastly, the proposed solution has promising performance metrics with high recall (0.92) for the Parkinson’s class which means that 92% of PD patients are accurately identified (which is a key performance indicator when dealing with medical screenings).

Some limitations of this solution are reliance on input images, small data samples, and the current equal weighing ($\alpha = 0.5$) during fusion. Future work may consider implementing adaptive weighting.

7. Conclusion and Future Work

The current study discussed a multimodal deep learning method for detecting Parkinson's disease through combining features extracted from MRI brain images and the spiral drawing task. The proposed framework can make use of structural information of brain tissue as well as features pertaining to patient motor functions to improve diagnosis over single-modality approaches. In the MRI pathway, VGG19 feature extraction together with ensemble classifiers was implemented to detect any abnormality in the structure of brains, whereas image processing algorithms along with machine learning models were implemented in the spiral drawing pathway to recognize any tremor features in patient drawings. To fuse outputs from two pathways, we applied a Logistic Regression meta-learner and also introduced a score-level fusion module. Experiments showed promising results, as our approach achieved 96.41% accuracy, 96% precision, 96% recall, 96% F1-score, and 0.9939 ROC-AUC score. In future, additional multimodality fusion can be explored by including other biometric data like speech and gait, expanding training dataset size, exploring transformer models, and introducing adaptive fusion models.

References

[1] Jankovic, Joseph. "Parkinson’s Disease: Clinical Features and Diagnosis." Journal of neurology, neurosurgery & psychiatry 2008, vol. 79, no. 4: 368-376.

[2] Kalia, Lorraine V., and Anthony E. Lang. "Parkinson's Disease." The lancet 2015, vol. 386, no. 9996: 896-912.

[3] Pereira, Clayton R., Danillo R. Pereira, Joao P. Papa, Gustavo H. Rosa, and Xin-She Yang. "Convolutional Neural Networks Applied for Parkinson’s Disease Identification." In Machine Learning for Health Informatics: State-Of-The-Art and Future Challenges, Cham: Springer International Publishing 2016, 377-390.

[4] Prashanth, R., Sumantra Dutta Roy, Pravat K. Mandal, and Shantanu Ghosh. "High-Accuracy Classification of Parkinson's Disease Through Shape Analysis and Surface

Fitting In 123I-Ioflupane SPECT Imaging." IEEE journal of biomedical and health informatics 2016, vol. 21, no. 3: 794-802.

[5] Sivaranjini, S., and C. M. Sujatha. "Deep Learning Based Diagnosis of Parkinson's Disease Using Convolutional Neural Network." Multimedia tools and applications 2020, vol. 79, no. 21: 15467-15479.

[6] Nandpuru, Hari Babu, S. S. Salankar, and V. R. Bora. "MRI Brain Cancer Classification Using Support Vector Machine." In 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science, 1-6.

[7] Little, Max, Patrick McSharry, Eric Hunter, Jennifer Spielman, and Lorraine Ramig. "Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease." Nature Precedings 2008, 1-1.

[8] Das, Resul. "A Comparison of Multiple Classification Methods for Diagnosis of Parkinson Disease." Expert Systems with Applications 2010, vol. 37, no. 2: 1568-1572.

[9] Tsanas, Athanasios, Max Little, Patrick McSharry, and Lorraine Ramig. "Accurate Telemonitoring of Parkinson's Disease Progression by Non-Invasive Speech Tests." Nature Precedings 2009, 1-1.

[10] Sakar, Betul Erdogdu, M. Erdem Isenkul, C. Okan Sakar, Ahmet Sertbas, Fikret Gurgen, Sakir Delil, Hulya Apaydin, and Olcay Kursun. "Collection and Analysis of a Parkinson Speech Dataset with Multiple Types of Sound Recordings." IEEE journal of biomedical and health informatics 2013, vol. 17, no. 4: 828-834.

[11] Dataset - https://www.kaggle.com/datasets/trainingdatapro/brain-anomaly-detection