Early Detection of Parkinson's Disease Using Explainable AI on Speech and Handwritten Drawing Features

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE (COMPUTER SCIENCE)

SYAM KUMAR
Reg. No: **2300705028**

UNDER THE GUIDANCE OF
Dr. THASLEEMA T M

DEPARTMENT OF COMPUTER SCIENCE
CENTRAL UNIVERSITY OF KERALA
Tejaswini Hills, Periye - 671320, Kasaragod, Kerala, India

APRIL 2025

Department of Computer Science, Central University of Kerala
Tejaswini Hills, Periye - 671320, Kasaragod

CERTIFICATE

This is to certify that the thesis entitled, ” **EARLY DETECTION OF PARKINSON’S DISEASE USING EXPLAINABLE AI ON SPEECH AND HANDWRITTEN DRAWING FEATURES** ” submitted by **SYAM KUMAR (REG. NO: 2300705028)** in partial fulfillment of the requirements for the award of M.Sc in Computer Science at the Central University of Kerala, is an authentic work carried out by him under my supervision and guidance.

DATE :

Dr. THASLEEMA T M
Assistant Professor
Department of Computer Science
Central University of Kerala
Kasaragod, Kerala - 671320

Department of Computer Science, Central University of Kerala
Tejaswini Hills, Periye - 671320, Kasaragod

CERTIFICATE

This is to certify that the thesis entitled, **" EARLY DETECTION OF PARKINSON'S DISEASE USING EXPLAINABLE AI ON SPEECH AND HANDWRITTEN DRAWING FEATURES "** is a bonafide work carried out by **SYAM KUMAR (REG. NO: 2300705028)** in partial fulfillment of the requirements for the award of M.Sc in Computer Science at the Central University of Kerala, Kasaragod, during the academic year **2023–2025**.

The work is satisfactory to award Master's Degree in Computer Science.

DATE :

EXAMINER 1:

EXAMINER 2:

Dr. JAYASUDHA J S
Professor
Head of the Department
Department of Computer Science
Central University of Kerala

DECLARATION

I, **SYAM KUMAR**, Reg No: **2300705028**, student of Fourth Semester M.Sc Computer Science, Central University of Kerala, do hereby declare that the thesis entitled, **" EARLY DETECTION OF PARKINSON'S DISEASE USING EXPLAINABLE AI ON SPEECH AND HANDWRITTEN DRAWING FEATURES "**, submitted to the Department of Computer Science is an original record of studies and bonafide work carried out by me during the period from **December 2024 to April 2025**.

To my knowledge, the matter embodied in the report has not been submitted to any other University or Institute for the award of any degree.

DATE:	**SYAM KUMAR**
PLACE:	**2300705028**

ACKNOWLEDGEMENT

Words play a heralding role in expressing our gratitude, if considered as a token of approval and symbol of acknowledgment. In this project, I have taken efforts. However, without the kind support and help of many individuals, it would not have been possible. I would like to extend my sincere thanks to all of them.

First and foremost, I express my profound gratitude to my guide, Dr. THASLEEMA T M, Assistant Professor in the Department of Computer Science at Central University of Kerala.I am indebted to her for her invaluable support, constant encouragement, and constructive criticism that have been instrumental in shaping the course of this project. Her guidance has been a driving force that propelled me forward, and I am deeply grateful for the motivation and direction he provided. Without her insightful mentorship, this project would not have come to fruition.

I also convey my heartful thanks to the Head of the Department, Dr. JAYASUDHA J S, for providing access to the laboratory and research facilities. Without her precious support it would not be possible to conduct this research.

I also thank our Teachers, Scholars, and Friends of the Department of Computer Science for sharing their knowledge and suggestions with me. I would like to thank various websites, research papers, and resources from the internet that have contributed to my understanding and informed my work. While too numerous to list individually, I deeply appreciate their collective impact.

I am truly grateful for the divine grace that guided and empowered me to complete this work. Without God's blessings, this achievement would not have been possible.

SYAM KUMAR

2300705028

Abstract

Parkinson's Disease (PD) is a chronic neurodegenerative disorder that progressively impairs both motor and speech functions. Early prediction of PD is vital to initiate timely treatment and improve patient outcomes. This thesis presents an explainable artificial intelligence (XAI) approach for early PD detection using two complementary data modalities: spiral and wave hand-drawn images, and speech-engineered features.

For the speech-based analysis, raw acoustic feature data comprising 755 features was initially used to train a suite of machine learning models, including Decision Tree, K-Nearest Neighbors, Support Vector Machine, Random Forest, Bagging, and Gradient Boosting. To reduce feature redundancy and improve model efficiency, Principal Component Analysis (PCA) and ANOVA-based feature selection were applied, and the models were re-evaluated on the reduced feature sets. This two-phase approach allowed for comprehensive insight into model behavior across both original and reduced-dimensional datasets, enabling better understanding of which features contribute most significantly to disease prediction. For image-based analysis, multiple transfer learning models were first fine-tuned for classification, and then also employed for deep feature extraction. These features were subsequently used with classical machine learning models to enhance performance and interpretability. Data augmentation techniques such as flipping, zooming, rotation, and shearing were applied to increase generalization across a limited dataset.

To promote model transparency and clinical trust, Local Interpretable Model-Agnostic Explanations (LIME) was employed for both the image and speech models. LIME facilitated the interpretation of model outputs by identifying key image regions and speech features influencing predictions. The proposed dual-modal framework combining transfer learning, classical machine learning, dimensionality reduction, and explainable AI techniques demonstrates promising accuracy and interpretability for early PD detection. This work serves as a robust foundation for building clinically reliable AI-assisted diagnostic tools.

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Introduction

Parkinson's Disease (PD) is a progressive neurological disorder that primarily affects movement and speech. Characterized by motor symptoms such as tremors, bradykinesia, rigidity, and postural instability, as well as non-motor symptoms including vocal degradation, PD poses significant diagnostic challenges, especially in its early stages. In recent years, artificial intelligence (AI) has shown considerable promise in supporting the early and objective detection of PD by analyzing diverse data modalities.

Among the various biomarkers studied, speech impairments and hand-drawing abnormalities (e.g., spiral and wave sketches) are non-invasive, easily collectible, and highly informative indicators of PD. Acoustic signals often reveal subtle vocal impairments long before overt motor symptoms appear, while drawing tasks can capture tremors and motor control deficiencies. Leveraging these two modalities, this research aims to develop an explainable AI framework for the early detection of Parkinson's Disease.

In this study, speech recordings were processed to extract a comprehensive set of 755 engineered features, which were initially used to train multiple machine learning classifiers. To enhance performance and reduce dimensionality, Principal Component Analysis (PCA) and ANOVA were later applied, and the classifiers were re-evaluated on the transformed datasets. Parallel to this, spiral and wave images were analyzed using transfer learning models, which were both fine-tuned for classification and used for deep feature extraction. These features were then passed into classical ML models for additional evaluation. To ensure transparency and foster clinical trust, Local Interpretable Model-Agnostic Explanations (LIME) was used to explain predictions for both speech and image data. This dual-modal, explainable approach demonstrates a practical and effective path toward robust, interpretable AI systems for PD diagnosis.

1.2 Background

Parkinson's Disease (PD) is the second most common neurodegenerative disorder, affecting over 10 million people worldwide. Its progression leads to significant motor dysfunctions and speech impairments, which are key indicators used in diagnosis. However, early-stage symptoms are often subtle and overlap with other conditions, making timely diagnosis difficult through clinical examination alone. As a result, there is a growing interest in leveraging machine learning and artificial intelligence to develop objective, non-invasive diagnostic tools.

Two commonly studied modalities for early PD detection are speech and handwriting tasks. Speech alterations such as monotonicity, reduced articulation, and vocal tremors can be detected using acoustic analysis. Similarly, drawing tasks like spirals and waves are used to assess motor impairments, providing visual biomarkers for the disease. Previous studies have explored each modality independently, but combining them can potentially improve diagnostic accuracy and robustness.

The emergence of explainable AI (XAI) further supports the adoption of these tools in clinical practice by offering interpretability and transparency in model decision-making. By integrating traditional machine learning with deep learning and XAI techniques, this project aims to create a comprehensive, interpretable framework for the early detection of Parkinson's Disease using both speech and drawing data.

1.3 Motivation

The early diagnosis of Parkinson's Disease (PD) is essential for timely treatment and better management of disease progression, yet current clinical methods often fall short due to their subjective nature and reliance on visible motor symptoms. Speech impairments and abnormalities in simple hand-drawn patterns, such as spirals and waves, have shown promise as early, non-invasive indicators of PD. These modalities are easy to collect and cost-effective, making them practical for scalable screening. However, building accurate models that can interpret these features and simultaneously offer transparency remains a significant challenge. This project is driven by the need to develop an explainable, AI-based framework that integrates speech-engineered features and drawing analysis using both classical machine learning and deep learning techniques. By incorporating dimensionality reduction methods like PCA and ANOVA, along with interpretability tools such as LIME, the proposed approach aims to deliver not only high diagnostic accuracy but also meaningful insights into the model's decision-making process—bridging the gap between performance and trust in clinical AI applications.

1.4 Organization of Thesis

This thesis is structured to provide a comprehensive understanding of the problem and the proposed solution. The first section introduces the background of Parkinson's Disease and the motivation for early detection, highlighting the role of artificial intelligence in overcoming current diagnostic challenges. The following section provides a literature review, discussing previous research on Parkinson's Disease diagnosis using speech and drawing data, as well as relevant machine learning and deep learning techniques, including the importance of explainable AI for model interpretability. The methodology section details the datasets used, preprocessing steps, feature extraction methods, and the application of dimensionality reduction techniques such as PCA and ANOVA, along with the use of LIME for model interpretability. The experimental setup section describes the implementation of transfer learning models, classical machine learning models, training configurations, and evaluation metrics. The results and discussion section presents the performance outcomes of the models, compares the impact of feature selection, and provides insights into model interpretability. The thesis concludes by summarizing the key findings, discussing the limitations of the study, and proposing future directions for research and improvements. References to the relevant literature are provided throughout the document to support the methodologies and findings presented.

Chapter 2

LITERATURE REVIEW

2.1 Introduction

Parkinson's Disease (PD) is a progressive neurodegenerative disorder that affects various aspects of motor and cognitive functions, with early diagnosis being crucial for better management and treatment outcomes. Traditionally, PD diagnosis relies heavily on clinical observation, which often identifies the disease only after substantial neuronal degeneration. However, recent advancements in machine learning (ML) and artificial intelligence (AI) have introduced novel methods for early PD detection, particularly through the analysis of motor and speech patterns. Various studies have demonstrated the potential of these technologies in providing faster, more accurate, and non-invasive diagnostic tools. These advancements highlight the increasing role of computational techniques in complementing traditional clinical approaches, offering the possibility of more reliable early-stage diagnosis. This chapter reviews key studies in the field that have utilized machine learning models, speech signal processing, and deep learning methods to improve the detection of Parkinson's Disease, presenting a comprehensive overview of the current state of research and the potential future direction of these innovative diagnostic approaches.

2.2 A Review of EXAI Models for Early Prediction of Parkinson's Disease on Speech and Handwritten Features

Handwriting analysis has become a crucial tool for the early detection of Parkinson's Disease (PD), as motor impairments commonly manifest in writing patterns. **Drotár et al.** advanced traditional methods by extracting dynamic and frequency-domain features such as velocity, acceleration, jerk, and pressure from digitized handwriting samples. They applied machine learning classifiers like Support Vector Machines, Random Forests, and k-Nearest Neighbors to distinguish between PD patients and healthy controls. Using the **PaHaW dataset**, their models achieved a maximum accuracy of approximately 80%, highlighting both the potential and challenges of handwriting-based PD diagnosis. Their research demonstrated that incorporating dynamic and frequency features significantly improved performance over basic feature sets, emphasizing the importance of feature selection techniques such as Principal Component Analysis (PCA) to enhance model generalization. This study laid a strong foundation for the future development of PD diagnostic tools based on handwriting analysis.

In the context of PD diagnosis, early detection is paramount for better clinical outcomes, yet traditional diagnostic methods often rely on clinical observation, identifying PD only after considerable neuron loss. In their study, **Aditi Govindu and Sushila Palwe** explored machine learning approaches for earlier and more accurate PD detection, focusing on voice features, as vocal impairments are early indicators of PD. The authors applied machine learning models such as Support Vector Machines (SVM), Random Forests, k-Nearest Neighbors (k-NN), and Decision Trees to classify PD patients and healthy individuals. Their experiments, particularly on the widely used **UCI Parkinson's dataset**, showed that these machine learning models could achieve high classification accuracy, with **Random Forests** and **SVM** performing notably well, reaching accuracies over 82%. The study underscores the potential of machine learning to provide fast, cost-effective, and non-invasive tools for early PD detection using simple voice-based data.

Further enhancing speech signal analysis, the study by **C. Okan Sakar et al.** investigates the use of various speech signal processing algorithms for PD classification. The authors introduce the **Tunable Q-factor Wavelet Transform (TQWT)** to extract detailed features from speech signals, improving the model's ability to capture frequency characteristics altered by PD. By comparing traditional algorithms with advanced techniques, the study demonstrated how TQWT enhanced classification performance. The work highlights the potential of TQWT-based models to provide more accurate and reliable speech-based biomarkers for PD diagnosis, offering significant improvements over conventional approaches. This study contributes to refining speech analysis methodologies for the early detection of Parkinson's Disease, showcasing the importance of advanced signal processing techniques.

In recent advancements in deep learning, **Saravanan et al. (2023)** presented a hybrid deep learning model combining the **VGG19** and **GoogleNet** architectures for PD detection based on spiral and wave drawings. To enhance the interpretability of the model, the authors integrated **LIME (Local Interpretable Model-agnostic Explanations)**, providing valuable visual explanations that pinpoint the drawing features most influential in the model's predictions. This approach not only boosts the accuracy of PD diagnosis but also improves the transparency of deep learning models, ensuring their clinical applicability. By combining high accuracy with explainability, this study bridges the gap between complex neural networks and actionable clinical insights, offering a trustworthy and transparent tool for PD diagnosis.

2.3 Summary

Recent advancements in machine learning (ML) and artificial intelligence (AI) have significantly enhanced the early detection of Parkinson's Disease (PD), a progressive neurodegenerative disorder. Traditionally diagnosed through clinical observation, PD can now be identified earlier using innovative approaches that analyze motor and speech patterns. Handwriting analysis, voice-based features, speech signal processing, and deep learning techniques have all emerged as promising tools for PD detection. Studies reviewed in this chapter demonstrate the potential of these methods to provide more accurate, fast, and non-invasive diagnostic solutions. By combining traditional clinical practices with advanced computational techniques, these approaches offer a more reliable means of diagnosing PD, facilitating early intervention and better management of the disease. The reviewed literature highlights the evolving role of AI and ML in clinical diagnostics and sets the stage for future innovations in the field.

Chapter 3

METHODOLOGY

3.1 Introduction

This section outlines the methodology adopted for the classification of Parkinson's Disease using a multi-modal framework that incorporates speech-engineered features and hand-drawn wave and spiral images. The goal is to build predictive models that are not only accurate but also interpretable using Explainable AI (XAI) techniques. For the speech modality, six traditional machine learning algorithms were initially trained to establish baseline performance. Subsequently, dimensionality reduction using Principal Component Analysis (PCA) and feature selection through ANOVA were applied to optimize the feature set. In the case of image data, transfer learning was employed in two stages: initially, pre-trained convolutional neural networks (CNNs) were fine-tuned for direct classification; later, the same networks were used to extract deep image features, which were fed into machine learning classifiers for further evaluation. To ensure transparency in model predictions, the LIME (Local Interpretable Model-agnostic Explanations) framework was utilized for both speech and image-based models. This methodology enables a comprehensive and interpretable analysis of Parkinson's Disease indicators across different data modalities.

3.2 Dataset

This study utilizes two distinct datasets to support the multi-modal classification of Parkinson's Disease: one comprising engineered speech features, and the other consisting of hand-drawn wave and spiral images.

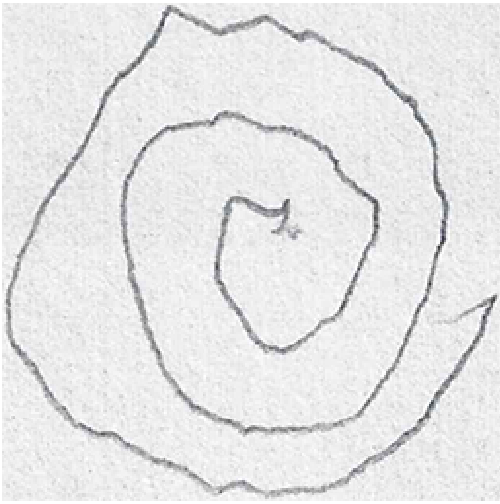
3.2.1 Speech Feature Dataset

The speech dataset consists of a total of 756 voice recordings, collected from 252 participants, including 188 individuals diagnosed with Parkinson's Disease (107 males and 81 females) and 64 healthy controls (23 males and 41 females). Each participant provided three repetitions of sustained phonation of the vowel sound /a/, recorded using a microphone at a sampling frequency of 44.1 kHz. From these recordings, various features were extracted through signal processing techniques, including jitter, shimmer, harmonics-to-noise ratio (HNR), pitch-related metrics, and Mel-frequency cepstral coefficients (MFCCs). The dataset is structured in a tabular format, where each row corresponds to a recording sample and each column represents a numerical feature, serving as input for traditional machine learning models.

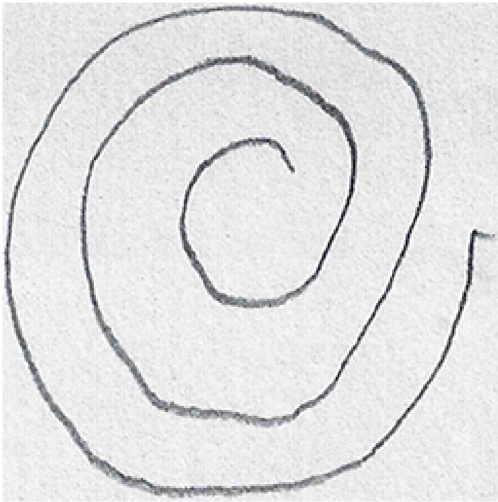
3.2.2 Wave and Spiral Drawing Dataset

The second dataset consists of scanned images of hand-drawn wave and spiral patterns, commonly used in clinical assessments to evaluate motor impairments characteristic of Parkinson's Disease. These drawings were collected from both Parkinson's patients and healthy individuals, with 36 images of each category (Parkinson's patients and healthy controls) used for training and 15 images of each category for testing. Thus, both spiral and wave drawings have 102 images each.

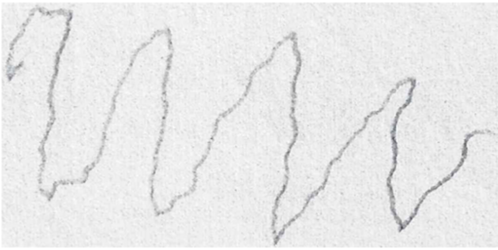
The images were preprocessed and resized to meet the input requirements of pre-trained CNNs, which were then used for feature extraction and classification. Corresponding labels indicate whether the drawing was created by a Parkinson's patient or a healthy control, serving as the target labels for the classification task.



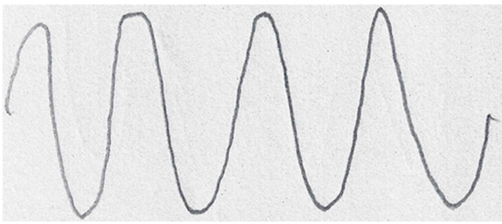
(a) Spiral - Parkinson's Disease



(b) Spiral - Healthy



(c) Wave - Parkinson's Disease



(d) Wave - Healthy

Figure 3.1: Examples of spiral and wave drawings.

3.3 Data Preprocessing

Preprocessing was performed independently for the two types of data used in this study: the structured numerical features extracted from speech recordings, and the image data comprising spiral and wave drawings.

For the speech-based features, initial data cleaning involved checking for missing or anomalous values, which were either imputed or removed as appropriate. The dataset was then partitioned into training and testing sets using an 80:20 split, with stratified sampling to preserve the class distribution across both sets. To address the issue of class imbalance in the training data, the Synthetic Minority Oversampling Technique (SMOTE) was applied exclusively to the training set. SMOTE synthetically generates new examples for the minority class by interpolating between existing minority instances, thereby helping the classifier learn a more balanced decision boundary. The resulting resampled training set had an equal representation of Parkinson's Disease and

Healthy Control samples. Following this, the features were standardized using StandardScaler, which transforms the data to have a mean of zero and unit variance. This transformation was applied separately to the training and testing data to prevent data leakage.

For the image-based data, comprising spiral and wave drawings, several preprocessing steps were conducted to prepare the images for use in convolutional neural networks (CNNs). All images were converted to grayscale, ensuring uniformity in input channels regardless of the original color format. They were then resized to a fixed resolution to match the input requirements of the pre-trained CNN models used in subsequent stages. Normalization was applied to scale pixel intensities to the [0, 1] range. To enhance model robustness and reduce overfitting, data augmentation techniques—such as slight rotations, zoom operations, and translations—were incorporated during the training process. This also served to mitigate any residual class imbalance in the image dataset. The preprocessed images were organized and stored in structured directories for streamlined loading during training and evaluation.

3.4 Feature Engineering

Feature engineering plays a crucial role in enhancing the performance of machine learning models by transforming raw data into meaningful inputs. In this study, feature engineering techniques were tailored separately for the two types of data sources—speech-based numerical features from CSV files and image-based features from hand-drawn wave and spiral drawings.

3.4.1 Speech Feature Selection and Dimensionality Reduction

The speech feature dataset includes various acoustic measurements such as jitter, shimmer, and MFCCs. To reduce redundancy and improve model efficiency, two widely used techniques were applied: *Principal Component Analysis (PCA)* and *Analysis of Variance (ANOVA)*.

- **Principal Component Analysis (PCA)**: PCA was used to reduce the high-dimensional feature space into a smaller set of uncorrelated components while retaining maximum variance. This helps in eliminating multicollinearity and reducing the computational burden during model training. The number of components was selected based on the cumulative variance explained, typically aiming to retain over 95% of the total variance.

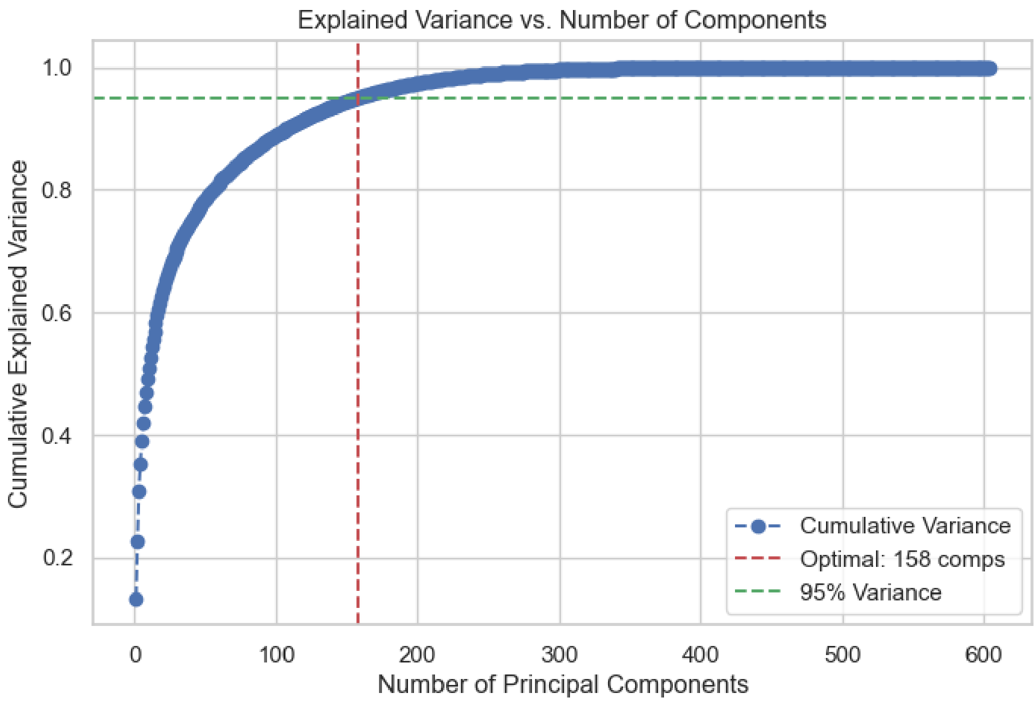


Figure 3.2: Graph representing number of principal components capturing 95% variance

- **Analysis of Variance (ANOVA)**: ANOVA was used as a feature selection method to identify the most statistically significant features that differentiate between Parkinson's and Healthy groups. Features were ranked based on their P-values, and the top-performing features were retained for further model development.

These dimensionality reduction and selection techniques were applied independently, and separate models were trained using (a) raw features, (b) PCA-transformed features, and (c) ANOVA-selected features. This allowed for comparative analysis of their impact on model performance.

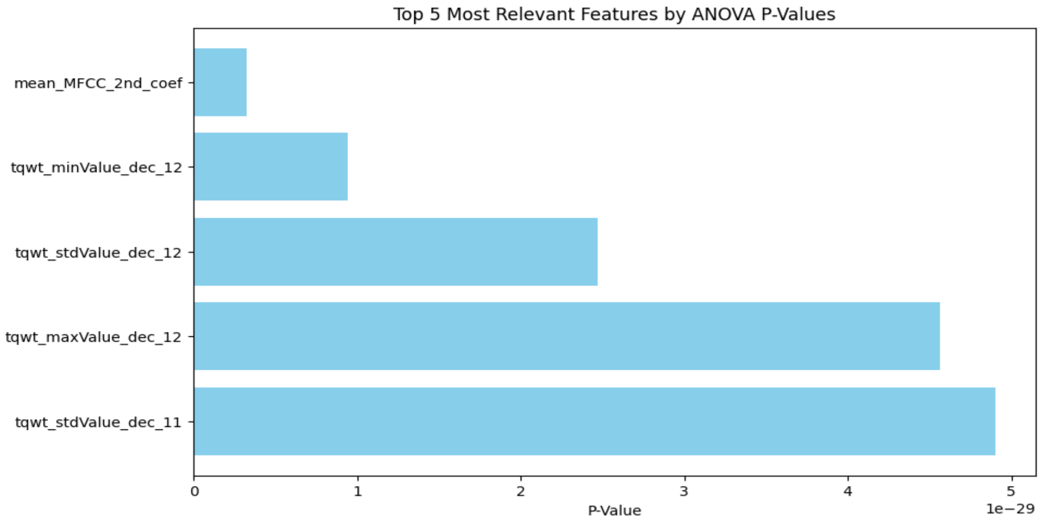


Figure 3.3: Top 5 most relevant features by ANOVA P-values

3.4.2 Feature Extraction from Hand-Drawn Images

For image-based data, deep learning models, particularly Convolutional Neural Networks (CNNs), are well-suited for capturing complex visual patterns. This study leveraged *pre-trained CNN models* both for *feature extraction* and for direct classification tasks.

- **Pre-trained CNNs for Feature Extraction**: Architectures such as *VGG19*, *InceptionV3*, *DenseNet121* and *VGG-INC* were employed. The classification layers were removed, and features were extracted from intermediate or penultimate layers.

- **Use with Traditional ML Models**: The extracted deep features were then used as input to traditional machine learning classifiers such as Support Vector Machines, Random Forest and KNearestNeighbors. This approach leverages the powerful representation capability of CNNs while maintaining flexibility and interpretability through classical ML models.

This two-step pipeline ensured a comprehensive analysis of CNN-based methods for image-based Parkinson's classification—both as end-to-end classifiers and as feature generators for downstream ML algorithms.

3.5 Model Design

This section presents the design of the models used for Parkinson's Disease classification. Two broad categories of models were employed based on the data modality: traditional machine learning models for structured features and transfer learning models for image-based classification. The machine learning models were used on both speech features and CNN-extracted features from wave and spiral images, while transfer learning models were used for direct classification of image data.

3.5.1 Machine Learning Models

Traditional machine learning algorithms were applied to both speech features (raw, PCA-transformed, and ANOVA-selected) and image features extracted from pre-trained CNNs. These models were chosen for their interpretability, efficiency, and strong performance on tabular data.

Random Forest

Random Forest is an ensemble-based learning technique that constructs a multitude of decision trees using bootstrap samples and random feature selection. The final prediction is made through majority voting. Its robustness to overfitting, high accuracy, and the ability to provide feature importance scores make it well-suited for both high-dimensional speech data and deep image features. It performs well even when the data contain non-linear relationships or noise.

Decision Tree

The Decision Tree is a flowchart-like model that splits the data based on feature thresholds to form branches and leaves leading to a class label. Although simple and prone to overfitting when used alone, it offers excellent interpretability and serves as a good baseline model. It is particularly useful when quick insights into feature influence are required.

K-Nearest Neighbors (KNN)

KNN is a distance-based algorithm that assigns a class to a sample based on the majority label among its k nearest neighbors in the feature space. It is easy to implement and effective for problems where the class distributions are spatially separated. Despite its computational cost during inference, KNN performs well when the data are scaled appropriately and when local decision boundaries dominate the classification task.

Support Vector Machine (SVM)

SVM seeks to find the optimal hyperplane that maximizes the margin between data classes. It is powerful in high-dimensional spaces and particularly effective in cases where the decision boundary is not linearly separable — in which case kernel tricks (like RBF) are applied. SVM's ability to generalize and its resistance to overfitting make it a strong choice for both speech and CNN-extracted features.

Bagging Classifier

Bagging is an ensemble method that trains multiple base learners (typically decision trees) on different subsets of the training data created via bootstrapping. The predictions are aggregated using majority voting to improve stability and reduce variance. It performs well on noisy datasets and mitigates overfitting, especially when using high-variance models like decision trees.

Gradient Boosting Classifier

Gradient Boosting is a sequential ensemble technique that builds additive models by optimizing a loss function through gradient descent. Each new model corrects the residual errors of the previous one, making the ensemble highly accurate and adaptive to complex patterns. It often outperforms other traditional ML methods in classification tasks involving heterogeneous or imbalanced data, as is often the case in medical datasets.

3.5.2 Transfer Learning Models

Deep learning through transfer learning was used to classify wave and spiral drawings. Instead of training CNNs from scratch, pre-trained models were leveraged to benefit from the rich feature hierarchies learned on large datasets like ImageNet. These models were modified by replacing the top layers with custom dense layers suitable for binary classification.

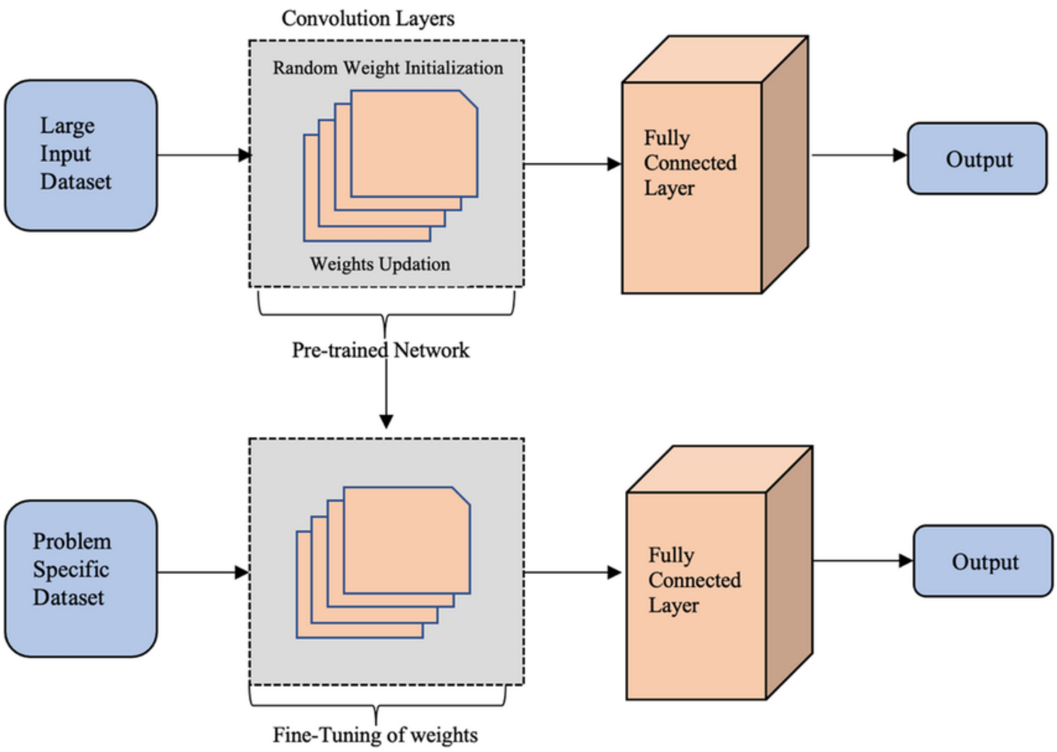


Figure 3.4: Transfer learning process of general CNN architecture

VGG19

VGG19 is a deep CNN with 19 layers consisting of small 3×3 convolutional filters and ReLU activations. It follows a uniform architecture where depth is increased by stacking more convolutional layers. Its simplicity and ability to generalize well make it a strong baseline for transfer learning. It is effective in capturing low- and mid-level patterns such as edges and curves found in hand-drawn waveforms and spirals.

InceptionV3

InceptionV3 incorporates factorized convolutions and parallel inception modules that allow the model to capture multi-scale spatial features. Its architectural optimizations reduce computational cost while improving representational power. It is particularly adept at recognizing complex structures and patterns, making it highly effective in capturing spatial distortions and irregularities in medical drawings.

DenseNet121

DenseNet121 introduces dense connectivity, where each layer receives inputs from all previous layers and passes its output to all subsequent layers. This promotes feature reuse and alleviates the vanishing gradient problem, enabling efficient and deeper learning. DenseNet121 is especially useful for extracting fine-grained features from sparse or intricate patterns, as seen in spiral and wave images of patients.

VGG19-Inception Hybrid (VGG19-INC)

The VGG19-Inception hybrid is a custom architecture combining the simplicity and depth of VGG with the multi-scale feature extraction capabilities of Inception modules. This model leverages the strengths of both frameworks to provide enhanced feature representations. It is tailored for the classification of medical images where both local and global feature detection is crucial, offering improved performance over using either architecture alone.

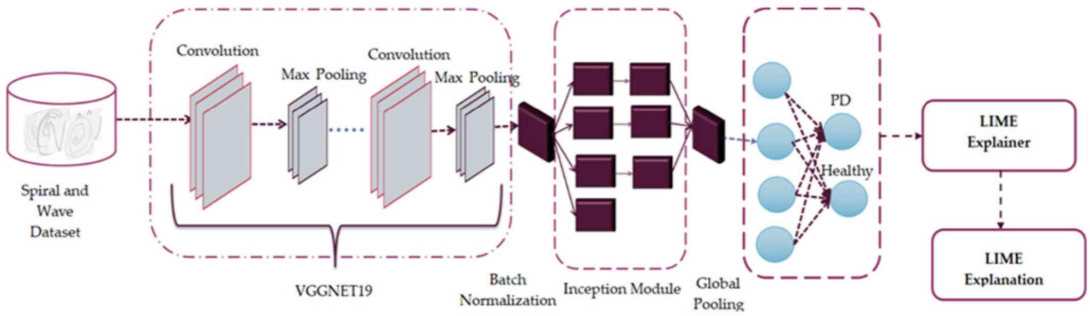


Figure 3.5: Proposed architecture of hybrid deep transfer learning VGG19-INC model

3.6 Training the Models

In this section, we describe the training process for both machine learning and deep learning models.

3.6.1 Data Splitting

For the CSV data, the dataset was initially split into training and testing subsets using a `test_size` of 0.2. This means 20% of the data was set aside for testing, and the remaining 80% was used for training. After this, SMOTE (Synthetic Minority Oversampling Technique) was applied to handle class imbalances in the training data.

For the image dataset, the images were organized into two directories: one for training and the other for testing. The training directory contains the images used for model training, while the testing directory holds the images used for final evaluation.

3.6.2 Machine Learning Models

Several machine learning models were trained using standard procedures. These models were applied to both the feature-extracted image data (from pre-trained models) and the CSV data.

3.6.3 Transfer Learning Models

For the deep learning models, pre-trained convolutional neural networks (CNNs) such as VGG19, InceptionV3, DenseNet121, and a custom VGG-Inception hybrid were employed for classification as well as feature extraction. After extracting features from the images, a custom classifier was trained using these features.

The transfer learning models were fine-tuned by modifying the last layers to suit the binary classification task. The training was conducted using the Adam optimizer, and the models were trained with binary cross-entropy as the loss function.

3.6.4 Optimizers and Loss Functions

For all transfer learning models, the Adam optimizer was used due to its efficiency and popularity in training deep learning models. The models were trained with binary cross-entropy as the loss function, which is ideal for binary classification tasks.

3.7 Explainability Integration using LIME

To enhance the interpretability of the machine learning models, the Local Interpretable Model-agnostic Explanations (LIME) framework was employed. LIME is a model-agnostic technique that provides local explanations by approximating the behavior of a complex model using an interpretable surrogate model around a given prediction.

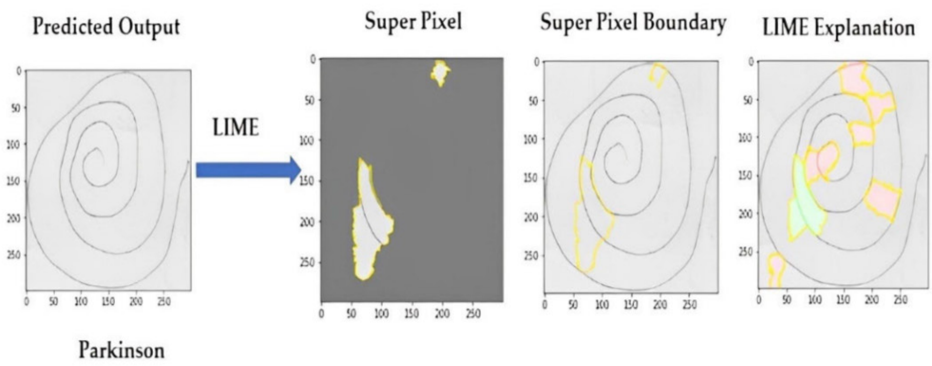


Figure 3.6: LIME explanation for PD prediction based on spiral drawing

In this study, LIME was applied primarily to the traditional machine learning models trained on both speech features and image features. For each prediction, LIME generated a set of feature contributions that explained the model's decision. These explanations allowed for a better understanding of which features (e.g., specific MFCCs or shimmer/jitter values in speech) were influential in classifying a sample as either Parkinson's-positive or negative.

The process involved the following steps:

- Selecting representative instances from the test set for explanation.
- Generating perturbed samples around the selected instance.

- Using the black-box model to make predictions on the perturbed data.
- Fitting a local, interpretable model (such as a linear regression) to approximate the predictions.
- Visualizing the weights assigned to each feature to understand their influence.

LIME was integrated using the `lime` Python library and was particularly useful in validating the model's reliance on medically meaningful features, enhancing trust in the results.

3.8 Experimental Setup

All experiments were conducted using Python 3.11 in Visual Studio, with the following system specifications:

3.8.1 Hardware Specifications

The experiments were carried out on a local machine with the following configuration:

- Operating System: Windows 11 Home 64-bit
- Processor: Intel Core i5
- RAM: 8 GB
- GPU: NVIDIA GeForce MX110

3.8.2 Software and Libraries

The following libraries and tools were used during the experimentation:

- **NumPy, Pandas** – For data manipulation and analysis
- **Matplotlib, Seaborn** – For data visualization

- **Scikit-learn** – For implementing traditional machine learning models (Random Forest, Decision Tree, etc.), feature selection (PCA, ANOVA), and evaluation metrics
- **TensorFlow and Keras** – For building and applying deep learning models, particularly transfer learning (VGG19, InceptionV3, DenseNet121)
- **LIME** – For explaining model predictions and generating local interpretability

3.9 Summary

This chapter describes the approach used for Parkinson's Disease classification, integrating both traditional machine learning (ML) and deep learning (DL) techniques. The study utilized speech feature data and hand-drawn wave/spiral drawings, with preprocessing steps including normalization, image resizing, augmentation, and label encoding. Feature extraction for the image data was performed using pre-trained convolutional neural networks (CNNs) such as VGG19, InceptionV3, DenseNet121, and a custom VGG-Inception hybrid. These models were also used for classification, alongside traditional ML models (Random Forest, Decision Tree, K-Nearest Neighbors, SVM, Bagging, and Gradient Boosting) applied to both speech and image data. For feature selection, Principal Component Analysis (PCA) and ANOVA were applied to the speech data. The models were trained using the Adam optimizer and binary cross-entropy loss. LIME was incorporated to provide model explainability.

Chapter 4

RESULTS AND DISCUSSION

4.1 Results

This section presents the results obtained from experiments conducted on both CSV-based speech features and image-based data using various machine learning and deep learning models. Performance metrics such as accuracy, precision, recall, F1-score, and confusion matrix were used to evaluate the effectiveness of the models across different tasks.

- **Accuracy:** Accuracy is the ratio of correctly predicted instances to the total number of instances. It provides an overall measure of how often the classifier is correct.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Precision:** Precision measures the proportion of positive predictions that are actually correct. It is useful when the cost of false positives is high.

$$Precision = \frac{TP}{TP + FP}$$

- **Recall:** Also known as sensitivity or true positive rate, recall measures the proportion of actual positives that were identified correctly.

$$Recall = \frac{TP}{TP + FN}$$

- **F1-Score:** The F1-score is the harmonic mean of precision and recall. It provides a balance between precision and recall, especially useful in imbalanced datasets.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

- **Confusion Matrix:** A confusion matrix is a summary table that shows the number of correct and incorrect predictions made by the classifier, categorized by actual and predicted class labels. It helps visualize performance and identify types of errors.

4.1.1 Speech-Based Analysis using Machine Learning

This section explores the performance of machine learning models applied to speech-derived features for Parkinson's Disease classification. Multiple analyses were performed to evaluate the effect of raw features, dimensionality reduction, and statistical feature selection on model accuracy and interpretability.

Raw Feature Analysis

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Decision Tree	0.770	0.868	0.814	0.840
KNN	0.763	0.964	0.708	0.816
SVM	0.829	0.866	0.912	0.888
Random Forest	**0.875**	0.912	0.920	0.916
Bagging	0.868	0.912	0.912	0.912
Boosting (Gradient Boosting)	0.829	0.899	0.867	0.883

Table 4.1: Performance Metrics Comparison of Machine Learning Models (on raw data)

In this analysis, machine learning models were trained directly on the original set of speech features without any preprocessing. The objective was to establish a performance baseline for comparison.

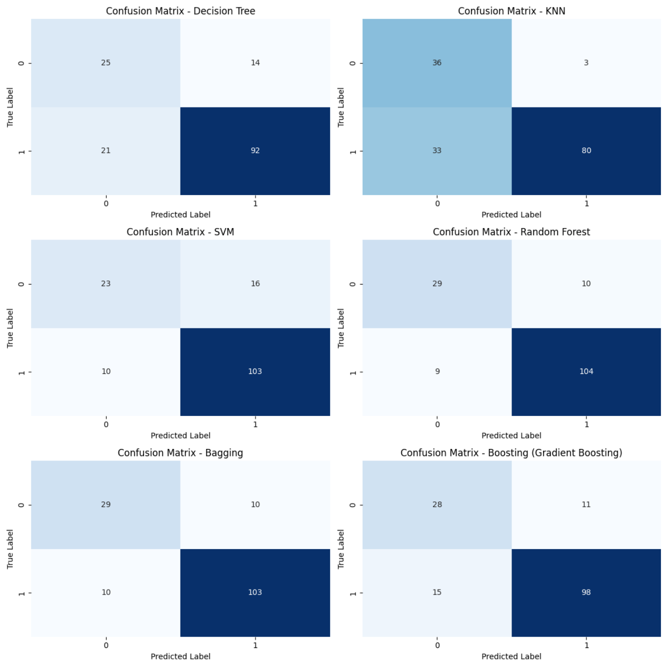


Figure 4.1: Confusion matrix for the six models (on raw data)

PCA-Based Dimensionality Reduction Analysis

Principal Component Analysis (PCA) was applied to reduce the dimensionality of the feature space. The models were then trained and evaluated on the transformed features to determine whether PCA enhanced performance or model stability.

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Decision Tree	0.724	0.809	0.823	0.816
KNN	**0.836**	0.844	0.956	0.896
SVM	0.797	0.853	0.876	0.865
Random Forest	0.809	0.800	0.991	0.885
Bagging	0.809	0.844	0.912	0.877
Boosting (Gradient)	0.829	0.848	0.938	0.891

Table 4.2: Performance Metrics Comparison of Machine Learning Models after PCA

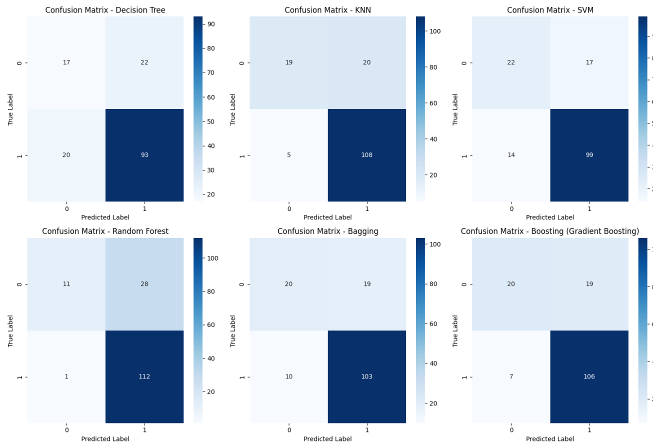


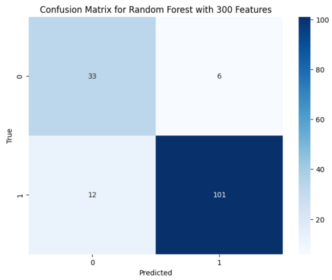
Figure 4.2: Confusion matrix for the six models (using PCA)

ANOVA-Based Feature Selection Analysis

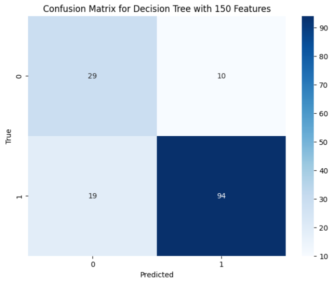
In this analysis, Analysis of Variance (ANOVA) was used to identify the most statistically significant features. The models were trained using different subsets of features, where the number of selected features (k-values) varied across 50, 100, 150, 200, 250, and 300. The performance of the classification models was evaluated for each k-value to determine the optimal feature set for improving accuracy and interpretability.

k	**Bagging**	**Boosting**	**Decision Tree**	**KNN**	**Random Forest**	**SVM**
50	0.829	0.789	0.770	0.750	0.829	0.763
100	0.855	0.816	0.803	0.809	0.842	0.796
150	**0.868**	0.829	**0.809**	**0.836**	0.875	0.803
200	0.862	0.862	0.796	0.803	0.862	**0.816**
250	0.849	0.836	0.796	0.836	0.875	0.803
300	0.849	**0.875**	0.789	0.836	**0.882**	0.816

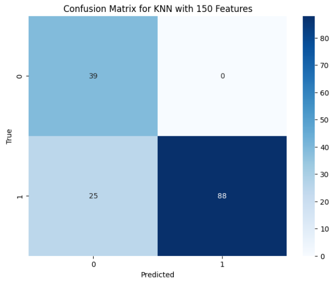
Table 4.3: Model Accuracies Across Different Training Set Sizes (k) using ANOVA



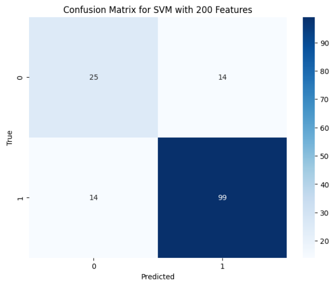
(a) Random Forest (k=300)



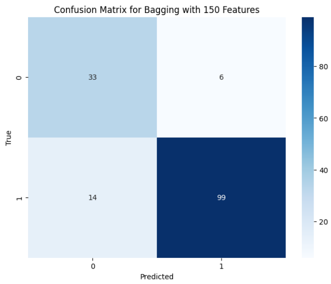
(b) Decision Tree (k=150)



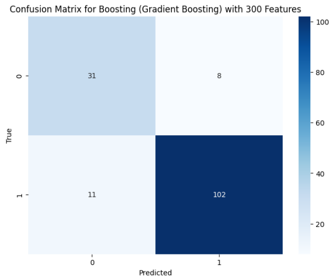
(c) KNN (k=150)



(d) SVM (k=200)



(e) Bagging (k=150)



(f) Boosting (k=300)

Figure 4.3: Confusion matrices for six models using ANOVA-selected features

4.1.2 Spiral and Wave Image Classification

Classification using Transfer Learning

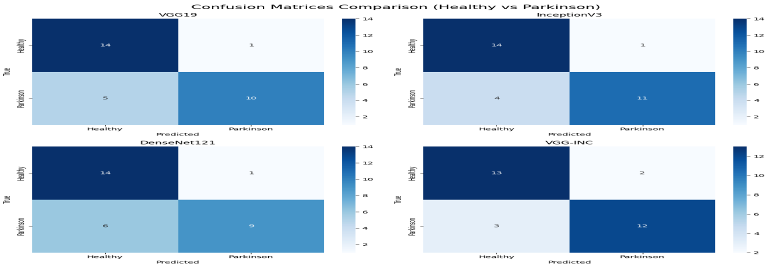
The models were fine-tuned directly using transfer learning. The accuracy results for both the spiral and wave datasets are summarized in Table 4.4.

Model	Spiral Accuracy (%)	Wave Accuracy (%)
VGG19	84.5	84.6
InceptionV3	**87.9**	**92.6**
DenseNet121	84.4	88.0
VGG19-INC	84.1	92.5

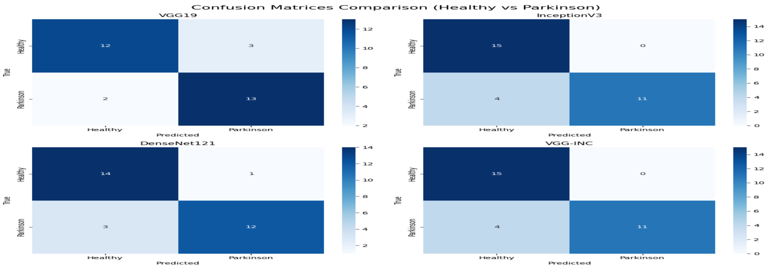
Table 4.4: Accuracy of Models on Spiral and Wave Image Classification (Transfer Learning)

Transfer Learning as Feature Extractor

In this approach, transfer learning was used as a feature extractor, and the features were passed to traditional machine learning classifiers for prediction. The results for both spiral and wave datasets are shown in Table 4.5.



(a) Spiral Drawings

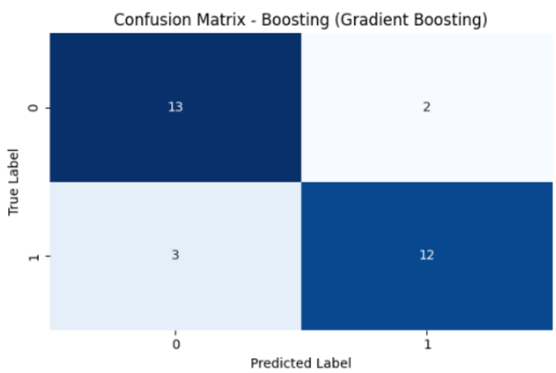


(b) Wave Drawings

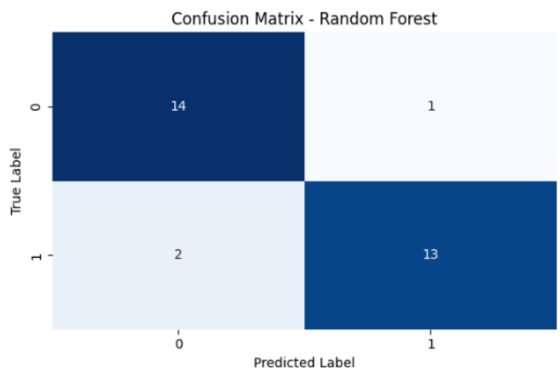
Figure 4.4: Confusion metrics of different transfer learning models on Spiral and Wave dataset

Model	**Spiral Accuracy (%)**	**Wave Accuracy (%)**
CNN + Decision Tree	73.3	76.7
CNN + KNN	**83.3**	80.0
CNN + SVM	76.7	73.3
CNN + Random Forest	73.3	**90.0**
CNN + Bagging	76.7	86.7
CNN + Gradient Boosting	**83.3**	83.3

Table 4.5: Accuracy of Models using Transfer Learning as Feature Extractors with ML Classifiers



(a) Spiral Drawing: Gradient Boosting



(b) Wave Drawing: Random Forest

Figure 4.5: Transfer Learning as Feature Extractor: Best models for both spiral and wave drawings

4.1.3 Model Explainability using LIME

To interpret and explain the model decisions, Local Interpretable Model-Agnostic Explanations (LIME) was utilized on both speech features (tabular data) and image-based inputs (spiral and wave drawings). This helped visualize which features or regions contributed most to the model predictions.

Speech Feature-based Explanations (CSV): LIME was applied to the traditional machine learning models trained on engineered speech features such as jitter, shimmer, and MFCCs. Figure 4.6 shows the contribution of each feature towards the prediction of a sample instance.

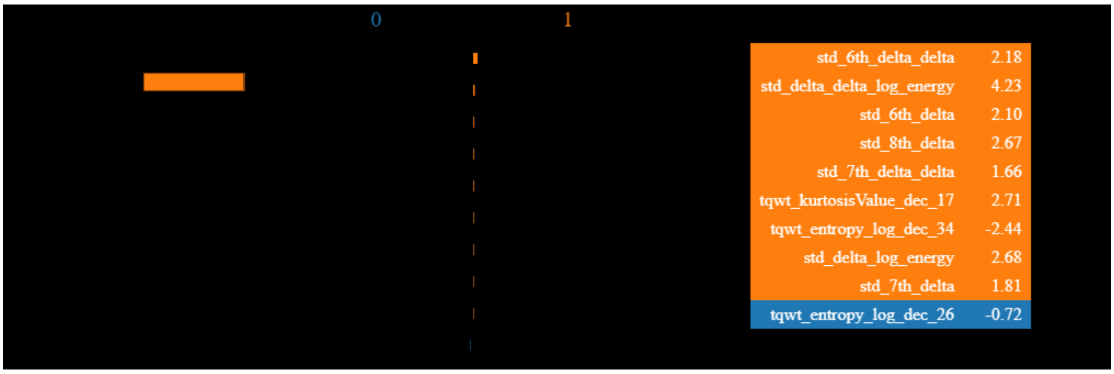


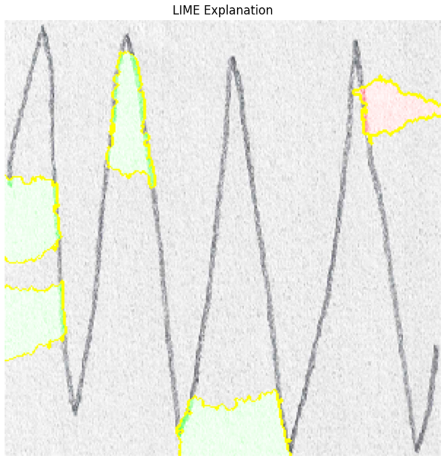
Figure 4.6: LIME explanation for a prediction on speech features (CSV data)

Each bar in the chart reflects how a particular speech feature influenced the model's decision, aiding in identifying clinically relevant predictors for Parkinson's Disease.

Image-based Explanations (Spiral and Wave Drawings): LIME was also used on CNN models to visualize the spatial regions of spiral and wave images that most influenced classification. The visualizations provided a pixel-level understanding of what regions in the drawings led to predictions. These are shown in Figure 4.7.

These interpretations validate that the models focused on regions of visible distortion and tremor—common indicators in Parkinson's Disease—when making decisions.

(a) Spiral Drawing



(b) Wave Drawing

Figure 4.7: LIME-based visualizations for CNN model predictions on image data

4.2 Discussion

This chapter presented the results of applying machine learning and deep learning techniques for Parkinson's Disease classification, utilizing both speech feature-based and image-based data. The models were evaluated using various performance metrics, including accuracy, precision, recall, F1-score, and confusion matrices, with particular emphasis on interpretability through LIME (Local Interpretable Model-agnostic Explanations).

4.2.1 Interpretation of Results

Experiments demonstrated that both speech feature-based and image-based data can be effectively utilized for Parkinson's Disease classification. Traditional machine learning models applied to speech features, such as jitter, shimmer, and MFCCs, achieved high accuracy across algorithms including Random Forest, Support Vector Machine (SVM), and Gradient Boosting. These results are consistent with previous studies that emphasize the significance of speech biomarkers in Parkinson's Disease detection. Analysis of confusion matrices indicated that the models performed particularly well in distinguishing between Parkinson's patients and healthy controls, although slight misclassifications occurred, especially among patients in intermediate stages of the disease.

In contrast, deep learning models trained on images of spiral and wave drawings, using transfer learning techniques (e.g., VGG19, InceptionV3, DenseNet121), achieved even higher classification performance. Convolutional neural networks effectively captured subtle distortions in the drawings, such as tremor-induced irregularities. The use of transfer learning facilitated better generalization from a relatively small dataset, addressing a common challenge in medical image classification tasks.

4.2.2 LIME Interpretability Insights

LIME was used to interpret the predictions made by both the machine learning and deep learning models. For the speech feature-based models, LIME revealed which specific speech features (such as jitter and MFCC coefficients) were most influential in predicting the presence of Parkinson's Disease. This is important from a clinical perspective, as it highlights key biomarkers that can be used for diagnostic purposes. The feature importance scores, represented through LIME's bar plots, show that jitter and shimmer were the most significant factors in the models' decision-making process.

For the image-based models, LIME provided a spatial interpretation of the regions in spiral and wave drawings that contributed to the predictions. This explanation helped identify regions where tremors or other motor impairments manifested, thus making the decision-making process more transparent. These findings validate the models' focus on clinically relevant patterns in the drawings, which could serve as early indicators of Parkinson's Disease. The heatmap visualizations obtained from LIME confirm that the models were sensitive to critical areas of the drawings, offering valuable insights into how these models were able to classify images correctly.

4.2.3 Clinical Implications

The findings from this study have significant implications for the early detection and monitoring of Parkinson's Disease. The ability to classify Parkinson's patients with high accuracy using non-invasive methods, such as speech recordings or simple drawings, could provide a cost-effective and accessible screening tool. Moreover, the interpretability provided by LIME can help clinicians understand which features or regions of the data are influencing the model's predictions, making these models more actionable in a clinical setting.

Chapter 5

FUTURE WORK AND CONCLUSION

5.1 Future Work

Future work could focus on expanding the dataset by including more diverse and larger samples, which would enhance the model's ability to generalize across different populations and disease stages. Employing advanced data augmentation techniques could further address data scarcity and variability, particularly for image-based datasets, leading to improved model robustness. In addition, future studies could move beyond manual feature engineering for speech data by utilizing end-to-end deep learning models capable of automatically extracting rich, complex features directly from raw signals. This approach could better capture the intricate characteristics of speech impairments associated with Parkinson's Disease. Another promising direction is the integration of multimodal data sources, such as combining speech features with clinical markers like movement patterns, handwriting dynamics, or genetic information. Such multimodal fusion could provide a more comprehensive view of the disease and significantly boost predictive performance. Also, longitudinal studies tracking patients over time could be incorporated to assess how the models perform in monitoring disease progression, opening the door to early intervention strategies and personalized treatment plans.

5.2 Conclusion

This thesis presented a dual-modality approach for Parkinson's Disease (PD) classification using both speech-based features and image-based spiral and wave drawings. Through the application of traditional machine learning algorithms on speech features, and convolutional neural networks (CNNs) via transfer learning on image data, the strengths and challenges of both data domains were explored. The experiments demonstrated that both modalities have the potential to effectively support early PD diagnosis with high accuracy.

In the image domain, four pre-trained CNN models — VGG19, InceptionV3, DenseNet121, and a custom VGG19-Inception hybrid—were evaluated on spiral and wave drawings. These models, when used directly for classification as well as for feature extraction followed by machine learning models, showcased strong performance across the board. Meanwhile, the speech feature analysis identified key acoustic patterns and demonstrated that models such as Random Forest and Gradient Boosting offered robust classification results. These combined efforts highlight the versatility of both data types in capturing PD-related characteristics.

To improve model transparency, LIME (Local Interpretable Model-agnostic Explanations) was employed for both speech and image-based models. LIME offered interpretable visual and feature-level insights into model predictions, supporting explainability which is crucial in medical AI systems. Overall, the study contributes a reliable, interpretable, and data-efficient framework for PD classification and sets the foundation for further research using multimodal data and real-world clinical deployment.

Bibliography

References

[1] S. Saravanan et al., “Explainable Artificial Intelligence (EXAI) Models for Early Prediction of Parkinson’s Disease Based on Spiral and Wave Drawings,” *IEEE Access*, 2023.

[2] A. Govindu and S. Palwe, ”Early detection of Parkinson’s disease using machine learning,” International Conference on Machine Learning and Data Engineering, ScienceDirect, 2021.

[3] C. Okan Sakar, Gorkem Serbes, Aysegul Gunduz, Hunkar C. Tunc, Hatice Nizam, Betul Erdogdu Sakar, Melih Tutuncu, Tarkan Aydin, M. Erdem Isenkul, and Hulya Apaydin, ”A comparative analysis of speech signal processing algorithms for Parkinson’s disease classification and the use of the tunable Q-factor wavelet transform,” *Applied Soft Computing Journal*, vol. 101, pp. 106888, 2021, doi: 10.1016/j.asoc.2020.106888.

[4] W. Wang, J. Lee, F. Harrou and Y. Sun, “Early Detection of Parkinson’s Disease Using Deep Learning and Machine Learning,” *IEEE Access*, vol. 8, pp. 147635-147646, 2020, doi: 10.1109/ACCESS.2020.3016062.

[5] H. Braak and E. Braak, “Pathoanatomy of Parkinson’s disease,” *J Neurol*, vol. 247, pp. II3-II10, 2000, doi: 10.1007/PL00007758.

[6] C. Váradi, “Clinical features of Parkinson’s disease: The evolution of critical symptoms,” *Biology*, vol. 9, no. 5, p. 103, May 2020.

[7] A. Rana et al., “An efficient machine learning approach for diagnosing Parkinson’s disease by utilizing voice features,” *Electronics*, vol. 11, no. 22, p. 3782, Nov. 2022.

[8] R. Yuvaraj, U. R. Acharya, and Y. Hagiwara, “A novel Parkinson’s disease diagnosis index using higher-order spectra features in EEG signals,” *Neural Comput. Appl.*, vol. 30, no. 4, pp. 1225-1235, Aug. 2018.

[9] C. Kotsavasiloglou et al., “Machine learning-based classification of simple drawing movements in Parkinson’s disease,” *Biomed. Signal Process. Control*, vol. 31, pp. 174-180, Jan. 2017.

[10] Ferdib-Al-Islam and L. Akter, “Early identification of Parkinson’s disease from hand-drawn images using histogram of oriented gradients and machine learning techniques,” in *Proc. Emerg. Technol. Comput., Commun. Electron. (ETCCE)*, Dec. 2020, pp. 1-6.

[11] P. Raundale, C. Thosar, and S. Rane, ”Prediction of Parkinson’s disease and severity of the disease using Machine Learning and Deep Learning algorithm,” 2021 2nd International Conference for Emerging Technology (INCET), pp. 1-5, doi: 10.1109/INCET51464.2021.9456292.

[12] T. J. Wroge, Y. Özkanca, C. Demiroglu, D. Si, D. C. Atkins, and R. H. Ghomi, ”Parkinson’s Disease Diagnosis Using Machine Learning and Voice,” 2018 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), pp. 1-7, doi: 10.1109/SPMB.2018.8615607.

[13] D. Yadav and I. Jain, ”Comparative Analysis of Machine Learning Algorithms for Parkinson’s Disease Prediction,” 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1334-1339, doi: 10.1109/ICICCS53718.2022.9788354.

[14] M. Alissa, M. A. Lones, J. Cosgrove, J. E. Alty, S. Jamieson, S. L. Smith, and M. Vallejo, ”Parkinson’s disease diagnosis using convolutional neural networks and

figure-copying tasks," Neural Comput. Appl., vol. 34, no. 2, pp. 1433–1453, Jan. 2022, doi: 10.1007/s00521-021-06469-7.

[15] M. Kamble, P. Shrivastava, and M. Jain, "Digitized spiral drawing classification for Parkinson's disease diagnosis," Meas., Sensors, vol. 16, Aug. 2021, Art. no. 100047.

[16] A. Lumini and L. Nanni, "Deep learning and transfer learning features for plankton classification," Ecol. Informat., vol. 51, pp. 33–43, May 2019.