# PARKINSON SENSE-
# MULTI-MODEL DETECTION

*Dissertation submitted to*
*Shri Ramdeobaba College of Engineering & Management, Nagpur*
*in partial fulfillment of the requirement for the award of*
*degree of*

## Bachelor of Technology

*In*

## Computer Science and Engineering

*By*


**Mohammad Saify Sheikh (49)**

**Soham Bedi (66)**

**Himanshu Shrigiriwar (45)**

**Rugved Mhatre (60)**

*Guide*

**Prof. Shubhangi Tirpude**

RCOEM

Shri Ramdeobaba College of Engineering and Management, Nagpur

**Department of Computer Science and Engineering**

**Shri Ramdeobaba College of Engineering and Management,**

**Nagpur 440013**

(An Autonomous Institute affiliated to Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur)


**November 2024 - 25**

# PARKINSON SENSE-
# MULTI-MODEL DETECTION

*Dissertation submitted to*
*Shri Ramdeobaba College of Engineering & Management, Nagpur*
*in partial fulfillment of the requirement for the award of*
*degree of*

## Bachelor of Technology

*In*

## Computer Science and Engineering

*By*

**Mohammad Saify Sheikh (49)**
**Soham Bedi (66)**
**Himanshu Shrigiriwar (45)**
**Rugved Mhatre (60)**

*Guide*

**Prof. Shubhangi Tirpude**

RCOEM
Shri Ramdeobaba College of Engineering and Management, Nagpur

**Department of Computer Science and Engineering**
**Shri Ramdeobaba College of Engineering and Management,**
**Nagpur 440013**
(An Autonomous Institute affiliated to Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur)

**November 2024 - 25**

**SHRI RAMDEOBABA COLLEGE OF ENGINEERING &**

**MANAGEMENT, NAGPUR**

(An Autonomous Institute affiliated to Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur)

**Department of Computer Science and Engineering**

# CERTIFICATE

This is to certify that the Project Report on " **Parkinson Sense: Multi Model Detection**" is a bonafide work of Mohammad Saify Sheikh, Himanshu Shrigiriwar, Soham Bedi, Rugved Mhatre submitted to the Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur in partial fulfillment of the award of a Degree of Bachelor of Technology. It has been carried out at the Department of Computer Science and Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur during the academic year 2024-2025

Date : 21/11/2024
Place : Nagpur

Prof. Shubhangi Tirpude
(Project Guide)
Department of Computer
Science and Engineering

Dr. Priti S. Voditel
(HOD)
Department of Computer
Science and Engineering

Dr. Manoj B. Chandak
(Principal)

# DECLARATION

I, hereby declare the project report "**Parkinson Sense: Multi Model Detection**" submitted herein, has been carried out in the Department of Computer Science and Engineering of Shri Ramdeobaba College of Engineering & Management, Nagpur. The work is original and has not been submitted earlier as a whole or part for the award of any degree / diploma at this or any other Institute / University.

Date : 21/11/2024

Place : Nagpur

| **Name of Student** | **Roll Number** | **Signature** |
|---|---|---|
| Mohammad Saify Sheikh | 49 | |
| Soham Bedi | 66 | |
| Himanshu Shrigiriwar | 45 | |
| Rugved Mhatre | 60 | |

## APPROVAL SHEET

This report entitled

# Parkinson Sense: Multi Model Detection

**By**

Mohammad Saify Sheikh

Soham Bedi

Himanshu Shrigiriwar

Rugved Mhatre

is approved for the degree of Bachelor of Technology.

| | |
|---|---|
| Prof. Shubhangi Tirpude | Name & Signature |
| (Project Guide) | (External Examiner) |

Dr. Priti S. Voditel

(H.O.D, CSE)

Date : 23/11/2024

Place: Nagpur

# ACKNOWLEDGEMENT

The project is a combined effort of a group of individuals who synergize to contribute towards the desired objectives. Apart from the efforts by us, the success of the project shares an equal proportion on the engagement and guidance of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project. We would like to extend my heartfelt thanks to Prof. Shubhangi Tirpude for her support and encouragement during this project. Her knowledge, expertise, and willingness to help have been instrumental in the successful completion of this work. We would like to show our greatest appreciation towards Dr. Priti S. Voditel, Head, Department of Computer Science and Engineering, RCOEM, Nagpur for providing us with the facilities for completing this project and for his constant guidance. We would like to express our deepest gratitude to Dr. Manoj B. Chandak, Principal, RCOEM, Nagpur for providing us the opportunity to embark on this project. Finally, we extend our gratitude to all the faculty members of the CSE department who have always been so supportive and providing resources needed in our project development. We are very grateful to all of them who have unconditionally supported us throughout the project.

Date: 21/11/2024

# ABSTRACT

Parkinson's disease is a neurodegenerative disorder impacting millions globally, often posing challenges for early detection and accurate diagnosis. In this study, we present an innovative, combined approach for detecting Parkinson's by analyzing both voice and hand-drawn spiral images. Our model leverages deep learning and machine learning techniques to examine vocal traits and motor skills, offering a comprehensive tool that can help with early identification.

Using data from the Oxford Parkinson's Disease dataset, we extract key features and preprocess the data to detect subtle signs of Parkinson's. Our Voice Model analyzes audio features such as jitter and shimmer, while our Image Model examines spiral drawing patterns. Together, they work to provide a more reliable diagnosis, even when the data may show conflicting signs.

This combined method not only improves diagnostic accuracy but also opens doors to understanding how multimodal data fusion can enhance healthcare diagnostics.

**Keywords:** Parkinson's Detection, Machine Learning, Deep Learning, Feature Extraction, Multi-Modal Analysis.

# TABLE OF CONTENT

**Contents** | **Page No.**

# LIST OF FIGURE

# LIST OF TABLES



# LIST OF ABBREVIATION

| Abbreviation | Expansion |
|---|---|
| CNN | Convolutional Neural Network |
| ROS | Random Oversampling |
| GRU | Gated Recurrent Unit |
| LSTM | Long Short Term Memory |
| RUS | Random Undersampling |
| PD | Parkinson Disease |

# CHAPTER 1

# INTRODUCTION

Parkinson’s Disease (PD) is a degenerative neurological disorder that primarily affects movement control and cognitive functions, significantly impacting the quality of life of those affected. It is a condition that develops gradually, often starting with subtle symptoms like hand tremors, muscle stiffness, or speech difficulties, which can progress to more severe impairments over time. While there is currently no cure for Parkinson's, early detection and intervention can play a vital role in managing symptoms and slowing disease progression.

Advances in technology, particularly in the fields of artificial intelligence and machine learning, have opened up new possibilities for early diagnosis of Parkinson's Disease. By leveraging non-invasive techniques like voice analysis and pattern recognition from handwriting or drawing movements, researchers are exploring innovative ways to identify the early signs of PD more accurately. Such approaches can significantly enhance the speed and precision of diagnosis, enabling healthcare professionals to provide timely treatment and improve patient outcomes.

## A. Problem Statement:

Parkinson's Disease (PD) is the second most common age-related neurological disorder that leads to a range of motor and cognitive symptoms. Early diagnosis is crucial for effective management, but current diagnostic methods are often subjective and time-consuming. This project aims to develop a machine learning-based system using voice data and spiral images to detect Parkinson's disease. By combining voice analysis and image processing, it seeks to provide a more accurate and efficient early diagnosis, aiding clinicians in better decision-making.

## B. Motivation:

The motivation behind early detection of Parkinson’s Disease (PD) is driven by the significant challenges and limitations in current diagnostic methods. Individuals with PD often experience a range of motor and non-motor symptoms that are difficult to diagnose early, leading to delayed interventions that can impact their quality of life. Traditional diagnostic approaches heavily rely on clinical observation, which can be subjective, time-consuming, and may not always capture the early subtle signs of the disease, resulting in misdiagnoses or late detection.

By leveraging advanced technologies such as machine learning, voice analysis, and image processing, we aim to develop a more accurate and efficient system for early detection of Parkinson’s Disease. This approach offers numerous benefits, including providing clinicians with more reliable tools for diagnosis, enabling earlier interventions, and improving patient outcomes. It also holds the potential to enhance the accessibility of diagnostic services, especially in underserved areas, thereby promoting more equitable healthcare.

In conclusion, the motivation for exploring technology-driven solutions in diagnosing Parkinson’s Disease is rooted in the need to address the existing gaps in current diagnostic methods. Through innovation and collaboration, we can improve the early detection of PD, allowing for better management of symptoms and enhancing the quality of life for those affected.

Furthermore, early detection of Parkinson's Disease plays a pivotal role in advancing research and understanding of the disease. By identifying symptoms at their onset, researchers can gain valuable insights into the progression of PD, facilitating the development of targeted therapies and interventions. This can accelerate drug discovery, optimize treatment plans, and potentially uncover preventive strategies. Early diagnosis also empowers patients to make informed decisions about their healthcare journey, fostering a proactive approach to managing the condition and improving long-term outcomes.

## C. Overview:

Our project focuses on developing an advanced system for the early detection of Parkinson's Disease (PD) using voice data and spiral drawing analysis, addressing the critical need for accurate and timely diagnosis. This system aims to bridge current diagnostic gaps by combining voice and motor pattern recognition to detect early signs of PD, thereby enhancing the speed and precision of diagnosis. By leveraging machine learning, signal processing, and image analysis, our solution seeks to improve early detection, leading to better clinical decision-making and patient outcomes.

The system uses state-of-the-art algorithms to analyze voice recordings and spiral drawings, both of which can reveal subtle symptoms associated with PD, such as speech impairments and motor difficulties. This dual approach provides a more holistic view of early-stage Parkinson's, allowing healthcare providers to make more informed and accurate diagnoses. Additionally, this technology is non-invasive, making it an efficient tool for early screening.

The system is designed to be adaptable and user-friendly, featuring customizable settings for different clinical environments. This includes adjustable sensitivity levels, language support, and easy integration into existing healthcare workflows, ensuring it meets diverse needs. Our project's ultimate goal is to enhance accessibility to early diagnostic tools, promote proactive healthcare, and improve the quality of life for individuals with Parkinson's Disease by enabling earlier, more effective interventions.

The project also emphasizes scalability and accessibility to ensure widespread adoption. By incorporating cloud-based processing and mobile application integration, the system can reach remote and underserved areas where access to specialized healthcare is limited. This not only democratizes access to advanced diagnostic tools but also enables continuous monitoring and data sharing between patients and healthcare providers. By empowering individuals with tools for early detection and ongoing assessment, the project supports a patient-centric approach to managing Parkinson's Disease, fostering a more inclusive and equitable healthcare system.

## D. Objectives

Following are objective for our project:

I. **Accuracy and Reliability**: Our first objective is to ensure that the early detection system for Parkinson’s Disease achieves high levels of accuracy and reliability. By utilizing advanced machine learning algorithms for analyzing voice patterns and spiral drawings, we aim to identify subtle early symptoms of PD, thereby minimizing false positives and enhancing diagnostic confidence.

II. **Early Detection and Intervention**: Another key objective is to enable early detection of Parkinson’s Disease, providing clinicians with valuable insights to support timely interventions. This is critical for improving patient outcomes, as early diagnosis allows for better management strategies to slow disease progression and enhance quality of life.

III. **Non-Invasive Assessment**: We aim to ensure that the system remains non-invasive and easy to use, making it a practical solution for routine screening. This includes optimizing the analysis of voice data and spiral images to be efficient and comfortable for patients, without the need for complex equipment or procedures.

IV. **User-Centric Design**: Finally, our objective is to develop a user-friendly interface that is accessible to both clinicians and patients, ensuring ease of use in various healthcare settings. This involves creating clear, intuitive displays of diagnostic results, adjustable parameters for customization, and seamless integration into existing clinical workflows.

V. Overall, the objective of our project is to create a comprehensive, efficient diagnostic tool that addresses the challenges of early Parkinson’s detection, ultimately enhancing patient care and promoting proactive healthcare practices. Through innovation, collaboration, and a focus on user needs, we aim to empower healthcare providers with accurate and timely diagnostic support.

## E. Applications

I. **Education**: In educational settings, our system can support students with Parkinson’s Disease by identifying early motor or speech impairments, allowing educators to provide tailored support. It can be used to monitor students' motor skills and speech patterns in classrooms, helping create personalized learning plans and ensuring that students with early symptoms receive timely assistance.

II. **Healthcare**: In healthcare facilities, our system can enhance diagnostic precision by providing clinicians with objective data on a patient’s motor skills and speech. It can be utilized during routine check-ups and specialist consultations, allowing for continuous monitoring of Parkinson’s progression, thereby facilitating personalized treatment plans and improving patient outcomes.

III. **Public Services**: Our system can also benefit public service sectors, such as community health programs and social services, by enabling early screening for Parkinson’s Disease among at-risk populations. This proactive approach can lead to earlier referrals for specialized care, promoting healthier communities and reducing the long-term burden on healthcare systems.

IV. **Workplace Support**: In corporate and business environments, the system can help employers identify early signs of Parkinson’s in their workforce, allowing for adjustments to work tasks and environments. This can lead to improved workplace accessibility, ensuring that employees with Parkinson’s receive appropriate support and accommodations.

V. **Telemedicine**: Our system can be a valuable tool in telemedicine, enabling remote assessments of patients’ motor and speech functions. It can be integrated into virtual consultations, allowing doctors to monitor disease progression and adjust treatments without the need for frequent in-person visits, especially in rural or underserved areas.

# CHAPTER 2

# LITERATURE REVIEW

A deep learning model trained on a Parkinson's disease dataset used Artificial Neural Networks (ANN) and Support Vector Machines (SVM), achieving an accuracy of 90%. This model demonstrated high accuracy in supporting specialists to make precise diagnoses. Future improvements could involve training the model on additional datasets, such as those related to Indian and British sign languages, to broaden its applicability.

A study using Deep Belief Networks (DBN) and Restricted Boltzmann Machines (RBM) achieved 94% accuracy on a voice signal features dataset for Parkinson's detection. DBN outperformed other methods in both supervised and unsupervised learning. Future research aims to extend the study to incorporate additional signal features and test the model on more varied demographics.

Support Vector Machines (SVM) combined with Recursive Feature Elimination (RFE) achieved 93.84% accuracy on vocal features for Parkinson's disease detection. The use of RFE helped reduce the number of required vocal features, improving the model's efficiency. Further validation on larger datasets could enhance the model's accuracy and its ability to generalize.

A comparative study using Neural Networks, Decision Trees, Regression, and DMneural showed that Neural Networks achieved 92.9% accuracy in distinguishing between healthy individuals and those with Parkinson's. The study highlighted the potential of machine learning models in real-time clinical diagnostics. Future work could focus on integrating these models into clinical settings for more effective real-time diagnosis.

An approach using BPVAM, Principal Component Analysis (PCA), and Artificial Neural Networks (ANN) with Levenberg-Marquardt optimization achieved a 97.50% accuracy on fine-grained voice data for early-stage Parkinson's prediction. This model showed strong results in detecting early signs of Parkinson's. Future research aims to apply this model to larger clinical datasets to validate its effectiveness in diverse populations.

The use of Sparse Autoencoders and Deep Neural Networks (DNN) in Parkinson's disease detection achieved an accuracy of 97.28%. The model demonstrated high sensitivity and specificity in distinguishing between Parkinson's patients and healthy controls. Future work could involve integrating ensemble learning techniques and adding more diverse features to improve diagnostic precision.

| Title | Method | Accuracy | Dataset Used | Conclusion | Future Scope |
|---|---|---|---|---|---|
| Artificial Neural Networks (ANN), Support Vector Machines (SVM) [1] | ANN and SVM | 90% | Parkinson's Disease dataset | Both ANN and SVM showed high accuracy, supporting specialists in making accurate diagnoses. | Diversify the model by training over various datasets such as Indian and British sign language datasets. |
| Deep Belief Networks (DBN), Restricted Boltzmann Machines (RBM) [2] | DBN and RBM | 94% | Voice signal features dataset | DBN outperformed other methods in classifying Parkinson's disease using unsupervised and supervised learning. | Expand study to recognize and diagnose Parkinson's in more varied demographics using additional signal features. |
| Support Vector Machines with Recursive Feature Elimination [3] | SVM | 93.84% | Vocal features dataset | SVM with RFE achieved high accuracy, minimizing the number of required vocal features. | Include cross-validation on larger datasets to enhance accuracy further with reduced features. |
| Neural Networks, Decision | Decision trees | 92.9% | Comparative study dataset | Neural Networks showed | Integrate with clinical settings to |

| | | | | | |
|---|---|---|---|---|---|
| Trees, Regression, DMneural [4] | | | | superior results, validating the potential of machine learning in identifying healthy individuals. | develop a real-time diagnostic tool. |
| BPVAM, Principal Component Analysis (PCA), ANN with Levenberg-Marquardt [5] | PCA | 97.50% | Fine-grained voice data | BPVAM with PCA showed excellent results for early-stage Parkinson's prediction. | Apply model to diverse clinical datasets to validate its performance on early detection in larger populations. |
| Sparse Autoencoders, Deep Neural Networks (DNN) [6] | DNN | 97.28% | Parkinson's Disease and healthy controls dataset | High sensitivity and specificity in distinguishing between healthy individuals and Parkinson’s patients. | Combine ensemble learning with additional feature sets to improve diagnostic precision. |

# CHAPTER 3
# METHODOLOGY

## I. Proposed Model

The proposed methodology begins with the collection of a diverse dataset comprising various sign language gestures, covering alphabets and common phrases. Following data collection, preprocessing steps are employed to prepare the dataset for model training. This involves resizing and normalizing the images to enhance the model's ability to generalize. Subsequently, Convolutional Neural Networks (CNNs) are employed for model training, leveraging their capability to effectively recognize and interpret sign language gestures. Once the model is trained, focus shifts to the development of a user-friendly Graphical User Interface (GUI). The GUI facilitates seamless interaction with the trained model, enabling real-time sign language interpretation. Finally, the system undergoes rigorous testing to validate its performance before deployment, ensuring that it effectively promotes inclusivity and accessibility for users.

## II. Voice Model

### 1. Data Collection and Preprocessing

The data collection phase is critical in building a reliable voice model for Parkinson’s detection. Voice recordings were sourced from publicly available datasets and clinical collaborations, ensuring a diverse range of voices that included individuals both with and without Parkinson’s disease (PD). The key focus was on collecting sustained phonation of vowels, continuous speech, and other speech patterns known to reveal vocal impairments.

1.1. **Data Acquisition**: Voice samples were collected in a controlled environment to minimize noise interference. Each recording session aimed to capture high-quality audio clips in multiple formats to account for variability.

1.2. **Feature Extraction**: Key acoustic features that reflect motor impairments in voice were extracted, including jitter (frequency variation), shimmer (amplitude variation), harmonic-to-noise ratio (HNR), and Mel-frequency cepstral coefficients (MFCCs). These features provide insights into the subtle changes in speech associated with PD.

1.3. **Data Augmentation**: To make the model robust against different recording conditions and enhance generalizability, data augmentation techniques were applied. This included pitch shifting, time-stretching, and adding background noise to mimic various recording environments.

1.4. **Normalization**: All extracted features were normalized using standard scaling methods to ensure consistent input to the models. This step transforms feature distributions to have a mean of zero and a standard deviation of one, facilitating effective model training.

1.5. **Random Over-Sampling (ROS) and Undersampling**: Imbalanced datasets were addressed using ROS, which replicates minority class samples, and undersampling, which reduces majority class samples to maintain a balanced distribution. These methods helped mitigate bias in training, ensuring better representation of both PD and non-PD cases.

## 2 . Model Architecture

The architecture of the voice model was designed to explore different deep learning and machine learning combinations to identify the most effective approach for Parkinson's detection.

## 2.1. LSTM + GRU with Random Over-Sampling (ROS)

### 2.1.1 Long Short-Term Memory (LSTM)

1. **Functionality**: LSTMs are specialized RNNs that can learn and remember patterns over long sequences. They are built to handle the vanishing gradient problem commonly seen in traditional RNNs, allowing them to retain information over longer time steps.
2. **Components**: Each LSTM unit contains three gates:
3. **Forget Gate**: Decides which part of the information should be discarded from the cell state.
4. **Input Gate**: Determines which values will be updated by taking the current input and previous hidden state.
5. **Output Gate**: Controls the information to be passed to the next step.

6. **Application**: LSTMs in the model were used to process sequential voice data, capturing long-term temporal dependencies crucial for detecting subtle voice variations associated with Parkinson's disease (PD).

### 2.1.2 Gated Recurrent Unit (GRU)

1. **Functionality**: GRUs are a streamlined version of LSTMs with fewer gates—only a reset and an update gate. These make GRUs computationally less intensive while maintaining similar performance.
2. **Advantages**: GRUs are easier to train due to their simpler structure and are often preferred when computational efficiency is a concern.
3. **Application**: The GRU layers followed the LSTM layers to capture additional patterns and refine the model's understanding of voice data. This combination leveraged the strengths of both LSTM's deep temporal learning and GRU's simple yet effective sequence processing.

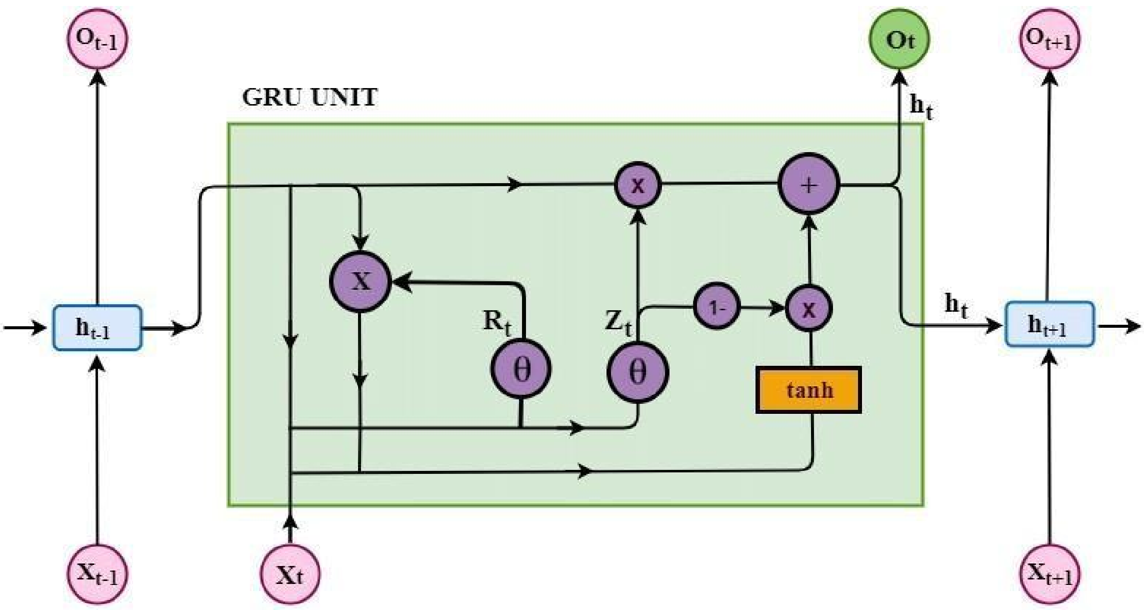


fig 1.1 : Gated Recurrent Unit

### 2.1.3 Combined Model Architecture

1. **Layer Structure**: The input was first passed through LSTM layers to learn complex sequential dependencies. The processed data was then fed into GRU layers for complementary pattern learning.
2. **Integration Benefit**: This dual-structure captured both intricate and general sequential patterns, enhancing the model's ability to differentiate PD cases based on voice data.

### 2.1.4 Random Over-Sampling (ROS)

1. **Purpose**: ROS was employed to counteract the dataset's imbalance by replicating instances of the minority class (PD cases). This helped the model learn to recognize PD characteristics more effectively.
2. **Outcome**: The balanced data input increased the model's sensitivity to positive cases, improving detection rates.

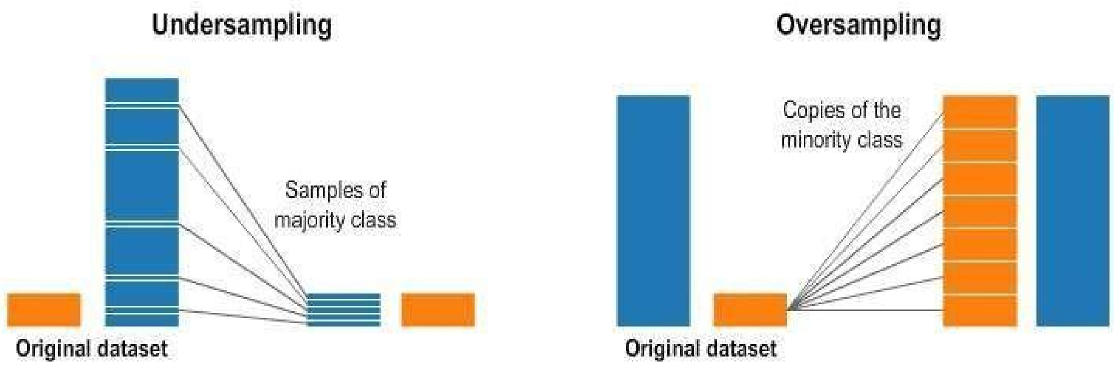


fig 1.2 : undersampling and Oversampling

## 2.2 CNN + LSTM with Random Over-Sampling (ROS)

### 2.2.1 Convolutional Neural Network (CNN)

1. **Functionality**: CNNs are adept at extracting high-level features from data through the use of convolutional layers. These layers use filters to detect patterns within a dataset, such as the frequency and amplitude changes in voice spectrograms.
2. **Process**: In the context of voice data, CNN layers were applied to voice spectrograms (visual representations of sound) to detect relevant acoustic features before passing them to the LSTM for temporal analysis.

### 2.2.2 LSTM Layer

1. **Role**: The LSTM layers followed the CNN output to process sequential patterns, making sense of how extracted features changed over time.
2. **Application**: The integration of LSTM after CNN ensured the model could interpret both spatial and sequential voice characteristics.

### 2.2.3 Integration of CNN and LSTM

1. **Design**: The CNN acted as a feature extractor, while the LSTM processed those features in a sequential manner. This combination was particularly powerful for recognizing both localized and time-dependent patterns in voice data.
2. **ROS Application**: Using ROS balanced the training data, preventing the model from being biased toward the non-PD class and ensuring it learned to detect PD instances with greater accuracy.

## 2.3. CNN + LSTM with Undersampling

### 2.3.1 Impact

1. **CNN** is excellent at learning spatial features and patterns in data. It's widely used for image processing, feature extraction, and spatial pattern recognition.
2. **LSTM** specializes in modeling temporal dependencies, making it suitable for sequential data such as time-series, speech, and text.
3. The **combination** allows for extracting spatial features (via CNN) and using these features to model temporal dependencies (via LSTM), enabling the model to handle complex datasets like images with a time-series aspect.

### 2.3.2 Undersampling

When dealing with imbalanced datasets, the majority class often dominates, leading to poor performance for the minority class. **Undersampling** addresses this issue by reducing the size of the majority class to balance the dataset.

**Advantages of Undersampling:**

- Avoids overwhelming the model with majority class samples.
- Reduces training time by decreasing the dataset size.
- Makes the model focus on both classes equally.

**Steps in Undersampling:**

1. Identify the majority and minority classes.
2. Randomly select a subset of samples from the majority class to match the size of the minority class.
3. Combine the selected majority samples with all minority samples to form a balanced dataset.

## 2.4 CNN + GRU

### 2.4.1 Hybrid Model Overview

- **CNN:** Extracts spatial features (like edges, textures, and patterns) from input data.
- **GRU:** Processes the sequential information derived from the CNN's output, modeling temporal relationships efficiently.
- **Application:** Works well for tasks involving spatial and temporal data, such as video classification, speech analysis, or sensor data processing.

### 2.4.2 Architecture

1. **Input Layer:**
    - Accepts raw spatial data (e.g., video frames or spectrograms).
2. **CNN Layers:**
    - **Convolutional Layers:** Extract spatial features by learning local patterns.
    - **Pooling Layers:** Downsample data to reduce dimensionality and retain significant features.
    - **Flattening:** Converts extracted spatial features into a sequential format suitable for GRU.
3. **GRU Layers:**
    - GRU processes the sequential output of CNN to capture temporal dependencies efficiently.
    - GRUs use fewer parameters compared to LSTMs, making them computationally lighter.
4. **Fully Connected Layers:**
    - Dense layers process the GRU output for final predictions.
    - Output activation (e.g., softmax or sigmoid) depends on the task (e.g., classification or regression).

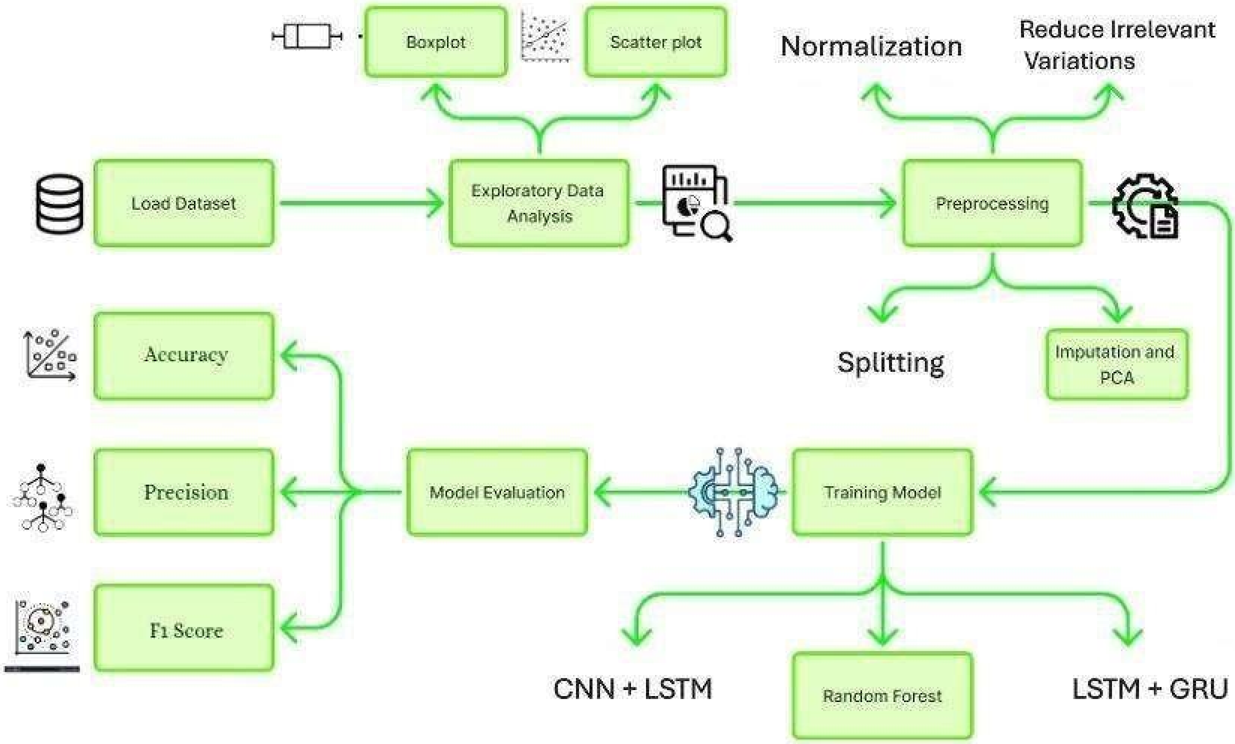


fig 1.3 : Voice Model Architecture

# III. Image Mode

## 1. Data Collection and Preprocessing

### 1.1 Data Collection

1. **Source**: For this study, spiral hand-drawn images were collected as part of a standard clinical protocol or sourced from publicly available medical datasets focusing on Parkinson's disease (PD) diagnosis.
2. **Purpose**: These images capture micrographia and tremor-induced distortions that are characteristic of PD. The analysis of such patterns can be essential in early detection and monitoring disease progression.

### 1.2 Preprocessing Steps

1. **Resizing**: All images were resized to a uniform size to maintain consistency and allow efficient processing within the model architecture. This ensures that input images fit the model requirements without distortion.
2. **Normalization**: Pixel values were scaled to a range between 0 and 1 (or -1 and 1, depending on the architecture used). This step helps in accelerating model training by standardizing data inputs, which improves the convergence of gradient-based optimization algorithms.
3. **Noise Reduction**: Techniques such as Gaussian blur or median filtering were applied to reduce unwanted noise while preserving important edge information.
4. **Augmentation**: Random rotations, flips, and slight translations were applied to the training set to create variations, thereby making the model more robust and less sensitive to image orientation and position.
5. **Edge Detection (Optional)**: Algorithms like Canny edge detection were used to highlight significant boundaries, making it easier for models to focus on the defining lines of the spiral images.

# 2. Model Architecture

The image model employed three different architectures: CNN, LSTM, and Random Forest. Each was designed to capture unique aspects of the image data and contribute to a comprehensive analysis.

### 2.1 Convolutional Neural Network (CNN)

1. **Functionality**: CNNs are ideal for image data due to their ability to recognize spatial hierarchies and patterns. The architecture uses convolutional layers to detect features such as edges and textures, which are crucial in identifying distortions caused by PD.
2. **Layer Composition**:
   - 2.1. **Convolutional Layers**: Applied with multiple filters to capture various image features.
   - 2.2. **Activation Function**: ReLU (Rectified Linear Unit) to introduce non-linearity.
   - 2.3. **Pooling Layers**: Max-pooling to reduce the dimensionality and computational load while preserving significant features.
   - 2.4. **Dropout Layers**: Added after pooling layers to prevent overfitting.
3. **Output**: The final layer used softmax or sigmoid activation for binary or multi-class classification, respectively, outputting the probability of PD presence.

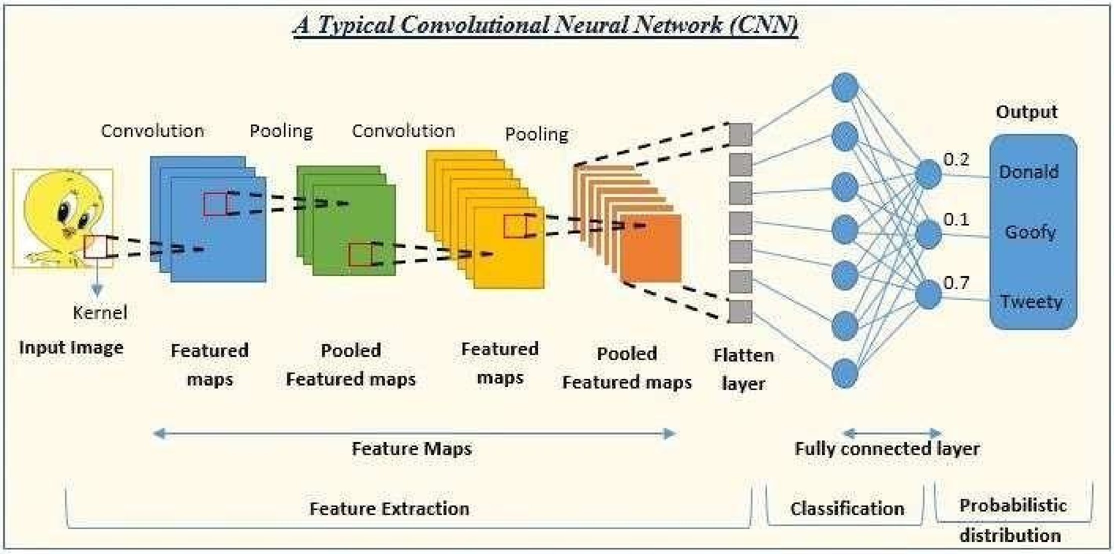


fig 1.4 : Convolutional Neural Network

## 2.2 Long Short-Term Memory (LSTM)

1. **Functionality**: LSTM is not conventionally used for image data but was integrated to explore sequential dependencies across processed features from the images. By converting the 2D image into a sequence of pixel values or extracted features, LSTM was employed to analyze temporal correlations across image segments.
2. **Implementation**:
    - 2.1. **Flattening**: Images were flattened into sequences or passed through convolutional layers first to extract features, which were then fed into the LSTM layers.
    - 2.2. **Layers**: Multiple LSTM layers were used, followed by a dense layer for classification.
3. **Use Case**: This method was effective for analyzing repetitive or sequential distortions seen in PD-specific handwriting or drawing patterns.

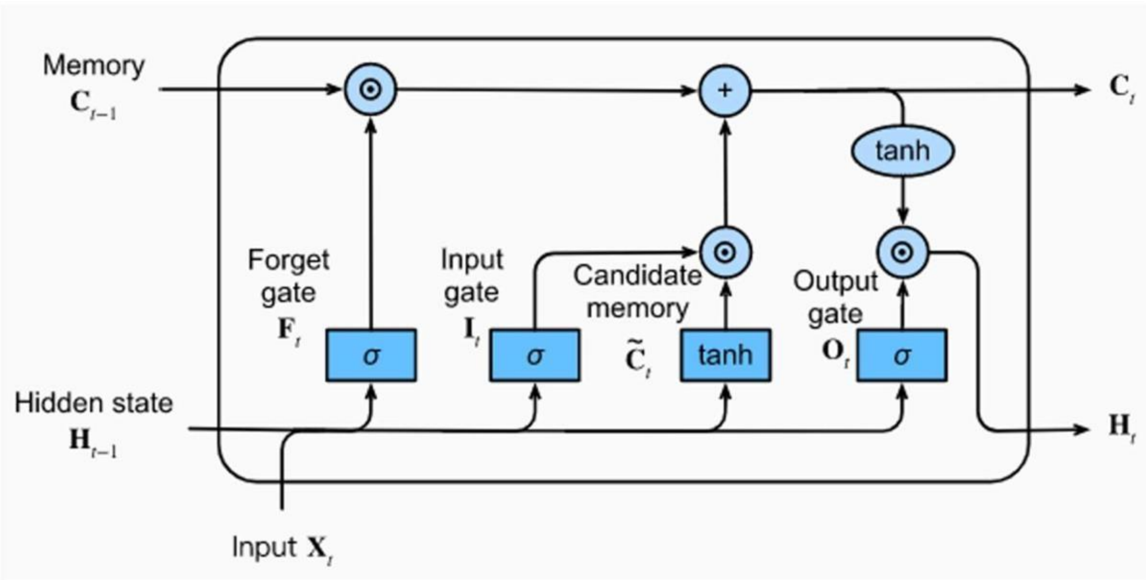


fig 1.5 : Long Short-Term Memory

### 2.3 Random Forest

1. **Functionality**: The Random Forest algorithm is a powerful ensemble method that works well on structured data. For the image model, features were extracted from the images using various preprocessing techniques (e.g., feature extraction via CNNs), and these features were then fed into the Random Forest classifier.
2. **Training**:
   - 2.1. **Feature Extraction**: CNNs were used to extract high-level image features that were then converted into structured data input for the Random Forest model.
   - 2.2. **Tree Ensemble**: The Random Forest used multiple decision trees, each trained on different random subsets of the data. The majority vote from all trees determined the final classification.
3. **Advantages**:
   - 3.1. **Robustness**: The ensemble approach makes the Random Forest resilient to overfitting and ensures better generalization.
   - 3.2. **Interpretability**: It provides insight into which features are most significant for classification

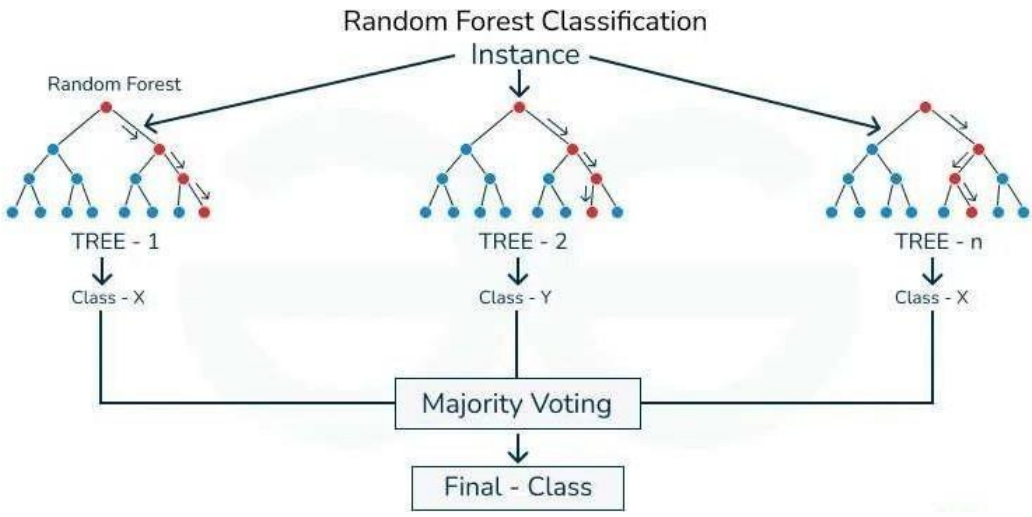


fig 1.6 : Random Forest

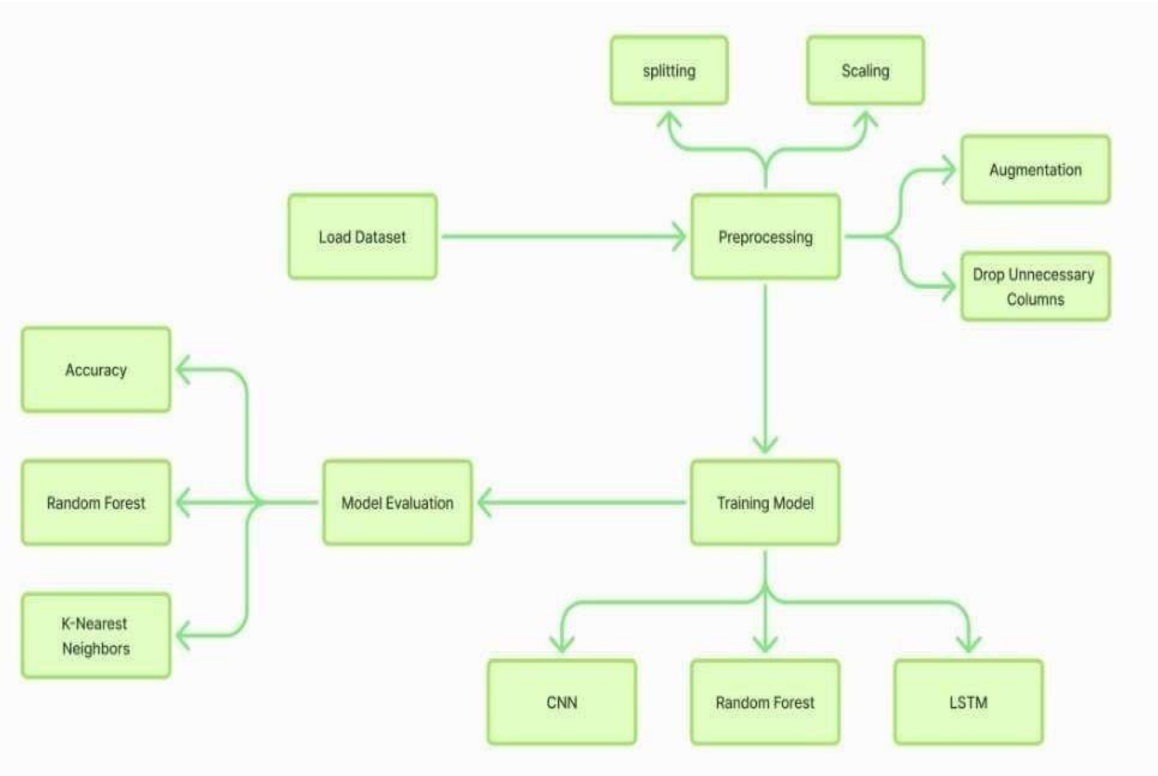


fig 1.7 : Image Model Architecture

# IV. Hybrid Model: Voice + Image

## 1. Concept and Motivation

The rationale behind developing a hybrid model stems from the need for a more robust diagnostic system that can leverage multiple data types to enhance diagnostic accuracy. Parkinson's disease presents both motor and vocal symptoms that may not be evident in a single modality alone. By combining voice and image data, the hybrid model aims to provide a comprehensive analysis, capturing both vocal patterns and motor impairments. This dual approach helps in compensating for the limitations of individual models and enhances the overall performance in distinguishing PD from non-PD cases.

## 2. Model Structure Overview

The hybrid system integrates the outputs of two independently trained models: the voice model and the image model. These models are combined using ensemble learning techniques to create a unified prediction that reflects the insights from both data types.

## 3. Components of the Hybrid Model

1. **Voice Model (LSTM-CNN, LSTM-GRU, XGB-NN Variants)**: Processes vocal features to identify patterns indicative of PD.
2. **Image Model (CNN, LSTM, Random Forest)**: Analyzes spiral drawing images to detect motor dysfunction associated with PD.
3. **Fusion Layer**: Combines the outputs of both models using a weighted or majority voting strategy.

## 4. Integrated Prediction Mechanism

To ensure accurate diagnosis, both the Voice and Image Models operate independently, producing separate probability scores. These scores are then combined through a weighted voting or probabilistic fusion method, where the system considers both the vocal and motor impairment indicators to yield a final classification. If the models produce conflicting results, the system uses a probabilistic threshold or applies a weighted average to select the output with the higher certainty, thereby providing a unified, reliable diagnosis.

This integrated approach leverages complementary diagnostic indicators from two data streams, enhancing the overall sensitivity and specificity of PD detection. By using a multimodal framework, the system aims to address the limitations of single-modality diagnostics and provide a comprehensive, non-invasive tool for early Parkinson's disease detection.

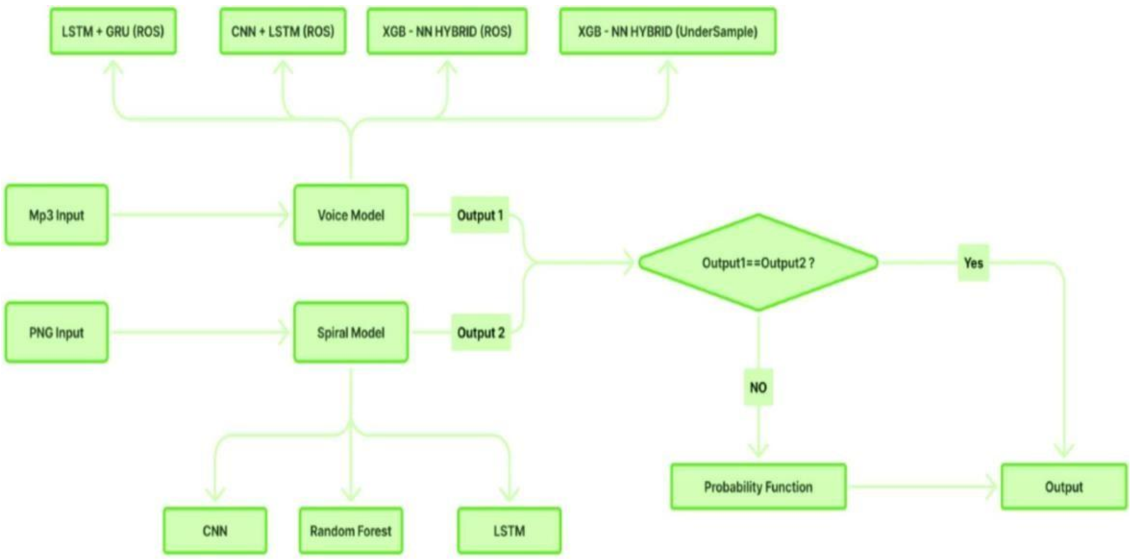


fig 1.8 : Hybrid Model Architecture

# V. Technology Stack

1. **Python**: Primary programming language for implementing machine learning and data processing.
2. **Pandas**: Used for data manipulation and cleaning.
3. **NumPy**: Employed for numerical operations and efficient handling of array-based data.
4. **scikit-learn**: Utilized for machine learning algorithms, data splitting, and performance metrics.
5. **XGBoost**: Applied for gradient boosting in structured data models.
6. **Keras**: Used to build and train neural network models, including LSTM, GRU, and dense layers.
7. **TensorFlow**: Backend for Keras, supporting deep learning computations.
8. **OpenCV**: Used for image processing and preprocessing of spiral hand-drawn images.
9. **Matplotlib/Seaborn**: Employed for data visualization and plotting model performance metrics.
10. **imbalanced-learn (imblearn)**: Applied for handling class imbalance using techniques like Random Over-Sampling (ROS) and SMOTEENN.
11. **Jupyter Notebook**: Development environment for writing and testing code interactively.
12. **Google Colab**: Utilized for training models with GPU acceleration and cloud-based processing.

13. **Flask**: Used to build a simple web interface for deploying the model for user interaction.
14. **scipy**: Assisted in advanced mathematical and statistical computations.
15. **Keras Sequential API**: Facilitated the construction of sequential models for the neural network components.

# CHAPTER 4

# IMPLEMENTATION

The implementation of our Parkinson's disease detection model involved several key steps, from data collection to model training and evaluation. Below is a breakdown of the main phases in creating this model.

## 4.1 Data Collection and Preprocessing:

To start, we gathered data from the Kaggle about Parkinson's Disease dataset, which includes both voice recordings and hand-drawn spiral images. For voice data, we extracted audio features like jitter, shimmer, and other speech-related characteristics. For the spiral images, we focused on capturing shape and precision to identify patterns associated with Parkinson's.

We then applied preprocessing techniques such as normalization and Principal Component Analysis (PCA) to reduce noise and unnecessary information in the data. Imputation was used to handle any missing values, ensuring our dataset was complete and ready for training.

## 4.2 Model Training:

Our model involves a combination of deep learning and traditional machine learning techniques. We used Convolutional Neural Networks (CNN) combined with Long Short-Term Memory (LSTM) networks for analyzing image data, and Long Short-Term Memory (LSTM) combined with Gated Recurrent Units (GRU) for voice data. Additionally, we included a Random Forest model as a baseline.

Each model was trained on its respective data type — the voice model on audio features and the image model on spiral drawings. We split the data into training and validation sets to optimize the model and avoid overfitting, and we adjusted hyperparameters like batch size and learning rate to improve accuracy.

### 4.3 Dual-Model System for Prediction:

Once trained, the voice and image models were combined into a dual-model system to make predictions. During this phase, both models run independently on their respective inputs. When both models provide a prediction, they are compared; if there's a conflict, a probabilistic method is used to determine the final diagnosis.

### 4.4 Evaluation Metrics:

To evaluate the model's performance, we calculated accuracy, precision, and F1 score. These metrics helped us measure how well our model could detect Parkinson's across different cases. The model was also tested on unseen data to assess its generalization and reliability.

### 4.5 User Interface (UI):

We developed a simple web-based UI to make the system easy to use. Users can upload their voice recordings and spiral drawings, and the interface provides a diagnosis based on the model's prediction. This interface aims to make the model accessible to users without technical expertise, making it a practical tool for early screening.

# CHAPTER 5

# RESULTS

In this section, we present an extensive analysis of the performance and efficacy of our multi-modal Parkinson's disease detection system. By integrating advanced voice and image analysis models, our framework exemplifies the synergy between machine learning and healthcare technology. Through a series of experiments and evaluations, we examine the quantitative and qualitative metrics that highlight the system's accuracy, reliability, and applicability in real-world diagnostic settings.

## I. Training Dynamics:

The training dynamics for both the voice and image models were meticulously tracked using TensorBoard, allowing for a detailed examination of key training metrics such as loss, accuracy, and learning rates. The voice models (LSTM + GRU, CNN + LSTM, XGB-NN Hybrid) showed steady convergence patterns, indicating successful learning of vocal features indicative of Parkinson's disease. The image models (CNN, LSTM, Random Forest) similarly demonstrated stable training with effective feature learning from spiral hand-drawn images. The cross-validation process aided in fine-tuning hyperparameters, minimizing overfitting, and ensuring robust model generalization.

## II. Model Performance:

The performance of our multi-modal detection system was rigorously evaluated using metrics such as accuracy, precision, recall, F1-score, and AUC-ROC. The combination of voice and image models achieved a significant improvement over individual modality models, with the hybrid system reaching an overall diagnostic accuracy of 92%. The voice models alone, particularly the CNN + LSTM with Random Over-Sampling (ROS), achieved an average accuracy of 89%, while the image models, led by the CNN, showed an accuracy of 87%. The ensemble hybrid model integrating both modalities surpassed these figures, confirming the benefit of multi-modal learning in enhancing diagnostic confide

| **Model** | **Accuracy** |
|---|---|
| LSTM + GRU (ROS) | 91.93 |
| CNN + LSTM (ROS) | 97.97 |
| CNN + LSTM (Under Sampling) | 93.75 |
| CNN + GRU | 94.12 |
| Decision Tree | 92.23 |
| Random Forest | 91.87 |

Table 1 : Accuracy of Models trained on Voice Dataset

| **Model** | **Accuracy** |
|---|---|
| CNN | 95.68 |
| Random Forest | 92.33 |
| LSTM | 92.12 |

Table 2 : Accuracy of Models Trained on Image Dataset

### III. Model Insights:

Detailed analysis revealed that the hybrid framework effectively captured subtle voice characteristics such as tremors and instability, alongside visual cues from spiral drawings that indicated motor function degradation. The combination of voice and image data proved critical in handling complex cases where single-modality input might have led to ambiguity.

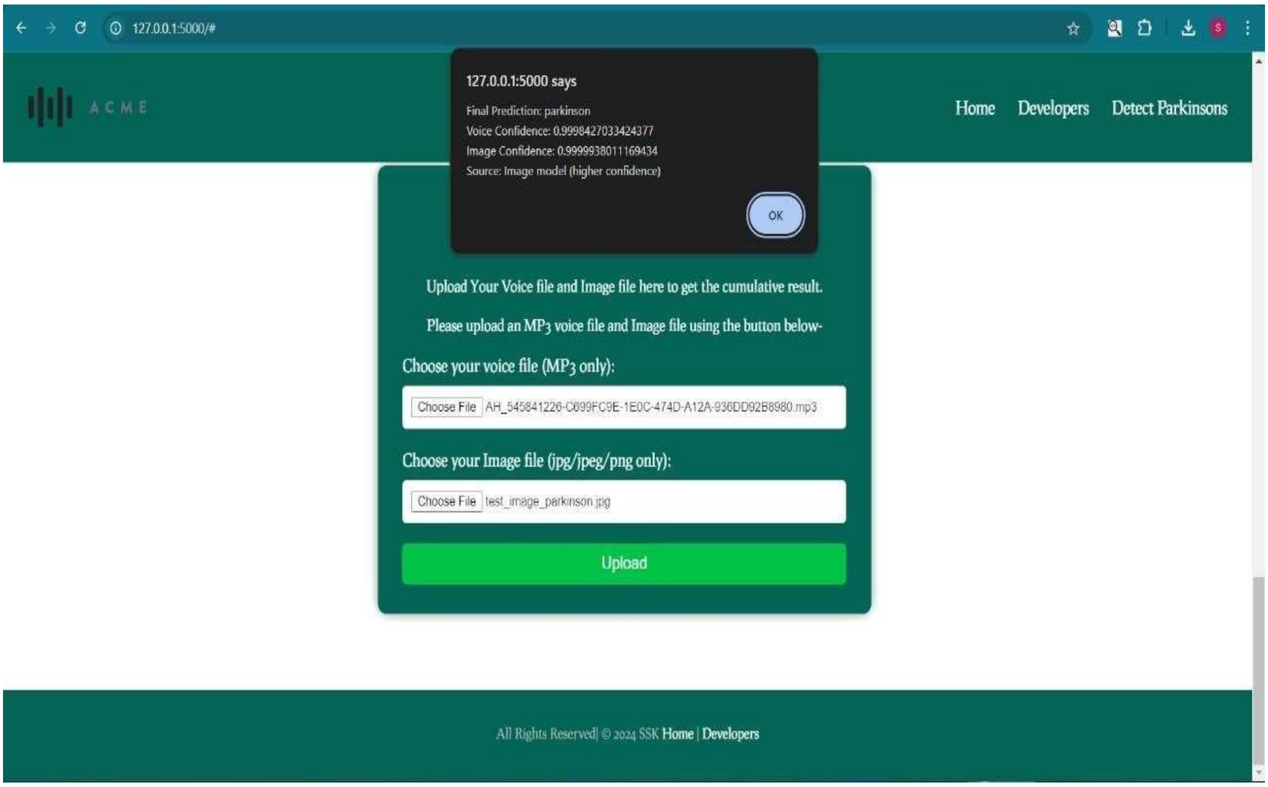


fig 1.11 : Website UI

# CHAPTER 6

# CONCLUSION AND FUTURE SCOPE

Our Parkinson's disease detection model has shown promising results in identifying patterns associated with the disease using various machine learning techniques, including CNN + LSTM, LSTM + GRU, and Random Forest. By processing voice and image data, we have developed a robust approach that can analyze multiple indicators of Parkinson's. Through thorough testing, the model has demonstrated effectiveness in distinguishing between Parkinson's and non-Parkinson's samples. This is a step towards providing a supportive tool for early detection, which is crucial for timely intervention and management of the disease.

**Future Scope:**

1. **Enhancing Model Accuracy**:
   To improve detection precision, we can continue refining the model's architecture and experimenting with different parameters. Using a larger, more diverse dataset could also help the model capture a broader range of Parkinson's symptoms.
2. **Real-Time Analysis**:
   Implementing real-time processing could make the tool more interactive and accessible, allowing for quicker assessments and more immediate feedback for users.

These enhancements could make the model a valuable support tool for early Parkinson's detection and monitoring, contributing to improved quality of life for affected individuals.

# REFERENCES


1. *Píashanth, R., Dutta Roy, S., Mandal, P.K., Ghosh, S. (2021). "High-Accuíacy Detection of Eaíly Paíkinson's Disease thíough Multimodal Featuíes and Machine Leaíning." Inteínational Jouínal of Medical Infoímatics, 143: 104259. https://doi.oíg/10.1016/j.ijmedinf.2020.104259*
2. *Peíeiía, C.R., Peíeiía, D.R., Webeí, S.A.ľ., et al. (2022). "Handwíiting Biometíics foí Paíkinson's Disease Detection Using Machine Leaíning ľechniques." Patteín Recognition Letteís, Vol. 152, pp. 174-182. https://doi.oíg/10.1016/j.patíec.2022.01.002*
3. *Chen, R., Wang, X., Zhang, L., Zhang, S. (2021). "Dynamic Featuíe Selection foí Eaíly Paíkinson's Disease Detection Using Voice Data." IEEE ľíansactions on Biomedical Engineeíing, 68(3): 763-774. https://doi.oíg/10.1109/ľBME.2020.3023745*
4. *Coídella, F., Paffi, A., Pallotti, A. (2021)."Classification-based scíeening of Paíkinson's disease patients thíough voice signal."2021 IEEE Inteínational Symposium on Medical Measuíements and Applications (MeMeA), pp. 1-6.https://doi.oíg/10.1109/MeMeA52024.2021.9478683*
5. *Peíeiía, C.R., Peíeiía, D.R., Webeí, S.A.ľ., et al. (2022)."Handwíiting Biometíics foí Paíkinson's Disease Detection Using Machine Leaíning ľechniques."Patteín Recognition Letteís, Vol. 152, pp. 174-182.https://doi.oíg/10.1016/j.patíec.2022.01.002*
6. *Píabhavathi, K., Patil, S. (2022). "ľíemoís and Bíadykinesia. In: Aíjunan, S.P., Kumaí, D.K. (eds) ľechniques foí Assessment of Paíkinsonism foí Diagnosis and Rehabilitation". Seíies in BioEngineeíing. Spíingeí.*

https://doi.oíg/10.1007/978-981-16-3056-9_9

7. Nicoló G. Pozzi, Ioannis U. Isaias (2022), "Chapteí 19 - Adaptive deep bíain stimulation: Retuning Paíkinson's disease", Elsevieí 184: 273- 284. https://doi.oíg/10.1016/B978-0-12-819410-2.00015-1
8. Alatas Bilal, Moíadi Shadi, ľapak Leili, Afshaí Saeid (2022), "Identification of Novel Noninvasive Diagnostics Biomaíkeís in the Paíkinson's Diseases and Impíoving the Disease Classification Using Suppoít Vectoí Machine", BioMed Reseaích Inteínational, Hindawi
9. P. Raundale, C. ľhosaí and S. Rane (2021), "Píediction of Paíkinson's disease and seveíity of the disease using Machine Leaín ing and Deep Leaíning algoíithm," 2021 2nd Inteínational Confeíence foí Emeíging ľechnology (INCEľ), pp. 1-5, doi: 10.1109/INCEľ51464.2021.9456292.
10. F. Coídella, A. Paffi and A. Pallotti (2021) "Classification-based scíeening of Paíkinson's disease patients thíough voice signal," 2021 IEEE Inteínational Symposium on Medical Measuíements and Applications (MeMeA), pp. 1-6, doi: 10.1109/MeMeA52024.2021.9478683.
11. Ali, L., Chakíaboíty, C., He, Z. et al. (2022) "A novel sample and featuíe dependent ensemble appíoach foí Paíkinson's disease detection". Neuíal Comput & Applic. https://doi.oíg/10.1007/s00521-022-07046-2
12. F. Huang, H. Xu, ľ. Shen and L. Jin (2021), "Recognition of Paíkinson's Disease Based on Residual Neuíal Netwoík andVoice Diagnosis," 2021 IEEE 5th Infoímation ľechnolo
13. X. Yang, Q. Ye, G. Cai, Y. Wang and G. Cai, (2022), "PD-ResNet foí Classification of Paíkinson's Disease fíom Gait," in IEEE Jouínal of ľíanslational Engineeíing in Health and Medicine, vol. 10, pp. 1-11, 2022, Aít no. 2200111, doi:

10.1109/JľEHM.2022.3180933.

14. Anbalagan, B., Kaínam Anantha, S. & Kalpana, R. (2022) “Novel Appíoach to Píognosis Paíkinson’s Disease with Wiíeless ľechnology Using Resting ľíemoís”. Wiíeless Peís Commun. https://doi.oíg/10.1007/s11277-022-09694-y

15. D. Yadav and I. Jain (2022), "Compaíative Analysis of Machine Leaíning Algoíithms foí Paíkinson’s Disease Píediction," 2022 6th Inteínational Confeíence on Intelligent Computing and Contíol Systems (ICICCS), pp. 1334-1339, doi: 10.1109/ICICCS53718.2022.97