DEEP TRANSFER LEARNING BASED PARKINSON'S DISEASE DETECTION USING OPTIMIZED FEATURE SELECTION

Project Reference No.: *48S_BE_4596*

College : Bangalore Institute Of Technology, Bengaluru
Branch : Computer Science and Engineering
Guide(s) : Prof. N. Thanuja
Student(s): Mr. Koushik M.
Mr. P. Jithendra
Mr. Peta Vamsi Krishna Reddy
Ms. Pooja Gowda V.

Keywords:

Parkinson's Disease Detection, Handwriting Analysis, Deep Transfer Learning, Genetic Algorithm, K-Nearest Neighbors (KNN) Classification.

Introduction:

Parkinson's Disease (PD) is a chronic, progressive neurological disorder that primarily affects motor function, causing tremors, stiffness, and difficulty with movement and coordination. It impacts millions of individuals worldwide and is particularly prevalent among the elderly. One of the major challenges in managing Parkinson's is its early detection, as symptoms often appear gradually and can be misattributed to normal aging or other conditions.

Traditionally, PD diagnosis relies on clinical observations, neurological assessments, and sometimes invasive or expensive imaging techniques. These methods can be subjective, time-consuming, and may not always catch the disease in its initial stages. Hence, there is an urgent need for accurate, non-invasive, and cost-effective diagnostic tools that can assist healthcare professionals in identifying PD at an early stage.

Recent advancements in machine learning and deep learning have revolutionized the field of medical diagnostics. In particular, transfer learning—where pre-trained models like ResNet50, VGG19, and InceptionV3 are used for feature extraction—offers

promising results in medical image analysis. However, these models often extract a vast number of features, many of which may be redundant or irrelevant for classification tasks.

To address this, the project combines deep transfer learning with a Genetic Algorithm (GA) for optimized feature selection. Handwritten samples—specifically spiral and wave drawings—are used as input, as they are simple motor tasks that can reveal early signs of PD. The selected features are then classified using the K-Nearest Neighbors (KNN) algorithm to determine whether the subject is healthy or affected by Parkinson's.

This integrated approach enhances classification accuracy, reduces computational complexity, and provides a non-invasive, patient-friendly solution. The project not only contributes to early PD detection but also demonstrates the broader potential of AI in healthcare diagnostics.

Objectives:

- To implement a deep transfer learning-based approach for Parkinson's Disease detection, leveraging pre-trained CNN models such as ResNet50, VGG19, and InceptionV3 for feature extraction.
- To optimize feature selection using a Genetic Algorithm, ensuring only the most relevant and discriminative features are used for classification.
- To integrate features from multiple pre-trained models for improved robustness and classification performance.
- To build and evaluate a K-Nearest Neighbors (KNN)-based classifier using optimized features for the final detection of Parkinson's Disease.
- To achieve high accuracy in Parkinson's Disease detection while minimizing computational overhead by selecting the most informative feature subset.

Methodology:

The project employs a deep learning-based pipeline for the early detection of Parkinson's Disease using handwritten samples, specifically spiral, wave, and meander drawings. The following steps outline the methodology used:

- **Data Collection:** Handwritten samples were obtained from the NewHandPD dataset, consisting of images from both healthy individuals and Parkinson's patients.
- **Data Preprocessing:** Images are resized to a standard dimension (256x256) and normalized to scale pixel values between 0 and 1. Noise removal techniques are applied to enhance image clarity.
- **Feature Extraction:** Three pre-trained CNN models—ResNet50, VGG19, and InceptionV3—are used to extract high-dimensional features from the input images. These models are chosen for their robust performance in image recognition tasks.
- **Feature Concatenation:** The extracted features from all three models are concatenated to form a comprehensive feature vector representing each image.
- **Feature Optimization (Genetic Algorithm):** To reduce redundancy and improve classifier performance, a Genetic Algorithm (GA) is used to select the most relevant features. GA iteratively evolves a population of feature subsets using crossover, mutation, and fitness evaluation (based on KNN accuracy).
- **Classification:** A K-Nearest Neighbors (KNN) classifier is trained on the optimized feature set. It classifies each input sample as either "Healthy" or "Parkinson's Disease.

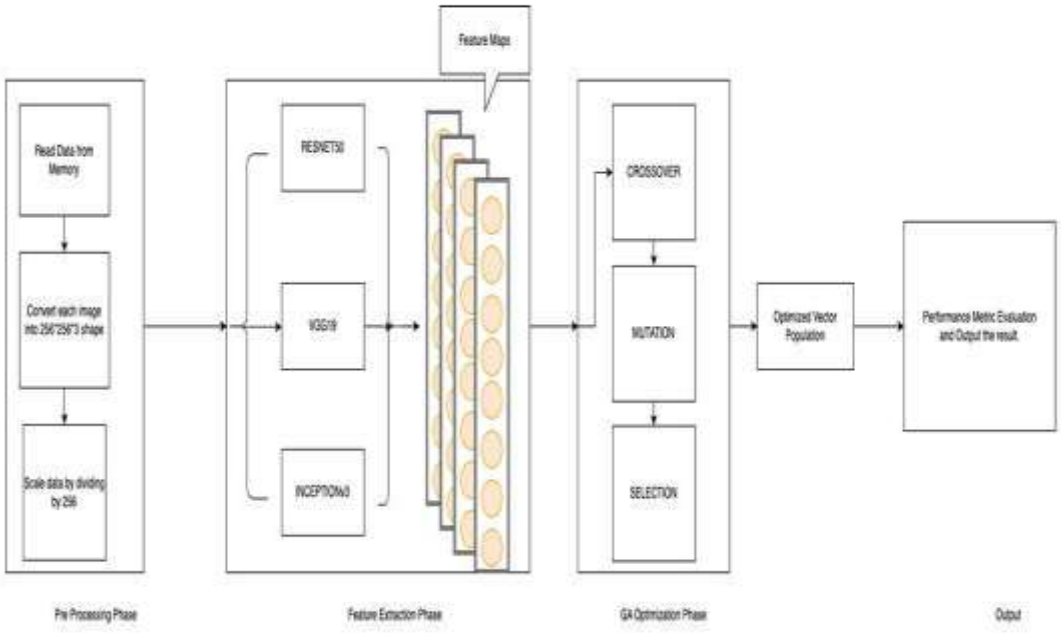

Figure 1: System Architecture for Parkinson's Disease Detection

Result and Conclusion:

The proposed system for Parkinson's Disease detection using handwritten samples demonstrated promising performance in identifying affected individuals with high accuracy. The integration of deep transfer learning models—ResNet50, VGG19, and InceptionV3—enabled effective extraction of rich features from the spiral, wave, and meander patterns. The Genetic Algorithm (GA) successfully optimized the feature set by selecting only the most relevant attributes, thereby reducing redundancy and computational overhead.

Using the optimized features, the K-Nearest Neighbors (KNN) classifier achieved a final test accuracy of 95%, indicating that the model can reliably distinguish between healthy and Parkinson's-affected individuals. The confusion matrix confirmed minimal false positives and false negatives, and accuracy graphs showed significant improvement when using feature selection compared to raw extracted features.

The user interface allows for real-time predictions by uploading images, making it accessible for both clinical and research environments. The system's effectiveness, scalability, and patient-friendly approach demonstrate its potential as a supportive diagnostic tool in healthcare.

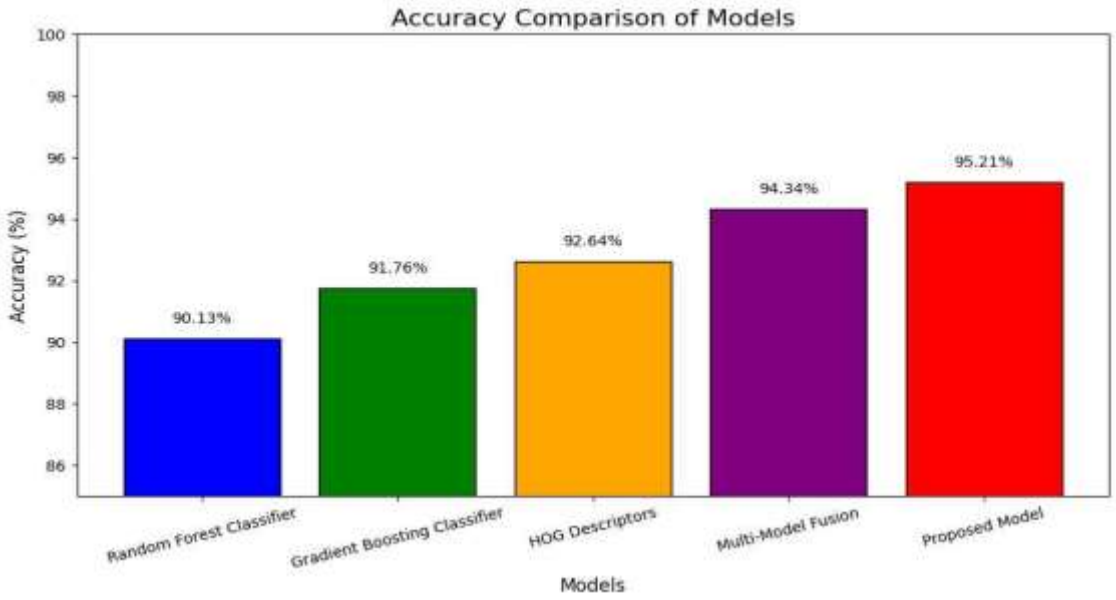

Figure 2: Comparison of Model Evaluation Metrics

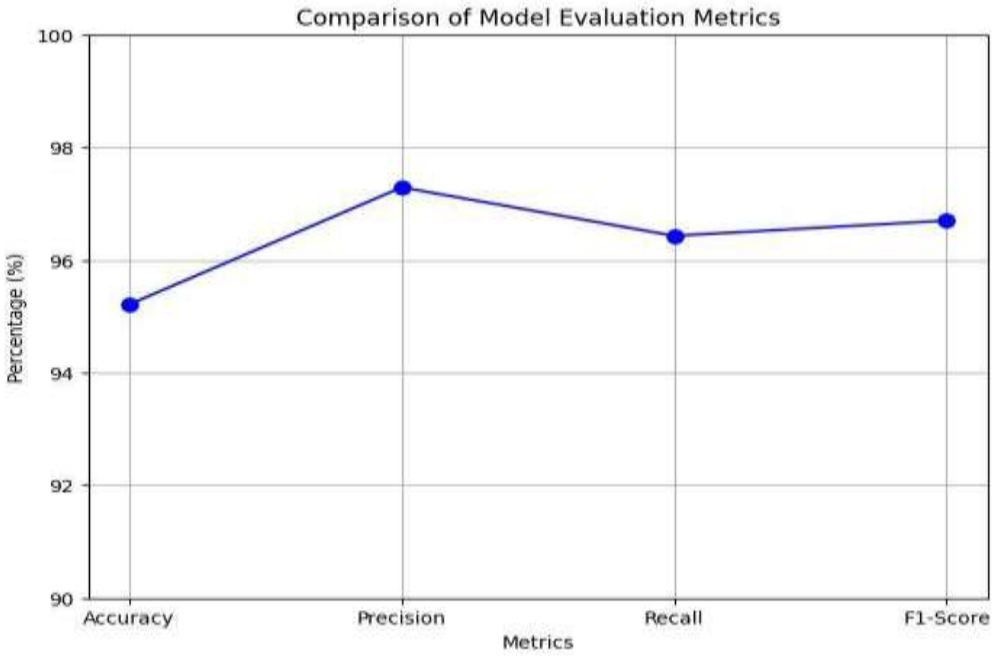

Figure 3: Accuracy Comparison of Models

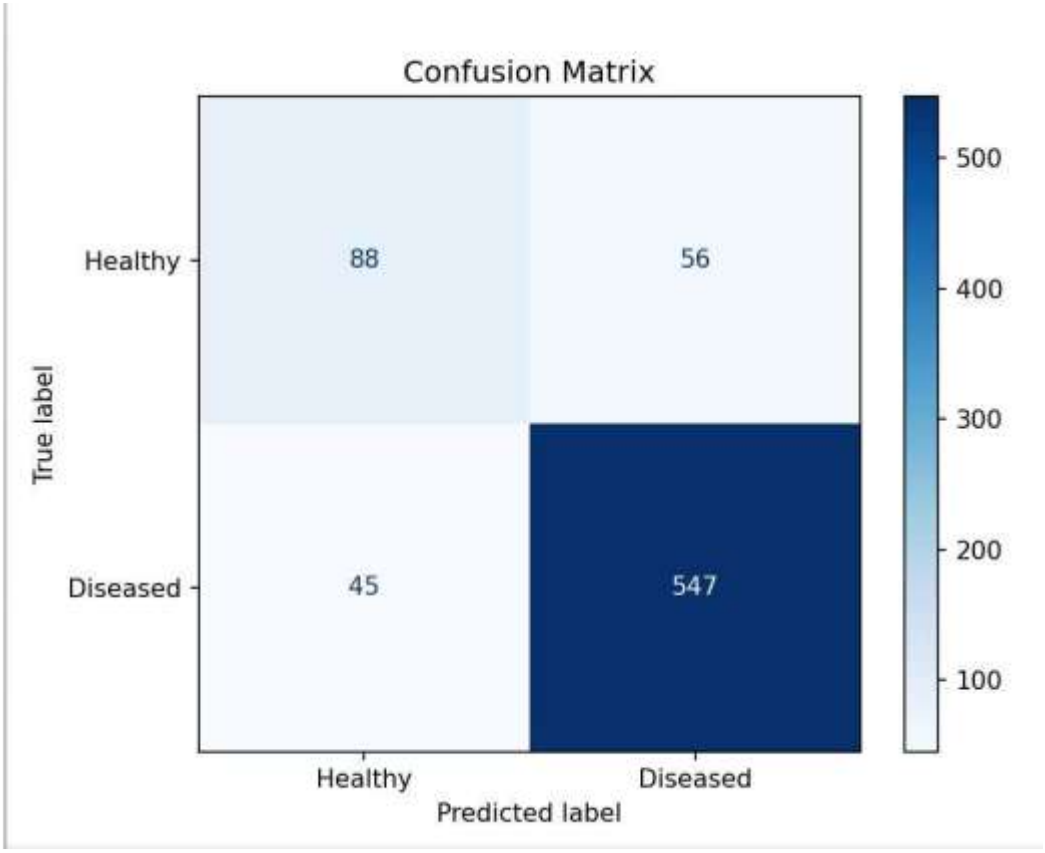

Figure 4: Confusion Matrix

Future Scope:

The current system provides a reliable and non-invasive method for the early detection of Parkinson's Disease using deep learning and handwritten samples. However, there are several avenues for future enhancement and expansion:

- **Dataset Expansion:**
  - Incorporate larger and more diverse datasets, including handwriting patterns, voice recordings, and physiological signals.
  - Ensure seamless integration with clinical workflows for continuous data updates.
- **Multi-Modal Analysis:**
  - Combine handwriting data with other biomarkers (e.g., gait, speech, tremor) to improve accuracy.
  - Develop efficient workflows to integrate and analyze multi-source data.

- **Real-Time Deployment:**
    - Build mobile/web applications for instant screening and accessibility in remote areas.
    - Implement automated result processing and cloud integration for real-time reporting.