Project Proposal: Parkinson's Disease classification using Bio Markers

By: Jahnavi Gujju
CIS 691 02 – Medical and Bioinformatics Capstone
Advisor: Dr. Suhila Sawesi
College Of Computing
Grand Valley State University

CHAPTER 1:

Abstract:

Parkinson's Disease Detection Using MRI Data The early diagnosis of Parkinson's disease is critical for effective management and intervention. This project aims to explore the use of machine learning techniques for the classification of Parkinson's disease based on MRI images. The primary objectives are to develop binary and multi-class classification models and identify key imaging biomarkers specific to Parkinson's disease.

Aim/Objectives:

Objective 1: Binary Classification.

To distinguish MRI images of individuals with Parkinson's disease from healthy controls. A Convolutional Neural Network (CNN) will be trained using augmented data. Key steps include: 1. Data Augmentation: Techniques such as rotation, shearing, and brightness adjustments were applied to balance the dataset. 2. Model Design: A CNN will be developed and optimized for binary classification to separate Parkinson's cases from healthy controls. 3. Evaluation Metrics: The model will be evaluated using metrics such as accuracy, precision, recall, and F1-score.

Objective 2: Multi-Class Classification

It extends the binary classification to a multi-class problem, categorizing MRI images into Parkinson's disease, healthy controls, and other unrelated conditions. Transfer learning techniques will be employed. Key steps include: 1. Model Selection: Pre-trained models like VGG16 will be adapted for multi-class outputs by modifying their classification layers. 2. Fine-Tuning: The weights of the pre-trained models will be selectively adjusted to detect Parkinson's-specific features. 3. Performance Analysis: Metrics like AUC and F1-score will be used to evaluate multi-class performance and handle class imbalances.

Objective 3: Identifying Parkinson's Biomarkers

The final objective is to identify imaging biomarkers associated with Parkinson's disease. The trained models' feature extraction layers will be analyzed to detect patterns that could indicate early stages of the disease. Impact This study aims to demonstrate the potential of machine learning in Parkinson's disease detection using MRI images. The models under development will provide a foundation for scalable, non-invasive diagnostic tools and aim to offer insights into early biomarkers, ultimately improving the understanding and management of Parkinson's disease.

Dataset:

Neurodegenerative Diseases

Problem Statement

Despite advances in medical imaging, the early detection of Parkinson’s disease remains challenging. Traditional diagnostic methods rely heavily on clinical assessments and subjective evaluations, leading to delays in diagnosis and treatment. This project seeks to leverage deep learning models for automatic classification of MRI images to aid in early detection and enhance diagnostic accuracy.

Research Questions: The following research questions will be addressed in this project:

1. How effectively can CNN-based models classify Parkinson’s disease from MRI images compared to traditional methods?
2. Can transfer learning improve the performance of multi-class classification in distinguishing Parkinson’s disease from other neurodegenerative conditions?
3. What key imaging biomarkers can be identified through deep learning models for early-stage Parkinson’s disease?
4. How does data augmentation impact the robustness and generalization of the classification models?

Goals: The goals of this project are to:

- Develop an accurate machine learning model for Parkinson’s disease classification using MRI images.
- Identify key imaging biomarkers that can aid in early diagnosis.
- Explore the use of transfer learning for multi-class classification.
- Improve model performance by applying advanced augmentation and fine-tuning techniques.
- Provide a scalable, non-invasive diagnostic tool for Parkinson’s disease detection.

Methodology:

This project employs a structured methodology to develop deep learning models for the automatic classification of MRI images in the early detection of Parkinson's disease (PD). The methodology consists of several key phases: data collection and preprocessing, model development, model training and evaluation, biomarker identification, and deployment and validation.

1. Data Collection & Preprocessing

- <u>Data Sources</u>: MRI datasets will be collected from established repositories, ensuring a comprehensive and diverse dataset that includes various stages of Parkinson's disease.
- <u>Image Normalization</u>: The collected images will undergo normalization to ensure uniformity in pixel intensity values across different scans. This step is crucial for reducing variability due to differences in imaging equipment and protocols.
- <u>Noise Reduction</u>: Techniques such as Gaussian filtering or non-local means will be applied to reduce noise in the images, enhancing the clarity of features relevant to PD.
- <u>Standardization of Dimensions</u>: All images will be resized to a consistent dimension to facilitate batch processing during model training.
- <u>Data Augmentation</u>: To address class imbalance within the dataset, data augmentation techniques such as rotation, flipping, and zooming will be employed. This enhances the robustness of the model by providing varied training examples.

2. Model Development

- <u>Convolutional Neural Network (CNN)</u>: A CNN will be trained from scratch for binary classification (PD vs. healthy controls), allowing the model to learn features directly from the MRI images.

- Transfer Learning: For multi-class classification (differentiating between PD stages), pre-trained models such as VGG16 and ResNet50 will be utilized. These models have been trained on large datasets and can extract high-level features effectively.
- Fine-Tuning: Pre-trained models will be fine-tuned to focus on Parkinson's-specific features by adjusting their final layers and retraining them on the MRI dataset.

3. Model Training & Evaluation

- Dataset Splitting: The dataset will be divided into training, validation, and test sets to ensure unbiased evaluation of model performance.
- Hyperparameter Optimization: Key hyperparameters such as learning rate, batch size, and activation functions will be optimized using techniques like grid search or random search. This process aims to enhance model performance during training.
- Performance Metrics: The models will be evaluated based on accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). These metrics provide a comprehensive view of model effectiveness in classifying MRI images.

4. Biomarker Identification

- Feature Extraction Analysis: The feature extraction layers of the trained models will be analyzed to identify imaging biomarkers specific to Parkinson's disease. This step is crucial for understanding which aspects of the MRI data are most indicative of PD.
- Explainable AI Techniques: Techniques like Grad-CAM (Gradient-weighted Class Activation Mapping) will be utilized to visualize important features within the MRI scans that contribute to the model's predictions. This enhances interpretability and aids clinical understanding.
- Statistical Validation: Statistical analyses will be conducted on the extracted features to validate their significance as potential biomarkers for Parkinson's disease.

Expected Outcomes:

- Development of an accurate CNN model for Parkinson's disease classification.
- Identification of imaging biomarkers specific to Parkinson's disease.
- Improved multi-class classification through transfer learning.
- A potential non-invasive tool for early detection and diagnosis.
- Published research findings that contribute to the understanding of Parkinson's biomarkers

Conclusion:

This study aims to demonstrate the potential of machine learning in Parkinson's disease detection using MRI images. By developing robust classification models and identifying early biomarkers, the research contributes to non-invasive diagnostic methodologies. The findings could serve as a foundation for further studies in medical imaging and AI-driven healthcare solutions.

Important Milestones:

- Project Proposal and Presentation: 5th February 2025
- Literature Review and Provide existing evidence: 19th February 2025
- Proposed System Design: 12th March 2025
- Meeting on Implementation of the Project: 2nd April 2025
- Final Project Report: 16th April 2025
- Final Project Presentation: 23rd April 2025

Références:

1. PPMI: Parkinson's Progression Markers Initiative. Available at: https://www.ppmi-info.org
2. Alzheimer's Disease Neuroimaging Initiative (ADNI). Available at: https://adni.loni.usc.edu
3. OASIS Brains Dataset. Available at: https://www.oasis-brains.org
4. Khan, M., et al. "Deep learning for neurodegenerative disease classification using MRI data." *IEEE Transactions on Neural Networks and Learning Systems*, 2021.
5. Litjens, G., et al. "A survey on deep learning in medical image analysis." *Medical Image Analysis*, 42, 2017.