Alzheimer’s MRI Predictor

Team Members : Nivithasri A (22PD25) , Valincia John (22PD39)

Problem Statement

Alzheimer’s disease (AD) and its precursor, Mild Cognitive Impairment (MCI), pose significant diagnostic challenges due to subtle MRI and clinical feature differences, particularly for MCI, which is often misclassified as AD or Cognitively Normal (CN).The project aims to develop a robust multimodal deep learning system to classify AD, MCI, and CN using MRI images and clinical features (Age, Sex), improve MCI recall to ~0.75–0.80, and provide an accessible interface with educational support for users.

Abstract

This project presents Alzheimer’s Assistant, a Streamlit-based application integrating an ensemble of ResNet50, DenseNet121, and EfficientNet-B0 models to classify AD, MCI, and CN from MRI scans and clinical data. Using a SMOTE-balanced dataset (~1197 samples, ~33.33% per class), the ensemble employs soft voting to achieve macro F1=0.8008, with AD recall=0.8875 and MCI recall=0.7037, improving over individual models (EfficientNet: MCI recall=0.6049). A chatbot provides Alzheimer’s education, enhancing user engagement. Future work includes weighted voting and additional clinical features (e.g., MMSE) to reach MCI recall ~0.75–0.80 and macro F1 ~0.85.

About Dataset

This dataset is part of the ADNI project, aimed at studying the progression of Alzheimer’s disease through neuroimaging and clinical assessments. It includes data from 369 subjects categorized into three main groups:

1. AD (Alzheimer's Disease): Individuals diagnosed with Alzheimer’s disease.
2. MCI (Mild Cognitive Impairment): Individuals exhibiting mild cognitive impairment, at risk of developing Alzheimer’s.
3. CN (Cognitively Normal): Individuals with no cognitive impairment symptoms.

Link : [ADNI 1 Standardized 1.5T List: Complete 1 Year]

Literature Study

Link : [Automated Alzheimer's disease classification using deep learning models with Soft-NMS and improved ResNet50 integration]

Summary :

The research paper introduces an advanced deep learning framework for automated classification of Alzheimer's Disease (AD) using MRI data. The proposed method integrates an enhanced ResNet50 network for deep feature extraction, Soft Non-Maximum Suppression (Soft-NMS) into a Faster R–CNN model for improved object detection, and a Bidirectional Gated Recurrent Unit (Bi-GRU) to capture sequential patterns in MRI slices. This combination enables more precise detection of AD-related brain abnormalities. Evaluated on the ADNI dataset, the model achieved high accuracy, notably 98.91% in distinguishing AD from cognitively normal individuals. The approach outperformed existing methods in multiple classification tasks (AD vs CN, AD vs MCI, MCI vs CN), demonstrating its robustness and effectiveness for early AD diagnosis. The paper also highlights current limitations such as dataset size and reliance on manual annotations, proposing future work involving multimodal data integration and longitudinal studies for better generalization and clinical relevance.

Methods Used and Implementation

Methods :

Data Preprocessing:

- MRI Images: .nii files converted to 224×224 grayscale PNGs (using the middle slice, e.g., slice 90).
    - Normalization:
        - Training: Mean = 0.5, Std = 0.5

        - Application (inference): Mean = 0, Std = 1 (to be aligned with training)

- Clinical Features:
    - Age: Normalized to [0, 1] range (original range: 50–90)
    - Sex: One-hot encoded — Male: [1, 0], Female: [0, 1]
- SMOTE (Synthetic Minority Over-sampling Technique):
    - Applied to training data to balance classes (~239 samples/class)
    - Generated ~223 synthetic AD MRI samples

Model Evaluation and Results

Evaluation Metrics

- Metrics: Precision, recall, F1-score, macro F1, accuracy, confusion matrix (hypothesized for ensemble).
- Test Set: 242 samples (80 AD, 81 MCI, 81 CN).
- Classification Reports:

  ResNet50:

	precision	recall	f1-score	support
AD	0.9000	0.9000	0.9000	80
MCI	0.8696	0.7407	0.8000	81
CN	0.7957	0.9136	0.8506	81
accuracy			0.8512	242
macro avg	0.8551	0.8514	0.8502	242

DenseNet121:

	precision	recall	f1-score	support
AD	0.8256	0.8875	0.8554	80
MCI	0.8060	0.6667	0.7297	81
CN	0.7640	0.8395	0.8000	81
accuracy			0.7975	242
macro avg	0.7985	0.7979	0.7951	242

EfficientNet-B0:

	precision	recall	f1-score	support
AD	0.7957	0.9250	0.8555	80
MCI	0.8596	0.6049	0.7101	81
CN	0.7283	0.8272	0.7746	81
accuracy			0.7851	242
macro avg	0.7945	0.7857	0.7801	242

Ensemble (Soft Voting):

	precision	recall	f1-score	support
AD	0.8452	0.8875	0.8659	80
MCI	0.8261	0.7037	0.7600	81
CN	0.7416	0.8148	0.7765	81
accuracy			0.8017	242
macro avg	0.8043	0.8020	0.8008	242

Results

Achievements:

- The ensemble model improved MCI recall to 0.7037, surpassing EfficientNet's 0.6049 by leveraging ResNet50's stronger performance (0.7407).

- Achieved AD recall of 0.8875, aligning well with the clinical benchmark of ~0.90.

- Successfully developed a Streamlit web app that integrates model predictions and a chatbot, offering a user-friendly interface for both MRI classification and Alzheimer's-related Q&A.

Areas for Improvement:

- MCI Recall: Although improved to 0.7037, it still falls short of the target range (0.75–0.80). Misclassifications include ~10 cases labeled as AD and ~14 as CN.

- Macro F1 Score: Ensemble F1 score (0.8008) lags behind ResNet50's individual performance (0.8502), indicating potential trade-offs in ensemble blending.

- CN Precision: At 0.7416, suggests some misclassification confusion between CN and other classes.

- Chatbot Integration: While effective at answering ~22 Alzheimer's-related questions, it currently lacks integration with prediction outputs or follow-up explanation features.

App Performance:

- Efficiently processes .nii MRI files, converts them to PNG format, performs predictions using the ensemble model, and displays results in a clean, two-column UI.

- Provides an interactive chatbot experience for user queries, enhancing educational support alongside diagnostic predictions.

Learning

1. Handling MCI Classification Remains Challenging

    - MCI samples showed lower recall due to subtle MRI features and clinical overlap with AD and CN.

- While Focal Loss ($\gamma = 2.0$) and weighted sampling (MCI weight = 1.5) helped, further tuning (e.g., $\gamma = 2.5$, weight ≈ 2.0) and richer inputs (e.g., MMSE, CDR scores) are recommended.

- ResNet50 with Cross-Entropy loss outperformed other models on MCI, suggesting simpler loss functions may be more effective in some cases.

2. Ensemble Design Can Be Optimized

- Soft voting improved MCI recall but equal weighting limited its potential. A weighted voting strategy (e.g., 0.5 ResNet50, 0.3 DenseNet121, 0.2 EfficientNet) could better reflect individual model strengths.

- Ensemble effectiveness heavily depends on aligning with the best-performing base learner.

3. Implementation & Preprocessing Gaps Affect Robustness

- Using a single MRI slice (slice 90) restricts spatial information; incorporating multi-slice or averaged inputs may enhance classification.

- Inconsistent normalization (training vs app) can degrade model accuracy—aligning these settings is crucial for reliable deployment.

GitHub link : [Alzheimer's-MRI-Predictor]()