Abstract

This presentation provides an overview of the research and methodologies employed in the classification of Alzheimer's Disease (AD) using MRI image biomarkers. The study focuses on the use of realistic data sourced from ADNI, prioritizing imaging data for its extensive historical availability compared to genetic and proteomic markers like APOE, Tau, and Beta-Amyloid. The Axial angle was selected for the MRI's acquisitions plane to enhance data quality, particularly for analyzing the Lateral Ventricle, a key indicator in Alzheimer's detection. A custom CNN was used for targeted image selection, emphasizing the "butterfly" shape characteristic of the Lateral Ventricle. Image enhancement techniques, including gamma correction and adaptive thresholding, were applied to improve contrast and edge clarity. Our Inception Net model achieved an accuracy of 81.7%, demonstrating the potential of our approach with the data collected. Future plans include exploring additional datasets and advanced image processing techniques to further improve diagnostic accuracy.

Data Preprocessing for the Classification of Alzheimer's Disease

Farhaan Pishori

Biomarkers

Biomarkers are biological measures that indicate the presence or state of a disease. Provide information for disease diagnosis, prognosis, and monitoring

Types of Biomarkers:

- Genetic: DNA, RNA variations (APOE)
- Proteomic: Protein levels & patterns (Beta-Amyloid & Tau proteins)
- Image: MRI, CT, PET scans

Biomarker Selection

Genetic and Proteomic Biomarkers:

- Presence of **APOE ε4 Allele** increases Alzheimer's risk and influences age of onset
- Increased levels of **Tau & Beta-Amyloid proteins** in the brain lead to a higher risk.
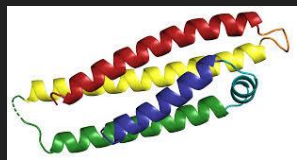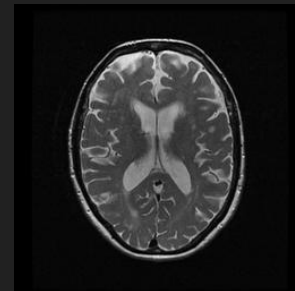- Limited historical data compared to imaging biomarkers
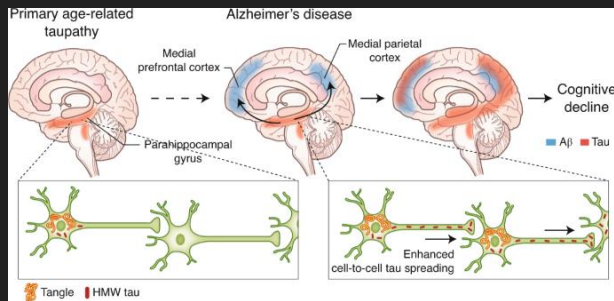
Image Biomarkers (MRI):

- Displays **brain shrinkage/atrophy** which leads to higher Alzheimer's risk
- Enlargement of the **Lateral Ventricle** suggests increased risk due to surrounding brain tissue loss
- Extensive Historical Data

Primary age-related taupathy / Alzheimer's disease diagram with labels: Medial prefrontal cortex, Medial parietal cortex, Parahippocampal gyrus, Cognitive decline, Aβ, Tau, Enhanced cell-to-cell tau spreading, Tangle, HMW tau
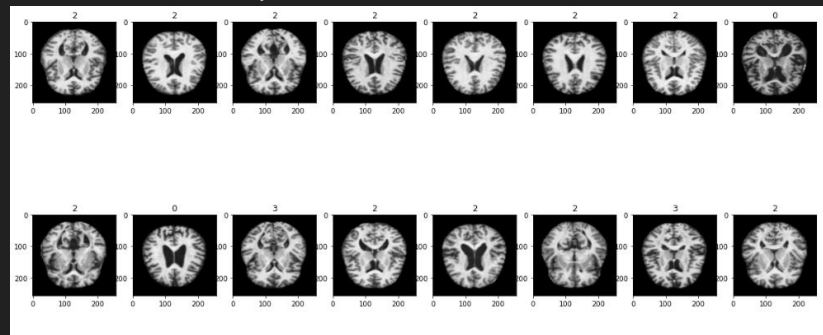
Background

- <u>Classification:</u> Alzheimer's Disease (AD), Cognitively Normal (CN), Mild Cognitive Impairment (MCI)
- <u>Data Source:</u> ADNI
- <u>Acquisition Plane:</u> Axial
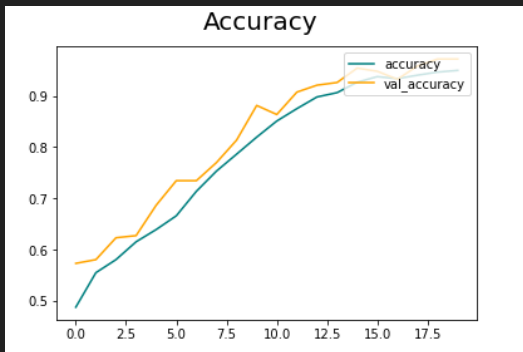- <u>Series Description:</u> Axial PD/T2 FSE

Importance of Data Quality

Kaggle Dataset - MRI images of 4 categories (96% Accuracy)

- Non Demented
- Very Mild Demented
- Mild Demented
- Moderate Demented

High accuracy mainly due to unrealistic data

Classification Report:

	precision	recall	f1-score	support
0	0.94	0.94	0.94	120
1	1.00	0.75	0.86	8
2	0.94	0.99	0.96	450
3	0.97	0.91	0.94	350
accuracy			0.95	928
macro avg	0.96	0.90	0.93	928
weighted avg	0.95	0.95	0.95	928

Acquisition Plane & Series Description
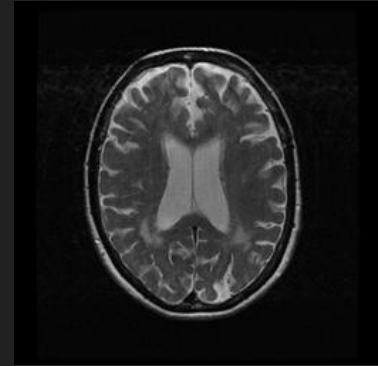

Axial PD/T2 FSE

Acquisition Plane: Axial Angle

- Chosen for having the **most data**
- Recommended by **Dr. Stevens** for providing more **variation in brain sections** relating to Alzheimer's
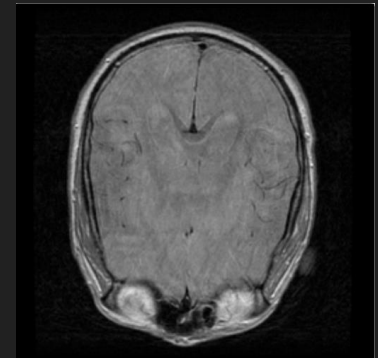- Specifically enhances the **Lateral Ventricle**

Series Description: Axial PD/T2 FSE

- Color differentiation
- Had the **most data** available

3D plane localizer

Data Collection

Data Overview

- Collected data from **1,716 MRI scans** (totaling **84,423 images**)

Python Scripts

- Converted **DICOM** images to **JPEG**
- Renamed and reorganized the images into a machine readable format

Dicom to JPG

```python
import os
import pydicom
import numpy as np
from PIL import Image

def convert_dcm_to_jpg(dcm_file_path, jpg_file_path):

    dicom_image = pydicom.dcmread(dcm_file_path)
    pixel_array = dicom_image.pixel_array

    pixel_array = pixel_array - np.min(pixel_array)
    pixel_array = (pixel_array / np.max(pixel_array) * 255).astype(np.uint8)

    image = Image.fromarray(pixel_array)

    image.save(jpg_file_path)

def traverse_and_convert(root_folder):
    for root, _, files in os.walk(root_folder):
        for file in files:
            if file.lower().endswith('.dcm'):
                dcm_file_path = os.path.join(root, file)
                jpg_file_path = os.path.splitext(dcm_file_path)[0] + '.jpg'
                convert_dcm_to_jpg(dcm_file_path, jpg_file_path)
                print(f'Converted: {dcm_file_path} to {jpg_file_path}')
                os.remove(dcm_file_path)
                print(f'Deleted: {dcm_file_path}')

if __name__ == "__main__":
    root_folder = 'ADNI'
    traverse_and_convert(root_folder)
```

Renaming Files

```python
import os
import re

def rename_jpg_files(directory):

    pattern = re.compile(r'^.*?_\d{17}_')

    for root, dirs, files in os.walk(directory):
        for file in files:
            if file.endswith('.jpg'):
                # Check if the filename matches the pattern
                match = pattern.search(file)
                if match:

                    new_name = file[match.end():]

                    old_path = os.path.join(root, file)
                    new_path = os.path.join(root, new_name)

                    os.rename(old_path, new_path)
                    print(f"Renamed: {old_path} -> {new_path}")

base_directory = "butterflymodel"

rename_jpg_files(base_directory)
```
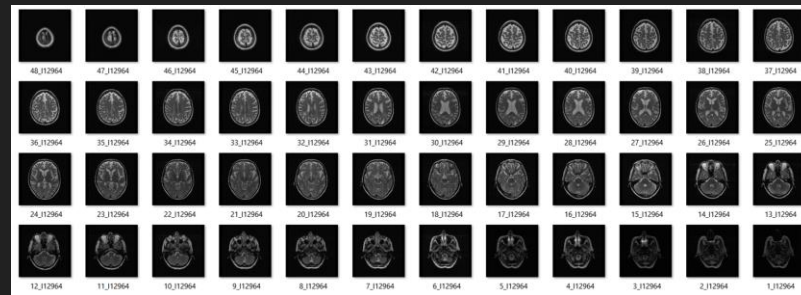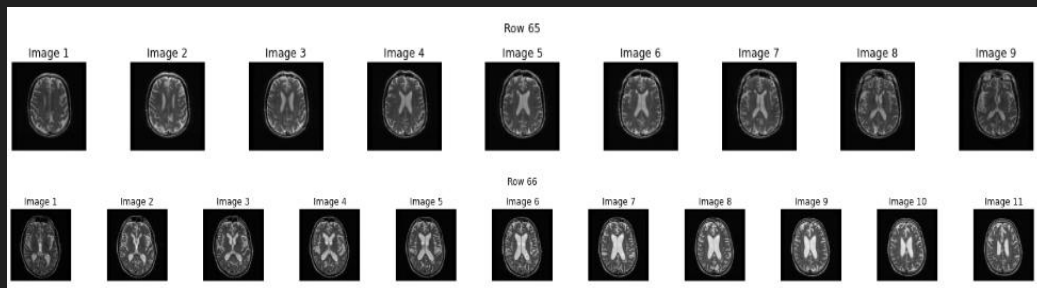
Image Selection

- Initial Model Approach:
  - Used all 50 MRI slices per scan, resulting in excessive noise and a **35% accuracy** (random guessing)
- Key Insights:
  - Realized the need to reduce the number of slices
  - Consulted with our mentor and identified the **Lateral Ventricle** as a critical factor in Alzheimer's diagnosis

Image Selection Cont...

- Focused on the **Lateral Ventricle** using a custom Convolutional Neural Network (CNN)
- Selected the middle 4 images with the characteristic "butterfly" shape

Image Enhancement

- Zoomed in on MRI scans to focus on the Lateral Ventricle
- Applied filters for clarity:
    - **Gamma correction** for better contrast
    - **Adaptive thresholding** to sharpen edges
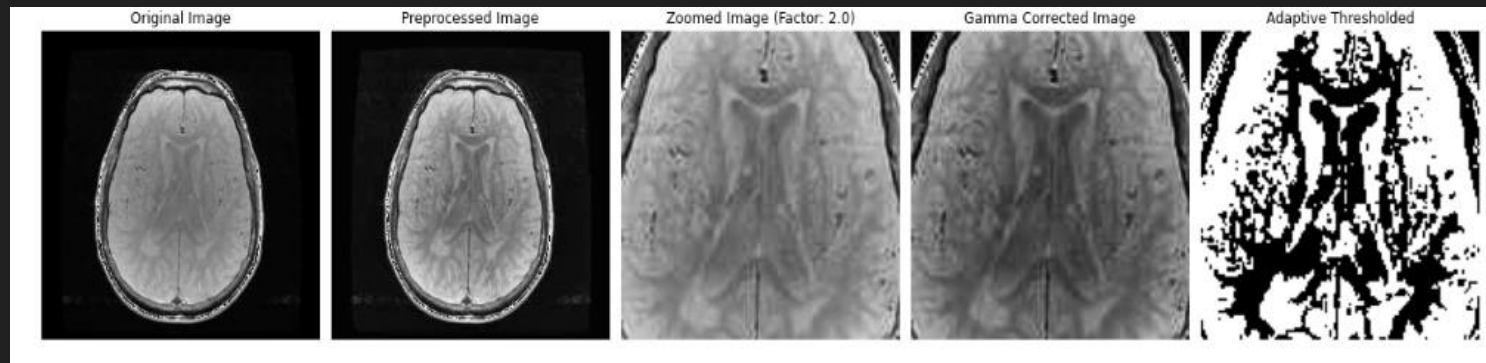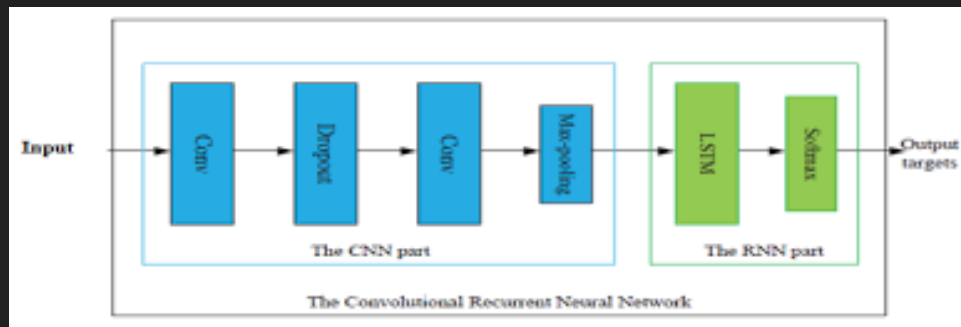
CRNN and Inception Net
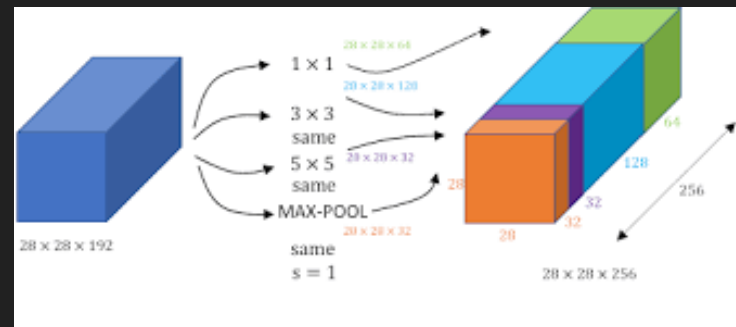
CRNN

Parameters: 2.5 million
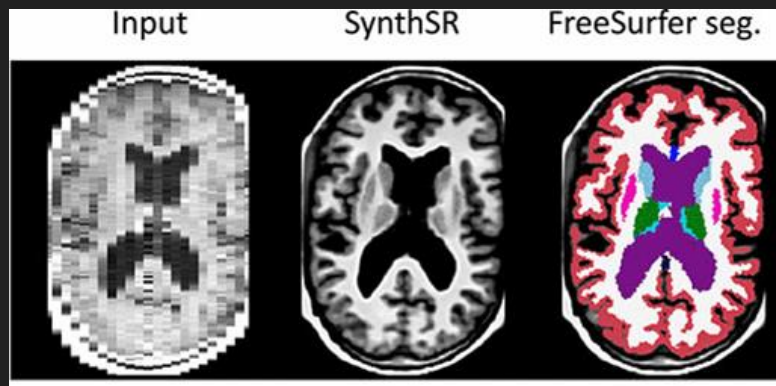
Accuracy: 75%

Inception Net

Parameters: 22 million

Accuracy: 81.7%

Future Plans

- Explore additional reputable datasets like **Oasis-1** or **Oasis-2** for improved accuracy
- Investigate using **Freesurfer** for image enhancement
- Further down the line, include other biomarkers such as APOE e4 allele and Tau/Beta-Amyloid levels

Acknowledgements

Thank you! Questions?