Alzheimer's Detection from MRI Scans

Medical Image Analysis (Semester Project)

Group: Group #7

MSCSF24M006 **Muhammad Haweras**
Supervised By: **Dr. Zobia Suhail**

Department of Computer Science, FCIT, Allama Iqbal Campus, Lahore

Contents

1 Introduction **2**

2 Related Work **3**

3 Dataset **4**
3.1 Dataset Structure 4
3.2 Disease Severity Classification 4

4 Methodology **5**
4.1 Class Balancing 5
4.2 Pre Processing 5
4.3 Features Extraction 5
4.4 ML Classifier Selection 5
4.5 Applying Deep Learning 6
4.5.1 About the AlexNet Architecture 6
4.5.2 AlexNet Architecture Implementation 6

5 Results **7**
5.1 RandomForestClassifier 7
5.1.1 Binary classification 7
5.1.2 Multi-class classification 8
5.2 AlexNet Architecture 9

6 Conclusion **11**

7 Future Directions **11**

Alzheimer's Detection from MRI Scans

Abstract

Alzheimer's disease is the most common cause of dementia. A person with dementia shows symptoms like memory loss,impaired reasoning, struggling with daily activities etc. The effect of Alzheimer on the brain is the accumulation of toxic proteins Plaques and tangles, which destroys nerve cells and eventually shrink the size of the brain. It starts with the hippocampus and in the later stages it spread over the whole brain shrinking it thoroughly. Alzheimer has been classified into 3 stages based on the severity of the disease. Different mechanisms can be applied for diagnosis and detection of Alzheimer's disease, among which is the detection through the MRI scans of the brain. Different image processing techniques have been employed for the detection of Alzheimer disease from the MRI Scans. For the past several years, Deep Learning Models have outdated all the previous techniques and now deep learning models just need dataset and they themselves learn and extracts feature from that data and are able to correctly classify any MRI Scan into 1 of the 4 possible classes relating to Alzheimer's detection. We manually extracted first and second-order statistical features from MRI scan images for binary and multiclass classification. These features were used to train various machine learning classifiers, and the best classifier was selected based on their performance. Additionally, we implemented a deep learning-based approach using the AlexNet architecture for multiclass classification. The classification results of the AlexNet model were compared with traditional machine learning approaches to assess its effectiveness in accurately identifying different stages of Alzheimer's disease.

1 Introduction

Dementia is a syndrome that can be caused by a number of diseases that affect memory, thinking, and the ability to perform daily activities. According to WHO's 2023 Report, currently more than 55 million people have dementia worldwide, **Alzheimer disease** is the most common form of dementia and may contribute to 60–70% of cases. To understand what happens to our brain in Alzheimer's disease, we need to understand the functioning of our brain. The brain comprises of 100 billion nerve cells which are connected through branches and connect at more than 100 trillions points called Synapses. At Synapses, bursts of chemicals called neurotransmitters are released. Now in case of Alzheimer's disease, 2 type of toxic proteins; **Plaques**(which are deposits of a protein fragment called beta-amyloid that build up in the spaces between nerve cells) and **Tangles**(which are twisted fibers of another protein called tau that build up inside cells) start accumulating in our brain. These proteins disrupt the working of electrical charges in our nerve cells as well as the working of neurotransmitters at the synapses, eventually resulting in the death of some nerve cells.

In the beginning, this mechanism start happening in the part of our brain responsible for memory functions (**hippocampus**) but since Alzheimer is a progressive disorder, it doesn't stay limited to the hippocampus, it keeps growing and affecting our parts, and end up destroying all parts of our cortex and can even cause death of the patient.

Based on the severity, Alzheimer's disease is classified into 3 stages; Mild Dementia, Moderate Dementia, Severe Dementia(can also be labeled as Very Mild Dementia). The initial changes in the brain can be observed many years before the actual dementia symptoms start showing up in the person's behavior. That's why Alzheimer's early detection is one of the primary concerns of scientists.

A lot of medical tests are conducted to diagnose Alzheimer and differentiate between Alzheimer cases and other Dementia cases. These tests include blood and urine tests, different cognitive functionalities tests, different senses tests etc and most importantly Brain Imaging mechanisms like MRI and CT scans are used to study Alzheimer. For the scope of our research, we're attempting to contribute towards the cause by providing state of the art mechanisms for studying **MRI Scans** to classify them into either non-demented or one of the 3 categories of Alzheimer.

In the past, several techniques and methods relating to image processing were used to determine Alzheimer, but now over the past few years, with the advent of advanced deep learning models, all other techniques are considered outdated and deep learning models are considered the go to way for detecting Alzheimer in the MRI Scans because of the level of precision and accuracy they hold. These models can learn and extract features from given datasets on their own and can develop the capability to classify any MRI Scan into one of the 4 possible classes associated with our problem.

2 Related Work

Shrikant Patro et al.(2019) [1] explores methods to detect Alzheimer's Disease (AD) at an early stage using MRI-based image processing. The methodology used in this study to detect AD involves preprocessing of MRI images, including resizing and intensity adjustments, followed with segmentation techniques such as bicubic interpolation and watershed segmentation to find areas of interest. These techniques enhance the visibility of brain regions to identify hippocampus atrophy and brain atrophy, which are the key indicators of AD progression. The methods used for detection of AD using segmentation and thresholding, gave 91.6% accuracy rate. These image processing techniques discussed in this study limited to single image processing but future improvements are suggested, including neural networks and machine learning models like SVM and CNN, to enable automated and more precise classification across larger image datasets.

Amir Ebrahimi et al.(2021) [2] implemented and compared several deep models and configurations, including two-dimensional (2D) and three-dimensional (3D) CNNs and recurrent neural networks (RNNs). In the first approach, a 2D CNN was trained on MRI slices in the single-view mode. Then for classification, the CNN model made decisions on all slices of one patient on the particular view. A majority voting mechanism was applied in multi-view mode to make the final decision on three views of an MRI volume. In the second approach, an LSTM model was used to classify a sequence of MRI slices in multi-view and single-view modes. The multi-view models were slightly more robust and accurate compared with single-view models. In the third approach, a 3D CNN was employed to classify MRI volumes, each in a single decision. These approaches demonstrated that the voxel-based method with transfer learning from ImageNet to MRI datasets using 3D CNNs considerably improved the results compared with the others, achieving 96.88% accuracy, 100% sensitivity, and 94.12% specificity.

Fatemah H Alghamedy et al.(2022) [3] discussed about multimodel computing approach with the goal of classifying and and locating Alzheimer's disease. With a collection of datasets CLAHE algorithm used as preprocessor of images. K-map technique used to segmentation and find area of interest. Three different type of classifier were used and it was concluded that SVM-SF has higher accuracy, specificity and precision other then ANN and ID3. When it comes to sensitivity and recall ANN took the lead.

EM Muhammad et al.(2024) [4] did research for Arab countries; absence of comprehensive medical records to train that specific model faced challenges. Used a pre trained CNN model and given a practical implementation of detecting AD. Given research about the major datasets ADNI vs OASIS for training models. It was concluded that during this era reach on different platforms like Science Direct, IEEE and Springer; VGG model is vastly used for best-achieved performance.

PK Pandey et al.(2019) [5] demonstrated the significant potential of using advanced deep learning techniques with transfer learning for early diagnosis of Alzheimer's Disease and MCI. The findings emphasized the effectiveness of ResNet-101 for multi-class classification, aiding healthcare professionals in making better diagnostic and treatment decisions

Anita et al.(2016) [6] argues that for early detection of Alzheimer using MRI Scans, we should focus on the Hippocampus region of the brain because that's where Alzheimer start putting an action in the initial stages. The authors first discussed different existing segmentation techniques in medical image processing literature and then purposed their own segmentation method to segment the hippocampus and classify it. The brain images converted into binary form using two approaches.The first approach is block mean, mask and labeling concepts and in the second approach top hat, mask and labeling concepts. However it is found that some part of the image contains holes which interrupt the segmentation process. To overcome this problem image hole filling techniques are implemented and related components are grouped into connected components. The shape analysis of hippocampus structure will result in classifying the Alzheimer's disease.

Shanthi et al.(2013) [7] claims that they came up with a succesfull method for segmenting the brain from MRI Scans without human intervention. They argue that their method can be used to detect any variation in the total volume or size of the brain(since in Alzheimer detection, brain volume holds a significant importance). They discussed an algorithm which takes care of intensity variations of pixels hence helps in accurate determination of volume changes of white matter, gray matter and CSF. Existing skull stripping algorithms are discussed and role of human intervention in them is shown. Dynamic Thresholding techniques are discussed and a new mechanism for providing the starter values of the Seed Region Growing(SRG) automatically instead of setting them manually. New mechanisms introduced for volumetric calculations of the White Matter, Grey Matter and CSF Regions. Eventually, this talks about considering the volumetric changes in the brain for the detection of Alzheimer disease.

3 Dataset

The Alzheimer's Disease Multiclass Dataset contains approximately 44,000 MRI images categorized into four distinct classes based on the severity of Alzheimer's disease. All images are skull-stripped and clean of non-brain tissue.

3.1 Dataset Structure

The dataset is organized into the following four directories, each representing a different class of disease severity:

NonDemented: Contains 12,800 MRI images of subjects with no signs of dementia.
VeryMildDemented: Contains 11,200 MRI images of subjects with very mild symptoms of dementia.
MildDemented: Contains 10,000 MRI images of subjects with mild dementia.
ModerateDemented: Contains 10,000 MRI images of subjects with moderate dementia.

Total Number of Images: 44,000
Image Format: MRI scans as .JPG files

3.2 Disease Severity Classification

The data set follows a severity ranking system for Alzheimer's disease:

NonDemented: No dementia.
VeryMildDemented: Early signs of dementia, very mild symptoms.
MildDemented: Clear signs of dementia, but still mild.
ModerateDemented: More pronounced symptoms of dementia, moderate severity.

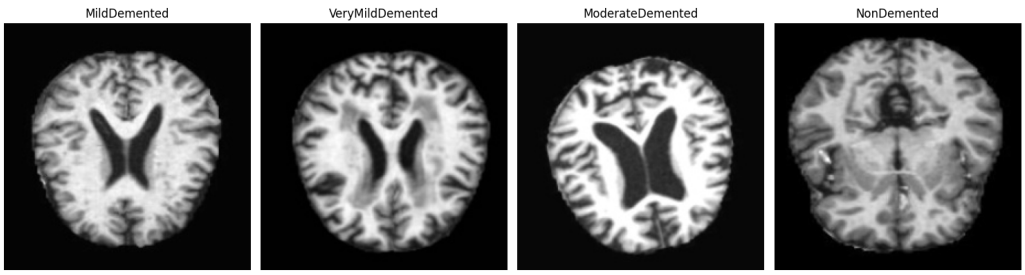

Figure 1: Different Stages of Dementia

4 Methodology

4.1 Class Balancing

To ensure a balanced dataset and reduce biasness in our Alzheimer's classification model, we equalized the number of images across all classes by selecting 3,000 images per class. The process was automated by a script programmed in python that created a new directory structure for the balanced dataset, and copied the selected images. This method ensured equal representation of all classes, promoting fairness, and improving the generalization of the classification model. Now, our new dataset contains 3000 images in each of the 4 folders, resulting in balanced classes.

4.2 Pre Processing

The images in our dataset were 3D and in jpg format, the size of the images was very uneven which was not suitable for performing image analysis. We convert the images into grayscale and resize the images to size of 227,227 using the CV Library. To perform these activities on a bulk level and handle all images within the directory together, we used glob function and used cv functions to write the output images in the new directory and obtained a new dataset which contains 3000 images in each folder with all of the images in grayscale with a size of 227,227.

4.3 Features Extraction

To prepare the dataset for Alzheimer's classification, we extracted both first-order and second-order statistical features from grayscale images. The dataset consisted of four classes (MildDemented, ModerateDemented, NonDemented, and VeryMildDemented), with each class mapped to a unique label. Images were processed as follows:

First-Order Statistical Features:

For each image, we calculated features such as mean, variance, standard deviation, median, skewness, kurtosis, mean absolute deviation, and median absolute deviation.

Second-Order Statistical Features:

Using the Gray Level Co-occurrence Matrix (GLCM), we computed energy, contrast, correlation, entropy, and inverse difference moment (homogeneity). Additionally, Local Binary Patterns (LBP) were used to calculate texture variance.

Results Storage:

Extracted features were stored in a structured format with columns representing feature names and the corresponding class label. The final dataset was saved as an Excel file by using pandas library, for further analysis.

4.4 ML Classifier Selection

To prepare the dataset for model training and evaluation, the extracted features were separated into input features (X) and their corresponding class labels (y). The dataset was then split into training (80%) and testing (20%) by using train_test_split model from sklearn library.
Then we train the dataset on different classifiers e.g. SVM, KNN, GradientBoosting, RandomForestClassifier. By comparing their accuracy **RandomForestClassifier** gave best accuracy among these classifiers.

4.5 Applying Deep Learning

As suggested by results (in Section 5.1.2), we need to achieve more accuracy. So for this purpose we used a deep learning architecture called AlexNet.

4.5.1 About the AlexNet Architecture

AlexNet consists of 5 convolution layers, 3 max-pooling layers, 2 Normalized layers, 2 fully connected layers and 1 SoftMax layer. Each convolution layer consists of a convolution filter and a non-linear activation function called "ReLU". The pooling layers are used to perform the max-pooling function and the input size is fixed due to the presence of fully connected layers. The input size is mentioned at most of the places as 224x224x3 but due to some padding which happens it works out to be 227x227x3. Above all this AlexNet has over 60 million parameters and 650,000 neurons.

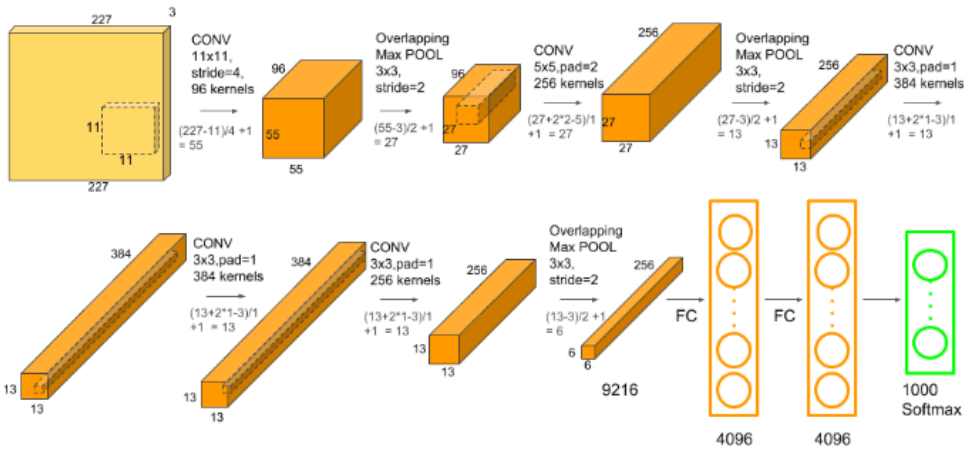

Figure 2: AlexNet Architecture

4.5.2 AlexNet Architecture Implementation

We did class balancing similar as defined in section 4.1 and took 3000 images for each class. The image dataset was preprocessed by resizing images to 227×227 pixels and normalizing pixel values to the range 0,1. AUTOTUNE from tensorflow was used to automatically adjust performance optimizations, ensuring efficient data pipeline management without manual tuning. Dataset caching and pre-fetching techniques were applied to enhance training efficiency. The model architecture consisted of five convolutional layers with ReLU activations, batch normalization, and max-pooling operations, followed by fully connected layers with dropout regularization to prevent overfitting. The model was trained using the Adam optimizer with a sparse categorical cross-entropy loss function over 50 epochs.

5 Results

5.1 RandomForestClassifier

5.1.1 Binary classification

In binary classification (e.g. whether the image is moderate demented or non demented) accuracy for:

- First Order Statistical Features was 82.21%
- Second Order Statistical Features was 82.33%
- After combining both first and second order statistical features accuracy was 87.25%.

Confusion Matrix:

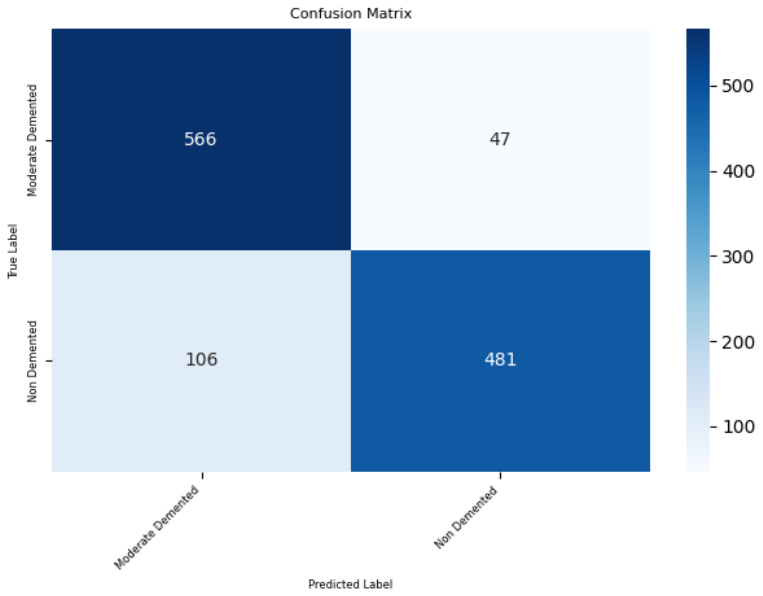

Figure 3: Confusion Matrix for Binary Classification.

Classification Report:

Class	Precision	Recall	F1-Score	Support
Moderate Demented	0.84	0.92	0.88	613
Non Demented	0.91	0.82	0.86	587
Accuracy			0.87	1200
Macro avg	0.88	0.87	0.87	1200
Weighted avg	0.88	0.87	0.87	1200

Table 1: ML Binary Classification Results

5.1.2 Multi-class classification

In multi-class classification(e.g either image is Mild Demented, Very Mild Demented, Moderate Demented or Non Demented) accuracy for:

- First Order Statistical Features was 47.16%
- Second Order Statistical Features was 48.66%
- After combining both first and second order statistical features accuracy was 56.83%.

Confusion Matrix:

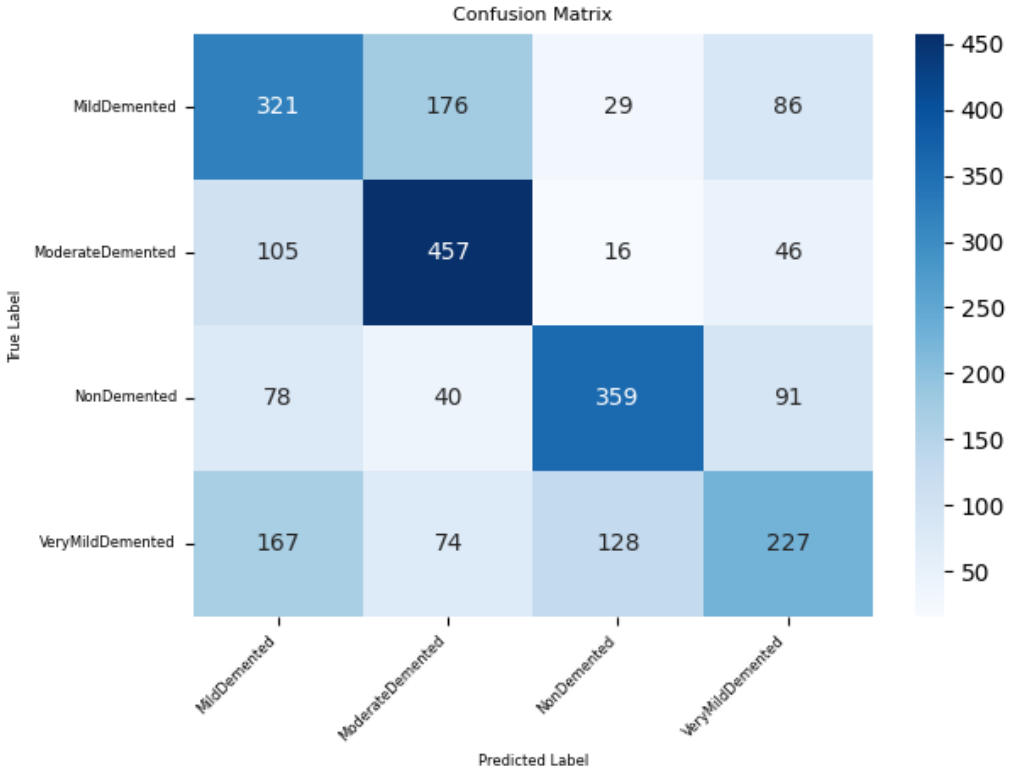

Figure 4: Confusion Matrix for multi-class Classification.

Classification Report:

Class	Precision	Recall	F1-Score	Support
Mild Demented	0.48	0.52	0.50	612
Moderate Demented	0.61	0.73	0.60	624
Non Demented	0.67	0.63	0.65	568
Very Mild Demented	0.50	0.38	0.43	596
Accuracy			0.57	2400
Macro avg	0.57	0.57	0.56	2400
Weighted avg	0.57	0.57	0.56	2400

Table 2: ML Multi-class Classification Results

5.2 AlexNet Architecture

Training and Validation accuracy

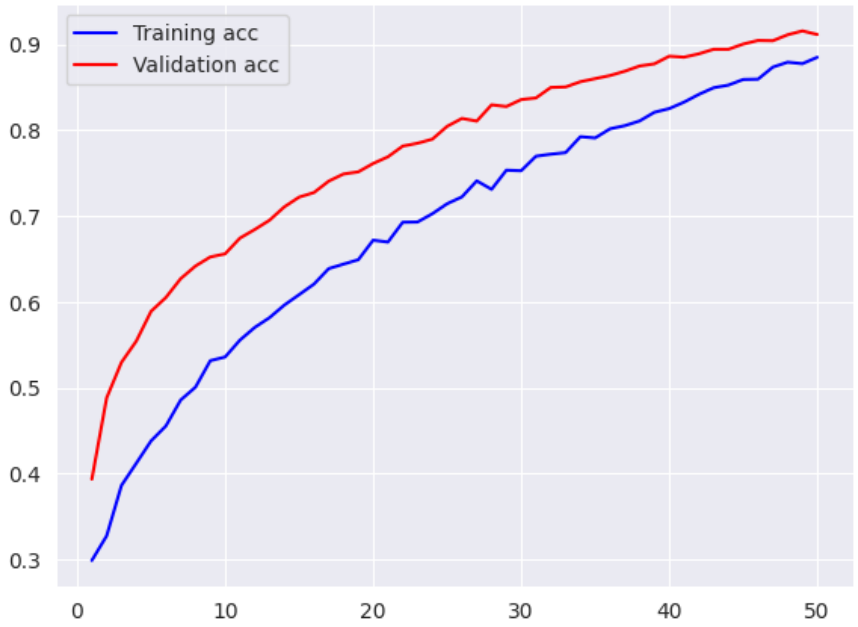

Figure 5: Training and Validation Accuracy

Both training and validation accuracy steadily increase over time. The validation accuracy surpasses the training accuracy, which could suggest that the model generalizes well without significant overfitting.

Training and Validation loss

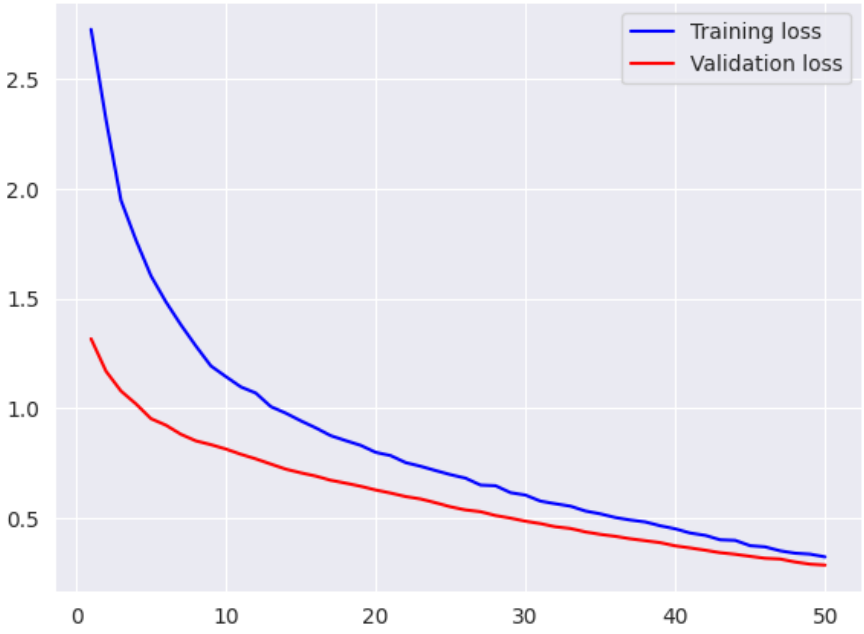

Figure 6: Training and Validation Loss

The training loss steadily decreases, indicating that the model is learning and fitting the training data well. The validation loss also decreases and closely follows the training loss curve, which further supports good generalization.

Accuracy

Accuracy for the classes of MildDemented, ModerateDemented, NonDemented and VeryMildDemented was **98.97%**, **100%**, **95.10%** and **94.80%** respectively with **overall accuracy of 97.23%**.

Confusion Matrix

Label 0,1,2,3 representing MildDemented, ModerateDemented, Non Demented and VeryMildDemented classes respectively.

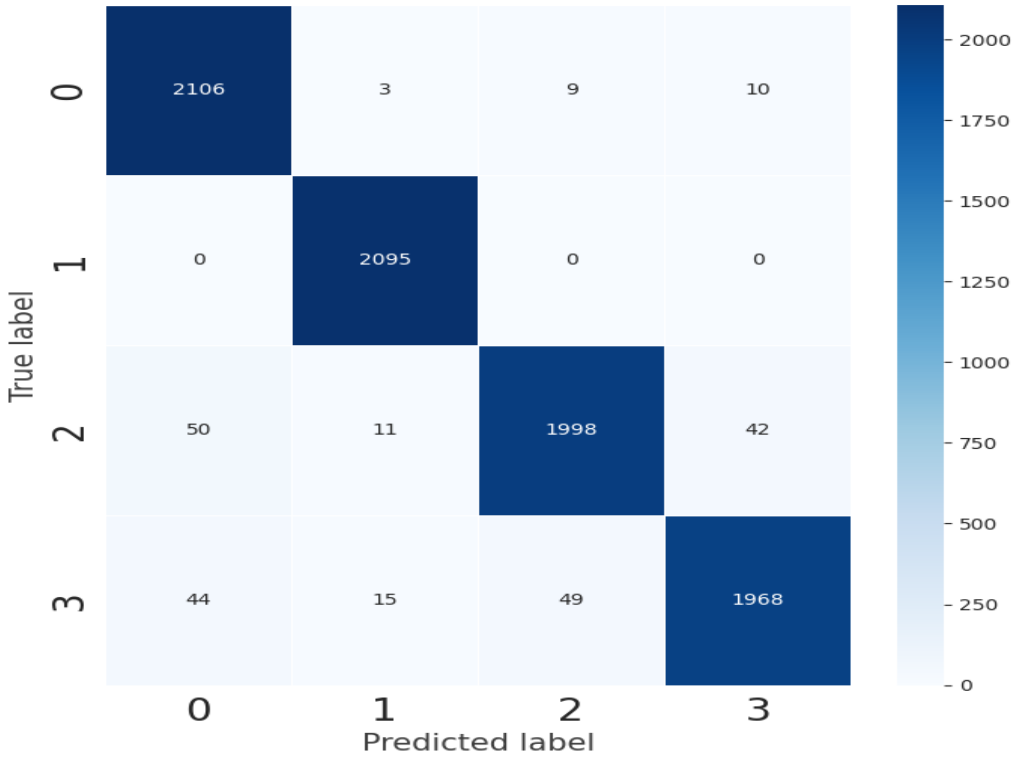

Figure 7: Consusion Matrix of AlexNet

Classification Report:

Class	Precision	Recall	F1-Score	Support
Mild Demented	0.96	0.99	0.97	2128
Moderate Demented	0.99	1.00	0.99	2095
Non Demented	0.97	0.95	0.96	2101
Very Mild Demented	0.97	0.95	0.96	2076
Accuracy			0.97	8400
Macro avg	0.97	0.97	0.97	8400
Weighted avg	0.97	0.97	0.97	8400

Table 3: AlexNet Results

6 Conclusion

For binary classification of Alzheimer disease detection, machine learning classifier gave good results with accuracy of 87.25% but when we applied it on multi-classes, it gave accuracy of only 56.83%.
After applying deep learning, accuracy for multi-class classification of Alzheimer detection improved drastically to 97.23% which clearly indicates that deep learning architectures are more accurate and reliable over traditional machine learning classifiers.

7 Future Directions

Take other clinically relevant data along with MRI scans e.g., patient demographics, and cognitive scores to improve classification accuracy with diverse and larger dataset containing more images in each class to enhance generalization and robustness. Consider more sophisticated data augmentation methods for additional performance improvements and using latest deep learning architectures like EfficientNet, ResNet, or Vision Transformers for multiclass classification. Consider deploying the model on resource-constrained devices and create real-time prediction systems for clinical environments.

References

[1] Shrikant Patro and VM Nisha. Early detection of alzheimer's disease using image processing. *International Journal of Engineering Research & Technology (IJERT)*, 8(05):2278–20181, 2019.

[2] Amir Ebrahimi, Suhuai Luo, and for the Alzheimer's Disease Neuroimaging Initiative. Convolutional neural networks for alzheimer's disease detection on mri images. *Journal of Medical Imaging*, 8(2):024503–024503, 2021.

[3] Fatemah H Alghamedy, Muhammad Shafiq, Lijuan Liu, Affan Yasin, Rehan Ali Khan, and Hussien Sobahi Mohammed. Machine learning-based multimodel computing for medical imaging for classification and detection of alzheimer disease. *Computational Intelligence and Neuroscience*, 2022(1):9211477, 2022.

[4] Eqtidar M Mohammed, Ahmed M Fakhrudeen, and Omar Younis Alani. Detection of alzheimer's disease using deep learning models: A systematic literature review. *Informatics in Medicine Unlocked*, 50:101551, 2024.

[5] Purushottam Kumar Pandey, Jyoti Pruthi, Saeed Alzahrani, Anshul Verma, and Benazeer Zohra. Enhancing healthcare recommendation: transfer learning in deep convolutional neural networks for alzheimer disease detection. *Frontiers in Medicine*, 11:1445325, 2024.

[6] R Anitha, S Jyothi, et al. A segmentation technique to detect the alzheimer's disease using image processing. In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, pages 3800–3801. IEEE, 2016.

[7] KJ Shanthi, DK Ravish, and M Sasikumar. Image segmentation an early detection to alzheimer's disease. In *2013 Annual IEEE India Conference (INDICON)*, pages 1–4. IEEE, 2013.