Multi-Class Alzheimer Classification using Hybrid Models

Arya Dhorajiya
Northeastern University
dhorajiya.a@northeastern.edu

Ayush Patel
Northeastern University
patel.ayushj@northeastern.edu

Spandan Maaheshwari
Northeastern University
maaheshwari.s@northeastern.edu

Abstract

Alzheimer's Disease is the most common form of dementia and the fifth-leading cause of death among people over the age of 65 leading to a neurological brain disorder causing progressive memory loss by damaging the brain cells and the ability to perform daily activities. In addition, based on official records, cases of death from Alzheimer's disease have increased significantly. Although there is currently no cure, it is possible to treat symptoms of dementia. The progression of dementia demands the urgency for the detection and classification of the tumors. Early diagnosis is paramount to the development and success of interventions, and neuroimaging represents one of the most promising areas for the early detection of AD. Machine learning methods on magnetic resonance imaging have been used in the diagnosis of Alzheimer's disease to accelerate the diagnosis process and assist radiologists. However, in conventional machine learning techniques, using handcrafted feature extraction methods on MRI images is complicated, requiring the involvement of an expert user. Therefore, implementing deep learning as an automatic feature extraction method could minimize the need for feature extraction and automate the process. In this report, we have implemented the evaluation metrics of manually trained CNN with pre-trained CNNs such as ResNet50, DenseNet121 and VGG16. Hence, proposing a system performing efficient detection and classification by using various Machine learning and Deep Learning Algorithms would be helpful to doctors all around the world.

1. Introduction

Over the past few decades, image processing and computer vision have helped medical personnel for the identification of various diseases through automated diagnostic processes. While the CDC attributed about 84,000 deaths in 2010 to Alzheimer's, the report estimated that number to be 503,400 among people 75 and older. That puts it in a close third place, behind heart disease and cancer. [2] Alzheimer's disease is the most common type of dementia. The disease is gradual, starting with mild memory loss and potentially progressing to the loss of communication and environmental awareness. The brain regions that are responsible for thought, memory, and language are affected by Alzheimer's disease. The severity of each stage of Alzheimer's dementia, such as Non-Demented, Very Mild, Mild, and Moderate, will be predicted. Since there is no treatment for this condition, it can be used to diagnose the symptoms early and to prevent the initial use of medications that can cause side effects.

A healthy brain has an appropriate amount of the chemical Acetylcholine. Reduced levels of the chemical itself as well as a loss of the nerve cells that respond to and use Acetylcholine are two problems in an Alzheimer's patient's brain. One of the treatments recommended for persons with mild to severe Alzheimer's disease is Memantine. According to studies, between 40 and 70% of those who take the medications have some sort of relief, with symptoms temporarily improving for 6 to 12 months. Cognitive stimulation treatment is an alternative diagnostic that is used. It is intended for those with mild to severe dementia. We want to create a model that can identify Alzheimer's at an early stage and anticipate the disease's progression because there are currently few diagnostic options for this condition.

In this project, we are classifying the type of dementia and comparing the result of this classification of dementia using various machine learning models such as ResNet50, DenseNet121, VGG16 as well as the model we made from scratch. In the literature, there are other algorithms and different modifications of the pre-trained networks that are used for image analysis, classification, and segmentation. Different approaches have been tested on other medical databases, both on MRI images of the brain and different parts of the human body. Many researchers have tried to solve the problem without pre-processing by using state-of-the-art models such as ResNet50, InceptionV3, MobileNet, etc.

The most common method of image pre-processing often involves image augmentation.[3] Image augmentation is the method of increasing the dataset by applying one or more transformations on the images (such as zooming, tilting over the axes and rotation. Most of the methods above use Otsu’s method to separate the background and use wavelet transform. The noise in MRI images is due to the fluctuations of the magnetic field in the coil. Magnetic Resonance Imaging is corrupted by Rician noise. Rician noise makes quantitative measurements difficult. Therefore, obtaining high-quality denoising images is the most important task of pre-processing.

2. Preliminaries

2.1 Dataset

The dataset is taken from Kaggle [1]. The dataset contains 6400 images of Alzheimer’s disease, which we have further divided into two sections, Training and Testing, where each of them has four sections, 'Non-Demeted', 'Very Mild Demented', 'Mild Demented' and 'Moderate Demented' 'Non-Demeted' contains MRI images of the brain that is not affected by Dementia, whereas other categories contain MRI images of the brain that is affected by various levels of Dementia. The training dataset has 5121 images, and the test dataset has 1279 images, all of which belong to one of the four classes before pre-processing.

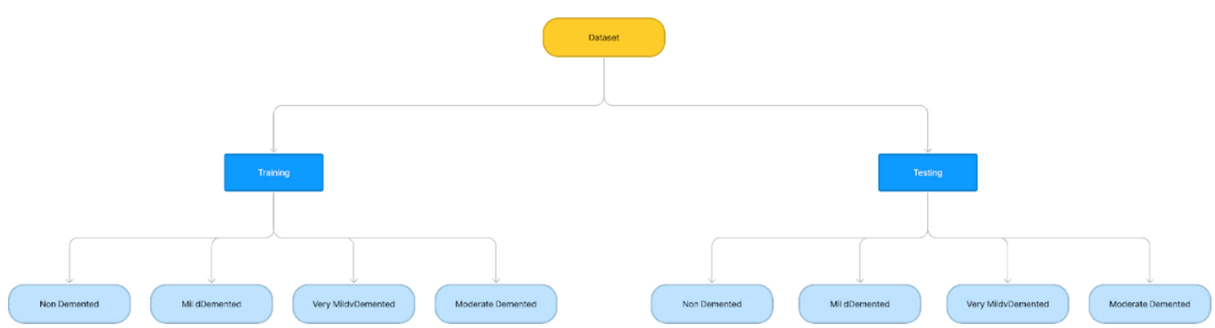
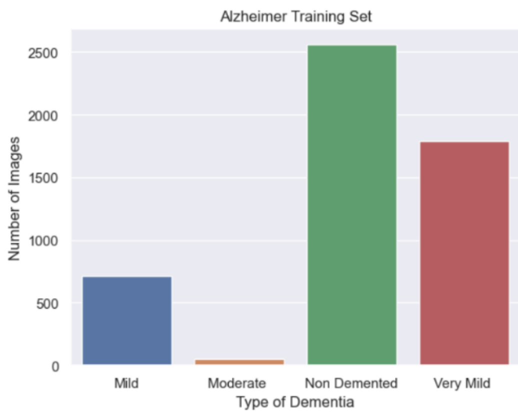


Figure 1: Structure of Dataset



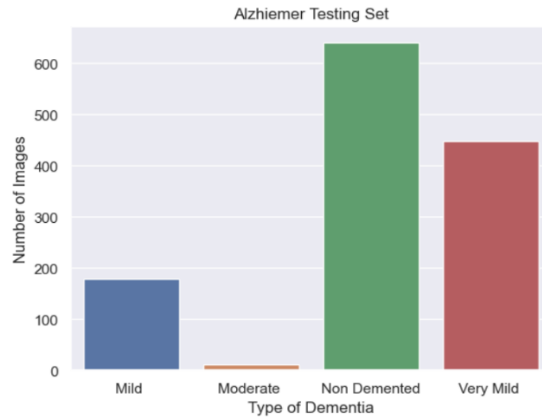


Figure 2: Number of Images in the Training Set; Figure 3: Number of Images in the Testing Set

2.2 Data Preprocessing

The images are usually stored in an RGB (Red Green Blue) format. In this format, the image is represented as a three-dimensional (or three-channel) array. One dimension is for channels (red, green, and blue colors) and two other dimensions are spatial dimensions. Thus, every pixel is encoded through three numbers. Each number is usually stored as an 8-bit unsigned integer type (0 to 255).

Rescaling is an operation that moves data from one numerical range to another by simple division using a predefined constant. In deep neural networks we want to restrict our input to the range from 0 to 1, due to possible overflow, optimization and stability issues.

Gray scaling:

Another type of transformation is gray scaling, which turns a color RGB image into images with only shades of gray representing colors. This pair of transformations can throw away noisy pixels and detect shapes in the picture.

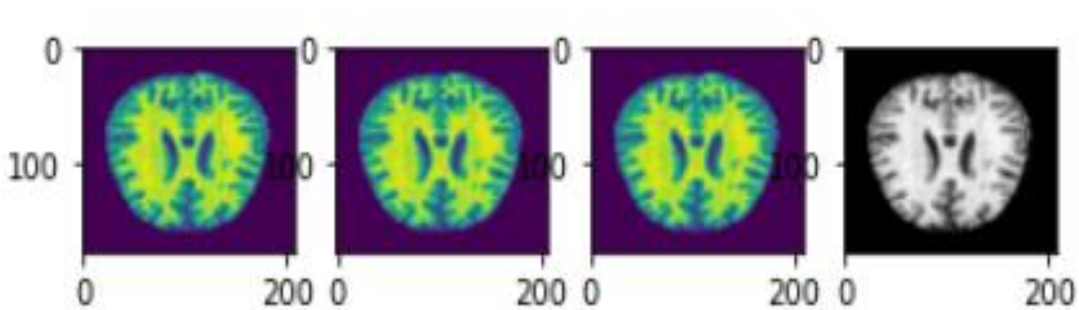


Figure 4: Gray Scaling and Rescaling

Rotation:

This transformation rotates the image in a certain direction clockwise or counterclockwise. The parameter that allows the rotations is called rotation range. It specifies the range of rotations in degrees from which the angle is chosen uniformly to do a rotation.

Shearing:

Shearing displaces each point in the vertical direction by an amount proportional to its distance from an edge of the image. In general, the direction does not have to be vertical and can be arbitrary. The parameter that controls the displacement rate is called shear range and corresponds to the deviation angle (in radians) between a horizontal line in the original picture and in the transformed image.

Zoom:

This transformation zooms the initial image in or out. The zoom range parameter controls the zooming factor and is chosen from the range [0.5, 1.5].

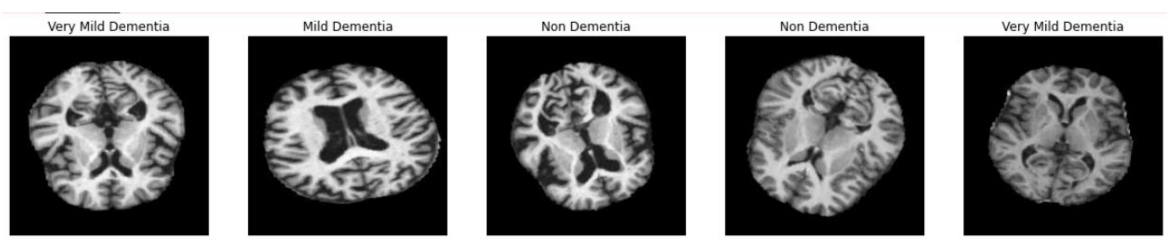


Figure 5: Rotation, Shearing and Zoom

2.3 Models used

We implemented a CNN model from scratch and used it to compare with other state-of-the-art models based on how they perform when trained on the given MRI Images. This aided us in determining the best approach to the problem and selecting the best model for it. The following models were used for evaluation:

1. ResNet50
2. VGG16
3. DenseNet121
4. Convolutional Neural Network

2.3.1 ResNet50

ResNet50 is a convolutional neural network that is 50 layers deep. It's a type of deep-learning model that can be used for image classification and other tasks. ResNet50 consists of several types of layers, including convolutional layers, pooling layers, and fully connected layers. The convolutional layers are used to extract features from the input data, and the pooling layers are used to reduce the spatial dimensions of the data. The fully connected layers are used to make predictions based on the extracted features. In addition to these standard layers, ResNet50 also uses residual blocks, which are groups of layers that are designed to learn the residuals between the output of the current layer and the expected output. This allows the network to learn more complex patterns and helps to improve its performance.

The 50-layer ResNet architecture includes the following elements, as shown in the table below:

- **A 7×7 kernel convolution** alongside 64 other kernels with a 2-sized stride.
- **A max pooling layer** with a 2-sized stride.
- **9 more layers** 3×3,64 kernel convolution, another with 1×1,64 kernels, and a third with 1×1,256 kernels. These 3 layers are repeated 3 times.
- **12 more layers** with 1×1,128 kernels, 3×3,128 kernels, and 1×1,512 kernels, iterated 4 times
- **18 more layers** with 1×1,256 cores, and 2 cores 3×3,256 and 1×1,1024, iterated 6 times
- **9 more layers** with 1×1,512 cores, 3×3,512 cores, and 1×1,2048 cores iterated 3 times
- **Average pooling**, followed by a fully connected layer with 1000 nodes, using the SoftMax activation function.

2.3.2 VGG16

VGG16 is a convolutional neural network model trained on the ImageNet dataset. It is known for its simplicity and good performance in image classification tasks. VGG-16 is one of the biggest networks that has 138 million parameters. The VGG16 model has 16 layers, including 13 convolutional layers and 3 fully connected layers. The convolutional layers are used to extract features from the input images, while the fully connected layers are used for classification. VGG16 is widely used as a benchmark for comparing the performance of different image classification algorithms. The 2nd, 4th, 7th, 10th, and 13th convolution layers are followed by 2×2 Max-pooling layers. VGG-16 accepts colored images with dimensions 224×224 and generates output according to the number of classes.

VGG-16, a deep CNN architecture with 16 layers:

- Total 13 convolution layers with **Kernel size 3×3**, each having **ReLU** as an Activation Function.
- The First Convolutional Layer contains **64 filters**, Second Convolutional Layer has **128 filters**, the Third Convolutional Layer has **256 filters**, and while Fourth and Fifth Convolutional Layer has **512 filters**.
- 5 layers of Max Pooling with **Filter size 2x2** and **stride size 2x2**.
- Three fully connected layers, where Two layers contain **4096 channels**, while the Third layer contains **1000 channels**.
- A **SoftMax Activation Layer** for classification purposes.

2.3.3 DenseNet121

DenseNet121 is a Convolutional Neural Network, which is known for solving the infamous 'Vanishing Gradient' problem by modifying the standard CNN architecture and simplifying the connectivity pattern between layers. Despite having a lot of layers, DenseNet121 contains only around 8 million parameters as the given model mostly uses Convolutional Layers, while using only one Fully Connected Layer. In a DenseNet architecture, each layer is connected directly with every other layer, instead of using a skip connection method, which is used in Resnet.

DenseNet121, a deep CNN architecture with 124 layers:

- Basic convolution layer with **64 filters** of **size 7X7** and a **stride of 2.**
- Basic pooling layer with **3x3 max pooling** and a **stride of 2.**
- **6 Convolutional Layers** with two different Filter sizes (**1x1 and 3x3**).
- Transition layer 1 (1 Conv Layer + 1 AvgPool Layer)
- **12 Convolutional Layers** with two different Filter sizes (**1x1 and 3x3**).
- Transition layer 2 (1 Conv Layer + 1 AvgPool Layer)
- **24 Convolutional Layers** with two different Filter sizes (**1x1 and 3x3**).
- Transition layer 3 (1 Conv Layer + 1 AvgPool Layer)
- **16 Convolutional Layers** with two different Filter sizes (**1x1 and 3x3**).
- Global Average Pooling layer- accepts all the feature maps of the network to perform classification.
- Fully connected Layer containing **1000 channels**, combined with **SoftMax activation function** and **Cross-Entropy Loss function**.

2.3.4 Convolutional Neural Network

The group has designed a Convolutional Neural Network model from scratch, consisting of 10 Convolutional Layers, 5 Max Pooling Layers, 5 Dropout Layers, 1 Flatten Layer, and 3 Dense Layers.

Model from Scratch, a CNN architecture with 24 Layers:

- 2 Convolutional Layers contain **16 filters** of **size 3x3** each.
- 1 Max Pooling Layer of **size 2x2**.
- 2 Convolutional Layers contain **32 filters** of size 3x3 each.
- 1 Max Pooling Layer of **size 2x2**.
- 2 Convolutional Layers contain **64 filters** of size 3x3 each.

- 1 Max Pooling Layer of **size 2x2**.
- 2 Convolutional Layers contain **128 filters** of size 3x3 each.
- 1 Max Pooling Layer of **size 2x2**.
- 1 Dropout Layer with **Drop rate 0.2**.
- 2 Convolutional Layers contain **256 filters** of size 3x3 each.
- 1 Max Pooling Layer of **size 2x2**.
- 1 Dropout Layer with **drop rate 0.2**.
- 1 Flatten Layer with **7680 channels**.
- 3 Dense Block containing **dense units (512, 128 and 64)** and **dropout rate (0.7, 0.5 and 0.3).**
- Dense Layer with **SoftMax** Activation Function.

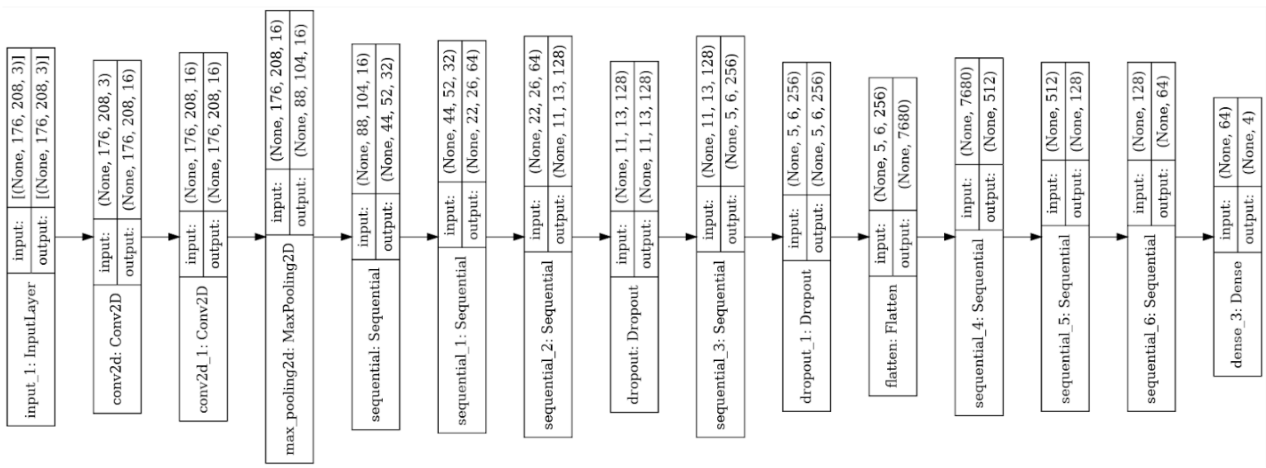


Figure 6: Architecture for the CNN model

We have used Adam as our optimizer and Categorical Cross Entropy as our Loss Function. The results of the Adam optimizer are generally better than every other optimization algorithm, have faster computation time, and require fewer parameters for tuning. Because of all that, Adam is recommended as the default optimizer for most of the applications. After describing the models used in our project, we will discuss the results that we achieved by comparing multiple performance metrics relevant to our problem set between our basic CNN model and other State-Of-The-Art models, i.e., VGG16, ResNet50 and DenseNet121.

3. Results

Figure 7 depicts the prediction of the CNN model which we made from scratch. As you can see the model has difficulty in predicting between Non-Demented and Very Mild Demented MRI images since the Posterior cortical atrophy (PCA) is almost the same. In layman’s terms the gaps or holes in the brain as seen below.

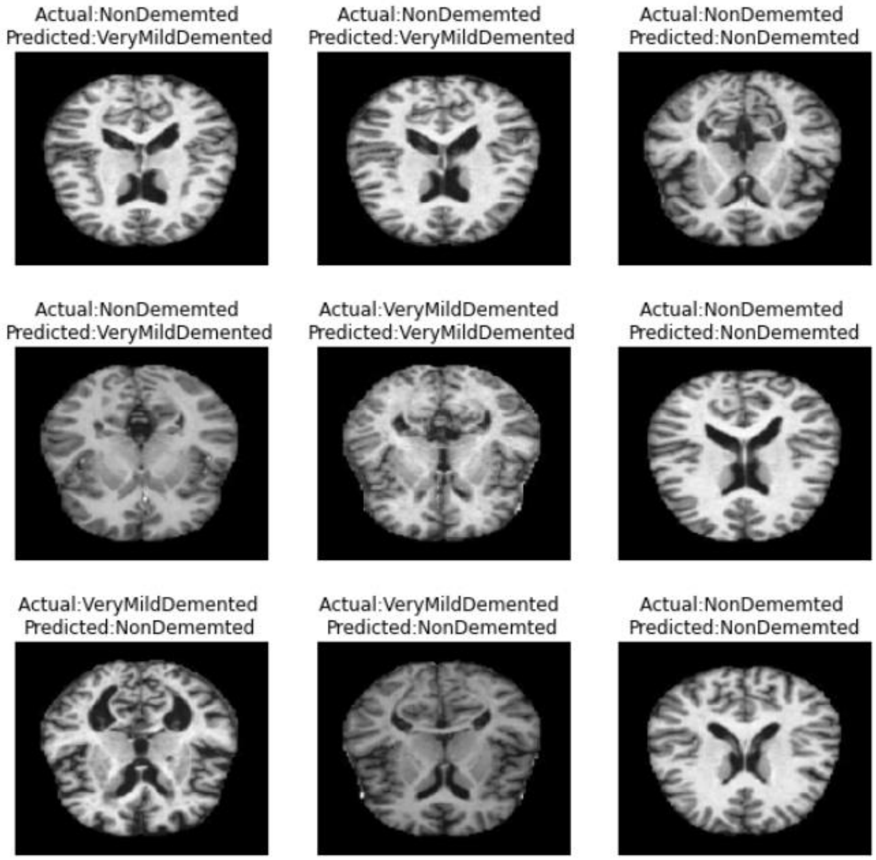


Figure 7: Prediction of the CNN model

Table 1 represents the results of the state-of-the-art model and the Convolution Neural Network which we made from scratch. The advantage of this approach is that it produces much better results and drastically reduces noise because of the pre-processing done on the data.

Model Type	Training			Testing		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Resnet50	0.67	0.33	0.45	0.60	0.38	0.46
DenseNet121	0.80	0.75	0.78	0.65	0.62	0.63
VGG16	0.75	0.68	0.71	0.67	0.62	0.64
Convolutional Neural Network	0.95	0.95	0.95	0.86	0.86	0.86

Table 1: Classification Report of all the models

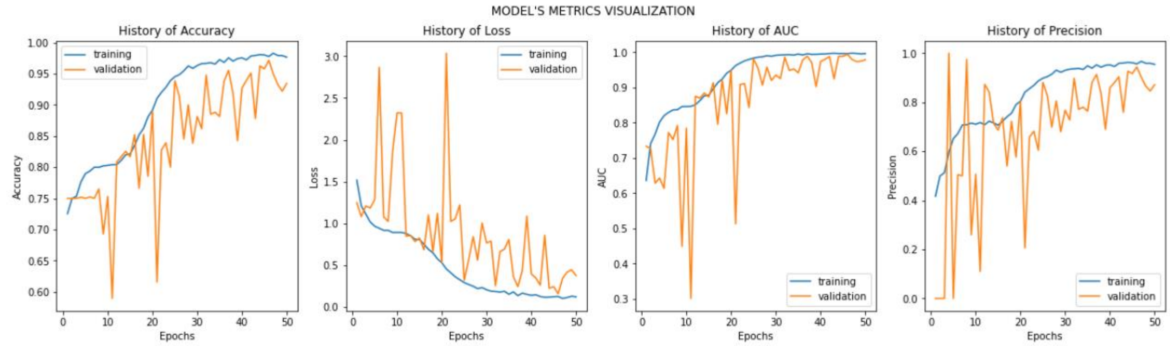


Figure 8: History of Accuracy, Loss, AUC and Precision for the CNN Model

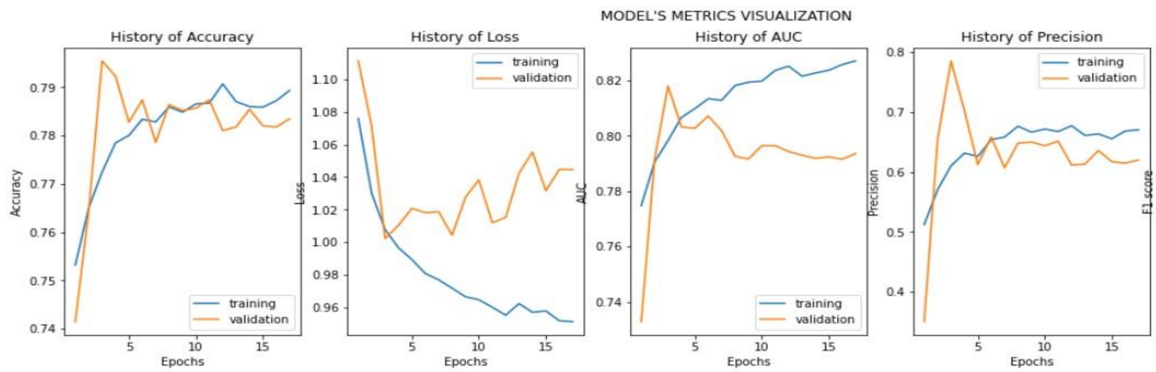


Figure 9: History of Accuracy, Loss, AUC and Precision for ResNet50

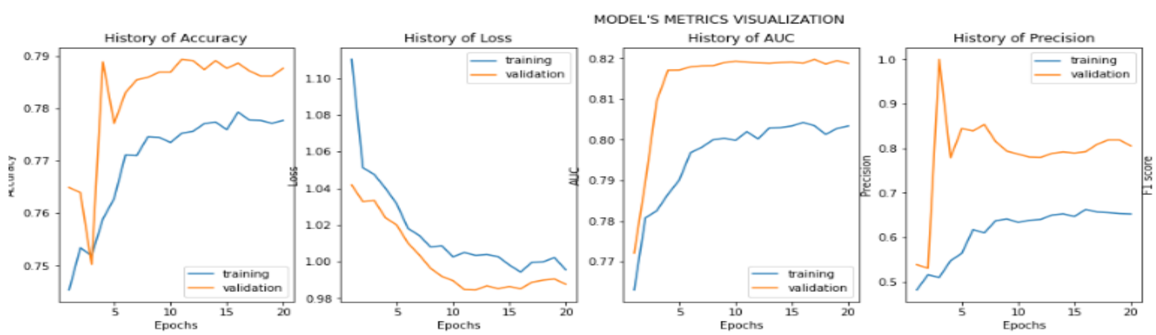


Figure 10: History of Accuracy, Loss, AUC and Precision for DenseNet121

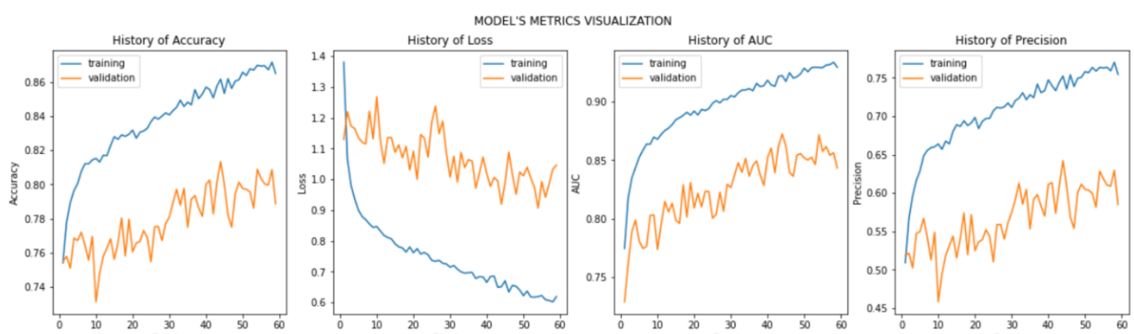


Figure 11: History of Accuracy, Loss, AUC and Precision for VGG16

4. Discussion

The highest accuracy obtained on the CNN model for training is 97.48 %, validation is 93.45 % and the testing accuracy was 93.05 %. Although, the model had difficulty predicting the Non-Demented and Very Mild Demented MRI images. As seen in Table 1, the CNN model outperforms other state-of-the-art models by achieving a 95% Precision and Recall rate on the Training Set, and an 86% Precision and Recall Rate on the Testing Set. We are using the F1 score as a metric to compare the models over accuracy as it is:

- More robust towards imbalanced data
- More relevant when False Positives and False Negatives are considered crucial in a problem set.

Our dataset is imbalanced for performing detection. False positives and false negatives also carry more significant risks in the medical field. Therefore, it makes sense for the F1 score to be one of our primary metrics. We now evaluate the plots for our CNN model and a state-of-the-model with good performance and discuss their findings. The architecture of our CNN model resonates the architecture of InceptionV3 model, we have also added dropout layers to further enhance the model thus leading to better results.

In the given above plots, we can see that our model, despite containing imbalanced data and noisy data, fits the data achieving good accuracy and low loss value. While DenseNet121 also fits the data better, it doesn’t achieve high accuracy, compared to our model, and it also overfits the data after some epochs, as we can infer from the loss plot of DenseNet121. DenseNet121 also achieves a very low Precision Rate of around 65% in the Validation Dataset, while our model achieves a very good Precision Rate of 86% in the Validation Dataset.

In the VGG16 model, the number of trainable parameters is above 100 million, so we need to increase the number of epochs to avoid convergence and overfitting. This is the downside for VGG16 as compared to DenseNet121. Despite having 100 million parameters, VGG16 still gives a bad performance as the images are very similar and deep features need to be extracted in order to get a better performance on the given images. In our CNN model, we added a limited number of CNN layers and Dense Layers, therefore keeping the number of parameters the same, while getting a better performance than VGG16.

In the ResNet50 model, we have found that in this case, the deeper the number of layers of the ResNet network that is used does not increase the model’s performance and even lowers the accuracy obtained. This is because the model takes fewer trainable parameters as compared to the other models.

5. Conclusion

From the above-stated results, we infer that we secured the highest precision and accuracy from our CNN model. CNN outperforms in the multiclass classification of Alzheimer’s. We observed a decent accuracy & F-1 score from CNN as compared to other state-of-the-art models. The proposed work was able to give a training accuracy of 97.48%, validation accuracy of 93.45%, and testing accuracy of 93.05% with very small misclassifications on non-demented and very mild demented.

We would like to know what would happen if the images passed to pre-trained models were pre-processed the same way as that for CNN. We would like to experiment and find out. In the future, we plan to use the extracted features from our CNN model and for prediction, integrate them into powerful machine learning models such as SVM and Random Forest. Also, we would integrate our deep learning model that discriminates Alzheimer's from other brain illnesses and brain abnormalities that are associated with the risk of conversion to Alzheimer's from mild cognitive impairment and other behavioral outcomes.

6. References

1. https://www.kaggle.com/datasets/yasserhessein/dataset-alzheimer
2. Brain MRI Analysis for Alzheimer's Disease Diagnosis Using CNN-Based Feature Extraction and Machine Learning.
3. New study ranks Alzheimer's as third leading cause of death, after heart disease and cancer - The Washington Post.
4. Sultan, H.H., Salem, N.M., & Al-Atabany, W. (2019). Multi-Classification of Brain Tumor Images Using Deep Neural Network. IEEE Access, 7, 69215-69225.
5. H. Fuse, K. Oishi, N. Maikusa, T. Fukami and J. A. D. N. Initiative, "Detection of Alzheimer's Disease with Shape Analysis of MRI Images," 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS), 2018, pp. 1031-1034, doi: 10.1109/SCIS-ISIS.2018.00171.
6. Y. Yamashita et al., "Computerized classification of patients with alzheimer's disease based on arterial spin-labeled perfusion MRI," 2010 World Automation Congress, 2010, pp. 1-4.
7. H. Li and Y. Fan, "Early Prediction Of Alzheimer's Disease Dementia Based On Baseline Hippocampal MRI and 1-Year Follow-Up Cognitive Measures Using Deep Recurrent Neural Networks," 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp. 368-371, doi: 10.1109/ISBI.2019.8759397.
8. C. Zhang et al., "ResNet or DenseNet? Introducing Dense Shortcuts to ResNet," 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 2021, pp. 3549-3558, doi: 10.1109/WACV48630.2021.00359.
9. M. Mahyoub, M. Randles, T. Baker and P. Yang, "Effective Use of Data Science Toward Early Prediction of Alzheimer's Disease," 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS), 2018, pp. 1455-1461, doi: 10.1109/HPCC/SmartCity/DSS.2018.00240.
10. D. Manzak, G. Çetinel and A. Manzak, "Automated Classification of Alzheimer's Disease using Deep Neural Network (DNN) by Random Forest Feature Elimination," 2019 14th International Conference on Computer Science & Education (ICCSE), 2019, pp. 1050-1053, doi: 10.1109/ICCSE.2019.8845325.