Classification of Alzheimer's Disease based on MRI Image Processing using Convolutional Neural Network (CNN) with Transfer Learning

Boram Lee, UID : 121301338
Principles of Data Science
December 10, 2024

Abstract

This project explores the most efficient methodologies for training relatively structured datasets by applying and comparing techniques such as transfer learning, data augmentation, and data balancing across varying amounts of data. The dataset used in this study consists of MRI brain images categorized into four classes based on Alzheimer's severity. The primary model utilized was a CNN, with additional tuning through transfer learning, data augmentation, and data balancing methods.

The highest performance was achieved by the most basic baseline CNN model, which demonstrated an accuracy of 99.99%. The findings suggest that for structured image data like MRI scans, simple CNN models are significantly more efficient and effective than transfer learning models, which are better suited for diverse and large-scale datasets. This project highlights the potential of basic CNN models in achieving high performance for medical imaging tasks.

1. Introduction

Alzheimer's Disease (AD) is a progressive and irreversible neurodegenerative disorder that primarily affects memory, thinking, and behavior. While there is currently no cure for AD, early detection and diagnosis are crucial for managing symptoms and improving the quality of life for patients. With the global prevalence of AD increasing due to aging populations, the development of advanced tools and models for early diagnosis has become a critical area of research.

This report focuses on detecting and classifying four stages of AD (Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented) using MRI image data and Convolutional Neural Networks (CNN). Given the relatively structured nature of MRI image data compared to other general image data, there is an expectation of achieving high classification performance, despite the subtle challenges posed by medical imaging data.

The classification task in this project was primarily performed using a basic CNN, enhanced with techniques such as data augmentation, data upsampling to address class imbalances, and transfer learning. The classification results were evaluated based on accuracy, recall, and precision. Interestingly, the most basic CNN model achieved the highest accuracy of 99.99% for classifying the four stages, outperforming approaches that incorporated additional techniques.

This project highlights the efficacy of using a basic CNN for efficient and accurate classification of AD stages. However, the findings also emphasize that distinguishing between Non-Demented, Very Mild Demented, and Mild Demented cases remains the most challenging aspect of the task.

2. **Materials and Methods**

2.1. Dataset

The dataset used in this project was sourced from Kaggle, where it is freely available and originates from the Open Access Series of Imaging Studies (OASIS). It provides MRI brain images categorized into four classes based on Alzheimer's severity: Non-Demented (ND), Very Mild Demented (VMD), Mild Demented (MLD), and Moderate Demented (MDD). The MRI images of 461 patients were sliced along the z-axis into 256 sections, with slices ranging from 100 to 160 selected for each patient. The original image dimensions are 496 pixels in width, 248 pixels in height, and three channels. The dataset comprises 67,222 ND, 13,725 VMD, 5,002 MLD, and 488 MDD images.

For this project, an appropriate number of training and test datasets were selected based on various training methods. The basic dataset was derived from the original OASIS dataset, with the number of MDD images multiplied across all categories. Additionally, 10% of either the original or basic dataset was set aside for testing purposes.

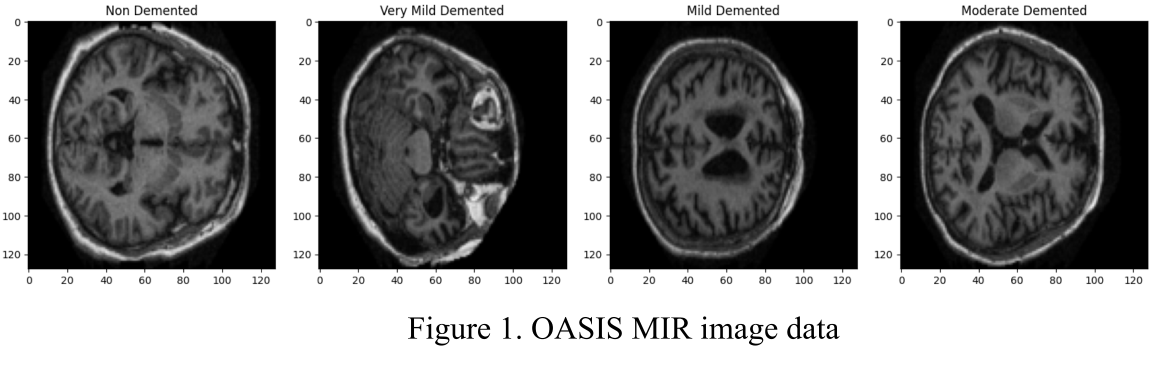

Figure 1. OASIS MIR image data

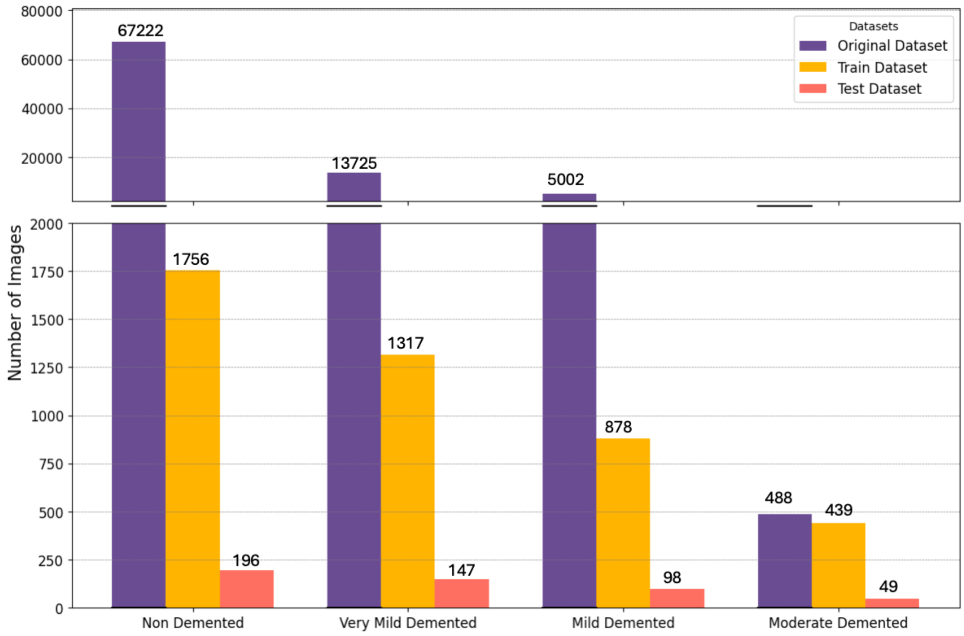

Figure 2. The number of datasets

2.2. Data Preprocessing

2.2.1. Basic CNN

The entire original dataset was resized to 128 x 128 pixels while retaining three channels. Although MRI image data is inherently grayscale and theoretically requires only one channel, the original OASIS dataset format with three channels was preserved. The second and third channels were created by duplicating the first channel.

2.2.2. Data Augmentation

To evaluate the impact of data augmentation, this project incorporated a data augmentation step. The dataset was augmented with various transformations, as detailed in Table 1. This process was implemented to mitigate overfitting issues and enhance model generalization.

Table 1. Data Augmentation Transformations

No.	Data augmentation
1	Horizontally flip with 30% probability
2	Vertically flip with 30% probability
3	Blur the images using a 3x3 Gaussian kernel
4	Randomly rotate the images between -45 and 45 degrees
5	Convert image to a PyTorch tensor (float32 automatically)
6	Normalize for 3-channel RGB (ImageNet mean)
7	Standard deviation for ImageNet

2.2.3. Oversampling

To address data imbalance issues, I utilized oversampling techniques through the Synthetic Minority Over-sampling Technique (SMOTE). While the original dataset contains a substantial amount of data for training the model, the MMD class has a significantly smaller number of samples compared to the other classes. Furthermore, even when I created new datasets based on the MMD class's size, the imbalance between the MMD class and the other classes remained a critical issue.

SMOTE was applied with the strategy of balancing all classes relative to the ND class, which has the largest number of samples. This approach aimed to mitigate the impact of data imbalance and improve classification performance.

2.3. Deep Learning Models

A basic CNN with a simple structure implemented in PyTorch was primarily used for this project. This basic CNN consists of two convolutional blocks: the first block utilizes 32 filters to extract features, and the second block uses 64 filters. Additionally, two fully connected (FC) layers were added, and batch normalization was applied in the initial forward pass to improve training stability.

To explore the effect of transfer learning on MRI image data, ResNet34, ResNet152, and AlexNet models pre-trained on ImageNet were also employed. The pre-trained parameters of these models remained unchanged, except for the final FC layer, which was fine-tuned using the OASIS MRI dataset.

ResNet is a deep CNN that utilizes shortcut connections between layers to enable residual learning. ResNet34 and ResNet152 refer to architectures with 34 and 152 layers, respectively, each consisting of multiple convolutional layers and shortcut connections. In contrast, AlexNet has a simpler structure with eight layers, including five convolutional layers, two fully connected layers, and a softmax layer. Despite having fewer layers compared to ResNet, AlexNet introduced key innovations such as ReLU activation and dropout regularization techniques, which have been widely adopted in modern neural networks.

2.4. Training Methodology

The experiments focused on classifying four different stages of Alzheimer's Disease (AD). GPU acceleration was employed using the Metal Performance Shaders (MPS) framework on an Apple M1 Max with 32 GB of memory. The training parameters were fixed, with the number of epochs set to 30, the learning rate set to 0.001, and the batch size set to 32. In total, 10 different case combinations were tested, and the performance of each case is detailed in Table 2.

Table 2. Accuracy of each model combination

No.	Model	Dataset			
		Amount (total/classes)	Augmentation	Upsampling	Accuracy (Training / Test)
1	baselineCNN	86437 / [67222, 13725, 5002, 488]	-	-	0.9776 / 0.9709 (epoch : 30) 1 / 0.9999 (epoch : 15)
2	baselineCNN	4390 / [1756, 1317, 878, 439]	-	-	1 / 0.9633
3	baselineCNN	4390 / [1756, 1317, 878, 439]	Transformation	-	1 / 0.5449
4	baselineCNN	7024 / [1756, 1756, 1756, 1756]	-	SMOTE	1 / 0.9612
5	ResNet34	86437 / [67222, 13725, 5002, 488]	-	-	0.81 / 0.65
6	ResNet34	7024 / [1756, 1756, 1756, 1756]	-	SMOTE	0.7893 / 0.6735
7	ResNet152	4390 / [1756, 1317, 878, 439]	-	-	0.8467 / 0.7245
8	ResNet152	7024 / [1756, 1756, 1756, 1756]	-	SMOTE	0.8818 / 0.7163
9	AlexNet	4390 / [1756, 1317, 878, 439]	-	-	0.6695 / 0.6102
10	AlexNet	7024 / [1756, 1756, 1756, 1756]	-	SMOTE	0.8438 / 0.6857

3. Results

The experiments were conducted using various pre-trained models, leading to several key observations. Notably, the baseline CNN trained on the largest amount of the original dataset achieved the highest accuracy (Table 2: No. 1). Contrary to my expectation that transfer learning would outperform other methods, the baseline CNN with the largest dataset delivered the best performance, even surpassing ResNet34 trained on the same dataset size (Table 2: No. 5).

In general, the transfer learning models did not demonstrate significant performance improvements in these experiments. However, when the training dataset was upsampled using SMOTE, the transfer learning models performed better than when trained on unbalanced datasets (Table 2: No. 5–10). Additionally, all baseline CNN models achieved accuracy above 0.95, except for the case with data augmentation. The data augmentation experiment resulted in the lowest test set performance, even though the training set accuracy reached 100% (Table 2: No. 3).

Interestingly, the baseline CNN trained for 15 epochs on the original dataset outperformed the model trained for 30 epochs (Table 2: No. 1). This phenomenon is illustrated in the left graph of Figure 3, where the loss curve exhibits intermittent spiking around the 17th, 24th, and 30th epochs. This can likely be attributed to batch-level differences in the data.

Finally, the confusion matrices for low-performance cases indicate that distinguishing between ND, VMD, and MLD is particularly challenging. This suggests that the difficulty in classification arises from the pathological similarity among these three classes (Figure 4).

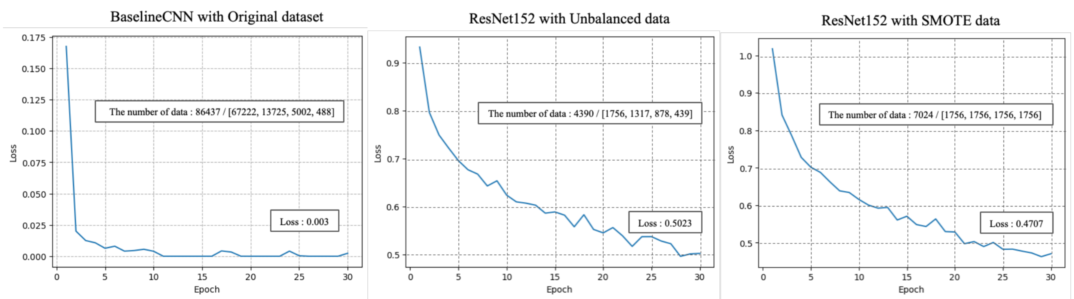

Figure 3. Loss graph based on different models and different technique

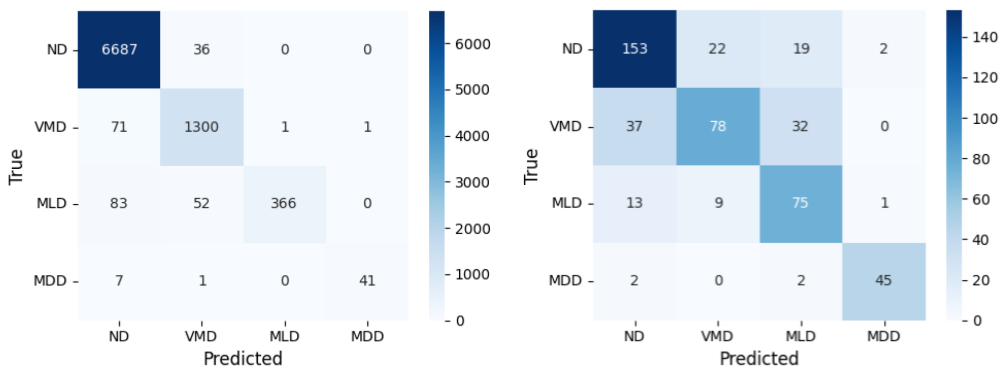

Figure 4. left : the confusion matrix of No. 8, right : the confusion matrix of No.1

4. Conclusion

Through this project, we confirmed the high potential of the healthcare and medical machine learning field, particularly in using relatively structured data, such as MRI image data, compared to everyday images. Specifically, we demonstrated that image analysis can be applied to one of the most challenging areas in modern medicine—early diagnosis of Alzheimer's disease.

As mentioned earlier, highly structured and formatted data, such as MRI and CT images, can achieve excellent performance even with a simple CNN without relying on techniques like transfer learning, data augmentation, or upsampling. Since the data is already highly structured, excessive data preprocessing to prevent data imbalance or overfitting can actually degrade model performance. Thus, increasing the amount of data and opting for relatively lightweight models is a more efficient approach.

While we achieved high performance using a large dataset, distinguishing between early-stage dementia classes like ND, VMD, and MLD becomes less accurate as the dataset size decreases. To address this limitation, the next step in this project will aim to improve performance by training the model not only with image data but also with various biomarker data.

The whole jupyter notebook code is on the link : https://drive.google.com/file/d/1iUQYoUE2nveXkJUMTh4Ia0O1GKlDYyOw/view?usp=sharing