Diagnosis of Alzheimer's Disease via Convolutional Neural Networks

Edward Zhou
University of Texas at Austin
Austin, TX 78705
EdwardZhou538@gmail.com

Abstract

The issue of detecting Alzheimer's Disease (AD) within its early stages has garnered much research in the world of medicine. Recently, deep learning frameworks utilizing magnetic resonance imaging (MRI) scans of the brain have shown to have increasingly more favorable results in detecting signs of AD within patients. In this paper, I introduce a Convolutional Neural Network for the purposes of using MRI scans of the brain to detect Alzheimer's disease within patients. Although the model's results are seemingly, its overall accuracy seems to be boosted by the unrealistic nature of the dataset in providing sufficient levels of variability. As such, I performed two types of data augmentation on the initial data in order to introduce more variation and "increase" the amount of data being used. Additionally, I also tested this method on the resnet18 model in order to analyze the difference between my model and a pretrained one.

1 Introduction

Alzheimer's Disease (AD) is a neurological disease most commonly found within the elderly that is characterized by symptoms of Dementia and reduced brain function. It is an incredibly common occurrence, affecting 5.8 million people across America, a number that is projected to over 14 million by 2060 according to the CDC[1]. Not only this, but it is also one of the deadliest conditions within the US, being the 6th leading cause of death among US adults, and has no cure. However, people with AD usually undergo a preclinical stage known as Mild Cognitive Impairment (MCI), which can be treated and potentially reversed before patients devolve into more severe levels of Dementia. As such, it is critical for patients to be able to identify the potential severity of the disease as soon as possible, in order to allow for more optimal treatment measures.

Due to recent advances in neuroimaging technology as well as the increasing prevalence of computer vision models in tackling real-world problems, neural networks have been used to increasingly higher success in aiding Doctors in AD diagnosis. Although models have used Computational Tomography (CT), Positron Emission Tomography (PET), and Medical Resonance Imaging (MRI) as image input for classification purposes, the model I am using will primarily be focused on MRI data due to their accuracy in constructing brain tissue imagery. Thus, the overall paper will focus on the following:

1. A Convolutional Neural Network is introduced that will aim to detect the severity of Alzheimer's Disease using Brain MRI scans.
2. Transfer learning from ResNet is used to analyze the difference between ResNet and CNN classifications.
3. Data augmentation is done to introduce variance and test the models' ability to recognize relative features over absolute ones.

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

2 Related Works

2.1 Image Classification

Image Classification refers to the use of computational models to classify an image into a distinct category. This is typically done using Convolutional Neural Networks (CNN) that can convolve across images to extract relevant features for categorization.

2.2 Transfer Learning

Transfer Learning refers to the application of pre-trained networks to solve a new and related problem. This has shown to be particularly useful in saving computational resources and improving efficiency of training new models, especially when the given dataset is not incredibly large. In image classification problems, architectures such as ResNet and VGG16 are commonly used for transfer learning purposes, but several state of the art models, such as classifiers involving visual transformers, could also be used for image classification purposes.

2.3 Data Augmentation

Data Augmentation is a technique to increase the amount of training data via making slight to moderate transformations to the input dataset, such as cropping, rotation, and gaussian blur. Overall, this technique is effective at reducing overfitting as it forces the model to learn the features of the randomly transformed data, rather than simply memorizing it.

3 Dataset

The dataset used for this model was the Kaggle Alzheimer's Classification Dataset (KACD)[4], shown in Figure 1. The KACD is a dataset comprised of 6400 brain MRI Scans. Each MRI scan is a 208x176 grayscale jpg image of an individual's brain, which is split into 4 different classification classes, No Dementia, Very Mild Dementia, Mild Dementia, and Moderate Dementia, depending on the individual's level of disease progression. Additionally, the initial dataset was already pre-split into a training and testing set; however, when training on these preexisting sets, the model would perform extremely well on both the training and validation set but do significantly worse on the test set. As such, it appeared that the test set was created on a different distribution to the training set, and I instead pooled both sets together and split into a ratio of 40/30/30 for the train/validation/test sets.

4 Model Architecture

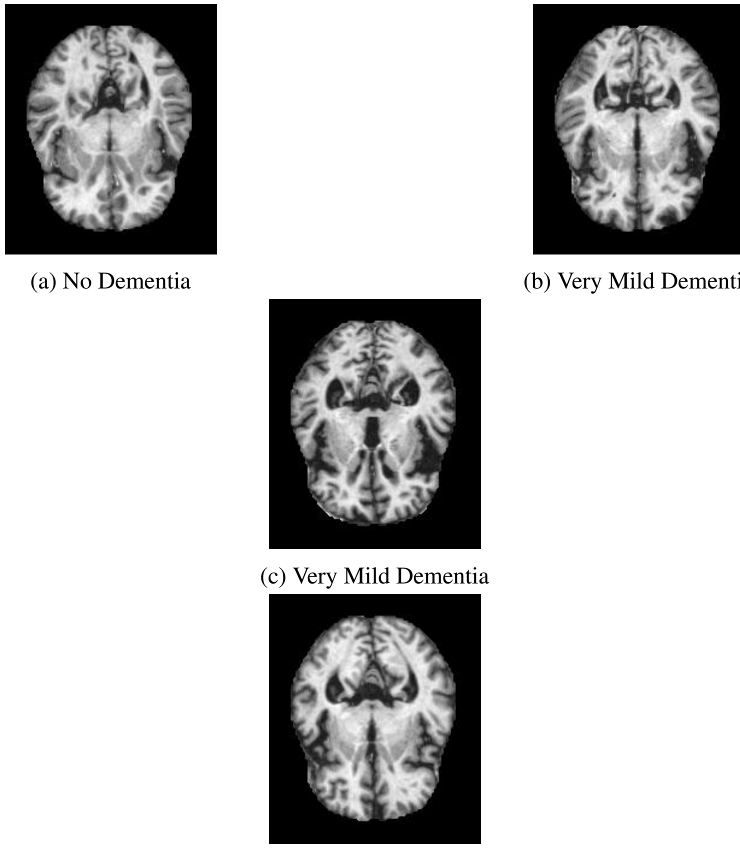
The proposed framework is a multi-layer convolutional neural network consisting of 4 CNN blocks and a fully-connected layer. Each CNN block is constructed in the following manner:

1. Convolutional Layer that takes i channel inputs and o channel outputs with kernel size 3
2. 2D Batch-Normalization Layer
3. 2D Max Pooling layer with kernel size 2 and stride 2
4. Rectified Linear Unit (ReLU) layer

The overall model architecture consists of the following layers:

1. CNN block, 3 input channels and 8 output channels
2. CNN block, 8 input channels and 16 output channels
3. CNN block, 16 input channels and 32 output channels
4. CNN block, 32 input channels and 64 output channels
5. FC layer, output size 4

Despite the fact that the images were in grayscale and technically could have been done in 1-channel input, they were instead read in a 3-channel "RGB" input to minimize the difference in input data



(a) No Dementia

(b) Very Mild Dementia

(c) Very Mild Dementia

(d) Very Mild Dementia

Figure 1: Sample MRI's from each class

between this model and the ResNet model, which will be explained in greater detail within the next section.

4.1 Computational Experiment using ResNet

As my first computational experiment, I wanted to compare the results of the model to those of a preexisting image classification model, namely resnet. ResNet is a neural network that distinctly uses skip connections between layers in order to decrease the effect of vanishing gradients on model training. ResNet has been one of the most successful image classification models to date, achieving top1 error of 19.38% and top5 error on 4.49% on the ImageNet dataset. As such, adopting the ResNet model could prove useful in decreasing training loss.[5]

For the overall architecture, I used the resnet18 model with pre-trained weights, consisting of 17 convolutional layers and 1 fully connected layer, as it was the only model my device could handle without exceeding maximum GPU memory. However, I also appended multiple fully-connected and dropout layers to the end of the model in order to refit the model to only having 4 classes, as well as reducing overfitting. As such, the final model consisted of the following:

1. ResNet, output size=1000
2. Dropout layer, probability=0.2
3. ReLU layer
4. FC layer, output size=500
5. Dropout layer, probability=0.2
6. ReLU layer

7. FC layer, output size=250
8. ReLU layer
9. FC layer, output size=100
10. ReLU layer
11. FC layer, output size=4

4.2 Loss function and training

Due to the classification nature of this problem, I used Cross-Entropy Loss as the metric for gradient-based learning. Cross-Entropy Loss is given by the following formula:

$$l(x, y) = L = \{l_1, ..., l_N\}^\top, l_n = -w_{y_n} \log \frac{\exp(x_{n,y_n})}{\sum_{c=1}^{C} \exp(x_n, c)} 1\{y_n \neq \text{ignore\_index}\} \tag{1}$$

where x is the input, w is the weight, y is the target, C is the number classes, and N is the number of samples in the minibatch[7].

During the optimization step, I used Adam Optimizer[6] with a learning rate of 10^{-3} and trained each model over 40 epochs. However, to preserve features of the original ResNet model, I froze the original resnet architecture and only allowed backpropagation to be done on the added FC layers on the first 30 epochs, and unfroze the layers for the remaining 10 epochs.

5 Results

Table 1: Results of CNN vs ResNet

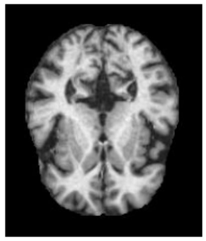
CNN				ResNet			
	Precision	Recall	F1 Score		Precision	Recall	F1 Score
Non	1.00	1.00	1.00	Non	1.00	1.00	1.00
Very Mild	1.00	1.00	1.00	Very Mild	.99	1.00	.99
Mild	1.00	1.00	1.00	Mild	.99	.99	.99
Moderate	1.00	1.00	1.00	Moderate	.98	.97	.98

Overall, we can see that the results, despite having incredibly high Precision, recall, and F1 Scores, are incredibly unrealistic and essentially meaningless. As such, it is important to understand how both models could output near 100% or 100% accuracy on the test set.

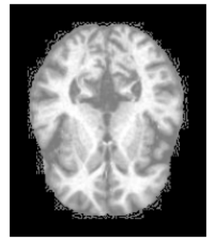
The initial assumption is that I somehow messed up, and that the train/test/validation sets have some kind of overlap between them. However, this is 100% not the case, as when training the model, I explicitly checked for overlap between each set, and found no overlap whatsoever between each set. This means that the high level of accuracy is likely coming from issues with the data itself. As such, the most likely assumption of this may be that the data itself is too simplistic, with too little variation within classes and too much variation between classes, which may not be very indicative of a real-world dataset. A possible solution to this may be to find a better real-world dataset, such as the Open Access Series of Imaging Studies (OASIS) or Alzheimer's Disease Neuroimagining Initiative (ADNI) dataset, but each of these datasets either had issues with accessing the data, or the MRI's themselves were too massive to be processed properly on my machine. As such, the only other possibility to create more "realistic" data is through data augmentation, which introduces the next computational experiment.

5.1 Computational Experiment using Data Augmentation

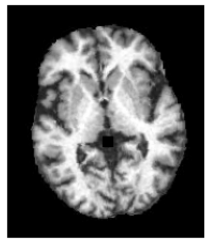
For this computational experiment, I applied 2 different types of data augmentation on the dataset. The first augmentation used was AutoAugment. AutoAugment is a form of data augmentation that automatically searches for improved data augmentation methods to boost accuracy. This form of data augmentation has proven relatively successful in improving test accuracy, achieving a 3.57% error on the imagenet dataset when combined with residual learning networks. framework.[3]



(a) Base MRI Image



(b) AutoAugment Transformation



(c) Manual Augment Transformation

Figure 2: Sample Transformation for Training Image

The second type of augmentation was a manual augment that employs 5 different types of transformations to randomly alter the image. The transformations are listed as below:

1. Color Jitter transformation to modify brightness
2. Horizontal flip with probability 0.5
3. Vertical flip with probability 0.5
4. Rotation in range [-10, 10] degrees
5. Randomized 10x10 cutout

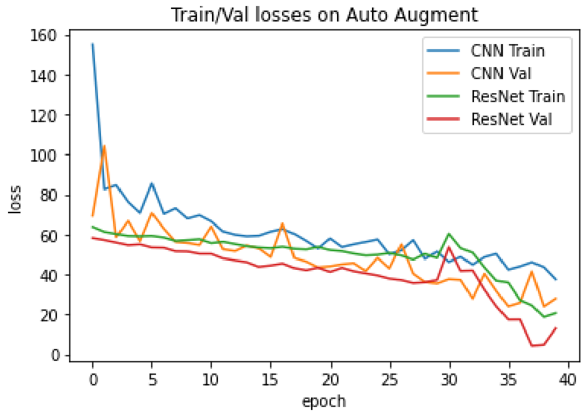
Overall, the purposes of these augments were to "distort" the image to force the model to learn specific types of relative features, rather than static features that may be dependent on absolute position. Additionally, cutout was applied to the manual augment to force the model to exclude specific features in its evaluation; however the impact of this on the final predictions is not incredibly clear.

The results of each transformation can be seen in the below table:

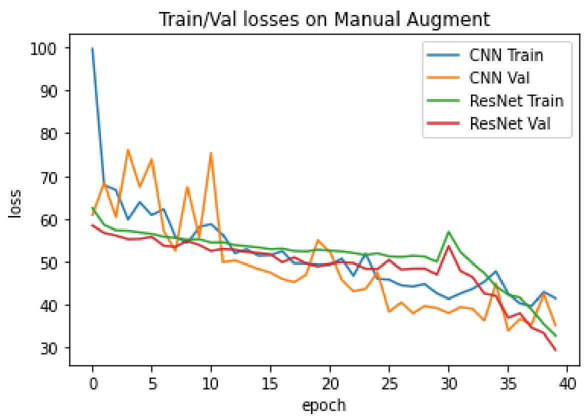
Table 2: Results of CNN vs ResNet

	CNN				ResNet			
		Precision	Recall	F1 Score		Precision	Recall	F1 Score
Auto	Non	.80	.97	.87	Non	.99	.98	.98
	Very Mild	.90	.70	.79	Very Mild	.96	.97	.97
	Mild	.92	.73	.81	Mild	.95	.97	.96
	Moderate	.96	1.00	.98	Moderate	1.00	.98	.99
Manual	Non	.86	.77	.82	Non	.95	.86	.90
	Very Mild	.62	.84	.72	Very Mild	.78	.89	.83
	Mild	.86	.44	.58	Mild	.83	.80	.81
	Moderate	1.00	.73	.85	Moderate	.77	.78	.78

As such, we can see more realistic scores on the test set when data augmentation is applied. We can see an overall trend within the table that ResNet will generally outperform the standard CNN, and AutoAugment will generally outperform the Manual Augment. This is not an incredibly surprising



(a) AutoAugment Transformation



(b) Manual Augment Transformation

Figure 3: Loss of Each Augment Over Time

outcome, as ResNet18 not only has more sophisticated architecture than the standard CNN, but also contains many more computational layers, allowing it to identify significantly more features than with an ordinary CNN. AutoAugment having higher accuracy than the Manual Augmentation may be due to the algorithm's nature of choosing more favorable augmentations on the image. However, it could also simply be the result of augmentation not making sufficient modifications to the original dataset.

6 Conclusion

This project proposed investigated the issue of Image classification of Alzheimer's Disease and proposed a CNN model that would classify its severity based on MRI data. However, the results of the model were worse than that of the more established architecture of ResNet. Additionally, the project investigated the effects of Data Augmentation on the model's ability to identify features and accurately classify disease severity. Additionally, we can see from figure 3 that both models have relatively consistent train/validation losses across both forms of Data Augmentation. As such, the possibility of overfitting (at least on this dataset) is relatively low.

For potential future work on this area, it could be useful to alter the model to use 3-Dimensional MRI data rather than a grayscale image for input. Although the actual MRI would likely contain unrealistically large volumes of tensor data making standard image classification near impossible, there are several methods to condense the tensor into more realistic sizes to improve the efficiency of image classification algorithms. For example, Bae, et al. used a series of 2d slices constructed from 3d brain imagery as data to feed into a CNN[2]. Additionally, Tensor Sketching methods could also be used to reduce the dimensionality of the input tensor, which could greatly speed up the training process for 3d input. Future work within this area could potentially improve both the accuracy and the speed of image classification, and as such there is still much to be done to improve the lives of current and potential Alzheimer's patients across the globe.

References

[1] What is alzheimer's disease?, Oct 2020.

[2] Jong Bin Bae, Subin Lee, Wonmo Jung, Sejin Park, Weonjin Kim, Hyunwoo Oh, Ji Won Han, Grace Eun Kim, Jun Sung Kim, Jae Hyoung Kim, and et al. Identification of alzheimer's disease using a convolutional neural network model based on t1-weighted magnetic resonance imaging. *Scientific Reports*, 10(1), 2020.

[3] Ekin D. Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, and Quoc V. Le. Autoaugment: Learning augmentation strategies from data. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[4] Sarvesh Dubey. Alzheimer's dataset (4 class of images), Dec 2019.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[6] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.

[7] Zhilu Zhang and Mert R. Sabuncu. Generalized cross entropy loss for training deep neural networks with noisy labels, 2018.