Predicting Alzheimer's Disease with Deep Learningkari

Edison Murairi

December 9, 2023

1 Introduction

1.1 Overview

Alzheimer's disease is a progressive neurodegenerative disorder characterized by cognitive decline and memory loss. It is the most common cause of dementia among the elderly population, affecting millions of individuals globally [5]. The disease is characterized by the accumulation of abnormal protein aggregates, such as beta-amyloid plaques and tau tangles, in the brain [11]. Early detection and accurate diagnosis of Alzheimer's disease are crucial for implementing effective interventions and improving patient outcomes. Detecting Alzheimer's disease, however, is challenging because most patients experience a sporadic form, characterized by a late onset [1].

In recent times, the efficacy of deep learning approaches in addressing various challenges within medical diagnosis, particularly in the early detection of Alzheimer's Disease (AD), has become evident. Our final project is geared towards achieving this objective. Numerous researchers have leveraged deep learning models to confront this diagnostic challenge. For example, Liu and colleagues [2] used a 3D convolutional neural network to distinguish between mild Alzheimer's disease dementia from mild cognitive impairment and cognitively normal individuals using structural MRIs. Moreover, Ref. [7] provides a comparative study of the diagnostic performance of 2D, 3D convolutional neural networks and recurrent neural networks when applied to MRI scan of the brain. Building upon these recent advancements, our study will further explore this issue utilizing various neural network architectures, as elaborated in the subsequent sections of this proposal.

1.2 Research Statement

The objective of our project is to leverage the OASIS MRI scans to determine the stage of dementia of a given patient.

1.3 Dataset

For our study, we will utilize the OSAIS MRI dataset [10], which comprises over 80,000 brain MRI images categorized into four stages representing the progression of Alzheimer's disease. The substantial volume and detailed classification granularity of this dataset render it highly suitable for training a deep neural network.

The role of medical imaging, with a specific focus on magnetic resonance imaging (MRI), has grown increasingly significant in the diagnosis of Alzheimer's disease. MRI offers a comprehensive view of the brain's structural details, enabling clinicians to observe changes associated with neurodegeneration. Therefore, we hypothesize that MRI scans can have markers associated with dementia, which can then be learned by a deep neural network.

1.4 Outline

The outline of the shared work is as follows:

- Exploratory Data Analysis (EDA): Karina with assistance by Edison

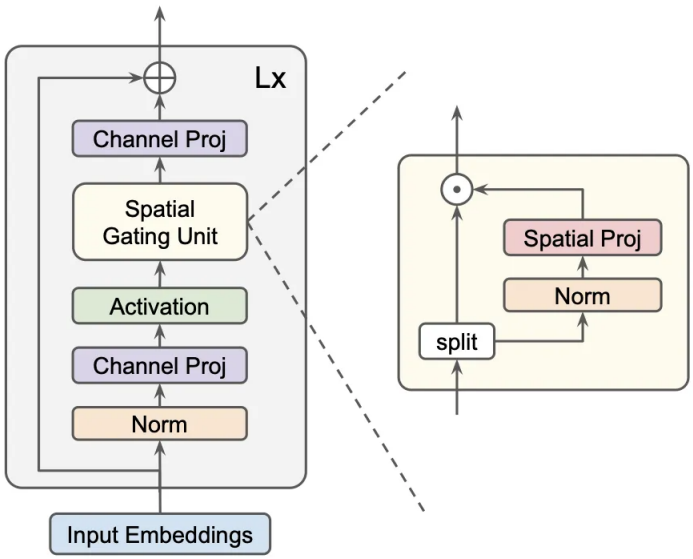

Figure 1: Gated MLP Architecture from Ref. [9]

- Modeling
  - Gated MLP (GMLP): Edison
  - ResNet Model with Attention: Karina
  - Convolutional Neural Network (CNN) with Attention and Explanability: Nina
- Combining all modeling codes into a single package: Karina
- Streamlit Application: Edison

Ultimately, all members contributed to all areas of the project through help debugging, suggestions, and codes. As for my contribution, I will focus onthe GMLP model and the Streamlit application.

2 Background on GMLP

GMLP is an architecture introduced in Ref. [9] as an alternative to transformers. Transformers are characterized by the inductive bias allowing a parametrization of spatial interactions based on the inputs [4]. The purpose of GMLP is to provide an alternative to this inductive bias solely with with mulilayer perceptrons (MLPs).

Figure 1 in Ref. [9]. In this architecture, the input is first normalized and thereafter passed to a channel projector. Let X be the output of the normalization channel. A channel projection performs the operation XU where U is a projector operator[1], and the output is passed to the activation function σ. The activation function the authors suggest is the so-called Gaussian Error Linear Unit (GeLU) [8], which has shown to perform well [9]. Arguably, the most important element of the network is the Gating Unit. First, the output (Z) of the activation function is split into Z_1 and Z_2. Z_2 is passed through the normalizer, and then through the spatial projection unit. We will denote the output of the normalizer by Z_2 as well. The spatial projection unit is problem-dependent. The simplest choice is to use a linear operation such as $Z_2' = W\,Z_2 + b$ where W is a weight matrix and b is a bias. Finally, Z_1 and Z_2' are combined through a Hadamard product as $s(Z) = Z_1 \odot Z_2'$. Finally, $s(Z)$ is passed through a projection unit, yielding $s(Z)V$ where V is a projector, and the input to the network is then added to the output of the last channel projection. To finish, we note that the whole block has to be repeated L times.

[1] A linear operator U is a projector if and only if $U^2 = U$.

3 My codes contribution

In general, we have all contributed codes to different phases of the project. I will describe the codes I have contributed in each phase.

3.1 Initial Codes and EDA

In the initial phase of the project, my first code focused on establishing a well-documented README markdown file on GitHub. The primary objective was to enhance the project's accessibility, ensuring that anyone with repository access could easily comprehend its scope and purpose. To achieve this, I incorporated lines of code to visualize images corresponding to each dementia class within our dataset. The resulting figure was prominently featured in the README file. Although I've made updates to the README on my personal branch, it is currently pending a merge due to a conflict.

Moving on to the Exploratory Data Analysis (EDA) phase, Karina made substantial contributions to the foundational codes. I took on the task of refining the functions, adding indicative labels to the generated figures, and ensuring overall readiness for presentation or potential publication.

3.2 Model Implementation

The subsequent stage involved the implementation of the gMLP (gated Multilayer Perceptron) architecture, guided by the pseudo-codes provided in a referenced paper [9]. While I successfully coded the architecture, I opted to compare it with an alternative implementation from Ref. [6]. After numerous trials and debugging attempts, I settled on specific parameters for the model, including a patch size of 16 for the embedding and six repetitions for the number of blocks.

I observed that the model in Ref. [6] performed constantly a bit better than my own implementation. Although I spent a lot of time on it, I believe there is still some debugging to perform. So, I proceeded with the implementation in Ref. [6] because it also seemed to train faster than my implementation.

3.3 Streamlit Application Implementation

My other role in the project was to develop the Streamlit application for the presentation. For this purpose, I wrote all the codes necessary. Specifically, I wrote the codes for the design of the presentations and I implemented the demos that we performed. For the demos, I wrote the codes to pull an image randomly from the test set once the user presses the 'Run Demo' button. Then, when the user presses the 'Correct Label' button, the demo will show the correct label of the image. Then, the user may press the 'Model Label' button to show what the model output.

3.4 Training and Testing Codes

From this project, we all decided to implement training and testing codes from scratch. I implemented my all dataset class to manage my data and perform data augmentation. I then wrote the codes for training and testing the model. At the end, Karina was able to combine all the models, training and testing codes for uniformity. This file is called 'combined_models.py', and this is the file that we chose to continue working with.

4 Results

The model was trained using ten epochs using the stochastic gradient descent as optimizer with learning 0.001. The loss function was a focal loss with parameters $\alpha = 1$ and $\gamma = 2$ to attempt to give more weights to the hard examples to classify. This way, the model may have a better chance at learning to classify them correctly. Table 1 shows the performance metrics, demonstrating that the model did not perform well. In addition, Figure 2 shows the loss and accuracy on training set and the test set for ten epochs. The plot clearly shows that although the loss on the training set is decreasing, it is increasing on the validation set. This result is concerning because it is suggesting that the model does

Metric	Value
Accuracy	0.51
F1 Score	0.37
COH Score	0.06

Table 1: Performance Metrics of the GMLP model

not generalize well. Moreover, the accuracy is also increasing on the training set while it decreases (or almost remains constant on the validation set). These two results strongly indicates that this model overfits the training set, and does not generalize.

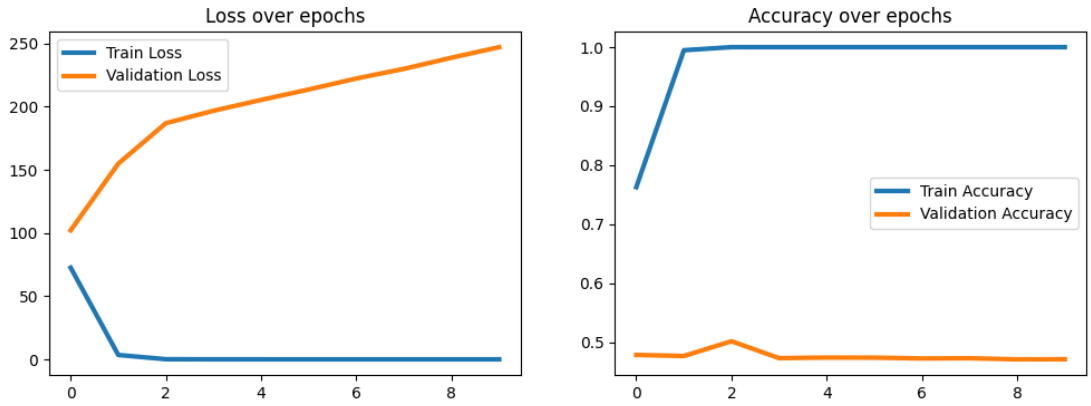

Figure 2: Loss and Accuracy on training set and validation set over 10 epochs

This overfitting result is probably due to the class in-balance. This diagnostic is made even clearer when examining the confusion matrix, Fig. 3. In this confusion matrix, 0 refers to 'No dementia', 1 refers to 'Very mild dementia' and 2 refers to 'Mild dementia'. In this confusion matrix, we can see that the model overpredicts 'No dementia'. More specifically, the model is clearly not capable of distinguishing between 'No dementia' and 'Very mild dementia'. There are at least two reasons explaining these results. The first is that it is possible that the MRI scans for 'No dementia' and 'Very mild dementia' are structurally similar enough that the model is not complex enough to distinguish between the two. The other scenario is that the class inbalance skews the model towards the 'No dementia' class since there are a lot more examples of this class. To address this issue, we could have use a Generative Adversarial Network (GAN) to generate more examples of the two under-represented classes so that the model can have a better chance at learning to classify them correctly. In fact, I implemented a GAN but it was not trained long enough, and did not produce realistic images of the two classes.

Another source of the poor result stems from how we are treating the MRI scans. MRI scans provided are 3D images, arranged into 2D slices. For each patient, we were provided slice 100 to slice 160, or 61 slices. In our modeling, we have treated each slice as independent image while they belong to a unique patient. Therefore, it is possible that our model will output different classes for the same patient depending on the slice that was given. This issue should have been addressed possibly by taking majority.

Another reason that the model did not perform very well is that by considering the individual slices of a patient's brain as independent image, we have lost the spatial relationship between the slices. Therefore, it is possible that the model is not able to capture very well relevant details about the brain in order to make correct classifications. There are papers that address this issue in the literature with the same dataset. For example, in Ref. [3], achieves $\sim$ 90% using Reccurent Visual

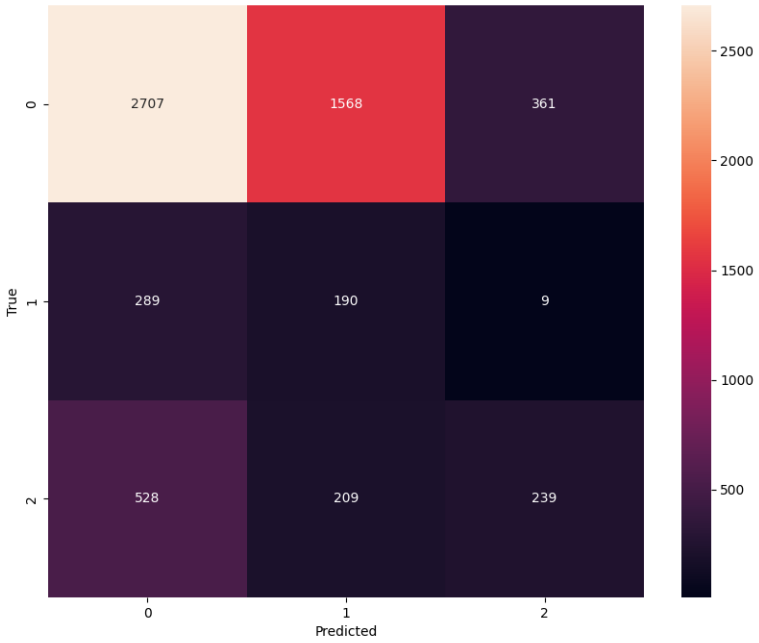

Figure 3: Confusion Matrix. 0 refers to 'No dementia', 1 refers to 'Very mild dementia' and 2 refers to 'Mild dementia'.

Networks (RVN) and Transformers based models by considering all the slices together as a 3D image.

5 Codes Percentage

For this calculation, I have considered only the codes that I have submitted. The modules are not counted since they are packaged codes, and it is not clear how many lines of codes they have. Then, after final calculation, the result is 93%, rounded.

6 Summary and Conclusion

Our project centered on predicting Alzheimer's disease using deep learning, with a focus on diverse neural network architectures. While each team member played a role in different phases, my primary contributions includedc contributing to the EDA, implementing the gMLP model and developing a user-friendly Streamlit application.

I found that the gMLP model did not perform well with accuracy $\sim 51\%$, for instancce. The model suffered from generalization issues most likely due to overfitting. The issue may stems from the class in-balance, which could be addressed by generating examples of the under-represented classes. Moreover, we could improve our model in general by considering the fact that different slices belong to a unique patient. Therefore, the final label for a given patient could be for example the majority over all the 61 different slices. Moreover, a model that directly laverages the 3D images, such as 3D CNNs or RVNs could yield a better performance as is demonstrated in the literatutre.

References

[1] Alzheimer's disease. *Nature Reviews Disease Primers*, 1, Oct. 2015.

[2] Generalizable deep learning model for early alzheimer's disease detection from structural mris. *Scientific Reports*, 12(1):17106, 2022.

[3] F. Altay, G. R. Sanchez, Y. James, S. V. Faraone, S. Velipasalar, and A. Salekin. Preclinical stage alzheimer's disease detection using magnetic resonance image scans. 2020.

[4] D. Bahdanau, K. Cho, and Y. Bengio. Neural machine translation by jointly learning to align and translate. 2014.

[5] C. Ballard, S. Gauthier, A. Corbett, C. Brayne, D. Aarsland, and E. Jones. Alzheimer's disease. *Lancet*, 377(9770):1019–1031, Mar. 2011.

[6] P. Bo. Blinkdl/rwkv-lm: 0.01, Aug. 2021.

[7] A. Ebrahimi, S. Luo, and Alzheimer's Disease Neuroimaging Initiative. Convolutional neural networks for alzheimer's disease detection on MRI images. *J Med Imaging (Bellingham)*, 8(2):024503, Apr. 2021.

[8] D. Hendrycks and K. Gimpel. Gaussian error linear units (GELUs). 2016.

[9] H. Liu, Z. Dai, D. R. So, and Q. V. Le. Pay attention to MLPs. 2021.

[10] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner. Open access series of imaging studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J. Cogn. Neurosci.*, 19(9):1498–1507, Sept. 2007.

[11] P. Scheltens, K. Blennow, M. M. B. Breteler, B. de Strooper, G. B. Frisoni, S. Salloway, and W. M. Van der Flier. Alzheimer's disease. *Lancet*, 388(10043):505–517, July 2016.