DIAGNOSIS OF ALZHEIMERS DISEASE USING DEEP LEARNING WITH CYCLEGAN FOR DATA AUGMENTATION

Sunny Wang

***Abstract*— Alzheimer's disease is a progressive disease that causes deterioration of neurons in the brain, leading to dementia and eventually death. Diagnosis of Alzheimer's conventionally consists of a combination of neuropsychological tests and laboratory tests, and clinical diagnosis accuracy lies at around 77%, As Alzheimer's is associated with loss in brain mass, which can be discerned from MRI scans, it is a suitable task for deep learning and computer vision. However, deep learning typically requires large amounts of data, and medical data is often scarce. A recent breakthrough in machine learning, the generative adversarial network (GAN), allows for generation of realistic images, providing a potential solution to lack of data. In this study, we construct ResNet50-based convolutional neural networks to perform Alzheimer's disease classification using MRI scans, achieving an F-1 score of 89%. Furthermore, by generating synthetic samples using CycleGAN, we demonstrate that GANs can significantly improve classification accuracy when used for data augmentation, achieving an F-1 score of 95%.**

I. Introduction

Alzheimer's disease (AD) is a progressive disease characterized by loss of cognitive ability and is the sixth leading cause of death in the United States. AD is typically classified and diagnosed based on a variety of factors, including cognitive tests and laboratory tests. According to a study conducted by Beach et. al., the overall clinical diagnosis accuracy was 77% with a low true negative rate [1]. This is far from perfect, raising the demand for a computer assisted tool to reinforce physicians' diagnosis.

As AD causes the breakdown and death of neurons, these changes in brain mass can be observed through technology such as magnetic resonance imaging (MRIs). These scans are suitable for computer vision and deep learning algorithms, such as the convolutional neural network (CNN), which has achieved impressive results in classification.

CNNs typically require large datasets to perform effectively. However, medical data is often scarce and limited in size, largely due to the high standards of consistency and organization required for medical data, and the cost and time required for data collection. This raises a demand for data augmentation techniques to improve medical machine learning models. One recently introduced technique is the generative adversarial network (GAN) [2], which achieves promising results in image generation.

The use of CNNs for Alzheimer's disease diagnosis has become more prominent in recent years. Hosseini-Asl et al. [3] used a deeply adaptive 3D convolutional neural network (DSA 3D-CNN) achieving a 94.8% accuracy in task specific classification. Glozman et al. [4] proposed a network of 2D CNNs, applied to each of three images extracted from each sample in the ADNI dataset, which achieved 83% accuracy using PET scans. However, there have been relatively few applications of GANs in classifying AD, with none of them using it for the purpose of data augmentation.

In this study, we investigate the potential for using deep learning in Alzheimer's disease classification by creating a convolutional neural network model. We also test the feasibility of using GANs for data augmentation, specifically using the CycleGAN architecture [5].

II. Methods

Fig. 1 outlines the model pipeline, consisting of dataset acquisition, preprocessing (including GAN augmentation), and classification using a CNN.

Data Acquisition. Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. For the purposes of this study, we trained a network to classify between Alzheimer's disease (AD) samples and normal cognition (NC). The dataset contained 705 samples labeled as NC and 476 samples labeled as AD.

Data Preprocessing. The data was stored in NIFTI volumes, which were converted into three dimensional NumPy arrays using nibabel. We extract three slices from the volume, one each from the axial, coronal, and sagittal orientations. The slices are taken by retrieving the midpoint of each axis length and were resized to 224 x 224. Skull stripping was applied, which is the process of removing the skull from the MRI images, allowing for more consistency among samples. This was done using the Extractor function from the deepbrain library.

CycleGAN based Data Augmentation. We constructed the models using the implementation from [4]. The model architecture is shown in Fig. 2.

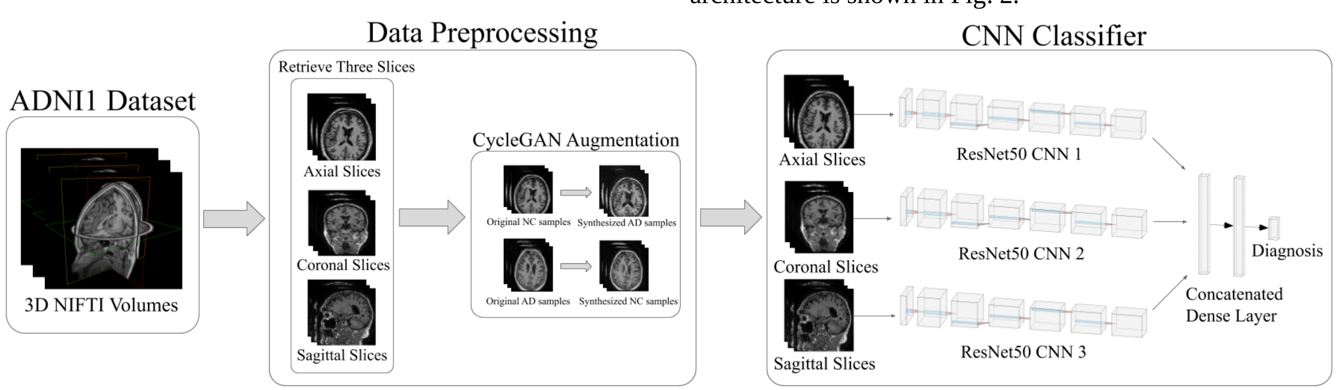

Figure 1. Model Pipeline.

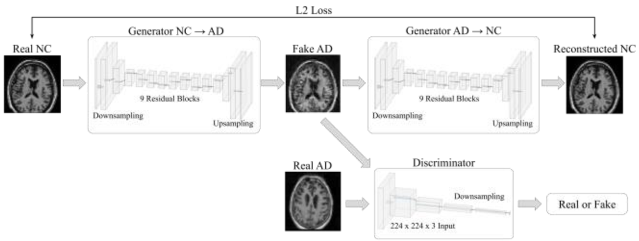

Figure 2. CycleGAN Model Architecture.

The original dataset was split according to the label and randomly paired. Three individual CycleGAN models were created, where each one was trained on data from a different MRI slice. Each model used the Adam optimizer with learning rate of 2e-4 and were trained for 100 epochs with a batch size of 1, as specified in the CycleGAN paper. The trained model was used to generate sufficient samples to create a balanced dataset. A total of 705 AD samples and 476 NC samples of each orientation were generated, for a total of 1181 images of each class.

Convolutional Neural Network Classifier. We used a transfer learning approach to create the model architecture as it would save training time and is generally effective when datasets are small. We used the ResNet50 convolutional neural network (CNN) as our pretrained model and modified the CNN to utilize three inputs to better encapsulate volumetric data. The model architecture is shown in Fig. 3.

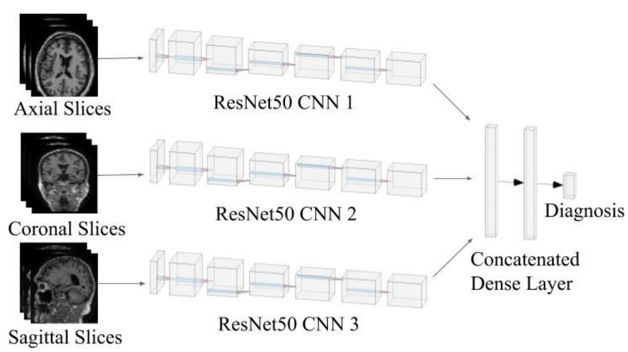

Figure 3. CNN Model Architecture

The neural network was fine-tuned using the Adam optimizer with a learning rate of 1e-4 and trained for 50 epochs with a batch size of 32. Models were evaluated using accuracy, precision, recall, and F1 score, with F1 score being the primary performance indicator.

III. Results and Discussion

TABLE I. Comparison between ResNet50 models. SS denotes models with skull stripping preprocessing.

Metric	ResNet50	ResNet50 + GAN	ResNet50 + SS	ResNet50 + SS + GAN
Accuracy	0.891	**0.954**	0.908	0.949
Precision	0.932	**0.951**	0.882	0.944
Recall	0.804	0.942	0.900	**0.959**
F1 Score	0.863	0.946	0.891	**0.951**

Table I indicates that skull stripping was effective as a method of preprocessing, as it increased the F1 score of the model from 0.863 to 0.891. It also demonstrates that the addition of CycleGAN substantially improves CNN classification performance, increasing all performance metrics for each model it was applied to. The F1 score increased from 0.863 to 0.946 for the original model and 0.891 to 0.951 for the skull stripped model.

From this, it is reasonable to infer that the synthesized images had meaningful features that benefited the model. The increased size and balance among classes in the CycleGAN augmented dataset are also factors that are potentially responsible for the increase in performance. Overall, these results demonstrate the effectiveness of GANs in data augmentation.

IV. Conclusions

In this study, we constructed CNN models utilizing the ResNet50 architecture to diagnose Alzheimer's disease while also addressing the problem of size limitations in medical datasets with the use of generative adversarial networks (GANs). Specifically, we used CycleGAN to generate images of one class using the other, balancing the dataset and increasing its overall size. Our results show that classification accuracy improved substantially, increasing the F1 score from 0.863 to 0.946 for the standard model and from 0.891 to 0.951 for the model utilizing skull stripping. Due to the lack of large datasets in many medical fields, the results obtained in this study can be generalized to many other fields as well. Overall, with promising results in data augmentation, GANs have potential to significantly improve upon classification tasks across a wide variety of applications.

References

[1] Thomas G Beach et al. "Accuracy of the clinical diagnosis of Alzheimer disease at National Institute on Aging Alzheimer Disease Centers, 2005–2010". In: Journal of neuropathology and experimental neurology 71.4 (2012), pp. 266–273.

[2] Ian Goodfellow et al. "Generative adversarial nets". In: Advances in neural information processing systems. 2014, pp. 2672–2680.

[3] Tanya Glozman and Orly Liba. Hidden cues: Deep learning for Alzheimer's disease classification CS331B project final report. 2016.

[4] Ehsan Hosseini-Asl, Robert Keynton, and Ayman El-Baz. "Alzheimer's disease diagnostics by adaptation of 3D convolutional network". In: 2016 IEEE International Conference on Image Processing (ICIP). IEEE. 2016, pp. 126–130.

[5] Jun-Yan Zhu et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks". In: Proceedings of the IEEE international conference on computer vision. 2017, pp. 2223–2232.