Medical Image Classification Using Deep Learning for Alzheimer's Detection and Limitation Analysis

Shiqi Hu
Biomedical Engineering
Columbia University
sh4193@columbia.edu

Yanting Yang
Biomedical Engineering
Columbia University
yy3189@columbia.edu

Wenjie Lin
Mechanical Engineering
Columbia Univeristy
wl2789@columbia.edu

Matthew Ho
Computer Engineering
Columbia University
mh4195@columbia.edu

***Abstract*—Nowadays, deep learning is frequently used in the field of medical diagnosis. Since the Alzheimer's disease (AD) is one of the most common neurodegenerative diseases in the elderly, there is a great clinical benefit in realizing automated diagnosis of AD without any prior feature analysis and regardless of the variability of image protocols and scanners. This paper involves our project that from the very beginning achieves the detection of AD using magnetic resonance images (MRI) with deep learning algorithms. We performed experiments on two CNN-based models for the classification of AD. The best classification model was AlexNet, which classified AD patients with healthy controls with an accuracy of 70.4%.**

Keywords—Alzheimer's Disease, Cnvolutional Neural Network, AlexNet, Medical Image Processing, MRI, Deep Learning

I. Introduction

A variety of techniques are used to detect Alzheimer's disease (AD). One technique uses magnetic resonance images (MRI) to detect brain abnormalities and predict which patients may develop Alzheimer's disease. Several studies have shown that structural MRI estimates of tissue damage or loss in characteristically vulnerable brain regions, such as the hippocampus and entorhinal cortex, are predictive of progression of mild cognitive impairment (MCI) to AD [1].

MRI images can be combined with machine learning algorithms to try and detect MCI to AD progression. Previous work has shown success with deep learning analysis for MCI to AD progression [2]. Deep learning differs from other machine learning algorithms since it requires little pre-processing and can detect abstract and complex patterns, making it useful for detecting subtle differences and abnormalities in anatomical images.

In this project we extend work using deep learning to build our own convolution neural network that is able to detect MCI to AD progression from a set of MRI images for a patient. The neural network is trained on data from Open Access Series of Imaging Studies (OASIS). We discuss our methods, process, results, and conclusions for potential future expansion in this project.

II. Approach

The roadmap is as follows. The raw scans were processed by OASIS and FreeSurfer and we will describe some of the techniques used as examples of extensions that could be done from this project. First the images were anonymized by removing dates, names, and IDs. Next the images were motion corrected and the white matter and deep gray matter volumetric structures in the subcortex region were segmented. This is done so that regional volumes can be then corrected for head size in order to normalize comparisons. Additional processing included smoothing to achieve a common spatial resolution to minimize scanner differences.

From the brain mask files we downloaded, we first visually examined them using SPM-12. SPM is an fMRI analysis software tool that can be run in MATLAB to analyze brain imaging data. SPM tutorials show how processing techniques mentioned above can be applied in SPM [3]. After we graped a general idea of out image dataset, we load our images to python to do further data handling.

After s series steps of data preprocessing, our input image data were coverted to Numpy arrays and were fed into our CNN-based models. Our deep learning models were based on the conventional CNN model which is shown in Fig. 1. Specifically, the inputs were normalized 2D/3D images and outputs were numerical numbers representing the possibility that the input belonged to each group. The architecture contains: convolutional block that comprised of the convolutional layer followed by a pooling layer which achieved rotation and translation invariance, fully-connected layer, output layer.

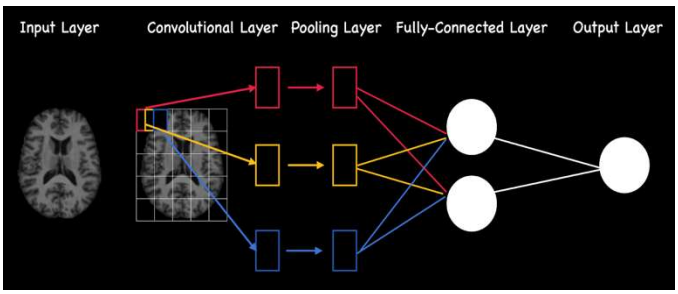


Fig. 1. Basic architecture of our convolutional neural network

To get ideal results, we tried different kinds of model structures and modified the parameters, which will be shown in the Results Session.

Additionally, Fig. 2. shows how the experiments performed and how we validate our model.

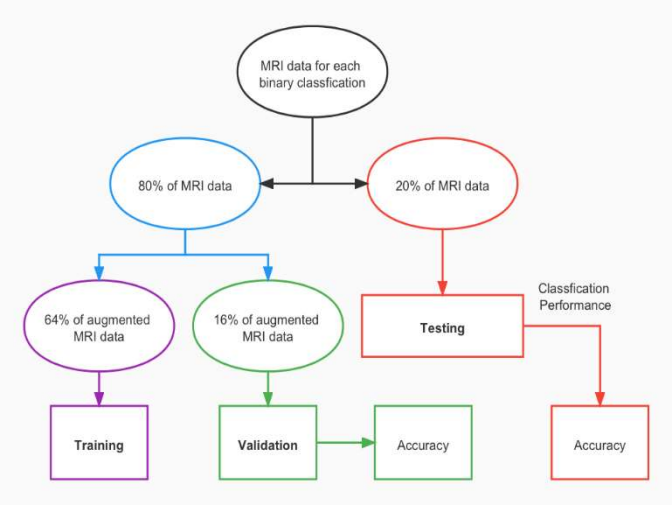


Fig. 2. Flowchart of main steps of the experiments performed.

III. Experiments

A. Database

Our input data is from the OASIS database [4], which is a collection of >1000 participants generated by the Knight ADRC and its affiliated studies. OASIS provides open access to a significant database of neuroimaging and processed imaging data across a broad demographic, cognitive, and genetic spectrum for use in neuroimaging, clinical, and cognitive research on normal aging and cognitive decline. In this project we use OASIS-1 which is specifically for Alzheimer's disease.

This data set consists of a cross-sectional collection of 416 subjects covering the adult life span aged 18 to 96 including individuals with early-stage Alzheimer's Disease (AD). For each subject, 3 or 4 individual T1-weighted MRI scans obtained within a single imaging session are included. The subjects are all right-handed and include both men and women. 100 of the included subjects over the age of 60 have been diagnosed with very mild to mild AD[4]. We are particularly interested in the Mini-Mental State Examination (MMSE) (Rubin et al., 1998) and Clinical Dementia Rating (CDR; 0 = nondemented; 0.5 – very mild dementia; 1 = mild dementia; 2 = moderate dementia) (Morris, 1993). All participants with dementia (CDR >0) were diagnosed with probable AD.

B. Data Selection and Data Preprocessing

All the raw data that contains 416 subjects are downloaded from OASIS-1 for preprocessing. The txt file in each subject folder includes acquisition details and anatomic measures derived from the scan images. The subjects that do not provide clinical diagnosis labels (CDR) are removed from the datasets. For each subject, all images are in 16-bit big-endian Analyze 7.5 format. A Python package NiBabel is used to read this format and convert it to a Numpy array for training.

After removing unusable subjects, only 233 subjects have left. Among those samples, there are only two subjects with CDR=2 in the database, the sample size of this group is too small to be representative thus these two subjects were deleted from the dataset. In addition, the imbalance in the size of the dataset across classes was quite evident in our problem. In the original training set, 88 samples had CDR=0, 45 samples had CDR=0.5 and 16 samples had CDR=1, the sample size of the class with CDR=0 was greater than the sum of the other two classes. Using such a training set greatly increased the difficulty of learning. The model tended to exhibit a bias toward the majority group while incorrectly ignoring the minority group. To address this issue, we simplified our classification task from predicting MCI to AD conversion with multiple labels to a binary classification problem with only HC (CDR=0) and others who are likely to develop AD (CDR>0). Therefore, subjects with CDR=0 are labeled as nondemented and those with CDR=0.5 and 1 are labeled as probable AD.

Our initial attempt was to use 3D images as input data. However, our current hardware cannot handle such a large dataset due to limited computing power and memory space (google colab RAM space). In addition, the available training data size was only 233, which was far from enough for the model to learn the feature space of the images. Therefore, 2D slices were extracted from the original 3D images. In details, we selected 30 transverse slices from each MRI scan image, and the input size was changed from (176, 208, 176) to (176, 208). Then, all 2D images and their corresponding labels were combined together and randomly shuffled to form the inputs.

C. Model Architecture

In this study, two CNN-based models were used to conduct the classification task. First, a basic CNN was constructed. It contained three basic convolutional blocks with number of filters 128, 64 and 32, kernel sizes of 5×5 and strides of 1, followed by a flatten layer and two dense layers (64, 32 nodes) with ReLU activation function. The output layer is activated by softmax with output nodes 2/3 depending on the classify classes.

AlexNet is well known for its great ability in image recognition and classification and was also used in our study. We proposed a AlexNet-based deep learning model with 5 CNN blocks. In AlexNet's first layer, the convolution window shape is 11×11 with a filter of 96 and a stride of 4 to greatly reduce the height and width of the output. The second layer is reduced to 5×5 with a filter of 256 and paddings for consistent height and width across the input and output. Both CNN layers are followed by a max pooling layer with size of 4 and a stride of 2. Then, three successive convolutional layers with a smaller window of 3×3 and filters of 384, 384 and 256 respectively are used, followed by another max pooling layer. The outputs of CNN blocks are flattened and connected to two fully-connected layer. Dropout layer is used to mitigate overfitting. The shape of output layer is 2 or 3 depends on the number of classes. All the layers use ReLU as activation function except that the output layer uses Sigmoid for classification.

D. Model train and preformance evaluation

All our models were trained for 20 epochs with a batch size of 32 and an earlystopping callback option to prevent

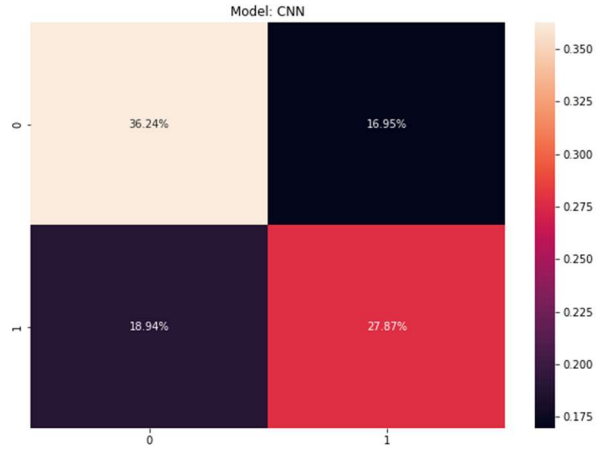
overfitting. The performance of each model was evaluated by loss function and dataset accuracy.

IV. Results

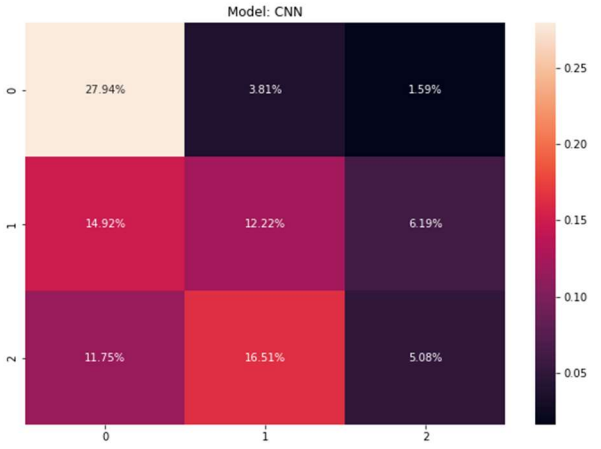
Based on different parameters modified, we got the results shown in TABLE I. For both models, the classification preformance of two labels outperformed that of three labels, which is consistent with our expectation. The best structure is AlexNet with binary output, which is of greater practical importance because it achieves relatively high test accuracy, although its training accuracy is slightly lower than that of the CNN model.

TABLE I. Training Results

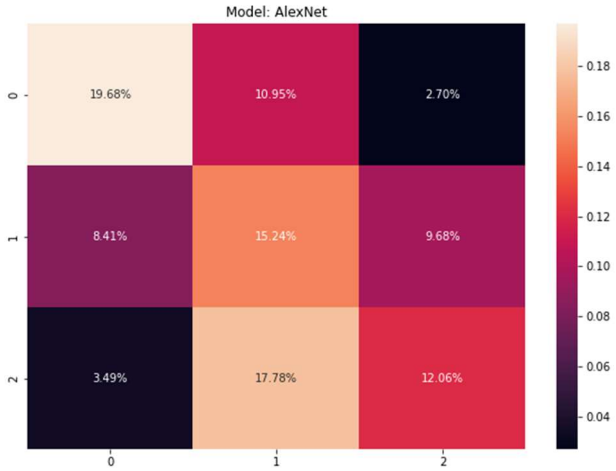
Model	Result		
	Train Accuracy	***Train Loss***	***Test Accuracy***
CNN—two outputs	77.4%	0.4683	64.1%
CNN—three outputs	57.3%	0.8924	45.2%
AlexNet—two outputs	72.2%	0.5407	70.4%
AlexNet—three outputs	63.1%	0.7830	57.8%



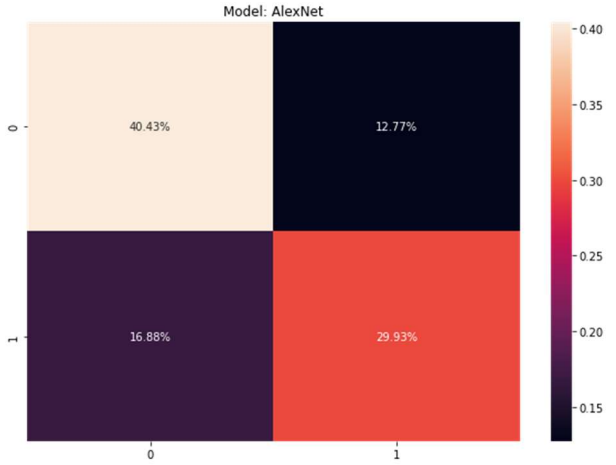
(a)



(b)



(c)



(d)

Fig. 3. Confusion matrix plotted in seaborn plot style. Plots illustrates the classification performance of test data on four model structures. The x-axis denotes the predicted classesnd the y-axis representing the true classes of the input data. The percentage numbers in each colour rectangle represents the ratio of data belongs to that area to total data. (a): CNN for two labels; (b): CNN for three labels; (c) AlexNet for two labels; (d) AlexNet for three labels.

V. Discussion

High accuracy of prediction for AD MRI is of great importance. At the beginning, we tried a 3D CNN-based deep learning model for 3D input images, and the result is not ideal. Later, after we converted the input into 2D images, it started to learn normally and the results were much better than those whose input was 3D.

And as we anticipated in the beginning, there were a lot of challenges when implementing our deep neural network. The problems are as follows:

- Limited dataset size

The size of the dataset matters. If it's too small, it is difficult for the model to learn th whole feature space, and the accuracy of the neural network will converge quickly and stay unchanged. In our project, we converted 3D images into 2D images as input data, and the results became better. In detail, for each 3D subject, we extract 30 slices 2D images from the z axis perspective.

- RAM size limitation

Because all software is written in Python using Colab Pro, the RAM size is still not enough for storage when we using 3D mages. We have tried to utilize Generator to solve this challenge. Generator is specifically designed to handle large datasets in Keras with similar functionality to setting batch size, enabling smaller data sizes to be processed at once and allowing for less storage space requirements. However, the performance of using 3D images with generator still far from satisfactory, so we turned to train our models with 2D images with larger sample size.

- Imbalance class

The dataset was divided into four groups baed on the patients CDR ratings. Because the amount of data in the fourth class (CDR=2) is too small and not representative, this class is discarded. The first class with CDR=0 has the largest amount of data, even more than the combined amount of data of the other two classes (CDR=0.5 and CDR=1). Therefore, this could lead to a huge bias in the model training, producing a classification result completely biased towards the group with CDR=0. This assumption was verified when we utilized 3D images to perform the binary classification. The output accuracy of our model was always 0.5946, which corresponded exactly to the ratio between group 0 (healthy controls with CDR=0) and group 1 (CDR=0.5 or 1). After we expanded our dataset by extracting 2D slices from each 3D MRI image, we randomly selected the same number of samples from each classes to form the final inputs, thus eliminating the imbalance.

- Inherent error of label

The label we obtain from the OASIS is Clinical Dementia Rating (CDR), which represents probable AD if greater than 0. It is an important assessment in the early and accurate diagnosis of dementia. Patients' cognitive and functional performance is quantified through a structured-interview protocol. However, the assessment is ultimately subjective in nature. In our dataset, the numbers of moderate and severe symptoms are extremely small, which make the overall samples not representative.

REFERENCES

[1] S. Basaia, F. Agosta, L. Wagner, E. Canu, G. Magnani, R. Santangelo, M. Filippi Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks Neuroimage Clin., 21 (2018), pp. 1-8, 10.1016/j.nicl.2018.101645

[2] S. Vieira, W.H. Pinaya, A. Mechelli Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications NEUROSCI. BIOBEHAV. REV., 74 (2017), pp. 58-75

[3] https://andysbrainbook.readthedocs.io/en/latest/SPM/SPM_Overview.html

[4] https://www.oasis-brains.org/

.