Convolutional Neural Networks to Distinguish Age-Related Macular Degeneration and Normal Images Obtained Through Optical Coherence Tomography

Ria Kim and Jonathan Medley

***Abstract*— Diagnosis of age-related macular degeneration can be performed via manual inspection of optical coherence tomography eye images. Recent analysis of ophthalmology clinical practice has found that physicians are limited in their ability to process the full extent of data OCT eye images provide, and their diagnoses are subject to differences in interpretation. Faster diagnostic methods also lead to faster provision of treatment, which can improve patient health and vision outcomes. To improve speed and standardization in AMD diagnosis, an image processing algorithm using a convolutional neural network was applied to a large set of OCT images. The algorithm attempted to independently classify 100 depth-varied OCT images between AMD and control images of known diagnosis for 269 AMD patients and 115 control patients, totaling 38,400 images. The algorithm correctly classified 74.1% of control images and 85.7% of AMD images, indicating early diagnostic success and suggesting further algorithm development is necessary to develop an automated classification system.**

I. INTRODUCTION

Age-related macular degeneration (AMD) is a disease that affects between twenty and twenty-five million people worldwide [1]. AMD deteriorates the macula of the eye, which is the central area of the retina. The retina is the layer of nerve tissue that receives light and outputs visual signals to be interpreted by the brain. The macula is the region of the retina that provides the greatest color vision and the highest resolution of the signals generated throughout the retina [2]. As the number and proportion of people in the world over the age of 60 continues to increase, the incidence of age-related diseases including AMD is expected to increase [1]. The progression of both forms of AMD, wet and dry, can both be slowed with treatment. Because its primary symptom, vision loss, does not occur after the initial onset of AMD, early detection and treatment can allow patients to retain their visual abilities for significant periods and lessen the severity of their vision loss [3]. Fast, reliable early AMD detection methods are therefore essential to early AMD identification and treatment as a crucial component of adequate medical care for elderly patients.

Optical Coherence Tomography (OCT) has recently emerged as a common imaging tool used in the early diagnosis of AMD. These images can indicate the presence of fatty deposits on the macula called drusen, a key indicator of AMD progression. Many adults have one drusen and are still considered healthy [4]. The number and size of drusen can be used to differentiate stages of AMD from early to intermediate to severe [2]. Only in severe cases do significant vision problems occur, while OCT images can detect early and intermediate forms of AMD. Because of this, OCT imaging can be used to screen for AMD during adult eye exams, which are recommended by the American Academy of Ophthalmology for all adults beginning at age forty. Yearly exams are widely recommended for adults age sixty and older [1]. Quickly processing and analyzing OCT scans to identify the presence and severity of drusen improves and accelerates the process of diagnosing and treating AMD.

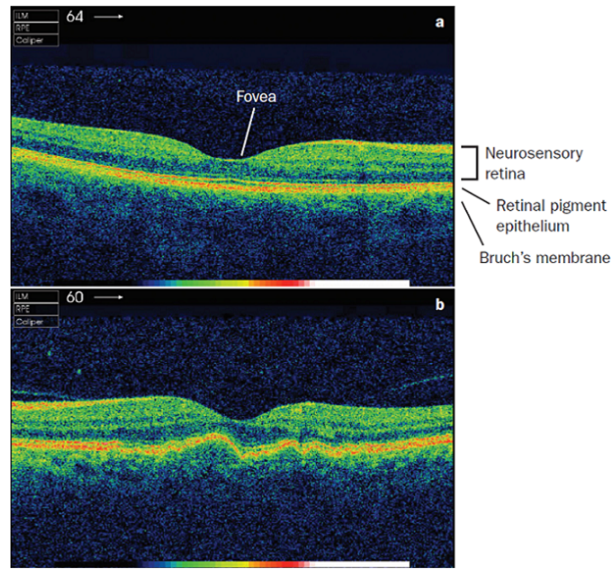

Fig. 1. Foveal SD-OCT images with and without drusen with indicated retinal layers [5]

The use of OCT analysis over traditional photographs represents a current area of biomedical research. Recent studies have suggested that cross-sectional spectral domain OCT (SD-OCT) images may offer more useful image features than traditional eye photographs, which have historically served as the gold standard for AMD diagnosis [3]. Through cross-sectional SD-OCT, the characteristics of various layers of the eye can be more readily observed and analyzed. These images can be readily interpreted by certified readers [3]. The presence of drusen in the layer of the retina in which they

*This work was not supported by any organization

[1]Ria Kim is pursing her Masters of Biomedical Engineering, George Washington University, 2121 I St NW, Washington, DC 20052 USA ria_kim@gwu.edu

[2]Jonathan Medley is pursing his Masters of Biomedical Engineering, George Washington University, 2121 I St NW, Washington, DC 20052 USA medley@gwu.edu

occur, the retinal pigment epithelium (RPE), can be observed in patients classified with AMD.

However, manual classification OCT eye images is subject to a number of risks and shortcomings. OCT systems generate images that contain more data than ophthalmologists are able to utilize clinically. With increased use of OCT systems in diagnostic practice, there is an urgent need for image processing systems to help clinicians make use of the data available to them within clinical time constraints [6]. In addition, manual diagnosis via OCT images is subject to the interpretation of the physician. Automated processing in diagnosis can help minimize subjectivity in diagnosis to provide increased standardization [7].

One promising method used to analyze and interpret medical images is convolutional neural networks. A convolutional neural network is a type of deep neural network, a system that perform tasks using multiple layers of non-linear operations to recognize patterns, and improve their ability to perform these tasks successfully with training. Specifically, convolutional neural networks incorporate convolution operations into at least one layer rather than simple matrix multiplication operations [8].

II. SOLUTION

A. VGG19

Through neural networks, image patterns that are not observable or even interpretable to human observers can be detected automatically, and can be used in determining the characteristics of an image [9]. In clinical context, convolutional neural networks can be used to detect disease patterns in images to calculate a diagnosis. Because of this, convolutional neural networks represent a promising approach to improving the speed and standardization of diagnosing AMD from OCT images. In this study, a convolutional neural network called VGG-19 was applied to a dataset of OCT images classified as presenting or not presenting AMD by experts using manual image analysis for drusen. The VGG-19 algorithm, developed by the Visual Geometry Group in the Department of Engineering Science at Oxford University, contains 19 weight layers and has demonstrated superior image recognition capabilities [10]. A version of this system performed best at localization and second-best in classification tests out of all algorithms submitted to ImageNet's Large Scale Visual Recognition Challenge in 2014 [11]. The goal of this study was to use VGG-19 to correctly classify AMD in OCT images.

B. Image Dataset

The image dataset to be used for this project is a subset of the AREDS 2 Ancillary Spectral Domain Optical Coherence Tomography Study (A2ASDOCT) [3]. 269 patients were previously diagnosed with AMD and 115 patients were identified as control subjects based on traditional eye photographs examined by trained graders [3]. For each patient, 100 OCT images were taken centered around the fovea, which is in focus in the 50th image. These images are taken with axial and lateral resolution of 3.24 and 6.70 micrometers, respectively [3].

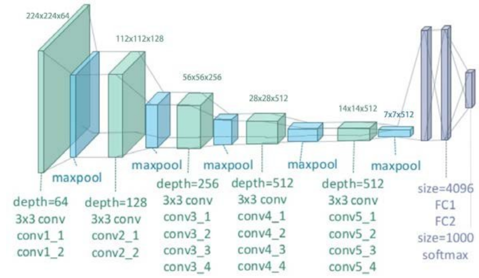

Fig. 2. Structure of the VGG19 Model. This model is characterized by the use of a convolution kernel size of 3x3. The structure allows the for the detection of lower-level to high-level feature patterns in order to classify different images with a high accuracy [9].

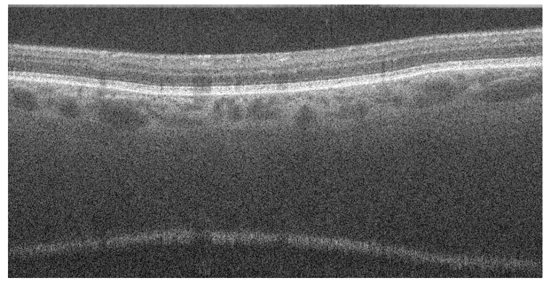

Fig. 3. A sample OCT image showing AMD.

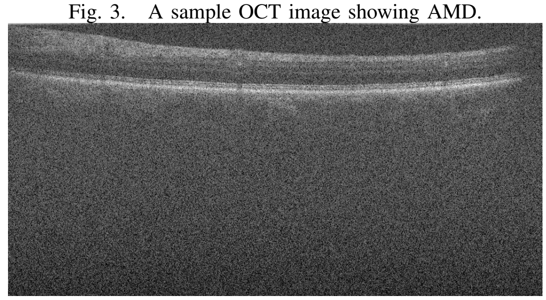

Fig. 4. A sample OCT image of a normal eye

The VGG19 model has been successfully implemented in order to understand lower level feature patterns of images, ones that humans may not necessarily understand [9]. This basic model was used in order to recognize the low to higher level features of OCT images, to classify them into two categories: AMD and normal. The data used was from SD-OCT Dataset [3].

C. Retraining

The model was based off of the pretrained VGG19 model, where the top four layers were unfrozen and retrained with the training set. In order to retrain the model, the image size was reduced by four times from 512x1000 to 128x250, using the nearest neighbour method. This reduction was made as the original VGG19 model was trained off of an image sized at 224x224, so a similarly sized image would allow for the lower-level features to be extracted with more accuracy. In addition to that, if the image is too large, the number of parameters would be too large and lead to poor results due

```
Layer (type)                 Output Shape              Param #
=================================================================
vgg19 (Model)                (None, 4, 7, 512)         20024384
_________________________________________________________________
flatten_3 (Flatten)          (None, 14336)             0
_________________________________________________________________
dense_5 (Dense)              (None, 1024)              14681088
_________________________________________________________________
dropout_3 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_6 (Dense)              (None, 2)                 2050
=================================================================
Total params: 34,707,522
Trainable params: 21,762,562
Non-trainable params: 12,944,960
_________________________________________________________________
```

Fig. 5. Retraining model of the VGG19, where the top 3 layers are unfrozen and trained on the OCT images. The number of trainable and frozen parameters can be seen.

Fig. 6. Training progress of the CNN over a period of 10 epochs, showing the training accuracy, validation accuracy, training loss, and validation loss. It can be seen that both the training and validation loss drops over the training period and the training accuracy continually increases towards 100%, while the validation accuracy increases and begins to plateau around 80% accuracy.

	Classified as AMD	Classified as Control
AMD	74.1%	25.9%
Control	14.3%	85.7%
Validation Accuracy	79.9%	
Training Accuracy	95.1%	
Training Validation Accuracy	88.6%	

Fig. 7. The confusion matrix and training, validation, and training validation accuracy for the CNN, showing its practical performance as an OCT image classifier.

to the lack of data for the large number of parameters. The training set consisted of 5460 AMD and 5820 control images, and was validated during training with 3000 AMD and 1200 control images. The model was trained over a period of 10 epochs, where the total number of parameters in the model that needed to be fine-tuned was approximately 22 million. Due to the limited amount of data, the entire model was not retrained. After the model was trained, the CNN was validated on a separate dataset that consisted of 1164 AMD, 1482 control images. By validating on a completely separate set of images, the classifier's practical accuracy and overfitting can be evaluated.

III. RESULTS

The training accuracy of the CNN can be seen in Figure 6- as the model continues to train and be fine-tuned, the training accuracy increases greatly- up to 95.1%, and the training validation continues to rise as well, up to 88.6%. Once tested on the external validation set, the accuracy is shown to be 79.9%, which indicates some overfitting of the model. This was to be expected, however, as the data set was limited and a larger number of images would contribute to a finer tuning of the model. As indicative of a good model, the loss also continues to drop throughout the training period, and would be expected to have a more exponential shape if more layers of the model were to be trained with a larger dataset. As can be seen in Figure 7, the confusion matrix for the model and its accuracy can be seen. The model classified control images at a higher rate than AMD images, at a rate of 85.7% compared to 74.1%. It is more important, however, to minimize false positives, and this is a limitation of the current model.

IV. CONCLUSIONS

The VGG-19 convolutional neural network algorithm was successfully applied to the dataset used in this study. The algorithm used only the pre-training data used to optimize the algorithm for general use, as well as the subset of training OCT images selected for this study. Because this procedure was able to correctly classify 85.7% of AMD images and 74.1% of control images, it represents in its current form a useful classification method for any given OCT retinal image. However, this study used each image independently, rather than analyzing the set of 100 images associated with each patient and making a further classification. Further algorithm development and analysis is necessary to diagnose patients from their associated series of OCT images rather than identifying AMD in individual images. In addition, regions of the eye in view at varying depths within the 100 images for each patient may impact the model. In particular, the fovea is in sharpest focus at the fiftieth image for each patient. Other patterns may emerge in the distribution of these image sets for individual patients. A more complex neural network design that compares entire sets of images when classifying AMD may provide far superior results than the independent analysis method used in this study.

Another challenge in training the model during this study was the extreme training time due to lack of high processing power- on a MacBook Pro, it took about 12 hours to train 10 epochs, limiting the ability to further fine-tune the model to the desired accuracy due to time constraints. Dedicated systems used in a clinical setting, however, would not be subject to such restrictions and could process an equivalent quantity of data quickly.

Other CNN pretrained models may function better for minimizing the number of false positives classifications, however it can be seen that the VGG19 model is an adequate basis for using CNNs to classify OCT images.

REFERENCES

[1] A. Chopdar, "Age related macular degeneration," *Bmj*, vol. 326, p. 485–488, Jan 2003.

[2] M. Haddrill and C. Slonim, "A guide to age-related macular degeneration," Aug 2018.

[3] S. Farsiu, S. J. Chiu, R. V. Oconnell, F. A. Folgar, E. Yuan, J. A. Izatt, and C. A. Toth, "Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography," *Ophthalmology*, vol. 121, no. 1, p. 162–172, 2014.

[4] J. Cafasso, "Drusen in eyes: Causes, treatment, and macular degeneration," Jan 2019.

[5] S. Horton and C. Guly, "Prevention and treatment of age-related macular degeneration," *Prescriber*, vol. 28, no. 1, p. 37–41, 2017.

[6] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169–208, 2010.

[7] O. Oyedotun and A. Khashman, "Iris nevus diagnosis: convolutional neural network and deep belief network," *Turkish Journal Of Electrical Engineering Computer Sciences*, vol. 25, p. 1106–1115, Apr 2017.

[8] A. Pai, "Cnn vs. rnn vs. ann - analyzing 3 types of neural networks," Mar 2020.

[9] Y. Zheng, C. Yang, and A. Merkulov, "Breast cancer screening using convolutional neural network and follow-up digital mammography," p. 4, 05 2018.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014.

[11] ImageNet, "Large scale visual recognition challenge 2014 (ilsvrc2014)," 2014.