Automated Age-Related Macular Degeneration and Diabetic Macular Edema Detection on OCT Images using Deep Learning

Abstract—**Age-related macular degeneration (AMD) is an eye disease that damages the retina, causing vision loss. Diabetic macular edema (DME) is also a form of vision loss for diabetic people. It is therefore crucial to detect AMD and DME in the early stages for the timely treatment of the eye and the prevention of any vision impairment. Automatic detection of DME and AMD on optical coherence tomography (OCT) images are presented in this paper. The method used is based on training a deep learning algorithm to classify them into healthy, dry AMD, wet AMD and DME categories. This method outperforms a transfer learning based method proposed recently in the literature for classification of OCT images into AMD and DME categories.**

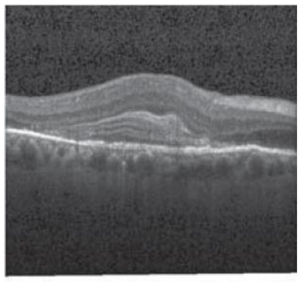
I. Introduction

Age-related macular degeneration (AMD) occurs as a result of the degeneration of the macula and is the prime reason of vision loss for people over 50. Diabetic macular edema (DME) is another eye condition that affects people with diabetes and may also cause blindness. AMD has three stages, namely early, intermediate and late. Dry and wet forms of AMD are the two types of late stage AMD. Dry form of AMD occurs as a result of the degeneration of the retina, whereas wet form of AMD is characterized by the degeneration of retinal blood vessels. Treatment of dry form might not be possible, however, there are different treatments for the wet form of AMD.

Diagnosis of AMD can be done by an eye doctor using either fundus or OCT images. These images can be acquired using different devices, such as a retinal fundus camera. Figure 1 shows two samples of captured fundus and OCT images.

Automated analysis of DME and AMD can be done using deep learning approaches. Deep learning methods have mostly been used for high classification and object detection of images. One of the most well known deep convolutional neural network is AlexNet [8]. This network was used to provide the best classification results on ImageNet 2012. Later, several deep convolutional neural networks have also been introduced, such as GoogLeNet [14], [15], VGG [12], RestNet [17] and DenseNet [6]. These networks contain more layers than AlexNet. As the number of layers are higher than AlexNet, they perform better when many images are available for training.

(a) Fundus Image



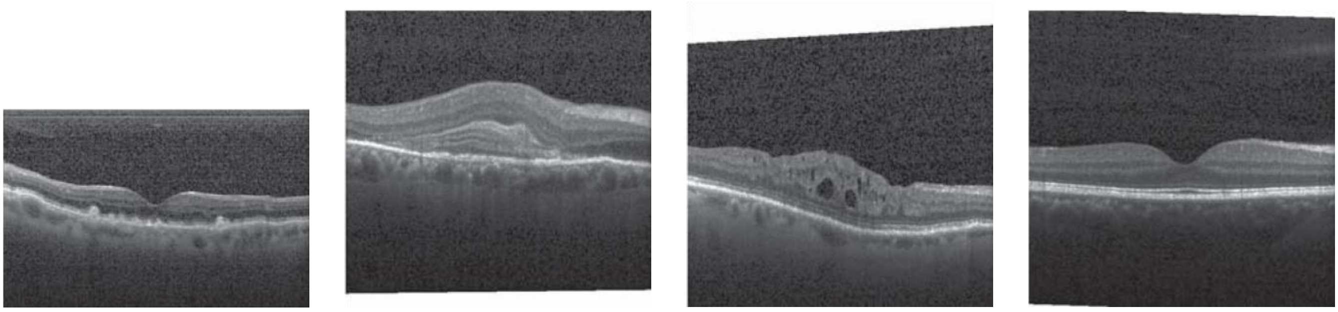
(b) OCT Image

Fig. 1: Fundus and OCT Images.

Automated detection of DME and AMD using deep learning methods have several advantages. One of the key advantages is that the disease can be detected at the early stages. With a carefully designed deep learning algorithm, small symptoms might be detected accurately even when they are unnoticed by an eye doctor during a check-up. Hence, the eye doctor can start the related treatment before the advancement of the disease.

The aim of this paper is to classify OCT images into healthy, dry AMD, wet AMD and DME types. An image of each type is shown in Figure 2. The method used here is based on training the AlexNet architecture and is different than the transfer learning based method of [7]. In [7], pre-trained network parameters on ImageNet dataset were used to fine tune the new network parameters for classification. However, AlexNet deep learning architecture is used here to train all network parameters using only OCT images.

The organization of this paper is as follows. First, related work on disease classification using fundus as well as OCT images are given. Then, the details of the deep learning method as well as the image database used for model generation and performance evaluation are explained. Finally, the performance of this method for four classes as well as various two classes of classifications are given and discussed.



(a) Dry AMD (b) Wet AMD (c) DME (d) Healthy

Fig. 2: Dry AMD, wet AMD, DME and healthy categories of captured images.

II. Related Work

Several studies researched automated detection of AMD on fundus images. Priya and Aruna [11] proposed automated detection of dry and wet forms of AMD using color fundus images. Their approach was based on extracting the features of images and using theses features as input to the probabilistic neural networks. Zheng [20] used an image-based classification method to show the existence of AMD on fundus images. Mookiah et. al. [10] proposed an automated detection of AMD using grayscale features of fundus images. These features were classified using several classifiers such as Naive Bayes, k-Nearest Neighbour, probabilistic neural networks, decision trees and support vector machines. Arabi [1] used the percentage of white pixels to total number of pixels in the fundus eye images in order to automatically detect dry and wet macular degeneration.

Deep learning methods for automated detection of AMD on fundus images were also proposed. Burlina [2], [3] used deep convolutional neural networks and their features in AMD classification of fundus images. Ting et. al. [18] setup a deep learning system to classify images into referable diabetic retinopathy (DR), vision-threatening DR, glaucoma and AMD. Govindaiah et. al. [5] proposed a deep learning method for AMD detection on fundus images. Their method allowed classification of four AMD stages, which were no AMD, early stage, intermediate stage and late stage AMD, using a modified VGG16 neural network architecture [12]. This study showed that training fundus images using this network outperformed the transfer learning based neural network. Tan et. al. [16] proposed the detection of AMD using fourteen layer deep convolutional neural networks. This layered network allowed the classification of healthy as well as dry and wet AMD fundus images.

Several other methods for automated detection of AMD using OCT images were also researched. A study by Srinivasan et. al. [13] proposed an automated classification of ME and dry AMD on OCT images. The approach was based on extracting multi-scale histogram of oriented gradient descriptors and using them as inputs to the support vector machines. Deng et. al. [4] also suggested automated AMD detection on OCT

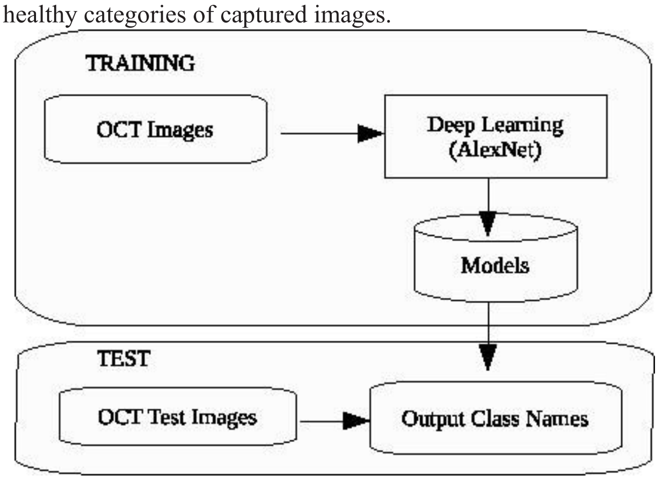


Fig. 3: Overview of the deep learning method used.

images. Their method was based on extracting certain features and inputting theses features into the classifiers. These features were extracted using a Gabour filter bank and non-linear energy transformation. The classifiers were random forests, support vector machines and neural networks. Their proposed method allowed the classification of healthy, dry and wet forms of AMD. Wang [19] used a multi-class model and linear configuration patterns in order to extract and select features of OCT images for the classification of AMD, DME and healthy macula.

Few deep learning based methods for the automated detection of AMD on OCT images were also used. Lee [9] utilized a modified VGG16 neural network [12] to classify healthy and AMD diagnosed OCT images. A recent study by Kermany et. al. [7] also made use of deep learning for classification of images. The new network was generated by adapting the existing network parameters to their OCT image database. They used network parameters which were generated using ImageNet database and showed the performance results for four AMD class identification.

III. Method

The method used here is based on a deep learning algorithm. Overview of this method is given in Figure 3. First, a set of images is used to train the AlexNet architecture [8]. This architecture is then used to model the OCT images for classification into healthy, dry AMD, wet AMD and DME types.

Finally, generated deep learning model is used to evaluate the system performance. A new set of OCT images is used during this evaluation.

AlexNet architecture contains eight layers and its parameter settings are as follows. First five layers are convolutional and the last three layers are fully connected layers. The first convolutional layer takes a 224×224×3 color image as input and convolves it using 96 kernels (11×11×3). The second convolutional layer takes the output of the first layer and convolves it using 256 kernels (5×5×48). Similarly, the third convolutional layer obtains the output of the second layer and convolves it using 384 kernels (3×3×192). The fourth convolutional layer applies 384 kernels to the output of the third layer. The fifth convolutional layer takes the output of the fourth convolutional layer and convolves it using 256 kernels (3×3×192). Finally, fully convolutional layers are obtained, with each fully connected layer containing 4090 neurons.

IV. Implementation

Original AlexNet architecture is used during training. AlexNet deep learning framework is trained using a single NVIDIA GeForce GTX 1080Ti GPU with Caffe deep learning framework. Network parameters used during training are as follows: learning rate is 0.01, gamma is 0.1, momentum is 0.9 and weight decay is 0.0005. Model generation is achieved using 800 epochs.

V. Database

Public image database provided by Kermany et. al. [7] is used for performance evaluation. This database contains 83484 OCT images. These RGB images have different resolutions and are all resized to 256×256×3.Theses images are categorized as healthy (normal in [7]), dry AMD (drusen in [7]), wet AMD (choroidal neovascularization in [7]) and DME. Total number of images in each category are 26315, 8616, 37205 and 11348, respectively. This database also provides 1000 additional images from which 250 images are used during the testing of each category. Table I lists the number of training and testing images for healthy, AMD and DME classes.

TABLE I: Number of OCT images used.

	Train	**Test**
Wet AMD	37205	250
Dry AMD	8616	250
DME	11348	250
Healthy	26315	250
Total	83484	1000

VI. Performance Evaluation

Deep learning based method is evaluated using various classification tasks. First, its performance is evaluated for classifying images into healthy, dry AMD, wet AMD and DME categories. Second, the performance is evaluated for classifying images into healthy and wet AMD categories. Third, the performance is evaluated for classifying images into healthy and dry AMD categories. Finally, it is evaluated for classifying images into healthy and DME categories. AlexNet deep learning architecture with the aforementioned network parameters is used for all of the classification tasks. It is trained using the training data of each class and tested on the available test data.

A. Multiclass Classification

First, classification performance is evaluated for healthy, dry AMD, wet AMD and DME image categories. AlexNet architecture is trained using 83484 OCT images. These consist of 26315 healthy images, 8616 dry AMD images, 37205 wet AMD images and 11348 DME images. The performance is tested using 250 images of each category.

Here, the accuracy of the AlexNet model is **97.1%**, sensitivity is **99.6%** and specificity is **98.4%**. Confusion matrix of this classification is given in Figure 4. These results are better than those produced by the transfer learning method of Kermany et. al. [7]. Note that Kermany et. al. [7] used two different model generation strategies. First, a large database was used to generate a transfer learning based model. For this model, performance accuracy was **96.6%**, sensitivity was **97.8%** and specificity was **97.4%**. Second, limited data was used to generate another transfer learning based model. For this model, performance accuracy was **93.4%**, sensitivity was **96.6%** and specificity was **94.0%**. These results are all reported in Table II.

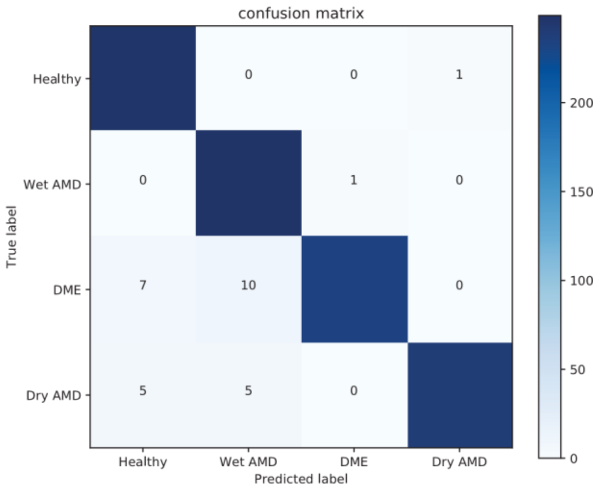
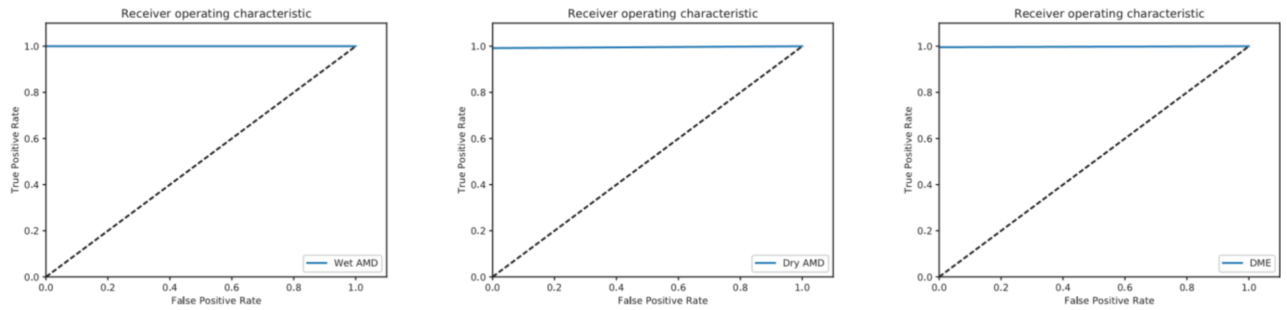


Fig. 4: Confusion matrix of multiclass classification.



(a) Healthy and Wet AMD (b) Healthy and Dry AMD (c) Healthy and DME

Fig. 5: Receiver operating characteristic (ROC) curves for two-class categorizations.

TABLE II: Multiclass classification performance.

Method	Accuracy	Sensitivity	Specificity
Kermany (large)	96.6%	97.8%	97.4%
Kermany (limited)	93.4%	96.6%	94.0%
Our model	**97.1%**	**99.6%**	**98.4%**

B. Healthy and Wet AMD Classification

Classification performance is also evaluated for healthy and wet AMD image categories. AlexNet architecture is trained using 26315 healthy OCT images and 37205 wet AMD OCT images. The performance is tested using 250 healthy and 250 wet AMD OCT images.

Here, the accuracy of the AlexNet model is **100%**, sensitivity is **100%** and specificity is **100%**. These results are exactly the same as those provided by Kermany et. al. [7] (see Table III). The ROC curve of this classification is given in Figure 5(a) and the confusion matrix is given in Figure 6(a).

TABLE III: Healthy and wet AMD classification performance.

Method	Accuracy	Sensitivity	Specificity
Kermany et. al.	100%	100%	100%
Our model	**100%**	**100%**	**100%**

C. Healthy and Dry AMD Classification

In this task, deep learning performance on OCT images is evaluated for classifying images into healthy and dry AMD categories. AlexNet architecture is trained using 26315 healthy OCT images and 8616 dry AMD OCT images. Testing is done using 250 healthy OCT images and 250 dry AMD OCT images.

Here, the accuracy of the AlexNet model is **99.6%**, sensitivity is **100%** and specificity is **99.2%**. Again, these results are equal to or better than those of Kermany et. al. [7]. Results comparison is given in Table IV. The ROC curve of this classification is given in Figure 5(b) and the confusion matrix is given in Figure 6(b).

TABLE IV: Healthy and dry AMD classification performance.

Method	Accuracy	Sensitivity	Specificity
Kermany et. al.	99.0%	98.0%	99.2%
Our model	**99.6%**	**100%**	**99.2%**

D. Healthy and DME Classification

In this final task, deep learning performance is evaluated for healthy and DME image classification. AlexNet architecture is trained using 26315 healthy OCT images and 11348 DME OCT images. Again, the performance is tested using 250 healthy and 250 DME OCT images.

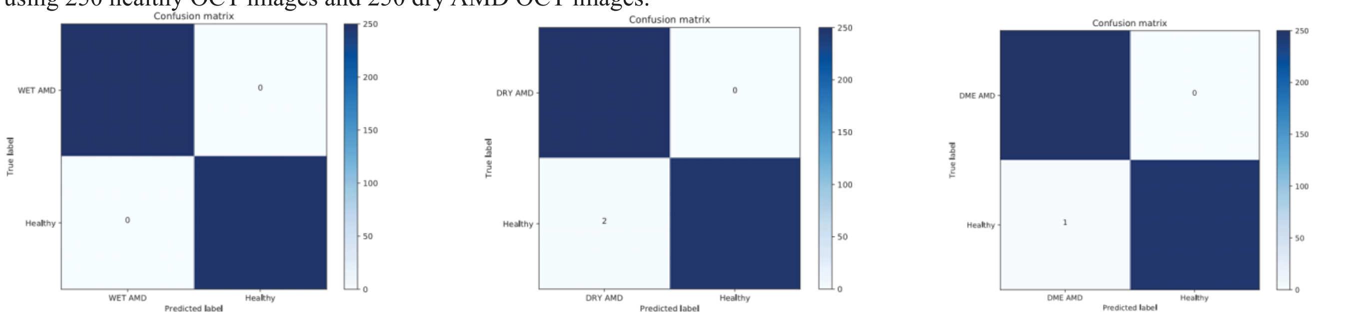
Accuracy of the AlexNet model used here is **99.8%**, sensitivity is **100%** and specificity is **99.6%**. These results are equal to or better than those of Kermany et. al. [7] (see Table V). The ROC curve of this classification is given in Figure 5(c) and the confusion matrix is given in Figure 6(c).

TABLE V: Healthy and DME classification performance.

Method	Accuracy	Sensitivity	Specificity
Kermany et. al.	98.2%	96.8%	99.6%
Our model	**99.8%**	**100%**	**99.6%**

VII. Discussions

Automated detection of AMD and DME is performed on OCT



(a) Healthy and Wet AMD (b) Healthy and Dry AMD (c) Healthy and DME

Fig. 6: Confusion matrices of two-class categorizations.

images using a deep learning method. Results show that training the network using only OCT images provides better performance than a transfer learning based deep network. In two-class categorizations, best performances are obtained from the classification of healthy and wet AMD images and the classification of healthy and DME images. High accuracy of these two classification tasks can be explained by the high number of available training images for parameter estimation of the deep neural network. The performance of healthy and dry AMD classification task might also be increased by increasing the number of training images. Still, AlexNet based classification method performs equal to or better in all four classification tasks than the previously proposed transfer learning based method.

VIII. Conclusions

A deep learning based method, AlexNet, is used for the automated detection of AMD and DME on OCT images. This method allows the classification of dry and wet forms of AMD as well as DME. Its performance is evaluated using various classification tasks. First, it is evaluated for classification into healthy, dry AMD, wet AMD and DME categories. Then, it is evaluated for healthy and wet AMD classification. Third, it is evaluated for classification into healthy and dry AMD categories. Finally, it is evaluated for healthy and DME classification. Results show that the method used here outperforms the transfer learning based method of recent literature for the classification of AMD and DME diagnosed OCT images.

References

[1] P. M. Arabi, V. Deepa, T. S. Naveen, and D. Samanta. Machine vision for screening of age-related macular degeneration using fundus images. In *8th International Conference on Computing, Communication and Networking Technologies*, pages 1–4, July 2017.

[2] P. Burlina, D. E. Freund, N. Joshi, Y. Wolfson, and N. M. Bressler. Detection of age-related macular degeneration via deep learning. In *IEEE 13th International Symposium on Biomedical Imaging*, pages 184–188, April 2016.

[3] P. Burlina, K. D. Pacheco, N. Joshi, D. E. Freund, and N. M. Bressler. Comparing humans and deep learning performance for grading amd: A study in using universal deep features and transfer learning for automated amd analysis. *Computers in Biology and Medicine*, 82:80–86, 2017.

[4] J. Deng, X. Xie, L. Terry, A. Wood, N. White, T. H. Margrain, and R. V. North. Age-related macular degeneration detection and stage classification using choroidal OCT images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9730:707–715, 2016.

[5] A. Govindaiah, M. A. Hussain, R. T. Smith, and A. Bhuiyan. Deep convolutional neural network based screening and assessment of agerelated macular degeneration from fundus images. In *IEEE 15th International Symposium on Biomedical Imaging*, pages 1525–1528, April 2018.

[6] G. Huang, Z. Liu, and K. Q. Weinberger. Densely connected convolutional networks. *CoRR*, abs/1608.06993, 2016.

[7] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, J. Dong, M. K. Prasadha, J. Pei, M. Ting, J. Zhu, C. Li, S. Hewett, J. Dong, I. Ziyar, A. Shi, R. Zhang, L. Zheng, R. Hou, W. Shi, X. Fu, Y. Duan, V. A. Huu, C. Wen, E. D. Zhang, C. L. Zhang, O. Li, X. Wang, M. A. Singer, X. Sun, J. Xu, A. Tafreshi, M. A. Lewis, H. Xia, and K. Zhang. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, 172(5):1122–1124.e9, 2018.