



Published in final edited form as:

Appl AI Lett. 2020 October ; 1(1): . doi:10.1002/ail2.16.

Visualizing and understanding inherent features in SD-OCT for the progression of age-related macular degeneration using deconvolutional neural networks

Sajib Saha^{1,2}, Ziyuan Wang^{1,3}, Srinivas Satta^{1,3}, Yogesan Kanagasiam², Zhihong Hu¹

¹Doheny Eye Institute, Los Angeles, California

²Australian e-Health Research Centre, CSIRO, Perth, Australia

³The University of California, Los Angeles, California

Abstract

To develop a convolutional neural network visualization strategy so that optical coherence tomography (OCT) features contributing to the evolution of age-related macular degeneration (AMD) can be better determined. We have trained a U-Net model to utilize baseline OCT to predict the progression of geographic atrophy (GA), a late stage manifestation of AMD. We have augmented the U-Net architecture by attaching deconvolutional neural networks (deconvnets). Deconvnets produce the reconstructed feature maps and provide an indication regarding the inherent baseline OCT features contributing to GA progression. Experiments were conducted on longitudinal spectral domain (SD)-OCT and fundus autofluorescence images collected from 70 eyes with GA. The intensity of Bruch's membrane-outer choroid (BMChoroid) retinal junction exhibited a relative importance of 24%, in the GA progression. The intensity of the inner retinal pigment epithelium (RPE) and BM junction (InRPEBM) showed a relative importance of 22%. BMChoroid (where the AMD feature/damage of choriocapillaris was included) followed by InRPEBM (where the AMD feature/damage of RPE was included) are the layers which appear to be most relevant in predicting the progression of AMD.

Keywords

age-related macular degeneration; AMD; CNN; convolutional neural network; deconvnet; deconvolutional neural network; FAF; fundus autofluorescence; SD-OCT

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Correspondence : Zhihong Hu, Doheny Eye Institute, 1355 San Pablo Street, Los Angeles, CA 90033, jhu@doheny.org.

AUTHOR CONTRIBUTIONS

Sajib Saha developed the system, performed experiments, and wrote the main manuscript text. Ziyuan Wang generated en face images and performed registration of the images. Yogi Kanagasiam and Srinivas Satta wrote part of the manuscript. Zhihong Jewel Hu cultivated the idea of the development, and also wrote part of the manuscript.

CONFLICT OF INTEREST

The authors declared no potential conflicts of interest.

This work was designed while S. Saha was with the Doheny Eye Institute, Los Angeles, CA 90033. He is currently with Australian e-Health Research Centre, CSIRO, Perth, Australia.

1. | INTRODUCTION

Age-related macular degeneration (AMD) is the leading cause of severe vision impairment among elderly individuals in the developed world. One in seven, who are over 50 years old, experiences a worsening of vision due to AMD that may affect one or both eyes. The incidence of AMD increases with age.¹ It is anticipated that about 8 million people in the United States who are 55 years of age and older are suffering vision impairment due to monocular or binocular intermediate AMD or monocular advanced AMD.² Advanced-stage AMD is manifest by the development of geographic atrophy (GA) or macular neovascularization (MNV) and may be associated with severe visual loss or blindness. The likelihood of progression to advanced AMD over 5 years period is 27% for patients with intermediate AMD.³ For patients, who already have advanced AMD in one eye, the chance that the other eye progresses to advanced AMD can be as high as 43%.³ Although there are now effective treatments for patients with neovascular AMD, these patients commonly progress to develop atrophy over time, and as result ultimately go on to lose vision. No proven treatment is presently available in the setting of non-neovascular disease to stop the development and progression of atrophy. Development of better treatment approaches, however, requires a more comprehensive understanding of crucial biomarkers contributing to the development and progression of AMD.

Historically, color fundus photography was the gold standard to diagnose early AMD. Large drusen, increased total drusen area, hyperpigmentation, and depigmentation are some of the important biomarkers identified from color fundus photographs,^{3,4} to determine both the presence AMD and its risk of progression. Fundus autofluorescence (FAF) imaging, which is another noninvasive, 2D imaging technique of metabolic mapping of naturally or pathologically occurring fluorophores of the ocular fundus also gained popularity in identification of atrophic lesions.⁵ In recent years, optical coherence tomography (OCT) has largely supplanted color fundus photography and FAF imaging as the primary imaging modality in clinical practice. OCT produces 3D cross-sectional imaging of the tissue and provides detailed anatomic information on the axial location of retinal abnormalities, which color fundus photography and FAF imaging cannot provide.

A number of studies have identified several OCT-based features to determine the level of AMD and its progression.⁶ Higher central drusen volume,⁷ intraretinal hyperreflective foci,⁸ heterogeneous internal reflectivity within drusenoid lesions,⁹ and reticular pseudodrusen or subretinal drusenoid deposits,¹⁰ are some of the most promising OCT based features that are found to be associated with a higher risk for progression to advanced AMD.⁷

A number of approaches have already been proposed for automated and semiautomated analysis of AMD-related features on OCT, including drusen,¹¹ GA,^{12,13} pigment epithelial detachment,¹⁴ and intra-/subretinal fluid.^{15,16} Aside from Sisternes et al,¹⁷ most of these methods focused on detecting and/or analyzing AMD features from a single time point. Sisternes et al focused on analyzing the likelihood of progression from early and intermediate AMD to exudative AMD and utilized 11 drusen specific features for making this determination.

Recently, there has been an increasing interest in applying deep learning based techniques for automated diagnosis of eye disease including AMD.^{18,19} In comparison to traditional machine learning approaches that still rely on handcrafted features to analyze images to detect pathologies, deep learning learns these features from the data itself. Deep learning is capable of characterizing hidden complex features by applying neural network architectures with multiple hidden layers. Deep convolutional neural networks (CNN) exhibit feature hierarchy and learn features of increasing complexity and abstraction with increasing network depth. While CNN-based models have achieved state-of-the-art performance in several AMD assessment tasks, they are often criticized for their “black box” nature. In particular, they may be difficult to interpret and it is not inherently obvious how they reach their final predictions. In the context of AMD, no attempt has been made so far to visualize the inherent patterns that contribute to a certain decision by the CNN and to justify its clinical relevance. On that perspective, as a preliminary study, this paper reports a deep learning system designed to better visualize and understand the intrinsic features in the data that contributes to the development and progression of AMD, particularly progression of GA. To justify the relevance of these features in disease evolution, extensive experiments are conducted on a longitudinal dataset of progressive AMD cases.

2. | BACKGROUND

2.1. | CNN visualization

Visualization of image features that contribute to CNN decision is of significant interest to gain intuition about why and how a certain decision is made. While a number of visualization methods are available, the majority of them are limited to visualize image features only at the initial layer where projections to pixel space are plausible.²⁰ Only a few methods are capable to effectively visualize the contributing image features at the different layers of the CNN. Among this short list of methods, the method proposed by Erhan et al²¹ finds the stimulus that maximizes the unit’s activation. An optimal stimulus is determined for each unit. The method explicitly performs gradient descent operation in image space, which requires vigilant initialization. Another drawback of the method is that it does not provide any insight about the unit’s invariances. To resolve the shortcomings of Erhan et al’s method, Ngiam et al²² focused on computing the Hessian of a given unit around the optimal response to provide a parametric view of unit’s invariances. The method works well for lower layers, however, for higher layers, the invariances are very multifaceted and are thus poorly apprehended by simple quadratic approximation. Simonyan et al²³ proposed two techniques to visualize CNN. The first method creates an artificial image, which maximizes the class score and visualizes the notion of the class. The second method produces a saliency map, specific to a given image and class. Girshick et al²⁴ produces visualizations by identifying patches that are accountable for strong activations at the higher layers of the model. The method proposed by Zeiler et al²⁰ visualizes which patterns in the training set activates the feature map. It also provides a non-parametric view of invariance. While the method is similar to Simonyan et al, it provides better insight into the contributing patterns. In contrast to Girshick et al,²⁴ the visualizations produced by Zeiler et al²⁰ are not just crops of input images, but rather top-down estimates that disclose structures inside each patch that excite a particular feature map.

2.2. | U-Net

U-Net²⁵ is a state-of-the-art deep learning architecture for semantic segmentation. U-Net can be trained end-to-end from very few images. Structure-wise U-Net uses a contracting path of convolutional blocks (encoders), followed by an expansive path of convolutional blocks (decoders). The encoders learn the contextual features and their hierarchy. The decoders perform semantic segmentation. In the original U-Net, the classical stochastic gradient descent optimization approach was used to interactively update network weights. However, in this work, we used a more efficient adaptive moment estimation (Adam) optimizer.⁶ In this work, we optimized a joint loss function²⁶ which consists of weighted logistic regression loss and Dice overlap loss, in comparison to the original weighted loss. Each encoder block of U-net consists of four key layers, that include a convolution layer, batch normalization layer, ReLU activation layer and a max pooling layer. The size of the convolution kernels in this work is kept static across encoder blocks which are defined to be 7×7 . The feature maps are appropriately zero padded so that the dimension before and after the convolution layer remains the same. Each decoder block consists of five key layers, that include an unpooling layer, concatenation layer, convolution layer, batch normalization layer and a ReLU activation layer.

3. | METHODOLOGY

3.1. | Generation of en face images

In non-exudative AMD, the areas affected by GA as evident in 3D SD-OCT images are characterized by loss of the photoreceptor layer, retinal pigment epithelium (RPE) layer, and choriocapillaris layer. The involvement of multiple layers of the retina make 2D en face images generated from the 3D layers a useful tool for the study of GA.²⁷

More specifically, an in-house developed automated 3D graph-based program²⁸ was used to segment the 3D OCT retinal layers and to generate the 2D OCT en face images between two retinal surfaces L_i and L_j potentially affected by AMD by calculating the mean pixel intensities among them. A set of seven en face images were generated taking into account different depth profile. Among them, 4 en face images were generated from retinal layers extended from inner retina to outer retina. The four en face images included IPLINLELM en face image generated between the retinal surface located in inner plexiform layer (IPL) and inner nuclear layer (INL) junction, and the retinal surface of external limiting membrane (ELM); the EMLInRPE en face image (where the AMD feature/damage of photoreceptor was included) generated between the retinal surface of EML, and the retinal surface of inner RPE; the InRPEBM en face image (where the AMD feature/damage of RPE was included) generated between the retinal surface of inner RPE and the retinal surface of Bruch's membrane (BM); and the BMChoroid en face image (where the AMD feature/damage of choriocapillaris was included) generated between the retinal surface of BM and the retinal surface of outer choroid. Additionally, three en face images with different division of the photoreceptor and RPE retinal layers were also included. The three en face images were the ELMISOS en face image generated between the retinal surface of ELM and the retinal surface located in the photoreceptor inner segment and outer segments junction; the ELMOutRPE en face image generated between the retinal surface of ELM and the retinal

surface of outer PRE; as well as the ELMBM en face image generated between the retinal surface of ELM and the retinal surface of BM.

The resultant en face images had a resolution of 512 pixels horizontally and 128 pixels vertically; therefore, each en face image was stretched vertically by a factor 4 to obtain a square image of 512×512 pixels.

Figure 1 shows some example en face images. From Figure 1, it is observable that some vessel artefacts in the en face images are highly likely. InRPEBM, the apparent artefacts are because of the thin retinal layer due to severe AMD, and the presence of retinal vessel shadows in the layer. For BMChoroid, the artefacts are due to the choroid vessels present in the layer.

3.2. | Ensembled CNN for pathology segmentation

Figure 1 shows the flowchart of the ensembled CNN architecture used to predict the GA progression based on baseline en face OCT images. Each of the CNNs here is an augment U-Net²⁵ which has square instead of rectangular sized filters and has a visualization block attached. FAF imaging has served as the gold standard modality for measuring and monitoring the enlargement of GA lesions due its high contrast. The training of the CNNs utilizes the seven baseline OCT en face images as input images and the corresponding manually segmented GA regions from FAF images at a future visit as the ground truth data.

While training the CNN, we optimized the weighted two-class logistic loss, dice loss and an additional weight decay term for regularization, likewise in²⁶ and defined as bellow:

$$\mathcal{L}_{\text{Overall}} = \lambda_1 \mathcal{L}_{\text{Logloss}} + \lambda_2 \mathcal{L}_{\text{Dice}} + \lambda_3 \|\mathbf{W}(\cdot)\|_F^2.$$

Here, λ_1 , λ_2 , λ_3 are the weight terms, $\|\mathbf{W}(\cdot)\|_F^2$ represents the Frobenius norm (also called Euclidean norm) on the weights \mathbf{W} of the CNN, and $\mathcal{L}_{\text{Logloss}}$ and $\mathcal{L}_{\text{Dice}}$ are, respectively, the weighted multiclass logistic loss and dice loss.

Weighted two-class logistic loss is defined as bellow:

$$\mathcal{L}_{\text{Logloss}} = - \sum_{\mathbf{x} \in \mathcal{S}} \omega(\mathbf{x}) g_l(\mathbf{x}) \log p_l(\mathbf{x}),$$

where $p_l(\mathbf{x})$ is the probability estimates of pixel \mathbf{x} belonging to class 1, $\omega(\mathbf{x})$ is the associated weight, and where $g_l(\mathbf{x})$ is a vector of ones and zeros representing the ground truth probability of pixel at location \mathbf{x} to belong to class 1.

With the same definition of $p_l(\mathbf{x})$ and $g_l(\mathbf{x})$, the dice loss is defined as bellow:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_{\mathbf{x} \in \mathcal{S}} p_l(\mathbf{x}) g_l(\mathbf{x})}{\sum_{\mathbf{x} \in \mathcal{S}} p_l^2(\mathbf{x}) + \sum_{\mathbf{x} \in \mathcal{S}} g_l^2(\mathbf{x})}.$$

A total of seven CNNs were ensembled and each of them were independently trained to segment GA using different en face images. The final segmentation is performed based upon the strongest output of the seven CNNs.

A pixel is classified as “GA” or “non-GA” based on the following equation²⁹:

$$l = \underset{l = 1, 2}{\operatorname{argmax}} O_l$$

$$O_l = \max_{j = 1, 2, \dots, 7} \text{Output}_{lj}; l = 1, 2, \quad (1)$$

where $j = 1, 2, \dots, 7$, represents the different neural networks, $l = 1$ and 2 , respectively, characterizes “non-GA” and “GA” pixel label. Equation (1) indicates that a pixel is classified to “GA” or “non-GA,” based on the strongest response between the two labels from the strongest output of the seven trained CNNs.

3.3. | Visualizing inherent en face features using deconvolutional neural network

To gain an understanding of which image features are influencing the CNN’s predictions for GA progression, we have adopted the visualization strategy proposed by Zeiler et al.²⁰ We attach a deconvolutional network (in short “deconvnet”) to each convnet block of the U-Net architecture (Figure 2). It should be noted that, similar to a convolutional network, a deconvnet also uses filtering, rectification, and pooling, however, in reverse order and thus maps the feature activity back to the input pixel space.

To examine a given convnet activation, we set all other activations in the layer to zero and pass the feature maps as input to the attached deconvnet layer. Then we successively (a) unpool, (b) rectify, and (c) filter, to reconstruct the activity in the layer beneath that gave rise to the chosen activation. This procedure is then repeated until the input pixel space is reached.

4. | EXPERIMENTS AND RESULTS

4.1. | Dataset

A total of 70 eyes that were diagnosed as having AMD with evidence of progressive GA from the Doheny Image Reading Center database were utilized in this study. These eyes did not show any evidence of choroidal neovascularization, or other ocular diseases or atrophy due to disease aside from AMD. For each eye both FAF and spectral domain (SD)-OCT images were available. FAF images were captured using a Heidelberg confocal scanning laser ophthalmoscopy (cSLO) (Spectralis HRA+OCT, Heidelberg Engineering, Heidelberg, Germany). The FAF image resolution varied from 496×596 pixels to 1536×1536 pixels with a physical size of $9 \text{ mm} \times 9 \text{ mm}$. All the images were resized to a consistent dimension. SD-OCT images were captured using a Cirrus HD-OCT camera (Carl

Zeiss Meditec, Dublin, California) with 1024 (depth) 512×128 cube ($2 \times 6 \times 6$ mm) centered on the fovea. Follow-up image were captured at 6 to 12 months intervals.

All images were deidentified according to Health and Insurance Portability and Accountability Act Safe Harbor prior to analysis. Ethics review and institutional review board approval from the University of California - Los Angeles were obtained. The research was performed in accordance with relevant guidelines/regulations, and informed consent was obtained from all participants.

A certified expert Doheny Image Reading Center (DIRC) FAF grader manually delineated all the follow-up FAF images. A labeled GA mask (an image with the GA region highlighted in white and the background of “non-GA” region in black) was created for each FAF image. The labeled GA masks of follow-up FAF images were used to generate GA masks as ground truths for the OCT baseline en face images for the GA progression prediction. A semiautomated software platform was developed to register^{30,31} FAF (as well as label masks) and OCT en face images. An experienced grader individually aligned the FAF image with the different en face images using that platform. The software platform then automatically generated the GA mask for the en face image under consideration. Generating GA masks likewise for the en face images, not only saved time but also produced more robust delineation, since GA has better visibility and contrast in FAF compared with SD-OCT.

4.2. | Experiments

4.2.1. | Training the CNNs and visualizing the inherent image features—Each of the CNN in Figure 2 is independently trained to segment GA using a mutually exclusive set of en face images. In line with Reference 26, in this work, we empirically set $\lambda_1 = 1$, $\lambda_2 = 0.5$, $\lambda_3 = 0.0001$, $\omega_1 = 10$ and $\omega_2 = 5$. Seventy en face images and their corresponding manual annotations of GA were available to train each CNN. En face images of the baseline capture were used, however, GA masks of the latest follow-up visits were used; since, the aim is to identify early imaging features responsible for disease progression. A total of 490 (ie, 70×7) en face images and 490 GA masks were available to train the ensemble. We performed data augmentation, specifically, rotation (in the range of -5° to 5°), shearing (in the range 0.2), scaling (in the range 0.2), and flipping, to increase the number of images by 30 times. Thus, a total of 14 700 en face images were used to in the experiment. At the start of the training, the learning rate was set to 0.1 and was gradually reduced by an order of magnitude at every 30 epochs. The training was performed with a momentum of 0.9. Ninety percent of the images were used for training and rest 10% were used for validation.

For the quantitative evaluation of the segmentation performance, we computed intersection over union (IoU) score, and accuracy of the ensembled CNN. IoU was defined as $\frac{(A \cap B)}{(A \cup B)}$, where A and B are, respectively, the ground truth and CNN produced atrophy region. Accuracy defined as $\frac{(TP + TN)}{(P + N)}$, measured the proportion of the enface image with actual atrophy (positive [P]) and without actual atrophy (negative [N]) which are correctly identified as with atrophy (true positive [TP]) and without atrophy (true negative [TN]). We obtained an average IoU score of 0.76 ± 0.07 and accuracy of 0.78 ± 0.06 .

Once the ensembled CNN is trained, the reconstructed features maps were generated at three different layers as shown in Figure 2, for each input en face image. All of the 490 original en face images (ie, en face images without augmentation) were passed through the network in batch and reconstructed features maps were generated. Figure 3 shows some sample visualizations. For any given feature map, we only show the top-most activation, projecting down to pixel space to reveal the inherent structures that excite a given feature map. In addition, with these visualizations we also show the corresponding input en face images.

4.2.2. | Understanding and determining SD-OCT features responsible for disease progression—To better understand the SD-OCT features related to disease progression, at first, the GA mask from the baseline visit was subtracted by the GA mask from a follow-up visit, and the difference mask was computed. The difference mask was then applied on the reconstructed feature maps of the baseline OCT en face images. Figure 4 shows some sample visualizations of the reconstructed feature maps, showing only the regions of interests defined by difference GA masks.

In order to better understand which SD-OCT layers³² were most relevant to predicting GA progression, for each filter we determined the best three reconstructed feature map differences. There are 64 filters in each layer of the U-Net. Thus, we determined a total of 192 reconstructed feature map differences in each of the three layers of the U-Net (Figure 3). We then grouped those map differences based on en face image types and counted the number of cases within each group. For each layer we computed the relative importance of the seven en face images defined as the ratio of the number of cases a particular en face image category listed in the best three and the total number of cases (ie, 192). Figure 5 summarizes the findings.

From the results, it is observable that BMChoroid followed by InRPEBM31 are the layers contributing the most to the prediction of GA progression. In layer 1 and layer 2, BMChoroid and InRPEBM clearly outnumber other en face images. However, in layer 3, along with BMChoroid and InRPEBM, other en face images specially IPLINLELM85 also contributes meaningfully.

5. | DISCUSSIONS AND CONCLUSION

AMD is the leading cause of severe vision impairment among elderly individuals in the developed world. Early detection is the key in AMD for providing timely treatment and minimizing vision loss. If patients who are more likely to progress or progress rapidly can be identified, they could be targeted for therapeutic trials of novel agents. CNN-based methods have achieved state-of-the-art performance in AMD detection. However, they have primarily considered CNN as a black-box and no attempts have been made to unbox its decision-making. In principle unboxing a CNN has the potential to determining more disease specific pertinent image features. In this study, we have developed a visualization strategy so that the interworking of CNNs for AMD evolution can be visualized and the inherent image features contributing to disease progression can be understood. As a representative of CNN family, we have trained a U-Net model to segment AMD pathology specifically GA, which is one of the best performing semantic segmentation models. We

have amended the visualization strategy of Zeiler et al²⁰ in this context and determined the most contributing retinal layers through visualization of corresponding en face images. We used transfer learning¹⁸ and used the pretrained weights from previous study⁵ by our group. The training took 28 hours to complete on an Intel(R) Core(TM) i7-7700 CPU @ 3.60 GHz machine with 16.0 GB RAM and NVIDIA Quadro P4000 GPU.

We conducted a series of experiments on eyes with AMD and progressive GA obtained from an anonymized Doheny Image Reading Center database. We observed that BMChoroid (where the AMD feature/damage of choriocapillaris was included) followed by InRPEBM (where the AMD feature/damage of RPE was included) are the layers which appear to be most relevant in predicting the progression of AMD.

Future work could involve applying our method to a larger and more diverse set of images. It would also be interesting to determine more specific imaging features responsible for disease evolution.

ACKNOWLEDGMENT

Research reported in this publication was partially supported by the National Eye Institute of the National Institutes of Health under Award Number R21EY030619.

Funding information

National Eye Institute of the National Institutes of Health, Grant/Award Number: R21EY030619

DATA AVAILABILITY STATEMENT

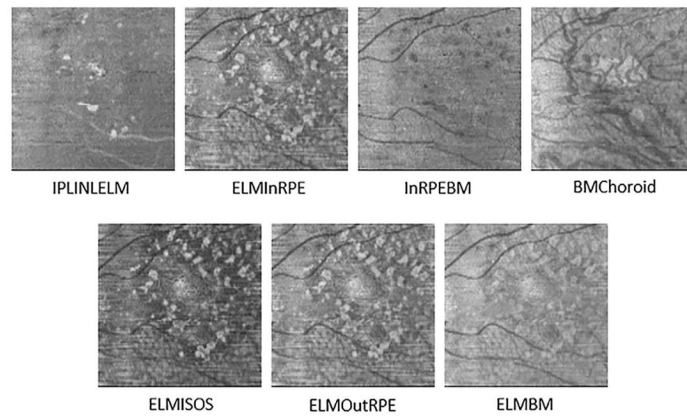
The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

REFERENCES

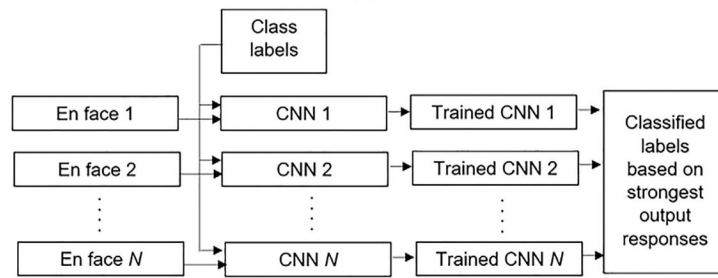
1. VanNewkirk MR, Nanjan MB, Wang JJ, Mitchell P, Taylor HR, McCarty CA. The prevalence of age related Maculopathy: the visual impairment project. *Ophthalmology*. 2000;107:1593–1600. [PubMed: 10919916]
2. Bressler NM, Bressler SB, Congdon NG, et al. Potential public health impact of age-related eye disease study results: AREDS report no. 11. *Arch Ophthalmol*. 2003;121:1621–1624. [PubMed: 14609922]
3. Age-Related Eye Disease Study Research Group. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. *Arch Ophthalmol*. 2001;119:1417–1436. [PubMed: 11594942]
4. Mohaimin SM, Saha SK, Khan AM, Arif AS, Kanagasingam Y. Automated method for the detection and segmentation of drusen in colour fundus image for the diagnosis of age-related macular degeneration. *IET Image Processing*. 2018;12(6):919–927.
5. Wang Z, Sadda SR, Hu Z. Deep learning for automated screening and semantic segmentation of age-related and juvenile atrophic macular degeneration. *Medical Imaging 2019: Computer-Aided Diagnosis*. Vol 10950, San Diego, CA: International Society for Optics and Photonics; 2019:109501Q.

6. Lei J, Balasubramanian S, Abdelfattah NS, Nittala MG, Sadda SVR. Proposal of a simple optical coherence tomography-based scoring system for progression of age-related macular degeneration. *Graefes Arch Clin Exp Ophthalmol*. 2017;255(8):1551–1558. [PubMed: 28534244]
7. Abdelfattah NS, Zhang H, Boyer DS, et al. Drusen volume as a predictor of disease progression in patients with late age-related macular degeneration in the fellow eye. *Invest Ophthalmol Vis Sci*. 2016;57:1839–1846. [PubMed: 27082298]
8. Nassisi M, Fan W, Shi Y, et al. Quantity of intraretinal hyperreflective foci in patients with intermediate age-related macular degeneration correlates with 1-year progression. *Invest Ophthalmol Vis Sci*. 2018;59(8):3431–3439. [PubMed: 30025092]
9. Ouyang Y, Heussen FM, Hariri A, Keane PA, Sadda SVR. Optical coherence tomography-based observation of the natural history of drusenoid lesion in eyes with dry age-related macular degeneration. *Ophthalmology*. 2013;120:2656–2665. [PubMed: 23830761]
10. Finger RP, Wu Z, Luu CD, et al. Reticular pseudodrusen: a risk factor for geographic atrophy in fellow eyes of individuals with unilateral choroidal neovascularization. *Ophthalmology*. 2014;121:1252–1256. [PubMed: 24518615]
11. Jain N, Farsiu S, Khanifar AA, et al. Quantitative comparison of drusen segmented on SD-OCT versus drusen delineated on color fundus photographs. *Invest Ophthalmol Vis Sci*. 2010;51(10):4875–4883. [PubMed: 20393117]
12. Chen Q, de Sisternes L, Leng T, Zheng L, Kutzscher L, Rubin DL. Semi-automatic geographic atrophy segmentation for SD-OCT images. *Biomed Opt Express*. 2013;4:2729–2750. [PubMed: 24409376]
13. Hu ZH, Medioni GG, Hernandez M, Hariri A, Wu X, Sadda SR. Segmentation of the geographic atrophy in spectral-domain optical coherence tomography and fundus autofluorescence images. *Invest Ophthalmol Vis Sci*. 2013;54:8375–8383. [PubMed: 24265015]
14. Penha FM, Rosenfeld PJ, Gregori G, et al. Quantitative imaging of retinal pigment epithelial detachments using spectral-domain optical coherence tomography. *Am J Ophthalmol*. 2012;153:515–523. [PubMed: 22030354]
15. Pilch M, Stieger K, Wenner Y, et al. Automated segmentation of pathological cavities in optical coherence tomography scans. *Invest Ophthalmol Vis Sci*. 2013;54:4385–4393. [PubMed: 23737469]
16. Zheng Y, Sahni J, Campa C, Stangos AN, Raj A, Harding SP. Computerized assessment of intraretinal and subretinal fluid regions in spectral-domain optical coherence tomography images of the retina. *Am J Ophthalmol*. 2013;155:277–286. [PubMed: 23111180]
17. de Sisternes L, Simon N, Tibshirani R, Leng T, Rubin DL. Quantitative SD-OCT imaging biomarkers as indicators of age-related macular degeneration progression. *Invest Ophthalmol Vis Sci*. 2014;55(11):7093–7103. [PubMed: 25301882]
18. Saha S, Nassisi M, Wang M, Lindenberg S, Sadda S, Hu ZJ. Automated detection and classification of early AMD biomarkers using deep learning. *Sci Rep*. 2019;9(1):1–9. [PubMed: 30626917]
19. Saha SK, Xiao D, Fernando B, Tay-Kearney ML, An D, Kanagasingam Y. Deep learning based decision support system for automated diagnosis of age-related macular degeneration (AMD). *Invest Ophthalmol Vis Sci*. 2017;58(8):25.
20. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. *European Conference on Computer Vision*. Cham, Switzerland: Springer; 2014:818–833.
21. Erhan D, Bengio Y, Courville A, Vincent P. Visualizing higher-layer features of a deep network. *University of Montreal*. 2009;1341(3):1.
22. Ngiam J, Chen Z, Chia D, Koh PW, Le QV, Ng AY. Tiled convolutional neural networks. *Advances in Neural Information Processing Systems*; 2010:1279–1287. Vancouver, Canada: Neural Information Processing Systems Foundation, Inc.
23. Simonyan K, Vedaldi A, Zisserman A. Deep inside convolutional networks: visualising image classification models and saliency maps. *arXiv:1312.6034*; 2013.
24. Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proc IEEE Conf Comput Vis Pattern Recognit*. 2014;580–587.

25. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham, Switzerland: Springer; 2015:234–241.
26. Roy AG, Conjeti S, Karri SP, et al. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomed Opt Express. 2017;8(8):3627–3642. [PubMed: 28856040]
27. Heiferman MI, Simonett J, Fawzi A. En face OCT imaging in Retinal disorders. Retin Physician. 2015;12(Jun 2015):45.
28. Hu Z, Wang Z, Sadda SR. Automated choroidal segmentation in spectral optical coherence tomography images with geographic atrophy using multimodal complementary information. J Med Imag. 2019;6(2):024009.
29. Hu Z, Wang Z, Sadda SR. Automated segmentation of geographic atrophy using deep convolutional neural networks. Medical Imaging 2018: Computer-Aided Diagnosis. Vol 10575, Houston, TX: International Society for Optics and Photonics; 2018:1057511.
30. Saha SK, Xiao D, Bhuiyan A, Wong TY, Kanagasingam Y. Color fundus image registration techniques and applications for automated analysis of diabetic retinopathy progression: a review. Biomed Signal Process Control. 2019;47:288–302.
31. Saha SK, Xiao D, Frost S, Kanagasingam Y. A two-step approach for longitudinal registration of retinal images. J Med Syst. 2016;40(12):277. [PubMed: 27787783]
32. Mishra Z, Ganegoda A, Selicha J, Wang Z, Sadda SR, Automated Retinal HZ. Layer segmenation using graph-based algorithm incorporating deep-learning-derived information. Sci Rep. 2020;10(1):1–8. [PubMed: 31913322]



(a)



(b)

FIGURE 1.

A, Sample en face images (generated using different depth profile of the SD-OCT scans) showing geographic atrophy of the same eye. B, Flowchart the ensemble CNN architecture

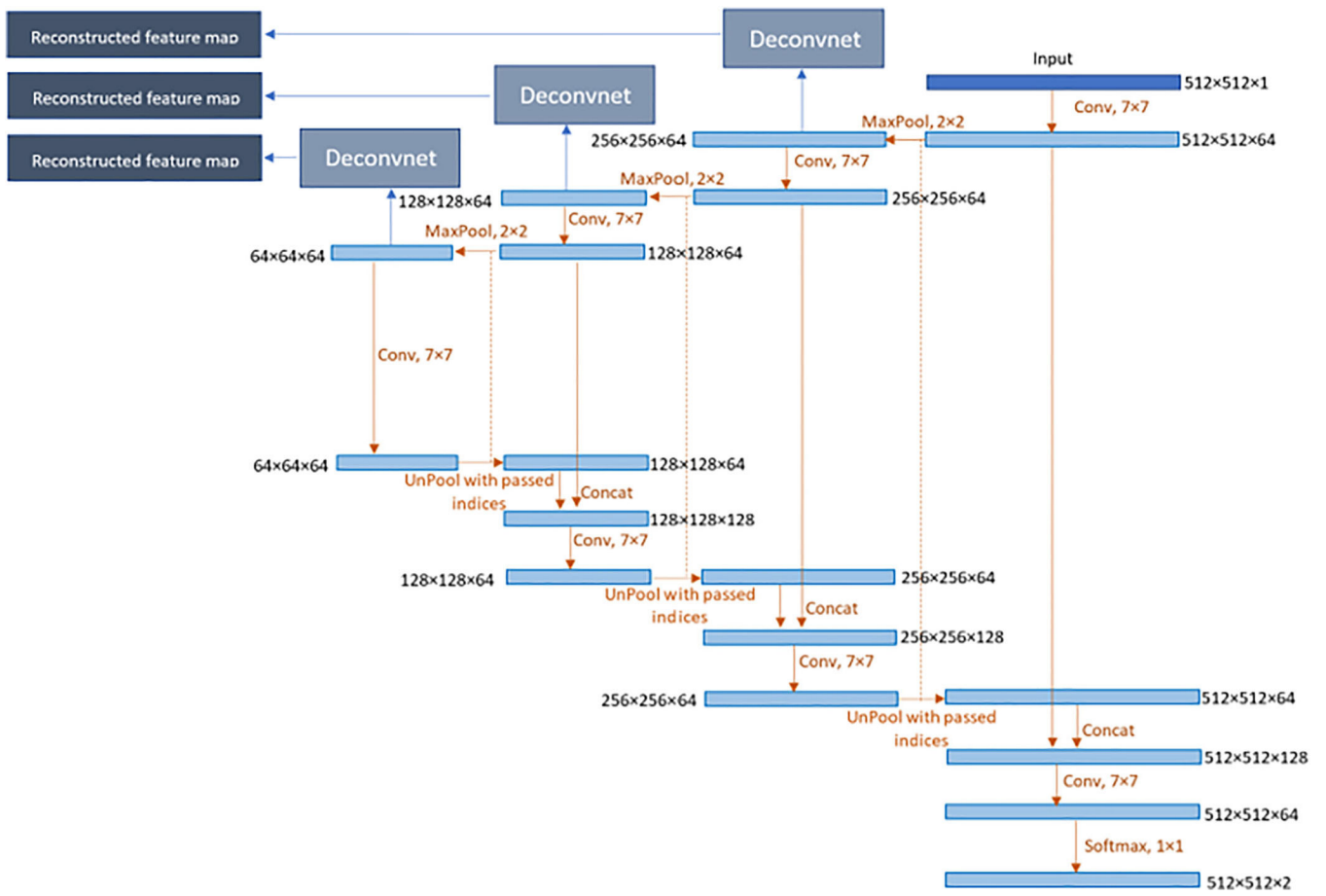


FIGURE 2.
Visualizing and understanding U-Net

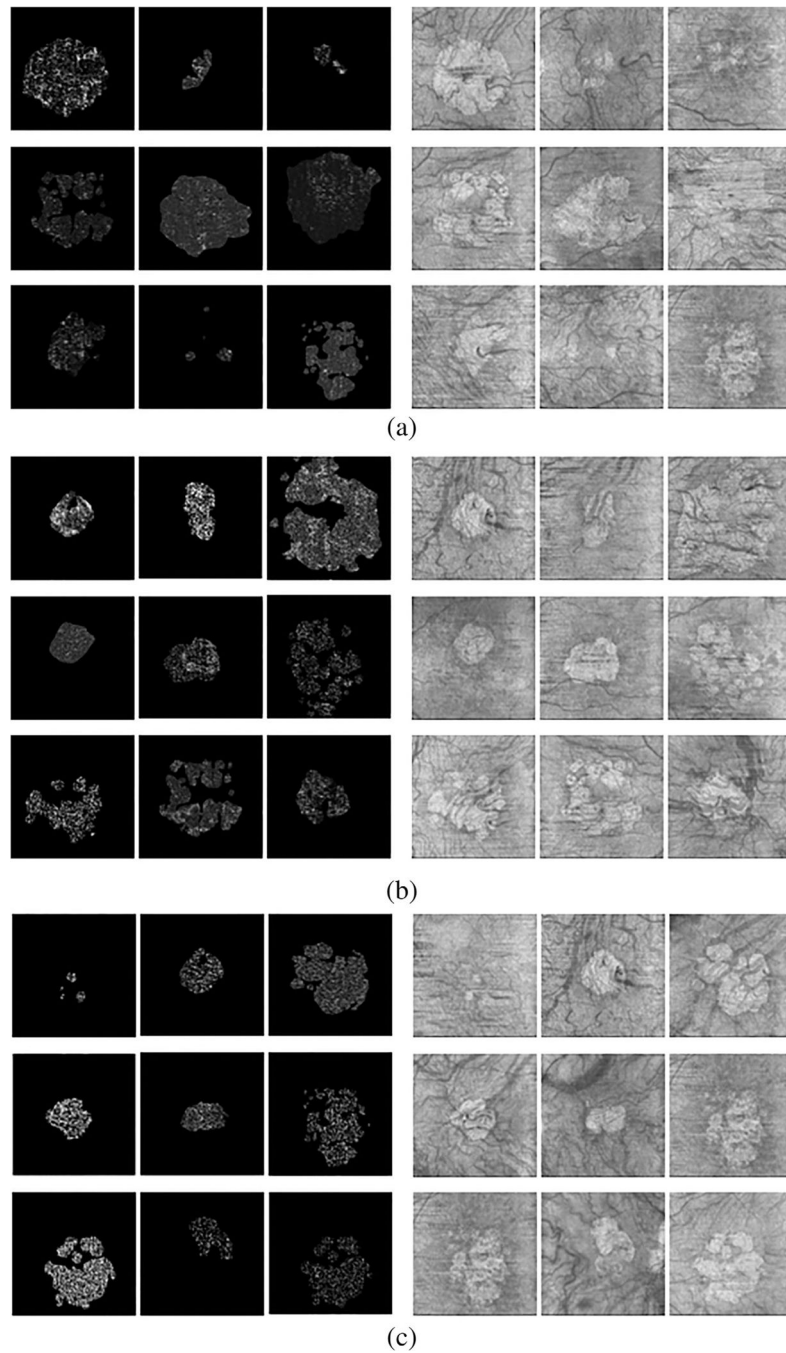


FIGURE 3. Visualization of features in (A) layer 1, (B) layer 2 and layer (3), in the trained models. We show the top nine activations in a random subset of reconstructed feature images. Reconstructed feature maps are shown in the left and the actual en face maps are shown in the right

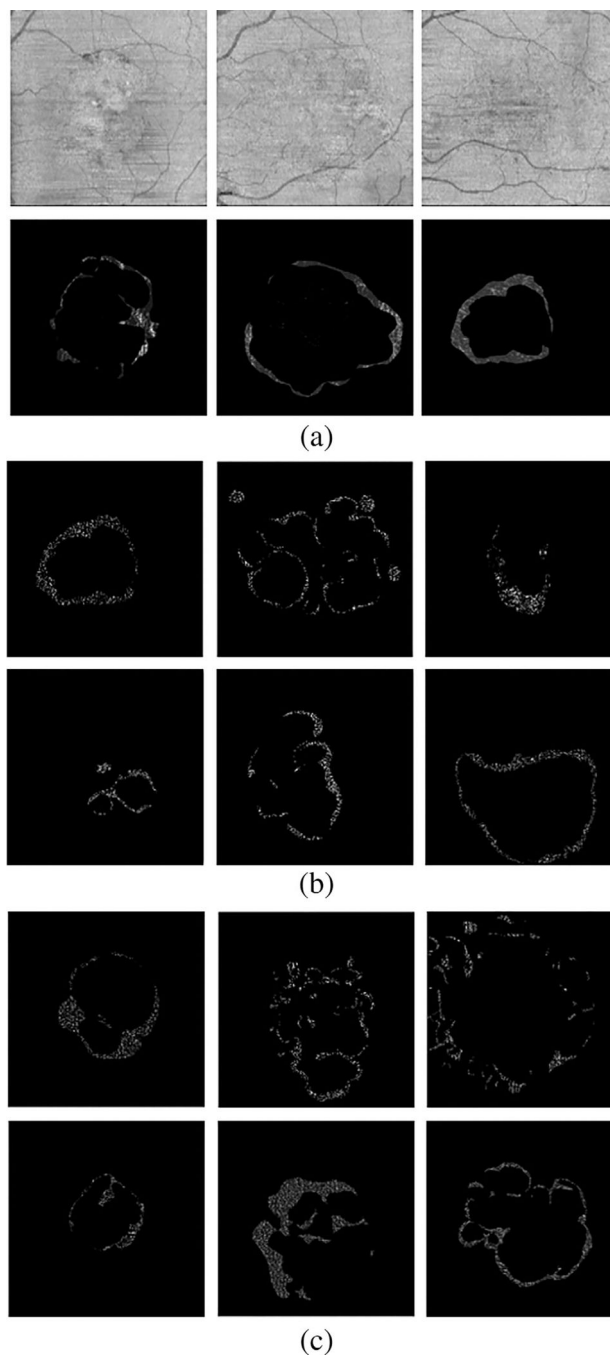


FIGURE 4.

Sample visualizations of the reconstructed feature maps. A, Layer 1, showing only the regions of interests (as determined by the GA evolution). We show the top 3 activations in a random subset of reconstructed feature maps (bottom row) and their corresponding en face images (top row). B, In layer 2, showing only the regions of interests (as determined by the GA evolution). We show the top six activations in a random subset of reconstructed feature maps. C, In layer 3, showing only the regions of interests

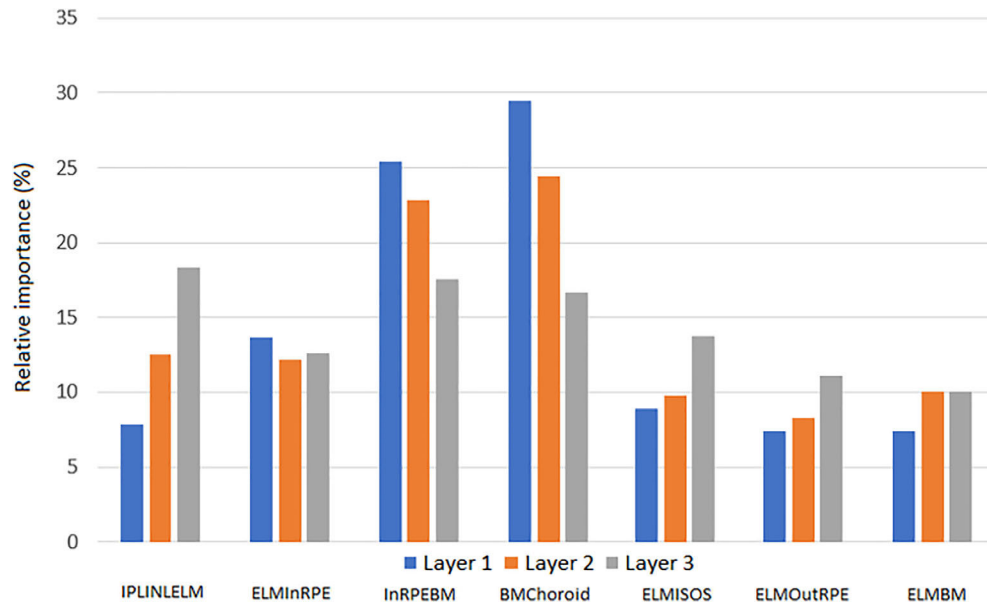


FIGURE 5.
Relative importance of different en faces in GA progression