Deep-learning-aided Diagnosis of DR, AMD, and Glaucoma based on Structural and Angiographic Optical Coherence Tomography

Pengxiao Zang, MS,1,2 Tristan T. Hormel, PhD,1 Thomas S. Hwang, MD,1 Steven T. Bailey, MD,1 David Huang, MD, PhD,1 Yali Jia, PhD1,2

Meeting Presentation: N/A.

Financial Support: National Institutes of Health (R01 EY027833, R01 EY024544, R01 EY031394, R01 EY023285, T32 EY023211, UL1TR002369, P30 EY010572); Unrestricted Departmental Funding Grant and William & Mary Greve Special Scholar Award from Research to Prevent Blindness (New York, NY); Bright Focus Foundation (G2020168).

The sponsor or funding organization had no role in the design or conduct of this research.

Conflict of Interest: Oregon Health & Science University (OHSU), Yali Jia and David Huang have a significant financial interest in Visionix, Inc. Yali Jia has a significant financial interest in Optos, Inc. These potential conflicts of interest have been reviewed and managed by OHSU.

Running head: DR, AMD, and glaucoma diagnosis based on deep learning

Address for reprints: 515 SW Campus Dr., CEI 3154, Portland, Oregon 97239-4197

Key words: Deep learning, optical coherence tomography, diabetic retinopathy, age-related macular degeneration, glaucoma.

Abbreviations and Acronyms: AMD = age-related macular degeneration; AREDS = Age-Related Eye Disease Study; AUC = area under the curve; CAMs = class activation maps; DME = diabetic macular edema; DR = diabetic retinopathy; ETDRS = Early Treatment of Diabetic Retinopathy Study; OCT = optical coherence tomography; OCTA = optical coherence tomography angiography. ROC = Receiver operating characteristic.

1. Casey Eye Institute, Oregon Health & Science University, Portland, OR, USA
2. Department of Biomedical Engineering, Oregon Health & Science University, Portland, OR, USA

Correspondence: Yali Jia, PhD, Casey Eye Institute & Department of Biomedical Engineering, Oregon Health & Science University, 515 SW Campus Dr., CEI 3154, Portland, Oregon 97239-4197. E-mail: jiaya@ohsu.edu
Abstract

Purpose: Timely diagnosis of eye disease is paramount to obtaining the best treatment outcomes. Optical coherence tomography (OCT) and its angiography (OCTA) have several advantages that lend themselves to the early detection of ocular pathology, including that the techniques produce large, feature-rich data volumes. However, the full clinical potential of both OCT and OCTA is stymied when the complex data they acquire must be manually processed. Here we propose an automated diagnostic framework based on structural OCT and OCTA data volumes that could substantially support the clinical application of these technologies.

Design: Cross-sectional study.

Participants: Five hundred and twenty-six OCT and OCTA volumes were scanned from the eyes of 91 healthy participants, 161 patients with diabetic retinopathy (DR), 95 patients with age-related macular degeneration (AMD), and 108 patients with glaucoma.

Methods: The diagnosis framework was constructed based on semi-sequential 3D convolutional neural networks. The trained framework classifies a combined structural OCT and OCTA scan as normal, DR, AMD, or glaucoma. Five-fold cross-validation was performed, with 60% of the data reserved for training, 20% for validation, and 20% for testing. Training, validation, and testing data sets were independent, with no shared patients. For the scans diagnosed as DR, AMD, or glaucoma, 3D class activation maps were generated to highlight the subregions which were considered important by the framework for the automated diagnosis.

Main outcome measures: Area under the curve (AUC) of the receiver operating characteristic curve and quadratic-weighted kappa were used to quantify the diagnostic performance of the framework.

Results: For the diagnosis of DR, the framework achieved a 0.95 ± 0.01 AUC. For the diagnosis of AMD, the framework achieved a 0.98 ± 0.01 AUC. For the diagnosis of glaucoma, the framework achieved a 0.91 ± 0.02 AUC.

Conclusion: A deep learning framework can provide reliable, sensitive, interpretable, and fully automated eye disease diagnosis.
Introduction

Diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma each represent a leading cause of blindness.\(^1\)\(^5\) While the pathophysiologic processes behind vision loss in each of these diseases are unique, they share qualities that make early diagnosis essential. Each is usually asymptomatic during early development.\(^2\)\(^5\)\(^6\) In the case of DR and glaucoma, treatment during the early stages is effective for slowing disease progression and preventing otherwise incurable vision loss.\(^7\)\(^8\) In AMD, conversion to the exudative form of the disease can also lead to rapid, catastrophic vision loss; diagnosis of wet AMD is consequently a major treatment indicator.\(^6\) For DR, AMD, and glaucoma, then, effective screening and early diagnosis are key to preventing poor visual outcomes.

However, current diagnostic protocols face important challenges. Among these is a reliance on qualitative traits that may instill subjectivity into diagnoses. Also problematic are protocols that recommend multiple imaging modalities (for example, fundus photography supplemented with optical coherence tomography (OCT) to confirm the presence of edema or exudation),\(^8\) which increases screening cost, requires more training for instrument technicians and can encourage patient non-compliance with clinician recommendations.\(^9\)

These issues can be alleviated through the use of OCT and OCT angiography (OCTA), which together provide depth-resolved (3D), micrometer-scale-resolution structural and vascular images of the retina.\(^10\)\(^-\)\(^13\) Numerous studies from multiple investigators have confirmed the ability of combined OCT and OCTA imaging to diagnose and detect pathology related to DR, AMD, and glaucoma using quantitative measurements.\(^14\)\(^-\)\(^22\) Additionally, combined structural OCT and OCTA have several advantages as a screening technology. Since 2014 OCT has been the most common procedure in ophthalmic practice and is cost-effective relative to allied modalities such as color fundus photography or dye-based angiography.\(^23\) Furthermore, since OCTA can be acquired from structural OCT through alternative data processing, only one procedure is required to obtain both structural and vascular information.\(^24\) And finally, since procedures are non-invasive, combined OCT and OCTA imaging can be performed at will.

Despite these advantages for diagnosing DR, AMD, and glaucoma, a diagnostic platform based on combined structural OCT and OCTA imaging will still require innovation before it can be translated into the clinic. Combined structural OCT and OCTA datasets are large, and manual review of these datasets can be prohibitively time-consuming. Manual review is nonetheless often required, particularly in analytic frameworks that rely on en face images since retinal slab segmentation errors (which are common in more pathologic retinas) can introduce
In underserved areas, the clinical infrastructure to meet these image analysis demands may not be available. To resolve these issues automated image analysis is required. Deep learning is a data-driven technique that is currently the most powerful tool for medical image classification tasks. Previously, diagnostic deep learning networks have been proposed for DR, AMD, and glaucoma. However, none of these methods can be used for the automated diagnosis of all three of these diseases simultaneously, which means each must be applied sequentially. This has the net effect of undermining generality and will require technicians to be familiar with several algorithms. Here, we instead present a deep-learning-based platform using combined structural OCT and OCTA data volumes as inputs capable of simultaneously diagnosing DR, AMD, and glaucoma. By relying on data volumes this platform avoids mis-segmentation artifacts in en face images (which are difficult to correct). Providing a unified diagnostic framework also ensures that each of these important diseases will be screened for, and also saves computational resources by checking for each disease type simultaneously. In addition, the network outputs 3D class activation maps (CAMs) to highlight the disease-related biomarkers which are helpful for treatment decisions and management, as well as for verifying the algorithm’s predictions.

**Methods**

**Data acquisition**

In this study, 102 eyes of 91 healthy participants, 161 eyes of 161 DR patients, 142 eyes of 95 AMD patients and 121 eyes of 108 glaucoma patients were examined at the Casey Eye Institute, Oregon Health & Science University, USA. Each patient had one or both eyes scanned; the entire data set used in this study included 526 volumetric scans. For each eye, the macular region was scanned using a commercial 70-kHz spectral-domain OCT system (Avanti RTVue-XR, Optovue Inc) with an 840-nm central wavelength. The scan depth was 1.6 mm in a 6.0 × 6.0 mm region (640 × 400 × 400 pixels) centered on the fovea. Blood flow was detected using the split-spectrum amplitude-decorrelation angiography (SSADA) algorithm based on the speckle variation between two repeated B-frames. The OCT structural images were obtained by averaging two repeated B-frames. For each data set, two volumetric raster scans (one x-fast scan and one y-fast scan) were registered and merged through an orthogonal registration algorithm to reduce motion artifacts. In addition, the projection-resolved OCTA algorithm was applied to all OCTA scans to remove flow projection artifacts. According to manufacturer...
recommendations and our experience, OCT/OCTA scans with a signal strength index lower than 50 are generally low quality and were excluded.

A masked trained retina specialist (TSH) graded 7-field color fundus photographs based on the Early Treatment of Diabetic Retinopathy Study (ETDRS) scale to generate the positive ground truth labels for the DR data volumes. Diabetic macular edema (DME) was identified using the central subfield thickness from structural OCT based on the Diabetic Retinopathy Clinical Research Network standard. The eyes with an ETDRS score of 14 or worse or any stage with DME were graded as DR cases. Another masked trained retina specialist (STB) generated the positive AMD ground truth labels by grading 7-field color fundus photographs based on the Age-Related Eye Disease Study (AREDS) scale. The eyes with AREDS simplified score of 1 or worse were graded as AMD cases. Glaucomatous eyes were determined through clinical diagnosis, and the inclusion criteria for this study were an optic disc rim defect (thinning or notching) or nerve fiber layer defect visible on slit-lamp biomicroscopy (DH). Participants were enrolled after informed consent in accordance with an Institutional Review Board approved protocol, and this study was conducted in compliance with the Declaration of Helsinki and Health Insurance Portability and Accountability Act.

Data inputs

While 3D OCT and OCTA scans can provide much more detailed information than 2D data projections, it is also much more challenging to train a network to extract the relevant information from data volumes than images. This difficulty is compounded in our work by the need to extract relevant features for three different diseases. In order to improve the computational and space efficiency of the framework, each volumetric OCT and OCTA is resized to 160 × 224 × 224 voxels and normalized to voxel values between 0 and 1. Combining the structural OCT and OCTA volumes, the final input dimensions were 160 × 224 × 224 × 2 pixels (Fig. 1).

DR, AMD, and glaucoma diagnostic framework

The proposed automated DR, AMD, and glaucoma diagnostic framework use a semi-sequential classifier which includes two parts (Fig. 1). The first part is a classifier used to diagnose DR and AMD in parallel. This part was trained based on the whole data set with a ground truth label of three classes (DR, AMD, and neither). The second part diagnoses glaucoma from data that was not diagnosed as DR or AMD by the first part, which means the glaucoma was sequentially diagnosed after the DR and AMD diagnosis. Therefore, the combination of these two
parts was named as semi-sequential classifier since it contained both parallel and sequential diagnoses. The reason for using a semi-sequential structure to diagnose glaucoma is that the difference between normal and glaucoma in macular OCT/OCTA is much smaller than the difference between normal and DR or AMD. In our experience, glaucoma could not be accurately detected if only one part was used for the diagnosis of DR, AMD, and glaucoma at the same time (see results). To ensure the second part only focused on the difference between normal and glaucoma, it was trained only based on the normal and glaucoma data with two-class labels. Therefore, the two parts were trained separately. Data not diagnosed as DR, AMD, or glaucoma could be considered as normal during our training framework, which relied on healthy eyes being distinguished from these three diseases. However, we note that other eye diseases could still be present in a clinical context. The classifier of each part use a customized 3D convolutional neural network architecture with 16 convolutional layers (Supplemental Fig. S1). For the first part, two output layers were designed for DR and AMD diagnosis, respectively. For the second part, only one output layer was used to classify each input as normal or glaucoma. Each output layer is a fully-connected layer with a softmax function. For the scans diagnosed as DR, AMD, or glaucoma by the full semi-sequential classifier, 3D CAMs are generated by projecting the weight parameters from the corresponding output layer back to the feature maps of the last convolutional layer before global average pooling.

**Evaluation and statistical analysis**

The area under the curve (AUC) for receiver operating characteristic (ROC) and precision-recall curves were used as the primary evaluation metrics to quantify diagnostic accuracy for each disease. Quadratic-weighted Cohen’s kappa was used as the metric to evaluate multiple disease diagnostic performance. In addition, the overall accuracy, sensitivity, and specificity were also calculated. Five-fold cross-validation with a 60/20/20 training/validation/testing data distribution was used to assess performance reliability. Data from a single participant was included in only one of either the training, validation or testing data sets. The parameters and hyperparameters in our framework were trained and optimized only using the training and validation data set. The test data set was used exclusively for evaluation to guarantee performance was not biased. In addition, adaptive label smoothing was used during training to reduce overfitting. To evaluate the performance improvement brought by the semi-sequential structure, a parallel classifier with three output layers was constructed to classify each input as normal, DR, AMD, or glaucoma. The parallel classifier was
trained, validated, and evaluated based on the same data set as the semi-sequential classifier. But unlike the semi-sequential classifier, glaucoma would be parallelly classified with DR and AMD by the parallel classifier.

**Results**

The framework achieved reliable performance as indicated by the AUCs of ROC curves on the test data set, which exceeded 0.9 for each disease in this study (Table 1, Fig. 2). For the precision-recall curves, both DR and AMD diagnoses achieved high AUCs (above 0.9). Though a separate part in the semi-sequential classifier was used to diagnose glaucoma, the AUCs of both ROC and precision-recall curves for glaucoma diagnosis were still lower than the other two eye diseases (Fig. 2). The overall accuracy of the multiple eye disease diagnosis (normal, DR, AMD, and glaucoma) was about 80%.

We also constructed two confusion matrices (for the first part of the semi-sequential classifier and the full semi-sequential classifier) using the overall results from a 5-fold cross-validation (Fig. 3). In the first part that only diagnosis DR and AMD, most misdiagnoses were between normal/glaucoma and DR. In the full semi-sequential classifier (which also includes glaucoma and normal diagnoses), normal eyes were most often misdiagnosed, and when diseased eyes were misdiagnosed, it was most often as normal eyes.

To quantify the performance improvement brought by the semi-sequential structure, a comparison between the glaucoma classification performances of semi-sequential and parallel classifiers was performed (Table 2). The comparison was performed only based on the normal and glaucoma testing data. With the semi-sequential classifier, the sensitivity, specificity, and AUC of the ROC curve were respectively improved by 12.00%, 6%, and 0.1. This improvement was because the semi-sequential diagnosis makes more sense than the parallel diagnosis of multiple diseases in this context, given the fact that the difference between normal and glaucoma data is much smaller than the differences between normal and AMD or DR data. In the training, the parallel classifier mostly focused on the learning of unique features of DR and AMD and ignored the glaucoma features. The improvement brought by the semi-sequential structure was critical for the glaucoma diagnosis performance of the proposed diagnosis framework.

In order to aid confirmation of the model’s outputs and interpret its decision-making, our framework also produces 3D class activation maps (CAM) (Fig. 4 and 5). We found that the CAMs frequently highlight pathology that is known to be associated with the diseases in this study, for example, non-perfusion and low-perfusion areas around...
the fovea were highly weighted for decision making of DR (highlighted regions in Fig. 4(A)). In AMD data, the CAMs highlighted most of the drusen areas (Fig. 5(C) and 5(D)).

From glaucoma classification (Fig. 6), we can see that the semi-sequential classifier was mostly focused on the vanished nerve fiber layer, which is consistent with known glaucoma pathophysiology⁴⁸,⁴⁹ (Fig. 6(D)). In addition, the low perfusion area was also highlighted by the CAM (Fig. 6(A)). These attention maps offer many opportunities for us to validate the performance of deep learning frameworks and discover new potential biomarkers for disease understanding and diagnosis.

Discussion

In this study, we proposed an automated diagnostic framework based on volumetric OCT/OCTA data that diagnoses DR, AMD, and glaucoma. The framework uses a semi-sequential classifier which consists of two parts with identical architecture, one of which diagnoses DR and AMD, and the other of which diagnoses glaucoma. We found that this semi-sequential structure, which uses separate parts (classifiers) for AMD/DR and glaucoma, outperformed a single parallel classifier that learns to diagnose all three diseases. The framework achieved an AUC of ROC curve over 0.9 for the diagnosis of each disease. These results indicate that our automated framework achieved reliable DR, AMD, and glaucoma diagnosis performance using only a single ophthalmic imaging modality.

Compared to current deep-learning-aided eye disease diagnosis methods based on OCT/OCTA, our framework also includes several advantages. The first advantage is our framework can be used to diagnose DR, AMD, and glaucoma simultaneously, which could reduce the time and financial costs of screening. In addition, ophthalmologists could have a more comprehensive understanding of the eye condition of the referred patients based on our diagnosis results. The second advantage is the use of the whole 3D volume. Other approaches which rely on en face images are prone to segmentation errors and may miss important features without access to cross-sectional information (such as small drusen or retinal fluid). And traditional frameworks that perform diagnosis based on the presence or absence of known pathologic features may not account for undiscovered relevant features or information, and fail to utilize all of the information available in combined structural OCT and OCTA data volumes. In the contrast, our approach is biomarker/feature-agnostic, which means that correlations or structures within the data volume that may be difficult for a human to identify can still be incorporated into decision-making. As a corollary, our framework
may also have a greater capacity to improve with more training data since a full data volume contains far more information than an image formed by projection.

Another significant advantage of our framework is the inclusion of 3D CAMs. Deep learning algorithms are often likened to “black-box”, since their decision-making is difficult to interpret. This is problematic, since opaque decision-making may hide important biases that could prove to be disadvantageous for certain groups. The interpretability provided by the 3D CAMs would allow clinicians to verify and understand the diagnosis decisions and ensure they are correct, an essential requirement in any diagnostic framework. Compared to 2D CAMs, 3D CAMs indicate which retinal layer in each B-scan is relevant for each diagnosis. We verified that the CAM output by our model highlighted features known to be associated with each of the diseases examined in this study: non-perfusion areas in DR (Figs 4(A)), drusen in AMD (Fig. 5(D)), and nerve fiber layers with abnormal structure in glaucoma diagnosis (Fig. 6(D)). Although the 3D CAMs did not demonstrate all features used for diagnosing eye diseases, they found many key features, indicating that our framework has successfully learned relevant features and that 3D CAMs could be useful in clinical review and sanity checks. In addition, the 3D CAMs were used to highlight the biomarkers which were selected by our framework, but not all the biomarkers were selected. That only some of the biomarkers were highlighted means, these biomarkers were already sufficient for our framework to make the diagnosis decision.

There are three aspects of the diagnosis performance of our framework that could be improved in future work. Firstly, our data set only contained healthy eyes or eyes that had one of the three diseases (DR, AMD, or glaucoma), whereas in clinical practice an eye may suffer from different condition (e.g., branch retinal vein occlusion) or even multiple diseases simultaneously (e.g., AMD with DR or AMD with retinitis pigmentosa). This limitation may lead to performance loss in our model if it were attempted on an eye with conditions that were not included in our data set. Secondly, the use of a semi-sequential structure increased the glaucoma diagnosis accuracy but also limited the framework for diagnosing eyes with both glaucoma and DR or glaucoma and AMD. This framework solved the multiple classification problem for a single diagnosis among three eye diseases, but future work will need to generalize our strategy in order to make multiple simultaneous diagnoses for more diseases or eyes with multiple diseases. Finally, some of the design choices that led to the second limitation were to improve glaucoma diagnostic performance (Table 2), but even so, the sensitivity for glaucoma diagnosis (71.67 ± 4.08%) was lower than the other two grades (90.00% ± 2.34% for DR and 88.28% ± 5.60% for AMD). Because only scans on macula were
used in this study information from the optic disc, where glaucoma pathology is more prominent,\textsuperscript{50} was unavailable for decision making. Training on a larger dataset with cases of multiple diseases would likely improve performance for not only glaucoma but the other diseases in this study as well. In particular, the accuracy of the parallel classifier could probably be similar to the semi-sequential classifier in the main module if more glaucoma data for training was available. The framework limitation could therefore be solved by using a better-trained parallel classifier.

In addition to diagnosis performance, there are also limitations if we use our framework in real-world clinical applications right now. Our framework can only be used in clinics with both OCT and OCTA available. But this limitation will gradually disappear as OCTA applications become more widespread. In addition, the data set used in this study were all scanned by Avanti RTVue-XR in Casey Eye Institute, Oregon Health & Science University, and only scans with signal strength index above 50 were preserved. The diagnosis performance may be lower on external or lower quality data, or data scanned on other OCT devices. Therefore, to improve the clinical utility of our framework, data without these limitations will also be included in the future.

**Conclusion**

We proposed a deep-learning-aided DR, AMD, and glaucoma diagnostic framework that takes combined 3D structural OCT and OCTA data as inputs. Our framework achieved reliable performance on the diagnosis of each disease for which it was designed, and produces 3D CAMs that can be used to interpret the model’s decision-making. By using our framework, the number of scanning procedures and eye exams required for the diagnosis of the three different eye diseases was reduced to just a single OCT/OCTA procedure. In addition, by using 3D data as inputs, our framework can totally avoid the influences from unstable retinal layer segmentation. At last, our results show that the biomarker-agnostic framework based on 3D OCT and OCTA could be beneficial for clinical practice.

**Reference**


23 Centers for Medicare & Medicaid Services. Medicare Provider Utilization and Payment Data: Physician and Other Supplier.


Structural OCT and OCTA data volumes are resampled and combined to form the input to a semi-sequential classifier. The first part of the classifier then diagnoses DR and AMD. Data not diagnosed as DR and AMD by the first part is fed to the second part for glaucoma. Eyes not diagnosed with DR, AMD or glaucoma can be considered normal or other diseases. For any disease diagnosis, the network also outputs 3D CAMs.

The area under the curve (AUC) was calculate for both curves. The models achieve AUCs of 0.95 ± 0.01 and 0.91 for ROC and precision-recall on DR diagnosis, AUCs of 0.98 ± 0.01 and 0.95 for ROC and precision-recall on AMD diagnosis, and AUCs of 0.91 ± 0.02 and 0.71 for ROC and precision-recall on glaucoma diagnosis. In addition, the glaucoma precision-recall curve look different from the other two curves since the glaucoma prediction was the combination of two parts in the semi-sequential classifier.

Based on the overall results of 5-fold cross-validation.

Figure 4. Class activation map (CAM) based on the DR output layer of the semi-sequential classifier for a correctly classified DR eye. (A) OCTA en face projection of the superficial vascular complex (SVC; inner 80% of the ganglion cell complex). The non-perfusion and low-perfusion areas were highlighted by the CAM. (B) Corresponding B-scan at the position of the red line in (A). (C) Structural OCT en face image of the ellipsoid zone (EZ; outer nuclear layer / ellipsoid zone boundary to ellipsoid zone / retinal pigment epithelium boundary). (D) Corresponding B-scan at the location of the red line in (C).

Figure 5. Class activation map (CAM) based on the AMD output layer of the semi-sequential classifier for a correctly classified AMD eye. (A) OCTA en face projection of the superficial vascular complex (SVC; inner 80% of the ganglion cell complex). (B) Corresponding B-scan at the position of the red line in (A). (C) Structural OCT en face image of the ellipsoid zone (Outer nuclear layer / ellipsoid zone boundary to ellipsoid zone / retinal pigment epithelium boundary). (D) Corresponding B-scan at the location of the red line in (C). The drusen area was highlighted by the CAM.

Figure 6. Class activation map (CAM) based on the glaucoma output layer of the semi-sequential classifier for a correctly classified glaucoma eye. (A) OCTA en face projection of the superficial vascular complex (SVC; inner 80% of the ganglion cell complex). The low perfusion area was highlighted (B) Corresponding B-scan at the position of the red line in (A). (C) Structural OCT en face image of the inner retina (Vitreous / inner limiting membrane boundary to outer plexiform layer / outer nuclear layer boundary). (D) Corresponding B-scan at the location of the red line in (C). The region of the vanished nerve fiber layer was highlighted.
Table 1. Automated disease diagnosis performance

<table>
<thead>
<tr>
<th>Metric</th>
<th>DR diagnosis</th>
<th>AMD diagnosis</th>
<th>Glaucoma diagnosis</th>
<th>Eye diseases diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall accuracy</td>
<td>90.19% ± 2.03%</td>
<td>94.53% ± 0.71%</td>
<td>89.25% ± 1.75%</td>
<td>79.43% ± 2.01%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>90.00% ± 2.34%</td>
<td>88.28% ± 5.60%</td>
<td>71.67% ± 4.08%</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>90.27% ± 1.99%</td>
<td>96.88% ± 1.76%</td>
<td>94.39% ± 1.98%</td>
<td></td>
</tr>
<tr>
<td>AUC of ROC</td>
<td>0.95 ± 0.01</td>
<td>0.98 ± 0.01</td>
<td>0.91 ± 0.02</td>
<td></td>
</tr>
<tr>
<td>AUC of precision-recall</td>
<td>0.91</td>
<td>0.95</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>Quadratic-Weighted Kappa</td>
<td>0.78 ± 0.05</td>
<td>0.86 ± 0.02</td>
<td>0.68 ± 0.05</td>
<td>0.57 ± 0.05</td>
</tr>
</tbody>
</table>

AMD = age-related macular degeneration; AUC = area under the curve; DR = diabetic retinopathy; ROC = receiver operating characteristic.
Table 2. The comparison between the glaucoma classification performances of the semi-sequential and parallel classifiers

<table>
<thead>
<tr>
<th></th>
<th>Overall accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC of ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semi-sequential</td>
<td>77.33% ± 3.82%</td>
<td>78.33% ± 5.53%</td>
<td>76.19% ± 6.73%</td>
<td>0.78 ± 0.03</td>
</tr>
<tr>
<td>Parallel classifier</td>
<td>63.11% ± 4.35%</td>
<td>56.67% ± 6.24%</td>
<td>70.48% ± 3.56%</td>
<td>0.68 ± 0.03</td>
</tr>
</tbody>
</table>

AUC = area under the curve; ROC = receiver operating characteristic.
We proposed a deep learning framework for the diagnosis of diabetic retinopathy, age-related macular degeneration, and glaucoma based on volumetric optical coherence tomography and its angiography. Three-dimensional class activation maps were also generated.