AMERICAN ACADEMY OF OPHTHALMOLOGY®

Deep-Learning–Aided Diagnosis of Diabetic Retinopathy, Age-Related Macular Degeneration, and Glaucoma Based on Structural and Angiographic OCT

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, PhD*[1,2]

Purpose: Timely diagnosis of eye diseases is paramount to obtaining the best treatment outcomes. OCT and OCT angiography (OCTA) have several advantages that lend themselves to early detection of ocular pathology; furthermore, the techniques produce large, feature-rich data volumes. However, the full clinical potential of both OCT and OCTA is stymied when complex data acquired using the techniques 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 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 semisequential 3-dimensional (3D) convolutional neural networks. The trained framework classifies combined structural OCT and OCTA scans as normal, DR, AMD, or glaucoma. Fivefold cross-validation was performed, with 60% of the data reserved for training, 20% for validation, and 20% for testing. The training, validation, and test data sets were independent, with no shared patients. For scans diagnosed as DR, AMD, or glaucoma, 3D class activation maps were generated to highlight subregions that were considered important by the framework for automated diagnosis.

Main Outcome Measures: The 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 an AUC of 0.95 $\pm$ 0.01. For the diagnosis of AMD, the framework achieved an AUC of 0.98 $\pm$ 0.01. For the diagnosis of glaucoma, the framework achieved an AUC of 0.91 $\pm$ 0.02.

Conclusions: Deep learning frameworks can provide reliable, sensitive, interpretable, and fully automated diagnosis of eye diseases.

Financial Disclosure(s): Proprietary or commercial disclosure may be found after the references. *Ophthalmology Science 2023;3:100245 © 2022 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).*

Supplemental material available at www.ophthalmologyscience.org.

Diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma each represents a leading cause of blindness.[1–5] Although the pathophysiologic processes of 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 early stages is effective for slowing disease progression and preventing otherwise incurable vision loss.[7,8] In patients with AMD, conversion to the exudative form of the disease can also lead to rapid, catastrophic vision loss; the diagnosis of wet AMD is, consequently, a major treatment indicator.[6] Therefore, for DR, AMD, and glaucoma, effective screening and early diagnosis are key to preventing poor visual outcomes. However, current diagnostic protocols face important challenges. Among these is reliance on qualitative traits that may instill subjectivity into diagnoses. Additionally, protocols that recommend multiple imaging modalities (e.g., fundus photography supplemented with OCT to confirm the presence of edema or exudation)[8]—which increases the screening cost, requires more training for instrument technicians, and can encourage patient noncompliance with clinician recommendations—are problematic.[9]

https://doi.org/10.1016/j.xops.2022.100245
ISSN 2666-9145/22

These issues can be alleviated through the use of OCT and OCT angiography (OCTA), which together provide depth-resolved, 3-dimensional (3D), micrometer-scale–resolution structural and vascular images of the retina.[10–13] Numerous studies by multiple investigators have confirmed the ability of combined OCT and OCTA imaging to diagnose and detect the pathology of 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, because OCTA images can be acquired from structural OCT scans via alternative data processing, only 1 procedure is required to obtain both structural and vascular information.[24] Finally, because these procedures are noninvasive, combined OCT and OCTA imaging can be performed at will.

Despite these advantages for the diagnosis of DR, AMD, and glaucoma, a diagnostic platform based on combined structural OCT and OCTA imaging will still require innovation before it can be translated to clinics. Combined structural OCT and OCTA data sets 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 because retinal slab segmentation errors (which are common in more pathologic retinas) can introduce artifacts.[25] In underserved areas, the clinical infrastructure required to meet these image analysis demands may not be available.[26] To resolve these issues, automated image analyses are required. Deep learning is a data-driven technique that is currently the most powerful tool for medical image classification tasks.[27] Previously, diagnostic deep learning networks were proposed for DR,[28–33] AMD,[34,35] and glaucoma.[36–39] However, none of these methods can be used for automated diagnosis of all 3 of these diseases simultaneously, which means that each must be applied sequentially. This has the net effect of undermining generality and requires technicians to be familiar with several algorithms. Therefore, here, we 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 missegmentation artifacts on en face images (which are difficult to correct). Providing a unified diagnostic framework also ensures that each of these important diseases is screened for and saves computational resources by checking for each disease type simultaneously. In addition, the network outputs 3D class activation maps (CAMs) to highlight disease-related biomarkers that are helpful for treatment decisions and management as well as verifying the algorithm's predictions.

Methods

Data Acquisition

In this study, 102 eyes of 91 healthy participants, 161 eyes of 161 patients with DR, 142 eyes of 95 patients with AMD, and 121 eyes of 108 patients with glaucoma were examined at the Casey Eye Institute, Oregon Health & Science University, Portland, Oregon. Each patient had 1 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 a central wavelength of 840 nm. The scan depth was 1.6 mm in a $6.0 \times 6.0\text{-mm}^2$ region ($640 \times 400 \times 400$ pixels) centered on the fovea. Blood flow was detected using the split-spectrum amplitude-decorrelation angiography algorithm based on the speckle variation between 2 repeated B frames.[13] OCT structural images were obtained by averaging 2 repeated B frames. For each data set, 2 volumetric raster scans (1 x-fast scan and 1 y-fast scan) were registered and merged using an orthogonal registration algorithm to reduce motion artifacts.[40] In addition, a projection-resolved OCTA algorithm was applied to all OCTA scans to remove flow projection artifacts in the deeper plexuses.[41,42] According to the manufacturer's recommendations and our experience, OCT or OCTA scans with a signal strength index of < 50 are generally low quality and were excluded.

A masked, trained retina specialist (T.S.H.) graded 7-field color fundus photographs based on the ETDRS scale[43,44] to generate positive ground truth labels for the DR data volumes. Diabetic macular edema was identified based on the central subfield thickness using structural OCT based on the Diabetic Retinopathy Clinical Research Network standard.[45] Eyes with an ETDRS score of 14 or worse or any stage with diabetic macular edema were graded as DR cases. Another masked, trained retina specialist (S.T.B.) generated positive AMD ground truth labels by grading 7-field color fundus photographs based on the Age-Related Eye Disease Study scale.[46] Eyes with an Age-Related Eye Disease Study simplified score of 1 or worse were graded as AMD cases. Glaucomatous eyes were determined using clinical diagnosis, and the inclusion criteria for this study were an optic disc rim defect (thinning or notching) or nerve fiber layer defect visible using slit-lamp biomicroscopy (D.H.). The 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

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

Diagnostic Framework for DR, AMD, and Glaucoma

The proposed automated diagnostic framework for DR, AMD, and glaucoma uses a semisequential classifier, which includes 2 parts (Fig 1). The first part is a classifier used to diagnose DR and AMD in parallel. This part was trained based on the entire data set with a ground truth label of 3 classes (DR, AMD, and neither). The second part is used to diagnose glaucoma using data that were not diagnosed as DR or AMD by the first part, which means that glaucoma was sequentially diagnosed after the diagnosis of DR and AMD. Therefore, the combination of these 2 parts was

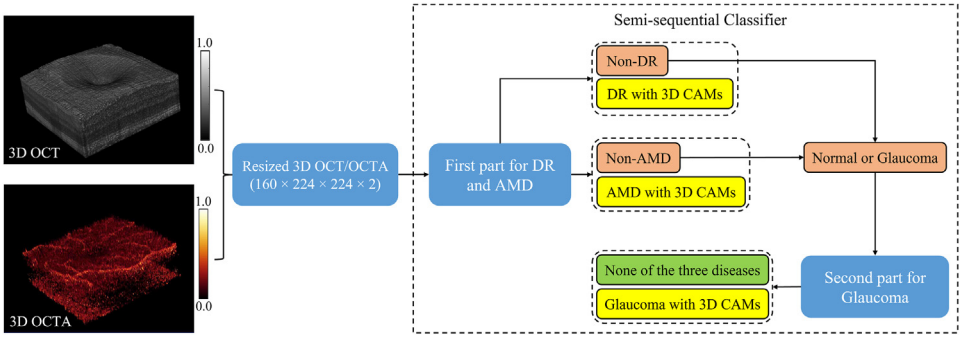

Figure 1. Automated diagnostic framework for diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma using combined volumetric OCT and OCT angiography (OCTA) data as inputs. Structural OCT and OCTA data volumes are resampled and combined to form the input for a semi-sequential classifier. The first part of the classifier then diagnoses DR and AMD. Data not diagnosed as DR and AMD by the first part are fed to the second part for glaucoma. Eyes not diagnosed with DR, AMD, or glaucoma can be considered normal or as having other diseases. For the diagnosis of any disease, the network also generates 3-dimensional (3D) class activation maps (CAMS).

named as a semisequential classifier because it contained both parallel and sequential diagnoses. The reason for using a semisequential structure to diagnose glaucoma is that the difference between normal and glaucoma in macular OCT or OCTA scans is much smaller than the difference between normal and DR or AMD. In our experience, glaucoma could not be accurately detected if only 1 part was used for the diagnosis of DR, AMD, and glaucoma at the same time (see the Results section). To ensure that the second part only focused on the difference between normal and glaucoma, it was trained based on only the normal and glaucoma data with 2-class labels. Therefore, the 2 parts were trained separately. Data that were not diagnosed as DR, AMD, or glaucoma could be considered as normal in our training framework, which relied on healthy eyes being distinguished from eyes with these 3 diseases. However, we noted that other eye diseases could still be present in a clinical context. The classifier of each part used a customized 3D convolutional neural network architecture with 16 convolutional layers (Fig S1, available at www.ophthalmologyscience.org). For the first part, 2 output layers were designed for the diagnosis of DR and AMD. For the second part, only 1 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 scans diagnosed as DR, AMD, or glaucoma by the full semisequential classifier, 3D CAMs are generated by projecting 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 was used as the primary evaluation metric to quantify the diagnostic accuracy for each disease. Quadratic-weighted Cohen kappa[47] was used as the metric to evaluate the diagnostic performance for multiple diseases. In addition, the overall accuracy, sensitivity, and specificity were calculated. Fivefold cross-validation, with 60% of the data reserved for training, 20% for validation, and 20% for testing, was used to assess the performance reliability. Data from a single participant were included in only the training, validation, or test data set. The parameters and hyperparameters in our framework were trained and optimized using only the training and validation data sets. The test data set was used exclusively for evaluation to guarantee that performance was not biased. In addition, adaptive label smoothing was used during training to reduce overfitting.[33]

To evaluate the performance improvement caused by the semisequential structure, a parallel classifier with 3 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 semisequential classifier. However, unlike the semisequential 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 the ROC curves of the test data set, which exceeded 0.9 for each disease in this study (Table 1 and Fig 2). For the precision-recall curves, the diagnosis of both DR and AMD achieved high AUCs (> 0.9). Although a separate part in the semisequential classifier was used to diagnose glaucoma, the AUCs of both the ROC and precision-recall curves for the diagnosis of glaucoma were still lower than those for the other 2 eye diseases (Fig 2). The overall accuracy of the diagnosis of multiple eye diseases (normal, DR, AMD, and glaucoma) was $\sim 80\%$.

We also constructed 2 confusion matrices (for the first part of the semisequential classifier and the full semisequential classifier) using the overall results from the fivefold cross-validation (Fig 3). In the first part, which only diagnoses DR and AMD, most misdiagnoses were between normal or glaucoma and DR. In the full semisequential 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 caused by the semisequential structure, a comparison between the glaucoma classification performances of the semisequential and parallel classifiers was performed (Table 2). The comparison was performed only based on the normal and glaucoma test data. With the semisequential classifier, the sensitivity, specificity, and AUC of the ROC curve were improved by

Table 1. Automated Disease Diagnosis Performance

Metric	DR Diagnosis	AMD Diagnosis	Glaucoma Diagnosis	Eye Disease Diagnosis
Overall accuracy	90.19 ± 2.03%	94.53 ± 0.71%	89.25 ± 1.75%	79.43 ± 2.01%
Sensitivity	90.00 ± 2.34%	88.28 ± 5.60%	71.67 ± 4.08%	
Specificity	90.27 ± 1.99%	96.88 ± 1.76%	94.39 ± 1.98%	
AUC of ROC	0.95 ± 0.01	0.98 ± 0.01	0.91 ± 0.02	
AUC of precision recall	0.91	0.95	0.71	
Quadratic-weighted kappa	0.78 ± 0.05	0.86 ± 0.02	0.68 ± 0.05	0.57 ± 0.05

AMD = age-related macular degeneration; AUC = area under the curve; DR = diabetic retinopathy; ROC = receiver operating characteristic.

12.00%, 6%, and 0.1, respectively. This improvement was because the semisequential diagnosis of multiple disease 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. During training, the parallel classifier mostly focused on learning the unique features of DR and AMD and ignored those of glaucoma. The improvement caused by the semisequential structure was critical for the glaucoma diagnostic performance of the proposed diagnosis framework.

To aid confirmation of the model's outputs and interpret its decision making, our framework also produced 3D CAMs (Figs 4, 5). We found that the CAMs frequently highlighted a pathology that is known to be associated with the diseases in this study, for example, nonperfusion and low-perfusion areas around the fovea were highly weighted

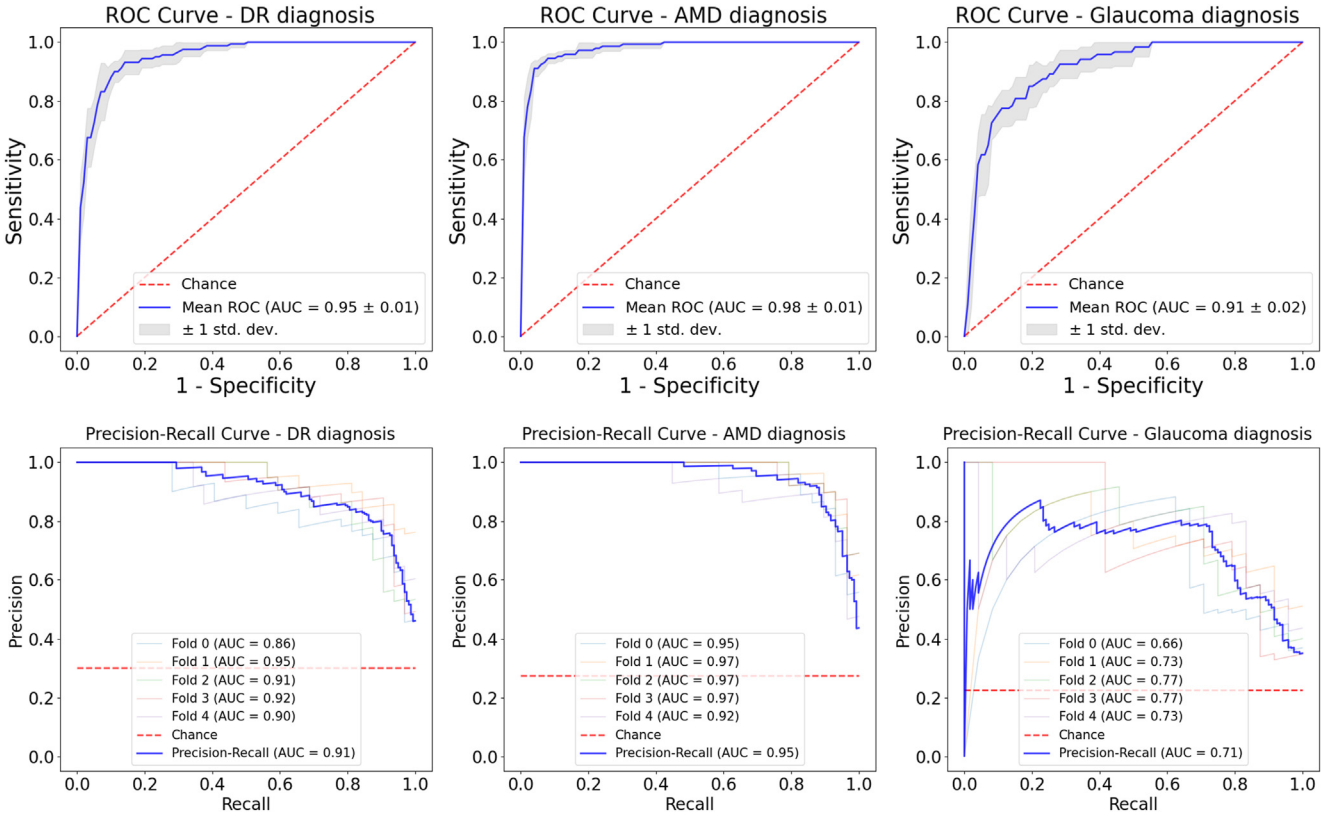

Figure 2. Receiver operating characteristic (ROC) and precision-recall curves derived from fivefold cross-validation for the diagnosis of diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma based on the full framework. The area under the curve (AUC) was calculated for both the curves. The models achieved AUCs of 0.95 ± 0.01 and 0.91 for ROC and precision recall, respectively, for the diagnosis of DR, AUCs of 0.98 ± 0.01 and 0.95 for ROC and precision recall, respectively, for the diagnosis of AMD, and AUCs of 0.91 ± 0.02 and 0.71 for ROC and precision recall, respectively, for the diagnosis of glaucoma. In addition, the precision-recall curve for glaucoma looked different from that of the other 2 diseases because glaucoma prediction was a combination of the 2 parts of the semisequential classifier. std. dev. = standard deviation.

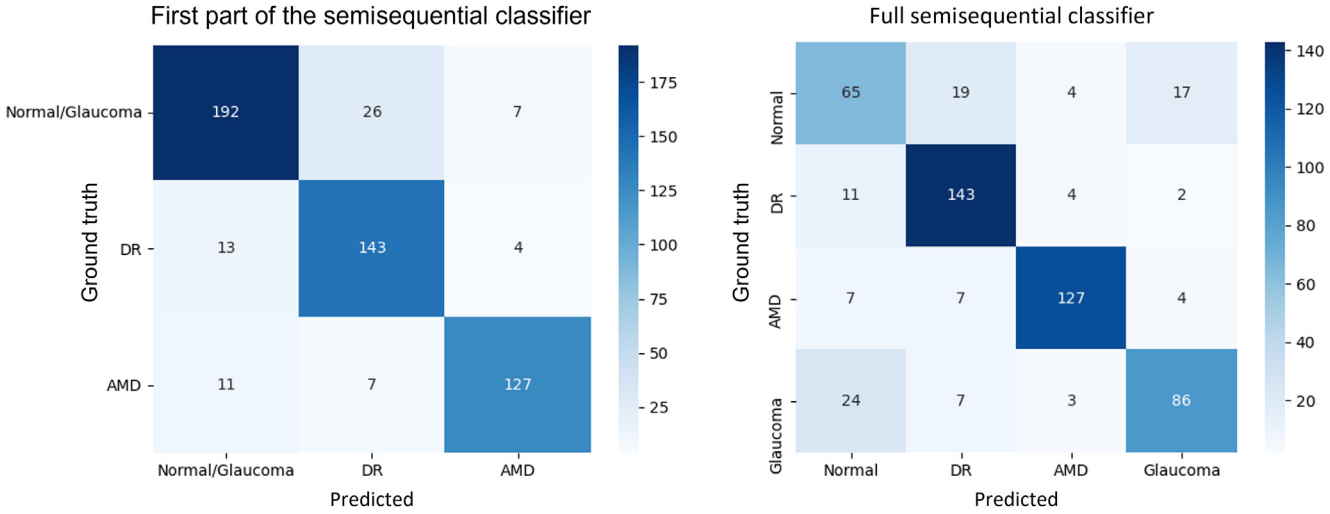

Figure 3. Confusion matrices for the first part of the semisequential classifier (left) and the full semisequential classifier (right) based on the overall results of fivefold cross-validation. AMD = age-related macular degeneration; DR = diabetic retinopathy.

for decision making for DR (highlighted regions in Fig 4A). In the AMD data, the CAMs highlighted most of the drusen areas (Fig 5C, D).

From the classification of glaucoma (Fig 6), we observed that the semisequential classifier was mostly focused on the vanished nerve fiber layer, which is consistent with the known pathophysiology of glaucoma (Fig 6D).[48,49] In addition, the low-perfusion area was also highlighted by the CAMs (Fig 6A). These attention maps offer many opportunities for us to validate the performance of deep learning frameworks and discover new potential biomarkers for understanding and diagnosing diseases.

Discussion

In this study, we proposed an automated diagnostic framework based on volumetric OCT or OCTA data that diagnoses DR, AMD, and glaucoma. The framework uses a semisequential classifier, which consists of 2 parts with identical architecture, one of which diagnoses DR and AMD and the other of which diagnoses glaucoma. We found that this semisequential structure, which uses separate parts (classifiers) for AMD or DR and glaucoma, outperformed a single parallel classifier that learns to diagnose all 3 diseases. The framework achieved an AUC of ROC curve of > 0.9 for the diagnosis of each disease. These results indicate that our automated framework achieved reliable diagnostic performance for DR, AMD, and glaucoma using only a single ophthalmic imaging modality.

Compared with current deep-learning–aided diagnostic methods based on OCT or OCTA for eye diseases, our framework also includes several advantages. The first advantage is that our framework can be used to diagnose DR, AMD, and glaucoma simultaneously, which can reduce the time and financial costs of screening. In addition, ophthalmologists can have a more comprehensive understanding of the eye condition of referred patients based on our diagnosis results. The second advantage is the use of the entire 3D volume. Other approaches that 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). Furthermore, 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 contrast, our approach is biomarker or 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 greater capacity to improve with more training data because a full data volume contains far more information than an image formed by projection.

Table 2. Comparison between the Glaucoma Classification Performances of the Semisequential and Parallel Classifiers

	Overall Accuracy	Sensitivity	Specificity	AUC of ROC
Semisequential classifier	77.33 ± 3.82%	78.33 ± 5.53%	76.19 ± 6.73%	0.78 ± 0.03
Parallel classifier	63.11 ± 4.35%	56.67 ± 6.24%	70.48 ± 3.56%	0.68 ± 0.03

AUC = area under the curve; ROC = receiver operating characteristic.

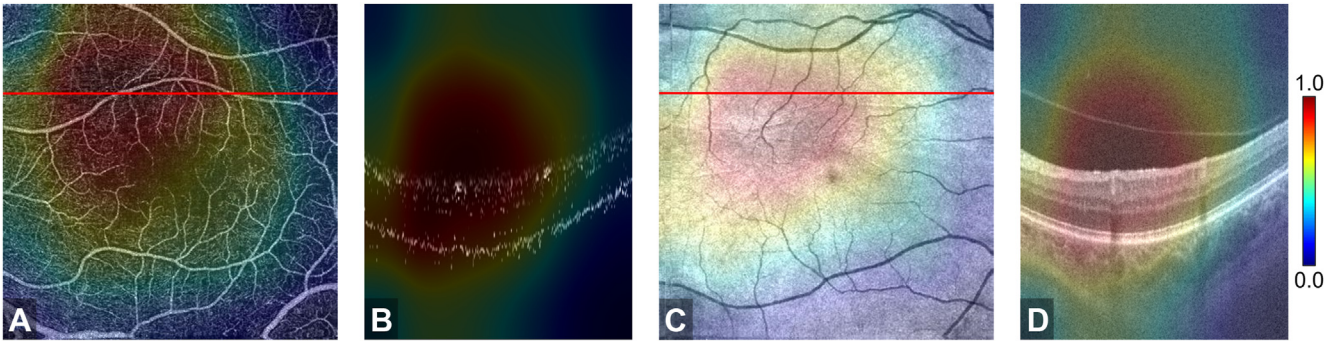

Figure 4. Class activation map based on the diabetic retinopathy output layer of the semisequential classifier for an eye with correctly classified diabetic retinopathy. **A,** OCT angiography (OCTA) en face projection of the superficial vascular complex ([SVC] inner 80% of the ganglion cell complex). The nonperfusion and low-perfusion areas were highlighted by the class activation map. **B,** Corresponding B scan at the position of the red line in (**A**). **C,** Structural OCT en face image of the ellipsoid zone ([EZ] the boundary between the outer nuclear layer and the EZ to the boundary between the EZ and the retinal pigment epithelium). **D,** Corresponding B scan at the location of the red line in (**C**).

Another significant advantage of our framework is the inclusion of 3D CAMs. Deep learning algorithms are often likened to a "black box" because their decision making is difficult to interpret. This is problematic because opaque decision making may hide important biases that could prove to be disadvantageous for certain groups. The interpretability provided by 3D CAMs would allow clinicians to verify and understand diagnosis decisions and ensure that they are correct, an essential requirement for any diagnostic framework. Compared with 2-dimensional 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: nonperfusion areas in the diagnosis of DR (Fig 4A), drusen in the diagnosis of AMD (Fig 5D), and nerve fiber layers with abnormal structure in the diagnosis of glaucoma (Fig 6D). Although the 3D CAMs did not demonstrate all features used for the diagnosis of eye diseases, they found many key features, indicating that our framework successfully learned relevant features and that 3D CAMs could be useful for clinical review and sanity checks. In addition, the 3D CAMs were used to highlight biomarkers that were selected by our framework; however, not all biomarkers were selected. The fact that only some of the biomarkers were highlighted means that these biomarkers were already sufficient for our framework to make the diagnosis decision.

There are 3 aspects of the diagnostic performance of our framework that could be improved in future studies. First, our data set contained only healthy eyes or eyes that had 1 of the 3 diseases (DR, AMD, or glaucoma), whereas in clinical practice, an eye may experience a 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. Second, the use of the semisequential structure increased the diagnosis accuracy for glaucoma but also limited the framework for the

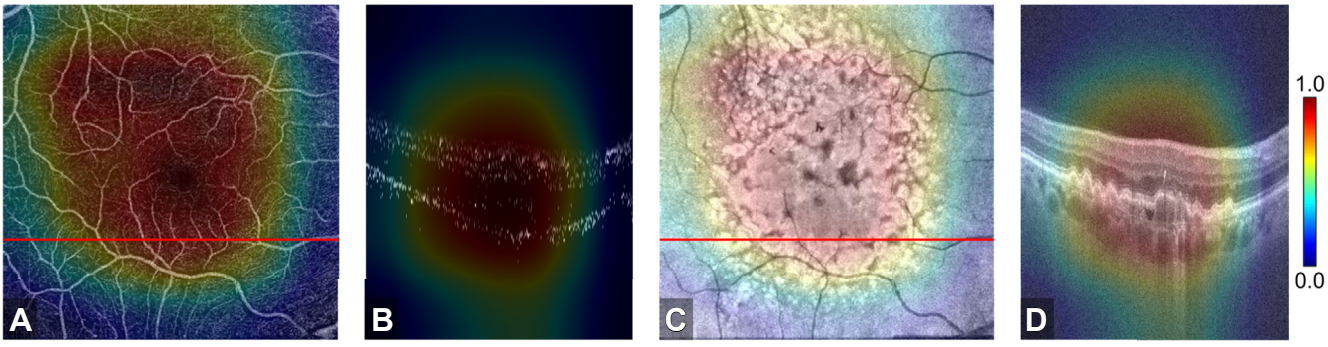

Figure 5. Class activation map based on the age-related macular degeneration output layer of the semisequential classifier for an eye with correctly classified age-related macular degeneration. **A,** OCT angiography (OCTA) en face projection of the superficial vascular complex ([SVC] inner 80% of the 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 ([EZ] the boundary between the outer nuclear layer and the EZ to the boundary between the EZ and the retinal pigment epithelium). **D,** Corresponding B scan at the location of the red line in (**C**). The drusen area was highlighted by the class activation map.

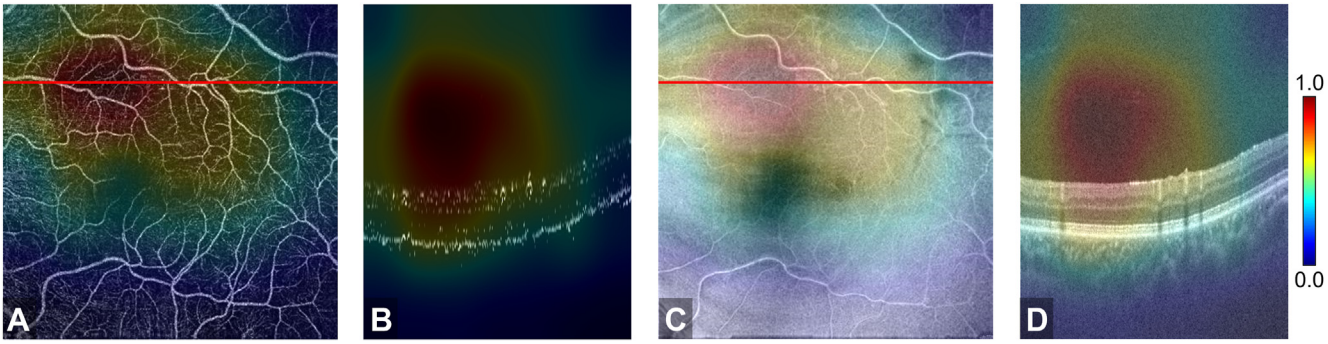

Figure 6. Class activation map based on the glaucoma output layer of the semisequential classifier for an eye with correctly classified glaucoma. **A,** OCT angiography (OCTA) en face projection of the superficial vascular complex ([SVC] inner 80% of the ganglion cell complex). The low-perfusion area is highlighted in (**B**). Corresponding B scan at the position of the red line in (**A**). **C,** Structural OCT en face image of the inner retina (the boundary between the vitreous and the inner limiting membrane to the boundary between the outer plexiform layer and the outer nuclear layer). **D,** Corresponding B scan at the location of the red line in (**C**). The region of the vanished nerve fiber layer was highlighted.

diagnosis of eyes with both glaucoma and DR or glaucoma and AMD. This framework solved the multiple classification problem for a single diagnosis among 3 eye diseases; however, future studies need to generalize our strategy to make multiple simultaneous diagnoses of more diseases or eyes with multiple diseases. Finally, some of the design choices that led to the second limitation were to improve the diagnostic performance for glaucoma (Table 2); however, even so, the sensitivity for the diagnosis of glaucoma (71.67 ± 4.08%) was lower than that for the other 2 diseases (90.00 ± 2.34% for DR and 88.28 ± 5.60% for AMD). Because only scans of the macula were used in this study, information from the optic disc, where the pathology of glaucoma is more prominent,[50] was unavailable for decision making. Training on a larger data set with cases of multiple diseases would likely improve performance for not only glaucoma but also the other diseases in this study. In particular, the accuracy of the parallel classifier could probably be similar to that of the semisequential classifier in the main module if more glaucoma data for training were available. This limitation of the framework could, therefore, be solved using a better-trained parallel classifier.

In addition to diagnostic performance, currently, there are also limitations if we use our framework in real-world clinical applications. Our framework can only be used in clinics with both OCT and OCTA available. However, this limitation will gradually disappear as OCTA applications become more widespread. In addition, all the data sets used in this study were scanned using Avanti RTVue-XR at Casey Eye Institute, Oregon Health & Science University, and only scans with a signal strength index of > 50 were preserved. The diagnostic performance may be lower with 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.

Conclusions

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

Footnotes and Disclosures

Originally received: June 30, 2022.
Final revision: October 21, 2022.
Accepted: October 28, 2022.
Available online: November 9, 2022.Manuscript no. XOPS-D-22-00153R2.

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

Disclosures:

All authors have completed and submitted the ICMJE disclosures form.

The authors have made the following disclosures:

Y.J.: Financial interest – Visionix, Inc, Optos, Inc. These potential conflicts of interest have been reviewed and managed by Oregon Health & Science University.

D.H.: Financial interest – Visionix, Inc.

Supported by National Institutes of Health (R01 EY027833, R01 EY024544, R01 EY031394, R01 EY023285, T32 EY023211, UL1TR002369, P30 EY010572).

Supported by Unrestricted Departmental Funding Grant and William & Mary Greve Special Scholar Award from Research to Prevent Blindness (New York, New York).

Supported by Bright Focus Foundation (G2020168).

HUMAN SUBJECTS: Human subjects were included in this study. Participants were enrolled after informed consent in accordance with an institutional review board-approved protocol. This study was conducted in compliance with the Declaration of Helsinki and Health Insurance Portability and Accountability Act.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Zang, Jia

Data collection: Zang

Analysis and interpretation: Zang, Hwang, Bailey, Huang

Obtained funding: Jia, Hwang; study was performed as part of regular employment duties at Oregon Health & Science University. No additional funding was provided

Overall responsibility: Zang, Hormel, Hwang, Bailey, Huang, Jia

Abbreviations and Acronyms:
3D = 3-dimensional; **AMD** = age-related macular degeneration; **AUC** = area under the curve; **CAM** = class activation map; **DR** = diabetic retinopathy; **OCTA** = OCT angiography; **ROC** = receiver operating characteristic.

Keywords:
Deep learning, OCT, Diabetic retinopathy, Age-related macular degeneration, Glaucoma.

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

References

1. Teo ZL, Tham YC, Yu M, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. *Ophthalmology*. 2021;128:1580−1591.
2. Beagley J, Guariguata L, Weil C, Motala AA. Global estimates of undiagnosed diabetes in adults. *Diabetes Res Clin Pract*. 2015;103:150−160.
3. Klaver CC, Wolfs RC, Assink JJ, et al. Genetic risk of age-related maculopathy: population-based familial aggregation study. *Arch. Ophthalmol*. 1998;116:1646−1651.
4. Rosenfeld PJ, Brown DM, Heier JS, et al. Ranibizumab for neovascular age-related macular degeneration. *N Engl J Med*. 2006;355:1419−1431.
5. Tham YC, Li X, Wong TY, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014;121:2081−2090.
6. Mitchell P, Liew G, Gopinath B, Wong TY. Age-related macular degeneration. *Lancet*. 2018;392:1147−1159.
7. Kass MA, Heuer DK, Higginbotham EJ, et al. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. *Arch Ophthalmol*. 2002;120:701−713.
8. Wong TY, Sun J, Kawasaki R, et al. Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology*. 2018;125:1608−1622.
9. Paz SH, Varma R, Klein R, et al. Noncompliance with vision care guidelines in Latinos with type 2 diabetes mellitus: the Los Angeles Latino Eye Study. *Ophthalmology*. 2006;113:1372−1377.
10. Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science*. 1991;254:1178−1181.
11. Makita S, Hong Y, Yamanari M, et al. Optical coherence angiography. *Opt Express*. 2006;14:7821−7840.
12. An L, Wang RK. In vivo volumetric imaging of vascular perfusion within human retina and choroids with optical micro-angiography. *Opt Express*. 2008;16:11438−11452.
13. Jia Y, Tan O, Tokayer J, et al. Split-spectrum amplitude-decorrelation angiography with optical coherence tomography. *Opt Express*. 2012;20:4710−4725.
14. Jia Y, Bailey ST, Hwang TS, et al. Quantitative optical coherence tomography angiography of vascular abnormalities in the living human eye. *Proc Natl Acad Sci*. 2015;112:E2395−E2402.
15. Hwang TS, Jia Y, Gao SS, et al. Optical coherence tomography angiography features of diabetic retinopathy. *Retina*. 2015;35:2371.
16. Adhi M, Duker JS. Optical coherence tomography−current and future applications. *Curr Opin Ophthalmol*. 2013;24:213−221.
17. Yehoshua Z, Rosenfeld PJ, Gregori G, Penha F. Spectral domain optical coherence tomography imaging of dry age-related macular degeneration. *Ophthalmic Surg Lasers Imaging Retina*. 2010;41:S6−S14.
18. Jia Y, Bailey ST, Wilson DJ, et al. Quantitative optical coherence tomography angiography of choroidal neovascularization in age-related macular degeneration. *Ophthalmology*. 2014;121:1435−1444.
19. You QS, Wang J, Guo Y, et al. Detection of reduced retinal vessel density in eyes with geographic atrophy secondary to age-related macular degeneration using projection-resolved optical coherence tomography angiography. *Am J Ophthalmol*. 2020;209:206−212.
20. Mwanza JC, Oakley JD, Budenz DL, et al. Ability of cirrus HD-OCT optic nerve head parameters to discriminate normal from glaucomatous eyes. *Ophthalmology*. 2011;118:241−248.
21. Liu L, Jia Y, Takusagawa HL, et al. Optical coherence tomography angiography of the peripapillary retina in glaucoma. *JAMA Ophthalmol*. 2015;133:1045−1052.
22. Liu L, Tan O, Ing E, et al. Sectorwise visual field simulation using optical coherence tomographic angiography nerve fiber layer plexus measurements in glaucoma. *Am J Ophthalmol*. 2020;212:57−68.
23. Centers for Medicare & Medicaid Services. Medicare provider utilization and payment data: physician and other supplier. https://data.cms.gov/provider-summary-by-type-of-service/medicare-physician-other-practitioners; 2022.
24. Hormel TT, Jia Y, Jian Y, et al. Plexus-specific retinal vascular anatomy and pathologies as seen by projection-resolved optical coherence tomographic angiography. *Prog Retin Eye Res*. 2021;80:100878.
25. Hormel TT, Huang D, Jia Y. Artifacts and artifact removal in optical coherence tomographic angiography. *Quant Imaging Med Surg*. 2021;11:1120.

26. Hazin R, Barazi MK, Summerfield M. Challenges to establishing nationwide diabetic retinopathy screening programs. *Curr Opin Ophthalmol*. 2011;22:174−179.
27. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436−444.
28. Sandhu HS, Eltanboly A, Shalaby A, et al. Automated diagnosis and grading of diabetic retinopathy using optical coherence tomography. *Investig Ophthalmol Vis Sci*. 2018;59:3155−3160.
29. Sandhu HS, Eladawi N, Elmogy M, et al. Automated diabetic retinopathy detection using optical coherence tomography angiography: a pilot study. *Brit J Ophthalmol*. 2018;102:1564−1569.
30. Alam M, Zhang Y, Lim JI, et al. Quantitative optical coherence tomography angiography features for objective classification and staging of diabetic retinopathy. *Retina*. 2020;40:322−332.
31. Heisler M, Karst S, Lo J, et al. Ensemble deep learning for diabetic retinopathy detection using optical coherence tomography angiography. *Transl Vis Sci Technol*. 2020;9:20.
32. Le D, Alam M, Yao CK, et al. Transfer learning for automated OCTA detection of diabetic retinopathy. *Transl Vis Sci Technol*. 2020;9:35−35.
33. Zang P, Gao L, Hormel TT, et al. DcardNet: diabetic retinopathy classification at multiple levels based on structural and angiographic optical coherence tomography. *IEEE Trans Biomed Eng*. 2021;68:1859−1870.
34. Venhuizen FG, van Ginneken B, van Asten F, et al. Automated staging of age-related macular degeneration using optical coherence tomography. *Investig Ophthalmol Vis Sci*. 2017;58:2318−2328.
35. Peng Y, Dharssi S, Chen Q, et al. DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs. *Ophthalmology*. 2019;126:565−575.
36. Thompson AC, Jammal AA, Medeiros FA. A review of deep learning for screening, diagnosis, and detection of glaucoma progression. *Transl Vis Sci Technol*. 2020;9:42.
37. Thompson AC, Jammal AA, Berchuck SI, et al. Assessment of a segmentation-free deep learning algorithm for diagnosing glaucoma from optical coherence tomography scans. *JAMA Ophthalmol*. 2020;138:333−339.
38. Asaoka R, Murata H, Hirasawa K, et al. Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. *Am J Ophthalmol*. 2019;198:136−145.
39. Mariottoni EB, Jammal AA, Berchuck SI, et al. An objective structural and functional reference standard in glaucoma. *Sci Rep*. 2021;11:1−10.
40. Kraus MF, Liu JJ, Schottenhamml J, et al. Quantitative 3D-OCT motion correction with tilt and illumination correction, robust similarity measure and regularization. *Biomed Opt Express*. 2014;5:2591−2613.
41. Zhang M, Hwang TS, Campbell JP, et al. Projection-resolved optical coherence tomographic angiography. *Biomed Opt Express*. 2016;7:816−828.
42. Wang J, Zhang M, Hwang TS, et al. Reflectance-based projection-resolved optical coherence tomography. *Biomed Opt Express*. 2017;8:1536−1548.
43. Early Treatment Diabetic Retinopathy Study Research Group. Fundus photographic risk factors for progression of diabetic retinopathy: ETDRS report number 12. *Ophthalmology*. 1991;98:823−833.
44. American Academy of Ophthalmology. International clinical diabetic retinopathy disease severity scale detailed table. 2022. https://scholar.archive.org/work/2l7wnyubzzcd5ki4z3ugjgjqru/access/wayback/http://www.icoph.org:80/dynamic/attachments/resources/diabetic-retinopathy-detail.pdf.
45. Flaxel CJ, Adelman RA, Bailey ST, et al. Diabetic retinopathy preferred practice pattern. *Ophthalmology*. 2020;127:66−145.
46. Marmor MF, Ravin JG. Fluorescein angiography: insight and serendipity a half century ago. *Arch Ophthalmol*. 2011;129:943e948.
47. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37−46.
48. Takusagawa HL, Liu L, Ma KN, et al. Projection-resolved optical coherence tomography angiography of macular retinal circulation in glaucoma. *Ophthalmology*. 2017;124:1589−1599.
49. Chen CL, Bojikian KD, Wen JC, et al. Peripapillary retinal nerve fiber layer vascular microcirculation in eyes with glaucoma and single-hemifield visual field loss. *JAMA Ophthalmol*. 2017;135:461−468.
50. Weinreb RN, Khaw PT. Primary open-angle glaucoma. *Lancet*. 2004;363:1711−1720.