AMERICAN ACADEMY OF OPHTHALMOLOGY®

Deep Learning Approaches Predict Glaucomatous Visual Field Damage from OCT Optic Nerve Head En Face Images and Retinal Nerve Fiber Layer Thickness Maps

Mark Christopher, PhD,[1] Christopher Bowd, PhD,[1] Akram Belghith, PhD,[1] Michael H. Goldbaum, MD,[1] Robert N. Weinreb, MD,[1] Massimo A. Fazio, PhD,[2] Christopher A. Girkin, MD,[2] Jeffrey M. Liebmann, MD,[3] Linda M. Zangwill, PhD[1]

Purpose: To develop and evaluate a deep learning system for differentiating between eyes with and without glaucomatous visual field damage (GVFD) and predicting the severity of GFVD from spectral domain OCT (SD OCT) optic nerve head images.

Design: Evaluation of a diagnostic technology.

Participants: A total of 9765 visual field (VF) SD OCT pairs collected from 1194 participants with and without GVFD (1909 eyes).

Methods: Deep learning models were trained to use SD OCT retinal nerve fiber layer (RNFL) thickness maps, RNFL en face images, and confocal scanning laser ophthalmoscopy (CSLO) images to identify eyes with GVFD and predict quantitative VF mean deviation (MD), pattern standard deviation (PSD), and mean VF sectoral pattern deviation (PD) from SD OCT data.

Main Outcome Measures: Deep learning models were compared with mean RNFL thickness for identifying GVFD using area under the curve (AUC), sensitivity, and specificity. For predicting MD, PSD, and mean sectoral PD, models were evaluated using R^2 and mean absolute error (MAE).

Results: In the independent test dataset, the deep learning models based on RNFL en face images achieved an AUC of 0.88 for identifying eyes with GVFD and 0.82 for detecting mild GVFD significantly ($P < 0.001$) better than using mean RNFL thickness measurements (AUC = 0.82 and 0.73, respectively). Deep learning models outperformed standard RNFL thickness measurements in predicting all quantitative VF metrics. In predicting MD, deep learning models based on RNFL en face images achieved an R^2 of 0.70 and MAE of 2.5 decibels (dB) compared with 0.45 and 3.7 dB for RNFL thickness measurements. In predicting mean VF sectoral PD, deep learning models achieved high accuracy in the inferior nasal ($R^2 = 0.60$) and superior nasal ($R^2 = 0.67$) sectors, moderate accuracy in inferior ($R^2 = 0.26$) and superior ($R^2 = 0.35$) sectors, and lower accuracy in the central ($R^2 = 0.15$) and temporal ($R^2 = 0.12$) sectors.

Conclusions: Deep learning models had high accuracy in identifying eyes with GFVD and predicting the severity of functional loss from SD OCT images. Accurately predicting the severity of GFVD from SD OCT imaging can help clinicians more effectively individualize the frequency of VF testing to the individual patient. *Ophthalmology 2020;127:346-356 © 2019 by the American Academy of Ophthalmology*

Glaucoma is a leading cause of blindness that is characterized by retinal ganglion cell death along with associated structural changes of the optic nerve head (ONH) and macula and loss of visual function.[1] Early detection and monitoring of glaucoma are critical to prevent irreversible loss of vision.[2] Over the past decade, spectral domain OCT (SD OCT) imaging has become the standard modality for evaluating glaucomatous structural damage of the ONH and parapapillary region because it can provide clinicians with objective, quantitative measurements of glaucoma-related retinal structures. Circumpapillary retinal nerve fiber layer (RNFL) thickness is commonly used by clinicians to diagnose glaucoma and estimate the rate of disease progression. Standard automated perimetry visual field (VF) testing is the standard of care for monitoring visual function in glaucoma. However, the subjectivity of the test, variability of results, and confounding effect of age-related changes in visual function can limit the utility of VF testing to detect glaucoma and accurately measure functional loss.[3-5] In addition, administering VF tests can be a time-consuming process in which patient fatigue and inattention can contribute to unreliable results and the need for additional testing.[6]

Over the past several years, the application of deep learning techniques to prediction tasks in ophthalmology has led to advances in automated disease detection. These

https://doi.org/10.1016/j.ophtha.2019.09.036
ISSN 0161-6420/19

include the development of models to detect diabetic retinopathy and glaucoma using fundus images.[7-10] There have also been a number of recent reports describing the application of deep learning models to SD OCT in diagnosis and segmentation tasks.[11-14] However, there is little work in applying deep learning models to predict function from structure in glaucoma. Given their success in identifying disease from fundus and SD OCT imaging, deep learning approaches may help to improve our understanding of the relationship between structure and function in glaucoma. Previous work has investigated models that predict function from structure. However, the accuracy has been limited, and model performance can be highly dependent on model assumptions about the linearity of the relationship.[15-22] Furthermore, deep learning—based techniques that accurately predict the severity of VF loss from SD OCT would also help clinicians more effectively target the frequency of VF testing to the individual patient with the possibility of reducing unnecessary and time-consuming VF testing in eyes that are predicted to be stable.

The aim of this study is to develop and evaluate a deep learning system to identify eyes with glaucomatous visual field damage (GFVD) and predict the severity of GFVD using SD OCT imaging of the parapapillary retina. With a large database of SD OCT images and VF testing, we trained deep learning models to use SD OCT imaging to (1) identify eyes with GFVD and (2) estimate the severity of glaucomatous damage as measured by mean deviation (MD), pattern standard deviation (PSD), and sectoral pattern deviation (PD).

Methods

Data Collection

The cohort included 1081 eyes from 665 participants without repeatable GFVD (GFVD−) and 828 eyes from 529 participants with repeatable GFVD (GFVD+). Participants were followed over the course of several years with semi-annual visits that included SD OCT imaging and VF testing. Study participants were selected from 2 longitudinal studies designed to evaluate structural and functional changes in glaucoma: the African Descent and Glaucoma Evaluation Study (ADAGES clinicaltrials.gov identifier: NCT00221923) and the University of California, San Diego—based Diagnostic Innovations in Glaucoma Study (DIGS, clinicaltrials.gov identifier: NCT00221897).[23] All participants gave written informed consent, and institutional review boards at all sites approved the study methods. All methods adhere to the tenets of the Declaration of Helsinki and to the Health Insurance Portability and Accountability Act. Inclusion in DIGS and ADAGES required that participants met the following criteria at study entry: 20/40 or better best-corrected visual acuity, at least 2 consecutive reliable standard automated perimetry VF tests, spherical equivalent < -6.0 dpt, and intraocular pressure <22 mmHg for healthy participants.[23] For this analysis, inclusion in the GFVD+ group was based on reliable, repeated abnormal VF results with PSD $\geq$ 5% of normal or glaucoma hemifield test outside of normal limits. Participants with GFVD+ were required to have 3 consecutive abnormal results in 24-2 standard automated perimetry testing. The GFVD− group consisted of participants without 3 repeated abnormal VF results. Table 1 summarizes the dataset characteristics.

Table 1. Characteristics of Participants with and without Glaucomatous Visual Field Damage

Parameter	GFVD−	GFVD+	*P* Value
No. of participants	665	529	
No. of eyes	1081	828	
No. SD OCT VF pairs	4261	5504	
VF MD (dB)	−0.04±1.6	−5.2±6.5	<0.001
Age (yrs)	54.8±20.6	58.0±26.1	0.02
Race n (%)			<0.001
European descent	442 (67)	293 (55)	
African descent	173 (26)	203 (38)	
Other	40 (6)	33 (6)	

dB = decibels; GVFD = glaucomatous visual field damage; MD = mean deviation; SD OCT = spectral domain OCT; VF = visual field.

At each visit, ONH-centered cube and circle scans were collected using a Spectralis SD OCT (Heidelberg Engineering GmbH, Heidelberg, Germany). The cube scans consisted of 73 B-scans composed of 768 A-scans, each captured in a 4.5×4.5-mm square centered on the ONH. Images were processed and RNFL segmentation was performed using our custom-designed San Diego Automated Layer Segmentation Algorithm (SALSA) tool.[24,25] By using the SALSA segmentations, RNFL thickness maps and RNFL en face images were extracted from each scan. Although the RNFL does not exist within the optic disc, SALSA does compute a segmentation within this region (Fig 1). For this analysis, we opted to retain the optic disc region rather than masking it to avoid errors or artificial biases that could be introduced by automated optic disc masking and to allow the deep learning models to determine which regions of the RNFL thickness maps and en face images were informative. Confocal scanning laser ophthalmoscopy (CSLO) images captured during SD OCT imaging were also extracted. Each SD OCT scan resulted in a set of three 2-dimensional images: an RNFL thickness map, RNFL en face image, and CSLO image. Figure 1 provides example images for a GVFD+ and GVFD− eye. For comparison, high-resolution RNFL circle scans consisting of 1536 A-scans around a 3.5-mm circle centered on the ONH were processed using standard Spectralis software (version 6.8.1) and evaluated for quality by the Imaging Data Evaluation and Analysis Reading Center according to standard protocols.[23]

The VF testing was performed at each visit for all participants using the Humphrey Field Analyzer II (Carl Zeiss Meditec, Inc, Dublin, CA) standard 24-2 testing pattern using the Swedish interactive thresholding algorithm. Tests that had more than 33% fixation losses, 33% false-negative errors, or 15% false-positive errors were excluded. The VFs were processed and evaluated for quality according to standard protocols by the University of California, San Diego Visual Field Assessment Center.[26] Quantitative VF metrics including MD, PSD, and sectoral PD were exported for all VFs using standard Humphrey Field Analyzer software. The sectors considered were based on the Garway-Heath map and consisted of the central, inferior, inferior nasal, superior, superior nasal, and temporal VF sectors.[27] Figure 2 details these VF sectors.

For each eye, the SD OCT images and VF tests acquired within 30 days were identified and resulted in a set of 9765 SD OCT VF pairs. This dataset was used to perform all subsequent analyses of predicting visual function from SD OCT imaging.

Preprocessing and Data Augmentation

The RNFL thickness maps extracted from each ONH cube scan consisted of a 73×768 matrix where each numeric value indicated

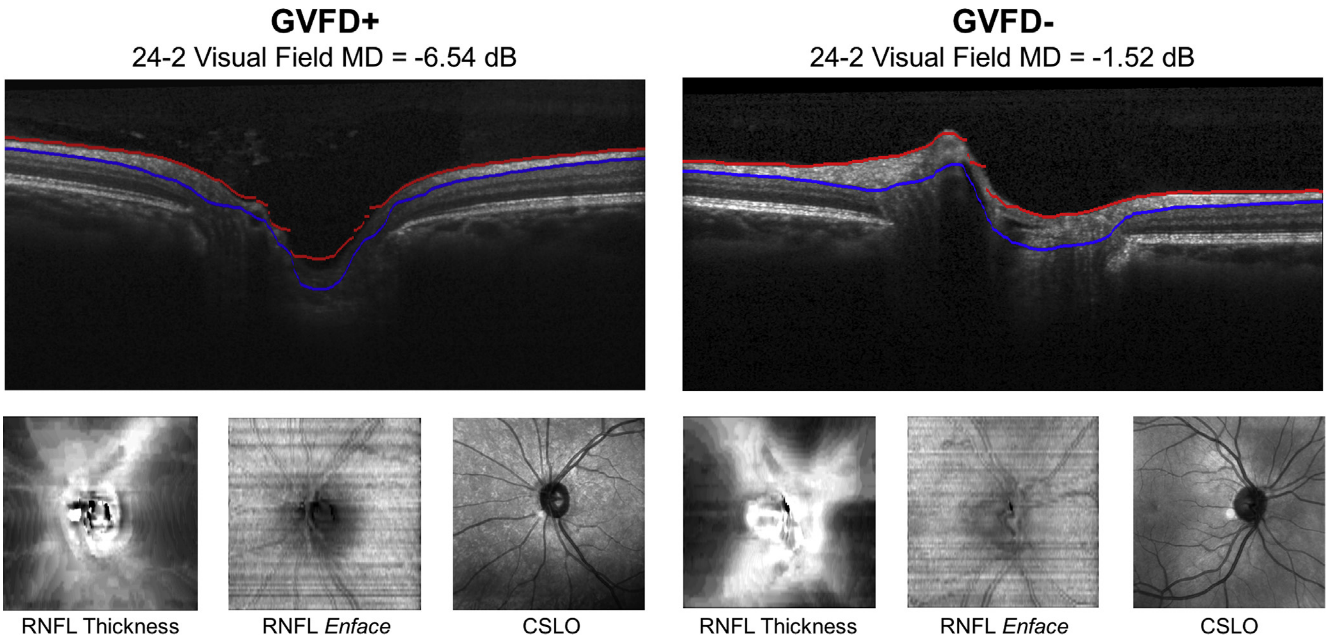

Figure 1. Example images for an eye with glaucomatous visual field (VF) damage and an eye without this damage. A single B-scan from the optic nerve head (ONH) cube scan is shown with the San Diego Automated Layer Segmentation Algorithm (SALSA) retinal nerve fiber layer (RNFL) segmentation illustrated **(top)**. The RNFL thickness map, RNFL en face image, and confocal scanning laser ophthalmoscopy (CSLO) image are also shown **(bottom)**. GVFD = glaucomatous VF damage; MD = mean deviation.

the distance between the inner limiting membrane and the RNFL at the corresponding location in the SD OCT scan. The RNFL en face image also consisted of a 73×768 matrix, but the numeric values in this case were computed by averaging the intensity values of voxels between the inner limiting membrane and RNFL in the corresponding SD OCT location. The CSLO images were captured during image acquisition and simply extracted from the raw image data.

After extracting each of these image types, preprocessing steps were applied to prepare them for application of the deep learning model. In the transfer learning approach adopted, the deep learning model was pretrained on a general image dataset containing images of a canonical size and pixel values scaled to a given numeric range. Each RNFL thickness map, RNFL en face image, and CSLO image was resized to 224×224 pixels and had its pixel values rescaled to the range of 0 to 255 to better match the expected model input.

An augmentation procedure in the form of horizontal mirroring was applied to the RNFL thickness, RNFL en face, and CSLO images. This mirroring was performed to mimic both right and left eye orientations for each image.

In addition to the extracted images, the mean RNFL thickness (mRNFLt) in the ONH region was computed for each cube scan and the mean circumpapillary RNFL thickness (cpRNFLt) was used for each circle scan.

Deep Learning Models and Training

The deep learning architecture used was ResNet50.[28] A transfer learning approach was adopted by initializing model weights by training on a large, general image recognition dataset (ImageNet).[29] Model weights were then further trained and fine-tuned on the SD OCT VF training dataset. A binary classification model was constructed to distinguish between GVFD+ and GVFD− eyes using each image type (RNFL thickness map, RNFL en face image, CSLO image). In addition, deep learning models were constructed to predict quantitative VF measurements (MD, PSD, central, temporal, inferior, inferior nasal, superior, and superior nasal PD) from each image type. To construct independent datasets for training, validation, and testing, the cohort was randomly divided by participant in an 85-5-10 percent split. Splitting by participant (instead of by image) meant that the validation and testing sets did not contain images from any eyes or individuals that were used to train the model. Training consisted of a total of 50 000 iterations with a batch size of 25. Model selection was performed by evaluating the models on the validation set after every 1000 iterations. For each of the 3 image types, the model with the best validation set performance was selected as the final model for evaluation on the testing set.

Caffe tools were used to define the model architecture and perform model training and evaluation. Training and evaluation were performed on a machine running CentOS 6.6 using NVIDIA Tesla K80 graphics cards.[30]

Model Evaluation

Evaluation of the models was performed on the independent test dataset. Performance in distinguishing between GVFD+ and GVFD− eyes was evaluated using sensitivity, specificity, and area under the curve (AUC). To evaluate the models in detecting different severities of GVFD, we defined eyes with an MD > -6.0 decibels (dB) as mild GVFD and those with MD ≤ -6.0 as moderate-to-severe GVFD. For each model, AUCs for detecting any GVFD, mild GVFD, and moderate-to-severe GVFD were computed. The AUCs of different models were statistically compared using a clustered approach[31] to control for multiple images collected from the same participant/eye. To help evaluate clinical utility, the sensitivity of each model at fixed levels of specificity (80%, 85%, 90%, and 95%) was also evaluated. Performance in predicting quantitative measurements (MD, PSD, sectoral PD) was performed using R^2 and mean absolute error (MAE). In all cases, a linear regression model that predicted VF

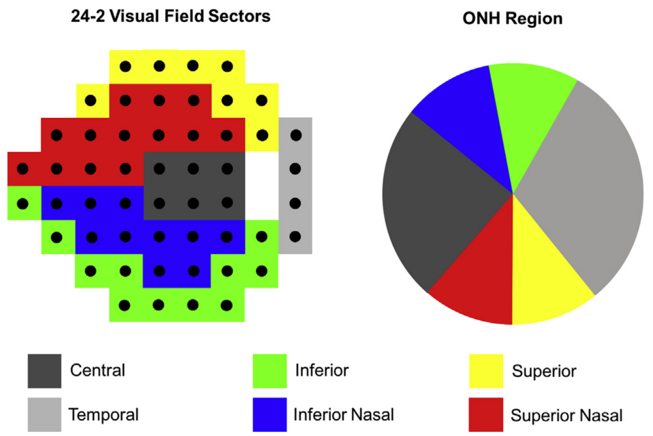

Figure 2. An illustration of the visual field (VF) sectors used in this analysis **(left)** and their corresponding mapping on the optic nerve head (ONH) **(right)**. These sectors are taken from Garway-Heath et al.[27]

quantitative measurements from mRNFLt and cpRNFLt measurements was used as a basis for comparison.

In addition to quantitative performance metrics, model inputs and outputs were also evaluated to help understand the model decision-making process. This was done through the use of deep learning visualization techniques (occlusion testing) and qualitative review of images resulting in correct and incorrect model predictions. The occlusion testing process placed a blank window (20×20 pixels) over a small region of an input image before applying the model. The effect of blanking each image region on model predictions could then be quantified. By repeating this process for all locations, the impact on model predictions could be mapped across the entire input image. By using 100 randomly selected test set images, average occlusion testing maps were generated for each trained deep learning model.

Results

Table 1 provides a summary of the study population included in the analysis. The 665 GVFD− participants (1081 eyes) had an average age of 54.8 years, which was younger than the average age of 58.0 years for the 529 GVFD+ participants (828 eyes) ($P = 0.02$). The GVFD+ eyes had a mean ± standard deviation VF MD of −5.2±6.5 dB compared with −0.04±1.6 dB for the GVFD− eyes ($P < 0.001$).

Identifying Glaucomatous Visual Field Damage

Table 2 and Figure 3 summarize the performance of the deep learning models in identifying GVFD eyes. Individually, the deep learning models based on the RNFL en face image achieved an AUC of 0.88 (95% confidence interval [CI], 0.86−0.90), the RNFL thickness map achieved 0.82 (95% CI, 0.80−0.95), and the CSLO image achieved 0.81 (95% CI, 0.79−0.84). The best performing deep learning model (RNFL en face image) significantly ($P < 0.001$) outperformed the global RNFL thickness measures mRNFLt (AUC, 0.82; 95% CI, 0.79−0.84) and cpRNFLt (AUC, 0.80; 95% CI, 0.77−0.83) in detecting eyes with any level of GVFD. Table 3 summarizes the performance of the models in identifying eyes with different severities of GVFD. The RNFL en face image deep learning model also achieved an AUC of 0.82 (95% CI, 0.79−0.85) in detecting mild GVFD and 0.97 (95% CI, 0.95−0.99) in detecting moderate-to-severe GVFD. The RNFL en face image deep learning model significantly ($P < 0.05$) outperformed both other deep learning models in identifying any GVFD and mild GVFD.

Table 4 provides the full results showing the sensitivity of each model at fixed levels of specificity. At 90% and 95% specificity, the RNFL en face image model achieved sensitivities of 0.72 and 0.68, respectively, the RNFL thickness map model achieved 0.64 and 0.5, respectively, and the CSLO image model achieved 0.58 and 0.48, respectively. The RNFL thickness and en face models achieved better sensitivities at all specificity levels than mRNFLt and cpRNFLt.

Predicting Quantitative Visual Field Measurements

The performance results of the deep learning models in predicting global quantitative VF measurements (MD and PSD) are presented in Table 5. In predicting MD, the best deep learning model was based on RNFL en face images ($R^2 = 0.70$; 95% CI, 0.64−0.74) followed by deep learning models using RNFL thickness maps ($R^2 = 0.63$; 95% CI, 0.57−0.68) and CSLO images ($R^2 = 0.48$; 95% CI, 0.41−0.54). The best performing deep learning model (RNFL en face images) significantly ($P < 0.001$) outperformed RNFL thickness measure predictions mRNFLt ($R^2 = 0.40$; 95% CI, 0.35−0.44) and cpRNFLt ($R^2 = 0.45$; 95% CI, 0.40−0.50). The MAEs in predicting MD for the deep learning models based on RNFL en face, thickness map, and CSLO images were 2.5 dB (95% CI, 2.3−2.7), 2.8 dB (95% CI, 2.6−3.0), and 3.1 dB (95% CI, 2.9−3.4), respectively. In all cases, the errors were significantly ($P < 0.05$) lower than those of mRNFLt (3.8 dB, 95% CI, 3.6−4.1) and cpRNFLt (3.7 dB, 95% CI, 3.4−3.9).

In predicting PSD, the en face deep learning model ($R^2 = 0.61$, 95% CI, 0.55−0.66) again outperformed deep learning models based on thickness maps ($R^2 = 0.56$, 95% CI, 0.48−0.62) and CSLO images ($R^2 = 0.48$, 95% CI, 0.42−0.54). The best performing deep learning model was significantly ($P < 0.05$) better than predictions based on mRNFLt ($R^2 = 0.49$, 95% CI, 0.44−0.54) and cpRNFLt measurements ($R^2 = 0.51$, 95% CI, 0.46−0.56). The MAEs in predicting PSD for the deep learning models based on RNFL en face, thickness map, and CSLO images were 1.5 dB (95% CI, 1.4−1.6), 1.5 dB (95% CI, 1.4−1.6), and 1.9 dB (95% CI, 1.8−2.0), respectively, and were significantly ($P < 0.05$) lower than those based on mRNFLt (2.1 dB, 95% CI, 2.0−2.2) and cpRNFLt (2.1 dB, 95% CI, 2.0−2.2).

Table 2. Diagnostic Accuracy of Deep Learning Model Performance in Identifying Eyes with Glaucomatous Visual Field Damage

Model	AUC (95% CI)	P Value
Deep Learning Models		
RNFL thickness map	0.82 (0.80−0.85)	0.59
RNFL en face image	**0.88 (0.86−0.90)**	<0.001*
CSLO image deep	0.81 (0.79−0.84)	0.70
RNFL Thickness		
mRNFLt	0.82 (0.79−0.84)	−
cpRNFLt	0.80 (0.77−0.83)	0.26

Boldface indicates highest performing model(s) by AUC.
− = not applicable; AUC = area under the curve; CI = confidence interval; cpRNFLt = circumpapillary retinal nerve fiber layer thickness; CSLO = confocal scanning laser ophthalmoscopy; mRNFLt = mean retinal nerve fiber layer thickness; RNFL = retinal nerve fiber layer.
*Significantly ($P < 0.05$) better than mRNFLt AUC.

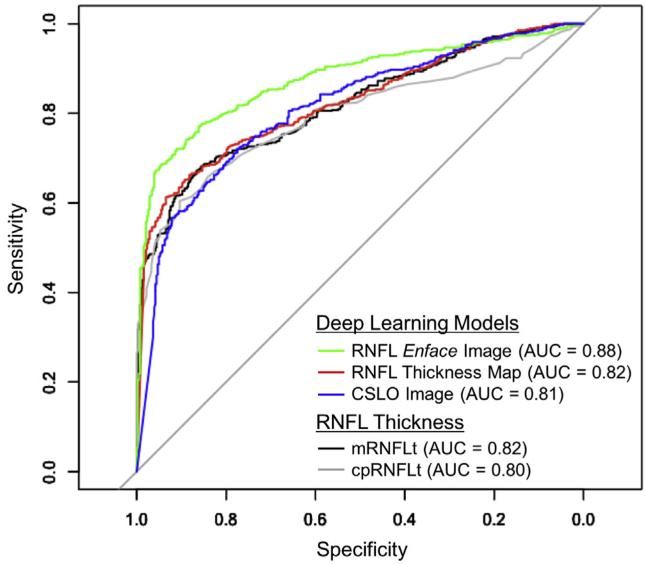

Figure 3. Receiver operating characteristic curves in identifying glaucomatous visual field damage (GVFD) eyes. The deep learning model based on retinal nerve fiber layer (RNFL) en face images achieved the highest area under the curve (AUC) of 0.88, significantly ($P < 0.05$) higher than the any other model. CSLO = confocal scanning laser ophthalmoscopy; cpRNFLt = circumpapillary retinal nerve fiber layer thickness; mRNFLt = mean retinal nerve fiber layer thickness.

The strongest sectoral associations for predicting VF PD from SD OCT were found for the en face deep learning models in the VF superior nasal ($R^2 = 0.67$) and inferior nasal sectors ($R^2 = 0.61$) (Table 6). Moderate associations were achieved by the en face deep learning models for the VF superior ($R^2 = 0.35$) and inferior ($R^2 = 0.26$) sectors. Weaker associations were achieved by en face deep learning models in the central ($R^2 = 0.09$) and temporal ($R^2 = 0.12$) sectors. In all cases, the deep learning models outperformed mRNFLt and cpRNFLt in predicting sectoral PD. In all but 1 sector, the en face deep learning model had the highest performance; the CSLO deep learning model achieved the highest performance in predicting the central sector PD ($R^2 = 0.15$).

Visualizing Models

Occlusion maps highlighted the regions that had the greatest impact on model predictions. An average occlusion map is shown for each image type and VF measurement pair (Fig 4). In the case of RNFL thickness maps, the models predicting MD, PSD, and glaucoma seemed to focus on regions surrounding superior and inferior nerve fiber bundles. In the case of RNFL en face images, models also focused on these regions. In the en face images, however, models seemed to give more weight to inferior regions. For CSLO images, models tended to give weight to inferior regions in glaucoma classification, to superior regions in MD prediction, and to the entire ONH region in PSD prediction. For all image types, model predictions of individual sector PD seemed to be based on known structure–function relationships; inferior ONH region predicted superior VF PD, superior ONH region predicted inferior VF PD, and so forth.

Qualitative review of correct and incorrect model predictions was also performed to help understand model performance. Example RNFL thickness maps, RNFL en face images, and CSLO images for correct and incorrect predictions of GVFD are shown in Figure 5. In the case of the example correct prediction of GVFD+ (Fig 5A), the images show clear thinning/loss of RNFL in the ONH region that could have led to the correct prediction. In the false-positive case (Fig 5C), the images show some asymmetry in the RNFL thickness (thinner RNFL in the superior vs. inferior sector) that may have contributed to the incorrect prediction in this case. In the false-negative example, there did appear to be diffuse thinning; however, no clear focal loss was present. This may have contributed to the model failing to detect GVFD.

Discussion

These results show that deep learning models applied to SD OCT data accurately identify GVFD+ eyes and predict global and sectoral VF quantitative measurements. The deep learning model based on RNFL en face images had a diagnostic accuracy of 0.88 for differentiating between eyes with and without GVFD, which was significantly better than mRNFLt and cpRNFLt measurements (AUCs of 0.82 and 0.80, respectively) and had consistently higher sensitivity at fixed specificities. For detection of mild GVFD, the diagnostic accuracy of the RNFL en face image

Table 3. Diagnostic Accuracy of Deep Learning Models and Retinal Nerve Fiber Layer Thickness for Identifying Glaucomatous Visual Field Damage by Severity of Functional Loss

	AUC in Detecting GVFD		
Model	*All (n = 948)*	*Mild (n = 735)*	*Moderate-to-Severe (n = 595)*
Deep Learning Models			
RNFL thickness map	0.82 (0.80–0.85)	0.74 (0.70–0.77)	**0.97 (0.95–0.98)**
RNFL en face image	**0.88 (0.86–0.90)**	**0.82 (0.79–0.85)**	**0.97 (0.95–0.99)**
CSLO image deep	0.81 (0.79–0.84)	0.75 (0.72–0.79)	0.92 (0.89–0.94)
RNFL Thickness			
mRNFLt	0.82 (0.79–0.84)	0.73 (0.69–0.76)	**0.97 (0.96–0.98)**
cpRNFLt	0.80 (0.77–0.83)	0.70 (0.66–0.74)	**0.97 (0.96–0.98)**

P value compares with mRNFLt AUC. Boldface indicates highest performing model(s) by AUC.
AUC = area under the curve; cpRNFLt = circumpapillary retinal nerve fiber layer thickness; CSLO = confocal scanning laser ophthalmoscopy; GFVD = glaucomatous visual field damage; mRNFLt = mean retinal nerve fiber layer thickness; RNFL = retinal nerve fiber layer.
Mild was defined as GVFD+ eyes with a mean deviation (MD) > −6.0, and moderate-to-severe was defined as GVFD+ eyes with MD ≤ −6.0.

Table 4. Sensitivity of Deep Learning Models and Retinal Nerve Fiber Layer Thickness for Identifying Eyes with Glaucomatous Visual Field Damage at Fixed Levels of Specificity

	Sensitivity in Detecting GVFD			
Model	*80% Specificity*	*85% Specificity*	*90% Specificity*	*95% Specificity*
Deep Learning Models				
RNFL thickness map	0.71	0.68	0.64	0.57
RNFL en face image	**0.80**	**0.78**	**0.72**	**0.68**
CSLO image deep	0.69	0.64	0.58	0.48
RNFL Thickness				
mRNFLt	0.71	0.69	0.62	0.53
cpRNFLt	0.69	0.63	0.60	0.54

Boldface indicates highest performing model(s).
cpRNFLt = circumpapillary retinal nerve fiber layer thickness; CSLO = confocal scanning laser ophthalmoscopy; GVFD = glaucomatous visual field damage; mRNFLt = mean retinal nerve fiber layer thickness; RNFL = retinal nerve fiber layer.

deep learning model (AUC of 0.82) was also significantly ($P < 0.001$) higher than mRNFLt and cpRNFLt measurements (AUCs of 0.73 and 0.70, respectively). Moreover, the SD OCT deep learning models explained a large proportion of the variation in VF MD and PSD with R^2 values of 0.70 and 0.61, respectively, with relatively small errors. Here again, the deep learning models significantly outperformed predictions based on mRNFLt and cpRNFLt measurements.

The ability of these deep learning models to accurately predict the severity of VF damage from SD OCT scans suggests that deep learning models may be used to better individualize the frequency of VF testing to each patient. In some cases, this personalized medicine approach may result in a reduction in the frequency of VF testing, whereas in other cases it may lead to more frequent monitoring of visual function. For example, if the SD OCT deep learning algorithm predicts VF damage to be similar to a patient's last VF test, then the clinician may opt to postpone VF testing to a later visit to save patient/technician time and expense. Postponing VF testing in even a small proportion of patients can lead to large savings for the healthcare system. Alternatively, if the SD OCT deep learning algorithm predicts VF damage to be worse than the patient's last VF test, then the clinician may opt to have VF testing done more frequently, thereby increasing the likelihood of detecting progression and adjusting treatment sooner.

The relationship between structure and function in glaucoma has been extensively studied.[15-20,32,33] Previous results have found anywhere from no relationship to moderate-high correlation between SD OCT–based structural measurements and VF metrics. These previous results include approaches that use machine learning techniques and sophisticated structure–function models defined a priori.[34,35] These approaches also predict VF measurements based on SD OCT segmentations, but do not use the deep learning approaches we have described. These approaches often include assumptions about the linearity of the structure–function relationship, whereas the deep learning methods make no such assumptions. In the case of Guo et al,[34] multilayer segmentations of wide-field SD OCT data were used as input to traditional machine learning classifiers to predict individual VF test points. Specific classifier features were also informed by a priori knowledge of spatial relationships between structure and function data.[35] Even given these advantages (additional layer segmentations, wide-field SD OCT data, incorporation of existing domain knowledge), their results were similar to ours in terms of R^2 (Guo et al $R^2 = 0.74$ vs. our $R^2 = 0.70$), whereas our model resulted in lower mean error across sectors (Guo et al MAE = 5.24 dB vs. our MAE = 2.5 dB). The performance of our approach (as measured by R^2) is higher for our deep learning models than the majority of previous reports.[15,18,32,33] Our deep learning models resulted in significant ($P < 0.05$) associations with global VF metrics (MD, PSD) and all VF sectors (central, temporal, inferior, inferior nasal, superior, superior nasal). The approach described also has an additional advantage compared with previous models: Fewer assumptions about the

Table 5. Performance of Deep Learning Models and Retinal Nerve Fiber Layer Thickness for Predicting Global Visual Field Mean Deviation and Pattern Standard Deviation Measured by R^2 and Mean Absolute Error

	MD		PSD	
Model	R^2 *(95% CI)*	*MAE (dB) (95% CI)*	R^2 *(95% CI)*	*MAE (dB) (95% CI)*
Deep Learning Models				
RNFL thickness map	0.63 (0.57–0.68)	2.8 (2.6–3.0)	0.56 (0.48–0.62)	**1.5 (1.4–1.6)**
RNFL en face image	**0.70 (0.64–0.74)**	**2.5 (2.3–2.7)**	**0.61 (0.55–0.66)**	**1.5 (1.4–1.6)**
CSLO image	0.48 (0.41–0.54)	3.1 (2.9–3.4)	0.48 (0.42–0.54)	1.9 (1.8–2.0)
RNFL Thickness				
mRNFLt	0.40 (0.35–0.44)	3.8 (3.6–4.1)	0.49 (0.44–0.54)	2.1 (2.0–2.2)
cpRNFLt	0.45 (0.40–0.50)	3.7 (3.4–3.9)	0.51 (0.46–0.56)	2.1 (2.0–2.2)

Boldface indicates highest performing model(s).
CI = confidence interval; cpRNFLt = circumpapillary retinal nerve fiber layer thickness; CSLO = confocal scanning laser ophthalmoscopy; dB = decibels; MAE = mean absolute error; MD = mean deviation; mRNFLt = mean retinal nerve fiber layer thickness; PSD = pattern standard deviation; RNFL = retinal nerve fiber layer.

Table 6. Performance of Deep Learning Models and Retinal Nerve Fiber Layer Thickness for Predicting Garway-Heath Visual Field Sectoral Mean Pattern Deviation Measured by R^2 (95% Confidence Interval)

Model	Central	Temporal	Inferior	Inferior Nasal	Superior	Superior Nasal
Deep Learning Models						
RNFL thickness map	0.08 (0.03–0.14)	0.11 (0.08–0.16)	0.06 (0.01–0.24)	0.45 (0.32–0.54)	0.31 (0.22–0.38)	0.52 (0.43–0.59)
RNFL en face image	0.09 (0.01–0.19)	**0.12 (0.01–0.24)**	**0.26 (0.09–0.39)**	**0.60 (0.51–0.67)**	**0.35 (0.25–0.43)**	**0.67 (0.60–0.72)**
CSLO image	**0.15 (0.04–0.25)**	0.08 (0.01–0.17)	0.22 (0.08–0.34)	0.10 (0.01–0.21)	0.19 (0.12–0.26)	0.26 (0.12–0.37)
RNFL Thickness						
mRNFLt	0.07 (0.04–0.10)	0.02 (0.00–0.03)	0.01 (0.00–0.18)	0.28 (0.20–0.34)	0.14 (0.06–0.21)	0.28 (0.24–0.32)
cpRNFLt	0.07 (0.03–0.10)	0.01 (0.00–0.05)	0.02 (0.00–0.22)	0.36 (0.29–0.41)	0.17 (0.08–0.24)	0.31 (0.27–0.35)

Boldface indicates highest performing model(s).
cpRNFLt = circumpapillary retinal nerve fiber layer thickness; CSLO = confocal scanning laser ophthalmoscopy; mRNFLt = mean retinal nerve fiber layer thickness; RNFL = retinal nerve fiber layer.

relationship between structure and function are needed. Structure–function models commonly make assumptions about what structural measurements are relevant (e.g., global vs. local RNFL thicknesses) and the form of the relationship between structure and function (e.g., linear, piecewise, logarithmic).[15,21] Our deep learning approach does not require these assumptions; rather, it identifies informative features and learns their relationship to visual function through training.

Compared with SD OCT imaging, VF testing is a time-consuming and subjective process that often generates noisy, highly variable results.[5,6] Variability in VF testing can mean that multiple testing appointments over the course of several years should be routine to diagnose and monitor glaucoma.[36] Our approach uses deep learning to directly model structure–function relationships and identify eyes with likely functional loss. Because we have built our models using data collected as part of standard glaucoma care (SD OCT imaging and VF results), our tools provide clinicians with predictions of familiar VF metrics (e.g., MD, PSD). These tools can enhance clinicians' understanding of the deep learning predictions and how to incorporate them into their workflow. Tools to estimate visual function from SD OCT scans may be especially relevant because the number of VF tests performed on patients with glaucoma and glaucoma suspects has substantially decreased, whereas the number of SD OCT images collected on these patients has substantially increased over the past several years.[37]

It is interesting to note that the deep learning models based on RNFL en face images outperformed models based on RNFL thickness maps and CSLO images in most cases for identifying both VF defects and predicting quantitative VF measurements. A possible explanation for the performance of the en face images is the additional information provided by the intensity information. En face images are computed by averaging the voxel intensity values within the RNFL. These images encode information that is not available through thickness alone. Previous work has shown that features based on SD OCT voxel intensity and texture features can aid in identifying glaucomatous damage.[38-42] These results show that going beyond thickness measurements, which include both neural and non-neural tissue and incorporating voxel intensity and texture measurements, can help improve a model's ability to identify glaucomatous damage and predict function.

The finer-scale predictions of visual function produced by training models to predict average PD in individual sectors based on the Garway-Heath map varied greatly depending on the sector under consideration. It is not surprising that predictions for sectors with relatively few VF testing points (central and temporal) were poor (R^2 = 0.12–0.15), whereas predictions for sectors with more testing points and in areas in which GFVD is more likely to occur (inferior nasal, superior nasal) were more accurate (R^2 = 0.60–0.67). This discrepancy is likely due to noisier mean sectoral PD in sectors with few VF test points. For all VF sectors, though, deep learning models again outperformed mRNFLt and cpRNFLt predictions (Table 6). Altering the test pattern to include more test points or including another program such as the 10-2 along with the 24-2 program would likely enhance our results.

Occlusion testing of the deep learning models confirmed some expected sectoral relationships between structure and function (Fig 4). For example, to predict function in the inferior and inferior nasal VF sectors, the models relied on structure in the superior ONH region. Likewise, inferior ONH regions were used by deep learning models to predict superior and superior nasal VF sectors.

Study Limitations

This study does have some limitations. One issue is the unknown generalizability of the results presented to other populations. The study population collected as part of ADAGES and DIGS may not be representative of other datasets in terms of age, race, recruitment/collection protocols, or some other unknown confounding variable, and the models may have learned structure–function relationships that are specific to these data. The control participants were recruited using many commonly used methods (e.g., advertisements, staff referrals, nonrelated family member), and this may not represent the same population from whom cases were recruited. However, the ADAGES and DIGS participants do represent a relatively diverse population largely of individuals of European and African descent

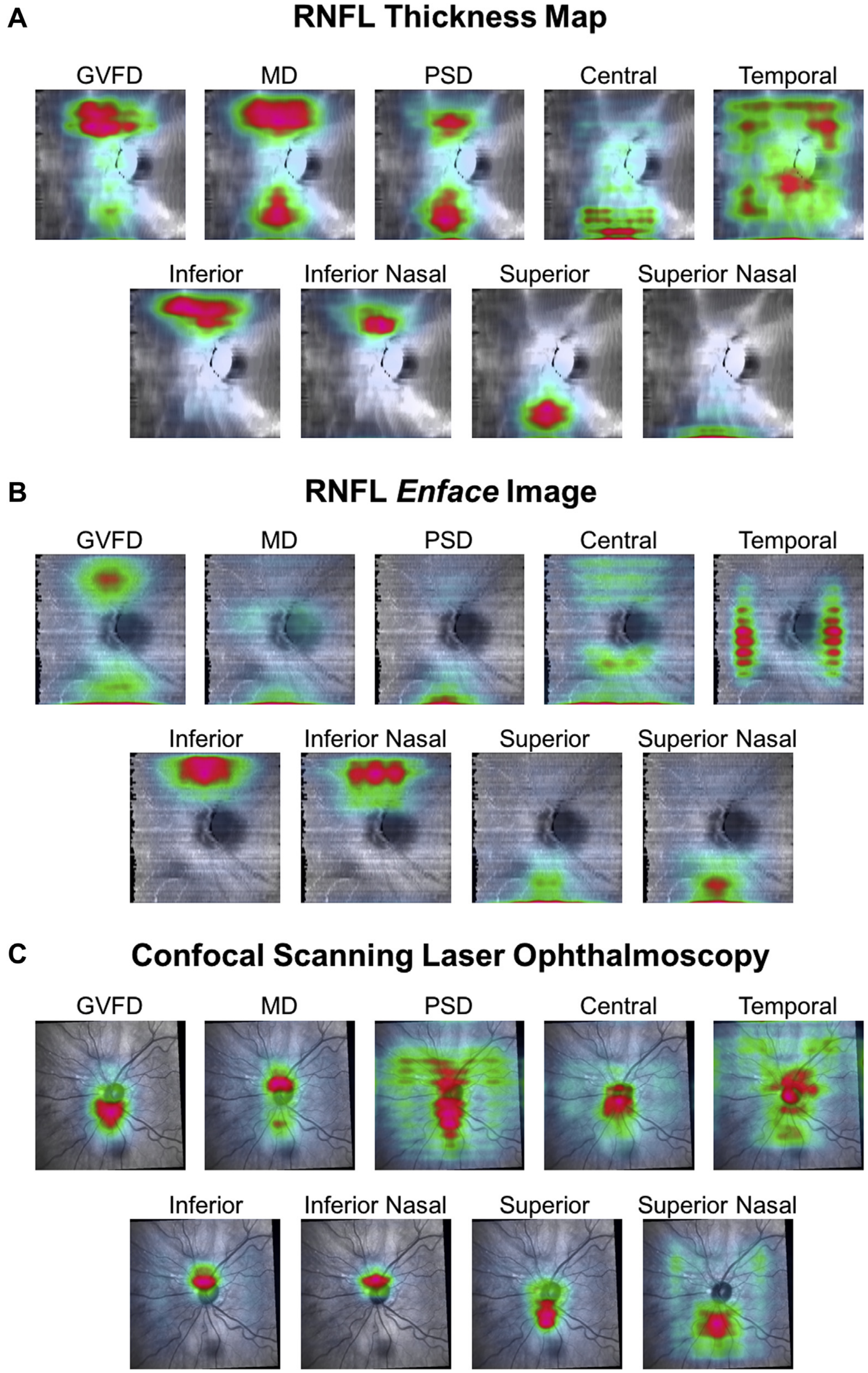

Figure 4. Heat maps created by occlusion testing that highlight informative image regions are shown for deep learning models based on retinal nerve fiber layer (RNFL) thickness maps (**A**), RNFL en face images (**B**), and confocal scanning laser ophthalmoscopy (CSLO) images (**C**). Color intensity indicates the amount of contribution to model classification of glaucomatous visual field damage (GVFD), prediction of global visual field (VF) metrics (mean deviation [MD], pattern standard deviation [PSD]), and sectoral VF pattern deviation (PD) (central, temporal, inferior, inferior nasal, superior, superior nasal).

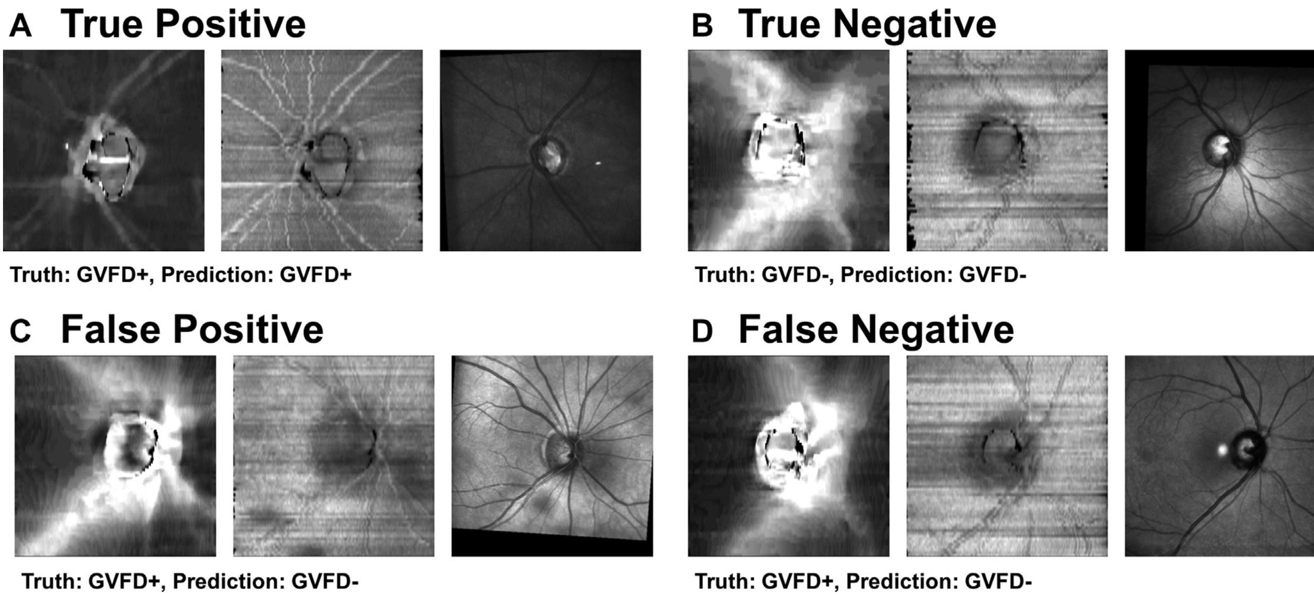

Figure 5. Example retinal nerve fiber layer (RNFL) thickness maps, RNFL en face images, and confocal scanning laser ophthalmoscopy (CSLO) images for which the deep learning models produce predictions resulting in true-positive (**A**), true-negative (**B**), false-positive (**C**), and false-negative (**D**). GVFD = glaucomatous visual field damage.

recruited from 3 different geographic locations in the United States (San Diego, CA; New York City, NY; and Birmingham, AL). This training set diversity should aid the models in generalizing well to other datasets. Although it is unlikely that the relative performance of the various ONH images used as input to the deep learning will be differentially affected by this possible confounding, we cannot rule out the possibility that the models are basing their predictions on some unmeasured confounder that may affect the estimate of the diagnostic performance. Replication on external data will allow us to determine the generalizability of these methods. Another issue is that the input to the models consisted of RNFL thickness maps and en face images extracted from ONH cube scans. The particular ONH scans used for this work were collected as part of ADAGES and DIGS to capture 3-dimensional measurements from across the entire ONH region. Collected from 2009 to 2015, a large number of scans were available for the data-intensive deep learning approach described. These scans have relatively large spacing between B-scans (~61 μm). Using SD OCT scans with higher-density imaging and more densely packed B-scans would likely further improve predictions. In addition, the SD OCT data used here required segmentation of the SD OCT volume before model application. The result is that the models can be sensitive to segmentation failures. If the models are presented with poor-quality RNFL images, they may not produce accurate estimates of visual function. We have previously validated the accuracy of SALSA compared with other segmentation tools.[24,25] However, no tool is perfect, and segmentation errors also exist in commercial instrument software. These errors will continue to be an issue for any approach that relies on these segmentations. Training deep learning models on appropriate datasets (i.e., those that look like real-world clinical data) may help make the models less sensitive to these segmentation errors. The occlusion testing maps (Fig 4) help provide some insight into how these deep learning models make predictions, but much about the model decision making remains unclear. Visualization of deep learning models is an active, ongoing area of research.[43,44] Applying newly developed visualization methods could help reveal detailed, fine-scale information about structure−function relationships on an individual patient basis.

In conclusion, the deep learning models based on SD OCT images had high accuracy in identifying eyes exhibiting GFVD and predicting global VF metrics to estimate the severity of functional loss. The deep learning model based on RNFL en face images did particularly well in both identifying GVFD and predicting VF summary metrics. Their high accuracy suggests that these models may help clinicians estimate visual function from SD OCT imaging and individualize the frequency of VF testing to the individual patient. By predicting VF loss from SD OCT scans, deep learning approaches may help clinicians continue to reduce their reliance on highly variable VF testing. The adoption of SD OCT imaging has transformed the clinical care of glaucoma. Deep learning techniques provide an opportunity to continue this transformation by enhancing the extraction of clinically relevant information from objective, reproducible SD OCT scans.

References

1. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA*. 2014;311:1901−1911.

2. Shaikh Y, Yu F, Coleman AL. Burden of undetected and untreated glaucoma in the United States. *Am J Ophthalmol*. 2014;158:1121−1129.
3. Wu Z, Medeiros FA. Impact of different visual field testing paradigms on sample size requirements for glaucoma clinical trials. *Sci Rep*. 2018;8:4889.
4. Russell RA, Crabb DP, Malik R, Garway-Heath DF. The relationship between variability and sensitivity in large-scale longitudinal visual field data. *Investig Ophthalmol Vis Sci*. 2012;53:5985−5990.
5. Wall M, Woodward KR, Doyle CK, Artes PH. Repeatability of automated perimetry: a comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. *Investig Ophthalmol Vis Sci*. 2009;50: 974−979.
6. Phu J, Khuu SK, Yapp M, et al. The value of visual field testing in the era of advanced imaging: clinical and psychophysical perspectives. *Clin Exp Optom*. 2017;100:313−332.
7. Christopher M, Belghith A, Bowd C, et al. Performance of deep learning architectures and transfer learning for detecting glaucomatous optic neuropathy in fundus photographs. *Sci Rep*. 2018;8:16685.
8. Abramoff M, Erginay A, Folk JC, Niemeijer M. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning improved automated detection of diabetic retinopathy on a publicly available. *Invest Ophthalmol Vis Sci*. 2016;57: 5200−5206.
9. Shibata N, Tanito M, Mitsuhashi K, et al. Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. *Sci Rep*. 2018;8:14665.
10. Li Z, He Y, Keel S, et al. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology*. 2018;125:1199−1206.
11. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med*. 2018;24:1342−1350.
12. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172:1122−1131.
13. Muhammad H, Fuchs TJ, De Cuir N, et al. Hybrid deep learning on single wide-field optical coherence tomography scans accurately classifies glaucoma suspects. *J Glaucoma*. 2017;26:1086−1094.
14. Devalla SK, Chin KS, Mari J-M, et al. A deep learning approach to digitally stain optical coherence tomography images of the optic nerve head. *Invest Ophthalmol Vis Sci*. 2018;59:63−74.
15. Yohannan J, Boland MV. The evolving role of the relationship between optic nerve structure and function in glaucoma. *Ophthalmology*. 2017;124(12S):S66−S70.
16. Russell RA, Malik R, Chauhan BC, et al. Improved estimates of visual field progression using Bayesian linear regression to integrate structural information in patients with ocular hypertension. *Investig Ophthalmol Vis Sci*. 2012;53: 2760−2769.
17. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res*. 2007;26:688−710.
18. Leite MT, Zangwill LM, Weinreb RN, et al. Structure-function relationships using the cirrus spectral domain optical coherence tomograph and standard automated perimetry. *J Glaucoma*. 2012;21:49−54.
19. Rao HL, Zangwill LM, Weinreb RN, et al. Structure-function relationship in glaucoma using spectral-domain optical coherence tomography. *Arch Ophthalmol*. 2011;129: 864−871.
20. Pollet-Villard F, Chiquet C, Romanet J-P, et al. Structure-function relationships with spectral-domain optical coherence tomography retinal nerve fiber layer and optic nerve head measurements: structure-function relationships with SD-OCT. *Invest Ophthalmol Vis Sci*. 2014;55:2953−2962.
21. Chu F-I, Marín-Franch I, Ramezani K, Racette L. Associations between structure and function are different in healthy and glaucomatous eyes. *PLoS One*. 2018;13:1−14.
22. Medeiros FA, Leite MT, Zangwill LM, Weinreb RN. Combining structural and functional measurements to improve detection of glaucoma progression using Bayesian hierarchical models. *Investig Ophthalmol Vis Sci*. 2011;52:5794−5803.
23. Sample PA, Girkin CA, Zangwill LM, et al. The African Descent and Glaucoma Evaluation Study (ADAGES): design and baseline data. *Arch Ophthalmol*. 2009;127:1136−1145.
24. Belghith A, Bowd C, Medeiros FA, et al. Automated segmentation of anterior lamina cribrosa surface: how the lamina cribrosa responds to intraocular pressure change in glaucoma eyes? 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). New York, NY. 2015:222−225.
25. Belghith A, Bowd C, Medeiros FA, et al. Does the location of Bruch's membrane opening change over time? Longitudinal analysis using San Diego Automated Layer Segmentation Algorithm (SALSA). *Investig Opthalmology Vis Sci*. 2016;57:675.
26. Racette L, Liebmann JM, Girkin CA, et al. African Descent and Glaucoma Evaluation Study (ADAGES): III. Ancestry differences in visual function in healthy eyes. *Arch Ophthalmol*. 2010;128:551−559.
27. Garway-Heath DF, Poinoosawmy D, Fitzke FW, Hitchings RA. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology*. 2000;107: 1809−1815.
28. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV; 2016:770−778.
29. Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge. *Int J Comput Vis*. 2015;115: 211−252.
30. Jia Y, Shelhamer E, Donahue J, et al. *Caffe: convolutional architecture for fast feature embedding*. Orlando, FL: *ACM Int Conf Multimed*; 2014:675−678.
31. Obuchowski NA. Nonparametric analysis of clustered ROC curve data. *Biometrics*. 1997;53:567−578.
32. Lopes FS, Matsubara I, Almeida I, et al. Structure-function relationships in glaucoma using enhanced depth imaging optical coherence tomography-derived parameters: a cross-sectional observational study. *BMC Ophthalmol*. 2019;19:52.
33. Asaoka R, Russell RA, Malik R, et al. A novel distribution of visual field test points to improve the correlation between structure-function measurements. *Investig Ophthalmol Vis Sci*. 2012;53:8396−8404.
34. Guo Z, Kwon YH, Lee K, et al. Optical coherence tomography analysis based prediction of Humphrey 24-2 visual field thresholds in patients with glaucoma. *Investig Ophthalmol Vis Sci*. 2017;58:3975−3985.
35. Zhu H, Crabb DP, Schlottmann PG, et al. Predicting visual function from the measurements of retinal nerve fiber layer structure. *Investig Ophthalmol Vis Sci*. 2010;51:5657−5666.
36. Wu Z, Saunders LJ, Daga FB, et al. Frequency of testing to detect visual field progression derived using a longitudinal cohort of glaucoma patients. *Ophthalmology*. 2017;124: 786−792.

37. Stein JD, Talwar N, Laverne AM, et al. Trends in use of ancillary glaucoma tests for patients with open-angle glaucoma from 2001 to 2009. *Ophthalmology*. 2012;119:748–758.
38. Leung CK-S. Retinal Nerve Fiber Layer (RNFL) Optical Texture Analysis (ROTA) for evaluation of RNFL abnormalities in glaucoma. *Invest Ophthalmol Vis Sci*. 2018;59:3497.
39. Belghith A, Medeiros FA, Bowd C, et al. A novel texture-based OCT enface image to detect and monitor glaucoma. *Invest Ophthalmol Vis Sci*. 2016;57(12).
40. Hood DC, De Cuir N, Blumberg DM, et al. A single wide-field OCT protocol can provide compelling information for the diagnosis of early glaucoma. *Transl Vis Sci Technol*. 2016;5:4.
41. Hood DC, Chen MF, Lee D, et al. Confocal adaptive optics imaging of peripapillary nerve fiber bundles: implications for glaucomatous damage seen on circumpapillary OCT scans. *Transl Vis Sci Technol*. 2015;4:12.
42. Fortune B, Burgoyne CF, Cull GA, et al. Structural and functional abnormalities of retinal ganglion cells measured in vivo at the onset of optic nerve head surface change in experimental glaucoma. *Investig Ophthalmol Vis Sci*. 2012;53:3939–3950.
43. Zhang Q, Zhu S. Visual interpretability for deep learning: a survey. *Front Inf Technol Electron Eng*. 2018;19:27–39.
44. Yosinski J, Clune J, Nguyen AM, et al. Understanding neural networks through deep visualization. Deep Learning Workshop, International Conference on Machine Learning (ICML). Paris, France. 2015:12.

Footnotes and Financial Disclosures

Originally received: March 7, 2019.
Final revision: August 5, 2019.
Accepted: September 23, 2019.
Available online: September 30, 2019. Manuscript no. 2019-496.

[1] Hamilton Glaucoma Center, Shiley Eye Institute, Viterbi Family Department of Ophthalmology, UC San Diego, La Jolla, California.

[2] School of Medicine, University of Alabama-Birmingham, Birmingham, Alabama.

[3] Bernard and Shirlee Brown Glaucoma Research Laboratory, Edward S. Harkness Eye Institute, Department of Ophthalmology, Columbia University Medical Center, New York, New York.

Financial Disclosure(s):

The author(s) have made the following disclosure(s): R.N.W.: Consultant – Aerie Pharmaceuticals, Allergan, Eyenovia, Implantdata, Unity; Research support – Heidelberg Engineering, Carl Zeiss Meditec, Centervue, Bausch & Lomb, Genentech, Konan Medical, National Eye Institute, Optos, Optovue, Research to Prevent Blindness.

M.A.F.: Research support – National Eye Institute, EyeSight Foundation of Alabama, Research to Prevent Blindness, Heidelberg Engineering.

C.A.G.: Research support – National Eye Institute, EyeSight Foundation of Alabama, Research to Prevent Blindness, Heidelberg Engineering.

J.M.L.: Consultant – Alcon, Allergan, Bausch & Lomb, Carl Zeiss Meditec, Heidelberg Engineering, Reichert, Valeant Pharmaceuticals, Bausch & Lomb, Carl Zeiss Meditec, Heidelberg Engineering, National Eye Institute, Novartis, Optovue, Reichert Technologies, Research to Prevent Blindness.

L.M.Z.: Research support – Carl Zeiss Meditec, Heidelberg Engineering, National Eye Institute, Optovue, Topcon Medical System Inc.

Financial Support: National Eye Institute: EY11008, P30 EY022589, EY026590, EY022039, EY021818, EY023704, EY029058, T32 EY026590, R21 EY027945, Genentech, Inc. Unrestricted grant from Research to Prevent Blindness (New York, NY).

HUMAN SUBJECTS: Human subjects were included in this study. The human ethics committees at all sites approved the study. All research adhered to the tenets of the Declaration of Helsinki. All participants provided informed consent.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Christopher, Bowd, Belghith, Goldbaum, Weinreb, Fazio, Girkin, Liebmann, Zangwill

Data collection: Christopher, Weinreb, Fazio, Girkin, Liebmann, Zangwill

Analysis and interpretation: Christopher, Bowd, Belghith, Goldbaum, Weinreb, Fazio, Girkin, Liebmann, Zangwill

Obtained funding: Weinreb, Fazio, Girkin, Liebmann, Zangwill

Overall responsibility: Christopher, Bowd, Belghith, Goldbaum, Girkin, Liebmann, Zangwill

Abbreviations and Acronyms:

ADAGES = African Descent and Glaucoma Evaluation Study; **AUC** = area under the receiver operating characteristic curve; **CI** = confidence interval; **cpRNFLt** = circumpapillary retinal nerve fiber layer thickness; **CSLO** = confocal scanning laser ophthalmoscopy; **dB** = decibels; **DIGS** = Diagnostic Innovations in Glaucoma Study; **GFVD** = glaucomatous visual field damage; **GVFD–** = without repeatable glaucomatous visual field damage; **GVFD+** = repeatable glaucomatous visual field damage; **MAE** = mean absolute error; **mRNFLt** = mean retinal nerve fiber layer thickness; **PD** = pattern deviation; **PSD** = pattern standard deviation; **RNFL** = retinal nerve fiber layer; **SALSA** = San Diego Automated Layer Segmentation Algorithm; **SD OCT** = spectral domain OCT; **VF** = visual field.

Correspondence:
Linda M. Zangwill, PhD, University of California, San Diego, 9500 Gilman Drive, #0946, La Jolla, CA 92093-0946. E-mail: lzangwill@ucsd.edu.