Title

Deep Learning Image Analysis of Optical Coherence Tomography Angiography Measured Vessel Density Improves Classification of Healthy and Glaucoma Eyes

Authors

Bowd, Christopher
Belghith, Akram
Zangwill, Linda M
et al.

Publication Date

2022-04-01

DOI

10.1016/j.ajo.2021.11.008

Peer reviewed

HHS Public Access

Author manuscript

Am J Ophthalmol. Author manuscript; available in PMC 2023 March 27.

Published in final edited form as:

Am J Ophthalmol. 2022 April ; 236: 298–308. doi:10.1016/j.ajo.2021.11.008.

Deep learning image analysis of optical coherence tomography angiography measured vessel density improves classification of healthy and glaucoma eyes

Christopher Bowd,

Akram Belghith,

Linda M. Zangwill,

Mark Christopher,

Michael H. Goldbaum,

Rui Fan,

Jasmin Rezapour,

Sasan Moghimi,

Alireza Kamalipour,

Huiyuan Hou,

Robert N. Weinreb

Hamilton Glaucoma Center, Shiley Eye Institute and The Viterbi Family Department of Ophthalmology, University of California, San Diego, La Jolla, CA

Abstract

Purpose: To compare convolutional neural network (CNN) analysis of *en face* vessel density images to gradient boosting classifier (GBC) analysis of instrument provided, feature-based optical coherence tomography angiography (OCTA) vessel density measurements and OCT RNFL thickness measurements for classifying healthy and glaucomatous eyes.

Design: Comparison of diagnostic approaches.

Methods: 130 eyes of 80 healthy individuals and 275 eyes of 185 glaucoma patients with optic nerve head (ONH) OCTA and OCT imaging were included. Classification performance of a VGG16 CNN trained and tested on entire *en face* 4.5 mm × 4.5 mm radial peripapillary capillary OCTA ONH images was compared to performance of separate GBC models trained and tested on standard OCTA and OCT measurements. Five-fold cross-validation was used to test predictions for CNNs and GBCs. Areas under the precision recall curves (AUPRC) were calculated to control for training/test set size imbalance and were compared.

Results: Adjusted AUPRCs for GBC models were 0.89 (95% CI = 0.82, 0.92) for whole image vessel density GBC, 0.89 (0.83, 0.92) for whole Image capillary density GBC, 0.91 (0.88, 0.93) for combined whole Image vessel and whole image capillary density GBC, and 0.93 (0.91, 095) for RNFL thickness GBC. The adjusted AUPRC using CNN analysis of *en face* vessel density

Corresponding Author: Christopher Bowd, Ph.D., Hamilton Glaucoma Center and Shiley Eye Institute, The Viterbi Family Department of Ophthalmology, 9500 Gilman Drive 0946, University of California, San Diego, La Jolla, CA., 92093-0946.

images was 0.97 (0.95, 0.99) resulting in significantly improved classification compared to GBC OCTA-based results and GBC OCT-based results ($P \leq 0.01$ for all comparisons).

Conclusion: Deep learning *en face* image analysis improves on feature-based GBC models for classifying healthy and glaucoma eyes.

Optical coherence tomography angiography (OCTA) is a non-invasive optical imaging technology that provides information about retinal vasculature in the form of vessel density measurements calculated automatically from two subsequent aligned OCT images.[1–3] Several studies have reported good OCTA measurement-based classification of healthy and glaucoma eyes.[4–8] In addition, vessel density measurements are lower in glaucoma suspect and glaucoma eyes compared to healthy eyes. Moreover, vessel density decreases with increasing glaucoma severity.[9] It also has been reported that baseline parapapillary and macular vessel density measurements can predict rapid glaucoma-related thinning of the retinal nerve fiber layer (RNFL) and that vessel density dropout is faster than ganglion cell thinning over two years of follow-up in glaucoma eyes.[9–11]

Over the past two decades, there has been increasing interest in employing machine learning (ML) analyses of structural and functional measurements in ophthalmology. In fact, an automated ML-based photographic diabetic retinopathy diagnostic system recently has gained FDA approval (IDx Technologies IDx-DR).[12] Traditional ML methods have been applied to various imaging and visual function measurements to improve the classification of healthy, glaucoma suspect and glaucomatous eyes compared to instrument based parameters in a large number of studies.[13–35] In addition, several studies also have used traditional and deep ML analyses of OCTA measurements to detect diabetic retinopathy and other ocular pathologies.[36–40] However, relatively few studies have investigated ML analysis of OCTA measurements for detecting glaucoma.

Recently, deep learning approaches, including convolutional neural networks (CNNs),have been employed to classify fundus photographs and optical coherence tomography images from glaucoma eyes and to predict structural and visual field defects in those eyes.[20, 21, 41–54] The current study sought to employ and compare CNNs trained and tested on entire radial peripapillary capillary (RPC) *en face* OCTA vessel density images (the RPC network extends from the internal limiting membrane to the lower boundary of the RNFL) to supervised gradient boosting classification (GBC) models trained and tested on individual OCTA measurements to correctly classify healthy and glaucoma eyes. The current study builds on a previous study from our group that determined that GBCs trained and tested on OCTA vessel density measurements improved disease classification compared to vessels density measurements alone.[55]

Because deep learning models tend to outperform traditional ML models for classification performance[56], we hypothesized that CNN analysis of *en face* OCTA images would outperform GBC analysis of OCTA measurements. We also compared the performance of CNNs to GBC model predictions based on OCT RNFL thickness measurements and standard feature-based (i.e., instrument specific) univariate OCTA and OCT measurements.

MATERIALS AND METHODS

This cross-sectional comparison of diagnostic techniques involved a group of patients with primary open-angle glaucoma (POAG, as defined below) and a group of healthy control participants from the Diagnostic Innovations in Glaucoma Study (DIGS) (ClinicalTrials.gov identifier: NCT00221897). The DIGS is an ongoing prospective, longitudinal study at the Hamilton Glaucoma Center, University of California, San Diego, designed to evaluate anatomical structure in glaucoma. Details of the DIGS protocol have been described elsewhere.[57] All methods adhered to the tenets of the Declaration of Helsinki and the Health Insurance Portability and Accountability Act and were approved by the Institutional Review Board of the University of California, San Diego.

Participants

Eligible participants had best corrected visual acuity of 20/40 or better and open angles on gonioscopy at study entry. All participants were older than 40 years. Participants were excluded if they had a history of intraocular surgery (except for uncomplicated cataract or uncomplicated glaucoma surgery). Eyes with coexisting retinal disease, uveitis, or non-glaucomatous optic neuropathy also were excluded. Diabetic participants with no evidence of retinal involvement were included.

Study group inclusion was determined at the participant level. Healthy participants had healthy appearing optic discs and RNFL OU based on masked assessment of digital stereoscopic photographs with no history of repeatable abnormal VF results (HFA II with 24–2 testing using the Swedish Interactive Thresholding Algorithm, Carl Zeiss Meditec Inc., Dublin, CA) and no history of elevated intraocular pressure (IOP) (all IOP ≤ 21 mm Hg) in either eye. Normal VFs were defined as those with MD and pattern standard deviation (PSD) > 5% and a Glaucoma Hemifield Test (GHT) result within normal limits.

Glaucoma patient participants had ≥ 2 consecutive and reliable (defined below) VF examinations with either PSD ≤ 5% or a GHT result outside of the 99% normal limits with similar patterns of glaucoma-related defects in consecutive exams as determined by investigator assessment in at least one eye.

Visual Field Testing

All VFs were evaluated by UC San Diego Visual Field Assessment Center (VisFACT) personnel based on a standardized protocol.[57] VF usability was assessed in a masked fashion from de-identified research participant data. VisFACT personnel includes Research Associates with 10+ years of VisFACT experience, MD glaucoma Research Fellows and Ph.D. glaucoma Research Scientist faculty. Visual fields with more than 33% fixation losses or more than 33% false-positive errors were automatically excluded. Visual fields exhibiting a learning effect (i.e., initial tests with reduced sensitivity followed by consistent improvement in a series of tests) also were excluded. Visual fields were further reviewed for lid and rim artifacts, fatigue effects, evidence that the visual field results were due to a disease other than glaucoma (e.g., homonymous hemianopia), and inattention. Test results indicating these characteristics were excluded.

Optical Imaging

The commercially available Avanti Angiovue OCTA (Optovue Inc. Fremont, CA, software version 2017.1.0.151) was used to obtain images of the optic nerve head used for analysis in the current study. The Avanti system for measuring vessel density and tissue thickness has been described previously.[6] Briefly, vessel density is defined as the proportion of measured area occupied by flowing blood vessels defined as pixels having decorrelation values acquired by the split-spectrum amplitude-decorrelation angiography (SSADA) algorithm above the threshold level.[58] Vessel density in the peripapillary RNFL was assessed within a 4.5 × 4.5 mm field of view centered on the ONH. Vessel density within the RNFL was measured from the internal limiting membrane (ILM) to the RNFL posterior boundary using standard AngioVue software, and ONH *en face* whole image vessel density and circumpapillary vessel density were obtained over the entire 4.5 × 4.5 mm scan. According to the DIGS protocol, OCTA images were obtained from either undilated or dilated eyes. This is because annual visits include a dilated exam in addition to optical imaging and semi-annual visits do not.

OCTA image quality review was completed according to the UC San Diego Imaging Data Evaluation and Analysis (IDEA) Reading Center standard protocol. Image usability was assessed in a masked fashion from de-identified research participant data. IDEA Center personnel includes Research Associates with 10+ years of IDEA Center experience and M.D. glaucoma Research Fellows. Images with a quality index (QI) < 4, poor clarity, residual motion artifacts visible as irregular vessel patterns or disc boundaries on the en face angiogram, image cropping or local weak signal due to vitreous opacity, or segmentation errors that could not be corrected were excluded.

The Spectralis SD OCT (Spectralis HRA+OCT, software version 5.4.7.0; Heidelberg Engineering, Inc.) was used to measure circumpapillary RNFL (cpRNFL) thickness. Spectralis OCT incorporates a real-time eye-tracking system that couples a confocal laser-scanning ophthalmoscope and SD OCT scanners to adjust for eye movements and to ensure that the same location of the retina is scanned over time. The cpRNFL thickness was measured using a high-resolution RNFL circle scan composed of 1,536 A-scan points from a 12-degree circle centered on the optic disc. OCT images were obtained from either undilated or dilated eyes.

Quality review of Spectralis images was completed according to IDEA Center standard protocol. To be included, all OCT images required a signal strength > 15 dB and were deemed acceptable quality for use based on subjective assessment.

Convolutional Neural Network (CNN)

We fine-tuned the VGG16 model[59] to detect glaucoma using OCTA-derived *en face* vessel density images. The CNN architecture VGG16 was utilized in the present study with weights pre-trained on the ImageNet dataset for fundus feature extraction using TensorFlow 1.15.0 as a back-end for Keras 2.1.6 on Python 3.6.5 for Linux. The VGG16 network weights were frozen in the first four convolution blocks. Weights in the final convolutional block were allowed to update during fine-tuning. The two fully connected (FC) layers (FC1

and FC2) were initialized with new values and allowed to update during fine-tuning. FC1 consisted of 256 units using the ReLU activation function. FC2 used the softmax function to provide the final glaucoma likelihood prediction. An Adam optimizer with a learning rate of 1e−5 and dropout rate of 50% for FC1 was used to train the network.

Gradient Boosting Classifier (GBC) Model

The GBC is an ensemble classifier that attempts to decrease error by resampling and varying the weights for individual weak learners in order to increase classification accuracy.[60] In an empirical comparison study of supervised learning algorithms[61] comparing random forests and boosted decision trees, gradient boosting algorithms had the best overall performance with random forests being close second.

Because our CNN was trained and tested on ONH images only, the current study only considered OCTA-based GBCs composed of whole image vessel density (wiVD) and whole image capillary density (wiCD) parameters for comparison. We trained and tested 3 vessel density-based GBC models to compare to CNN analysis of *en face* OCTA images. The first GBC model, wiVD GBC, included whole image, superior hemiretina and inferior hemiretina vessel densities. The second GBC model, wiCD GBC, included whole image, superior hemiretina and inferior hemiretina capillary densities. The third GBC model, wiVD and CD GBC, included whole image, superior hemiretina and inferior hemiretina vessel and capillary densities.

To compare the classification success of GBCs trained and tested on vessel density measurements to GBCs trained and tested on retinal tissue thickness measurements, we also employed a fourth GBC model that included global cpRNFL, nasal RNFL, inferior RNFL, superior RNFL, and temporal RNFL thickness measurements (called cpRNFL GBC) obtained using Spectralis OCT. All described OCTA and OCT measurements are included in standard clinical printouts and can be automatically exported from the Avanti and Spectralis instruments.

Training and Evaluation

Five-fold cross-validation was used to provide out-of-sample predictions for deep learning CNN and GBC models to avoid over optimistic estimates of classification accuracy. Both healthy and glaucoma eyes were randomly divided at the patient level into 5 subsets. Then, for each model, we used four subsets to train the model and then we used the fifth subset to assess model performance. This sequence was repeated 5 times, with each subset serving as the test set one time, so that each tested eye was never part of its own training set and was tested only once. Each training subset included 230 (95% CI = 211, 240) eyes of 148 glaucoma patients and 102 (95% CI = 94, 110) eyes of 60 healthy individuals. The testing subset included 45 (95% CI = 42, 50) eyes of 37 glaucoma patients and 28 (95% CI = 26, 34) eyes of 20 healthy individuals.

Statistical Analyses

Descriptive statistics included mean and standard deviation for normally distributed variables and median, first quartile, and third quartile values for non-normally distributed

variables. Student's t-tests or Mann-Whitney tests were used to evaluate demographic and clinical differences between glaucoma patients and healthy individuals.

Areas under the precision-recall curves (AUPRC) curves and sensitivities at fixed specificities were used to assess the ability to discriminate eyes with glaucoma from healthy eyes to control for training/test set size imbalance[62]. As measurements from both eyes of the same subject are likely to be correlated, the cluster of data for the study subject were considered as the unit of resampling and bias corrected SEs were calculated. AUPRCs were covariate adjusted for inclusion of both eyes and for age, image quality, axial length and pseudophakia as suggested by Janes and Pepe[63], and compared statistically using the Wald test based on clustered bootstrap covariance.[64]

Statistical analyses were performed using Stata 14.2 (StataCorp LLC, College Station, TX). P values less than 0.05 were considered statistically significant.

RESULTS

Images from 130 eyes of 80 healthy individuals and 275 eyes of 185 glaucoma patients were included in the analyses. A summary of the demographic variables and measurements of the healthy participants and glaucoma patients are shown in Table 1. Glaucoma patients were significantly older ($P < 0.001$) than healthy individuals and glaucoma eyes had worse VF mean deviation (MD) ($P < 0.001$) than healthy eyes.

For univariable analyses, the adjusted AUPRCs for classifying healthy and glaucoma eyes were 0.85 (95% CI = 0.76, 0.88) for wiVD, 0.86 (95% CI = 0.79, 0.88) for wiCD and 0.86 (95% CI = 0.81, 0.87) for cpRNFL thickness.

Overall, GBC model performance was somewhat better than univariable analyses. The adjusted AUPRCs were 0.89 (95% CI = 0.82, 0.92) for the wiVD GBC model, 0.89 (95% CI = 0.83, 0.92) for the wiCD GBC model, 0.91 (95% CI = 0.88 0.93) for the combined wiVD and CD GBC model, and 0.93 (95% CI = 0.91, 0.95) for the Spectralis cpRNFL thickness GBC model.

The adjusted AUPRC using CNN analysis of RPC *en face* vessel density images was 0.97 (95% CI = 0.95, 0.99) resulting in improved classification compared to all GBC results and all univariable results ($P < 0.01$ for all comparisons). These values along with sensitivities at fixed specificities of 0.80, 0.85, 0.90 and 0.95 for all analyses are shown in Table 2.

Figure 1 shows examples of *en face* images from a glaucoma eye and a healthy eye classified correctly by CNN analysis and incorrectly by all investigated GBCs.

We also investigated performance of the same models and univariate measurements in a subset of 183 eyes with early glaucoma defined as MD ≥ −6 dB. In this subset analysis, patterns of results were similar but AUPRCs and sensitivities at fixed specificities were lower, as expected. For univariable analyses, the adjusted AUPRCs for classifying healthy and glaucoma eyes were 0.81 (95% CI = 0.77, 0.83) for wiVD, 0.82 (95% CI = 0.78, 0.85) for wiCD and 0.84 (95% CI = 0.79, 0.87) for cpRNFL thickness.

For GBC models, adjusted AUPRCs were 0.85 (95% CI = 0.81, 0.87) for the wiVD GBC model, 0.86 (95% CI = 0.82, 0.88) for the wiCD GBC model, 0.87 (95% CI = 0.84, 0.09) for the wiVD and CD GBC model and 0.89 (95% CI = 0.86, 0.92) for the cpRNFL GBC model. The adjusted AUPRC for CNN analysis of RPC *en face* vessel density images was 0.94 (95% CI = 0.92, 0.96) resulting in improved classification compared to all GBC results and all univariable results ($P \leq 0.002$ for all comparisons, Table 3).

Finally, an ablation-like analysis was conducted to investigate the relative strength of effect of different machine learning classifiers and the effect of transfer learning using four additional ML models described below:

1. LeNet-5: A classic shallow neural network that consists of two sets of convolutional and average pooling layers, followed by a flattening convolutional layer, then two fully connected layers and finally a softmax classifier.
2. VGG16 without transfer learning.
3. Resnet-50 without transfer learning. ResNet50 contains 50 layers (16 residual layers and 2 fully connected layers, each residual layer contains 3 convolutional layers).
4. Resnet-50 with transfer learning.

Results from these analyses are shown in Table 4 and indicate that using other ML models including DL models with and without transfer learning still significantly improved classification performance compared to that of all univariate variables and most GBCs.

DISCUSSION

For classifying measurements from healthy and glaucoma eyes with early to moderate disease, the current study compared CNNs trained and tested on *en face* radial peripapillary images to GBCs trained on combinations of whole image and regional OCTA optic nerve head measurements and global and regional RNFL thickness measurements. It also compared them to individual OCTA and OCT parameters. The CNN trained on OCTA images showed significantly better diagnostic accuracy than the four investigated GBCs and instrument-provided whole image vessel density, whole image capillary density and circumpapillary RNFL thickness measurements. Specifically, the *en face* vessel density image-based CNN resulted in an AUPRC of 0.97 compared to AUPRCS for GBCs ranging from 0.89 to 0.93 and AUPRCS from individual OCTA and OCT measurements ranging from 0.85 to 0.86. Similar improvement in diagnostic accuracy remained when only the early glaucoma group was included in the analysis (AUPRC of 0.94 for CNN compared to OCTA and OCT measurements ranging from 0.81 to 0.89. We also have demonstrated that other ML models trained and tested on en face images differentiate between healthy and glaucoma eyes significantly better than most GBCs investigated and all individual OCTA and OCT parameters investigated (Table 4). These results show promise for using deep learning image-based analyses for improved detection of early to moderate glaucoma using OCTA. Moreover, they provide general support for using deep learning image-based analyses of optical imaging measurements in the broader context.

To the best of our knowledge, this is the first study to compare CNN-based image analysis to shallow learning analyses of OCTA measurements. In the current sample, sensitivity at 95% specificity from the investigated CNN was 72.1% while the highest sensitivities at this specificity were 66.7% and 62.5% for GBC models and for individual parameters, respectively. Sensitivities at various specificities were provided because depending on the clinical application, target specificities may be different. For instance, a very high specificity may be important for glaucoma screening where false positives are undesirable and a relaxed specificity may be important when attempting to detect glaucoma in a young glaucoma suspect where a false negative may result in delayed treatment, leading to a preventable decrease in visual function.

Published ML analyses of OCTA measurements are sparse. A recent study by Meng et al[65] used local phase quantization, a blur insensitive texture descriptor[66], to extract features from *en face* OCTA images of healthy individuals and glaucoma patients. Principle component analysis was then used for dimensionality reduction and the most significant features used in the classification task were identified using a decision tree classifier. Classification based on the best identified features using AdaBoost resulted in a best sensitivity/specificity trade-off of 94.4%/94.0% in 157 *en face* optic disc images from healthy individuals and 52 images from glaucoma patients using only 6 features. Reported accuracy was 94.3%. These results are not directly comparable to the current results because *en face* images from several retinal layers (superficial, deep, and outer) and the choriocapillaris were incorporated in their models. In addition, methods for classifying glaucoma and healthy eyes were not described and severity of disease information in patient eyes was not included.

Providing additional evidence that deep learning methods can outperform shallow learning methods for classifying glaucoma and healthy eyes, a recent publication by Maetschke and colleagues[48] investigated CNN analysis of raw OCT optic nerve head volumes and compared results to many feature-based supervised ML methods. Parameters included in feature-based ML classifiers were clock-hour RNFL thicknesses, quadrant thicknesses, average thickness, rim area, disc area, average cup-to-disc ratio, vertical cup-to-disc ratio, and cup volume. Average MD (range) in glaucoma eyes, defined based on two consecutive abnormal VFs, was −6.8 dB (−32.9, −2.17). CNN analysis of OCT volumes resulted in an area under the receiver operating characteristic curve (AUROC) of 0.94 for the classification task while feature-based traditional machine learning analysis AUROCS ranged from 0.82 for support vector machine and gradient boosting classifiers to 0.89 for logistic regression.

The current study has several possible limitations. First, significantly more glaucoma patients than healthy participants had hypertension and diabetes (Table 1), both of which can damage micro-vessels and influence blood flow. In a previous study[55] we performed an exploratory analysis of differences in vessel density measures among healthy subjects with and without hypertension. We did not find a significant difference in any ONH vessel density measures and thus did not pursue an adjusted analysis in the current study. Due to the low number of diabetic healthy subjects, we were unable to adjust for this covariate in any analysis herein. In addition, more glaucoma eyes than healthy eyes were pseudophakic. For this reason, all models were adjusted for this variable. Second, it has been reported that parapapillary OCTA measurements are not ideal for discriminating between healthy and

glaucoma eyes[8, 67, 68] possibly due, in part, to vascular crowding of large vessels in the optic disc decreasing visibility of microvasculature in this region.[69] Third, results of CNN and GBC models were not tested on an external, independent test dataset so results may be somewhat worse in other populations, although five-fold cross-validation with independent test and training sets was used to control for overfitting of models. Fourth, our use of strict OCTA image quality control likely reduced the available number of usable images available from the DIGS cohort. Previous reports indicate that a significant number of OCTA images include artifacts that make them potentially unusable.[70, 71] Such images were excluded based on IDEA Center image assessment protocol. It also is possible that reliance on previously suggested VF reliability criteria (percent fixation loss and false positive cut-offs) decreased the available number of study participants[72], although a retrospective analysis of 63,000 VisFACT processed VFs indicates that less than 1% were excluded based on these criteria.

In conclusion, the current results indicate that convolutional neural networks trained and tested on *en face* OCTA images can improve diagnostic accuracy and sensitivities at fixed specificities in glaucoma eyes compared to gradient boosting classifier models that combine structural measurements from OCTA and OCT measurements independently and single OCTA and OCT parameters. Such techniques could be incorporated into instrument software to improve clinical utility.

ACKNOWLEDGEMENTS/DISCLOSURES

Funding/Support:

National Institutes of Health/National Eye Institute, Bethesda MA: R01 EY029058, R21 EY027945, P30 EY022589, R01 EY027510, K12 EY024225.

An unrestricted grant from Research to Prevent Blindness, New York NY.

Participant retention incentive grants in the form of glaucoma medication at no cost from Novartis/Alcon Laboratories Inc, Allergan, Akorn and Pfizer Inc.

Financial Disclosures:

C. Bowd - None

A. Belghith - None

L.M. Zangwill - Research Support: Carl Zeiss Meditec, Heidelberg Engineering, Optovue, Topcon Medical Systems

M. Christopher - None

M.H. Goldbaum - None

J. Rezapour - None

S. Moghimi - None

A. Kamalipour – None

H. Hou

R.N. Weinreb - Consultant: Aerie Pharmaceuticals, Allergan, Bausch & Lomb, Eyenovia, Implantdata, Novartis; Research Support: Bausch and Lomb, Carl Zeiss Meditec, Centervue, Heidelberg Engineering, Konan Medical, Optovue, Research to Prevent Blindness

REFERENCES

1. Venugopal JP, Rao HL, Weinreb RN, et al. Repeatability of vessel density measurements of optical coherence tomography angiography in normal and glaucoma eyes. Br J Ophthalmol. Mar 2018;102(3):352–357. doi:10.1136/bjophthalmol-2017-310637 [PubMed: 28739645]
2. Manalastas PIC, Zangwill LM, Saunders LJ, et al. Reproducibility of Optical Coherence Tomography Angiography Macular and Optic Nerve Head Vascular Density in Glaucoma and Healthy Eyes. J Glaucoma. Oct 2017;26(10):851–859. doi:10.1097/IJG.0000000000000768 [PubMed: 28858159]
3. Hollo G. Intrasession and Between-Visit Variability of Sector Peripapillary Angioflow Vessel Density Values Measured with the Angiovue Optical Coherence Tomograph in Different Retinal Layers in Ocular Hypertension and Glaucoma. PloS one. 2016;11(8):e0161631. doi:10.1371/journal.pone.0161631 [PubMed: 27537500]
4. Chen HS, Liu CH, Wu WC, Tseng HJ, Lee YS. Optical Coherence Tomography Angiography of the Superficial Microvasculature in the Macular and Peripapillary Areas in Glaucomatous and Healthy Eyes. Invest Ophthalmol Vis Sci. Jul 1 2017;58(9):3637–3645. doi:10.1167/iovs.17-21846 [PubMed: 28728171]
5. Geyman LS, Garg RA, Suwan Y, et al. Peripapillary perfused capillary density in primary open-angle glaucoma across disease stage: an optical coherence tomography angiography study. Brit J Ophthalmol. Sep 2017;101(9):1261–1268. doi:10.1136/bjophthalmol-2016-309642 [PubMed: 28148529]
6. Liu L, Jia Y, Takusagawa HL, et al. Optical Coherence Tomography Angiography of the Peripapillary Retina in Glaucoma. JAMA ophthalmology. Sep 2015;133(9):1045–52. doi:10.1001/jamaophthalmol.2015.2225 [PubMed: 26203793]
7. Rao HL, Pradhan ZS, Weinreb RN, et al. Regional Comparisons of Optical Coherence Tomography Angiography Vessel Density in Primary Open-Angle Glaucoma. American journal of ophthalmology. Nov 2016;171:75–83. doi:10.1016/j.ajo.2016.08.030 [PubMed: 27590118]
8. Rao HL, Pradhan ZS, Weinreb RN, et al. A comparison of the diagnostic ability of vessel density and structural measurements of optical coherence tomography in primary open angle glaucoma. PloS one. 2017;12(3):e0173930. doi:10.1371/journal.pone.0173930 [PubMed: 28288185]
9. Hou H, Moghimi S, Proudfoot JA, et al. Ganglion Cell Complex Thickness and Macular Vessel Density Loss in Primary Open-Angle Glaucoma. Ophthalmology. Aug 2020;127(8):1043–1052. doi:10.1016/j.ophtha.2019.12.030 [PubMed: 32085875]
10. Hou H, Moghimi S, Zangwill LM, et al. Progressive ganglion cell complex thickness and vessel density loss in healthy, pre-perimetric glaucoma and primary open angle glaucoma eyes. Invest Ophthalmol Vis Sci. 2019;59:E-Abstract 3218.
11. Moghimi S, Zangwill LM, Penteado RC, et al. Macular and Optic Nerve Head Vessel Density and Progressive Retinal Nerve Fiber Layer Loss in Glaucoma. Ophthalmology. Nov 2018;125(11):1720–1728. doi:10.1016/j.ophtha.2018.05.006 [PubMed: 29907322]
12. Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. Npj Digit Med. Aug 28 2018;1doi:UNSP 39 10.1038/s41746-018-0040-6
13. Bizios D, Heijl A, Bengtsson B. Trained artificial neural network for glaucoma diagnosis using visual field data: a comparison with conventional algorithms. J Glaucoma. Jan 2007;16(1):20–8. [PubMed: 17224745]
14. Bowd C, Chan K, Zangwill LM, et al. Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser ophthalmoscopy of the optic disc. Invest Ophthalmol Vis Sci. Nov 2002;43(11):3444–54. [PubMed: 12407155]
15. Bowd C, Chan K, Zangwill LM, et al. Comparison of learning neural networks and linear discriminate functions to discriminate between glaucomatous and non-glaucomatous eyes using HRT optic disc topograhy parameters. Invest Ophthalmol Vis Sci. 2001;42:s118.
16. Bowd C, Medeiros FA, Zhang Z, et al. Relevance vector machine and support vector machine classifier analysis of scanning laser polarimetry retinal nerve fiber layer measurements. Invest Ophthalmol Vis Sci. Apr 2005;46(4):1322–9. [PubMed: 15790898]

17. Brigatti L, Hoffman D, Caprioli J. Neural networks to identify glaucoma with structural and functional measurements. Am J Ophthalmol. May 1996;121(5):511–21. [PubMed: 8610794]
18. Burgansky-Eliash Z, Wollstein G, Chu T, et al. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. Invest Ophthalmol Vis Sci. Nov 2005;46(11):4147–52. [PubMed: 16249492]
19. Chan K, Lee TW, Sample PA, Goldbaum MH, Weinreb RN, Sejnowski TJ. Comparison of machine learning and traditional classifiers in glaucoma diagnosis. IEEE Trans Biomed Eng. Sep 2002;49(9):963–74. [PubMed: 12214886]
20. 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. Nov 12 2018;8(1):16685. doi:10.1038/s41598-018-35044-9 [PubMed: 30420630]
21. Christopher M, Belghith A, Weinreb RN, et al. Retinal Nerve Fiber Layer Features Identified by Unsupervised Machine Learning on Optical Coherence Tomography Scans Predict Glaucoma Progression. Invest Ophthalmol Vis Sci. Jun 1 2018;59(7):2748–2756. doi:10.1167/iovs.17-23387 [PubMed: 29860461]
22. Goldbaum MH, Sample PA, Chan K, et al. Comparing machine learning classifiers for diagnosing glaucoma from standard automated perimetry. Invest Ophthalmol Vis Sci. Jan 2002;43(1):162–9. [PubMed: 11773027]
23. Goldbaum MH, Sample PA, White H, et al. Interpretation of automated perimetry for glaucoma by neural network. Invest Ophthalmol Vis Sci. Aug 1994;35(9):3362–73. [PubMed: 8056511]
24. Hothorn T, Lausen B. Bagging tree classifiers for laser scanning images: a data- and simulation-based strategy. Artif Intell Med. Jan 2003;27(1):65–79. [PubMed: 12473392]
25. Li AN, Cheng J, Wong DWK, Liu J. Integrating Holistic and Local Deep Features for Glaucoma Classification. 2016 38th Annual International Conference of the Ieee Engineering in Medicine and Biology Society (Embc). 2016:1328–1331.
26. Lietman T, Eng J, Katz J, Quigley HA. Neural networks for visual field analysis: how do they compare with other algorithms? J Glaucoma. Feb 1999;8(1):77–80. [PubMed: 10084278]
27. Liu S, Graham S, Shulz A, et al. A Deep Learning-Based Algorithm Identifies Glaucomatous Discs Using Monoscopic Fundus Photographs. Opthalmol Glaucoma. 2018;1(1):15–22.
28. Madsen EM, Yolton RL. Demonstration of a neural network expert system for recognition of glaucomatous visual field changes. Mil Med. Aug 1994;159(8):553–7. [PubMed: 7824147]
29. Mardin CY, Hothorn T, Peters A, Junemann AG, Nguyen NX, Lausen B. New glaucoma classification method based on standard Heidelberg Retina Tomograph parameters by bagging classification trees. J Glaucoma. Aug 2003;12(4):340–6. [PubMed: 12897579]
30. Medeiros FA, Jammal AA, Thompson AC. From Machine to Machine: An OCT-Trained Deep Learning Algorithm for Objective Quantification of Glaucomatous Damage in Fundus Photographs. Ophthalmology. Apr 2019;126(4):513–521. doi:10.1016/j.ophtha.2018.12.033 [PubMed: 30578810]
31. Mutlukan E, Keating D. Visual field interpretation with a personal computer based neural network. Eye. 1994;8 (Pt 3):321–3. [PubMed: 7958038]
32. Phan S, Satoh S, Yoda Y, et al. Evaluation of deep convolutional neural networks for glaucoma detection. Jpn J Ophthalmol. May 2019;63(3):276–283. doi:10.1007/s10384-019-00659-6 [PubMed: 30798379]
33. Swindale NV, Stjepanovic G, Chin A, Mikelberg FS. Automated analysis of normal and glaucomatous optic nerve head topography images. Invest Ophthalmol Vis Sci. Jun 2000;41(7):1730–42. [PubMed: 10845593]
34. Wroblewski D, Francis BA, Chopra V, et al. Glaucoma detection and evaluation through pattern recognition in standard automated perimetry data. Graefes Arch Clin Exp Ophthalmol. Nov 2009;247(11):1517–30. doi:10.1007/s00417-009-1121-7 [PubMed: 19579030]
35. Zangwill LM, Chan K, Bowd C, et al. Heidelberg retina tomograph measurements of the optic disc and parapapillary retina for detecting glaucoma analyzed by machine learning classifiers. Invest Ophthalmol Vis Sci. Sep 2004;45(9):3144–51. doi:10.1167/iovs.04-0202 [PubMed: 15326133]

36. Alam M, Le D, Lim JI, Chan RVP, Yao XC. Supervised Machine Learning Based Multi-Task Artificial Intelligence Classification of Retinopathies. J Clin Med. Jun 2019;8(6)doi:ARTN 872 10.3390/jcm8060872
37. Mirshahi R, Anvari P, Riazi-Esfahani H, Sardarinia M, Naseripour M, Falavarjani KG. Foveal avascular zone segmentation in optical coherence tomography angiography images using a deep learning approach. Sci Rep. Jan 13 2021;11(1):1031. doi:10.1038/s41598-020-80058-x [PubMed: 33441825]
38. Nagasato D, Tabuchi H, Masumoto H, et al. Automated detection of a nonperfusion area caused by retinal vein occlusion in optical coherence tomography angiography images using deep learning. PloS one. Nov 7 2019;14(11)doi:ARTN e0223965 10.1371/journal.pone.0223965
39. Sandhu HS, Elmogy M, Sharafeldeen AT, et al. Automated Diagnosis of Diabetic Retinopathy Using Clinical Biomarkers, Optical Coherence Tomography, and Optical Coherence Tomography Angiography. American journal of ophthalmology. Aug 2020;216:201–206. doi:10.1016/j.ajo.2020.01.016 [PubMed: 31982407]
40. Heisler M, Karst S, Lo JL, et al. Ensemble Deep Learning for Diabetic Retinopathy Detection Using Optical Coherence Tomography Angiography. Translational Vision Science & Technology. Jan 2020;9(2)doi:ARTN 20 10.1167/tvst.9.2.20
41. Ahn JM, Kim S, Ahn KS, Cho SH, Lee KB, Kim US. A deep learning model for the detection of both advanced and early glaucoma using fundus photography. PloS one. Nov 27 2018;13(11)doi:ARTN e0207982 10.1371/journal.pone.0207982
42. 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. American journal of ophthalmology. Feb 2019;198:136–145. doi:10.1016/j.ajo.2018.10.007 [PubMed: 30316669]
43. Asaoka R, Tanito M, Shibata N, et al. Validation of a Deep Learning Model to Screen for Glaucoma Using Images from Different Fundus Cameras and Data Augmentation. Ophthalmol Glaucoma. Jul - Aug 2019;2(4):224–231. doi:10.1016/j.ogla.2019.03.008 [PubMed: 32672542]
44. Chai YD, Liu HY, Xu J. Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. Knowl-Based Syst. Dec 1 2018;161:147–156. doi:10.1016/j.knosys.2018.07.043
45. Christopher M, Bowd C, Belghith A, et al. Deep Learning Approaches Predict Glaucomatous Visual Field Damage from OCT Optic Nerve Head En Face Images and Retinal Nerve Fiber Layer Thickness Maps. Ophthalmology. Mar 2020;127(3):346–356. doi:10.1016/j.ophtha.2019.09.036 [PubMed: 31718841]
46. Christopher M, Nakahara K, Bowd C, et al. Effects of Study Population, Labeling and Training on Glaucoma Detection Using Deep Learning Algorithms. Transl Vis Sci Technol. Apr 2020;9(2):27. doi:10.1167/tvst.9.2.27
47. Jammal AA, Thompson AC, Mariottoni EB, et al. Human Versus Machine: Comparing a Deep Learning Algorithm to Human Gradings for Detecting Glaucoma on Fundus Photographs. American journal of ophthalmology. Mar 2020;211:123–131. doi:10.1016/j.ajo.2019.11.006 [PubMed: 31730838]
48. Maetschke S, Antony B, Ishikawa H, Wollstein G, Schuman J, Garnavi R. A feature agnostic approach for glaucoma detection in OCT volumes. PloS one. 2019;14(7):e0219126. doi:10.1371/journal.pone.0219126 [PubMed: 31260494]
49. Shibata N, Tanito M, Mitsuhashi K, et al. Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. Sci Rep-Uk. Oct 2 2018;8doi:ARTN 14665 10.1038/s41598-018-33013-w
50. Asano S, Asaoka R, Murata H, et al. Predicting the central 10 degrees visual field in glaucoma by applying a deep learning algorithm to optical coherence tomography images. Sci Rep. 202i;11:2214. [PubMed: 33500462]
51. Medeiros FA, Jammal AA, Thompson AC. From Machine to Machine An OCT-Trained Deep Learning Algorithm for Objective Quantification of Glaucomatous Damage in Fundus Photographs. Ophthalmology. Apr 2019;126(4):513–521. doi:10.1016/j.ophtha.2018.12.033 [PubMed: 30578810]

52. Phene S, Dunn RC, Hammel N, et al. Deep Learning and Glaucoma Specialists The Relative Importance of Optic Disc Features to Predict Glaucoma Referral in Fundus Photographs. Ophthalmology. Dec 2019;126(12):1627–1639. doi:10.1016/j.ophtha.2019.07.024 [PubMed: 31561879]
53. Thompson AC, Jammal AA, Medeiros FA. A Deep Learning Algorithm to Quantify Neuroretinal Rim Loss from Optic Disc Photographs. Invest Ophth Vis Sci. Jul 2019;60(9)
54. Urata C, Mariottoni EB, Jammal AA, Thompson AC, Medeiros FA. A Deep Learning Algorithm can Predict Retinal Nerve Fiber Layer Thickness with Lower Long-term Variability than Conventional Automated Segmentation. Invest Ophth Vis Sci. Jul 2019;60(9)
55. Bowd C, Belghith A, Proudfoot JA, et al. Gradient-boosting classifiers combining vessel density and tissue thickness measurements for classifying early to moderate glaucoma. Am J Ophthalmol. Sep 2020;217:131–139. doi:10.1016/j.ajo.2020.03.024 [PubMed: 32222368]
56. Shen D, Wu G, Suk HI. Deep Learning in Medical Image Analysis. Annu Rev Biomed Eng. Jun 21 2017;19:221–248. doi:10.1146/annurev-bioeng-071516-044442 [PubMed: 28301734]
57. Sample PA, Girkin CA, Zangwill LM, et al. The African Descent and Glaucoma Evaluation Study (ADAGES): design and baseline data. Arch Ophthalmol. Sep 2009;127(9):1136–45. doi:127/9/1136 [pii] 10.1001/archophthalmol.2009.187 [PubMed: 19752422]
58. Jia Y, Tan O, Tokayer J, et al. Split-spectrum amplitude-decorrelation angiography with optical coherence tomography. Opt Express. Feb 13 2012;20(4):4710–25. doi:10.1364/OE.20.004710 [PubMed: 22418228]
59. Simonyan K, Zisserman A. Very deep convolutional networks for large scale image recognition. 2014;https://arxiv.org/abs/1409.1556.
60. Schapire RE, Freund Y, Bartlett P, Lee WS. Boosting the margin: A new explanation for the effectiveness of voting methods. Ann Stat. Oct 1998;26(5):1651–1686.
61. Caruana R, Niculescu-Mizil A. An empirical comparison of supervised learning algorithms. 2006:161–168.
62. Saito T, Rehmsmeier M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLoS ONE 10(3): e0118432 10.1371/journalpone0118432.2015 10(3) e0118432 10.1371/journal.pone.0118432.
63. Janes H, Pepe MS. Adjusting for covariate effects on classification accuracy using the covariate-adjusted receiver operating characteristic curve. Biometrika. Jun 2009;96(2):371–382. doi:10.1093/biomet/asp002 [PubMed: 22822245]
64. Johnson WD, Romer JE. Hypothesis Testing of Population Percentiles via the Wald Test with Bootstrap Variance Estimates. Open J Stat. Feb 2016;6(1):14–24. doi:10.4236/ojs.2016.61003 [PubMed: 27034909]
65. Meng CY, Ng EYK, Jahmunah V, et al. Automated detection of glaucoma using optical coherence tomography angiogram images. Comput Biol Med. Dec 2019;115doi:ARTN 103483 10.1016/j.compbiomed.2019.103483
66. Castanedo F. A review of data fusion techniques. TheScientificWorldJournal. 2013;2013:704504. doi:10.1155/2013/704504
67. Chihara E, Dimitrova G, Amano H, Chihara T. Discriminatory Power of Superficial Vessel Density and Prelaminar Vascular Flow Index in Eyes With Glaucoma and Ocular Hypertension and Normal Eyes. Invest Ophthalmol Vis Sci. Jan 1 2017;58(1):690–697. doi:10.1167/iovs.16-20709 [PubMed: 28134965]
68. Wang XL, Jiang CH, Ko T, et al. Correlation between optic disc perfusion and glaucomatous severity in patients with open-angle glaucoma: an optical coherence tomography angiography study. Graef Arch Clin Exp. Sep 2015;253(9):1557–1564. doi:10.1007/s00417-015-3095-y
69. Van Melkebeke L, Barbosa-Breda J, Huygens M, Stalmans I. Optical Coherence Tomography Angiography in Glaucoma: A Review. Ophthalmic Res. 2018;60(3):139–151. doi:10.1159/000488495 [PubMed: 29794471]
70. Kamalipour A, Moghimi S, Hou H, et al. OCT Angiography Artifacts in Glaucoma. Ophthalmology. Apr 2 2021;doi:10.1016/j.ophtha.2021.03.036

71. Holmen IC, Konda SM, Pak JW, et al. Prevalence and Severity of Artifacts in Optical Coherence Tomographic Angiograms. JAMA ophthalmology. Feb 1 2020;138(2):119–126. doi:10.1001/jamaophthalmol.2019.4971 [PubMed: 31804666]

72. Asaoka R, Fujino Y, Aoki S, Matsuura M, Murata H. Estimating the Reliability of Glaucomatous Visual Field for the Accurate Assessment of Progression Using the Gaze-Tracking and Reliability Indices. Ophthalmol Glaucoma. Mar - Apr 2019;2(2):111–119. doi:10.1016/j.ogla.2019.02.001 [PubMed: 32672604]

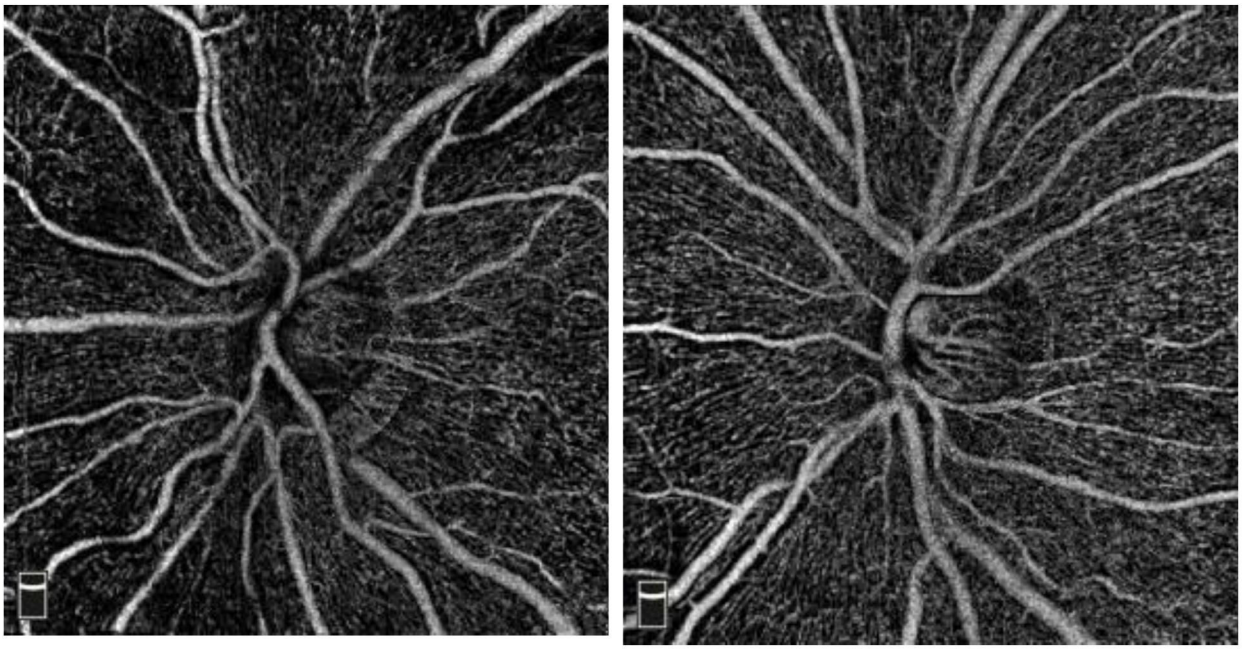

Figure 1:
Example enface vessel density maps of a glaucoma eye (left) and a healthy eye (right) correctly classified by CNN analysis (probability of glaucoma = 0.72 and probability of glaucoma = 0.19, respectively) but incorrectly classified by whole image vessel density (wiVD) GBC (probability of glaucoma = 0.33 and probability of glaucoma = 0.72, respectively), whole image capillary density (wiCD) GBC (probability of glaucoma = 0.24 and, probability of glaucoma = 0.85, respectively), wiVD and CD GBC (probability of glaucoma = 0.19 and probability of glaucoma = 0.74, respectively), and cpRNFL GBC (probability of glaucoma = 0.43 and probability of glaucoma = 0.66, respectively). A cut-off of > 0.50 was used to classify eyes as glaucomatous.

Table 1:

Patient and eye characteristics by diagnosis. Mean values and 95% confidence intervals are shown for continuous variables. Statistical significance of differences in continuous and categorical variables are determined by two-sample t-tests and Fisher's exact test for patient level variables (respectively) and linear mixed effects models for eye level variables.

	Diagnosis		
	Healthy (n = 80, 130 Eyes)	**Glaucoma (n = 185, 275 Eyes)**	**p-value**
Age (years)	62.1 (57.6, 66.3)	71.8 (70.2, 74.5)	<0.001
Sex (% Female)	72.0%	51.5%	<0.001
Race (%)			
Non-White	41.0%	38.0%	0.15
White	59.0%	62.0%	
Hypertension (%)	41.5%	53.1%	0.031
Pseudophakia	10.1%	43.0%	<0.001
Diabetes (%)	5.0%	15.2%	0.022
MD (dB)	0.02 (−1.22, 1.34)	−5.40 (−6.32, −4.77)	<0.001
IOP (mmHg)	15.1 (13.7, 16.0)	14.1 (13.8, 15.1)	0.52
AL (mm)	23.8 (23.4, 24.4)	24.2 (24.1, 24.3)	0.049
CCT (μm)	547.4 (536.2, 558.7)	536.6 (529.1, 543.7)	0.08
BMO Area (mm^2)	1.9 (1.8, 2.1)	1.9 (1.8, 2.0)	0.54

MD = Mean Deviation, IOP = Intraocular Pressure, AL = Axial Length, CCT = Central Corneal Thickness, BMO = Bruch's Membrane Opening

Table 2:

Estimated areas under the receiver operating curves (AUROC) and sensitivities at fixed specificities for all investigated univariable OCTA and OCT parameters, gradient boosting classifier (GBC) models and convolutional neural network (CNN) analyses. (130 healthy eyes from 80 subjects and 275 glaucoma eyes from 185 patients).

		Sensitivity At				Adjusted AUROC p-value compared to Deep Learning
	AUROC (95% CI)	80% Spec.	85% Spec.	90% Spec.	95% Spec.	
Univariate variable						
Whole image vessel density (wiVD)	0.81 (0.75, 0.82)	81.4%	77.0%	72.1%	60.1%	<0.001
Whole image capillary density (wiCD)	0.82 (0.77, 0.84)	82.1%	80.2%	73.8%	61.7%	<0.001
Circumpapillary RNFL thickness (cpRNFL)	0.85 (0.81, 0.87)	84.3%	81.4%	75.9%	62.5%	<0.001
Gradient Boosted Classifiers						
WiVD GBC	0.83 (0.78, 0.84)	82.5%	80.1%	74.3%	62.1%	<0.001
WiCD GBC	0.84 (0.79, 0.86)	83.9%	82.2%	75.4%	64.6%	<0.001
WiVD VD and CD GBC	0.86 (0.82 0.88)	86.3%	81.3%	75.9%	65.2%	<0.001
cpRNFL GBC	0.88 (0.84, 0.89)	87.4%	83.2%	77.1%	66.7%	0.004
CNN						
En face RPC density image	0.92 (0.90, 0.95)	89.9%	86.6%	79.1%	72.1%	

RPC = radial peripapillary capillary

Table 3:

Estimated areas under the receiver operating curves (AUROC) and sensitivities at fixed specificities for all investigated univariable OCTA and OCT parameters, gradient boosting classifier (GBC) models and convolutional neural network (CNN) analyses for early glaucoma (130 healthy eyes from 80 subjects and 183 early glaucoma eyes, with mean deviation ≥ -6 dB, from 135 patients).

		Sensitivity At				Adjusted AUROC p-value compared to Deep Learning
	AUROC (95% CI)	80% Spec.	85% Spec.	90% Spec.	95% Spec.	
Univariate variable						
Whole image vessel density (wiVD)	0.73 (0.70, 0.76)	74.5%	73.1%	68.9%	50.2%	<0.001
Whole image capillary density (wiCD)	0.75 (0.71, 0.80)	76.1%	74.4%	69.6%	53.6%	<0.001
Circumpapillary RNFL thickness (cpRNFL)	0.77 (0.73, 0.81)	78.2%	75.4%	71.6%	58.5%	<0.001
Gradient Boosted Classifiers						
WiVD GBC	0.80 (0.74, 0.81)	78.5%	75.1%	70.7%	60.8%	<0.001
WiCD GBC	0.82 (0.77, 0.83)	76.9%	77.5%	72.2%	61.9%	<0.001
WiVD and CD GBC	0.83 (0.78, 0.85)	81.3%	79.7%	73.6%	62.5%	<0.001
cpRNFL GBC	0.85 (0.83, 0.87)	81.9%	80.4%	74.1%	63.1%	0.002
CNN						
En face RPC density image	0.90 (0.88, 0.94)	87.1%	85.3%	77.2%	69.3%	

RPC = radial peripapillary capillary

Table 4:

Estimated areas under the precision recall curves (AUPRC) and classification comparisons for all investigated univariable OCTA and OCT parameters, gradient boosting classifier (GBC) models and machine learning models (130 healthy eyes from 80 participants and 275 glaucoma eyes from 185 patients).

	AUPRC (95% CI)	Adjusted AUPRC p-value compared to Lenet-5	Adjusted AUPRC p-value compared to VGG16 without transfer learning	Adjusted AUPRC p-value compared to Resnet-50 without transfer learning	Adjusted AUPRC p-value compared to VGG16 with transfer learning	Adjusted AUPRC p-value compared to Resnet-50 with transfer learning
Univariate variable						
Whole image vessel density (wiVD)	0.85 (0.76, 0.88)	<0.001	<0.001	<0.001	<0.001	<0.001
Whole image capillary density (wiCD)	0.86 (0.79, 0.88)	<0.001	<0.001	<0.001	<0.001	<0.001
Circumpapillary RNFL thickness (cpRNFL)	0.86 (0.81, 0.87)	<0.001	<0.001	<0.001	<0.001	<0.001
Gradient Boosted Classifiers						
WiVD GBC	0.89 (0.82, 0.92)	0.08	<0.001	<0.001	<0.001	<0.001
WiCD GBC	0.89 (0.83, 0.92)	0.07	<0.001	<0.001	<0.001	<0.001
WiVD VD and CD GBC	0.91 (0.88 0.93)	0.12	0.03	0.04	<0.001	<0.001
cpRNFL GBC	0.93 (0.91, 0.95)	0.35	0.1	0.12	0.01	0.02
Lenet-5						
En face RPC density image	0.92 (0.90, 0.95)		0.06	0.07	0.02	0.03
VGG16 without transfer learning						
En face RPC density image	0.95 (0.94, 0.97)	0.06		0.42	0.18	0.21
Resnet-50 without transfer learning						
En face RPC density image	0.95 (0.94, 0.97)	0.07	0.42		0.36	0.38
VGG16 with transfer learning						
En face RPC density image	0.97 (0.95, 0.99)	0.02	0.18	0.36		0.64
Resnet-50 with transfer learning						
En face RPC density image	0.97 (0.95, 0.99)	0.03	0.21	0.38	0.64	