

Clinical Applications of Artificial Intelligence in Glaucoma

Siamak Yousefi, PhD

Department of Ophthalmology, University of Tennessee Health Science Center, Memphis, TN, USA
 Department of Genetics, Genomics, and Informatics, University of Tennessee Health Science Center, Memphis, TN, USA

ORCID:

Siamak Yousefi: <https://orcid.org/0000-0001-8633-5730>

Abstract

Ophthalmology is one of the major imaging-intensive fields of medicine and thus has potential for extensive applications of artificial intelligence (AI) to advance diagnosis, drug efficacy, and other treatment-related aspects of ocular disease. AI has made impressive progress in ophthalmology within the past few years and two autonomous AI-enabled systems have received US regulatory approvals for autonomously screening for mid-level or advanced diabetic retinopathy and macular edema. While no autonomous AI-enabled system for glaucoma screening has yet received US regulatory approval, numerous assistive AI-enabled software tools are already employed in commercialized instruments for quantifying retinal images and visual fields to augment glaucoma research and clinical practice. In this literature review (non-systematic), we provide an overview of AI applications in glaucoma, and highlight some limitations and considerations for AI integration and adoption into clinical practice.

Keywords: Artificial Intelligence; Convolutional Neural Network (CNN); Deep Learning; Glaucoma; Machine Learning; Ophthalmology

J Ophthalmic Vis Res 2023; 18 (1): 97–112

INTRODUCTION

Artificial Intelligence (AI) applications in ophthalmology have shown significant advancements due mainly to the availability of computational platforms, generation of large annotated ocular images, and emergence of AI algorithms. Several landmark studies have highlighted the effectiveness of AI applications in screening, referral, and diagnosis of different ocular conditions.^[1–3]

AI is a broad term encompassing a wide range of subfields including image processing and expert systems, in which models are preprogrammed and thus require domain knowledge (i.e., human expertise to guide the programmer). In contrast, another subfield of AI, machine learning, can learn from data and identify the outcome of new circumstances without being explicitly programmed. Machine learning models can be further sub-divided to include *supervised learning* in which the label of data is available

Correspondence to:

Siamak Yousefi, PhD. 930 Madison Ave., Suite 726, Memphis, TN 38163, USA.
 E-mail: siamak.yousefi@uthsc.edu

Received: 16-10-2022 Accepted: 05-11-2022

Access this article online

Website: <https://knepublishing.com/index.php/JOVR>

DOI: 10.18502/jovr.v18i1.12730

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

How to cite this article: Yousefi S. Clinical Applications of Artificial Intelligence in Glaucoma. *J Ophthalmic Vis Res* 2023;18:97–112.

and *unsupervised learning* in which the labels of the data are unknown. *Deep convolutional neural networks* (CNNs) are supervised machine learning models that utilize a stack of hidden layers composed of artificial neurons to emulate human brain in the learning and recognizing process.

Figure 1 shows a broad timeline of major retinal imaging instruments and AI applications in glaucoma. The ophthalmoscope was invented by Helmholtz in the 1850s, which revolutionized ophthalmology, as it allowed direct visualization of the retina and optic disc. Introduction of fundus photography in the 1910s allowed documentation of the status of the retina, thus enhancing the monitoring and management of glaucoma patients. AI was born in the 1940s and some researchers attempted to apply some AI techniques, including classical image processing, to locate the optic disc in retinal frames generated from a television ophthalmoscope in 1950s. However, it was not until the 1980s that some expert systems were applied to retinal images to quantify optic disc properties useful in detecting glaucoma. Subsequently, AI models have been broadly applied to different aspects of glaucoma including retinal and optic nerve image and visual field (VF) quantification, screening, referral, diagnosis, forecasting (prediction), prognosis, and monitoring.

Early machine learning models in glaucoma were based on neural networks and attempted to diagnose glaucoma from VFs in the 1990s.^[4, 5] Thereafter, various machine learning models were applied to diagnose glaucoma based on fundus photographs, optical coherence tomography (OCT), OCT angiography, and other ocular and demographic parameters, followed by various deep learning models in the 2010s.^{[6]–[31]}

METHODS

In this review, we used search combinations of “artificial intelligence”, “machine learning”, “neural networks”, “deep learning”, “glaucoma screening”, “glaucoma diagnosis”, “glaucoma progression”, “segmentation”, and “image annotation” in Google and PubMed to review broad applications of AI in glaucoma. We categorized AI applications in glaucoma into four major groups: (1) applications in retinal imaging and VF quantification; (2) applications in screening, referral, diagnosis, and forecasting (prediction); (3) applications in

monitoring and progression detection; and (4) applications in estimating functional parameters from structural factors. We then highlighted some of the limitations and challenges of integrating these AI models into clinical care.

AI in Glaucoma Image and Data Quantification, and Characterization

Retinal imaging and VF testing in conjunction with clinical examinations form the primary basis for assessment and diagnosis of glaucoma.^[32] While *color fundus photography* has long been used to document retinal status, recent *OCT imaging* provided three-dimensional views of retinal layers and optic nerve head structures.^[33] In addition to these modalities, functional assessment typically performed via *standard automated perimetry (SAP)*^[34] has remained a standard practice for diagnosis and prognosis of visual function in patients with glaucoma.^[34, 35] These three imaging modalities comprise the major components of glaucoma assessment. As such, improvements in the quantification and characterization of retinal images and VFs could promote objectivity, improve consistency in glaucoma assessment, and set a common ground for research and clinical practice. For instance, interpreting vertical cup-to-disc ratio (CDR) may facilitate glaucoma diagnosis^[36] as CDR is a major risk factor; likewise, monitoring RNFL thickness may facilitate prognosis as thinning of RNFL is a hallmark of glaucoma progression.^[37] Quantification of VFs in the form of glaucoma-induced patterns of VF loss could also facilitate diagnosis and assist therapy adjustment and prognosis plan optimization based on the shape, type, and depth of defect with consideration of the patient’s quality of the life.^[38]

AI models have been proposed to quantify retinal images as early as the 1950s [Figure 1 & Figure 2: top row]. In conventional AI models (image processing and expert systems), the role of human expertise in hand-crafting algorithms to quantify glaucoma-induced changes and lesions from retinal images was critical. For instance, optic disc and cup boundaries were automatically detected based on various classical image processing techniques that typically require human expertise in the process. Retinal fundus image processing usually requires pre-processing steps to prepare and enhance the image for feature

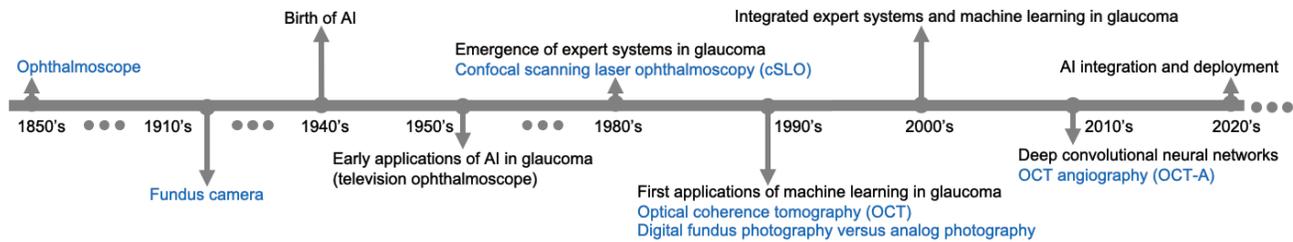


Figure 1. Timeline of major retinal imaging instruments and landmark artificial intelligence applications in glaucoma. Introduction of landmark imaging instruments are listed in blue and AI events are provided in black. AI, artificial intelligence; OCT, optical coherence tomography; OCTA, OCT angiography.

extraction (identifying landmarks). Many of the image processing techniques include histograms equalization and morphological (shape) filtering, and active contours. More involved processes such as gradient vector flow was used to delineate optic disc and cup boundaries in early AI models.^[6] As CDR is a major glaucoma risk factor, many of the follow-up AI models focused on localization and quantification of the optic disc and cup in fundus photographs to identify CDR. Hoover et al localized and quantified optic disc and cup based on the information derived from blood vessels.^[39] Chrastek et al suggested an automated model for ONH segmentation and quantification based on morphological operations, Hough transform, and active contours.^[11] Wong et al used several classical image processing steps to segment the optic cup and disc from retinal images then used a fusion network to combine quantified parameters and subsequently employed an SVM classifier to discriminate normal eyes from glaucomatous eyes.^[40] Follow-up studies on fundus photographs also have applied broad classical image processing techniques including edge detection, morphological filtering, adaptive deformable filters, and active contours to quantify optic disc characteristics to assist glaucoma diagnosis [Figure 2: third row].⁴¹⁻⁴⁵

Emerging deep CNN models however changed the paradigm from manual feature engineering to automatic end-to-end quantification of color fundus images (Fig. 2: fourth row). One of the first applications of deep learning in quantifying optic disc and cup from fundus photographs was introduced in 2015.^[12] They developed a deep learning model using two publicly available datasets of fundus images and segmented optic disc and cup and computed the degree of vessel kinking integrated with prior knowledge about

retinal structures to quantify fundus images. Other models have obtained AUCs up to 0.92 in detecting glaucoma from the quantified retina and ONH characteristics based on fundus images.^[46]

Fundus photographs were traditionally used to document retinal structure. However, with the introduction of OCT^[47] in the 1990s, this modality soon became popular and is now an indispensable component of glaucoma assessment.^[33] OCT quantification is thus highly rewarding yet challenging because OCT provides a significantly lower resolution compared to color fundus photographs, and lesions and characteristics are not typically as obvious as those in fundus photographs. Moreover, the shadows generated due to blood vessels pose additional quantification challenges.^[33, 48] Nevertheless, OCT provides substantial retinal structural information in three dimensions and its quantification can be highly useful. For these reasons, OCT quantification and interpretation has always been an active area of research since its invention. Like color fundus photographs, the conventional AI methods to quantify OCT images typically include classical image preprocessing techniques such as linear or non-linear filtering, edge detection, and local texture analysis. Koozekanani et al developed an algorithm for OCT retinal layer segmentation based on classical edge detection and Markov modeling. Based on 1450 OCT B-scans, the derived retinal thickness measurements deviated from the ground truth thicknessed by less than 10 microns for ~74% of the B-scans and by less than 25 micron for ~99% of the B-scans.^[49] Ishikawa et al proposed an algorithm based on adaptive thresholding technique to segment macular OCT images and subsequently used quantified parameters to diagnose glaucoma. Based on a dataset with about 60 OCT images, they obtained

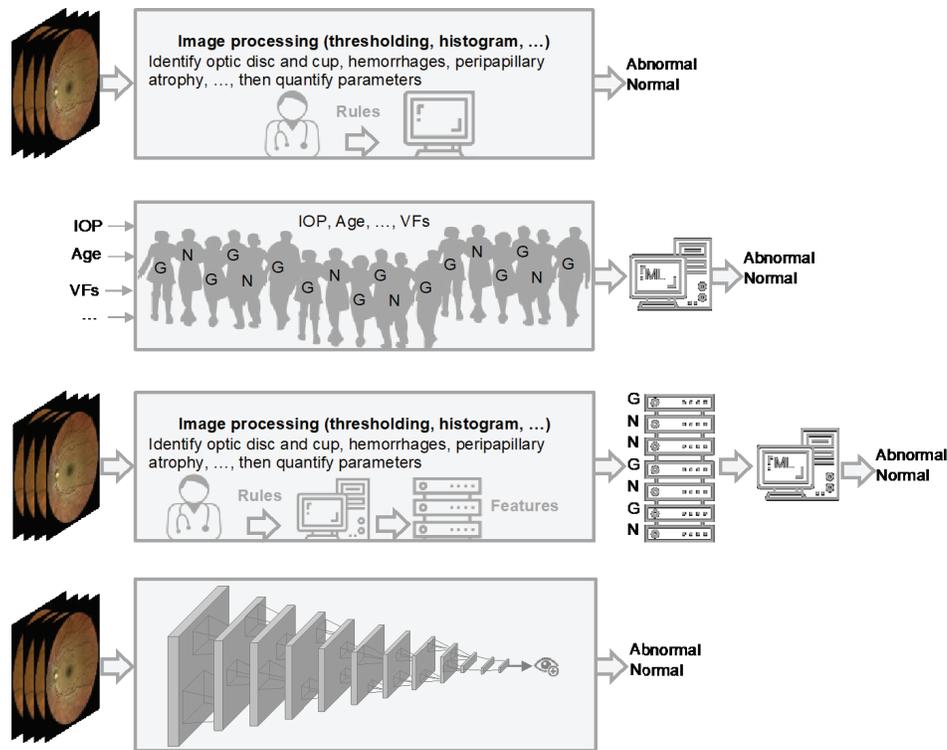


Figure 2. Evolution of AI in glaucoma. **First row:** Image processing and expert systems were used to identify glaucoma landmarks or features (such as cup-to-disc ratio or hemorrhages) from retinal images with the assistance of a glaucoma specialist and glaucoma landmarks are identified. **Second row:** Numerical parameters like raw visual fields (VFs), intraocular pressure (IOP), and age from normal and glaucomatous subjects (presented as *N* and *G*) are input to a conventional machine learning model (e.g., neural network) without glaucoma specialist assistance and diagnosis is made. **Third row:** Image processing and expert systems were used to quantify glaucoma landmarks (extract features) with the assistance of a glaucoma specialist then quantified parameters (features) from normal and glaucomatous subjects are fed to a conventional machine learning model to make diagnosis. **Fourth row:** Retinal image is fed to an end-to-end deep learning model and the diagnosis is made without assistance from a glaucoma specialist.

AUCs up to 0.97 for discriminating normal eyes from eyes with established glaucoma.^[50] Follow-up models further improved the segmentation accuracy. For example, Kafieh et al developed an OCT segmentation model based on local image textures and diffusion mapping to quantify retinal layers. They evaluated their model using 23 OCT images collected from normal and glaucomatous eyes and obtained retinal layer quantifications with lower than ~8 microns of thickness error.^[51]

Emerging deep CNN models however have transformed OCT image quantification from manual feature extraction and annotation to automatic end-to-end quantification. Recent deep learning models provide detailed quantifications of OCT layers as well as information regarding existing pathologies and underlying ocular condition.^{[13]–[17, 52, 53]}

In terms of VFs, various methods have been proposed to summarize, quantify, and annotate

VFs. Garway-Heath et al^[54] developed a model to map VF test locations on optic nerve structure to better quantify the relationship between localized VF and retinal nerve fiber layer (RNFL) loss. Some groups focused on identifying local patterns of glaucomatous VF loss then classifying and quantifying the severity levels based on subjective assessments.^[55, 56] However, manual identification and classification of VF patterns is labor-intensive and requires high levels of expertise that may be prone to inter-and intra-reader variability.^[57, 58] Subsequently, numerous automated models based on conventional machine learning approaches were proposed to identify and classify patterns of VF defect using unsupervised Gaussian mixture modeling (GMM), archetypal analysis, or deep archetypal analysis.^{[22]–[31]} Most of the AI models for quantifying and annotating VFs are based on conventional unsupervised learning. Figure 3 shows how classical archetypal analysis applied

can be used to decompose VFs into 18 prevalent patterns of VF loss [Figure 3: top panel] and decomposition of OCT circle scans to 16 prevalent patterns of RNFL loss [Figure 3: bottom panel]. Such a model can decompose VF or OCT data to a weighted combination of these prevalent patterns and even be used for subsequent detection of glaucoma progression.^{[22]–[31]}

In terms of clinical applications, most of the commercially available OCT imaging instruments provide some level of OCT image quantification, interpretation, and visualization. The widely used Humphrey VF analyzer provides several summary parameters including mean deviation (MD), pattern standard deviation (PSD), and visual field index (VFI), and regional parameters such as glaucoma hemifield test (GHT). However, this is not the case for most (if not all) fundus cameras. But, as OCT is predominantly used in glaucoma clinical care, color fundus photograph quantification may not be a major limitation in clinical applications. Nevertheless, fundus cameras are usually cheaper and more portable than current commercialized OCT instruments, and thus may be more appropriate for community-based glaucoma screening. Thus, while innovative AI models may be applied to OCT and VF data to provide more objective and consistent parameters in glaucoma clinical practice, AI may augment color fundus photography in quantifying more specific and sensitive parameters to enhance community-based and glaucoma screening programs.

Clinical considerations

The effectiveness of using optic nerve characteristics in detecting glaucoma has been investigated extensively. Damms et al observed that vertical CDR best suits glaucoma screening, whereas the rim area is more appropriate for detecting progression.^[59] A follow-up study, however, found that localized rim area led to the highest specificity of 90% and sensitivity of 91% for discriminating glaucoma from normal eyes based on computerized raster tomography.^[60] A recent study suggested vertical CDR as the most important feature for diagnosing glaucoma based on color fundus photographs.^[36] Therefore, more accurate quantification of parameters such as CDR and RNFL thickness profiles may augment clinical care and improve more objective glaucoma assessment and diagnosis. Moreover, tracking

the quantified parameters over time may facilitate a more objective and accurate monitoring and progression detection.

Applications of AI in Glaucoma Screening, Referral, Diagnosis, and Forecasting

While the first applications of image processing in glaucoma dates back to the 1950s, the first applications of machine learning models in glaucoma dates to the 1990s when several teams applied neural networks to VFs for glaucoma diagnose [Figure 2: second row].^[4, 5] Numerous follow-up neural network-based models were proposed to diagnose glaucoma based on VFs.^{[61]–[64]} As VFs were composed of numerical values of threshold sensitivity or total deviations, they provided the optimal input to neural networks which may explain the extensive utility of early neural network models for glaucoma diagnosis based on VFs. A study conducted by Chan and colleagues compared several machine learning models including multilayer perceptron (MLP), support vector machine (SVM), linear and quadratic discriminant analysis, mixture of Gaussian (MOG), and mixture of generalized Gaussian (MGG) in diagnosing glaucoma based on VFs and found that machine-learning-type classifiers provided higher accuracy compared to best VF indexes from the STATPAC software in diagnosing glaucoma.^[7] Other teams utilized various machine learning classifiers such as SVM, discriminant analysis, bagging, and ensemble learning to identify glaucoma based on VFs.^{[8]–[10]}

Raw VFs provide a small grid of numbers (typically fewer than 9*9), thus, in contrast to color fundus and OCT images, VFs are basically inappropriate for deep CNN analysis. As such, some researchers have applied deep CNN models on VF printouts (reports) rather than raw VF numbers. Li et al developed a deep CNN model based on over 4000 VF printouts and obtained an AUC up to about 0.87 in differentiating normal from glaucomatous VF while an SVM model achieved an AUC of 0.67 and glaucoma experts achieved an AUC up to 0.62.^[65] A recent deep CNN models utilized over 16,000 VFs and obtained AUCs up to 0.93 for diagnosing glaucoma.^[66]

Glaucoma is characterized by progressive structural loss of retinal ganglion cells (RGCs), therefore structural evaluation is a critical step

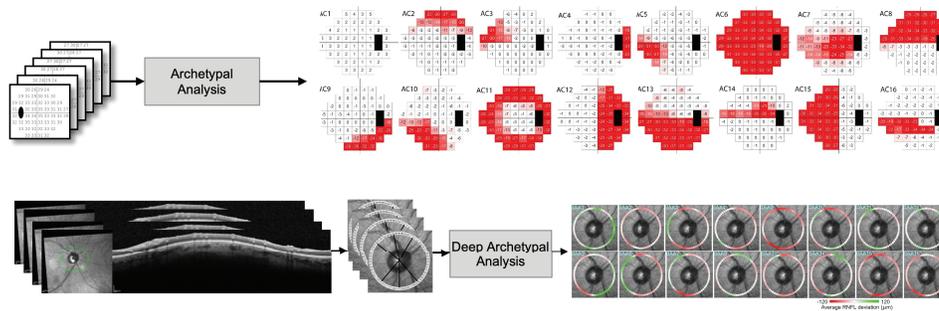


Figure 3. Visual field (VF) and optical coherence tomography (OCT) image quantification. Top: VFs were quantified to 18 prominent patterns of VF loss based on classical archetypal analysis. Bottom: OCT circle scans were quantified to 16 patterns of RNFL loss based on deep archetypal analysis.

in glaucoma assessment. Some early studies showed that RVM and SVM classifiers can discriminate glaucoma from normal eyes using RNFL thickness measurements derived from scanning laser polarimetry (SLP) instruments with AUROCs of up to 0.94.^[67] Others generated ONH parameters (after a manual outlining of the optic disk border) such as cup's volume, depth, and shape as well as rim's characteristics from CSLO instruments to diagnose glaucoma.^[68] While AI models were applied to SLP or cSLO-derived ONH and RNFL parameters (generated by the instruments), diagnostic AI models based on color fundus images were more involved as instruments typically did not provide quantified parameters. As a result, the AI models were required to first quantify characteristics (extract features) then learn those quantified features for glaucoma diagnosis. Bock et al developed image processing models (conventional AI) for glaucoma screening that first preprocessed color fundus photographs and generated different generic features and then performed dimension reduction to lower the number of features. These features were then combined, and a glaucoma risk index was generated, which achieved an AUC of 0.88 in screening glaucoma.^[69] Other AI models used a combination of texture and higher order spectra features from color fundus photographs, then employed numerous machine learning models including SVM, naive Bayesian, and random-forest and obtained an accuracy up to 91% based on the random-forest classifier.^[70] Cheng et al first segmented the optic disc and cup using histograms of pixel intensities, neighbor statistics, and incorporation of location pixel information to first compute CDR for glaucoma screening and

thus obtained AUCs up to 0.82 based on two independent datasets.^[71] More complex learning processes such as multi-task learning has been used to detect glaucoma and several other ocular conditions from fundus photographs.^[72]

Deep learning models could, however, learn complex glaucoma features using several layers of neurons in an end-to-end process. One of the first deep CNN models in glaucoma used a network with four convolutional layers and two fully connected layers. This model obtained AUCs up to 0.88 for detecting glaucoma from fundus photographs based on two different publicly available datasets.^[73] Some of the recent deep CNN models applied to fundus photographs have reached AUROCs up to 0.99 for glaucoma diagnosis.^[74–76, 78] Other deep learning models have obtained AUC up to about 0.97 for glaucoma screening and AUC up to 0.94 for glaucoma referral.^[36, 79] A recent meta-analysis paper analyzed the accuracy of seventeen deep learning-based studies that utilized 30 different patient cohorts and reported an AUC of 0.93 (95% CI 0.92–0.94) for diagnosing glaucoma based on color fundus photographs.^[80]

As OCT has become a dominant imaging modality for glaucoma assessment,^[33] numerous teams have explored the utility of glaucoma diagnosis based on OCT. Some studies have shown the usefulness of the OCT-derived RNFL parameters in distinguishing normal eyes from eyes with glaucoma without the utilization of AI models. Based on 94 normal subjects and patients with early glaucoma, Bowd et al obtained an AUC of 0.91 using COT parameters and showed the accuracy based on OCT was superior to

scanning laser polarimetry (SLP), frequency-doubling technology (FDT), and short-wavelength automated perimetry (SWAP) in discriminating normal eyes from eyes with early glaucoma.^[81] Another study used RNFL thickness parameters of 95 age-matched normal and glaucomatous eyes and obtained accuracies up to about 90% in distinguishing normal eyes from glaucomatous eyes based on the commercially available OCT instruments of Stratus and Cirrus (Carl Zeiss Meditec, Dublin, CA).^[82] These results were promising, and several follow-up studies showed that OCT can discriminate normal eyes from glaucomatous eyes with AUCs ranging from 0.89 to 0.96 based on RNFL or GCIPL thickness parameters derived from macula or ONH OCT images.^[37, 83–87] These studies showed the capability of OCT-derived retinal parameters in diagnosing glaucoma without employing machine learning models.

The capability of machine learning models in improving the diagnostic accuracy for glaucoma based on OCT-derived retinal parameters were investigated further. Burgansky-Eliash et al evaluated five conventional machine learning classifiers including linear discriminant analysis, SVM, recursive partitioning and regression tree, generalized linear model, and generalized additive model based on OCT-derived parameters of 89 normal and glaucomatous eyes and obtained the best AUC of 0.98 (specificity of 95% and sensitivity of 92.5%) for discriminating normal eyes from glaucomatous eyes using an SVM classifier.^[88] Follow-up AI models based on OCT-derived RNFL thickness measurements collected from 152 normal and glaucomatous eyes using SVM and ANN machine learning classifiers obtained AUCs up to about 0.99^[89] and another neural network-based model evaluated RNFL-derived OCT segmentation strategies and obtained an AUC up to 0.85 for glaucoma diagnosis.^[90]

Recent deep CNN models have also been applied to OCT-derived retinal parameters to diagnose glaucoma. Asaoka et al applied a deep learning model on over 4000 grids (8*8) of macular OCT-derived RNFL and ganglion complex layer (GCL) thickness profiles and obtained an AUC about 0.94.^[18] Follow-up deep learning models have obtained AUCs up to 0.99 based on OCT-derived retinal parameters.^[19–21] Deep learning approaches have also been applied to raw unsegmented OCT images for glaucoma diagnosis.

Ran et al developed a multi-task three-dimensional (3D) deep learning model to diagnose glaucoma based on over 8000 raw volumetric OCT scans collected from multiple institutes and obtained AUCs in the range of 0.86 to 0.90.^[91] A recent meta-analysis investigated five different deep learning studies that analyzed six cohorts of OCT and reported an AUC of about 0.96 (95% CI 0.94–0.99) for diagnosing glaucoma based on deep learning models. When averaged across the cohorts, the pooled sensitivity was 0.94 (95% CI 0.92–0.96) and pooled specificity was 0.95 (95% CI 0.91–0.97).^[80]

Machine learning models have been applied to other retinal imaging modalities as well. OCT angiography (OCTA), a recent imaging technology in ophthalmology, provides high-resolution images of retinal vasculature structure and function that are appropriate for deep learning models.^[92] OCTA-derived vasculature parameters have shown great promise in discriminating normal eyes from glaucomatous eyes without using any AI model.^[93–95] As OCAT is a newer technology, a limited number of AI models have explored this modality to date. A recent study investigated the capability of deep learning and conventional machine learning classifiers to diagnose glaucoma based on 405 OCTA images and quantified parameters. The best AUC of the gradient boosting classifier (GBC) model based on quantified OCTA parameters was 0.89 while a deep learning model based on a VGG16 architecture achieved an AUC of 0.93 based on radial peripapillary capillary en face OCTA images of the ONH.

Most applications of AI models have been centered around glaucoma detection for screening and diagnosis purposes, while forecasting glaucoma could play an important role in identifying those with future disease development and potential vision loss. Thakur et al developed a deep learning model based on over 60,000 fundus photographs to forecast glaucoma before disease development. They achieved AUCs up to approximately 0.77 and 0.88 for forecasting glaucoma four to seven years and one to three years before onset, respectively. Their model achieved an AUC of about 0.95 once tested to diagnose glaucoma.^[77] Other forecasting models are usually centered around predicting future VF or OCT parameters. Wen et al developed a model to forecast future VF tests (up to 5.5 years) from current VF tests using deep learning based on more than 30,000 VFs and obtained average

point-wise mean absolute errors (MAE) of about 2.5 dB.^[96] Sedai and colleagues developed a deep learning model to forecast RNFL thickness measurements from raw OCT and quantified RNFL thickness measurements, VFs, and clinical data collected from multiple visits, and reached mean MAEs as low as about 1.8 micron in estimating global RNFL thickness across normal eyes, and eyes with suspect and established glaucoma.^[97] Such validated models may facilitate personalized patient care by determining the most appropriate inter-visit schedule for timely interventions.

Clinical considerations

While two autonomous AI-enabled models have received US FDA approvals for screening diabetic retinopathy and macular edema,^[98, 99] no autonomous AI models have yet received FDA approval in glaucoma screening, diagnosis, or prognosis. It should be noted that most of the commercially available VF and OCT instruments already include some AI-enabled quantification and interpretation tools, and the evidence summarized above, strongly suggests that autonomous AI models would be warranted for glaucoma screening, diagnosis, and forecasting. Assistive AI models may benefit glaucoma clinical practice and augment clinical assessment while autonomous AI models may provide greater benefit to population-based screening. Nevertheless, various hurdles remain for full development and integration of assistive and autonomous AI models in glaucoma, as discussed in Section 6.

Applications of AI in Glaucoma Prognosis and Monitoring

Detecting glaucoma-induced structural and functional loss is critical for preserving vision and maintaining quality of life of patients with glaucoma. However, identifying glaucoma-induced vision changes by inspecting a sequence of fundus photographs, OCT images, and VFs can be perplexing at both ends of the glaucoma spectrum - in the early stages of the disease, where structural and functional deficits are subtle; or in the late stage of the disease, where OCT is unable to provide required dynamic ranges (flooring effect) and VF presents significant VF

variability.^[100, 101] Early methods for detecting glaucoma progression introduced in the 1990s include: Advanced Glaucoma Intervention Study (AGIS) criteria;^[102] Collaborative Initial Glaucoma Treatment Study (CIGTS) criteria,^[103] and the widely used Guided Progression Analysis (GPA).^[35] All these models provided event-based approaches utilizing ad-hoc rules to detect VF progression.

Point-wise linear regression (PLR)^[104] and Permutation of PLR (PoPLR)^[105] have provided trend-based approaches utilizing linear regression to identify VF progression. Other follow-up models used statistical analysis of summary or regional parameters such as VFI or mean deviation (MD) to detect glaucoma progression based on VFs.^[106–108] More complex models based on structure–function relationship using dynamic estimates of the current glaucoma state and velocity of progression over time showed improved accuracy over the ordinary linear regression approaches.^[108] These methods have used mathematical and static approaches to detect progression.

Lin et al introduced one of the first applications of machine learning models in detecting glaucomatous progression based on VFs.^[109] They obtained an AUC of 0.92 (average specificity and sensitivity of 88% and 86%, respectively) using a neural network with three hidden layers. Sample et al introduced one of the first unsupervised machine learning-based models to detect glaucoma progression.^[110] They developed an unsupervised variational Bayesian model and identified several prominent patterns of VF loss. They identified the progression of glaucoma across these patterns. Most of the follow-up machine learning models also used unsupervised machine learning models to analyze VFs.^[25, 28, 111, 112] As VF testing was an older technology compared to OCT, most of the early glaucoma progression models have been applied to VFs.

Wollstein et al investigated the utility of OCT-derived RNFL thickness measurements in detecting glaucoma progression and reported that OCT was more sensitive than VF in detecting glaucomatous progression.^[113] Similarly, other studies also showed the utility of OCT in detecting glaucoma progression.^[114] Yousefi et al evaluated the usefulness of several supervised machine learning models to detect glaucoma based on VFs and OCT parameters and reported the superiority of OCT in detecting glaucoma progression compared with VFs.^[115] While unsupervised

archetypal analysis has been used for detecting glaucoma progression,^[116] a recent study used deep archetypal analysis to identify patterns of VF loss and then used some of those patterns for detecting ocular hypertensive patients with future rapid glaucoma progression (rate of MD loss faster than -1 dB/year).^[117] The application of deep CNN models in detecting glaucoma progression, however, has been limited. A recent study showed the effectiveness of a convolutional long short-term memory (LSTM) neural network that was trained and tested on over 670,000 VFs in identifying glaucoma progression from VFs with AUCs from 0.79 to 0.82.^[118]

Clinical considerations

Most of the early applications of AI in detecting glaucoma progression have been focused on VFs. This may be explained by two facts. First, the longer existence of VF testing technology in glaucoma care has resulted in the availability of more datasets with longer follow ups (compared to OCT). Second, VFs are already in numeric format and appropriate for most conventional machine learning models which is not the case for fundus images. AI-enabled models for detecting glaucoma progression based on fundus images do exist, but they are rare. This may reflect the fact that quantification of fundus images to provide appropriate input for most conventional machine learning models was challenging. In contrast to fundus imaging, OCT imaging already includes quantification such as retinal thickness profiles that are appropriate for most conventional machine learning models. With the advancement of deep learning models however, more innovative AI models would be desirable in order to fully exploit color fundus photographs as well as raw OCT images for detecting glaucoma progression in clinical practice.

Most current progression detection models utilize statistical approaches based on linear regression, and assume that glaucoma progresses linearly, while there is evidence that glaucomatous progression may be non-linear and rapid, particularly during the later stages.^[119, 120] Additionally, most of these methods provide only information of whether the eye is progressed or not, without supplemental information on the type of pattern of loss. Therefore, unsupervised machine learning models may offer unbiased analysis of progression and provide explainable

outcomes with information on local patterns of loss, rather than a sole binary outcome. As glaucoma is a multifactorial disease caused by a complex interaction of multiple factors, using a comprehensive set of input information may facilitate detection of glaucoma progression. However, no instrument/device yet provides a comprehensive analysis of glaucoma progression based on combined imaging and VF data along with ocular, clinical, and demographic factors. This is an unmet need and future AI models may facilitate detection of glaucoma using multiple sources of information.

Applications of AI in Estimating Functional Parameters from Structural Factors

Advancements in AI models in glaucoma have posed critical questions regarding the feasibility of using objective OCT measurements to assess and monitor visual functional loss. A successful solution may replace subjective and tedious VF testing with objective and quick OCT imaging for glaucoma assessment. To that end, numerous teams have attempted to estimate global, regional, and point-wise VF parameters from raw OCT or OCT-derived measurements based on statistical models or machine learning approaches.^[121–123] Some recent studies have utilized deep learning models to estimate global and local VF damage from raw OCT scans and quantified thickness measurements.^[124–126] These studies have used scanning laser polarimetry (SLP)-derived RNFL thickness measurements to estimate VF threshold sensitivities based on linear and non-linear regression, and obtained approximately 3.9 dB errors (Zhu et al),^[121] OCT-derived retinal parameters to estimate VF sensitivities based on support vector regressor machines and achieved a root mean square error (RMSE) of about 3.7 dB,^[122] OCT-derived RNFL to estimate VF sensitivities based on deep learning models and obtained RMSE of about 6.1 dB,^[124] OCT-derived RNFL thickness measurements to estimate global VF mean deviation (MD) based on deep learning and achieved MAE of about 2.9 dB,^[125] raw OCT images from macula and optic disc to estimate VF global parameters based on 3-D deep learning models and obtained RMSE about 2.4 dB and MAE of about 2.3 dB.^[126] A follow-up model used an artificial neural network (ANN) model to estimate MD from OCT-derived RNFL parameters based

on four independent large datasets from different races, different instruments, and different scanning types and obtained MAE of about 4.0 dB and RMSE of approximately 5.2 dB with reasonable generalizability on other datasets.^[127]

Clinical considerations

Validated AI models may replace VF testing that can be tedious, subjective, and highly variable at later stages of the disease, with OCT imaging that is more objective, quick, and reproducible. Such models could generate quick outcomes and provide more objective glaucoma assessment. However, most of the current models for predicting VF parameters from OCT have several shortcomings. Some of these models underestimate or overestimate global or local VF parameters at both ends of the glaucoma spectrum. Additionally, while the overall error rate of these models may fall within the VF variability, still lower local error rates are required to reach a reasonable level for clinical applications. Finally, the generalizability of these models needs to be evaluated based on representative clinical data to gain clinical utility.

Unresolved Challenges and Future Directions

As discussed in previous sections, AI models may perform a wide range of tasks such as retinal image and data annotation and interpretation, diagnosis, and prognosis, in order to enhance glaucoma research and clinical practice. Some of the AI models could generate outcomes more quickly, accurately, and consistently than more standard approaches. However, some of the AI models, particularly deep learning, have several limitations as listed below.

Unstable: Deep learning models could be fragile and sometimes even with slightly modifying (unrelated) regions of retina (e.g., by flipping pixels), the diagnosis outcome of the model may change.

Biased: Embedded bias in data could simply become integrated into the model. An AI model that has been trained on retinal images from glaucoma subjects at later stages will be biased and may simply miss patients at early stages of the disease.

Memory: Models may lose their previous ability if retrained on new data. An AI model that was trained

on retinal images from glaucoma patients at the early stages of the disease may lose its capability if retrained on new data from subjects in later stages of the disease. The model may simply forget its previous capabilities.

Unexplainable: While there have been efforts in explaining the outcome of deep learning models, still explainability is a critical challenge in applications in ophthalmology. Clinicians better trust models that explain reasons why they have made a decision rather than providing a sole binary diagnosis.

Uncertain: Except in rare cases,^[128] most deep learning models only provide the likelihood of diagnosis and not certainty. In contrary to human expertise that may be wrong on challenging cases, a deep learning model be wrong on simple cases (provide a high likelihood on a definitely wrong decisions).

Foolish: Despite remarkable outcomes in some applications, deep learning models may make foolish mistakes. For instance, a model that has been developed to diagnose glaucoma based on fundus photographs may say a subject has glaucoma from an irrelevant input picture of the lung, while human experts won't make such simple mistakes.

Addressing some of these challenges are the bases of currently active research areas and innovative solutions are becoming increasingly available. In addition to technical challenges, other limitations have hindered widespread clinical utility of AI and deep learning models in glaucoma clinical practice. Some of these challenges are as follows.

Challenges related to glaucoma definitions:

Inconsistent definition of glaucoma: There is not yet a widespread consensus on glaucoma definition and different studies and guidelines have used different definitions.^[129, 130]

Inconsistent definition of glaucoma progression: The problem is even worse for progression as there is no consensus on glaucoma-induced changes as well as a level that constitute a real change.

Challenges related to training and testing AI models:

Most of these reference datasets are annotated by non-ophthalmologists, ophthalmologists, or glaucoma experts with diverse levels of expertise and considerable intra-and inter-rater variability rates. Evaluating AI models based on

different datasets thus leads to different levels of performance.

The datasets may be selective and may not represent a diverse group of patients with different ethnicities, required phenotypes, or different severity levels thus generating selective bias.

Datasets are not from the targeted AI model. For instance, the AI model is for proposed for glaucoma screening, but training data is collected from subjects visiting a tertiary eye hospital.

Challenges related to integration of AI in glaucoma clinical practice:

Performance evaluation: Some of the AI model have reported studies that are not based on AI-related study design and reporting guidelines^[131–133] leading to challenges in validating those models.

Acceptable performance (specificity and sensitivity): The required performance level of a model for glaucoma screening is different from a model targeted for glaucoma diagnosis or prognosis. These levels are not well-defined in glaucoma screening and diagnosis. For instance, as glaucoma is a low- prevalence disease, highly specific and sensitive models are required for screening.

In some AI models, it is confusing whether the model is assistive or autonomous and subsequently whether the AI model has been validated logically. While most of the AI models for retinal image quantification and interpretation are assistive AI tools, most of the diagnostic AI models are autonomous in generating a diagnosis. For instance, if a model is assistive, the evaluation phase requires involvement of the physicians/glaucoma specialists as well. Thus, evaluating the AI model alone limits clinical utility.

Other broad challenges related to integration of AI in glaucoma clinical practice:

Standards for oversight of Software: In different countries, it is still challenging to understand who will oversight an AI system in clinical practice? Developers, physicians, clinics, or providers? These need to be elucidated and each country's regulations may impact this differently.

Liability: It is vital to determine who is (are) responsible for a misdiagnosis or missed diagnosis. These are highly dependent on local regulations and legal systems as well.

Ethical considerations: The minimum requirement for AI models is to not to harm

patient. The broader view is that these models would also need to benefit patients and improve clinical and patient outcome.

Reimbursement: Issues need to be resolved for reimbursement and revenue sharing with those who are involved in clinical care, if these AI systems are to receive widespread clinical utility.

Sharing and privacy: The minimum requirement for AI models is that they not violate patient safety and privacy. There are several additional aspects to be clarified, including dataset sharing and who owns the datasets used to train AI models.

Summary

AI has shown tremendous potential in both research and clinical treatment of glaucoma. Various conventional AI and emerging deep learning models have been proposed to quantify retinal images and VFs in order to screen, diagnose, forecast, and prognose glaucoma. Some of the AI assistive models have already been integrated in some glaucoma imaging and VF instruments; however, no autonomous AI model has yet received US regulatory approvals to be used in glaucoma care. While there are many challenges regarding integration of AI in glaucoma clinics, a major challenge is the lack of a widely used reference standard for glaucoma, as most of the AI models are trained based on datasets that are subjectively evaluated based on different definitions of glaucoma or its progression. Other challenges include lack of standardized evaluation and reporting of the performance of AI models, targeted patient populations, and liability and ethical issues. Nevertheless, AI applications can provide major improvements in several important areas including *glaucoma research* by setting common grounds for reproducible factors, screening programs with highly specific and sensitive autonomous models for detecting glaucoma, clinical care with establishing assistive and autonomous glaucoma models for delineating hallmarks and diagnosis, and in clinical trial design by identifying subjects and even offering novel digital endpoints.

Financial Support and Sponsorship

This work was supported by the National Institute of Health (NIH) grants EY033005, EY031725, and

a Challenge Grant from Research to Prevent Blindness (RPB), New York. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

The author has no relevant conflict of interest to disclose.

REFERENCES

- Ting DS, Cheung GC, Wong TY. Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges: A review. *Clin Exp Ophthalmol* 2016;44:260–277.
- Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016;316:2402–2410.
- De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018;24:1342–1350.
- Kelman SE, Perell HF, D'Autrechy L, Scott RJ. A neural network can differentiate glaucoma and optic neuropathy visual fields through pattern recognition. In: Mills RP, Heijl A, editors. *Perimetry Update 1990/1991, Proceedings of the IXth International Perimetric Society Meeting*. Amsterdam/New York: Kugler Publications; 1991. p. 291–295.
- Nagata S, Kani K, Sugiyama A. A computer assisted visual field diagnosis system using neural networks. In: Mills RP, Heijl A, editors. *Perimetry Update 1990/1991, Proceedings of the IXth International Perimetric Society Meeting*. Amsterdam/New York: Kugler Publications; 1991. p. 291–295.
- Mendels F, Heneghan C, Thiran J-P. Identification of the optic disk boundary in retinal images using active contours. 1999.
- 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* 2002;49:963–974.
- Bengtsson B, Bizios D, Heijl A. Effects of input data on the performance of a neural network in distinguishing normal and glaucomatous visual fields. *Invest Ophthalmol Vis Sci* 2005;46:3730–3736.
- Bizios D, Heijl A, Bengtsson B. Trained artificial neural network for glaucoma diagnosis using visual field data: A comparison with conventional algorithms. *J Glaucoma* 2007;16:20–28.
- Wroblewski D, Francis B, Chopra V, Kawji AS, Quiros P, Dustin L, et al. Glaucoma detection and evaluation through pattern recognition in standard automated perimetry data. *Graefes Arch Clin Exp Ophthalmol* 2009;247:1517–1530.
- Chrastek R, Wolf M, Donath K, Niemann H, Paulus D, Hothorn T, et al. Automated segmentation of the optic nerve head for diagnosis of glaucoma. *Med Image Anal* 2005;9:297–314.
- Lim G, Cheng Y, Hsu W, Lee ML. Integrated optic disc and cup segmentation with deep learning. Paper presented at: 2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI); 2015 Nov 9-11 Nov.
- Lee CS, Tying AJ, Deruyter NP, Wu Y, Rokem A, Lee AY. Deep-learning based, automated segmentation of macular edema in optical coherence tomography. *Biomed Opt Express* 2017;8:3440–3448.
- Kugelman J, Alonso-Caneiro D, Read SA, Vincent SJ, Collins MJ. Automatic segmentation of OCT retinal boundaries using recurrent neural networks and graph search. *Biomed Opt Express* 2018;9:5759–5777.
- Masood S, Fang R, Li P, Li H, Sheng B, Mathavan A, et al. Automatic choroid layer segmentation from optical coherence tomography images using deep learning. *Sci Rep* 2019;9:3058.
- Zhang H, Yang J, Zhou K, Li F, Hu Y, Zhao Y, et al. Automatic segmentation and visualization of choroid in OCT with knowledge infused deep learning. *IEEE J Biomed Health Inform* 2020;24:3408–3420.
- Wu Q, Zhang B, Hu Y, Liu B, Cao D, Yang D, et al. Detection of morphologic patterns of diabetic macular edema using a deep learning approach based on optical coherence tomography images. *Retina* 2020.
- Asaoka R, Murata H, Hirasawa K, Fujino Y, Matsuura M, Miki A, et al. Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. *Am J Ophthalmol* 2019;198:136–145.
- Lee J, Kim YK, Park KH, Jeoung JW. Diagnosing glaucoma with spectral-domain optical coherence tomography using deep learning classifier. *J Glaucoma* 2020;29:287–294.
- Ran AR, Cheung CY, Wang X, Chen H, Luo LY, Chan PP, et al. Detection of glaucomatous optic neuropathy with spectral-domain optical coherence tomography: A retrospective training and validation deep-learning analysis. *Lancet Digit Health* 2019;1:e172–e182.
- Thompson AC, Jammal AA, Berchuck SI, Mariottoni EB, Medeiros FA. Assessment of a segmentation-free deep learning algorithm for diagnosing glaucoma from optical coherence tomography scans. *JAMA Ophthalmol* 2020;138:333–339.
- Sample PA, Chan K, Boden C, Lee TW, Blumenthal EZ, Weinreb RN, et al. Using unsupervised learning with variational bayesian mixture of factor analysis to identify patterns of glaucomatous visual field defects. *Investig Ophthalmol Vis Sci* 2004;45:2596–2605.
- Goldbaum MH, Sample PA, Zhang Z, Chan K, Hao J, Lee TW, et al. Using unsupervised learning with independent component analysis to identify patterns of glaucomatous visual field defects. *Investig Ophthalmol Vis Sci* 2005;46:3676–3683.
- Bowd C, Weinreb RN, Balasubramanian M, Lee I, Jang G, Yousefi S, et al. Glaucomatous patterns in frequency doubling technology (FDT) perimetry data identified by unsupervised machine learning classifiers. *PLoS One* 2014;9:e85941.

25. Yousefi S, Goldbaum MH, Balasubramanian M, Medeiros FA, Zangwill LM, Liebmann JM, et al. Learning from data: Recognizing glaucomatous defect patterns and detecting progression from visual field measurements. *IEEE Trans Biomed Eng* 2014;61:2112–2124.
26. Yousefi S, Goldbaum MH, Zangwill LM, Medeiros FA, Bowd C. Recognizing patterns of visual field loss using unsupervised machine learning. *Proc SPIE Int Soc Opt Eng* 2014;2014:90342M.
27. Elze T, Pasquale LR, Shen LQ, Chen TC, Wiggs JL, Bex PJ. Patterns of functional vision loss in glaucoma determined with archetypal analysis. *J R Soc Interface* 2015;12:20141118.
28. Yousefi S, Balasubramanian M, Goldbaum MH, Medeiros FA, Zangwill LM, Weinreb RN, et al. Unsupervised Gaussian mixture-model with expectation maximization for detecting glaucomatous progression in standard automated perimetry visual fields. *Transl Vis Sci Technol* 2016;5:2.
29. Wang M, Shen LQ, Pasquale LR, Boland MV, Wellik SR, De Moraes CG, et al. Artificial intelligence classification of central visual field patterns in glaucoma. *Ophthalmology*.
30. Thakur A, Goldbaum M, Yousefi S. Convex representations using deep archetypal analysis for predicting glaucoma. *IEEE J Transl Eng Health Med* 2020;8:3800107.
31. Gupta K, Thakur A, Goldbaum M, Yousefi S. Glaucoma precognition: Recognizing preclinical visual functional signs of glaucoma. Paper presented at: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2020 June 14–19.
32. Nadler Z, Wollstein G, Ishikawa H, Schuman JS. Clinical application of ocular imaging. *Optom Vis Sci* 2012;89:E543–E553.
33. Stein JD, Talwar N, Laverne AM, Nan B, Lichter PR. Trends in use of ancillary glaucoma tests for patients with open-angle glaucoma from 2001 to 2009. *Ophthalmology* 2012;119:748–758.
34. Alencar LM, Medeiros FA. The role of standard automated perimetry and newer functional methods for glaucoma diagnosis and follow-up. *Indian J Ophthalmol* 2011;59:S53–S58.
35. Bengtsson B, Heijl A. A visual field index for calculation of glaucoma rate of progression. *Am J Ophthalmol* 2008;145:343–353.
36. Phene S, Dunn RC, Hammel N, Liu Y, Krause J, Kitade N, et al. Deep learning and glaucoma specialists: The relative importance of optic disc features to predict glaucoma referral in fundus photographs. *Ophthalmology* 2019;126:1627–1639.
37. Bussel, Il, Wollstein G, Schuman JS. OCT for glaucoma diagnosis, screening and detection of glaucoma progression. *Br J Ophthalmol* 2014;98:ii15–19.
38. Brusini P, Johnson CA. Staging functional damage in glaucoma: Review of different classification methods. *Surv Ophthalmol* 2007;52:156–179.
39. Hoover A, Goldbaum M. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Trans Med Imaging* 2003;22:951–958.
40. Wong DW, Liu J, Lim JH, Tan NM, Zhang Z, Lu S, et al. Intelligent fusion of cup-to-disc ratio determination methods for glaucoma detection in ARGALI. *Annu Int Conf IEEE Eng Med Biol Soc* 2009;2009:5777–5780.
41. Aquino A, Gegundez-Arias ME, Marin D. Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques. *IEEE Trans Med Imaging* 2010;29:1860–1869.
42. Haleem MS, Han L, Hemert JV, Li B, Fleming A, Pasquale LR, et al. A novel adaptive deformable model for automated optic disc and cup segmentation to aid glaucoma diagnosis. *J Med Syst* 2017;42:20.
43. Hu M, Zhu C, Li X, Xu Y. Optic cup segmentation from fundus images for glaucoma diagnosis. *Bioengineered* 2017;8:21–28.
44. Wang J, Wang Z, Li F, Qu G, Qiao Y, Lv H, et al. Joint retina segmentation and classification for early glaucoma diagnosis. *Biomed Opt Express* 2019;10:2639–2656.
45. Zhou W, Yi Y, Gao Y, Dai J. Optic disc and cup segmentation in retinal images for glaucoma diagnosis by locally statistical active contour model with structure prior. *Comput Math Methods Med* 2019;2019:8973287.
46. Fu H, Cheng J, Xu Y, Zhang C, Wong DWK, Liu J, et al. Disc-aware ensemble network for glaucoma screening from fundus image. *IEEE Trans Med Imaging* 2018;37:2493–2501.
47. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, et al. Optical coherence tomography. *Science (New York, NY)* 1991;254:1178–1181.
48. Hood DC, Fortune B, Arthur SN, Xing D, Salant JA, Ritch R, et al. Blood vessel contributions to retinal nerve fiber layer thickness profiles measured with optical coherence tomography. *J Glaucoma* 2008;17:519–528.
49. Koozekanani D, Boyer K, Roberts C. Retinal thickness measurements from optical coherence tomography using a Markov boundary model. *IEEE Trans Med Imaging* 2001;20:900–916.
50. Ishikawa H, Stein DM, Wollstein G, Beaton S, Fujimoto JG, Schuman JS. Macular segmentation with optical coherence tomography. *Invest Ophthalmol Vis Sci* 2005;46:2012–2017.
51. Kafieh R, Rabbani H, Abramoff MD, Sonka M. Intra-retinal layer segmentation of 3D optical coherence tomography using coarse grained diffusion map. *Med Image Anal* 2013;17:907–928.
52. Gu Z, Cheng J, Fu H, Zhou K, Hao H, Zhao Y, et al. CE-Net: Context encoder network for 2D medical image segmentation. *IEEE T Med Imaging* 2019;38:2281–2292.
53. Wilson M, Chopra R, Wilson MZ, Cooper C, MacWilliams P, Liu Y, et al. Validation and clinical applicability of whole-volume automated segmentation of optical coherence tomography in retinal disease using deep learning. *JAMA Ophthalmol* 2021;139:964–973.
54. 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.
55. Brusini P. Clinical use of a new method for visual field damage classification in glaucoma. *Eur J Ophthalmol* 1996;6:402–407.
56. Keltner JL, Johnson CA, Cello KE, Edwards MA, Bandermann SE, Kass MA, et al. Classification of visual field abnormalities in the ocular hypertension treatment study. *Arch Ophthalmol* 2003;121:643–650.

57. Lichter PR. Variability of expert observers in evaluating the optic disc. *Trans Am Ophthalmol Soc* 1976;74:532–572.
58. Jampel HD, Friedman D, Quigley H, Vitale S, Miller R, Knezevich F, et al. Agreement among glaucoma specialists in assessing progressive disc changes from photographs in open-angle glaucoma patients. *Am J Ophthalmol* 2009;147:39–44 e31.
59. Damms T, Dannheim F. Sensitivity and specificity of optic disc parameters in chronic glaucoma. *Invest Ophthalmol Vis Sci* 1993;34:2246–2250.
60. Gundersen KG, Heijl A, Bengtsson B. Sensitivity and specificity of structural optic disc parameters in chronic glaucoma. *Acta Ophthalmol Scand* 1996;74:120–125.
61. Goldbaum MH, Sample PA, White H, Colt B, Raphaelian P, Fechtner RD, Weinreb RN. Interpretation of automated perimetry for glaucoma by neural network. *Invest Ophthalmol Vis Sci* 1994;35:3362–3373.
62. Madsen EM, Yolton RL. Demonstration of a neural network expert system for recognition of glaucomatous visual field changes. *Mil Med* 1994;159:553–557.
63. Spenceley SE, Henson DB, Bull DR. Visual field analysis using artificial neural networks. *Ophthalmic Physiol Opt* 1994;14:239–248.
64. Lietman T, Eng J, Katz J, Quigley HA. Neural networks for visual field analysis: How do they compare with other algorithms? *J Glaucoma* 1999;8:77–80.
65. Li F, Wang Z, Qu G, Song D, Yuan Y, Xu Y, et al. Automatic differentiation of glaucoma visual field from non-glaucoma visual field using deep convolutional neural network. *BMC Med Imaging* 2018;18:35.
66. Huang X, Jin K, Zhu J, Xue Y, Si K, Zhang C, et al. A structure-related fine-grained deep learning system with diversity data for universal glaucoma visual field grading. *Front Med* 2022;9:832920.
67. Bowd C, Medeiros FA, Zhang ZH, Zangwill LM, Hao JC, Lee TW, et al. Relevance vector machine and support vector machine classifier analysis of scanning laser polarimetry retinal nerve fiber layer measurements. *Investig Ophthalmol Vis Sci* 2005;46:1322–1329.
68. Burgansky-Eliash Z, Wollstein G, Bilonick RA, Ishikawa H, Kagemann L, Schuman JS. Glaucoma detection with the Heidelberg retina tomograph 3. *Ophthalmology* 2007;114:466–471.
69. Bock R, Meier J, Nyul LG, Hornegger J, Michelson G. Glaucoma risk index: Automated glaucoma detection from color fundus images. *Med Image Anal* 2010;14:471–481.
70. Acharya UR, Dua S, Du X, Sree SV, Chua CK. Automated diagnosis of glaucoma using texture and higher order spectra features. *IEEE Trans Inf Technol Biomed: A publication of the IEEE Eng Med Biol Soc* 2011;15:449–455.
71. Cheng J, Liu J, Xu Y, Yin F, Wong DW, Tan NM, et al. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE Trans Med Imaging* 2013;32:1019–1032.
72. Chen X, Xu Y, Yin F, Zhang Z, Wong DW, Wong TY, et al. Multiple ocular diseases detection based on joint sparse multi-task learning. *Annu Int Conf IEEE Eng Med Biol Soc* 2015;2015:5260–5263.
73. Xiangyu C, Yanwu X, Damon Wing Kee W, Tien Yin W, Jiang L. Glaucoma detection based on deep convolutional neural network. *Annu Int Conf IEEE Eng Med Biol Soc* 2015;2015:715–718.
74. Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA* 2017;318:2211–2223.
75. Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology* 2018;125:1199–1206.
76. Liu H, Li L, Wormstone IM, Qiao C, Zhang C, Liu P, et al. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs. *JAMA Ophthalmol* 2019;137:1353–1360.
77. Thakur A, Goldbaum M, Yousefi S. Predicting glaucoma before onset using deep learning. *Ophthalmol Glaucoma* 2020;3:262–268.
78. Li F, Su Y, Lin F, Li Z, Song Y, Nie S, et al. A deep-learning system predicts glaucoma incidence and progression using retinal photographs. *J Clin Invest* 2022;132:e157968.
79. Liu S, Graham SL, Schulz A, Kalloniatis M, Zangerl B, Cai W, et al. A deep learning-based algorithm identifies glaucomatous discs using monoscopic fundus photographs. *Ophthalmol Glaucoma* 2018;1:15–22.
80. Aggarwal R, Sounderajah V, Martin G, Ting DSW, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis. *NPJ Digit Med* 2021;4:65.
81. Bowd C, Zangwill LM, Berry CC, Blumenthal EZ, Vasile C, Sanchez-Galeana C, et al. Detecting early glaucoma by assessment of retinal nerve fiber layer thickness and visual function. *Invest Ophthalmol Vis Sci* 2001;42:1993–2003.
82. Chang RT, Knight OJ, Feuer WJ, Budenz DL. Sensitivity and specificity of time-domain versus spectral-domain optical coherence tomography in diagnosing early to moderate glaucoma. *Ophthalmology* 2009;116:2294–2299.
83. Xu J, Ishikawa H, Wollstein G, Bilonick RA, Folio LS, Nadler Z, et al. Three-dimensional spectral-domain optical coherence tomography data analysis for glaucoma detection. *PLoS One* 2013;8:e55476.
84. Mwanza JC, Oakley JD, Budenz DL, Anderson DR, Cirrus Optical Coherence Tomography Normative Database Study G. Ability of cirrus HD-OCT optic nerve head parameters to discriminate normal from glaucomatous eyes. *Ophthalmology* 2011;118:241–248 e241.
85. Sung KR, Na JH, Lee Y. Glaucoma diagnostic capabilities of optic nerve head parameters as determined by Cirrus HD optical coherence tomography. *J Glaucoma* 2012;21:498–504.
86. Lisboa R, Paranhos A, Jr., Weinreb RN, Zangwill LM, Leite MT, Medeiros FA. Comparison of different spectral domain OCT scanning protocols for diagnosing preperimetric glaucoma. *Invest Ophthalmol Vis Sci* 2013;54:3417–3425.
87. Hatanaka Y, Muramatsu C, Sawada A, Hara T, Yamamoto T, Fujita H. Glaucoma risk assessment based on clinical data and automated nerve fiber layer defects detection. *Conference proceedings: Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Conf* 2012;2012:5963–5966.
88. Burgansky-Eliash Z, Wollstein G, Chu T, Ramsey JD, Glymour C, Noecker RJ, et al. Optical coherence tomography machine learning classifiers for glaucoma

- detection: A preliminary study. *Invest Ophthalmol Vis Sci* 2005;46:4147–4152.
89. Bizios D, Heijl A, Hougaard JL, Bengtsson B. Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by Stratus OCT. *Acta Ophthalmol* 2010;88:44–52.
 90. Larrosa JM, Polo V, Ferreras A, Garcia-Martin E, Calvo P, Pablo LE. Neural network analysis of different segmentation strategies of nerve fiber layer assessment for glaucoma diagnosis. *J Glaucoma* 2015;24:672–678.
 91. Ran AR, Wang X, Chan PP, Chan NC, Yip W, Young AL, et al. Three-dimensional multi-task deep learning model to detect glaucomatous optic neuropathy and myopic features from optical coherence tomography scans: A retrospective multi-centre study. *Front Med* 2022;9:860574.
 92. WuDunn D, Takusagawa HL, Sit AJ, Rosdahl JA, Radhakrishnan S, Hoguet A, et al. OCT angiography for the diagnosis of glaucoma: A report by the American Academy of Ophthalmology. *Ophthalmology* 2021;128:1222–1235.
 93. Lee EJ, Lee KM, Lee SH, Kim TW. OCT angiography of the peripapillary retina in primary open-angle glaucoma. *Invest Ophthalmol Vis Sci* 2016;57:6265–6270.
 94. Rao HL, Kadambi SV, Weinreb RN, Puttaiah NK, Pradhan ZS, Rao DAS, et al. Diagnostic ability of peripapillary vessel density measurements of optical coherence tomography angiography in primary open-angle and angle-closure glaucoma. *Br J Ophthalmol* 2017;101:1066–1070.
 95. Yarmohammadi A, Zangwill LM, Diniz-Filho A, Suh MH, Manalastas PI, Fatehee N, et al. Optical coherence tomography angiography vessel density in healthy, glaucoma suspect, and glaucoma eyes. *Invest Ophthalmol Vis Sci* 2016;57:OCT451–OCT459.
 96. Wen JC, Lee CS, Keane PA, Xiao S, Rokem AS, Chen PP, et al. Forecasting future Humphrey visual fields using deep learning. *PLoS One* 2019;14:e0214875.
 97. Sedai S, Antony B, Ishikawa H, Wollstein G, Schuman JS, Garnavi R. Forecasting retinal nerve fiber layer thickness from multimodal temporal data incorporating OCT Volumes. *Ophthalmol Glaucoma* 2020;3:14–24.
 98. 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* 2018;1:39.
 99. Perez-Rueda A, Jimenez-Rodriguez D, Castro-Luna G. Diagnosis of subclinical keratoconus with a combined model of biomechanical and topographic parameters. *J Clin Med* 2021;10:2746.
 100. Henson DB, Chaudry S, Artes PH, Faragher EB, Ansons A. Response variability in the visual field: Comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. *Invest Ophthalmol Vis Sci* 2000;41:417–421.
 101. Mikelberg FS, Parfitt CM, Swindale NV, Graham SL, Drance SM, Gosine R. Ability of the Heidelberg retina tomograph to detect early glaucomatous visual field loss. *J Glaucoma* 1995;4:242–247.
 102. Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability. *Ophthalmology* 1994;101:1445–1455.
 103. Katz J. Scoring systems for measuring progression of visual field loss in clinical trials of glaucoma treatment. *Ophthalmology* 1999;106:391–395.
 104. Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. *Invest Ophthalmol Vis Sci* 2002;43:1400–1407.
 105. O'Leary N, Chauhan BC, Artes PH. Visual field progression in glaucoma: Estimating the overall significance of deterioration with permutation analyses of pointwise linear regression (PoPLR). *Invest Ophthalmol Vis Sci* 2012;53:6776–6784.
 106. Gardiner SK, Demirel S. Detecting change using standard global perimetric indices in glaucoma. *Am J Ophthalmol* 2017;176:148–156.
 107. Zhu H, Russell RA, Saunders LJ, Ceccon S, Garway-Heath DF, Crabb DP. Detecting changes in retinal function: Analysis with non-stationary Weibull error regression and spatial enhancement (ANSWERS). *PLoS One* 2014;9:e85654.
 108. Hu R, Marin-Franch I, Racette L. Prediction accuracy of a novel dynamic structure-function model for glaucoma progression. *Invest Ophthalmol Vis Sci* 2014;55:8086–8094.
 109. Lin A, Hoffman D, Gaasterland DE, Caprioli J. Neural networks to identify glaucomatous visual field progression. *Am J Ophthalmol* 2003;135:49–54.
 110. Sample PA, Boden C, Zhang Z, Pascual J, Lee TW, Zangwill LM, et al. Unsupervised machine learning with independent component analysis to identify areas of progression in glaucomatous visual fields. *Invest Ophthalmol Vis Sci* 2005;46:3684–3692.
 111. Goldbaum MH, Lee I, Jang G, Balasubramanian M, Sample PA, Weinreb RN, et al. Progression of patterns (POP): A machine classifier algorithm to identify glaucoma progression in visual fields. *Invest Ophthalmol Vis Sci* 2012;53:6557–6567.
 112. Yousefi S, Kiwaki T, Zheng Y, Sugiura H, Asaoka R, Murata H, et al. Detection of longitudinal visual field progression in glaucoma using machine learning. *Am J Ophthalmol* 2018;193:71–79.
 113. Wollstein G, Schuman JS, Price LL, Aydin A, Stark PC, Hertzmark E, et al. Optical coherence tomography longitudinal evaluation of retinal nerve fiber layer thickness in glaucoma. *Arch Ophthalmol* 2005;123:464–470.
 114. Na JH, Sung KR, Lee JR, Lee KS, Baek S, Kim HK, et al. Detection of glaucomatous progression by spectral-domain optical coherence tomography. *Ophthalmology* 2013;120:1388–1395.
 115. Yousefi S, Goldbaum MH, Balasubramanian M, Jung TP, Weinreb RN, Medeiros FA, et al. Glaucoma progression detection using structural retinal nerve fiber layer measurements and functional visual field points. *IEEE Trans Biomed Eng* 2014;61:1143–1154.
 116. Wang M, Shen LQ, Pasquale LR, Petrakos P, Formica S, Boland MV, et al. An artificial intelligence approach to detect visual field progression in glaucoma based on spatial pattern analysis. *Invest Ophthalmol Vis Sci* 2019;60:365–375.
 117. Yousefi S, Pasquale LR, Boland MV, Johnson CA. Machine-identified patterns of visual field loss and an association

- with rapid progression in the ocular hypertension treatment study. *Ophthalmology* 2022;129:1402–1411.
118. Dixit A, Yohannan J, Boland MV. Assessing glaucoma progression using machine learning trained on longitudinal visual field and clinical data. *Ophthalmology* 2020;128:1016–1026.
 119. Pathak M, Demirel S, Gardiner SK. Nonlinear trend analysis of longitudinal pointwise visual field sensitivity in suspected and early glaucoma. *Transl Vis Sci Technol* 2015;4:8.
 120. Chen A, Nouri-Mahdavi K, Otarola FJ, Yu F, Affifi AA, Caprioli J. Models of glaucomatous visual field loss. *Invest Ophthalmol Vis Sci* 2014;55:7881–7887.
 121. Zhu H, Crabb DP, Schlottmann PG, Lemij HG, Reus NJ, Healey PR, et al. Predicting visual function from the measurements of retinal nerve fiber layer structure. *Invest Ophthalmol Vis Sci* 2010;51:5657–5666.
 122. Bogunovic H, Kwon YH, Rashid A, Lee K, Critser DB, Garvin MK, et al. Relationships of retinal structure and Humphrey 24-2 visual field thresholds in patients with glaucoma. *Invest Ophthalmol Vis Sci* 2014;56:259–271.
 123. Guo Z, Kwon YH, Lee K, Wang K, Wahle A, Alward WLM, et al. Optical coherence tomography analysis based prediction of Humphrey 24-2 visual field thresholds in patients with glaucoma. *Invest Ophthalmol Vis Sci* 2017;58:3975–3985.
 124. Sugiura H, Kiwaki T, Yousefi S, Murata H, Asaoka R, Yamanishi K. Estimating glaucomatous visual sensitivity from retinal thickness with pattern-based regularization and visualization. *Kdd'18: Proc 24th Acm Sigkdd Int Conf Knowl Discov Data Min* 2018:783–792.
 125. Christopher M, Bowd C, Belghith A, Goldbaum MH, Weinreb RN, Fazio MA, 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* 2020;127:346–356.
 126. Yu HH, Maetschke SR, Antony BJ, Ishikawa H, Wollstein G, Schuman JS, et al. Estimating global visual field indices in glaucoma by combining macula and optic disc OCT scans using 3-dimensional convolutional neural networks. *Ophthalmol Glaucoma* 2021;4:102–112.
 127. Huang X, Sun J, Majoor J, Vermeer KA, Lemij H, Elze T, et al. Estimating the severity of visual field damage from retinal nerve fiber layer thickness measurements with artificial intelligence. *Transl Vis Sci Technol* 2021;10:16.
 128. Huang X, Sun J, Gupta K, Montesano G, Crabb DP, Garway-Heath DF, et al. Detecting glaucoma from multimodal data using probabilistic deep learning. *Front Med* 2022;9:923096.
 129. Prum BE, Jr., Rosenberg LF, Gedde SJ, Mansberger SL, Stein JD, Moroi SE, et al. Primary open-angle glaucoma preferred practice pattern((R)) guidelines. *Ophthalmology* 2016;123:P41–P111.
 130. European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Chapter 2: Classification and terminology. Supported by the EGS Foundation: Part 1: Foreword; Introduction; Glossary; Chapter 2 Classification and terminology. *Br J Ophthalmol* 2017;101:73–127.
 131. Sounderajah V, Ashrafian H, Golub RM, Shetty S, De Fauw J, Hooft L, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: The STARD-AI protocol. *BMJ Open* 2021;11:e047709.
 132. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension. *Nat Med* 2020;26:1351–1363.
 133. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI Extension. *BMJ* 2020;370:m3164.