AMERICAN ACADEMY OF OPHTHALMOLOGY®
In partnership with the AMERICAN GLAUCOMA SOCIETY

A Deep Learning-Based Algorithm Identifies Glaucomatous Discs Using Monoscopic Fundus Photographs

Sidong Liu, PhD,[1,2] Stuart L. Graham, FRANZCO, PhD,[3] Angela Schulz, PhD,[3] Michael Kalloniatis, PhD,[4] Barbara Zangerl, PhD, DVM,[4] Weidong Cai, PhD,[5] Yang Gao, PhD,[6] Brian Chua, FRANZCO,[7] Hemamalini Arvind, FRANZCO, PhD,[1] John Grigg, MD, FRANZCO,[1,7] Dewei Chu, PhD,[8] Alexander Klistorner, MD, PhD,[1,3] Yuyi You, MD, PhD[1,3]

Purpose: To develop and test the performance of a deep learning-based algorithm for glaucomatous disc identification using monoscopic fundus photographs.

Design: Fundus photograph database study.

Participants: Four thousand three hundred ninety-four fundus photographs, including 3768 images from previous Sydney-based clinical studies and 626 images from publicly available online RIM-ONE and High-Resolution Fundus (HRF) databases with definitive diagnoses.

Methods: We merged all databases except the HRF database, and then partitioned the dataset into a training set (80% of all cases) and a testing set (20% of all cases). We used the HRF images as an additional testing set. We compared the performance of the artificial intelligence (AI) system against a panel of practicing ophthalmologists including glaucoma subspecialists from Australia, New Zealand, Canada, and the United Kingdom.

Main Outcome Measures: The sensitivity and specificity of the AI system in detecting glaucomatous optic discs.

Results: By using monoscopic fundus photographs, the AI system demonstrated a high accuracy rate in glaucomatous disc identification (92.7%; 95% confidence interval [CI], 91.2%−94.2%), achieving 89.3% sensitivity (95% CI, 86.8%−91.7%) and 97.1% specificity (95% CI, 96.1%−98.1%), with an area under the receiver operating characteristic curve of 0.97 (95% CI, 0.96−0.98). Using the independent online HRF database (30 images), the AI system again accomplished high accuracy, with 86.7% in both sensitivity and specificity (for ophthalmologists, 75.6% sensitivity and 77.8% specificity) and an area under the receiver operating characteristic curve of 0.89 (95% CI, 0.76−1.00).

Conclusions: This study demonstrated that a deep learning-based algorithm can identify glaucomatous discs at high accuracy level using monoscopic fundus images. Given that it is far easier to obtain monoscopic disc images than high-quality stereoscopic images, this study highlights the algorithm's potential application in large population-based disease screening or telemedicine programs. *Ophthalmology Glaucoma 2018;1:15-22 © 2018 by the American Academy of Ophthalmology*

Supplemental material available at www.ophthalmologyglaucoma.org.

Glaucoma has been recognized as one of the leading causes of blindness worldwide, predicted to affect 76 million people in 2020 and more than 110 million in 2040.[1] Given the irreversible nature of glaucomatous optic nerve degeneration combined with increased longevity of the population, early diagnosis is important to prevent severe visual morbidity and the associated healthcare and society burdens. Current open-angle glaucoma detection has been established in a community-based healthcare system, and it is heavily dependent on accessibility to qualified ophthalmologists, optometrists, or general practitioners for clinical examination as well as dedicated ophthalmic testing equipment, including tonometers, automated perimetry, and disc photography or OCT. However, it is recognized that there are a substantial number of undiagnosed glaucoma cases in the population.[2–4] Open-angle glaucoma screening is problematic because of the high cost of case identification, the availability of specialized facilities, and the need for testing to maintain high specificity while preserving sensitivity. Access problems are particularly relevant to many rural areas and developing countries.[5] Compared with automated perimetry and OCT, fundus photography can be obtained in a more cost-efficient manner, and qualified ophthalmologists have achieved reasonably good results in glaucomatous disc identification based on stereoscopic discs photographs.[6,7] It is far easier to obtain monoscopic fundus photographs than it is to obtain high-quality stereoscopic discs images. Automated assessment of high-quality fundus

https://doi.org/10.1016/j.ogla.2018.04.002
ISSN 2589-4196/18

photographs obtained with smart phones could be used to screen a large population for ophthalmic diseases.[8]

Automated glaucomatous disc identification using fundus images is a challenging task owing to the variability in the disc appearance and factors such as magnification, angle, color, focus, and lighting. Conventional image-based classification methods were limited in their ability to process fundus images at the pixel level and require handcrafted feature descriptors to transform the raw pixel arrays to a suitable representation (i.e., a vector in a feature space), which amplifies the aspects that are important for discrimination and suppresses irrelevant variations. Based on the feature vectors, the classification system could detect and classify patterns in the images (i.e., using a hyperplane to separate feature vectors in the same feature space). However, this usually requires domain knowledge to design such feature descriptors and a considerable amount of engineering work to implement them.

Deep learning[7] methods show that good representations of data can be learned automatically using a multilayer convolutional neural network with multiple levels of abstraction. These methods have improved dramatically the state of the art in many visual tasks, such as object recognition,[9,10] object detection,[11] and playing visual games.[12] Deep learning discovers intricate structure in large datasets by automatically adjusting the internal parameters of the system that are used to compute the representations of data. A typical deep convolutional neural network model uses only pixel value categorical labels as inputs and can be trained end to end from images directly.

The deep learning-empowered artificial intelligence (AI) system has demonstrated outstanding performance in skin cancer classification, better than United States board-certified dermatologists.[13] The deep learning technology also has been applied recently in pilot investigations in the field of ophthalmology. It has been shown that the use of deep learning-based algorithms can improve the assessment and grading of diabetic retinopathy (DR) significantly.[14–17] Unlike DR, glaucoma does not have a definitive disease cause, and therefore a screening program will need to target the general population (in contrast to DR screening only in diabetic patients) for blindness prevention. There was an attempt to use deep learning for analysis of visual fields, and preperimetric glaucomatous visual fields were found to be distinguishable from normal fields.[18] However, as stated previously, access to perimetric testing, testing duration, and cost could limit the potential of using automated perimetry as a glaucoma screening test in the general community. Furthermore, it is known that in most cases, retinal nerve fiber layer or structural changes occur before detectable functional defects in glaucoma; however, in the Ocular Hypertension Treatment Study (OHTS) study, in which disc photography, but not OCT, was used as an end point, a significant number of patients converted on visual fields first.[19] As such, early detection based on structural changes may be a more beneficial screening approach. Therefore, in this study, we developed a deep learning-based AI system using monoscopic fundus images from glaucoma patients and normal participants. The AI system performance was assessed and compared with a panel of practicing ophthalmologists from Australia, New Zealand, Canada, and the United Kingdom.

Table 1. Demographic Data of the Fundus Photographs from Sydney-Based Studies

	Glaucoma	Normal
No. of images	2369	1419
No. of participants	627	403
No. of eyes	1254	806
Mean age ± SD (yrs)	60.9±12.8	55.6±12.2
Gender (male/female)	362/265	196/207
Race (white/Asian)	586/41	363/40

SD = standard deviation.

Methods

Dataset

Fundus photographs were obtained from 3 tertiary ophthalmic clinic centers in Sydney, Australia (Eye Associates, Macquarie University Ophthalmology Clinic, and University of New South Wales Optometry Clinic) using images from several previous prospective clinical studies, including 3768 images (both eyes from 627 glaucoma patients and 403 healthy participants). All the studies were approved by the human ethics committees at the University of Sydney, Macquarie University, and University of New South Wales. Demographic data of the Sydney-based participants are shown in Table 1. All the research adhered to the tenets of the Declaration of Helsinki and all the participants provided informed consent. All cases of glaucoma were open-angle glaucoma (including primary open-angle glaucoma and

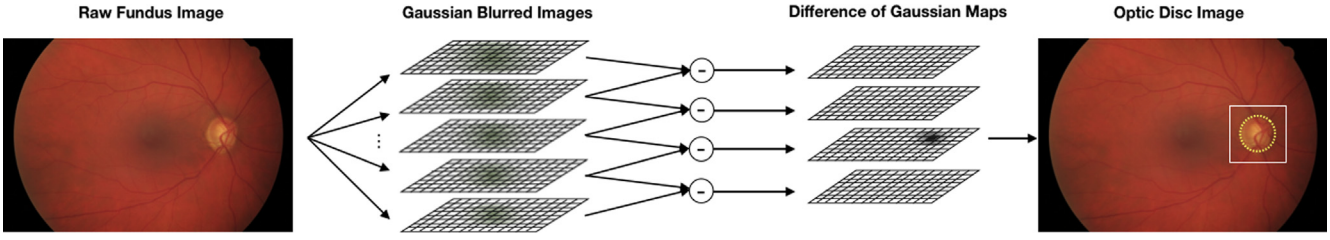

Figure 1. Difference-of-Gaussian-based optic disc detection. Multiple Gaussian filters with different kernel sizes were applied to the input image to construct the Gaussian scale-space of the image. The optic disc then can be identified by searching the local maxima in the scale-space.

normal-tension glaucoma), confirmed by Royal Australia and New Zealand College of Ophthalmologists-certified glaucoma subspecialists based on history, clinical examination, perimetry results, and OCT imaging. Images from 2 online databases also were incorporated in this study, including 596 images (40 glaucomatous eyes and 536 healthy eyes) from the RIM-ONE database[20] with definitive diagnoses and 30 images (15 glaucomatous eyes and 15 healthy eyes) from the online High-Resolution Fundus (HRF) database.[21] Because these photographs were sourced from different locations with several different clinicians confirming the diagnosis at each site, it was not possible to assign an exact definition of glaucoma across the entire dataset. For the Sydney-based clinical sites only, glaucoma cases confirmed by the clinician and manifesting either typical glaucomatous optic disc rim changes, visual field defects, or both were included. The online data sets include their own definitions and clinical confirmations.

Data Preparation

We merged all databases except the HRF database, and then partitioned the dataset into a training set (80% of all cases, to train the AI system) and a testing set (20% of all cases, to test the performance of the AI system). The training–testing partition was performed strictly on the patient level to ensure that the images of the same patient were assigned to the same dataset to avoid the system memorizing the images themselves instead of learning patterns in them. We used the 30 HRF images as an additional testing set. These images showed different characteristics in image resolution, magnification, angle, lighting, and color tone that were unseen previously by the system; therefore, these images were ideal to test the robustness and capability of the system.

All training and testing images were preprocessed using the same procedure. The difference-of-Gaussian blob detector[22] was used to locate and estimate the size of the optic disc in fundus images, as shown in Figure 1. The difference-of-Gaussian detector sequentially applied a set of n isotropic Gaussian filters with enlarging kernel sizes (r) to the input image and computed the pixel-wise differences of 2 consecutive Gaussian blurred images, resulting in $n - 1$ difference-of-Gaussian maps, which formed a 3-dimentional Gaussian scale-space of the image (2 dimensions of the image space plus 1 dimension of the scale variable). The maxima (x, y, r) in the scale space indicated that a blob-shaped structure (i.e., optic disc) was identified that matched both the center coordinates (x, y) and kernel size (r) of a Gaussian filter.

In this study, we used 16 Gaussian filters (kernel size ranging from 50 to 104 voxels) to detect optic discs with varying sizes and magnification factors. Then we expanded the detected region by cropping the optic disc and surrounding tissue using a square whose side was twice the radius of the detected blob. The cropped square images then were resized to the same image resolution (512×512) for further analysis.

For eyes with multiple images, all images were included in the analysis to allow the AI system to learn invariance to angular changes in focus and illumination levels, thus increasing the robustness of the system.[23] After preprocessing, a total of 4364 optic disc images were acquired from the merged dataset. Images that were blurry or with partial black background were removed from the training and testing sets, resulting in 4000 images (3200 in the training set and 800 in the testing set). No images were removed from the HRF dataset.

Deep Learning Model

We adopted the ResNet50 (online codes are available at: https://github.com/KaimingHe/deep-residual-networks) architecture to

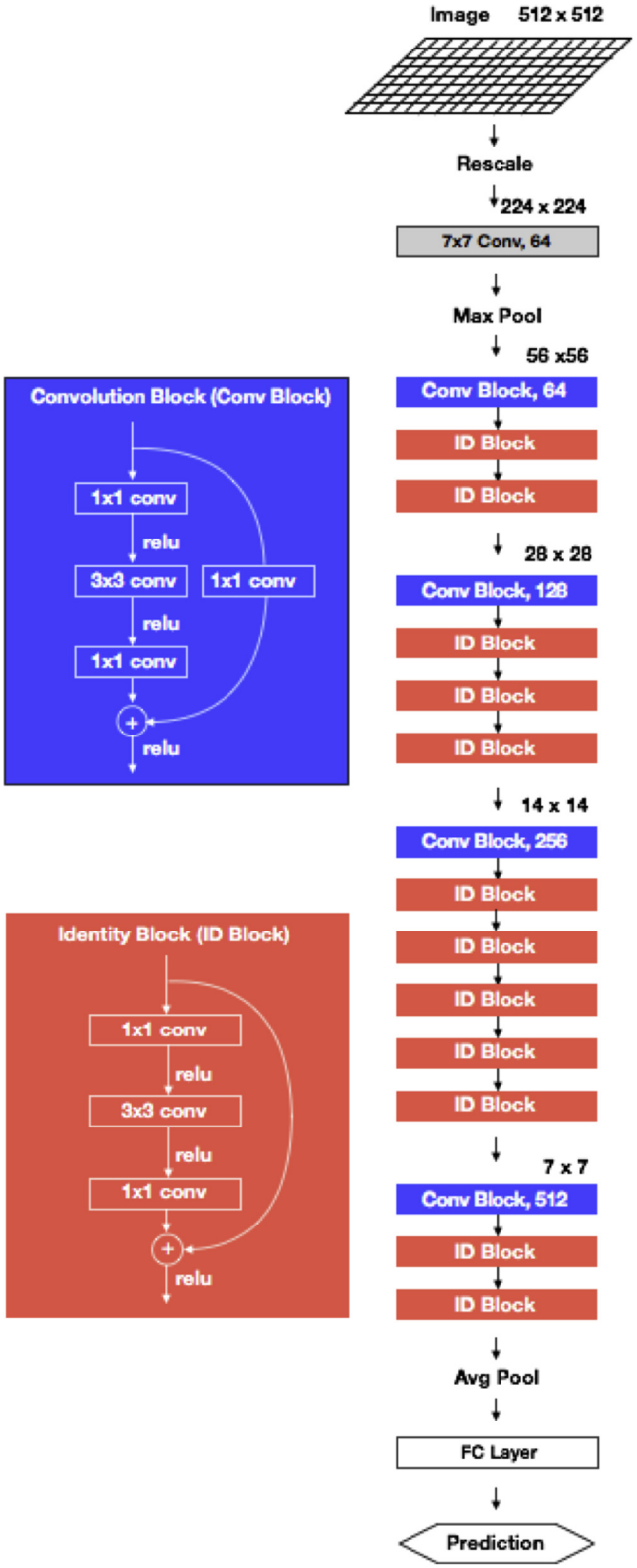

Figure 2. The ResNet50 architecture with 50 layers in the neural network. The hidden layers in the middle were grouped into convolutional blocks (Conv Block) and identify blocks (ID Block) to avoid showing the repeated layers and to visualize this network in a concise way. Avg = average; conv = convolutional; FC = fully connected; relu = rectified linear units.

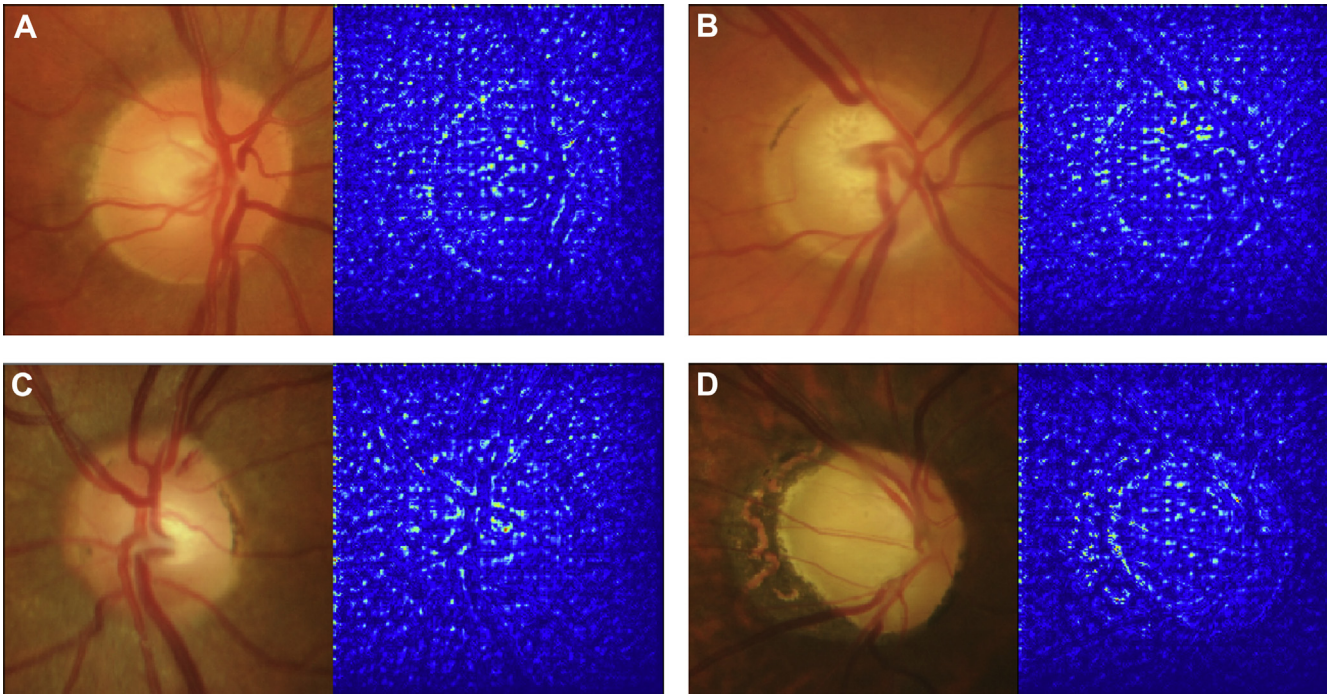

Figure 3. Representative input images and their corresponding saliency maps detected by the artificial intelligence system: (**A**) normal, (**B**) advanced glaucoma, (**C**) glaucoma with disc margin hemorrhage, and (**D**) glaucoma with peripapillary atrophic crescent.

build our AI system. Figure 2 shows the architecture of the ResNet50 model, with a convolutional layer on the top (64 channel; kernel size, 7×7; stride, 2), 48 convolutional neural network layers (grouped into 16 convolutional blocks and identification blocks) in the middle and a fully connected layer as the output layer at the bottom. A previous model based on the same architecture, which was trained on 1.28 million images from the ImageNet database, achieved 94.75% top 5 accuracy on 1000 object categories in the Large Scale Visual Recognition Challenge 2015.[10] We adopted this pretrained model and replaced its output layer (with 1000 nodes indicating the probability distribution of 1000 categories) with our customized output layer (with 2 nodes indicating the probability of glaucomatous and normal discs); therefore, the natural image features learned by the ImageNet pretrained model could be transferred seamlessly to our model. We then optimized this model for glaucomatous disc identification using our 3200 training images.

During training, categorical cross entropy was used as the loss function to assess the rate of misclassification. Stochastic gradient descendent method was used to optimize the parameters of the model with a uniform learning rate of 0.001, a decay factor of 1×10^{-6}, and momentum of 0.9. The rectified linear unit was used as the activation function to accelerate the training process. During inference, the system could output a probability, P, of being a glaucoma disc, per image, when a test image was given. We defined the prediction as glaucomatous ($\hat{y} = 1$) if P was more than a threshold probability, t; otherwise, the prediction was normal ($\hat{y} = 0$).

Performance Assessment

Sensitivity (ratio of correctly recognized glaucomatous discs to all glaucomatous discs in the testing set) and specificity (ratio of correctly predicted normal discs to total number of normal discs in the testing set) were used to evaluate the AI system and human experts. To compare the performance of the AI model with that of the glaucoma specialists and general ophthalmologists, we further created 16 testing subsets; each contained 80 images sampled randomly from the testing set (800 images). These datasets were sent to 11 glaucoma specialists and 7 general ophthalmologists. Among them, 2 glaucoma specialists and 1 general ophthalmologist received 3 datasets, and their repeated performances also were evaluated. The HRF dataset (30 images) also was used to test the robustness and generalization capability of the AI system, whose performance also was compared with that of 3 glaucoma specialists.

Results

Figure 3 shows the representative glaucomatous and normal disc photographs and the corresponding saliency maps generated by the algorithm during analysis. Bright areas with higher values reveal the pixels that most influence the prediction of the system. There is a correlation between the appearance of the discs (mainly on the disc's border and the blood vessels) and the pattern of the saliency maps.

We assessed the performance of the AI system against a panel of ophthalmologists using the testing datasets of 800 monoscopic disc photographs. Varying the threshold probability t in the interval [0,1] generated a curve of sensitivities and specificities of the model.

The sensitivity and specificity of glaucomatous optic disc identification of each test by the ophthalmologists are shown in Figure 4 and Table S1 (available at www.ophthalmologyglaucoma.org). Glaucoma subspecialists showed higher sensitivity than general

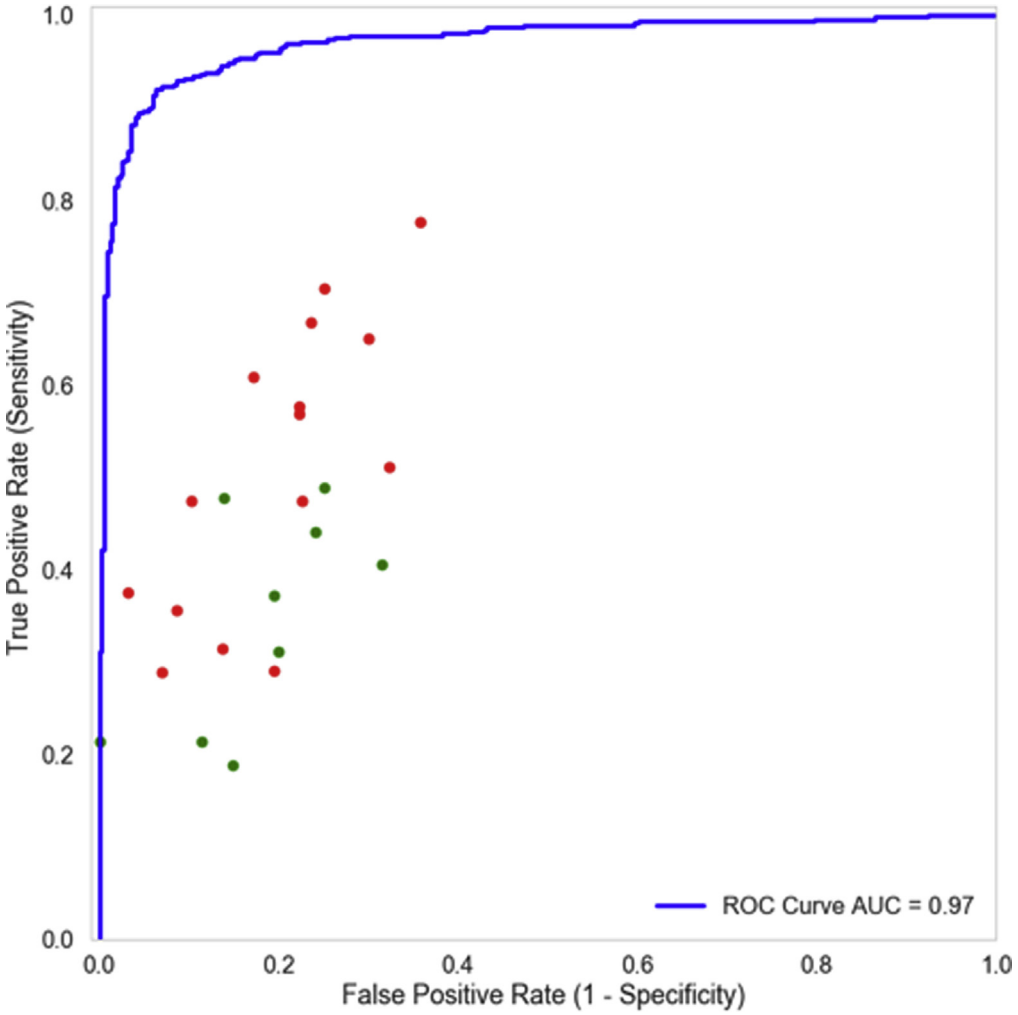

Figure 4. Graph showing a comparison of the performance between the artificial intelligence (AI) system (*blue curve*) and ophthalmologists (*green*, general ophthalmologists; *red*, glaucoma subspecialists) on random sets of 80 images obtained from 800 testing images. Each point represents a test performed by an ophthalmologist. AUC = area under the receiver operating characteristic curve; ROC = receiver operating characteristic.

ophthalmologists in glaucomatous disc identification. Twenty-four tests were completed by 11 glaucoma subspecialists and 7 were completed by general ophthalmologists (2 specialists and 1 ophthalmologist repeated 3 times on different testing sets), and the AI system demonstrated superior performance, achieving an accuracy rate of 92.7% (95% confidence interval [CI], 91.2%–94.2%), with 89.3% sensitivity (95% CI, 86.8%–91.7%) and 97.1% specificity (95% CI, 96.1%–98.1%). Based on all the testing images from the dataset, the AI system exhibited a similar accuracy rate (91.6%), achieving 87.9% sensitivity and 96.5% specificity with an area under the receiver operating characteristic curve of 0.97 (95% CI, 0.96–0.98).

We further tested the performance of the AI system based on an independent online database (HRF) containing 30 high-quality fundus images. The same images also were presented to 3 ophthalmologists for classification. Again, the AI system achieved a high accuracy rate with 86.7% in both sensitivity and specificity and a high area (0.89; 95% CI, 0.76–1.00). In comparison, the ophthalmologists' results in monoscopic photograph assessment were 75.6% in sensitivity and 77.8% in specificity (Fig 5).

Discussion

In this study, we described an algorithm developed by using the latest deep learning technology. By comparing its performance against that of a panel of ophthalmologists including glaucoma subspecialists, we found that, based on monoscopic disc photographs, the algorithm achieved superior sensitivity and specificity in glaucomatous disc identification. This level of accuracy was similar to that of ophthalmologists using stereoscopic photographs or ophthalmoscopy, as shown in previous studies.[6,7] Because it is far easier to obtain monoscopic disc photographs than it is to obtain high-quality stereoscopic photographs, this study demonstrated the potential of using this algorithm in glaucoma screening in a large population.

Traditional glaucoma optic nerve analysis by ophthalmologists is based on specific parameters or features of the optic nerve head, including disc size, cup-to-disc ratio, rim contour, pallor, disc hemorrhage, and peripapillary atrophy. Optic disc assessment using deep learning represents a break in

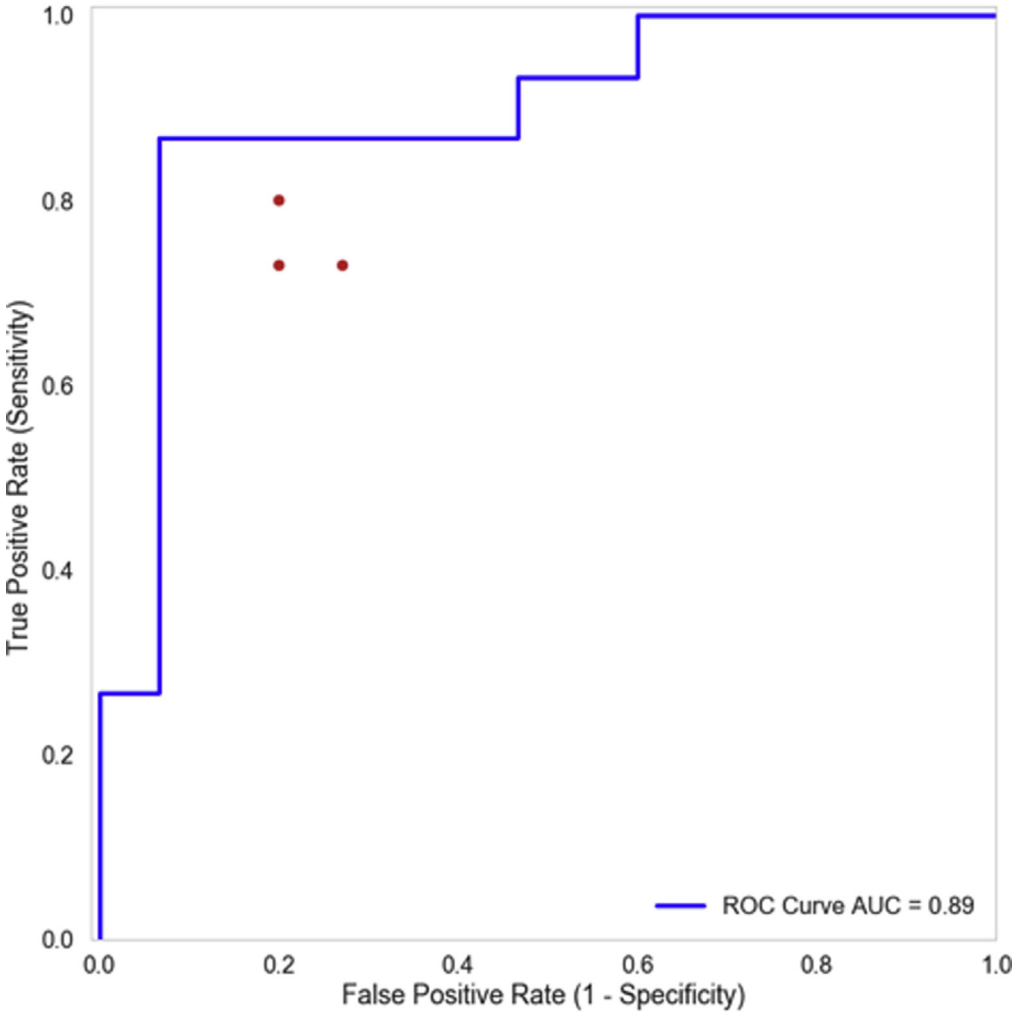

Figure 5. Graph showing a comparison of the performance of the artificial intelligence (AI) system (*blue curve*) and ophthalmologists (*red dots*) on 30 images from the publicly available High-Resolution Fundus online database. AUC = area under the receiver operating characteristic curve; ROC = receiver operating characteristic.

convention; training of the deep neural network does not rely on any human knowledge. Raw pixel-based disc images represent the input of training data, whereas on the terminal side of the training pipeline, the 2-way diagnosis, glaucoma or normal, was set as the output (Fig 2). Therefore, the AI system's judgment may be achieved based on completely different features of the optic discs that are as yet unknown. More detailed analysis of the training parameters, including the back propagation equations, may help to reveal those hidden components, and therefore potentially may help ophthalmologists to improve optic nerve assessment in the future.

In contrast to previous deep learning studies on DR, where the raw fundus images were used for algorithm development, we designed this unique 2-step automated pipeline—the AI system first automatically selected the optic disc region and then coregistered this peripapillary area onto a standard 512×512-pixel panel for further deep neural network evaluation. We found this can improve training efficacy significantly (data not shown). With approximately 4000 training images, we achieved a high accuracy rate in glaucomatous disc identification based on monoscopic photographs.

According to earlier studies, the accuracy of glaucomatous disc identification among ophthalmologists was approximately 80% using stereoscopic photographs.[6] In the current study, the accuracy rate using monoscopic photographs was lower at approximately 65% even for glaucoma subspecialists. This is likely because stereoscopic photographs tend to provide better interobserver and intraobserver reproducibility in optic disc analysis.[24] Although in one study the comparison of disc assessment performance between monoscopic and stereoscopic photographs by ophthalmologists did not reach statistical significance, both sensitivity and specificity were still slightly higher when using stereoscopic photographs.[25] In addition, in this study, to ensure the diversity of disc photographs, the testing cohort comprised 800 disc photographs with various image qualities and resolutions, and it also contained photographs of different stages of glaucoma and deliberately included (rather than excluded) a considerable number of anomalous optic discs. These potentially can increase the

overall difficulty of disc assessment for humans. Indeed, the doctors' performances were improved when tested using the 30 online HRF images (approximately 77% accuracy), although this was still not as good as the AI system (86% accuracy). In addition, because optic disc assessment is carried out in a subjective manner, some ophthalmologists tend to be more conservative than others in diagnosing glaucoma. This may lead to a more variable measure of sensitivity, whereas the specificity has been higher and consistent among all clinicians (>80%).

It is important to note that the AI system has achieved a general accuracy rate of more than 85% based on only monoscopic disc photographs, which is at least at the same level of ophthalmologists using stereoscopic photographs. This suggests that there may be potential for use of this algorithm in future glaucoma screening programs. Furthermore, this also raises another interesting question—if a new algorithm is trained based on stereoscopic images, will it reach an even higher level of accuracy rate? However, it is important to address that at this stage, the AI system still cannot replace ophthalmologists in disease diagnosis. It still will be the ophthalmologist's role to make the final diagnosis based on history; clinical examination with gonioscopy, stereo assessment of the optic disc, and OCT; and functional assessment with perimetry, pupil response, and color vision. Other comorbidities confounding the assessment will need to be excluded, including other potential causes of optic neuropathy. The AI systems usually are developed for specific tasks, such as fundus photograph assessment or perimetry analysis, whereas ophthalmologists are trained to diagnose a large spectrum of eye disorders ranging from those of the anterior segment to the optic nerve based on history, physical examination results, and clinical investigations. It remains very challenging for AI systems to be operative in such complex scenarios.

A limitation of this study is that the diagnosis associated with each of the photographs was based on different ophthalmologist's assessments, and the diagnostic criteria could not be standardized across all participants. It is still possible that some of those images labelled glaucoma were not, whereas some normal images actually may shown very early glaucoma. However, even if this were the case, both the AI system and the clinicians had not seen the new photograph set previously and were forced to make a call in the same scenario. The AI system always outperformed the clinicians regardless of the photographs used. It is also possible that patients were more likely to have undergone disc photography if their discs were borderline or atypical, which could influence the overall training of the AI system to predict that atypical discs were glaucomatous, whereas clinicians may be more conservative in their assessments. However, we see this as an overall advantage of the system: that it can interpret unusual discs more successfully from both normal and glaucoma participants.

In summary, this study demonstrated that the deep learning-based algorithm can detect glaucomatous optic discs at a high accuracy level by using only monoscopic fundus images, which are easier to obtain in a large population. This highlights its potential application in population-based disease screening or telemedicine programs.[26]

References

1. Tham YC, Li X, Wong TY, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*. 2014;121(11):2081−2090.
2. Chua J, Baskaran M, Ong PG, et al. Prevalence, risk factors, and visual features of undiagnosed glaucoma: the Singapore Epidemiology of Eye Diseases Study. *JAMA Ophthalmol*. 2015;133(8):938−946.
3. Mitchell P, Smith W, Attebo K, Healey PR. Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study. *Ophthalmology*. 1996;103(10):1661−1669.
4. Klein BE, Klein R, Sponsel WE, et al. Prevalence of glaucoma. The Beaver Dam Eye Study. *Ophthalmology*. 1992;99(10):1499−1504.
5. Wilson R, Leske C, Lee P, et al. Screening for open-angle glaucoma. In: Weinreb RN, Healey PR, Topouzis F, eds. *Glaucoma Screening—WGA Consensus Series*. Kugler; 2008:3−55.
6. Reus NJ, Lemij HG, Garway-Heath DF, et al. Clinical assessment of stereoscopic optic disc photographs for glaucoma: the European Optic Disc Assessment Trial. *Ophthalmology*. 2010;117(4):717−723.
7. Vessani RM, Moritz R, Batis L, et al. Comparison of quantitative imaging devices and subjective optic nerve head assessment by general ophthalmologists to differentiate normal from glaucomatous eyes. *J Glaucoma*. 2009;18:253−261.
8. Haddock LJ, Kim DY, Mukai S. Simple, inexpensive technique for high-quality smartphone fundus photography in human and animal eyes. *J Ophthalmol*. 2013;2013:518479.
9. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vis*. 2015;115(3):211−252.
10. He K, Zhang X, Ren S, Sun J. *Deep residual learning for image recognition. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, Nevada; 2016: 27−30. ISSN:1063-6919; https://doi.org/10.1109/CVPR.2016.90; IEEE.
11. He K, Gkioxari G, Dollar P, et al. Paper presented at: *The IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy; 2017:22−29. ISSN: 2380-7504; https://doi.org/10.1109/ICCV.2017.322; IEEE.
12. Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning. *Nature*. 2015;518(7540):529−533.
13. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115−118.
14. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402−2410.
15. Abramoff MD, Lou Y, Erginay A, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci*. 2016;57(13):5200−5206.
16. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*. 2017;124(7):962−969.
17. Takahashi H, Tampo H, Arai Y, et al. Applying artificial intelligence to disease staging: deep learning for improved staging of diabetic retinopathy. *PLoS One*. 2017;12(6):e0179790.
18. Asaoka R, Murata H, Iwase A, Araie M. Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. *Ophthalmology*. 2016;123(9):1974−1980.
19. Keltner JL, Johnson CA, Anderson DR, et al. The association between glaucomatous visual fields and optic nerve head

features in the Ocular Hypertension Treatment Study. *Ophthalmology*. 2006;113(9):1603−1612.
20. Fumero F, Alayon S, Sanchez JL, et al. RIM-ONE: an open retinal image database for optic nerve evaluation. Paper presented at: 24th International Symposium on Computer-Based Medical Systems (CBMS) 2011:1-6; Bristol, UK.
21. Budai A, Odstricilik J, Kollar R, et al. A public database for the evaluation of fundus image segmentation algorithms. Paper presented at: The Association of Research in Vision and Ophthalmology (ARVO) Annual Meeting, Fort Lauderdale, FL; 2011:1345.
22. Lowe DG. Distinctive image features from scale-invariant keypoints. *Int J Comput Vis*. 2004;60(2):91−110.
23. Dosovitskiy A, Fischer P, Springenberg JT, et al. Discriminative unsupervised feature learning with exemplar convolutional neural networks. *IEEE Trans Pattern Anal Mach Intell*. 2016;38(9):1734−1747.
24. Lehmann MV, Mardin CY, Martus P, Bergua A. 3D vs 2D qualitative and semiquantitative evaluation of the glaucomatous optic disc atrophy using computer-assisted stereophotography. *Eye (Lond)*. 2008;22(5):628−635.
25. Chan HH, Ong DN, Kong YX, et al. Glaucomatous optic neuropathy evaluation (GONE) project: the effect of monoscopic versus stereoscopic viewing conditions on optic nerve evaluation. *Am J Ophthalmol*. 2014;157(5):936−944.
26. Thomas SM, Jeyaraman MM, Hodge WG, et al. The effectiveness of teleglaucoma versus in-patient examination for glaucoma screening: a systematic review and meta-analysis. *PLoS One*. 2014;9(12):e113779.

Footnotes and Financial Disclosures

Originally received: November 30, 2017.
Final revision: February 14, 2018.
Accepted: March 12, 2018.
Available online: June 5, 2018. Manuscript no. 2017-2737.

[1] Save Sight Institute, Sydney Medical School, The University of Sydney, Sydney, Australia.

[2] Brain and Mind Centre, Sydney Medical School, The University of Sydney, Sydney, Australia.

[3] Department of Clinical Medicine, Faculty of Medicine and Health Sciences, Macquarie University, Sydney, Australia.

[4] Centre for Eye Health, and School of Optometry and Vision Science, The University of New South Wales, Kensington, Australia.

[5] School of Information Technologies, The University of Sydney, Sydney, Australia.

[6] School of Instrumentation Sciences and Opto-electronics Engineering, Beihang University, Beijing, China.

[7] Glaucoma Unit, Sydney Eye Hospital, Sydney, Australia.

[8] School of Materials Science and Engineering, The University of New South Wales, Kensington, Australia.

Financial Disclosure(s):
The author(s) have no proprietary or commercial interest in any materials discussed in this article.

Supported by the National Health and Medical Research Council (Y.Y.); and the Ophthalmic Research Institute of Australia (Y.Y.). The sponsor or funding organization had no role in the design or conduct of this research.

HUMAN SUBJECTS: Human subjects were included in this study. The human ethics committees at the University of Sydney, Macquarie University, and University of New South Wales approved the study. All research adhered to the tenets of the Declaration of Helsinki. All participants provided informed consent.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Liu, Graham, Klistorner, You

Analysis and interpretation: Liu, Kalloniatis, Zangerl, Cai, Gao, Chua, Arvind, Grigg, Chu, Klistorner, You

Data collection: Graham, Schulz, Kalloniatis, Zangerl, Chua, Arvind, Grigg, You

Obtained funding: None

Overall responsibility: Liu, Graham, Schulz, You

Abbreviations and Acronyms:
AI = artificial intelligence; **CI** = confidence interval; **DR** = diabetic retinopathy; **HRF** = High-Resolution Fundus.

Correspondence:
Yuyi You, MD, PhD, Save Sight Institute, Sydney Medical School, The University of Sydney, 8 Macquarie Street, Sydney 2000, Australia. E-mail: yuyi.you@gmail.com.