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;■:1–8 © 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

Published by Elsevier Inc.

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 normal-tension glaucoma), confirmed by Royal Australia and New Zealand College of Ophthalmologists-certified glaucoma

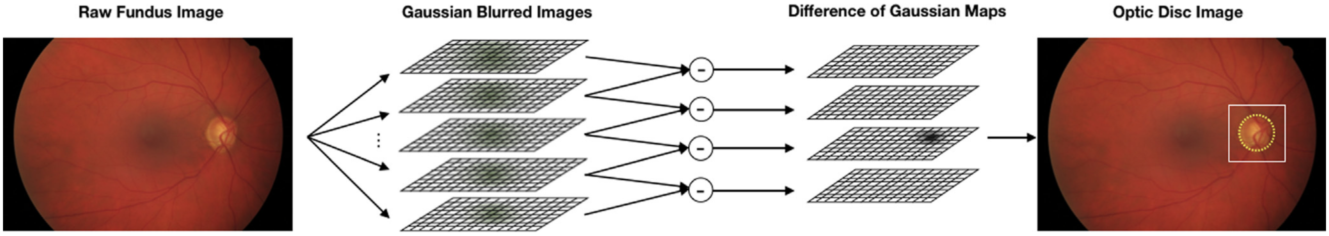

Figure 1. Difference-of-Gaussian-based optic disc detection. Multiple Gaussian filters with different kernel sizes were applied to the input image to construct the Gaussian scale-space of the image. The optic disc then can be identified by searching the local maxima in the scale-space.