Analyzing fundus images to detect diabetic retinopathy (DR) using deep learning system in the Yangtze River delta region of China

Li Lu[1,2#], Peifang Ren[1#], Qianyi Lu[3], Enliang Zhou[2], Wangshu Yu[1], Jiani Huang[1], Xiaoying He[1], Wei Han[1]

[1]Department of Ophthalmology, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China; [2]Department of Ophthalmology, The First Affiliated Hospital of University of Science and Technology of China, Hefei, China; [3]Department of Ophthalmology, The First Affiliated Hospital of Soochow University, Suzhou, China

Contributions: (I) Conception and design: L Lu, W Han, P Ren; (II) Administrative support: W Han; (III) Provision of study materials or patients: Q Lu, E Zhou; (IV) Collection and assembly of data: W Yu, J Huang; (V) Data analysis and interpretation: L Lu, X He; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Wei Han. Department of Ophthalmology, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310000, China. Email: hanweidr@hotmail.com.

Background: This study aimed to establish and evaluate an artificial intelligence-based deep learning system (DLS) for automatic detection of diabetic retinopathy. This could be important in developing an advanced tele-screening system for diabetic retinopathy.

Methods: A DLS with a convolutional neural network was developed to recognize fundus images of referable diabetic retinopathy. A total data set of 41,866 color fundus images were obtained from 17 cities in the Yangtze River Delta Urban Agglomeration (YRDUA). Five experienced retinal specialists and 15 ophthalmologists were recruited to verify images. For training, 80% of the data set was used, and the other 20% served as the validation data set. To effectively understand the learning process, the DLS automatically superimposed a heatmap on the original image. The regions utilized by the DLS were highlighted for diagnosis.

Results: Using the local validation data set, the DLS achieved an area under the curve of 0.9824. Based on the manual screening criteria, an operating point was set at about 0.9 sensitivity to evaluate the DLS. Specificity was recorded at 0.9609 and sensitivity was 0.9003. The DLSs showed excellent reliability, repeatability, and high efficiency. After analyzing the misclassification, it was found that 88.6% of the false-positives were mild non-proliferative diabetic retinopathy (NPDR) whereas, 81.6% of the false-negatives were intraretinal microvascular abnormalities.

Conclusions: The DLS efficiently detected fundus images from complex sources in the real world. Incorporating DLS technology in tele-screening will advance the current screening programs to offer a cost-effective and time-efficient solution for detecting diabetic retinopathy.

Keywords: Diabetic retinopathy; fundus image; deep learning system (DLS); artificial intelligence

Introduction

Diabetes mellitus (DM) and its associated complications pose a major global health threat. The latest edition of the International Diabetes Federation (IDF) diabetes atlas shows that 463 million adults aged 20–79 years had diabetes mellitus globally in 2019. This estimate is projected to rise to 578 million by 2030, and 700 million by 2045 (1). Pharmacologic therapy including metformin and insulin remains standard therapy for DM. Currently, for patients with atherosclerotic cardiovascular disease, chronic kidney

Page 2 of 10

Lu et al. Deep learning system for detecting diabetic retinopathy

disease, and heart failure, glucagon-like peptide-1 receptor agonist or sodium-glucose cotransporter-2 inhibitor are considered as the best choice for a second agent (2,3). Based on drug-specific effects and patient factors, personalized combination therapy is increasingly advocated. Diabetic retinopathy (DR), a serious complication that arises from DM, causing blindness and vision impairment in the working-age population across the globe (4,5), can be divided into two types: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). The global prevalence for DR ranges from 18% to 30% in type 2 diabetic patients, whereas, for PDR, the global prevalence ranges between 2.9% to 4.4% (6). Notably, China has the highest number of DM patients in the world with about 116.4 million cases (1). The prevalence rates are 18.45% for DR, 15.06% for NPDR, and 0.99% for PDR. In addition, DM patients from rural areas in China have been shown to have a higher risk of developing DR than those in urban areas (7,8).

The diagnosis of DR contains all features of the comprehensive adult medical eye evaluation, including history, eye clinical examination and a number of tests ancillary to the clinical examination (9). Particularly, the ancillary imaging tests can unveiling vital information not detectable to the clinical examination. Currently, the application of optical coherence tomography angiography (OCTA) has added a new perspective on our understanding of diabetic retinopathy by detecting preclinical microvascular changes, quantifying regions of macular nonperfusion and identify retinal neovascular tissue (8,10). However, fundus photography is the most widespread diagnosis and screening method by recording retinal images. Diabetic retinopathy is treatable at its early stages. Annual DR screening for diabetic patients is recommended by many guidelines (9,11). Governments and foundations have provided hospitals in China with screening services. However, a nationwide traditional screening system that relies on in-person dilated eye examination remains impractical. This is as a result of inadequate funds, access issues, and few trained eye care personnel. There is a need to devise new effective screening strategies to curb the rapidly increasing burden of diabetes.

Recent advances in telemedicine and machine learning (a branch of computer science that focuses on teaching machines to detect patterns in data) can provide solutions to these problems (12,13). Deep learning, a subclass of machine learning, mimics the way the human brain works and uses artificial neural networks to solve any feature expression problem. In medical practice, this technology has been used to automatically categorize massive medical images (14,15).

For DR screening, several deep learning systems (DLS) have been developed to grade images from multiple imaging techniques including fundus camera, optical coherence tomography (OCT), and OCTA (16-18). The DLS showed excellent performance similar to board-certified specialists. Therefore, integrating tele-screening with DLS provides a cost-effective solution. Retinal images for DM patients can be taken from the nearest primary care clinics without any trained ophthalmologist, this provides solutions to access problems. Through DLS, a few ophthalmologists can do scale screenings. Yangtze River Delta Urban Agglomeration (YRDUA) containing 26 cities, located in Yangtze River Delta Region of China is one of the highly populated and developed regions of China. It is also one of the six megalopolitan regions in the world (19).

This study aimed to create and train a DLS for referable DR detection using data set for about 41,866 retinal photographs obtained from departments of ophthalmology in hospitals from the 17/26 cities in YRDUA. We believe that the large volume and high complexity of the raw retinal fundus image data from real-world sources in a certain area can provide more characteristic original disease information and data complexity compared with the public databases, which ensures robust performance in future practical applications of our DLS. We present the following article in accordance with the STROBE reporting checklist (available at http://dx.doi.org/10.21037/atm-20-3275).

Methods

All the data used in this study were pseudonymized. The basic abstraction of our DLS and the structure of the artificial neural network are shown in *Figure 1*. Original fundus images from hospitals were pre-processed by cropping and resizing to obtain input images with a resolution of 224×224 pixels.

Data collection

A total data set of consecutive 41,866 color fundus images obtained from departments of ophthalmology in hospitals among 17/26 cities in YRDUA between January 1, 2018 and June 1, 2019 was created. From the total data set, 80% constituted the training data set (*Figure 2*). It was considered that consecutive images from different cameras in different
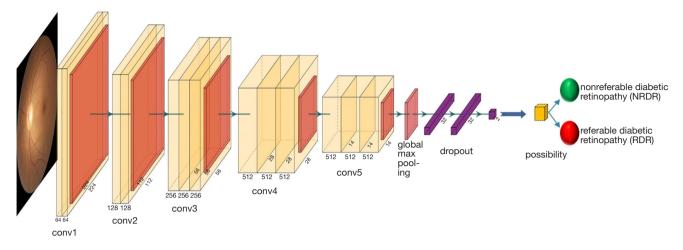
Figure 1 The basic convolutional neural network (CNN) architecture and workflow of our DLSs. Conv, convolution layers.

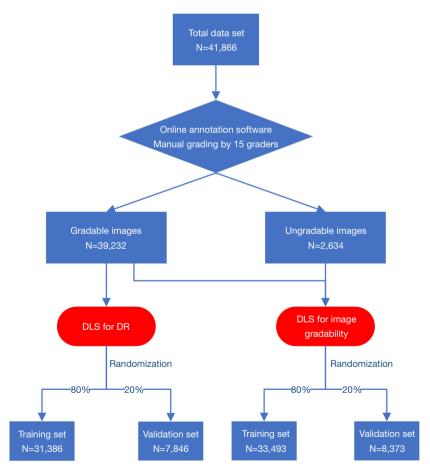**Figure 2** Workflow diagram showing the overview of developing deep learning systems to detect DR.

Page 4 of 10

Lu et al. Deep learning system for detecting diabetic retinopathy

cities were more valuable than those from public data sets. The quality of retinal images varied considerably and were perfect representatives for local patients. The DLS for this study thus was more suitable to local patients after training. Three different desktop retinal cameras and digital retinography systems (Canon, Topcon, and Heidelberg) were used in the 18 hospitals. Similar imaging protocols were applied for the 3 camera types. All images for the total data set were maculalutea-centered 45° color fundus photographs. Depending on the patient's condition, the doctor decides whether to dilate the pupils.

Definitions and the reference standard

According to the International Classification of Diabetic Retinopathy (ICDR), International Council of Ophthalmology (ICO) Guidelines for Diabetic Eye Care 2017, DR stages can be classified into 5 grades: no DR, mild NPDR, moderate NPDR, severe NPDR, and PDR (20). In this study, referable diabetic retinopathy (RDR) was defined as moderate NPDR, severe NPDR, and PDR (21), while none referable diabetic retinopathy (NRDR) was defined as fundus photographs of no DR(normal or other diseases) and mild NPDR. Many Chinese retinal specialists recommend that some moderate NDR patients and all patients with worse DR should receive pan-retinal photocoagulation (PRP). This has also been highlighted in the ICO Guidelines for Diabetic Eye Care as a significant criterion for screening RDR from DM patients

For manual grading, 15 licensed ophthalmologists and 5 experienced retinal specialists were recruited from the two eye centers. They were divided into 5 groups. Graders in the same group evaluated the same images. Each individual was blinded to the grading made by the other graders, and the results of the in-person dilated fundus exam. Then, grader would make an independent decision of the fundus photographs. Consistent results obtained from separate graders in one group were analyzed and used as the reference standard. Results that differed among same group of graders were cross-checked by an experienced retinal specialist for the final grading (22).

Besides, due to the complexity of our data sources, a DLS was trained to select quality images from the total data set for grading. All the graders assessed the quality and gradability of the images before they were classified as DR. The following criteria were used to determine a gradable image (21,23).

(I) The focus should be good enough for grading of smaller retinal lesions.

(II) Getting images with perfect exposure because dark and washed-out areas interfere with detailed grading.

(III) Image field definition: primary field must include the entire optic nerve head and macula.

(IV) Fewer artifacts: avoid dust spots, arc defects, and eyelash images.

(V) There should be no other errors in the fundus photograph, such as the absence of objects in the picture.

(VI) Images must be fundus photographs (For the few hospitals that did not equip anterior segment cameras, retinal cameras are used for anterior segment photography).

In general, for this research, we adopted a deep convolutional neural network for the two DLSs pre-trained on the ImageNet dataset named Visual Geometry Group 16 (VGG16) architecture. One DLS was used to classify the referable DR and the other DLS was used to assess the quality and gradability of images. All graders used online annotation software linked with the DLS.

Validation data set and statistical analyses

The remaining 20% of the total data set was used as the local validation data set and had the same data sources with the training data set (Figure 2). Retinal cameras, digital retinography systems, and associated protocols were consistent with the training data set. As with manual grading, DLSs performance was calculated based on sensitivity, specificity, and area under the receiver operating curve (AUC) (24). The receiver operating curves were plotted by varying the operating threshold (21). Based on the guidelines and criteria of Australia, UK, and Singapore (25-27), the results were evaluated at 0.900 sensitivity operating point. The false-positive and false-negative images of the validation data set were classified by 5 experienced retinal specialists (28). The Clopper-Pearson method was used to calculate the 95% CIs). To provide more detailed guidance for clinical analysis, a visualization heatmap highlighting strong prognostic regions of the fundus images was created using Rishab Gargeya's method (29). Stata version 14 (StataCorp) was used for all statistical analyses.

Ethical statement

The study was conducted in accordance with the

Table 1 Summarizing the training and local validation data set

	Training data set (%)	Validation data set (%)
Referable diabetic retinopathy		
Yes	4,206 (13.4)	1,044 (13.3)
No	27,180 (86.6)	6,802 (86.7)
Total	31,386	7,846
Gradability of images		
Gradable	31,555 (94.2)	7,677 (91.7)
Ungradable	1,938 (5.8)	696 (8.3)
Total	33,493	8,373

Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of First Affiliated Hospital, School of Medicine, Zhejiang University (Hangzhou, Zhejiang, China) (NO. 2019-1561) and individual consent for this retrospective analysis was waived.

Results

A total of 41,866 color fundus images obtained from departments of ophthalmology in hospitals from 17/26 cities in YRDUA between January 1, 2018 and June 1, 2019 were included in the training and validation data set. From that, 2,634 images were labeled as ungradable, and 39,232 images were used for DR severity grading. Each group graded between 7,508 and 9,204 (median 8,032) fundus photographs. About 10% of the graded photographs were submitted to experienced retinal specialists for final grading. After a simple random sampling, 31,386 images were assigned to the training data set and the remaining 7,846 images were used for validation. The proportion of referable diabetic retinopathy and gradable images are summarized in *Table 1*.

Performance and evaluation of the DLSs

Performance of the DLSs in validation data set was evaluated at an operating point close to 0.9 sensitivity. In the non-referable/referable diabetic retinopathy (NRDR/RDR) classification, the DLS achieved an AUC of 0.9824 (with 95% CI: 0.9733 to 0.9915), specificity of 0.9609 (with 95% CI: 0.9327 to 0.9796) and sensitivity of 0.9003 (with 95% CI: 0.8870 to 0.9125). Besides, for image gradability, AUC was recorded at 0.9945 (with 95% CI: 0.9918 to 0.9971), sensitivity, 0.9001 (with 95% CI: 0.8883 to 0.9110),

and specificity 0.9790 (with 95% CI: 0.9590 to 0.9909) (*Figure 3A,B*).

The DLSs also revealed excellent reliability, repeatability, and high efficiency. We selected 100 images from each study data set as the initial sample and transformed the images by random treatments (cutting less than 5% of the side length, 0–3-pixel random horizontal shift, turning left and right, rotating less than 15°) nine times. Thereafter, the DLSs were tested on the initial sample and the 9 treated samples. The outcomes for the two DLSs were consistent. Besides, it averagely took 8.7s seconds to select gradable images and 10.3 seconds to detect RDR.

Incorrect grading analysis

The analyses of false-negative and false-positive images were performed by experienced retinal specialists. The total number of false-negative classification was 38. The most common clinical feature was the undetected RDR with intraretinal microvascular abnormalities [n=31 (81.6%)]. Moreover, there were 4 RDR with retinal photocoagulation laser scars and 3 RDR characterized by massive retinal hemorrhage. Besides, among the 487 false-positive images, 431 (88.6%) mild NPDR images were characterized as RDR. The remaining images were other abnormalities in the fundus, for example, age-related macular degeneration, retinal vein occlusion, proliferative retinopathy, myopic maculopathy, and normal fundus photos with or without artifacts (*Table 2*).

Visualization heatmap analysis

Visualization analysis could present the learning procedure of our DLS and reveal the areas contributing most to

Page 6 of 10

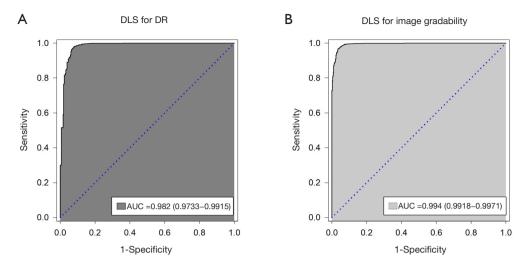Lu et al. Deep learning system for detecting diabetic retinopathy

Figure 3 Receiver operating characteristic (ROC) curves for our DLSs. (A) DLS for DR; (B) DLS for image gradability. AUC, area under the receiver operating curve.

Table 2 Analyses of false-negative and false-positive images in the local validation data set

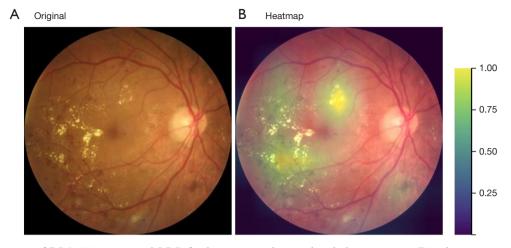| Feature	No	Proportion
False-negative		
IRMA	31	81.6%
PRP laser scar	4	10.5%
Massive retinal hemorrhage	3	7.9%
Total	38	100%
False-positive		
Mild NPDR	431	88.6%
AMD	10	2.1%
RVO	8	1.6%
Proliferative retinopathy	7	1.4%
Myopic maculopathy	7	1.4%
Normal fundus images	24	4.9%
Total	487	100.0%

IRMA, intraretinal microvascular abnormality; PRP, peripheral retinal photocoagulation; NPDR, non-proliferative diabetic retinopathy; AMD, age-related macular degeneration; RVO, retinal vein occlusions.

the DLS. At the end of the network, a convolutional visualization layer was implanted, and a visualization heatmap automatically generated. The original RDR fundus image has been displayed in *Figure 4A*. From the overlying fundus heatmap on *Figure 4A*, the regions that the DLS considered most significant in making its decision are highlighted in *Figure 4B*. Typical lesions were observed in such regions, for example, hard exudate, neovascularization, and retinal hemorrhage. The lesions were used by ophthalmologists to diagnose DR.

Discussion

Advanced computer science and the availability of big data

Figure 4 Visualization of DLS. (A) An original RDR fundus image with typical pathologic regions; (B) A heatmap generated from deep features overlaid on the original image, highlighting the valuable areas for prediction.

have improved artificial intelligence (AI) through the use of machine learning and deep learning techniques. The applications of these techniques in healthcare systems have improved disease screening and clinical diagnosis (30). Ophthalmologists require a variety of image data to help them in making the correct diagnosis of ocular diseases, particularly fundus diseases. The digital fundus photograph is the most basic and significant image used. Recent studies have shown that deep learning systems associated with fundus photographs are vital tools in identifying DR, glaucoma, retinopathy of prematurity (ROP), and age-related macular degeneration (AMD) (31-34).

China, one of the biggest developing countries in the world has made great progress in improving her health care systems. However, there is an increasing number of DM patients in the country. Patients from rural areas lack a basic understanding of DM and its complications. Besides, patients exhibiting DR symptoms, hardly seek any medical advice until when the disease has progressed enough to cause vision loss. Factors including limited financing resources and trained eye care personnel indicate that there is a need to develop a low-cost and effective screening method for early detection of the disease. In this study, a novel DLS designed to automatically recognize diabetic retinopathy in retinal fundus images achieved great success. All the original fundus images from desktop retinal cameras and digital retinography systems were obtained from hospitals of the Yangtze River Delta Urban Agglomeration. Thereafter, the training data set and the validation data set were constructed. In real-world screening conditions, the

rate of detecting ungradable images or poor-quality images has been reported at 20% (35,36). There was a demand to automatically assess the quality and gradability of retinal fundus images for DR screening (37). Hence, we developed another DLS to analyze the gradability of images captured by different examiners using different cameras to ensure each image in data sets was strictly a fundus image with the required quality and field definition. Other domestic and international studies have trained and validated the DLSs using high-quality photographs from public databases (23,29). Based on this, we created a real-world regional screening tool for local DM patients at a low cost. After training, the DLSs recorded high AUC, sensitivity and specificity performance in the local validation data set. The results showed high reliability and repeatability.

From the literature, the distribution of misclassification including false negatives and false positives was reported rarely. Analyzing such issues could optimize AI when managing medical image categorization tasks. Generally, the DLS from this study shows low rates of false-negative rate and false-positive. Most false-negative cases are caused by other complicated intraretinal microvascular abnormalities and signs, which suggest a more precise direction of optimization. Moreover, 88.6% of false-positive images are mild NPDR, which leads to unnecessary referrals, increase economic and psychological burden of patients and waste of resources. Future research should focus on upgrading the modified DLS from this study to solve the underlying drawbacks.

Although the designed DLS displayed a promising

Page 8 of 10

Lu et al. Deep learning system for detecting diabetic retinopathy

prospect, this study had limitations. First, the DLS has no function to detect diabetic macular edema (DME). Previous studies have, however, reported deep learning systems that identified RDR and referable diabetic macular edema (RDME) based on retinal images (24,28,38). According to the ICDR, DME is defined as any hard exudates within a one-disc diameter of the fovea or an area of hard exudates in the macular area that encompassed at least 50% of the disc area. OCT was considered the most sensitive method to identify DME and also provide a quantitative assessment of DME in determining DME severity (20). Unlike OCT, the definition of DME depending on fundus image is kind of out of date. Moreover, deep learning has been applied to analyze OCT images. For instance, Schlegl *et al.* developed a fully automated diagnostic method based on deep learning to detect and quantify macular fluid in conventional OCT images (39). Therefore, the ground truth about DME must include OCT imaging, and the DLS for DME recognition may apply a multi-modal method combined with fundus image and OCT image. We have been working on it. Second, since this study aimed at creating a real-world regional screening tool for local DM patients, the DLS only validated the local data set. Therefore, for its extensive application, massive external validation is needed. Third, the imaging protocols of our data sources required that examiners just take one-field photos for each patient. This is in contrast with standard seven-field stereoscopic images, one-field photos could decrease sensitivity to DR. Lastly the developed DLS cannot identify ocular diseases other than DR, this is not an automated comprehensive diagnostic platform to screen fundus diseases.

In future practical applications, we think there are some key issues worth mentioning. We suggest that the DLS should be integrated into every desktop retinal cameras and digital retinography systems at the screening sites, rather than acting as a terminal for processing data collected from the screening sites. Additionally, we need to monitor the DLS on a regular basis and iterate on it using the accumulated data to further improve the performance of the DLS.

In conclusion, this study demonstrates that the DLS created is capable of processing original images from different sources in the real world and achieves excellent outcomes in the local validation data set. This work provides a framework to further establish a regional telemedicine screening platform for detecting DR. This will greatly enlarge the scope of screening in a cost-effective and time-efficient way. Patients and ophthalmologists thus will significantly benefit from these advancements thereby reducing the rise in global DR cases.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at http://dx.doi.org/10.21037/atm-20-3275

Data Sharing Statement: Available at http://dx.doi.org/10.21037/atm-20-3275

Peer Review File: Available at http://dx.doi.org/10.21037/atm-20-3275

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (Available at http://dx.doi.org/10.21037/atm-20-3275). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Ethics Committee of First Affiliated Hospital, School of Medicine, Zhejiang University (Hangzhou, Zhejiang, China) (No. 2019-1561) and individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the

original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Williams R, Colagiuri S, Chan J, et al. IDF Atlas 9th Edition 2019.; 2019.
2. Wang W, Liu H, Xiao S, et al. Effects of Insulin Plus Glucagon-Like Peptide-1 Receptor Agonists (GLP-1RAs) in Treating Type 1 Diabetes Mellitus: A Systematic Review and Meta-Analysis. Diabetes Ther 2017;8:727-38.
3. Patoulias D, Imprialos K, Stavropoulos K, et al. SGLT-2 Inhibitors in Type 1 Diabetes Mellitus: A Comprehensive Review of the Literature. Curr Clin Pharmacol 2018;13:261-72.
4. Yau JWY, Rogers SL, Kawasaki R, et al. Global Prevalence and Major Risk Factors of Diabetic Retinopathy. Diabetes Care 2012;35:556-64.
5. Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. Lancet 2010;376:124-36.
6. Thomas RL, Dunstan FD, Luzio SD, et al. Prevalence of diabetic retinopathy within a national diabetic retinopathy screening service. Br J Ophthalmol 2015;99:64-8.
7. Song P, Yu J, Chan KY, et al. Prevalence, risk factors and burden of diabetic retinopathy in China: a systematic review and meta-analysis. J Glob Health 2018;8:010803.
8. Vujosevic S, Muraca A, Alkabes M, et al. Early microvascular and neural changes in patients with type 1 and type 2 diabetes mellitus without clinical signs of diabetic retinopathy. Retina 2019;39:435-45.
9. Flaxel CJ, Adelman RA, Bailey ST, et al. Diabetic Retinopathy Preferred Practice Pattern(R). Ophthalmology 2020;127:P66-145.
10. Russell JF, Shi Y, Hinkle JW, et al. Longitudinal Wide-Field Swept-Source OCT Angiography of Neovascularization in Proliferative Diabetic Retinopathy after Panretinal Photocoagulation. Ophthalmol Retina 2019;3:350-61.
11. Solomon SD, Chew E, Duh EJ, et al. Diabetic Retinopathy: A Position Statement by the American Diabetes Association. Diabetes Care 2017;40:412-8.
12. Liesenfeld B, Kohner E, Piehlmeier W, et al. A telemedical approach to the screening of diabetic retinopathy: digital fundus photography. Diabetes Care 2000;23:345-8.
13. Deo RC. Machine Learning in Medicine. Circulation 2015;132:1920-30.
14. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115-8.
15. Ehteshami Bejnordi B, Veta M, Johannes Van Diest P, et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. JAMA 2017;318:2199.
16. Guo Y, Hormel TT, Xiong H, et al. Development and validation of a deep learning algorithm for distinguishing the nonperfusion area from signal reduction artifacts on OCT angiography. Biomed Opt Express 2019;10:3257.
17. Kuwayama S, Ayatsuka Y, Yanagisono D, et al. Automated Detection of Macular Diseases by Optical Coherence Tomography and Artificial Intelligence Machine Learning of Optical Coherence Tomography Images. J Ophthalmol 2019;2019:6319581.
18. Tufail A, Rudisill C, Egan C, et al. Automated Diabetic Retinopathy Image Assessment Software. Ophthalmology 2017;124:343-51.
19. Xu M, He C, Liu Z, et al. How Did Urban Land Expand in China between 1992 and 2015? A Multi-Scale Landscape Analysis. PLoS One 2016;11:e0154839.
20. Muqit M. ICO Guidelines For Diabetic Eye Care 2017. Available online: http://www.icoph.org/enhancing_eyecare/diabetic_eyecare.html
21. 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;316:2402.
22. Verbraak FD, Abramoff MD, Bausch GC, et al. Diagnostic Accuracy of a Device for the Automated Detection of Diabetic Retinopathy in a Primary Care Setting. Diabetes Care 2019;42:651-6.
23. Yang WH, Zheng B, Wu MN, et al. An Evaluation System of Fundus Photograph-Based Intelligent Diagnostic Technology for Diabetic Retinopathy and Applicability for Research. Diabetes Ther 2019;10:1811-22.
24. Sahlsten J, Jaskari J, Kivinen J, et al. Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema Grading. Sci Rep 2019;9:10750.
25. Ting DSW, Cheung CY, Lim G, 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.
26. National Health Service (NHS) Diabetic Eye Screening Programme and Population Screening Programmes. Diabetic eye screening: commission and provide.; 2019.
27. Chakrabarti R, Harper C, Keeffe J. Diabetic retinopathy

Page 10 of 10

Lu et al. Deep learning system for detecting diabetic retinopathy

management guidelines. Exp Rev Ophthalmol 2014;7.

28. Li Z, Keel S, Liu C, et al. An Automated Grading System for Detection of Vision-Threatening Referable Diabetic Retinopathy on the Basis of Color Fundus Photographs. Diabetes Care 2018;41:2509-16.

29. Gargeya R, Leng T. Automated Identification of Diabetic Retinopathy Using Deep Learning. Ophthalmology 2017;124:962-9.

30. Ting DSW, Peng L, Varadarajan AV, et al. Deep learning in ophthalmology: The technical and clinical considerations. Prog Retin Eye Res 2019;72:100759.

31. Gulshan V, Rajan RP, Widner K, et al. Performance of a Deep-Learning Algorithm vs Manual Grading for Detecting Diabetic Retinopathy in India. JAMA Ophthalmol 2019;137:987.

32. Liu H, Li L, Wormstone IM, et al. Development and Validation of a Deep Learning System to Detect Glaucomatous Optic Neuropathy Using Fundus Photographs. JAMA Ophthalmol 2019;137:1353.

33. Burlina PM, Joshi N, Pacheco KD, et al. Assessment of Deep Generative Models for High-Resolution Synthetic Retinal Image Generation of Age-Related Macular Degeneration. JAMA Ophthalmol 2019;137:258.

34. Gupta K, Campbell JP, Taylor S, et al. A Quantitative Severity Scale for Retinopathy of Prematurity Using Deep Learning to Monitor Disease Regression After Treatment. JAMA Ophthalmol 2019;137:1029.

35. Scanlon PH, Malhotra R, Thomas G, et al. The effectiveness of screening for diabetic retinopathy by digital imaging photography and technician ophthalmoscopy. Diabetic Med 2003;20:467-74.

36. Scanlon PH, Foy C, Malhotra R, et al. The influence of age, duration of diabetes, cataract, and pupil size on image quality in digital photographic retinal screening. Am J Ophthalmol 2006;141:603.

37. Saha SK, Fernando B, Cuadros J, et al. Automated Quality Assessment of Colour Fundus Images for Diabetic Retinopathy Screening in Telemedicine. J Digit Imaging 2018;31:869-78.

38. Krause J, Gulshan V, Rahimy E, et al. Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy. Ophthalmology 2018;125:1264-72.

39. Schlegl T, Waldstein SM, Bogunovic H, et al. Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning. Ophthalmology 2018;125:549-58.