A deep learning system for identifying lattice degeneration and retinal breaks using ultra-widefield fundus images

Zhongwen Li[1#], Chong Guo[1#], Danyao Nie[2#], Duoru Lin[1], Yi Zhu[1,3], Chuan Chen[1,3], Li Zhang[1], Fabao Xu[1], Chenjin Jin[1], Xiayin Zhang[1], Hui Xiao[1], Kai Zhang[1,4], Lanqin Zhao[1], Shanshan Yu[1], Guoming Zhang[2], Jiantao Wang[2], Haotian Lin[1]

[1]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou 510060, China; [2]Shenzhen Ophthalmic Center, Jinan University, Shenzhen 518001, China; [3]Department of Molecular and Cellular Pharmacology, University of Miami Miller School of Medicine, Miami, Florida, USA; [4]School of Computer Science and Technology, Xidian University, Xi'an 710071, China

Contributions: (I) Conception and design: Z Li, C Guo, D Nie, J Wang, H Lin; (II) Administrative support: H Lin; (III) Provision of study materials or patients: J Wang, G Zhang, D Nie; (IV) Collection and assembly of data: Z Li, D Lin, L Zhang, F Xu, X Zhang, H Xiao, L Zhao, S Yu; (V) Data analysis and interpretation: Z Li, C Guo, D Nie, L Zhang, F Xu, H Lin, D Lin, Y Zhu, C Chen, C Jin, K Zhang, J Wang, G Zhang, H Xiao, L Zhao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Haotian Lin, Stata Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou 510060, China. Email: haot.lin@hotmail.com; Jiantao Wang. Shenzhen Ophthalmic Center, Jinan University, Shenzhen 518001, China. Email: wangjiantao65@126.com.

Background: Lattice degeneration and/or retinal breaks, defined as notable peripheral retinal lesions (NPRLs), are prone to evolving into rhegmatogenous retinal detachment which can cause severe visual loss. However, screening NPRLs is time-consuming and labor-intensive. Therefore, we aimed to develop and evaluate a deep learning (DL) system for automated identifying NPRLs based on ultra-widefield fundus (UWF) images.

Methods: A total of 5,606 UWF images from 2,566 participants were used to train and verify a DL system. All images were classified by 3 experienced ophthalmologists. The reference standard was determined when an agreement was achieved among all 3 ophthalmologists, or adjudicated by another retinal specialist if disagreements existed. An independent test set of 750 images was applied to verify the performance of 12 DL models trained using 4 different DL algorithms (InceptionResNetV2, InceptionV3, ResNet50, and VGG16) with 3 preprocessing techniques (original, augmented, and histogram-equalized images). Heatmaps were generated to visualize the process of the best DL system in the identification of NPRLs.

Results: In the test set, the best DL system for identifying NPRLs achieved an area under the curve (AUC) of 0.999 with a sensitivity and specificity of 98.7% and 99.2%, respectively. The best preprocessing method in each algorithm was the application of original image augmentation (average AUC =0.996). The best algorithm in each preprocessing method was InceptionResNetV2 (average AUC =0.996). In the test set, 150 of 154 true-positive cases (97.4%) displayed heatmap visualization in the NPRL regions.

Conclusions: A DL system has high accuracy in identifying NPRLs based on UWF images. This system may help to prevent the development of rhegmatogenous retinal detachment by early detection of NPRLs.

Keywords: Deep learning; fundus image; lattice degeneration; retinal breaks; ultra-widefield

Page 2 of 11

Li et al. Deep learning for identifying NPRLs

Introduction

Lattice degeneration and retinal breaks are clinically significant peripheral retinal lesions that predispose patients to the development of rhegmatogenous retinal detachment (RRD) (1,2). The prevalence of lattice degeneration is about 8% in the general population, and 16.9% in myopic patients (2,3). Approximately 18.7–29.7% of RRD is associated with lattice degeneration (1,4). Retinal breaks are present in about 6% of eyes in both clinical and autopsy studies (5,6). Remarkably, at least 50% of untreated retinal breaks with persistent vitreoretinal traction will lead to RRD (7,8).

RRD is an important cause of visual disability and visual loss (9). Despite the advent of sophisticated techniques and treatment advances, the prognosis remains poor, with 42% of patients achieving 20/40 vision, and only 28% if the macula is involved (10). Therefore, it is imperative to assess the conditions of the peripheral retina as a routine ophthalmologic examination, especially for myopic patients (8). Furthermore, notable peripheral retinal lesions (NPRLs), including lattice degeneration and/or retinal breaks, should be monitored regularly by retinal specialists, and prophylactic laser photocoagulation should be considered at an appropriate time to prevent RRD (8,11).

Screening NPRLs in the peripheral retina requires an experienced ophthalmologist to perform a dilated fundus examination, which is time-consuming, labor-intensive, and substantially impacts the deployment of screening on a large scale. In addition, a large amount of research has shown that automated image interpretation using deep learning (DL) algorithms can efficiently and accurately identify conditions such as diabetic retinopathy (DR) (12,13), age-related macular degeneration (AMD) (14,15), and glaucoma (16). However, most previous studies used traditional fundus images with limited visible scope (30° to 60°), which provide little information on the peripheral retina.

The emergence of the ultra-widefield fundus (UWF) imaging system, covering 200° panoramic images of the retina, compensates for the deficiency of traditional fundus cameras (17). In particular, the peripheral retina can be observed through UWF imaging with a single capture without requiring a dark setting, contact lens, or pupillary dilation (17). Employment of UWF images in conjunction with deep learning algorithms may provide accurate identification of NPRLs with significant benefits encompassing increased accessibility and affordability for high-risk populations. Subsequently, this technology could decrease the incidence of RRD through an appropriate precautionary intervention. In this study, we aimed to develop a DL algorithm to detect NPRLs from UWF images and verify its performance in an independent dataset.

Methods

Image collection

The initial UWF images were obtained from individuals undergoing routine ophthalmic health evaluation between November 2016 and January 2019 at Zhongshan Ophthalmic Center and Shenzhen Ophthalmic Hospital using an OPTOS nonmydriatic camera and 200° fields of view. Patients underwent this examination without mydriasis. All images were de-identified prior to transfer to study investigators. The study was approved by the Ethics Committee of Zhongshan Ophthalmic Center and followed the tenets of the Declaration of Helsinki.

Classification and reference standard

The features of NPRLs (lattice degeneration and/or retinal breaks) were determined according to the Preferred Practice Pattern® Guidelines from the American Academy of Ophthalmology Retina/Vitreous Panel (8). The images were classified into two categories, NPRLs and non-NPRLs, according the criteria shown in the *Table 1*. The image quality was included in the classification as follows: excellent quality referred to images without any problems; adequate quality referred to images with deficiencies in focus, illumination or other artifacts, but part of NPRLs could still be identified; and poor quality referred to images that were insufficient for any interpretation (over one-third of the image was obscured). Images of poor quality were excluded from the study.

Training a DL system requires a robust reference standard (18,19). For this purpose, all anonymous images were independently classified by 3 ophthalmologists who had over 5 years of experience in retina specialty. The final classification was determined when agreement was achieved among the 3 ophthalmologists. Any level of disagreement was adjudicated by another retina specialist with 20 years of experience in retinal examinations. The process of image classification is described in *Figure 1*.

Imaging preprocessing and DL system development

To obtain the best model, four state-of-the-art convolutional neural network (CNN) architectures,

Table 1 Reference standards of notable peripheral retinal lesions

Classification	Presence of clinical features
I NPRLs	Inclusion of lattice degeneration and/or retinal breaks
Lattice degeneration	Sharply demarcated, circumferentially oriented, ovoid or linear patches of the atrophic retina, with/without pigmentation, crisscrossing fine white lines, glistening white (frost-like) areas or punched out areas of extreme retinal thinning
Retinal breaks	The shape of a full-thickness defect in retina including atrophic hole, operculated hole, horseshoe tear and retinal dialysis
Lattice degeneration with retinal breaks	Including both of the above manifestations
II Non-NPRLs	None of the above manifestations
Excellent quality	No problems with image
Adequate quality	Some NPRLs could be identified despite image deficiency
Poor quality	Insufficient for any interpretation (over one-third of the image was obscured)

NPRLs, notable peripheral retinal lesions, defined as the presence of lattice degeneration and/or retinal breaks.
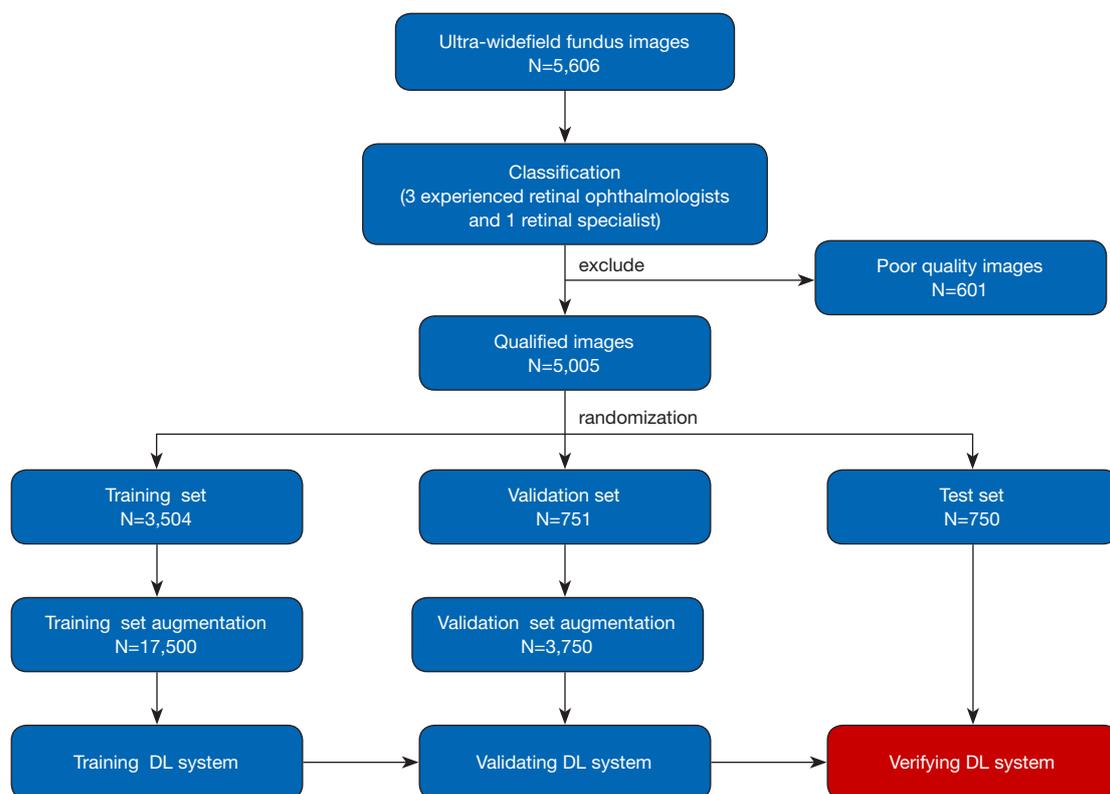

Figure 1 Workflow diagram showing the overview of developing a deep learning system to identify notable peripheral retinal lesions (lattice degeneration and/or retinal breaks).

including InceptionResNetV2, InceptionV3, ResNet50, and VGG16, were investigated in this study. Their architectural characteristics are summarized in *Table 2* (20). Weights pretrained for ImageNet classification were used to initialize the CNN architectures (21).

To determine a suitable preprocessing technique that

Page 4 of 11

Li et al. Deep learning for identifying NPRLs

Table 2 Architectural characteristics of convolutional neural networks

Item	Size (MB)	Parameters	Depth*
InceptionResNetV2	215	55,873,736	572
InceptionV3	92	23,851,784	159
ResNet50	98	25,636,712	50
VGG16	528	138,357,544	23

*, depth is the topological depth of the network, including activation layers and batch normalization layers. MB, Mbyte.

Table 3 Proportions of each type of notable peripheral retinal lesions in the training, validation and test datasets

Item	Training set, No. (%)	Validation set, No. (%)	Test set, No. (%)
Lattice degeneration	437 (12.5)	92 (12.3)	100 (13.3)
Retinal breaks	118 (3.4)	23 (3.1)	24 (3.2)
Lattice generation with retinal breaks	150 (4.3)	28 (3.7)	32 (4.3)
NPRLs	705 (20.1)	143 (19.1)	156 (20.8)
Non-NPRLs	2,799 (79.9)	608 (80.9)	594 (79.2)
Total (origin)	3,504 (100.0)	751 (100.0)	750 (100.0)
Total (augmentation)*	17,500	3,750	NA

*, augmentation is approximately 5 times the original size. NPRLs, notable peripheral retinal lesions, defined as the presence of lattice degeneration and/or retinal breaks; NA, not applicable.

can improve the performance of the DL algorithm, we investigated the three following methods:

(I)   Original images without any augmentation.

(II) Original images were augmented using brightness shift with a factor ranging from 0.8 to 1.6, rotation up to 45° and horizontal and vertical flipping on both the training set and validation set to approximately 5 times the original size.

(III) Histogram equalizations were applied to all images, including the test set, to balance the brightness of the image. Horizontal flipping, vertical flipping and rotation up to 45° were also applied to the training set and validation set to increase their size by five times.

All image pixel values were normalized to range between 0 to 1 and the images were resized to 512 by 512. The adaptive moment estimation (ADAM) optimizer with an initial learning rate of 0.001, beta 1 of 0.9, beta 2 of 0.999, fuzz factor of 1e-7 and zero learning rate decay were applied. For the VGG16 models, the same optimizer with an initial learning rate of 0.0001 was used. Each model was trained up to 180 epochs. During the training process, the validation loss was evaluated using the validation set after each epoch, and was used as a reference for model selection. Early stopping was applied, and if the validation loss did not improve for 60 consecutive epochs, the training process was stopped. The model state where the validation loss was the lowest was saved as the final state of the model.

All eligible images were randomly divided into 3 sets, with 70% (3,504 images) as a training set, 15% (751 images) as a validation set and 15% (750 images) as a test set (with no participants overlap among these sets). The images in the training and validation sets were used to establish and determine the models, respectively. Then, the selected model was verified through images in the test set. The number of images was augmented to 17,500 in the training set and 3,750 in the validation set to improve DL efficiency. *Table 3* provides further information on each set.

Features of misclassification and heatmaps of positive images

Reasons for errors made by the optimal DL system were analyzed by checking all the misclassified images. Heatmaps were generated using the Gradient weighted Class Activation Mapping (Grad-CAM) algorithm for all

true-positive images and all false-positive images in the test set. Grad-CAM calculates the gradient of the output of the penultimate convolutional layer (the layer before the fully connected layers and is usually the last convolutional layer) with respect to each pixel in the input image. Image pixel with higher impact on the model's prediction has heatmap color closer to the red spectrum in the Jet color map, while those with less impact has color closer to the blue spectrum.

General ophthalmologist comparisons

To evaluate our DL system in the context of screening NPRLs, we recruited 2 general ophthalmologists who had 3 and 5 years of experience respectively in UWF image analysis from a physical examination center, and then compared the performance of the system with that of general ophthalmologists in detecting NPRLs in the test set.

Statistical analyses

The performance of the DL system and general ophthalmologist were evaluated using three critical outcome measures: accuracy, sensitivity and specificity. Additionally, a receiver operating characteristic (ROC) curve was used to evaluate the efficiency of the DL system. The area under the curve (AUC) of ROC with 95% confidence intervals was also applied to assess the DL system. In the test set, unweighted Cohen's kappa coefficients were employed to compare the results of the best DL model and the general ophthalmologists to the reference standard respectively. All statistical analyses were conducted using python 2.7.15.

Results

A total of 5,606 UWF images from 2,566 participants aged 15–76 years (mean age 44.7 years, 45.8% female) were labeled for NPRLs. After filtering out 601 poor quality images due to cataract or artifacts (e.g., arc defects, dust spots and serious eyelash images), 5,005 images were used to develop the DL system, among which 1,004 images had been classified by ophthalmologists as NPRLs and 4,001 as non-NPRLs (e.g., retinal hemorrhage, exudation and epiretinal membrane).

Four algorithms (InceptionResNetV2, InceptionV3, ResNet50, and VGG16) were used to train models with the aforementioned 3 preprocessing methods. Thus, a total of 12 models were trained. Their performance is presented in *Figure 2*, which indicates that the best algorithm for each

preprocessing method was InceptionResNetV2 (average AUC =0.996), and the best preprocessing approach in each algorithm was applying augmentation of original images in training and validation sets (average AUC =0.996). Based on the optimal preprocessing method, InceptionV3 achieved an AUC of 0.997 (95% CI, 0.994–0.999), ResNet50 achieved an AUC of 0.989 (95% CI, 0.978–0.997), InceptionResNetV2 achieved an AUC of 0.999 (95% CI, 0.997–1.00), and VGG16 achieved 0.998 (95% CI, 0.995–1.00) in detecting NPRLs (*Figure 2*), and the accuracies were 98.7% (740/750), 97.5% (731/750), 99.1% (743/750) and 98.5% (739/750), respectively. *Table 4* presents further information describing the performance of these 4 DL algorithms.

The performance of the best DL model and general ophthalmologists in detecting NPRLs is shown in *Table 5*. The general ophthalmologist with 5 years of experience had a 93.6% sensitivity and a 98.7% specificity, and the general ophthalmologist with 3 years of experience had an 85.9% sensitivity and a 96.8% specificity, while the best model had a 98.7% sensitivity and a 99.2% specificity. Compared with the reference standard, the unweighted Cohen's kappa coefficients were 0.927 (95% CI, 0.893–0.960), 0.833 (95% CI, 0.783–0.883) and 0.972 (95% CI, 0.951–0.993) for the 5-year experience general ophthalmologist, the 3-year experience general ophthalmologist and the DL model, respectively.

To analyze errors made by the optimal DL system, we checked all misclassified images. Among these images, 2 (29%) were NPRL images misclassified as non-NPRLs (*Figure 3*), and 5 (71%) were non-NPRL images misclassified as NPRLs (*Figure 4*). In addition, of all 154 true-positive NPRL images in the test set, 150 (97.4%) displayed heatmap visualization of NPRLs (*Figure 5*). Of the 5 false-positive NPRL images, all showed redder regions in areas that were partly similar to lattice degeneration. Among these images, 2 (40%) showed heatmap visualization of a retinal pigmented nevus, 1 (20%) showed highlighted regions with peripheral retinal hyperpigmentation, and the remaining 2 (40%) presented heatmap regions with proliferative vitreous membrane (*Figure 4*).

Discussion

With the utilization of color fundus photographs, DL systems have achieved unprecedented success in detecting many retinal diseases (22-24). However, most of these studies focused on lesions that appear in the posterior pole
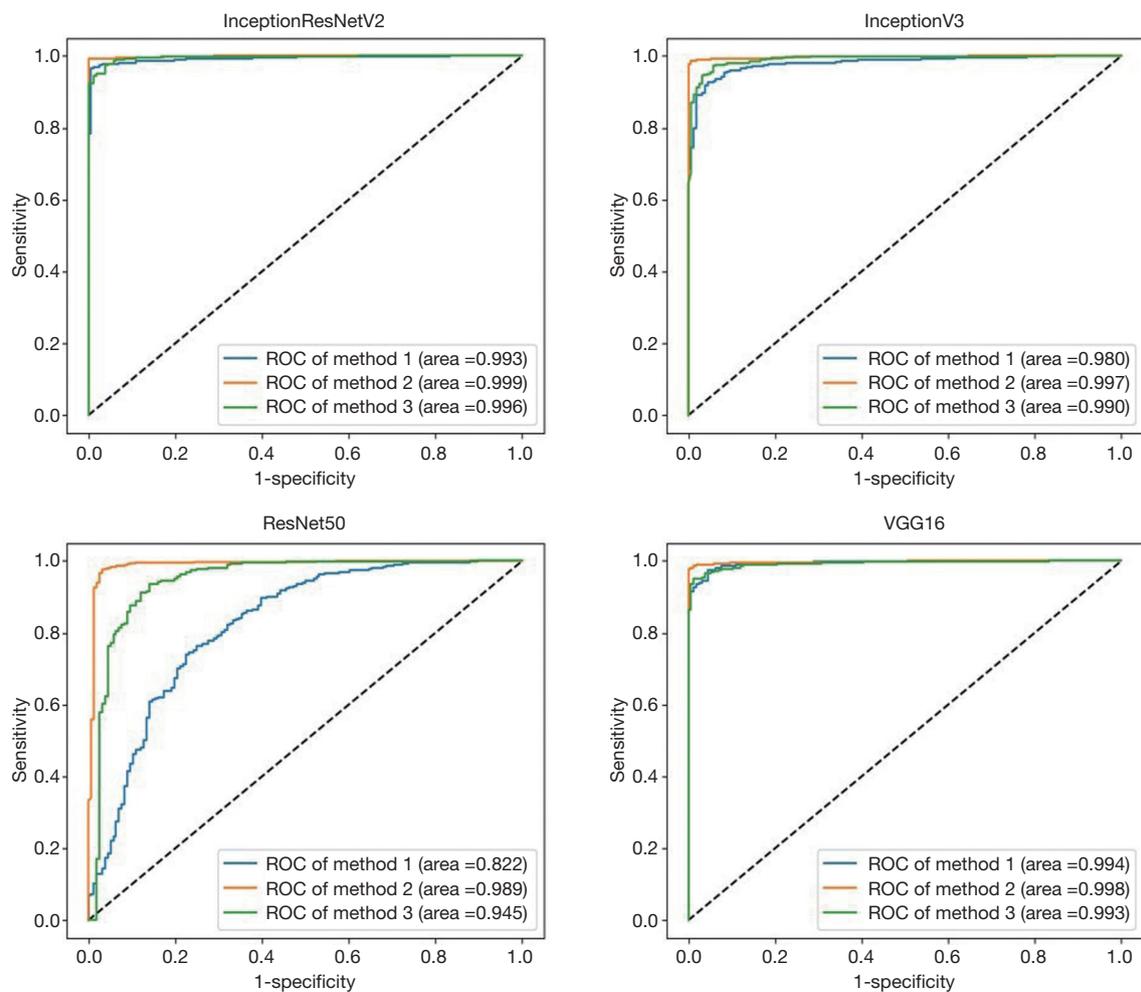
Page 6 of 11

Li et al. Deep learning for identifying NPRLs

Figure 2 The test set performance for the detection of notable peripheral retinal lesions. Method 1 uses original images, method 2 uses augmented images, method 3 uses augmented histogram-equalized images.

area of the retina. In this study, by focusing on peripheral retina, we successfully established DL systems based on UWF images with high accuracy in detecting NPRLs. The AUC of the best DL system achieved 0.999 with 98.7% sensitivity and 99.2% specificity. Based on these results, the DL system exhibited a remarkable performance in discerning NPRLs. Moreover, the performance of the DL system is better than both 5- and 3-year experience general ophthalmologists (*Table 5*). The agreement between the system and the reference standard is higher than that of the general ophthalmologists according to the unweighted Cohen's kappa coefficients. It further validates the effectiveness of our DL system and indicates that the system could be used as a potential screening tool. To the best of our knowledge, this study was the first to use DL to detect

NPRLs based on a large number of UWF images.

In our study, to obtain the most accurate DL system, 12 models established by 4 different algorithms were assessed using 3 types of preprocessed UWF images. According to the results shown in *Figure 2*, the best preprocessing method is applying brightness, rotation, mirror flipping augmentation to approximately 5 times the original size in training and validation sets. Hwang *et al.* (25) also showed that the application of augmented images in the training and validation sets can enhance the performance of DL models in the detection of AMD. A possible explanation is that augmentation turns each image into several images of various conditions, therefore the sample size is increased, which enables the generalization of the DL system to unseen data. In addition, the accuracy of the DL system built using

Table 4 Performance of four DL algorithms trained by three preprocessing methods in the test set

Item	NPRLs		
	Sensitivity (95% CI), %	Specificity (95% CI), %	Accuracy (95% CI), %
Method 1			
Inception V3	90.4 (85.5–95.3)	96.0 (94.4–97.6)	94.8 (93.2–96.4)
ResNet50	37.2 (24.8–49.6)	97.5 (96.2–98.8)	84.9 (82.1–87.7)
InceptionResNetV2	98.1 (95.9–100)	97.0 (95.6–98.4)	97.2 (96.0–98.4)
VGG16	95.5 (92.2–98.8)	94.8 (93.0–96.6)	94.9 (93.3–96.5)
Method 2			
Inception V3	98.7 (96.9–100)	98.7 (97.8–99.6)	98.7 (97.9–99.5)
ResNet50	96.8 (94.0–99.6)	97.6 (96.4–98.8)	97.5 (96.4–98.6)
InceptionResNetV2	98.7 (96.9–100)	99.2 (98.5–99.9)	99.1 (98.4–99.8)
VGG16	97.4 (94.9–99.9)	98.8 (97.9–99.7)	98.5 (97.6–99.4)
Method 3			
Inception V3	84.0 (77.7–90.3)	98.5 (97.5–99.5)	95.5 (94.0–97.0)
ResNet50	71.8 (63.5–80.1)	98.1 (97.0–99.2)	92.7 (90.8–94.6)
InceptionResNetV2	93.6 (89.6–97.6)	98.8 (97.9–99.7)	97.7 (96.6–98.8)
VGG16	92.9 (88.7–97.1)	97.3 (96.0–98.6)	96.4 (95.0–97.8)

Method 1, training based on original images; Method 2, training based on augmented original images; Method 3, training based on augmented histogram-equalized images. DL, deep learning system; NPRLs, notable peripheral retinal lesions, defined as the presence of lattice degeneration and/or retinal breaks; CI, confidence interval.

Table 5 The performance of the deep learning system *vs.* general ophthalmologists in detecting notable peripheral retinal lesions

Item	Sensitivity (95% CI), %	Specificity (95% CI), %	Accuracy (95% CI), %
Ophthalmologist A	93.6 (89.6–97.6)	98.7 (97.8–99.6)	97.6 (96.5–98.7)
Ophthalmologist B	85.9 (80.0–91.8)	96.8 (95.4–98.2)	94.5 (92.8–96.2)
Deep learning system	98.7 (96.9–100)	99.2 (98.5–99.9)	99.1 (98.4–99.8)

CI, confidence intervals; A, the general ophthalmologist with 5-year working experience in physical examination center; B, the general ophthalmologist with 3-year working experience in physical examination center.

augmented histogram-equalized images is slightly lower than that of the systems built using augmented original images. Although histogram equalization can equalize the brightness of images and increase clarity, some information may be altered during this process. Consequently, the performance of the DL systems using histogram-equalized images is not as good as the performance of systems based on augmented original images.

InceptionResNetV2 is the best DL algorithms in detecting NPRLs when compared to other ones. Among all the models, InceptionResNetV2 has the most layers

(*Table 2*). Therefore, it can represent a more complex relationship between the input (UWF image) and output (the label we attempt to predict). Normally, a larger network is more prone to overfitting. Nevertheless, InceptionResNetV2 reduces this tendency by mimicking the skip connections from Residual Network (ResNet). We also speculate that InceptionResNetV2 performs well on the task because NPRLs manifest in a wide variety of forms and patterns, which could appear similar to other lesions at times, and as a result, require a larger network to capture the complexity.
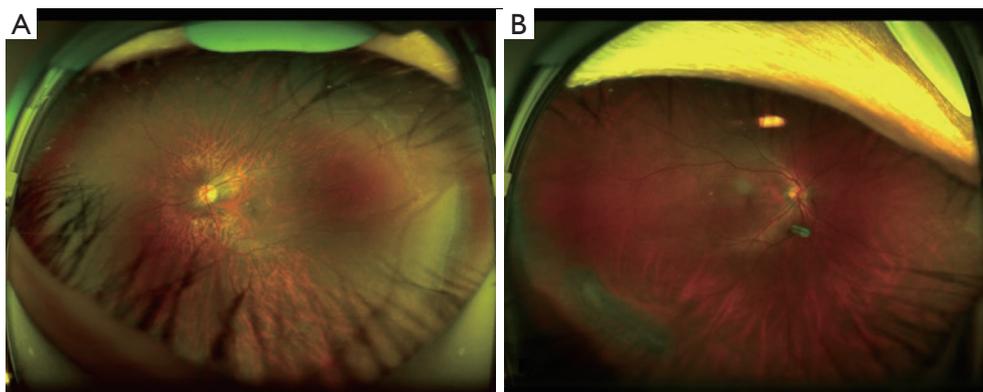
Page 8 of 11

Li et al. Deep learning for identifying NPRLs

Figure 3 Ultra-widefield fundus images showing false-negative cases. (A) Lattice degeneration on the bottom right, partly coved by eyelashes; (B) lattice degeneration with small atrophic holes on the bottom left, partly coved by eyelashes.
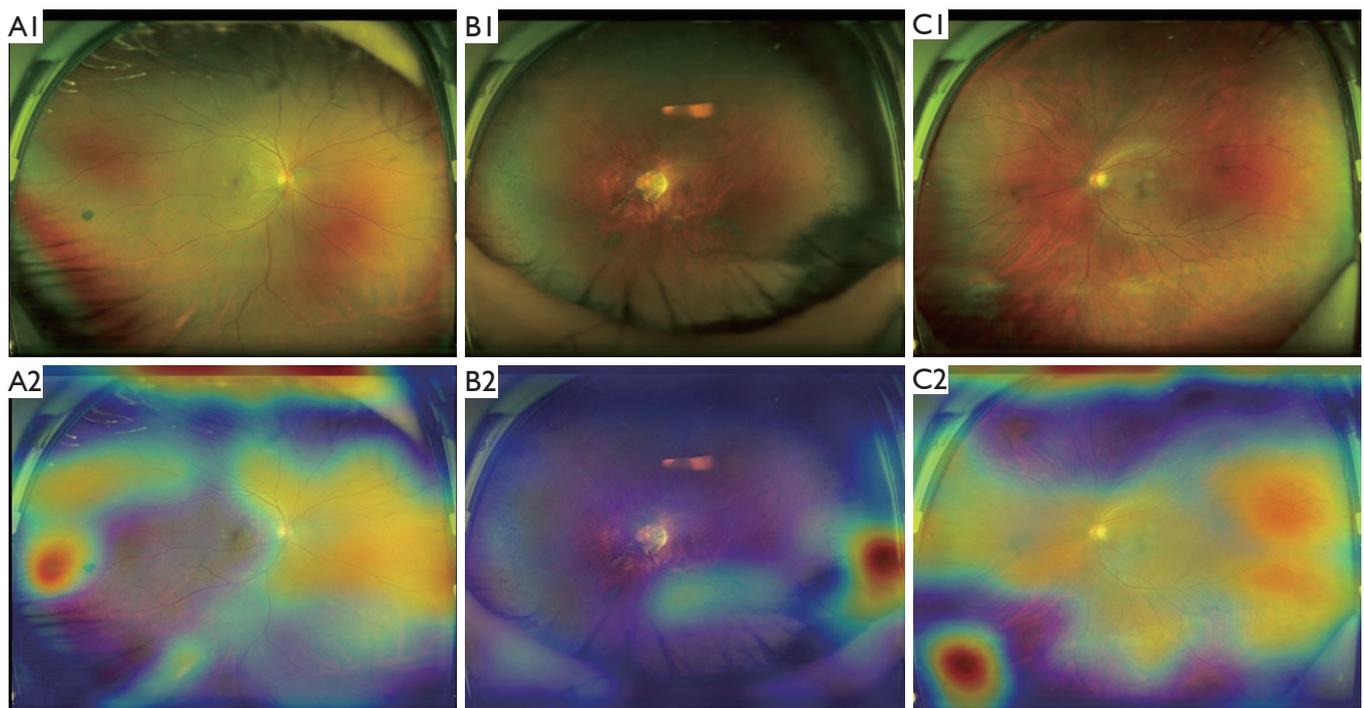

Figure 4 Ultra-widefield fundus images and corresponding heatmaps showing typical false-positive cases. (A) Retinal pigmented nevus shown on the left of A1 is the reddest region displayed in heatmap A2; (B) Dense retinal hyperpigmentation manifested on the bottom right of B1 is the reddest region visualized in heatmap B2; (C) proliferative vitreous membrane presented on the bottom left of C1 is the reddest region shown in heatmap C2.

DL systems are often regarded as a "black box" because they utilize millions of image features to identify diseases. Although various high-accuracy DL systems have been developed for the automated classification of retinopathies, the rationale for the outputs generated by these systems is unclear to clinicians. In an attempt to explain this rationale, our study visualized the DL systems in the detection of NPRLs, with heatmaps generated for all true-positive images and all false-positive images. Lesions typically seen in NPRLs were identified as the important regions
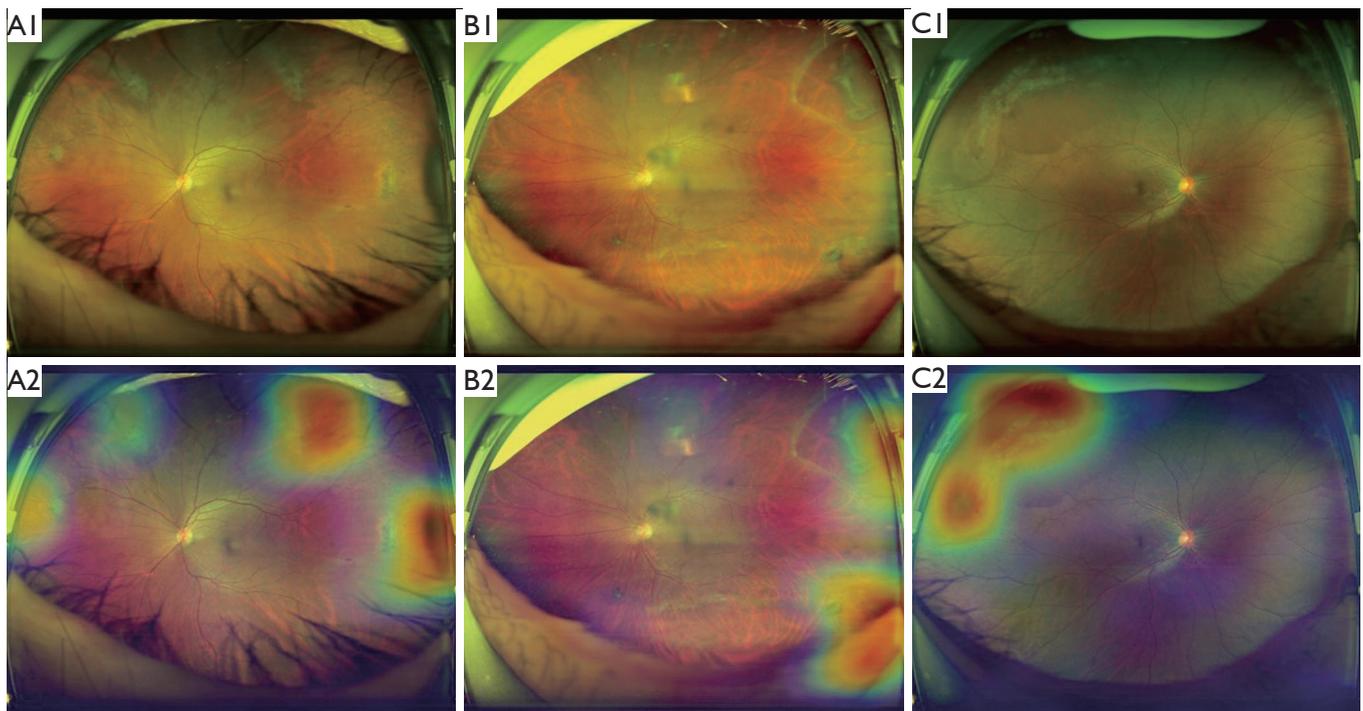
Figure 5 Ultra-widefield fundus images and corresponding heatmaps showing typical true-positive cases. (A) Lattice degeneration shown in A1 corresponds to the highlighted regions displayed in heatmap A2; (B) retinal breaks manifested in the B1 are the highlighted regions visualized in heatmap B2; (C) both lattice degeneration and retinal breaks presented in the C1 are the highlighted regions shown in heatmap C2.

in 150 of 154 true-positive NPRL images (*Figure 5*), which further substantiates the validity of this DL system. Similarly, Kermany *et al.* (26) used the occlusion test to identify the areas of greatest importance used by the DL model in assigning a diagnosis of AMD. This test successfully identified the most clinically significant regions of pathology in 94.7% of images. In addition, Keel *et al.* (27) created heatmaps highlighting localized landmarks on images of DR and glaucoma, with an accuracy of 90% and 96%, respectively.

Although our DL system had high accuracy, misclassification still existed. To analyze errors made by the best DL system, we checked all the misclassified images carefully. In the false-negative group, only 2 images were misclassified due to lesions that were unclear and partly covered by eyelashes (*Figure 3*). In all 5 false-positive NPRL images, the heatmap appeared in an area that was similar to lattice degeneration (*Figure 4*). Increasing the number of these error-prone images in the training set could potentially minimize both false-positive and false-negative results.

Due to the high sensitivity, our DL system can

be qualified for serving two purposes in the clinic: first, screening NPRLs as part of ophthalmic health evaluations in organizations such as physical examination centers or community hospitals which lack professional ophthalmologists; second, detecting peripheral RRD precursors in patients who cannot tolerate a dilated fundus examination, such as those with shallow anterior chamber angles. If a patient with a positive result is identified by the DL algorithm, then that patient can be referred to a retinal specialist to further determine whether retinal traction is involved in NPRLs or whether prophylactic treatment should be conducted for the prevention of RRD, and to validate the follow-up time. In addition, screeners could educate patients with positive findings on symptoms that might be early warning signs of RRD, such as flashes, peripheral visual field loss, increased floaters and decreased visual acuity. Moreover, these patients would be advised to contact their ophthalmologist promptly if they have any of these symptoms. Ideally, our system can improve early detection and timely treatment of RRD.

Our study has several limitations. First, we used two-dimensional images lacking stereoscopic qualities rather

Ann Transl Med 2019;7(22):618 | http://dx.doi.org/10.21037/atm.2019.11.28

Page 10 of 11

Li et al. Deep learning for identifying NPRLs

than three-dimensional images to train the DL system, thus making the identification of elevated lesions such as retinal traction involving NPRLs challenging. Second, the lattice degeneration and retinal breaks were not classified independently. Establishing a DL system to precisely differentiate retinal breaks from lattice degeneration was difficult due to the small retinal breaks that often emerged within lattice degeneration. Therefore, our current system mainly applies to screen people with dangerous RRD precursors in the peripheral retina and then refer them to ophthalmologists in a timely manner. Future improvement of the system will help to distinguish the specific types of these precursors. Third, although UWF imaging can capture the largest retinal view compared to other existing technologies, this method still does not cover the whole retina. Hence our DL system may miss a few NPRLs diagnoses which are not captured by the UWF imaging. Moreover, a missed diagnosis would occur if NPRLs appear in an obscured area of the image. A multi-center study with large sample sizes is needed to investigate the generalizability of the DL system for detecting NPRLs.

In summary, our DL system is able to achieve high sensitivity and specificity for identifying NPRLs using UWF images. Future studies will be dedicated to investigating the feasibility of using this algorithm as a screening approach to detect NPRLs in different clinical settings.

Footnote

Conflicts of Interest: 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 approved by the Institutional Review Board of Zhongshan Ophthalmic Center (Guangzhou, Guangdong, China) and adhered to the tenets of the Declaration of Helsinki.

References

1. Mitry D, Singh J, Yorston D, et al. The predisposing pathology and clinical characteristics in the scottish retinal detachment study. Ophthalmology 2011;118:1429-34.
2. Wilkinson CP. Evidence-based analysis of prophylactic treatment of asymptomatic retinal breaks and lattice degeneration. Ophthalmology 2000;107:12-15, 15-18.
3. Chen DZ, Koh V, Tan M, et al. Peripheral retinal changes in highly myopic young asian eyes. Acta Ophthalmol 2018;96:e846-51.
4. Mitry D, Charteris DG, Fleck BW, et al. The epidemiology of rhegmatogenous retinal detachment: geographical variation and clinical associations. Br J Ophthalmol 2010;94:678-84.
5. Wilkinson CP. Evidence-based medicine regarding the prevention of retinal detachment. Trans Am Ophthalmol Soc 1999;97:397-404, 404-406.
6. Wilkinson CP. Interventions for asymptomatic retinal breaks and lattice degeneration for preventing retinal detachment. Cochrane Database Syst Rev 2014;(9):CD003170.
7. Jalali S. Retinal detachment. Community Eye Health 2003;16:25-26.
8. American academy of ophthalmology retina/vitreous panel. Posterior vitreous detachment, retinal breaks, and lattice degeneration. Available online: www.aao.org/ppp; 2014. Accessed February 10, 2019.
9. Gonzales CR, Gupta A, Schwartz SD, et al. The fellow eye of patients with rhegmatogenous retinal detachment. Ophthalmology 2004;111:518-21.
10. Pastor JC, Fernandez I, Rodriguez De La Rua E, et al. Surgical outcomes for primary rhegmatogenous retinal detachments in phakic and pseudophakic patients: the retina 1 project - report 2. Br J Ophthalmol 2008;92:378-82.
11. Tsai CY, Hung KC, Wang SW, et al. Spectral-domain optical coherence tomography of peripheral lattice degeneration of myopic eyes before and after laser photocoagulation. J Formos Med Assoc 2019;118:679-85.
12. 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;316:2402-10.

13. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. Ophthalmology 2017;124:962-9.

14. Ting DS, 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;318:2211-23.

15. Keel S, Li Z, Scheetz J, et al. Development and validation of a deep-learning algorithm for the detection of neovascular age-related macular degeneration from colour fundus photographs. Clin Exp Ophthalmol 2019. [Epub ahead of print].

16. Phene S, Dunn RC, Hammel N, et al. Deep learning and glaucoma specialists: the relative importance of optic disc features to predict glaucoma referral in fundus photographs. Ophthalmology 2019. [Epub ahead of print].

17. Nagiel A, Lalane RA, Sadda SR, et al. Ultra-widefield fundus imaging: a review of clinical applications and future trends. Retina 2016;36:660-78.

18. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.

19. 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.

20. Keras-Documentation. Models for image classification with weights trained on imagenet. Available online: https://keras.io/applications/; 2016. Accessed March 1, 2019.

21. Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. Int J Comput Vision 2015;115:211-52.

22. 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. [Epub ahead of print].

23. Brown JM, Campbell JP, Beers A, et al. Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks. JAMA Ophthalmol 2018;136:803-10.

24. Burlina PM, Joshi N, Pekala M, et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. JAMA Ophthalmol 2017;135:1170-6.

25. Hwang DK, Hsu CC, Chang KJ, et al. Artificial intelligence-based decision-making for age-related macular degeneration. Theranostics 2019;9:232-45.

26. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 2018;172:1122-1131.e9.

27. Keel S, Wu J, Lee PY, et al. Visualizing deep learning models for the detection of referable diabetic retinopathy and glaucoma. JAMA Ophthalmol 2019;137:288-92.