Heliyon 10 (2024) e33813

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Deep learning for predicting circular retinal nerve fiber layer thickness from fundus photographs and diagnosing glaucoma

Yongli Xu [a,b,1], Hanruo Liu [c,d,1], Run Sun [b], Huaizhou Wang [c], Yanjiao Huo [c], Ningli Wang [c,e,**], Man Hu [f,*]

[a] College of Statistics and Data Science, Beijing University of Technology, Beijing, China
[b] Department of Mathematics, Beijing University of Chemical Technology, Beijing, China
[c] Beijing Institute of Ophthalmology, Beijing Tongren Hospital, Capital Medical University, Beijing Ophthalmology & Visual Science Key Lab, Beijing, China
[d] School of Information and Electronics, Beijing Institute of Technology, Beijing, China
[e] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University & Capital Medical University, Beijing Tongren Hospital, Beijing, China
[f] Department of Ophthalmology, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing, China

ARTICLE INFO

Keywords:
Deep learning
Retinal nerve fiber layer thickness
Fundus photography
Optical coherence tomography
Glaucoma

ABSTRACT

Purpose: This study aimed to propose a new deep learning (DL) approach to automatically predict the retinal nerve fiber layer thickness (RNFLT) around optic disc regions in fundus photography trained by optical coherence tomography (OCT) and diagnose glaucoma based on the predicted comprehensive information about RNFLT.

Methods: A total of 1403 pairs of fundus photographs and OCT RNFLT scans from 1403 eyes of 1196 participants were included. A residual deep neural network was trained to predict the RNFLT for each local image in a fundus photograph, and then a RNFLT report was generated based on the local images. Two indicators were designed based on the generated report. The support vector machines (SVM) algorithm was used to diagnose glaucoma based on the two indicators.

Results: A strong correlation was found between the predicted and actual RNFLT values on local images. On three testing datasets, we found the Pearson r to be 0.893, 0.850, and 0.831, respectively, and the mean absolute error of the prediction to be 14.345, 17.780, and 19.250 μm, respectively. The area under the receiver operating characteristic curves for discriminating glaucomatous from healthy eyes was 0.860 (95 % confidence interval, 0.799–0.921).

Conclusions: We established a novel local image-based DL approach to provide comprehensive quantitative information on RNFLT in fundus photographs, which was used to diagnose glaucoma. In addition, training a deep neural network based on local images to predict objective detail information in fundus photographs provided a new paradigm for the diagnosis of ophthalmic diseases.

* Corresponding author. Beijing Children's Hospital, 56 Nanlishi Road, Xicheng District, Beijing 100045, China.
** Corresponding author. Beijing Tongren Hospital, 8 Chongwenmennei Street, Dongcheng District, Beijing, 100005, China.
E-mail addresses: wningli@vip.163.com (N. Wang), human1125@163.com (M. Hu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e33813
Received 15 March 2024; Received in revised form 27 June 2024; Accepted 27 June 2024
Available online 28 June 2024

1. Introduction

Glaucoma is the leading cause of irreversible blindness in the world [1], and the number of patients with glaucoma is estimated to reach 110 million by 2040 [2]. Glaucoma causes structural changes to the optic disc (OD) and retinal nerve fiber layer (RNFL), resulting in corresponding defects in visual function [3]. Glaucoma progression is irreversible, but early diagnosis and treatment can usually delay glaucomatous visual impairment [4]. Assessing structural changes to OD and RNFL is essential for the early diagnosis of glaucoma.

Fundus photography and optical coherence tomography (OCT) are the primary examinations to assess structural changes to OD and RNFL. Fundus photography has the advantages of low cost and easy operation for diagnosing glaucoma. However, a high degree of subjectivity exists in the evaluation of fundus photographs by ophthalmologists. Subjective evaluation of fundus photographs suffers from low reproducibility, even when performed by senior ophthalmologists [5–7]. In addition, ophthalmologists frequently underestimate or overestimate glaucoma likelihood when evaluating fundus photographs [8]. OCT can accurately measure the RNFL thickness (RNFLT) around the OD, which is a sensitive indicator for evaluating the structural damage caused by glaucoma [9]. However, the cost of OCT is much higher than fundus photography, and hence OCT equipment is generally unavailable in primary hospitals and physical examination centers in developing countries.

The development of artificial intelligence (AI) provides a powerful tool for establishing the correlation between fundus photographs and objective physiological indicators. The deep learning (DL) methods can accurately predict human physiological indicators such as age, blood pressure, body mass index, and hemoglobin concentration based on fundus photographs [10–13]. In 2019, Medeiros et al. trained a DL convolutional neural network to assess the fundus photographs and predict the mean RNFLT in OCT [14]. Although this study made great progress in predicting objective indicators in OCT based on fundus photographs, using only the mean RNFLT is far from revealing the rich details in OCT and thus does not provide the necessary information for diagnosing glaucoma.

This study aimed to design a DL model based on local photographs (DL-LP) to predict the OCT-measured RNFLT curve around the OD. Furthermore, based on this predicted RNFLT curve, a report sheet similar to the OCT examination was automatically generated to comprehensively evaluate the status of RNFL (Fig. 1). Finally, the scale and morphological information of this curve were extracted as features, and support vector machines (SVM) algorithm [15] was used to make an automatic diagnosis of glaucoma.

2. Methods

A total of 1403 pairs of fundus photographs and OCT data were included in this study between 2016 and 2020. The study was conducted according to the tenets of the Declaration of Helsinki and approved by the institutional review board of Beijing Tongren Hospital (identifier, 1212017BJTR519). The medical ethics committee exempted the need for the informed consent of patients as this

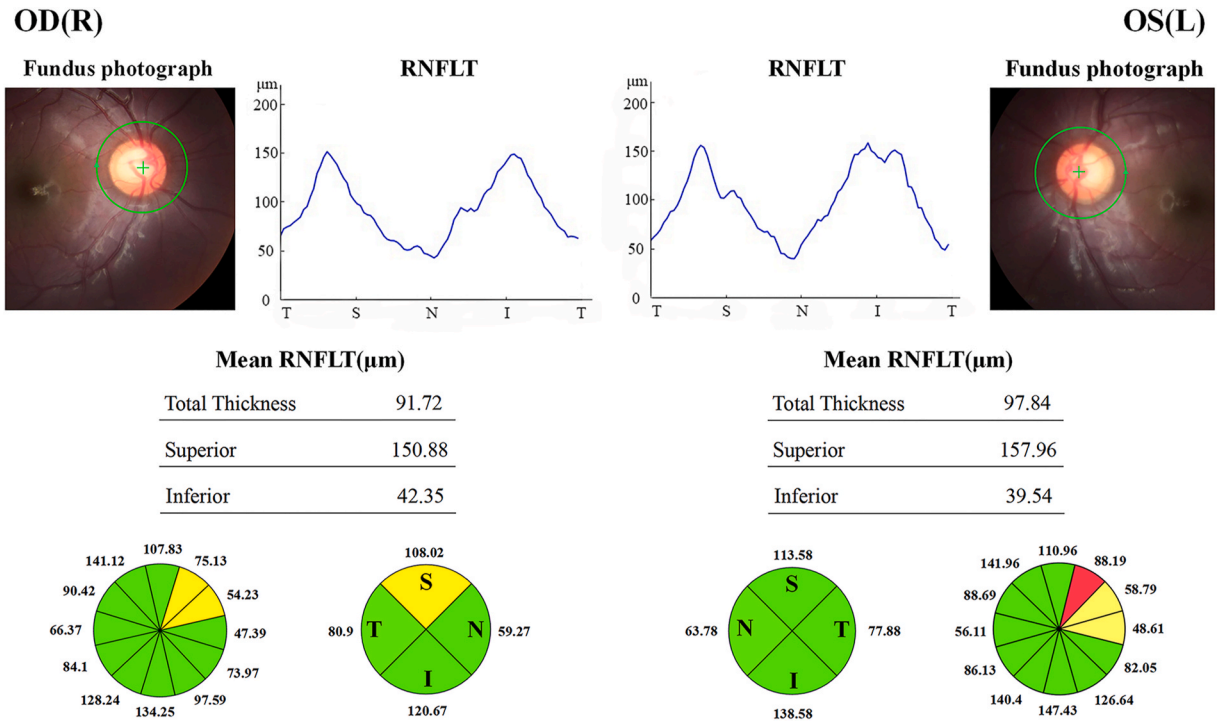

Fig. 1. An example of an automatically generated RNFL examination report from fundus photographs.
RNFL retinal nerve fiber layer, I inferior, S superior, N nasal, T temporal.

study was a retrospective review and analysis of fully anonymized color retinal fundus photographs and OCT reports.

2.1. Participants

The 1403 pairs of fundus photographs and OCT data were divided into (1) a training dataset comprising 220 pairs of fundus photographs (CR-DGi; Canon) with 1572 × 1308 pixels and OCT (SD-OCT; Heidelberg Engineering) data collected from Beijing Tongren Hospital; (2) an internal testing dataset comprising 55 pairs of fundus photographs (CR-DGi; Canon) with 1572 × 1308 pixels and OCT (SD-OCT; Heidelberg Engineering) data collected from Beijing Tongren Hospital; (3) an internal testing dataset comprising 264 pairs of fundus photographs (CR-DGi; Canon) with 1716 × 1404, 1924 × 1556, and 2196 × 1740 pixels and OCT (SD-OCT; Heidelberg Engineering) data collected from Beijing Tongren Hospital; and (4) an external testing dataset comprising 864 pairs of fundus photographs (VK-2; Kowa) with 2032 × 1934 pixels and OCT (DRI-OCT; Topcon) data collected from Handan First Hospital. The details of the training dataset and three testing datasets are presented in Table 1. The purpose of including the external testing dataset is to verify the generalization ability of proposed algorithm.

The inclusion criteria were as follows: (1) age 18 years or older; (2) complete records of fundus photographs and OCT and medical history; and (3) for participants with the same name or ID number, only one of them was included. The exclusion criteria were as follows: (1) fundus photographs containing severe resolution reductions or significant artifacts; (2) illumination of fundus photographs too dark or too light; (3) OCT images having poor quality, motion artifacts, and segmentation errors; and (4) participants with a history of other ocular or systemic diseases affecting the optic nerve other than glaucoma.

Two senior glaucoma specialists (M.H. and H.W.) used the fundus photographs and OCT to detect glaucomatous optic neuropathy (GON). These two glaucoma specialists independently diagnosed each eye as non-GON, probable GON, and definite GON according to the criteria described in a previous study [16]. When an eye was judged as probable GON by at least one glaucoma specialist, or the two glaucoma specialists could not reach an agreement, that eye was judged as probable GON. The training dataset comprised 31 fundus photographs of GON, 35 of probable GON, and 154 of non-GON. The testing dataset 1 comprised 7 fundus photographs of GON, 9 of probable GON, and 34 of non-GON. The testing dataset 2 consisted of 41 fundus photographs of GON, 57 of probable GON, and 166 of non-GON. Considering that the data in the testing dataset 3 were collected from the population examined physically, the two glaucoma specialists did not perform glaucoma diagnosis on these fundus photographs.

2.2. Development of the DL system

We designed a DL system to predict the local RNFLT in color fundus photographs. The goal was to use a deep neural network (DNN) to predict the RNFLT at the centers of a series of square images (patches) surrounding the OD in color fundus photographs. In order to obtain a DNN with accurate predictions, we need to construct a training dataset to train the DNN (modulate the parameters in the DNN). After completion of the training of this DNN, we used it to predict the local RNFLT of the new fundus photographs in the testing dataset. The workflow of this DL system is shown in Fig. 2. The input X of each sample was a patch in the color fundus photograph, and the output Y was the RNFLT at the center of the patch. Registering the color fundus photograph and OCT to construct the input–output pair (X, Y) is needed because the input came from the color fundus photograph and the output from the OCT.

The original RNFLT data were obtained by circular scanning of the OD region using OCT. This circular scanning was used to generate a customized one-page report to show the scanned location and RNFLT. In the OCT report in Fig. 2, the upper left diagram is a low-resolution fundus photograph. In this fundus photograph, the green circle indicates the path of the circular scanning (with a diameter of 3.45 mm). In the lower right diagram of this OCT report, the black curve represents the RNFLT within the circumference of the scan path. We designed a MATLAB program to extract the value of this black curve within 360° and stored it as an RNFLT vector.

Performing image registration of the color fundus photograph and the low-resolution fundus photograph in the OCT report is necessary to obtain the corresponding position of the RNFLT curve in the OCT report in the color fundus photograph. We used Photoshop software for manual registration: the blood vessels in the two fundus photographs were completely overlapped by scaling, translating, and rotating the low-resolution fundus photographs. After image registration, the corresponding circumference of the OCT scanning path in the color fundus photograph was recorded. This registration operation was implemented to all the images in the training and testing datasets. In the color fundus photograph, 90 points were selected at equal intervals on this circle. With each point as the center and 60 pixels as the side length, 90 patches in the color fundus photograph were intercepted. Our goal was to predict the RNFLT at the center of the patch based on the red–green–blue (RGB) intensity values within each patch.

We established a training dataset consisting of 19,800 (220 × 90) patch-level training samples based on the results of image

Table 1
Baseline characteristics of participants.

Datasets	Pairs of images(N)	Eyes (N)	RNFLT (um) Mean ± SD	Individuals (N)	Age(y) Mean ± SD	Female (N)Total (%)	Source Datasets
Training	220	220	91.2 ± 17.8	131	49.4 ± 18.1	77(58.8)	Beijing Tongren Hospital
Testing 1	55	55	90.6 ± 18.2	33	48.5 ± 17.6	19 (57.6)	Beijing Tongren Hospital
Testing 2	264	264	91.7 ± 17.2	168	47.1 ± 17.4	87 (51.8)	Beijing Tongren Hospital
Testing 3	864	864	104.45 ± 11.72	864	41.7 ± 12.3	298 (34.5)	Handan First Hospital

RNFLT, retinal nerve fiber layer thickness; SD, standard deviation.

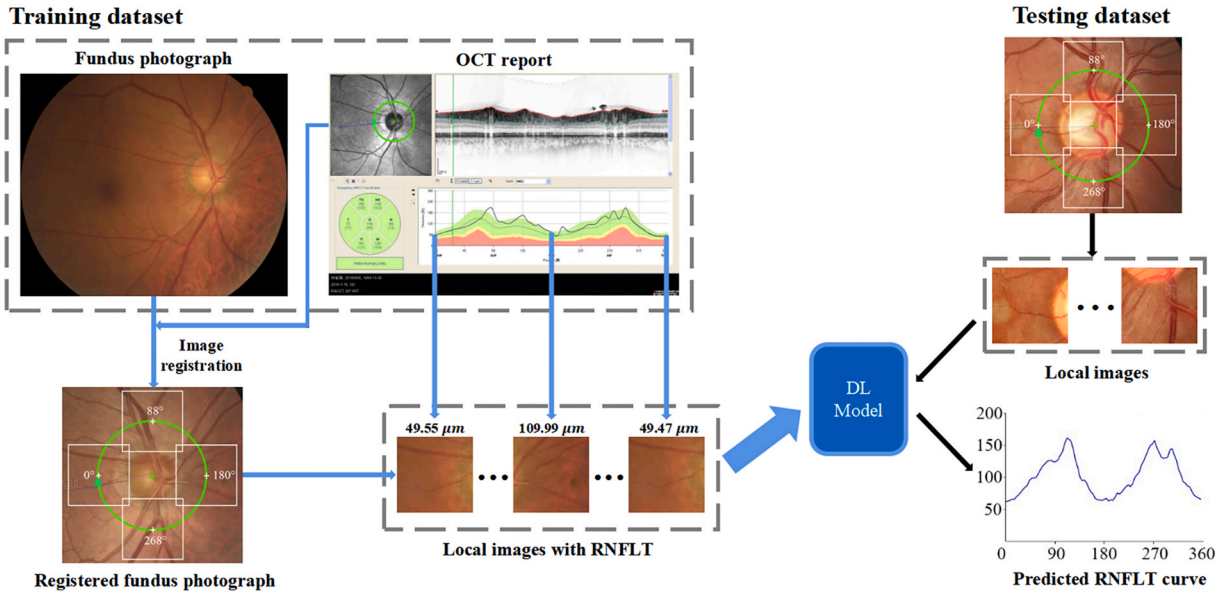

Fig. 2. Workflow of the deep learning model based on the local photographs. RNFLT retinal nerve fiber layer thickness.

registration. The input of each sample was a square RGB image of size 60 × 60 × 3 (the value of each pixel represents the intensity of the image in the corresponding channel), and the output was a real value (RNFLT at the center of the square). Before training the DNN on the training dataset, the pixel values were scaled to range from 0 to 1.

We used a standard residual deep neural network (ResNet-50) [17] as the backbone architecture for RNFLT prediction. The parameters of ResNet-50 were trained on the ImageNet dataset [18]. However, further training was performed on the training dataset because the regression task of the present study differed largely from that of the ImageNet. We made two adjustments to the structure of ResNet-50. First, considering that the standard ResNet-50 is used for classification problems while RNFLT prediction is a regression problem, we replaced the cross-entropy loss with square loss. Second, we added a global pooling layer before the fully connected layer of ResNet-50 to reduce the number of parameters of the model. Before training the ResNet-50, we performed normalization preprocessing on the pixel intensity values of fundus images. The pixel intensity values in each fundus image were linearly transformed, and the range of values was changed from [0, 255] to [−1, 1]. The network was trained with a stochastic gradient descent method with a learning rate of 10^{-4} and minibatches of size 64.

Following the method described in a previous study [19], we calculated the mean RNFLT and inferior, superior, nasal, temporal (ISNT) score of the RNFLT curve, which reflected the scale and shape information of the RNFLT curve, respectively. Then, a two-dimensional classification line was established by SVM based on the two aforementioned indicators. The mean value of the RNFLT of 90 patches in a fundus photograph was calculated as the mean RNFLT. The ISNT score of the RNFLT curve was calculated as follows. First, the original RNFLT vector (with 90 dimensions) was normalized by subtracting each component from the average of this vector and dividing by the standard deviation of this vector. Then, the inner product of this normalized vector and a baseline vector was calculated as the ISNT score. The baseline vector was the average RNFLT vector of 50 typical normal eyes. The fundus photographs of these typical normal eyes were collected in the physical examination center of Tongren Hospital and were not included in the training and testing datasets.

2.3. Statistical analyses

The performance of the DL-LP was evaluated by comparing the predictions based on color fundus photographs with the actual OCT local RNFLT in testing datasets. We calculated the mean absolute error (MAE) and the mean absolute percentage error (MAPE) of the predictions as well as the coefficient of determination and the Pearson correlation coefficient. In testing dataset 2, the receiver-operating characteristic (ROC) curve and area under the ROC curve (AUC) were used to assess and compare the ability of the SVM algorithms based on predicted RNFLT in discriminating eyes with glaucoma from healthy eyes. Two SVMs were compared based only on the predicted average RNFLT and on both predicted average RNFLT and ISNT score. Besides the ROC curve and AUC, the sensitivity and specificity of the operating point with a threshold of 0.5 in ROC curves were also measured with two-sided 95 % confidence interval (CI). These CIs were calculated as Clopper–Pearson intervals, which were exact intervals based on cumulative probabilities. All statistical tests were performed using the Statistical Product and Service Solutions (SPSS) software.

3. Results

A strong correlation was observed between patch-level DL predictions of RNFLT from fundus photographs and the actual OCT measurements in the testing datasets (Fig. 3a–c). Table 2 shows that, in patch-level, the MAE between the DL predictions of RNFLT from fundus photographs and the actual OCT measurements were all less than 20 μm in the testing datasets 1, 2, and 3. Especially, the MAE was less than 15 μm in the testing dataset 1. In patch-level, the MAPE between the DL predictions of RNFLT from fundus photographs and the actual OCT measurements were all less than 0.28. The R^2 was greater than 0.47 (range, 0.477–0.699; $P < 0.001$), and the Pearson r was greater than 0.83 (range, 0.831–0.893; $P < 0.001$).

In image-level (the mean value of all the 90 patch-level RNFLT values), in testing datasets 1, 2, and 3, the MAE between the DL predictions of RNFLT from fundus photographs and the actual OCT measurements were all less than 11 μm (range, 8.83–10.73 μm). In image-level, the MAPE between the DL predictions of RNFLT from fundus photographs and the actual OCT measurements were all less than 0.13. A strong correlation was found between the predicted and the observed image-level RNFLT values. The R^2 was greater than 0.42 (range, 0.427–0.619; $P < 0.001$), and the Pearson r was greater than 0.46 (range, 0.468–0.824; $P < 0.001$).

Fig. 4 shows the prediction of RNFLT by DL-LP in one eye with GON and one eye with non-GON. Fig. 4b shows that the range of RNFLT predicted by DL-LP was roughly equivalent to the true value, and the MAE of 90 patches was 11.06 μm for the fundus photograph of the eye with GON in Fig. 4a. The maximum absolute error was 32.15 μm, and more than 70 % of patches had an absolute error of less than 15 μm. Fig. 4d shows that the change trend of the RNFLT curve predicted by DL-LP was highly consistent with the real curve, and the MAE of 90 patches was 11.46 μm for the fundus photograph of the eye with non-GON in Fig. 4c.

Fig. 5(a–h) shows several local images in the two fundus photographs in Fig. 4, and the RNFLT predicted by DL-LP on these local images. Figs. 4 and 5 show that DL-LP had smaller errors in predicting RNFLT on the nasal and temporal sides and larger errors in predicting RNFLT on the inferior and superior sides. As shown in Fig. 4d, the areas with the largest prediction errors were near the peaks of the real RNFLT curve. The absolute error was the largest, with a maximum value of 43 μm, at one of the peaks of the real RNFLT curve. One possible reason for this phenomenon is that the proportion of training samples at peaks is small.

We established two indicators, mean RNFLT and ISNT score [19], for each fundus photograph based on the 90 patch-level RNFLTs predicted by DL-LP to characterize the scale and morphology of RNFLT curves. First, the mean RNFLT and ISNT score were calculated using the predicated RNFLT from the fundus images. Then SVM was used to diagnose glaucoma based on the two indictors above. The Mean ± standard deviation of predicted ISNT scores of eyes in the training dataset and testing datasets 1, 2 and 3 were 28.77 ± 7.86, 29.82 ± 6.39, 29.27 ± 7.51, and 30.8 ± 4.51, respectively.

In testing dataset 2, we evaluated the accuracy of applying SVM to discriminate GON from non-GON based on the aforementioned two indicators. The AUC was found to be 0.860 (95 % CI, 0.799–0.921) for the SVM based on both mean RNFLT and ISNT score, and 0.809 (95 % CI, 0.736–0.882) for the SVM based only on mean RNFLT (Fig. 6). When the threshold was set as 0.5, the sensitivity and specificity reached 0.7805 (95 % CI, 0.620–0.889) and 0.8494 (95 % CI, 0.784–0.898), respectively, for the SVM based on both mean RNFLT and ISNT score. For the SVM based only on mean RNFLT, the sensitivity and specificity reached 0.732 (95 % CI, 0.568–0.853) and 0.783 (95 % CI, 0.711–0.842), respectively. The aforementioned results indicated that predicting the local RNFLT in fundus photographs to obtain the morphological indicator ISNT score of the RNFLT curve played an essential role in improving the accuracy of glaucoma diagnosis.

4. Discussion

Fundus photographs contain a wealth of human health information, and AI has been used to establish a correlation between fundus photographs and human physiological indicators [10–13]. The RNFLT curve can reflect a variety of eye diseases [20]. Ophthalmologists usually cannot obtain comprehensive information on RNFL based on fundus photographs. OCT is required for this, which is expensive compared with fundus photographs. Therefore, using AI to establish the correlation between fundus photographs and the

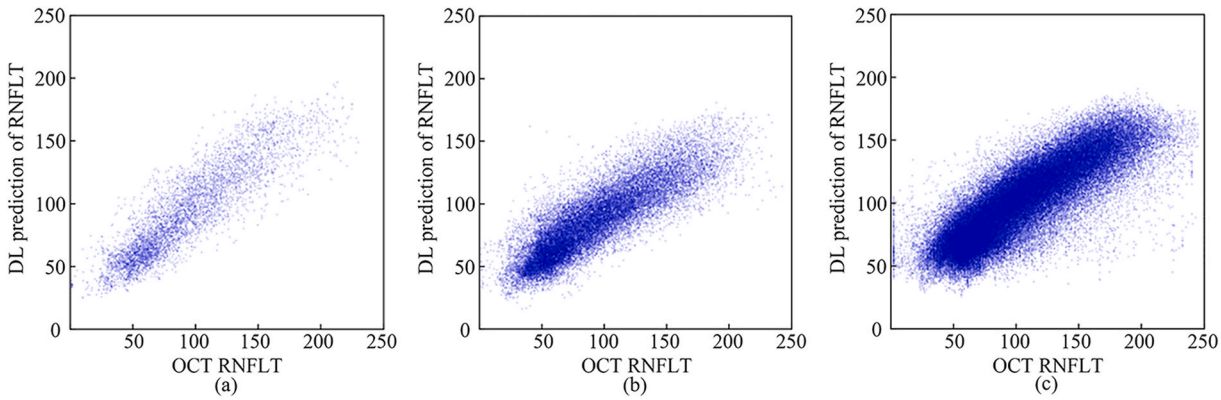

Fig. 3. Scatter diagram illustrating the relationship between patch-level predictions obtained by the deep learning algorithm evaluating fundus photographs and actual retinal nerve fiber layer thickness measurements from OCT. Data are from the testing datasets 1, 2, and 3.

Table 2
Performance evaluation of DL-LP for predicting RNFLT.

	Patch-level				Image-level			
	MAE/μm	MAPE	R^2	r	MAE/μm	MAPE	R^2	r
Testing dataset 1	14.345	0.244	0.699	0.893	9.142	0.110	0.427	0.824
Testing dataset 2	17.780	0.234	0.477	0.850	10.730	0.122	0.457	0.766
Testing dataset 3	19.250	0.278	0.491	0.831	8.830	0.084	0.619	0.468

DL-LP, deep learning based on local photographs; RNFLT, retinal nerve fiber layer thickness; MAE, mean absolute errors; MAPE, mean absolute percentage error.

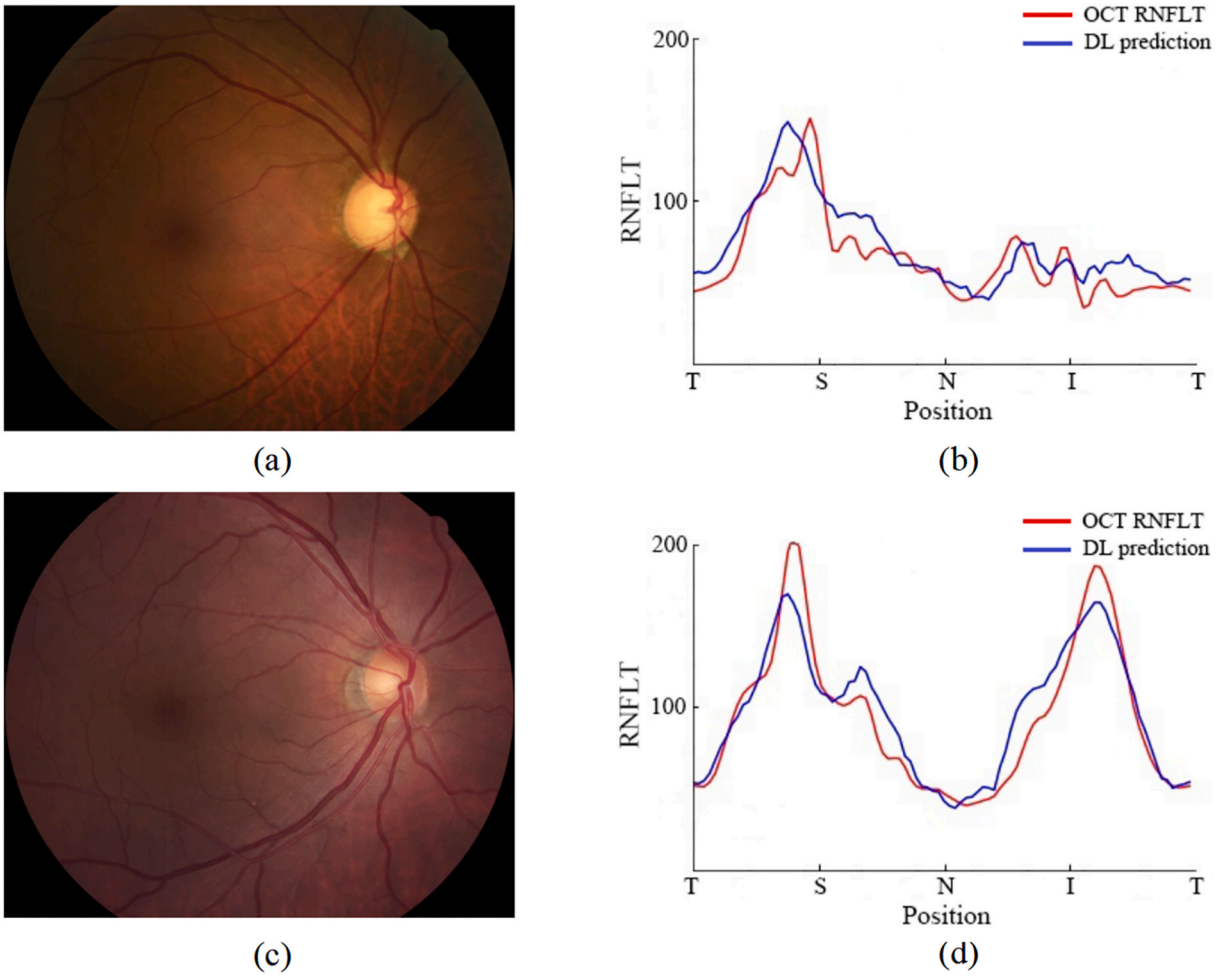

Fig. 4. Two examples of fundus photographs and the predicted and actual patch-level RNFLT curve. (a) An example of a fundus photograph with GON; (b) predicted and actual patch-level RNFLT curves of the fundus photograph in (a); (c) example of a fundus photograph with non-GON; and (d) predicted and actual patch-level RNFLT curve of the fundus photograph in (c). RNFLT retinal nerve fiber layer thickness, GON glaucomatous optic neuropathy.

RNFLT curve has important clinical practicality. AI has been applied to predict the mean RNFLT and diagnose glaucoma based on fundus photographs [14]. However, Mean RNFLT only shows the average level of RNFLT without showing details of the RNFLT curve. The method proposed in this study can predict RNFLTs in 90 local areas around the optic disc. Therefore, this method provided richer and more comprehensive RNFLT information, which is essential for glaucoma diagnosis. The experimental results showed that compared to the diagnosis of glaucoma only based on the predicted mean RNFLT, the predicted mean RNFLT and ISNT score can obtain higher prediction accuracy (Table 3). This study also provided an economical and powerful auxiliary tool for the diagnosis of other RNFLT-related diseases.

In the field of automatic diagnosis of glaucoma based on fundus photographs, the DL methods have achieved satisfactory sensitivity and specificity. However, most DL methods are black-box methods without interpretability, which limits the practical application of

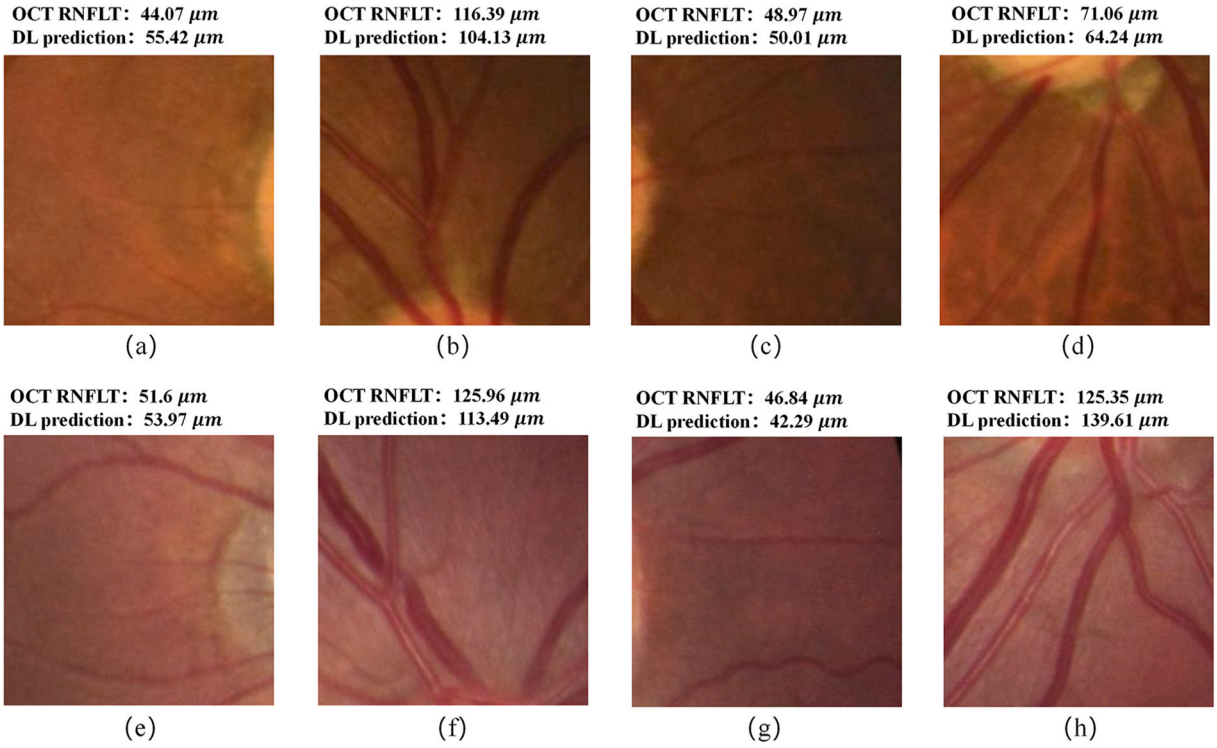

Fig. 5. Examples of local images in fundus photographs. (a–d) Local images from a fundus photograph of an eye with GON; (e–h) Local images from a fundus photograph of an eye with non-GON. GON glaucomatous optic neuropathy; RNFLT, retinal nerve fiber layer thickness.

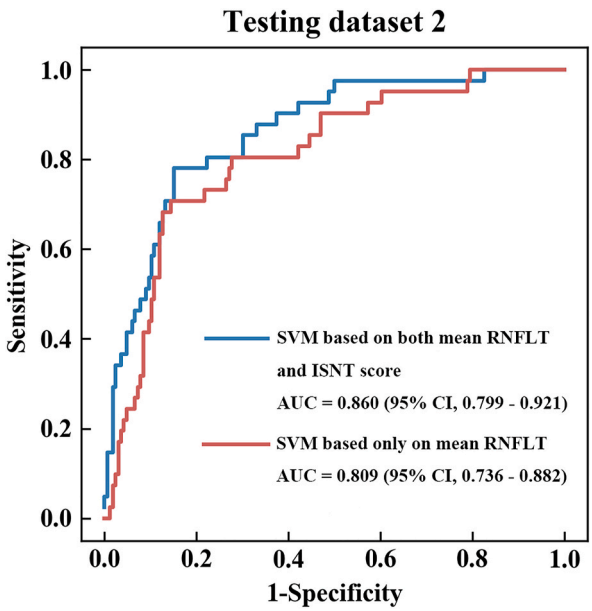

Fig. 6. Receiver-operating characteristic curve in diagnosing glaucoma in testing dataset 2 using SVM based on predicted RNFLT curve by DL-LP. AUC area under the ROC curve, SVM support vector machines, RNFLT retinal nerve fiber layer thickness, DL-LP deep learning based on local photographs, ISNT inferior, superior, nasal, temporal.

such methods in clinical practice [21]. At present, the quantitative indicators for glaucoma diagnosis obtained using DL methods are still very limited, including VCDR, mean CDR and ISNT score of rim [16,22]. In this study, we predicted the RNFLT curve around the OD based on fundus photographs, and then calculated the mean RNFLT and INST score of the RNFLT curve, expanding the quantitative

Table 3
Performance of the SVM with predicted RNFLT by DL-LP on testing dataset 2.

Method	Number of images				% (95 % CI)	
	True positive	False negative	True negative	False positive	Sensitive	Specificity
SVM based only on the predicted mean RNFLT	30	11	130	36	0.732 (0.568–0.853)	0.783 (0.711–0.842)
SVM based on the predicted mean RNFLT + ISNT score	32	9	141	25	0.781 (0.620–0.889)	0.8494 (0.784–0.898)

SVM, support vector machines; RNFLT, retinal nerve fiber layer thickness; DL-LP, deep learning based on local photographs; ISNT, inferior, superior, nasal, temporal.

indicators for diagnosing glaucoma. In the future, we will design more interpretable indicators based on the predicted RNFLT curve and combine these indicators with rim-related indicators (VCDR, etc.) to further improve the accuracy of the algorithm while maintaining the interpretability.

In addition, this study established a new paradigm for using AI to mine local information in medical images. For the training of DL models, thousands or even tens of thousands of well-labeled training datasets are often required to obtain accurate predictions. In this study, we only used 220 pairs of fundus photographs and OCT data to build the DL-LP model and obtained accurate prediction results for RNFLT. This was due to our localization technology, where each pair of image-level data was converted into 90 pairs of patch-level local data. This DL paradigm greatly reduced the cost of constructing training data and enabled the DL model to learn richer detailed information in fundus photographs. This paradigm can be used to explore the correlation between other pairs of medical data, such as optical image–CT pair and optical image–MRI pair.

Despite the promising results, our study has several limitations. First, every local fundus photograph was selected as a square with resolution of 60×60 given the limitations in the GPU's computational power. As a result, the texture details of RNFL at large scales were partially lost in the local fundus images, which affected the accuracy of the DNN in predicting RNFLT. In addition,the proposed DL-LP was only evaluated in the medical records collected in two hospitals. It is of great significance to evaluate our model in multiple medical centers, especially in primary eye care providers. Furthermore, only Chinese patients were included. It is need to test the algorithm using data from other ethnicities to investigate its performance.

Funding statement

This study was partially supported by the National Natural Science Foundation of China (Grant number: 82171051).

Data availability statement

The data analyzed during the study are available from the corresponding author upon reasonable request.

CRediT authorship contribution statement

Yongli Xu: Writing – original draft, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Hanruo Liu:** Writing – review & editing, Resources, Data curation, Conceptualization. **Run Sun:** Visualization, Software, Formal analysis. **Huaizhou Wang:** Resources, Formal analysis. **Yanjiao Huo:** Resources, Data curation. **Ningli Wang:** Writing – review & editing, Supervision. **Man Hu:** Writing – original draft, Project administration, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:HanruoLiu reports financial support was provided by National Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R.R. Bourne, G.A. Stevens, R.A. White, et al., Causes of vision loss worldwide, 1990-2010: a systematic analysis, Lancet Glob Health 1 (6) (2013) e339–e349.
[2] Y.C. Tham, X. Li, T.Y. Wong, et al., Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis, Ophthalmology 121 (11) (2014) 2081–2090.
[3] R.N. Weinreb, T. Aung, F.A. Medeiros, The pathophysiology and treatment of glaucoma: a review, JAMA 311 (18) (2014) 1901–1911.
[4] D.F. Garway-Heath, D.P. Crabb, C. Bunce, et al., Latanoprost for open-angle glaucoma (UKGTS): a randomised, multicentre, placebo-controlled trial, Lancet 385 (9975) (2015) 1295–1304.
[5] J.M. Tielsch, J. Katz, H.A. Quigley, et al., Intraobserver and interobserver agreement in measurement of optic disc characteristics, Ophthalmology 95 (3) (1988) 350–356.

[6] R. Varma, W.C. Steinmann, I.U. Scott, Expert agreement in evaluating the optic disc for glaucoma, Ophthalmology 99 (2) (1992) 215–221.
[7] H.D. Jampel, D. Friedman, H. Quigley, et al., Agreement among glaucoma specialists in assessing progressive disc changes from photographs in open-angle glaucoma patients, Am. J. Ophthalmol. 147 (1) (2009) 39–44.
[8] H.H. Chan, D.N. Ong, Y.X. Kong, et al., Glaucomatous optic neuropathy evaluation (GONE) project: the effect of monoscopic versus stereoscopic viewing conditions on optic nerve evaluation, Am. J. Ophthalmol. 157 (5) (2014) 936–944.
[9] A.J. Tatham, F.A. Medeiros, Detecting structural progression in glaucoma with optical coherence tomography, Ophthalmology 124 (12S) (2017) S57–S65.
[10] T.H. Rim, G. Lee, Y. Kim, et al., Prediction of systemic biomarkers from retinal photographs: development and validation of deep-learning algorithms, Lancet Digit Health 2 (10) (2020) e526–e536.
[11] A. Mitani, A. Huang, S. Venugopalan, et al., Detection of anaemia from retinal fundus images via deep learning, Nat. Biomed. Eng. 4 (1) (2020) 18–27.
[12] C.Y. Cheung, D. Xu, C.Y. Cheng, et al., A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre, Nat. Biomed. Eng. 5 (6) (2021) 498–508.
[13] B. Babenko, A. Mitani, I. Traynis, et al., Detection of signs of disease in external photographs of the eyes via deep learning, Nat. Biomed. Eng. 6 (12) (2022) 1370–1383.
[14] F.A. Medeiros, A.A. Jammal, A.C. Thompson, From machine to machine: an OCT-trained deep learning algorithm for objective quantification of glaucomatous damage in fundus photographs, Ophthalmology 126 (4) (2019) 513–521.
[15] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (1995) 273–297.
[16] Y. Xu, M. Hu, H. Liu, et al., A hierarchical deep learning approach with transparency and interpretability based on small samples for glaucoma diagnosis, NPJ Digit Med 4 (1) (2021) 48.
[17] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
[18] J. Deng, W. Dong, R. Socher, et al., Imagenet: a large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition, Ieee, 2009.
[19] Y. Xu, Y. Yang, Y. Huo, M. Hu, A group of novel indexes of optical coherence tomography for computer-aided diagnosis of glaucoma, J Nonlinear Convex Anal 23 (10) (2022) 2437–2447.
[20] M. Adhi, J.S. Duker, Optical coherence tomography–current and future applications, Curr. Opin. Ophthalmol. 24 (3) (2013) 213–221.
[21] H. Liu, L. Li, I.M. Wormstone, et al., Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs, JAMA Ophthalmol 137 (12) (2019) 1353–1360.
[22] R. Hemelings, B. Elen, A.K. Schuster, et al., A generalizable deep learning regression model for automated glaucoma screening from fundus images, NPJ Digit Med 6 (1) (2023) 112.