tvst

Article

A Deep Learning Model for Screening Multiple Abnormal Findings in Ophthalmic Ultrasonography (With Video)

Di Chen[2,*], Yi Yu[1,*], Yiwen Zhou[1,*], Bin Peng[1], Yujing Wang[4], Shan Hu[3], Miao Tian[1], Shanshan Wan[1], Yuelan Gao[1], Ying Wang[1], Yulin Yan[1], Lianlian Wu[2], LiWen Yao[2], Biqing Zheng[3], Yang Wang[1], Yuqing Huang[1], Xi Chen[1], Honggang Yu[2], and Yanning Yang[1]

[1] Department of Ophthalmology, Renmin Hospital of Wuhan University, Wuhan, Hubei Province, China
[2] Department of Gastroenterology, Renmin Hospital of Wuhan University, Wuhan, Hubei Province, China
[3] School of Resources and Environmental Sciences of Wuhan University, Wuhan, Hubei Province, China
[4] Department of Ophthalmology, Zhongnan Hospital of Wuhan University, Wuhan, Hubei Province, China

Correspondence: Yanning Yang, Department of Ophthalmology, Renmin Hospital of Wuhan University, 99 Zhangzhidong Road, Wuhan 430060, Hubei Province, China. e-mail: ophyyn@163.com

Purpose: The purpose of this study was to construct a deep learning system for rapidly and accurately screening retinal detachment (RD), vitreous detachment (VD), and vitreous hemorrhage (VH) in ophthalmic ultrasound in real time.

Methods: We used a deep convolutional neural network to develop a deep learning system to screen multiple abnormal findings in ophthalmic ultrasonography with 3580 images for classification and 941 images for segmentation. Sixty-two videos were used as the test dataset in real time. External data containing 598 images were also used for validation. Another 155 images were collected to compare the performance of the model to experts. In addition, a study was conducted to assess the effect of the model in improving lesions recognition of the trainees.

Results: The model achieved 0.94, 0.90, 0.92, 0.94, and 0.91 accuracy in recognizing normal, VD, VH, RD, and other lesions. Compared with the ophthalmologists, the modal achieved a 0.73 accuracy in classifying RD, VD, and VH, which has a better performance than most experts ($P < 0.05$). In the videos, the model had a 0.81 accuracy. With the model assistant, the accuracy of the trainees improved from 0.84 to 0.94.

Conclusions: The model could serve as a screening tool to rapidly identify patients with RD, VD, and VH. In addition, it also has potential to be a good tool to assist training.

Translational Relevance: We developed a deep learning model to make the ultrasound work more accurately and efficiently.

Introduction

This study showed that 10.8 million people who are blind and 35.1 million people who are visually impaired were caused by cataract. Cataract remains a major public health problem in the world.[1] With opaque dioptric media, ultrasonography has been the first-choice tool to evaluate the posterior segment and assess the structural changes in eyes.[2,3] Thus, using ultrasound to assess the posterior segment of all patients with dense cataract prior to surgery is large examinational volumes.

Retinal detachment (RD) has been reported to had a incidence ranged from 6.3 to 17.9 per 100,000 people.[4] Rapid diagnosis is critical in these patients, as an irreversible loss of vision may be caused if diagnosis is delayed.[5] Vitreous hemorrhage (VH) with vitreous detachment (VD) also should be timely diagnosed to avoiding further serious consequences, which are associated with a high incidence of retinal tears and detachment.[6] The ability to accurately recognize VD

translational vision science & technology

and VH processes is associated with the urgency of the patients who would need an ophthalmologist for further examination. Therefore, prompt recognition and appropriate treatment of these three symptoms are essential in the primary care setting.[7] Use of ocular ultrasonography may be effective for early detection of them.[8] However, a prospective study conducted by Kim et al. showed that the sensitivity and specificity in diagnosing RD were 75% and 94%.[9] Although the result of the study conducted by Shinar et al. found a 97% sensitivity and a 92% specificity.[10] Studies also reported the accuracy of VD and VH varied from 91% to 99%[11,12] and 84% to 100%.[12,13] The accuracy of using ultrasound in diagnosing ocular diseases seems to vary among ophthalmologists. Additionally, an ophthalmologist may not be available on call in rural settings. The use of ultrasound to accurately and efficiently identify ocular emergencies in these settings is urgent.

Deep learning has greatly improved the diagnostic accuracy and efficiency of some medical diseases. An artificial intelligence (AI) system was used to accurately classifying the presence of any diabetic retinopathy in fundus images and has dramatically reduced the time-consuming in diagnosing diabetic retinopathy.[14] Ding et al. used a deep learning model to help gastroenterologists analyzing small bowel capsule endoscopy images more efficiently and more accurately.[15] Even with the large volumes and poor consistency among the ophthalmologists, there is still a lack of reliable deep learning ophthalmic ultrasound screening systems to rapidly diagnose patients with RD, VD, and VH. An efficient and accurate AI-assisted diagnostic system has the potential to reduce visual impairment due to misdiagnosis and delayed diagnosis in clinical practice.

In our present study, we used a deep learning model to screen RD, VD, and VH in ultrasound images, which should be referred to an ophthalmologist for further evaluation and treatment. We evaluated the model in both internal and external datasets. In addition, a study was used to assess the effect of the model in improving lesions' recognition of the trainees.

Methods

The study was approved by the Ethics Committee of Renmin Hospital of Wuhan University and under trail registration number ChiCTR2000036326 of the Primary Registries of the World Health Organization (WHO) Registry Network, and adherence to the Declaration of Helsinki. The training ultrasonic images were collected from Renmin Hospital of Wuhan University from June 2017 to August 2020. The instrument used in this study was the SW-2100 (Tianjin Sauvy Electronic Technology Co. LTD, China). The model had five functions. First, the deep convolutional neural network 1 (DCNN1) was used to filter out unqualified images; the DCNN2 to segment the eyeball; and the DCNN3 to classify abnormal and normal images. Then, the DCNN4 was used to recognize VD. Meanwhile, DCNN5, DCNN6, and DCNN7 were used to recognize VH, RD, and others, respectively. Figure 1 shows the workflow of the model.

Datasets and Preprocessing

Five hundred ninety-three images from 326 eyes of 244 patients from Renmin Hospital of Wuhan University were collected to filter out unqualified images. Another 941 images from 607 eyes of 532 patients were annotated using image labeling software (VGG, Visual Geometry Group Department of Engineering Science, University of Oxford) by one experienced expert and used to segment the eyeball. There were 3580 images (train:test = 2812:768) from 1668 eyes of 1416 patients that were used to train and test DCNN3 to distinguish normal and abnormal images. The 1980 training images from 1012 eyes of 853 patients and 594 testing images from 293 eyes of 275 patients were collected to recognize VD. As the DCNN5 was used to recognize VH, 2436 images (train:test = 1842:594) from 1210 eyes of 1032 patients were used. In order to train the DCNN6 to recognize RD, 511 RD images (train:test = 421:90) from 269 eyes of 261 patients and 2004 not-RD images (train:test = 1500:504) from 1008 eyes of 834 patients were collected. For recognizing other lesions, 1574 not-other lesion images (RD, VD, and VH) from 746 eyes of 672 patients and 1275 other lesions' images from 667 eyes of 567 patients were collected to train and test DCNN7. We also used an external dataset of 598 images from 215 eyes of 154 patients from Zhongnan Hospital of Wuhan University to evaluate the performance of the system under different environments (Fig. 2). Supplementary Table S1 showed the baseline information and sample distribution. Supplementary Figure S1 showed the representative images predicted by the system. Images from the same eyes or same patients were not split among the data sets, and were placed independently in the testing set or training set. And all images were reviewed by two experts with discussion. In order to train the model to recognize RD, VD, VH and other lesions, the following images were excluded: two or more of RD, VD, VH, or other lesions co-exist in one image.
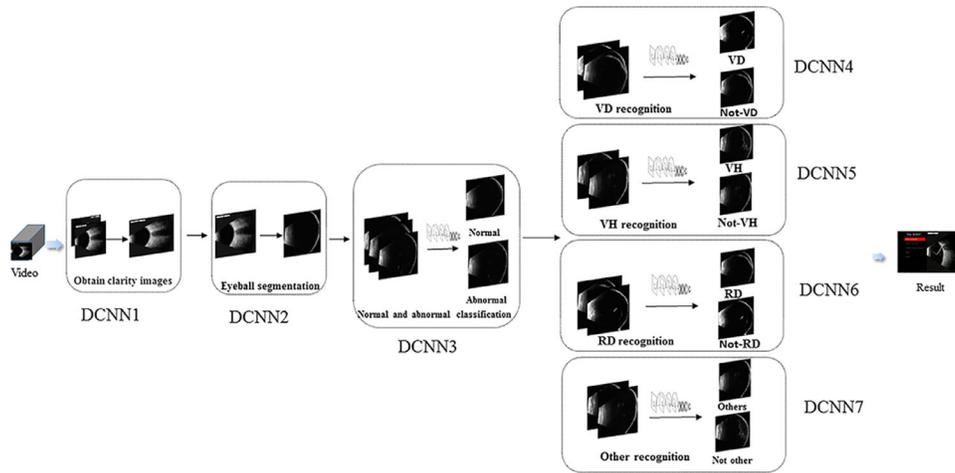
Figure 1. **The flowchart of the model.** Images from videos were put into the proposed architectures, and firstly screened by DCNN1 to obtain clear images, and then segmented the eyeball by DCNN2. Next, the images would be classified to abnormal and normal by DCNN3. Finally, the abnormal images will be further classified to VD, VH, RD, and others by DCNN4 to DCNN7, respectively.
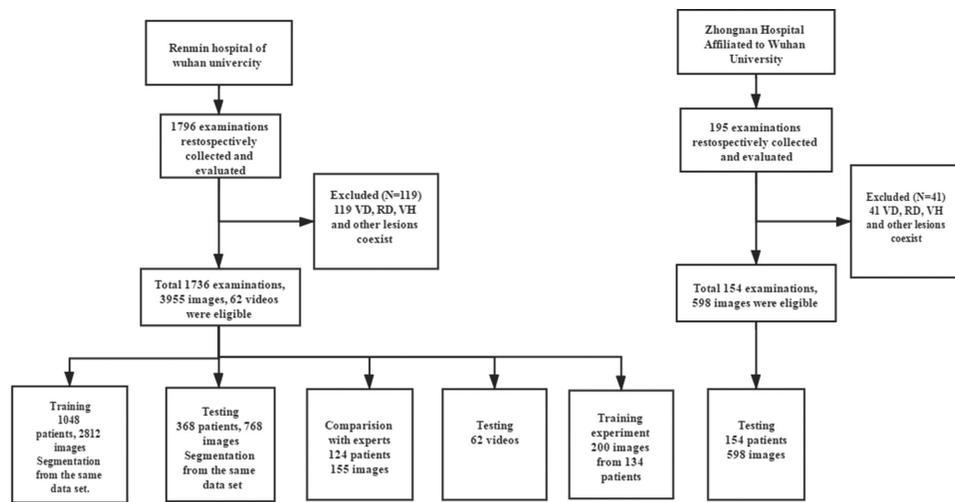

Figure 2. Flowchart of the model development and validation. RD, Retinal detachment; VD, vitreous detachment; VH, vitreous hemorrhage.

Development of the Model

We used U-net++[16] to segment the eyeball and ResNet-50[17] for image classification. First, with transfer learning,[18] we used our data to retrain and replace the final classification layer of the architectures. Dropout,[19] early stopping,[20] and data augmentation[21] were used to minimize the overfitting risk in our model. During the training phase, the input images were randomly resized to 224 degrees × 224 degrees pixels in classification and 512 degrees × 512 degrees pixels in segmentation.

The flow of the system was: (1) images were put into the proposed architectures, and first filtered out unqualified images by DCNN1; (2) then it segmented the eyeballs by DCNN2; (3) then images would be classified to abnormal and normal by DCNN3; and (4) next, the abnormal images will be further recognized by DCNN4 to DCNN7. Models 4 to 7 were parallel in the system. Images from DCNN3 will be recognized by models 4 to 7, respectively. In doing so, the inputs to each DCNN were dependent on the outputs of the previous DCNNs (e.g. DCNN1 to DCNN4). Supplementary Figure S2 showed the confusion matrices

of each DCNN. Supplementary Figure S3 shows the receiver operating characteristic (ROC) of the DCNNs.

Meanwhile, we also established class activation maps[22] (CAMs) to indicate suspicious lesion regions. A global average pooling was performed on the convolutional feature map, which was used as the feature of the full connection layer to produce the desired output. The confidence of the prediction is positively correlated with the color depth of the CAMs.

The parameters of segmentation were: 2 batch size, 0.0001 learning rate, and 0.5 threshold to distinguish between background and positive samples. For the parameters of classification models, the batch size was 64, the learning rate was 0.0001, and the convergent epochs was 30.

Python software (version 3.6.5) was used to write the algorithms. The open-source Keras (version 2.1.5) library and TensorFlow (version 1.12.2) library were used as the backend. A server with four NVIDIA Geforce GTX 1080 (with 8 GB GPU memory) was used to train the model.

Evaluation of the Model

The internal test dataset (Renmin Hospital of Wuhan University) and one external dataset from Zhongnan Hospital of Wuhan University were used to evaluate the performance of our model.

Testing the DCNNs on Videos

The time per frame in the videos to output a prediction (including segmentation and classification) was 240 ms on a GPU. Sixty-two videos from 68 eyes of 62 patients from Renmin Hospital of Wuhan University were collected to test the performance of the model. These videos were cut into images with one frame per second (fps) and reviewed by two experts with discussion. The accuracy of these videos was defined as the number of correctly identified frames divided by the total number of frames. In addition, in Supplementary Visualization S1, images were captured at 5 fps. The rule - "the prediction of the most frames in consecutive 12 frames as the final result of this fragment" - was used to smoosh noises. Thus, each 12 frames will give a result.

Comparison Between DCNNs and Ophthalmologists in Still Images

A set of test images (155 images of RD, VD, VH, others, and normal) from 131 eyes of 124 patients were prepared to assess the classified ability of the DCNNs

and the ophthalmologists. We compared the classification performance of the model to three experts from Renmin Hospital of Wuhan University. All of them were asked to classify images into RD, VD, VH, others, and normal. The three experts had at least 10 years of experience in ophthalmic ultrasound.

Comparison the Performance of the Ophthalmologists With and Without Assistance of Deep Learning

In order to assess the effect of the model in improving lesions recognition of the trainees, we included 10 trainees, none who had any prior experience or training in ultrasound in this study. First, the trainees were asked to read the images without the model assistant. Then, after 2 weeks to washout, the trainees were asked to read the images with the model assistant in the same images. Two hundred images from 143 eyes of 134 patients were collected (RD = 40; VD = 40; VH = 40; normal = 40; and others = 40) to assess.

Statistical Analysis

Accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and areas under the ROC curve (AUC) with 95% confidence interval (CI) were described. The criteria for evaluating the performance of each model were optimized from 0 to 1 with values greater than 0.90, ranging between 0.8 and 0.9, between 0.5 and 0.79, and less than 0.5 defined as the highest, good, moderate, and poor performance in the present study.[23] A χ^2 test was performed to analyze the difference in accuracy between the model and experts. The Mann-Whitney U test was applied to compare the accuracy of the trainees with or without model assistant. A value of $P < 0.05$ was considered a statistically significant difference. All results were analyzed with SPSS software, version 23.0 (IBM, Chicago, IL) and MedCalc, version 19.1-64 bit (MedCalc Software Ltd., Ostend, Belgium).

Results

Test in Both Internal and External Datasets

DCNN1 achieved a 0.95 (95% CI = 0.91–0.99) accuracy to filter out unqualified images. The Intersection over Union (IoU) of DCNN2 was 0.93. In classifying the abnormal and normal images, DCNN3 achieved both highest accuracy of 0.94 (95% CI = 0.92–0.96) in the internal dataset and 0.97 (95% CI =

Table 1.　Performance of the DCNNs in Classification

		Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
DCNN3	Internal	0.94 (0.93–0.96)	0.94 (0.92–0.96)	0.95 (0.91–0.98)	0.99 (0.97–0.99)	0.82 (0.76–0.87)	0.95 (0.93–0.96)
	External	0.97 (0.95–0.98)	0.99 (0.97–0.99)	0.89 (0.81–0.94)	0.98 (0.96–0.99)	0.93 (0.86–0.97)	0.94 (0.92–0.96)
DCNN4	Internal	0.90 (0.88–0.93)	0.88 (0.79–0.94)	0.91 (0.88–0.93)	0.64 (0.55–0.72)	0.98 (0.96–0.99)	0.89 (0.87–0.92)
	External	0.81 (0.77–0.84)	0.91 (0.80–0.97)	0.79 (0.75–0.83)	0.36 (0.28–0.45)	0.99 (0.96–0.99)	0.88 (0.86–0.91)
DCNN5	Internal	0.92 (0.89–0.94)	0.79 (0.69–0.86)	0.94 (0.92–0.96)	0.73 (0.63–0.81)	0.96 (0.93–0.97)	0.86 (0.83–0.89)
	External	0.88 (0.85–0.91)	0.73 (0.65–0.81)	0.93 (0.90–0.95)	0.79 (0.71–0.86)	0.91 (0.87–0.93)	0.91 (0.88–0.93)
DCNN6	Internal	0.94 (0.93–0.96)	0.92 (0.84–0.96)	0.95 (0.92–0.96)	0.76 (0.67–0.84)	0.99 (0.97–0.99)	0.93 (0.91–0.95)
	External	0.88 (0.85–0.91)	0.79 (0.72–0.85)	0.93 (0.89–0.95)	0.85 (0.79–0.90)	0.89 (0.85–0.92)	0.91 (0.89–0.93)
DCNN7	Internal	0.91 (0.89–0.94)	0.92 (0.88–0.95)	0.91 (0.87–0.94)	0.92 (0.88–0.94)	0.91 (0.87–0.94)	0.91 (0.89–0.93)
	External	0.92 (0.89–0.94)	0.79 (0.71–0.86)	0.96 (0.93–0.98)	0.88 (0.81–0.93)	0.93 (0.89–0.95)	0.94 (0.91–0.95)

DCNN3, normal and abnormal classification model; DCNN4, VD recognition model; DCNN5, VH recognition model; DCNN6, RD recognition model; DCNN7, other lesion recognition model; RD, retinal detachment; VD, vitreous detachment; VH, vitreous hemorrhage. CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; AUC, areas under the receiver operating characteristic curve.

Table 2.　Comparison Between the Model and Ophthalmologists in Still Images

	Accuracy of Classification (95% CI)			
	Model	Expert A	Expert B	Expert C
VD	0.97 (0.90–1.03)	0.29 (0.12–0.46)	0.29 (0.12–0.46)	0.61 (0.43–0.80)
VH	0.45 (0.27–0.64)	0.52 (0.33–0.70)	0.55 (0.36–0.73)	0.61 (0.43–0.80)
RD	0.58 (0.40–0.77)	0.87 (0.75–1.00)	0.81 (0.66–0.95)	0.81 (0.66–0.95)
Others	0.87 (0.75–1.00)	0.93 (0.43–1.03)	0.52 (0.33–0.70)	0.81 (0.66–0.95)
Normal	0.77 (0.62–0.93)	0.90 (0.79–1.01)	1.00	0.93 (0.43–1.03)
Average	0.73 (0.66–0.80)	0.70 (0.63–0.78)	0.63 (0.56–0.71)	0.75 (0.69–0.82)

RD, retinal detachment; VD, vitreous detachment; VH, vitreous hemorrhage. CI, confidence interval.

0.95–0.98) in the external dataset. To recognize VD, VH, RD, and other lesions, the DCNN4, DCNN5, DCNN6, and DCNN7 all had the highest accuracy with 0.90 (95% CI = 0.88–0.93), 0.92 (95% CI = 0.89–0.94), 0.94 (95% CI = 0.93–0.96), and 0.91 (95% CI = 0.89–0.94). As the AUCs of the model, DCNN3, DCNN6, and DCNN7 had the highest values with 0.95 (95% CI = 0.93–0.96), 0.93 (95% CI = 0.91–0.95), and 0.91 (95% CI = 0.89–0.93). In addition, DCNN4 and DCNN5 also had good AUCs (0.89 and 0.86). These results demonstrated our model had a high potential for accurately detecting these lesions in the future (Table 1).

Testing on the Videos

Sixty-two videos were used to test the performance of the model. On these videos, the model achieved an accuracy of 0.81 (95% CI = 0.79–0.83). The accuracy of the model to recognize RD, VD, VH, others, and normal were 0.79 (95% CI = 0.80–0.92), 0.80 (95% CI = 0.80–0.88), 0.77 (95% CI = 0.70–0.84), 0.82 (95% CI

= 0.78–0.86), and 0.88 (95% CI = 0.83–0.93), respectively.

Comparison Between the DCNNs and Ophthalmologists in Still Images

Three experts classified the images into RD, VD, VH, others, and normal with an average accuracy of 0.70 (95% CI = 0.63–0.78), 0.63 (95% C =: 0.56–0.71), and 0.75 (95% CI = 0.69–0.82), respectively. The model achieved an accuracy of 0.73 (95% CI = 0.66–0.80), which has a higher accuracy than 2 of the experts (both $P < 0.05$; Table 2).

Comparison of the Performance of the Ophthalmologists With and Without the Assistance of Deep Learning

The trainees had an accuracy of 0.84 (95% CI = 0.83–0.86) without the model assistant and a 0.94 (95% CI = 0.93–0.95) accuracy with the assistant. More details about the performances of individual trainees
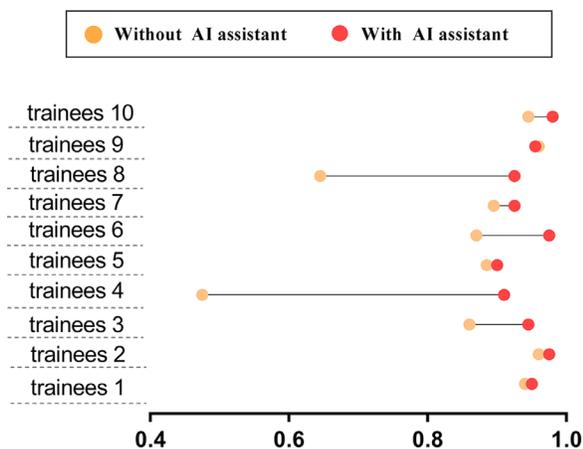
Figure 3. **The changes of the accuracy in the trainees.** Horizontal lines depict the change in accuracy for each trainee with and without model assistant. The *orange dot* represents performance without model assistant, and the *red dot* represents performance with the model assistant.

are shown (Supplementary Table S2). In addition, the accuracy of our model was 0.91 (95% CI = 0.86–0.95). With the model assistant, we could see the mean accuracy of the trainees had a statistically significant increase (0.10, 95% CI = 0.01–0.21, $P < 0.05$). The changes of the accuracy in the trainees ae shown in Figure 3.

Discussion

We introduced a system using deep learning to screen the abnormal findings and highlight the location of the lesions in ultrasonic images. The model achieved a good accuracy in recognizing RD, VD, VH, other lesions, and normal patients, which has a better performance than most experts.

The number of the patients with cataract was large in recent years. Ultrasound as an advisable and important tool has been widely used to screen patients with cataract prior to surgery.[2] However, this study showed that more than 30% patients were normal in the ultrasound examinations.[24,25] If we can screen out this subset of normal patients, the workload of ophthalmologist will be greatly reduced. In the present study, our model had a 0.94 accuracy in classifying abnormal and normal images, which supplied that the model can be a good tool to screen out most normal patients.

RD, VD, and VH are three common diagnoses encountered in the emergency department (ED).[26] Timely diagnosis and treatment of these lesions means less prevalence of blindness.[27] In our present study, the accuracy of our model to recognize RD, VD,

and VH were 0.94, 0.90, and 0.92, respectively. The highest accuracy suggested that this model may be a good choice to assist screening these parts of patients, which were the major patients in the ED. Meanwhile, the performance of our model in recognizing patients with RD, VD, VH, others, and normal was better than most experts. In the area of limited resources or experts, the model had a good ability to screen these lesions accurately and rapidly. Furthermore, our model had a highest 0.94 accuracy in recognizing RD. It was important to rapidly and accurately detect RD, as an irreversible vision loss may be caused by it.[28] Thus, the ability to rapidly detect RD can help these patients to receive timely consultation. Furthermore, VD often present with symptoms similar to those of RD.[29] In addition, VD can co-exist with RD,[30] making a definitive diagnosis more challenging. The highest accuracy of our model to differentiate RD and VD can assist the ophthalmologists to distinguish more accurately.

Ultrasound has become a reasonable and readily available tool in the ED, which can help bridge the ophthalmologic screening skills gap.[31] However, across the country, ultrasound training is not uniform.[32] One study showed that a highest confidence in ultrasound skills was associated with much higher accuracy of an emergency ultrasound diagnosis.[33] Thus, those using ultrasound in emergent situations should been trained to achieve a proper level to avoid making critical mistakes in the diagnosis.[34] In the present study, our model can rapidly recognize lesions and showed the locations of the lesions with a heatmap. With the assistance of the model, the accuracy of the trainees had improved from 0.84 to 0.94. The effectiveness of the model implied that it may be a good assistant training tool in the future. Furthermore, the prompt use of the heatmap could show the diagnostic basis of the model, which may give a good tip to the trainees.

Our model achieved a good accuracy on the videos. In addition, the time of per frame in the videos to output a prediction was 240 ms, which meant that the model may have potential to achieve real-time application in clinical practice.

As for the limitation, ophthalmologists do not incorporate patient clinical features and history in reviewing the images, which may have a slight effect on the accuracy.

In conclusion, we constructed a deep learning ophthalmic ultrasound screening system. The system could serve as a screening tool to exclude most normal patients and recognize RD, VD, and VH, which should be referred to an ophthalmologist for further evaluation and treatment. In addition, the system may be a good tool to assist training in the future.

Acknowledgments

References

1. Khairallah M, Kahloun R, Bourne R, et al. Number of people blind or visually impaired by cataract worldwide and in world regions, 1990 to 2010. *Invest Ophthalmol Vis Sci*. 2015;56:6762–6769.

2. Bello T, Adeoti C. Ultrasonic assessment in preoperative cataract patients. *Niger Postgrad Med J*. 2006;13:326–328.

3. Ahmed J, Shaikh FF, Rizwan A, Memon MF. Evaluation of vitreo-retinal pathologies using B-scan ultrasound. *Pak J Ophthalmol*. 2009;25:1–5.

4. Mitry D, Charteris DG, Fleck BW, Campbell H, Singh J. The epidemiology of rhegmatogenous retinal detachment: geographical variation and clinical associations. *Br J Ophthalmol*. 2010;94:678–684.

5. Pastor J, Fernandez I, De La, Rua ER, et al. Surgical outcomes for primary rhegmatogenous retinal detachments in phakic and pseudophakic patients: the Retina 1 Project—report 2. *Br J Ophthalmol*. 2008;92:378–382.

6. Sarrafizadeh R, Hassan TS, Ruby AJ, et al. Incidence of retinal detachment and visual outcome in eyes presenting with posterior vitreous separation and dense fundus-obscuring vitreous hemorrhage. *Ophthalmology*. 2001;108:2273–2278.

7. Pokhrel PK, Loftus SA. Ocular emergencies. *Am Fam Physician*. 2007;76:829–836.

8. Lahham S, Shniter I, Thompson M, et al. Point-of-care ultrasonography in the diagnosis of retinal detachment, vitreous hemorrhage, and vitreous detachment in the emergency department. *JAMA Network Open*. 2019;2:e192162.

9. Kim DJ, Francispragasam M, Docherty G, et al. Test characteristics of point-of-care ultrasound for the diagnosis of retinal detachment in the emergency department. *Acad Emerg Med*. 2019;26:16–22.

10. Shinar Z, Chan L, Orlinsky M. Use of ocular ultrasound for the evaluation of retinal detachment. *J Emerg Med*. 2011;40:53–57.

11. Genovesi-Ebert F, Rizzo S, Chiellini S, Di Bartolo E, Marabotti A, Nardi M. Reliability of standardized echography before vitreoretinal surgery for proliferative diabetic retinopathy. *Ophthalmologica*. 1998;212(Suppl. 1):91–92.

12. Parchand S, Singh R, Bhalekar S. Reliability of ocular ultrasonography findings for pre-surgical evaluation in various vitreo-retinal disorders. *Semin Ophthalmol*. 2014;29:236–241.

13. Kumar A, Verma L, Jha S, Tewari H, Khosla P. Ultrasonic errors in analysis of vitreous haemmorhage. *Indian J Ophthalmol*. 1990;38:162.

14. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology*. 2017;124:962–969.

15. Ding Z, Shi H, Zhang H, et al. Gastroenterologist-level identification of small-bowel diseases and normal variants by capsule endoscopy using a deep-learning model. *Gastroenterology*. 2019;157:1044–1054.e1045.

16. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: a nested U-net architecture for medical image segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Supp 2018*. 2018;11045:3–11.

17. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* Las Vegas, NV; 2016:770–778.

18. Shao L, Zhu F, Li X. Transfer learning for visual categorization: a survey. *IEEE Trans Neural Netw Learn Syst*. 2014;26:1019–1034.

19. Baldi P, Sadowski P. The dropout learning algorithm. *Artificial Intelligence*. 2014;210:78–122.

20. Prechelt L. Automatic early stopping using cross validation: quantifying the criteria. *Neural Networks*. 1998;11:761–767.

21. Tanner MA, Wong WH. The calculation of posterior distributions by data augmentation. *J Am Stat Assoc*. 1987;82:528–540.

22. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV; 2016:2921–2929.

23. Cao K, Verspoor K, Sahebjada S, Baird PN. Evaluating the performance of various machine learning algorithms to detect subclinical keratoconus. *Transl Vis Sci Technol*. 2020;9:24.

24. Eze K, Enock M, Eluehike S. Ultrasonic evaluation of orbito-ocular trauma in Benin-City, Nigeria. *Niger Postgrad Med J*. 2009;16:198–202.

25. Qureshi MA, Laghari K. Role of B-scan ultrasonography in pre-operative cataract patients. *Int J Health Sci*. 2010;4:31.

26. Tintinalli J. *Tintinallis emergency medicine: A comprehensive study guide*. New York, NY: McGraw-Hill Education; 2015.

27. Bagheri N, Mehta S. Acute vision loss. *Prim Care*. 2015;42:347–361.

28. Muth CC. Sudden vision loss. *JAMA*. 2017;318:584–584.

29. Schott ML, Pierog JE, Williams SR. Pitfalls in the Use of Ocular Ultrasound for Evaluation of acute vision loss. *J Emerg Med*. 2013;44:1136–1139.

30. Thimons J. Posterior vitreous detachment. *Optometry Clin the Official Publication of the Prentice Society*. 1992;2:1–24.

31. Connolly K, Beier L, Langdorf MI, Anderson CL, Fox JC. Ultrafest: a novel approach to ultrasound in medical education leads to improvement in written and clinical examinations. *Western J Emerg Med*. 2015;16:143–148.

32. Bahner DP, Goldman E, Way D, Royall NA, Liu YT. The state of ultrasound education in US medical schools: results of a national survey. *Acad Med*. 2014;89:1681–1686.

33. Davis DP, Campbell CJ, Poste JC, Ma G. The association between operator confidence and accuracy of ultrasonography performed by novice emergency physicians. *J Emerg Med*. 2005;29:259–264.

34. Kaye A, Fox C, Hymel B, et al. The importance of training for ultrasound guidance in central vein catheterization. *Middle East J Anaesthesiol*. 2011;21:61–66.

Supplementary Material

Visualization 1. A representative video shows how the model screen lesions. This video shows how the model recognize lesions in five cases (RD, VH, VD, other, and normal, respectively). Begin the model to detect the lesions, the five diagnosis in the left of the videos were gray. In case 1, when the model recognizes the lesion, the diagnosis of "retinal detachment" was activated and lighted. Meanwhile, the red rectangular box highlighted the location of the RD using heatmap. The time of the examination also was shown. The other four cases were similar. According to the risk level of the lesions, we give different colors to the diagnosis of "retinal detachment, vitreous hemorrhage, posterior vitreous detachment, others and normal." When the model recognizes a lesion, the corresponding color will be activated.