tvst

Artificial Intelligence

Deep Learning-Based System for Disease Screening and Pathologic Region Detection From Optical Coherence Tomography Images

Xiaoming Chen[1,2,*], Ying Xue[3,*], Xiaoyan Wu[3], Yi Zhong[2,4], Huiying Rao[3], Heng Luo[2,4,5], and Zuquan Weng[2,4]

[1] College of Mathematics and Computer Science, Fuzhou University, Fujian province, China
[2] The Centre for Big Data Research in Burns and Trauma, College of Mathematics and Computer Science, Fuzhou University, Fujian province, China
[3] Department of Ophthalmology, Fujian Provincial Hospital, Fuzhou, China
[4] College of Biological Science and Engineering, Fuzhou University, Fujian province, China
[5] MetaNovas Biotech Inc., Foster City, CA, USA

Correspondence: Heng Luo, MetaNovas Biotech Inc., Foster City, CA, USA. e-mail: hengluo88@gmail.com
Zuquan Weng, College of Biological Science and Engineering, Fuzhou University, Fujian province, China. e-mail wengzq@fzu.edu.cn

Received: June 22, 2022
Accepted: November 29, 2022
Published: January 30, 2023

Keywords: deep learning; optical coherence tomography; image classification; object detection; ensemble learning

Citation: Chen X, Xue Y, Wu X, Zhong Y, Rao H, Luo H, Weng Z. Deep learning-based system for disease screening and pathologic region detection from optical coherence tomography images. Transl Vis Sci Technol. 2023;12(1):29, https://doi.org/10.1167/tvst.12.1.29

Purpose: This study was designed to apply deep learning models in retinal disease screening and lesion detection based on optical coherence tomography (OCT) images.

Methods: We collected 37,138 OCT images from 775 patients and labelled by ophthalmologists. Multiple deep learning models including ResNet50 and YOLOv3 were developed to identify the types and locations of diseases or lesions based on the images.

Results: The model were evaluated using patient-based independent holdout set. For binary classification of OCT images with or without lesions, the performance accuracy was 98.5%, sensitivity was 98.7%, specificity was 98.4%, and the F1 score was 97.7%. For multiclass multilabel disease classification, the models was able to detect vitreomacular traction syndrome and age-related macular degeneration both with an accuracy of more than 99%, sensitivity of more than 98%, specificity of more than 98%, and an F1 score of more than 97%. For lesion location detection, the recalls for different lesion types ranged from 87.0% (epiretinal membrane) to 98.2% (macular pucker).

Conclusions: Deep learning-based models have potentials to aid retinal disease screening, classification and diagnosis with excellent performance, which may serve as useful references for ophthalmologists.

Translational Relevance: The deep learning-based models are capable of identifying and predicting different eye diseases and lesions from OCT images and may have potential clinical application to assist the ophthalmologists for fast and accuracy retinal disease screening.

Introduction

As of 2019, more than 2.2 billion people's lives were impacted by vision impairment, and one-half of these vision problems could be prevented or cured with proper treatment.[1] Retinal diseases without treatments may lead to vision loss and a quality decrease in the patients' lives, because the retina is an important part of the visual system.[2]

Optical coherence tomography (OCT)[3] is a useful technology that can help ophthalmologists to assess the retinal health conditions of patient.[4–6] OCT generates series of cross-sectional two-dimensional images (B-scans), which can be used to examine structural details and pathological changes on different retinal layers.[7]

Despite the usefulness, retinal lesions on each B-scan must be manually identified, which requires specialized expertise and time. Additionally, the diagnosis results are subject to individual ophthalmologist's experiences and may not be unanimous. In contrast, an automated disease screening system may have the potential to provide fast, consistent, and accurate results for diagnosis. Therefore, in this study, we aimed to develop an artificial intelligence system based on OCT images to aid the ophthalmologists for retinal disease screening.

In recent years, deep learning technology has substantial development and applications in the field of medical imaging. Deep learning[8] uses multilayer neural networks including convolutional layers to gain high accuracy and powerful learning performance in image classification.[9] Many neural networks were developed to assess the histopathological images of breast cancer,[10] malignant mesothelioma,[11] and coronary artery fibrous plaque detection.[12] For OCT images, deep learning models have also been studied, especially in the field of macular vision function. For example, Lee et al.[13] combined OCT images and electronic medical records to predict the occurrence of macular lesions in OCT images. Hwang et al.[14] constructed a deep learning model to predict outcomes and suggest further treatments for age-related macular degeneration (AMD), a specific type of disease of the macula. In addition to image classification, deep learning models were also developed for contour detection and layer segmentation of OCT images. For example, Orlando et al.[15] used a segmentation model to identify photoreceptor alteration in macular diseases. Recently, researchers developed models that can both classify the diseases and highlight the lesion areas on the images. For example, Fang et al.[16] designed the Lesion-Aware network, a neural network with attention modules to classify and highlight macular lesions. Although recent developments turned the spotlight on the automatic diagnosis of various eye conditions with improved performance, the majority of studies only focused on disease classification without highlighting the lesion regions from OCT images, which may not provide sufficient information for diagnosis. In contrast, the interpretability of deep learning model is also important for ophthalmologists to exam and affirm the prediction results. Although some studies used class activation map (CAM)-based technology[17] to visualize the model attention, this method does not classify or label different lesion types if they coexist on the same image. To help with clinical diagnosis, a better way to interpret the prediction results via highlighting the lesion regions along with type classification is preferred.

To solve these problems, in this study, we harvested OCT images and labelled them using 10 different lesion types belonging to two disease categories. An intelligent system was developed to classify retinal diseases and detect the lesion types and regions. For better prediction performance, we used transfer learning and ensemble learning techniques to develop deep learning models. Transfer learning[18] applies knowledge learned in relevant tasks in novel domains with better performance and fewer required data points. One important benefit of using transfer learning is to enable the model development based on a limited number of samples. Ensemble learning[19] combines multiple classifiers to achieve better performance than a single classifier, by integrating prediction results from individual models to vote for a final prediction. Because an individual model may be biased and have its own learning limitation, the combination of models can decrease the chance of errors and improve the generalization ability and accuracy of the final prediction.[20] Furthermore, we also developed an object detection model based on YOLOv3[21] to predict the lesion types and locations. We believe this study provides a comprehensive assessment of prediction and detection for eye diseases using different models based on OCT images.

Materials and Methods

This study was approved by the Ethics Committee of the Fujian Provincial Hospital, Fuzhou, Fujian, China (K2019-05-035).

Data Source

The overall workflow of this study is shown in Figure 1. We obtained 37,620 OCT B-scan images from Fujian Provincial Hospital between September 2017 and April 2019. There were 482 images (1.3%) with unclear retinal structures owing to poor quality that were excluded, which decreased the sample size to 37,138 images (775 patients). All OCT images were harvested using the SPECTRALIS system (Heidelberg Engineering, Heidelberg, Germany) and exported in a PNG format with 768 × 496 resolution.

Image Labeling and Training

Patient identification information on the OCT images was removed before the image labeling process. Then all OCT scans were diagnosed by two certified ophthalmologists in an internally developed web application. If there was any disagreement on the image label between the two, the image was sent to a senior ophthalmologist to finalize the diagnosis. When an OCT B-scan image contained retinal pathological

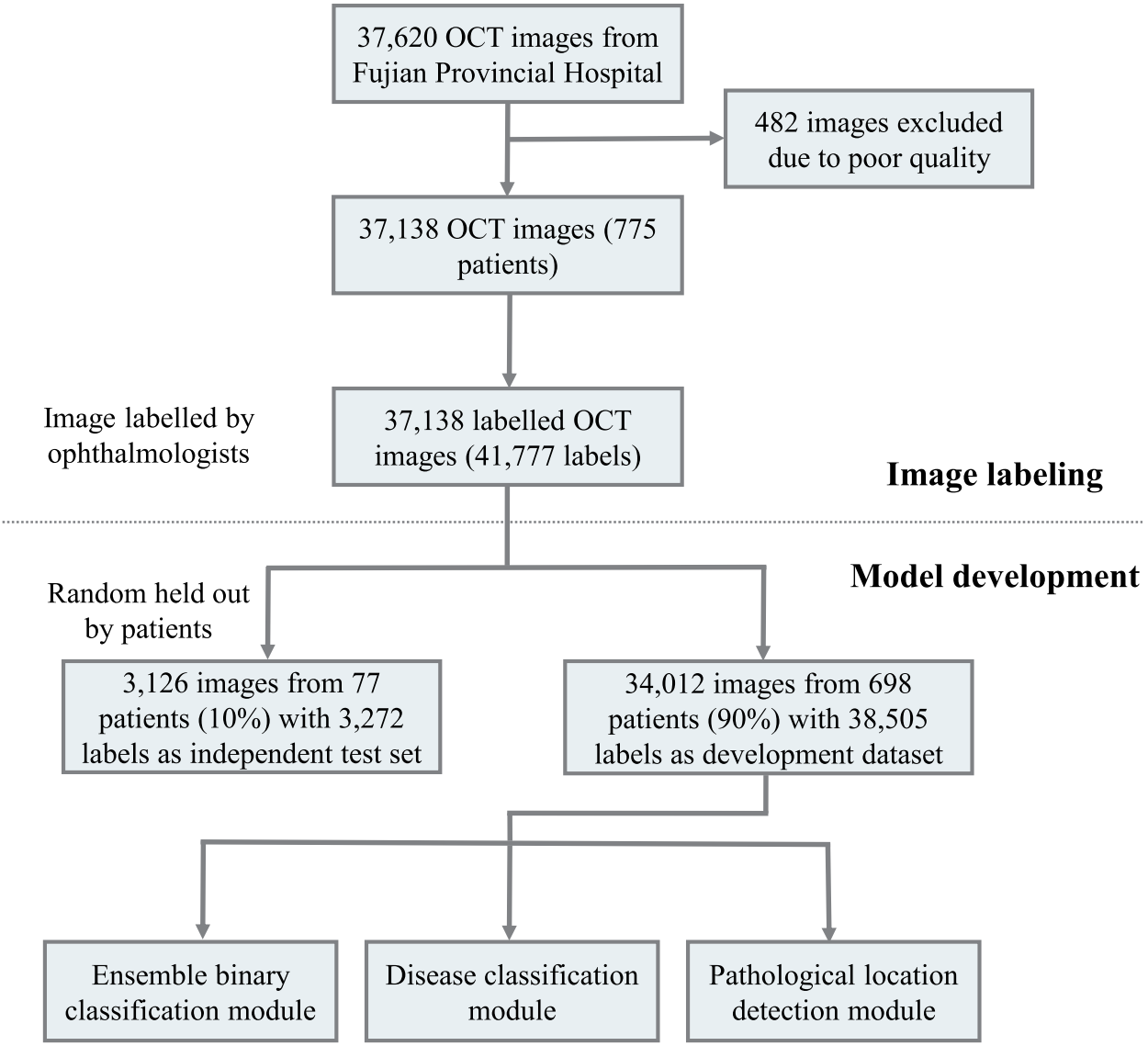

Figure 1. The flowchart of OCT image filtering, processing and separation for model development and test.

lesions, the ophthalmologists labelled both the type and location of the lesion on this B-scan with a rectangle box. Note that an OCT image may contain multiple types of retinal lesions (multiple labels) or no lesion at all (healthy control). We labelled 10 lesion types corresponding to two disease categories: (1) vitreomacular traction syndrome,[22] including lesions of cystoid macular edema, epiretinal membrane, full-thickness retinal eminence, macular pucker, and retinal thickness increased by detachment; and (2) AMD,[23] including lesions of atrophy of outer retina, choroid and retinal neovascularization, hemorrhage and exudation, pigment epithelial detachment, and retinal thickness increased by edema.

To create an independent test set for model evaluation, we randomly held out all images from 10% patients, which resulted in 3126 images from 77 patients (41.2% of these images were labeled with lesions). The rest images from the remaining 90% patients were randomly divided into training and validation sets using a 4:1 ratio. The training and validation sets were used to train and fine tune the models, and the independent test set was used to evaluate the model performance. Because the individual patients in the independent test set were different from those in the training or validation set, there was no information leakage when assessing the model performance on the independent test set.

Development of the Intelligent System

In this study, we developed an intelligent system that included three deep learning modules (Fig. 2). The first

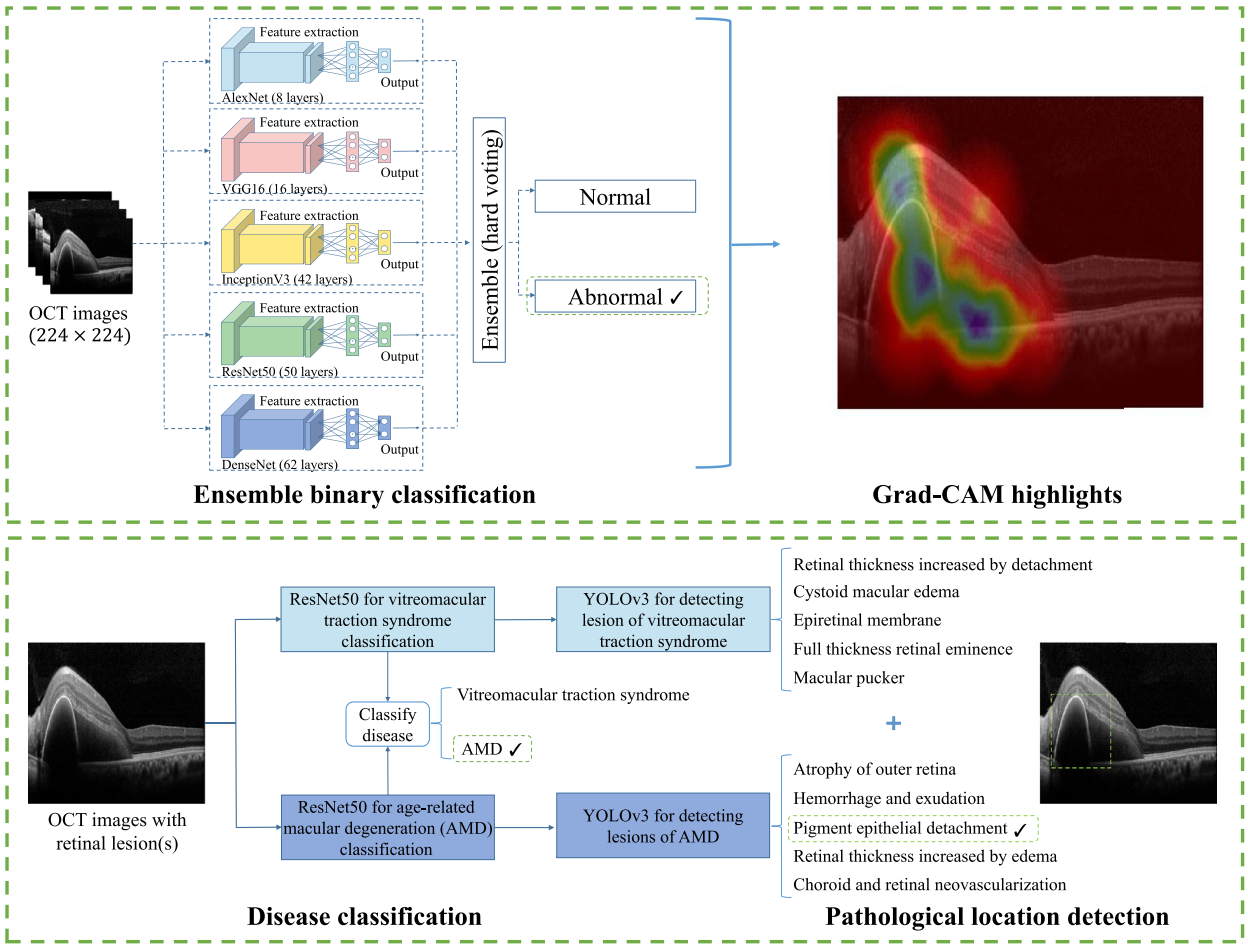

Figure 2. The overall framework of our deep learning-based system. It contains three modules: the ensemble binary classification module visualized by Grad-CAM, the disease classification module, and the pathology detection module.

part is an ensemble deep learning module of binary classifiers to identify OCT images as normal (healthy) versus abnormal (disease). The binary classifiers are five convolutional neural network–based deep learning models, namely, AlexNet,[24] DenseNet,[25] InceptionV3,[26] ResNet50,[27] and VGG16.[28] These deep learning classifiers include deep convolutional layers, pooling layers[29] (feature dimensionality reducer), and other functional parts such as dropout[30] and batch normalization[31] to extract and process complicated features from the image. Each of the classifiers was initiated with pretrained weights trained on ImageNet and then trained on our own data as a process of transfer learning, which enabled us to train the models on our dataset of a limited size. Ensemble learning was used to integrate all results from five models to vote for a final prediction to improve the generalization ability and prediction accuracy.[20] To visualize and explain these models, Grad-CAM[32] was used to generate CAMs for the test images, which highlighted the regions from the inputs that were important for the models to make predictions.

Because the binary classification module does not tell which disease or where lesion locates on an image, we further developed a multiclass multilabel disease classification module and a pathological location detection module. All OCT images identified as abnormal by the ensemble binary classification model were analyzed by the disease classification module to identify if they belong to any of the two disease categories, namely, vitreomacular traction syndrome and AMD. The disease classification module was developed based on ResNet50 model and trained using transfer learning like before. Furthermore, two YOLOv3 models[21] were developed for pathological location detection to identify the lesion types for vitreomacular traction syndrome and AMD, respectively. YOLOv3 consists of an image feature extraction module, DarkNet53, to extract feature maps and a detection module to locate the target object on the image. In this study, we

adopted the pretrained weights of DarkNet53 trained from ImageNet and trained the detection module with COCO dataset and adapt it on our own data. One image may contain more than one disease or pathology type, and the disease classification and pathological location detection modules were both designed for multiclass multilabel prediction (labels not exclusive). After the final prediction, all results generated by three different modules were combined for comprehensive image diagnoses.

We trained our deep learning models based on PyTorch platform and a server with two Intel Xeon Gold 6140 2.30GHz CPUs and four GeForce RTX 2080 Ti graphic cards. The learning rate was set to 0.001, the momentum was set to 0.9 and the batch size was set to 16. We applied early stopping to supervise model training and avoid overfitting.

Performance Evaluation

The prediction performance of the classifiers was evaluated using several different metrics including accuracy, sensitivity, specificity, F1 score and the area under the receiver operating characteristic curve. Additionally, the confusion matrices were provided as references. The pathological location detection module was assessed using recall and precision.

Results

The statistics of the patient demographics and labels of diseases and lesion types were listed in Table 1. Of all 37,138 OCT images from 775 patients, 13,106 images (35.3%) from 370 patients (47.7%) were labelled with

Table 1. Statistics of the OCT Image Data

Type	Development Dataset	Independent Test Set	Total
Patient information			
Patients	698	77	775
Female (%)	401 (57.4)	40 (51.9)	441 (56.9)
Images	34,012	3126	37,138
Labels	38,505	3272	41,777
Label information			
Vitreomacular traction syndrome			
Macular pucker (%)	6434 (16.8)	546 (15.2)	6980 (16.7)
Retinal thickness increased by detachment (%)	1262 (3.3)	22 (0.6)	1284 (3.1)
Epiretinal membrane (%)	498 (1.3)	23 (0.6)	521 (1.2)
Full-thickness retinal eminence (%)	495 (1.3)	74 (2.1)	569 (1.4)
Cystoid macular edema (%)	426 (1.1)	88 (2.5)	514 (1.2)
AMD			
Hemorrhage and exudation (%)	3283 (8.6)	154 (4.3)	3437 (8.2)
Retinal thickness increased by edema (%)	2067 (5.4)	74 (2.1)	2141 (5.1)
Choroid and retinal neovascularization (%)	1261 (3.3)	33 (1.0)	1294 (3.1)
Pigment epithelial detachment (%)	500 (1.3)	102 (2.8)	602 (1.4)
Atrophy of outer retina (%)	354 (0.9)	49 (1.4)	403 (1.0)
Normal (%)	21,925 (57.4)	2107 (58.8)	24,032 (57.5)

Table 2. Performance of the Deep Learning Binary Classifiers on the Independent Test Set

Model	No. of Layers	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC (%)
AlexNet	8	95.6	95.8	95.6	93.5	96.1
VGG16	16	95.9	97.7	95.0	93.9	96.1
InceptionV3	42	94.2	95.6	93.6	91.5	94.8
ResNet50	50	94.2	94.0	94.3	91.3	94.8
DenseNet	62	94.9	97.7	93.5	92.6	95.7
Ensemble model		98.5	98.7	98.4	97.7	98.1

ACU, area under the receiver operating characteristic curve.

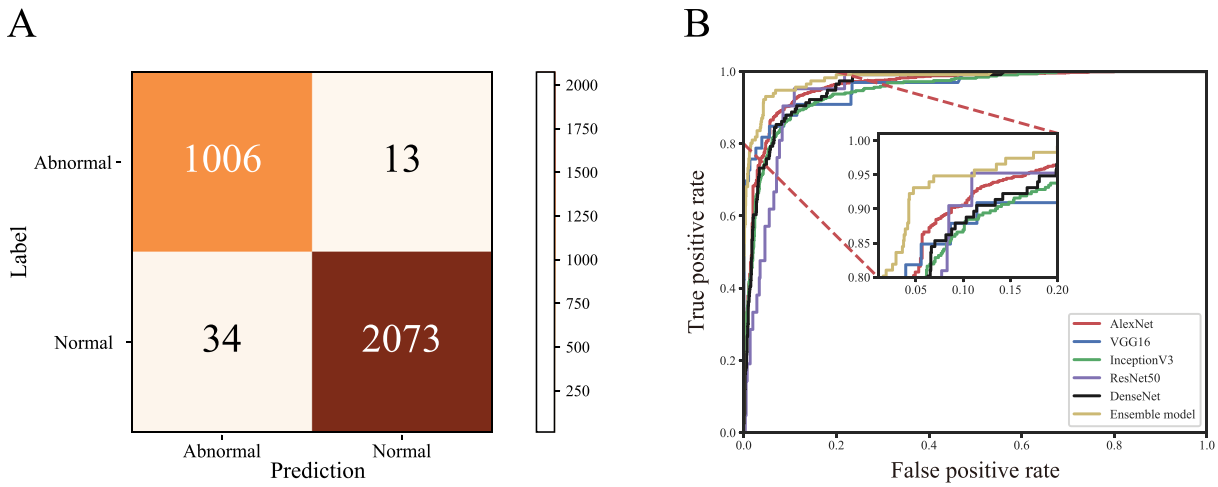

Figure 3. Performance evaluation of the binary classification module. (A) Confusion matrix of the ensemble binary classification module. (B) Receiver operating characteristic curve of individual classifiers and the ensemble classifier for binary classification.

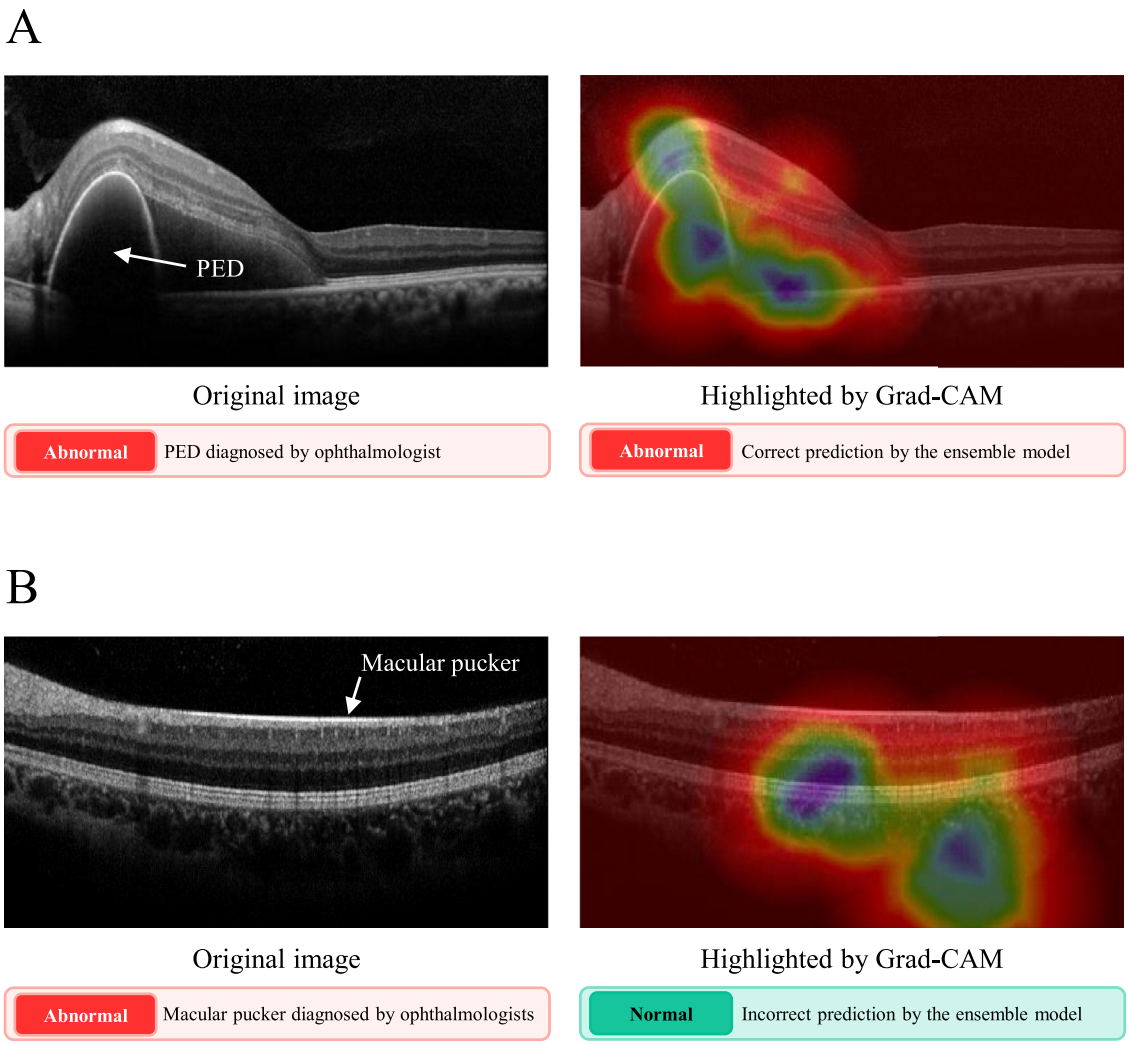

Figure 4. Example heatmap visualization of OCT images by Grad-CAM. (A) An example of the correct classification and highlight of the labeled pigment epithelial detachment pathology (PED) in an OCT image. (B) An example showing incorrect classification and prediction for an OCT image with the macular pucker pathology.

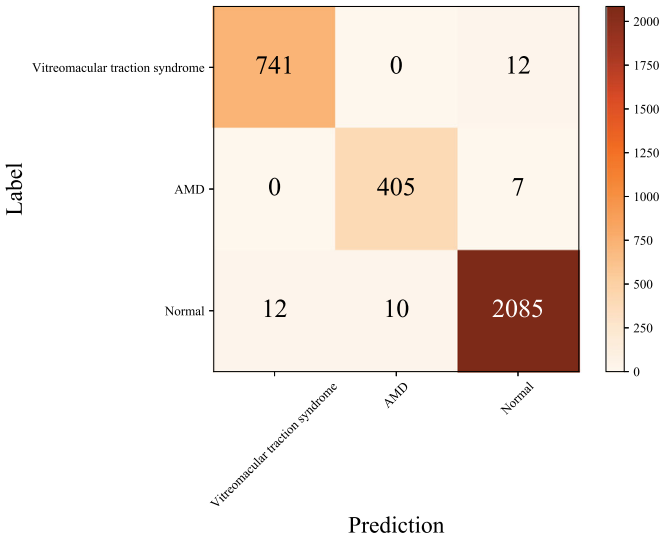

Figure 5. The confusion matrix of the disease classification module.

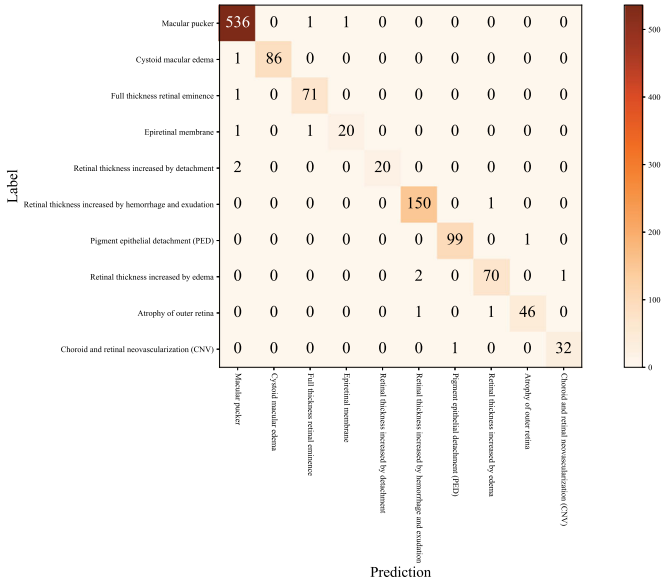

Figure 6. The confusion matrix of the pathology detection module.

lesions and the rest were normal. The most common type of lesion, macular pucker, was found 6980 times (16.7%), and the least common type of lesion, atrophy of outer retina, was seen 403 times (1.0%).

The performance of the binary classification module on the independent test set (3126 OCT images from 77 patients) is shown in Table 2, which separated disease images from healthy controls. The ensemble model of five models achieved an area under the receiver operating characteristic curve of 98.1%, accuracy of 98.5%, sensitivity of 98.7%, specificity of 98.4%, and F1 score of 97.7%, which is better than any individual model alone for all metrics. The confusion matrix of the ensemble model and the receiver operating characteristic curves of all models were presented in Figure 3. The confusion matrix indicated a total of 47 incorrect cases of 34 false positives and 13 false negatives were found among the 3126 predictions. The 13 false negatives included 8 examples of macular pucker, 2 examples of hemorrhage and exudation, 1 example of cystoid macular edema and 1 example of pigment epithelial detachment, and 1 example of retinal thickness increased by edema.

To interpret how the deep learning model make predictions, Grad-CAM was used to provide visual explanations for the neural network models. It generated heatmap to overlay on the OCT image and highlight spots where the model paid attention. Two examples of correct and incorrect predictions are shown in Figure 4. In Figure 4A, the model assigned higher weights in the correct lesion region on the OCT image, whereas in Figure 4B the model made false negative-predictions and Grad-CAM highlighted regions that were supposed to be healthy and normal.

The confusion matrix and performance metrics for the disease classification module on the independent test set are shown in Figure 5 and Table 3, which identified if the abnormal sample belongs to the disease category of vitreomacular traction syndrome or AMD. The prediction performance for vitreomacular traction syndrome was accuracy of 99.3%, sensitivity of 98.4%, specificity of 99.5%, and an F1 score of 98.4%; the detection for AMD obtained an accuracy of 99.5%, sensitivity of 98.3%, specificity of 98.3%, and F1 score of 97.9%. The combined accuracy of the binary classification along with disease classification modules is 98.7%.

Two YOLOv3 models were developed to detect the pathology types and locations for vitreomacular traction syndrome and AMD. The confusion matrix and performance metrics of the pathology detection on the independent test set were shown in Figure 6

Table 3. Performance of the Disease Classification Module on the Independent Test Set

	Vitreomacular Traction Syndrome				AMD Findings				
Evaluation Metrics	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	Overall Accuracy
Performance	99.3	98.4	99.5	98.4	99.5	98.3	99.7	97.9	98.7

Table 4. Performance of the Pathology Detection Module on the Independent Test Set Evaluated by the Pathology Level and Patient Level

	Pathology Based				Patient Based		
Pathology	No. of Labels	Percent/Total Labels (%)	Recall (%)	Precision (%)	No. of Individuals	Percent/Total Individuals (%)	Recall (%)
Vitreomacular traction syndrome							
Macular pucker	546	46.9	98.2	98.2	33	42.9	100
Cystoid macular edema	88	8.6	97.7	93.5	7	9.1	100
Full-thickness retinal eminence	74	6.4	95.9	92.2	10	13.0	100
Epiretinal membrane	23	2.0	87.0	67.0	2	2.6	100
Retinal thickness increased by detachment	22	1.9	90.9	74.1	5	6.5	100
AMD							
Retinal thickness increased by hemorrhage and exudation	154	13.2	97.4	94.3	8	10.4	100
Pigment epithelial detachment	102	8.8	97.1	95.2	10	13.0	100
Retinal thickness increased by edema	74	6.4	94.6	90.9	2	2.6	100
Atrophy of outer retina	49	4.2	93.9	85.2	4	5.2	100
Choroid and retinal neovascularization	33	2.8	97.0	80.0	5	6.5	100

and Table 4. High recall values were found for all types of pathology detections, ranging from 87.0% (epiretinal membrane) to 98.2% (macular pucker).

Discussion

Factors such as aging and unhealthy lifestyles can increase the risks of visual impairment, which may ultimately cause permanent vision loss and impact life quality.[33] As a result, early detection, interference, and prevention are critical. However, owing to insufficient numbers and uneven allocations of medical resources and ophthalmologists, many patients with retinal diseases cannot received timely diagnoses and treatments.[34] In recent years, the development of deep learning and artificial intelligence may provide a promising solution. Studies have shown that artificial intelligence systems have achieved excellent results in image classification and segmentation.[35–37] However, owing to the limited number of high-quality labelled OCT image data and concerns over prediction accuracy, artificial intelligence–based diagnosis has not yet been widely used in the real-world clinical environment.

In this study, we collected 37,138 OCT images and labelled them one by one with two disease categories and 10 types of retinal pathologies. Three submodules—the ensemble binary classification module, the disease classification module, and the pathology detection module—were developed to identify images of diseases, classify the disease types, and highlight the types and regions of pathologies. Because most of our training samples are healthy controls (64.7% images), we developed this three-module system instead of a one-step model to mitigate the impact from the negatives and leverage the advantage of ensemble classification.

In the first binary classification module, the ensemble learning model was used to combine the results from five individual models. Although requiring more time for training and a more complicated model architecture, the ensemble model outperformed all individual models. It is worth mentioning that the individual models with fewer layers exhibited a better performance than those with more layers. This finding is expected, because models with greater complexity are more likely to overfit on a smaller dataset.[38,39] To mitigate this problem, transfer learning and ensemble learning were leveraged to decrease the influence of overfitting, bring beneficial effects, and obtain better performance than individual models, as shown in previous research.[40–42] Additionally, the model attention at the lesion regions was visualized and highlighted, which provided important references for the ophthalmologists. The second disease classification

module classified OCT images into two retinal disease categories, namely, vitreomacular traction syndrome and AMD. Although this strategy is a further expansion of prediction details, the model still maintained high accuracy and relatively balanced false positives and false negatives. The last pathology detection module additionally extended the prediction categories into 10 types of pathologies. Although it was the most detailed analysis and prediction, owing to the uneven distributions of pathology types, the performance decreased compared with previous modules and varied across different pathologies. However, its accuracy was still at a good level, ranging from 87.0% to 98.2%, which may still have clinical implications. Our model contained three modules as a multistage deep learning system, as prediction at each stage could provide specialized analytical information about the OCT images. To test if the multistage deep learning system is better than a single combined model, we developed a single ResNet50 model for direct disease classification without the previous binary classification module. The results were shown in Supplementary Table 1 and Supplementary Figure 1, which were worse compared to Table 3 and Figure 5, respectively. We also developed a single YOLOv3 model for direct location detection of 10 retinal pathologies without the previous classification modules. The performance decreased as shown in Supplementary Table 2 compared with Table 4. We believe in our multistage system, each module was optimized its own specialized task, the influence of data imbalance was decreased at each step, and the overall performance was better.

Recent studies using deep learning for retinal disease diagnosis aimed at achieving high accuracy with interpretability. For example, Dong et al.[36] collected ocular fundus images of 10 retinal diseases and grouped them into three groups for deep learning classification; Sunija et al.[43] developed a deep learning model for the classification of choroidal neovascularization, diabetic macular edema, drusen and normal OCT images. Cen et al.[44] used deep neural networks to detect 39 fundus diseases and conditions and Lee et al.[45] used deep learning to screen fundus abnormalities including AMD, diabetic retinopathy, epiretinal membrane, retinal vascular occlusion, and suspected glaucoma. However, these studies only used Grad-CAM heatmaps to interpret the model, which may not be as informative as the pathology detection approach. Furthermore, Wang et al.[46] combined pathological and thickness features to detect 15 types of pathological locations from OCT images via deep learning, which has more types of pathologies compared with our work. However, our multistage system of classification and lesion detection on a limited dataset can still provide important validation and complement to the existing research.

There are several limitations to our study. First, our OCT data were collected from just one hospital, which may only represent a regional and small patient population. Additionally, the sample size may be insufficient and unevenly distributed among different retinal pathologies. The system was designed to recognize some major types of retinal problems to aid the ophthalmologists on regular tasks. Given the data and model limitation, the model performance on new and unseen types of lesions or diseases should be questioned. It still requires the assistance of the ophthalmologists to finalize the diagnosis. As an important future direction, it is very important to increase the number of patients, sources, and images to decrease model biases and improve its clinical applicability. Second, the pathological locations within our samples were labelled with rectangles, while retinal layer segmentation for pathological detection[47] has also been used and may be more reliable for specific lesion detection. Although the rectangular labeling requires less effort, it is a future direction to perform segmentation of the retinal layers for better and more reliable detection of particular lesions. Third, our system has not been tested in real-world clinical diagnosis. Although the dry experiments indicated great model performance, the clinical translatability has not been assessed. We believe this system may aid and provide helpful references for the ophthalmologists for diagnosis. However, the system is not meant to replace ophthalmologists and expert decisions are needed to finalize the diagnosis. Fourth, OCT images are only one dimension of clinical diagnosis, other clinical data including medical history and other examinations are also necessary for high-quality diagnosis and treatment recommendation. It is yet a future direction to integrate electronic medical records with OCT images for a more comprehensive diagnosis and recommendation system.

Conclusions

In this work, 37,138 OCT images from 775 patients were harvested and 10 pathology types corresponding to two disease types, vitreomacular traction syndrome and AMD, were labelled by ophthalmologists. An intelligent multistage system was developed for disease screening and pathology detection from OCT images. The system contained a binary classification module, a disease classification module, and a pathology location detection module to analyze the OCT image with different levels of granularity, identify the disease and

pathology types and locate the regions. All three modules showed good performance when evaluated on the independent test set. We demonstrated this system is capable to identify and predict eye diseases from the OCT images and may have potential clinical application to assist the ophthalmologists on decision-making.

Acknowledgments

The authors thank Zhaofan Xia for her help in supporting laboratory equipment.

Supported by National Natural Science Foundation of China (81971837), Natural Science Foundation of Fujian Province, China (2019J01507), and Funds of Scientific Research-Support Project, Fujian Provincial Department of Finance (83020008).

Disclosure: **X, Chen**, None; **Y. Xue**, None; **X. Wu**, None; **Y. Zhong**, None; **H. Rao**, None; **H. Luo**, None; **Z. Weng**, None

* XC and YX contributed equally to this work.

References

1. Bashshur R, Ross C. World report on vision. *Int J Eye Banking*. 2020;8(3):1–4.
2. Bourne RR, Stevens GA, White RA, et al. Causes of vision loss worldwide, 1990–2010: a systematic analysis. *Lancet Global Health*. 2013;1(6):e339–e349.
3. Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science*. 1991;254(5035):1178–1181.
4. Hee, MR, Baumal, CR, Puliafito, CA, et al. Optical coherence tomography of age-related macular degeneration and choroidal neovascularization. *Ophthalmology*. 1996;103(8):1260–1270.
5. Puliafito CA, Hee MR, Lin CP, et al. Imaging of macular diseases with optical coherence tomography. *Ophthalmology*. 1995;102(2):217–229.
6. Srinivasan VJ, Wojtkowski M, Witkin AJ, et al. High-definition and 3-dimensional imaging of macular pathologies with high-speed ultrahigh-resolution optical coherence tomography. *Ophthalmology*. 2006;113(11):2054–2065.e2053.
7. Nassif N, Cense B, Park B, et al. In vivo high-resolution video-rate spectral-domain optical coherence tomography of the human retina and optic nerve. *Optics Express*. 2004;12(3):367–376.
8. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–444.
9. Kattenborn T, Leitloff J, Schiefer F, Hinz S. Review on convolutional neural networks (CNN) in vegetation remote sensing. *ISPRS J Photogram Remote Sensing*. 2021;173:24–49.
10. Sohail A, Mukhtar MA, Khan A, Zafar MM, Khan S. Deep object detection based mitosis analysis in breast cancer histopathological images. 2020. arXiv:2003.08803.
11. Courtiol P, Maussion C, Moarii M, Pronier E, Clozel T. Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nat Med*. 2019;25(10):1519–1525.
12. Liu X, Du J, Yang J, Xiong P, Lin F. Coronary artery fibrous plaque detection based on multi-scale convolutional neural networks. *J Signal Processing Syst*. 2020;95(1–3):325–333.
13. Lee CS, Baughman DM, Lee AY. Deep learning is effective for classifying normal versus age-related macular degeneration OCT images. *Ophthalmol Retina*. 2017;1:322–327.
14. Hwang DK, Hsu CC, Chang KJ, et al. Artificial intelligence-based decision-making for age-related macular degeneration. *Theranostics*. 2019;9(1):232–245.
15. Orlando JI, Gerendas BS, Riedl S, Grechenig C, Schmidt-Erfurth U. Automated quantification of photoreceptor alteration in macular disease using optical coherence tomography and deep learning. *Sci Rep*. 2020;10(1):5619.
16. Fang L, Wang C, Li S, Rabbani H, Chen X, Liu Z. Attention to lesion: lesion-aware convolutional neural network for retinal optical coherence tomography image classification. *IEEE Trans Med Imaging*. 2019;38(8):1959–1970.
17. Zhou B, Khosla A, Lapedriza A, et al. Learning deep features for discriminative localization[C]. *Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition*. 2016:2921–2929.
18. Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. A Survey on deep transfer learning. In: *International Conference on Artificial Neural Networks*; Rhodes, Greece, October 4–7, 2018.
19. Liu Y, Yao X. Ensemble learning via negative correlation. *Neural Networks*. 1999;12(10):1399–1404.
20. Sagi O, Rokach L. Ensemble learning: a survey. *WIREs Data Mining and Knowledge Discovery*. 2018;8(4):e1249.
21. Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. *arXiv preprint* arXiv:1804.02767. 2018.

22. Bottós J, Elizalde J, Arevalo JF, Rodrigues EB, Maia M. Vitreomacular traction syndrome. *J Ophthalmic Vis Res*. 2012;7(2):148.
23. Nowak JZ. Age-related macular degeneration (AMD): pathogenesis and therapy. *Pharmacol Rep*. 2006;58(3):353.
24. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks[J]. *Communications of the ACM*. 2017;60(6);84–90.
25. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; Honolulu, Hawaii, July 21–26, 2017;4700–4708.
26. Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016;2818–2826.
27. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; Las Vegas, Nevada, June 27–30, 2016;770–778.
28. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. *arXiv preprint* arXiv:1409.1556. 2014.
29. Sun M, Song Z, Jiang X, Pan J, Pang Y. Learning pooling for convolutional neural network. *Neurocomputing*. 2017;224:96–104.
30. Baldi P, Sadowski PJ. Understanding dropout[J]. *Advances in Neural Information Processing Systems*. 2013;26.
31. Bjorck N, Gomes CP, Selman B, et al. Understanding batch normalization[J]. *Advances in Neural Information Processing Systems*. 2018;31.
32. Selvaraju RR, Cogswell M, Das A, et al. Gradcam: Visual explanations from deep networks via gradient-based localization[C]. *Proceedings of the IEEE International Conference on Computer Vision*. 2017;618–626.
33. Mabaso RG, Oduntan OA. Risk factors for visual impairment and blindness amongst black adult diabetics receiving treatment at Government healthcare facilities in Mopani District, Limpopo province, South Africa. *African Journal of Primary Health Care and Family Medicine*. 2014;6(1):1–8.
34. An L, Jan CL, Feng J, Wang Z, Zhan L, Xu X. Inequity in access: cataract surgery throughput of Chinese ophthalmologists from the China national eye care capacity and resource survey. *Ophthalmic Epidemiology*. 2020;27(1):29–38.
35. Lu W, Tong Y, Yu Y, Xing Y, Chen C, Shen Y. Deep learning-based automated classification of multi-categorical abnormalities from optical coherence tomography images. *Transl Vis Sci Technol*. 2018;7(6):41–41.
36. Dong L, He W, Zhang R, et al. Artificial intelligence for screening of multiple retinal and optic nerve diseases. *JAMA Network Open*. 2022;5(5):e229960–e229960.
37. Li Q, Li S, He Z, et al. DeepRetina: layer segmentation of retina in OCT images using deep learning. *Transl Vis Sci Technol*. 2020;9(2):61–61.
38. Lever J, Krzywinski M, Altman N. Model selection and overfitting. *Nat Methods*. 2016;13(9):703–704.
39. Ying X. An overview of overfitting and its solutions. *J Physics Conference Series*. 2019;2:1168.
40. Rajaraman S, Jaeger S, Antani SK. Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. *PeerJ*. 2019;7:e6977.
41. Barhate N, Bhave S, Bhise R, Sutar RG, Karia DC. Reducing overfitting in diabetic retinopathy detection using transfer learning. In: *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*; Greater Noida, UP, India, October 30–31, 2020;298–301.
42. Le D, Alam M, Yao CK, et al. Transfer learning for automated OCTA detection of diabetic retinopathy. *Transl Vis Sci Technol*. 2020;9(2):35.
43. Sunija AP, Kar S, Gayathri S, Gopi VP, Palanisamy P. OctNET: a lightweight CNN for retinal disease classification from optical coherence tomography images. *Computer Methods and Programs in Biomedicine*. 2021;200:105822.
44. Cen LP, Ji J, Lin JW, et al. Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. *Nat Commun*. 2021;12(1):4828.
45. Lee J, Lee J, Cho S, et al. Development of decision support software for deep learning-based automated retinal disease screening using relatively limited fundus photograph data. *Electronics*. 2021;10(2):163.
46. Wang L, Wang G, Zhang M, et al. An intelligent optical coherence tomography-based system for pathological retinal cases identification and urgent referrals. *Transl Vis Sci Technol*. 2020;9(2):46.
47. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med*. 2018;24(9):1342–1350.