EPMA Journal (2024) 15:39–51
https://doi.org/10.1007/s13167-024-00350-y

RESEARCH

Automated detection of nine infantile fundus diseases and conditions in retinal images using a deep learning system

Yaling Liu[1] · Hai Xie[2] · Xinyu Zhao[1] · Jiannan Tang[1] · Zhen Yu[1] · Zhenquan Wu[1] · Ruyin Tian[1] · Yi Chen[1,3] · Miaohong Chen[1,3] · Dimitrios P. Ntentakis[4] · Yueshanyi Du[1] · Tingyi Chen[1,3] · Yarou Hu[1] · Sifan Zhang[3,5] · Baiying Lei[2] · Guoming Zhang[1,3]

Received: 1 September 2023 / Accepted: 21 January 2024 / Published online: 15 February 2024
© The Author(s), under exclusive licence to European Association for Predictive, Preventive and Personalised Medicine (EPMA) 2024

Abstract

Purpose We developed an Infant Retinal Intelligent Diagnosis System (IRIDS), an automated system to aid early diagnosis and monitoring of infantile fundus diseases and health conditions to satisfy urgent needs of ophthalmologists.

Methods We developed IRIDS by combining convolutional neural networks and transformer structures, using a dataset of 7697 retinal images (1089 infants) from four hospitals. It identifies nine fundus diseases and conditions, namely, retinopathy of prematurity (ROP) (mild ROP, moderate ROP, and severe ROP), retinoblastoma (RB), retinitis pigmentosa (RP), Coats disease, coloboma of the choroid, congenital retinal fold (CRF), and normal. IRIDS also includes depth attention modules, ResNet-18 (Res-18), and Multi-Axis Vision Transformer (MaxViT). Performance was compared to that of ophthalmologists using 450 retinal images. The IRIDS employed a five-fold cross-validation approach to generate the classification results.

Results Several baseline models achieved the following metrics: accuracy, precision, recall, F1-score (F1), kappa, and area under the receiver operating characteristic curve (AUC) with best values of 94.62% (95% CI, 94.34%-94.90%), 94.07% (95% CI, 93.32%-94.82%), 90.56% (95% CI, 88.64%-92.48%), 92.34% (95% CI, 91.87%-92.81%), 91.15% (95% CI, 90.37%-91.93%), and 99.08% (95% CI, 99.07%-99.09%), respectively. In comparison, IRIDS showed promising results compared to ophthalmologists, demonstrating an average accuracy, precision, recall, F1, kappa, and AUC of 96.45% (95% CI, 96.37%-96.53%), 95.86% (95% CI, 94.56%-97.16%), 94.37% (95% CI, 93.95%-94.79%), 95.03% (95% CI, 94.45%-95.61%), 94.43% (95% CI, 93.96%-94.90%), and 99.51% (95% CI, 99.51%-99.51%), respectively, in multi-label classification on the test dataset, utilizing the Res-18 and MaxViT models. These results suggest that, particularly in terms of AUC, IRIDS achieved performance that warrants further investigation for the detection of retinal abnormalities.

Conclusions IRIDS identifies nine infantile fundus diseases and conditions accurately. It may aid non-ophthalmologist personnel in underserved areas in infantile fundus disease screening. Thus, preventing severe complications. The IRIDS serves as an example of artificial intelligence integration into ophthalmology to achieve better outcomes in predictive, preventive, and personalized medicine (PPPM / 3PM) in the treatment of infantile fundus diseases.

Keywords Deep learning; infant · Fundus disease · Retinal image · Predictive preventive personalized medicine (PPPM / 3PM)

Abbreviations

AI	Artificial intelligence
AUC	Area under the receiver operating characteristic curve
CNNs	Convolutional neural networks
CRF	Congenital retinal fold
DA	Depth attention
DL	Deep Learning
EPMA	European Association for Predictive, Preventive and Personalized Medicine
eps	Epsilon
F1	F1 score
IRIDS	Infant Retinal Intelligent Diagnosis System
Max ViT	Multi-Axis Vision Transformer
PPPM/3PM	Predictive, preventive, and personalized medicine
RB	Retinoblastoma

Yaling Liu, Hai Xie and Xinyu Zhao contributed equally and are considered co–first authors.

Sifan Zhang, Baiying Lei, and Guoming Zhang contributed equally to this work.

Extended author information available on the last page of the article

ROC	Receiver operating characteristic
ROP	Retinopathy of prematurity
RP	Retinitis pigmentosa
Res-18	ResNet-18
SGD	Stochastic gradient descent
T-SNE	T-distributed stochastic neighbor embedding

Introduction

Early diagnosis and treatment of infantile fundus diseases and conditions are key factors in improving the quality of life for infants and their families. Common infantile fundus diseases include retinopathy of prematurity (ROP) [1], Coats disease [2, 3], retinoblastoma (RB) [4], retinitis pigmentosa (RP) [5], coloboma of the choroid [6], congenital retinal fold (CRF) [7] and familial exudative vitreoretinopathy [8]. Delayed diagnosis and management of these conditions can lead to reduced treatment effectiveness and the potential for severe, sometimes irreversible structural and functional complications, such as refractive errors, night blindness, strabismus, and neovascular glaucoma [2, 9]. Identifying high-risk individuals is a vital step in preventing these serious complications, as emphasized in the European Association for Predictive, Preventive and Personalized Medicine (EPMA) white paper [10].

ROP, a typical example of an infantile fundus abnormality, is a leading cause of visual impairment and blindness in early life, requiring timely diagnosis and intervention [11]. Approximately 8000 new infants of infantile RB worldwide require enucleation surgery every year [12–14]. Due to the relative rarity of some fundus diseases in infants, even when encountered, making an accurate diagnosis in certain children's hospitals, and maternal and child health hospitals, may be challenging. This can have detrimental consequences not only for individuals and their families but also for society at large. However, numerous underserved areas worldwide lack access to experienced ophthalmologists [15], placing infants with fundus diseases at immediate risk. It is critical to develop a unique strategy that adheres to the PPPM/3PM principles for early detection of infants with fundus diseases [16]. Therefore, the latest technological innovations in digital computing can be applied to develop an effective automated diagnostic tool for the early diagnosis and management of infantile fundus diseases.

Deep learning (DL) is a mature yet continuously evolving technology, especially in the realm of computer-aided diagnosis of human diseases [17–19]. This frontier continues to expand into other areas of medicine, such as clinical practice, translational medical research, and basic biomedical research [20–24]. DL has demonstrated notable performances in automating the screening and diagnosis of various fundus diseases, some of these applications are already progressing towards clinical implementation [25–29]. Regarding infantile fundus diseases, our team successfully used DL to identify ROP, zone I of ROP, and A-ROP, effectively applying augmentation algorithms for retinal grading [30–34]. Most relevant studies have focused on detecting a single fundus disease using retinal images. Recently, DL-assisted systems have been developed to detect multiple fundus diseases [35, 36], including one that utilized a smartphone-based wide-field retinal imaging system for classifying multiple diseases in children [37]. It's worth noting that the above-mentioned studies have primarily focused on single fundus diseases or examined retinal images of adults and older children.

Working hypothesis and purpose of the study

No DL studies involving the detection of multiple infantile fundus diseases have been conducted thus far. In clinical practice, especially in remote areas lacking specialized ophthalmologists, such a multi-disease detection system for infantile fundus diseases would not only be necessary but also potentially highly beneficial. Many infantile fundus diseases are difficult for parents to detect, and they are often identified when a child seeks treatment for conditions such as amblyopia at a later age. Moreover, the diagnosis of infantile fundus diseases is challenging, leading to an increased incidence of severe complications. We hypothesized that utilizing DL models for automatic detection of infantile fundus diseases through retinal images may improve the diagnostic rate in infants, prevent the occurrence of severe complications, and ultimately enhance their quality of life. We believe that this approach may be effectively applied in clinical settings, providing more accurate and reliable tools for screening and managing infantile fundus diseases.

To achieve this, we developed an automated disease detection system for infantile fundus diseases called the Infantile Retinal Intelligence Diagnosis System (IRIDS). It can classify nine common infantile fundus diseases and conditions using 7679 retinal images (1089 infants) collected from hospitals throughout China. We anticipate significant benefits in improving individual outcomes for preventable infantile fundus diseases, accompanied by a positive cost-effectiveness impact on advanced medical services for the population. This includes the use of innovative DL screening technologies that utilize predictive disease modeling and treatment algorithms tailored to each patient's personalized profile [38].

Methods

Datasets and labeling

Retinal images from four diverse data sources were collected to develop and validate the IRIDS. The primary dataset for training, validation, and testing was collected from the Department of Ophthalmology, Shenzhen Eye Hospital, Shenzhen, Guangdong, China, between April 2004 and January 2021. A smaller portion of the data came from three other maternal and child health hospitals in Guangdong Province. This diagnostic study was approved by the Ethics Committee of Shenzhen Eye Hospital. All institutions abided by the tenets of the Declaration of Helsinki [39], and written informed consent was obtained from the parents of all enrolled infants.

The data acquisition and processing methods are shown in Fig. 1A and Supplementary Fig. 2. All preoperative images collected from the various data sources underwent quality screening by two junior ophthalmologists. Initially, we selected nine prevalent categories based on their high

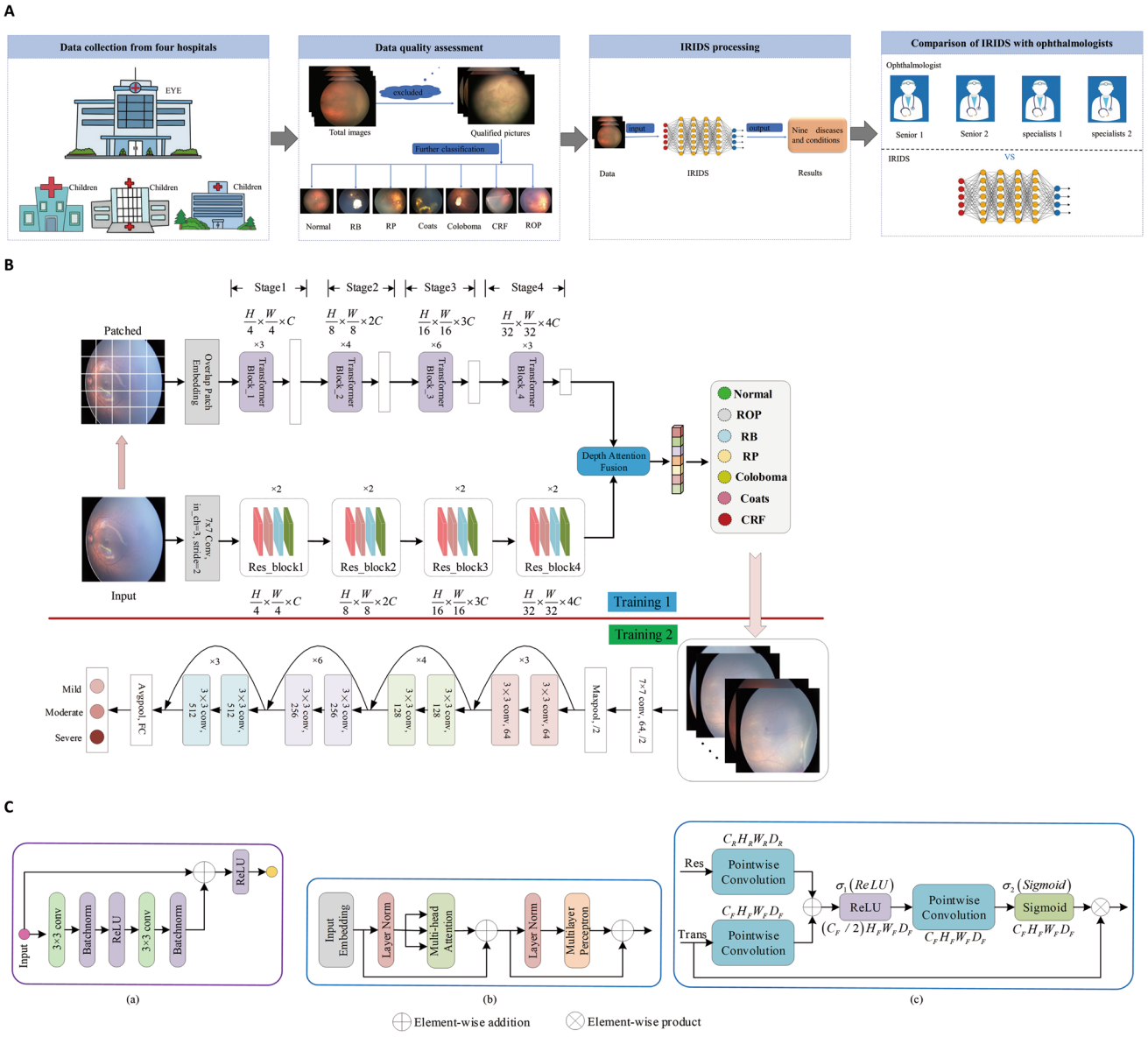


Fig. 1 Workflow of the overall study and architecture of the IRIDS. **A.** Overview of the study protocol. Images were collected from four clinical centers and filtered by quality assessment. Subsequently, the data was processed by IRIDS, and finally, an ophthalmologists-IRIDS comparison was conducted. **B**. The architecture of the IRIDS. The IRIDS consists of two branches: CNNs and Transformers. The model employs four-stage Transformer modules to extract global features, and the ResNet-34 with residual blocks is selected to extract local features. The extracted features are fused by the depth attention module to fully represent the specific features related to different diseases. **C.** Illustration of the used blocks. (a)-(c) represent the module of the Res_block, Transformer block, and Depth attention fusion module, respectively

incidence and substantial clinical data availability in our database. This ensured statistically robust and meaningful analysis while maintaining analytical validity by excluding categories that were less common or had inadequate data. To maintain a balance among different categories, not all the images were included in the database. For example, if there were too many ROP images, only a representative portion was included to ensure data diversity while still ensuring the total number of images was adequate. In the second phase, the included images were discarded if they were unclear, too dark, or too bright or if the image quality was poor. This rigorous selection process led to the exclusion of 9500 images. Consequently, a final dataset comprising 7697 images (1089 infants) was included for the study. The images included in our study were independently diagnosed (labeled) by two trained junior ophthalmologists. In our study, 'label' refers to the specific diagnosis assigned to each image. In instances of discrepancies between the labels provided by these two junior ophthalmologists, we sought the evaluation of a senior ophthalmologist to make the final say. The images with unanimous labels from the two junior ophthalmologists or recognized by the senior ophthalmologist were included in our study. This stringent selection process was implemented to ensure the accuracy of classification by the IRIDS algorithms. Images with conflicting labels underwent review by a senior ophthalmologist, and in cases of uncertainty, the final label was determined by an experienced retinal specialist. Finally, images with agreed-upon labels were retained for training, validation, and test datasets for the automated system.

Classification

All retinal images were classified into nine primary classes/subclasses of diseases and conditions (Table 1 and Supplementary Fig. 1) according to fundus signs described in Pediatric Retina [40]. Our categorization approach was inspired and informed by the methodology used by Cen LP et al. [36] in their study of adult fundus diseases, adapted here for pediatric conditions. Common fundus diseases with distinct retinal characteristics recognizable on retinal images were classified as independent major classes. To account for real-world scenarios, we included normal retinal images. These major classes included normal, Coats' disease, choroid coloboma, RB, RP, CRF, and ROP.

Some diseases share similar characteristics and are not easily distinguishable solely through fundus images. For instance, the symptoms of infants with stage 1 or 2 ROP are milder than those in other stages. To address this, we clustered them together and classified them as "mild" ROP. Likewise, infants with stage 3 ROP are more likely to require treatment with laser photocoagulation or intravitreal injection of anti-vascular endothelial growth factor than infants with mild ROP; thus, we classified this stage as "moderate" ROP. Infants with stage 4 and 5 ROP typically require vitrectomy and/or other surgical treatments, so we clustered them together and classified as "severe" ROP.

Following this clustering, the nine types of diseases and conditions were divided into seven main classes and three subclasses. Each class was assigned a unique class identity number (ID) (Table 1).

Architecture of the IRIDS

The workflow and architecture of the IRIDS are shown in Fig. 1B and C. Convolutional neural networks (CNNs) mainly focus on local features, which leads to an inadequate representation of the global features, whereas the transformer network pays more attention to the extraction of global features. As a result, the classification performance was limited when only

Table 1 Brief descriptions of features in retinal images of diseases and conditions

ID	Diseases/conditions	Brief descriptions of retinal images
0	Normal	The optic disc was red with clear borders, a C/D ratio of approximately 0.2, a normal foveal reflex in the macular region, a flat retina, and normal retinal vascular morphology
1	Coats	Visible distortion of retinal vessels, irregular cystic dilatation or beading, retinal punctate or patchy hemorrhages, and massive yellow-white fatty exudates in the deep and subretinal layers
2	Coloboma	Defects are visible in the retina, with clear boundaries and different shapes and sizes. Retinal blood vessels can be seen crawling in the defect area, and white sclera is exposed
3	RB	Clear borders, single or multiple lesions, white or yellow nodular bulges, uneven surface, and different sizes
4	RP	Mid-peripheral RPE atrophy with bone-spicule perivascular pigmentation, arteriorlar attenuation and waxy disc pallor
5	CRF	Raised retinal folds that start from the optic disc and often extend to the surrounding retina
6	ROP	
6.1	Mild ROP	ROP stage 1 and stage 2, the demarcation line is relatively flat and white, lies within the plane of the retina; the ridge may vary in height and its color may appear to range from white to pink
6.2	Moderate ROP	ROP stage 3, extraretinal neovascularization extends from the ridge to the vitreous and is continuous with the posterior ridge, or proliferates
6.3	Severe ROP	ROP stage 4 and stage 5, retinal detachment, include partial retinal detachment and total retinal detachment

RB Retinoblastoma, *RP* retinitis pigmentosa, *CRF* congenital retinal fold, *ROP* retinopathy of prematurity

one structure was used to extract features from the analyzed images. Hence, we applied a hybrid model that combines CNNs and transformer structures, as shown in Fig. 1B. Res-Blocks were derived from the ResNet-18 (Res-18) model [41], which is represented as a residual block. A residual block is defined as shown in Fig. 1C (a). The transformer blocks used were the same as those of the Multi-Axis Vision Transformer (MaxViT) [42], whose main transformer structure is shown in Fig. 1C (b). After obtaining high-level features from the Res-blocks and Transformer blocks, we used the depth-attention fusion module to fuse the local and global features from the paths of the ResNet and MaxViT modules, as illustrated in Fig. 1C (c). Thus, the entire network can express features related to each specific disease and focus on extracting key features for each condition.

The Pytorch library was used to train and test the datasets in all the models. An NVIDIA TITAN XP GPU was used to accelerate the training and testing. The Adam with the betas = (0.9, 0.997), Epsilon (eps) = 10^{-8} was chosen as the optimizer for the student network. The discriminators used stochastic gradient descent (SGD) as the optimizer with a momentum of 0.9 and a weight decay of 5×10^{-5}. The initial learning rates of the student and discriminator networks were set to 0.001 and then multiplied by 0.1 every 20 epochs. The maximum epoch was set to 100, and the size of the mini-batch was set to 16. All images were resized to 224 × 224 pixels and used as inputs for all networks.

Comparison of IRIDS performance with that of ophthalmologists

To compare the performance between the IRIDS and ophthalmologists with clinical experience with fundus diseases, two retinal specialists with 10 to 15 years of clinical experience and two senior ophthalmologists with 5 to 10 years of experience were invited to independently classify the same images in the reader study. For this comparison, we used an additional dataset of 450 retinal images (45 infants) obtained from Shenzhen Eye Hospital in 2022. This dataset included both subclass and main class categories, with 50 images (5 infants) in each category.

Statistical analysis

We evaluated the performance of the IRIDS model using several performance metrics, including accuracy, precision, recall, kappa, area under the receiver operating characteristic curve (AUC), and F1-score (F1). Accuracy was calculated as the ratio of correctly predicted images (true positives and true negatives) to the total number of predictions (true positives, false negatives, false positives, and true negatives). Precision represents the proportion of true positive predictions out of all positive predictions. Recall, also referred to as sensitivity, measures the proportion of true positive predictions out of all actual positive samples. By contrast, kappa, measures the level of agreement between the observed accuracy (p0) and accidental agreement (pe) that could occur by chance.

To calculate these metrics, we utilized the below formulas:

$$\text{Accuracy} : (TP + TN)/(TP + FN + FP + TN)$$

$$\text{Precision} : TP/(TP + FP)$$

$$\text{Recall} : TP/(TP + FN)$$

$$\text{Kappa} : (p0 - pe)/(1 - pe)$$

TP, TN, FP, FN, and TN represent true positives, true negatives, false positives, false negatives, true negatives, respectively. p0 is the accuracy of prediction and pe is the accidental consistency.

Specifically, we employed a fivefold cross-validation technique to obtain the classification results for all categories, and these metrics were calculated and analyzed for each fold to evaluate the overall performance of the model. N-out-of-N bootstrapping was used to estimate the 95% CIs of the performance metrics at the image level, providing a statistical measure of the reliability of these metrics.

Results

Data characteristics and prediction system architecture

The entire workflow, simplified architecture of the IRIDS, and primary dataset processing flow are shown in Fig. 1A, B, and Supplementary Fig. 2. A total of 7697 images (1089 infants) were collected for model training (80%) and testing (20%). Among the 7697 images, 3006 exhibited normal conditions, 326 exhibited Coats, 562 exhibited coloboma, 1013 exhibited RB, 179 exhibited RP, 561 exhibited CRF, 1028 exhibited mild ROP, 667 exhibited moderate ROP, and 355 exhibited severe ROP (Supplementary Table 1).

Performance in data test

First, we conducted several comparative experiments using baseline models (Res-18, Res-34, Res-50, ViT-small, RegionViT, and MaxViT) [41–44]. The classification results are listed in Table 2. The Res-18 model achieved the best classification performance among the CNN models (Table 2). Among the Transformer models, the MaxViT

Table 2 The classification results of different baseline methods (%). The values outside and inside the brackets represent the mean value and variance, respectively

Method	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1-score (95% CI)	Kappa (95% CI)	AUC (95% CI)
Res-18	**93.19 (92.41–93.97)**	**90.62 (84.54–96.70)**	**87.08 (82.99–91.16)**	**88.90 (86.71–91.09)**	**86.93 (84.63–89.23)**	**99.12 (99.12–99.12)**
Res-34	93.01 (92.43–93.59)	91.60 (89.77–93.43)	86.19 (81.50–90.88)	88.76 (86.98–90.54)	87.20 (86.34–88.06)	98.98 (98.98–98.98)
Res-50	92.69 (92.02–93.36)	90.73 (89.76–91.70)	87.68 (86.74–88.62)	89.21 (88.57–89.85)	87.01 (86.37–87.65)	98.27 (98.26–98.28)
ViT-small	92.37 (92.31–92.43)	90.19 (90.05–90.33)	89.97 (89.72–90.22)	90.02 (89.94–90.10)	88.56 (88.39–88.73)	98.82 (98.82–98.82)
RegionViT	90.25 (90.17–90.33)	87.69 (87.50–87.88)	83.38 (82.99–83.77)	85.48 (85.29–85.67)	83.51 (83.32–83.70)	98.04 (97.98–98.09)
MaxViT	**94.62 (94.34–94.90)**	**94.07 (93.32–94.82)**	**90.56 (88.64–92.48)**	**92.34 (91.87–92.81)**	**91.15 (90.37–91.93)**	**99.08 (99.07–99.09)**

model achieved the highest accuracy and precision for all classification metrics. Therefore, the Res-18 model was selected to extract the local contextual feature information, and considering its low complexity and high classification performance, the MaxViT model was utilized to extract the global features. In addition, we plotted receiver operating characteristic (ROC) curves for these baseline methods to compare the performance of the models. From this, we found that the Res-18 and MaxVit models achieved the best performance (Fig. 2A).

Performance in the advanced test

Through experiments comparing baseline methods, the Res-18 and MaxViT models were chosen as the main structures to build the IRIDS. To fuse the extracted features from the Res-blocks and Transformer blocks, we employed a depth attention (DA) fusion module to efficiently fuse high-level features. To verify the effectiveness of the used modules of IRIDS, we conducted ablation experiments in terms of sub-networks and modules. Hence, the IRIDS consisted of the Res-18, MaxViT, and DA modules. The experimental results are presented in Table 3. Our proposed method achieved the best classification performance, with accuracy, precision, recall, F1-score, Kappa, and AUC of 96.45% (95% CI, 96.37%-96.53%), 95.86% (95% CI, 94.56%-97.16%), 94.37% (95% CI, 93.95%-94.79%), 95.03% (95% CI, 94.45%-95.61%), 94.43% (95% CI, 93.96%-94.90%), and 99.51% (95% CI, 99.51%-99.51%), respectively.

We observed that the IRIDS achieved the best ROC performance and the highest AUC for each disease in terms of all evaluation metrics (Figs. 2B and 3A). Specifically, for RP and CRF, the proposed method's correct prediction probability exceeded that of other comparison methods by at least five percentage points. All images in the normal category were correctly detected using our method.

T-distributed stochastic neighbor embedding (T-SNE) visualization

We conducted t-SNE visualization to prove the effectiveness of the extracted features of our proposed method by applying the features of the last feature extraction layer to all the comparison methods, as shown in Fig. 3B. We observed that all predicted categories were better separated by our method than by the other methods (Fig. 3B).

Comparison of IRIDS with ophthalmologists

The results of the comparison between the IRIDS and ophthalmologists are shown in Fig. 2B. For the diagnosis of six main classes of fundus disease, the AUC of the IRIDS was higher than or equal to that of retinal specialists for three of the six diseases (Coats, Coloboma, and RB). The IRIDS achieved an AUC comparable to that of retinal specialists for other fundus diseases (RP, ROP, and CRF). The IRIDS showed a superior AUC compared to senior ophthalmologists for the six main classes.

In addition, we performed a human–machine comparison of the ROP staging model (Fig. 2C and Supplementary Table 2). The results of the IRIDS were better than those of the ophthalmologists for severe ROP, but for the other categories of ROP, the difference was insignificant.

Fig. 2 ROC curves of each of the deep learning methods analyzed in this study. **A.** ROC curves of the different baseline methods. (a)-(f) represent the Res-18, Res-34, Res-50, ViT-small, RegionViT and MaxViT models, respectively. **B.** The ROC curves of the comparison methods for the main classes of diseases and conditions. (a)-(f) represent the Coats, Coloboma, RB, ROP, RP, and CRF categories, respectively. **C.** The ROC curves of the comparison methods for ROP staging. (a)-(c) represent the Mild, Moderate, and Severe conditions, respectively

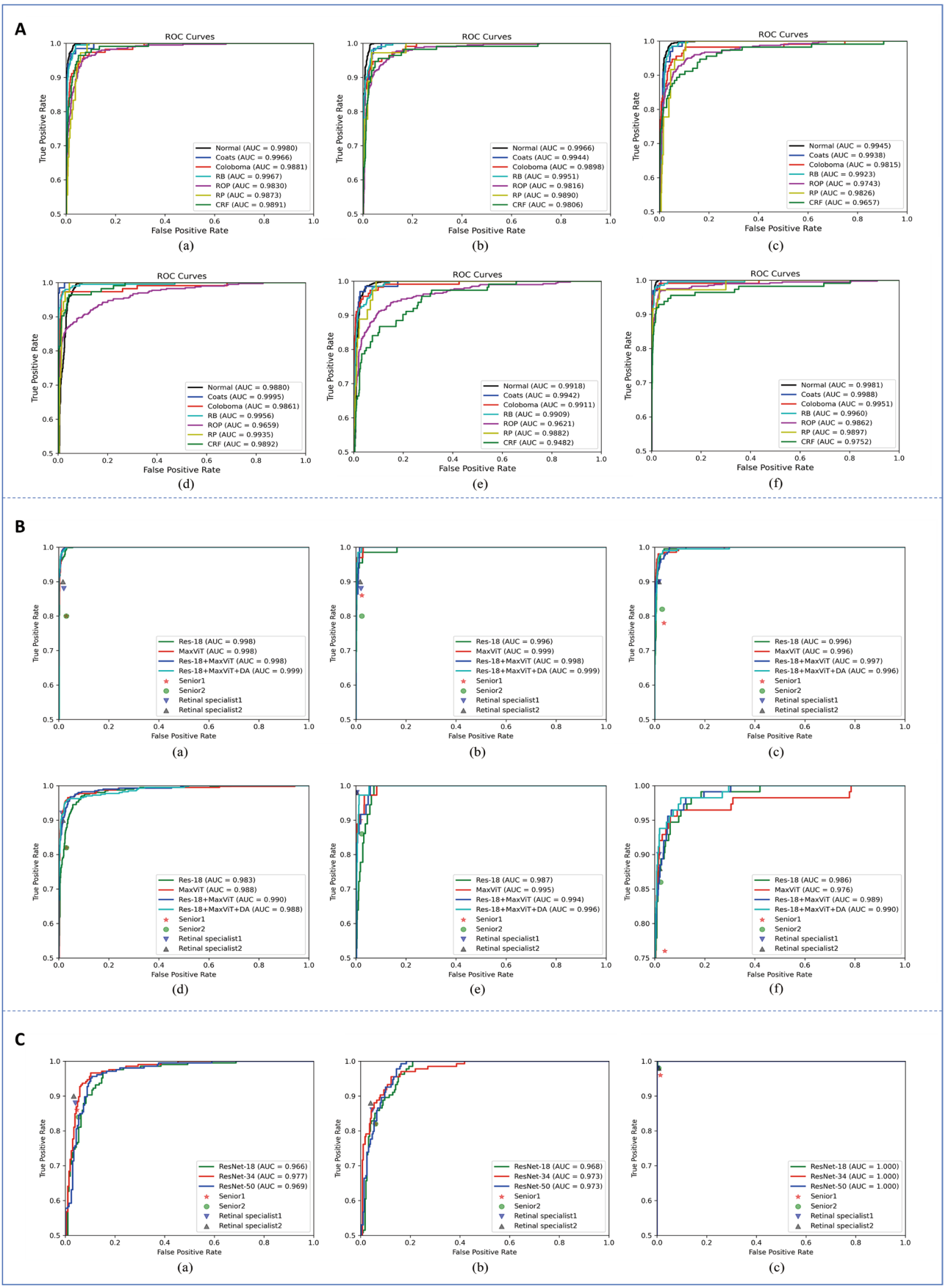

A
ROC Curves
True Positive Rate
False Positive Rate
Normal (AUC = 0.9980)
Coats (AUC = 0.9966)
Coloboma (AUC = 0.9881)
RB (AUC = 0.9967)
ROP (AUC = 0.9830)
RP (AUC = 0.9873)
CRF (AUC = 0.9891)
(a)
Normal (AUC = 0.9966)
Coats (AUC = 0.9944)
Coloboma (AUC = 0.9898)
RB (AUC = 0.9951)
ROP (AUC = 0.9816)
RP (AUC = 0.9890)
CRF (AUC = 0.9806)
(b)
Normal (AUC = 0.9945)
Coats (AUC = 0.9938)
Coloboma (AUC = 0.9815)
RB (AUC = 0.9923)
ROP (AUC = 0.9743)
RP (AUC = 0.9826)
CRF (AUC = 0.9657)
(c)
Normal (AUC = 0.9880)
Coats (AUC = 0.9995)
Coloboma (AUC = 0.9861)
RB (AUC = 0.9956)
ROP (AUC = 0.9659)
RP (AUC = 0.9935)
CRF (AUC = 0.9892)
(d)
Normal (AUC = 0.9918)
Coats (AUC = 0.9942)
Coloboma (AUC = 0.9911)
RB (AUC = 0.9909)
ROP (AUC = 0.9621)
RP (AUC = 0.9882)
CRF (AUC = 0.9482)
(e)
Normal (AUC = 0.9981)
Coats (AUC = 0.9988)
Coloboma (AUC = 0.9951)
RB (AUC = 0.9960)
ROP (AUC = 0.9862)
RP (AUC = 0.9897)
CRF (AUC = 0.9752)
(f)
B
Res-18 (AUC = 0.998)
MaxViT (AUC = 0.998)
Res-18+MaxViT (AUC = 0.998)
Res-18+MaxViT+DA (AUC = 0.999)
Senior1
Senior2
Retinal specialist1
Retinal specialist2
(a)
Res-18 (AUC = 0.996)
MaxViT (AUC = 0.999)
Res-18+MaxViT (AUC = 0.998)
Res-18+MaxViT+DA (AUC = 0.999)
(b)
Res-18 (AUC = 0.996)
MaxViT (AUC = 0.996)
Res-18+MaxViT (AUC = 0.997)
Res-18+MaxViT+DA (AUC = 0.996)
(c)
Res-18 (AUC = 0.983)
MaxViT (AUC = 0.988)
Res-18+MaxViT (AUC = 0.990)
Res-18+MaxViT+DA (AUC = 0.988)
(d)
Res-18 (AUC = 0.987)
MaxViT (AUC = 0.995)
Res-18+MaxViT (AUC = 0.994)
Res-18+MaxViT+DA (AUC = 0.996)
(e)
Res-18 (AUC = 0.986)
MaxViT (AUC = 0.976)
Res-18+MaxViT (AUC = 0.989)
Res-18+MaxViT+DA (AUC = 0.990)
(f)
C
ResNet-18 (AUC = 0.966)
ResNet-34 (AUC = 0.977)
ResNet-50 (AUC = 0.969)
Senior1
Senior2
Retinal specialist1
Retinal specialist2
(a)
ResNet-18 (AUC = 0.968)
ResNet-34 (AUC = 0.973)
ResNet-50 (AUC = 0.973)
(b)
ResNet-18 (AUC = 1.000)
ResNet-34 (AUC = 1.000)
ResNet-50 (AUC = 1.000)
(c)

Table 3 The ablation experiments of the proposed model (%). The values outside and inside of the brackets represent the mean value and variance, respectively

Method	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1-score (95% CI)	Kappa (95% CI)	AUC (95% CI)
Res-18	93.19 (92.41–93.67)	90.62 (84.54–96.70)	87.08 (83.00–91.16)	88.90 (86.71–91.09)	86.93 (84.63–89.23)	99.12 (99.12–99.12)
MaxViT	94.62 (94.34–94.90)	94.07 (93.32–94.82)	90.56 (88.64–92.48)	92.34 (91.87–92.81)	91.15 (90.37–91.93)	99.08 (99.07–99.09)
Res-18 + MaxViT	95.64 (0.02)	96.58 (0.09)	90.77 (0.30)	93.56 (0.08)	92.50 (0.16)	99.39 (99.39–99.39)
Res-18 + MaxViT + DA	**96.45 (96.37–96.53)**	**95.86 (94.56–97.16)**	**94.37 (93.95–94.79)**	**95.03 (94.45–95.61)**	**94.43 (93.96–94.90)**	**99.51 (99.51–99.51)**

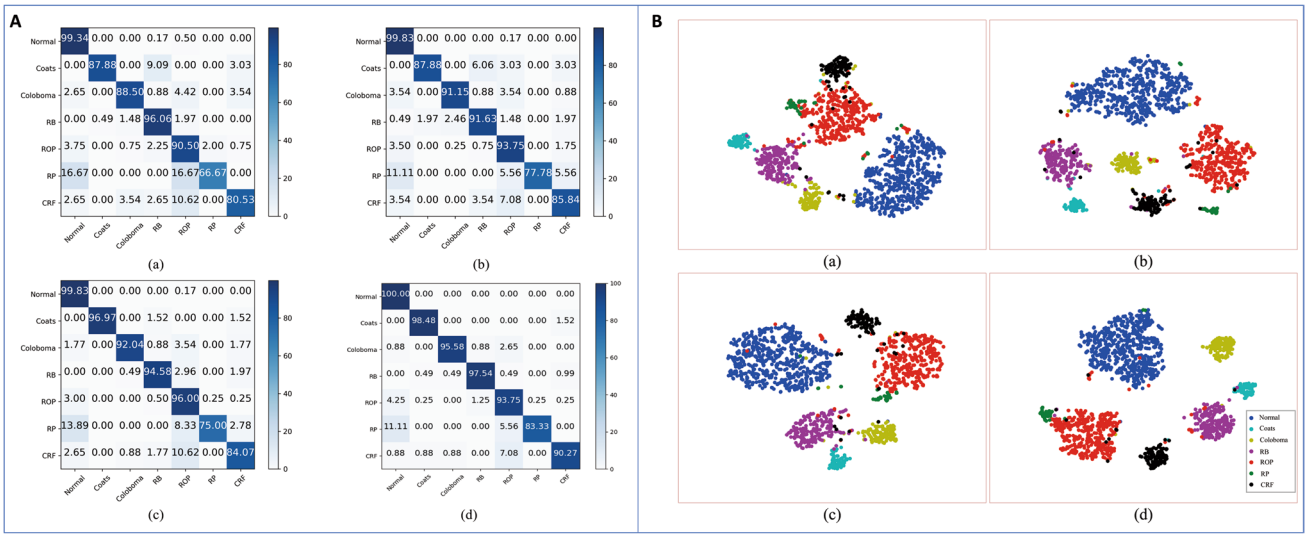


Fig. 3 Visualization of the classifier performances. **A.** Confusion matrixes of different methods. The confusion matrix is a specific matrix used to present the visualization effect of algorithm performance. Each column represents the predicted value and each row represents the actual category. (a)-(d) represent the Res-18, MaxViT, Res-18 + MaxViT and Res-18 + MaxViT + DA (Proposed) models, respectively. **B.** T-SNE visualization of different methods. (a)-(d) represent the Res-18, MaxViT, Res-18 + MaxViT and Res-18 + MaxViT + DA (Proposed) models, respectively

Discussion

In this study, we developed the IRIDS model, which automatically identified nine infantile fundus diseases and conditions. By comparing multiple algorithms, we established that our model had the highest performance, with an AUC of 99.51%. Compared to retinal specialists, the IRIDS exhibits superior diagnostic accuracy.

We first conducted several comparative experiments using baseline models, and their results are presented in Table 2. Finally, we selected the Res-18 and MaxViT models that combined DA modules as the main structures to build the IRIDS. The IRIDS was used to analyze data collected from four hospitals. Despite potential variations in image quality, imaging techniques, and machine types across these hospitals, the model achieves the accuracy of 96.45% (95% CI, 96.37–96.53), which is better than other methods and indicates its robust adaptability.

Many CNN-based DL methods focus on local feature extraction, while overlooking valuable information from the pan-retinal viewing field. To address this issue, a transformer model was adopted for vision tasks [45–47], which can be employed to extract long-distance dependent features from input images. Therefore, the integration of CNNs and transformer structures within a single model enables the extraction of both local and global features, ultimately improving the model's classification performance. The inherent opacity of CNNs in image analysis poses a potential obstacle that hinders clinicians from adopting DL techniques [48]. However, our model can

effectively extract key local and global features relevant to specific diseases, as demonstrated by the results of the t-SNE analysis. This characteristic allows for an improved classification performance and provides promising potential for its clinical applications.

In this study, we investigated the automated detection of infantile fundus diseases using the IRIDS. We have previously reported the use of ultrawide-field retinal images to diagnose ROP [31]. Studies have also been conducted on the staging of ROP and the diagnosis of RB and RP in infants [49–51]. However, these studies have mainly focused on detecting a single infantile fundus disease using retinal images. Subsequently, an increasing number of studies have shifted focus towards automating the diagnosis of multiple fundus diseases. Some studies have aimed to distinguish among several fundus diseases using a single algorithm [35]. Stevenson et al. [52] developed a multiclass DL model for detecting fundus diseases using a publicly available retinal image database. This study demonstrated the feasibility of developing artificial intelligence (AI) classifiers using standard laptops without prior programming or AI knowledge. Son et al. [53] used a larger dataset to build a DL model for detecting 12 retinal abnormalities. The external testing set in their study achieved an AUC of up to 98%, indicating high performance. Moreover, the model's interpretable and reliable classification outputs suggest its potential for clinical use as an automated screening system for retinal images. However, compared to IRIDS, their system primarily focuses on detecting retinal abnormalities rather than directly diagnosing specific diseases or conditions. Therefore, these models cannot be used directly for screening. Lin et al. [54] developed a DL-based system that prospectively evaluated 18,163 color fundus photographs and detected 14 fundus diseases, achieving an AUC of 0.98. Li et al. [35] developed and prospectively validated a retinal image-based DL algorithm to simultaneously identify ten fundus diseases in clinical practice. This is in line with general screening methods and is practical. Cen et al. [55] used a large database of 600,000 retinal images to establish a digital light processing system with a high F1 score, sensitivity, specificity, and AUC for the detection of 39 fundus diseases and conditions. Ju et al. [56] introduced a novel method that enabled deep neural networks to learn from a comprehensive fundus database with over one million images, facilitating the recognition of various fundus diseases. Gu et al. [57] evaluated the performance of an AI system in detecting multiple fundus diseases in real-world scenarios, specifically primary healthcare settings, and achieved high accuracy. However, these studies only focused on adult fundus diseases and did not include infantile fundus diseases. IRIDS utilizes non-invasive retinal images, which are easily obtained, to enable automatic screening for multiple infantile fundus diseases and conditions. The IRIDS demonstrates high performance, achieving notable classification results across multiple evaluation metrics. The accuracy is 96.45% and the AUC is 99.51% (Figs. 2B and 3A). These findings underscore the superior clinical applicability and adaptability of IRIDS compared to existing approaches.

Conclusions, expert recommendations, outlook in the framework of PPPM/3PM, and limitations of the study

Conclusions

In summary, the IRIDS utilizes advanced DL capabilities to accurately identify nine infantile fundus diseases and conditions, demonstrating high diagnostic precision and reliability. It serves as a pioneering example of integrating AI into ophthalmology, significantly aiding non-ophthalmologist personnel in underserved areas for effective infantile fundus disease screening. By promptly detecting retinal lesions, the IRIDS may effectively prevent severe complications, thus improving visual outcomes and the overall quality of life for children and their families worldwide, especially in regions lacking retinal specialists. The deployment of IRIDS holds great promise for enhancing early detection and intervention, aligning with the goals of PPPM/3PM. Ultimately, this leads to improved prognosis, better healthcare access, and outcomes for children in need. The IRIDS exemplifies the potential of AI to revolutionize healthcare and offers a model for future innovations in the field.

Predictive approach

By using the depth-attention fusion module to fuse the local and global features from the paths of the ResNet and MaxViT modules, the IRIDS not only achieved automated screening for multiple infantile fundus conditions and diseases, but also demonstrated high diagnostic accuracy comparable to or better than that of ophthalmologists, including retinal specialists. As shown in Table 1, the IRIDS may automatically detect the most common fundus diseases and conditions in infants, such as Coats' disease, RB, RP, choroid coloboma, and CRF, with AUC values close to 1.0, indicating excellent performance. The high accuracy of our model, which may be due to the strict dataset selection process [53], was further confirmed by the high AUC results achieved by the IRIDS, despite the high imbalance rate, as reported in a study using a multilabel setting [58]. Within the context of PPPM [38], it would be advantageous to strive to incorporate this new

modality into current clinical workflows because it has the potential to provide more personalized and quantifiable methods for the diagnosis of multiple infantile diseases. This further confirms the robustness and reliability of our model, making it a promising tool for the early detection and treatment of infantile fundus disease.

Targeted prevention

In the context of prevention [38], due to its outstanding performance in human–machine comparison, the IRDS system enables us to identify and promptly treat some infantile fundus diseases that are prone to be overlooked or misdiagnosed, even in the absence of retinal specialists. This helps to prevent further deterioration of the patient's condition and to halt the progression of the disease. The IRIDS is helpful in reducing the burden on ophthalmologists and improving screening efficiency, especially in less-developed regions.

Personalization of medical services

We included nine different infantile fundus diseases and conditions from four hospitals. To address the limitations of using a single structure to extract features from image analysis, we employed a hybrid model that combines CNNs and transformer structures. This hybrid model allows simultaneous attention to local and global features. As a result, the AUC for the classification of the nine disease categories reached an impressive 99.51%. Consequently, the entire network may capture disease-specific features and focus on extracting key features for each condition. Therefore, IRIDS may diagnose individual infants based on a single retinal image. Owing to its simplicity and ease of training, the IRIDS may also be applied in environments where there are no professional ophthalmologists, potentially helping to eliminate the problems of underdiagnosis, misdiagnosis, and overexamination.

Limitations and outlook in the framework of PPPM/3PM

This study has a few limitations. First, similar to many other fundus diseases, some infantile fundus diseases lack universal diagnostic criteria or consensus. Mitigating this issue can involve referencing multiple authoritative sources or relevant studies. Second, although we collected data on up to nine common infantile fundus diseases and conditions, we did not include all infantile fundus diseases or conditions. Diseases such as familial exudative vitreoretinopathy and retinal detachment were omitted due to insufficient image data. Adding an "others" class can be an effective way to address this limitation. Third, our research dataset was relatively small and limited to a single geographic region. Furthermore, this dataset includes only Chinese individuals, limiting its universal clinical application of the AI system. Therefore, images from databases in other regions worldwide should be included to increase diversity and expand the applicability of the database.

The IRIDS is the first of its kind to automatically recognize infantile fundus diseases in a non-invasive manner. It also aligns with a crucial aspect of the PPPM/3PM approach emphasized at the EPMA World Congress 2019, which focused on customizing and continuously monitoring patients' clinical parameters to improve treatment outcomes [59]. To facilitate the future application of DL in this field, it is strongly recommended to consider the following points:

1. Ensuring comprehensive detection of infantile fundus diseases or conditions across a wide range of categories, while also enabling intelligent clinical diagnosis applicable to each specific category.
2. Enabling comprehensive prediction of disease progression and facilitating treatment decision assessments related to clinical settings, enabling early application in clinical monitoring. Through the PPPM/3PM approach, IRIDS aims to provide better clinical outcomes for patients by offering a complete prediction from diagnosis to treatment decision.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13167-024-00350-y.

Author contributions GZ has full access to all data in the study and is responsible for the integrity of the data and the accuracy of the data analysis. YL, HX, and XZ contributed equally and are considered co-first authors. Concept and design: YL, HX, and XZ. Acquisition, analysis, or interpretation of data: YL, HX, XZ, and SZ. Drafting of the manuscript: YL and HX. Critical revision of the manuscript for important intellectual content: GZ, JT, XZ, and DPN. Statistical analysis: YL and JT. Obtained funding: GZ. Administrative, technical, or material support: ZY, ZW, RT, YC, MC, YD, TC, YH, and BL. Supervision: GZ.

Funding This study was supported by National Natural Science Foundation of China (No. 82271103, 82301269, 82301226, 62376164, 62106153, U22A2024), Sanming Project of Medicine in Shenzhen (No. SZSM202311018), Guangdong Basic and Applied Basic Research Foundation (No. 2022A1515012326, 22201910240002529), Shenzhen Medical Research Fund (No. C2301005), Shenzhen Key Medical Discipline Construction Fund (No. SZXK038), Shenzhen Fund for Guangdong Provincial High Level Clinical Key Specialties (No. SZGSP014), Shenzhen Science and Technology R&D Fund Project (No. JCYJ20220530153607015), China Ophthalmology New Technology Incubation Project.

Data availability These authors do not have permission to share data except the first author.

Code availability The source code used in this study is available upon request from the first author only.

Declarations

Conflicts interests The authors have no relevant financial or non-financial interests to disclose.

Ethics approval This diagnostic study was approved by the Ethics Committee of Shenzhen Eye Hospital. All institutions abided by the tenets of the Declaration of Helsinki.

Consent to participate Written informed consent was obtained from the parents of all enrolled infants.

Consent for publication This article has been approved for publication by the authors.

References

1. Chiang MF, Quinn GE, Fielder AR, Ostmo SR, Paul Chan RV, Berrocal A, et al. International Classification of Retinopathy of Prematurity, Third Edition. Ophthalmology. 2021;128(10):e51–68. https://doi.org/10.1016/j.ophtha.2021.05.031.
2. Shields JA, Shields CL, Honavar SG, Demirci H. Clinical variations and complications of Coats disease in 150 cases: the 2000 Sanford Gifford Memorial Lecture. Am J Ophthalmol. 2001;131(5):561–71. https://doi.org/10.1016/s0002-9394(00)00883-7.
3. Spitznas M, Joussen F, Wessing A, Meyer-Schwickerath G. Coat's disease. An epidemiologic and Fluorescein angiographic study. Albrecht Von Graefes Arch Klin Exp Ophthalmol. 1975;195(4):241–50. https://doi.org/10.1007/BF00414937.
4. Rao R, Honavar SG. Retinoblastoma. Indian J Pediatr. 2017;84(12):937–44. https://doi.org/10.1007/s12098-017-2395-0.
5. Pagon RA. Retinitis pigmentosa. Surv Ophthalmol. 1988;33(3):137–77. https://doi.org/10.1016/0039-6257(88)90085-9.
6. Giles K, Raoul C, Yannick B, Peter W. Uveal coloboma: about 3 cases at the University Teaching Hospital, Yaounde, Cameroon. Pan Afr Med J. 2016;24:201. https://doi.org/10.11604/pamj.2016.24.201.9770.
7. Nishina S, Suzuki Y, Yokoi T, Kobayashi Y, Noda E, Azuma N. Clinical features of congenital retinal folds. Am J Ophthalmol. 2012;153(1):81-7 e1. https://doi.org/10.1016/j.ajo.2011.06.002.
8. Liche F, Majji AB. Familial exudative vitreoretinopathy. Ophthalmology. 2012;119(5):1093. https://doi.org/10.1016/j.ophtha.2012.02.025.
9. Fielder A, Blencowe H, O'Connor A, Gilbert C. Impact of retinopathy of prematurity on ocular structures and visual functions. Arch Dis Child Fetal Neonatal Ed. 2015;100(2):F179–84. https://doi.org/10.1136/archdischild-2014-306207.
10. Golubnitschaja O, Costigliola V, Epma. General report & recommendations in predictive, preventive and personalised medicine 2012: white paper of the European Association for Predictive, Preventive and Personalised Medicine. EPMA J. 2012;3(1):14. https://doi.org/10.1186/1878-5085-3-14.
11. Good WV. Retinopathy of Prematurity Incidence in Children. Ophthalmology. 2020;127(4S):S82–3. https://doi.org/10.1016/j.ophtha.2019.11.026.
12. Dimaras H, Corson TW, Cobrinik D, White A, Zhao J, Munier FL, et al. Retinoblastoma. Nat Rev Dis Primers. 2015;1:15021. https://doi.org/10.1038/nrdp.2015.21.
13. Global Retinoblastoma Study G, Fabian ID, Abdallah E, Abdullahi SU, Abdulqader RA, Adamou Boubacar S, et al. Global Retinoblastoma Presentation and Analysis by National Income Level. JAMA Oncol. 2020;6(5):685–95. https://doi.org/10.1001/jamaoncol.2019.6716.
14. Chen HY, Lehmann OJ, Swaroop A. Genetics and therapy for pediatric eye diseases. EBioMedicine. 2021;67:103360. https://doi.org/10.1016/j.ebiom.2021.103360.
15. Coleman K, Coleman J, Franco-Penya H, Hamroush F, Murtagh P, Fitzpatrick P, et al. A New Smartphone-Based Optic Nerve Head Biometric for Verification and Change Detection. Transl Vis Sci Technol. 2021;10(8):1. https://doi.org/10.1167/tvst.10.8.1.
16. Golubnitschaja O, Potuznik P, Polivka J Jr, Pesta M, Kaverina O, Pieper CC, et al. Ischemic stroke of unclear aetiology: a case-by-case analysis and call for a multi-professional predictive, preventive and personalised approach. EPMA J. 2022;13(4):535–45. https://doi.org/10.1007/s13167-022-00307-z.
17. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56. https://doi.org/10.1038/s41591-018-0300-7.
18. Baek SU, Lee WJ, Park KH, Choi HJ. Health screening program revealed risk factors associated with development and progression of papillomacular bundle defect. EPMA J. 2021;12(1):41–55. https://doi.org/10.1007/s13167-021-00235-4.
19. Li S, Li M, Wu J, Li Y, Han J, Cao W, et al. Development and validation of a routine blood parameters-based model for screening the occurrence of retinal detachment in high myopia in the context of PPPM. EPMA J. 2023;14(2):219–33. https://doi.org/10.1007/s13167-023-00319-3.
20. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science. 1999;286(5439):531–7. https://doi.org/10.1126/science.286.5439.531.
21. Wang Y, Tetko IV, Hall MA, Frank E, Facius A, Mayer KF, et al. Gene selection from microarray data for cancer classification–a machine learning approach. Comput Biol Chem. 2005;29(1):37–46. https://doi.org/10.1016/j.compbiolchem.2004.11.001.
22. Yu KH, Levine DA, Zhang H, Chan DW, Zhang Z, Snyder M. Predicting Ovarian Cancer Patients' Clinical Response to Platinum-Based Chemotherapy by Their Tumor Proteomic Signatures. J Proteome Res. 2016;15(8):2455–65. https://doi.org/10.1021/acs.jproteome.5b01129.
23. Yu KH, Fitzpatrick MR, Pappas L, Chan W, Kung J, Snyder M. Omics AnalySIs System for PRecision Oncology (OASISPRO): a web-based omics analysis tool for clinical phenotype prediction. Bioinformatics. 2018;34(2):319–20. https://doi.org/10.1093/bioinformatics/btx572.
24. Check Hayden E. The automated lab. Nature. 2014;516(7529):131–2. https://doi.org/10.1038/516131a.
25. Chew EY. Age-related Macular Degeneration: Nutrition, Genes and Deep Learning-The LXXVI Edward Jackson Memorial Lecture. Am J Ophthalmol. 2020;217:335–47. https://doi.org/10.1016/j.ajo.2020.05.042.
26. Shon K, Sung KR, Shin JW. Can Artificial Intelligence Predict Glaucomatous Visual Field Progression? A Spatial-Ordinal Convolutional Neural Network Model. Am J Ophthalmol. 2022;233:124–34. https://doi.org/10.1016/j.ajo.2021.06.025.
27. Ee CL, Samsudin A. Comparison of Smartphone-Based and Automated Refraction with Subjective Refraction for Screening of Refractive Errors. Ophthalmic Epidemiol. 2022;29(5):588–94. https://doi.org/10.1080/09286586.2021.1986550.
28. Dai L, Wu L, Li H, Cai C, Wu Q, Kong H, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nat Commun. 2021;12(1):3242. https://doi.org/10.1038/s41467-021-23458-5.

29. Panwar N, Huang P, Lee J, Keane PA, Chuan TS, Richhariya A, et al. Fundus Photography in the 21st Century–A Review of Recent Technological Advances and Their Implications for Worldwide Healthcare. Telemed J E Health. 2016;22(3):198–208. https://doi.org/10.1089/tmj.2015.0068.
30. Zhao J, Lei B, Wu Z, Zhang Y, Li Y, Wang L, et al. A Deep Learning Framework for Identifying Zone I in RetCam Images. IEEE Access. 2019;7:103530–7. https://doi.org/10.1109/access.2019.2930120.
31. Zhang Y, Wang L, Wu Z, Zeng J, Chen Y, Tian R, et al. Development of an Automated Screening System for Retinopathy of Prematurity Using a Deep Neural Network for Wide-Angle Retinal Images. IEEE Access. 2019;7:10232–41. https://doi.org/10.1109/access.2018.2881042.
32. Rugang Zhang JZ. Hai Xie, Tianfu Wang, Automatic diagnosis for aggressive posterior retinopathy of prematurity via deep attentive convolutional neural network. Expert Syst Appl. 2022;187:115843.
33. Maji D, Sekh AA. Automatic grading of retinal blood vessel in deep retinal image diagnosis. J Med Syst. 2020;44(180). http://doi.org/10.1007/s10916-020-01635-1.
34. Xie HLH, Zeng X, He Y, Chen G. AMD-GAN: attention encoder and multi-branch structure based generative adversarial networks for fundus disease detection from scanning laser ophthalmoscopy images. Neural Netw. 2020;132:477–90.
35. Dong L, He W, Zhang R, Ge Z, Wang YX, Zhou J, et al. Artificial Intelligence for Screening of Multiple Retinal and Optic Nerve Diseases. JAMA Netw Open. 2022;5(5):e229960. https://doi.org/10.1001/jamanetworkopen.2022.9960.
36. Cen LP, Ji J, Lin JW, Ju ST, Lin HJ, Li TP, et al. Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. Nat Commun. 2021;12(1):4828. https://doi.org/10.1038/s41467-021-25138-w.
37. Patel TP, Kim TN, Yu G, Dedania VS, Lieu P, Qian CX, et al. Smartphone-Based, Rapid, Wide-Field Fundus Photography for Diagnosis of Pediatric Retinal Diseases. Transl Vis Sci Technol. 2019;8(3):29. https://doi.org/10.1167/tvst.8.3.29.
38. Golubnitschaja O, Baban B, Boniolo G, Wang W, Bubnov R, Kapalla M, et al. Medicine in the early twenty-first century: paradigm and anticipation - EPMA position paper 2016. EPMA J. 2016;7(1):23. https://doi.org/10.1186/s13167-016-0072-4.
39. World Medical A. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310(20):2191–4. https://doi.org/10.1001/jama.2013.281053.
40. Reynolds JD and Olitsky SE, pediatric retina. 1st ed. Heidelberg: Springer-Verlag Berlin. 2010. https://doi.org/10.1007/978-3-642-12041-11.
41. 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. 2016;770–778. https://doi.org/10.48550/arXiv.1512.03385.
42. Tu Z, Talebi H, Zhang H, Yang F, Peyman Milanfar, Alan Bovik, et al. Maxvit: Multi-axis vision transformer. arXiv preprint arXiv. 2022: p. 01697. https://doi.org/10.48550/arXiv.2204.01697.
43. Chen C, Rameswar P, Fan Q. Regionvit: Regional-to-local attention for vision transformers. arXiv preprint arXiv. 2022. https://doi.org/10.48550/arXiv.2106.02689.
44. Lee SH, Lee S, Song BC. Vision transformer for small-size datasets. arXiv preprint arXiv. 2021. https://doi.org/10.48550/arXiv.2112.13492.
45. Yu S, Ma K, Bi Q, Bian C. Mil-vt: Multiple instance learning enhanced vision transformer for fundus image classification[C]// Medical Image Computing and Computer Assisted Intervention-MICCAI 2021: 24th International Conference, Strasbourg, France, September 27-October 1, 2021, Proceedings, Part VIII 24. Springer International Publishing. 2021: p. 45-54. https://doi.org/10.1007/978-3-030-87237-3_5.
46. Sun R, Li Y, Zhang T, Mao Z, Wu F. Lesion-aware transformers for diabetic retinopathy grading[C]. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: p. 10938–10947. https://doi.org/10.1109/CVPR46437.2021.01079.
47. Playout C, Duval R, Boucher MC, Cheriet F. Focused Attention in Transformers for interpretable classification of retinal images. Med Image Anal. 2022;82:102608. https://doi.org/10.1016/j.media.2022.102608.
48. Wong TY, Bressler NM. Artificial Intelligence With Deep Learning Technology Looks Into Diabetic Retinopathy Screening. JAMA. 2016;316(22):2366–7. https://doi.org/10.1001/jama.2016.17563.
49. Ng WY, Zhang S, Wang Z, Ong CJT, Gunasekeran DV, Lim GYS, Ting DSW. Updates in deep learning research in ophthalmology. Clin Sci. 2021;135(20):2357–76.
50. Huang YPBH, Kang EYC, Chen KJ, Hwang YS, Lai CC, Wu WC. Automated detection of early-stage ROP using a deep convolutional neural network. Br J Ophthalmol. 2021;105(8):1099–103.
51. Durai C, Jebaseeli TJ, Alelyani S, Mubharakali A. Early Prediction and Diagnosis of Retinoblastoma Using Deep Learning Techniques. arXiv preprint arXiv. 2021. https://doi.org/10.48550/arXiv.2103.07622.
52. Stevenson CH, Hong SC, Ogbuehi KC. Development of an artificial intelligence system to classify pathology and clinical features on retinal fundus images. Clin Exp Ophthalmol. 2019;47(4):484–9. https://doi.org/10.1111/ceo.13433.
53. Son J, Shin JY, Kim HD, Jung KH, Park KH, Park SJ. Development and Validation of Deep Learning Models for Screening Multiple Abnormal Findings in Retinal Fundus Images. Ophthalmology. 2020;127(1):85–94. https://doi.org/10.1016/j.ophtha.2019.05.029.
54. Lin D, Xiong J, Liu C, Zhao L, Li Z, Yu S, et al. Application of Comprehensive Artificial intelligence Retinal Expert (CARE) system: a national real-world evidence study. Lancet Digit Health. 2021;3(8):e486–95. https://doi.org/10.1016/S2589-7500(21)00086-8.
55. Cen L-P, Ji J, Lin J-W, Ju S-T, Lin H-J, Li T-P, et al. Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. Nat Commun. 2021;12(1). https://doi.org/10.1038/s41467-021-25138-w.
56. Ju L, Yu Z, Wang L, Zhao X, Wang X, Bonnington P, et al. Hierarchical Knowledge Guided Learning for Real-world Retinal Disease Recognition. IEEE Trans Med Imaging. 2023;PP. https://doi.org/10.1109/TMI.2023.3302473.
57. Gu C, Wang Y, Jiang Y, Xu F, Wang S, Liu R, et al. Application of artificial intelligence system for screening multiple fundus diseases in Chinese primary healthcare settings: a real-world, multicentre and cross-sectional study of 4795 cases. Br J Ophthalmol. 2023. https://doi.org/10.1136/bjo-2022-322940.
58. Zhu H, Cheng C, Yin H, Li X, Zuo P, Ding J, et al. Automatic multilabel electrocardiogram diagnosis of heart rhythm or conduction abnormalities with deep learning: a cohort study. Lancet Digit Health. 2020;2(7):e348–57. https://doi.org/10.1016/S2589-7500(20)30107-2.
59. Golubnitschaja O, Topolcan O, Kucera R, Costigliola V, Epma. 10th Anniversary of the European Association for Predictive, Preventive and Personalised (3P) Medicine - EPMA World Congress Supplement 2020. EPMA J. 2020;11(Suppl 1):1–133. https://doi.org/10.1007/s13167-020-00206-1.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Yaling Liu[1] · Hai Xie[2] · Xinyu Zhao[1] · Jiannan Tang[1] · Zhen Yu[1] · Zhenquan Wu[1] · Ruyin Tian[1] · Yi Chen[1,3] · Miaohong Chen[1,3] · Dimitrios P. Ntentakis[4] · Yueshanyi Du[1] · Tingyi Chen[1,3] · Yarou Hu[1] · Sifan Zhang[3,5] · Baiying Lei[2] · Guoming Zhang[1,3]

✉ Sifan Zhang
12111244@mail.sustech.edu.cn

✉ Baiying Lei
leiby@szu.edu.cn

✉ Guoming Zhang
13823509060@163.com

[1] Shenzhen Eye Hospital, Shenzhen Eye Institute, Jinan University, Shenzhen 518040, China

[2] National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China

[3] Guizhou Medical University, Guiyang, Guizhou, China

[4] Retina Service, Ines and Fred Yeatts Retina Research Laboratory, Angiogenesis Laboratory, Department of Ophthalmology, Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA

[5] Southern University of Science and Technology School of Medicine, Shenzhen, China