Experimental Eye Research 265 (2026) 110893

Contents lists available at ScienceDirect

Experimental Eye Research

journal homepage: www.elsevier.com/locate/yexer

Axial length prediction Model based on screening fundus photography data in school-age children

Zixun Wang [a,2], Hua Rong [a,2], Jingtao Yu [a], Yifan Zhou [a], Xueshuo Xie [b,c], Weiping Lin [a], Zhiqing Li [a], Hua Yan [d,*,1], Bei Du [a,**,1], Ruihua Wei [a,1,***]

[a] Tianjin Key Laboratory of Retinal Functions and Diseases, Tianjin Branch of National Clinical Research Center for Ocular Disease, Eye Institute and School of Optometry, Tianjin Medical University Eye Hospital, Tianjin, China
[b] Haihe Lab of ITAI, Tianjin, China
[c] Nankai University, Tianjin, China
[d] Department of Ophthalmology, Tianjin Medical University General Hospital, Tianjin, China

ARTICLE INFO

Keywords:
Axial length
Color fundus photograph
Prediction
Deep learning
School-age children

ABSTRACT

This study developed deep learning (DL) models to predict axial length (AL) in 6–10-year-old schoolchildren using minimally abnormal color fundus photographs (CFPs), while evaluating the impact of integrating age, Diopter Sphere (DS), and sex. Following quality assessment of 5460 initial CFPs from 3840 children, 3840 images from 2779 children were utilized and partitioned into training (70 %), validation (20 %), and test (10 %) sets. ResNet101 served as the core architecture, with supplemental clinical parameters integrated into the fully connected layer for continuous AL prediction. Model interpretation employed Grad-CAM-generated heatmaps. Comparative analysis demonstrated that DS and age achieved moderate predictive accuracy ($R^2 = 0.37$), a CFP-only model showed significantly stronger performance ($R^2 = 0.70$), and combining CFPs with DS and age further improved accuracy ($R^2 = 0.75$). However, incorporating sex alongside CFPs, DS, and age substantially reduced efficacy ($R^2 = 0.41$). Heatmaps revealed that regions critical for AL predictions anatomically corresponded to retinal vasculature and immediate perivascular tissues. These findings collectively indicate that DL may leverage near-normal CFPs for pediatric AL prediction, with selective enhancement by age and DS, but degradation when categorical variables (such as sex) are included. Subtle changes in the fundus vasculature may help DL to identify the cause of CFP changes with AL.

1. Introduction

Globally, myopia is one of the leading causes of visual impairment and blindness. Uncorrected refractive error remains the leading cause of visual impairment in both pediatric and adult populations worldwide (Burton et al., 2021). By the year 2050, approximately 50 % of the global population will have myopia, with 10 % of this group exhibiting high myopia. This trend exhibits an upward trajectory (Holden et al., 2016). Meanwhile, visual impairment also imposes a significant financial burden worldwide, with an estimated annual loss in global productivity of $411 billion (Bullimore et al., 2021). Myopia, especially high myopia, is strongly associated with a significant risk of complications that can lead to blindness and impaired vision (Vision Loss Expert Group, 2021). Pathologic myopia complications include high myopic choroidal neovascularization, retinal choroidal atrophy, retinal cleavage, macular tear, and even retinal detachment (Vision Loss Expert Group, 2021). Early fundus monitoring in myopic patients and adolescents facilitates the detection of fundus lesions, enabling early prevention and intervention (Nasseri et al., 2024).

The axial length (AL) is an indicator of eyeball size (Kakita et al., 2011). It has emerged as a key predictor of myopia progression and control in children, where early intervention targeting excessive AL

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: phuayan2000@163.com (H. Yan), dubei1982@126.com (B. Du), rwei@tmu.edu.cn (R. Wei).
[1] Contributed equally to this work and were both considered Corresponding Authors.
[2] Contributed equally to this paper as First Authors.

https://doi.org/10.1016/j.exer.2026.110893
Received 10 July 2025; Received in revised form 23 January 2026; Accepted 25 January 2026
Available online 27 January 2026

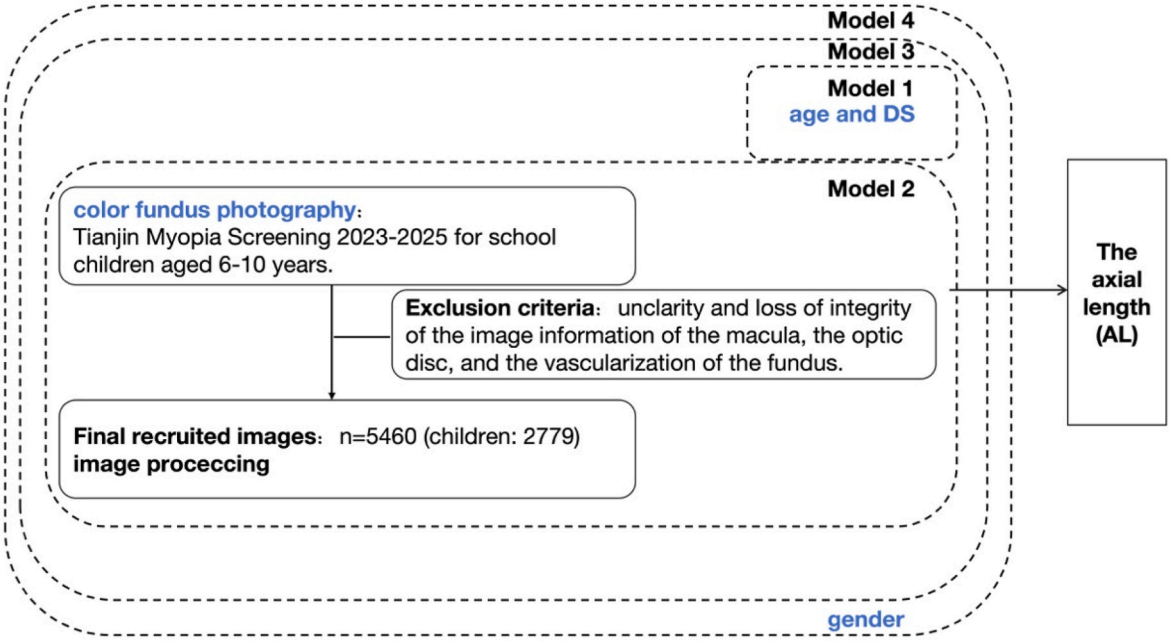

Fig. 1. The process of designing models and selecting CFPs. Exclusion criteria include any retinal or choroidal disease observed in the CFP and patients with other ocular surface diseases. The above exclusion process was performed by three experienced ophthalmologists. Model l contained structured data (age and DS), while Model 2 included CFP only. Model 3 included CFP, age, and DS. Model 4 included sex, age, CFP, and DS. The outputs are all AL.

elongation is vital for preventing high myopia and associated complications, such as posterior scleral staphyloma (Li et al., 2024). Crucially, AL elongation induces both refractive changes and distinct morphological alterations in the fundus. Eyes with longer ALs often exhibit discernible abnormalities on color fundus photographs (CFPs), such as fundus tessellation (FT), peripapillary atrophy (PPA), diffuse or patchy choroidal-retinal atrophy (Hayashi et al., 2010), as well as modifications around the optic nerve and straightening of retinal vessels. These established morphometric associations between AL and fundus structure provide a biological foundation for correlating AL with features captured in CFPs.

Deep learning (DL), particularly convolutional neural networks (CNNs), has been increasingly applied to analyze CFPs for myopia-related tasks (Qian et al., 2024; Meng et al., 2024; Prashar and Tay, 2024). As DL and artificial intelligence (AI) methods continue to be combined with CFP and other data, screening data can be used to classify and perform supervised learning (Liu et al., 2022; Wang et al., 2024). DL excels at identifying subtle, complex patterns within images that may elude conventional analysis. For instance, Wei et al. utilized an AI-assisted method to quantify choroidal macrovascular areas in CFPs and found a significant correlation with AL (Wei et al., 2023). Research has also examined incorporating demographic factors, such as age and sex, alongside CFP data to enhance the predictive accuracy of ocular parameters, including AL (Yamashita et al., 2020; Zhu et al., 2022). However, despite these advances, no studies have specifically developed CNN models to predict AL using CFPs obtained during routine myopia screening in school-aged children– a population where early AL monitoring is paramount, but tools like optical biometers are often inaccessible in screening settings (Cao et al., 2023; Zhang et al., 2024; Gu et al., 2024).

Thus, the main aim of this study was to develop a CNN-based model for predicting AL in 6–10-year-old schoolchildren using CFP and basic information including age and/or sex, and the results were used to explore the correlation between AL, sex and age-related features, thus broadening the role of normal fundus structure in predicting the AL values in the development of myopia, and discover more ocular information potentially contained in the CFP.

2. Methods

2.1. Data collection and preprocessing

The design of models and selection of CFPs is illustrated in Fig. 1. The study was approved by the Ethics Committee of Tianjin Medical University Eye Hospital and conducted in accordance with the principles outlined in the Declaration of Helsinki (2024KY-67). Patients with basic refractive information and CFP were selected from the Tianjin Myopia Screening 2023–2025 for school children aged 6–10 years. Children using myopia control interventions, including orthokeratology lenses or defocus glasses, were included in the study, as the model primarily focuses on the structural relationship between the fundus and AL. Refraction measurements were conducted after cycloplegia induced by three drops of 1 % cyclopentolate hydrochloride administered at 5-min intervals, with measurements taken 30 min after the final drop (Hu et al., 2025). The SE range for inclusion was chosen from −10D to +10D to ensure maximum coverage of refractive error in children. Exclusion criteria include any retinal or choroidal disease observable in the CFP, as well as patients with other ocular diseases affecting the ocular surface. CFP is included to visualize the optic disc and central macular area. In 7680 fundus photographs, three experienced ophthalmologists reviewed the images to ensure the clarity and integrity of the macular, optic disc, and vascular information. Finally, 5460 CFPs from 5460 eyes of 2779 patients were included. These 5460 images included the central macula and optic disc from all eyes. Afterwards, we extracted age, sex, cycloplegia refraction, and AL data for these children from the screening report. The Lenstar LS-900 optical biometer (Haag-Streit AG, Koeniz, Switzerland) was used for all AL measurements. All CFP images were examined using a 45° fundus camera (Canon Inc., 9-1, Kanagawa, Japan). Refraction measurements after cycloplegia were conducted by using the autorefraction (FKR.800, Topcon, Tokyo, Japan).

Our model employed two input sources: structured data (age, sex, and cycloplegic refraction) and CFPs. Regarding the sex information, we labeled males as 0 and females as 1 as sex-specific input values. Previous studies have highlighted the importance of age in predicting AL, so we included age as a factor (Che-Ning et al., 2024). In addition, we incorporated diopter sphere (DS) degrees as one of the input data. Regarding

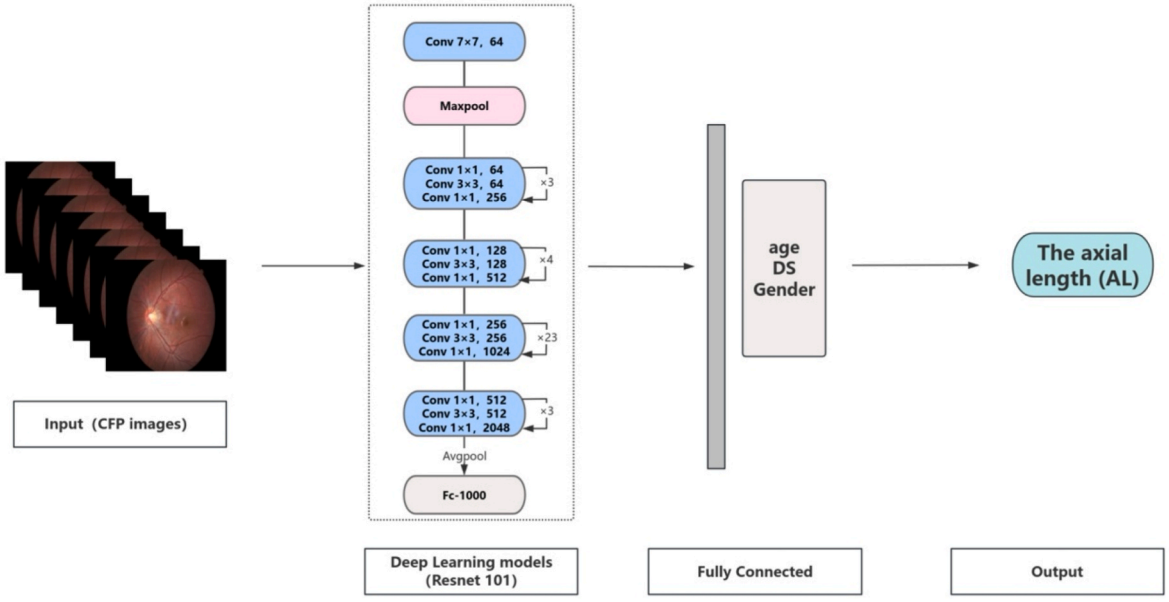

Fig. 2. The process for model development. Image input was preprocessed and pretrained with the deep learning model "Resnet101". The training was conducted for 100 epochs on a GPU-enabled system, utilizing the nn library. Data Parallel for multi-GPU training. At each epoch, the model underwent both training and validation. The training process included forward propagation, loss computation, backpropagation, and optimization updates. The validation loss was computed at the end of each epoch, and the model with the best performance was saved for later evaluation.

the CFPs, we preprocessed the images at 2736 × 1824 pixels and resized them to 224 × 224 pixels for model input. The images were resized to 224 × 224 pixels to standardize the input for the pre-trained ResNet101 model, ensuring efficient feature extraction while maintaining diagnostic integrity. To validate and compare the predictive efficacy of CFP for AL and determine whether including additional necessary information can further enhance prediction, we constructed four models. Model l contained structured data (age and DS), while Model 2 included CFP only. Model 3 included CFP, age, and DS. Model 4 included sex, age, CFP, and DS. The outputs are all AL.

2.2. Model development

The architecture for model development is illustrated in Fig. 2. A deep learning-based regression model was developed using ResNet101 to predict AL from fundus images. The dataset was split into training, validation, and test sets in a ratio of 7:2:1. The dataset was transformations (Converts an image to a tensor, scales pixel values to [0, 1] if they exceed 1, and normalizes using ImageNet mean & std (for pre-trained models) and loaded using PyTorch's DataLoader with a batch size of 192 for efficient batch processing. The ResNet101 model was initialized with pre-trained weights, and its fully connected layer was progressively added to incorporate additional features. (The pre-trained weights for ResNet-101 (10.48550/arXiv 0.1512.03385) were sourced from PyTorch's model zoo (torchvision.models), which were originally trained on the ImageNet-1k dataset (ILSVRC 2012).) These weights are the default ones provided by PyTorch when pre-trained = True is used. Model 2 used only image data as input, predicting AL solely from fundus images. In Model 3, DS and age were introduced as additional inputs alongside image features. The Model 4 further incorporated sex as an additional variable. A dropout rate of 0.25 was applied before the final prediction layer to enhance generalization and reduce overfitting. Previous studies have employed the Mean Absolute Error (MAE) method; however, in neural networks, the MAE update gradient remains constant, which is not conducive to model learning (Che-Ning et al., 2024). We use mean squared error (MSE) to improve the model's accuracy and better align with our training objectives. The model was trained using the Adam optimizer (learning rate = 0.001) and the MSE loss function. A learning rate scheduler (Step LR) was employed to reduce the learning rate by a factor of 0.5 every 10 epochs.

The training was conducted for 100 epochs on a GPU-enabled system (NVIDIA Tesla P100), utilizing the nn library. DataParallel for multi-GPU training. At each epoch, the model was trained and validated. This training strategy ensured optimal convergence and prevented overfitting, ultimately improving the model's predictive accuracy.

2.3. Heatmap generation

Gradient-weighted Class Activation Mapping (Grad-CAM) is a visualization technique that interprets the decision-making process of CNNs by highlighting the regions of the input image most influential to the model's predictions (Selvaraju et al., 2017). In this study, Grad-CAM was used to analyze the model's attention in a ResNet-based regression model trained to predict AL from fundus images. The target layer for Grad-CAM was selected as the last convolutional layer to capture high-level spatial features. During the forward pass, the activations of the target layer were extracted, and during the backward pass, the gradients of the output with respect to the activations were computed. These gradients were then globally averaged to produce a weight map, which was combined with the activations to generate heatmap. The heatmap was normalized to the range [0, 1] and superimposed on the original image using a color map (cv2.COLORMAP_JET) to visualize the regions of interest. The implementation ensured that the model was set to evaluation mode (model.eval ()) to disable dropout and batch normalization layers during inference. This approach provided interpretable insights into the model's behavior, demonstrating its ability to focus on clinically relevant regions of the fundus images.

2.4. Statistical analysis

Baseline characteristics were expressed as means ± standard deviations (SD) for continuous variables and as counts (proportions) for categorical variables. To evaluate the performance of the deep learning regression models (Models 2–4), we utilized the Coefficient of Determination (R^2) and MSE. R^2 represents the proportion of variance for AL explained by the model and is defined as:

Table 1
Baseline characteristics of the data set.

Characteristic	Value
Number of children	2779 (5460 eyes)
Sex, male, %	1306, 47.4 %
female,%	1473,52.6 %
Number of CFPs	5460
Axial length, mean ± SD, mm	23.18 ± 0.84
Minimum and maximum, mm	19.84, 26.84
<22 mm	360 (6.6 %)
≥22 mm and <24 mm	4218 (77.2 %)
≥24 mm	882 (16.2 %)
DS, mean ± SD, D	0.64 ± 1.37
Minimum and maximum, D	−5.75, +9.50
Age, mean ± SD, y	7.7 ± 0.96
Minimum and maximum, y	6, 10

CFP: color fundus photography; DS: diopter sphere; SD: Standard Deviation.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

The MSE, which measures the average squared difference between predicted and actual values, was calculated as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Statistical analysis and model implementation were performed using Python (version 3.11). A Bland-Altman analysis was performed to assess the agreement between the predicted and measured AL values, with the limits of agreement defined as the mean difference ±1.96 standard deviations. These plots and limits of agreement (mean difference ± 1.96 SD) were generated using the matplotlib and seaborn libraries.

3. Results

3.1. Data characteristics

Table 1 summarizes the eye characteristics. Initially, 3840 school-age children were included in this study. Of the 7680 images, 2220 (28.91 %) were deemed of relatively poor quality and excluded. Finally, this study included 5460 fundus images from 5460 eyes of 2779 children. Of the 2779 children, the mean age was 7.7 ± 0.96 years (range, 6–10 years). Of the 2779 children, 1473 were female (52.6 %), and 1306 were male (47.4 %). All the 5460 eyes had macular and optic disc-centered images included. The mean AL of the included eyes was 23.18 ± 0.84 mm (range, 19.84–26.84 mm).

3.2. Models for predicting AL using different input parameters

Fig. 3 analyzes the consistency evaluation of the true and predicted AL values in Models 2–4. The R^2 and MSEs for all models with different input parameters (CFP, age, DS, and/or sex) are presented in Table 2, labeled 1–4. In Model 1, R^2 is only 0.37, and MSE reaches 0.46 (mm^2). When only CFP is entered as an eigenvalue, as in Model 2, R2 can reach 0.70, but the MSE is 0.21 (mm^2). After adding DS and age to CFP, Model 3 R^2 improves to 0.75, and the MSE decreases to 0.17 (mm^2). Adding sex to Model 3 as Model 4 resulted in an R^2 of only 0.41 and an MSE of 0.40 (mm^2).

3.3. Heatmap generation

Figs. 4–6 show heatmaps of CFP predictions with varying accuracies for AL. In Fig. 4, we present the three patients with CFP who were accurately predicted to have AL with up to 100 % accuracy. Their ALs were 23.20 mm, 23.51 mm, and 23.77 mm, respectively. In Fig. 5, we selected the three patients with CFP who predicted AL with an accuracy near the median of the overall accuracies. Their predicted AL values were 22.89 mm, 23.01 mm, and 22.68 mm, respectively, while their true AL values were 23.06 mm, 22.87 mm, and 22.97 mm, respectively. Finally, we selected two heat maps: patients whose true AL values were greater than the predicted values (A and B), and those whose true AL values were less than the predicted values (C and D), as shown in Fig. 6. Their predicted AL values were 22.99 mm, 22.84 mm, 23.30 mm, and 22.98 mm, respectively. However, their true AL values were 25.04 mm, 26.76 mm, 22.00 mm, and 21.16 mm, respectively.

4. Discussion

Our study demonstrates that DL models can effectively predict AL in school-aged children using CFPs, even in eyes with a largely normal fundus appearance. Critically, CFP alone (Model 2, R^2 = 0.70) significantly outperformed predictions based solely on DS and age (Model 1, R^2 = 0.37). Furthermore, incorporating DS and age alongside CFP yielded the highest predictive accuracy (R^2 = 0.75). Conversely, adding sex information to this combined model (Model 4, R^2 = 0.41)

Table 2
R^2, MSE and SD for the 4 prediction models.

Model	Model descriptions	R2	MSE (mm2)	SD (mm2)
1	DS, age	0.37	0.46	±0.67
2	CFP	0.70	0.21	±0.45
3	CFP, DS and age	0.75	0.17	±0.41
4	CFP, DS, age and sex	0.41	0.40	±0.63

CFP: color fundus photography; DS: diopter sphere; SD: Standard Deviation.

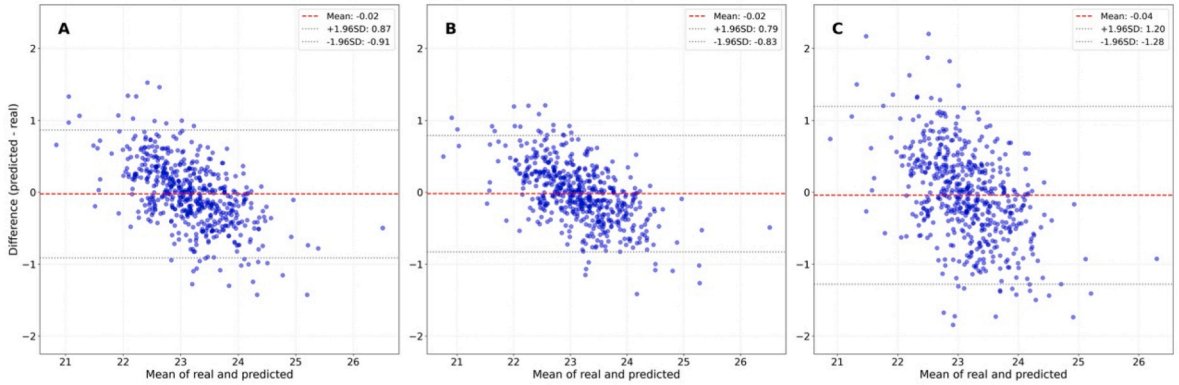

Fig. 3. The Bland-Altman plots show the distribution of predicted and actual AL in a test of Model 2 (A), Model 3 (B), and Model 4 (C). Eyes with long ALs tended to be underestimated, while those with short ALs tended to be overestimated.

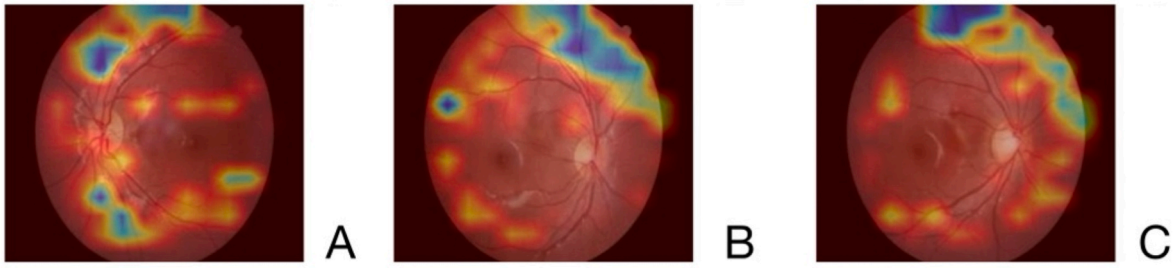

Fig. 4. Examples showing 3 of the best prediction eyes' heatmaps of integrated gradients from model 2. (A: Actual AL:23.20 mm, Predicted AL:23.20 mm; B: Actual AL:23.51 mm, Predicted AL:23.51 mm; C: Actual AL:23.77 mm, Predicted AL:23.76 mm) These heatmaps show that attention is focused on the retinal vasculature and the perivascular area.

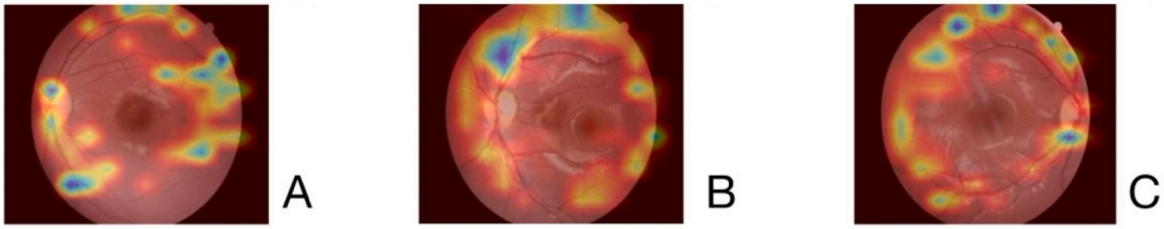

Fig. 5. Examples showing 3 of the average-accuracy-prediction eyes' heatmaps of integrated gradients from model 2. (A: Actual AL:23.06 mm, Predicted AL:22.89 mm; B: Actual AL:22.87 mm, Predicted AL:23.01 mm; C: Actual AL:22.97 mm, Predicted AL:22.68 mm) These heatmaps are characterized by a focus of attention on the area around the macular, away from the retinal vessels.

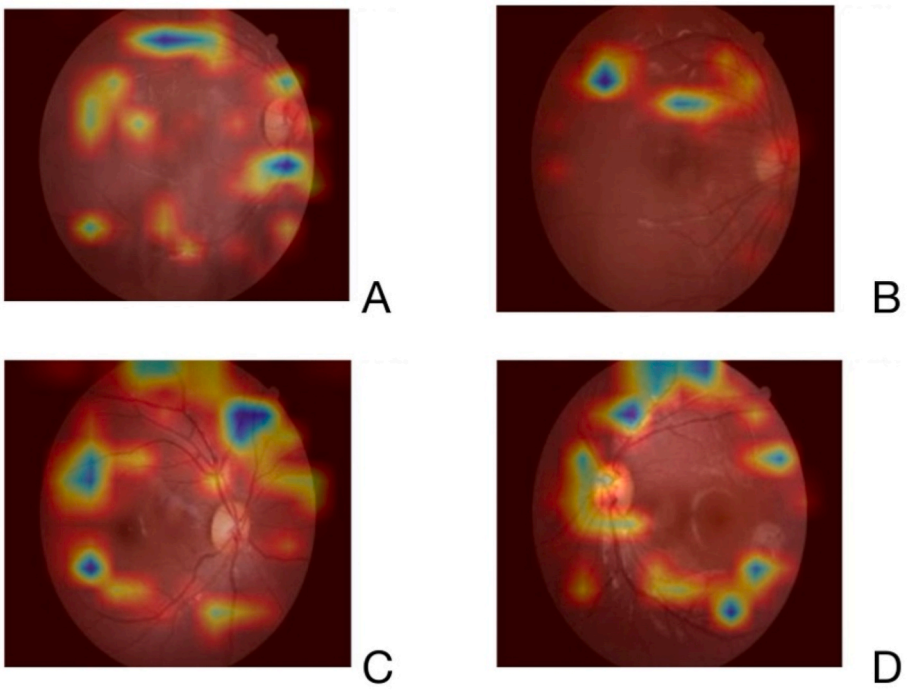

Fig. 6. Examples showing 2 heat maps of two patients whose true values of AL were greater than the predicted values (A: Actual AL:25.04 mm, Predicted AL:22.99 mm; B: Actual AL:26.76 mm, Predicted AL:22.84 mm) and whose true values of AL were less than the predicted values (C: Actual AL:22.00 mm, Predicted AL:23.30 mm; D: Actual AL:21.16 mm, Predicted AL:22.98 mm) The distraction and lack of attention can be observed in the heat maps, which may be related to the more extreme AL values and lack of training.

substantially reduced performance, yielding accuracy comparable to that of the DS and age-only models. These findings suggest the potential of CFP for estimating the ocular axis in pediatric myopia screening and for selecting input parameters for the DL model, both of which are necessary considerations.

In the present study, our DL models for AL prediction using only CFP achieved an R^2 of 0.70. To facilitate comparison with similar studies, we summarize the different studies in Table 3. Li et al. reported an R^2 of 0.59 for their deep learning model, which included participants aged 50 years or older (Dong et al., 2021). The cohort had a mean ± SD AL of

Table 3
Performance comparison of different CFP prediction models for AL research.

Author	Model descriptions	R^2	MAE	MSE
Che-NingYang et al.	CFP	/	0.831	/
Dong Li et al.	CFP	0.59	0.56	/
Yunzhe Wang et al.	CFP (Ultra-wide angle)	0.57	1.41	0.90

CFP: color fundus photography.

23.24 ± 1.15 mm and a low proportion of long eyes (3.1 % of eyes with an AL ≥26 mm). The predictive efficacy is high for myopia screening in school-aged children and is clinically relevant in real-world scenarios. The baseline eye axis and degree of refractive error in children aged 6–10 years have been reported as independent risk factors for the later development of high myopia and pathological changes (Han et al., 2022; Lanca et al., 2025). The latest cutting-edge myopia management tools, such as low-level red-light therapy and atropine therapy, are also centered on axial control (Zhang et al., 2024). Early monitoring and control of AL are crucial for preventing the progression of myopia in school-age children (Liu et al., 2024). Additionally, CFPs are a key component in screening programs for both children and adults, and an increasing number of portable fundus cameras are in use. However, tools for AL measurements are still not readily available, and our model offers the potential for a reduced workflow for myopia screening tasks with a large base. Recently, Richul Oh et al. developed a model to predict AL using both horizontal and vertical optical coherence tomography images, with MAE and R^2 values of 0.592 and 0.847 in the internal test set and 0.556 and 0.663 in the external test set (Oh et al., 2024). It worked well in images, such as scleral staphyloma, for larger fundus lesions with significant changes in fundus curvature. However, AL performs poorly on OCT images without obvious fundus abnormalities (Oh et al., 2024). In children with essentially normal funds, OCT is not as adaptable as CFPs, both in terms of ease of access and necessity.

In terms of model comparisons, the predictive efficacy of model 2 improved when age and DS were added relative to CFP alone (R^2 = 0.75). Yang et al. proposed that age is a significant predictor of AL (Yang et al., 2024). This is consistent with our results. Interestingly, the predictions also improved when we included DS. The reason we did not improve predictive efficacy for the age factor compared to previous studies may be due to the population selection. Previous studies have not focused on predicting fundus and AL in children; instead, they have discussed the performance of CFP in predicting AL in the elderly and in populations spanning a wide age range (Ilanchezian et al.; Korot et al., 2021). In fact, the potential advantages of this study lies in the complementary role of DL itself to medicine, as we selected children of similar age and with roughly similar fundus, and in actual clinical practice, such CFP may not bring ophthalmologists more diagnostic hints, but after DL, the possibility that the CFP can provide information about AL is a surprise for this study. Including the age factor improved the model's performance. However, because we included children aged 6–10 years, the age span was not significant, and it is possible that age did not play a significant role in the prediction.

In addition, we took sex into account. Indu Ilanchezian et al. proposed that sex is an important factor in the accuracy of CFP DL models (Ilanchezian et al.). Even Korot et al. used DL to distinguish sex from CFPs (Korot et al., 2021). This seems to prove that sex is one of the key messages in the CFP. Our study compared the accuracy of including a sex factor in AL prediction. This result differs significantly from previous studies. Model 4 R^2 decreased, but MSE increased after adding sex (R^2 = 0.4109, MSE = 0.4051). Takehiro et al. used CFP to predict sex and found that sex differences did not begin until about 10 years of age, which is consistent with our findings (Yamashita et al., 2023). We propose that the CFP may contain DL-recognizable information about sex characteristics. However, adding sex information to the CFP base instead reduces the accuracy of the CFP in predicting AL. This suggests that the correlation between myopia and DL is not significant in children aged 6–10 years. This suggests that we carefully select the input to incorporate into model training. Adding too much information may reduce the model's accuracy, rather than increasing it. In addition, the previous study focused on the sex prediction of CFP in adult pairs, considering that the adult fundus is more fully developed, while the CFP of children aged 6–10 years old is still in the process of development, in our model, it is important to note that the CFP information of children of a younger age is not a significant indicator of sex in future studies on the extraction of sex-related information from CFP.

Currently, CFP-related AI focuses on identifying and detecting individuals with myopic fundus abnormalities. Guo et al. quantified FT density and optic disc morphology using DL techniques to investigate the correlations between these fundus characteristics and refractive function in young patients with myopia (Guo et al., 2024). Gong et al. similarly found that FT density (FTD) is significant for the detection and prediction of myopic macular degeneration (Gong et al., 2024). In another study of school-aged children, the findings indicated that those exhibiting a macular FT pattern not only presented with longer AL but also more pronounced myopic refraction, compared to their peers with a peripapillary FT pattern (Huang et al., 2025). Wei et al. found a correlation between FT density and myopia progression in Chinese children aged 7–9 years. However, their study used ophthalmologists in a semi-supervised mode to calibrate FT density (Li et al., 2025). In our study population, more than 90 % of children had normal fundus without significant FT and PPA. We know that a phase of diffuse thinning and choroidal atrophy precedes the FT phase. Such a period is often difficult to detect in CFP (Zhao et al., 2024). We believe that it may be such changes processed by DL that can be predictive of the AL. In addition, our study was not only a dichoptic task but also a direct prediction of AL, which further reinforces the role of DL in processing information difficult for the human eye to recognize. We propose the hypothesis that even in a population of children with normal optic discs and maculae, subtle changes in detailed fundus-related parameters may occur with increasing AL. Such changes may result from variations in vascular morphology, including the angle of travel and the bifurcation angles of each vessel. Zhao et al. proposed a data analysis model that combines a DL model with an attention module to classify images, segment vascular structures, and measure vascular parameters, including the main angle, branching angle, bifurcation edge angle, and bifurcation edge coefficient (Zhao et al., 2024). They found that the progression of myopia is associated with a series of quantitative retinal vascular parameters, particularly the vascular angles (Zhao et al., 2024).

Subtle color differences in the CFP may be another important basis for DL to predict AL. Seiji et al. observed the morphology, area, and color of the optic disc in 75 students aged 8–9 years over a six-year follow-up period. They found a correlation between optic disc parameters and changes in the eye axis during this period (Sameshima et al., 2024). Wulff et al. recently found that optical differences between cameras can also lead to significant geometric differences between fundus photographs (Wulff et al., 2025). Our study used the same camera model, which helps avoid differences in CFP across devices. However, the color of fundus photographs is affected by many factors. In the future, it may be necessary to explore ways to improve the stability of the CFP prediction model across devices. Their results showed that both a smaller optic disc area and a gradual deepening of color were positively correlated with eye axis growth (Sameshima et al., 2024). In other words, it is likely that there is no smaller optic disc area, but that this area appears smaller due to the decreased magnification for myopic eyes. Both in terms of age and population, this study complements ours. It also provides another hypothesis to make the model interpretable. A magnification difference could cause changes in the above due to the different optical properties of the myopic eye.

Additionally, this study uses a heatmap to visualize the accuracy of the deep learning model's results. First, it is important to note that the accuracy remains dependent on the data scope. The model demonstrates reduced accuracy at AL values above 26 mm and below 22 mm, which

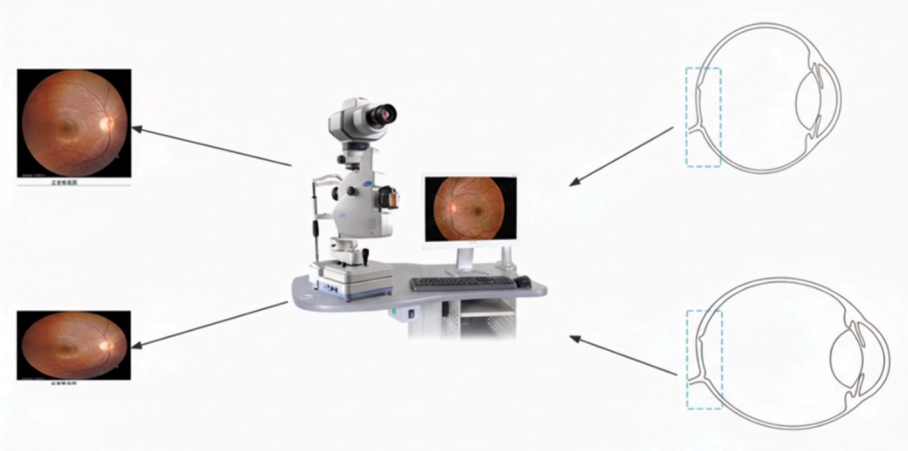

Fig. 7. Illustrative diagrams of possibilities for phenomena for which the CFP can explain heat maps.

corresponds to the distribution range of the training dataset. Interestingly, upon analyzing the heatmap, we observed that for school-age children with nearly normal CFP, the DL model did not focus solely on the macula and optic disc, as was the case in previous studies. Attention was also directed toward the retinal secondary vessels and surrounding areas, thereby improving the accuracy of AL predictions. We propose that, in young children with minimally abnormal CFPs, the macula and optic disc serve as morphological anchors ("static objects") for deep learning models, given their structural prominence and limited microstructural variance during the initial phase of AL elongation, as illustrated in the conceptual diagrams in Fig. 7. In contrast, retinal vasculature and perivascular tissues exhibit relative biomechanical displacement in response to AL increase, manifesting as quantifiable feature shifts detectable by models against the stabilized background. This vascular dynamism is further amplified by AL-induced alterations in vessel distribution patterns. Finally, incorporating the nasal optic nerve region degraded prediction accuracy—an effect potentially attributable to positional centrality bias in model attention, as evidenced by activation heatmaps centered on the optic disc.

This study has some limitations. First, although we strictly adopted a 7:2:1 data split for training, validation, and testing, we did not perform external validation across regions with different geographic locations. This may compromise the model's robustness. In addition, our predicted ALs were primarily for children screened at 6–10 years of age; therefore, most children had ALs between 19.84 mm and 26.84 mm, and there was insufficient data to train individual abnormal axial lengths using fundus images. In the future, we will further improve the model's performance by incorporating data from different age groups and high myopia populations. In addition, although the CFP-predicted AL in our study had an $R^2 > 0.7$, there is still a substantial discrepancy between the true and predicted values, and measuring the ocular axis remains indispensable in clinical testing at this time. We are currently focusing on the potential predictive value of specific annotations, such as CFP vascularity and AL. Future studies could further enhance the model's predictive power by incorporating additional variables such as visual acuity, parental history of myopia, and daily outdoor activity time, which were not included in the current screening-focused model.

5. Conclusions

In conclusion, we trained four AL prediction models using CFP, age, sex, and DS in 6- to 10-year-old schoolchildren. Among them, CFP alone was a potential predictor of AL, outperforming the model that used only age and DS. Adding age and DS as additional predictors to CFP would further improve the prediction accuracy and model stability. In addition, incorporating age, DS, and sex into the model along with CFP resulted in decreased predictive performance. These findings suggest that careful consideration should be given when selecting predictors for AL prediction in school-aged children, as adding extraneous factors does not necessarily improve model performance. Based on heat map observations, we found that in school-aged children with normal fundi, retinal vessels, and peripheries, retinal sensitivity to AL changes increased with increasing AL.

CRediT authorship contribution statement

Zixun Wang: Writing – original draft, Validation, Resources, Methodology, Investigation, Data curation. **Hua Rong:** Visualization, Software, Project administration, Investigation, Formal analysis. **Jingtao Yu:** Software, Project administration, Methodology, Data curation. **Yifan Zhou:** Investigation, Formal analysis. **Xueshuo Xie:** Supervision, Software, Resources. **Weiping Lin:** Validation, Software. **Zhiqing Li:** Writing – review & editing. **Hua Yan:** Writing – review & editing, Visualization. **Bei Du:** Writing – review & editing, Supervision, Funding acquisition. **Ruihua Wei:** Writing – review & editing, Funding acquisition, Conceptualization.

Ethics approval and consent to participate

This study was approved by the ethics committee of the Tianjin Medical University Eye Hospital [No. 2024KY-67]. All study procedures adhered to the tenets of the Declaration of Helsinki.

Consent for publication

Not applicable.

Funding

This study was supported by the Tianjin Key Medical Discipline Construction Project (TJYXZDXK-3-004A-2) and Haihe Lab of ITAI Science and Technology Project (25HHXCSS00007).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Not applicable.

Abbreviations

CFP	color fundus photography
AI	artificial intelligence
AL	axial length
FT	fundus tessellation
PPA	peripapillary atrophy
DL	deep learning
CNN	Convolutional neural network
DS	diopter sphere
MAE	mean absolute error
MSE	mean square error
Grad-CAM	gradient-weighted class activation mapping

Data availability

Data will be made available on request.

References

Bullimore, M.A., Ritchey, E.R., Shah, S., Leveziel, N., Bourne, R.R.A., Flitcroft, D.I., 2021. The risks and benefits of myopia control. Ophthalmology 128 (11), 1561–1579. https://doi.org/10.1016/j.ophtha.2021.04.032.

Burton, M.J., Ramke, J., Marques, A.P., et al., 2021. The lancet global health commission on global eye health: vision beyond 2020. Lancet Global Health 9 (4), e489–e551. https://doi.org/10.1016/S2214-109X(20)30488-5.

Cao, S., Zhang, R., Jiang, A., Kuerban, M., Wumaier, A., Wu, J., Xie, K., Aizezi, M., Tuersun, A., Liang, X., Chen, R., 2023. Application effect of an artificial intelligence-based fundus screening system: evaluation in a clinical setting and population screening. Biomed. Eng. Online 22 (1), 38. https://doi.org/10.1186/s12938-023-01097-9.

Dong, L., Hu, X.Y., Yan, Y.N., et al., 2021. Deep learning-based estimation of axial length and subfoveal choroidal thickness from color fundus photographs. Front. Cell Dev. Biol. 9, 653692. https://doi.org/10.3389/fcell.2021.653692.

GBD 2019 Blindness and Vision Impairment Collaborators; Vision Loss Expert Group of the Global Burden of Disease Study, 2021. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the right to sight: an analysis for the global burden of disease study. Lancet Global Health 9 (2), e144–e160. https://doi.org/10.1016/S2214-109X(20)30489-7. Epub 2020 Dec 1. Erratum in: Lancet Glob Health. 2021 Apr;9(4):e408. doi: 10.1016/S2214-109X(21)00050-4.

Gong, W., Wang, J., Deng, J., et al., 2024. Quantification of fundus tessellation reflects early myopic maculopathy in a large-scale population of children and adolescents. Transl. Vis. Sci. Technol. 13 (6), 22. https://doi.org/10.1167/tvst.13.6.22.

Gu, C., Wang, Y., Jiang, Y., Xu, F., Wang, S., Liu, R., Yuan, W., Abudureyimu, N., Wang, Y., Lu, Y., Li, X., Wu, T., Dong, L., Chen, Y., Wang, B., Zhang, Y., Wei, W.B., Qiu, Q., Zheng, Z., Liu, D., Chen, J., 2024. 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. 108 (3), 424–431. https://doi.org/10.1136/bjo-2022-322940.

Guo, Z., Chen, L., Wang, L., et al., 2024. Automated measurement and correlation analysis of fundus tessellation and optic disc characteristics in myopia. Sci. Rep. 14 (1), 28399. https://doi.org/10.1038/s41598-024-80090-1.

Han, X., Liu, C., Chen, Y., He, M., 2022. Myopia prediction: a systematic review. Eye Lond Engl 36 (5), 921–929. https://doi.org/10.1038/s41433-021-01805-6.

Hayashi, K., Ohno-Matsui, K., Shimada, N., et al., 2010. Long-term pattern of progression of myopic maculopathy. Ophthalmology 117 (8), 1595–1611.e4. https://doi.org/10.1016/j.ophtha.2009.11.003.

Holden, B.A., Fricke, T.R., Wilson, D.A., et al., 2016. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology 123 (5), 1036–1042. https://doi.org/10.1016/j.ophtha.2016.01.006.

Hu, X., Wang, Z., Sun, B., Wan, X., Chang, F., Lin, W., Wei, R., 2025. Orthokeratology for myopia control: a three-year longitudinal study on axial length modulation influenced by baseline age and spherical equivalent. BMC Ophthalmol. 25 (1), 466. https://doi.org/10.1186/s12886-025-04303-5.

Huang, D., Lin, X., Zhu, H., Ling, S., Dong, Z., Ke, X., Long, T., Qian, Y., Yan, Q., Li, R., Zhong, H., Liu, H., 2025. The associations between myopia and fundus tessellation in school children: a comparative analysis of macular and peripapillary regions using deep learning. Transl. Vis. Sci. Technol. 14 (1), 4. https://doi.org/10.1167/tvst.14.1.4.

Ilanchezian I, Kobak D, Faber H, Ziemssen F, Berens P, Ayhan MS. Interpretable Sex Classification from Retinal Fundus Images Using BagNets. Published online June 25, 2021. doi:10.1101/2021.06.21.21259243.

Kakita, T., Hiraoka, T., Oshika, T., 2011. Influence of overnight orthokeratology on axial elongation in childhood myopia. Investig Opthalmology Vis Sci. 52 (5), 2170. https://doi.org/10.1167/iovs.10-5485.

Korot, E., Pontikos, N., Liu, X., et al., 2021. Predicting sex from retinal fundus photographs using automated deep learning. Sci. Rep. 11 (1), 10286. https://doi.org/10.1038/s41598-021-89743-x.

Li, H., Gao, N., Li, R., Luodian, L., Sui, J., Bai, Y., Wu, D., He, Q., Wang, Y., Li, Z., Wei, R., 2024. Microcirculatory parameters as risk factors for predicting progression of posterior staphyloma in highly myopic eyes: a case-control study. Eye Vis (Lond). 11 (1), 45. https://doi.org/10.1186/s40662-024-00413-1.

Li, D, Lanca, C, Zhang, X, Grzybowski, A, He, X, Pan, C, 2025. Spatial frequency of environments and myopia: a systematic review on associated evidence and underlying mechanisms. Acta Ophthalmol. https://doi.org/10.1111/aos.17437. Published online January 3, :aos.17437.

Li, C., Wu, H., Xiao, O., et al., 2025. Eye shape deformity predicts myopic maculopathy progression among highly myopic individuals: a 4-year longitudinal study. Retina 45 (1), 52–60. https://doi.org/10.1097/IAE.0000000000004281.

Liu, R., Li, Q., Xu, F., Wang, S., He, J., Cao, Y., Shi, F., Chen, X., Chen, J., 2022. Application of artificial intelligence-based dual-modality analysis combining fundus photography and optical coherence tomography in diabetic retinopathy screening in a community hospital. Biomed. Eng. Online 21 (1), 47. https://doi.org/10.1186/s12938-022-01018-2.

Liu, Z., Sun, Z., Du, B., Gou, H., Wang, B., Lin, Z., Ren, N., Pazo, E.E., Liu, L., Wei, R., 2024. The effects of repeated low-level red-light therapy on the structure and vasculature of the choroid and retina in children with premyopia. Ophthalmol. Ther. 13 (3), 739–759. https://doi.org/10.1007/s40123-023-00875-x.

Meng, J., Song, Y., He, W., Lu, Z.L., Chen, Y., Wei, L., Zhang, K., Qi, J., Du, Y., Lu, Y., Zhu, X., 2024. A novel artificial intelligence-based classification of highly myopic eyes based on visual function and fundus features. Transl. Vis. Sci. Technol. 13 (9), 12. https://doi.org/10.1167/tvst.13.9.12.

Oh, R., Kang, M., Ahn, J., et al., 2024. Prediction of axial length from macular optical coherence tomography using deep learning model. Transl. Vis. Sci. Technol. 13 (9), 14. https://doi.org/10.1167/tvst.13.9.14.

Prashar, J., Tay, N., 2024. Performance of artificial intelligence for the detection of pathological myopia from color fundus images: a systematic review and meta-analysis. Eye (Lond). 38 (2), 303–314. https://doi.org/10.1038/s41433-023-02680-z.

Qian, B., Sheng, B., Chen, H., Wang, X., Li, T., Jin, Y., Guan, Z., Jiang, Z., Wu, Y., Wang, J., Chen, T., Guo, Z., Chen, X., Yang, D., Hou, J., Feng, R., Xiao, F., Li, Y., El Habib Daho, M., Lu, L., Ding, Y., Liu, D., Yang, B., Zhu, W., Wang, Y., Kim, H., Nam, H., Li, H., Wu, W.C., Wu, Q., Dai, R., Li, H., Ang, M., Ting, D.S.W., Cheung, C. Y., Wang, X., Cheng, C.Y., Tan, G.S.W., Ohno-Matsui, K., Jonas, J.B., Zheng, Y., Tham, Y.C., Wong, T.Y., Wang, Y.X., 2024. A competition for the diagnosis of myopic maculopathy by artificial intelligence algorithms. JAMA Ophthalmol. 142 (11), 1006–1015. https://doi.org/10.1001/jamaophthalmol.2024.3707.

Sameshima, S., Yamashita, T., Terasaki, H., et al., 2024. Longitudinal changes of funduscopic optic disc size, color and cup-to-disc ratio in school children. Int. J. Retina Vitr. 10 (1), 51. https://doi.org/10.1186/s40942-024-00570-4.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2017. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 618–626. https://doi.org/10.1109/ICCV.2017.74. Venice, Italy.

Wang, Y., Han, X., Li, C., Luo, L., Yin, Q., Zhang, J., Peng, G., Shi, D., He, M., 2024. Impact of gold-standard label errors on evaluating performance of deep learning models in diabetic retinopathy screening: nationwide real-world validation study. J. Med. Internet Res. 26, e52506. https://doi.org/10.2196/52506.

Wei R, Li J, Yang W, et al. Association of tessellation density with progression of axial length and refraction in children: an AI-assisted 4-year study. Retina. Published online November 7, 2023. doi:10.1097/IAE.0000000000003991.

Wulff, J., Koska, B., Giese, M., Bäumer, C., Richter, R., Foerster, A., Bechrakis, N.E., Timmermann, B., 2025. Evaluation of OPTOS wide-field fundus image projections for radiotherapy planning of uveal melanoma. J. Appl. Clin. Med. Phys. 26 (4), e70009. https://doi.org/10.1002/acm2.70009. Epub 2025 Feb 17.

Yamashita, T., Asaoka, R., Terasaki, H., Yoshihara, N., Kakiuchi, N., Sakamoto, T., 2023. Three-year changes in sex judgment using color fundus parameters in elementary school students. PLoS One 18 (11), e0295123.

Yamashita, T., Asaoka, R., Terasaki, H., et al., 2020. Factors in color fundus photographs that can be used by humans to determine sex of individuals. Transl. Vis. Sci. Technol. 9 (2), 4. https://doi.org/10.1167/tvst.9.2.4.

Yang, C.N., Chen, W.L., Yeh, H.H., Chu, H.S., Wu, J.H., Hsieh, Y.T., 2024. Convolutional neural network–based prediction of axial length using color fundus photography. Transl. Vis. Sci. Technol. 13 (5), 23. https://doi.org/10.1167/tvst.13.5.23.

Zhang, J., Lin, S., Cheng, T., et al., 2024. RETFound-enhanced community-based fundus disease screening: real-World evidence and decision curve analysis. npj Digit. Med. 7 (1), 108. https://doi.org/10.1038/s41746-024-01109-5.

Zhang, X.J., Zhang, Y., Yip, B.H.K., et al., 2024. Five-year clinical trial of the low-concentration atropine for myopia progression (LAMP) study: phase 4 report. Ophthalmology 131 (9), 1011–1020. https://doi.org/10.1016/j.ophtha.2024.03.013.

Zhao, Y., Zhao, Z., Yang, J., Li, L., Nasseri, M.A., Zapp, D., 2024. AI-based fully automatic analysis of retinal vascular morphology in pediatric high myopia. BMC Ophthalmol. 24 (1), 415. https://doi.org/10.1186/s12886-024-03682-5.

Zhu, Z., Hu, W., Chen, R., et al., 2022. Retinal age gap as a predictive biomarker of stroke risk. BMC Med. 20 (1), 466. https://doi.org/10.1186/s12916-022-02620-w.