Ophthalmol Ther (2024) 13:305–319
https://doi.org/10.1007/s40123-023-00842-6

ORIGINAL RESEARCH

Deep Transfer Learning for Ethnically Distinct Populations: Prediction of Refractive Error Using Optical Coherence Tomography

Rishabh Jain · Tae Keun Yoo · Ik Hee Ryu · Joanna Song · Nitin Kolte · Ashiyana Nariani

Received: September 27, 2023 / Accepted: October 20, 2023 / Published online: November 13, 2023

ABSTRACT

Introduction: The mismatch between training and testing data distribution causes significant degradation in the deep learning model performance in multi-ethnic scenarios. To reduce the performance differences between ethnic groups and image domains, we built a deep transfer learning model with adaptation training to predict uncorrected refractive errors using posterior segment optical coherence tomography (OCT) images of the macula and optic nerve.

Methods: Observational, cross-sectional, multicenter study design. We pre-trained a deep learning model on OCT images from the B&VIIT Eye Center (Seoul, South Korea) (N = 2602 eyes of 1301 patients). OCT images from Poona Eye Care (Pune, India) were chronologically sorted into adaptation training data (N = 60 eyes of 30 patients) for transfer learning and test data (N = 142 eyes of 71 patients) for validation. Deep learning models were trained to predict spherical equivalent (SE) and mean keratometry (K) values via transfer learning for domain adaptation.

Results: Both adaptation models for SE and K were significantly better than those without adaptation ($P < 0.001$). In myopia/hyperopia classification, the model trained on circular optic disc OCT images yielded the best performance (accuracy = 74.7%). It also performed best in estimating SE with the lowest mean absolute error (MAE) of 1.58 D. For classifying the degree of corneal curvature, the optic nerve vertical algorithm performed best (accuracy = 65.7%). The optic nerve horizontal model achieved the lowest MAE (1.85 D) when predicting the K value. Saliency maps frequently highlighted the retinal nerve fiber layers.

Conclusions: Adaptation training via transfer learning is an effective technique for estimating

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40123-023-00842-6.

R. Jain
Department of Biomedical Engineering, Duke University, Durham, NC, USA

T. K. Yoo (✉) · I. H. Ryu
Department of Refractive Surgery, B&VIIT Eye Center, B2 GT Tower, 1317-23 Seocho-Dong, Seocho-Gu, Seoul, Republic of Korea
e-mail: eyetaekeunyoo@gmail.com; fawoo2@yonsei.ac.kr

T. K. Yoo · I. H. Ryu · J. Song
Research and Development Department, VISUWORKS, Seoul, South Korea

N. Kolte
Poona Eye Care, Pune, Maharashtra, India

A. Nariani
Department of Ophthalmology, King Edward Memorial Hospital and Seth Gordhandas Sunderdas Medical College, Mumbai, Maharashtra, India

refractive errors and *K* values using macular and optic nerve OCT images from ethnically heterogeneous populations. Further studies with larger sample sizes and various data sources are needed to confirm the feasibility of the proposed algorithm.

Keywords: Adaptation training; Transfer learning; Ethnically distinct populations; OCT; Refractive errors

Key Summary Points

Why carry out this study?

Despite the known racial and ethnic differences in ocular measurements and pathological conditions, few studies have examined the application of deep learning in different races.

We aimed to employ a transfer learning technique for the adaptation process, initially train the algorithm on an East Asian dataset, and validate it in an Indian population to predict refractive error and corneal curvature using posterior segment optical coherence tomography (OCT) images.

What was learned from the study?

We successfully developed deep learning algorithms to predict refractive error and corneal curvature from macular and optic nerve OCT images.

Transfer learning successfully tuned the domain shift caused by differences in ethnicity and imaging devices with small data. This algorithm will be clinically useful for screening when further studies with larger sample sizes and various data sources are conducted.

INTRODUCTION

Deep learning is a type of artificial intelligence (AI) in which a model learns predictive features and patterns from raw data without requiring the features to be specified explicitly [1]. It has been applied in ophthalmology for identifying vision-related diseases [2, 3]. However, despite the known racial and ethnic differences in ocular measurements and pathological conditions [4, 5], few studies have examined the application of deep learning in different races. Especially the mismatch between training and testing data distribution causes significant degradation in the model performance in multi-ethnic scenarios. Although ethnic differences in retinal images have been reported [6], the application of deep learning models to different ethnicities has not yet been studied. It is necessary to study whether deep learning models developed for a specific race can be applied to other races or domains.

In transfer learning, information gleaned from one task is applied to a related problem in a different domain [7], which is typically used when training data are scarce or heterogeneous. The transfer learning approach could improve disease detection accuracy for data-disadvantaged ethnic groups and for intra-study domain adaptation [8]. In a previous study, a model trained on large-scale genome data was first adapted for a small ethnic group through fine-tuning and additional domain adaptation training [9]. Thus, a new task with a small adaptation dataset can be trained using pre-trained weights from a large dataset through transfer learning [10]. It is expected that the spread of AI models for various groups will become possible through intra-study adaptation training using ocular images.

Refractive errors (myopia, hyperopia, and astigmatism) occur when light is not accurately focused on the macula. Uncorrected refractive error (URE) remains the leading cause of visual impairment and the second leading cause of blindness globally, with 596 million people suffering from distance visual impairment and 43 million considered legally blind in 2020 [11, 12]. A link between URE and other ocular

morbidities has been observed, such as an increased risk of open-angle glaucoma and retinal detachment in myopia [13]. In addition, progressive elongation of the eyeball in high myopia also causes degenerative fundoscopic changes, including posterior staphyloma, lacquer cracks of the Bruch's membrane, chorioretinal atrophy, and vision-threatening choroidal neovascularization [14]. This difficulty is exacerbated by the fact that clinicians focus on pathological ocular lesions in patients with low vision. Therefore, the impact of refraction may be underappreciated in diagnosing and distinguishing multiple possible causes of visual impairment [15].

In this study, we evaluated retinal imaging-domain deep learning models in ethnically distinct populations. Recently, both fundus photography and posterior segment optical coherence tomography (OCT) have been used to train deep learning models to predict refractive errors [16]. We sought to tackle the domain difference in developing deep learning models to predict URE, classify the degree of myopia and hyperopia, and estimate the corneal curvature from macular and optic nerve OCT images. We aimed to employ a transfer learning technique for the adaptation process, initially train the algorithm on an East Asian dataset, and validate it in an Indian population.

METHODS

Study Design

The main objective of this study was to evaluate the performance of multi-ethnic machine learning schemes via transfer learning in ethnically distinct populations. Data were obtained from two sources: South Korea (for initial model development) and India (for model adaptation and validation). This observational, cross-sectional study was approved by the Ethics Committee of Sahyadri Hospital in Pune, Maharashtra, India, and the Korean National Institute for Bioethics Policy (KoNIBP, No. P01-202302-01-009) in Seoul, South Korea. All procedures adhered to the principles of the Declaration of Helsinki. Informed consent was obtained from all Indian patients in their preferred language (English, Hindi, or Marathi), and the KoNIBP waived the requirement for informed consent for retrospective data collection.

Dataset

We pre-trained the deep learning models to learn the features of OCT using a large Korean dataset (N = 2602 eyes of 1301 patients). We enrolled an independent Indian dataset for adaptation and external validation. The Indian dataset was chronologically sorted into adaptation training data (N = 60 eyes of 30 patients) for transfer learning and test (N = 142 eyes of 71 patients) data for validation.

To build the initial deep learning models using a large Korean dataset, we analyzed the preoperative ocular data of healthy patients aged between 21 and 50 years who underwent refractive surgery at the B&VIIT Eye Center between January 2018 and December 2020. We used the preoperative data to guarantee that the subjects have no past surgical history. Patients with poor image quality were excluded. Retinal cross sections passing through the center of the retina (fovea) along the horizontal and longitudinal axes were collected during a three-dimensional OCT examination (Topcon 3D OCT-1 Maestro, Tokyo, Japan). All subjects underwent visual acuity measurements, autorefractometer readings (Nidek ARK-1, Nidek Co Inc., Gamagori, Japan), and manifest refraction.

For the Indian dataset, preoperative patients aged > 18 years who sought LASIK, cataract surgery, or routine ophthalmic examination at Poona Eye Care (Pune, Maharashtra, India) between October 2022 and December 2022 were recruited. Each subject underwent a visual acuity test, autorefractometer readings (Topcon KR-8900 Auto Keratorefractometer, Tokyo, Japan), and manifest refraction. Horizontal and vertical OCT cross sections through the center of the macula and the center of the optic nerve, as well as OCT cross sections around the circumference of the optic nerve, were obtained using Cirrus HD-OCT 500 (Software Version 10.0.1.19039, Carl Zeiss Meditec, Jena, Germany) using the

built-in automatic focusing feature to select the highest-quality images of the macula.

We used the central pseudo-color B-scans images from each volume for the deep learning analysis. In this study, those with a self-reported history of any previous ocular surgery or those with retinal or optic nerve pathology, as identified on OCT, were excluded from both the Korean and Indian datasets. Supplementary Fig. 1 shows examples of OCT images from the B&VIIT and Poona Eye Care.

Algorithm Development

Spherical equivalent (SE) was used as the target metric for determining refractive error, which incorporates both refractive power (sphere) and astigmatism (cylinder). High myopia was defined as SE of − 6.00 D or below, mild-to-moderate myopia as SE between − 6.00 and 0.00 D, and hyperopia as SE of 0.00 D or greater. The *K* value was used as the target metric for determining corneal curvature, which averages the flat meridian (K1) and steep meridian (K2) measured by keratometry. A flat cornea was defined as a *K* value of 42.0 D or below, a regular cornea as *K* value between 42.0 and 45.0 D, and a steep cornea as *K* value of 45.0 or greater [17]. This categorization has been adopted in previous studies to predict different postoperative prognoses according to corneal curves [17, 18]. When analyzing the keratometry values, eight eyes were excluded from the original test set (N = 142 eyes) due to missing values.

As shown in Fig. 1, we employed stepwise transfer learning to predict URE (measured using SE) and corneal refractive power (*K* value) to overcome the ethnic and image domain differences between multicenter data sources. As the Indian dataset was too small to fully train the CNN, we adopted transfer learning for domain adaptation based on a large Korean dataset. In the prior-learning stage, we pretrained the models without freezing weights to predict the URE and *K* values from the Korean OCT dataset via transfer learning using a pretrained CNN from the ImageNet dataset (Supplementary Fig. 2). In the target-task learning stage, the fully connected layer of the trained model was fine-tuned using only small adaptation datasets from Indian subjects. This fine-tuning used the pre-trained weights as a starting feature-space point for further training on the small target dataset (Indian dataset) with the same architecture as the source model. Deep learning models were trained individually for each OCT imaging domain (macula horizontal and vertical OCT). We used the macular-horizontal OCT model pre-trained on the Korean dataset as a prior model for training the optic nerve OCT (horizontal, vertical, and circular) image domains.

A pre-trained ResNet50 architecture was used as the feature extractor (Supplementary Fig. 3). This architecture has shown robust predictive performance in OCT-based tasks [19, 20]. All ResNet50 weights of the feature extractor were trained using the Korean dataset. Through this pretraining, the feature extractor learns the basic features and patterns present in the OCT images, such as retinal layers, thickness, and shape. Once pretraining is completed, the CNN can be fine-tuned on a small task dataset. The weights of the feature extractor nets were fixed for the target task learning based on the Indian dataset. In our study, the last layer of the CNN architecture was replaced with a modified fully connected layer (2×2048 weights and 2×1 bias for ResNet50), a Soft-max function for classification tasks, and a linear regressor for regression tasks. The input images were resized according to the input tensor of ResNet50. To avoid overfitting, typical linear data augmentation methods were used in all training processes [21]. Augmentation using changes in magnification and axis was set widely to cover the magnification and scan area between the two domains random position translations (− 20% to + 20%), random rotations (− 15° to 15°), random magnifications (− 25% to 25%). The cross-entropy loss was used as the loss function for training the classification models. We used mean absolute error (MAE) as the loss function for all for regression tasks. All models were optimized using stochastic gradient descent (SGD) with momentum algorithm (learning rate = 0.0001) with a mini-batch size of 32. The experiments for optimizing the architecture were performed using MATLAB R2022a

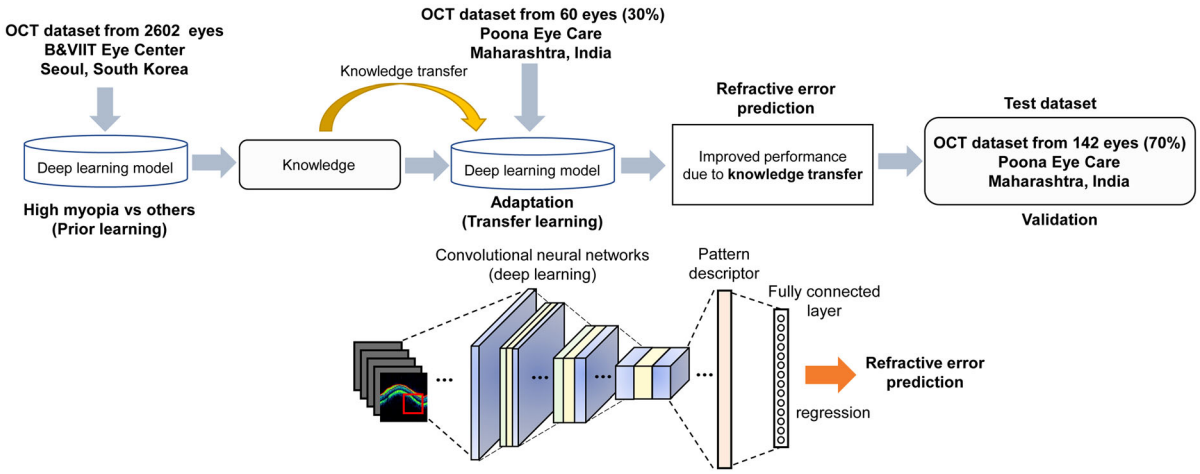


Fig. 1 Deep learning model development and workflow. Schematic of the training and knowledge transfer process for predicting refractive errors and keratometry. *OCT* optical coherence tomography

(MathWorks Inc., Natick, MA, USA). For interpretability, saliency maps were generated using the gradient-weighted class activation mapping (Grad-CAM) technique to visualize the areas on which the deep-learning model was focused. For hands-on experience with this study, the codes for practice and light-version models are provided at a publicly accessible source (https://data.mendeley.com/datasets/89z7h5gnpw). In this material, the deep learning models were developed based on Python and implemented in Keras with TensorFlow.

Statistical Analysis

To evaluate the predictive performance of the deep learning models, the MAE, median absolute prediction error (MedAE), and Pearson correlation coefficient were computed in the test dataset validation. The MAE values between the models were compared using a paired *t* test. To detect specific conditions using each binary deep learning model, the area under the receiver operating characteristic curve (ROC-AUC) was calculated. Youden's index, a widely used estimate of optimal thresholds that assigns equal weight to sensitivity and specificity, was adopted in this study [22]. Cohen's kappa and Matthew's correlation coefficients, which are common metrics in multiclass classification [23], were used to evaluate the multiclass prediction tasks [24]. These two metrics range from 0 to 1, and the closer they are to 1, the higher the classification performance.

RESULTS

The characteristics of the patients in the B&VIIT (N = 2602 eyes of 1301 subjects) and Poona Eye Care (adaptation set with N = 60 eyes of 30 subjects; validation set with N = 142 eyes of 71 subjects) are summarized in Table 1. The datasets showed different characteristic distributions. The mean keratometry was the only parameter for which no significant differences were observed among the development, adaptation, and validation datasets.

The experimental results of modulating the size of the adaptation dataset are shown in Fig. 2. As expected, the SE and mean keratometry prediction accuracies improved as the size of the adaptation dataset increased. However, this effect tapered off when the size of the adaptation dataset was > 40. When compared to the no-adaptation training (size of the adaptation dataset = 0), the performance was greatly improved after all adaptation processes ($P < 0.001$ for all paired comparisons). As can be seen here, models trained with the Korean dataset without adaptation training showed critically low accuracy on the unseen new Indian dataset in both the SE and mean

Table 1 Demographics and ocular data of subjects in training data (Korea) and validation data (India)

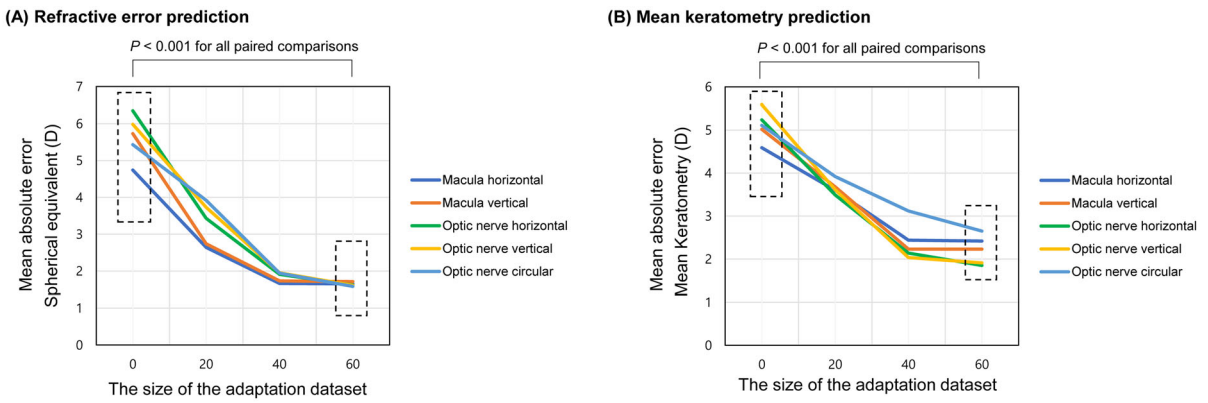
	Original development data (B&VIIT Eye Center, South Korea; N = 2602 eyes from 1301 patients)	Adaptation data (Poona Eye Care, India; N = 60 eyes from 30 patients)	Validation data (Poona Eye Care, India; N = 142 eyes from 71 patients)	P value for multiple comparison
Age (years)	26.23 ± 5.80	32.58 ± 11.21	30.41 ± 9.17	< 0.001
Sex, female:male	784:517	15:15	30:41	< 0.001
Axial length (mm)	25.47 ± 1.68	24.64 ± 1.03	23.97 ± 1.17	< 0.001
Manifest refraction				
Sphere (D)	− 3.77 ± 2.16	− 2.40 ± 2.67	− 2.40 ± 2.41	< 0.001
Cylinder (D)	− 1.91 ± 1.50	− 0.76 ± 1.08	− 0.83 ± 0.99	< 0.001
Spherical equivalent (D)	− 4.73 ± 2.38	− 2.78 ± 2.69	− 2.82 ± 2.53	< 0.001
Corneal refractive power				
Flattest keratometry (D)	42.53 ± 3.13	42.94 ± 1.30	43.47 ± 1.48	< 0.001
Steepest keratometry (D)	44.89 ± 3.82	44.01 ± 1.29	44.54 ± 1.58	< 0.001
Mean keratometry (D)	43.71 ± 3.42	43.48 ± 1.23	44.01 ± 1.49	0.698

keratometry prediction. Based on this analysis, a size of 60 is used for the results presented below.

The performance of the deep learning model in estimating URE and corneal curvature is presented in Table 2. The model trained on circular cross sections of the optic nerve performed significantly better than the other models in estimating spherical equivalent, with a MAE of 1.58 D and a MedAE of 1.38 D. For keratometry prediction, the models trained on horizontal and vertical cross sections of the optic nerve showed the best performance, with MAEs of 1.85 and 1.91 D and MedAEs of 1.59 and 1.67 D, respectively. The P values for the MAE comparison of the other three models were < 0.001, indicating a significantly worse performance.

Figure 3 shows the accuracy of the model in classifying eyes as having hyperopia, mild myopia, or high myopia. The model trained on circular cross sections of the optic nerve showed the best prediction accuracy of 74.7%, with a Cohen's kappa of 0.489 and Matthew's correlation of 0.494.

Figure 4 shows the accuracy of the model in classifying the eyes as having flat, regular, or steep corneas. The model trained on vertical cross sections of the optic nerve showed the best prediction accuracy of 65.7%, with a Cohen's



Spherical equivalent (D)	The size of the adaptation dataset			
	0	20	40	60
Macula horizontal	4.74	2.64	1.66	1.65
Macula vertical	5.73	2.74	1.73	1.71
Optic nerve horizontal	6.35	3.43	1.91	1.62
Optic nerve vertical	5.98	3.73	1.95	1.61
Optic nerve circular	5.43	3.92	1.94	1.58

Mean Keratometry (D)	The size of the adaptation dataset			
	0	20	40	60
Macula horizontal	4.59	3.6	2.44	2.42
Macula vertical	5.02	3.68	2.23	2.23
Optic nerve horizontal	6.24	3.49	2.137	1.85
Optic nerve vertical	5.59	3.61	2.035	1.91
Optic nerve circular	5.11	3.92	3.12	2.65

Fig. 2 Experimental results for deep learning model adaptation. Influence of the adaptation dataset size on **A** refractive error and **B** mean keratometry prediction accuracy

kappa of 0.364 and Matthew's correlation of 0.353. The overall performance in classifying the degree of corneal curvature was worse than that in classifying the degree of myopia/hyperopia, with the best accuracy in the former (65.7%) being still lower than the worst accuracy in the latter (68.8%).

The ROC curves are shown in Fig. 5. The optic nerve circular model had an AUC of 0.956 and 0.869 for detecting high myopia and hyperopia, respectively, demonstrating the best performance. The optic nerve vertical model was the best at detecting flat and steep corneas with ROC-AUCs of 0.848 and 0.879, respectively. We also performed the deep learning experiments on two additional learning tasks without adaptation training. Figure 6 shows that the transfer learning models with adaptation training outperformed the models without adaptation (including models trained with only the Korean dataset and those with only the Indian adaptation dataset) in all prediction tasks ($P < 0.001$ for all comparisons).

Figures 7 and 8 show examples of the Grad-CAM saliency mapping technique for the prediction of SE and *K* value and for the classification of hyperopia/myopia and flat/steep corneal curvature. The entire retinal layer of the local area around the fovea was frequently highlighted to determine URE in the macula horizontal, macula vertical, and optic nerve circular cross-sectional images. The Bruch's membrane opening and the internal limiting membrane at the optic nerve cupping were frequently highlighted for determining SE in the horizontal and vertical optic nerve cross-sectional images. A similar pattern was observed when predicting *K* values. The entire retinal layer around the fovea was frequently highlighted to determine *K* values.

DISCUSSION

In this study, we presented a deep transfer learning framework by adaptation training to build better models for the data-disadvantaged group. Transfer learning successfully tuned the domain shift caused by differences in ethnicity and imaging devices with small data. During adaptation learning, we tried to overcome the difference between the two image domains through traditional data augmentation. We developed deep learning algorithms to predict refractive error and corneal curvature from five types of OCT images: horizontal and vertical cross sections of the macula and horizontal,

Table 2 Validation performance of the developed model after adaptation for estimating spherical equivalent and keratometry in the Indian validation set

	Domain	MAE ± SD (D)	MedAE (D)	Pearson correlation	*P* value for correlation	*P* value for MAE comparison
SE prediction (ResNet50)	Macula horizontal	1.65 ± 1.14	1.41	0.715	< 0.001	< 0.001
	Macula vertical	1.71 ± 1.07	1.42	0.703	< 0.001	< 0.001
	Optic nerve horizontal	1.62 ± 1.16	1.40	0.720	< 0.001	< 0.001
	Optic nerve vertical	1.61 ± 1.10	1.40	0.734	< 0.001	0.003
	Optic nerve circular	1.58 ± 1.03	1.38	0.741	< 0.001	Reference
Keratometry prediction (ResNet50)	Macula horizontal	2.42 ± 1.71	2.18	0.627	< 0.001	< 0.001
	Macula vertical	2.23 ± 1.74	2.05	0.655	< 0.001	< 0.001
	Optic nerve horizontal	1.85 ± 1.70	1.59	0.689	< 0.001	Reference
	Optic nerve vertical	1.91 ± 1.72	1.67	0.680	< 0.001	0.141
	Optic nerve circular	2.65 ± 1.79	2.27	0.602	< 0.001	< 0.001

MAE mean absolute prediction error, *MedAE* median absolute prediction error, *SD* standard deviation, *SE* spherical equivalent

vertical, and circular cross sections of the optic nerve. We also used these images to train the algorithm to classify eyes as having high myopia, moderate myopia, or hyperopia and to classify corneal curvature as flat, normal, or steep. Attempts to predict refractive errors and corneal curvature from OCT images of the retina and optic nerve in various domains are new challenges. Although it is an experimental study, new relationships that have been clinically estimated [25, 26] were observed through deep learning. There was a limit to classification performance in prediction of corneal curvature, but we found a significant possibility.

This study demonstrated that AI could use posterior segment morphology, as shown on OCT, to estimate the ocular properties typically associated with anterior segment anatomy. It also underscores the value of the transfer learning technique in pre-training a model on one task and adapting it to a more specific task using heterogeneous training data. Although URE and keratometry may not immediately indicate ocular pathology and we can easily measure them, this study provides a new insight that OCT images have another piece of ocular information (Supplementary Table 1).

While macular OCT images have previously been used for predicting refractive error, this

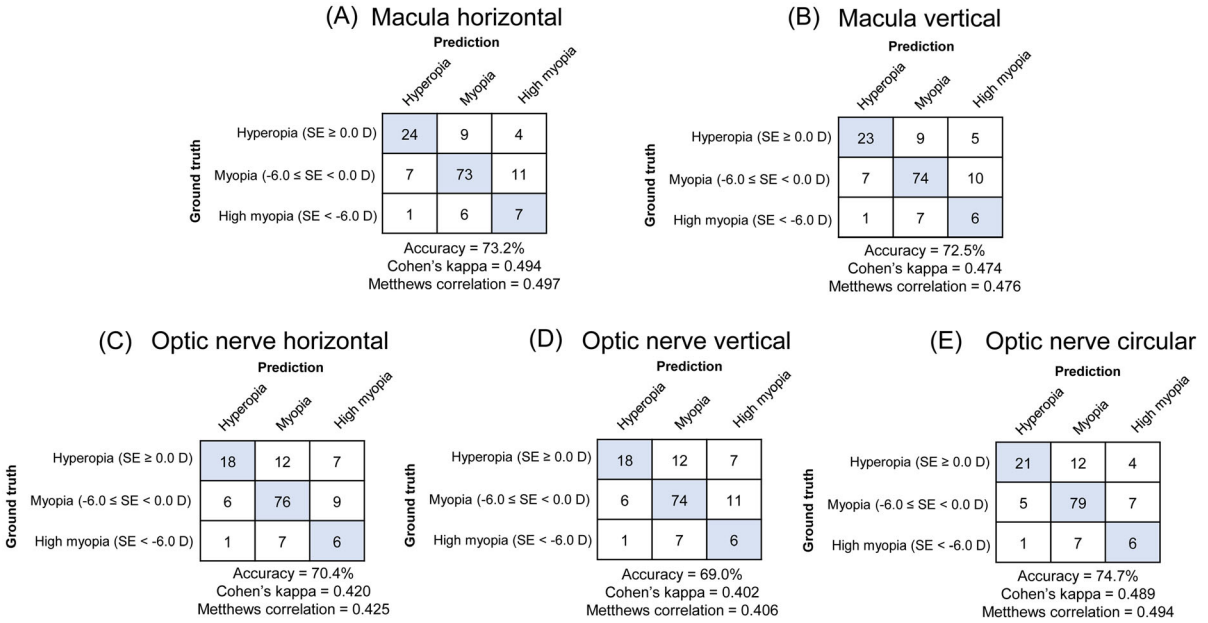


Fig. 3 Classification performance for identifying hyperopia/mild myopia/high myopia (N = 142 eyes). Ground truth vs. prediction of the degree of myopia/hyperopia for models trained on **A** horizontal and **B** vertical cross-sections of the macula and **C** horizontal, **D** vertical, and **E** circular cross-sections of the optic nerve. *SE* spherical equivalent

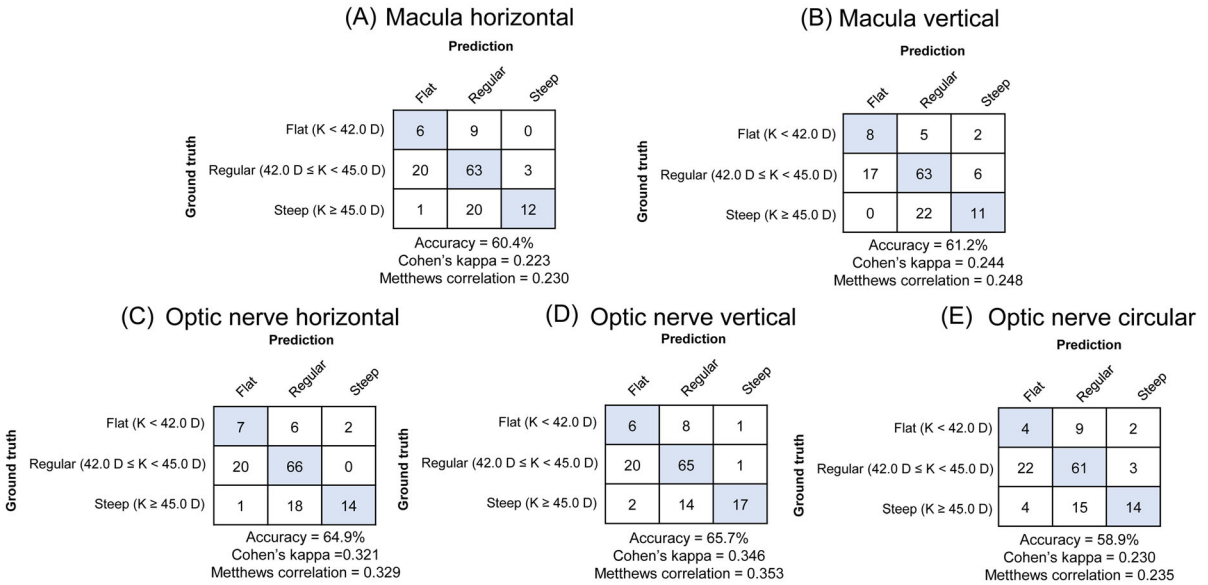


Fig. 4 Classification performance for identifying flat/regular/steep cornea (N = 134 eyes). Ground truth vs. prediction of the degree of corneal curvature for models trained on **A** horizontal and **B** vertical cross sections of the macula and **C** horizontal, **D** vertical, and **E** circular cross sections of the optic nerve. When analyzing the keratometry values, eight eyes were excluded from the original test set (N = 142 eyes) due to missing value. *K* mean keratometry

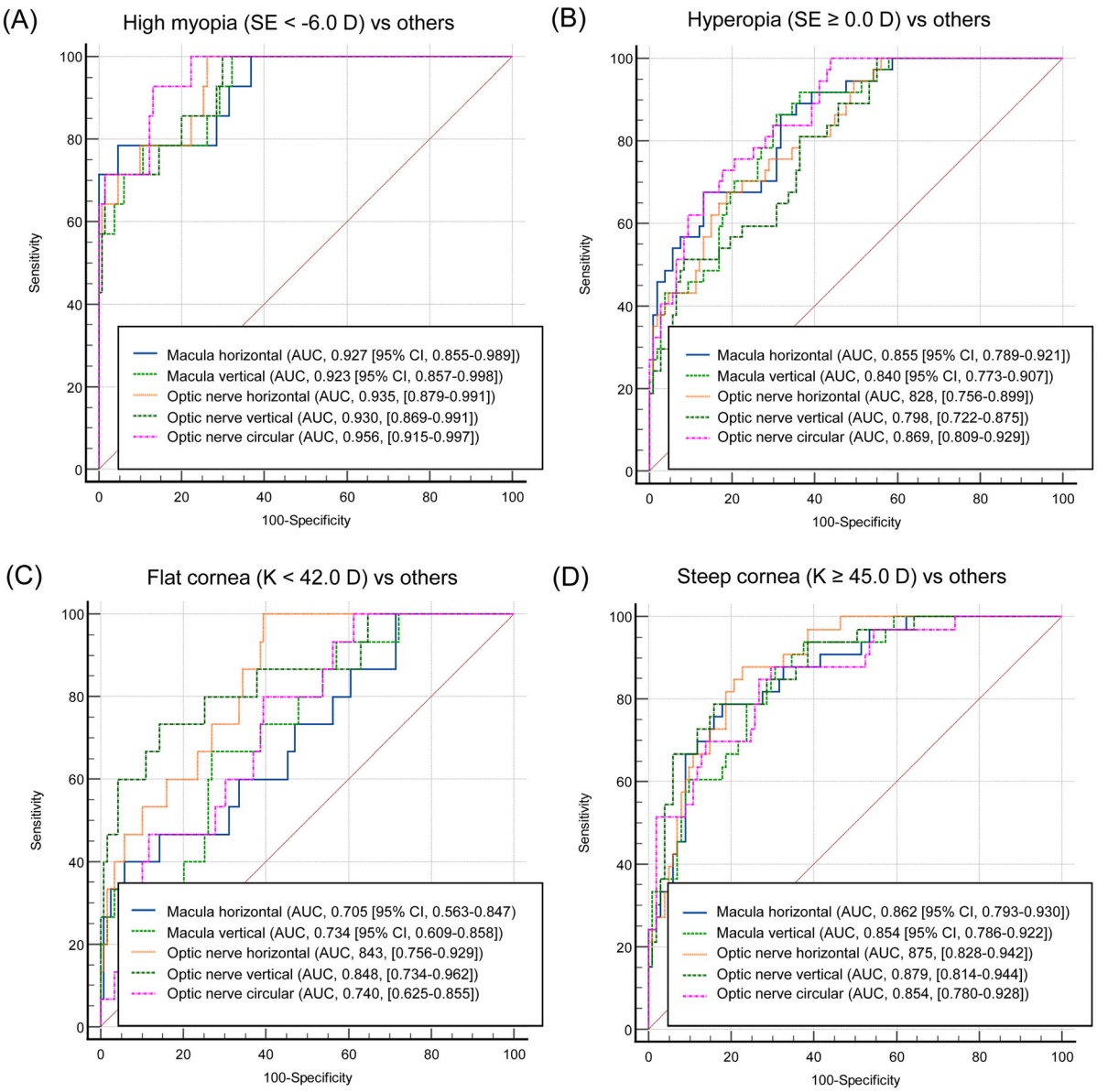


Fig. 5 ROC curves for uncorrected refractive errors and keratometry value prediction. ROC curves to distinguish high myopia, hyperopia, flat corneal curvature, and steep corneal curvature. *SE* spherical equivalent, *K* mean keratometry, *ROC* receiver operating characteristic curve, *AUC* area under the curve, *CI* confidence interval

work represents a significant improvement in performance (the lowest MAE improved from 2.66 D to 1.58 D) [27]. In contrast to the previous study, we relied on ethnically diverse training data and employed a stepwise transfer learning technique. Our work demonstrates how deep learning models trained on data from one center or a single ethnic group can be repurposed and applied elsewhere. In this study, deep learning models trained on a large Korean dataset were tuned relatively accurately using a small Indian adaptation set. To date, AI-based studies have suffered from performance decline during external validation, limiting their robustness in heterogeneous data [28]. However, fine-tuning the training process through transfer learning for a small adaptive dataset allows researchers to overcome this performance

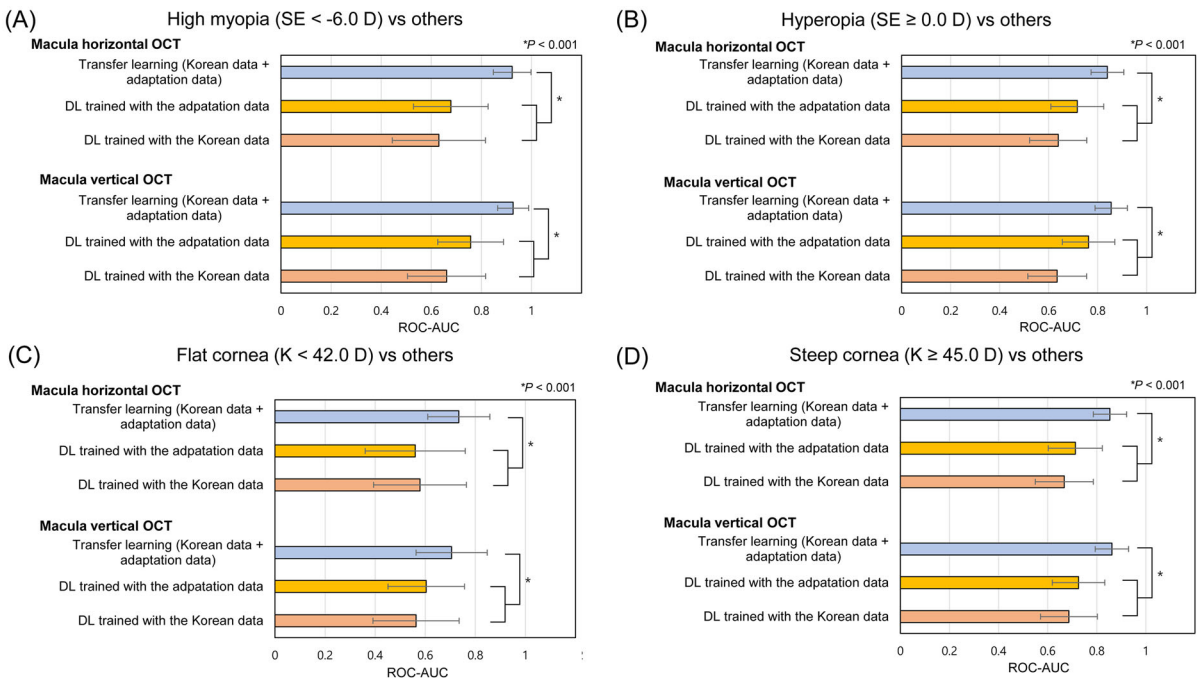


Fig. 6 Comparison of deep learning (DL) schemes in the test dataset. The tasks are to predict **A** high myopia, **B** hyperopia, **C** flat cornea, and **D** steep cornea. *SE* spherical equivalent, *K* mean keratometry, *OCT* optical coherence tomography

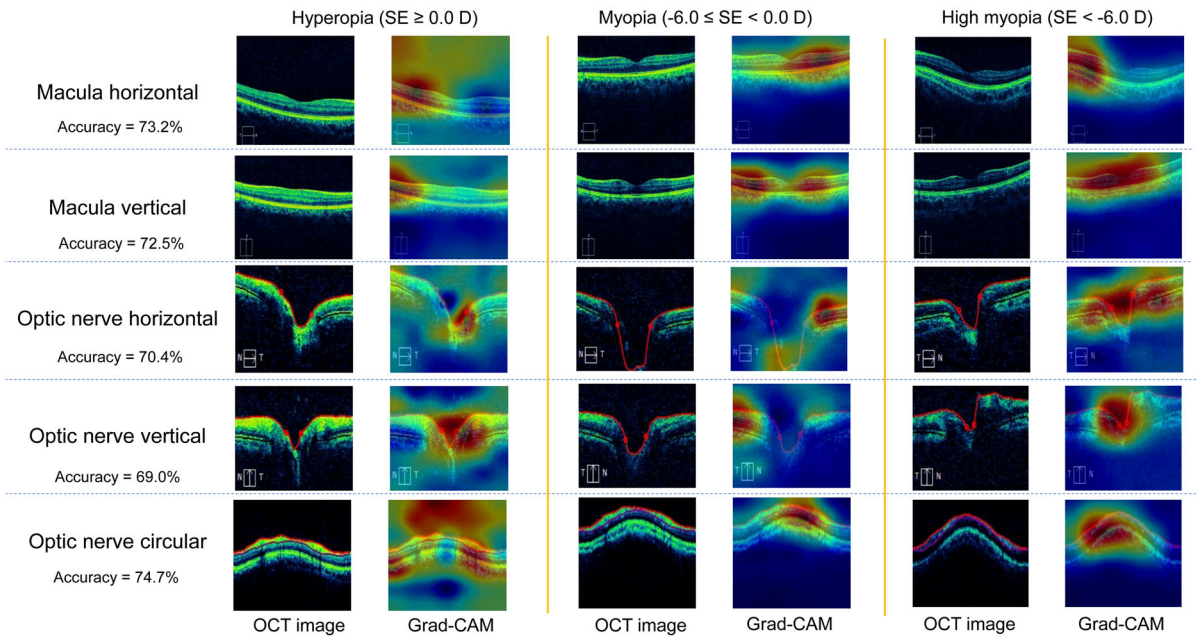


Fig. 7 Grad-CAM saliency maps for spherical equivalent (SE) prediction and hyperopia/mild myopia/high myopia classification. *OCT* optical coherence tomography

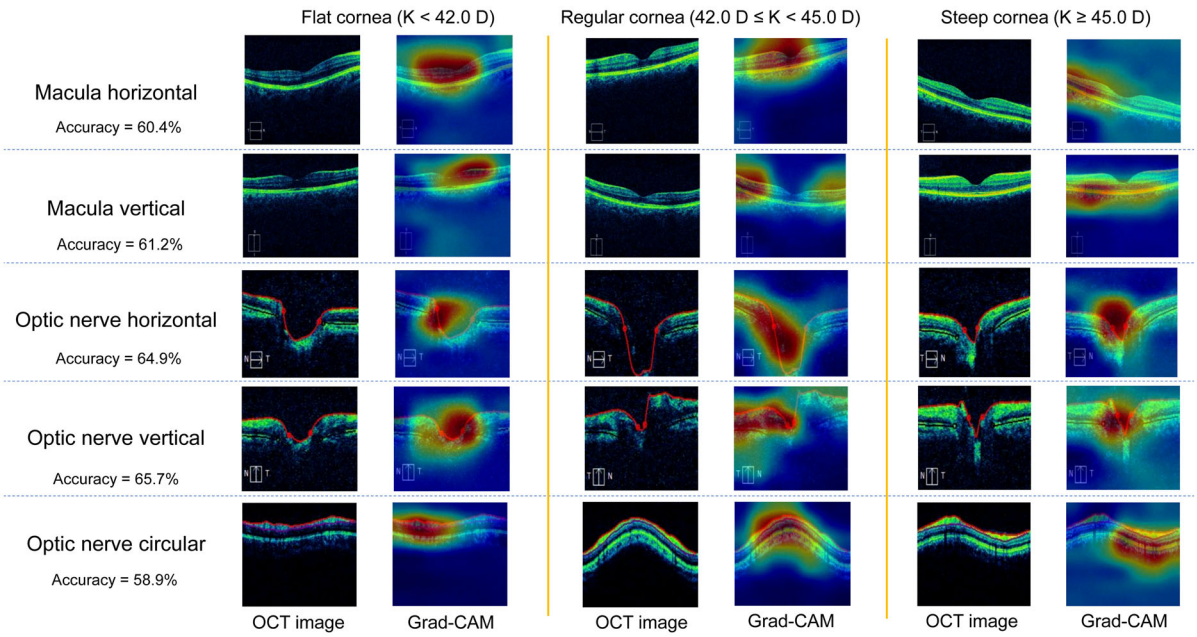


Fig. 8 Grad-CAM saliency maps for *K* value prediction and flat/regular/steep cornea classification. *K* mean keratometry, *OCT* optical coherence tomography

degradation in external validation sets [9]. Recently, the fine-tuning of specific tasks based on large deep learning models, known as foundation models, has become widespread and has been used as an adaptive training method for datasets with different data distributions [29]. Our study is consistent with this recent trend.

Through deep learning, it was confirmed that there was a significant relationship between the shape of the optic nerve and the refractive power. We are also unaware of previous reports using optic nerve OCT images to predict either refractive error or corneal curvature. However, it is worth noting that the algorithm's performance when estimating *K* value was lower than when estimating SE (highest accuracy of 65.7 vs. 74.3%, respectively). The refractive power of the cornea (*K* value) accounts for a large portion of the overall optical refractive power of the eye but is heavily influenced by external factors such as eyelid conditions [30]. However, because the overall refractive power (SE) and shape of the retina may interact with each other during growth and development, their relationship is expected to be stronger [31].

Using saliency maps, we also sought to understand which OCT image regions contributed the most to the model's predictive results to generate new insights into the relationship between posterior segment anatomy and refraction. The results of Grad-CAM saliency mapping highlighted the Bruch's membrane opening, suggesting a relationship between myopia and the morphological factors of the optic nerve head. A larger Bruch's membrane opening diameter is linearly associated with severe myopia [32]. All retinal and choroidal layers observed on OCT are associated with myopia [33]. Furthermore, highly myopic eyes may be related to inner retinal layer thinning and a decreased ratio of posterior choroidal thickness to Bruch's membrane thickness [34]. Deep-learning algorithms reflect these characteristics in SE prediction. Interestingly, this relationship was similar in predicting *K* value.

Our study had several limitations. The Korean OCT images were obtained on the Topcon

3D OCT-1 Maestro, whereas Indian OCT images were obtained using the Zeiss Cirrus HD-OCT 500; the images from the latter were of slightly higher quality on average than the former. The clinical profiles of the two populations also varied: Korean patients were overwhelmingly myopic, with no hyperopia in the initial training set, whereas 23% of Indian patients underwent routine ophthalmic examination, resulting in a much lower incidence of high myopia. Thus, the pre-training and validation data exhibited key differences, with the former being biased towards a more negative SE. Finally, the Indian dataset was limited in size. A larger number of study datasets may further improve the prediction accuracy of the algorithm. In addition, it is not sure that the developed model is clinically feasible to use OCT imaging for refractive error. We expect this study to be useful only when OCT is more widely available as a screening equipment in the future. It should be noted that the study data did not include the pathological eyes. Further research is needed to confirm whether the proposed algorithm properly works in eyes with glaucoma or keratoconus. Finally, analyzing pseudo-color images taken from OCT measurement results other than raw volume images is also a disadvantage of this study.

CONCLUSIONS

In conclusion, we have developed an algorithm to predict and classify the degree of refractive error and corneal curvature using macular and optic nerve OCT images. Using the deep transfer learning technique, we were able to train the model on data from individuals who were ethnically distinct from our validation dataset population and demonstrated robust performance. Thus, our proposed adaptation training had great potential to promote the deep learning screening in ethnically heterogeneous populations. Although we investigated the concept of deep transfer learning for adaptation, this algorithm will be clinically useful for screening when further studies with larger sample sizes and various data sources are conducted.

Author Contributions. Tae Keun Yoo and Rishabh Jain had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: Rishabh Jain, Ashiyana Nariani. Acquisition, analysis, or interpretation of data: Rishabh Jain, Tae Keun Yoo, Ik Hee Ryu, Joanna Song, Nitin Kolte, Ashiyana Nariani. Drafting of the manuscript: Rishabh Jain. Critical revision of the manuscript for important intellectual content: Ik Hee Ryu, Ashiyana Nariani. Statistical analysis: Rishabh Jain, Tae Keun Yoo. Administrative, technical, or material support: Tae Keun Yoo, Ik Hee Ryu, Joanna Song. Supervision: Tae Keun Yoo, Ik Hee Ryu, Ashiyana Nariani.

Funding. This project was supported by U.S. Fulbright-Nehru Scholar, and it did not sponsor the Rapid Service Fee. The Rapid Service Fee was funded by the authors.

Data Availability. The data used in this study cannot be made publicly accessible due to restrictions from the Ethics Committee of Sahyadri Hospital and the KoNIBP. Data can be accessed upon reasonable request pending approvals from the Ethics Committees. For hands-on experience with this study, the codes for practice and light-version models are provided at a publicly accessible source (https://data.mendeley.com/datasets/89z7h5gnpw). In this material, the deep learning models were developed based on Python and implemented in Keras with TensorFlow.

Declarations

Conflict of Interest. Rishabh Jain reports a grant from Fulbright Scholarship. Tae Keun Yoo, Ik Hee Ryu, and Joanna Song are employees of VISUWORKS. This company supplied the study tools. Nitin Kolte and Ashiyana Nariani have nothing to disclose.

Ethical Approval. This observational, cross-sectional study was approved by the Ethics Committee of Sahyadri Hospital in Pune, Maharashtra, India, and the Korean National Institute for Bioethics Policy (KoNIBP, No. P01-202302-01-009) in Seoul, South Korea. All

procedures adhered to the principles of the Declaration of Helsinki. Informed consent was obtained from all Indian patients in their preferred language (English, Hindi, or Marathi), and the KoNIBP waived the requirement for informed consent for retrospective data collection.

REFERENCES

1. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–44.

2. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA. 2016;316:2402–10.

3. De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med. 2018;24:1342–50.

4. Sekimitsu S, Zebardast N. Glaucoma and machine learning: a call for increased diversity in data. Ophthalmol Glaucoma. 2021;4:339–42.

5. Coyner AS, Singh P, Brown JM, Ostmo S, Chan RVP, Chiang MF, et al. Association of biomarker-based artificial intelligence with risk of racial bias in retinal images. JAMA Ophthalmol. 2023. https://doi.org/10.1001/jamaophthalmol.2023.1310.

6. Ho H, Tham Y-C, Chee ML, Shi Y, Tan NYQ, Wong K-H, et al. Retinal nerve fiber layer thickness in a multiethnic normal Asian population: the Singapore epidemiology of eye diseases study. Ophthalmology. 2019;126:702–11.

7. Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. J Big Data. 2016;3:9.

8. Wang R, Chaudhari P, Davatzikos C. Embracing the disharmony in medical imaging: a simple and effective framework for domain adaptation. Med Image Anal. 2022;76: 102309.

9. Gao Y, Cui Y. Deep transfer learning for reducing health care disparities arising from biomedical data inequality. Nat Commun. 2020;11:5131.

10. Yoo TK, Kim SH, Kim M, Lee CS, Byeon SH, Kim SS, et al. DeepPDT-Net: predicting the outcome of photodynamic therapy for chronic central serous chorioretinopathy using two-stage multimodal transfer learning. Sci Rep. 2022;12:18689.

11. Naidoo KS, Leasher J, Bourne RR, Flaxman SR, Jonas JB, Keeffe J, et al. Global vision impairment and blindness due to uncorrected refractive error, 1990–2010. Optom Vis Sci. 2016;93:227.

12. Burton MJ, Ramke J, Marques AP, Bourne RRA, Congdon N, Jones I, et al. The Lancet Global Health Commission on global eye health: vision beyond 2020. Lancet Glob Health. 2021;9:e489-551.

13. Shen L, Melles RB, Metlapally R, Barcellos L, Schaefer C, Risch N, et al. The association of refractive error with glaucoma in a multiethnic population. Ophthalmology. 2016;123:92–101.

14. Avila MP, Weiter JJ, Jalkh AE, Trempe CL, Pruett RC, Schepens CL. Natural history of choroidal neovascularization in degenerative myopia. Ophthalmology. 1984;91:1573–81.

15. Sunness JS, El Annan J. Improvement of visual acuity by refraction in a low-vision population. Ophthalmology. 2010;117:1442–6.

16. Varadarajan AV, Poplin R, Blumer K, Angermueller C, Ledsam J, Chopra R, et al. Deep learning for predicting refractive error from retinal fundus images. Invest Ophthalmol Vis Sci. 2018;59:2861–8.

17. Luft N, Siedlecki J, Reinking F, Mayer WJ, Schworm B, Kassumeh S, et al. Impact of extreme (flat and steep) keratometry on the safety and efficacy of

small incision lenticule extraction (SMILE). Sci Rep. 2021;11:17854.

18. Reitblat O, Levy A, Kleinmann G, Lerman TT, Assia EI. Intraocular lens power calculation for eyes with high and low average keratometry readings: comparison between various formulas. J Cataract Refract Surg. 2017;43:1149–56.

19. Wang J, Deng G, Li W, Chen Y, Gao F, Liu H, et al. Deep learning for quality assessment of retinal OCT images. Biomed Opt Express BOE. 2019;10:6057–72.

20. Choi KJ, Choi JE, Roh HC, Eun JS, Kim JM, Shin YK, et al. Deep learning models for screening of high myopia using optical coherence tomography. Sci Rep. 2021;11:21663.

21. Choi JY, Yoo TK, Seo JG, Kwak J, Um TT, Rim TH. Multi-categorical deep learning neural network to classify retinal images: a pilot study employing small database. PLoS ONE. 2017;12: e0187336.

22. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biom J. 2005;47:458–72.

23. Yoo TK, Choi JY, Kim HK, Ryu IH, Kim JK. Adopting low-shot deep learning for the detection of conjunctival melanoma using ocular surface images. Comput Methods Progr Biomed. 2021;205: 106086.

24. Chicco D, Warrens MJ, Jurman G. The Matthews correlation coefficient (MCC) is more informative than Cohen's Kappa and brier score in binary classification assessment. IEEE Access. 2021;9: 78368–81.

25. Chen X-Y, He H-L, Xu J, Liu Y-X, Jin Z-B. Clinical features of fundus tessellation and its relationship with myopia: a systematic review and meta-analysis. Ophthalmol Ther. 2023. https://doi.org/10.1007/s40123-023-00802-0.

26. Huang D, Qian Y, Yan Q, Ling S, Dong Z, Ke X, et al. Prevalence of fundus tessellation and its screening based on artificial intelligence in Chinese children: the Nanjing Eye Study. Ophthalmol Ther. 2023;12: 2671–85.

27. Yoo TK, Ryu IH, Kim JK, Lee IS. Deep learning for predicting uncorrected refractive error using posterior segment optical coherence tomography images. Eye. 2022;36:1959–65.

28. Rim TH, Lee AY, Ting DS, Teo K, Betzler BK, Teo ZL, et al. Detection of features associated with neovascular age-related macular degeneration in ethnically distinct data sets by an optical coherence tomography: trained deep learning algorithm. Br J Ophthalmol. 2021;105:1133–9.

29. Willemink MJ, Roth HR, Sandfort V. Toward foundational deep learning models for medical imaging in the new era of transformer networks. Radiol Artif Intell. 2022;4: e210284.

30. Dogan E, Akbas Kocaoglu F, Yalniz-Akkaya Z, Elbeyli A, Burcu A, Ornek F. Scheimpflug imaging in dermatochalasis patients before and after upper eyelid blepharoplasty. Semin Ophthalmol. 2015;30: 193–6.

31. Flitcroft DI. The complex interactions of retinal, optical and environmental factors in myopia aetiology. Prog Retin Eye Res. 2012;31:622–60.

32. Sung MS, Heo H, Piao H, Guo Y, Park SW. Parapapillary atrophy and changes in the optic nerve head and posterior pole in high myopia. Sci Rep. 2020;10:4607.

33. Sung MS, Heo MY, Heo H, Park SW. Bruch's membrane opening enlargement and its implication on the myopic optic nerve head. Sci Rep. 2019;9: 19564.

34. Jonas JB, Holbach L, Panda-Jonas S. Bruch's membrane thickness in high myopia. Acta Ophthalmol. 2014;92:e470-474.