



OPEN ACCESS

Deep learning algorithms for automatic detection of pterygium using anterior segment photographs from slit-lamp and hand-held cameras

Xiaoling Fang ^{1,2}, Mihir Deshmukh,¹ Miao Li Chee,¹ Zhi-Da Soh ¹,
Zhen Ling Teo ¹, Sahil Thakur ¹, Jocelyn Hui Lin Goh,¹ Yu-Chi Liu,^{1,3}
Rahat Husain,^{1,3} Jodhbir Mehta,^{1,3} Tien Yin Wong ^{1,3,4}, Ching-Yu Cheng ^{1,3,4},
Tyler Hyungtaek Rim ^{1,3}, Yih-Chung Tham ^{1,3}

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bjophthalmol-2021-318866>).

For numbered affiliations see end of article.

Correspondence to

Dr Yih-Chung Tham, Ocular Epidemiology Research Group, Singapore Eye Research Institute, The Academia, 20 College Road Discovery Tower, Level 6, Singapore 169856, Singapore; tham.yih.chung@seri.com.sg

XF and MD contributed equally. TR and Y-CT contributed equally.

XF and MD are joint first authors. TR and Y-CT are joint senior authors.

Received 17 January 2021
Accepted 25 June 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Fang X, Deshmukh M, Chee ML, et al. *Br J Ophthalmol* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bjophthalmol-2021-318866

ABSTRACT

Background/aims To evaluate the performances of deep learning (DL) algorithms for detection of presence and extent pterygium, based on colour anterior segment photographs (ASPs) taken from slit-lamp and hand-held cameras.

Methods Referable pterygium was defined as having extension towards the cornea from the limbus of >2.50 mm or base width at the limbus of >5.00 mm. 2503 images from the Singapore Epidemiology of Eye Diseases (SEED) study were used as the development set. Algorithms were validated on an internal set from the SEED cohort (629 images (55.3% pterygium, 8.4% referable pterygium)), and tested on two external clinic-based sets (set 1 with 2610 images (2.8% pterygium, 0.7% referable pterygium, from slit-lamp ASP); and set 2 with 3701 images, 2.5% pterygium, 0.9% referable pterygium, from hand-held ASP).

Results The algorithm's area under the receiver operating characteristic curve (AUROC) for detection of any pterygium was 99.5% (sensitivity=98.6%; specificity=99.0%) in internal test set, 99.1% (sensitivity=95.9%, specificity=98.5%) in external test set 1 and 99.7% (sensitivity=100.0%; specificity=88.3%) in external test set 2. For referable pterygium, the algorithm's AUROC was 98.5% (sensitivity=94.0%; specificity=95.3%) in internal test set, 99.7% (sensitivity=87.2%; specificity=99.4%) in external set 1 and 99.0% (sensitivity=94.3%; specificity=98.0%) in external set 2.

Conclusion DL algorithms based on ASPs can detect presence of and referable-level pterygium with optimal sensitivity and specificity. These algorithms, particularly if used with a handheld camera, may potentially be used as a simple screening tool for detection of referable pterygium. Further validation in community setting is warranted.

Synopsis/precis DL algorithms based on ASPs can detect presence of and referable-level pterygium optimally, and may be used as a simple screening tool for the detection of referable pterygium in community screenings.

INTRODUCTION

Pterygium is the most common degenerative, age-related ocular surface disease, characterised by an overgrowth of the bulbar conjunctiva that can

encroach to the central cornea at advanced stages, causing corneal scarring, irregular astigmatism and visual impairment.¹⁻³ Previous population-based studies showed that the prevalence of pterygium was higher in rural populations compared with urban, with the prevalence ranging up to 39.5%.⁴⁻¹⁰

In under-resourced communities/countries, where ophthalmologists are limited and not easily accessible, most advanced pterygium cases are typically detected late. This often results in delayed surgical intervention.^{5 6 11} Importantly, surgical removal of advanced pterygium also carries higher risk of postsurgery complications such as higher rates of recurrence, corneal scarring, postsurgery induced astigmatism, and thus poorer prognosis.¹²⁻¹⁴ Taken together, it is important to detect and refer moderate or advanced pterygium (ie, those that warrant surgical intervention) timely, especially in rural communities. Hence, to cater for under-resourced communities, a new screening method for referable pterygium cases (ie, moderate or advanced pterygium that warrant surgical intervention), and one that is non-reliant on physical consultation of ophthalmologists, is needed.

The advent of artificial intelligence and deep learning (DL) potentially provides new solutions to address clinical gaps. Past DL algorithms for detection of pterygium have relied on small datasets, without validation on external populations, which is critical for testing robustness of algorithms.^{15 16}

In an attempt to address the abovementioned gap of late detection of advanced pterygium in rural areas, we designed and evaluated the performance of newly developed DL algorithms for detection of the presence of any pterygium and referable type pterygium, using colour anterior segment photographs (ASPs).

METHODS

Study population/database description

We used clinical data and colour ASPs from the Singapore Epidemiology of Eye Diseases (SEED) cohort study.¹⁷ Sixteen thousand six hundred thirty-six eligible participants with pterygium grading data were initially included. From which, 1366 pterygium eyes and 1566 non-terygium eyes (randomly selected from an original pool of 15 192 non-terygium participants, for the purpose of data

balancing) were selected for algorithm development and internal testing. The detailed selection process was shown in online supplemental figure 1). The dataset was randomly distributed into a development set (N=1685; 2344 eyes) and an internal test set (N=421; 588 eyes) based on a 80:20 ratio at individual level, ensuring that there was no overlap of data of the same individual across the development and internal test set. The development set (80%) was further divided into training (70%) and tuning (10%). The internal test set was not accessed during model development.

We further used datasets from the two independent studies as external test sets. The first external test set was derived from a clinical study conducted at the Singapore Eye Research Institute, consisting of 1001 subjects (1947 study eyes). The second external test set was derived from the Outram Polyclinic study,¹⁸ consisting of 1904 subjects (3493 study eyes). All study subjects provided informed consent.

Inclusion and exclusion criteria

Across development and testing sets, poor quality ASPs such as those with image artefacts due to eye movement or blinking, defocused images, and incomplete view of the nasal of temporal conjunctiva, were excluded. Eyes with pinguecula or mild conjunctival naevus were also included in this study.

Acquisition of colour ASPs

ASPs in SEED and the external test set 1 were taken using a slit-lamp attached digital camera (Topcon model DC-1 with FD-21 flash attachment; Topcon, Tokyo, Japan). During acquisition of the images, the wide beam (diffuse illumination) was set at 45° angle from the viewing system, with magnification set at 16 times. The images were then stored in JPEG format (non-compressed format).

ASPs in the external set 2 were acquired using a hand-held digital camera (MEC-5-ASL-D7100-N85, Miles Research, California, USA). The camera setting was kept constant at aperture priority dial, the side lighting illuminators were angled at 60°. The images were stored as JPEG files (non-compressed format). Finally, we resized all images to dimension of 224 pixels by 224 pixels.

Definitions of pterygium and referable pterygium

Presence of pterygium was determined from colour ASPs by a single examiner (XLF), across all datasets. XLF also cross-referenced with the original clinical recordings of the study eyes, made based on slit lamp evaluations. In the event of ambiguity, further adjudication was performed by ST and C-YC. The intra-grader and inter-grader variability of pterygium grading were assessed, and discussed in detailed in our previous study.¹⁹ In brief, two study ophthalmologists (XLF and ST) performed the grading on 50 anterior segment photos. After 2 weeks, these photos were graded again by one of the study ophthalmologists (XLF). The results showed that our grading had a good intragrader agreement of 0.87 (95% CI 0.83 to 0.97), and inter-grader agreement of 0.80 (95% CI 0.65 to 0.95).

Pterygium was defined based on the appearance of a fibrovascular subepithelial growth extending across the limbus onto the cornea.^{20 21} On the other hand, the definition of referable pterygium was derived based on findings and justification from several past studies. For instance, previous studies reported that pterygia with encroachment onto the cornea >2.25 mm or base width at the limbus >5 mm, were more likely to have corneal astigmatism of ≥ 2 D, and thus ought to warrant for surgical

removal.²²⁻²⁴ In addition, previous study also showed that medium-sized pterygium (defined as extension onto the cornea with length between 2.03.5 mm and vertical length between 5.1 and 7.0 mm) also had higher ocular aberrations compared with small-sized pterygium (with encroachment of <2.00 mm in length or base width <5.00 mm).^{25 26} Building on from these previous findings, we further defined referable pterygium in our study, by using more stringent criteria of >2.50 mm encroachment onto the cornea (measured from the limbus) or base width (at limbus) of >5.00 mm. The size of pterygium was measured using the slit lamp's measurement graticule. For images taken using handheld cameras, the size of the pterygium was first measured in pixels, absolute value conversions (to mm) were done by taking into account the scale factors (microns/pixel) of the handheld camera's magnification. This method has been described in detailed previously.^{18 27}

Development of DL

The general design framework of the algorithm is provided in online supplemental figure 2). In this study, we developed two separate DL algorithms, one for detection of any pterygium, and the other for referable pterygium. The primary inputs to each of the DL algorithms were the ASPs, and the relevant clinical labels (ie, pterygium status). With this annotated data, convolutional neural networks (CNNs) in the form of VGG16 architecture coupled with batch normalisation layers were used.^{28 29} The convolutional layer weights were initialised based on the ImageNet pretrained model.³⁰ These CNNs were used to extract features from the ASPs. Specifically, the 'activation values' of the last CNN layer were 'extracted' and averaged pooled to get a single dimensional tensor for each image. These extracted features were then used to classify the image through a multilayer perceptron (MLP) neural network which is also a part of our DL network (as illustrated in online supplemental figure 2). The MLP network architecture consisted of an input layer, a latent layer and an output layer. The input and latent layer used Rectified Linear Unit activations and the output layer used Sigmoid activation to give an output probability of the image corresponding to a particular class. The number of output neurons in the input, latent and output layer is 4096, 4096 and 2, respectively.

To further reduce overfitting, input images were subjected to data augmentation processes (including random lateral flipping, random rotation of $\pm 10^\circ$ angle, random shearing of $\pm 10^\circ$ angle and random rescaling of between 0.8 and 1.2 times of original image size). Model parameters were optimised using a scheduled decreasing learning rate using the Adam optimiser along with a weighted cross-entropy loss layer. Early stopping regularising method was further used to prevent overfitting on the training set by monitoring validation loss from the internal tuning set.³¹ Due to the greater imbalance between the positive and negative samples of referable pterygium in the training dataset, class penalties were additionally applied while training the referable pterygium model. The final outputs of the two algorithms were the probability for presence of any and referable pterygium, respectively. Details of the optimised model parameters, training codes and the validation script were further described in the following links: https://github.com/SERI-EPI-DS/pterygium_detection/releases/tag/v1.0

https://github.com/SERI-EPI-DS/pterygium_detection#validation

Saliency MAP

To understand which regions of the ASPs were most likely used by the algorithm for prediction of any pterygium and referable

Table 2 Algorithm performances in detection of any pterygium

Testing sets	Detection of any pterygium				
	AUC % (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	Accuracy % (95% CI)	MCC
Internal:					
SEED	99.5 (99.0 to 99.9)	98.6 (97.0 to 99.7)	99.0 (97.7 to 100.0)	98.8 (98.0 to 99.7)	0.976
External:					
Set 1	99.1 (97.4 to 99.9)	95.9 (89.2 to 100.0)	98.5 (97.8 to 98.9)	98.4 (97.8 to 98.9)	0.782
Set 2	99.7 (99.4 to 99.8)	100.0 (100.0 to 100.0)	88.3 (87.2 to 89.4)	88.6 (87.5 to 89.7)	0.400

AUC, area under the curve; MCC, Matthew's correlation coefficient; SEED, Singapore Epidemiology of Eye Diseases.

pterygium, we generated saliency maps using the Grad-CAM technique,³² highlighting regions in the image which contributed more towards the predicted output (ie, hotter colour indicating greater contribution).

Statistical analysis

To evaluate the respective algorithm's performance, we used the metrics of area under the curve (AUC), sensitivity, specificity and accuracy. The optimal classification threshold was selected based on the Youden's index which denotes the maximal point for sensitivity and specificity values (). The threshold value determined from the internal tuning set was 0.361 for any pterygium, and 0.271 for referable pterygium. Additionally, we evaluated the precision (ie, positive predictive value) of the algorithm and plotted the precision–recall curves. The 95% CIs for these performance metrics were computed using non-parametric bootstrapping with 2000 bootstrap replicates. In addition, the Matthew's correlation coefficient values (MCC) were also calculated. The MCC can be interpreted as a discretisation of the Pearson correlation for binary variables.³³ MCC values between 0.81 and 1.00 indicate strong correlation; values between 0.61 and 0.80 indicate good correlation; and values between 0.41 and 0.60 indicate moderate correlation. Values less than 0.40 indicate poor correlation.

RESULTS

We developed the DL algorithms using 2503 colour ASPs (inclusive of training and tuning) from the SEED study. We further validated the performance of the algorithms using 629 ASPs from the internal test set, 2610 from external test set 1 and 3701 from external test set 2. The mean age was 63.0±9.7 years in the internal test set, 64.2±7.8 years in external test set 1 and

62.1±7.0 years in external test set 2. Additional study participant demographics and characteristics are summarised in table 1.

We first examined the performance of the algorithm for detection of any pterygium (table 2, online supplemental figure 3A). In the internal test set, the AUC for detection of pterygium was 99.5% (95% CI 99.0% to 99.9%) with sensitivity of 98.6% and specificity of 99.0%. In external test set 1, the AUC for detection of pterygium was 99.1% (95% CI 97.4% to 99.9%) with sensitivity of 95.9%, and specificity of 98.5%. In external test set 2, the AUC for detection of pterygium was 99.7% (95% CI 99.4% to 99.8%) with sensitivity of 100.0%, and specificity of 88.3%. The algorithm showed very strong correlation (MCC=0.976) in the internal test set, and moderate-to-strong correlation across the external test sets (MCC ranged between 0.40 and 0.782).

On the other hand, for detection referable pterygium (table 3, online supplemental figure 3B), the internal test set demonstrated AUC of 98.5% (95% CI 96.4% to 99.6%) with sensitivity of 94.0% and specificity of 95.3%. The external set 1 demonstrated AUC of 99.7% (95% CI 99.3% to 100.0%) with sensitivity of 87.2% and specificity of 99.4%. The external set 2 showed an AUC of 99.0% (95% CI 97.4% to 99.8%) with sensitivity of 94.3% and specificity of 98.0%. The MCC was good in internal test set (0.745), and moderate-to-good across external test sets (MCC ranged between 0.536 and 0.660).

In addition, we further evaluated the precision–recall curves and only the internal datasets demonstrated relatively good precision performance (precision of 97% for any pterygium, and 66.2% for referable pterygium, online supplemental table 1, online supplemental figure 4). The poorer precision values in external test sets could be due to the small number of positive cases.

Table 1 Characteristics of the development and testing datasets

Characteristics	Development set (SEED)		Internal test set (SEED)	External test set 1 (SNEC clinic study)	External test set 2 (polyclinic study)
	Training set	Tuning set			
Digital camera types	Slit-lamp attached	Slit-lamp attached	Slit-lamp attached	Slit-lamp attached	Hand-held
Number of subjects	1474	211	421	1001	1904
Number of eyes	2057	287	588	1947	3493
Number of images	2198	305	629	2610	3701
Age, years	64.2 (10.1)	64.3 (10.2)	63.0 (9.7)	64.2 (7.8)	62.1 (7.0)
Male gender, n	864 (58.6%)	114 (54.0%)	233 (55.3%)	375 (37.6%)	713 (37.5%)
Number of pterygium cases (by images):					
Any pterygium, n	1085 (49.4%)	160 (52.5%)	321 (51.0%)	73 (2.8%)	92 (2.5%)
Referable pterygium*, n	219 (10.0%)	27 (8.9%)	53 (8.4%)	20 (0.7%)	34 (0.9%)

Data presented as mean (SD) or number (percentage), where appropriate.

*Referable pterygium defined as presence of pterygium with >2.50 mm encroachment onto the cornea (measured from the limbus) or base width (at limbus) of >5.00 mm.

SEED, Singapore Epidemiology of Eye Diseases study; SNEC, Singapore National Eye Centre.

Table 3 Algorithm performances in detection of referable pterygium

Testing sets	Detection of referable pterygium*				
	AUC % (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	Accuracy % (95% CI)	MCC
Internal:					
SEED	98.5 (96.6 to 99.6)	94.0 (85.7 to 100.0)	95.3 (93.3 to 96.8)	95.2 (93.3 to 96.8)	0.745
External:					
Set 1	99.7 (99.3 to 100.0)	87.2 (63.6 to 100.0)	99.3 (98.9 to 99.6)	99.3 (98.9 to 99.6)	0.660
Set 2	99.0 (97.4 to 99.8)	94.3 (84.2 to 100.0)	98.0 (97.5 to 98.4)	98.0 (97.5 to 98.4)	0.536

*Referable pterygium defined as presence of pterygium with >2.50 mm encroachment onto the cornea (measured from the limbus) or base width (at limbus) of >5.00 mm. AUC, Area under receiver operating characteristic curve; MCC, Mathew's correlation coefficient; SEED, Singapore Epidemiology of Eye Diseases.

In figures 1 and 2, the saliency maps highlighted regions within the ASP which the DL algorithm likely focused on when predicting presence of any pterygium and referable pterygium. Generally, the highlighted regions corresponded well with the actual site of pterygium.

DISCUSSION

Using nearly 10 000 images from population-based and clinic-based datasets, we developed and tested a novel ASP-based DL algorithm for the detection of any and referable pterygium. The DL algorithms demonstrated optimal performances with high sensitivity and specificity. Furthermore, when applying this algorithm on ASPs taken from hand-held digital cameras, we observed a similarly good performance in the detection of referable pterygium as compared with ASPs from slit-lamp mounted digital cameras. Our proof-of-concept findings indicate that this algorithm coupled with a handheld camera may be deployed as a simple, automated and cost-saving alternative for the screening of referable pterygium.

A key strength of our study was the use of saliency maps to elucidate the algorithm's 'decision-making process' in making predictions for any and referable pterygium. The highlighted regions were congruent with the actual site of pterygium, showing that the algorithm was making predictions based on relevant and clinically appropriate features of pterygium. Building on these illustrations, it is also conceivable to incorporate these clinically informative saliency maps as part of the screening results during deployment, therefore further facilitating the clinical adoption of this algorithm.

Our study had substantially larger sample size (n=9443 images) compared with two previous studies which also developed a DL algorithm for the detection of pterygium, based on ASP (Zulkifley *et al*¹⁶ consisted of only 120 images; and Zhang *et al*¹⁵ consisted of 450 images from single data source). The development and

internal testing sets of these two past studies were limited in sample size, and external testing was not performed. Performance wise, Zulkifley *et al* reported an AUC of 0.97. On the other hand, Zhang *et al* developed reported an accuracy of 93%. In comparison, our current study demonstrated more superior performance than these past studies in detecting any pterygium (AUC=99.5%), and was able to further substantiate our findings with replication on external test sets. Furthermore, compared with these past studies, our algorithms were trained based on dataset from a multiethnic population-based study. In addition, external testing was also performed in two independent clinical studies, further demonstrating the generalisability of these algorithms. Furthermore, both past studies^{15 16} only focused on identifying presence of any pterygium. In this regard, an algorithm that detects any presence of pterygium would have poorer clinical utility and would result in unnecessary referrals because not all pterygium cases (especially mild types) need to have surgical interventions. For this reason, in our current study, we especially developed an algorithm which could also detect referable pterygium, identifying pterygium type with substantial extension or size which justifies for surgical removal. Although Zhang *et al*¹⁵ also attempted an algorithm that identified 'pterygium type which required treatment', the definition and criteria of this pterygium type was not clearly described in their article, thus limiting the interpretation of this algorithm's performance. Additionally, to evaluate whether variation in pupil size (ie, dilated and non-dilated eyes) affected the performance of the algorithm. Using the internal test set, we further performed subgroup analyses, stratified by dilated and non-dilated eyes (online supplemental table 2). We observed that the algorithm's performance was similar across the two groups, indicating that different pupil sizes unlikely affected the algorithm's performance.

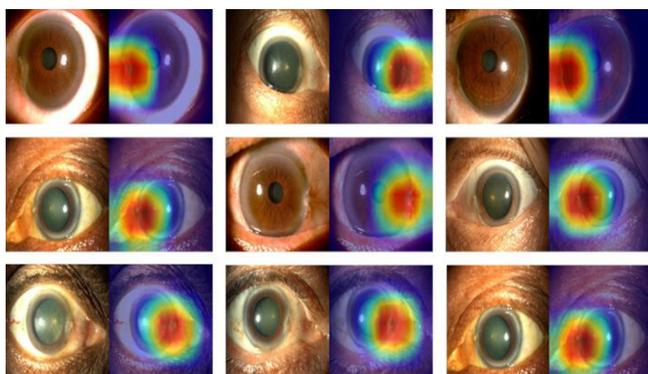


Figure 1 Saliency maps highlighting regions which the deep learning algorithm predicted for presence of any pterygium.

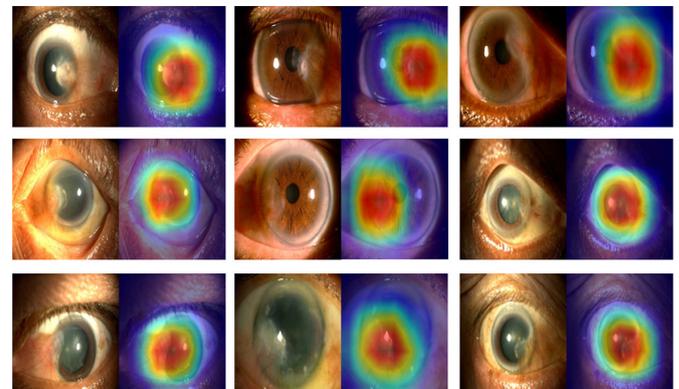


Figure 2 Saliency maps highlighting regions which the deep learning algorithm predicted for presence of referable pterygium.

We further investigated images which were misclassified by the algorithm for referable pterygium. Across the internal and external test sets, we observed a number of false negative misclassifications for pterygia that were more translucent in appearance (online supplemental figure 5). This indicates that the algorithm was more likely to miss pterygia of such semi-transparent type. Hence, additional refinement and training of the algorithm involving more of such cases may be needed to further improve the algorithm's performance. However, more transparent pterygium is usually associated with less underlying corneal scarring, and have a lower likelihood of sight-threatening consequence when not detected, compared with typical 'fleshy' type advanced pterygium.²⁰ On the other hand, the false positive misclassifications were mainly attributed to severe corneal arcus and iris atrophy (online supplemental figure 6). The saliency maps also further illustrated that the algorithm likely interpreted these appearance as the 'feature' responsible for the prediction of pterygium (online supplemental figure 6). This observation indicates that further training involving more cases of iris atrophy and corneal arcus are needed, to minimise the algorithm's false positive rate. In addition, it is worthy to note that among non-ptyerigium eyes with pinguecula and conjunctival naevus, the algorithm still correctly identified these eyes as 'non-referable pterygium cases', none of these cases were mistaken as false positive by the algorithm (data not shown in tables).

Despite the promising findings shown by both algorithms individually, we observed potential shortcomings of each algorithm if deployed solely 'on its own'. There were some instances whereby the same image was classified as 'absence of any pterygium' by any pterygium algorithm, but incorrectly identified as 'referable pterygium' by the referable pterygium algorithm (ground truth was non-ptyerigium). There were two of such misclassifications made by the referable pterygium algorithm in internal test set, 43 in external test set 1 and 140 in external test set 2 (data not shown in tables), indicating that the any pterygium algorithm was more accurate in identifying non-ptyerigium cases. On the other hand, there was another scenario, whereby 'positive cases' flagged up by the any pterygium algorithm were instead identified by the referable pterygium algorithm as 'non-referable' (and the ground truth was indeed non-referable pterygium). There were 157 of such 'classifications' made by the any pterygium algorithm in internal test set, 43 in external test set 1 and 140 in external test set 2 (data not shown in tables). This also indicates that, sole reliance on any pterygium algorithm would indeed result in unnecessary referrals of mild pterygium cases. In this instance, the addition of the referable pterygium algorithm would help to better determine if surgical referral was indeed needed. Hence, in order to better leverage on the merits of both models, a 'stacked approach' which integrates both algorithms sequentially may be viable for eventual real-world deployment (conceptually illustrated in online supplemental figure 7). In brief, model 1 (any pterygium algorithm) would first be deployed to analyse the image, if an 'absence of pterygium' output was generated, no further investigations/actions would be needed. On the other hand, if 'presence of any pterygium' was detected by model 1, then the same image would be further analysed by model 2 (referable pterygium algorithm) to determine if the pterygium is non-referable or referable type. Nevertheless, future real-world evaluation on the performance of this stacked approach in identifying referable pterygium is still needed.

Our study has several limitations. First, this study was trained and tested on Asian eyes only; the generalisability of this algorithm to other ethnic groups remains to be evaluated. Second, the external test sets in our study had limited cases of referable

ptyerigium. Hence, further validations in larger clinical studies, and subsequently in 'real world' community settings would be needed to test the algorithm's true clinical utility as a screening tool. Third, in this proof-of-concept study, to determine the best performance of the algorithm, we used Youden's index, which provides a threshold with balanced maximisation of sensitivity and specificity. However, for real-world deployment, other contextual factors need to be further taken into account when determining the eventual classification thresholds. These considerations include, the deployment site (ie, whether in rural community screening sites or primary care facilities), local regulatory requirements for implementation of health technology (which may require minimal levels of specificity and sensitivity to be achieved before rollout is granted), and availability of healthcare resources/ facilities for treatment (ie, communities with finite resources may opt for a more stringent threshold which would yield higher positive predictive value, to more strictly identify those which truly require treatment.³⁴ Fourth, it should be noted that the ground truth of referable pterygium was defined based on the presence of pterygium with >2.50 mm extension towards the cornea or with a base width of >5.00 mm, but did not take total area and fleshiness (ie, thickness)²⁰ of the pterygium into account. However, previous studies have indicated that horizontal extension along with the base width of the pterygium^{23–26} have the greatest influence on corneal astigmatism and ocular aberrations. Thus, the omission of total area and fleshiness from our definition criteria, unlikely had major bearing on our findings. Lastly, the current study did not include other limbal disorders such as phlyctenulosis, peripheral corneal ulcer, limbal tumour. However, it should be noted that these cases are rare, and difficult to be curated in sufficient numbers for DL purpose. Future work which involves detection of various ocular surface diseases using DL is viable. However, the development of such DL model would require further curation of such cases from large hospital-based studies.

In conclusion, we developed and validated novel ASP-based DL algorithms, showing robust performance in external datasets, including one with hand-held ASP images. This suggests our algorithm may potentially be used as a simple screening tool for the detection of referable pterygium. Nevertheless, further validation of the algorithm in community setting is required. If validated, this algorithm may also be useful for under-resourced rural communities, where ophthalmologists and primary physicians are scarcely available, and access to eye care is poor.

Author affiliations

¹Singapore Eye Research Institute, Singapore National Eye Centre, Singapore

²Department of Ophthalmology, Shanghai Eye Diseases Prevention and Treatment Center, Shanghai Eye Hospital, Shanghai, China

³Ophthalmology and Visual Sciences Academic Clinical Program, Duke-NUS Medical School, Singapore

⁴Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore

Contributors Conception and design: XLF, C-YC, TR and YCT. Analysis and interpretation: XLF, MLC, MD, Z-DS, ST and YCT. Data collection: XLF, MLC, MD, Z-DS, ST, Y-CL, TR, C-YC and YCT. Manuscript preparation and overall responsibility: XLF, MD, MLC, Z-DS, ZLT, Y-CL, JM, RH, TYW, C-YC, TR and YCT. All authors approved the final manuscript.

Funding This study is supported by the Singapore Ministry of Health's National Medical Research Council (NMRC/CIRG/1488/2018, NMRC/OFLCG/004a/2018). YCT is supported by the Singapore Ministry of Health's National Medical Research Council [NMRC/MOH-TA18nov-0002]. XLF is supported by the Natural Science Foundation of Shanghai (18ZR1435600). The sponsor or funding organization had no role in the design or conduct of this research.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval All study procedures adhered to the principles of the Declaration of Helsinki and informed consent was obtained from all study participants, ethical approval was obtained from the SingHealth Centralized Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data request can be made to corresponding author Dr Yih-Chung Tham.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Xiaoling Fang <http://orcid.org/0000-0003-1803-2580>

Zhi-Da Soh <http://orcid.org/0000-0002-1182-3489>

Zhen Ling Teo <http://orcid.org/0000-0003-3443-8601>

Sahil Thakur <http://orcid.org/0000-0002-7948-6992>

Tien Yin Wong <http://orcid.org/0000-0002-8448-1264>

Ching-Yu Cheng <http://orcid.org/0000-0002-4916-6074>

Tyler Hyungtaek Rim <http://orcid.org/0000-0001-6465-2620>

Yih-Chung Tham <http://orcid.org/0000-0002-6752-797X>

REFERENCES

- Tomidokoro A, Miyata K, Sakaguchi Y, et al. Effects of pterygium on corneal spherical power and astigmatism. *Ophthalmology* 2000;107:1568–71.
- Chu WK, Choi HL, Bhat AK, et al. Pterygium: new insights. *Eye* 2020;34:1047–50.
- Sutphin. Basic and clinical science course. *Section 8: External Disease And Cornea* 2017-2018;2018.
- Zhong H, Chen Q, Li J, et al. Ethnic variations in pterygium in a rural population in southwestern China: the Yunnan minority eye studies. *Ophthalmic Epidemiol* 2016;23:116–21.
- Rezvan F, Khabazkhoob M, Hooshmand E, et al. Prevalence and risk factors of pterygium: a systematic review and meta-analysis. *Surv Ophthalmol* 2018;63:719–35.
- Liu L, Wu J, Geng J, et al. Geographical prevalence and risk factors for pterygium: a systematic review and meta-analysis. *BMJ Open* 2013;3:e003787.
- Durkin SR, Abhary S, Newland HS, et al. The prevalence, severity and risk factors for pterygium in central Myanmar: the Meiktila eye study. *Br J Ophthalmol* 2008;92:25–9.
- Lu P, Chen X, Kang Y, et al. Pterygium in Tibetans: a population-based study in China. *Clin Exp Ophthalmol* 2007;35:828–33.
- McCarty CA, Fu CL, Taylor HR. Epidemiology of pterygium in Victoria, Australia. *Br J Ophthalmol* 2000;84:289–92.
- Cajucum-Uy H, Tong L, Wong TY, et al. The prevalence of and risk factors for pterygium in an urban Malay population: the Singapore Malay eye study (SIMES). *Br J Ophthalmol* 2010;94:977–81.
- Resnikoff S, Lansingh VC, Washburn L, et al. Estimated number of ophthalmologists worldwide (international Council of ophthalmology update): will we meet the needs? *Br J Ophthalmol* 2020;104:588–92.
- Mahar PS, Manzar N. Pterygium recurrence related to its size and corneal involvement. *J Coll Physicians Surg Pak* 2013;23:120.
- Hirst LW. The treatment of pterygium. *Surv Ophthalmol* 2003;48:145–80.
- Tomidokoro A, Oshika T, Amano S, et al. Quantitative analysis of regular and irregular astigmatism induced by pterygium. *Cornea* 1999;18:412–5.
- Zhang K, Liu X, Liu F, et al. An interpretable and expandable deep learning diagnostic system for multiple ocular diseases: qualitative study. *J Med Internet Res* 2018;20:e11144.
- Zulkifley MA, Abdani SR, Zulkifley NH. Pterygium-Net: a deep learning approach to pterygium detection and localization. *Multimed Tools Appl* 2019;78:34563–84.
- Majithia S, Tham Y-C, Chee M-L, et al. Cohort profile: the Singapore epidemiology of eye diseases study (seed). *Int J Epidemiol* 2021;50:41–52.
- Tun TA, Chua J, Shi Y, et al. Association of iris surface features with iris parameters assessed by swept-source optical coherence tomography in Asian eyes. *Br J Ophthalmol* 2016;100:1682–5.
- Fang XL, Chong CCY, Thakur S, et al. Ethnic differences in the incidence of pterygium in a multi-ethnic Asian population: the Singapore epidemiology of eye diseases study. *Sci Rep* 2021;11:501.
- Tan DT, Chee SP, Dear KB, et al. Effect of pterygium morphology on pterygium recurrence in a controlled trial comparing conjunctival autografting with bare sclera excision. *Arch Ophthalmol* 1997;115:1235–40.
- Saw SM, Tan D. Pterygium: prevalence, demography and risk factors. *Ophthalmic Epidemiol* 1999;6:219–28.
- Mohammad-Salih PA-K, Sharif AFMD. Analysis of pterygium size and induced corneal astigmatism. *Cornea* 2008;27:434–8.
- Kampitak K. The effect of pterygium on corneal astigmatism. *J Med Assoc Thai* 2003;86:16–23.
- Avisar R, Loya N, Yassar Y, et al. Pterygium-induced corneal astigmatism. *Isr Med Assoc J* 2000;2:14–15.
- Gumus K, Erkilic K, Topaktas D, et al. Effect of pterygia on refractive indices, corneal topography, and ocular aberrations. *Cornea* 2011;30:24–9.
- Gumus K, Guven A, Altinkaynak M, et al. Comparison of different measurement tools and dimensional parameters of pterygium to investigate its impact on refractive indices and ocular aberrations. *Eye Contact Lens* 2018;44:118–24.
- Sidhartha E, Gupta P, Liao J, et al. Assessment of iris surface features and their relationship with iris thickness in Asian eyes. *Ophthalmology* 2014;121:1007–12.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv* 2014;1556.
- Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv* 2015;1502.
- Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge. *Int J Comput Vis* 2015;115:211–52.
- Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. *advances in neural information processing systems* 2001.
- Grad-cam: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on computer vision* 2017.
- Boughorbel S, Jarray F, El-Anbari M, Fethi J. Optimal classifier for imbalanced data using Matthews correlation coefficient metric. *PLoS One* 2017;12:e0177678.
- Tham Y-C, Anees A, Zhang L, et al. Referral for disease-related visual impairment using retinal photograph-based deep learning: a proof-of-concept, model development study. *Lancet Digit Health* 2021;3:e29–40.