REVIEW ARTICLE

Artificial intelligence and external photographs in ophthalmology: a systematic review

Kenneth Ka Hei Lai[1,2], Carmen Sze Ching Lo[3*], Han Wang[1], Xiaoyan Hu[1], Fatema Mohamed Ali Abdulla Aljufairi[1,4], Jake Uy Sebastian[1,5], Chi Pui Pang[1], Kelvin Kam Lung Chong[1,2,6]*

[1]Department of Ophthalmology and Visual Sciences, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China
[2]Department of Ophthalmology and Visual Sciences, Prince of Wales Hospital, Hong Kong SAR, China
[3]Department of Ophthalmology, Department of the Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China
[4]Department of Ophthalmology, Salmaniya Medical Complex, Government Hospitals, Bahrain
[5]Department of Ophthalmology, Vicente Sotto Memorial Medical Center, Cebu City, Philippines
[6]Hong Kong Eye Hospital, Hong Kong SAR, China

*Contributed equally

Correspondence and reprint requests:
Dr Kelvin Kam Lung Chong, Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong, Hong Kong SAR, China. Email: chongkamlung@cuhk.edu.hk

Abstract

We systematically reviewed the literature regarding the use of artificial intelligence trained with external photographs, defined as unprocessed clinical images taken from cameras and slit-lamps, for measurement of eyelid/periorbital parameters and detection and classification of multiple ophthalmic diseases including blepharoptosis, thyroid eye disease, eyelid tumors, keratitis, trachoma, pterygium, diabetic retinopathy, cataract, strabismus, and other oculofacial disorders.

Key words: Artificial intelligence; Face; Ophthalmology

Introduction

Artificial intelligence (AI) is a type of machine learning that applies complex algorithms to analyze large amounts of data.[1,2] AI has demonstrated good performance in the management of diabetic retinopathy, glaucoma, age-related macular degeneration, and myopia.[3-6] Fundus photography and optical coherence tomography are commonly used in the field of AI ophthalmology. Facial images have been used in AI for facial recognition, emotion recognition, and medical diagnoses (eg, Down syndrome, fetal alcohol syndrome, and other genetic disorders).[7-10] In recent years, AI has been used to analyze external photographs, defined as unprocessed clinical images taken from cameras and slit-lamps, for multiple ophthalmic diseases.

Reliable screening tools can facilitate timely diagnoses and prompt treatment, thereby reducing morbidity and health burden. Technological advances enable the use external photographs in AI.[11,12] Facial images can readily be acquired using cameras, making them cost-effective and convenient screening tools for multiple ophthalmic diseases such as thyroid eye disease (TED) and eyelid tumors, especially among patients with poor access to healthcare services. Ophthalmology is one of the highest-demand healthcare fields; the use of AI and external photographs has great potential to revolutionize ophthalmology services by shortening waiting time, enhancing disease management, and improving patient outcomes.[13] We systematically reviewed the literature regarding clinical applications of AI and external photographs in the field of ophthalmology.

Methods

The PubMed, Scopus, Web of Science, and IEEE databases was systematically searched on 27 April 2023 using key words: 'artificial intelligence' OR 'machine learning' OR 'neural network' OR 'deep learning' and a combination of 'facial landmarking' OR 'external photo' OR 'face photo' OR 'eyelid' OR 'anterior segment photograph' OR 'anterior segment image' OR 'anterior segment photo'

OR 'pterygium'. The PRISMA guidelines were followed without language restrictions. The quality of included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies 2.

Inclusion criteria were (1) clinical trials, prospective and retrospective cohort studies, comparative studies, validation, or evaluation studies, and epidemiologic studies; (2) studies involving facial images or external photographs in AI for detecting ophthalmic conditions; and (3) original articles. Exclusion criteria were (1) studies for which the full text was unavailable; (2) studies reporting the use of AI for non-ophthalmic conditions; (3) studies involving images taken from optical coherence tomography (OCT), fundus photography, or oculus keratography; and (4) review articles without original data.

Two reviewers independently searched the literature and screened the abstracts of identified studies. Disagreements were resolved by discussion with four senior reviewers. The two reviewers then independently assessed the full text of eligible studies, and data were extracted and cross-validated. In cases of unresolved disagreement, a thorough discussion involving all reviewers was conducted until a consensus was reached. The study selection process based on the PRISMA flow chart[14] is shown in the **Figure**. A customized form was used to record year of publication, study population, types of ophthalmic conditions, number and focal region of external photographs, number of test sets, types of deep neural network, and key findings. The risk of bias was assessed by two reviewers using the Quality Assessment of Diagnostic Accuracy Studies 2.

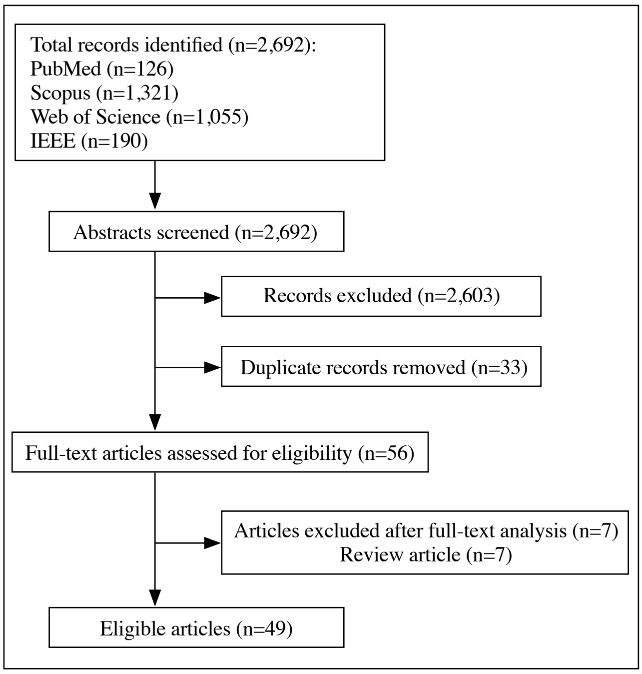

Figure. Flowchart showing the literature search process.

Results

In total, 49 studies were included in the systematic review (**Table 1**). Among the 49 studies, 13 exhibited low risk for all four risk of bias parameters and six displayed low risk for all three applicability concerns parameters (**Table 2**).

Measurement of eyelid/periorbital parameters

The eyelid is usually assessed during evaluation of other ophthalmic disorders. The eyelid position is measured by placing an opaque ruler in front of the patient's eye.[15] Moriyama et al[16] used an AI system trained with two-dimensional facial images to parameterize the fine structures of the eyes including the position of the iris, degree of eyelid opening, and the shape, complexity, and texture of the eyelid. The system also estimated the movement parameters of the eyelids and iris. This model failed to match well only five out of 577 images. Thomas et al[17] compared manual and AI-automated measurements of the vertical palpebral aperture and reported good agreement between the two methods (Pearson correlation coefficient=0.87); 94% of measurements were within two standard deviations of the mean. Van Brummen et al[18] used AI to assess eight periorbital measurements including marginal reflex distance (MRD) 1 and 2, medial and lateral canthal heights, medial and lateral brow heights, and medial and lateral intercanthal distances. The performance of AI closely agreed with that of human graders; the mean absolute differences were 0.5 mm for MRD 1 and 2 and lateral and medial canthal heights, 1.5 to 2 mm for lateral and medial brow heights, and 2 to 4 mm for medial and lateral intercanthal distances. Chen et al[19] described the first smartphone-based AI-assisted image-processing algorithm for measurements of MRD 1 and 2 and levator function; the respective Pearson correlation coefficients between the smartphone-based AI and gold standard AI-assisted software were 0.91, 0.84, and 0.73. Cao et al[20] compared AI and manual segmentations of eyelid features on two-dimensional digital photographs; the Dice coefficients for eyelid and corneal segmentation between the two methods were 0.922 and 0.973, respectively. Zhu et al[21] used an automated algorithm to measure eyelid position based on nine parameters and reported 91% accuracy. Manual measurement of eyelid parameters can be time-consuming, subjective, and unrepeatable. AI-assisted eyelid assessments may enhance efficiency and repeatability in busy clinical settings.

Blepharoptosis

Blepharoptosis refers to the drooping of inferior displacement of the upper eyelid obstructing the visual axis. Hung et al[22] used an AI system to identify blepharoptosis in clinical photographs and achieved 92% sensitivity and 88% specificity. Tabuchi et al[23] developed a smartphone application that used AI to automatically detect blepharoptosis in facial images; the accuracy was 83% and the area under the curve (AUC) was 0.9. Hung et al[24] used an AI model to identify referable blepharoptosis using facial images; the model performed better than a non-ophthalmic physician. Lou et al[25] used AI to automatically measure eyelid morphological properties before and after blepharoptosis surgery and then compared the results with those of manual measurements. The interclass correlation coefficient between the two methods for measurement of MRD 1 and 2 ranged from 0.93 to 0.97. Şimşek et al[26] analyzed surgical outcomes after upper eyelid surgery (blepharoplasty and Müller's muscle-conjunctival resection) using AI facial landmark detection based on palpebral distance, eye-opening area, and average brow height. Qu et al[27] used AI-based facial images to assess blepharoplasty outcomes in terms of lower eyelid skin wrinkles, eyelid lacrimal sulci, skin gloss, and aesthetic scores. Sun et al[28] used fully automatic postoperative appearance prediction based on facial images and reported an overall performance of 86% in predicting postoperative appearance after blepharoptosis surgery. AI based on facial images can detect blepharoptosis and predict postoperative appearance after upper eyelid surgery.

Thyroid eye disease

TED is an autoimmune inflammatory disease that occurs in up to 40% of patients with Graves' disease; it is the most common orbital disorder in adults.[29] Prompt diagnosis and treatment are essential because mild disease can progress to sight-threatening conditions such as exposure keratopathy and dysthyroid optic neuropathy.[30] Yoo et al[31] used a deep learning model to predict postoperative appearance (after orbital decompression) in patients with TED using external ocular images. Subjective quality assessment by three ophthalmologists showed promising results; AI can be a supportive tool for prediction of postoperative images and decision making. Moon et al[32] used a machine learning system to detect TED and predict the clinical activity score (CAS) using external photographs. The system achieved 88% sensitivity and 87% specificity for TED detection and predicted the CAS within one point of the reference in up to 85% of patients. Karlin et al[33] used AI to detect TED and differentiate severity using external photographs and reported 85% accuracy, with higher recall for active cases. Huang et al[34] used an AI network to detect TED based on nine eye positions; the network displayed an average AUC exceeding 0.85. Shao et al[35] used a deep learning system to automatically measure eyelid morphology in TED patients and reported strong agreement between automatic and manual measurements. The use of AI systems to detect and evaluate TED may enable timely diagnosis and management.

Eyelid tumors

Eyelid tumors constitute 5% to 10% of all skin tumors, and timely diagnosis is important to minimize both morbidity and mortality.[36] Typically, an eyelid tumor is detected during eyelid examination, and histopathological examination is required to confirm the diagnosis. Adamopoulos et al[37] used an AI system to identify eyelid basal cell carcinoma using external photographs and reported a maximum accuracy of 100%. Li et al[38] used AI to distinguish malignant and benign eyelid tumors based on external photographs and reported a maximum accuracy of 81.8%. Hui et al[39] used an AI system to differentiate malignant from benign eyelid tumors using facial images and exhibited a maximum accuracy of 81%; malignant tumor images comprised basal cell carcinoma (49%), sebaceous adenocarcinoma (26%), squamous cell carcinoma (10%), and melanoma (7%). Lee et al[40] used an AI network to differentiate malignant from benign/no eyelid lesions; the mean AUCs of the two convolutional neural network (CNN) models were 0.91 and 0.95. Patients with eyelid tumors may benefit from AI that uses facial images to identify high-risk lesions for early referral.

Keratitis

Keratitis is a major cause of visual impairment and can rapidly progress to corneal perforation and vision loss if left undiagnosed; early detection and management can halt disease progression and improve visual outcomes.[41] Infectious keratitis can be classified as bacterial, fungal, parasitic, and viral; identification of the causative agent is important for clinical management.[41] Xu et al[42] first described an AI system for keratitis diagnosis based on slit-lamp images solely and reported 80% accuracy, which was significantly higher than the 49.27%±11.5% by human ophthalmologists. Ji et al[43] developed a multi-attribute AI network based on anterior segment images and reported an average diagnostic accuracy of 84.89% and a maximum accuracy of 89.51%. Kuo et al[44] compared the performance of various deep learning algorithms in diagnosing bacterial keratitis using external eye photographs; the diagnostic accuracies of all models (69%-72%) were comparable to those of ophthalmologists (66%-74%), with the EfficientNet B3 exhibiting the best average AUC (74% sensitivity, 64% specificity, and 77% positive and 61% negative predictive values). Natarajan et al[45] used an AI system to detect viral keratitis based on slit-lamp images and reported a maximum accuracy of 72%. Hung et al[46] used an AI system based on slit-lamp images to differentiate bacterial from fungal keratitis and reported a maximum accuracy of 80.0% (79.6% to 95.9% for bacterial keratitis and 26.3% to 65.8% for fungal keratitis). Redd et al[47] developed a corneal photograph-based AI model to differentiate bacterial from fungal corneal ulcers; the best-performing model achieved a maximum AUC of 0.83, compared with 0.42 to 0.79 by human experts. Zhang et al[48] used a deep learning system and reported a maximum diagnostic accuracy of 77.08% for infectious keratitis (70.27% for bacterial keratitis, 77.71% for fungal keratitis, 83.81% for *Acanthamoeba* keratitis, and 79.31% for herpes simplex keratitis). Hu et al[49] evaluated six deep learning algorithms based on slit-lamp images and reported a maximum accuracy of 0.735 and a maximum AUC of 0.85 (0.87 for viral keratitis, 0.87 for fungal keratitis, and 0.64 for bacterial keratitis). Kogachi et al[50] used CNNs to differentiate culture-positive from culture-negative infectious keratitis based on corneal photographs; the CNNs failed to detect morphological differences between culture-positive and culture-negative statuses on external photographs. Loo et al[51] used AI to identify biomarkers of infectious keratitis using external photographs and reported good accuracy (Dice coefficient 0.62-0.95); these biomarkers were strongly correlated (0.84) with best-corrected visual acuity (BCVA).[52] AI can facilitate diagnosing infectious keratitis and identifying the causative pathogen based on external photographs.

Table 1. Summary of studies included in the systematic review

Study	Study aims	Outcome measures	Sample size	Origin of images	Algorithm/ architecture
Measurement of eyelid/periorbital parameters					
Moriyama et al,[16] 2006	To propose a system for detailed analysis of the eye region	Accuracy by comparing positions of model points for upper and lower eyelids, as well as iris center	577 images	-	Generative eye region model
Thomas et al,[17] 2020	To determine whether OpenFace (AI) can extract clinically useful measurements from images before and after ptosis correction	Agreement in interpupillary distance to vertical palpebral aperture ratio between AI and clinicians: Bland-Altman plots	32 patients	-	-
Van Brummen et al,[18] 2021	To develop an AI tool for standardized objective assessment of eyelid and periorbital soft tissue position	Accuracy of segmentation: Dice coefficient; agreement between graders: Bland-Altman plots, mean absolute difference, ICC	418 images (retrospective), 42 images (prospective); training/ validation size: 397 (80%/20%); test size: 21	United States	-
Chen et al,[19] 2021	To propose the first smartphone-based AI-assisted image processing algorithm for measurements of MRD 1 and 2 and levator function	Pearson correlation coefficients: agreement between model and gold standard in terms of ICC, mean absolute error, and limits of agreement	822 eyes in 411 participants; training size: 72%; validation size: 18%; test size: 10%	Taiwan	-
Cao et al,[20] 2021	To propose a deep learning approach for automatic and objective evaluation of morphologic eyelid features using two-dimensional digital photographs, and to assess agreement between automatic and manual measurements	Accuracy of model: Dice coefficient; agreement between manual and automatic MRDs: Spearman correlation coefficient, intraclass correlation coefficient, Bland-Altman plots	406 eyes in 203 normal participants; automatic morphologic measurement of eyelids: 203; automatic eyelid and cornea identification network: training size: 872; validation size: 220; test size: 276	China	-
Zhu et al,[21] 2022	To develop an automated algorithm for measuring ophthalmic parameters to facilitate diagnosis of ptosis, eyelid retraction, and other eye pathologies	Object detection: recall, precision, accuracy, IoU; parameter measurement results: recall, precision, accuracy, IoU	8,200 images; training size: 75%; validation size: 10%; test size: 15%	China	-
Blepharoptosis					
Hung et al,[22] 2021	To develop an algorithm for accurate identification of blepharoptosis from clinical photographs	Accuracy, sensitivity, specificity, PPV, NPV, F1 score, AUC	500 images; training size: 136 normal, 177 blepharoptosis; validation size: 34 normal, 45 blepharoptosis; test size: 18 normal, 24 blepharoptosis	Taiwan	-
Tabuchi et al,[23] 2022	To assess AI performance in automated classification of images taken using a tablet for individuals with blepharoptosis and with normal eyelid	Accuracy, sensitivity, specificity, AUC	1276 images of 606 participants (624 from 347 blepharoptosis cases, 652 from 367 normal participants)	Japan	-
Hung et al,[24] 2022	To develop an AI model that accurately and automatically identifies referable blepharoptosis, then compare its performance with that of non-ophthalmic physicians	Accuracy, sensitivity, specificity, AUC	782 images; training size: 587; validation size: 145; test size: 25 ptotic eyes, 25 healthy eyes	-	-
Lou et al,[25] 2021	To propose a novel deep learning image analysis method that automatically measures eyelid morphological properties before and after blepharoptosis surgery	Agreement between manual and automated measurements of MRD 1 and 2: ICCs; automated model reliability: ICCs	135 ptotic eyes in 103 patients; training size: 4.138 eyes in 2069 volunteers; test size: 135 eyes in 103 patients	China	-
Şimşek et al,[26] 2021	To evaluate postoperative changes using a computer vision algorithm for anterior full-face photographs of patients who underwent upper eyelid blepharoplasty surgery with or without Müller's muscle-conjunctival resection	Postoperative change in right eye palpebral distance and left eye palpebral distance; postoperative change in right eye-opening area, left eye-opening area ratio	55 patients	Turkey	-

Table 1. (cont'd)					
Study	**Study aims**	**Outcome measures**	**Sample size**	**Origin of images**	**Algorithm/ architecture**
Qu et al,[27] 2022	To explore the impact of a multichannel CNN-based eye model on eye plastic surgical repair and aesthetic effects	Reconstruction effects: similarity (overlapping volume), efficiency (required time); surgical results: ocular score, aesthetic score; complication rates	64 patients	China	-
Sun et al,[28] 2022	To automatically predict postoperative appearance after blepharoptosis surgery	Overall prediction performance: eye overlap ratio; local prediction performance: lid contour analysis by midpupil lid distances, absolute error between predicted and actual MRD1; agreement with clinicians: 5-point satisfaction scale, 10-point similarity survey	970 pairs of images of 450 eyes in 362 patients; training/validation size: 895; test size: 75	China	-
Thyroid eye disease					
Yoo et al,[31] 2020	To present a deep learning model for realistic prediction of postoperative appearance after orbital decompression surgery	Performance comparison between adversarial network models: similarity to postoperative images, image quality; TED O classification performance: accuracy, sensitivity, specificity, AUC; subjective quality assessment by ophthalmologists	109 pairs of pre- and post-operative images; training size: 76 pairs; validation size: 76 pairs and 500 synthesized pairs; test size: 33 pairs	-	-
Moon et al,[32] 2022	To develop a machine learning system that mimics an expert's CAS assessment using digital facial images, then evaluate its accuracy for predicting the CAS and diagnosing active TED	Diagnostic performance in identifying active TED: sensitivity, specificity, AUC; diagnostic performance in CAS prediction	1020 patients; training size: 920; test size: 100	Korea	-
Karlin et al,[33] 2022	To describe a deep learning classifier that identifies TED based on facial images	Accuracy, sensitivity, specificity	Training size: 1994 images; test size: 344 images	United States	-
Huang et al,[34] 2022	To establish an intelligent system for TED diagnosis based on facial images	Model ability to conduct binary classification (yes/no): AUC, sensitivity, specificity; accuracy of eye location, corneal and scleral segmentation: IoU	21 840 images of 3120 eyes in 1560 patients; training size: 70%; validation size: 10%; test size: 20%	China	-
Shao et al,[35] 2023	To propose a deep learning approach that automatically measures eyelid morphology in patients with TED	Accuracy of model: Dice coefficients, IoU; measurement agreement between automatic and manual measurements of MRDs: Bland-Altman plots, intraclass correlation coefficients	74 TED, 74 healthy eyes; training size: 18 000 facial images for eye detection, 3452 eye images for eye segmentation; validation size: 3000 facial images for eye detection, 272 eye images for eye segmentation; test size: 9000 and 148 facial images for eye detection, 148 eye images for eye segmentation	China	-
Eyelid tumors					
Adamopoulos et al,[37] 2021	To develop an intelligent system for pattern recognition, identification, and classification of eyelid basal cell carcinoma	Performance: accuracy score on classification	65 images of basal cell carcinoma, 78 images of healthy individuals; training size: 70%; validation size: 30%	Greece	-
Li et al,[38] 2022	To develop an AI system that automatically locates eyelid tumors and distinguishes between malignant and benign eyelid tumors	Diagnostic performance: accuracy, sensitivity, specificity, AUC; agreement with clinicians: unweighted Cohen's kappa coefficients	1533 images; training size: 883; validation size: 168; internal test size: 197; external test size: 285	China	-

Table 1. (cont'd)					
Study	**Study aims**	**Outcome measures**	**Sample size**	**Origin of images**	**Algorithm/ architecture**
Hui et al,[39] 2022	To develop a deep learning image analysis system for the automatic identification of benign and malignant eyelid tumors	Accuracy, sensitivity, specificity, AUC	309 images of 229 patients; development size: 309; validation size: 36	China	-
Lee et al,[40] 2023	To evaluate deep learning model performance, compared with that of human ophthalmologists, in differentiating eyelid lesions using clinical eyelid photographs	Diagnostic performance: accuracy, sensitivity, specificity, PPV, NPV, AUC	4954 images of 928 patients; training size: 4475; test size: 134	Korea	-
Keratitis					
Xu et al,[42] 2021	To propose a sequential-level deep model that effectively discriminates infectious corneal disease via classification of clinical images	Diagnostic performance: accuracy	2284 images of 867 patients; training size: 1922; test size: 362	China	-
Ji et al,[43] 2022	To propose a multi-task recognition method for automatic diagnosis of keratitis	Diagnostic performance: accuracy, precision, recall, specificity, F1 score	954 images; training size: 668; test size: 286	China	-
Kuo et al,[44] 2021	To compare the performances of various deep learning algorithms for diagnosing bacterial keratitis using external eye photographs	Diagnostic performance: sensitivity, specificity, PPV, NPV, AUC	1512 images of 1512 patients; 5 datasets of 185-186 images of bacteria keratitis and 116-117 images of non-bacterial keratitis; training size: 4 datasets; validation size: 1 dataset	Taiwan	-
Natarajan et al,[45] 2022	To use and study the application of AI-based deep learning algorithms for diagnosing viral keratitis	Diagnostic performance: sensitivity, specificity, AUC	307 images of 285 eyes; training size: 267; test size: 40	India	-
Hung et al,[46] 2021	To develop a deep learning model for identifying bacterial keratitis and fungal keratitis using slit-lamp images	Diagnostic performance: accuracy, AUC	1330 images of 580 patients; training size: 904; validation size: 212; test size: 214	Taiwan	-
Redd et al,[47] 2022	To develop computer vision models for image-based differentiation of bacterial and fungal corneal ulcers, then compare their performances with those of human experts	AUC	980 images of 980 patients; training size: 396; validation size: 50; test size: 100, 80	India	-
Zhang et al,[48] 2022	To construct a deep learning auxiliary diagnostic model for early diagnosis of infectious keratitis	Diagnostic performance: accuracy, AUC	4830 images; training size: 4347; test size: 483; external validation size: 200 images	China	-
Hu et al,[49] 2023	To propose a deep learning system based on slit-lamp images for automatic screening and diagnosis of infectious keratitis	Accuracy, recall, specificity, AUC	2757 images of 744 patients; training size: 1925; validation size: 301; test size: 531	China	VGG-16, ResNet34, Inception-v4, DenseNet121, ViT-Base, EfficientNetV2-M
Kogachi et al,[50] 2023	To determine whether CNNs can detect morphological differences between images of microbiologically positive and negative corneal ulcers	AUC	1970 images of 886 patients; training size: 1590; validation size: 200; test size: 211	India	-
Loo et al,[51] 2021	To propose a fully automatic deep learning algorithm for segmentation of ocular structures and microbial keratitis biomarkers on slit-lamp photography images	Performance: Dice similarity coefficient, Hausdorff distance	133 eyes; training size: 95; validation size: 19; test size: 19	United States, India	-
Loo et al,[52] 2021	To assess the clinical applicability of automatic image analysis in microbial keratitis by evaluating the relationship between BCVA and biomarker measurements on slit-lamp photographs	Performance of automatic segmentation measurement; correlation between measurements and BCVA	76 patients	United States, India	-

Table 1. (cont'd)					
Study	**Study aims**	**Outcome measures**	**Sample size**	**Origin of images**	**Algorithm/ architecture**
Trachoma					
Kim et al,[55] 2019	To test an automated algorithm in evaluating eyelid photographs for clinical signs of trachoma	Performance: accuracy, sensitivity, specificity; agreement between automated algorithm and expert graders: Cohen's kappa coefficients	1656 images; validation size: 100	Niger, Ethiopia	CNN with 3 stage convolutional layers, fully connected layers, and 2 hidden layers
Socia et al,[56] 2022	To evaluate the plausibility of an AI model for augmenting and replacing human image graders in the evaluation and diagnosis of TF	Diagnostic performance: sensitivity, specificity, PPV; agreement with clinicians: Cohen's kappa coefficient	2299 images; training/validation size: 44 TF; test size: 12 TF; external cohort: 1546; training/ validation size: 421 TF; test size: 106 TF	Tanzania	ResNet101, VGG-16
Pterygium					
Lopez et al,[60] 2019	To present a deep learning method for automatic classification of pterygium and non-pterygium images	Performance: AUC, sensitivity, specificity	3017 images (325 pterygium, 2692 non-pterygium)	Brazil	CNN
Zulkifley et al,[61] 2019	To propose a deep learning approach for automatic detection and localization of pterygium-infected tissues based on color images	Performance: accuracy, sensitivity, specificity	120 images (60 pterygium, 60 normal); training size: 80%; test size: 20%	Australia	Pterygium-Net based on CNN
Zhu et al,[62] 2022	To propose a two-category model and a segmentation model that assist with pterygium diagnosis based on anterior segment images	Performance: sensitivity, specificity, AUC, kappa value; comparison of model performance: mean IoU, IoU, mean average precision, and pixel accuracy	1034 images (517 pterygium, 517 normal); training size: 734 (367 pterygium); test size: 300 (150 pterygium)	China	AlexNet, VGG-16, ResNet18, ResNet50, PSPNet
Wan et al,[63] 2022	To propose a deep learning system for diagnosing and measuring the pathological progress of pterygium with anterior segment images	Performance: Dice coefficient, kappa consistency coefficient	489 images (245 pterygium, 244 normal); training size: 250; test size: 239	China	U-Net++
Fang et al,[64] 2022	To evaluate the performance of deep learning algorithms based on color anterior segment photographs from slit-lamp and hand-held cameras for detecting the presence and extent of pterygium	Detection: AUC, sensitivity, specificity	3132 images of 2932 eyes (1566 pterygium, 1566 non-pterygium); training size: 2503; internal test set size: 629; external test set size: 6311	Singapore	CNN with VGG-16, multilayer perceptron
Hung et al,[65] 2022	To evaluate the efficacy of a deep learning system based on slit-lamp photographs in pterygium grading and recurrence prediction	Grading: sensitivity, specificity, F1 score, accuracy; prediction: sensitivity, specificity, PPV, NPV	237 images of 237 eyes (176 pterygium, 61 non-pterygium); training size: 189 images; testing size: 48 images	Taiwan	DL network with U-Net, multilayer perceptron
Jais et al,[66] 2021	To propose a machine learning model for changes in BCVA after pterygium surgery according to morphological characteristics	Performance: accuracy, sensitivity, specificity	93 images; training size: 90%; test size: 10%	Malaysia	Support Vector Machine, Naive Bayes Classifier, decision tree, logistic regression
Liu et al,[67] 2023	To determine accuracy of a fusion training AI model in pterygium screening and detection using a smartphone	Detection: accuracy; segmentation: F1 score, sensitivity, specificity, AUC	20 987 slit-lamp images; 1094 smartphone images; training size: 15 456; validation size: 3080; test size: 6625	China	RFRC (Faster RCNN based on ResNet101), SRU-Net (U-Net based on SE-ResNet50)
Diabetic retinopathy					
Babenko et al,[68] 2022	To propose a deep learning model trained on external photographs of eyes for the detection of diabetic retinopathy, diabetic macular edema, and poor blood glucose control	Superiority of deep learning model predictions compared with baseline model predictions: AUC, DeLong method; sensitivity, specificity, PPV, NPV	Development size: training set 126 066 patients, tuning set 19 766 patients; validation size (4 sets): 27 415, 5058, 10 402, 6266 patients	United States	Inception-v3

Table 1. (cont'd)					
Study	**Study aims**	**Outcome measures**	**Sample size**	**Origin of images**	**Algorithm/ architecture**
Cataract					
Kiuchi et al,[70] 2022	To develop an AI system for preoperative safety management in cataract surgery based on facial recognition, laterality confirmation, and intraocular lens parameter verification	Accuracy and authentication rate of facial recognition on the first attempt; cumulative authentication rate after repeated attempts; false rejection rate, false acceptance rate	171 patients (162 patients in facial recognition test)	Japan	ISG-539
Strabismus					
Huang et al,[71] 2021	To provide an automatic strabismus screening method for people in remote areas with limited access to healthcare	Positional similarity estimates of normal and strabismus images (ratio)	60 images (30 strabismus, 30 normal)	Korea	-
Other oculofacial disorders					
Schulz et al,[72] 2023	To determine the feasibility, validity, reliability of clinically meaningful eyelid measurements automatically extracted from videos of individuals with oculofacial disorders	Feasibility, validity, reliability	77 facial nerve palsy, 33 ptosis, 33 TED, 65 controls; 7101 training images	-	Custom program VALID (video analysis of eyelids) based on U-Net
Greene et al,[73] 2019	To compare the sensitivity of a clinician-based tool (eFACE) to a well-established facial palsy intervention (eyelid weight placement) with the sensitivity of an automated facial measurement algorithm (Emotrics)	Palpebral fissure distance reduction	53 patients	United States	Emotrics

Abbreviations: AI=artificial intelligence, AUC=area under the curve, BCVA=best-corrected visual acuity, CAS=clinical activity score, CNN=convolutional neural networks, ICC=intraclass correlation coefficient, IoU=intersection over union, MBH=medial brow height, NPV=negative predictive value, PPV=positive predictive value, TED=thyroid eye disease, TF=trachomatous inflammation – follicular

Trachoma

Trachoma is an infection by ocular strains of *Chlamydia trachomatis*, which is readily treatable with azithromycin.[53,54] Clinical signs of trachoma often guide treatment decisions.[54] Kim et al[55] developed a CNN-based automated algorithm based on external eye images to assess clinical signs of trachoma: trachomatous inflammation–follicular (TF) and trachomatous inflammation–intense (TI). Maximum accuracy was 0.70 for TF and 0.85 for TI; agreements between AI and expert graders were $\varkappa$=0.44 and $\varkappa$=0.69 for TF and TI, respectively.[55] Socia et al[56] used an AI model to detect TF using external eye images and reported a maximum sensitivity of 95% and positive predictive value of 50% to 70%; the model reduced the burden of skilled image graders by 66% to 75%. AI can assist with clinical assessment of trachoma, reducing the burden on skilled image graders.

Pterygium

Pterygium, a common fibrovascular degeneration disease, is characterized by a wing-shaped conjunctival tissue growth that extends over the adjacent cornea.[57] It is readily treatable with surgery, and its recurrence risk is closely related to the size of the pterygium; thus, early identification and grading of pterygium are essential.[58,59] Lopez et al[60] developed a deep learning model for pterygium identification using external eye photographs in two color formats: RGB (ie, red, green, and blue) and grayscale. The model classified images as pterygium or non-pterygium with an AUC of 99.4%. Similarly, Zulkifley et al[61] constructed a deep learning model based on color photographs of the eye to detect pterygium, with 95% sensitivity and 98% sensitivity; the model simultaneously localized pterygium tissues with 81% accuracy. Zhu et al[62] developed a two-category and segmentation-based model of pterygium using anterior segment photographs; the model had 99.0% diagnostic accuracy, 98.7% sensitivity, and 99.3% specificity. In addition, the model can be used to stratify pterygium severity and risk. Wan et al[63] used deep learning techniques to measure pterygium severity in terms of the width of corneal invasion by pterygium and reported Dice coefficients of 0.902 to 0.962 and a kappa coefficient of 0.918. Fang et al[64] developed a deep learning algorithm based on color anterior segment photographs taken from slit-lamp and hand-held cameras to detect referable pterygium and reported AUCs of 98.5% to 99.5%. Hung et al[65] applied deep learning on slit-lamp photographs for pterygium grading and recurrence risk prediction and reported 86.7% to 91.7% accuracy for grading and 81.8% specificity and 66.7% sensitivity for prediction. Jais et al[66]

Table 2. Risk of bias assessment of included studies							
Study	**Risk of bias**				**Applicability concerns**		
	Patient selection	**Index test**	**Reference standard**	**Flow and timing**	**Patient selection**	**Index test**	**Reference standard**
Measurement of eyelid/periorbital parameters							
Moriyama et al,[16] 2006	Low	Low	Unclear	Unclear	Low	Low	Unclear
Thomas et al,[17] 2020	Unclear	Low	Unclear	Low	Unclear	Unclear	Unclear
Van Brummen et al,[18] 2021	Low	Low	High	Low	Unclear	Low	High
Chen et al,[19] 2021	Low	Low	Low	Low	Unclear	Low	Low
Cao et al,[20] 2021	Low	Low	Low	Low	High	Low	Low
Zhu et al,[21] 2022	Unclear	Low	Low	Low	Unclear	Low	Low
Blepharoptosis							
Hung et al,[22] 2021	Low	Low	Low	Low	High	Low	Low
Tabuchi et al,[23] 2022	Unclear	Low	Low	Low	High	Low	Low
Hung et al,[24] 2022	Low	Low	Low	Low	Unclear	Low	Low
Lou et al,[25] 2021	Low	Low	Low	Low	Unclear	Unclear	Low
Şimşek et al,[26] 2021	Low	Low	Low	Low	Low	Low	Low
Qu et al,[27] 2022	Low	Low	Unclear	Unclear	Unclear	Low	High
Sun et al,[28] 2022	Unclear	Unclear	High	Low	High	High	High
Thyroid eye disease							
Yoo et al,[31] 2020	Unclear	Low	Unclear	Low	Unclear	Low	Unclear
Moon et al,[32] 2022	Low	Low	Low	Unclear	Low	Low	Low
Karlin et al,[33] 2022	Low	Low	Low	High	Unclear	Low	Low
Huang et al,[34] 2022	Low	Low	Low	Low	Low	Unclear	Low
Shao et al,[35] 2023	Low	Low	Low	Low	Low	Low	Low
Eyelid tumors							
Adamopoulos et al,[37] 2021	Unclear	Low	High	High	Unclear	Low	Unclear
Li et al,[38] 2022	Low	Low	Low	High	High	Unclear	Low
Hui et al,[39] 2022	Low	Low	Low	High	High	Low	Low
Lee et al,[40] 2023	Low	Low	Low	Low	Unclear	Low	Low
Keratitis							
Xu et al,[42] 2021	Low	Low	Low	High	Unclear	Low	Low
Ji et al,[43] 2022	Low	Low	Low	Low	Unclear	Low	Low
Kuo et al,[44] 2021	Low	Low	High	High	Unclear	Low	Unclear
Natarajan et al,[45] 2022	Low	Low	Low	Low	Unclear	Low	Low
Hung et al,[46] 2021	Unclear	Low	Low	Low	Unclear	Unclear	Low
Redd et al,[47] 2022	Unclear	Low	Low	Unclear	Unclear	Low	Unclear
Zhang et al,[48] 2022	Unclear	Low	Low	Unclear	Unclear	Low	Low
Hu et al,[49] 2023	Low	Low	Low	Low	Unclear	Unclear	Low
Kogachi et al,[50] 2023	Low	Low	Unclear	Low	Unclear	Low	Unclear
Loo et al,[51] 2021	Low	Low	Unclear	Unclear	Low	Low	Low
Loo et al,[52] 2021	Low	Low	Low	Unclear	Low	Low	Low

Table 2. (cont'd)							
Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Trachoma							
Kim et al,[55] 2019	Low	Low	Low	Low	Low	Low	Low
Socia et al,[56] 2022	Unclear	Low	Low	Unclear	Unclear	Low	Low
Pterygium							
Lopez et al,[60] 2019	Unclear	Unclear	Low	Low	High	High	Low
Zulkifley et al,[61] 2019	Unclear	Low	Low	Unclear	High	Low	Low
Zhu et al,[62] 2022	Unclear	Low	Low	Unclear	Unclear	Low	Low
Wan et al,[63] 2022	Unclear	Low	Low	Unclear	Unclear	Low	Low
Fang et al,[64] 2022	Low	Low	Low	Low	Low	Low	Low
Hung et al,[65] 2022	Unclear	Low	Low	Low	Low	Unclear	Low
Jais et al,[66] 2021	Unclear	Low	High	Unclear	Unclear	Low	High
Liu et al,[67] 2023	Unclear	Low	Low	Low	Low	Low	Low
Diabetic retinopathy							
Babenko et al,[68] 2022	Unclear	High	Low	Low	Unclear	High	Low
Cataract							
Kiuchi et al,[70] 2022	High	Low	Unclear	Unclear	High	Unclear	Unclear
Strabismus							
Huang et al,[71] 2021	Unclear	Low	High	High	Unclear	Low	High
Other oculofacial disorders							
Schulz et al,[72] 2023	Low	Low	High	High	High	Low	High
Greene et al,[73] 2019	Low	Low	High	Low	Unclear	Low	Unclear

evaluated the performance of four machine learning models in postoperative classification of changes in BCVA according to pterygium characteristics on external photographs; the best-performing model (ie, Support Vector Machine) achieved 94.4% accuracy, 100% specificity, and 92.1% sensitivity. Liu et al[67] trained an AI fusion model using slit-lamp and smartphone images to detect and grade pterygium and reported 92.38% accuracy (AUC=0.94) with smartphone-based images and 94.29% accuracy (AUC=0.96) with slit-lamp images. These findings support the use of AI systems based on camera and slit-lamp-based images for pterygium screening and severity grading.

Diabetic retinopathy

Diabetic retinopathy screening usually requires a fundus examination or fundus photographs. Babenko et al[68] used an AI system that assessed pupil size, conjunctival vessel caliber, and tortuosity based on external photographs of the eyes to detect diabetic retinopathy, diabetic macular edema, and poor blood glucose control; its predictive performance was better than that of logistic regression models based on demographic and medical history data. Predictions could be generalized to patients in various diabetic retinopathy screening programs. Their AI model achieved superior prediction, compared with the baseline model.

Cataract

In ophthalmic surgery, common types of surgical confusion include implantation of the wrong intraocular lens, operation on the wrong eye, treatment of the wrong patient, and performance of the wrong procedure.[69] Kiuchi et al[70] developed an AI system for preoperative safety management in cataract surgery, utilizing external photographs to confirm patient identity. Authentication rates during initial and subsequent attempts were 92.0% and 96.3%, respectively. When combined with laterality confirmation and intraocular lens parameter verification by AI, the false rejection rate and false acceptance rate were both 0%. Implementation of AI in ophthalmic surgery can improve preoperative safety management and reduce preventable instances of surgical confusion.

Strabismus

Strabismus is a condition in which the eyes are not properly aligned with each other; this misalignment can cause double vision, loss of stereopsis, and suppression. Huang et al[71] trained an AI system with frontal facial images to calculate the distances from the pupil center to the medial and lateral canthi and measure deviation in the positional similarity of the two eyes for strabismus screening. The results were promising that the system could serve as a strabismus screening tool for patients living in remote areas.

Other oculofacial disorders

Schulz et al[72] measured eyelid parameters (MRD 1 and 2, blink lagophthalmos, and ocular surface area exposure) on facial videos in patients with facial nerve palsy or TED and reported good reliability and validity. Greene et al[73] used an automated facial measurement algorithm (Emotrics) for eyelid parameters in patients with unilateral facial palsy who received an eyelid weight; the agreement between measurements using the clinician-graded facial function scale (eFACE) and Emotrics measurements was good.

Discussion

This systematic review showed that AI using external photographs exhibits high accuracy in measuring periorbital parameters and in detecting a wide range of ophthalmic diseases such as TED, blepharoptosis, eyelid tumors, keratitis, trachoma, pterygium, strabismus, and diabetic retinopathy. AI can grade the severity of diseases such as TED, trachoma, and pterygium and can differentiate various eyelid tumor and keratitis types based on external photographs. Overall, external photographs provide a rich source of information for AI algorithms, which can be applied to a wide range of ophthalmic conditions.

AI has an increasing role in ophthalmology, mainly due to advances in multimodal imaging. Deep learning systems can identify vision-threatening diabetic retinopathy and age-related macular degeneration using fundus photographs and OCT images, thereby preventing treatment delays.[74,75] Diagnosing retinopathy of prematurity is challenging for ophthalmology trainees; AI can distinguish the disease using fundus photographs, thereby improving diagnostic accuracy and avoiding treatment delays.[76] AI can predict disease severity and determine disease progression in glaucoma through analyses of the visual field and the optic nerve structure's retinal nerve fiber on OCT or fundus photographs.[77,78] In recent years, external photographs have been increasingly used in ophthalmology; external photographs have great potential to serve as reliable screening tools, thereby reducing waiting time, enhancing management, and improving outcomes.

Deep-learning models mimic the diagnostic approaches of ophthalmologists, and classifiers can associate features of ophthalmic diseases at periorbital landmarks within image datasets for disease diagnosis. For example, eyelid signs are the most common presenting sign in TED. AI can detect these signs using external photographs based on eyelid parameters that correspond to areas emphasized by ophthalmologists. Although TED diagnosis is usually confirmed through clinical examination, a survey study in the United Kingdom revealed that the initial diagnosis of TED was incorrect in 58% of patients, and 25% of patients had a 1-year delay in diagnosis.[79] Deep learning systems reported thus far have demonstrated promising performance in identifying new cases of TED, enabling the prioritization of early-onset and urgent-care cases.

AI screening tools can reduce social inequality by facilitating access to healthcare services. Patients with low socioeconomic status and low health literacy may have difficulties accessing healthcare services. The use of AI in disease detection and severity grading may help prioritize patients who need early eye care to attend specialty clinics, especially those living in rural regions. A reliable screening tool enables timely detection and reduces morbidity, especially in cases of TED and eyelid tumors as well as other sight-threatening but readily treatable conditions such as keratitis, trachoma, and pterygium. Initial screening of photographs taken by smartphone can be conducted without sophisticated equipment, thereby enhancing convenience and accessibility of healthcare services. Additionally, AI can enhance the personalization of treatment plans. For example, patients undergoing blepharoptosis correction may receive AI prediction of their postoperative appearance; this can be an adjunctive tool for preoperative assessment in clinical settings. Taken together, deep learning systems offer a positive outlook regarding widespread clinical application of AI.

Despite encouraging results in the use of AI to detect ophthalmic conditions based on external photographs, several challenges must be addressed before its implementation in real-world settings. Major foreseeable issues include standardization, validation, applicability, and ethical concerns.[80] Standardization remains the greatest challenge because external photographs are highly operator-dependent and easily affected by lighting conditions[35]; this is compounded by the subjectivity of image quality and the technical demands of AI programs. Standardized protocols for photograph capture and analysis, as well as objective thresholds for quality control, should be established to reduce inter-operator variability and ensure consistency.[81-83] Further training of models with images of variable quality such as those captured at different angles and under different lighting conditions is recommended.[28,33] Furthermore, most studies have targeted Asian populations, and there is a lack of studies targeting Black or Hispanic populations. This may decrease generalizability in terms of applicability and feasibility within these populations. Large-scale multi-country collaborations are needed to ensure that AI algorithms are reliable across diverse settings in real-world environments.[23,24,38,40,47,64] Furthermore, ethical issues related to the use of facial photographs are expected. Compared with OCT and fundus photographs, external photographs are more recognizable and identifiable, leading to privacy concerns. Strict compliance with regulations on data collection, storage, and use transparency is essential.

To enhance clinical relevance of AI models as screening tools, training AI models with larger numbers of early-stage lesions can improve their sensitivity to more subtle disease processes.[32,40] Incorporation of clinical information (such as chief complaints and medical histories) into AI models may better mimic the clinical-reasoning process during identification and classification of ophthalmic diseases.[34,39,67] To expand the use of AI beyond clinical settings, incorporation of smartphones and cameras as less-specialized initial screening tools is suggested.[39,65,67] The fusion training model that combines high-quality slit-lamp images with external photographs enables high-accuracy detection and grading for initial screening by smartphone.[67] Future research may focus on the use of AI in community settings and beyond.

Overall, AI has demonstrated promising performance in the detection and classification of various ophthalmic conditions based on external photographs; it can serve as an inexpensive and convenient tool for patients to access healthcare services. Further studies are needed to confirm the role of deep learning methods in improving diagnostic accuracy and clinical applicability.

Contributors

All authors designed the study, acquired the data, analyzed the data, drafted the manuscript, and critically revised the manuscript for important intellectual content. All authors had full access to the data, contributed to the study, approved the final version for publication, and take responsibility for its accuracy and integrity.

Conflicts of interest

authors have disclosed no conflicts of interest.

Funding/support

This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

All data analyzed in this study are available from the corresponding author upon reasonable request.

References

1. *Shen Y, Xiao T, Yi S, Chen D, Wang X, Li H. Person re-identification with deep Kronecker-product matching and group-shuffling random walk. IEEE Trans Pattern Anal Mach Intell 2021;43:1649-65.*
2. *Huang Z, Yu Y, Xu J, Ni F, Le X. PF-Net: point fractal network for 3D point cloud completion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2020: 7662-70.*
3. *Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. Ophthalmology 2017;124:962-9.*
4. *Yousefi S, Elze T, Pasquale LR, et al. Monitoring glaucomatous functional loss using an artificial intelligence-enabled dashboard. Ophthalmology 2020;127:1170-8.*
5. *Schmidt-Erfurth U, Vogl WD, Jampol LM, Bogunović H. Application of automated quantification of fluid volumes to anti-VEGF therapy of neovascular age-related macular degeneration. Ophthalmology 2020;127:1211-9.*
6. *Du R, Xie S, Fang Y, et al. Deep learning approach for automated detection of myopic maculopathy and pathologic myopia in fundus images. Ophthalmol Retina 2021;5:1235-44.*
7. *Roomaney I, Nyirenda C, Chetty M. Facial imaging to screen for fetal alcohol spectrum disorder: a scoping review. Alcohol Clin Exp Res 2022;46:1166-80.*
8. *Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. Genome Med 2019;11:70.*
9. *Qin B, Liang L, Wu J, Quan Q, Wang Z, Li D. Automatic identification of Down syndrome using facial images with deep convolutional neural network. Diagnostics (Basel) 2020;10:487.*
10. *Mukhiddinov M, Djuraev O, Akhmedov F, Mukhamadiyev A, Cho J. Masked face emotion recognition based on facial landmarks and deep learning approaches for visually impaired people. Sensors (Basel) 2023;23:1080.*
11. *Debnath T, Reza MM, Rahman A, Beheshti A, Band SS, Alinejad-Rokny H. Four-layer ConvNet to facial emotion recognition with minimal epochs and the significance of data diversity. Sci Rep 2022;12:6991.*
12. *Kizilgul M, Karakis R, Dogan N, et al. Real-time detection of acromegaly from facial images with artificial intelligence. Eur J Endocrinol 2023;188:lvad005.*
13. *Goetz RK, Hughes FE, Duignan ES, et al. A template for reducing ophthalmology outpatient waiting times: community ophthalmic care. Ir J Med Sci 2018;187:237-41.*
14. *Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.*
15. *Starck WJ, Griffin JE Jr, Epker BN. Objective evaluation of the eyelids and eyebrows after blepharoplasty. J Oral Maxillofac Surg 1996;54:297-303.*
16. *Moriyama T, Kanade T, Xiao J, Cohn JF. Meticulously detailed eye region model and its application to analysis of facial images. IEEE Trans Pattern Anal Mach Intell 2006;28:738-52.*
17. *Thomas PBM, Gunasekera CD, Kang S, Baltrusaitis T. An artificial intelligence approach to the assessment of abnormal lid position. Plast Reconstr Surg Glob Open 2020;8:e3089.*
18. *Van Brummen A, Owen JP, Spaide T, et al. PeriorbitAI: artificial intelligence automation of eyelid and periorbital measurements. Am J Ophthalmol 2021;230:285-96.*
19. *Chen HC, Tzeng SS, Hsiao YC, Chen RF, Hung EC, Lee OK. Smartphone-based artificial intelligence-assisted prediction for eyelid measurements: algorithm development and*

observational validation study. JMIR Mhealth Uhealth 2021;9:e32444.

20. Cao J, Lou L, You K, et al. A novel automatic morphologic analysis of eyelids based on deep learning methods. Curr Eye Res 2021;46:1495-502.
21. Zhu X, Song X, Min X, et al. Calculation of ophthalmic diagnostic parameters on a single eye image based on deep neural network. Multimed Tools Appl 2022;81:2311-31.
22. Hung JY, Perera C, Chen KW, et al. A deep learning approach to identify blepharoptosis by convolutional neural networks. Int J Med Inform 2021;148:104402.
23. Tabuchi H, Nagasato D, Masumoto H, et al. Developing an iOS application that uses machine learning for the automated diagnosis of blepharoptosis. Graefes Arch Clin Exp Ophthalmol 2022;260:1329-35.
24. Hung JY, Chen KW, Perera C, et al. An outperforming artificial intelligence model to identify referable blepharoptosis for general practitioners. J Pers Med 2022;12:283.
25. Lou L, Cao J, Wang Y, et al. Deep learning-based image analysis for automated measurement of eyelid morphology before and after blepharoptosis surgery. Ann Med 2021;53:2278-85.
26. Bahçeci Şimşek İ, Şirolu C. Analysis of surgical outcome after upper eyelid surgery by computer vision algorithm using face and facial landmark detection. Graefes Arch Clin Exp Ophthalmol 2021;259:3119-25.
27. Qu Y, Lin B, Li S, et al. Effect of multichannel convolutional neural network-based model on the repair and aesthetic effect of eye plastic surgery patients. Comput Math Methods Med 2022;2022:5315146.
28. Sun Y, Huang X, Zhang Q, et al. A fully automatic postoperative appearance prediction system for blepharoptosis surgery with image-based deep learning. Ophthalmol Sci 2022;2:100169.
29. Bahn RS, Heufelder AE. Pathogenesis of Graves' ophthalmopathy. N Engl J Med 1993;329:1468-75.
30. Bartalena L, Kahaly GJ, Baldeschi L, et al. The 2021 European Group on Graves' orbitopathy (EUGOGO) clinical practice guidelines for the medical management of Graves' orbitopathy. Eur J Endocrinol 2021;185:G43-G67.
31. Yoo TK, Choi JY, Kim HK. A generative adversarial network approach to predicting postoperative appearance after orbital decompression surgery for thyroid eye disease. Comput Biol Med 2020;118:103628.
32. Moon JH, Shin K, Lee GM, et al. Machine learning-assisted system using digital facial images to predict the clinical activity score in thyroid-associated orbitopathy. Sci Rep 2022;12:22085.
33. Karlin J, Gai L, LaPierre N, et al. Ensemble neural network model for detecting thyroid eye disease using external photographs. Br J Ophthalmol 2023;107:1722-9.
34. Huang X, Ju L, Li J, et al. An intelligent diagnostic system for thyroid-associated ophthalmopathy based on facial images. Front Med (Lausanne) 2022;9:920716.
35. Shao J, Huang X, Gao T, et al. Deep learning-based image analysis of eyelid morphology in thyroid-associated ophthalmopathy. Quant Imaging Med Surg 2023;13:1592-604.
36. Deprez M, Uffer S. Clinicopathological features of eyelid skin tumors. A retrospective study of 5504 cases and review of literature. Am J Dermatopathol 2009;31:256-62.
37. Adamopoulos A, Chatzopoulos EG, Anastassopoulos G, Detorakis E. Eyelid basal cell carcinoma classification using shallow and deep learning artificial neural networks. Evol Syst (Berl) 2021;12:583-90.
38. Li Z, Qiang W, Chen H, et al. Artificial intelligence to detect malignant eyelid tumors from photographic images. NPJ Digit Med 2022;5:23.
39. Hui S, Dong L, Zhang K, Nie Z. Noninvasive identification of benign and malignant eyelid tumors using clinical images via deep learning system. J Big Data 2022;9:84.
40. Lee MJ, Yang MK, Khwarg SI, et al. Differentiating malignant and benign eyelid lesions using deep learning. Sci Rep 2023;13:4103.
41. Austin A, Lietman T, Rose-Nussbaumer J. Update on the management of infectious keratitis. Ophthalmology 2017;124:1678-89.
42. Xu Y, Kong M, Xie W, et al. Deep sequential feature learning in clinical image classification of infectious keratitis. Engineering 2021;7:1002-10.
43. Ji Q, Jiang Y, Qu L, Yang Q, Zhang H. An image diagnosis algorithm for keratitis based on deep learning. Neural Process Lett 2022;54:2007-24.
44. Kuo MT, Hsu BW, Lin YS, et al. Comparisons of deep learning algorithms for diagnosing bacterial keratitis via external eye photographs. Sci Rep 2021;11:24227.
45. Natarajan R, Matai HD, Raman S, et al. Advances in the diagnosis of herpes simplex stromal necrotising keratitis: a feasibility study on deep learning approach. Indian J Ophthalmol 2022;70:3279-83.
46. Hung N, Shih AK, Lin C, et al. Using slit-lamp images for deep learning-based identification of bacterial and fungal keratitis: model development and validation with different convolutional neural networks. Diagnostics (Basel) 2021;11:1246.
47. Redd TK, Prajna NV, Srinivasan M, et al. Image-based differentiation of bacterial and fungal keratitis using deep convolutional neural networks. Ophthalmol Sci 2022;2:100119.
48. Zhang Z, Wang H, Wang S, et al. Deep learning-based classification of infectious keratitis on slit-lamp images. Ther Adv Chronic Dis 2022;13:20406223221136071.
49. Hu S, Sun Y, Li J, et al. Automatic diagnosis of infectious keratitis based on slit lamp images analysis. J Pers Med 2023;13:519.
50. Kogachi K, Lalitha P, Prajna NV, et al. Deep convolutional neural networks detect no morphological differences between culture-positive and culture-negative infectious keratitis images. Transl Vis Sci Technol 2023;12:12.
51. Loo J, Kriegel MF, Tuohy MM, et al. Open-source automatic segmentation of ocular structures and biomarkers of microbial keratitis on slit-lamp photography images using deep learning. IEEE J Biomed Health Inform 2021;25:88-99.
52. Loo J, Woodward MA, Prajna V, et al. Open-source automatic biomarker measurement on slit-lamp photography to estimate visual acuity in microbial keratitis. Transl Vis Sci Technol 2021;10:2.
53. Burton MJ, Mabey DC. The global burden of trachoma: a review. PLoS Negl Trop Dis 2009;3:e460.
54. World Health Organization. Trachoma control. A guide for programme managers. Accessed 19 February 2024. Available from: http://apps.who.int/iris/bitstream/10665/43405/1/9241546905_eng.pdf.
55. Kim MC, Okada K, Ryner AM, et al. Sensitivity and specificity of computer vision classification of eyelid photographs for programmatic trachoma assessment. PLoS One 2019;14:e0210463.
56. Socia D, Brady CJ, West SK, Cockrell RC. Detection of trachoma using machine learning approaches. PLoS Negl Trop Dis 2022;16:e0010943.
57. Chu WK, Choi HL, Bhat AK, Jhanji V. Pterygium: new insights. Eye (Lond) 2020;34:1047-50.
58. Nejima R, Masuda A, Minami K, Mori Y, Hasegawa Y,

Miyata K. Topographic changes after excision surgery of primary pterygia and the effect of pterygium size on topographic restoration. Eye Contact Lens 2015;41:58-63.
59. Aidenloo NS, Motarjemizadeh Q, Heidarpanah M. Risk factors for pterygium recurrence after limbal-conjunctival autografting: a retrospective, single-centre investigation. Jpn J Ophthalmol 2018;62:349-56.
60. López YP, Aguilera LR. Automatic classification of pterygium-non pterygium images using deep learning. In: VipIMAGE 2019: 391-400.
61. Zulkifley MA, Abdani SR, Zulkifley NH. Pterygium-Net: a deep learning approach to pterygium detection and localization. Multimed Tools Appl 2019;78:34563-84.
62. Zhu S, Fang X, Qian Y, et al. Pterygium screening and lesion area segmentation based on deep learning. J Healthc Eng 2022;2022:3942110.
63. Wan C, Shao Y, Wang C, Jing J, Yang W. A novel system for measuring pterygium's progress using deep learning. Front Med (Lausanne) 2022;9:819971.
64. Fang X, Deshmukh M, Chee ML, et al. Deep learning algorithms for automatic detection of pterygium using anterior segment photographs from slit-lamp and hand-held cameras. Br J Ophthalmol 2022;106:1642-7.
65. Hung KH, Lin C, Roan J, et al. Application of a deep learning system in pterygium grading and further prediction of recurrence with slit lamp photographs. Diagnostics (Basel) 2022;12:888.
66. Jais FN, Che Azemin MZ, Hilmi MR, Mohd Tamrin MI, Kamal KM. Postsurgery classification of best-corrected visual acuity changes based on pterygium characteristics using the machine learning technique. ScientificWorldJournal 2021;2021:6211006.
67. Liu Y, Xu C, Wang S, et al. Accurate detection and grading of pterygium through smartphone by a fusion training model. Br J Ophthalmol 2023;108:336-42.
68. Babenko B, Mitani A, Traynis I, et al. Detection of signs of disease in external photographs of the eyes via deep learning. Nat Biomed Eng 2022;6:1370-83.
69. Parikh R, Palmer V, Kumar A, Simon JW. Surgical confusions in ophthalmology: description, analysis, and prevention of errors from 2006 through 2017. Ophthalmology 2020;127:296-302.
70. Kiuchi G, Tanabe M, Nagata K, Ishitobi N, Tabuchi H, Oshika T. Deep learning-based system for preoperative safety management in cataract surgery. J Clin Med 2022;11:5397.
71. Huang X, Lee SJ, Kim CZ, Choi SH. An automatic screening method for strabismus detection based on image processing. PLoS One 2021;16:e0255643.
72. Schulz CB, Clarke H, Makuloluwe S, Thomas PB, Kang S. Automated extraction of clinical measures from videos of oculofacial disorders using machine learning: feasibility, validity and reliability. Eye (Lond) 2023;37:2810-6.
73. Greene JJ, Tavares J, Guarin DL, Hadlock T. Clinician and automated assessments of facial function following eyelid weight placement. JAMA Facial Plast Surg 2019;21:387-92.
74. Bellemo V, Lim ZW, Lim G, et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. Lancet Digit Health 2019;1:e35-e44.
75. Dong L, Yang Q, Zhang RH, Wei WB. Artificial intelligence for the detection of age-related macular degeneration in color fundus photographs: a systematic review and meta-analysis. EClinicalMedicine 2021;35:100875.
76. Scruggs BA, Chan RVP, Kalpathy-Cramer J, Chiang MF, Campbell JP. Artificial intelligence in retinopathy of prematurity diagnosis. Transl Vis Sci Technol 2020;9:5.
77. Bojikian KD, Lee CS, Lee AY. Finding glaucoma in color fundus photographs using deep learning. JAMA Ophthalmol 2019;137:1361-2.
78. Zheng C, Johnson TV, Garg A, Boland MV. Artificial intelligence in glaucoma. Curr Opin Ophthalmol 2019;30:97-103.
79. Lim NC, Sundar G, Amrith S, Lee KO. Thyroid eye disease: a Southeast Asian experience. Br J Ophthalmol 2015;99:512-8.
80. He M, Li Z, Liu C, Shi D, Tan Z. Deployment of artificial intelligence in real-world practice: opportunity and challenge. Asia Pac J Ophthalmol (Phila) 2020;9:299-307.
81. Lee JG, Jun S, Cho YW, et al. Deep learning in medical imaging: general overview. Korean J Radiol 2017;18:570-84.
82. He J, Baxter SL, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019;25:30-6.
83. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019;380:1347-58.