Application of artificial intelligence in anterior segment ophthalmic diseases: diversity and standardization

Xiaohang Wu¹#, Lixue Liu¹#, Lanqin Zhao¹, Chong Guo¹, Ruiyang Li¹, Ting Wang¹, Xiaonan Yang¹, Peichen Xie², Yizhi Liu¹, Haotian Lin¹,³

¹State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, ²Zhongshan School of Medicine, ³Center for Precision Medicine, Sun Yat-sen University, Guangzhou 510080, China

Contributions: (I) Conception and design: H Lin, X Wu; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: None; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work.

Correspondence to: Haotian Lin, Yizhi Liu. Zhongshan Ophthalmic Center, Sun Yat-sen University, 54# Xianlie Road, Guangzhou 510060, China. Email: haot.lin@hotmail.com; yizhi_liu@aliyun.com.

Abstract: Artificial intelligence (AI) based on machine learning (ML) and deep learning (DL) techniques has gained tremendous global interest in this era. Recent studies have demonstrated the potential of AI systems to provide improved capability in various tasks, especially in image recognition field. As an image-centric subspecialty, ophthalmology has become one of the frontiers of AI research. Trained on optical coherence tomography, slit-lamp images and even ordinary eye images, AI can achieve robust performance in the detection of glaucoma, corneal arcus and cataracts. Moreover, AI models based on other forms of data also performed satisfactorily. Nevertheless, several challenges with AI application in ophthalmology have also arisen, including standardization of data sets, validation and applicability of AI models, and ethical issues. In this review, we provided a summary of the state-of-the-art AI application in anterior segment ophthalmic diseases, potential challenges in clinical implementation and our prospects.

Keywords: Artificial intelligence (AI); anterior eye segment; computer-assisted diagnosis; machine learning (ML)

Submitted Jan 21, 2020. Accepted for publication Apr 17, 2020.

doi: 10.21037/atm-20-976

View this article at: http://dx.doi.org/10.21037/atm-20-976

Introduction

First conceptualized in 1956 by Dartmouth scholar John McCarthy, artificial intelligence (AI) is a general term referring to hardware or software that exhibits intelligent behavior (1). As a subset of AI, machine learning (ML) was proposed subsequently in 1959 and stated that “the computer should have the ability to learn using various statistical techniques, without being explicitly programmed” (2). ML has been widely adopted in applications such as computer vision and predictive analytics using complex mathematical models. Deep learning (DL) is a class of state-of-the-art machine learning techniques that has sparked tremendous global interest in recent years (3).

With improved performance, the deployment of DL focuses on image recognition, speech recognition and natural language processing (4,5).

Healthcare has become one of the frontiers of AI application in recent years, particularly for those image-centric subspecialties such as radiology, dermatology, pathology and ophthalmology (6). For example, AI algorithms have proved effective in the detection of pulmonary tuberculosis from chest radiographs and melanoma on digital skin photographs (7,8). Similarly, in ophthalmology, AI has been primarily applied to image-based diagnosis of diabetic retinopathy, glaucoma, age-related macular degeneration and retinopathy of
prematurity, most of which are retinal diseases (9).

Unlike the diagnosis of retinal diseases, which is largely dependent on ophthalmoscopy, multiple examinations are required to diagnose anterior segment diseases given the complexity of its structure and physiological functions. Anterior eye segment is defined as the front third of the eye that includes conjunctiva, cornea, anterior chamber, iris, pupil, ciliary body, and lens. These structures constitute the path of light through the eye and the ocular refractive system. To establish a definite diagnosis of anterior segment eye diseases, evaluation of both anatomy and function are necessitated, including slit-lamp biomicroscopy, optical coherence tomography, corneal topography, tonometry, perimetry, etc. Accordingly, various forms of data besides images, such as videos, formatted parameters and texts, have been used in the AI-assisted detection of these diseases.

This review summarizes the application of AI systems in anterior segment ophthalmic diseases, potential challenges in clinical adoption and possible paths forward.

**Constructing models**

Multiple ophthalmic imaging modalities, such as slit-lamp images (10-17), AS-OCT images (18-23) and tear film lipid layer interference images (24-27), have been widely used in AI for the diagnosis of anterior segment ocular diseases. In addition, structured data, such as corneal parameters (28-33), have been used as well. The ophthalmic imaging modalities used in diagnosis are illustrated in Figure 1.

The steps for building an AI model include the following phases: data preparation, data partition, model optimization and evaluation of the system, as showed in Figure 2.

**Data preparation**

In order to improve the performance of models or meet

<table>
<thead>
<tr>
<th>Imaging modalities</th>
<th>Image features</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slit light</td>
<td></td>
<td>Cataract detection and grading</td>
</tr>
<tr>
<td>Slit-lamp images</td>
<td></td>
<td>Conjunctival hyperemia grading</td>
</tr>
<tr>
<td>Diffuse light</td>
<td></td>
<td>Cataract detection and grading</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High intraocular pressure detection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identification, localization of conjunctiva, cornea and lens diseases</td>
</tr>
<tr>
<td>Retro-illumination</td>
<td></td>
<td>Posterior capsule opacification progress prediction</td>
</tr>
<tr>
<td>AS-OCT images</td>
<td></td>
<td>Angle-closure detection</td>
</tr>
<tr>
<td>Tear film lipid layer interference images</td>
<td></td>
<td>Tear Film Lipid Layer Classification</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dry eye diagnosis</td>
</tr>
<tr>
<td>Frontal eye videos after fluorescein instillation</td>
<td></td>
<td>Dry eye detection, BUT calculation</td>
</tr>
<tr>
<td>Retinal fundus images</td>
<td></td>
<td>Cataract grading</td>
</tr>
<tr>
<td>Visible wavelength image by digital camera</td>
<td></td>
<td>Corneal arcus and cataract detection</td>
</tr>
</tbody>
</table>

**Figure 1** The ophthalmic imaging modalities in AI applications in anterior segment ocular diseases.
Figure 2 The workflow of constructing models.
Threshold determined with the analysis of ROC curve (31), decisions, such as the optimal classification or operating threshold determined with the analysis of ROC curve (31), are also made based on the performance on validation dataset.

**Evaluating the model performance**

Model evaluation aims to estimate the performance of the final model on future data. The most commonly used metrics in AI for disease diagnosis are AUC [the area under the receiver’s operator characteristics (ROC curve)], sensitivity, and specificity. AUC is more commonly used than accuracy because it isn’t affected by the classification threshold or the proportion of positive examples as accuracy is. AUC ranges from 0.5 (for a model with no predictive value) to 1 (for a perfect model), and a value greater than 0.9 indicates the model is outstanding, 0.8–0.9 indicates the model is good or excellent (38). ROC curve is graphical display of sensitivity (true positive rate, TPR) on y-axis and (1 – specificity) (false positive rate, FPR) on x-axis for varying cut-off points of probabilities predicted by the model. Except for examining the metrics on the testing dataset, comparing the performance of the model with human experts or other systems are also common, and ROC curve has been usually used to visually demonstrate the comparison of several test results simultaneously.

**AI Algorithms**

Machine Learning is a paradigm that allows computers to learn patterns from a large corpus of data, and use the learned knowledge to make prediction. It can be roughly divided into traditional machine learning algorithms and deep learning algorithms. Traditional machine learning algorithms use variables selected by experts as input, and usually does not involve large neural networks. They include algorithms such as Linear Regression, Logistic Regression, Support Vector Machine, Decision Tree, Random Forest and etc. Deep learning algorithms make use of multimedia datasets, such as images, videos and sound, and usually involves the use of large-scale neural networks such as Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). While conventional machine learning methods are still employed, most recent researches focus on using deep neural networks.

Deep learning tasks mainly include three use cases: classification, object detection and semantic segmentation. In classification tasks, an algorithm learns to classify an image into a set of predefined categories. VGG, ResNet,
Inception and InceptionResNet are some of the popular CNNs used for classification.

The task of object detection focuses not only on what is in the image, but also on the position of the object(s) of interest. Popular object detection algorithms include Faster-RCNN, You Only Look Once (YOLO) and Single Shot Multibox Detector (SSD).

In semantic segmentation, an algorithm tries to classify each pixel in an image into a set of predefined categories. Popular semantic segmentation algorithms include Fully Convolutional Network (FCN), U-Net and Google’s DeepLab.

**AI Application in anterior segment ophthalmic diseases**

A total of forty-five representative articles of AI application in anterior segment ophthalmic diseases are included in this review. Among them, the most intensively researched are conjunctiva and tear film, cornea, glaucoma and cataract. The distribution of the papers concerning these four topics in year of publication is shown in Figure 3. Details about aforementioned articles are summarized in http://fp.amegroups.cn/cms/8e23864133337a312ed32e11e7655d8/atm-20-976-1.pdf (39-53).

**Conjunctiva and tear film**

AI application in conjunctiva and tear film mainly focus on the diagnosis of dry eye. Dry eye is a multi-factorial disease characterized by unstable tear film causing a variety of symptoms and/or visual impairment, potentially accompanied by ocular surface damage (54). The diagnosis of dry eye is complicated and there is not a single perfect reference standard of it (55). Instead, a variety of examinations can provide supportive information, including fluorescein break up time test, tear film interferometry, tear film protein analysis and meibography.

Most investigations about dry eye were designed for the automatic analysis of relevant examinations (24-27,40,42,56). Among these studies, more emphasis is on tear film interferometry. The interference patterns of tear film lipid layer can be analyzed by MLP, SVM and other algorithms to achieve automatic classification (24-27). Meibography,
another imaging technology, has been investigated in another research (40). A combination of SIFT and SVM algorithms was applied to calculate the length and width of Meibomian glands in meibography images and then differentiate between healthy and unhealthy ones. Besides, fluorescein break up time test, a basic dry eye test, was also studied in (42,56). By analyzing the recorded videos during fluorescein break up time test, break-up time (BUT) can be calculated (42) and dry areas detected (56) using MRF and RANSAC algorithms respectively.

Other than direct observation, AI algorithm was also applied to screen for the diagnostic candidate molecules. Grus et al. proved that analysis of tear film protein patterns by ANN can serve as a diagnostic tool for detection of dry eyes (43,57). This artificial neural network showed an AUC of 0.93 and a specificity and sensitivity of approximately 90% for each (43), which provided a clue to accurate diagnosis of dry eyes.

Another research field lies in automatic conjunctival hyperemia grading, which is usually a subjective and time-consuming procedure for humans. Using slit-lamp images or even ordinary eye images as input, ML methods are able to perform an objective assessment of hyperemia grading, removing both intra- and inter-expert subjectivity while providing a reduction in computation time (37,39,41).

**Cornea**

Corneal ectasia, a group of eye disorders characterized by localized corneal thinning which leads to protrusion of the thinned cornea (58), seems to be the focus of AI research in cornea. Early detection of corneal ectasia is important since it is a contraindication for laser refractive surgery (59). A variety of AI models have been developed for the detection of corneal ectasia as a whole (28,31) or one or more specific forms of it (29,30,32,33,44), all of which were trained based on corneal parameters.

The most common form of corneal ectasia is keratoconus (60). Typically, keratoconus starts at puberty and may progress until the fourth decade or later (61). It’s essential that keratoconus is detected at an early stage, so corneal collagen crosslinking can be applied in time, preventing a major deformation of cornea and potentially the need for a corneal transplant. Souza et al. found that SVM, MLP and RBFNN classifiers, when trained on Orbscan II data, could represent useful techniques for keratoconus detection from other non-keratoconus patterns with no differences between their performances (44).

Forme fruste keratoconus, or preclinical keratoconus, is an early and subclinical form of keratoconus. Although it is relatively stable with subtle topographic characteristics, it can decompensate after laser refractive surgery. Kovacs et al. developed ML classifiers to discriminate eyes with preclinical keratoconus from normal eyes with an AUC of 0.96 (30), which is better than using unilateral single parameter.

More impressive works are from the multi-group classification of corneal conditions (29,32,33). Ruiz et al. developed a SVM algorithm to identify 5 different corneal patterns: keratoconus, forme fruste keratoconus, astigmatism, after refractive surgery and normal. For this 5-group classification, the accuracy was 88.8%, with a weighted average sensitivity of 89.0% and specificity of 95.2%, suggesting the potential of this model to screen patients before refractive surgery.

Except for corneal ectasia, automatic diagnosis of corneal arcus is also studied. S et al. used a SVM algorithm to detect corneal arcus and cataracts, both of which are common among elderly people, in visible-wavelength images and obtained satisfactory results (34).

Few studies paid attention to keratitis and other inflammatory or degenerative corneal diseases on the basis of images classification. The interpretable and expandable deep learning system developed by Zhang et al. can identify keratitis and other ocular surface disorders, distinguish different anatomical parts and foci, discern the diagnostic information relevant to the diagnosis, and provide treatment suggestions (45).

**Glaucoma**

Glaucoma is a group of ocular disorders with multifactorial etiology united by characteristic intraocular pressure-associated optic neuropathy and is a leading cause of irreversible blindness worldwide (62,63). Automatic detection of glaucoma-related features can contribute to timely diagnosis and thus improve patients’ prognosis (62).

The two main types of glaucoma are open-angle glaucoma and angle-closure glaucoma. They can be differentiated by the structure of anterior chamber angle. Patients of both types usually experience elevations of intraocular pressure (IOP) and visual field defect throughout the disease. Aloudat et al. used decision tree and SVM to identify the presence of high IOP from frontal eye images (35). This framework achieved an overall accuracy of 95.5%. As for analysis of visual defect, Wang et al. proposed
an archetype method to detect visual field progression based on spatial pattern analysis and this method outperformed clinician-based evaluation in terms of accuracy and agreement (46). These studies provided helpful techniques for diagnosing glaucoma and monitoring its progression.

AS-OCT scans can detect the presence of angle closure and thus contribute to the diagnosis of angle-closure glaucoma. ML as well as DL algorithms have been applied for the detection of angle-closure in OCT images or parameters (18-20,23,48). Fu et al. was the first to report the application of DL for this purpose (18). This DL system showed an AUC of 0.96 and proved a promising technology for the interpretation of OCT images. However, the usefulness of this system needs to be further validated in diverse population settings with the use of different devices. Additionally, two ML models trained on OCT images were also developed to classify the different mechanisms of angle-closure glaucoma (21,22).

Compared with angle-closure glaucoma, the diagnosis of open-angle glaucoma is more difficult to be established since patients rarely experience symptoms until central visual field defect occurs. Previous studies found that certain parameters extracted from the 24-hour contact lens sensor (CLS) have been associated with the visual field progression rate recorded in treated glaucoma patients (64). Martin et al. deployed a RF algorithm to discriminate between primary open-angle glaucoma and healthy eyes using CLS parameters and IOP (47). The results demonstrated that the combined feature set achieved a significantly better performance than the IOP feature set. Accordingly, CLS parameters may contain information complementary to IOP for the differentiation of primary open-angle glaucoma from healthy eyes.

**Cataract**

A cataract is the loss of lens transparency due to opacification of the lens. Cataracts are the leading cause of visual impairment worldwide, accounting for more than 50% of blindness in low-income and middle-income countries (65). Early detection and timely treatment of cataracts can improve patients’ quality of life and reduce healthcare costs (66). The current standard of management of a visually significant cataract is surgical removal of the cataractous lens and its replacement with an intraocular lens (67). Among the complications of cataract surgery, posterior capsule opacification is the most common one, which can cause secondary visual loss (68).

Multiple ML and DL algorithms, including SVM, DCNN and CRNN, have been applied for automatic diagnosis and grading of cataracts using different kinds of images such as slit-lamp images (10-15,36,45,49,50), visible wavelength images (34), and fundus images (16). The majority of aforementioned studies focused on a single specific cataract subtype, which can severely limit its application in real-world health-care settings. To address this problem, Wu et al. established a DL model to achieve the collaborative management of cataracts using a three-step strategy: capture mode recognition, cataract diagnosis and detection of referable cataracts with respect to etiology and severity (49). This AI agent showed satisfactory performance in the comprehensive diagnosis of age-related cataracts, pediatric cataracts and posterior capsule opacification. Such a universal AI platform can be utilized in multilevel clinical scenarios and thus greatly facilitate the collaborative management of cataracts.

Apart from the usage of images, several ML prediction models using clinical and biometric variables have also been built (51-53). These models can be deployed to optimize the postoperative refractive outcome (53) and estimate the risk of pediatric cataracts (51) or posterior capsule opacification (52).

**Challenges and future prospects**

Despite the reported successes of AI application in anterior segment ocular diseases, several limitations and hurdles must be appropriately addressed before widespread clinical implementation. Although vast amounts of data are generated in daily ophthalmology care, they need to be properly prepared before being included in usable data sets. Trained professionals are required in this process to accomplish labeling, annotations, segmentations and quality assurance, making this process expensive in both time and cost. This issue is aggravated by data-hungry algorithms such as CNN. Thus, the development of algorithms that can learn with less data is in urgent need. Actually, many self-supervised, weakly-supervised and unsupervised methods have achieved satisfactory performance with a small dataset (69,70), which can be a future trend in this field.

Besides the amount of data, the quality of data is also important and can greatly influence the performance of AI models (71,72). Unfortunately, availability of high-quality data still presents a key challenge in the implementation of AI technologies in medicine (73,74). Here, we propose a quality assessment system for images.

Before being included in data sets, images need to be scored in terms of the following items: being original or
not, resolution, features, anatomic structures, positions or projections for imaging, image capture devices, annotations, formats in naming and storing, description, and source information. In each item, there are different levels of quality, with 10 referring the best and 0 referring the worst. Details of assessment items are shown in Table 1.

Furthermore, data sets in most AI studies are collected from a homogenous population, which can be a barrier to broader application. Ideally, AI models should be validated in larger data sets from various patient cohorts under different conditions. For this purpose, established data sets should be made more accessible to promote intellectual
In addition to data issues, the characteristics of anterior segment ocular diseases also deserve more attention. For many anterior segment ocular diseases such as dry eye and glaucoma, it takes multiple examinations to add up to a definite diagnosis. However, current studies principally focus on automatic analysis of a single examination (like angle-closure detection in OCT images) rather than direct detection. For this reason, the integrative analysis of multimodal data (videos, images, formatting parameters, texts, etc.) to diagnose a complicated disease is an unsolved challenge in this field.

While it may take multiple examinations to diagnose one disease, different diseases can also be detected in a single examination and this is already partially realized in AI. For instance, a previous investigation showed that cataract and corneal arcus can both be detected in one slit-lamp image by a SVM algorithm (34). The efficiency and applicability of such a multitasking AI system is much better than traditional binary classification models. What’s more, slit-lamp images have also been used in other studies to accomplish conjunctiva hyperemia grading (37,39,41) and high IOP detection (35), indicating the potential of this system to complete even more tasks. Accordingly, more efforts should be made to establish this kind of multitasking systems.

In contrast to the diagnosis of posterior segment ocular diseases, which usually requires ophthalmoscopy, detection of anterior segment ocular diseases is much more convenient. Several AI models already used ordinary eye images rather than slit-lamp images to achieve conjunctival hyperemia grading (41), detection of high IOP (35), corneal arcus and cataract (34). Based on these studies, it’s possible that smartphone photography can serve as a diagnostic tool for anterior segment ocular diseases in the near future.

In conclusion, the applicability of AI in clinical settings can be improved by building systemic AI models trained on multimodal and heterogeneous data with a less data-hungry algorithm. Furthermore, home-based diagnosis and management of diseases may also be realized someday. Although ethical, legal and regulatory issues remain to be settled, AI will undoubtedly play a crucial role in the revolution of current health care patterns.

Acknowledgments

Funding: This study was supported by the National Key Research and Development Program (2018YFC0116500), the Key Research Plan for the National Natural Science Foundation of China in Cultivation Project (91846109), the Science Foundation of China for Excellent Young Scientists (81822010), the National Natural Science Foundation of China (81770967, 81873675, 81800810), the Science and Technology Planning Projects of Guangdong Province (2019B030316012, 2018B010109008, 2017B030314025), Guangdong Science and Technology Innovation Leading Talents (2017TX04R031), the Natural Science Foundation of Guangdong Province (2018A030310104). The sponsor or funding organization had no role in the design or conduct of this research.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Dr. Haotian Lin and Dr. Limin Yu) for the series “Medical Artificial Intelligent Research” published in Annals of Translational Medicine. The article was sent for external peer review organized by the Guest Editors and the editorial office.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-976). The series “Medical Artificial Intelligent Research” was commissioned by the editorial office without any funding or sponsorship. HL served as the unpaid Guest Editors of the series. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


