Artificial Intelligence for the Diagnosis and Screening of Retinal Diseases

Alessandro Arrigo, Emanuela Aragona and Francesco Bandello
IRCCS San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy

Artificial intelligence (AI) is becoming established as a new method for analysing ophthalmological data, and unveiling new clinical and pathogenic insights into retinal diseases. AI-based algorithms are largely employed in the field of the most prevalent retinal diseases, including diabetic retinopathy, age-related macular degeneration and myopia. Several research groups are also testing AI in other retinal diseases, including inherited retinal dystrophies, retinopathy of prematurity, central serous chorioretinopathy and retinal vein occlusion. AI models are mainly used in screening of the fundus and structural optical coherence tomography images. However, more advanced methodologies are under investigation to extract clinically relevant information regarding the biomarkers of disease activity and outcome measures. AI is a powerful tool for increasing the amount of information obtained in clinical and research contexts. Moreover, several ethical issues require debate, and specific rules are needed to govern the use of AI algorithms and check the quality of the analysed data. This article reviews the current use of AI in retinal diseases, unmet needs and future perspectives.

Article highlights

- Artificial intelligence (AI) represents a powerful way of analysing a large amount of data, extracting those features characterizing a given disease or condition.
- AI-based models have been largely tested in the main retinopathies, including age-related macular degeneration, diabetic retinopathy and myopia.
- AI algorithms have been highly effective in screening for retinopathies.
- The same approaches have unveiled potentially relevant biomarkers of disease progression and complications.
- Future efforts should focus on developing data quality checks and validating methodologies to reliably assess the contribution of AI-based models in clinical and research settings.

Background: Artificial intelligence and its application in ophthalmology

The term ‘artificial intelligence’ first appeared in 1956. One of the most frequently used definitions comes from a paper by Kaplan and Haeinlein, who defined AI systems as having “the ability to process external data systematically and learn from it to achieve specific goals and tasks”.3

The rapid growth of AI-based approaches in several disciplines stimulated the application of this technology in medicine. AI includes different possible approaches, which are usually categorized according to the method used to present a given problem to the machine. In particular, we can distinguish machine learning (ML) and deep learning (DL) technologies.

ML is defined as “computers’ ability to learn without being explicitly programmed”.3 In brief, ML uses raw data to extract learning information to train the algorithm and build models. There are different approaches in ML, the descriptions of which are beyond the scope of this review. Overall, ML approaches can be supervised or unsupervised, depending on the presence of human supervision.3

DL is an evolution based on a more complex extraction and analysis of data. Indeed, DL involves the use of deep neural networks, a complex multi-layer organization that is able to extract and elaborate complex patterns in data and to perform much more advanced and reliable analyses.4–6 The main differences between ML and DL are as follows:

- Data requirements: DL requires more data than ML to train the model on the basis of much more complex patterns.
- Extraction of the features: ML models usually manually extract the features from the data to train the model, whereas DL algorithms are able to automatically learn the features from the data.
- DL requires higher computational power and costs more than ML.
Interpretation: ML models are usually easier to interpret than DL models.

The main AI models are graphically summarized in Figure 1. A schematic of the main differences between ML and DL is shown in Figure 2.

Both ML and DL base data processing on artificial neural networks. That is, an imitation of the neural structure of the central nervous system. The information is then divided into different layers to facilitate data elaboration and the processing of inputs to outputs.

Considering all these characteristics, ophthalmology is well suited to the requirements of AI model development. Indeed, thanks to current diagnostic approaches, ophthalmologists can acquire very large datasets that are highly reliable, high quality and non-invasive way. This is particularly true in the medical retina field, where non-invasive multimodal retinal imaging represents the cornerstone of retinal diagnostics.7

In this article, we provide an updated overview of AI application for the most common retinal diseases.

**Methods**

We searched all English language and human subject articles using a keyword search of the MEDLINE library. Keywords included the following: artificial intelligence, machine learning, deep learning, macular disease, retinal disease, retinopathy and maculopathy. All the references were carefully examined by two expert researchers (FB, AA), who collected and ordered all the relevant information according to the main topic of the review.

**Artificial intelligence models in diabetic retinopathy**

Diabetic retinopathy (DR) represents a major cause of visual impairment in developed countries. The ever-growing incidence of DR and its complications necessitates new strategies to screen patients, prevent complications and optimize management. For this reason, DR is a major target of AI-based models, especially those that can be used for large-scale screening. AI models should recognize DR-related alterations, such as haemorrhages, exudates, cotton wool spots and neovascularization; and, starting from fundus images, establish the presence or absence of DR, and grade it according to the conventional DR grading scales.

Several AI systems have been developed to reach this goal. The IDx-DR system (Digital Diagnostics, Coralville, IA, USA) combines the results from DR-related biomarkers to grade DR. It was included in the Iowa Detection Program and tested in Caucasian, North African and sub-Saharan populations. This included images from 3,640 participants in the Nakuru Eye Study in Kenya, and reached a sensitivity of 87% and a specificity of 70%. The IDx-DR system was also tested in the Messidor-2 dataset (ADCIS, Saint-Contest, France) – which was established to facilitate the testing of computer-assisted diagnosis of DR – including fundus images.
of 1,748 eyes (874 patients), with a sensitivity of 97% and a specificity of 59%. Similar results have have also been achieved in real-life settings. Thanks to these good results, IDx-DR obtained approval from the United States Food and Drug Administration in 2018. The device can also work in conjunction with the Topcon NW4000 non-mydriatic fundus camera (Topcon Healthcare, Tokyo, Japan).

The RetmarkerDR software (Critical Health SA, Coimbra, Portugal) is a Conformité Européenne-marked class IIa medical device capable of distinguishing DR from non-DR eyes and measuring the microaneurysm turnover rate during follow-up. The system had a sensitivity of 73% for any DR, 85% for referable DR and 98% for proliferative DR. A similar system is EyeArt (Eyenkur Inc., Woodland Hills, CA, USA), which has a sensitivity of 95% for any DR, 94% for referable DR and 99.6% for proliferative DR. EyeArt was also tested on the Messidor-2 dataset, reaching a sensitivity of 94% and a specificity of 72%. EyeArt was implemented in smartphone-based DR screening of 296 patients, with a sensitivity of 96% for any DR, 99% for referable DR and 99% for sight-threatening DR. More recently, this tool has been employed on a dataset of more than 30,000 images, with a sensitivity of 96% for referable DR.

Other systems with similar sensitivities and specificities are a Google Inc. sponsored system (Mountain View, CA, USA), Bosch DR Algorithm (Robert Bosch GmbH, Gerlingen, Germany), a system developed by the Singapore National Eye Centre, Singapore Eye Research Institute and National University of Singapore School of Computing, RetinaLyze (Retinalyze A/S, Hørsholm, Denmark) and EyeWisdom® (Visionary Intelligence Ltd, Vistel, Beijing, China). Further AI-based DR screening systems are listed in the recent review by Grzybowski et al. The high reliability of AI-based screening has been further supported by a recent study testing a DL algorithm to analyse 14 common retinal alterations by Ting et al. in 2017 by analysing almost half a million images.

Another application of AI in the DR setting is in the grading and staging of the disease. Indeed, it is assumed that AI models may help ophthalmologists to optimize the detection and analysis of DR-related alterations, thus optimizing DR diagnostic work-up. Gulshan et al. tested AI for this purpose in 2016, achieving very high sensitivity and specificity in categorizing DR severity and the presence of diabetic macular oedema. Ting et al. supported these findings in 2017 by analysing almost half a million images.

Artificial intelligence models in age-related macular degeneration

Age-related macular degeneration (AMD) represents another leading cause of visual impairment in developed countries and a major target for AI-based approaches. The analysis of AMD data is mainly based on optical coherence tomography (OCT) images; thus, image segmentation is a fundamental step for accurately recognizing retinal structures. Many unsupervised AI-based segmentation algorithms have been developed, providing very high performance in detecting retinal features. These algorithms are useful not only for assessing the integrity of retinal structures, but also for localizing and quantifying retinal fluids during follow-up. Many approaches have been proposed in recent years, providing comparable results in terms of reliability, and these have been excellently reviewed by Schmidt-Erfurth et al.

Interestingly, AI has facilitated further progress in the AMD field, helping to assess treatment responses and factors predicting visual outcome. Moraes et al. used AI to assess distinctive features in first- and second-treated AMD eyes, highlighting differences in terms of AMD-related lesion prevalence and the course of the disease. Similarly, AI has proved useful in quantifying drusen and assessing their topographic distribution, both in fundus and OCT images, as well as for analysing hyper-reflective foci on OCT scans in early to intermediate AMD stages. These factors were used to calculate the risk of disease progression and onset of complications. In this context, Schmidt-Erfurth et al. provided important contributions regarding the use of AI in AMD:

- predicting the risk of AMD progression;
- highlighting the prognostic importance of intra-retinal cystoid fluid in neovascular AMD;
- developing automatic approaches for calculating fluid volumes to improve therapeutic management of neovascular AMD;
- assessing the efficacy profiles of two different anti-vascular endothelial growth factor drugs.

More recently, AI has successfully predicted visual outcomes for a treat-and-extend regimen, by using the changes in intra-retinal and sub-retinal fluids after single injections as predictive factors. AI-based approaches were useful for segmenting images of macular neovascularizations and for quantifying leakage on fluorescein angiography. However, although fluid changes have been investigated as important predictive biomarkers in neovascular AMD, a recent AI-based study found no significant relationship between OCT angiography characteristics of the macular neovascularizations and the response to treatments.

Moving beyond neovascular complications, AI has also been employed in the setting of geographic atrophy, based on the high reliability of AI-based algorithms to detect atrophic margins and to follow the expansion over time. Recently, AI has been tested alongside a novel treatment for geographic atrophy, namely pegecetacoplan. AI-based models were useful in calculating the topography and progression rate of geographic atrophy, as well as in quantifying photoreceptor thinning and loss over time.

A very interesting application of AI-based technology is in self-monitoring of macular status. One such system is the Notal Vision Home OCT (Notal Vision, Manassas, VA, USA). In a recent study, patients performed daily self-imaging at home using this system for 3 months, with the aim of daily monitoring neovascular activity by OCT images and AI-based elaboration. This system turned out to be feasible and fast, showing very good agreement with expert-based grading of disease activity. A similar self-monitoring strategy was tested in dry AMD with the ForeseeHome system (Notal Vision). This device collected clinical data from 2,123 patients over a mean duration of 3 years, and was highly feasible and useful in detecting changes in visual acuity and producing alert signals regarding the risk of disease progression.

Artificial intelligence models in myopia

Myopia is a global epidemic that is predicted to affect almost 50% of the world’s population by 2050. High myopia, defined as a spherical equivalent of 5.00 D or more, is a common finding, leading to potentially vision-threatening complications, such as myopic maculopathy and retinal detachment. Therefore, the development of even more powerful diagnostic approaches is fundamental to guaranteeing optimal management of myopic complications, and AI can play a major role in this field. Several AI algorithms have been tested to detect signs of pathological myopia both in fundus and OCT images. All these algorithms showed high reliability and sensitivity in detecting myopic alterations, with area under the curve values in most cases being ≥0.9.
AI has been tested in inherited retinal dystrophies to segment retinal structures on OCT, to detect atrophy changes and to follow all these alterations over time. In particular, AI models have been used in retinitis pigmentosa, choroideremia and Stargardt disease. The intriguing goal in this field is to improve characterization of patients with inherited retinal dystrophies and to study the genotype-phenotype correlations in very large cohorts of patients.

Another field of interest for AI models is retinopathy of prematurity (ROP), which is a major ophthalmological problem worldwide and a diagnostic clinical challenge. Many AI-based algorithms have been tested to screen for ROP and/or monitor disease progression. The diagnostic improvement of ROP provided by AI models could allow earlier identification and prompt treatment, which would reduce the impact on visual function and improve patients’ quality of life. Recently, a large study validated a vascular severity score as an appropriate output for AI-based models, finding a good correlation with the International Classification of Retinopathy of Prematurity staging system and paving the way for expanded use of AI algorithms in diagnosing ROP. Although the data available in the current literature should not be considered conclusive, more extensive use of AI-based approaches and clinical practice is likely.

Current limitations of artificial intelligence models: The other side of the medal

Although the contribution of AI-based algorithms is clear to the entire scientific community, the extensive use of AI technology is affected by several issues. First, the clinical validation of AI findings is mandatory to definitively assess the use of AI models in ophthalmology. Indeed, up to now, no standardized methods for evaluating the accuracy and robustness of AI methodology and findings have been developed. This limitation means that AI technologies are currently only used in research.

Major AI-related issues concern the usability, accessibility and cost demands of AI technologies. Indeed, AI models are expensive and require advanced technical skills and computational power, which strongly limits their use in areas with limited resources. In addition, AI algorithms require big data, which are difficult to collect in clinical practice. The feasibility of collecting data in daily practice, and the potential for patient distress in obtaining the data, should be carefully considered. The requirement for large amounts of data also leads irremediably to high heterogeneity in data quality. Indeed, we may assume that not all the data included in AI models are characterized by the same level of quality. For this reason, new ways of checking and classifying data quality are mandatory to improve the accuracy and relevance of the output from AI models.

Another major pitfall relates to data privacy and cybersecurity, given that AI technologies are often connected to the web. In addition, AI models are strongly associated with telemedicine strategies, with the potential for ethical biases, which should be carefully addressed. Although telemedicine has the undoubted advantage of increasing access to diagnostic procedures and reducing the burden for hospitals, it is impersonal and could significantly undermine the doctor–patient relationship and the inter-human emphatic exchanges that are the basis of care. Moreover, automating data analysis and interpretation provided by AI models is not synonymous with competence. AI technologies do not have the necessary expertise, intuition or critical thinking that are fundamental to properly interpreting findings and to reaching a diagnosis and considering therapeutic solutions. Therefore, human input will always be needed to process the output of AI technologies.

Conclusions

The current article provides an up-to-date review of AI-based technologies in retinopathies. AI has huge potential to improve the diagnostic work-up and management of retinopathies. Moreover, some studies have shown promising results in using AI-based approaches for the daily self-monitoring of patients. Currently, AI algorithms are mainly confined to research settings, especially in screening for the main retinopathies and identifying possible biomarkers of progression. However, the rapid growth of AI-based models is likely to result in wider adoption of these approaches both in clinical and research settings in the future. Caution is advised as power without control is potentially dangerous and many unmet needs remain. Novel methods for checking data quality and for validation are warranted to guarantee appropriate and responsible employment of AI technology in ophthalmology and healthcare systems.  

References

5. GUCON 2018. 8657997.


