Artificial Intelligence in Ophthalmology: Present and Future Directions

Orly Gal-Or MD[1,2,3], Alon Tiosano MD[1,2,3], Inbar Perchik BSc[2], Yogev Giladi MD[1,2], and Irit Bahar MD[1,2,3]

[1]Department of Ophthalmology and [2]Ophthalmology Artificial Intelligence Center, Rabin Medical Center (Beilinson Campus), Petah Tikva, Israel
[3]Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

ABSTRACT Artificial intelligence in ophthalmology is used for automatic diagnosis, data analysis, and predicting responses to possible treatments. The potential challenges in the application and assimilation of artificial intelligence include technical challenges of the algorithms, the ability to explain the algorithm, and the ability to diagnose and manage the medical course of patients. Despite these challenges, artificial intelligence is expected to revolutionize the way ophthalmology will be practiced. In this review, we compiled recent reports on the use and application of deep learning in various fields of ophthalmology, potential challenges in clinical deployment, and future directions.

IMAJ 2024; 26: 90–96

KEY WORDS: artificial intelligence (AI), deep learning, ophthalmology

Artificial intelligence (AI) is increasingly prevalent in the field of medicine. The new technology can enhance patient care and analyze large quantities of medical data, such as electronic records and imaging tests, to identify patterns and make predictions. These tools can assist physicians and researchers in the early diagnosis of diseases and help avoid potential risk factors in developing personalized treatment plans.

AI was initially described in 1956 as a technology capable of autonomous thinking and replicating human behavior after training [1]. Machine learning was introduced in 1959 as an algorithm exposed to multiple inputs, successfully adapting its behavior automatically [2]. Deep learning, a subset of machine learning, is defined as a system that utilizes a programmable neural network, enabling the machine to make decisions without human assistance. Convolutional neural networks structured to mimic the neural network of the human brain, *learns* to extract features from data through unique filters simulating the convolution process, requiring data processing to obtain the necessary information [3] [Figure 1].

Ophthalmology is among the pioneering medical fields that have embraced AI, with deep learning taking a central role in ophthalmic practice. This integration is primarily due to the widespread use of imaging tools for diagnostic and therapeutic decision-making and access to vast amounts of digital data [Figure 2]. Numerous AI researchers in ophthalmology have utilized ocular imaging to construct deep-learning models for data analysis [Figure 3]. Deep learning technology analyzes and filters data, enables automatic diagnosis, and predicts possible outcomes.

Figure 1. A comparative view of artificial intelligence

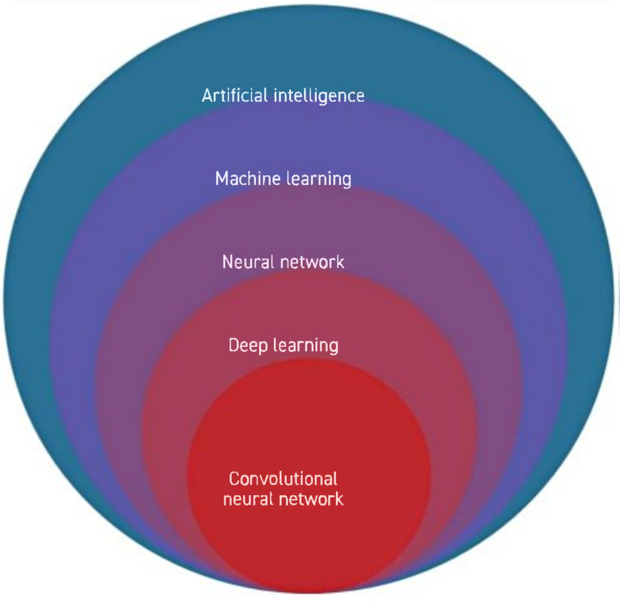
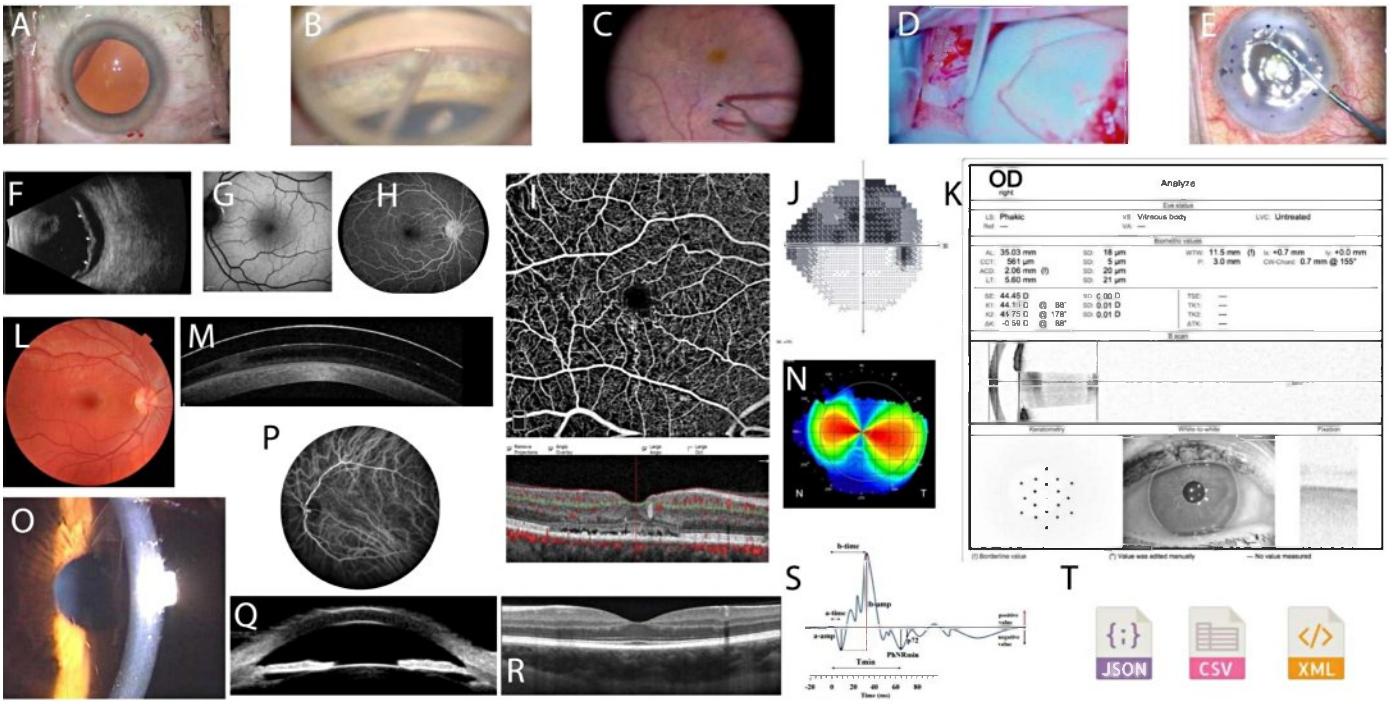


Figure 2. Types of visual and tabular data in ophthalmology

OCT = optical coherence tomography

[A-E] Surgical view of cataract, glaucoma, retina, optic nerve sheet and cornea surgeries, respectively, **[F]** B-scan ultrasonography of the eyeball, **[G]** Fundus autofluorescence, **[H]** Wide field fluorescein angiography, **[I]** OCT angiography, **[J]** Visual field, **[K]** Biometry printout, **[L]** Color fundus photography, **[M]** Anterior segment OCT, **[N]** Corneal topography, **[O]** Slit-lamp view of anterior segment, **[P]** Indocyanine green angiography, **[Q]** Ultrasound biomicroscopy, **[R]** OCT, **[S]** Electroretinogram, **[T]** Different types of tabular data

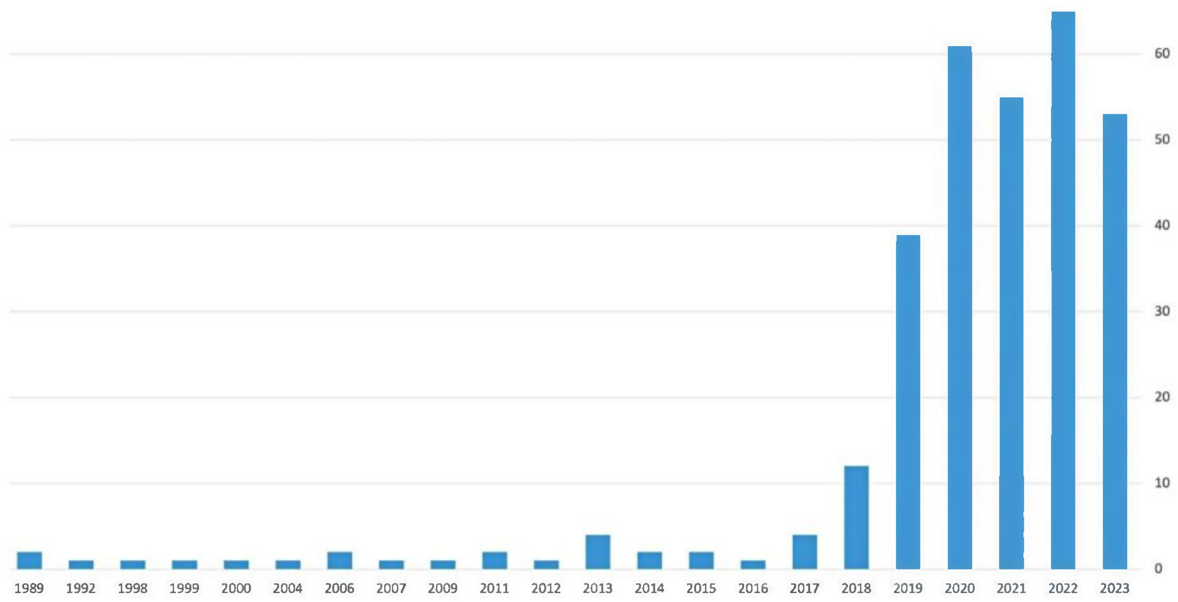


The visual information in ophthalmology is rich, diverse, and tailored to various subspecialties. Resembling other medical fields, this visual information is accompanied by textual and tabular data, providing an additional layer of rich context. Different imaging modalities can be classified according to their relevance to specific subspecialties. The field of oculoplastics focuses on visual information from images of eyelids. The anterior segment relies on visual information from the cornea, iris, and crystalline lens such as anterior segment optical coherence tomography (OCT), corneal topography, and ultrasound biomicroscopy.

OPHTHALMOLOGY STANDS OUT AS ONE OF THE PIONEERING MEDICAL FIELDS TO EMBRACE ARTIFICIAL INTELLIGENCE DUE TO ITS ACCESS TO LARGE AMOUNTS OF DIGITAL DATA AND THE NEED FOR DIAGNOSTIC AND THERAPEUTIC DECISION-MAKING TOOLS.

In addition, numeric data from measurements for calculating the power of intraocular lenses intended for cataract surgery and mappings illustrate numeric and visual data on corneal thickness and curvature structure. In the posterior segment domain, a wide variety of images includes fundus photographs and various angiographies employing contrasting agents (indocyanine green angiography, fluorescein angiography) to demonstrate pathologies in different layers of the retina and choroid, high-resolution OCT displaying a cross-section of retinal layers, and OCT-based angiography modeling blood vessels at different depths of the retina through the detection of red blood cell movement. Retinal imaging also encompasses ultrasonographic tests of the eyeball. The fields of glaucoma and neuro-ophthalmology similarly incorporate a wide range of visual images, including OCT for measuring the thickness of optic nerve fibers or the ganglion cell layer, visual field tests for functional assessment, and magnetic resonance

Figure 3. Ophthalmology publications focused on artificial intelligence based on PubMed search by MeSH Terms, Nov 2023



imaging and computed tomography for demonstrating the structure of the orbit or the visual pathways.

Today, the primary advantage of deep learning in ophthalmology lies in conducting screening tests for the diagnosis of common vision-threatening diseases such as diabetic retinopathy [4], glaucoma [5], age-related macular degeneration (AMD) [4,6], and retinopathy of prematurity [7]. In these cases, the models have achieved a level of accuracy comparable to that of an ophthalmology specialist.

UTILIZING DEEP LEARNING ON VISUAL DATA, INCLUDING FUNDUS PHOTOGRAPHS, OPTICAL COHERENCE TOMOGRAPHY SCANS, AND VISUAL FIELDS, DEMONSTRATES ROBUST PERFORMANCE IN DIAGNOSING A WIDE RANGE OF DISEASES INCLUDING DIABETIC RETINOPATHY, GLAUCOMA, RETINOPATHY OF PREMATURITY, AND AGE-RELATED MACULAR DEGENERATION.

In this review, we summarized the implementations of deep learning systems in ophthalmology, highlighting the clinical challenges and future potential.

DIABETIC RETINOPATHY

Screening tests for diabetic retinopathy and immediate referral for treatment are a global necessity to prevent blindness. These screening tests and the triage process can be conducted by various medical professionals, including ophthalmologists, optometrists, and medical technologists.

The use of AI for screening and grading the severity of diabetic retinopathy based on fundus images is among the most promising applications in medicine. In recent years, it has revolutionized the diagnostic performance of diabetic retinopathy. Recent studies have shown that AI systems have achieved diagnostic performance comparable to experts in the field and, in some cases, even demonstrated superior performance while proving to be more cost-effective for the healthcare system [8].

Among the groups demonstrating excellent diagnostic performance were Abràmoff and colleagues [9], who showed that their deep learning system achieved an area under the receiver operating characteristics curve (AUC) of 0.980 with sensitivity and specificity values of 96.8% and 87%, respectively, in detecting diabetic retinopathy requiring referral for treatment (defined as non-proliferative diabetic retinopathy of moderate severity or worse and including referable macular edema). Their results were tested and validated on publicly available data.

Gulshan and co-authors [10] from Google AI Healthcare trained a deep learning system for grading diabetic retinopathy and referable macular edema compared by 54 ophthalmologists in the United States on 10,000 images from publicly available datasets (EyePACS-1 and Messidor-2), achieving AUC performance of 0.991 and 0.990, respectively.

Ting et al. [11] published a diagnostic system developed and evaluated on 10 external datasets, which was tested over 5 years in various countries (Singapore, China, Mexico, United States, and Australia). Their model demonstrated an AUC performance of 0.936, sensitivity of 90.5%, and specificity of 91.6%, emphasizing reliable diagnosis in different ethnic groups.

In addition, deep learning technologies have been used in diabetic retinopathy patients to determine the prevalence and risk factors for cardiovascular diseases and predict the development of diabetic macular edema based on OCT images and fundus photographs, achieving AUC values of 0.89, sensitivity of 85%, and specificity of 85% [12]. AI systems have also received U.S. Food and Drug Administration approval. IDX-DR was the first automated system approved for diabetic retinopathy screening, followed by the EyRIS SELENA system, which was approved for clinical use in the European Union [13].

THE IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE MODELS IN CLINICAL PRACTICE POSES TECHNICAL AND CLINICAL CHALLENGES.

GLAUCOMA

Patients with glaucoma may present with irreversible damage to their visual fields if diagnosis and therapeutic intervention are not performed in a timely manner. This situation represents a clinical need that can benefit from using AI. Clinicians and patients would notice improvements in disease diagnosis at early stages, assess structural and functional damage, and customize optimal treatment to prevent visual impairment and enable functional ability in the long term. Several studies have employed AI for glaucoma diagnosis based on structural changes in fundus images and OCT scans [14,15]. The algorithmic results demonstrated sensitivity and specificity of 95.6% and 92%, respectively, with an AUC of 0.986 in detecting secondary optic neuropathy due to glaucoma.

Xu et al. [16] identified angle-closure and primary angle-closure glaucoma by implementing an algorithm on 4000 angle anterior-segment OCT images, achieving AUC values of 0.928 and 0.964, respectively.

Elze and colleagues [17] developed an algorithm that identifies patterns of visual field damage, proving helpful in detecting early visual field damage secondary to glaucoma. Lichter and co-authors [18] developed a clinical prediction tool that utilized intraocular pressure and visual field data to predict disease progression based on various target pressure scenarios. Machine learning provided various tools for glaucoma diagnosis and progression prediction. In an ideal future scenario, these tools might even demonstrate to patients the prognosis of their expected clinical course, predicting the likelihood of visual field damage progression while optimizing various therapeutic strategies.

AGE-RELATED MACULAR DEGENERATION

Age-related macular degeneration (AMD) is the primary cause of irreversible vision loss in the elderly. The most common screening method involves using color fundus photographs to identify findings such as drusen, geographic atrophy, macular thickening, and hemorrhage. The scanning process takes only a few seconds and is non-invasive. With the aging population and increased life expectancy, there is a clinical need for a deep learning system for early diagnosis and referral of patients requiring treatment at tertiary centers. Burlina and colleagues [6] developed a deep learning system based on the Age-Related Eye Disease Study data with an accuracy of 91.6% and an AUC of 0.96.

In addition to diagnosing AMD based on fundus images, some studies have focused on predicting the risk of disease progression. Burlina and colleagues [19] built a regression model based on deep learning that estimates a patient's presumed risk of progressing to advanced AMD within 5 years.

OCT is a breakthrough in managing patients with macular retinal diseases such as macular aging and secondary macular edema due to diabetes. An OCT examination allows microscopic visualization of retinal layers and pathological findings. From an AI perspective, OCT has several attractive features. First, it makes available a vast amount of OCT images required for model training. Second, successive scans allow volumetric reconstruction, and third, imaging of structural details not demonstrated by other imaging modalities can be incorporated. Recently, there has been an increase in the use of AI algorithms for quantitative analysis of findings in OCT images. Schlegl and colleagues [20] developed a deep learning system for patients with neovascular AMD that automatically identifies and quantifies subretinal fluids, like a retina specialist. Moraes and colleagues [21] expanded the algorithm beyond quantitative fluid processing to detect additional biomarkers, such as hyper-reflective foci and subretinal hyper-reflective material, demonstrating high clinical value. In addition, Zhang et al. [22] successfully developed an algorithm that identifies geographic atrophy (GA) based on OCT images in dry AMD patients. The algorithm detects loss of the retinal pigment epithelium layer, disappearance of photoreceptors, and areas of increased choroidal reflectivity associated with GA. Integrating multiple imaging modalities is a hot topic in AI research.

DeepMind and Moorfields Eye Hospital recently combined AI systems for filtering and classification tasks. Their system was the first to identify 15 different morphological features of the retina successfully. The system's output was fed into a classification network that performed triage into various categories based on the urgency of the findings [23]. In the longer term, such a system could be used in community settings where OCT devices are available, applying the system for initial triage without the intervention of an ophthalmologist.

RETINOPATHY OF PREMATURITY

Retinopathy of prematurity (ROP) is the leading cause of blindness in infancy. Clinical examination or digital fundus photography is required for the identification of early signs of ROP at a severity level requiring therapeutic intervention. Timely treatment can prevent most cases leading to blindness due to ROP [24]. There are two main barriers to ROP screening. First, there is significant variability among diagnosticians when diagnosing ROP, leading to inconsistency in therapeutic interventions. Second, there is a shortage of skilled diagnosticians. Telemedicine implementation addresses these challenges, allowing a single physician to virtually screen infants over large areas. Brown and colleagues [7] developed a deep learning system that they compared to the conventional diagnostic process of clinical examination and fundus photography. The system successfully identified severe stage plus disease of ROP with an AUC of 0.98. In addition, the system provided a severity rating for ROP.

ARTIFICIAL INTELLIGENCE IN OTHER EYE DISEASES

Beyond retinal diseases and glaucoma, AI has potential applications in automatically diagnosing and grading cataract based on slit-lamp examination or anterior segment images. Li et al. [25] developed a deep learning system for diagnosing anterior segment diseases such as cataract, keratitis, and pterygium by segmenting anatomical structures and marking pathological findings. Yoo et al. [26] integrated preoperative data with anterior segment images, successfully predicting postoperative spectacle independence with an accuracy of 93.4% and an AUC of 0.97. A recent Cochrane review by Vandevenne and colleagues [27] reviewed 63 studies to assess the diagnostic ability of AI in keratoconus. They found AI to be a promising tool in diagnosing manifest keratoconus.

CHALLENGES

Despite the high accuracy of AI-based models, their daily practice application is hindered by technical and clinical limitations. First, many published studies were trained on homogeneous datasets. Training and evaluating an AI model based on retinal images are subject to significant variations due to differences in image quality, resolution, magnification, or field of view. Diverse data representation is crucial for the model to handle this challenge. Second, limited data availability for rare diseases, such as choroidal or retinal lesions/genetic retinal diseases, or significant variability among examiners in diseases like glaucoma or ROP, poses a challenge. AI algorithms learn based on the presented data, and if the training dataset is small or does not represent true patient populations, it can impact accuracy of the results. Third, clinicians and patients are still concerned about the *black box* component in AI models. To overcome this limitation, researchers commonly incorporate an explanatory component in the model, which is represented by heatmaps highlighting the region that helped the algorithm reach its conclusion or alternatively presenting the different weights of model components [28]. Researchers at Moorfields Eye Hospital and DeepMind developed an alternative approach to address the *black box* issue. They demonstrated an intermediate stage where a filtering network is presented, emphasizing the relevant pathology location for the physician [23].

Another challenge is the diverse regulatory approvals required between different countries, with the legal aspects of clinical management governed by technology-based AI. The implementation of AI-based medicine depends not only on physicians but also on patient willingness to accept it, posing a significant challenge in its integration.

FUTURE DIRECTIONS

AI is poised to become an integral part of the diagnostic process, treatment decision-making, and clinical research in the not-so-distant future. To achieve this goal, there are several clinical domains where researchers need to enhance diagnostic capabilities through improving test sensitivity and specificity is evident. These areas include evaluating the optic nerve head for the presence of optic neuropathy through fundus photography [29], automatic identification and management of progressing visual field loss in glaucoma patients, myopia progression prediction in children, and more.

AI may also play an important role in surgeries by guiding surgeons and alerting them before a complication is recognized. In addition, the growing phenomenon of robot-assisted surgery will perform based on different advanced AI models.

There is a need to adapt AI networks focusing on multi-modality for diagnosis, disease progression prediction, and treatment customization. The importance of external validation of algorithms that have undergone training on small and homogeneous populations and have shown promising results is crucial, as they may perform less accurately when tested in a heterogeneous population in real-world settings. Beyond improving diagnostic accuracy, the effectiveness and capability of AI tools to enhance patient care need to be assessed through prospective clinical research. Reporting guidelines for models based on AI will need to adhere to international agreements, making them more transparent and allowing for an evaluation of their effectiveness and utility.

CONCLUSIONS

AI research has achieved a breakthrough in the field of ophthalmology. The wealth of data, advancements in digital development, and the demand for high-quality medical services have led to the advanced use of AI systems to maximize performance in routine clinic work, treatment optimization, and prediction of ophthalmic disease progression. Collaborative efforts among ophthalmologists, AI researchers, and engineers must find solutions to clinical problems and implement these technologies. There are still barriers to implementing and integrating these algorithms into clinical practice. Future research is crucial in evaluating the cost-benefit of deep learning systems in patient care. Despite the numerous challenges, AI in ophthalmology appears promising and is expected to influence the nature of our engagement and treatment provision for patients.

Correspondence

Dr. O. Gal-Or
Dept. of Ophthalmology, Rabin Medical Center (Beilinson Campus), Petah Tikva 49100, Israel
Phone: (972-3) 937-6101
Email: wataya@gmail.com

References

1. McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Mag* 2006; 27 (4): 12.
2. Samuel AL. Some studies in machine learning using the game of checkers. *IBM J Res Dev* 1959; 3 (3): 210-29.
3. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; 521 (7553): 436-44.
4. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA* 2017; 318 (22): 2211-23.
5. Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology* 2018; 125 (8): 1199-206.
6. Burlina PM, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol* 2017; 135 (11): 1170-6.
7. Brown JM, Campbell JP, Beers A, et al. Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks. *JAMA Ophthalmol* 2018; 136 (7): 803-10.
8. Xie Y, Nguyen QD, Hamzah H, , et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. *Lancet Digit Health* 2020; 2 (5): e240-9.
9. Abràmoff MD, Lou Y, Erginay A, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci* 2016; 57 (13): 5200-6.
10. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016; 316 (22): 2402-10.
11. Ting DSW, Cheung CYL, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA* 2017; 318 (22): 2211-23.
12. Varadarajan AV, Bavishi P, Ruamviboonsuk P, et al. Predicting optical coherence tomography-derived diabetic macular edema grades from fundus photographs using deep learning. *Nat Commun* 2020; 11 (1): 130.
13. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018; 1: 39.
14. Liu H, Li L, Wormstone IM, et al. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs. *JAMA Ophthalmol* 2019; 137 (12): 1353-60.
15. Thompson AC, Jammal AA, Berchuck SI, Mariottoni EB, Medeiros FA. Assessment of a segmentation-free deep learning algorithm for diagnosing glaucoma from optical coherence tomography scans. *JAMA Ophthalmol* 2020; 138 (4): 333-9.
16. Xu BY, Chiang M, Chaudhary S, Kulkarni S, Pardeshi AA, Varma R. deep learning classifiers for automated detection of gonioscopic angle closure based on anterior segment OCT images. *Am J Ophthalmol* 2019; 208: 273-80.
17. Elze T, Pasquale LR, Shen LQ, Chen TC, Wiggs JL, Bex PJ. Patterns of functional vision loss in glaucoma determined with archetypal analysis. *J R Soc Interface* 2015; 12 (103): 20141118.
18. Lichter PR, Musch DC, Gillespie BW, et al. Interim clinical outcomes in the Collaborative Initial Glaucoma Treatment Study comparing initial treatment randomized to medications or surgery. *Ophthalmology* 2001; 108 (11): 1943-53.
19. Burlina PM, Joshi N, Pacheco KD, Freund DE, Kong J, Bressler NM. Use of deep learning for detailed severity characterization and estimation of 5-year risk among patients with age-related macular degeneration. *JAMA Ophthalmol* 2018; 136 (12): 1359-66.

20. Schlegl T, Waldstein SM, Bogunovic H, et al. Fully automated detection and quantification of macular fluid in oct using deep learning. *Ophthalmology* 2018; 125 (4): 549-58.

21. Moraes G, Fu DJ, Wilson M, et al. Quantitative analysis of OCT for neovascular age-related macular degeneration using deep learning. *Ophthalmology* 2021; 128 (5): 693-705.

22. Zhang G, Fu DJ, Liefers B, et al. Clinically relevant deep learning for detection and quantification of geographic atrophy from optical coherence tomography: a model development and external validation study. *Lancet Digit Health* 2021; 3 (10): e665-75.

23. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018; 24 (9): 1342-50.

24. Elsas FJ, Collins M, Jones J, et al. Multicenter trial of cryotherapy for retinopathy of prematurity: ophthalmological outcomes at 10 years. *Arch Ophthalmol* 2001; 119 (8): 1110-8.

25. Li Z, Jiang J, Chen K, et al. Preventing corneal blindness caused by keratitis using artificial intelligence. *Nat Commun* 2021; 12 (1).

26. Yoo TK, Ryu IH, Lee G, et al. Adopting machine learning to automatically identify candidate patients for corneal refractive surgery. *NPJ Digit Med* 2019; 2: 59.

27. Vandevenne MM, Favuzza E, Veta M, et al. Artificial intelligence for detecting keratoconus. *Cochrane Database Syst Rev* 2023; 11 (11): CD014911.

28. Lam C, Yu C, Huang L, Rubin D. Retinal lesion detection with deep learning using image patches. *Invest Ophthalmol Vis Sci* 2018; 59 (1): 590-6.

29. Milea D, Najjar RP, Jiang Z, et al. Artificial intelligence to detect papilledema from ocular fundus photographs. *N Engl J Med* 2020; 382 (18): 1687-95.