The Algorithms of Power

Gurus of ophthalmic AI celebrate the development of artificial intelligence technologies and presents questions to be addressed before the AI-based medical devices are introduced to real life

Hazards and Potential Problems of AI Medical Devices

By Andrzej Grzybowski, Professor of Ophthalmology and Chair of Department of Ophthalmology, University of Warmia and Mazury, Olsztyn, Poland, and Head of Institute for Research in Ophthalmology, Foundation for Ophthalmology Development, Poznań, Poland

Prof. Andrzej Grzybowski

We have recently seen a significant development of many AI-related technologies and applications, and there's a lot of enthusiasm about the promises of AI in healthcare, including improving patient and population outcomes, making medical teams' work easier, as well as reducing costs by avoiding errors and unnecessary procedures. We have entered the fourth stage of the Industrial Revolution and AI is its most important theme.

Ambitious expectations for AI in healthcare include outperforming doctors, helping to diagnose what is presently undiagnosable to treat what is currently untreatable, predicting the unpredictable, and classifying the unclassifiable. AI might help preserve the doctor-patient relationship and move it from

the present “shallow medicine” into “deep medicine” based on deep empathy and connection. Currently, the average time of a clinic visit in the US for an established patient is seven minutes, and for a new patient – 12 minutes, and in many Asian countries it is down to two minutes per patient. To make this even worse, part of this time must be devoted to completing the electronic health record.

AI-based “deep medicine” might give us more time for crucial relations with our patients – and those cannot be replaced by technology. AI-based technologies using the deep-learning (DL) approach have been shown to support decisions in many medical specialties, including radiology, cardiology, oncology, dermatology, and ophthalmology. AI/DL models have reduced waiting times, improved medication adherence, customized insulin dosages, and helped interpret magnetic resonance images. AI/DL algorithms were shown to detect disease states based on image analysis, including retinal diseases from fundus photos and OCT scans, lung diseases from chest radiographs, and skin disorders from skin photos. Two autonomous AI-based medical devices are registered in the US for detection of diabetic retinopathy, and a few more are available in the EU; AI algorithms have been used for DR screening in many parts of the world.

In Poland, I started an AI-based DR screening project in 2017, and since 2018, my team has been conducting a big project aiming to screen 40,000 diabetic patients in the Wielkopolska region, funded by the EU. A new and very promising application is to use eye images to identify risk of cardiovascular or neurodegenerative disorders.

However, when talking about the increasing enthusiasm around and prospects of AI in ophthalmology, we must also mention rising problems and questions that need to be addressed before the AI-based medical devices are introduced to real life.

One of the main problems is the lack of clarity of what constitutes the evidence of impact and demonstrable benefit for the many AI-medical devices, and who can assess the evidence.

The future development of the AI field depends on an easier – and, preferably, unlimited – access to the medical data stored within the electronic health records. This access, however, cannot constitute privacy overuse of this very sensitive data. According to the US National Institute of Standard and Technology, biometric data, including retina images, are personally identifiable information and should be protected from inappropriate access.

Although present AI models were shown to diagnose and stage some ocular diseases from images, including fundus photos, OCT and visual field data, most AI algorithms were tested on a dataset not corresponding well to real-life conditions. Patient populations were usually homogeneous regarding ethnicity, age, lack of comorbidities, and poor-quality images.

Moreover, some algorithms were shown to misrepresent and exacerbate health problems in minority groups. Future datasets should better describe who is represented and in what way, to avoid structural biases (please see one of the recent initiatives at www.datadiversity.org).

Source: https://www.canva.com/

Thus, future studies on validating algorithms on real life ocular images from heterogenous populations are needed, including both good- and poor-quality images. Otherwise, we may face a "good-AI-gone-bad." Cherry-picking best results might make the situation even worse. It should be highlighted that AI-based algorithms might behave unpredictably when applied in the real world. It has been shown that performance of the algorithm degrades when applied to images generated by a different device or in a different clinical environment to those of the training set. All these problems might lead to misdiagnosis and erroneous treatment suggestions, breaching the trust in AI technologies. Finally, we should be able to imagine that if an AI system made an error, it could harm hundreds or even thousands of patients. Thus, I like to repeat Tetlock & Gardner (Superforecasting) who said: "If you do not get feedback, your confidence grows much faster than your accuracy."

One of the recent independent studies comparing seven different algorithms, reported that one of the tested algorithms was significantly worse than human graders at all levels of DR severity – it missed

25.58 percent of advanced retinopathy cases, which could potentially lead to serious consequences (1). This study showed possible problems and patient safety hazards related to clinical use of some algorithms. They include limitations related to training of an algorithm on particular demographic group, including ethnicity, age, sex, and its further use on a different population. Moreover, many studies exclude low-quality images, treated as ungradable images, and patients with comorbid eye diseases, which makes them less reflecting the conditions of real life.

It should be also remembered that AI algorithms can be designed to perform in unethical ways. For example, Uber's software, Greyball, was designed to allow the company to identify and circumvent local regulations and Volkswagen's algorithm that allowed vehicles to pass emission tests by reducing their nitrogen oxide emission when they were being tested. Moreover, clinical decision-support algorithms could be designed to generate increased profits for their owners, such as recommending particular drugs, tests, and more, without clinical users' awareness. Finally, AI systems are vulnerable to cybersecurity attacks that could cause the algorithm to misclassify medical information. For more on this subject, reach for our recent publication, Artificial Intelligence in Ophthalmology (2).

1. Lee AY, et al. Multicenter, Head-to-Head, Real-World Validation Study of Seven Automated Artificial Intelligence Diabetic Retinopathy Screening Systems. Diabetes Care. 2021 May;44(5):1168-1175.

2. Artificial Intelligence in Ophthalmology, A.Grzybowski (ed), Springer 2021. https://link.springer.com/book/10.1007/978-3-030-78601-4

Our virtual AI in Ophthalmology Meeting in **June 2022**, sponsored by the Polish Ministry of Science and Education, turned out to be a great success, with over 600 registrations from over 20 countries, and lectures delivered by world-leading specialists in this field. I have received many requests to repeat the event next year.

As collaboration and networking of people interested in future applications of AI in ophthalmology is vitally important, I decided to start building the foundations for the International AI in Ophthalmology Society (IAIOph).

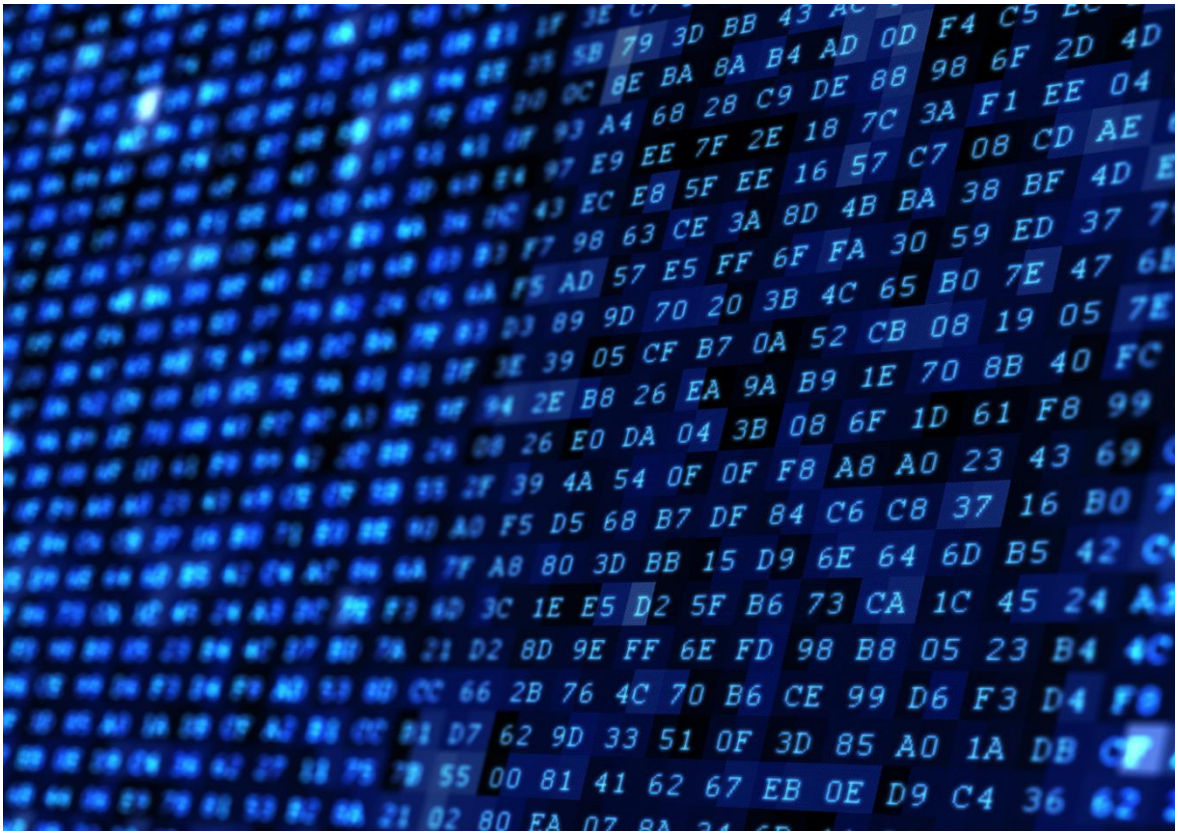
Everyone is welcome to join it directly at iaisoc.com or by emailing me: ae.grzybowski@gmail.com.

All the lectures from the 2022 AI in Ophthalmology Meeting are available at **aiinophthalmology.com.**

What are the major challenges to developing AI further in the near future?

***Linda Zangwill**, Professor of Ophthalmology in Residence, Richard K. Lansche and Tatiana A. Lansche Endowed Chair, Co-Director of Clinical Research, Hamilton Glaucoma Center Director, Data Coordinating Center, Shiley Eye Institute, UC San Diego, California, USA*

Development of AI algorithms to detect glaucoma is now relatively straightforward if one has appropriate datasets and computational resources. One of the major challenges to the implementation of AI in clinical settings is to ensure that the algorithm is generalizable to the targeted populations and not biased due to limitations of the training set. Evaluating the generalizability of the results requires extensive testing of the AI algorithm on external datasets from diverse populations. Another challenge is determining how to integrate the AI system and results into clinical practice. Where and how should the AI algorithm results be placed in the electronic health record or PACS system that the clinician uses in their routine management of glaucoma patients? What type of summary information and/or visualization of the AI results should be provided? It is essential to determine how the AI results can be provided in a way that is easy and fast to use so that it provides added value and does not slow down the busy clinical workflow. One can develop the best AI algorithm, but if clinicians are not willing or able to use it, it will not improve clinical care. Other challenges for the development and implementation of AI include how best to open the black box to provide information on what the algorithm used to make its decision, as well as medical, legal, ethical, and privacy issues.



Source: https://www.canva.com/

***Michael F. Chiang**, Director, National Eye Institute, National Institutes of Health, Bethesda, Maryland, USA*

I will articulate a few challenges: first, we are losing many opportunities to utilize ophthalmic image data for developing AI systems because those data are locked in proprietary standards and inaccessible to researchers and clinicians. Second, we need to improve the culture of data sharing, standards for data representation, and methods for establishing ground truth to take full advantage of building large, AI-ready datasets for knowledge discovery. Third, AI systems are best at addressing discrete questions (such as "Is there plus disease in this retinal image from a baby undergoing ROP screening?"), whereas real-world scenarios require addressing numerous questions in parallel. Fourth, AI systems are typically trained and validated in fairly narrow populations and specific imaging devices, whereas real-world applications will need to be rigorously validated to ensure they work across broad populations and devices without bias.

***Damien Gatinel**, Head of the Anterior Segment and Refractive Surgery Department, Rothschild Foundation Hospital, Paris, France*

The limits of AI development mainly concern data collection because the common point of any project is to use a large volume of quality data. It is common that even when a large data set has been compiled, it is necessary to reduce its size drastically.

We can also foresee certain ethical problems insofar as we sometimes do not know by what mechanism(s) certain results are obtained in terms of classification or prediction.

***Paisan Ruamviboonsuk**, Clinical Professor of Ophthalmology, College of Medicine, Rangsit University, Assistant Hospital Director for Centers of Medical Excellence, Center of Excellence for Vitreous and Retinal Disease, Rajavithi Hospital, Bangkok, Thailand*

I think we can take advantage of multimodal images in ophthalmology to develop AI models that are more efficient in screening or detecting diseases or detecting disease progression. There are countless AI models for different kinds of tasks today; however, the major challenges for me rest on how useful these models are in reducing the risk of blindness; how useful they are to be deployed in the real-world. Many AI models work well in internal validation but fall short in real-world deployment. The other challenges would rest on the "prediction" of treatment outcome and disease progression. The models for these tasks now have accuracy around 70 percent, we look forward to better predictions in the future.

***Michael D. Abràmoff**, The Robert C. Watzke, Professor of Ophthalmology, Professor of Electrical and Computer Engineering, and Biomedical Engineering, Department of Ophthalmology and Visual Sciences, University of Iowa Hospital and Clinics, Iowa, USA*

Theoretical challenges that I see: in healthcare, training data will always be sparse, so how can we build AIs that use limited amounts of training data and how do we use proxies under deep learning conditions? Under what conditions can an AI be changed “somewhat” without requiring full (and often expensive) validation? We have to be able to figure out how we expand reimbursement for AIs that meet some but not all of the criteria above, and how we deal with the information loss that comes with repeated examination of existing datasets, such as an expensive validation dataset. Practical challenges that I predict include, but are not limited to: the need for better education and adoption of highly validated AI systems that are integrated into clinical workflows and sustainably reimbursed. AI in healthcare needs to focus on solutions that offer the greatest benefit to patients. How do we regulate vernacular AIs that are safe and effective in certain subpopulations but not others? While there may be AI technologies that sound exciting, if they aren’t positively impacting patient outcomes, they won’t bring any real benefit to healthcare and could slow the adoption of the solutions having a positive impact. Of course, all of this is dependent on having access to appropriately diverse and reliable data sets with which to train new AI systems.

Prof. Linda Zangwill

AI and Glaucoma

Linda Zangwill, Professor of Ophthalmology in Residence, Richard K. Lansche and Tatiana A. Lansche Endowed Chair, Co-Director of Clinical Research, Hamilton Glaucoma Center Director, Data Coordinating Center, Shiley Eye Institute, UC San Diego, California, USA

Why should AI be used in glaucoma? What are the benefits of AI in this field?

There are numerous benefits to using AI to assist in clinical decision making for glaucoma detection and management. AI can improve the accuracy and consistency of glaucoma detection across all levels of ophthalmic care. It can also be used to detect individuals with progressive glaucoma who are in need of closer follow-up and conversely suggest that a patient's glaucoma is stable and requires less frequent follow-up. In addition, by providing information on the probability of glaucoma, the clinician can integrate this information into their patient management decisions. Moreover, AI can help to screen high-risk individuals for glaucoma in primary care and community settings, so that the disease can be diagnosed in its earliest treatable stages.

Source: https://www.canva.com/

Do we know which AI algorithms are better: those based on OCT or those developed on fundus images?

AI algorithms for glaucoma detection based on OCT and fundus images both have high diagnostic accuracy (1, 2). Both are useful as they can be valuable in different settings. OCT is the standard of care for the clinical management of glaucoma in most ophthalmology clinics. However, in many communities, particularly in underserved areas, fundus photography is much more likely to be available than OCT imaging. Moreover, glaucoma detection using fundus photography can be integrated with screening for other eye diseases such as diabetic retinopathy and macular degeneration more easily in primary care settings. It is therefore important to develop accurate AI algorithms for both OCT and fundus photography.

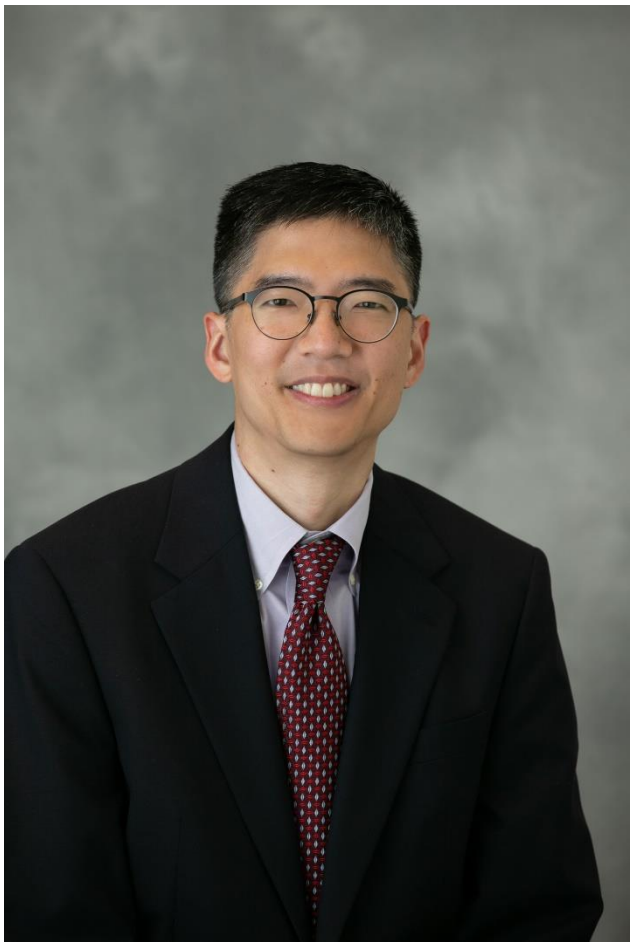
Might AI-based glaucoma detection algorithms be used in glaucoma screening in the future?

I believe that AI-based glaucoma detection algorithms will be used for targeted glaucoma screening of high-risk individuals in primary care and/or community settings. Glaucoma screening will likely be integrated with existing algorithms for detection of diabetic retinopathy and other eye diseases. As AI algorithms provide a probability of glaucoma, the cut-off used to refer for follow-up ophthalmic examination can be set to a high specificity needed for screening tests. It should be noted that the US Preventive Services Task Force on Screening for Primary Open Angle Glaucoma recently concluded in the general population of asymptomatic adults 40 years and older "the current evidence is insufficient to assess the balance of benefits and harms of screening for primary open-angle glaucoma" (3). For these reasons, integration of the glaucoma screening with other eye diseases, and targeted screening to high-risk populations should be considered.

1. JH Wu, "Performances of machine learning in detecting glaucoma using fundus and retinal optical coherence tomography images: A meta-analysis," Am J Ophthalmol, 237, 1 (2022). PMID: 34942113.

2. AK Chaurasia, "Diagnostic accuracy of artificial intelligence in glaucoma screening and clinical practice," J Glaucoma, 31, 285 (2022). PMID: 35302538.

3. US Preventive Services Task Force et al., "Screening for primary open-angle glaucoma: US Preventive Services Task Force Recommendation Statement," JAMA, 27, 1992 (2022). PMID: 35608574.



Prof. Michael Chiang

Retinopathy of Prematurity Perspective

Michael F. Chiang, Director, National Eye Institute, National Institutes of Health, Bethesda, Maryland, USA

Why was there a need for a new retinopathy of prematurity (ROP) classification?

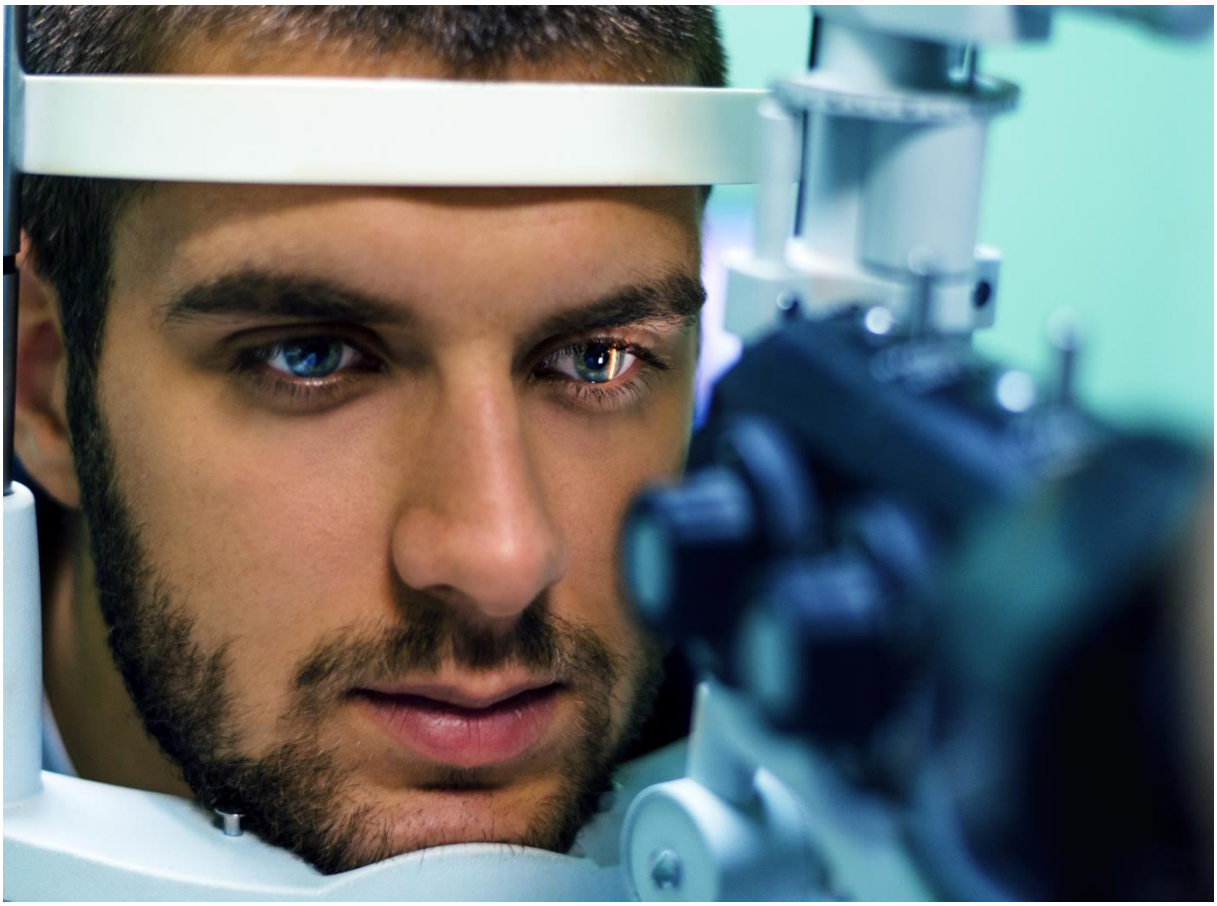
Interestingly, the early studies on ROP, then called retrolental fibroplasia, were conducted by Thaddeus S. Szewczyk, an American ophthalmologist of Polish origin, who is mostly forgotten today. He was the first to indicate a relationship between the development of ROP and high exposure to oxygen in an incubator or by withdrawing oxygen too rapidly. What has changed in understanding of the ROP pathophysiology in recent 50 years?

We've recently published the third version of the international ROP classification system because a number of new challenges since the previous in 2005 have arisen: (i) concerns about subjectivity in critical elements of ROP disease classification such as plus disease and zone, (ii) innovations

in ophthalmic imaging and artificial intelligence, (iii) novel pharmacologic therapies (such as anti-VEGF agents) with unique regression and reactivation features post-treatment compared with laser photocoagulation, and (iv) recognition that patterns of ROP in some regions of the world did not neatly fit into the previous classification system.

Why is ROP such a promising area for AI-based medical devices?

First, there are unmet needs in ROP care worldwide, such as workforce challenges, subjectivity of clinical diagnosis, and significant medical legal liability. Second, there is an existing international ROP classification system so there is a standard method of clinical diagnosis that is used worldwide. Third, there is a standard approach to clinical management based on decades of multicenter collaborative clinical trials such as CRYO-ROP, ETROP, and BEAT-ROP. Finally, there are pediatric retinal imaging devices coupled with infrastructure to capture clinical diagnosis and outcome data. Taken together, these factors create a good environment for artificial intelligence implementation and evaluation research studies.



Source: https://www.canva.com/

What do we know about problems related to differences between graders?

Classically, we diagnose ophthalmic diseases by examining morphology of the eye. These clinical observations are typically qualitative, and we often convert those morphological observations into structured classifications (for example, “stage 1” or “plus disease” in ROP, “neovascularization elsewhere” in diabetic retinopathy, and similar). We and many others have shown that these clinical diagnoses and classifications are subjective, and that there are often significant differences, even among experts, in making these diagnostic distinctions. This is a fundamental challenge that limits the accuracy and consistency of clinical ophthalmic diagnosis.

Prof. Damien Gatinel

Anterior Segment and Artificial Intelligence

Damien Gatinel, Head of the Anterior Segment and Refractive Surgery Department, Rothschild Foundation Hospital, Paris, France

What are the differences between supervised and unsupervised learning in AI? Is the second option safe enough since we do not understand how the algorithm works?

In both cases, the prerequisite is identical; have a large amount of good quality data. In the case of supervised learning, we use labeled data, and we train an algorithm to classify data as inputs in the most efficient way or make predictions. In the pathology of the anterior segment, screening for keratoconus is an obvious application. To develop an effective algorithm, one must have training data from distinct groups (keratoconus versus normal cornea). Whatever the type of algorithm used (logistic regression, decision trees, neural networks), it will use input data for which its origin is clearly indicated. In the case of supervised learning, the approach is significantly different; the problem is usually to uncover hidden and unknown relationships present within a disparate dataset or to search for unknown patterns. It is less a question of predicting than discovering links between certain data that allow them to be grouped together, making it possible to classify large volumes of data. The algorithms reduce the dimensionality of the data entered in the system and estimate the distance in a smaller residual space between the data that one seeks to group. We used this process to evaluate the possibility of automatically classifying large volumes of topographic examinations, which can be of great interest for quickly finding specific categories (eyes operated on for refractive surgery, keratoconus, and similar). In any case, it is important to clear up confusion; if the algorithms are built according to a well-identified approach, the variables used to establish the groupings are not always easily identifiable. It is always necessary to be careful and have methods to limit the risk of overfitting and ensure that the system one develops remains generalizable.

The development of the PEARL-DGS Formula – the AI-based IOL calculation formula – is an important achievement. What are its parameters and how can it be used and tested today by practitioners?

The PEARL-DGS formula is based on an optical model using thick lens formulas, AI algorithms for the prediction of the anatomical position of the implant, and the curvature of the face posterior surface of the cornea (when this surface is not measured). The methods used correspond to supervised learning, made possible by obtaining a large set of quality data from pseudophakic eyes containing preoperative biometric data and the refractive result obtained. It also uses an axial length value per approximate or exact segment sum if the biometer provides this value. It makes it possible to consider a history of corneal refractive surgery and the results obtained for the first eye surgery to improve the precision of the power calculation. All the steps used to calculate the implant's power have been published and the code was deposited in an online directory. The formula is available under the following link: www.iolsolver.com.

Source: https://www.canva.com/

What will be the next AI-based applications in the anterior segment?

Anterior segment pathologies are a wide range of areas where one can consider using AI. The calculation of the implant's power is of course already part of the landscape, but we are working on using neural networks and high-resolution OCT images of the anterior segment to establish an objective diagnosis of lens opacity. For a more objective distinction between proven cataracts (visually significant opacities) and non-dysfunctional lenses in terms of transparency. The same process has also been used to characterize the presence of corneal stromal edema. In both cases, it is possible to obtain an enriched image where the probability of lens opacity or corneal edema is shown for each pixel. Other interesting applications include improving the adaptation of contact lenses on irregular or reshaped corneas by quickly predicting the parameters of the lens, better predicting the size of phakic implants from biometric and refractive data. It is also possible to envisage diagnostic aid devices based on image banks of the anterior segment using conventional image recognition methods. Finally, we were interested in predicting subjective refraction from the objective measurement of the wavefront and high-degree optical aberrations. These are just examples, but the limits are those of your imagination!

Dr. Paisan Ruamviboonsuk

On Transfer Learning, GANs, and More

Paisan Ruamviboonsuk, Clinical Professor of Ophthalmology, College of Medicine, Rangsit University, Assistant Hospital Director for Centers of Medical Excellence, Center of Excellence for Vitreous and Retinal Disease, Rajavithi Hospital, Bangkok, Thailand

What is transfer learning and why do you think it can bring benefits to healthcare and ophthalmology?

Transfer learning (TL) is a type of deep learning model made use of other, already available, deep learning (DL) models or other datasets. TL may be used for the easier development of a DL model or to improve the accuracy of a DL model. For example, many models today were developed from information transferred from ImageNet, which is an open-source model available on the internet. In ophthalmology, information from OCT datasets, for example, can be transferred to corresponded datasets of color fundus images (CFI) to develop a DL model for making analysis on CFI, which may provide better accuracy than traditional DL developed from only data from CFI. This is because the model learns more from both CFI and OCT datasets. The benefits would include more AI models developed, with better performance.

What are the limitations of traditional AI models?

Traditional AI models may require a very large dataset developed to achieve high enough performance. In addition, there are a lot of data available today from multimodal imaging in ophthalmology, but traditional AI may be able to make use of only one type of data at a time.

What are GANs and how can they help ophthalmologists?

GANs is a DL model developed to create new images from existing images, and therefore, GANs is a TL model in nature. There are many uses of GANs in movie and advertisement industries, for example when creating an image of a zebra from an image of a horse. In medicine, GANs are used to create images of a modality that may be less commonly used, such as MRI, from images of a modality that may be more commonly used, such as a CAT scan. The new images may be used in AI research or to guide clinicians in clinical studies. In our study, we used GANs to create ultrasound biomicroscopy (UBM) images of anterior segment from corresponded anterior segment OCT images to detect plateau iris. In another study, researchers used GANs to create fundus images to unlock the black box of DL. The researchers in this study developed a DL model to detect where the nerve fiber layer or optic disc neuroretinal rim was thinning in fundus images of glaucomatous eyes. They used GANs to create a fundus image in which that thin area had normal thickness and another image in which that area was extremely thin. These new images highlighted where in fundus images the DL model used for making diagnosis of glaucoma and ophthalmologists could use these images created by GANs to judge if the model pointed out the correct areas.

Prof. Michael D. Abràmoff

AI: Autonomous or Assistive

Michael D. Abràmoff, The Robert C. Watzke, Professor of Ophthalmology, Professor of Electrical and Computer Engineering, and Biomedical Engineering, Department of Ophthalmology and Visual Sciences, University of Iowa Hospital and Clinics, Iowa, USA

What is the difference between autonomous and assistive AI medical devices?

The term "assistive" is for AI systems where the clinician makes the ultimate medical decision, and the clinician (the user) is liable for the AI performance, while the term "autonomous" is reserved for those systems where the AI makes the ultimate medical decision, and it is the AI creator who carries the liability for the AI performance, not the user. If someone claims autonomy for an AI, the next question should be whether the liability lies with the user (1).

Who should be responsible for a potential mistake made by an AI medical device?

Along with my colleagues, we previously proposed that creators of autonomous AI systems assume liability for harm caused by the diagnostic output of the device when used properly and on label (2). The article states that this is essential for adoption: it may be inappropriate for clinicians using an autonomous AI to make a clinical decision they are not comfortable making themselves, to nevertheless have full medical liability for harm caused by that autonomous AI. This view was recently endorsed by the American Medical Association in its 2019 AI Policy. Such a paradigm for responsibility is more complex for assistive AI, where medical liability may fall only on the provider using it, because they are ultimately responsible for the medical decision, or on a combination of both, where even the relative balance of liability of the AI user and the AI creator come into play (3).

What are the major concerns regarding AI and how can they be addressed?

All stakeholders in the healthcare system have valid concerns about AI that need to be addressed. Stakeholders include patients, patient organizations, physicians and other providers, bioethicists, medicolegal experts, regulators such as US FDA and US FTC and Joint Commission, and payers such as CMS (Medicare and Medicaid) and private payers. Is there patient or population benefit, such as outcome improvement, from the use of the AI? I have called AI that is technologically cool but offers no patient benefit "glamour AI." Does it increase health disparities, or otherwise negatively affect some populations? Is there racial, ethnic or other bias in the safety or efficacy of the AI? Who is liable if something goes wrong? What happens with a patient's data when AI is used, and how is patient data used in development and usage?

Source: https://www.canva.com/

There may be other, not yet anticipated concerns out there. The only way to address these known and unknown concerns is with an ethical framework for AI, which starts with the basic millennia-old bioethical principles such as Autonomy, Justice, Beneficence and Non-maleficence, and Responsibility. By measuring how much a given AI system meets each of these bioethical principles, AI creators can build systems that address all concerns in a provable (falsifiable) manner; this is called metrics for ethics. I and others have published extensively on these subjects, including an ethical framework for AI that has itself been used to create regulatory consideration for AI with US FDA, and reimbursement considerations for US CMS and other payers, and these have all been applied successfully, leading to regulatory approval and reimbursement for autonomous AI in the US (1, 2, 3).

From the ethical framework, the following can be derived: AI technology also needs to be validated through a preregistered, peer reviewed clinical trial that is conducted in the intended clinical setting, with outcomes that meet or exceed all superiority endpoints. For example, IDx-DR exceeded all superiority endpoints at 87 percent sensitivity, 91 percent specificity, with a valid diagnostic result for 96 percent of subjects, and was proven to have no racial and ethnic bias all of which exceeds human specialist performance. These outcomes led to FDA clearance and helped establish trust with all

industry stakeholders facilitating the adoption of autonomous AI into the Standards of Care for diabetes, reimbursement through CPT code 92229, and widespread system adoption. The ultimate goal of AI advancement is to improve patient outcomes by increasing access, lowering costs, and improving the quality of care that is available to the people who need it most.

1. MD Abràmoff et al., "Foundational Considerations for Artificial Intelligence Using Ophthalmic Images," Ophthalmology, 129, e14 (2021). PMID: 34478784.

2. MD Abramoff et al., "Diagnosing Diabetic Retinopathy with Artificial Intelligence: What Information Should Be Included to Ensure Ethical Informed Consent?" Front Med (Lausanne), 8 (2021). PMID: 34901083.

3. DS Char et al., "Identifying Ethical Considerations for Machine Learning Healthcare Applications," The American Journal of Bioethics, 20, 7 (2020).

4. MD Abramoff et al., "Lessons Learned About Autonomous Ai: Finding a Safe, Efficacious, and Ethical Path through the Development Process," Am J Ophthalmol, 214, 314 (2020).

5. MD Abramoff et al., "A Reimbursement Framework for Artificial Intelligence in Healthcare," NPJ Digit Med, 5 (2022). PMID: 35681002.

What You Need to Know About AutoML…

By Tomasz Krzywicki, Data Scientist, Uniwersytet Warmińsko-Mazurski, Olsztyn, Poland

What is AutoML and how can it be used?

AutoML is a cloud computing software or service that aims to automatically produce predictive models that solve certain problems based on a transferred dataset. To use the AutoML tools, all we need is access to a computer with the appropriate software installed or access to the cloud computing that offers AutoML services. We should also remember to have a properly marked dataset for solving a given problem. The use of these tools is limited only to indicating the dataset and starting the process of finding optimal model architectures, which can be lengthy. With a bit of patience, we can use the created model in any form, for example deploy it in another server service or on any device, and analyze prediction metrics prepared with AutoML tools.

MEng. Tomasz Krzywicki

What are the available Auto ML types and how do they differ?

The most prominent division of AutoML tools is the computing environment. In the case of software installed on local computers, the calculations are performed on them. The analogy is with cloud computing services. It is worth noting that AutoML tools running on local computers may require vital hardware resources in the form of a graphics processor and a reasonable amount of operating memory, preferably at least 16 GB. This most apparent division is also related to the cost of these tools. AutoML software installed on local computers is mostly free, while cloud-based AutoML services come at a cost.

What is the cost of these ML devices?

The cost depends on factors such as the location of the server room where the computation is performed, the type and complexity of the problem being solved, the target location for model deployment, and the scale of computational resources used in the time spent on model creation. When planning the costs, the server space for storing the datasets should also be considered. Some services do not allow downloading the created models to the computer disk, but only deploy them in other server services, which involves additional costs. For example, one hour of the AutoML service running in Ohio, US, costs US$1 per node (virtual machine). However, cloud computing providers often offer a free trial period, which is enough to test the capabilities of AutoML services.

What are the major challenges to developing AI further in the near future?

Currently, the most popular method for creating intelligent systems is machine learning, which is a heuristic that involves fitting a mathematical function, or a group of mathematical functions, to a certain dataset to obtain an optimal solution in the form of predictions close to the labels in the dataset. Therefore, at present, artificial intelligence can learn certain patterns, but it cannot think, and it requires continuous monitoring. Some researchers believe that AI will soon come to the end of its development capabilities. Major players in the world of technology, however, are already researching an entirely new form of this field, taking inspiration from the human brain and, from my point of view, this is the biggest challenge for the near and more distant future.