

# Artificial Intelligence

AI is poised to revolutionize medicine. An overview of the field, with selected applications in ophthalmology.

By Linda Roach, Contributing Writer

**F**ROM THE BACK OF THE EYE TO THE front, artificial intelligence (AI) is expected to give ophthalmologists new automated tools for diagnosing and treating ocular diseases. This transformation is being driven in part by a recent surge in attention to AI's medical potential from big players in the digital world like Google and IBM. But, in ophthalmic AI circles, computerized analytics are being viewed as the path toward more efficient and more objective ways to interpret the flood of images that modern eye care practices produce, according to ophthalmologists involved in these efforts.

## Starting With Retina

The most immediately promising computer algorithms are in the field of retinal diseases. For instance, researchers from the Google Brain initiative reported in 2016 that their "deep learning" AI system had taught itself to accurately detect diabetic retinopathy (DR) and diabetic macular edema in fundus photographs.<sup>1</sup>

And AI is being applied to other retinal conditions, notably including age-related macular degeneration (AMD),<sup>2,3</sup> retinopathy of prematurity (ROP),<sup>4</sup> and reticular pseudodrusen.<sup>5</sup>

But retina is just the beginning: Researchers are developing AI-based systems to better detect or evaluate other ophthalmic conditions, including pediatric cataract,<sup>6</sup> glaucoma,<sup>7</sup> keratoconus,<sup>8</sup> corneal ectasia,<sup>9</sup> and oculoplastic reconstruction.<sup>10</sup>

"There's a whole spectrum, all the way from

screening to full management, where these algorithms can make things better and make things more objective. There are a lot of times when clinicians just disagree, but an AI system gives the same answer every time," said Michael D. Abràmoff, MD, PhD, a leading figure in the exploration of AI for ophthalmology who is at the University of Iowa in Iowa City.

## Where Do MDs Fit In?

One of the most common concerns clinical ophthalmologists and other physicians express about AI is that it will replace them. But Renato Ambrósio Jr., MD, PhD, who has been working on a machine learning algorithm to predict the risk of ectasia after refractive surgery,<sup>9</sup> said he encourages his colleagues to regard AI as just another tool in their diagnostic armamentarium.

"I use the tools developed by our group for enhanced ectasia susceptibility characterization in my daily practice," said Dr. Ambrósio, at the Federal University of São Paulo in São Paulo, Brazil. "It has allowed me to improve not only the sensitivity to detect patients at risk but also the specificity, allowing me to proceed with surgery in patients who may have been considered at high risk by less sophisticated approaches."

**High-powered software.** When these tools are ready for widespread clinical use, physicians won't need to become AI experts, because the software is likeliest to reside within devices like optical coherence tomography (OCT) machines, said Ursula

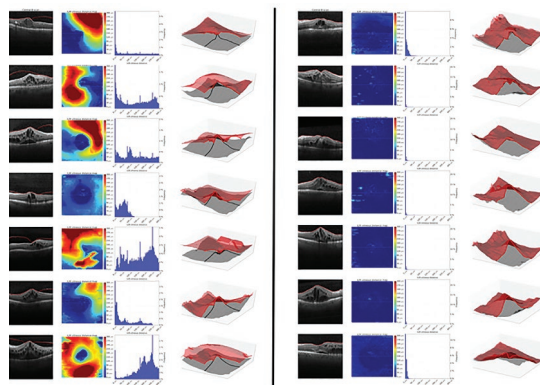
Schmidt-Erfurth, MD, at the Medical University of Vienna in Vienna, Austria. Her group is working on several AMD-analyzing algorithms.

“An automated algorithm is just a software tool, and ours are all based on routine OCT images—[using] the same OCT machines that are available in thousands and thousands of hospitals and private offices,” Dr. Schmidt-Erfurth said. “Ideally, this software would be built into each machine that is being sold, or eye doctors could buy it as an add-on.”

### AI Basics

Although the term artificial intelligence originated in the 1950s, the concept was still languishing on the fringes of computer science as recently as two decades ago, Dr. Abramoff said. He and others wanted to try to echo the human brain’s mechanisms with “neural networks,” but the available computers could not handle the complexity, he said. Furthermore, anyone who talked about neural networks at that time was regarded as a little crazy, Dr. Abramoff said. As a result, he and other researchers started working on expert systems and automated image analysis.

Today, there are a variety of approaches to building AI systems to automatically detect and measure pathologic features in images of the eye.



**3D OCT.** Dr. Schmidt-Erfurth’s group developed a fully automated segmentation algorithm for the posterior vitreous boundary. These images are of the boundary in patients with (left panel) and without (right panel) vitreomacular adhesion.

The labels are sometimes used interchangeably; all of them in some way analyze pixels and groups of pixels in fundus photographs, or 3-dimensional “voxels” in OCT images.

**Simple automated detectors.** In the simplest form of AI, programmers give the software mathematical descriptors of the features to detect, and a rules-based algorithm looks for these patterns on incoming images (“pattern recognition”). Positive “hits” are combined to produce a diagnostic indicator.

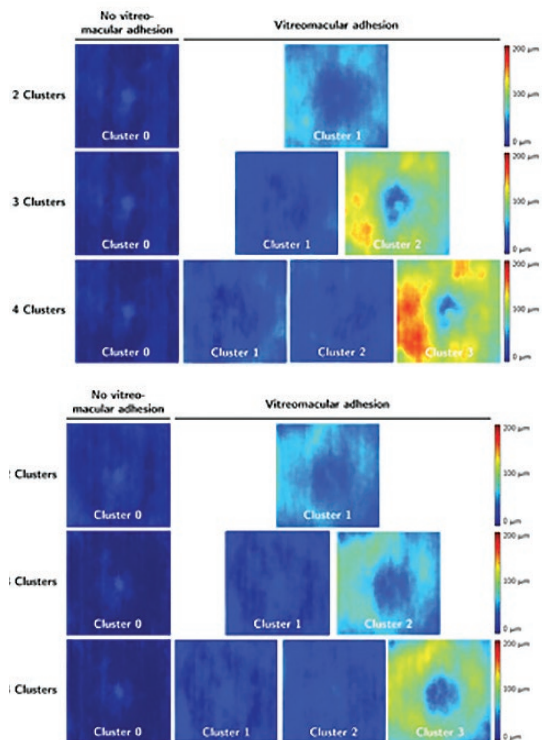
“Because there might be multiple types of lesions you are looking for, you have many lesion detectors, and then you combine them into a diagnostic output, saying this patient is suspected of having [a specific] disease,” Dr. Abramoff said.

**Basic machine learning.** Early on, researchers realized “that it was hard to write rules that tell a computer algorithm how to ‘see,’” Dr. Abramoff said. As a result, they turned to machine learning.

In this approach, the algorithm is given some basic rules about what disease features look like, along with a “training set” of images from affected and unaffected eyes. The algorithms examine the images to learn about the differences.

The earliest machine learning systems resembled large-scale regression analysis. Initially, operators adjusted the analytic parameters to improve the system’s performance. Later, machine learning algorithms were instructed to improve their accuracy with “neighbor networks”—that is, by looking at neighboring pixels or groups of pixels to judge whether together they indicated disease.

**Advanced machine learning.** This type of machine learning structure consists of one or two interconnected layers of small computing units called “neurons,” which mimic the multilayered structure of the visual cortex.



**AUTOMATED ASSESSMENT.** Unsupervised clustering of vitreomacular interface configurations in patients with (top panel) branch retinal vein occlusion and (bottom panel) central RVO.

The inputs to the first layer are the same disease feature detectors used in basic machine learning. However, the outputs of the neurons in each layer feed forward into the next layer, and the final layer yields the diagnostic output. Thus, the neural network learns to associate specific outputs of disease feature detectors with a diagnostic output.

**Deep learning with convolutional neural networks (CNNs).** The term “deep learning” is used because there are multiple interconnected layers of neurons—and because they require new approaches to train them. This latest iteration of AI comes closer to resembling “thinking,” because CNNs learn to perform their tasks through repetition and self-correction.

A CNN algorithm teaches itself by analyzing pixel or voxel intensities in a labeled training set of expert-graded images, then providing a diagnostic output at the top layer. If the system’s diagnosis is wrong, the algorithm adjusts its parameters (which are called weights and which represent synaptic strength) slightly to decrease the error. The network does this over and over, until the system’s output agrees with that of the human graders.

This process is repeated many times for every image in the training set. Once the algorithm optimizes itself, it is ready to work on unknown images.

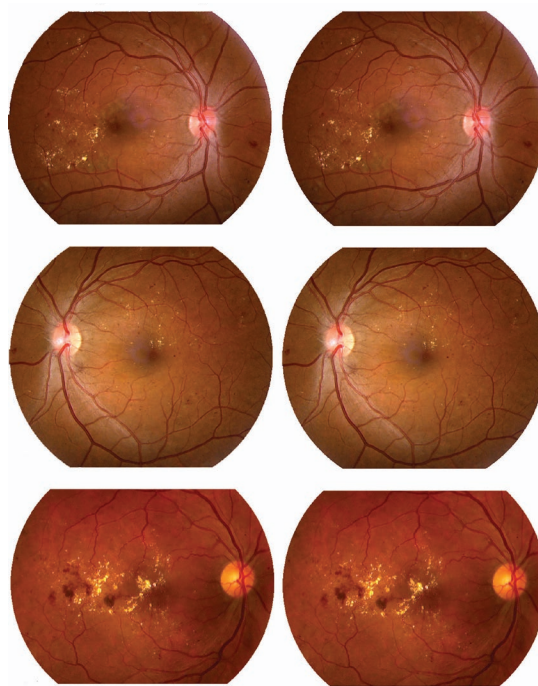
**Disease feature-based versus image-based (“black box”) learning.** Many ophthalmic AI researchers prefer to design their machine learning algorithms based on clinically known characteristics of disease, such as hemorrhages or exudates. So, when a supervised learning algorithm works, scientists can verify that its output is based on the presence of the same image characteristics that a human would identify, and they can adjust the algorithm if necessary.

However, the successful system that Google Brain reported in 2016 was an example of an unsupervised black box system—an approach that unsettles some MDs and intrigues others.

Google’s deep learning algorithm taught itself to correctly identify diabetic lesions in photographs even though it was not told what the lesions look like, said Peter A. Karth, MD, MBA, a vitreoretinal subspecialist from Eugene, Oregon, who is a consultant to the Google Brain project. “What’s so exciting with deep learning is we’re not actually yet sure what the system is looking at. All we know is that it’s arriving at a correct diagnosis as often as ophthalmologists are,” he said.

### Research Spotlight: DR

Computerized algorithms for detection and management of DR are the main focus for many teams



**FALSE NEGATIVES?** Dr. Abràmoff and his colleagues selected disease images (left column), minimally altered them via the process known as adversarialization, then gave the altered images (right column) to image-based black box systems to evaluate. While clinicians correctly identified the altered ones as DR images, the image-based systems were more likely to classify them as normal.

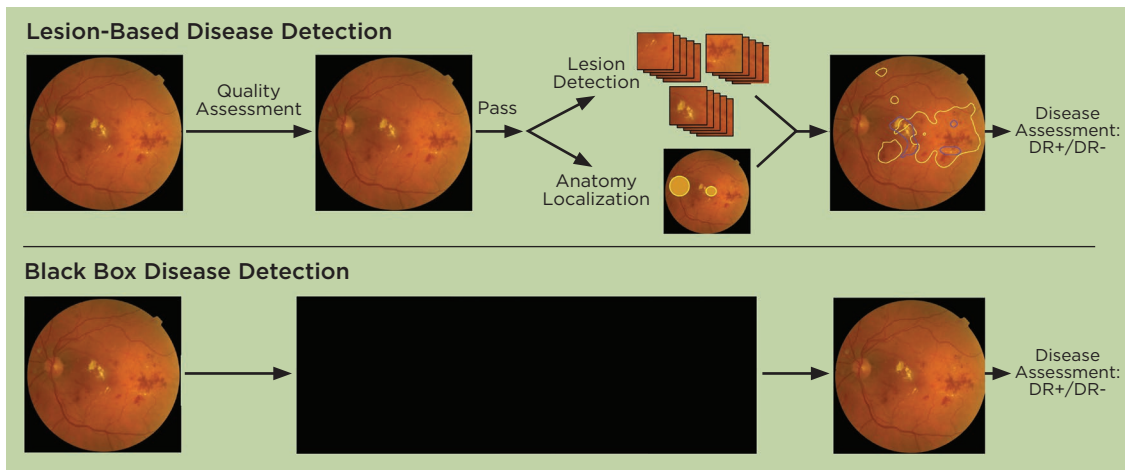
of ophthalmologists, computer scientists, and mathematicians around the world.

**An outgrowth of telemedicine.** Telemedicine for DR helped lay the groundwork for AI, said Michael F. Chiang, MD, at Oregon Health & Science University in Portland.

In a typical telemedicine setup, patient data and photographs are collected at a primary care clinic and then sent to an ophthalmologist elsewhere for interpretation, Dr. Chiang said. “It’s not that big of an extrapolation to say that, instead of feeding the fundus photographs to a human expert somewhere else, the clinic could feed them to an AI machine and figure out whether this patient needs to see an eye doctor or not. The algorithms could potentially provide an initial level of disease screening,” he said.

**A hybrid screening algorithm.** Over the course of the last decade, Dr. Abràmoff’s group developed a proprietary DR screening algorithm, called IDx-DR (IDx), which it is commercializing in partnership with IBM Watson Health in Europe. European physicians have been able to access a secure server-based version since 2014.

When the secure server-based software was released, IDx-DR consisted of a supervised machine



**NEURAL NETS.** How do unsupervised deep learning systems (bottom) know what they know and come to conclusions? This opacity troubles many AI researchers.

learning algorithm that relied on mathematical descriptors to recognize lesions. More recently, the designers sought to improve the algorithm’s performance by adding CNNs—and it worked.<sup>11</sup>

“Before this, we had been using machine learning to combine outputs of all the different detectors into diagnostic outputs. We wanted to see if—instead of using mathematical equations as detectors—we could use small CNNs to find these lesions. And that led to a statistically significant improvement in performance,” Dr. Abramoff said.

**Sensitivity and specificity.** The hybrid system’s

sensitivity (the primary measure of safety) was 96.8% (95% confidence interval [CI]: 93.3% to 98.8%). This was not significantly different from that of the previously published results with the unenhanced algorithm (CI of 94.4% to 99.3%), the scientists reported.<sup>11</sup> But the specificity level was much better: 87.0% (95% CI: 84.2% to 89.4%, vs. a CI of 55.7% to 63.0% previously).

This higher specificity is important, because IDx envisions its AI algorithm as a screening tool, and high specificity means fewer false-positives, Dr. Abramoff said. Patients flagged by IDx could

## What Will AI Mean for You?

**Integrating details from thousands of images.** “The retina is the perfect playground for AI, because of the availability of excellent high-resolution retinal scanners and because of the retinal structure.

“The retina has an anatomy that is much like the brain, with many structural neuro-sensory layers. This delicate morphology requires precise high-resolution diagnosis based on imaging.

“[The process of detecting] details in this microenvironment in thousands and thousands of scans—this is a typical task for pattern recognition, which is exactly what AI is doing.”

—Dr. Schmidt-Erfurth

**Help for overburdened MDs.**

“You can imagine that individual ophthalmologists might be a little worried that this is going to reduce the [numbers of] patients in their clinics. But we think it’ll go the other way, by bringing more people into the health care system. So those patients who are slipping through the cracks—AI is going to make them aware that they have a problem and get them into doctors’ offices. It will help overburdened physicians get more of the patients they need to see and fewer of the ones they don’t.”

—Dr. Karth

**Adding consistency to subjective diagnoses.** “Ophthalmic

disease is often very subjective and very qualitative. We basically look at the eye and describe what we see.

“Classically we describe what we see by using words, images, and drawings. Then we assign diagnoses to these words, images, and drawings—and follow them over time.

“Something we’ve seen in ROP—which has been seen in many ophthalmic diseases—is that when it comes to looking at things that are very qualitative and subjective, even experts are not very consistent. So, one reason I’m interested in AI is to use computing systems to analyze data and make diagnoses more consistently. Because ophthalmology

then be referred to an ophthalmologist for manual confirmation of the automated diagnosis and for possible treatment, he said. He added, “IDx-DR is still investigational in the United States until FDA clearance, which we are hoping to obtain after they evaluate the results from our recently completed FDA pivotal study.”

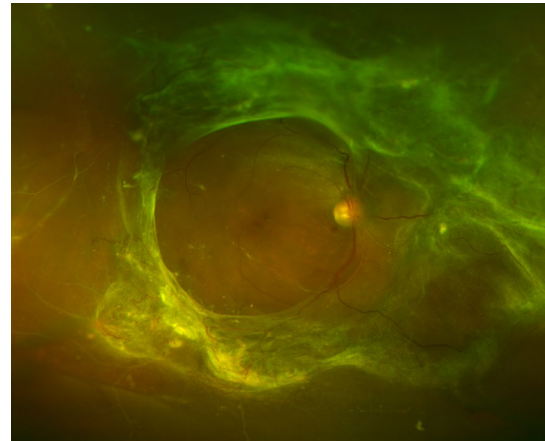
### Research Spotlight: AMD

In Austria, Dr. Schmidt-Erfurth assembled a computational imaging research team for pragmatic reasons. Expert OCT graders at her department’s Vienna Reading Center were being overwhelmed by the task of manually grading images from a series of large international clinical trials of anti-AMD drugs, she said.

“When the better OCT instruments came, with thousands and thousands of scans, it was pretty clear that manual analysis was not possible anymore,” she said.

**True breakthroughs?** Their efforts have produced deep learning algorithms that she believes constitute “true breakthroughs” in the evaluation and treatment of eyes with AMD.

**Monitoring therapy’s effectiveness.** “First, we have developed algorithms that can not only recognize disease activity on OCT scans but also can assess this activity<sup>2</sup> precisely,” Dr. Schmidt-Erfurth said. “Each time we see a patient we can say ‘yes’ or ‘no’ [that] there is fluid there. We also can quantify



**TELEMEDICINE TO AI.** Using telemedicine to transmit images of DR (shown here) has helped lay the groundwork for AI.

the amount of fluid, to determine whether the disease is more active or less active than before. This is unique in AMD.”

**Predictions based on drusen.** The second breakthrough is an AI algorithm for making individualized predictions about eyes with drusen underneath the retina, the earliest stage of AMD, Dr. Schmidt-Erfurth said. The algorithm does this by quantifying drusen volume on OCT and tracking how the volume changes over time.<sup>12</sup>

“Our algorithm can predict the course of the disease. It can identify exactly which patients will

is heavily image-based, a lot of that boils down to the use of computers for analyzing images and trying to quantify various features in the images.

“I would argue that this is one potential benefit of AI. By putting numbers on qualitative images, these systems could potentially give ophthalmologists more information about whether the patient is getting worse or getting better or staying the same over time.”

—Dr. Chiang

**A better means for predicting ectasia.** “The need for an enhanced test to screen for ectasia risk prior to refractive surgery is supported by the fact that we have cases that develop ectasia with no identifiable risk factors. On the other

hand, we have cases in which the cornea is stable long after LASIK despite the presence of 1 or more risk factors. We need to improve the sensitivity and the specificity of our approach for screening patients.

“I have integrated corneal tomography and biomechanics into my practice since 2004. Learning how to read those data is a challenge for the clinician [especially considering the volume of data]. This makes it very hard for a human being to make a decision.

That is why I started exploring new methods to analyze and integrate the data to allow for more accurate clinical decisions. Machine learning techniques have been fundamental for such an approach.”

—Dr. Ambrósio

**Machines might notice what ophthalmologists can’t.** “Ophthalmologists do a very good job at diagnosing diabetic retinopathy. But there are tasks—and this is one—that machines may be better at: highly repetitive, very detailed tasks, where ideally you look at every single square millimeter of the retina, every pixel in the image.

“There also are minor or overlooked features of diabetic eyes that ophthalmologists don’t often look for.

“[These features] are so minute, so rare, that we just don’t have our radar up for them. But a deep learning system could be looking at those minute, rare features on every image every time.”

—Dr. Karth



**INTEREXPERT VARIABILITY.** AI may help provide greater diagnostic clarity and promote increased agreement among experts in certain diseases, notably ROP (shown here).

develop which type of advanced disease, whether it may be wet or dry [AMD], and it allows us to identify at-risk patients just by using the non-invasive in vivo imaging. You can do it precisely for each individual, and it's all based on routine OCTs," she said.

**Individualizing treatment intervals.** Moreover, the group has developed an algorithm that can make an individualized prediction of recurrence intervals after intravitreal injection therapy for neovascular AMD.<sup>3</sup> This information can help physicians avoid over- or undertreatment when the therapy is provided on an as-needed basis. The algorithm bases its predictions on the subretinal fluid volume in the central 3 mm during the first 2 months after therapy initiation, and it has an accuracy of 70% to 80%.<sup>3</sup>

**Next up: A hunt for new biomarkers.** Dr. Schmidt-Erfurth said that, as scientists refine and study their deep learning algorithms, she expects that her group's algorithms will discover new, previously unsuspected biomarkers that will help ophthalmologists treat patients. This is because deep learning systems notice details that are not readily apparent to the human eye, she said.

"These algorithms are not limited to what we as traditional clinicians believe is a pathological feature," she said. "They are searching by themselves and identifying entirely new biomarkers. And this unsupervised learning will really help us to understand disease beyond the already conventional knowledge."

### Current Limitations

The groundswell of research interest in AI can't mask the fact that the field is grappling with some significant challenges.

**Quality of the training sets.** If the training set of images given to the AI tool is weak, the software is unlikely to produce accurate outcomes. "The systems are only as good as what they're told. It's important to come up with robust reference standards," Dr. Chiang said.

Dr. Abramoff agreed. "You need to start with datasets that everyone agrees are validated. You cannot just take any set from a retinal clinic and say, well, here's a set of bad disease and here's a group of normal," he said.

**Problems with image quality.** "The state-of-the-art systems are very good at finding diabetic eye disease. But one thing they're not very good at recognizing is when they're not seeing diabetic eye disease. For example, these systems will often get confused by a patient who has a central retinal vein occlusion instead of diabetic retinopathy," Dr. Chiang said.

He added, "Another challenge is that a certain percentage of images aren't very good. They're blurry or don't capture enough of the retina. It's really important to make sure that these systems recognize when images are of inadequate quality."

**The black box dilemma.** When a CNN-based system analyzes a new image or data, it does so based upon its own self-generated rules. How, then, can the physician using a deep learning algorithm really know that the outcome is correct? This is the "black box" problem that haunts some medical AI researchers and is downplayed by others, Dr. Abramoff said.

**Wrong answers.** Dr. Abramoff concocted an experiment that he believes illustrates why there is reason for concern. His team changed a small number of pixels in fundus photographs of eyes with DR and then gave these "adversarial" images to image-based black box CNN systems for evaluation. The changes in the images were minor, undetectable to an ophthalmologist's eye. However, when these CNNs evaluated the altered images, more than half the time they judged them to be disease-free, Dr. Abramoff said.<sup>13</sup>

"To any physician looking at the adversarial photo it would still look like disease. But we tested the images with different black box CNN algorithms and they all made the same mistake," he said. "So, it's easy for this type of algorithm to make these kinds of mistakes, and we don't know why that is the case. I believe feature-based algorithms are much less prone to these mistakes."

### Looking Ahead

Despite these challenges, it's clear that AI will occupy an increasingly critical role in medicine.

**A valuable research tool.** "There is definitely a huge role for neural networks in research, for hy-

pothesis generation and discovery,” Dr. Abràmoff said. “For instance, to find out whether associations exist between some retinal disease and some image feature, such as in hypertension. There, it does not matter initially that the neural network cannot be fully explained. Once we know an association exists, we then explore what the nature of that association is.”

**Augmentation, not replacement, of MDs.** Dr. Chiang, who is helping to develop AI techniques to assess ROP, said that he believes automated systems can and should complement what physicians do.

“Machines can help the doctor make a better diagnosis, but they are not good at making medical decisions afterward,” he said. “Doctors and patients make management decisions by working together to weigh the various risks and benefits and treatment alternatives. The role of the doctor will continue to [involve] the art of medicine—which is a uniquely human process.”

- 1 Gulshan V et al. *JAMA*. 2016;316(22):2402-2410.
- 2 Klimesha S et al. *Invest Ophthalmol Vis Sci*. 2017;58(10):4039-4048.
- 3 Bogunovic H et al. *Invest Ophthalmol Vis Sci*. 2017;58(7):3240-3248.
- 4 Campbell JP et al. *JAMA Ophthalmol*. 2016;134(6):651-657.
- 5 van Grinsven MJ et al. *Invest Ophthalmol Vis Sci*. 2015;56(1):633-639.
- 6 Liu X et al. *PLoS One*. 2017;12(3):e0168606.
- 7 Yousefi S et al. *Transl Vis Sci Technol*. 2016;5(3):2.
- 8 Ruiz Hidalgo I et al. *Cornea*. 2017;36(6):689-695.
- 9 Ambrósio R Jr. et al. *J Refract Surg*. 2017;33(7):434-443.
- 10 Tan E et al. *J Eur Acad Dermatol Venereol*. 2017;31(4):717-723.
- 11 Abràmoff MD et al. *Invest Ophthalmol Vis Sci*. 2016;57(13):5200-5206.
- 12 Schlanitz FG et al. *Br J Ophthalmol*. 2017;101(2):198-203.
- 13 Lynch SK et al. Catastrophic failure in image-based convolutional neural network algorithms for detecting diabetic retinopathy. Presented at: ARVO 2017 Annual Meeting; May 10, 2017; Baltimore.

## Further Reading

Abràmoff MD. **Image processing.** In: Schachat AP, Wilkinson CP, Hinton DR, Sadda SR, Wiedemann P, eds. *Ryan's Retina*. 6th ed. Amsterdam: Elsevier. 2017;196-223.

Abràmoff MD et al. **Retinal imaging and image analysis.** *IEEE Rev Biomed Eng*. 2010;3:169-208.

Lewis-Krauss G. **The great AI awakening.** *The New York Times*, Dec. 14, 2016.

Mukherjee S. **The algorithm will see you now.** *The New Yorker*. April 3, 2017.

## Meet the Experts



**Michael D. Abràmoff, MD, PhD** Retina specialist and the Robert C. Watzke, MD, Professor of Ophthalmology and Visual Sciences, Electrical and Computer Engineering, and Biomedical Engineering in the Department of Ophthalmology and Visual Sciences at the University of Iowa and the VA Medical Center, both in Iowa City; founder and president of IDx in Iowa City. *Relevant financial disclosures:*

*Alimera Biosciences: S; IDx: E,O,P; NEI: S; Research to Prevent Blindness: S.*



**Renato Ambrósio Jr., MD, PhD** Director of cornea and refractive surgery at the Instituto de Olhos Renato Ambrósio, and associate professor at the Federal University of São Paulo in São Paulo, Brazil. *Relevant financial disclosures:*

*Alcon: C; Allergan: L; Carl Zeiss: L; Mediphacos: L; Oculus: C.*



**Michael F. Chiang, MD** The Knowles Professor of Ophthalmology & Med-

ical Informatics and Clinical Epidemiology at Oregon Health & Science University in Portland, Oregon, and a member of the Academy's IRIS Registry Executive Committee. *Relevant financial disclosures:* *Clarity Medical Systems: C (unpaid); Novartis: C; NIH: S; National Science Foundation: S.*

**Peter A. Karth, MD, MBA** Vitreoretinal subspecialist at Oregon Eye Consultants in Eugene, Oregon, and physician consultant to Google Brain. *Relevant financial disclosures:* *Google Brain: C; Carl Zeiss Meditec: C; Peregrine Surgical: C; Visio Ventures: Cofounder.*



**Ursula Schmidt-Erfurth, MD** Professor and chair of ophthalmology at the Medical University of Vienna in Vienna, Austria. *Relevant financial disclosures:* *Alcon/Novartis: C; Bayer Healthcare Pharmaceuticals: C; Boehringer Ingelheim Pharmaceuticals: C; Genentech: C.*



**See disclosure key, page 12.**