Google's AI program: Building better algorithms for detecting eye disease

March 13 2018

Credit: AI-generated image ([disclaimer](#))

The ability of artificial intelligence (AI) to help screen patients for a common diabetic eye disease gains momentum with a new study published online today in *Ophthalmology*, the journal of the American Academy of Ophthalmology. Lily Peng, M.D., Ph.D., and her colleagues at Google AI research group, show that they could improve their disease

detecting software by using a small subset of images adjudicated by ophthalmologists who specialize in retinal diseases. The specialists' input was then used to improve their computer's performance so that it was roughly equal to that of individual retinal specialists.

More than 29 million Americans have diabetes, and are at risk for diabetic retinopathy, a potentially blinding eye disease. People often don't notice changes in their vision in the disease's early stages. But as it progresses, diabetic retinopathy usually causes vision loss that in many cases cannot be reversed. That's why it's so important that people with diabetes have yearly screenings.

In earlier research, Dr. Peng and her team used neural networks—complex mathematical systems for identifying patterns in data—to recognize diabetic retinopathy. They fed thousands of retinal scans into these neural networks to teach them to "see" tiny hemorrhages and other lesions that are early warning signs of retinopathy. Dr. Peng showed the software worked roughly as well as human experts.

But Dr. Peng is interested in developing a system that would be good enough for her grandmother. So, to improve the accuracy of the software, she included the input of retina specialists, ophthalmologists who specialize in diseases of the retina.

"For my grandma, I would love to have a panel of subspecialists who actually treat the disease, to sit and debate her case, giving their opinion," Dr. Peng said. "But that is really expensive and it's hard to do. So how do you build an algorithm that gets close to this?"

To tease out how this could be done, Dr. Peng compared the performance of the original algorithm with manual image grading by either a majority decision of three general ophthalmologists, or a consensus grading by three retinal specialists.

The grading of diabetic retinopathy can be a complex process that requires the identification and quantification of fine features such as small aneurysms and hemorrhages. As a result, there can be a fair amount of variability among physicians examining images, looking for disease.

The retina specialists graded the images separately, then worked together to resolve any disagreements. Their review and subsequent consensus diagnosis offered considerable insight into the grading process, helping to correct errors such as artifacts caused by dust spots, distinguishing between different types of hemorrhages, and creating more precise definitions for "gray areas" that make it difficult to make a definitive diagnosis. At the end of the process, the retina specialists indicated that the precision used in the decision process was above that typically used in everyday clinical practice.

Using these specialist-graded images, Dr. Peng could then fine-tune the software, which improved their model's performance and improved detection of disease.

"We believe this work provides a basis for further research and raises the bar for reference standards in the field of applying machine learning to medicine," Dr. Peng said.

  More information: Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy, DOI: 10.1016/j.ophtha.2018.01.034 , www.aaojournal.org/article/S01 … (17)32698-2/fulltext

Provided by American Academy of Ophthalmology