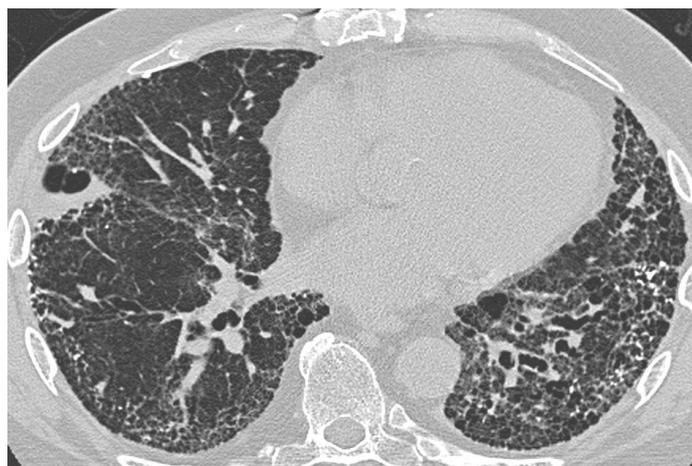**Intelligent Machines**

Million-Dollar Prize Hints at How Machine Learning May Someday Spot Cancer

Chinese researchers have developed an algorithm that could help make lung cancer diagnosis less error-prone.

by Will Knight   May 9, 2017

2

f

𝕏

reddit

Machine learning often requires massive data sets to develop an effective algorithm, but for this contest, teams

✉

were provided with only 2,000 images.

A contest aimed at automating the detection of lung cancer shows how machine learning may be poised to overhaul medical imaging.

The challenge offered $1 million in prizes for the algorithms that most accurately identified signs of lung cancer in low-dose computed tomography images. The winning algorithms won't necessarily be adopted by clinicians, but they could inspire algorithmic innovations that find their way into medical imaging.

Low-dose CT scans have shown great potential in recent years for detecting lung cancer earlier. They use less radiation and do not require a contrast dye to be injected into the body. But diagnosis is very difficult, meaning a high number of false positives and too many unnecessary medical procedures.

A machine-learning technique known as deep learning has proven especially effective for finding patterns in images

in recent years (see "10 Breakthrough Technologies 2013: Deep Learning"). There is now growing hope that this and other machine-learning methods may help improve standards of diagnosis in medicine by automatically recognizing patterns that indicate disease—including ones that are too subtle for the human eye to catch.

Deep learning has already been used to detect skin cancer in images with roughly the same number of errors as made by professional dermatologists. And the technique has proven effective for detecting a common cause of blindness in retinal images. There is now growing interest, among doctors and entrepreneurs, in deploying the technique more broadly. As this happens, however, more effort may be needed to make such algorithms explainable (see "The Dark Secret at the Heart of AI").

Related Story

Keyvan Farahani, a program director at the National Cancer

The Dark Secret at the Heart of AI

No one really knows how the most advanced algorithms do what they do. That could be a problem.

Institute, which supplied the imaging data used in the contest, says reducing the number of false lung cancer diagnoses made from low-dose CT scans would make a real difference for patients. There are about 222,500 new cases of lung cancer in the U.S. each year, according to the American Cancer Society.

Farahani says existing software for identifying signs of lung cancer are unreliable. "Preliminary results suggest [the top algorithms] are better than what's available already," he says. Farahani does not foresee algorithms taking the place of medical experts, though. "Deep learning will help digest large amounts of data," he says. "I don't think they're going to replace doctors or radiologists."

One of the

Subscribe to **Weekend Reads**

challenges in this contest was the fact that only 2,000 images were made available to teams. Machine learning often requires very large data sets in order to develop an effective algorithm. But other data, like details of the equipment used, were included.

The winning team employed a neural network and put extra effort into annotating images to provide more data points. It also used an additional data set, and broke the challenge into two parts: identifying nodules and then diagnosing cancer. It isn't yet clear how the best algorithm might measure up to a doctor, because each algorithm provides a probability rather than a definitive outcome.

"We think that explicitly

Should

dividing this problem into two stages is critical, which seems also to be what human experts would do," says Zhe Li, a member of the winning team and a student at Tsinghua University, one of China's foremost academic institutes.

algorithms help diagnose disease?
Share your thoughts.

Besides hinting at the potential for deep learning in medical imaging, the lung cancer contest highlights the growing reputation of Chinese AI researchers.

The contest, held on the data science site Kaggle, was organized by Booz Allen Hamilton, a management consulting firm that has arranged several other major data science contests before. The $1 million in prize money came from the Laura and John Arnold Foundation.

Kaggle was founded in 2010 and acquired earlier this year by Google. The site has proven to be a powerful

Recommended for You

01   A Single Autonomous Car Has a Huge Impact on Alleviating Traffic

02   Nvidia CEO: Software Is Eating the World, but AI Is Going to Eat Software

way of crowdsourcing the development of machine-learning algorithms, and is also a popular way to identify talent.

Read Next

Deep Learning

With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.

Josh Sullivan, who leads the data science team at Booz Allen Hamilton, says one motivation for the contest is talent acquisition, noting that 238 entrants have also applied for jobs at the company. He adds that the company is making the winning algorithms available for free to maximize the potential benefits to the medical community.

Li, of the winning team, says developing something that might save people's lives is gratifying, but the real reason for taking part was a bit less altruistic. "To be honest, the major motivation is to win the prize money," he says.

Get stories like this before