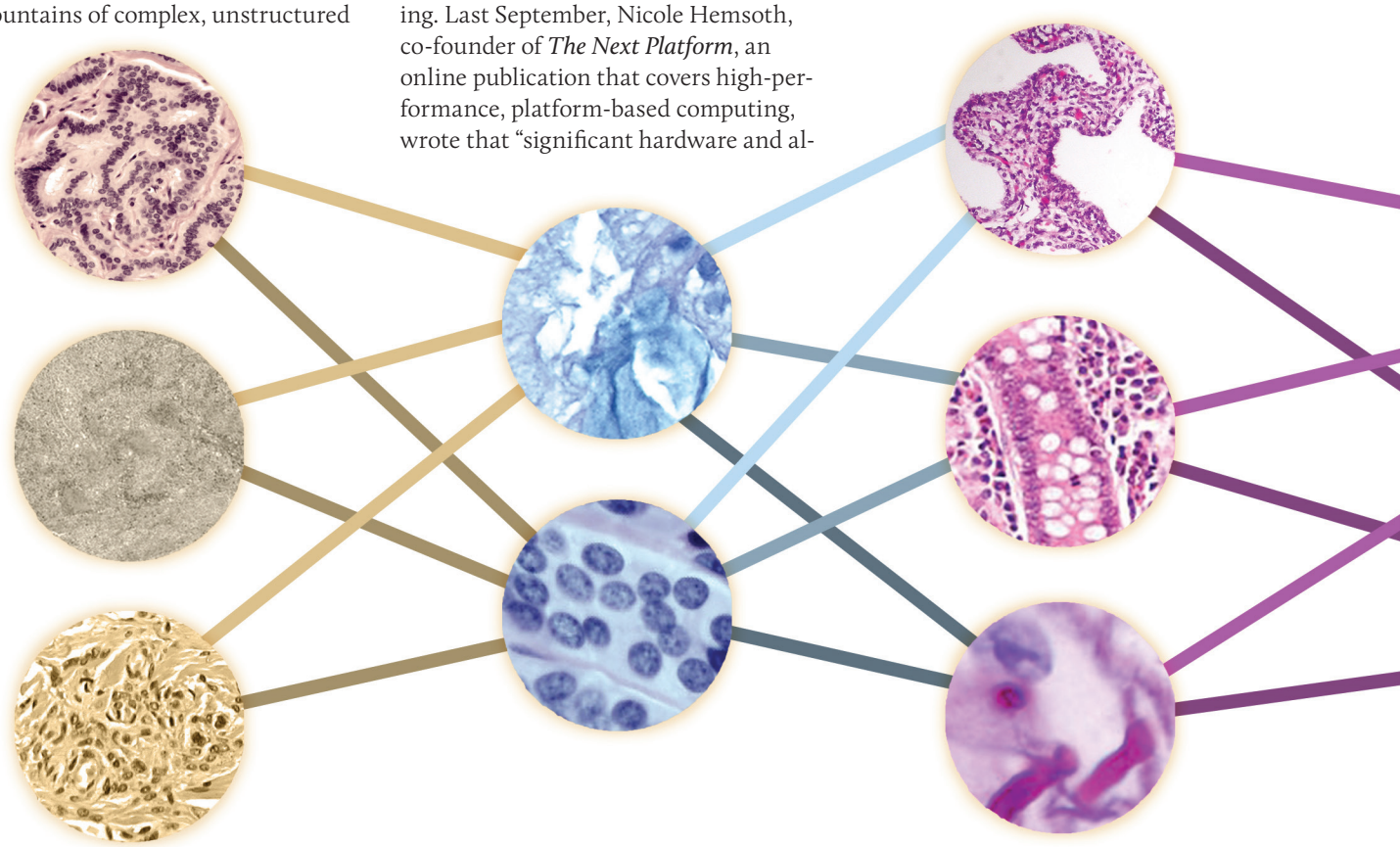
AI's Forthcoming Transformation of Medicine

Life-altering changes — in the form of new products and services — are on the horizon as researchers rush to apply deep-learning technology to solve a wide range of complex, and previously elusive, medical problems.

By Jon Reisfeld, SWE Contributor

A new age of medicine is dawning fast, as deep learning takes hold. Deep learning, a smarter, more versatile, and far more powerful version of artificial intelligence (AI) than its predecessor, machine learning, excels at making sense of mountains of complex, unstructured data, the kind modern medicine creates by the exabyte. Everyone, from well-heeled tech giants to small start-up firms, is racing to exploit its impressive data-mining and predictive capabilities.

The pace of advances has been dizzying. Last September, Nicole Hemsoth, co-founder of *The Next Platform*, an online publication that covers high-performance, platform-based computing, wrote that "significant hardware and algorithmic developments" in the prior 12 months were "propping up what appears to be an initial Cambrian explosion of new applications for deep-learning frameworks in areas as diverse as energy, medicine, physics and beyond."



What is the science driving these innovations? Machine learning and deep learning are two forms of AI modeled on how the human brain works. AI enables computers to learn without being explicitly programmed. The computers use algorithms, instead of groups of neurons, to separately analyze and then compare patterns in data. Given enough examples of inputs and desired outputs, these systems can learn to accurately predict new results from new information.

The primary difference between machine learning and deep learning is a matter of degree. Machine learning processes information with no more than two algorithms. Deep learning, theoretically, has no limits. Some deep- learning systems use 1,000 algorithms or more. The more algorithms used, the greater the degree of complexity the thinking network can handle.

Machine learning's computing power was limited, initially, to the number-crunching ability of a single central processing unit (CPU) – the type of processor that runs a typical laptop or work station. But in 2012, graphic processing units (GPUs), the kind that run data-intense video games, became available in parallel and expandable configurations. That faster, more robust computing power led to the development of deep-learning networks that could, just as quickly, handle a virtually unlimited number of algorithms.

This article will give you a feel for what's possible, by offering a close-up look at three AI medical-related projects.

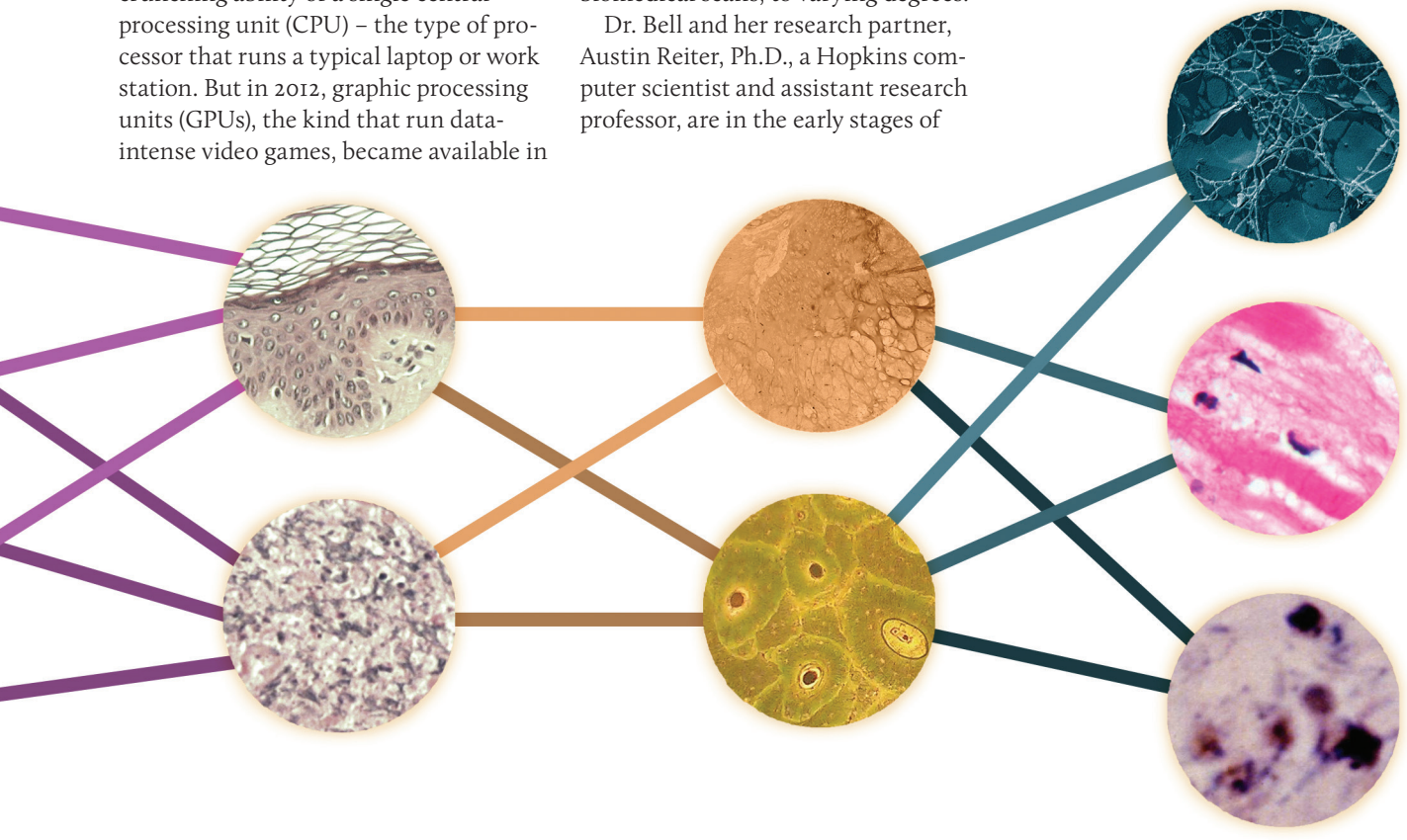
LESS IS MORE? A NEW ROLE FOR AI IN MEDICINE

Biomedical engineer and Johns Hopkins University Assistant Professor Muyinatu Bell, Ph.D., whose name appeared on *MIT Technology Review's* 2016 list of the world's top "35 Innovators Under 35," is at it again — innovating. This time, Dr. Bell, who founded and runs Hopkins' new Photoacoustic and Ultrasonic Systems Engineering (PULSE) Lab, has turned her attention to eliminating the noise, acoustic clutter, and reflective artifacts that consistently plague ultrasound and photoacoustic biomedical scans, to varying degrees.

Dr. Bell and her research partner, Austin Reiter, Ph.D., a Hopkins computer scientist and assistant research professor, are in the early stages of training their machine-learning model to perform an odd assignment: They want it to identify and extract noise, rather than useful information, from a biomedical scanner's raw signal feed before sending it onto the image display.

"The machine-learning network will learn what images should look like by studying a large body of simulated images that we create," Dr. Bell explained. "And then, it will detect and remove artifacts when they appear."

The smarter the network gets, the clearer the final scans should look. In effect, Drs. Bell and Reiter will be training the network to act as a smart signal processor, bypassing the signal processors supplied by scanner manufacturers. That's necessary, Dr. Bell explained, because both ultrasound and photoacoustic imaging processors rely on a shared set of faulty assumptions that will inevitably produce random noise, acoustic clutter, and reflective artifacts under the right conditions.



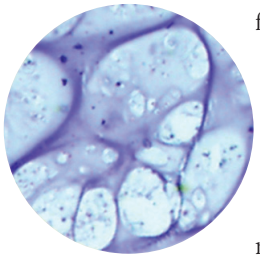
The first faulty assumption is that sound waves originating in the body will always travel directly from their source to an externally placed transducer, without encountering any reflective bodies, such as bones, along the way. The assumption doesn't allow for the possibility of echoes.

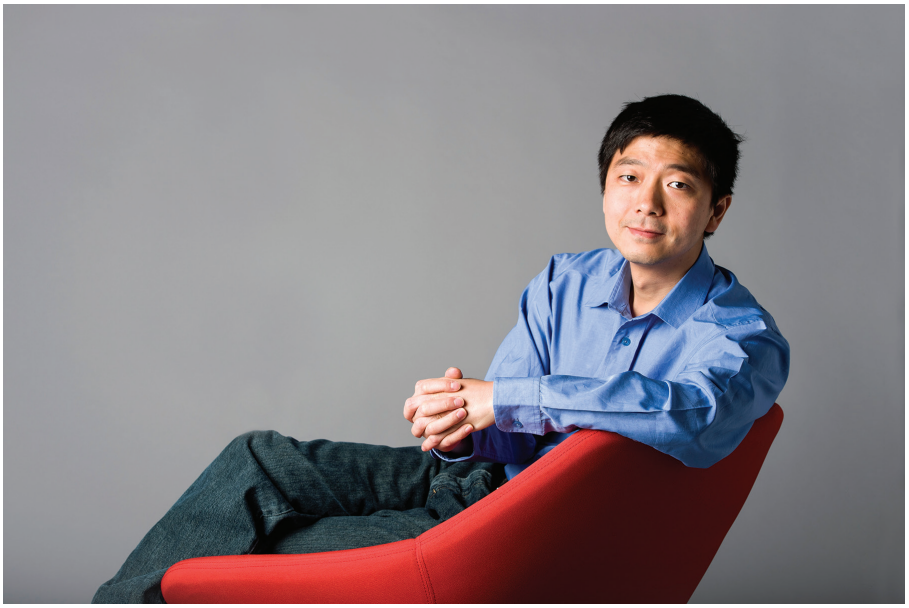
The second assumption is that sound waves move consistently through the body at standard, predictable rates, for all people. "In our bodies, all those parameters are highly variable," Dr. Bell explained. "That's another flaw: It's geometrically based on varying medium properties."

Poor-quality scans can have serious consequences. "Overweight patients tend to have very poor ultrasound image quality," Dr. Bell said, because fat deposits scatter ultrasound signals. The resulting noise can make scans far less useful for early detection of cancer lesions and tumors.

Photoacoustic imaging, on the other hand, primarily produces reflective artifacts. These vary, depending on where the scan originates within the body. Areas with lots of highly reflective bone surfaces will produce more echoes than areas without them. These duplicate images, she said, are especially problematic. "They can cloud the image and make it appear as if you have two photoacoustic signals where there should only be one."

Dr. Bell deals with these kinds of image abnormalities every day at the PULSE Lab, a highly interdisciplinary research program that brings biomedical, mechanical, and electrical engineers; computer scientists; and robotics experts together to design advanced imaging systems for surgical guidance and other uses. "The field is really suffering from reflection artifacts," she said. And finding a way to eliminate them is what first led her to consider machine learning.





Hoifung Poon, Ph.D., computer scientist and leader of Microsoft's Project Hanover, described the effort: "We're imagining a service where you can input 500 mutations found in a specific cancer, push a button, and have it generate a clinical report right away…That's the intelligence-gathering part that occupies tumor boards today, only now it will take seconds to do."

For the initial round of machine-learning training, Drs. Bell and Reiter used an extremely simple reflective artifact training model. They used a special simulation software program to create 30,000 data pairs. Each pair contained the location of a real, surgical tool tip and its reflective artifact. The pairings were created by establishing the surrounding environmental conditions and a host of other factors. They labeled the first 20,000 pairs correctly and turned the machine-learning network's off-the-shelf algorithm loose, looking for patterns in the data. Meanwhile, they held the remaining 10,000 unlabeled pairs back — for use testing the trained network's predictive powers.

"The idea," Dr. Bell said, "is to train the network to know where the true signal is and to identify anything else as a false signal and remove it."

When they ran the test on the unlabeled pairs, they were pleasantly surprised. "So far, preliminary results are looking good," Dr. Bell said. "We got submillimeter errors in point locations." The next step is to test different off-the-shelf algorithms to see whether one outperforms another.

If the concept ultimately proves viable, she said, the technology could make future photoacoustic images echo free. That would encourage surgeons to adopt photoacoustic-based surgical guidance systems.

What's the long-term potential?

"We're starting with simple models but hoping it can scale to more complicated cases that would be representative of ultrasound," Dr. Bell said. "If it does, then it's cleaning up ultrasound images, which everybody would love, and ultrasound is already a widely available tool in the clinic with known limitations. So, if we can get rid of those limitations, which I'm attempting to do, then it will certainly have large-scale impact, especially for patients known to produce poor-quality images."

Then, Dr. Bell took a breath and smiled. "I'm still not too sure how well it will work," she said, "because it's still in the middle; it's not an end-product yet. The research is ongoing. Although, I'm very excited and I believe in the results;

Muyinatu Bell, Ph.D., biomedical engineer, assistant professor, and founder and director of the new Photoacoustic and Ultrasonic Systems Engineering (PULSE) Lab at Johns Hopkins University, is developing reflective artifact models with the aim of cleaning up ultrasound images. This, she says, "will certainly have large-scale impact, especially for patients known to produce poor-quality images."

there's no telling what's going to happen as we continue on this path."

PRECISION CANCER TREATMENTS FOR ALL! NOT TOMORROW, BUT SOON

Before more than a tiny percentage of cancer patients can enjoy the enormous benefits of precision medicine, someone, somewhere, must do the impossible: They must find a way to eliminate the weeks of medical research currently required at the outset of each new cancer case.

Hoifung Poon, Ph.D., computer scientist and leader of Microsoft's Project Hanover, is on the job. Dr. Poon and his small team of computer scientists, aided by collaborators both inside and outside Microsoft, are using deep learning and machine reading to automate and replace today's time-consuming tumor board research with "an artificial intelligence-powered tumor board for the future."

When they're done, Dr. Poon said, the cloud-based data gathering and decision support system they're building will be able to simultaneously accept in-depth cancer patient disease profiles from oncologists everywhere, and in seconds rather than weeks, return comprehensive, up-to-date medical research reports. "We're imagining a service where you can input 500 mutations found in a specific cancer, push a button, and have it generate a clinical report right away," Dr. Poon said. "That report would tell you, 'Here is all we know about these gene mutations and here is what we know about the relevant drugs.' That's the intelligence-gathering part that occupies tumor boards today, only now it will take seconds to do."

PHASE ONE: TURNING INFORMATION INTO KNOWLEDGE

The first challenge, and it's huge, is to build a self-directed, curated, structured database to house all known information on gene pathways, gene regulation, gene expression, cancer mutations, drug action mechanisms — and more. The database must automatically update itself by scanning new published reports, starting with those contained in PubMed, a database maintained by the National Center for Biotechnology Information.

"PubMed currently contains 26 million medical research papers," Dr. Poon said, "but the most problematic part is that each year it adds over 1 million more. That's why we want to automate this reading process."

The other challenge is language ambiguity. "You can have the same kind of facts expressed differently," Dr. Poon explained. In addition, the same terms can, depending on the context, mean completely different things. Dr. Poon and his team are using the Literome database, which they developed before Project Hanover launched in the summer of 2015, to do the machine reading, aided by natural language processing (NLP) technology and Microsoft's Statistical Parsing and Linguistics Analysis Toolkit (SPLAT).

They've expanded Literome's mission to include tracking and storing data on individual drugs and their potential impacts on genes and gene mutations. That way, Literome will be able to predict which drugs will prove most effective against specific cancers. They're also continuing to refine Literome's machine-reading capabilities. "We need to use machine learning here," he said, "because you don't want to manually code all these patterns."

Years ago, Dr. Poon said, the National Cancer Institute teamed up with *Nature* magazine to develop a human-curated database. "They came out with this database — very high quality — but they gave up after two years. The database contained information on about 10,000 unique gene regulation pairs."

In contrast, Literome, using machine-reading technology, already has extracted data on 1 million unique gene pairs. "Once we develop the machine-reading mechanism and finish the

machine learning, it will run automatically and be able to scan all 26 million PubMed papers in a day."

BENEFITS TO CANCER PATIENTS

Freed of their research responsibilities, tumor boards will be able to offer precision cancer treatments to far more patients. "Instead of treating 100 cancer patients a year," Dr. Poon speculated, "we may be able to help them treat 500 a year or 1,000 a year or, eventually, 10,000 a year."

The impact could be dramatic: no more chemotherapy, today's standard of care. Chemotherapy indiscriminately targets all cells in the body undergoing division — the time when they're most vulnerable to attack. The idea is that because cancer cells divide far more frequently than normal cells do, they would be disproportionately destroyed. Dr. Poon compares this approach to "carpet bombing," because both methods produce "lots and lots of collateral damage."

Each patient would receive a customized cancer treatment plan that specifies the best drug to target the genetic mutations driving the patient's cancer. That's critical, Dr. Poon added, because of what we now know about the true nature of cancer. "Cancer is caused by our normal cells' DNA acquiring a bunch of mutations and then going haywire and growing infinitely … eventually becoming lethal."

And, he explained, no single set of gene mutations causes cancer. "Every cancer, every tumor is different. When you compare cancers from two lung cancer patients, each might have a couple dozen mutations, but they might have almost no overlap in those mutations. So, they might actually be completely different diseases."

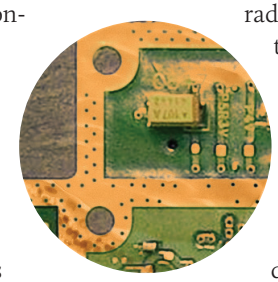
Precision cancer treatments can achieve dramatic results. In 2001, when the first drug targeting the genetic mutation behind chronic Myelomal leukemia — a cancer of white blood cells — was introduced, survival rates tripled, from 30 percent to about 90 percent.

PHASE TWO: PREDICTING THE BEST ONE–TWO PUNCH

Still, single drug treatments may not be enough, owing to genetic variability within tumors and to the incredible resilience of some cancer mutations to use gene-signaling pathways to devise workarounds to the drugs. After initial success, Dr. Poon said, "the cancer can come roaring back."

To prevent this, the best strategy would be to develop a two-drug therapy plan that simultaneously attacks the cancer on multiple fronts. The best combinations create synergies in which the drugs perform better individually when prescribed together. But this creates a "combinational explosion," which presents Dr. Poon's team with an enormous computational and predictive challenge.

The hundreds of known cancer drugs that need to be evaluated in combination produce "tens of thousands of combinations, at least," Dr. Poon said. But, to figure out the most promising of those combinations, he explained, you also must look at how they interact with the 20,000 genes in the human cancer genome. That puts the number of combinations to consider in the trillions.



Dr. Poon and his team are using the research findings culled from Literome, along with paired cancer drug research results from their collaborators at the Knight Cancer Center in Oregon, to develop their machine-learning network's predictive powers. Each round of anti-cancer drug test results takes two years to complete. Dr. Poon said several more iterations may be necessary to fully tune the machine-learning model and make Project Hanover ready for prime time.

"This is still in a completely research phase," Dr. Poon said. "but there are promising signs that we probably are onto something meaningful."

EXTREMELY FAST, EXTREMELY ACCURATE MEDICINE

Igor Barani, M.D., CEO of Enlitic, a San Francisco-based deep-learning and AI startup, was excited — and maybe a bit nervous. In a matter of days, he would unveil his company's first two fully functional deep-learning-powered clinical decision support tools to radiologists at the Radiology Society of North America's (RSNA) annual meeting, in Chicago.

The stakes were high. The two products Dr. Barani would introduce — Patient Triage, an X-ray interpretation and sorting tool, and a computerized tomography (CT) lung screening scan that makes lung cancer detection more effective — represented the first in a series of deep-learning diagnostic tools that Enlitic expected would transform radiology practices. They would be ready for market once Enlitic applied for, and received, approval from the Food and Drug Administration.

Radiologists tended to react to the Enlitic AI systems in two ways: Though impressed with their capabilities, radiologists were either uncomfortable trusting findings from deep-learning networks whose inner workings remained mysteries, or they were fearful that these powerful tools might replace them.

"They're seeing these networks do things that computers couldn't do before," said Dr. Barani, a radiology oncologist in his own right. "So, there's a tremendous level of anxiety."

Dr. Barani said he hopes Enlitic's decision support tools would ultimately win radiologists over because of their ability to increase the accuracy and speed of diagnoses and free radiologists to expand their practices.

To address the "black box" aspect of the technology, Enlitic has led the

Igor Barani, M.D., CEO of Enlitic, right, with radiologist Ben Covington Jr. Of the screening tool he's developed, Dr. Barani notes: "All the normal chest X-rays — they don't really need to look at or could glance at them very quickly. It should be very helpful for them to know that our algorithm, when it took a pass at it, didn't find anything. It can allow them to focus their skills on things that are complex, difficult to interpret, or that may require physician input."

industry in enhancing its products with interactive tools, such as heat map overlays, that highlight areas of interest on scans, lists of possible diagnoses, and probability scores indicating their likelihood. "We spend a lot of time figuring out how to help radiologists visualize the output of these networks," Dr. Barani said, adding that those visualization tools help build trust in the networks' findings.

Enlitic's diagnostic products have impressive capabilities. The Patient Triage tool not only can read and interpret chest X-rays, a first for any computer-based application, but it can do it in milliseconds — about 10,000 times faster than a radiologist. The system stores normal X-rays and automatically generates patient reports. Meanwhile, it interprets the abnormal X-rays and routes them to a reading radiologist along with its findings.

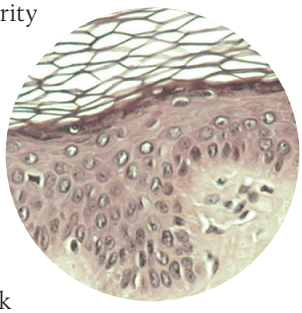
Patient Triage was designed to enable radiology practices to process large volumes of chest X-rays far more efficiently. (Chest X-rays are the most commonly used diagnostic images globally.) "We want to take the mundane out of their workflow," Dr. Barani explained. "All the normal chest X-rays — they don't really need to look at or could glance at them very quickly. It should be very helpful for them to know that our algorithm, when it took a pass at it, didn't find anything. It can allow them to focus their skills on things that are complex, difficult to interpret, or that may require physician input."

The second product Dr. Barani will introduce performs highly accurate and efficient lung cancer screenings. Each year, approximately 9 million Americans with prior histories of smoking become eligible for lung cancer screenings. Lung cancer is a highly lethal illness, claiming the lives of 80 to 90 percent of those diagnosed with it. It's also one of the most difficult cancers to detect on a medical scan, leading radiologists to err on the side of caution. The false positive rate of diagnoses can be very high.

"They're doing it for a good reason," Dr. Barani said. "You don't want to miss a nodule that might turn out to be malignant, no matter how small. But down the line, this leads to a lot of unnecessary biopsies, which are invasive, very expensive, and that cause patients a tremendous amount of anxiety."

Enlitic's CT lung cancer screening tool reduces the likelihood of false positives. Its unique features, including a 3-D deep-learning engine, enable it to detect lung cancer extremely early. In benchmarked tests, it detected cancer nodules in chest CT scans 50 percent more accurately than an expert panel of thoracic radiologists.

The lung screening CT scan tool examines lung CT scans slice by slice in its search for nodules. When it finds one with potentially malignant imaging features, it can further characterize the nodule by extracting and comparing its features against comparable features in a database of hundreds of thousands of lung cancer cases. It also calculates a similarity score based on the degree that the nodule's features match those from the database. That similarity score can then be used to estimate the risk of malignancy.



Dr. Barani had received considerable advance interest from radiologists who would be at the RSNA meeting. He said he expected to be in "nonstop meetings throughout the conference" and hoped to form new strategic partnerships and commercial commitments.

"I'm looking for partners who are willing to deploy this technology, so we can explore, not just how it works in real, clinical settings, but consider different pricing models and help us understand the value we're bringing to radiologists and hospitals.

"I think we'll be able to achieve massive synergies for anybody who deploys our technology to meet the needs of their patient populations."