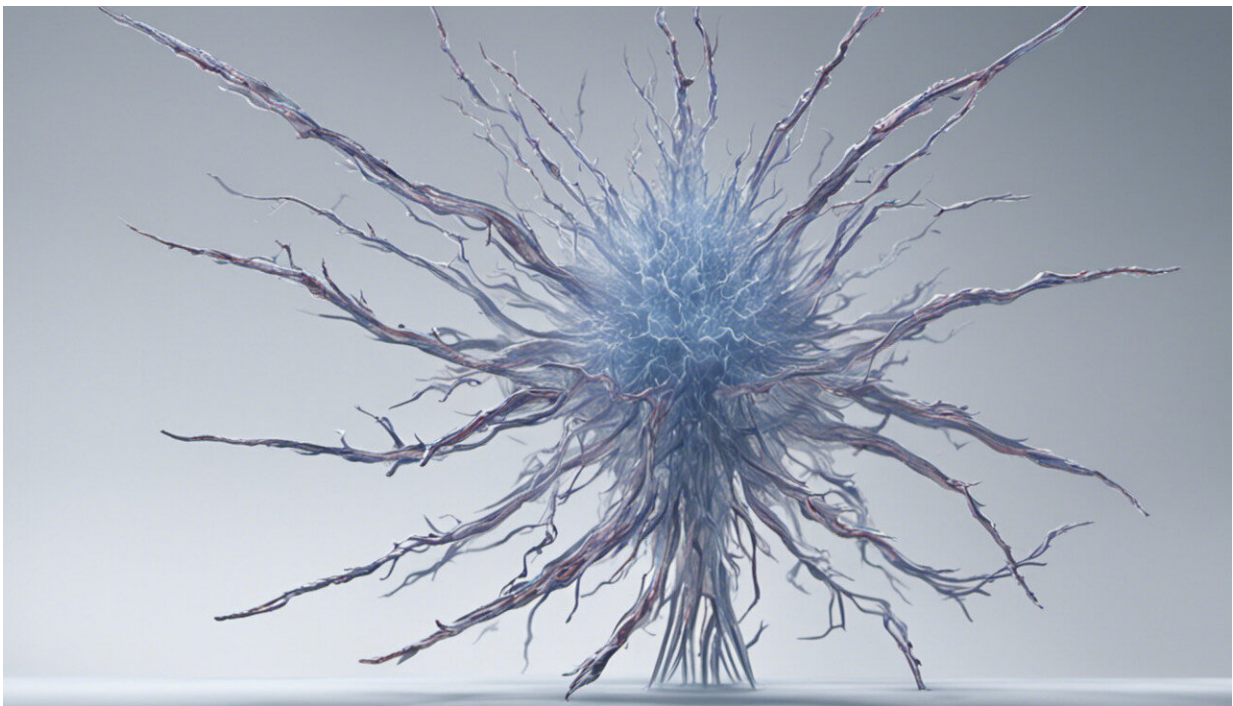# An algorithm could play a major role in helping radiologists diagnose cancer early, accurately

August 16 2019, by Avni Shah

Credit: AI-generated image (disclaimer)

Breast cancer is the leading cause of cancer-related death among women, and it is difficult to diagnose. Nearly 1 in 10 cancers is misdiagnosed as not cancerous; on the other hand, the more mammograms a woman has, the greater the chance she will see a false positive result and face an

unnecessary invasive procedure—most likely a biopsy.

More accurate diagnostic techniques are emerging. But what if instead we relied on the guidance of an [algorithm](link)?

Assad Oberai, Hughes Professor in the Aerospace and Mechanical Engineering Department at the USC Viterbi School of Engineering, asked this exact question in a recent paper published in ScienceDirect. Along with a team of researchers—including USC Viterbi Ph.D. student Dhruv Patel—Oberai specifically considered the following: Can you train a machine to interpret real-world images using synthetic data and streamline the steps to diagnosis?

The answer, he said, is most likely yes.

In the case of breast ultrasound elastography—an emerging imaging technique that provides information about a potential breast lesion by evaluating its stiffness in a noninvasive way—Oberai sought to determine if they could skip the most complicated steps of the process. Instead, he created physics-based models that showed varying levels of key properties. He then used thousands of data inputs derived from those models to train the machine learning algorithm.

Synthetic versus real-world data

But why would you use synthetically derived data to train the algorithm? Wouldn't real data be better?

"If you had enough data available, you wouldn't," Oberai said. "But in the case of medical imaging, you're lucky if you have 1,000 images. In situations like this, where data is scarce, these kinds of techniques become important."

Oberai and his team used about 12,000 synthetic images to train their machine learning algorithm. That process is similar in many ways to how photo identification software works, learning through repeated inputs how to recognize a particular person in an image or how our brain learns to classify a cat versus a dog. Through enough examples, the algorithm is able to glean different features inherent to a benign tumor versus a malignant tumor and make the correct determination.

Oberai and his team achieved nearly 100% classification accuracy on other synthetic images. Once the algorithm was trained, they tested it on real-world images to determine how accurate it could be in providing a diagnosis, measuring these results against biopsy-confirmed diagnoses associated with these images.

"We had about an 80% accuracy rate," Oberai said. "Next, we continue to refine the algorithm by using more real-world images as inputs."

Changing the method for diagnosing cancer

There are two prevailing points that make machine learning an important tool in advancing the landscape for cancer detection and diagnosis. First, machine learning algorithms can detect patterns that might be opaque to humans. Through manipulation of many such patterns, the algorithm can produce an accurate diagnosis. Secondly, machine learning offers a chance to reduce operator-to-operator error.

Would this then replace a radiologist's role in determining diagnosis? Oberai does not foresee an algorithm that serves as a sole arbiter of cancer diagnosis, but instead a tool that helps guide radiologists to more accurate conclusions.

"The general consensus is these types of algorithms have a significant role to play, including from imaging professionals whom it will impact

the most. However, these algorithms will be most useful when they do not serve as black boxes," Oberai said. "What did it see that led it to the final conclusion? The algorithm must be explainable for it to work as intended."

Adapting the diagnosis algorithm for other cancers

Because cancer causes different types of changes in the tissue it impacts, its presence can ultimately lead to a change in its physical properties, such as a change in density or porosity. These changes are can be discerned as a signal in medical images. The role of the machine learning algorithm is to pick out this signal and use it to determine whether a given tissue that is being imaged is cancerous.

Using these ideas, Oberai and his team are working with Vinay Duddalwar, professor of clinical radiology at the Keck School of Medicine of USC, to better diagnose renal cancer through contrast-enhanced CT images. Using the principles identified in training the [machine learning algorithm](#) for [breast cancer](#) [diagnosis](#), they are looking to train the algorithm on other features that might be prominently displayed in renal cancer cases, such as changes in tissue that reflect cancer-specific changes in a patient's microvasculature, the network of microvessels that help distribute blood within tissues.

Provided by University of Southern California

provided for information purposes only.