

# Scientists trained a computer to classify breast cancer tumors

November 19 2018

---



Three-dimensional culture of human breast cancer cells, with DNA stained blue and a protein in the cell surface membrane stained green. Image created in 2014 by Tom Misteli, Ph.D., and Karen Meaburn, Ph.D. at the NIH IRP.

Using technology similar to the type that powers facial and speech recognition on a smartphone, researchers at the University of North Carolina Lineberger Comprehensive Cancer Center have trained a computer to analyze breast cancer images and then classify the tumors with high accuracy.

In a study published in the journal *NPJ Breast Cancer*, researchers reported they used a form of artificial intelligence called [machine learning](#), or [deep learning](#), to train a computer to identify certain features of breast [cancer](#) tumors from images. The computer also identified the [tumor](#) type based on complex molecular and genomic features, which a pathologist can't yet identify from a picture alone. They believe this approach, while still in its early stages, could eventually lead to cost savings for the clinic and in breast cancer research.

"Your smartphone can interpret your speech, and find and identify faces in a photo," said the study's first author Heather D. Couture, a graduate research assistant in the UNC-Chapel Hill Department of Computer Science. "We're using similar technology where we capture abstract properties in images, but we're applying it to a totally different problem."

For the study, the researchers used a set of 571 images of breast cancer tumors from the Carolina Breast Cancer Study to train the computer to classify tumors for grade, estrogen receptor status, PAM50 intrinsic subtype, histologic subtype, and risk of recurrence score. To do this, they created software that learned how to predict labels from images using a training set, so that new images could be processed in the same way.

They then used a different set of 288 images to test the computer's ability to distinguish features of the tumor on its own, comparing the computer's responses to findings of a pathologist for each tumor's grade

and subtype, and to separate tests for gene expression subtypes. They found the computer was able to distinguish low-intermediate versus high-grade tumors 82 percent of the time. When they had two pathologists review the tumor grade for the low-intermediate grade group, the pathologists agreed with each other about 89 percent of the time, which was slightly higher than the computer's accuracy.

In addition, the computer identified estrogen receptor status, distinguished between ductal and lobular tumors, and determined whether each case had a high or low risk of recurrence high levels of accuracy. It also identified one of the molecular subtypes of breast cancers—the basal-like subtype—which is based on how genes within the tumor were expressed—with 77 percent accuracy.

"Using artificial intelligence, or machine learning, we were able to do a number of things that pathologists can do at a similar accuracy, but we were also able to do a thing or two that pathologists are not able to do today," said UNC Lineberger's Charles M. Perou, Ph.D., the May Goldman Shaw Distinguished Professor of Molecular Oncology, professor of genetics, and of pathology and laboratory medicine in the UNC School of Medicine. "This has a long way to go in terms of validation, but I think the accuracy is only going to get better as we acquire more images to train the computer with."

The computer's ability to identify the basal-like subtype was exciting to researchers, and could have applications in cancer research. They also believe the technology could have applications in communities that do not have pathology resources as well as in helping to validate pathologists' findings.

"We were surprised that the computer was able to get a pretty [high accuracy](#) in estimating biomarker risk just from looking at the pictures," said UNC Lineberger's Melissa Troester, Ph.D., a professor in the UNC

Gillings School of Global Public Health. "We spend thousands of dollars measuring those biomarkers using molecular tools, and this new method can take the image and get 80 percent accuracy or better at estimating the tumor phenotype or subtype. That was pretty amazing."

Couture said deep learning technology has been used in a range of applications, including [speech recognition](#) and autonomous vehicles.

"Humans can look at one or two examples of something, and be able to generalize when they see other objects," Couture said. "For example, chairs come in so many different forms, but we can recognize it as something we sit on. Computers have a much harder time generalizing from small amounts of data. But on other hand, if it you provide enough labeled data, they can learn concepts that are much more complex than humans can assess visually—such as identifying the basal-like subtype from an image alone."

The unique aspect of their work, researchers said, was that they were able to use the technology to see features of the tumors that humans cannot. They want to figure out what the computer is seeing, as well as to study whether the technology could predict outcomes.

"The [computer](#) extracted a lot of information from the [images](#)," Troester said. "We would like to test how well these features predict outcomes, and if we can use these features together with things like molecular data to do even better at giving patients a precise view of what their disease course looks like, and what treatments might be effective."

**More information:** Heather D. Couture et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype, *npj Breast Cancer* (2018). [DOI: 10.1038/s41523-018-0079-1](#)

Provided by UNC Lineberger Comprehensive Cancer Center

Citation: Scientists trained a computer to classify breast cancer tumors (2018, November 19)  
retrieved 1 May 2024 from

<https://medicalxpress.com/news/2018-11-scientists-breast-cancer-tumors.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--