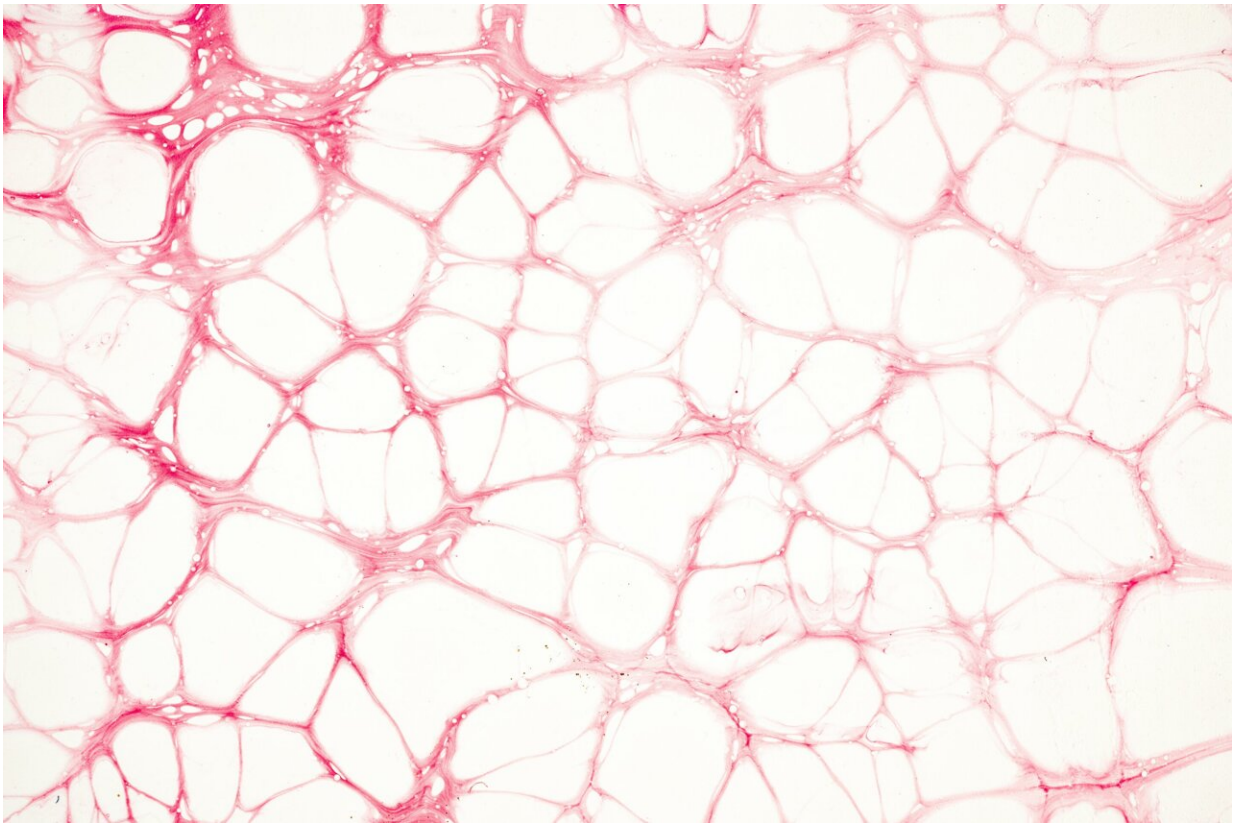# Medical Xpress

How X/Twitter trained an AI tool for pathologists

August 29 2023, by Nina Bai

Credit: Unsplash/CC0 Public Domain

The most impressive uses of artificial intelligence rely on good data—and lots of it. Chatbots, for example, learn to converse from millions of web pages full of text. Autonomous vehicles learn to drive

from sensor data recorded on millions of road trips.

For highly technical tasks, like understanding [medical images](#), however, good data sets are harder to find. In a new study, Stanford Medicine researchers have trained an AI-powered algorithm on a treasure trove of high-quality, annotated medical images from a surprising source—Twitter, now known as X.

The [new algorithm](#), which processed more than 200,000 pathology images, typically high-resolution views of cells or tissues stained with dye, and discussions from X, is capable of reading images for a wide range of conditions, such as melanoma, breast cancer or a parasitic infection. It can then retrieve similar images based on an image or text search. The algorithm is not meant to provide diagnoses on its own, but to serve as a powerful reference tool for clinicians and students.

"The main application is to help human pathologists look for similar cases to reference," said James Zou, Ph.D., assistant professor of biomedical data science and senior author of the study published Aug. 17 in *Nature Medicine*.

Not your average tweet

"It might be surprising to some folks that there is actually a lot of high-quality medical knowledge that is shared on Twitter," Zou said. In fact, the social media platform has become a popular forum for pathologists to share interesting images—so much so that the community has widely adopted a set of 32 hashtags to identify subspecialties.

"It's a very active community, which is why we were able to curate hundreds of thousands of these high-quality pathology discussions from Twitter," he said.

A typical pathology-related tweet might include an image from an unidentified patient, a brief description and relevant hashtags.

For the researchers, these pairings of image and [natural language](link) (in this case, the clinicians' written commentary) provide a valuable opportunity to teach the algorithm to recognize and link both kinds of data. "The goal here is to train a model that can understand both the [visual image](link) and the text description," Zou said. This would, in a way, attach meaning to a pathology image.

First, the researchers had to build a sizeable training data set. To separate wheat from chaff, the team used the 32 hashtags to retrieve relevant tweets in English from 2006 to 2022. They removed retweets, sensitive tweets and non-pathology images. They included replies with the most likes, which suggest quality, and excluded those with question marks, which imply doubt.

After this filtering process and the addition of 32,000 more annotated images from public data sets sourced from the web, the researchers had just over 200,000 image-text pairs. They called this collection OpenPath, one of the largest public datasets of human-annotated pathology images.

Next, the researchers trained an AI model on the OpenPath data set. The model uses a technique known as language-image contrastive learning, which identifies features of images and text, then maps them onto each other. These features might be the size of cell nuclei or certain phrases, for example, but are not explicitly provided by the researchers.

"The power of such a model is that we don't tell it specifically what features to look for. It's learning the relevant features by itself," Zou said.

The trained AI model, which the researchers have named PLIP, for

Pathology Language-Image Pre-training, would allow a clinician to input a new image or text description to search for similar annotated images in the database—a sort of Google Image search customized for pathologists, Zou said.

PLIP can also match a new image to a choice of disease descriptions—discerning, for example, whether an image shows normal tissue or a malignant tumor.

An intelligent reference

When PLIP was put to the test on new data sets, it handily outperformed existing models. On a measure of accuracy known as the F1 score, with lowest performance at 0 and highest at 1, the new model achieved scores of 0.6 to 0.8 compared with a previous model's 0.3 to 0.6. Zou emphasized that PLIP is not meant to compete with human pathologists but to support them.

"Maybe a pathologist is looking at something that's a bit unusual or ambiguous," he said. "They could use PLIP to retrieve similar images, then reference those cases to help them make their diagnoses."

The central innovation of the new study—harnessing high-quality medical knowledge from social media—could be extended to other specialties such as radiology and dermatology, which also rely on visual inspection, Zou suggested.

As for PLIP, the researchers have been continuously collecting new pathology data from X and other sources. "The more data you have, the more it will improve," Zou said.

  More information: Zhi Huang et al, A visual–language foundation model for pathology image analysis using medical Twitter, *Nature*

Medicine (2023). DOI: 10.1038/s41591-023-02504-3