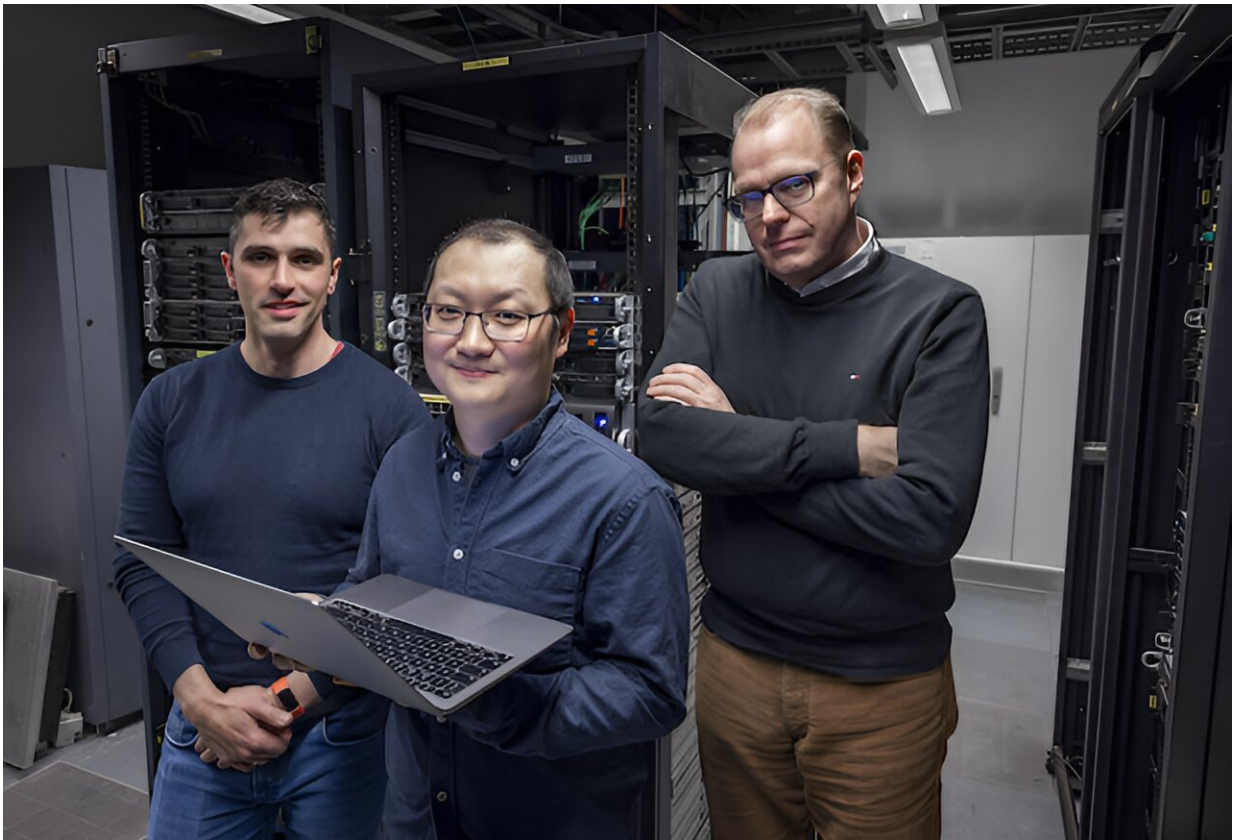


AI system learns to speak the language of cancer to enable improved diagnosis

June 11 2024



Credit: University of Glasgow

A computer system which harnesses the power of AI to learn the language of cancer is capable of spotting the signs of the disease in biological samples with remarkable accuracy, its developers say.

An international team of AI specialists and cancer scientists are behind the breakthrough development, which can also provide reliable predictions of patient outcomes.

Currently, pathologists examine and characterize the features of tissue samples taken from [cancer patients](#) on slides under a microscope. Their observations on the tumor's type and stage of growth help doctors determine each patient's course of treatment and their chances of recovery.

The new system, which the researchers call 'Histomorphological Phenotype Learning' (HPL), could aid human pathologists to provide faster, more accurate diagnoses of the disease, potentially helping to improve [cancer care](#) in the future.

The team, led by researchers from the University of Glasgow and New York University, outline how they developed and trained the HPL system in a new paper [published in the journal *Nature Communications*](#).

They began by collecting thousands of high-resolution images of tissue samples of lung adenocarcinoma taken from 452 patients stored in the United States National Cancer Institute's Cancer Genome Atlas database. In many cases, the data is accompanied by additional information on how the patients' cancers progressed.

Next, they developed an algorithm which used a training process called self-supervised [deep learning](#) to analyze the images and spot patterns based solely on the visual data in each slide.

The algorithm broke down the slide images into thousands of tiny tiles, each representing a small amount of human tissue. A deep neural network scrutinizes the tiles, teaching itself in the process to recognize and classify any visual features shared across any of the cells in each

tissue sample.

Dr. Ke Yuan, of the University of Glasgow's School of Computing Science, supervised the research and is the paper's senior author. He said, "We didn't provide the algorithm with any insight into what the samples were or what we expected it to find. Nonetheless, it learned to spot recurring visual elements in the tiles which correspond to textures, cell properties and tissue architectures called phenotypes.

"By comparing those visual elements across the whole series of images it examined, it recognized phenotypes which often appeared together, independently picking out the architectural patterns that human pathologists had already identified in the samples."

When the team added analysis of slides from squamous cell lung cancer to the HPL system, it was capable of correctly distinguishing between their features with 99% accuracy.

Once the algorithm had identified patterns in the samples, the researchers used it to analyze links between the phenotypes it had classified and the clinical outcomes stored in the database, including how long patients lived after having cancer surgery.

The algorithm discovered that certain phenotypes, such as tumor cells which are less invasive, or lots of inflammatory cells attacking the tumor, were more common in patients who lived longer after treatment. Others, like aggressive [tumor cells](#) forming solid masses, or regions where the immune system was excluded, were more closely associated with the recurrence of tumors.

The predictions made by the HPL system correlated well with the real-life outcomes of the patients stored in the database, correctly assessing the likelihood and timing of cancer's return 72% of the time. Human

pathologists tasked with the same prediction drew the correct conclusions with 64% accuracy.

When the research was expanded to include analysis of thousands of slides across 10 other types of cancers, including breast, prostate and bladder cancers, the results were similarly accurate despite the increased complexity of the task.

Professor John Le Quesne, from the University of Glasgow's School of Cancer Sciences, is one of the co-senior authors of the paper and supervised the research. He said, "We were surprised but very pleased by the effectiveness of machine learning to tackle this task. It takes many years to train human pathologists to identify the cancer subtypes they examine under the microscope and draw conclusions about the most likely outcomes for patients. It's a difficult, time-consuming job, and even highly-trained experts can sometimes draw different conclusions from the same slide.

"In a sense, the algorithm at the heart of the HPL system taught itself from first principles to speak the language of cancer—to recognize the extremely complex patterns in the slides and 'read' what they can tell us about both the type of cancer and its potential effect on patients' long-term health. Unlike a human pathologist, it doesn't understand what it's looking at, but it can still draw strikingly accurate conclusions based on mathematical analysis.

"It could prove to be an invaluable tool to aid pathologists in the future, augmenting their existing skills with an entirely unbiased second opinion. The insight provided by human expertise and AI analysis working together could provide faster, more accurate cancer diagnoses and evaluations of patients' likely outcomes. That, in turn, could help improve monitoring and better-tailored care across each patients' treatment."

Dr. Adalberto Claudio Quiros, a research associate in the University of Glasgow's School of Cancer Sciences and School of Computing Science, is a co-first author of the paper. He said, "This research shows the potential that cutting-edge machine learning has to create advances in cancer science which could have significant benefits for patient care.

"This kind of self-learning algorithm will only become more accurate as additional data is added, helping it become more fluent in the language of cancer. Unlike humans, it brings no pre-conceived ideas to its work, so it may even find patterns across the datasets that haven't been fully explored before.

"Ultimately, our aim is to provide doctors and patients with a tool that can help provide them with an improved understanding of their prognosis and treatment."

The team's paper, titled "Mapping the landscape of histomorphological cancer phenotypes using self-supervised learning on unlabeled, unannotated pathology slides," is published in *Nature Communications*.

More information: Adalberto Claudio Quiros et al, Mapping the landscape of histomorphological cancer phenotypes using self-supervised learning on unannotated pathology slides, *Nature Communications* (2024). [DOI: 10.1038/s41467-024-48666-7](https://doi.org/10.1038/s41467-024-48666-7)

Provided by University of Glasgow

Citation: AI system learns to speak the language of cancer to enable improved diagnosis (2024, June 11) retrieved 18 June 2024 from <https://phys.org/news/2024-06-ai-language-cancer-enable->

[diagnosis.html](#)

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.