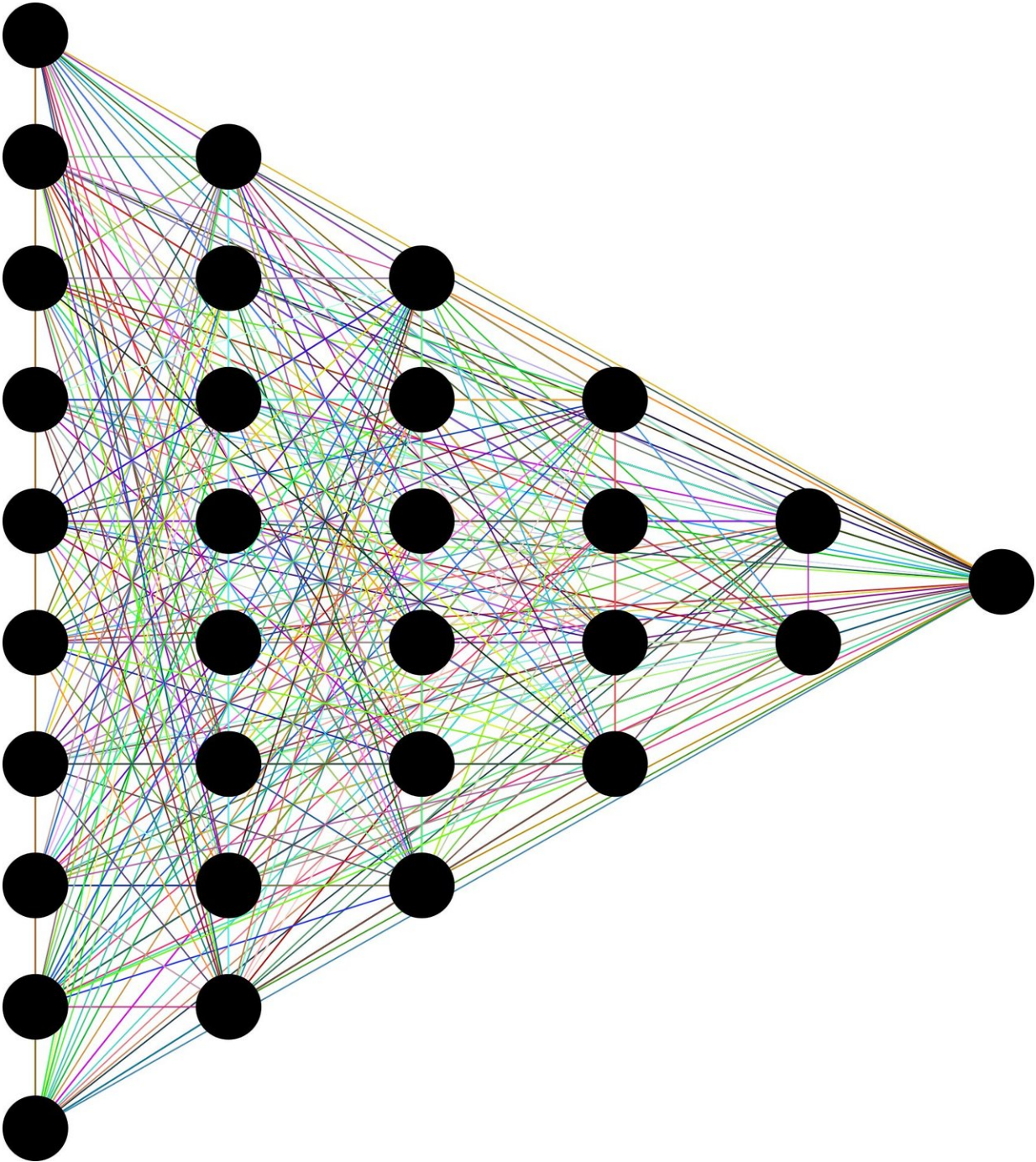


# Deep learning helps tease out gene interactions

December 10 2019

---



Credit: CC0 Public Domain

Carnegie Mellon University computer scientists have taken a deep

learning method that has revolutionized face recognition and other image-based applications in recent years and redirected its power to explore the relationship between genes.

The trick, they say, is to transform massive amounts of gene expression data into something more image-like. Convolutional neural networks (CNNs), which are adept at analyzing visual imagery, can then infer which [genes](#) are interacting with each other. The CNNs outperform existing methods at this task.

The researchers' report on how CNNs can help identify disease-related genes and developmental and genetic pathways that might be targets for drugs is being published today in the *Proceedings of the National Academy of Science*. But Ziv Bar-Joseph, professor of computational biology and [machine learning](#), said the applications for the new method, called CNNC, could go far beyond gene interactions.

The new insight described in the paper suggests that CNNC could be similarly deployed to investigate causality in a wide variety of phenomena, including [financial data](#) and social networking, said Bar-Joseph, who co-authored the paper with Ye Yuan, a post-doctoral researcher in CMU's Machine Learning Department.

"CNNs, which were developed a decade ago, are revolutionary," Bar-Joseph said. "I'm still in awe of Google Photos, which uses them for facial recognition," he added as he scrolled through photos on his smartphone, showing how the app could identify his son at different ages, or identify his father based on an image of the rear right side of his head. "We sometimes take this technology for granted because we use it all the time. But it's incredibly powerful and is not restricted to images. It's all a matter of how you represent your data."

In this case, he and Yuan were looking at gene relationships. The

approximately 20,000 genes in humans work in concert, so it's necessary to know how genes work together in complexes or networks to understand human development or diseases.

One way to infer these relationships is to look at [gene expression](#)—which represents the activity levels of genes in cells. Generally, if gene A is active at the same time gene B is active, that's a clue that the two are interacting, Yuan said. Still, it's possible that this is a coincidence or that both are activated by a third gene C. Several previous methods have been developed to tease out these relationships.

To employ CNNs to help analyze gene relationships, Yuan and Bar-Joseph used single-cell expression data—experiments that can determine the level of every gene in a single cell. The results of hundreds of thousands of these single-cell analyses were then arranged in the form of a matrix or histogram so that each cell of the matrix represented a different level of co-expression for a pair of genes.

Presenting the data in this way added a spatial aspect that made the data more image-like and, thus, more accessible to CNNs. By using data from genes whose interactions already had been established, the researchers were able to train the CNNs to recognize which genes were interacting and which weren't based on the visual patterns in the data matrix, Yuan said.

"It's very, very hard to distinguish between causality and correlation," Yuan said, but the CNNC method proved statistically more accurate than existing methods. He and Bar-Joseph anticipate CNNC will be one of several techniques that researchers will eventually deploy in analyzing large datasets.

"This is a very general method that could be applied to a number of analyses," Bar-Joseph said. The main limitation is data—the more data

there is, the better CNNs work. Cell biology is well-suited for using CNNC, as a typical experiment can involve tens of thousands of cells and generate a massive amount of data.

**More information:** *Proceedings of the National Academy of Science* (2019). [DOI: 10.1073/pnas.1911536116](https://doi.org/10.1073/pnas.1911536116)

Provided by Carnegie Mellon University

Citation: Deep learning helps tease out gene interactions (2019, December 10) retrieved 21 June 2024 from <https://techxplore.com/news/2019-12-deep-gene-interactions.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.