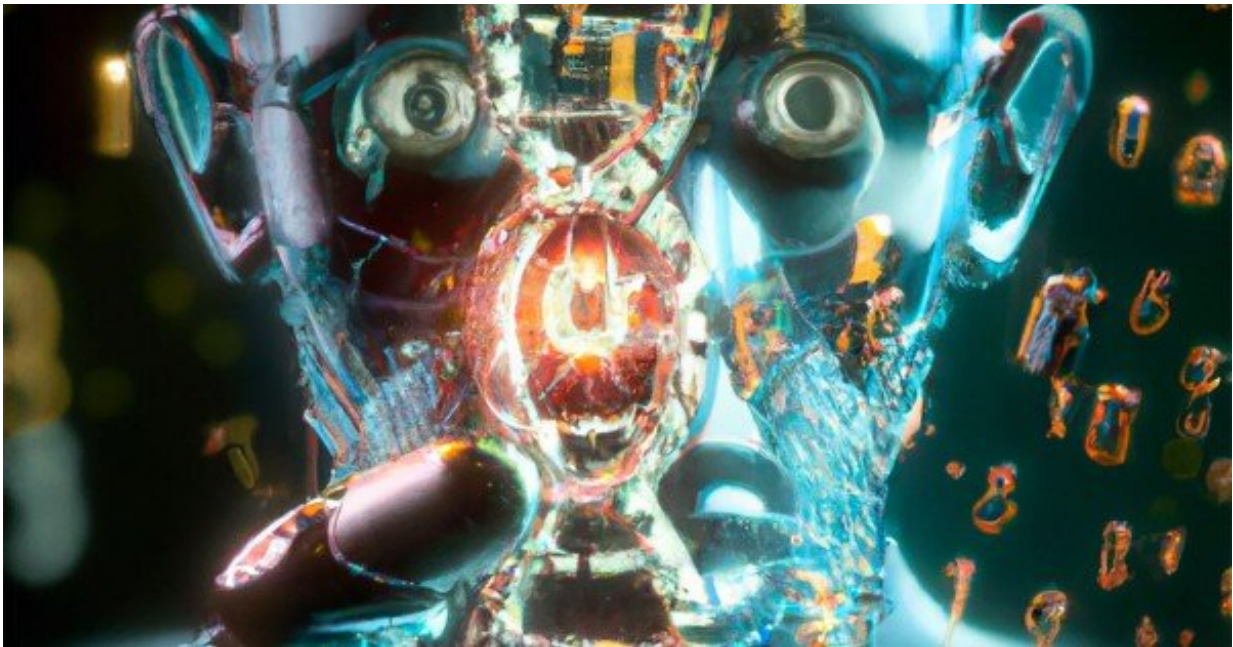


New computer program 'learns' to identify mosaic mutations that cause disease

January 3 2023



This image was generated by artificial intelligence based upon a request for 3D art using a natural language interpretation of "artificial intelligence that detects mutations in the genome." Credit: Joseph Gleeson/UC San Diego

Genetic mutations cause hundreds of unsolved and untreatable disorders. Among them, DNA mutations in a small percentage of cells, called mosaic mutations, are extremely difficult to detect because they exist in a tiny percentage of the cells.

Current DNA mutation software detectors, while scanning the 3 billion bases of the human genome, are not well suited to discern mosaic [mutations](#) hiding among normal DNA sequences. Often medical geneticists must review DNA sequences by eye to try to identify or confirm mosaic mutations—a time-consuming endeavor fraught with the possibility of error.

Writing in the January 2, 2023, issue of *Nature Biotechnology*, researchers from the University of California San Diego School of Medicine and Rady Children's Institute for Genomic Medicine describe a method for teaching a computer how to spot mosaic mutations using an artificial intelligence approach termed "[deep learning](#)."

Deep learning, sometimes referred to as artificial neural networks, is a machine learning technique that teaches computers to do what comes naturally to humans: learn by example, especially from large amounts of information. Compared with traditional statistical models, deep learning models use [artificial neural networks](#) to process visually represented data. The models function in ways similar to human visual processing, with much greater accuracy and attention to detail, leading to major advances in computational abilities, including mutation detection.

"One example of an unsolved disorder is [focal epilepsy](#)," said senior study author Joseph Gleeson, MD, Rady Professor of Neuroscience at UC San Diego School of Medicine and director of neuroscience research at the Rady Children's Institute for Genomic Medicine.

"Epilepsy affects 4% of the population, and about one-quarter of focal seizures fail to respond to common medication. These patients often require surgical excision of the short-circuited focal part of the brain to stop seizures. Among these patients, mosaic mutations within the brain can cause epileptic focus.

"We have had many epilepsy patients where we were not able to spot the cause, but once we applied our method, called 'DeepMosaic,' to the genomic data, the mutation became obvious. This has allowed us to improve the sensitivity of DNA sequencing in certain forms of epilepsy, and had led to discoveries that point to new ways to treat brain disease."

Gleeson said accurate detection of mosaic mutations is the first step in [medical research](#) toward developing treatments for many diseases.

Co-first and co-corresponding author Xiaoxu Yang, Ph.D., a postdoctoral scholar in Gleeson's lab, said DeepMosaic was trained on almost 200,000 simulated and biological variants across the genome until, "finally, we were satisfied with its ability to detect variants from data it had never encountered before."

To train the computer, the authors fed examples of trustworthy mosaic mutations as well as many normal DNA sequences and taught the computer to tell the difference. By repeatedly training and retraining with ever-more complex datasets and selection between a dozen of models, the computer was eventually able to identify mosaic mutations much better than [human eyes](#) and prior methods. DeepMosaic was also tested on several independent large-scale sequencing datasets that it had never seen, outperforming prior approaches.

"DeepMosaic surpassed traditional tools in detecting mosaicism from genomic and exonic sequences," said co-first author Xin Xu, a former undergraduate research assistant at UC San Diego School of Medicine and now a research data scientist at Novartis. "The prominent visual features picked up by the deep learning models are very similar to what experts are focusing on when manually examining variants."

DeepMosaic is freely available to scientists. It is not a single [computer](#) program, but rather an open-source platform that can enable other

researchers to train their own neural networks to achieve a more targeted detection of mutations using a similar image-based setup, the researchers said.

More information: Xiaoxu Yang et al, Control-independent mosaic single nucleotide variant detection with DeepMosaic, *Nature Biotechnology* (2023). [DOI: 10.1038/s41587-022-01559-w](https://doi.org/10.1038/s41587-022-01559-w)

Provided by University of California - San Diego

Citation: New computer program 'learns' to identify mosaic mutations that cause disease (2023, January 3) retrieved 22 February 2023 from <https://medicalxpress.com/news/2023-01-mosaic-mutations-disease.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.