

AI with swarm intelligence learns to detect cancer, lung diseases and COVID-19

26 May 2021



Credit: CC0 Public Domain

Communities benefit from sharing knowledge and experience among their members. Following a similar principle—called "swarm learning"—an international research team has trained artificial intelligence algorithms to detect blood cancer, lung diseases and COVID-19 in data stored in a decentralized fashion. This approach has advantage over conventional methods since it inherently provides privacy preservation technologies, which facilitates cross-site analysis of scientific data. Swarm learning could thus significantly promote and accelerate collaboration and information exchange in research, especially in the field of medicine. Experts from the DZNE, the University of Bonn, the information technology company Hewlett Packard Enterprise (HPE) and other research institutions report on this in the scientific journal *Nature*.

Science and medicine are becoming increasingly digital. Analyzing the resulting volumes of information—known as "[big data](#)"—is considered a key to better treatment options. "Medical research data are a treasure. They can play a decisive role in developing personalized therapies that are

tailored to each individual more precisely than conventional treatments," said Joachim Schultze, Director of Systems Medicine at the DZNE and professor at the Life & Medical Sciences Institute (LIMES) at the University of Bonn. "It's critical for science to be able to use such data as comprehensively and from as many sources as possible."

However, the exchange of medical research data across different locations or even between countries is subject to data protection and data sovereignty regulations. In practice, these requirements can usually only be implemented with significant effort. In addition, there are technical barriers: For example, when huge amounts of data have to be transferred digitally, data lines can quickly reach their performance limits. In view of these conditions, many [medical studies](#) are locally confined and cannot utilize data that is available elsewhere.

Data Remains on Site

In light of this, a research collaboration led by Joachim Schultze tested a novel approach for evaluating research data stored in a decentralized fashion. The basis for this was the still young "Swarm Learning" technology developed by HPE. In addition to the IT company, numerous research institutions from Greece, the Netherlands and Germany—including members of the "German COVID-19 OMICS Initiative" (DeCOI) - participated in this study.

Swarm Learning combines a special kind of information exchange across different nodes of a network with methods from the toolbox of "machine learning," a branch of artificial intelligence (AI). The linchpin of machine learning are algorithms that are trained on data to detect patterns in it—and that consequently acquire the ability to recognize the learned patterns in other data as well. "Swarm Learning opens up new opportunities for

collaboration in medical research, as well as in business. The key is that all participants can learn from each other without having to share confidential data," said Dr. Eng Lim Goh, Senior Vice President and Chief Technology Officer for artificial intelligence at HPE.

In fact, with Swarm Learning, all research data remains on site. Only algorithms and parameters are shared—in a sense, lessons learned. "Swarm Learning fulfills the requirements of data protection in a natural way," Joachim Schultze emphasized.

Collaborative Learning

Unlike "federated learning," in which the data also remains locally, there is no centralized command center, the Bonn scientist explained. "Swarm Learning happens in a cooperative way based on rules that all partners have agreed on in advance. This set of rules is captured in a blockchain." This is a kind of digital protocol that regulates [information exchange](#) between the partners in a binding manner, it documents all events and all parties have access to it. "The blockchain is the backbone of Swarm Learning," Schultze said. "All members of the swarm have equal rights. There is no central power over what happens and over the results. So there is, in a sense, no spider controlling the data web."

Thus, the AI algorithms learn locally, namely on the basis of the data available at each network node. The learning outcomes of each node are collected as parameters through the blockchain and smartly processed by the system. The outcome, i. e. optimized parameters, are passed on to all parties. This process is repeated multiple times, gradually improving the algorithms' ability to recognize patterns at each node of the network.

Lung Images and Molecular Features

The researchers are now providing practical proof of this approach through the analysis of X-ray images of the lungs and of transcriptomes: The latter are data on the gene activity of cells. In the current study, the focus was specifically on immune cells circulating in the blood—in other words, white blood cells. "Data on the gene activity of blood cells

are like a molecular fingerprint. They hold important information about how the organism reacts to a disease," Schultze said. "Transcriptomes are available in large numbers just like X-ray images, and they are highly complex. This is exactly the kind of information you need for artificial intelligence analysis. Such data is perfect for testing Swarm Learning."

The research team addressed a total of four infectious and non-infectious diseases: two variants of blood cancer (acute myeloid leukemia and acute lymphoblastic leukemia), as well as tuberculosis and COVID-19. The data included a total of more than 16,000 transcriptomes. The swarm learning network over which the data were distributed typically consisted of at least three and up to 32 nodes. Independently of the transcriptomes, the researchers analyzed about 100,000 chest X-ray images. These were from patients with fluid accumulation in the lung or other pathological findings as well as from individuals without anomalies. These data were distributed across three different nodes.

A High Rate of Success

The analysis of both the transcriptomes and the X-ray images followed the same principle: First, the researchers fed their algorithms with subsets of the respective data set. This included information about which of the samples came from patients and which from individuals without findings. The learned pattern recognition for "sick" or "healthy" was then used to classify further data, in other words it was used to sort the data into samples with or without disease. The accuracy, i.e. the ability of the algorithms to distinguish between healthy and diseased individuals, was around 90 percent on average for the transcriptomes (each of the four diseases was evaluated separately); in the case of the X-ray data, it ranged from 76 to 86 percent.

"The methodology worked best in leukemia. In this disease, the signature of gene activity is particularly striking and thus easiest for artificial intelligence to detect. Infectious diseases are more variable. Nevertheless, the accuracy was also very high for tuberculosis and COVID-19. For X-ray data, the rate was somewhat lower, which is due to the lower

data or image quality," Schultze commented on the results. "Our study thus proves that Swarm Learning can be successfully applied to very different data. In principle, this applies to any type of information for which pattern recognition by means of artificial intelligence is useful. Be it genome data, X-ray images, data from brain imaging or other complex data."

The study also found that Swarm Learning yielded significantly better results than when the nodes in the network learned separately. "Each node benefits from the experience of the other nodes, although only local data is ever available. The concept of Swarm Learning has thus passed the practical test," Schultze said.

A Vision for the Future

"I am convinced that swarm learning can give a huge boost to medical research and other data-driven disciplines. The current study was just a test run. In the future, we intend to apply this technology to Alzheimer's and other neurodegenerative diseases," Schultze said. "Swarm Learning has the potential to be a real game changer and could help make the wealth of experience in medicine more accessible worldwide. Not only research institutions but also hospitals, for example, could join together to form such swarms and thus share information for mutual benefit."

More information: Warnat-Herresthal et al., Swarm Learning for decentralized and confidential clinical machine learning, *Nature* (2021), [DOI: 10.1038/s41586-021-03583-3](https://doi.org/10.1038/s41586-021-03583-3)

Provided by German Center for
Neurodegenerative Diseases

APA citation: AI with swarm intelligence learns to detect cancer, lung diseases and COVID-19 (2021, May 26) retrieved 3 June 2022 from <https://medicalxpress.com/news/2021-05-ai-swarm-intelligence-cancer-lung.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.