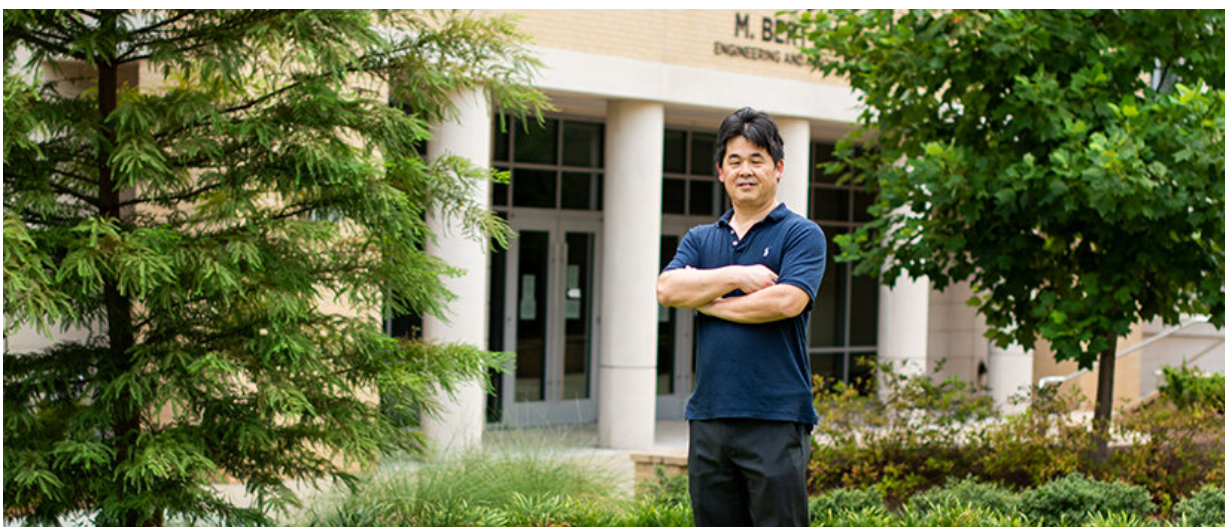# Deep learning algorithm to speed up materials discovery in emerging tech industries

August 25 2020

Credit: University of South Carolina

Solid-state inorganic materials are critical to the growth and development of electric vehicle, cellphone, laptop battery and solar energy technologies. However, finding the ideal materials with the desired functions for these industries is extremely challenging. Jianjun Hu, an associate professor of computer science at the University of South Carolina is the lead researcher on a project to generate new hypothetical materials.

Due to the vast chemical design space and the high sparsity of candidates, experimental trials and first-principle computational simulations cannot be used as screening tools to solve this problem. Instead, researchers have developed a deep learning-based smart algorithm that uses a technique called generative adversarial network (GAN) model to dramatically improve the material search efficiency up to two orders of magnitude. It has the potential to greatly speed up the discovery of novel functional materials.

The work, published in *NPJ Computational Materials*, was a collaboration between researchers at the University of South Carolina College of Engineering and Computing and Guizhou University, a research university located in Guiyang, China.

Inspired by the deep learning technique used in Google's AlphaGo, which learned implicit rules of the board game Go to defeat the game's top players, the researchers used their GAN neural network to learn the implicit chemical composition rules of atoms in different elements to assemble chemically valid formulas. By training their deep learning models using the tens of thousands of known inorganic materials deposited in databases such as ICSD and OQMD, they created a generative machine learning model capable of generating millions of new hypothetical inorganic material formulas.

"There is almost an infinite number of new materials that could exist, but they haven't been discovered yet," said Jianjun Hu. "Our algorithm, it's like a generation engine. Using this model, we can generate a lot of new hypothetical materials that have very high likelihoods to exist."

Without explicitly modeling or enforcing chemical constraints such as charge neutrality and electronegativity, the deep learning-based smart algorithm learned to observe such rules when generating millions of hypothetical materials' formulas. The predictive power of the algorithm

has been verified both by known materials and recent findings in materials discovery literature."One major advantage of our algorithm is the high validity, uniqueness and novelty, which are the three major evaluation metrics of such generative models," said Shaobo Li, a professor at Guizhou University who was involved in this study.

This is not the first time that an algorithm has been created for materials discovery. Past algorithms were also able to produce millions of potential new materials. However, very few of the materials discovered by these algorithms were synthesizable due to their high free energy and instability. In contrast, nearly 70 percent of the [inorganic materials](#) identified by Hu's team are very likely to be stable and then possibly synthesizable.

"You can get any number of formula combinations by putting elements' symbols together. But it doesn't mean the physics can exist," said Ming Hu, an associate professor of mechanical engineering at UofSC also involved in the research. "So, our algorithm and the next step, structure prediction [algorithm](#), will dramatically increase the speed to screening new function materials by creating synthesizable compounds."

These new materials will help researchers in fields such as electric vehicles, green energy, solar energy and cellphone development as they continually search for new materials with optimized functionalities. With the current materials discovery process being so slow, these industries' growth has been limited by the materials available to them.

The next major step for the team is to predict the crystal structure of the generated formulas, which is currently a major challenge. However, the team has already started working on this challenge along with several leading international teams. Once solved, the two steps can be combined to discover many potential materials for energy conversion, storage and other applications.

Provided by University of South Carolina