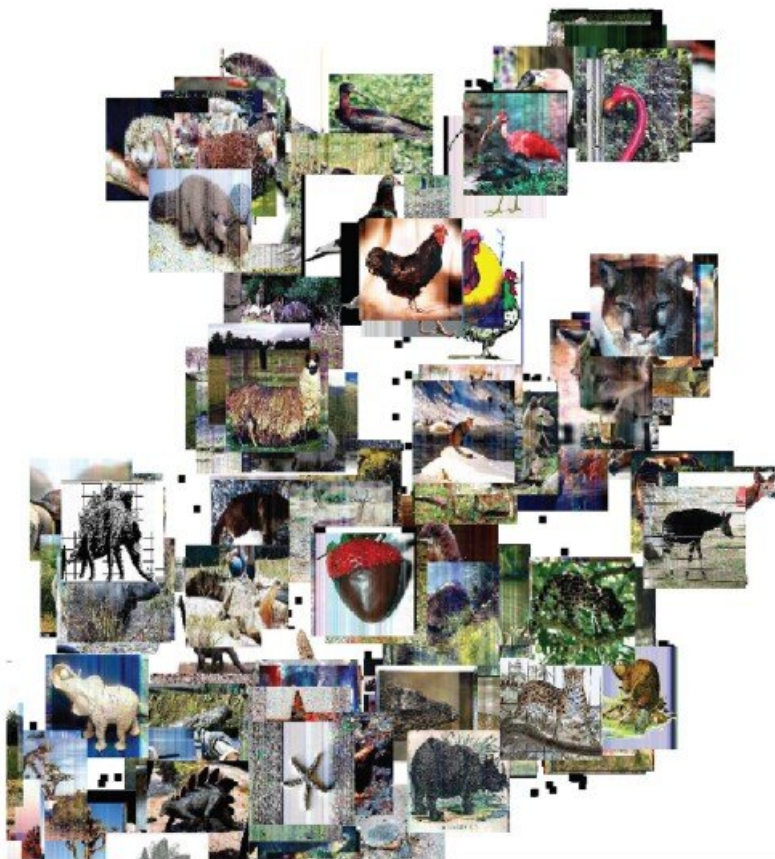


A novel neural network to understand symmetry, speed materials research

October 8 2021



Low-dimensional uniform manifold approximation projection to visualize how neural networks learn semantic similarity of natural images. Credit: Joshua

Agar/Lehigh University

Understanding structure-property relations is a key goal of materials research, according to Joshua Agar, a faculty member in Lehigh University's Department of Materials Science and Engineering. And yet currently no metric exists to understand the structure of materials because of the complexity and multidimensional nature of structure.

Artificial neural networks, a type of machine learning, can be trained to identify similarities—and even correlate parameters such as structure and properties—but there are two major challenges, says Agar. One is that the majority of vast amounts of data generated by materials experiments are never analyzed. This is largely because such images, produced by scientists in laboratories all over the world, are rarely stored in a usable manner and not usually shared with other research teams. The second challenge is that [neural networks](#) are not very effective at learning symmetry and periodicity (how periodic a material's structure is), two features of utmost importance to [materials researchers](#).

Now, a team led by Lehigh University has developed a novel machine learning approach that can create similarity projections via machine learning, enabling researchers to search an unstructured image database for the first time and identify trends. Agar and his collaborators developed and trained a neural network model to include symmetry-aware features and then applied their method to a set of 25,133 piezoresponse force microscopy images collected on diverse materials systems over five years at the University of California, Berkeley. The results: they were able to group similar classes of material together and observe trends, forming a basis by which to start to understand structure-property relationships.

"One of the novelties of our work is that we built a special neural network to understand symmetry and we use that as a feature extractor to make it much better at understanding images," says Agar, a lead author of the paper where the work is described: "Symmetry-Aware Recursive Image Similarity Exploration for Materials Microscopy," published today in *npj Computational Materials*. In addition to Agar, authors include, from Lehigh University: Tri N. M. Nguyen, Yichen Guo, Shuyu Qin and Kylie S. Frew and, from Stanford University: Ruijuan Xu. Nguyen, a lead author, was an undergraduate at Lehigh University and is now pursuing a Ph.D. at Stanford.

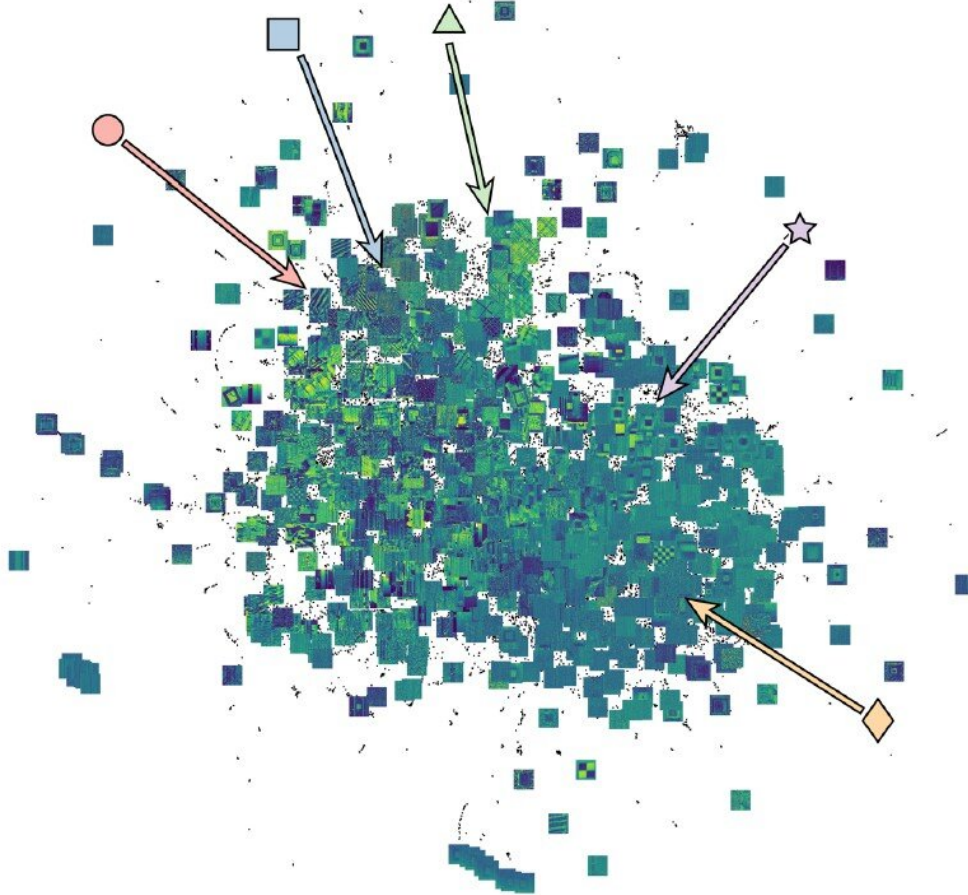
The team was able to arrive at projections by employing Uniform Manifold Approximation and Projection (UMAP), a non-linear dimensionality reduction technique. This approach, says Agar, allows researchers to learn "...in a fuzzy way, the topology and the higher-level structure of the data and compress it down into 2D."

"If you train a neural network, the result is a vector, or a set of numbers that is a compact descriptor of the features. Those features help classify things so that some similarity is learned," says Agar. "What's produced is still rather large in space, though, because you might have 512 or more different features. So, then you want to compress it into a space that a human can comprehend such as 2D, or 3D—or, *maybe*, 4D."

By doing this, Agar and his team were able to take the 25,000-plus images and group very similar classes of material together.

"Similar types of structures in material are semantically close together and also certain trends can be observed particularly if you apply some metadata filters," says Agar. "If you start filtering by who did the deposition, who made the material, what were they trying to do, what is the material system... you can really start to refine and get more and more similarity. That similarity can then be linked to other parameters

like properties."



Low-dimensional uniform manifold approximation projection showing symmetry-aware image similarity from a database of greater than 25,000 piezoresponse force microscopy images. Credit: Joshua Agar/Lehigh University

This work demonstrates how improved data storage and management could rapidly accelerate materials discoveries. According to Agar, of particular value are images and data generated by failed experiments.

"No one publishes failed results and that's a big loss because then a few years later someone repeats the same line of experiments," says Agar. "So, you waste really good resources on an experiment that likely won't work."

Instead of losing all of that information, the data that has already been collected could be used to generate new trends that have not been seen before and speed discovery exponentially, says Agar.

This study is the first "use case" of an innovative new data-storage enterprise housed at Oak Ridge National Laboratory called [DataFed](#). DataFed, according to its website is "...a federated, big-[data storage](#), collaboration, and full-life-cycle management system for computational science and/or data analytics within distributed high-performance computing (HPC) and/or cloud-computing environments."

"My team at Lehigh has been part of the design and development of DataFed in terms of making it relevant for scientific use cases," says Agar. "Lehigh is the first live implementation of this fully-scalable system. It's a federated database so anyone can pop up their own server and be tied to the central facility."

Agar is the machine learning expert on Lehigh University's Presidential Nano-Human Interface Initiative team. The interdisciplinary initiative, integrating the social sciences and engineering, seeks to transform the ways that humans interact with instruments of scientific discovery to accelerate innovations.

"One of the key goals of Lehigh's Nano/Human Interface Initiative is to put relevant information at the fingertips of experimentalists to provide actionable information that allows more informed decision-making and accelerates scientific discovery," says Agar. "Humans have limited capacity for memory and recollection. DataFed is a modern-day Memex;

it provides a memory of scientific information that can easily be found and recalled."

DataFed provides an especially powerful and invaluable tool for researchers engaged in interdisciplinary team science, allowing researchers who are collaborating on team projects located in different/remote locations to access each other's raw data. This is one of the key components of our Lehigh Presidential Nano/Human Interface (NHI) Initiative for accelerating scientific discovery," says Martin P. Harmer, Alcoa Foundation Professor in Lehigh's Department of Materials Science and Engineering and Director of the Nano/Human Interface Initiative.

More information: Symmetry-aware recursive image similarity exploration for materials microscopy, *npj Computational Materials*, [DOI: 10.1038/s41524-021-00637-y](https://doi.org/10.1038/s41524-021-00637-y)

Provided by Lehigh University

Citation: A novel neural network to understand symmetry, speed materials research (2021, October 8) retrieved 26 April 2024 from <https://phys.org/news/2021-10-neural-network-symmetry-materials.html>

<p>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.</p>
--