

A machine learning breakthrough: Using satellite images to improve human lives

July 20 2021, by Edward Lempinen



Deep streams of data from Earth-imaging satellites arrive in databases every day, but advanced technology and expertise are required to access and analyze the data. Now a new system, developed in research based at the University of California, Berkeley, uses machine learning to drive low-cost, easy-to-use technology that one person could run on a laptop, without advanced training, to address their local problems. Credit: NASA

More than 700 imaging satellites are orbiting the earth, and every day they beam vast oceans of information—including data that reflects climate change, health and poverty—to databases on the ground. There's just one problem: While the geospatial data could help researchers and policymakers address critical challenges, only those with considerable wealth and expertise can access it.

Now, a team based at UC Berkeley has devised a [machine learning system](#) to tap the problem-solving potential of [satellite](#) imaging, using low-cost, easy-to-use technology that could bring access and analytical power to researchers and governments worldwide. The study, "A generalizable and accessible approach to machine learning with global satellite imagery," was published today (Tuesday, July 20) in the journal *Nature Communications*.

"Satellite images contain an incredible amount of data about the world, but the trick is how to translate the data into usable insights without having a human comb through every [single image](#)," said co-author Esther Rolf, a final-year Ph.D. student in computer science. "We designed our system for accessibility, so that one person should be able to run it on a laptop, without specialized training, to address their local problems."

"We're entering a regime in which our actions are having truly global impact," said co-author Solomon Hsiang, director of the Global Policy Lab at the Goldman School of Public Policy. "Things are moving faster than they've ever moved in the past. We're changing resource allocations faster than ever. We're transforming the planet. That requires a more responsive management system that is able to see these things happen, so that we can respond in a timely, effective way."

The project was a collaboration between the Global Policy Lab, which Hsiang directs, and Benjamin Recht's research team in the department of Electrical Engineering and Computer Sciences. Other co-authors are

Berkeley Ph.D. graduates Tamma Carleton, now at University of California, Santa Barbara; Jonathan Proctor, now at Harvard's Center for the Environment and Data Science Initiative; Ian Bolliger, now at the Rhodium Group; and Vaishaal Shankar, now at Amazon; and Berkeley Ph.D. student Miyabi Ishihara.

All of them were at Berkeley when the project began. Their collaboration has been remarkable for bringing together disciplines that often look at the world in different ways and speak different languages: computer science, environmental and climate science, statistics, economics and public policy.

But they have been guided by a common interest in creating an open access tool that democratizes the power of technology, making it usable even by communities and countries that lack resources and advanced technical skill. "It's like Ford's Model T, but with machine learning and satellites," Hsiang said. "It's cheap enough that everyone can now access this new technology."

MOSAICS: Improving lives, protecting the planet

The system that emerged from the Berkeley-based research is called MOSAICS, short for Multi-Task Observation using Satellite Imagery & Kitchen Sinks. It ultimately could have the power to analyze hundreds of variables drawn from satellite data—from soil and water conditions to housing, health and poverty—at a global scale.

The research paper details how MOSAICS was able to replicate with reasonable accuracy reports prepared at great cost by the U.S. Census Bureau. It also has enormous potential in addressing development challenges in low-income countries and to help scientists and policymakers understand big-picture environmental change.

"Climate change is diffuse and difficult to see at any one location, but when you step back and look at the broad scale, you really see what is going on around the planet," said Hsiang, who also serves as co-director of the multi-institution Climate Impact Lab.

For example, he said, the satellite data could give researchers deep new insights into expansive rangeland areas such as the Great Plains in the U.S. and the Sahel in Africa, or into areas such as Greenland or Antarctica that may be shedding icebergs as temperatures rise.



In the Indian state of Andhra Pradesh, a satellite image shows hundreds of green aquaculture ponds where local farmers grow fish and shrimp. Geospatial imaging holds enormous potential for developing nations to address challenges related to agriculture, poverty, health and human migration, scholars at UC Berkeley say. But until now, the technology and expertise needed to efficiently access and

analyze satellite data usually has been limited to developed countries. Credit: NASA Earth Observatory images by Joshua Stevens, using Landsat data from the U.S. Geological Survey

"These areas are so large, and to have people sitting there and looking at pictures and counting icebergs is really inefficient," Hsiang explained. But with MOSAIKS, he said, "you could automate that and track whether these glaciers are actually disintegrating faster, or whether this has been happening all along."

For a government in the developing world, the technology could help guide even routine decisions, such as where to build roads.

"A government wants to build roads where the most people are and the most economic activity is," Hsiang said. "You might want to know which community is underserved, or the condition of existing infrastructure in a community. But often it's very difficult to get that information."

The challenge: Organizing trillions of bytes of raw satellite data

The growing fleet of imaging satellites beam data back to Earth 24/7—some 80 terabytes every day, according to the research, a number certain to grow in coming years.

But often, imaging satellites are built to capture information on narrow topics—supplies of fresh water, for example, or the condition of agricultural soils. And the data doesn't arrive as neat, orderly images, like snapshots from a photo shop. It's raw data, a mass of binary information. Researchers who access the data have to know what they're looking for.

Merely storing so many terabytes of data requires a huge investment. Distilling the layers of data embedded in the images requires additional computing power and advanced human expertise to tease out strands of information that are coherent and useful to other researchers, policymakers or funding agencies.

Inevitably, exploiting [satellite images](#) is largely limited to scholars or agencies in wealthy nations, Rolf and Hsiang said.

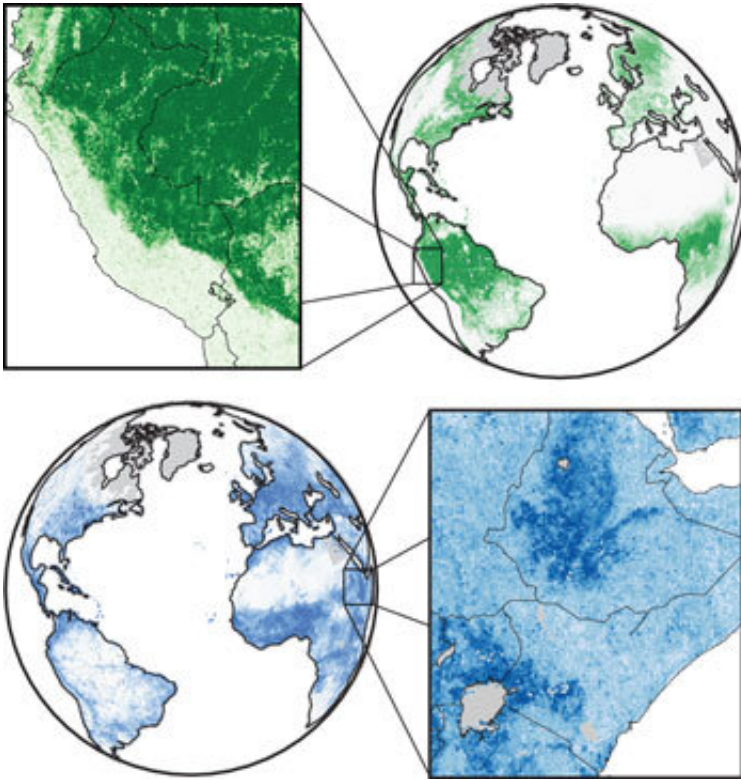
"If you're an elite professor, you can get someone to build your satellite for you," said Hsiang. "But there's no way that a conservation agency in Kenya is going to be able to access the technology and the experts to do this work.

"We wanted to find a way to empower them. We decided to come up with a Swiss Army Knife—a practical tool that everyone can access."

Like Google for satellite imagery, sort of

Especially in low-income countries, one dimension of poverty is a poverty of data. But even communities in the U.S. and other developed countries usually don't have ready access to [geospatial data](#) in a convenient, usable format for addressing local challenges.

Machine learning opens the door to solutions.



The illustrations show how the MOSAIKS machine learning system developed at UC Berkeley predicts, in fine detail, forest cover (above, in green) and population (below). Credit: Esther Rolf, Jonathan Proctor, Tamma Carleton, Ian Bolliger, Miyabi Ishihara, Vaishaal Shankar, Benjamin Recht and Solomon Hsiang

In a general sense, machine learning refers to computer systems that use algorithms and statistical modeling to learn on their own, without step-by-step human intervention. What the new research describes is a system that can assemble data delivered by many satellites and organize it in ways that are accessible and useful.

There are precedents for such systems: Google Earth Engine and Microsoft's Planetary Computer are both platforms for accessing and analyzing global geospatial data, with a focus on conservation. But, Rolf said, even with these technologies, considerable expertise is often

required to convert the data into new insights.

The goal of MOSAIKS is not to develop more complex machine learning systems, Rolf said. Rather, its innovation is in making [satellite data](#) widely useable for addressing global challenges. The team did this by making the algorithms radically simpler and more efficient.

MOSAIKS starts with learning to recognize minuscule patterns in the images—Hsiang compares it to a game of Scrabble, in which the algorithm learns to recognize each letter. In this case, however, the tiles are minuscule pieces of satellite image, 3 pixels by 3 pixels.

But MOSAIKS doesn't conclude "this is a tree" or "this is pavement." Instead, it recognizes patterns and groups them together, said Proctor. It learns to recognize similar patterns in different parts of the world.

When thousands of terabytes from hundreds of sources are analyzed and organized, researchers can choose a village or a country or a region and draw out organized data that can touch on themes as varied as soil moisture, health conditions, human migration and home values.

In a sense, Hsiang said, MOSAIKS could do for satellite databases what Google in the early days did for the Internet: map the data, make it accessible and user-friendly at low cost, and perhaps make it searchable. But Rolf, a [machine learning](#) scholar based in the Berkeley Electrical Engineering and Computer Sciences department, said the Google comparison goes only so far.

MOSAIKS "is about translating an unwieldy amount of data into usable information," she explained. "Maybe a better analogy would be that the system takes very dense information—say, a very large article—and produces a summary."

Creating a living atlas of global data

Both Hsiang and Rolf see the potential for MOSAIKS to evolve in powerful and elegant directions.

Hsiang imagines the data being collected into computer-based, continually evolving atlases. Turn to any given "page," and a user could access broad, deep data about conditions in a country or a region.

Rolf envisions a system that can take the stream of data from humanity's fleet of imaging satellites and remote sensors and transform it into a flowing, real-time portrait of Earth and its inhabitants, continually in a state of change. We could see the past and the present, and so discern emerging challenges and address them.

"We've sent so much stuff to space," Hsiang says. "It's an amazing achievement. But we can get a lot more bang for our buck for all of this [data](#) that we're already pulling down. Let's let the world use it in a useful way. Let's use it for good."

More information: Esther Rolf et al, A generalizable and accessible approach to machine learning with global satellite imagery, *Nature Communications* (2021). [DOI: 10.1038/s41467-021-24638-z](https://doi.org/10.1038/s41467-021-24638-z)

Provided by University of California - Berkeley

Citation: A machine learning breakthrough: Using satellite images to improve human lives (2021, July 20) retrieved 2 May 2024 from <https://techxplore.com/news/2021-07-machine-breakthrough-satellite-images-human.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.