

# Deep learning underlies geographic dataset used in hurricane response

November 1 2022



This image from Sept. 30, 2022, shows how the Federal Emergency Management Agency used ORNL's USA Structures data along with new satellite images to identify structures that were destroyed in Lee County, Florida, during Hurricane Ian. Credit: ORNL, U.S. Dept. of Energy

As Hurricane Fiona made landfall as a Category 1 storm in Puerto Rico on Sept. 18, 2022, some areas of the island were inundated with nearly 30 inches of rain, and power to hundreds of thousands of homes was knocked out. Only 10 days later, Hurricane Ian, a Category 4 storm and one of the strongest and most damaging storms on record, landed in Lee County, Florida, leveling homes and flooding cities before moving up the coast and making landfall again as a Category 1 storm in South Carolina.

Extreme weather and natural disasters are happening with increasing frequency across the United States and its territories. Accurate and detailed maps are critical in [emergency response](#) and recovery.

Even before the hurricanes made landfall, the Federal Emergency Management Agency was working with researcher Lexie Yang and her team at the Department of Energy's Oak Ridge National Laboratory to forecast potential damage and accelerate on-the-ground response using U.S. Structures, a massive dataset of building outlines and attributes covering more than 125 million structures.

Over the past seven years, researchers in ORNL's Geospatial Science and Human Security Division have mapped and characterized all structures within the United States and its territories to aid FEMA in its response to disasters. This dataset provides a consistent, nationwide accounting of the buildings where people reside and work.

The agency requested two new attributes for the data the same day Fiona made landfall: occupancy types and addresses, critical information in speeding federal emergency funds to households and businesses.

"We encountered some language barriers when we were adding the new data: the [limited information](#) that was available to us was in Spanish. In addition, there are many different ways of documenting Puerto Rico's

addresses. Having to unify those data and validate the attribution information was a unique challenge for us," Yang said.

Even with that challenge, Yang's team was able to translate, validate and conflate the new attributes to the U.S. Structures data in about 50 hours. This is the result of having a scalable information pipeline and database in place built from years of effort. FEMA began planning for its response using the baseline U.S. Structures maps of areas likely to be impacted. FEMA staff added layers of data as the disasters unfolded, allowing the agency to prioritize response to the most heavily impacted areas.

"FEMA has GIS [[geographic information systems](#)] analysts that take our data and integrate it with post-disaster satellite imagery, aerial imagery and information that first responders are collecting in the field," said ORNL's Carter Christopher, section head for Human Dynamics in the Geospatial Science and Human Security Division.

The existing dataset, paired with real-time impact information, can speed recovery by supporting damage assessments that property owners need in order to receive funds for rebuilding in days rather than weeks or months.

"Our team is extremely proud to be part of this project," Yang said. "We see how our technical capabilities and knowledge can transform the dataset used by FEMA and local stakeholders."

U.S. Structures got its start in 2015, when former ORNL researchers Mark Tuttle and Melanie Laverdiere were working on a FEMA project to map mobile home parks in the U.S. Mobile homes are particularly vulnerable to [natural disasters](#), and little data existed identifying the location of these at-risk structures.

The team used deep learning, a subset of machine learning, to process images and compile the data. Machine learning uses computers to detect patterns in massive amounts of data, then makes predictions based on what the computer learns from those patterns. In deep learning, the computing system creates its own algorithms rather than using algorithms developed and input by a human.

After the national mobile homes parks database was compiled, FEMA requested a more comprehensive structures database.

The process began with a stream of high-resolution images from a commercial satellite imagery provider and some preprocessing. The raw imagery needed to be matched to actual terrain variations—a process called orthorectification—and sharpened to improve the resolution. That process took the image from a [spatial resolution](#) of 2 to 3 meters to the 0.3 meters needed for feature extraction.

The spatial resolution is similar to that seen on Google Maps; items that are a few meters in size are recognizable to the human eye. Once prepped, the images entered a feature extraction pipeline hosted by a GPU cluster within ORNL's Compute and Data Environment for Science, or CADES, which offers high-performance data services for researchers throughout the lab.

To get the deep learning model started, scientists gave the system a range of marked-up images, or training data, to study. Working as a deep neural network, the machine learning model trained itself to analyze similar inputs.

To date, more than 59,000 training examples representing a broad and diverse range of geographic features have been incorporated into the U.S. Structures model. As the team began work on a new state, they prepped the training set with new, region-specific examples in addition



to the cumulative training data for the states that came before it.

The gains in output over the last few years came from ORNL's continuously improved hardware and compute power, advancements made in deep learning, and a growing volume of training data informing the artificial intelligence-based model. As the project progressed, the maps became more accurate, requiring less human intervention, and the time it took to process the images got shorter and shorter.

Convolutional neural networking compressed a process that would have taken many years by human hand into minutes. To date, the team has processed 1.1 petabytes of imagery—stitching together and describing the equivalent of a billion digital photographs.

After the feature extraction was complete, the researchers drew from commercial parcel data vendors to conflate land-use information directly onto the U.S. Structures building features.

"That additional information, when available, makes the structures data more powerful. Is that square a house, a warehouse or a church? Each of those has different implications in a disaster," said Christopher.

If no reliable land-use data was available, the team used a separate machine learning model to distinguish residential from non-residential structures. Structures also are described with other attributes such as a unique building identifier, square footage, longitude and latitude.

"We take a lot of time verifying that whatever we're handing off to FEMA is the highest quality that we can provide," Yang said.

This powerful [open-source dataset is publicly available](#) from the U.S. government's GeoPlatform. Additionally, the U.S. Geological Survey has added the data to [the National Map](#), a collaborative effort among U.S.

agencies and partners to deliver topographic information. The ORNL team hopes open access to the data will be useful to academic institutions for research and to small municipal agencies for risk planning.

"A lot of rural counties and small jurisdictions may not have the budget to collect or purchase this kind of data otherwise," Christopher said. "It could be used by first responders or basic services providers. It could also be applied to needs at a county level for town planning or property appraisals."

ORNL researchers on the project include Taylor Hauser, Benjamin Swan, Andrew Reith and Matthew Whitehead. Other contributors include Brad Miller, Matthew Crockett and Katie Heying.

In the project's next phase, the team expects to populate the two key attributes—occupancy types and addresses—for the rest of the states and tackle height and elevation information needed for flood modeling.

Building out a sustainable process to detect and incorporate changes over time will be key to extending the lifetime of the dataset. Additionally, this powerful model could be used for similar purposes across the world in [disaster planning](#) and response or paired with other sensing technology to extract other useful information.

Chris Vaughan, Yang's project partner at FEMA, has been an enthusiastic advocate for U.S. Structures, promoting its use and touting the data's consistent scheme and accessibility.

"Disaster operations require a standardized and accessible structure dataset to help streamline assistance to survivors. ORNL's work on U.S. Structures has helped us share incident data with our interagency partners like never before," Vaughan said. "In addition, they are helping

us close long-standing data gaps related to vulnerable populations, which is a top priority for our team."

Yang has seen growing interest from federal agencies, research organizations, local governments and practitioners not only in using the data set, but also in contributing and incorporating data from smaller local projects.

"This project is still evolving, and we expect to continue to have major updates to the current data," she said. "We hope that more communities will use the data. It's already proven to be valuable through FEMA's work, but there may be other applications that are even more impactful."

Provided by Oak Ridge National Laboratory

Citation: Deep learning underlies geographic dataset used in hurricane response (2022, November 1) retrieved 12 May 2024 from <https://phys.org/news/2022-11-deep-underlies-geographic-dataset-hurricane.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.