LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MIT

LLSC News
Lincoln Laboratory Supercomputing Center

April 16, 2024

An AI Dataset Carves New Paths to Tornado Detection

Tornadoes are violent mysteries. A public artificial intelligence dataset could help models reveal when and why they form, improving forecasters' ability to issue warnings.

Kylie Foy | Communications & Community Outreach Office

The return of spring touches off tornado season. A tornado's twisting funnel of dust and debris seems an unmistakable sight. But that sight can be obscured to radar, the tool of meteorologists. It's hard to know exactly when a tornado has formed, or even why.

A new dataset could hold answers. It contains radar returns from thousands of tornadoes that have hit the United States in the past 10 years. Storms that spawned tornadoes are flanked by other severe storms, some with nearly identical conditions, that never did. Lincoln Laboratory researchers who curated the dataset, called TorNet, have now released it open source. They hope to enable breakthroughs in detecting one of nature's most mysterious and violent phenomena.

"A lot of progress is driven by easily available, benchmark datasets. We hope TorNet will lay a foundation for machine learning algorithms to both detect and predict tornadoes," says Mark Veillette, the project's co-principal investigator with James Kurdzo. Both researchers work in the Air Traffic Control Systems Group.

Along with the dataset, the team is releasing models trained on it. The models show promise for machine learning's ability to spot a twister. Building on this work could open new frontiers for forecasters, helping them provide more accurate warnings that might save lives.

Swirling uncertainty
About 1,200 tornadoes occur in the United States every year, causing millions to billions in economic damage and

Reflectivity Factor (dBZ)   **Radial Velocity (m/s)**
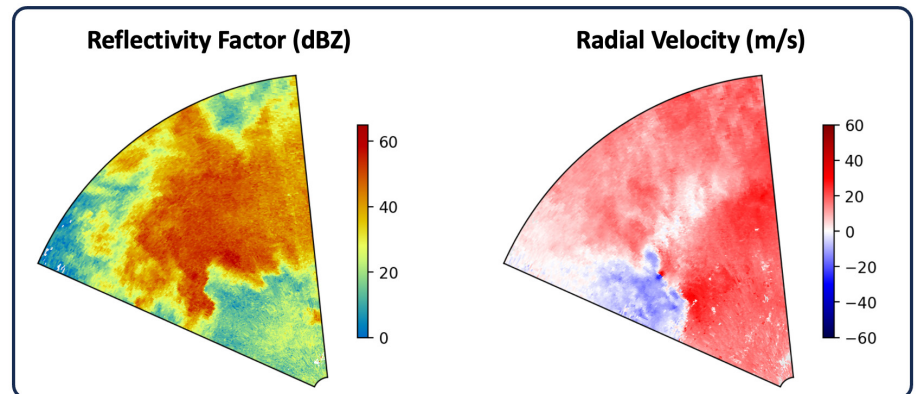
Each sample in the TorNet dataset includes six types of radar images, portraying different radar data products. The images shown here are two of those products, including reflectivity factor and radial velocity from a tornadic example in the dataset.

claiming 71 lives on average. Last year, one unusually long-lasting tornado killed 17 people and injured at least 165 others along a 59-mile path in Mississippi.

Yet tornadoes are notoriously difficult to forecast because scientists don't have a clear picture of why they form. "We can see two storms that look identical, and one will produce a tornado and one won't. We don't fully understand it," Kurdzo says.

A tornado's basic ingredients are thunderstorms with instability caused by rapidly rising warm air and wind shear that causes rotation. Weather radar is the primary tool used to monitor these conditions. But tornadoes lay too low to be detected, even when moderately close to the radar. As the radar beam with a given tilt angle travels further from the antenna,

it gets higher above the ground, mostly seeing reflections from rain and hail carried in the "mesocyclone," the storm's broad, rotating updraft. A mesocyclone doesn't always produce a tornado.

With this limited view, forecasters must decide whether or not to issue a tornado warning. They often err on the side of caution. As a result, the rate of false alarms for tornado warnings is more than 70%. "That can lead to boy-who-cried-wolf syndrome," Kurdzo says.

In recent years, researchers have turned to machine learning to better detect and predict tornadoes. However, raw datasets and models have not always been accessible to the broader community, stifling progress. TorNet is filling this gap. The dataset contains more than

200,000 radar images, 13,587 of which depict tornadoes. The rest of the images are non-tornadic, taken from storms in one of two categories: randomly selected severe storms or false-alarm storms (those that led a forecaster to issue a warning but that didn't produce a tornado).

A challenge in curating the dataset was first finding tornadoes. Within the entire corpus of weather radar data, tornadoes are extremely rare events. The team then had to balance those tornado samples with difficult non-tornado samples. If the dataset were too easy, say by comparing tornadoes to snowstorms, an algorithm trained on the data would likely over-classify storms as tornadic.

"What's beautiful about a true benchmark dataset is that we're all working with the same data, with the same level of difficulty, and can compare results," Veillette says. "It also makes meteorology more accessible to data scientists, and vice versa. It becomes easier for these two parties to work on a common problem."

Both researchers represent the progress that can come from cross-collaboration. Veillette is a mathematician and algorithm developer who has long been fascinated by tornadoes. Kurdzo is a meteorologist by training and a signal processing expert. In grad school, he chased tornadoes with custom-built mobile radars, collecting data to analyze in new ways.

"This dataset also means that a grad student doesn't have to spend a year or two building a dataset. They can jump right into their research," Kurdzo says.

Chasing answers with deep learning

Using the dataset, the researchers developed baseline artificial intelligence (AI) models. They were particularly eager to apply deep learning, a form of machine learning that excels at processing visual data. On its own, deep learning can extract features (key observations that an algorithm uses to make a decision) from images across a dataset. Other machine learning approaches require humans to first manually label features.

"We wanted to see if deep learning could rediscover what people normally look for in tornadoes and even identify new things that typically aren't searched for by forecasters," Veillette says.

The results are promising. Their deep learning model performed similar or better than all tornado-detecting algorithms known in literature. The trained algorithm correctly classified 50% of weaker EF-1 tornadoes and over 85% of tornadoes rated EF-2 or higher, which make up the most devastating and costly occurrences of these storms.

They also evaluated two other types of machine learning models, and one traditional model to compare against. The source code and parameters of all these models are freely available. The models and dataset are also described in a paper submitted to a journal of the American Meteorological Society (AMS). Veillette presented this work at the AMS Annual Meeting in January.

"The biggest reason for putting our models out there is for the community to improve upon them and do other great things," Kurdzo says. "The best solution could be a deep learning model, or someone might find that a non–deep learning model is actually better."

TorNet could be useful in the weather community for others uses too, such as for conducting large-scale case studies on storms. It could also be augmented with other data sources, like satellite imagery or lightning maps. Fusing multiple types of data could improve the accuracy of machine learning models.

This project was funded by Lincoln Laboratory's Climate Change Initiative, which aims to leverage the Laboratory's diverse technical strengths to help address climate problems threatening human health and global security.

Taking steps toward operations

On top of detecting tornadoes, Kurdzo hopes that models might help unravel the science of why they form.

"As scientists, we see all these precursors to tornadoes — an increase in low-level rotation, a hook echo in reflectivity data, specific differential phase (KDP) foot and differential reflectivity (ZDR) arcs. But how do they all go together? And are there physical manifestations we don't know about?" he asks.

Teasing out those answers might be possible with explainable AI. Explainable AI refers to methods that allow a model to provide its reasoning, in a format understandable to humans, of why it came to a certain decision. In this case, these explanations might reveal physical processes that happen before tornadoes. This knowledge could help train forecasters, and models, to recognize the signs sooner.

"None of this technology is ever meant to replace a forecaster. But perhaps someday it could guide forecasters' eyes in complex situations, and give a visual warning to an area predicted to have tornadic activity," Kurdzo says.

Such assistance could be especially useful as radar technology improves and future networks potentially grow denser. Data refresh rates in a next-generation radar network are expected to increase from every five minutes to approximately one minute, perhaps faster than forecasters can interpret the new information. Because deep learning can process huge amounts of data quickly, it could be well-suited for monitoring radar returns in real time, alongside humans. Tornadoes can form and disappear in minutes.

But the path to an operational algorithm is a long road, especially in safety-critical situations, Veillette says. "I think the forecaster community is still, understandably, skeptical of machine learning. One way to establish trust and transparency is to have public benchmark datasets like this one. It's a first step."

The next steps, the team hopes, will be taken by researchers across the world who are inspired by the dataset and energized to build their own algorithms. Those algorithms will in turn go into test beds, where they'll eventually be shown to forecasters, to start a process of transitioning into operations.

In the end, the path could circle back to trust.

"We may never get more than a 10- to 15-minute tornado warning using these tools. But if we could lower the false-alarm rate, we could start to make headway with public perception," Kurdzo says. "People are going to use those warnings to take the action they need to save their lives."

MIT LINCOLN LABORATORY
SUPERCOMPUTING CENTER