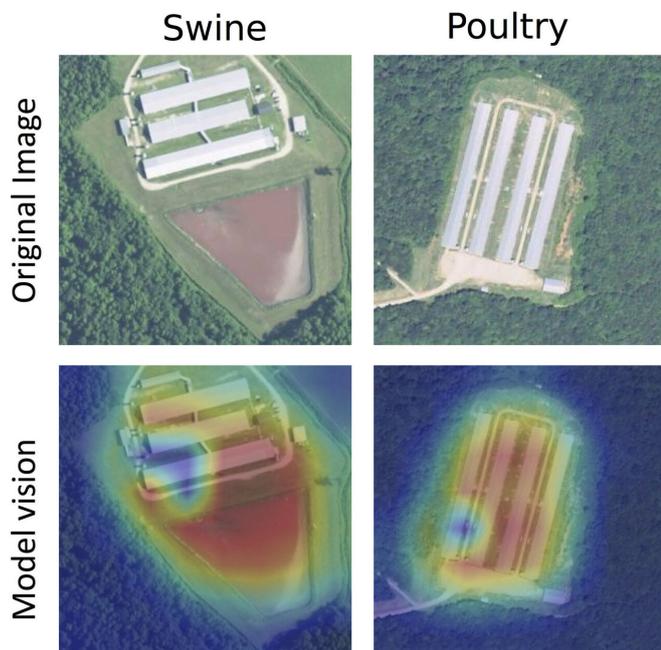


How machine learning can help regulators

9 April 2019, by Melissa De Witte



Sample swine (left) and poultry (right) facilities, with the original image (top) and a heat map of the way the algorithmic models processed the image (bottom). The red areas show where the model detected the likelihood of facility locations. Credit: National Agriculture Imagery Program / U.S. Department of Agriculture

How to locate potentially polluting animal farms has long been a problem for environmental regulators. Now, Stanford scholars show how a map-reading algorithm could help regulators identify facilities more efficiently than ever before.

Law Professor Daniel Ho, along with Ph.D. student Cassandra Handan-Nader, have figured out a way for [machine learning](#) – teaching a computer how to identify and analyze patterns in data – to efficiently locate industrial animal operations and help regulators determine each facility's environmental risk. The researchers' findings are set to publish April 8 in *Nature Sustainability*.

"Our work shows how a [government agency](#) can

leverage rapid advances in computer vision to protect [clean water](#) more efficiently," said Ho, the William Benjamin Scott and Luna M. Scott Professor of Law, and a senior fellow at the Stanford Institute for Economic Policy Research.

A basic problem, with complex consequences

According to the Environmental Protection Agency (EPA), agriculture is the leading contributor of pollutants into the nation's water supply, with substantial [pollution](#) believed to be emanating from large-scale, concentrated animal feeding operations, known also as CAFOs.

But environmental monitoring efforts have been stymied by a basic problem: Regulators have no systematic way of determining where CAFOs are located, Ho said. The United States Government Accountability Office reports that no federal agency has reliable information on the number, size and location of large-scale agricultural operations.

While the Clean Water Act does require some federal permitting, it only applies to operations that actually discharge pollutants into U.S. waterways – not facilities that could potentially cause contamination – intentionally or not, Ho said.

With no definite list to turn to, efforts to monitor potentially polluting facilities are difficult and, in some cases, impossible.

"This information deficit stifles enforcement of the environmental laws of the United States," Ho said.

Some environmental and public interest groups have tried to identify facilities themselves by scanning terrain manually or poring over aerial photos, but they have found it an incredibly time-intensive task. It took one environmental group over three years to look at images from just one state. Monitoring efforts like these could never scale or be done in real time, Ho said.

Using big data to fill in the gaps

Ho and Handan-Nader, then a research fellow at Stanford Law School and now pursuing a doctorate in political science, turned their attention to a type of artificial intelligence called deep learning. A subset of machine learning, deep learning algorithms have revolutionized the ability to detect complex objects in imagery.

With the help of several open source tools and a team of students in economics and computer science to assist with data analysis, Ho and Handan-Nader were able to retrain an existing image-recognition model to recognize large-scale animal facilities by using information collected by two nonprofit groups and publicly available satellite images from the USDA's National Agricultural Imagery Program (NAIP). The researchers focused on trying to identify poultry facilities in North Carolina because most are not required to obtain permits, Ho said.

The model, already savvy in scanning images based on an enormous corpus of digital images, was retrained to pick up on similar clues that the environmental organizations had been manually monitoring. For example, swine farms were identifiable by compact rectangular barns abutted by large liquid manure pits, and poultry by long rectangular barns and dry manure storage. By homing in on these prominent features, the model was also able to provide size estimates for the facilities.

The researchers found that their algorithm was able to identify 15 percent more poultry farms than what was originally found through manual endeavors. And because their approach could scale across years of NAIP imagery, their algorithm was able to accurately estimate growth within the vicinity of a recently constructed feed mill.

"The model detected 93 percent of all poultry CAFOs in the area, and was 97 percent accurate in determining which ones appeared after the feed mill opened," Handan-Nader and Ho write in the paper.

Complementary, interdisciplinary approach

Ho and Handan-Nader hope that machine learning

can complement the human monitoring efforts of environmental agencies and interest groups.

"Now all kinds of researchers with programming ability can harness these open-source tools for novel applications," said Handan-Nader, a co-author on the paper. "You can stand on the shoulders of giants and expand on what experts in these kinds of machine learning techniques have done."

Using machine learning for rote tasks can free people to do more complex ones, such as determining the possible environmental hazards of a facility, Handan-Nader said. The researchers estimated that their algorithm could capture 95 percent of existing large-scale facilities using fewer than 10 percent of the resources required for a manual census.

Ho and Handan-Nader hope that, eventually, advances in aerial imagery will enable a computer model to detect actual discharge into [waterways](#).

"Increasingly, complex social problems cannot be solved from the confines of a narrow discipline alone, and the ability to leverage innovation cross-campus can help address core problems of law and public policy," Ho said.

More information: Deep learning to map concentrated animal feeding operations. *Nature Sustainability*. doi.org/10.1038/s41893-019-0246-x

Provided by Stanford University

APA citation: How machine learning can help regulators (2019, April 9) retrieved 19 June 2019 from <https://phys.org/news/2019-04-machine.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.