Citizen science projects have a surprising new partner—the computer

February 6 2019

After being shown thousands of images, the computer starts to recognize the patterns, edges, and parts of the animal, like this elephant trunk. Credit: Snapshot Serengeti

For more than a decade, citizen science projects have helped researchers

use the power of thousands of volunteers who help sort through datasets that are too large for a small research team. Previously, this data generally couldn't be processed by computers because the work required skills that only humans could accomplish.

Now, computer machine learning techniques that teach the computer specific image recognition skills can be used in crowdsourcing projects to deal with massively increasing amounts of data—making computers a surprising new partner in citizen science projects.

The research, led by the University of Minnesota-Twin Cities, was chosen as the cover story for the most recent issue of the British Ecological Society's scientific journal *Methods in Ecology and Evolution*.

In this study, data scientists and citizen science experts partnered with ecologists who often study wildlife populations by deploying camera traps. These camera traps are remote, independent devices, triggered by motion and infrared sensors that provide researchers with images of passing animals. After collection, these images have to be classified according to the study's goals to produce useful ecological data for analysis.

"In the past, researchers asked citizen scientists to help them process and classify the images within a reasonable time-frame," said the study's lead author Marco Willi, a recent graduate of the University of Minnesota master's program in data science and researcher in the University's School of Physics and Astronomy. "Now, some of these recent camera trap projects have collected millions of images. Even with the help of citizen scientists, it could take years to classify all of the images. This new study is a proof of concept that machine learning techniques can help significantly reduce the time of classification."

Recent camera trap projects have collected millions of images like this image of a giraffe. Without the help of computers, it could take researchers years to classify all of the images, even with the help of citizen scientists. Credit: Snapshot Serengeti

Researchers used three datasets of images collected from Africa—Snapshot Serengeti, Camera CATalogue, and Elephant Expedition—and one dataset from Snapshot Wisconsin with images collected in North America. The datasets each featured between nine and 55 species and exhibited significant differences in how often various species were photographed. These datasets also differed in aspects such as dataset size, camera placement, camera configuration, and species coverage which allows for drawing more general conclusions.

The researchers used machine learning techniques that teach the computer how to classify the images by showing the computer datasets of images already classified by humans. For example, the machine would be shown full and partial images that are known to be images of zebras from various angles. The computer then would start to recognize the patterns, edges, and parts of the animal, and learn how to identify the image as a zebra. The researchers can also build upon some of these skills to help computers identify other animals, such as a deer or squirrel, with even fewer images.

The computer also learns to identify empty images, which are images without animals where the cameras were usually set off by vegetation blowing in the wind. In some cases, these empty images make up about 80 percent of all camera trap images. Eliminating all the empty images can greatly speed the classification process.

The computer's accuracy rates for identifying empty images across projects range between 91.2 percent and 98.0 percent, while accuracies for identifying specific species are between 88.7 percent and 92.7 percent. While the computer's classification accuracy is low for rare species, the computer can also tell researchers how confident it is in its predictions. Removing low-confidence predictions increases the computer's accuracies to the level of citizen scientists.

"Our machine learning techniques allow ecology researchers to speed up the image classification process and pave the way for even larger citizen science projects in the future," Willi said. "Instead of every image having to be classified by multiple volunteers, one or two volunteers could confirm the computer's classification."

The datasets used by researchers each featured between nine and 55 species and exhibited significant differences in how often various species were photographed. Credit: Panthera

While this study focused on ecology camera trap programs, Willi said the same techniques can also be used in other citizen science projects such as classifying images from space.

"Data in a wide range of science areas is growing much faster than the number of citizen science project volunteers," said study co-author Lucy Fortson, a University of Minnesota physics and astronomy professor and co-founder of Zooniverse, the largest citizen science online platform that hosted the projects in the study. "While there will always be a need for human effort in these projects, combining these efforts with the help of

Big Data techniques can help researchers process more data even faster and allows the volunteers to focus on the harder, rarer classifications."

Led by Fortson, the Zooniverse team at the University of Minnesota, including Willi, is working to integrate machine learning techniques into the platform so the hundreds of researchers from astronomy to zoology using the platform can take advantage of them.

In addition to researchers at the University of Minnesota, the international team on this study included researchers from the University of Oxford, Wisconsin Department of Natural Resources, Institute for Communities and Wildlife in Africa, Adler Planetarium, and the conservation organization Panthera.

To read the full research study entitled "Identifying Animal Species in Camera Trap Images using Deep Learning and Citizen Science," visit the *Methods in Ecology and Evolution* website.

Provided by University of Minnesota