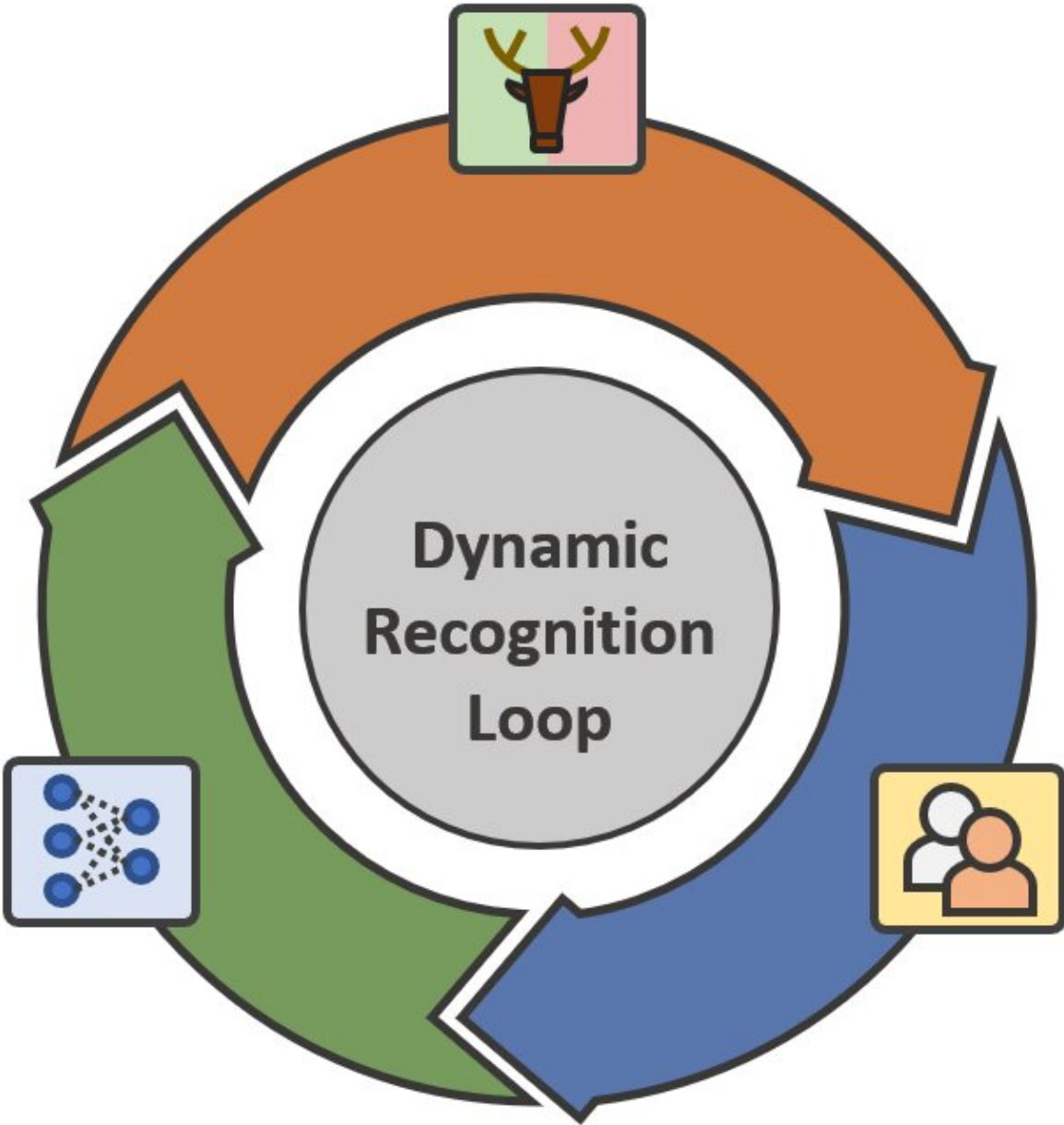


A framework to automatically identify wildlife in collaboration with humans

November 2 2021, by Ingrid Fadelli



In real-world applications, AI models do not stop at one training stage. As data collection progresses over time, there is a continuous cycle of inference, annotation, and model updating. When there are novel and difficult samples, human annotation is inevitable. Credit: Miao et al.

Over the past few decades, computer scientists have developed numerous machine learning tools that can recognize specific objects or animals in images and videos. While some of these techniques have achieved remarkable results on simple animals or items (e.g., cats, dogs, houses), they are typically unable to recognize wildlife and less renowned plants or animals.

Researchers at University of California, Berkeley (UC Berkeley) have recently developed a new wildlife identification approach that performs far better than techniques developed in the past. The approach, presented in a paper published in *Nature Machine Intelligence*, was conceived by Zhongqi Miao, who initially started exploring the idea that artificial intelligence (AI) tools could classify wildlife images collected by movement-triggered [camera traps](#). These are cameras that wildlife ecologists and researchers often set up to monitor species inhabiting specific geographic locations and estimate their numbers.

The [effective use](#) of AI for identifying species in wildlife images captured by camera traps could significantly simplify the work of ecologists and reduce their workload, preventing them from having to look through hundreds of thousands of images to generate maps of the distribution of species in specific locations. The framework developed by Miao and his colleagues is different from other methods proposed in the past, as it merges machine learning with an approach dubbed

'humans in the loop' to generalize better on real-world tasks.

"An important aspect of our 'humans in the loop innovation' is that it addresses the 'long-tailed distribution problem,'" Wayne M. Getz, one of the researchers who carried out the study, told TechXplore. "More specifically, in a set of hundreds of thousands of images generated using camera traps deployed in an area over a season, images of common species may appear hundreds or even thousands of times, while those of rare species may appear just a few times. This produces a long-tailed distribution of the frequency of images of different species."

If all species were captured by camera traps with equal frequency, their distribution would be what is known as 'rectangular.'" On the other hand, if these frequencies are highly imbalanced, the most common frequencies (plotted first down the y-axis) would be far larger than least common frequencies (plotted at the bottom of the graph), resulting in a long-tailed distribution.

"If standard AI image recognition software were applied to long-tailed distributional data, then the method would fail miserably when it comes to identifying rare species," Getz explained. "The primary purpose of our study was to find a way to improve the identification of rare species by incorporating humans into the process in an iterative manner."

When trying to apply conventional AI tools in real-world settings, computer scientists can encounter several challenges. As mentioned by Getz, the first is that data collected in the real world often follows a long-tail distribution and current state-of-the-art AI models do not perform as well on this data, compared to data with a rectangular or normal distribution.

"In other words, when applied to data with a long-tailed distribution, large or more frequent categories always lead to much better

performance than smaller and rare categories," Miao, lead author of the paper, told TechXplore. "Furthermore, instances of rare categories (especially images of rare animals) are not easy to collect, making it even harder to get around this long-tail distribution issue through data collection."

Another challenge of applying AI in real-world settings is that the problems they are meant to solve are typically open-ended. For instance, wildlife monitoring projects can continue indefinitely and span across long periods of time, during which new camera traps will be set up and a variety of new data will be collected.

In addition, new animal species might suddenly appear in the sites monitored by the cameras due to several possible factors, including unexpected invasions, animal reintroduction projects or recolonizations. All of these changes will be reflected in the data, ultimately impairing the performance of pre-trained machine learning techniques.

"So far, the human contribution to the training of AI has been inevitable," Miao said. "As real-world applications are open-ended, ensuring that AI models learn and adapt to new content requires additional human annotations, especially when we want the models to identify new animal species. Thus, we think there is a loop of AI recognition system of new data collection, human annotation on new data and model update to the novel categories."

In their previous research, the researchers tried to address the factors impairing the performance of AI in real-world settings in several different ways. While the approaches they devised were in some ways promising, their performance was not as good as they had hoped, achieving a classification accuracy below 70 percent when tested on standardized long-tailed datasets.

"It's hard for people to trust an AI model that could only produce ~70 percent accuracy," Miao said. "Overall, we think a deployable AI model should: achieve a balanced performance across imbalanced distribution (long-tailed recognition), be able to adapt to different environments (multi-domain adaptation), be able to recognize novel samples (out-of-distribution detection), and be able to learn from novel samples as fast as possible (few-shot learning, life-long learning, etc.). However, each one these characteristics have proved difficult to realize, and none of them have been fully solved yet, let alone combining them together and coming up with a perfect AI solution."

Instead of using renowned and existing AI tools or trying to develop an 'ideal' method, therefore, Miao and his colleagues decided to create a highly performing tool that relies on a certain amount of input from humans. As so far human annotations on data have proved to be particularly valuable for enhancing the performance of deep learning-based models, they focused their efforts on maximizing their efficiency.

"The goal of our project was to minimize the need for human intervention as much as possible, by applying human annotation solely on difficult images or novel species, while maximizing the recognition performance/accuracy of each model update procedure (i.e., update efficiency)," Miao said.

By combining machine learning techniques with human efforts in an efficient way, the researchers hoped to achieve a system that was better at recognizing animals in real-world wildlife images, overcoming some of the issues they encountered in their past studies. Remarkably, they found that their method could achieve 90 percent accuracy on wildlife image classification tasks, using 1/5 of the annotations that standard AI approaches would require to achieve this accuracy.

"Putting AI techniques into practice has always been significantly

challenging, no matter how promising theoretical results are in previous studies on standard datasets," Miao said. "We thus tried to propose an AI recognition framework that can be deployed in the field even when the AI models are not perfect. And our solution is to introduce efficient human efforts back into the recognition system. And in this project, we use wildlife recognition as a practical use case of our framework."

Instead of evaluating AI models using a single dataset, the framework devised by Miao and his colleagues focuses on how efficiently a previously trained model can analyze newly collected datasets containing images of previously unobserved species. Their approach incorporates an active learning technique, which uses a prediction confidence metric to select low-confidence predictions, so that they can be annotated further by humans. When a model identifies animals with high levels of confidence, on the other hand, their framework stores these predictions as pseudo labels.

"Models are then updated according to both human annotations and pseudo labels," Miao explained. "The model is evaluated based on: the overall validation accuracy of each category after the update (i.e., update performance); percentage of high-confidence predictions on validation (i.e., saved human effort for annotation); accuracy of high-confidence predictions; and percentage of novel categories that are detected as low-confidence predictions (i.e., sensitivity to novelty)."

The overall aim of the optimization algorithm used by Miao and his colleagues is to minimize human efforts (i.e., to maximize a model's high-confidence percentage), while maximizing performance and accuracy. Technically speaking, the researchers' framework is a combination of active learning and semi-supervised learning with humans in the loop. All of the codes and data used by Miao and his colleagues are publicly available and [can be accessed online](#).

"We proposed a deployable human-machine recognition framework that is also applicable when the models are not perfectly performing by themselves," Miao said. "With the iterative human-machine updating procedure, the framework can keep updated be deployed when new data are continuously collected. Furthermore, each technical component in this framework can be replaced with more advanced methods in the future to achieve better results."

The experimental setting outlined by Miao and his colleagues is arguably more realistic than those considered in previous works. In fact, instead of focusing on a single cycle of model training, validation and testing, it focuses on numerous cycles or stages, which allows models to better adapt to changes in the data.

"Another unique aspect of our work is that we proposed a synergistic relationship between humans and machines," Miao said. "Machines help relieve the burden of humans (e.g., ~80 percent annotation requirements), and humans help annotate novel and challenging samples, which are then used to update the machines, such that the machines are more powerful and more generalized in the future. This is a continuous and long-term relationship."

In the future, the framework introduced by this team of researchers could allow ecologists to monitor animal species in different places more efficiently, reducing the time they spend examining images collected by trap cameras. In addition, their framework could be adapted to tackle other real-world problems that involve the analysis of data with a long-tailed distribution or that continuously changes over time.

"Miao is now working on the problem of trying to identify species from satellite or aerial images which present two challenges compared with camera trap images: the resolution is much lower because cameras are much more distant from the subjects that are capturing and the

individual being imaged may be one of many in the overall frame; images generally show only a 1-d projection (i.e., from the top) rather than the 2-d projections (front/back and leftside/rightside) of camera trap data," Getz said.

Miao, Getz and their colleagues now also plan to deploy and test the framework they created in real-world settings, such as camera trap wildlife monitoring projects in Africa organized by some of their collaborators. Meanwhile, Miao is working on other deep learning tools for the analysis of aerial images and audio recordings, as these could be particularly useful for identifying birds or marine animals. His overall goal is to make deep learning more accessible for ecologists and researchers analyzing wildlife images.

"On a broader scale, we think that the synergistic relationship between humans and machines is an exciting topic and that the goal of AI research should be to develop tools that augment people's abilities (or intelligence), rather than to eliminate the existence of humans (e.g., looking for perfect machines that can handle everything without the need for humans)," Miao added. "It is more like a loop where machines make humans better, and humans make machines more powerful in return, just like in the iterative framework we proposed in the paper. We call this Artificial Augmented Intelligence (A2I or A-square I), where ultimately, people's intelligence is augmented with artificial intelligence and vice versa. In the future, we want to keep exploring the possibilities of A2I."

More information: Zhongqi Miao et al, Iterative human and automated identification of wildlife images, *Nature Machine Intelligence* (2021). [DOI: 10.1038/s42256-021-00393-0](https://doi.org/10.1038/s42256-021-00393-0)

Ziwei Liu et al, Large-scale long-tailed recognition in an open world. arXiv:1904.05160v2 [cs.CV], arxiv.org/abs/1904.05160

Ziwei Liu et al, Open compound domain adaptation.
arXiv:1909.03403v2 [cs.CV], arxiv.org/abs/1909.03403

© 2021 Science X Network

Citation: A framework to automatically identify wildlife in collaboration with humans (2021, November 2) retrieved 3 May 2024 from <https://techxplore.com/news/2021-11-framework-automatically-wildlife-collaboration-humans.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.