Using machine learning to provide automatic image annotation for wildlife camera traps in the Arctic

Håvard Thom[a], John Markus Bjørndalen[a], Eivind Flittie Kleiven[b], Eeva Marjatta Soininen[b], Siw Turid Killengreen[b], Dorothee Ehrich[b], Rolf A. Ims[b], Otto Anshus[a], Alexander Horsch[a]

[a] Department of Computer Science, University of Tromsø - The Arctic University of Norway

[b] Department of Arctic and Marine Biology, University of Tromsø - The Arctic University of Norway

Abstract

The arctic tundra is considered the terrestrial biome expected to be most impacted by climate change, with temperatures projected to increase as much as 10 °C by the turn of the century. The Climate-ecological Observatory for Arctic Tundra (COAT) project monitors the climate and eco-systems using several sensor types. We report on results from projects that automate image annotations from two of the camera traps used by COAT: an artificial tunnel under the snow for capturing information about small mammals, and an open-air camera trap using bait that captures information of a range of larger sized birds and mammals. These traps currently produce over two million pictures per year.

We have developed and trained several Convolutional Neural Network (CNN) models to automate annotation of images from these camera traps. Results show that we get a high accuracy: 97.84% for tunnel traps, and 94.1% for bait traps. This exceeds previous state of the art in animal identification on camera trap images, and is at a level where we can already relieve experts from manual annotation of images.

Introduction

The arctic tundra is expected to be more challenged by climate change than any other terrestrial biome, with models projecting an average temperature increase in the Arctic as large as 10 °C by the turn of the century [1]. The Climate-ecological Observatory for Arctic Tundra (COAT) [1] [2] is a long-term research initiative for real time detection, documentation and understanding of climate impacts on terrestrial arctic ecosystems. COAT is a collaboration of several Norwegian research institutions under the umbrella of FRAM - High North Centre for Climate and Environment. The study areas include the bioclimatic extremes of the terrestrial Arctic, the low-arctic coast of the Norwegian mainland and the high-arctic Svalbard archipelago.

COAT uses several types of sensors for observations, including satellites, snow observations, meteorological observations, camera traps, and microphones. Manual annotation and analytics of images from the camera traps is by now already a demanding and time consuming task, and this burden will further increase as more traps are deployed. This paper reports on results from using machine learning techniques on two separate sets of camera trap images: camera traps used in small mammal tunnels underneath the snow and baited camera traps above the snow. These traps are expected to capture more than two million pictures per year with the current and near future set of cameras.

Small mammal camera traps

Figure 1: Image of a tundra vole from one of the COAT camera traps under the snow.

One of the camera traps used is an artificial tunnel (see Figure 1) with a motion sensitive camera that records mammals that run through all year round [3]. These traps give brand new information on a whole community of different species (shrews, voles, lemming, least weasel and ermine) during the long arctic winter, since the traps become part of tunnel systems used by mammals in the winter. The resulting image sequences can be used to analyze animal population and community dynamics. Correct species identification is however essential for this purpose. An advantage in this regard is that the images are similar with respect to background and distance to the animal.

We trained three different Convolutional Neural Network (CNN) models and used them to predict and label images (see [4] for more details). To train the models, we had to manually annotate a dataset selected from 74,429 unlabeled camera trap images. The images were pre-processed to accommodate the algorithms we used, and to remove unnecessary information that could influence training, such as the black borders displaying image metadata.

Results show that we achieve 97.84% accuracy, 97.81% precision and 93.45% recall on a dataset with 10000 camera trap images and 11 classes. This exceeds previous state of the art in animal identification on camera trap images [5] [6] [7].

Figure 2: Image from one of the COAT open-air camera traps showing a golden eagle.

Figure 3: Image from baited trap with a partially obstructed red fox surrounded by two hooded crows and one raven.

Open-air baited camera traps

These camera traps take images every five minutes, both day and night, overlooking an area with bait on the ground. Automatic annotation of these images is a harder problem than for the tunnel traps described above, as there will be more variation in weather, area around the camera, types of animals and distances to the animals in the image. Many images contain animals that are far away from the bait, making them hard to see, even by human experts. Images with animal presence will usually contain several different species, making the problem significantly harder than with single animal images. Furthermore, accumulating snow covers the bait, which partially obstructs the view of animals digging down to the bait, as seen in Figure 3.

This project created a system that use three different CNN methods (Faster Region-based CNN, Single Shot Multibox Detector and You Only Look Once v2) for training and annotation, and compared the results of each method. See [8] for more details about the system.

Results from the project show that we can detect and classify animals in the images with 94.1% accuracy at 21 frames per second, exceeding the performance of previous work. Reindeer is the most challenging species to detect, as they are frequently far away from the camera. Many of the other animal classes had an accuracy close to 100%.

Conclusions

The promising results from these projects show that we are close to automating the annotation of images from the camera traps. We have not yet done a study that compares the accuracy of the automatic annotation systems to human experts that are currently annotating the images. We expect that the automatic system is close to the accuracy level of experts, as it is easy to make subtle mistakes in the software they are currently using, and humans are prone to make mistakes in tiresome and repetitive work.

We believe that the results are promising and expect that further improvements to automatic annotation of images will provide the means for scaling up the number of camera traps. This will enable observations with more spatial resolution and of larger geographic areas without the limitations currently introduced by limited human resources.

References

[1] R. A. Ims, J. U. Jepsen, A. Stien and N. G. Yoccoz, "Science Plan for COAT: Climate-Ecological Observatory for Arctic Tundra," Fram Centre, Norway, 2013.

[2] "COAT Web Page," [Online]. Available: http://www.coat.no/. [Accessed 08 2017].

[3] E. M. Soininen, I. Jensvoll, S. T. Killengreen and R. Ims, "Under the snow: a new camera trap opens the white box of subnivean ecology.," *Remote Sensing in Ecology and Conservation,* vol. 1, no. 1, pp. 29-38, 2015.

[4] H. Thom, "Automatic rodent identification in camera trap images using deep convolutional neural networks.," UiT - The Arctic University of Norway, Capstone Project, 2016.

[5] X. Xu, J. Wang, R. Kays, P. A. Jansen, T. Wang and T. Huang, "Automated identification of animal species in camera trap images," *EURASIP Journal on Image and Video Processing,* vol. 2013, no. 1, p. 52, 2013.

[6] G. Chen, T. X. Han, Z. He, R. Kays and T. Forrester, "Deep convolutional neural network based species recognition for wild animal monitoring," in *ICIP*, 2014, pp. 856-862.

[7] A. Gomez, A. Salazar and F. Vargas, "Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks," vol. abs/1603.06169, 2016.

[8] H. Thom, "Unified Detection System for Automatic, Real-Time, Accurate Animal Detection in Camera Trap Images from the Arctic Tundra," UiT - The Arctic University of Norway, Master Project, 2017.

[9] Å. Ø. Pedersen, A. Stien, E. Soininen and R. A. Ims, "Climate-ecological observatory for arctic tundra-status 2016," *Fram Forum,* pp. 36-43, Mar 2016.