Photo recognition

for 4-poster tick management system

Coral M. Salort Nogueras[1], Timothy Green[2]

[1]Computer Science and Biology, University of Puerto Rico, Río Piedras Campus, San Juan PR 00925,
[2]Environmental Protection Division, Upton, NY 11973

ABSTRACT

The 4-Poster Tick Management System® works as a feeding station for deer to mitigate tick dispersion by forcing deer to make contact with rollers containing a tickicide when trying to eat. Time lapse photography is used to monitor deer visitation, producing about 1 million images per year which must be manually classified by researchers. The successful implementation of MLWIC: Machine Learning for Wildlife Image Classification in R package by Mikey Tabak could provide a less time-consuming and unsupervised data analysis. At the moment, the package has been successfully installed and can be used to classify and train a new model. Although the first step is to enable the package to identify presence or absence of deer, the end goal is to identify deer, turkeys, raccoons, other animals or no animals in the pictures. Monitoring deer visitation rates to demonstrate the efficiency of the bait stations is crucial to control the widespread distribution of ticks thus diminishing Lyme and other tick-borne diseases in the laboratory area.

REFERENCES

Gosselin, T., & Green, T. *The efficiency of 4-poster tick management stations on three tick species populations within Brookhaven National Laboratory.*

Tabak, M. A., Norouzzadeh, M. S., Wolfson, D. W., Sweeney, S. J., Vercauteren, K. C., Snow, N. P., . . . Miller, R. S. (2018). Machine learning to classify animal species in camera trap images: Applications in ecology. *Methods in Ecology and Evolution, 10*(4), 585-590. doi:10.1101/346809

INTRODUCTION

Controlling the tick population in Long Island is a serious matter. On Brookhaven National Laboratory, the Environmental Protection Division set up fourteen *4-Poster Tick Management System®* in 2013 and has been monitoring the efficiency of the system. Each station is monitored by a camera trap. This method produces about a million pictures a year that the scientist had to manually sort into one of these five categories: deer, turkey, raccoon, no animals or other animals. This project is meant to find a solution to ease the hand work for the researchers.

Figure 1. Example of images obtained from the trap cameras represented as a stack of photos.

A potential method is using a Computer Vision code that could sort the images automatically into one of the above mentioned categories confidently with little or no help from researchers. In 2018, Tabak, M. *et. al* presented the MLWIC R package (Machine Learning for Wildlife Image Classification). Internally, these package uses the Tensorflow framework in which they trained a ResNet-18 convolutional neural network to identify certain species with up to 97.6% accuracy. MLWIC is free and available on GitHub.

This project seeks to habilitate this code for the computer located at the research laboratory. This involves installing it, writing up a detailed easy to follow guide and testing the efficiency and viability of using it. This involves the possibility of training a model with the pictures gathered throughout the years thanks to the camera traps.

PROGRESS

After downloading the package from GitHub, follwing the installation steps provided and encountering many problems, the MLWIC R package was successfully installed. A guide was developed for Mac and Windows, including posible problems and installation tips.

In order to run the classification command, we had to provide a two column CSV with all the *unique* image names in the first column and another column filled with zeroes. For this, we developed an R script that would automatically produce the CSV and then we had to convert it to Unix Linebreaks using Notepad++. We were then ready to run the classify command and we noticed the built in model would consitently identify species that were not present on the images, neither on site.

Figure 2. Pictures taken in the same station during the day (a) and during the night (b).

We decided to prepare to train our own model using the train function provided by the R package. For this, we had to create another CSV but this time, the second column had to have meaningful values. These values would represent a label of a class (or group) that we wanted. The labels in our case would go as follows: (0) No_Animal, (1) Deer, (2) Turkey, (3) Raccoon, (4) Other_Animal. A training set should also be chosen from the huge repository stored so far. We wanted to include a sample nighttime and daytime (Figure 2), trying to get an equal amount of samples for each group and images from all the 4-Poster stations. After chosing the sample and writing the code to produce the CSV, we started the training.

Figure 3. Turkey (a), raccoon (b) and other animal (c) image examples taken in the same station. Both images in Figure 2 contain deer.

NEXT STEPS

After the training has finalized, we need to test out model using our brand new model using the classify function on a different set of images and measure the accuracy. Also, another R script can be developed to sort classified images into folders

U.S. DEPARTMENT OF ENERGY | Office of Science

BROOKHAVEN NATIONAL LABORATORY