Data Documentation for DRUM   Marco Willi - 30. September 2018

Camera Trap Images used in *"Identifying Animal Species in Camera Trap Images using Deep Learning and Citizen Science"*

The following sections describe the camera-trap image dataset and the corresponding meta-data about them as used in the study *"Identifying Animal Species in Camera Trap Images using Deep Learning and Citizen Science"*. Removed were images classified as "Human" due to privacy reasons. Images and meta-data for the dataset "Elephant Expedition" are not included and are published outside DRUM.

Camera Trap Images

The following files contain the camera-trap images used in the study. All images were downloaded from Zooniverse and have been resized to 330x330 pixels.

camera_catalogue_images.tar
snapshot_serengeti_images.tar
snapshot_wisconsin_images.tar

Unpacking the files

The files can be unpacked using the following unix-command.

tar -xf camera_catalogue_images.tar

Other programs can be used to unpack such files as well.

File Content

Each file has the following directory / file structure once unpacked:

all/
   class1/
      image1.jpeg
      image2.jpeg
      imageN.jpeg
   class2/
   classN/

The images are stored in class specific folders (e.g. all zebra images are stored in all/zebra/)**.** The following figure shows the class folders of camera_catalogue_images.tar:

```
aardvark      buffalo        duikersteenbok  hippopotamus   klipspringer  nyala       reptile     waterbuck
aardwolf      bushbaby       eland           honeyBadger    kudu          otter       rhino       wildcat
africancivet  bushbuck       elephant        hyaenabrown    leopard       pangolin    roansable   wilddog
bat           bushpig        gemsbock        hyaenaspotted  lion          polecat     rodent      wildebeest
batEaredFox   caracal        genet           impala         MACAQUE       porcupine   serval      zebra
bird          cheetah        giraffe         insect         mongoose      rabbithare  vehicle
blank         domesticanimal hartebeest      jackal         monkeybaboon  reedbuck    warthog
```

The following figure shows some of the 'aardvark' images of camera_catalogue_images.tar

```
4430853_0.jpeg  4514689_0.jpeg  5261928_0.jpeg  5373605_0.jpeg  5439470_0.jpeg  9337613_0.jpeg  9713526_0.jpeg
4431159_0.jpeg  4515114_0.jpeg  5261932_0.jpeg  5373607_0.jpeg  5439471_0.jpeg  9337614_0.jpeg  9713527_0.jpeg
4431727_0.jpeg  4515115_0.jpeg  5261939_0.jpeg  5383917_0.jpeg  5440100_0.jpeg  9337854_0.jpeg  9713977_0.jpeg
4431839_0.jpeg  4515117_0.jpeg  5262874_0.jpeg  5388791_0.jpeg  5440101_0.jpeg  9337855_0.jpeg  9713978_0.jpeg
4436036_0.jpeg  4515118_0.jpeg  5262887_0.jpeg  5406256_0.jpeg  5442042_0.jpeg  9337857_0.jpeg  9713979_0.jpeg
4436353_0.jpeg  4515147_0.jpeg  5264493_0.jpeg  5423220_0.jpeg  5443826_0.jpeg  9338223_0.jpeg  9713980_0.jpeg
4438832_0.jpeg  4515870_0.jpeg  5265455_0.jpeg  5423473_0.jpeg  5443827_0.jpeg  9338224_0.jpeg  9713985_0.jpeg
4449044_0.jpeg  4515874_0.jpeg  5266738_0.jpeg  5423474_0.jpeg  5444184_0.jpeg  9338441_0.jpeg  9714048_0.jpeg
4449280_0.jpeg  4516186_0.jpeg  5267724_0.jpeg  5423512_0.jpeg  5445128_0.jpeg  9339263_0.jpeg  9714049_0.jpeg
```

Meta-Data

Each camera-trap dataset has a meta-data file which contains for each image all the necessary information. The files are contained in this file:

meta_data.tar.gz

Unpacking the files

The files can be unpacked using the following unix-command.

tar xvzf meta_data.tar.gz

Other programs can be used to unpack such files as well.

File Content

Each dataset contains one single file:

camera_catalogue/subject_set_used.json

Each json-file (a common human and machine-readable file) contains one entry per capture event which consists of 1 or more images. Each entry is identified by its capture-id and has the following fields:

- *fnames*: a list with the image file names
- *img_ids*: a list with the image names
- *file_paths*: a relative path to where the images are save*d*
- *urls:* a list of urls of the source (original) images downloaded via Zooniverse
- *meta_data*: key/value pairs with the following keys (not always available):
  - location: the location identifier of where the images were taken
  - pielou: pielous evenness index of the citizen science annotations (a value of 0 means complete agreement among all volunteers, a value of 1 means complete disagreement).

- time: the time of the day HHMMSS of when the photos were taken (000000 if unknown)
- n_species: the number of species / classes in the images
- n_users: the number of volunteer annotations used to determine the label / class
- date: the date of when the images were taken (YYYYMMDD)
- datetime: the date and time of when the images were taken (YYYMMDDHHMMSS)
- label: a list of class labels for the images

Example:

"9909433": {"img_ids": ["9909433_0"], "meta_data": {"pielou": 0, "time": "000000", "n_species": 1.0, "location": "C1055864", "datetime": "20010101000000", "date": "20010101", "n_users": 3}, "urls": ["https://s3-eu-west-1.amazonaws.com/pantherabucketleopard1/25_2017/C1055864.JPG"], "label": ["elephant"], "file_paths": ["all/elephant/"], "fnames": ["9909433_0.jpeg"]},

The "9909433" refers to the capture-id. The capture contains one image with name "9909433_0.jpeg" that is stored in "all/elephant/"and is of label "elephant".

File Size

Note that the .json files can be quite large. The largest file is 1.1 GB which may be too large to open in a text editor or to load into memory using a program like Python on computers with low memory. The files are intended to be processed with a program, for example, to train a model. Memory is normally not an issue in such cases because a powerful computer is required for training models anyway (which presumably has also enough memory). The exact memory requirement to open the larger .json files is unclear, however, we recommend at least 8 GB RAM, preferably 16 GB or more.

Dataset Descriptions

Snapshot Serengeti

This dataset was the largest used in the study *"Identifying Animal Species in Camera Trap Images using Deep Learning and Citizen Science"*, containing over seven million images of which 1.7 million were used for model training. The cameras were located in Tanzania, at various places in the Serengeti National Park. The cameras were triggered by a combination of infrared and motion sensors and took three images, after which the trigger was blocked for one minute. Volunteers on Zooniverse (citizen scientists) could choose from 54 different species, a number which has been increased slightly over time. In the study 51 of the species were modeled, of which 35 were also covered by the Camera CATalogue dataset. This project featured a large proportion of empty images (81.8%).
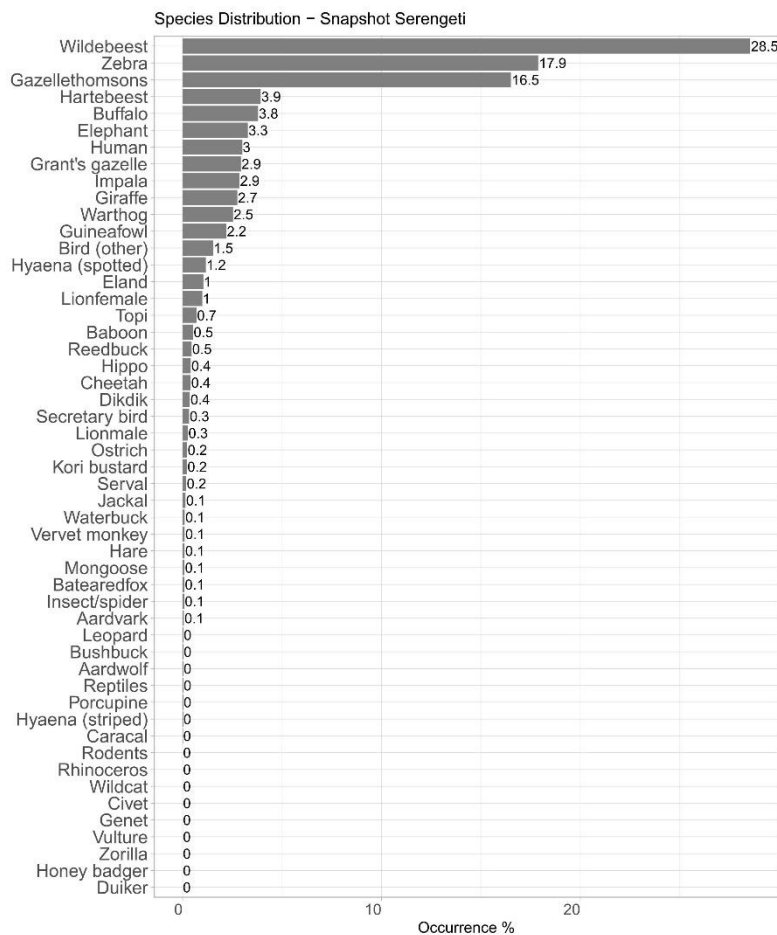

Figure 1: Species distribution of Snapshot Serengeti as modeled in the study "Identifying Animal Species in Camera Trap Images using Deep Learning and Citizen Science" (some very rare species not shown).

Camera CATalogue

This project featured species limited to South Africa, as part of a larger monitoring network implemented by Panthera (www.panthera.org). The camera traps were positioned along roads, animal paths, and drainage lines to maximise capture probabilities for target species and took one image per capture event. The placement along roads caused a large fraction of vehicle images (47.8%). However, the chosen locations ensured captured animals were often quite close to the camera, making it easier to identify species. Volunteers could choose from 54 different categories to annotate images. In the study 48 species (excluding rare species) were modeled, of which 35 were also covered by the Snapshot Serengeti dataset. This dataset contained roughly 0.5 million images.
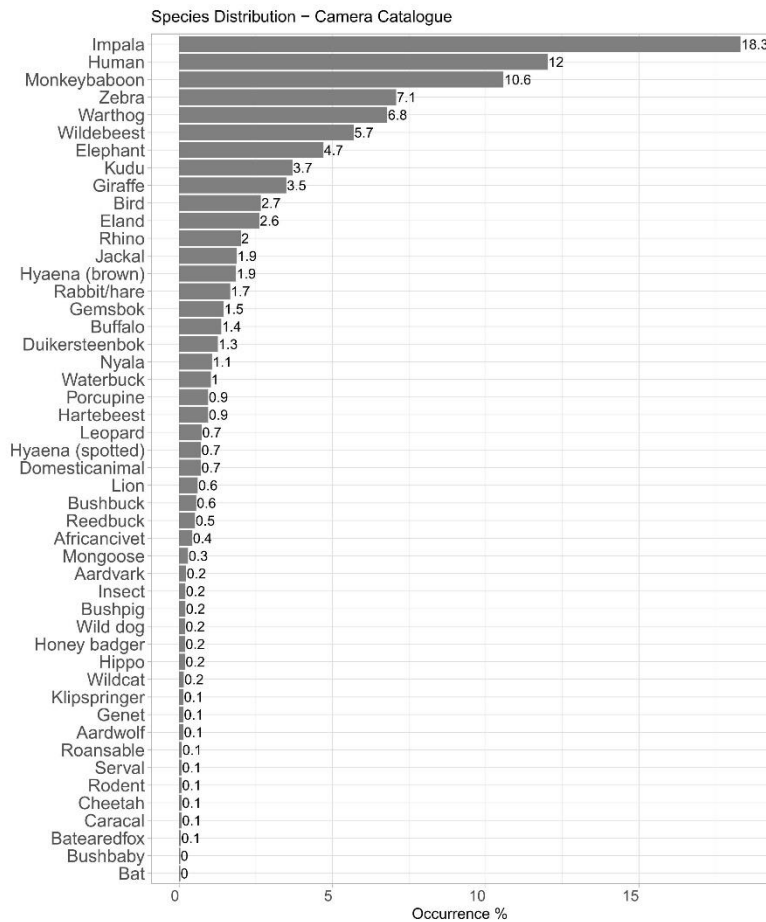

Figure 2: Species distribution of Camera CATalogue as modeled in the study "Identifying Animal Species in Camera Trap Images using Deep Learning and Citizen Science" (some very rare species not shown).

Snapshot Wisconsin

This was the only USA-based project used in the cited study and thus had no species overlap with the other projects. The cameras were placed in forests along trails throughout Wisconsin to monitor wildlife populations. The cameras were triggered by heat and movement, and took three images in a short interval when triggered after which there was a 15 second lag time before the cameras would trigger again if there was still motion. In total there were 43 different species that the volunteers could choose from, of which 31 were modeled (excluding rare species). This project had the largest imbalance in species occurrence with deer accounting for over 65% of all non-empty capture events. This dataset contained roughly 0.5 million images.
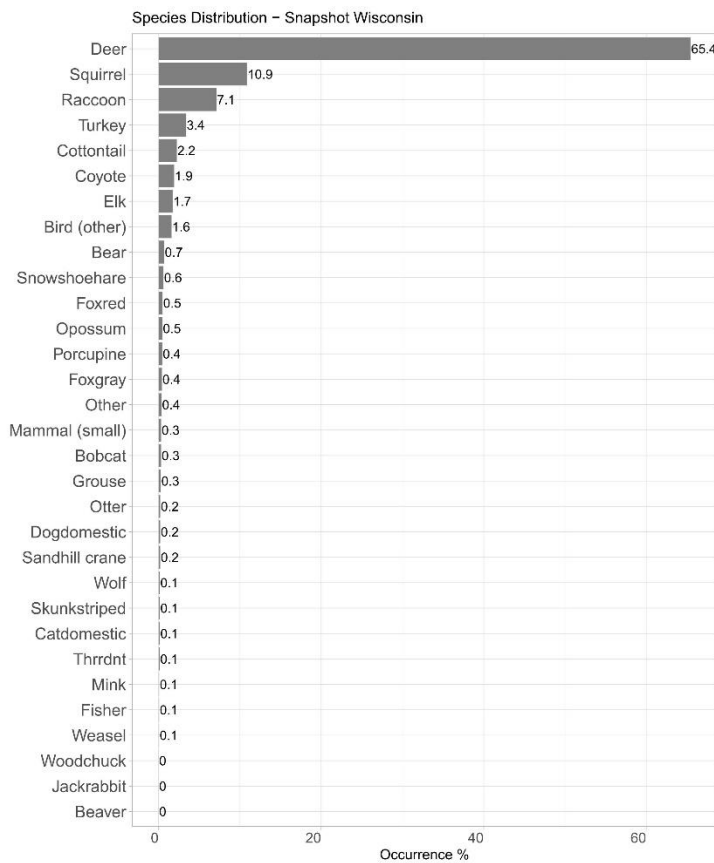

Figure 3: Species distribution of Snapshot Wisconsin as modeled in the study "Identifying Animal Species in Camera Trap Images using Deep Learning and Citizen Science" (some very rare species not shown).