

# ReadMe - V\_rodents\_cameratraps\_image\_classification\_lemming\_blocks

Dataset responsible: Eeva Soininen  
([eeva.soininen@uit.no](mailto:eeva.soininen@uit.no))

02.12.2021

## 1 Protocol

Camera trapping of small mammals has been conducted using the COAT protocol 'protocol\_camera\_trapping\_small\_mammals\_varanger'.

### 1.1 Spatial layout of sampling and changes in that.

Camera trapping of small rodents in Komagdalen was started in summer 2015. A pilot study was conducted from summer 2014 to 2015 and the data is available in dataset 'V\_rodents\_cameratraps\_pilot'.

Some camera traps have been moved in 2016 and 2018 to avoid flooding or filling with snow. In 2018, cameras have been set up in a second locality (Vestre Jakobselv).

More detailed information about which sites were included in the study design can be found in the auxiliary file 'V\_rodents\_cameratraps\_image\_classification\_lemming\_blocks\_aux.txt'

### 1.2 Timing of sampling

The camera traps are deployed the whole year and the images are downloaded once a year, usually in July. Therefore, the data files contain data from July to July. For example, the file for 2020 contains data from July 2019 to July 2020

## 2 Description of the dataset

The dataset includes three different types of files and all files are saved as ;-separated txt-files:

- One file per year and locality with data on presence or absence of small mammals and image quality (`_komagdalen_2016.txt`)
- One coordinate file with coordinates of all sites (`_coordinates.txt`)
- One auxiliary file with additional information about the sites (`_aux.txt`)

In addition to this readme-file, there is one file that describes the model used for automatic image classifications and one file per year that summarizes the quality check after automatic classification.

### 2.1 V\_rodents\_cameratraps\_image\_classification\_lemming\_blocks.txt

These files contain presence (1) or absence (0) of the following small mammal species:

```
[1] "aves = Bird"
[2] "cricetidae = Vole (either Grey-sided vole or tundra vole, Red-backed vole very rare)"
[3] "lem_lem = Lemmus lemmus (Norwegian lemming)"
[4] "mus_erm = Mustela erminea (Stoat)"
[5] "mus_niv = Mustela nivalis (Least weasel)"
[6] "sor_sp = Sorex sp (Shrew)"
```

In addition, there is a category for empty images (empty = 1) and for images with bad quality. If it is not possible to recognize whether there is an animal on the image or not, the image is classified as bad quality (bad quality = 1). All other categories are set to NA if an image is classified as bad quality.

The images are classified automatically by a machine learning model (see 'small\_mammal\_classification\_model\_v2021\_summary' for more information about the model).

After automatic classification, a subset of images was annotated manually for a quality check (see 'V\_rodents\_cameratraps\_image\_classification\_lemming\_blocks\_quality\_check' for more information and the results of the quality check). The quality check showed that the model performance for the categories 'least\_weasel' and 'stoat' and for classifications with low confidence is not satisfying. Therefore, all images classified as least weasel or stoat as well as all images classified with a confidence below 0.8 were annotated manually.

The manual annotations are included in the data files and can be used instead of the automatic classifications. We recommend to use the manual classifications for all images classified as least weasel or stoat and all images classified with a confidence below 0.8.

Example of the first rows of the classification file:

```
sn_region sn_locality sn_section sc_type_of_sites_ecological sn_site
1 varanger komagdalen gargas hummock_mire ko_ga_hu_1
2 varanger komagdalen gargas hummock_mire ko_ga_hu_1
3 varanger komagdalen gargas hummock_mire ko_ga_hu_1
t_date t_time v_image_name v_class_id
1 2015-08-24 08:52:04 ko_ga_hu_1_2015-08-24_0001.JPG aves
2 2015-08-24 08:52:04 ko_ga_hu_1_2015-08-24_0001.JPG bad_quality
3 2015-08-24 08:52:04 ko_ga_hu_1_2015-08-24_0001.JPG cricetidae
v_presence_automatic v_presence_manual v_confidence_automatic
1 NA 0 NA
2 1 1 0.9999613
3 NA 0 NA
v_observer_manual v_type_manual_classification v_comment
1 efk manual_classification_2016 <NA>
2 efk manual_classification_2016 <NA>
3 efk manual_classification_2016 <NA>
```

## Description of the columns included in the classification file:

Column name	Description	Possible values
sn_region	Study region	varanger
sn_locality	Locality (within region)	komagdalen, vestre_jakobselv
sn_section	Section (within locality)	gargas, hubejohka, kjoltindan, ryggfjellet, gaasevannan, reinhaugen, skoarrajohka, tranemyra

sc_type_of_sites_ecological	Habitat type	hummock mire, snowbed
sn_site	Unique Site ID	e.g. ko_ga_hu_1, ko_ga_sn_1, ko_hu_hu_3, ko_hu_sn_3, ko_kj_sn_23
t_date	Date when the image was taken	YYYY-MM-DD
t_time	Time when the image was taken	HH:MM:SS
v_image_name	Image name	e.g. ko_ga_hu_1_2015-08- 24_0001.JPG, ko_hu_sn_1_2015-09- 05_0004.JPG
v_class_id	Class ID (species/functional group or image quality)	aves, bad_quality, cricetidae, empty, lem_lem, mus_erm, mus_niv, sor_sp
v_presence_automatic	Presence of species/functional group and image quality (results from automatic classification of the images by a machine learning model)	NA, 1, 0
v_presence_manual	Presence of species/functional group and image quality (results from manual classification for a quality check and of images with unsatisfying results from automatic classification)	0, 1, NA
v_confidence_automatic	Confidence of the machine learning model that the images belongs to the category	[numeric] values between 0 and 1
v_observer_manual	Initials of the person who manually classified the images	e.g. es (Eeva Soininen)
v_type_manual_classification	Reason for manual classification	komagdalen, vestre_jakobselv
v_comment	Comments	[character]

---

## 2.2 V\_rodents\_cameratraps\_image\_classification\_lemming\_blocks\_coordinates.txt

This file contains the coordinates of all camera sites included in the study design. Coordinates are given in decimal degrees and UTM 33 (WGS 84).

Example of the first rows of coordinate files:

```
      sn_site      e_dd      n_dd e_utm33 n_utm33
1 ko_ga_hu_1 29.90737 70.34465 1054721 7872741
2 ko_ga_hu_1b 29.90745 70.34474 1054721 7872751
3 ko_ga_sn_1 29.92304 70.36851 1054643 7875475
```

## 2.3 V\_rodents\_cameratraps\_image\_classification\_lemming\_blocks\_aux.txt

This file contains further information about the dataset such as old site names and the years when sites were first included in the study design and when sites were excluded from the study design.

Example of the first rows of auxiliary files:

```
      sn_region sn_locality sn_section sc_type_of_sites_ecological      sn_site
1 varanger komagdalen gargas hummock_mire ko_ga_hu_1
2 varanger komagdalen gargas hummock_mire ko_ga_hu_1b
3 varanger komagdalen gargas snowbed ko_ga_sn_1
      sn_site_old date_first date_last
1 GNY1 2015-08-24 2016-07-04
2 GNY1-2016 2016-07-04 <NA>
3 GNY10 2015-08-26 2016-07-04
      v_comment
1 <NA>
2 <NA>
3 no rockfield but hummocks that make alot of microtopography in the snowbed.
```

\* year\_last is NA if the site is still included in the study design

## 3 Related datasets

### 3.1 V\_rodents\_cameratraps\_image\_metadata\_lemming\_blocks

This dataset contains image metadata for each image such as trigger mode (motion sensor or timelapse) and temperature.

### 3.2 V\_rodents\_cameratraps\_annual\_metadata\_visits

This datasets contains information of the annual visits of the camera sites when the cameras are checked and images are downloaded such as the serial number of the camera and whether the camera was functioning.

### 3.3 V\_rodents\_cameratraps\_manual\_image\_classification\_lemming\_blocks

This dataset contains presence (1) or absence (0) of small mammal species from manual classification for all images from Komagdalen between summer 2015 and summer 2018. In this period, all images were annotated

manually. The manual classifications in the dataset 'V\_rodents\_cameratrap\_manual\_image\_classification\_lemming\_blocks' differ slightly from the manual classifications included in the dataset 'V\_rodents\_cameratrap\_image\_classification\_lemming\_blocks' because different persons annotated the images.

### **3.4 V\_rodents\_cameratrap\_pilot**

This dataset contains data from a pilot study performed in Komagdalen in 2014 and 2015.

## **4 Data processing**

1. Metadata is extracted and the images are renamed with a name consisting of the site-ID, the date when the image was taken and a number (e.g. ko\_ga\_sn\_1\_2020-01-01\_0001.JPG)
2. The images are classified automatically using a machine learning model
3. A random subset of images is annotated manually for a quality check
4. Images belonging to categories with unsatisfying model performance are annotated manually
5. Information from metadata files, automatic and manual classifications are combined and formatted according to the requirements of the COAT data portal

All data processing is performed in R.