

# Quality check of automatic classification of images from small mammal cameras 2019

Hanna Boehner

25.11.2021

The images from July 2018 to July 2019 from small mammal camera traps included in the data set ‘V\_rodents\_camperatraps\_image\_classification\_lemming\_blocks’ were classified automatically using a machine learning model. See the protocol ‘protocol\_camera\_trapping\_small\_mammals\_varanger’ and the document ‘small\_mammal\_classification\_model\_v2021\_summary’ for more information about the study design and the machine learning model. The model returns the confidence that the image belongs to the following classes: empty, bad quality, vole, lemming, stoat, least weasel, shrew and bird. The class with the highest confidence is then selected as the image label.

## Summary of the automatic classification

In total, 189027 images were taken in Komagdalen and 171180 images were taken in Vestre Jakobselv. The number of images per class based on automatic classification as well as the number of images selected for the quality check are shown in table 1. Figure 1 shows the distribution of registered species/classes over a year (from August 2018 until June 2019).

Table 1: Number of images in Komagdalen and Vestre Jakobselv based on automatic classification and number of images selected for the quality check.

Class	Komagdalen	Vestre Jakobselv	Quality check (Komagdalen)	Quality check (Vestre Jakobselv)
Bad quality	17518	22178	136	161
Bird	90967	82012	47	66
Empty	1282	359	413	389
Least weasel	50821	41261	55	66
Lemming	357	133	157	177
Shrew	23423	21106	116	87
Stoat	4071	3486	96	90
Vole	588	645	280	259
TOTAL	189027	171180	1300	1295

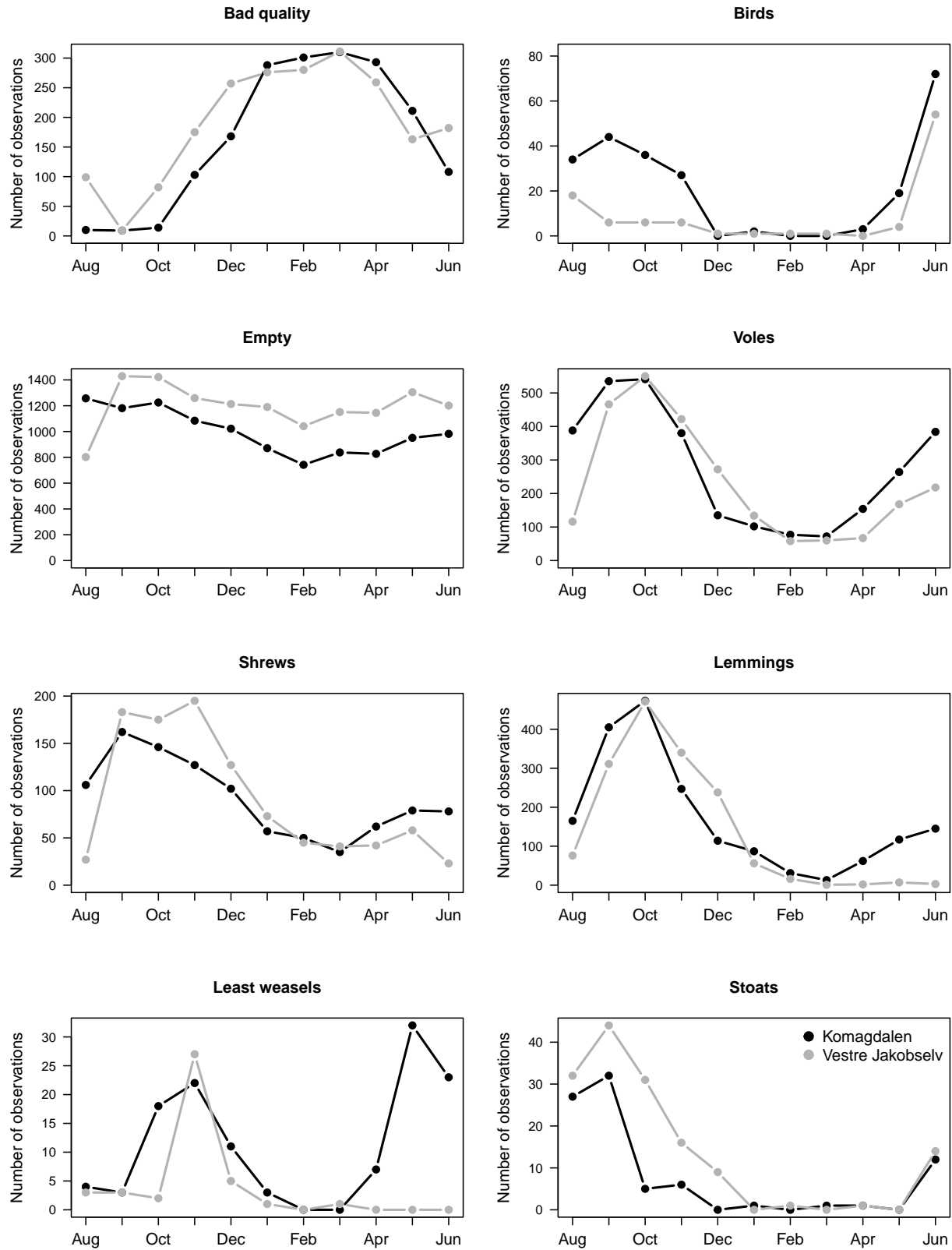


Figure 1: Distribution of species/classes registered by small mammal cameras over a year. The numbers are based on automatic classification and calculated as the number of days and sites with images labeled as a certain class per month.

## Quality check - part 1

Figure 2 shows a histogram with the number of images per confidence class and the cumulative density curve for Komagdalen and Vestre Jakobselv.

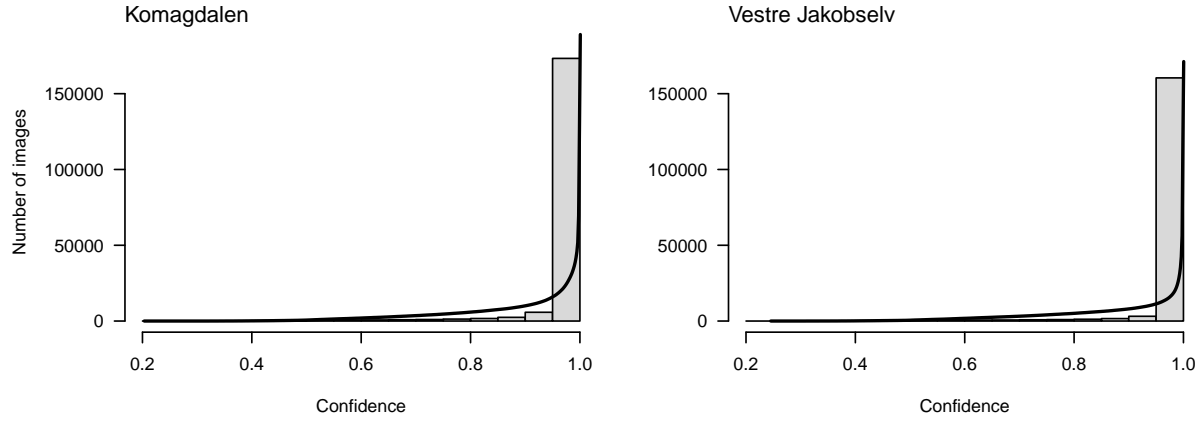


Figure 2: Histogram and cumulative density (black line) of the classification confidences of images from Komagdalen and Vestre Jakobselv.

For a quality check, 500 randomly selected images of each of the localities were labeled manually to calculate prediction accuracy of the model. Accuracy was calculated as the number of correct predictions divided by the number of all predictions. Accuracy was 0.98 for Komagdalen and 0.97 for Vestre Jakobselv.

In addition, 100 randomly selected images per confidence class (0-0.1, 0.1-0.2, ..., 0.9-1.0) and locality were also labeled manually and prediction accuracy was calculated for each confidence class (Figure 3).

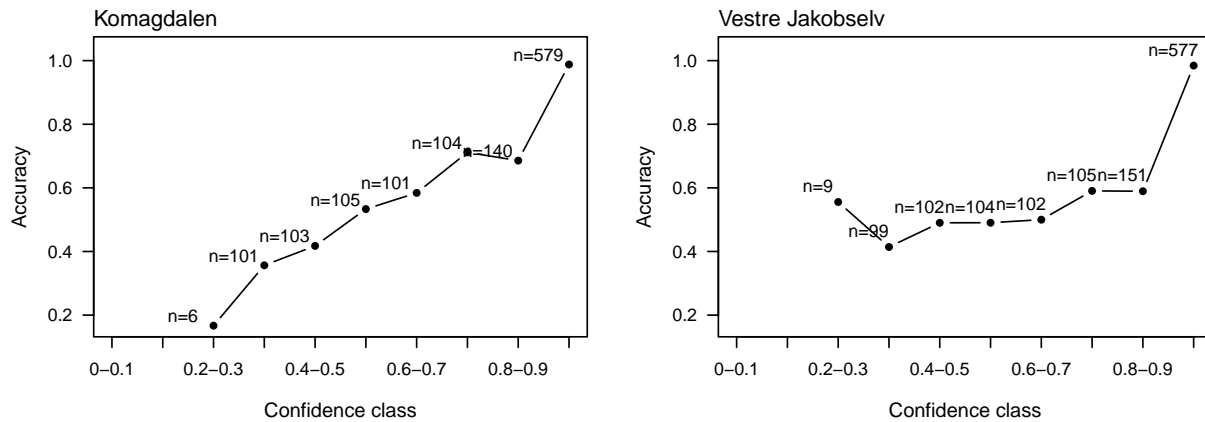


Figure 3: Prediction accuracy of images that were classified with a confidence between 0 and 0.1, between 0.1 and 0.2, ..., and between 0.9 and 1.0.

## Quality check - part 2

The image data set is usually unbalanced with a lot of empty, vole and lemming images but only a few images of stoats, least weasels and birds and thus, only a few images of these classes will be labeled manually when selecting images randomly. In order to increase the sample size of rare species or classes, 100 randomly selected images per locality and class (based on model classification) were annotated manually. Precision, recall and F1 score were calculated for each class including the 500 randomly selected images and the 100 images per class (Table 2). Since including 100 images of each class in the quality check data set increased proportion of rare species, the number of true positives, false positives and false negatives was corrected for the proportion of images of each class in the complete data set

$$Precision = \frac{TP}{TP + FP} \quad (TP = \text{True positives})$$

$$Recall = \frac{TP}{TP + FN} \quad (FP = \text{False positives})$$

$$F1 = 2 * \frac{precision * recall}{precision + recall} \quad (FN = \text{False negatives})$$

Table 2: Precision, recall and F1 score for images from Komagdalen (KO) and Vestre Jakobselv (VJ)

Class	Precision (KO)	Recall (KO)	F1 (KO)	Precision (VJ)	Recall (VJ)	F1 (VJ)
empty	0.99	0.99	0.99	0.98	0.98	0.98
shrew	0.97	0.93	0.95	0.77	0.93	0.85
vole	0.98	0.98	0.98	0.99	0.98	0.99
least_weasel	0.55	1.00	0.71	0.69	1.00	0.82
lemming	0.99	0.97	0.98	0.98	0.95	0.96
bad_quality	0.93	0.96	0.94	0.97	0.98	0.98
bird	0.46	1.00	0.63	0.65	1.00	0.79
stoat	0.95	1.00	0.97	0.89	1.00	0.94

Figure 4 shows confusion matrices for Komagdalen and Vestre Jakobselv including the 500 randomly selected images and the 100 images per class.

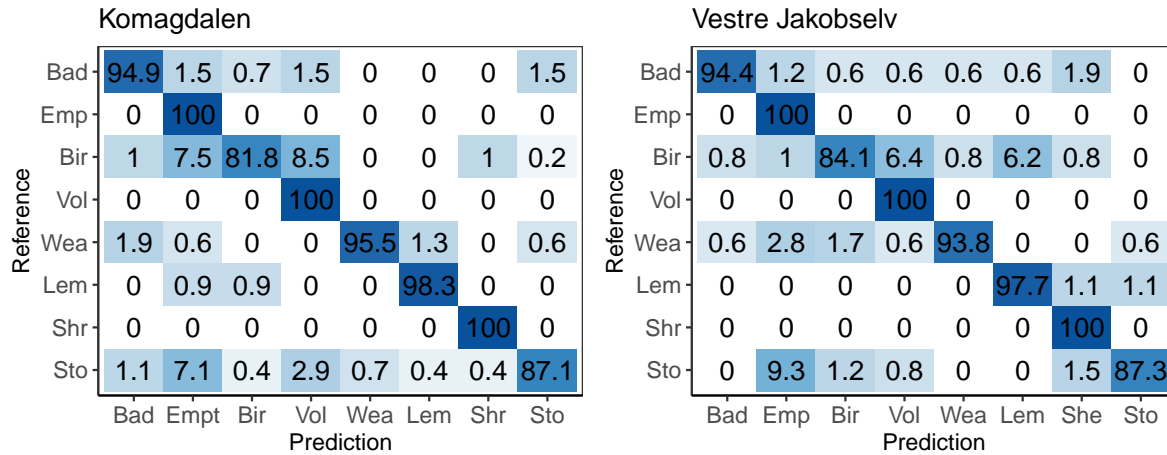


Figure 4: Confusion matrix (percentage of correct labels for each class) for Komagdalen and Vestre Jakobselv. (Bad = Bad quality, Emp = Empty, Bir = Bird, Vol = Vole, Wea = Least weasel, Lem = Lemming, Shr = Shrew, Sto = Stoat)