

Quality check of automatic classification of images from small mammal cameras 2020

Hanna Boehner

25.11.2021

The images from July 2019 to July 2020 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, 207660 images were taken in Komagdalen and 154337 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 2019 until June 2020).

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	12181	26675	148	203
Bird	142768	87828	82	91
Empty	2395	1455	475	438
Least weasel	28794	34520	76	91
Lemming	755	551	156	98
Shrew	17145	2392	99	73
Stoat	2838	431	93	69
Vole	784	485	170	230
TOTAL	207660	154337	1299	1293

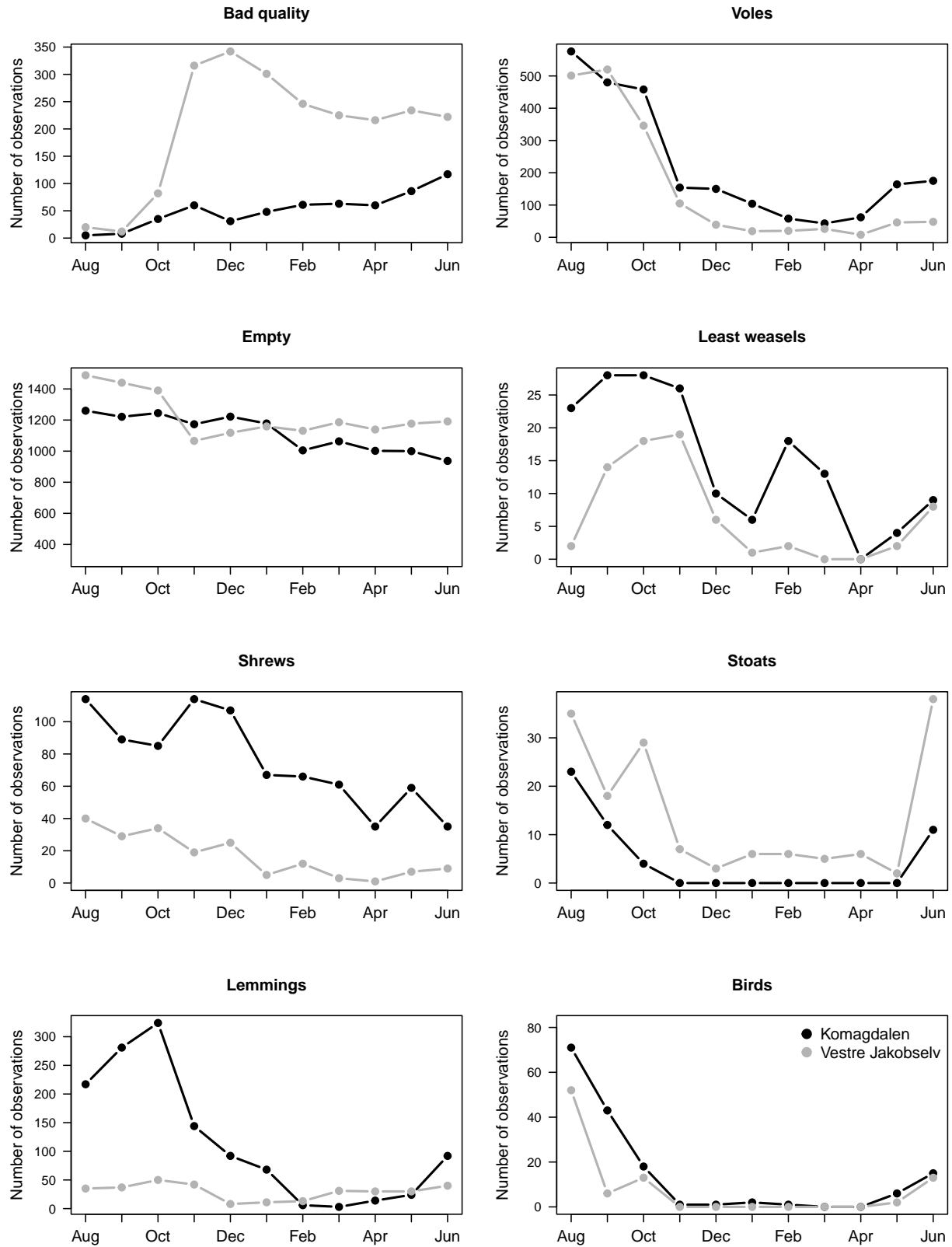


Figure 1: Distribution of species/classes registered by small mammal cameras over a year. The numbers are based on automatic classification and calculated as the number of days and sites with images labeled as a certain class per month.

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.

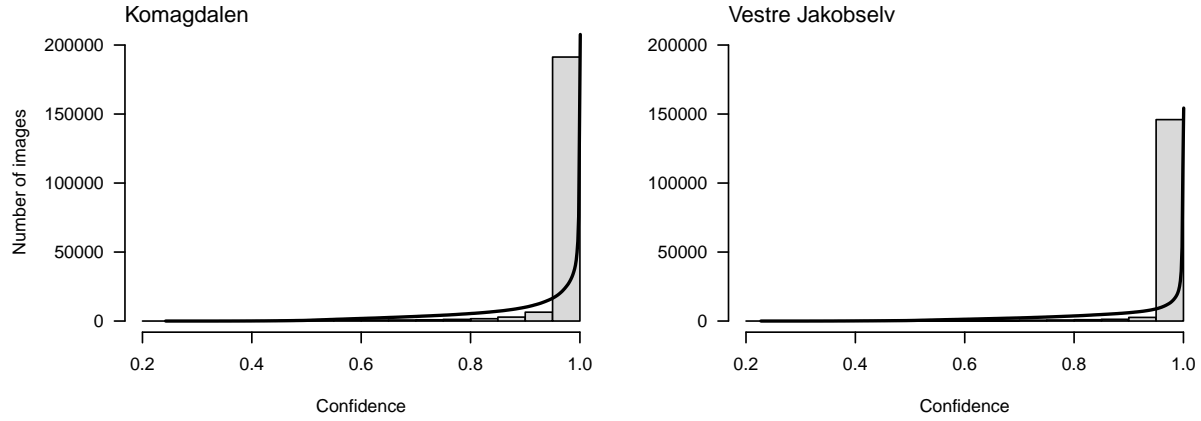


Figure 2: Histogram and cumulative density (black line) of the classification confidences of images from Komagdalen and Vestre Jakobselv.

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 0.99 for Komagdalen and 0.99 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).

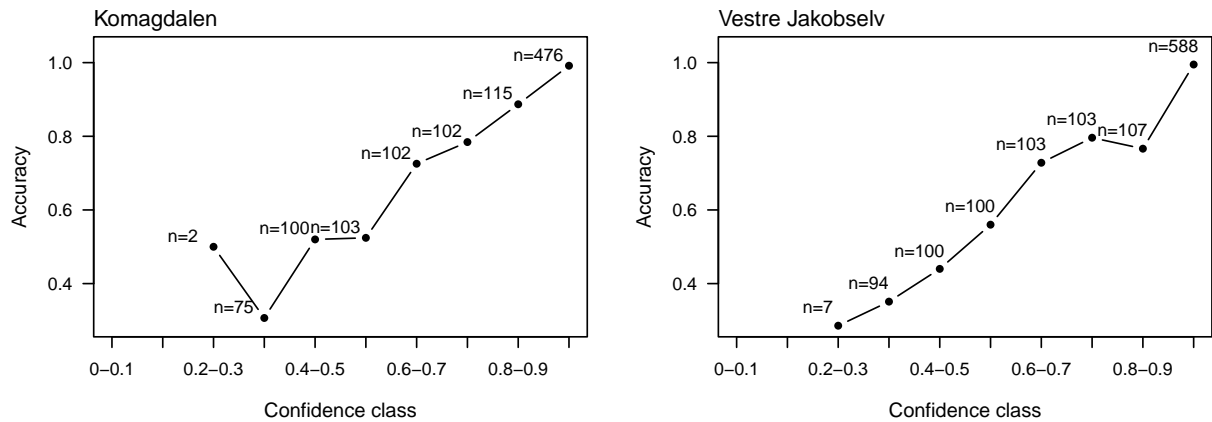


Figure 3: Prediction accuracy of images that were classified with a confidence between 0 and 0.1, between 0.1 and 0.2, ..., and between 0.9 and 1.0.

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)
least_weasel	0.74	0.99	0.85	0.90	1.00	0.95
vole	0.98	0.99	0.98	1.00	0.99	0.99
empty	0.99	0.99	0.99	1.00	0.99	0.99
shrew	0.93	0.84	0.88	0.74	1.00	0.85
lemming	0.98	0.93	0.96	0.87	1.00	0.93
stoat	0.90	0.99	0.94	0.69	1.00	0.82
bad_quality	0.98	0.94	0.96	0.98	1.00	0.99
bird	0.80	1.00	0.89	0.92	1.00	0.96

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

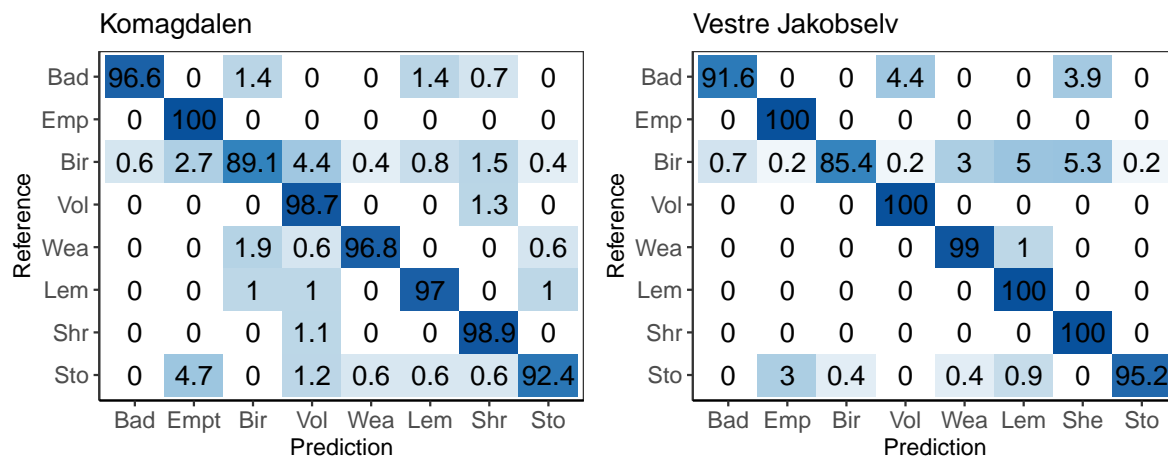


Figure 4: Confusion matrix (percentage of correct labels for each class) for Komagdalen and Vestre Jakobselv. (Bad = Bad quality, Emp = Empty, Bir = Bird, Vol = Vole, Wea = Least weasel, Lem = Lemming, Shr = Shrew, Sto = Stoat)