# Small mammal classification model version 2021

A machine learning model for automatic classification of images from small mammal camera traps

Hanna Boehner

25.11.2021

## Model training

A machine learning model was trained with 53636 images from Komagdalen, Vestre Jakobselv, Porsanger, Håkøya, Kirkesdalen and Valdres. Images from 6 animal classes (Voles, Lemmings, Shrews, Least weasels, Stoat and Birds) as well as empty and bad quality images have been included in the model.

An image was classified as 'bad quality' if it was not possible to decide whether the image was empty or not. For example, if the camera was not working properly and the image was completly blurry or black, if the box was full of snow or vegetation or if the image was taken while the camera was set up or checked (image of humans or landscape).

Images were annotated with a species if it was possible to identify the species without help from previous or following images. This means images where only a part of the animal is visible (e.g. the tail) were also included in the training dataset.

Table 1: Number of images per class used for training the model. Abbreviations used in the datasets are given in parenthesis.

| Class ID | Number of training images | Number of internal validation images |
|---|---|---|
| Bad quality | 6453 | 350 |
| Bird (Aves) | 3382 | 144 |
| Empty | 9444 | 350 |
| Least weasel (mus_niv) | 1725 | 74 |
| Lemming (lem_lem) | 9449 | 350 |
| Shrew (sor_sp) | 9265 | 350 |
| Stoat (mus_erm) | 4024 | 318 |
| Vole (cricetidae) | 9894 | 350 |
| TOTAL | 53636 | 2286 |

The model was trained using the R package 'keras' (Chollet et al. 2017) with a tensorflow backend. The ResNet50 architecture (He et al. 2016) was used to train a model from scratch for 55 epochs with a one-cycle learning rate policy with a minimum learning rate of 0.000001 and maximum learning rate of 0.001 (Smith 2018). Previous to training, the images were resized to 224x224 pixels and image augmentation (shifts, horizontal flips, rotations, zooms and shears) was applied to expand the training dataset.

A subset of 2286 was selected randomly from all images available for training and was used for internal model validation during training. After each epoch, the model was evaluated on the internal validation dataset by calculating accuracy and loss of the predictions after each epoch. Accuracy was calculated as the number of correct predictions divided by the number of all predictions. The loss describes the error of the model during training and was calculated using a cross-entropy loss function. Training accuracy of the model after 55 epochs was 0.99 and validation accuracy was 0.97 (Figure 1).

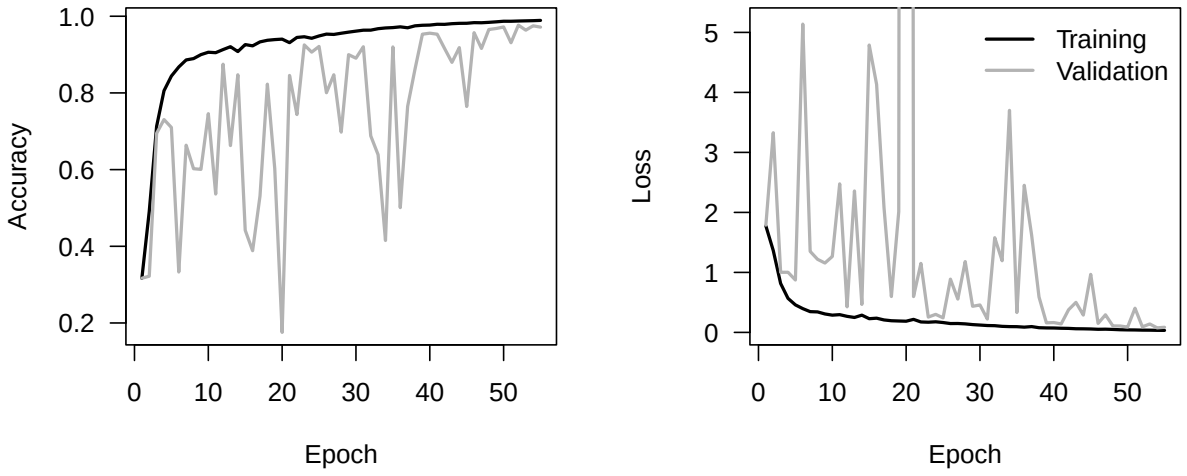


Figure 1: Accuracy and loss of the training and the internal validation dataset.

## External model validation

A separate test dataset consisting of 4425 images from Komagdalen, Vestre Jakobselv and Porsanger was used for external model validation. The test images were selected and annotated manually.

The model correctly classified 98.3 % of the images. 96.2 % of the images were classified with a confidence higher than 0.95 and 99.6 % of these images were classified correctly.

Precision, Recall and F1 score were calculated to evaluate model performance for each class (Table 1). Figure 2 shows a confusion matrix for the test dataset.

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

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

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

Table 2: Number of images per class used for external validation of the model and precision, recall and F1 score for each of the 8 classes included in the test dataset.

| Class ID | Number of images | Precision | Recall | F1 score |
|---|---|---|---|---|
| Bad quality | 748 | 0.98 | 1.00 | 0.99 |
| Empty | 1081 | 0.97 | 0.99 | 0.98 |
| Bird | 75 | 0.87 | 1.00 | 0.93 |
| Vole | 1038 | 0.99 | 0.97 | 0.98 |
| Least weasel | 24 | 0.92 | 1.00 | 0.96 |
| Lemming | 629 | 1.00 | 0.99 | 0.99 |
| Shrew | 639 | 0.99 | 0.97 | 0.98 |
| Stoat | 191 | 0.99 | 0.97 | 0.98 |

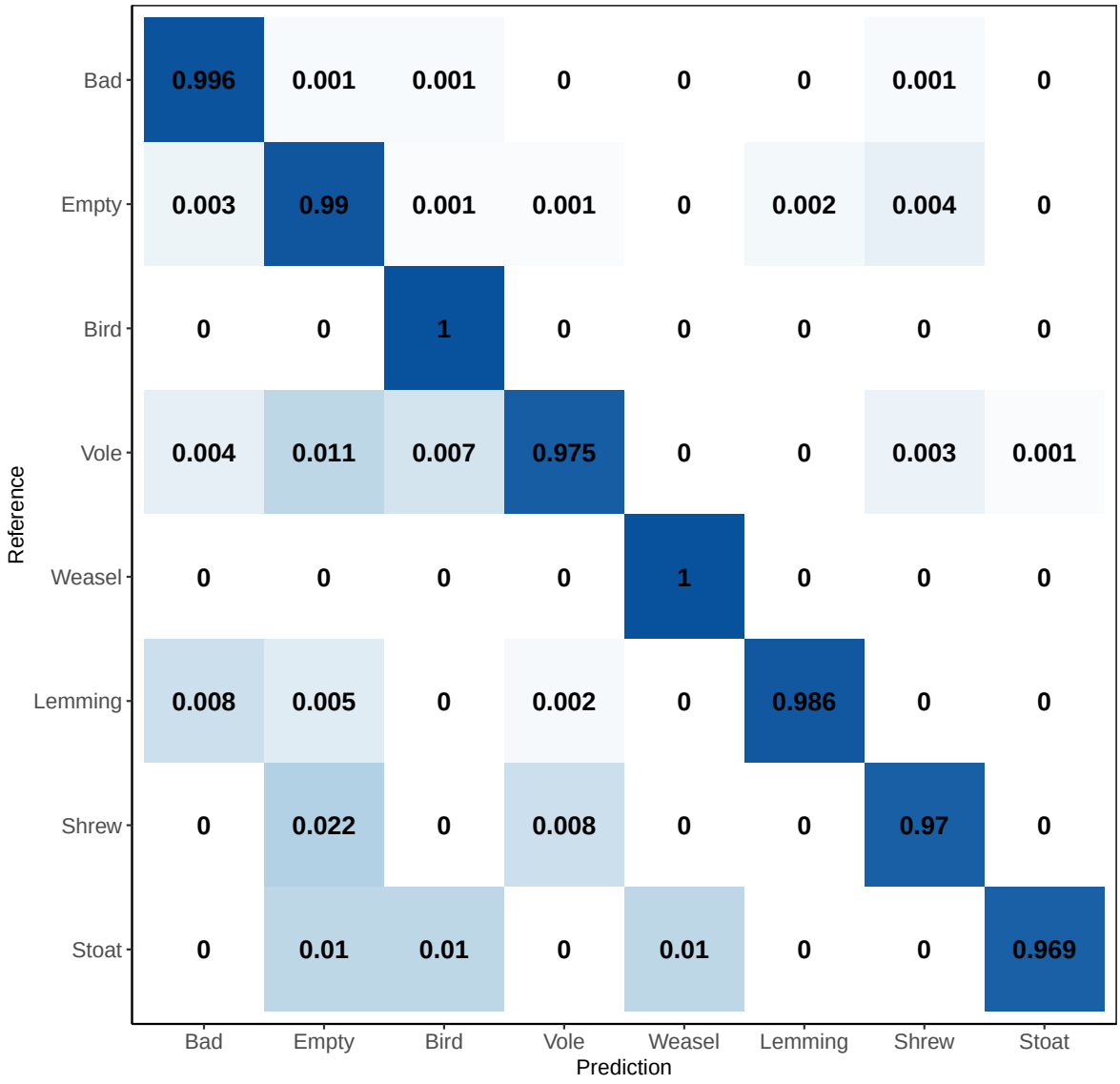


Figure 2: Confusion matrix for the test dataset.

# References


Chollet, François et al. (2017). *R Interface to Keras*. https://github.com/rstudio/keras.

He, Kaiming et al. (2016). "Identity Mappings in Deep Residual Networks". In: *Computer Vision – ECCV 2016*. Ed. by Bastian Leibe et al. Cham: Springer International Publishing, pp. 630–645. ISBN: 978-3-319-46493-0.

Smith, Leslie N. (2018). *A disciplined approach to neural network hyper-parameters: Part 1 – learning rate, batch size, momentum, and weight decay*. arXiv: 1803.09820 [cs.LG].