Article

Can transfer learning overcome the challenge of identifying lemming species in images taken in the near infrared spectrum?

Davood Kalhor [1,*], Mathilde Poirier [2,3], Xavier Maldague [1] and Gilles Gauthier [2,3]

1   Department of Electrical and Computer Engineering, Université Laval, Quebec City (Quebec), G1V 0A6, Canada; xavier.maldague@gel.ulaval.ca (X.M.)

2   Department of Biology, Université Laval, Quebec City (Quebec), Canada; mathildepoirier05@gmail.com (M.P.); gilles.gauthier@bio.ulaval.ca (G.G.)

3   Centre for Northern Studies, Université Laval, Quebec City (Quebec), G1V 0A6, Canada

*   Correspondence: davood.kalhor.1@ulaval.ca (D.K.)

Abstract: Using a camera system developed earlier for monitoring the behavior of lemmings under the snow, we are now able to record a large number of short image sequences from this rodent which plays a central role in the Arctic food web. Identifying lemming species in these images manually is wearisome and time-consuming. To perform this task, we present a deep neural network which has several million parameters to configure. Training a network of such an immense size with conventional methods requires a huge amount of data but a sufficiently large labeled dataset of lemming images is currently lacking. Another challenge is that images are obtained in darkness in the near infrared spectrum, causing the loss of some image texture information. We investigate whether these challenges can be tackled by a transfer learning approach in which a network is pretrained on a dataset of visible spectrum images that does not include lemmings. We believe this work provides a basis for moving toward developing intelligent software programs that can facilitate the analysis of videos by biologists.

Keywords: transfer learning; deep neural network; species identification; lemmings)

1. Introduction

Lemmings are small mammals native to the arctic and subarctic zones, where they remain active throughout the long winter under the snow [1]. Their dramatic population fluctuations [2] affect the tundra food web, as several arctic predators are dependent on them for their survival [3]. Despite the crucial role of winter reproduction in these population dynamics, our understanding of subnivean animals during this long season remains very limited due to the difficulty of studying them in harsh and remote winter conditions of the Arctic. These animals are often studied using methods such as live trapping and post-winter nest surveys during summer due to significant logistical and financial constraints of direct winter studies. Camera trapping offers a promising alternative for continuous winter monitoring, yet its application in the Arctic, particularly for subnivean animals, is rare [4–6] due to technical challenges.

In earlier studies [5,7], we developed an autonomous camera system to monitor lemming activity and behavior under the snow. Several units of the proposed system have been deployed in the field, capturing thousands of short videos of lemming activity. However, manually analyzing this growing dataset is labor-intensive and time-consuming. One valuable application of this system is to estimate relative lemming abundance from recorded videos, which requires identifying species in the images.

Over the past decade, deep learning, particularly Convolutional Neural Networks (CNNs), has emerged as a powerful alternative to traditional image classification methods that rely heavily on manual feature engineering (e.g., see [8]). A key advantage of CNNs is their ability to automatically extract features directly from raw data, often achieving higher accuracy and better generalization (see [9] for a recent review of these two approaches).

Various CNN architectures, such as AlexNet, VGGNet, and InceptionNet, have been developed for image classification (see [10] for a review). Despite outperforming traditional methods, CNNs typically require substantial computational resources and large labeled datasets, limiting their applicability to real-world problems like species identification in our dataset. Transfer learning offers a practical solution to this challenge. In this approach, a CNN pretrained on a large dataset like ImageNet is used as a base model. All layers are frozen, except the last few, depending on the size of the custom dataset, and only the unfrozen layers are retrained. The core idea is that low-level visual features (e.g., edges, textures) are generally transferable from one image classification task to another.

To the best of our knowledge, no existing model addresses lemming species identification, and most prior studies have focused on classifying large animals using high-quality images captured under favorable lighting conditions. In contrast, this study investigates whether a transfer learning approach can effectively tackle the challenge of identifying lemming species from low-quality images taken in darkness using near-infrared imaging. Among the top-performing architectures for image classification, ResNet has gained widespread popularity due to its robust performance [11,12]. Accordingly, we adopt a pretrained ResNet model and fine-tune it on our dataset.

2. Data and Method

We built a dataset from videos recorded by several units of a custom camera system [5] deployed on Bylot Island in the Canadian Arctic to monitor subnivean animal behavior between August 2018 to June 2020. Non-informative and low-quality images (previously extracted from videos featuring animals) were removed in two stages: automatic filtering based on image metrics (e.g., mean intensity, Laplacian variance), followed by a manual fine-selection process. Two lemming species were identified: brown lemming (*Lemmus trimucronatus*) and collared lemming (*Dicrostonyx groenlandicus*). Since collared lemmings turn white in winter, they were labeled separately as gray collared lemming and white collared lemming. The final dataset included 1909 images of brown lemming, 2466 images of gray collared lemming, and 387 images of white collared lemming (Figure 1). Each animal was also annotated with a bounding box using custom interactive software we developed to define the region of interest for cropping.

To build a lemming species classifier, we adapted the ResNet50architecture by replacing its final layer with a three-unit dense layer to match the number of classes. Instead of training the entire network, we applied transfer learning by fine-tuning a ResNet50 model pretrained on ImageNet while freezing all layers except the last. The model was trained for 10 epochs using the Adam optimizer (learning rate = 0.001) and cross-entropy loss.

Model performance was evaluated using two variations of k-fold cross-validation: Stratified K-Fold (SKFCV) and Stratified Group K-Fold (SGKFCV). Both preserve class balance across folds, but SGKFCV also ensures that groups do not overlap. Training was conducted on both the original and cropped datasets (DS1 and DS2), resulting in four model variants (M1, M2, M3, and M4).
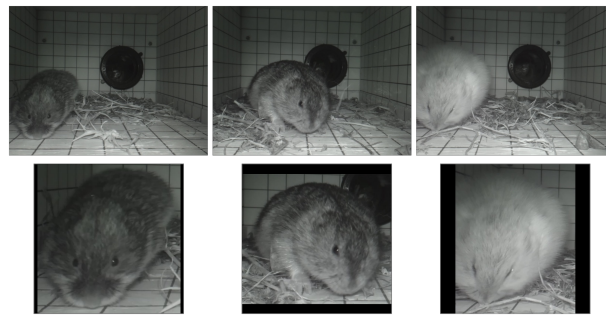
Figure 1. Examples from the lemming dataset DS1 (top row) and its cropped version DS2 (bottom row). From left: brown lemming, gray collard lemming, and white collared lemming.
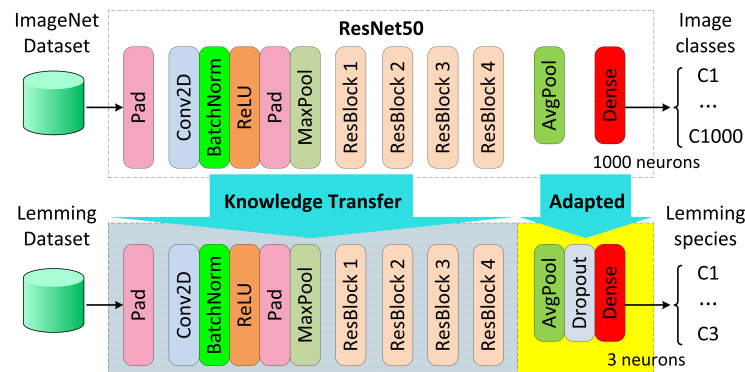

Figure 2. Lemming species classification pipeline using transfer learning from a pretrained ResNet50 model on ImageNet.

3. Results

The average accuracy across folds for the four variations of the model (M1 to M4) was 99%, 99%, 85%, 94%, respectively. The model performance in identifying each species is presented in Table 1. Models M1 and M2 performed unexpectedly well. We suspect that data leakage between the training and validation sets may have artificially inflated the results. The performance decline observed for models M3 and M4, where SGKFCV was used to ensure that images from the same sequence did not appear in both training and validation sets, supports this hypothesis. A comparison between results of M3 and M4 indicates that training on cropped images can significantly enhance model performance, likely by minimizing the influence of features from irrelevant objects such as background vegetation, and promoting focus on features related to the animals themselves.

Table 1. Confusion matrices for four model variants (M1, M2, M3, and M4) trained on datasets DS1 and DS2 using Stratified K-Fold Cross-Validation (SKFCV) and Stratified Group K-Fold Cross-Validation (SGKFCV). Each matrix shows predicted classes (columns) vs. true classes (rows). C1, C2, and C3: brown lemming, gray collared lemming, and white collared lemming, respectively.

		Predicted class											
		M1 (DS1 & SKFCV)			**M2 (DS2 & SGKFCV)**			**M3 (DS1 & SKFCV)**			**M4 (DS2 & SGKFCV)**		
		C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
	C1	**0.998**	0.001	0.001	**0.991**	0.007	0.002	**0.851**	0.147	0.002	**0.952**	0.045	0.003
True class	C2	0.016	**0.983**	0.001	0.009	**0.988**	0.003	0.179	**0.816**	0.005	0.047	**0.948**	0.005
	C3	0	0	**1**	0	0	**1**	0.007	0.003	**0.990**	0	0.008	**0.992**

4. Discussion and Future Work

Our results support the suitability of transfer learning for our image classification task. We demonstrated that features extracted by a ResNet architecture pretrained on a large dataset of generic object categories in the visible spectrum can be effectively repurposed. Specifically, these features proved useful for classifying new categories that were not part of the original training data and were captured in the near-infrared spectrum, where color and some texture information are lost. Several directions can be explored in future work: (1) comparing other architectures (e.g., VGGNet, InceptionNet) on our dataset; (2) expanding the dataset using recordings from our new system [6]; and (3) evaluating transfer learning on tasks like animal localization.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Poirier, M.; Gauthier, G.; Domine, F. What guides lemmings movements through the snowpack? *Journal of Mammalogy* **2019**, *100*, 1416–1426.
2. Krebs, C.J. *Population fluctuations in rodents*; The University of Chicago Press, Chicago, 2013.
3. Schmidt, N.M.; Ims, R.A.; Høye, T.T.; Gilg, O.; Hansen, L.H.; Hansen, J.; Lund, M.; Fuglei, E.; Forchhammer, M.C.; Sittler, B. Response of an arctic predator guild to collapsing lemming cycles. *Proc. R. Soc. B.: Biological Sciences* **2012**, *279*, 4417–4422.
4. Soininen, E.M.; Jensvoll, I.; Killengreen, S.T.; Ims, R.A. Under the snow: a new camera trap opens the white box of subnivean ecology. *Remote Sensing in Ecology and Conservation* **2015**, *1*, 29–38.
5. Kalhor, D.; Poirier, M.; Pusenkova, A.; Maldague, X.; Gauthier, G.; Galstian, T. A Camera Trap to Reveal the Obscure World of the Arctic Subnivean Ecology. *IEEE Sensors Journal* **2021**, *21*, 28025–28036.
6. Kalhor, D.; Poirier, M.; Gauthier, G.; Ibarra-Castanedo, C.; Maldague, X. An Autonomous Monitoring System with Microwatt Technology for Exploring the Lives of Arctic Subnivean Animals. *Electronics* **2024**, *13*.
7. Pusenkova, A.; Poirier, M.; Kalhor, D.; Galstian, T.; Gauthier, G.; Maldague, X. Optical design challenges of subnivean camera trapping under extreme Arctic conditions. *Arctic Science* **2022**, *8*, 313–328.
8. Low, D.G. Distinctive image features from scale-invariant keypoints. *Journal of Computer Vision* **2004**, *60*, 91–110.
9. Pagire, V.; Chavali, M.; Kale, A. A comprehensive review of object detection with traditional and deep learning methods. *Signal Processing* **2025**, *237*, 110075.
10. Chen, L.; Li, S.; Bai, Q.; Yang, J.; Jiang, S.; Miao, Y. Review of Image Classification Algorithms Based on Convolutional Neural Networks. *Remote Sensing* **2021**, *13*.
11. Zhang, R.; Zhu, Y.; Ge, Z.; Mu, H.; Qi, D.; Ni, H. Transfer Learning for Leaf Small Dataset Using Improved ResNet50 Network with Mixed Activation Functions. *Forests* **2022**, *13*.
12. Norouzzadeh, M.S.; Nguyen, A.; Kosmala, M.; Swanson, A.; Palmer, M.S.; Packer, C.; Clune, J. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc Natl Acad Sci* **2018**, *115*, E5716–E5725.