Automated counting of large vertebrate species using AutoML technology

Vladislav Sobolevskii[1,*], *Leonid* Kolpaschikov[2], *Sophia* Rosenfeld[3], and *Vladimir* Mikhailov[1]

[1]St. Petersburg Federal Research Center of the Russian Academy of Science, 14 line V.O., 39, St. Petersburg, Russia
[2]United Directorate of Taimyr Nature Reserves, Talnakhskaya 22, Norilsk, Russia
[3]Institute of Ecology and Evolution A. N. Severtsov of the Russian Academy of Science, Leninsky prospect, 33, Moscow, Russia

Abstract. The purpose of the presented work is to develop an automation system for synthesizing models of automatic recognition of different animal species in photo and video images. The paper presents a system for recognizing and counting two large vertebrate species - reindeer (Rangifer tarandus) and white-cheeked goose (Branta bernicla) on aerial images. The AutoGenNet recognition system is based on a convolutional neural network (CNN) of Mask R-CNN architecture using the concept of automatic machine learning (AutoML). The created system is able to automate a number of stages of model creation for recognizing objects in images. In particular, the presented system utilizes transfer learning. This approach significantly reduces the amount of training data required. The CNN model is synthesized automatically based on the images marked up by AutoGenNet system. To learn the Mask R-CNN model and to test the recognition accuracy, we used the images of reindeer herds obtained during aerial surveys in Taimyr and the images of brant goose flocks taken in different regions of the Arctic zone of the Russian Federation. On average, the trained software correctly recognised 82% of reindeer on the test array. Correctly recognizable brant geese accounted for 65% across the entire data set tested. Considering that this model of different animal species recognition was created automatically, with minimal involvement of machine learning specialists, this result indicates the successful application of the AutoML approach.

1 Introduction

The development of monitoring software systems for the entire natural resource sector is a critical task, given the increasing challenges and issues associated with ecosystems and natural resources. Monitoring software systems enable the assessment of the current state of ecosystems and natural resources, including air, water, soil, and biodiversity. This is essential for understanding ecosystem health and identifying changes resulting from human activities or natural factors. Additionally, these systems assist in the development of strategies for

* Corresponding author: arguzd@yandex.ru

sustainable management of natural resources, such as forests, freshwater, mineral resources, and agricultural land. This is particularly important for preventing resource depletion and ensuring availability for future generations.

In the context of livestock farming, monitoring software systems facilitate the tracking of animal health within herds. This includes disease diagnosis, identification of stress indicators, and monitoring physical conditions. Early detection of health issues helps to prevent disease outbreaks and improve overall productivity. Furthermore, these systems enable efficient management of pastures and feed, which is crucial for sustainable livestock production. This involves optimizing feeding practices, assessing pasture quality, and minimizing feed losses. To date, one of the most common methods for monitoring large herds or groups of animals is the analysis of aerial photographs. Further processing involves recognising the animals in the photographs and counting their numbers. Manual processing of photos requires quite a lot of time. For example, processing data from one season of aerial surveys of domestic reindeer herds in Taimyr takes about three months. The delay prevents prompt use of monitoring data for management decisions. The need to reduce the time required obtaining monitoring results increases the relevance of automating procedures for recognising and counting animals in photographs.

This paper presents a technology and a software system for automating the creation of animal counting systems. This will increase the speed of implementation of smart technologies in animal husbandry. We have developed an automatic object recognition system for photographs based on convolutional neural networks (CNN). CNN are a powerful tool in the field of machine learning, particularly applicable to computer vision tasks. They can be used for the automatic recognition of individual animals based on their unique characteristics, such as shape, colour, and texture of their fur. This capability can be beneficial for monitoring the health and behaviour of animals. With the aid of cameras connected to CNN, it is possible to observe the living conditions of animals, including their habitat, availability of food, water, temperature, and humidity. Moreover, the evolving of the Internet of Things (IoT) opens new opportunities for data collection regarding animals and their living conditions, allowing CNN to operate in real time and provide more accurate monitoring. CNN can be applied in various areas of livestock management — from wildlife monitoring to domestic animal husbandry. This opens new opportunities for public health protection and ecosystem management.

To date, software complexes based on CNN have been developed and are actively used in agriculture [1, 2]. In the presented work, we propose a system for automating the creation of CNN models. The automated architecture selected is the Mask Regions with Convolutional Neural Networks (MRCNN) [3]. There are already successful precedents for utilizing this architecture for monitoring tasks, including applications in agriculture [4, 5]. To demonstrate the versatility of the presented solution, the system was tested on two fundamentally different species of animals.

1.1 Reindeer

Reindeer, an important element of polar ecosystems with circumpolar distribution, is the most important object of monitoring in the Arctic zone of the Russian Federation. Some populations of reindeer have been practically exterminated and are under protection now. The populations preserved on the Taimyr Peninsula (R. t. taimyrensis), in Yakutia (R. t. taimyrensis and R. t. sibiricus) and in Chukotka (R. t. Philarchus) are main and important hunting resource [6, 7].

Monitoring data on abundance and sex-age structure of populations are the information basis for determining conservation measures and commercial take rates. Current methods for counting reindeer are based on the observation that during hot weather, when blood-sucking

insects are active and reindeer gather in large herds in a confined area in the northern part of their summer range [8]. This ecological peculiarity has been used for aerial counts of the Taimyr population since 1966.

1.2 Brant goose

This is the most highly northern species of geese, inhabiting areas in the Arctic up to the almost vegetation-free polar deserts on oceanic islands, with a circumpolar range, which makes it an indicator of the state of Arctic ecosystems [9]. It is the only species of geese that is closely associated with sea coasts throughout the annual cycle. There are three subspecies in Russia: nominate (Branta bernicla bernicla), Atlantic (Branta bernicla hrota) and Pacific (Branta bernicla nigricans). The latter two are included in the Red Book of Endangered Species of Russia. Strong fluctuations in the abundance of the species and its breeding success from year to year are known to be related to weather conditions in the Arctic [10]. The brant goose nests in colonies and forms large aggregations during moulting and migration, and does not mix with other geese species. This, combined with its almost uniform dark colouration, making it clearly distinguishable from all other waterfowl species, as well as the lack of sexual dimorphism, makes the bird an ideal object for developing methods of automatic recognition in photographs.

2 Materials and Methods

The methodology for conducting aerial surveys of reindeer populations involves several stages. First, reconnaissance flights are used to identify the locations of the largest reindeer groups. Subsequently, survey flights with light aircraft are carried out, during which concentrations of animals are detected, and direct photographic documentation of the herds takes place, while the remaining territory is covered with flight paths to account for smaller groups and individual animals. Following this, the number of reindeers within the herds is counted manually by operators on the obtained photographs, allowing for the estimation of the majority of the population directly. To assess the number of animals not included in the herds, an area approximation method is utilized [11]. Additionally, the photographs obtained aid in determining the sex and age structure of the herds based on visually distinguishable characteristics of the individuals [12, 13].

The acquired photographs exhibit a variety of landscapes, lighting conditions, and angles of view of the groups of animals in relation to the camera. Obstacles, such as rocks, hills, and depressions, make it challenging to detect the animals. The reindeer in the photographs vary in size and coloration, influenced by their physiology, lighting conditions, and the positioning of the herds. In dense summer herds, images of the reindeer may overlap, resulting in only partial visibility of their bodies. A single photograph may capture anywhere from hundreds to tens of thousands of reindeer. Moreover, the number of herd photographs taken during aerial surveys can reach several thousand. All these described features significantly complicate the task of automated individual recognition using CNN models.

Manual processing of photographs, animal counts and the determination of the sex-age structure by biologists is time-consuming. However, the final monitoring data must be provided as soon as possible for wise use, protection and/or determination of commercial take rates. For example, in order to adjust the limits of commercial take rates of reindeer for the next hunting season, data on the population status should be obtained 1-2 weeks after the aerial survey.

A similar methodology is used for aerial surveys of geese, and was described in detail previously [14].

This paper presents the AutoGenNet recognition software system, which is based on the MRCNN architecture and built using the concept of automated machine learning - AutoML. The main purpose of AutoML is to automate the processes of creating and using CNN models.

In the proposed work, the AutoML approach was utilized for several tasks: for structural-parameter synthesis of CNN models and optimization of their hyperparameters, for transfer learning of CNN models, and for the development of a user-friendly graphical interface for the software complex that allows for the automation of CNN model synthesis.

One of the key innovations of the proposed software is the application of transfer learning. This process involves several stages of training the CNN model. In the first stage, the network is trained on a large dataset of images featuring various objects (such as animals, humans, vehicles, etc.). The second stage involves fine-tuning the model on images of specific objects that need to be recognized. This significantly reduces the number of labelled images required for training the target objects: from hundreds of thousands to just a few hundred with the transfer approach. Furthermore, this method allows for the same template to be used for creating recognition systems for different animals, eliminating the need to recreate a model from scratch when tasked with recognizing a new species.

In the AutoGenNet system, the main data array for the first stage of transfer learning of the MRCNN model is the MS COCO (Microsoft Common Objects in Context) dataset, comprising 328,000 labelled images designed for detection and segmentation tasks. As previously mentioned, this training stage needs to be conducted only once. However, the MS COCO dataset does not include images of reindeer, so the MRCNN cannot differentiate them from other animals such as sheep, cattle, or horses after the first stage.

In the second stage, two versions of MRCNN are further trained: the first version on labelled images of reindeer herds and the second on images of brant goose. This stage is carried out by zoologists using the tools and user interface of the AutoGenNet system. A special program called Markup has been developed for annotating animals in images and preparing training datasets. Upon completion of the annotation, the resulting files serve as the training dataset for the second stage of transfer learning for the MRCNN.

Subsequently, the AutoGenNet system enables automated training of the MRCNN for object recognition. The user uploads the annotated training dataset files and initiates the transfer learning process. The model trained in this manner can then be used to recognize target objects in new images that were not included in the MS COCO dataset.

Images processed by the MRCNN model are returned to the user in a format compatible with editing and viewing through the Markup program. Zoologists no longer need to manually count all the animals in such images; they only need to correct any recognition errors made by the model. The corrected images can also be leveraged for subsequent iterations of transfer learning, allowing for gradual improvement of the object recognition accuracy. Additional training can be conducted on each user-edited image to enhance the precision of the initial MRCNN model, or a new MRCNN model can be trained once a sufficient quantity of training images has been gathered.

The transfer learning process in the presented system is fully automated, and the user interface is designed with non-programming users, allowing for the creation of models for recognizing and counting animals without the involvement of machine learning specialists. This level of automation makes the AutoGenNet software accessible to professionals unfamiliar with machine learning and programming.

3 Results and Discussion

To learn the MRCNN model and to test the recognition accuracy, we used the images of reindeer herds obtained during aerial surveys in Taimyr and the images of brant goose flocks taken in different regions of the Arctic zone of the Russian Federation.

3.1 Reindeer

The initial dataset for training the model consisted of 109 labelled photographs of reindeer herds. An independent validation dataset contained 30 photographs of herds. This dataset is insufficient for comprehensive training of a CNN model from scratch; however, it was adequate for creating a functional prototype of the model using the transfer learning approach.

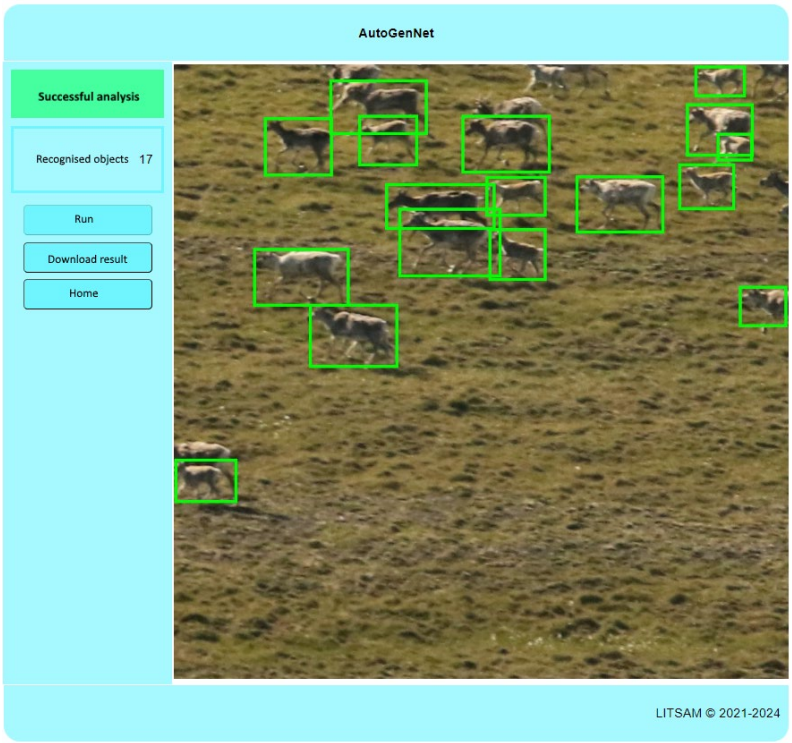

Fig. 1. Reindeer recognition on a homogeneous background.

Figure 1 shows a fragment of an aerial image of a reindeer herd on a homogeneous field from the verification dataset. Such large-scale images are used in aerial surveys to determine the sex-age structure of herds. Out of 20 fully distinguishable reindeer in the herd, three reindeer were not recognised (also, the fragments of four more reindeer on the image boundaries were not recognised). The error in counting the number of animals was 15%.

Fig. 2. Recognising and counting reindeer in a large herd against a homogeneous background.

Figure 2 shows the result of recognising reindeer in a large herd. From images at this scale, the number of animals in the population is mainly determined by direct counting. The example shows that MRCNN works well with this kind of images. The network effectively recognises animals in dense aggregations and is able to distinguish fawns even in situations where females are blocking them. The recognition error for this image was 13%.

The result of recognising a herd of reindeer shot in a boggy area with a lot of glare and inhomogeneous terrain revealed that the chosen model can work with complex inhomogeneous backgrounds. The results indicated that the recognition system did not make any false negatives (Type II errors) on the given photograph. The recognition error in this example was 14%.

On average, the trained software correctly recognised 82% of reindeer on the test array, which is quite a high result for automatic animal recognition systems in natural conditions. Considering that this model of reindeer recognition was created automatically, with minimal involvement of machine learning specialists, this result indicates the successful application of the AutoML approach.

3.2 Brant goose

The result of recognising a flock of brant geese captured against a complex water background with many glares showed that MRCNN can work with such a background. Although two brant geese were counted twice, the background objects were never mistaken by the software recognition system as images of geese. The recognition error in this example was 18%.

Fig. 3. A flock of brant geese on a homogeneous sandy background.

Figure 3 shows a flock of 30 brant geese on a homogeneous background. The machine estimated the number of birds is 33 (10% error). However, five birds were marked twice, two were not recognised or marked by the system. Thus, the real recognition error is 20%.

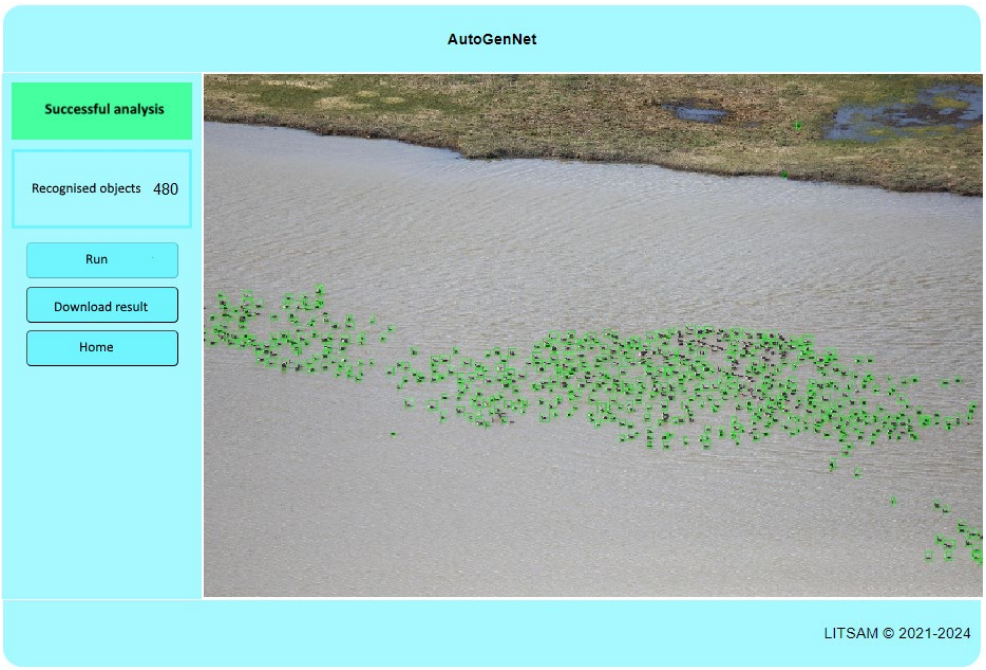

Fig. 4. A large flock of brant geese on water.

Figure 4 shows a flock of 550 brant geese on water. Machine estimation is 480 marked objects, counting error - 14.8%. The result was quite good, which confirms the possibility of using MRCNN for counting populations of brant geese on water in aggregations during the moulting period.

On average, the recognition error of brant geese in the images was about 35% for the whole tested dataset. This result is almost twice as bad as the average recognition error for reindeer.

The probable reasons that led to the decrease in accuracy are as follows:

• Brant geese can be photographed on land, on water and in the air, while reindeer are on land in the vast majority of images. With such a variety of bird photographs, the number of photographs in the training sample must be increased by at least three times the number of reindeer photographs.

• Brant geese on land and water, like reindeer, are in a two-dimensional space, while flying birds are in a three-dimensional space. As a result, the photographs of brant geese in the air may have multiple overlapping images, making recognition difficult. The paradox is that the operator can successfully recognise birds in clusters by "collecting" their images from scattered fragments, but it is quite difficult to correctly mark up such images and train a MRCNN to recognise birds on them with acceptable accuracy and this requires an additional increase in the amount of learning data.

• Open wings of birds in the air significantly complicate the animal contour geometry. In practice, learning to recognise objects with significantly changing contours requires several times more training data.

Nevertheless, the test results showed that it is possible to recognise brant geese in photographs. However, in order to increase the accuracy to the level achieved in reindeer recognition, the training sample size should be much bigger. Due to the empirical nature of the process of creating machine learning models in general and CNN in particular, it is difficult to predict the required training sample size. Based on our previous study, it can be

assumed that 300 to 2 000 of brant geese may be required to achieve the accuracy comparable to that of the reindeer recognition model.

4 Conclusion

The use of the presented system in practice will significantly simplify and accelerate the creation of systems for automatic recognition of different animal species in photo and video images. The use of the AutoGenNet software suite for recognizing and counting animals, along with the developed system for annotating and marking images called Markup, does not require in-depth knowledge of machine learning or programming. The developed software complex is not limited to specific recognition objects and can be applied for counting various subjects in photographs, such as sheep, mountain goats, geese, and more. In this case, no modifications to the existing program will be required, which makes it universal.

The number of images of specific wildlife species is usually limited. This often significantly increases the complexity of creating recognition models, which is evident from the example of brant geese. Therefore, the direction of data augmentation was chosen for further work. This is an approach that describes methods for artificially expanding the training data set by modifying the original array of images of the training sample. The use of such methods leads to an increase in the quality of training. The modification may involve color operations (changing color channels, increasing contrast or brightness), or geometric operations on the image as a whole or on individual objects in the image (image rotation, mirroring, etc.). Data augmentation methods are expected to be used in the next version of the AutoGenNet recognition system.

Peripheral tasks related to the automation of CNN model training remain unresolved. In particular, there is a need to develop a methodology or technology that can determine the required size of the training dataset in the second stage of transfer learning. The complexity of this task arises from the unique characteristics of neural network training processes, which can be highly variable and unpredictable. Currently, this problem is often addressed through empirical assumptions based on the existing experience of creating similar models.

There is a certain pattern indicating that recognition error decreases with an increasing volume of the dataset. However, this pattern only holds up to a certain threshold. If the amount of training data becomes excessive, overfitting may occur, leading to an increase in errors. The volume of data at which overfitting occurs is specific to each particular task and can only be predicted with a certain degree of conditionality. Therefore, for model development, it is recommended to conduct experiments to determine the optimal volume of training data. This, in turn, implies the need to increase the size of the training image set and the associated effort required for processing those images.

Acknowledgments

The research is supported by the grant of Russian Science Foundation №24-16-20017, https://rscf.ru/en/project/24-16-20017/, and St. Petersburg Science Foundation.

References

1. T. Thorat, B.K. Patle, S.K. Kashyap, Smart Agricultural Technology, **3**, 00114 (2023).
2. M.A. Bhatti, M.S. Syam, H. Chen, Y. Hu, L.W. Keung, Z. Zeeshan, Y.A. Ali, N. Sarhan, Big Data Research, **36**, 100465 (2024).

3. K. He, G. Gkioxari, P. Dollar, R. Girshick, Mask R-CNN. Computer Vision and Patter Recognition. arXiv:1703.06870 [cs] (2017).
4. P. Ganesh, K. Volle, T.F. Burks, S.S. Mehta, IFAC-PapersOnLine, **52 (30),** 70-75 (2019).
5. G. Zhao, J. Hu, W. Xiao, J. Zou, Chinese Journal of Aeronautics, **34 (12),** 214-226 (2020).
6. Materials to the formation of a strategy for the conservation of wild reindeer in the Arctic zone of the Russian Federation. Ed. A.L. Vaisman, O.B. Pereladova (World Wildlife Fund, Moscow, 2022).
7. V. Mikhailov, L. Kolpashikov, International Journal of Environmental Studies, **74 (5),** 862-883 (2017).
8. V.A. Zyryanov, B.M. Pavlov, G.D. Yakushkin, *Ecological bases of population counting of commercial animals in the tundra zone of the Taymir*, Ecological bases of accounting of the number of commercial animals in the tundra zone of Taymyr. Problems of hunting economy of the Krasnoyarsk region, 70-72 (1971).
9. M.V. Gavrilo, *Organisation of ornithofauna monitoring in the system of environmental monitoring in the Northern Sea Route water area*, Theses of the Second International Scientific Online Seminar "Actual issues of studying Arctic and subarctic ecosystems in the conditions of global changes in the natural environment and climate" (Salekhard, 2023).
10. E.E. Syroechkovsky, *Black goose. Field identifier of goose-like birds of Russia* (Moskow, 2011).
11. N.G. Chelintsev, *Mathematical bases of animal accounting* (Moscow, 2000).
12. L.A. Kolpashchikov, J.L. Kokorev, G.D. Jakushkin, A.L. Kolesnikov, A.M. Shapkin, I.A. Vasiliev, B.V. Shilin, V.V. Mikhailov, *Temporary methodological recommendations for aerial census of wild reindeer in Taimyr using thermal imaging and digital aerial photography equipment* (R&P, Norilsk, 2008).
13. M.G. Bondar, L.A. Kolpashchikov, Estimation of the number and summer location of the Taimyr population of wild reindeer in 2017. Scientific Proceedings of the Federal State Budgetary Institution United Directorate of Taimyr Nature Reserves, 27-45 (2018).
14. S.B. Rosenfeld, M.Y. Soloviev, G.V. Kirtaev, N.V. Rogova, M.N. Ivanov, Zoological Journal, **96 (2),** 201-221 (2017).