Reindeer Recognition and Counting System Based on Aerial Images and Convolutional Neural Networks

Vladimir Mikhailov
St. Petersburg Federal Research Centre
Russian Academy of Sciences
Saint Petersburg, Russia
mwwcari@gmail. com

Vladislav Sobolevskii
St. Petersburg Federal Research Centre
Russian Academy of Sciences
Saint Petersburg, Russia
arguzd@yandex. ru

Abstract. **Described animal recognition and counting system based on convolutional neural networks with MRCNN architecture. Initial training of the network is performed using a basic array of MS COCO images, and additional training is performed using an array of aerial photographs of reindeer herds. A web-interface of the system has been developed. The error of counting reindeer in the image from the verification sample is about 13%.**

Key words: **reindeer, recognition, convolutional neural networks, aerial imagery**

I. INTRODUCTION

The development of an automatic system for recognition and counting the number of reindeers was driven by the following reasons. The currently used methodology for counting the number of wild reindeer in tundra populations (Taimyr, Yakutia, Chukotka reindeer, and migrating herds of reindeer from Canada and Alaska) is based on the ecological characteristics of the species, consist in the fact that in hot weather, during the flight of blood-sucking insects, reindeer gather in herds of many thousands in a limited area in the northern part of their summer range (subzones of arctic and typical tundra) [1, 2]. The herds in the aggregations are photographed and the number of animals in them is counted directly, "by head".

The advantage of this approach over purely approximation-based population estimation methods [3] is the significantly greater accuracy of the results, as the vast majority of animals in the population (up to 90%) are counted directly from herd photographs and only a small number are estimated by area-based approximation. However, manual processing of the survey results for large populations takes about three months, whereas for the ecologically based management of population dynamics, the non-depletion of biological resources of the species and the determination of the norms of commercial reindeer harvesting, it is desirable to have the population data in the second half of August, that is, 10-15 days after the end of the aerial count.

The task was therefore to automate the processing of aerial photographs in order to reduce the time it takes to obtain aerial survey results.

II. SELECTING A CALCULATION MODEL

The technology of convolutional neural networks (CNN) is adopted as the intellectual basis of the recognition system. This class of architectures is a highly specialized tool, suitable primarily for images and other data that can be represented in matrix form. As images store all information as two-dimensional matrices (i.e., as pixels), it is necessary to consider not only values from the neurons themselves but also values from a group of nearest neurons when working with images. To this end, besides neurons there is another type of elements in convolutional layers of CNN that apply certain linear operations to all input data of each neuron of the layer - the convolution core. The convolutional kernel is a grid that "slides" across the image (or previous layer convolutional layer) and looks for patterns and patterns in the data. If it finds a part of the image that matches a kernel pattern, it passes a large positive value to the current layer's computational neuron. If there is no match, the kernel will pass a small value or zero.

Because the convolution kernel is applied to every position in the image, the convolution layer of CNN is extremely effective in image processing tasks because features or patterns in the images can appear anywhere in these images. That is, CNN is able to analyze context-dependent data.

The Mask Regions with Convolution Neural Networks (MRCNN) architecture [4] was chosen for this task. This architecture is a subset of the classical CNN. Due to the complexity of the architecture, it more successfully copes with the tasks of semantic and object segmentation of images. It is a modification of the existing architecture Fast Region-based Convolution Neural Network (FRCNN), in which was added a module responsible for recognition and generation of object masks.

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.