
Artificial intelligence for automatically detecting animals in camera trap images: a combination of MegaDetector and YOLOv5

Margarita Mulero-Pázmány^{*1}, Sandro Hurtado², Cristian Cardas-Ezeiza², María Luisa Antequera-Gómez², Cristóbal Barba-González², David Romero-Pacheco¹, Francisco Díaz-Ruiz³, Ismael Navas-Delgado², Raimundo Real-Giménez¹.

¹ Department of Animal Biology, University of Málaga, 29071 Málaga, Spain

² KHAOS Research group, ITIS Software, University of Málaga, 29071 Málaga, Spain

³ Conservation Biology Research Group, Department of Anatomy, Cellular Biology and Zoology, University of Extremadura, 06006 Badajoz, Spain

Abstract

Camera traps have gained high popularity for collecting animal images in a cost-effective and non-invasive manner, but manually examining these large volumes of images to extract valuable data is a laborious and costly process. Deep learning, specifically object detection techniques, constitutes a powerful tool for automating this task. Here, we describe the development and result of a deep-learning workflow based on MegaDetector and YOLOv5 for automatically detecting animals in camera trap images. For the development, we first used MegaDetector, which automatically generated bounding boxes for 93.2% of the images in the training set, differentiating animals, humans, vehicles, and empty photos. This annotation phase allowed to discard useless images. Then, we used the images containing animals within the training dataset to train four YOLOv5 models, each one built for a group of species of similar aspects as defined by a human expert. Using four expert models instead of one reduces the complexity and variance between species, allowing for more precise learning within each of the groups. The final result is a workflow where the end-user enters the camera trap images into a global model. Then, this global model redirects the images towards the appropriate expert model. Finally, the final animal classification into a particular species is based on the confidence rates provided by a weighted voting system implemented among the expert models. We validated this workflow using a dataset of 120.000 images collected by 100 camera traps over five years in Andalusian National Parks (Spain) with a representation of 24 mammal species. Our workflow approach improved the global classification F1-score from 0.92 to 0.96. It also increased the precision for distinguishing similar species, for example from 0.41 to 0.96 for *Capreolus capreolus*; and from 0.24 to 0.73 for *Dama dama*, often confounded with other ungulate species, which demonstrates its potential for reliable animal detection in camera trap images.

Keywords: automatically detecting animals, mammals, deep learning

*Corresponding author: muleromara@uma.es