Using AI for Review of Automated Remote Wildlife Camera Data

Arash Shamseddini [a,*], Dennis Thideman [a], Logan Thideman [a]

[a] *EDI Environmental Dynamics Inc., 3810 18th Ave., Prince George, BC, Canada*

Abstract: There is growing potential to collect large amounts of data using automated remote technologies during field-based programs. These data are used to analyze project-related effects and to establish baseline conditions upon which to compare future monitoring data. New opportunities for using AI in processing and analyzing arise often, yet effectively using, verifying, and ultimately integrating these techniques into effective workflows takes time and diligence.

As one example, we will present our successful experience in using object recognition AI in processing remote, automated wildlife imagery data. Object recognition was available in its early stages in 2008 to aid in photo tagging. The potential to use facial recognition in processing wildlife camera data was recognized, tested, and verified in the years that followed. Some of the challenges that were overcome to integrate this tool fully into the analysis workflow include inaccuracies, incomplete data processing, and potential for bias towards certain species or environmental conditions. Overcoming these challenges required a patient and cautious approach to reap the benefits of accurate and reliable data. Use of AI is a factor that allows automated data programs to increase in scale to obtain defensible, quantitative data analysis for effects assessment and monitoring, given the use of an effective study design.

1. Introduction

In recent decades, there has been an exponential growth in the advancement of technology incorporated into wildlife conservation efforts. From Unmanned Aerial Vehicles (UAVs) such as drones to camera traps, Acoustic Recoding Units (ARUs) and more, all of these have become standard tools in wildlife studies [1-2]. These technological innovations have immensely addressed various conservation needs and facilitated better insight and effective management of wildlife resources [3]. In fact, the recent increase in public and academic interest in preserving biodiversity has led to the growth of the field of conservation technology [4]. The term ‘Conservation Technology’ was first proposed by Berger-Tal in 2018 to broadly describe the use of technology to manage and conserve wildlife [5] which aims to increase accessibility to tools and modern technology to address conservation problems in entirely new ways [4].

Accurate detection of individual animals is integral to the management of vulnerable wildlife species [6]. A 2022 paper reports that out of more than 120,000 species monitored by the International Union for the Conservation of Nature (IUCN) Red List of Threatened Species, up to 17,000 have a ‘Data deficient’ status which emphasizes the importance of data collection in wildlife management initiatives [7].

All of these efforts have conventionally gravitated the focus of developments toward building high-precision devices which are now able to inexpensively capture many types of data from an ecosystem. Among these, camera traps have quickly transformed the way in which many ecologists study the distribution of wildlife species, their activity patterns and interactions among members of the same ecological community [8]. The camera trapping method has its roots in the 1890s. George Shiras developed a technique for photographing wildlife by incorporating a tripwire that an animal triggered; this innovation gained him recognition for introducing a novel approach to wildlife photography [9]. In the modern age,

* Corresponding author
Email address: ashamseddini@edynamics.com

camera traps are remote devices equipped with a motion or infrared sensor that automatically record images or videos [10]. Some typical applications of camera traps are to describe activity patterns, foraging, social behaviour, denning and antipredator behaviours. Changes to these behaviours can affect individual survivorship and fitness and given sufficient frequency and effect size, population dynamics [11].

These cameras naturally generate a huge amount of graphic data which often makes manual analysis, that mainly revolves around detection and/or classification of wildlife species, inefficient. Hence, recently the focus of attention has shifted toward developing automated tools to facilitate these analyses in the pipeline, with Artificial Intelligence (AI) which uses Deep Learning with Neural Network algorithms behind the scenes, being at the forefront of these developments. These algorithms can perform image classification and object detection after being trained using a pre-labelled dataset that uniquely identifies each species (or category) of interest [8]. Some of the use cases of the application of AI in this context are removing empty images (i.e., images without animals, also referred to as blanks [12]), species identification [13-16], species classification [14-15], or counting of individuals when there is a single species in an image [14].

These advancements in technology have led to the development and launch of AI-powered platforms such as Conservation AI [17], MegaDetector [18], MLWIC2: Machine Learning for Wildlife Image Classification [19], Wildlife Insights [20] and more. Some of these tools are under continuous development and in their paper, Juliana Vélez et al., have done a comprehensive review of some of these platforms and their applications [8]. Although AI makes it feasible to process camera trap images in a short period of time with a decent accuracy, these models require a huge amount of diverse data for training. The performance of these models may suffer when developed on a limited training data and then applied more broadly e.g., when trained on the data from a certain ecosystem and used in a different habitat or when applying the models to low resolution images [8, 13, 18-19, 21].

In this paper, we present our experience in analyzing camera trap images with MegaDetector v5.0 (referred to as MD hereafter). MD is an open-source object detection Deep Learning model developed by Microsoft specifically for the processing of camera trap data [18]. The model has been trained on millions of images from a wide range of locations and contexts and is able to detect three object classes within images, namely, humans, animals, and vehicles, and can thus be implicitly used to detect images that are blank [22]. MD is often used as a "coarse filter" in processing camera trap images where experts are only interested in reviewing images in which there are animal(s). The processing of images using MD is far faster compared to human analysis. One research reports a 500% increase in processing speed [22]. In terms of performance metrics, very promising numbers are reported in the literature [8, 22].

2. Methods

We decided to benchmark MD for processing time and performance using the data from a wildlife study that includes a multi-year remote camera trap program for one of our clients. The objective of this work is to improve understanding of seasonal wildlife use within a mining project area, particularly caribou use of lichen habitats and movement patterns of large mammals along the local trail networks. The trails generally run northeast-southwest parallel to the valley at this location. This information is used to support effective mitigation for wildlife during exploration activities in support of the wildlife mitigation plan and to provide data on wildlife habitat for consideration during the planning phase.

A total of 24,595 images were first analyzed by MD. For this analysis, a machine with an NVIDIA T1000 Graphics Processing Unit (GPU) was used. Next, and to have a baseline, a human expert was asked to analyze the same images using Timelapse, a software used by wildlife scientists for management and processing of camera trap images and videos [21]. The results of the two approaches were then compared.

To simplify the benchmarking, the following assumptions were made:

- The "animal" class was assumed as the target class for both analyses.
- Only detections with confidence ≥ 0.8 were preserved and low-confidence detections were excluded from the pool. Based on MD's distribution of detection confidence (image below), with the 25th percentile of the distribution a little over 0.8, this is a fair assumption.
- For images where MD generated multiple detections, only the detection with the highest confidence (maximum confidence detection) was retained.

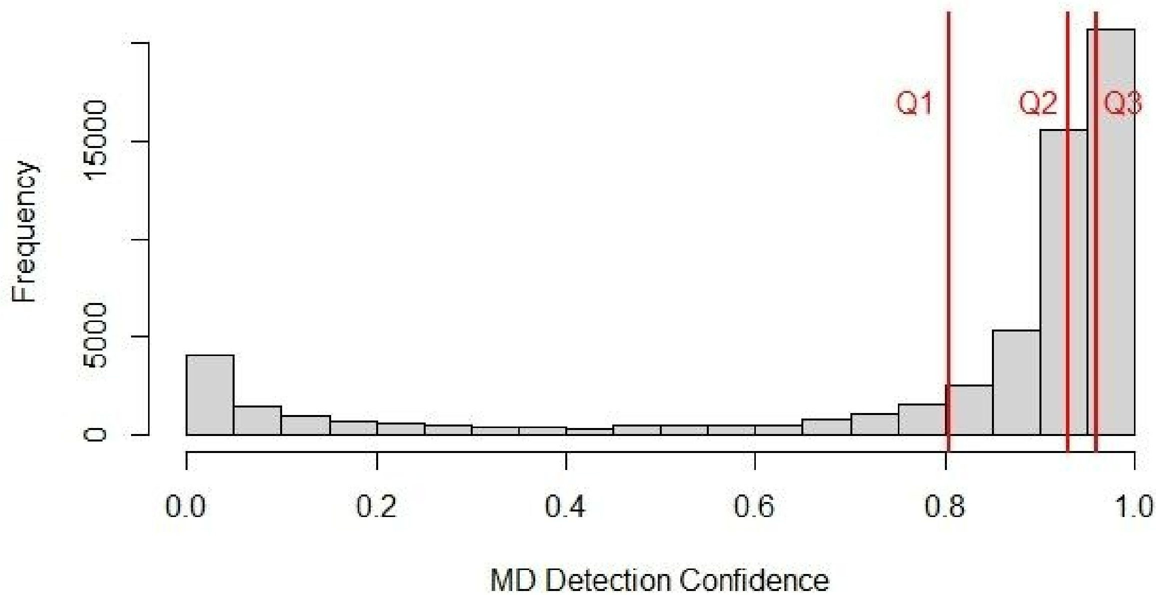

Figure 1: Distribution of MegaDetector Detection Confidence

- The comparison is made at the episode level. In camera trap terminology, an episode is the sequence of images that capture an activity/motion when the camera is triggered [21]. It was assumed that for the images within an episode, if MD gets at least one prediction right, it would be considered that it has made the correct prediction for that episode.

3. Results

3.1. Processing speed

Our computer was able to process the images at approximately 1.25 images per second. This result is in alignment with the benchmark timings reported on MD's website [23]. When compared, we concluded that processing with MD is roughly six times faster than manual processing.
In manual analysis we varied in time to review images. On average users required up to 5 seconds per image. Similar values were observed across multiple projects and common literature.

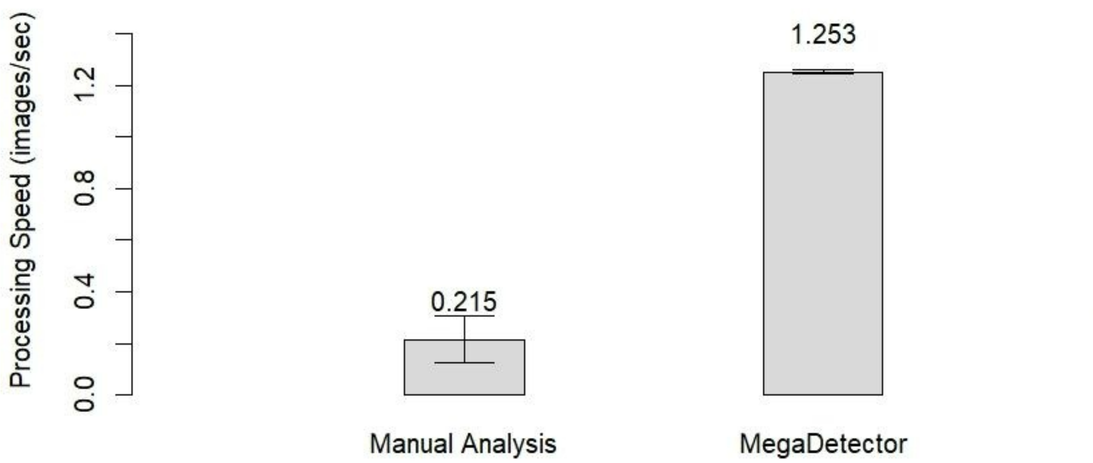

Figure 2: Comparison of Processing Speed between MegaDetector and Manual Analysis

3.2. Performance

To evaluate MD's performance, we used precision ($\frac{True\ Positive}{True\ Positive+False\ Positive}$) and recall ($\frac{True\ Positive}{True\ Positive+False\ Negative}$) as performance metrics.

Based on our results both precision and recall were calculated to be 100%, i.e., MD did not miss any "animal" and did not falsely predict an image as an "animal" class either which was quite impressive. This information was used as a benchmark for later application of MD in our workflow.

3.3. Beyond benchmarking

We have successfully integrated MD into our camera trap data management workflow and used it, in several other projects, mainly to filter out the blank, i.e., noise, from the source images which has continued to be both efficient and precise. Our dedicated image data analysis workstation is equipped with an NVIDIA RTX A4000 GPU which is approximately 6.65 times faster than our benchmark machine.

4. Conclusion

Camera traps have been used for quite a while in ecological research to study wildlife behaviour in a non-invasive fashion. The main issue with this technology is the huge amount of visual data being generated which often makes manual analysis a formidable challenge. The application of Artificial Intelligence approaches using deep learning algorithms to process the outputs has recently gained traction in the field. Using a computer that runs on GPU, these models are able to process hundreds of thousands if not millions of images in a single day, some with decent accuracy but still with room for improvement given the limitations in the training phase.

In this paper we went over our experience with MegaDetector (MD), an object detection AI tool developed by Microsoft. We started with a benchmark test run of MD to have a better understanding of its performance. Our benchmark results showed that MD is on average close to six times faster than our routine manual analysis (baseline). Also, for the "animal" class as the target class, MD's predictions on our dataset were practically flawless (precision = recall = 100%). The impressive performance of MegaDetector in this test run, encouraged us to continue using it as an integral piece in our camera trap image data management workflow.

References

[1] Elmore, Jared A., et al. "Evidence on the efficacy of small unoccupied aircraft systems (UAS) as a survey tool for North American terrestrial, vertebrate animals: a systematic map." Environmental Evidence 12.1 (2023): 3.
[2] Lahoz-Monfort, José J., and Michael JL Magrath. "A comprehensive overview of technologies for species and habitat monitoring and conservation." BioScience 71.10 (2021): 1038-1062.
[3] Mathur, Pradeep K., Bilal Habib, and Prateek Mathur. "Technology advancement and integration in the context of wildlife conservation." Human Bond Communication: The Holy Grail of Holistic Communication and Immersive Experience (2017): 115-130.
[4] Schulz, Andrew K., et al. "Conservation tools: the next generation of engineering–biology collaborations." Journal of the Royal Society Interface 20.205 (2023): 20230232.
[5] Berger-Tal, Oded, and José J. Lahoz-Monfort. "Conservation technology: The next generation." Conservation Letters 11.6 (2018): e12458.
[6] Corcoran, Evangeline, et al. "Automated detection of wildlife using drones: Synthesis, opportunities and constraints." Methods in Ecology and Evolution 12.6 (2021): 1103-1114.
[7] Tuia, Devis, et al. "Perspectives in machine learning for wildlife conservation." Nature communications 13.1 (2022): 1-15.
[8] Vélez, Juliana, et al. "An evaluation of platforms for processing camera-trap data using artificial intelligence." Methods in Ecology and Evolution 14.2 (2023): 459-477.
[9] Brower, Matthew. "George Shiras and the circulation of wildlife photography." History of photography 32.2 (2008): 169-175.
[10] Nazir, Sajid, et al. "WiseEye: Next generation expandable and programmable camera trap platform for wildlife research." PLoS one 12.1 (2017): e0169758.

[11] Caravaggi, Anthony, et al. "A review of factors to consider when using camera traps to study animal behavior to inform wildlife ecology and conservation." Conservation Science and Practice 2.8 (2020): e239.
[12] Beery, Sara, Grant Van Horn, and Pietro Perona. "Recognition in terra incognita." Proceedings of the European conference on computer vision (ECCV). 2018.
[13] Villa, Alexander Gomez, Augusto Salazar, and Francisco Vargas. "Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks." Ecological informatics 41 (2017): 24-32.
[14] Norouzzadeh, Mohammad Sadegh, et al. "Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning." Proceedings of the National Academy of Sciences 115.25 (2018): E5716-E5725.
[15] Schneider, Stefan, Graham W. Taylor, and Stefan Kremer. "Deep learning object detection methods for ecological camera trap data." 2018 15th Conference on computer and robot vision (CRV). IEEE, 2018.
[16] Tabak, Michael A., et al. "Machine learning to classify animal species in camera trap images: Applications in ecology." Methods in Ecology and Evolution 10.4 (2019): 585-590.
[17] Chalmers, Carl, et al. "Conservation AI: Live stream analysis for the detection of endangered species using convolutional neural networks and drone technology." arXiv preprint arXiv:1910.07360 (2019).
[18] Beery, Sara, Dan Morris, and Siyu Yang. "Efficient pipeline for camera trap image review." arXiv preprint arXiv:1907.06772 (2019).
[19] Tabak, Michael A., et al. "Improving the accessibility and transferability of machine learning algorithms for identification of animals in camera trap images: MLWIC2." Ecology and evolution 10.19 (2020): 10374-10383.
[20] Ahumada, Jorge A., et al. "Wildlife insights: A platform to maximize the potential of camera trap and other passive sensor wildlife data for the planet." Environmental Conservation 47.1 (2020): 1-6.
[21] Greenberg, Saul, Theresa Godin, and Jesse Whittington. "Design patterns for wildlife-related camera trap image analysis." Ecology and Evolution 9.24 (2019): 13706-13730.
[22] Fennell, Mitchell, Christopher Beirne, and A. Cole Burton. "Use of object detection in camera trap image identification: Assessing a method to rapidly and accurately classify human and animal detections for research and application in recreation ecology." Global Ecology and Conservation 35 (2022): e02104.
[23] www.github.com/agentmorris/MegaDetector. Accessed March 2024.