Improving automatic wildlife distance estimation in camera trap images

Adam M. Kosinski, Shiyue Wang

Introduction

In wildlife conservation, it is very important to know population density.

Camera trap distance sampling (Howe et al. 2017) is an effective way of measuring population density, but it requires knowing the distance from the camera to the animal in each camera trap photo.

Goal: Improve on automated methods to estimate animal-to-camera distance in camera trap images, and create a user interface to run the method.

Existing methods

Haucke et al. 2022 proposed an automatic distance estimation method by combining an AI-generated depth map and AI-located animal bounding boxes to extract distance information.

Johanns et al. 2022 improved upon Haucke et al. 2022's method by using a fully automated calibration method and incorporating a segmentation model.

Data

Our camera trap images are from camera sites in Duke Forest, taken from March to May 2023.

https://pixabay.com/photos/camera-trap-danger-forest-grass-4058597/

Our Distance Estimation Method

Arrow color coding

Calculate Depth

Locate Animal

To obtain animal distance from the camera (depth), we need
1) the depth of pixels in the image, and
2) which pixels in the image are the animal.

INPUT - Camera Trap Image

Use robust principal component analysis (RPCA) to mark pixels that differ between images due to animal movement; these pixels are likely to be the animal

Run MegaDetector v5a Model

Bounding Boxes

Clip to bounding box

Segmentation Mask
white / black = animal / not animal

Crop around animal

Run ZoeD-M12-NK Depth Model

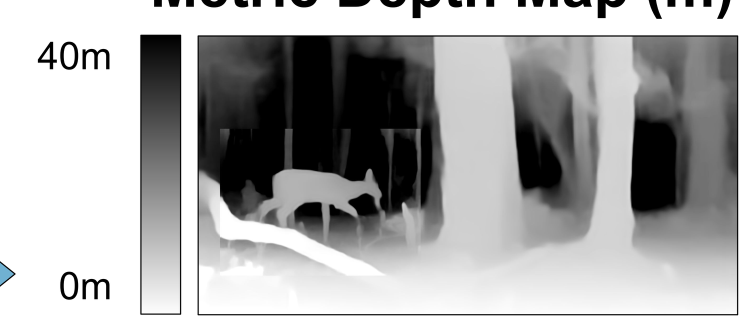
Relative Depth Map

Crop produces more accurate animal depth

Align and paste into relative depth map

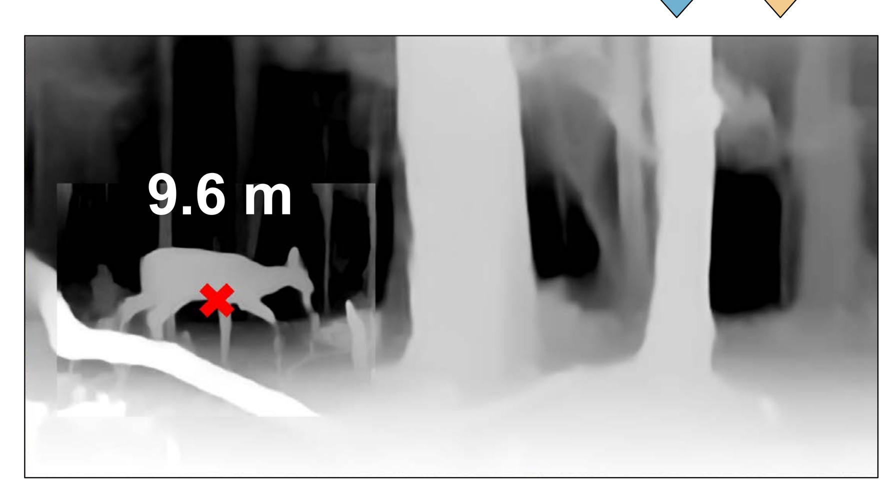
Use a manually calibrated image from this site to convert from unitless relative depth to depth in meters

Metric Depth Map (m)



Get median depth of pixels in the segmentation mask

OUTPUT
Animal distance from camera

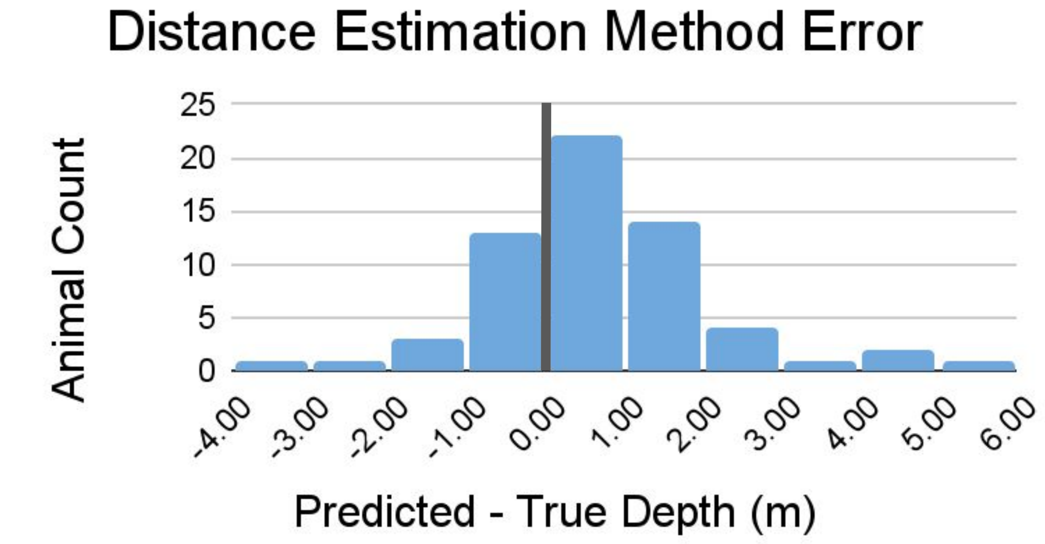


Method adapted from Haucke et al. (2022) and Johanns et al. (2022)

Results

We evaluated our method on a test set of 51 images from 8 camera sites, which contained 62 animals.

We estimated true distances manually using known distances to scene locations.



For context, humans can usually only estimate distance to within 1m.

Improvements

Both previous methods use an older depth model (DPT), and don't use cropped animal depth. Other differences:

- Haucke et al. 2022 does not use segmentation.
- Johanns et al. 2022 uses an automated calibration method, and the DINOv1 segmentation model.

Method	Mean Error (m)	Mean Absolute Error (m)	RMSE (m)
Haucke et al. 2022	2.23	2.82	4.69
Johanns et al. 2022	-1.67	2.43	3.41
Our method	**0.67**	**1.19**	**1.65**

Metrics were evaluated on our test set of 51 images.

Qualitative observations:

- Haucke et al. 2022: most error from several outliers with hugely overestimated distances
- Johanns et al. 2022: automatic calibration predicted depths 3-7m regardless of true depth → large underestimates at far distances

References

Beery, S., Morris, D., & Yang, S. (2019). *Efficient Pipeline for Camera Trap Image Review* (arXiv:1907.06772). arXiv. http://arxiv.org/abs/1907.06772

Bhat, S. F., Birkl, R., Wofk, D., Wonka, P., & Müller, M. (2023). *ZoeDepth: Zero-shot Transfer by Combining Relative and Metric Depth* (arXiv:2302.12288). arXiv. http://arxiv.org/abs/2302.12288

Giraldo-Zuluaga, J.-H., Salazar, A., Gomez, A., & Diaz-Pulido, A. (2019). Camera-trap images segmentation using multi-layer robust principal component analysis. *The Visual Computer*, *35*(3), 335–347. https://doi.org/10.1007/s00371-017-1463-9

Haucke, T., Kühl, H. S., Hoyer, J., & Steinhage, V. (2022). Overcoming the distance estimation bottleneck in estimating animal abundance with camera traps. *Ecological Informatics*, *68*, 101536. https://doi.org/10.1016/j.ecoinf.2021.101536

Howe, E. J., Buckland, S. T., Després-Einspenner, M., & Kühl, H. S. (2017). Distance sampling with camera traps. *Methods in Ecology and Evolution*, *8*(11), 1558–1565. https://doi.org/10.1111/2041-210X.12790

Johanns, P., Haucke, T., & Steinhage, V. (2022). Automated distance estimation for wildlife camera trapping. *Ecological Informatics*, *70*, 101734. https://doi.org/10.1016/j.ecoinf.2022.101734

Acknowledgements

Project lead: John Poulsen, PhD
Project manager: Ankita Gupta
Duke Forest camera trap images provided by Libby White.
Data+ coordinators: Ariel Dawn, Paul Bendich