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.

Our Distance Estimation Method
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

1. 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
2. Use a manually calibrated image from this site to convert from unitless relative depth to depth in meters
3. Align and paste into relative depth map
4. Get median depth of pixels in the segmentation mask

OUTPUT - Animal distance from camera

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.

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.

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

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

References