

CAMERA-TRAP CLASSIFICATION WITH DEEP LEARNING UNDER GROUND TRUTH UNCERTAINTY

Leonard Hockerts, Peter Stewart, and Tiffany Vlaar

NeurIPS 2025 Tackling Climate Change with Machine Learning Workshop



PRICKLY PEAR PROJECT KENYA

Publicly available dataset (Stewart et al., 2024):

- Camera trap images with associated volunteer classifications from Zooniverse.
- Subset of images (~14%) classified by expert.
- >40 species of East African mammal, as well as >60 species of bird.

Substantial class imbalance.

We focus on the 17 most commonly observed species.

CLIMATE CHANGE IMPACT

Climate change is rapidly affecting biodiversity across the globe. Camera traps are an indispensable tool for monitoring biodiversity. Deep learning can aid timely analysis of camera trap data, but label ground truth uncertainty may affect (species-specific) accuracy of models.

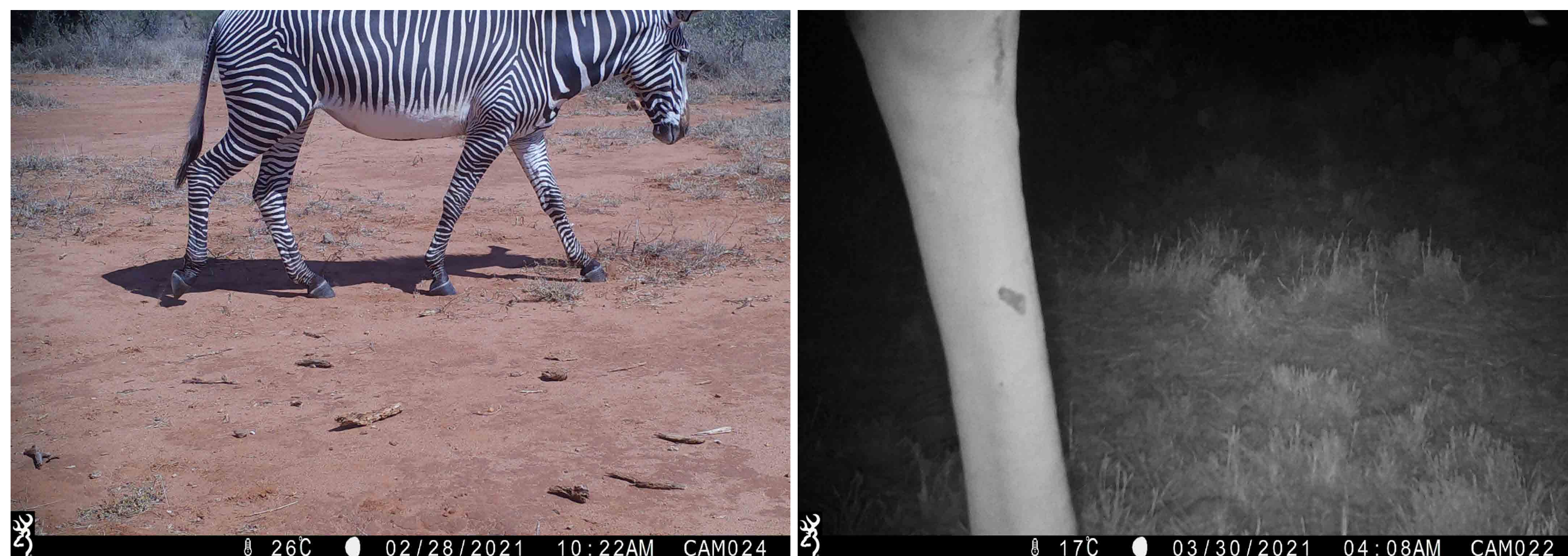
We show that:

- Including training images with a moderate degree of volunteer disagreement can, in combination with pre-training, improve performance while avoiding discarding valuable data.
- There are differences in species-specific classification biases between volunteers and models, suggesting avenues for future model refinement and potential of combined citizen science and deep learning approaches to produce more accurate performance.

The involvement of volunteers can support these efforts, as well as deepen the public's connection with the ongoing biodiversity crisis (Green et al., 2020).

GROUND TRUTH UNCERTAINTY

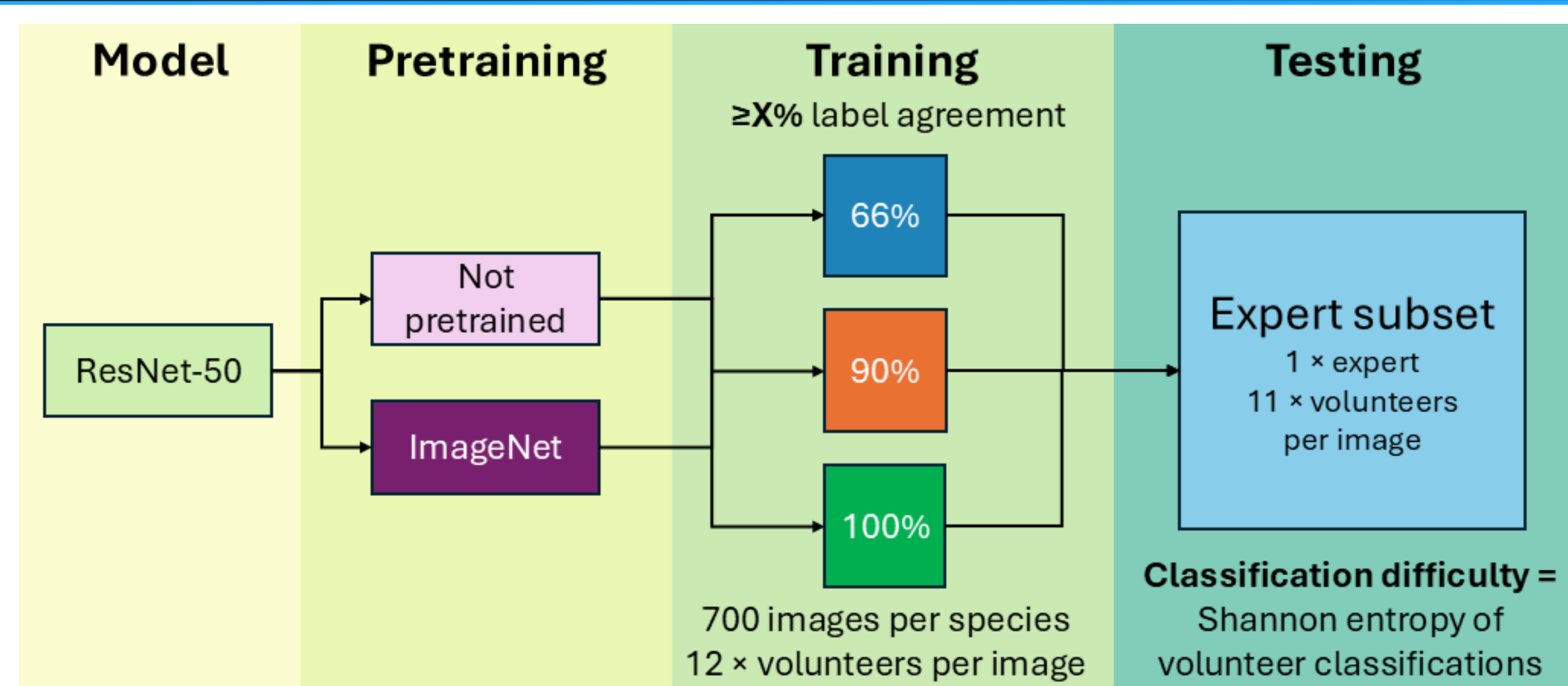
Labels for camera trap data obtained through citizen science approaches typically feature disagreement amongst volunteers, i.e. label ground truth uncertainty.



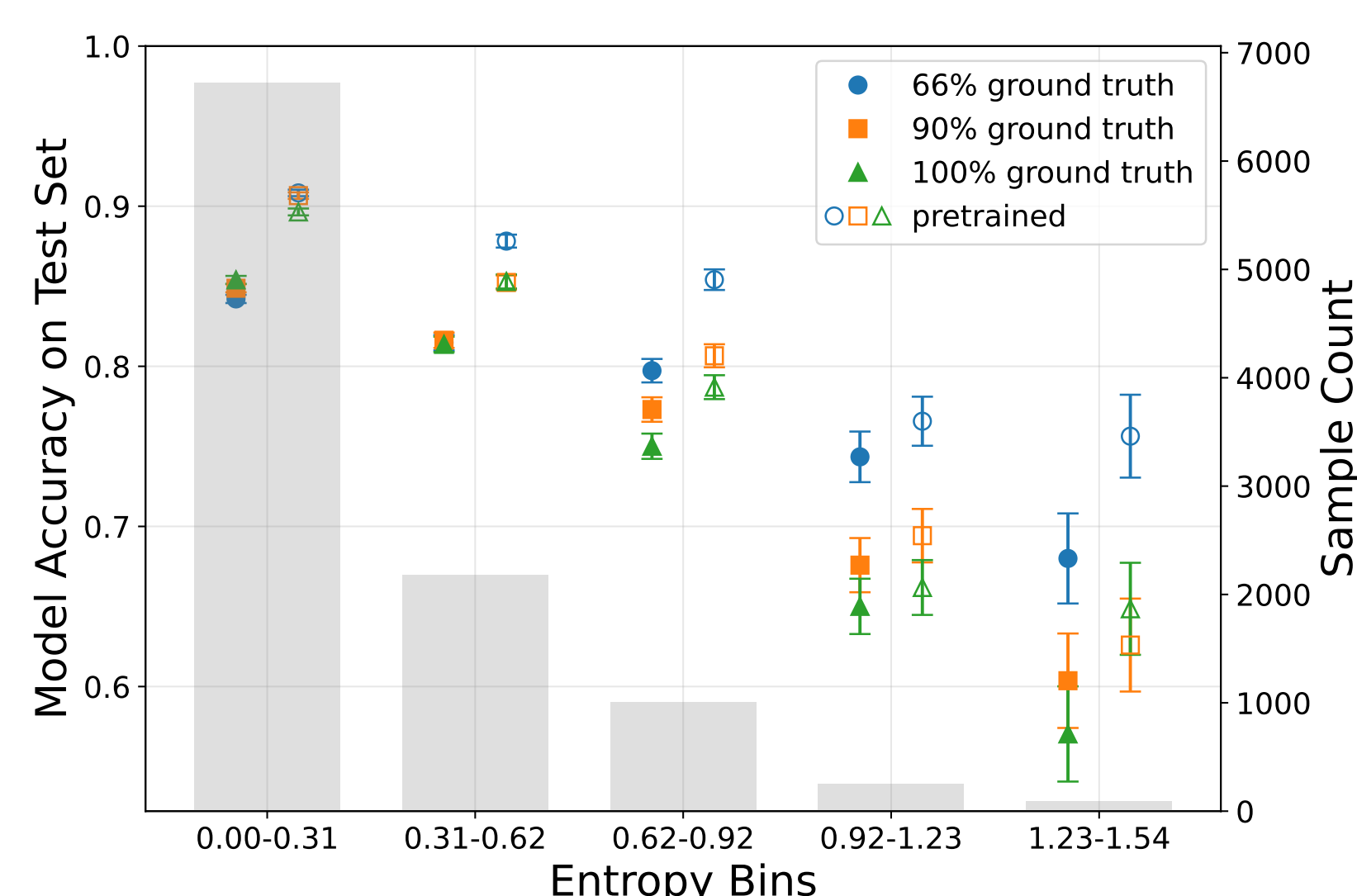
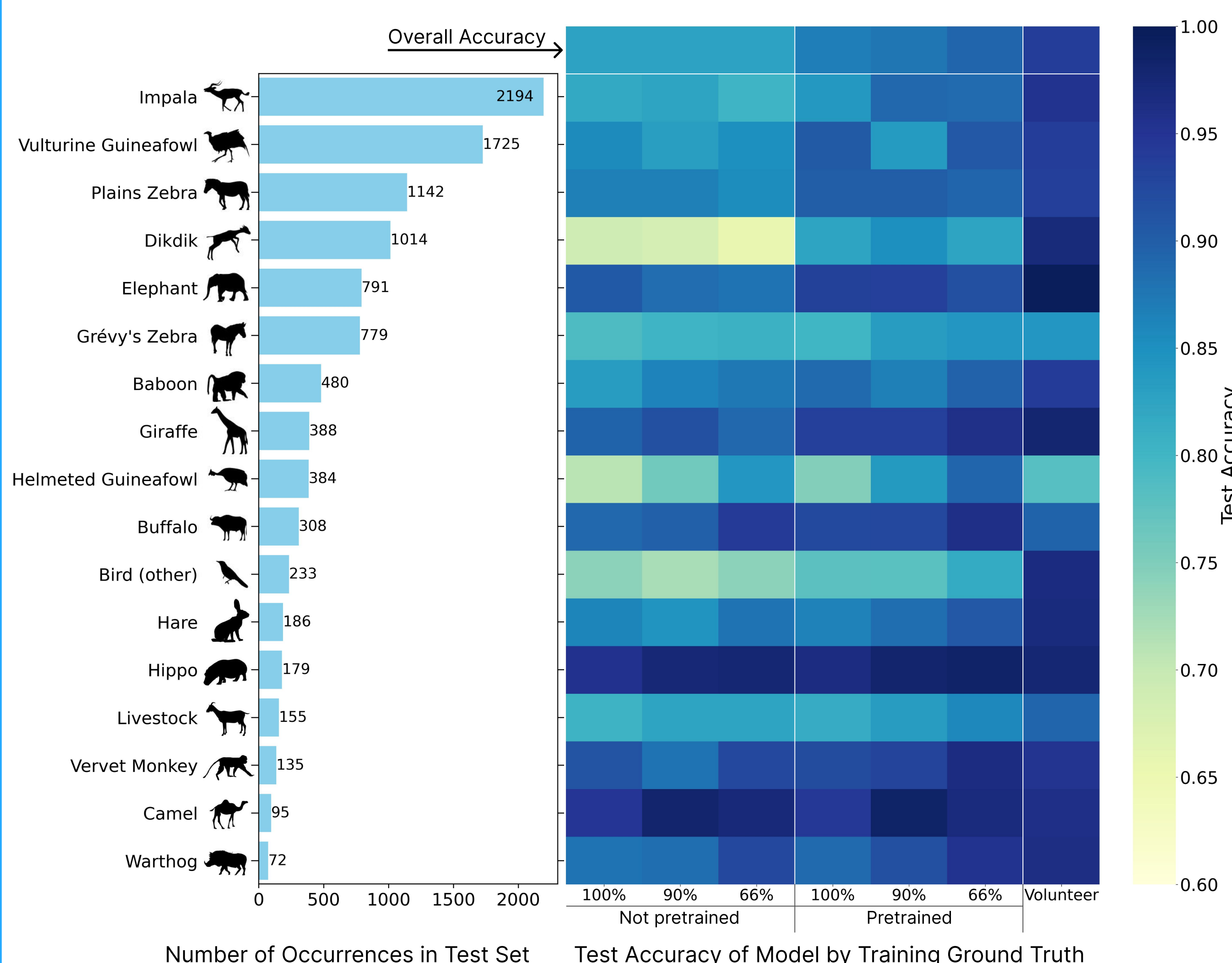
Volunteer Classifications:
11× Grévy's zebra, 1× plains zebra

Volunteer Classifications:
8× Giraffe, 2× Dikdik, 1× Camel, 1× Other

METHODOLOGY



RESULTS



- We find fundamental differences in the test accuracy of species classifications from camera trap data made by deep learning models to those made by human classifiers.
- We show that model design choices can enhance species-specific performance.
- We demonstrate that training on data with higher levels of ground truth uncertainty can improve performance at test time on images that are difficult to classify for humans.
- We observe that model pre-training consistently improves test accuracy, and amplifies the effects of training on data with higher levels of ground truth uncertainty.

FUTURE WORK

- Addressing severe class imbalance.
- Study how model design choices impact performance under ground truth uncertainty and resulting carbon footprint.
- Consider other classification difficulty metrics.
- Model vs. volunteer behaviour.

REFERENCES

- Dataset: <https://doi.org/10.5281/zenodo.15584756>
- P. Stewart, R.A. Hill, A.M.O. Oduor, P.A. Stephens, M.J. Whittingham, and W. Dawson. Multi-species impacts of invasive opuntia cacti on mammal habitat use. Ecology Letters, 2025.
- S.E. Green, J.P. Rees, P.A. Stephens, R.A. Hill, A.J. Giordano. Innovations in Camera Trapping Technology and Approaches: The Integration of Citizen Science and Artificial Intelligence. Animals, 10(1):132, 2020.
- C. Shannon and W. Weaver. The mathematical theory of communication. University of Illinois Press, 1949.

Questions/comments reach out at Tiffany.Vlaar@glasgow.ac.uk