Understanding the Impact of Training Set Size on Animal Re-identification

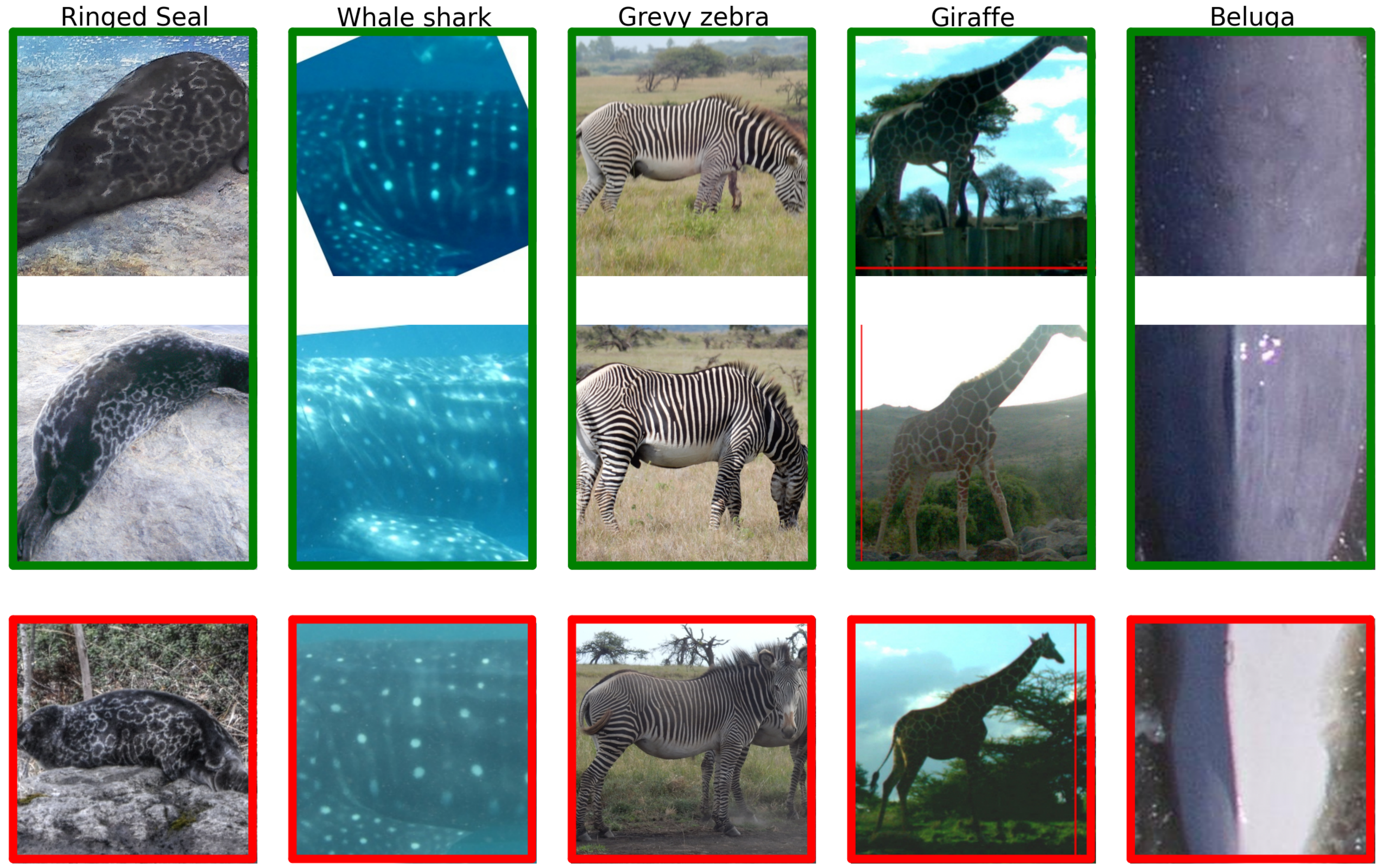
Aleksandr Algasov[1], Ekaterina Nepovinnykh[1], Tuomas Eerola[1], Lasha Otarashvili[3], Charles V. Stewart[2], Heikki Kälviäinen[1], and Jason A. Holmberg[3]

[1]Computer Vision and Pattern Recognition Laboratory, Lappeenranta-Lahti University of Technology LUT, Finland
[2]Department of Computer Science, Polytechnic Institute, Troy, NY, USA
[3] Wild Me, Conservation X Labs, Portland, OR, USA

Motivation

Re-identification and tracking individual animals provides various possibilities for ecological research and conservation, but traditional methods like tagging and GPS collars have limitations. Re-identification methods using image processing offer a non-invasive alternative, allowing for the identification of animals based on unique physical traits. However, these methods face challenges with the labor-intensive task of collecting and annotating large datasets. This study investigates the impact of the reduced number of samples in training data on the accuracy of re-identification methods.

Data



Sample images from datasets. Images connected with green borders are from the same individual, while images with red borders are from different individuals

Re-identification Methods

- Local feature-based methods:
  - **Hotspotter[1]**. SIFT-based re-identification algorithm that highlights distinctive keypoints or "hot spots" on body markings and ensures geometric consistency with RANSAC-based reranking.
  - **ALFRE-ID[2]**. Animal re-identification framework that extracts pelage patterns using pre-trained CNN-based feature extractors, aggregates them with Fisher vectors, and performs spatial reranking. It doesn't need training on the target dataset, using extractors suited to different species.
- End-to-end learning-based methods:
  - **MiewID**. An end-to-end learning-based re-identification method that utilizes the ArcFace loss with an EfficientNet CNN architecture as a backbone.
  - **TransReID[3]**. A transformer-based re-identification method originally developed for persons and vehicles. It uses overlapping image patches and a Jigsaw Patch Module to enhance robustness against occlusions and misalignments.
  - **MegaDescriptor[4]**. A swin Transformer-based end-to-end algorithm specifically designed and trained for animal re-identification.
  - **BioCLIP[5]**. A CLIP model for general biology tasks, trained on TreeOfLife-10M, with over 10 million images covering 454 thousand taxa. For re-identification purposes internal vision transformer is used as an end-to-end method.

Datasets and Re-identification Features

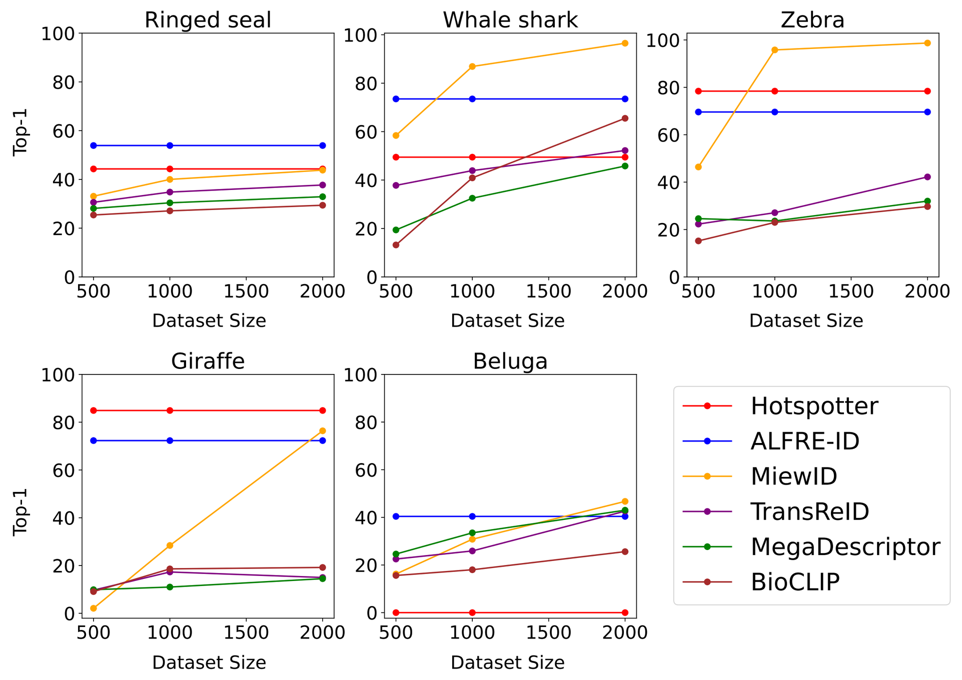
Datasets and data splits used in the experiments.

	Saimaa ringed seal			Whale shark			Grevy's zebra			Masai giraffe			Beluga whale		
	Train	Test	Val	Train	Test	Val	Train	Test	Val	Train	Test	Val	Train	Test	Val
Total annotations	2000	480	462	4621	933	806	2535	690	631	2227	1471	583	2290	582	579
Number of individuals	277	146	146	306	119	120	215	345	174	455	695	89	286	188	187
Average annots per class	14	4.8	4.6	15.1	7.8	6.7	11.8	2	3.6	5.0	2.1	6.6	8	3.1	3.1
Unseen classes in test	-	0.54	0.48	-	0.55	0.57	-	0.4	1.0	-	1.0	1.0	-	0.5	0.45
Identifiable feature	unique and permanent pelage ringed patterns			distinctive spot patterns around dorsal fins, and gills			unique stripe patterns all over body			spot patterns and neck markings			may have unique scars or pigmentation variations		

Acknowledgements

The authors would like to thank Finnish Cultural Foundation for funding the research. In addition, authors would like to thank Department of Environmental and Biological Sciences at the University of Eastern Finland (UEF), the Zoological Society of San Diego d/b/a San Diego Zoo Wildlife Alliance and Wild Me, a lab of Conservation X Labs, for providing the data and identifying each individual.

Results



Dependency of Top-1 accuracies of methods on dataset size for each of five species.

- **HotSpotter** provides high performance for Grevy's Zebra and Masai Giraffe, but fails on Beluga Whale.
- **ALFRE-ID** excels when amount of training data is low or does not exist.
- **MiewID** is generally the best method when the amount of training data is large. The performance degrades with less training data.
- **TransReID** is inconsistent for Saimaa Ringed Seal and Masai Giraffe, but provides reasonable good performence for other datasets.
- **MegaDescriptor** suffers from the small amount of training data.
- **BioCLIP** performs well on Whale Sharks, but has low accuracy for other species. The performance affected by the small input size.

Discussion

- Large training datasets favor end-to-end learning-based methods, particularly CNN-based models like MiewID.
- Transformer-based methods require even more training data due to higher parameter counts rendering them unsuitable for species-specific models on most species.
- Local feature-based methods are superior on small training datasets.
- Intra-individual variance and species-specific pattern characteristics significantly impact the performance of re-identification methods.

References

[1] Crall, J.P., Stewart, C.V., Berger-Wolf, T.Y. et al. HotSpotter - Patterned species instance recognition, *WACV* (2013).

[2] Nepovinnykh, E., Chelak, I., Eerola, T. et al. Species-Agnostic Patterned Animal Re-identification by Aggregating Deep Local Features, Int J Comput Vis (2024).

[3] S. He, H. Luo, P. Wang, F. Wang, H. Li and W. Jiang, TransReID: Transformer-based Object Re-Identification, *ICCV* (2021).

[4] Čermák, Vojtěch, et al. WildlifeDatasets: An open-source toolkit for animal re-identification, WACV (2024).

[5] Stevens, Samuel, et al. BioCLIP: A Vision Foundation Model for the Tree of Life, CVPR (2024).