Can Poachers Use Image Geolocalization Models to Find Out Sensitive Species Locations from Open-source Camera Trap Wildlife Images?

Author: Wenxin Zhao (Rose)
Contact: wenxin.zhao.gr@dartmouth.edu

2024/05/11

Introduction and Related Work

There are ample open-source wildlife image datasets taken by camera traps to be harnessed by ML researchers for biology and conservation [16, 3**?** , 1], many of which are used as benchmarks for training large models [9, 14, 11]. Camera traps effectively capture many species, including sensitive ones, in the same location over long periods [17]. The geolocation of the photos has shown useful in providing additional context when performing classification and improving identification accuracy [7, 2]. Therefore, it became common for datasets to add such metadata into their annotations [3, 5, 10, 1]. However, this information poses substantial risks of exposing sensitive species' living habitat locations [15], and there have been cases of informing poachers to locate them [20]. To protect this privacy information, two notable contributors of wildlife datasets, Wildlife Insight by Google, and iNaturalist, attempted to protect against poachers by obfuscating the geolocation of endangered species by averaging with other locations and truncating decimals [1] or adding a radius of location accuracy [10]. However, little or no research has been conducted to systematically study if wildlife geolocation information is safely protected.

Image geolocation prediction model predicts location by matching the input image features to the GPS data features [6]. Camera trap images can be particularly vulnerable to image geolocalization models because the datasets include multiple images taken in the same locations over different times, with a large proportion of the images consisting of the surrounding environment, providing more information about a location than a single image of a close-up view of an animal alone. In one attempt to evaluate this geolocation privacy concern, Beery [4] assessed with PlaNet [21], which used a CNN structure, and concluded privacy was preserved. However, as of 2024, with the development of ML models from CNN to Transformers[18], image geolocalization models have since improved drastically [12, 6, 8]. In particular, GeoCLIP, using clip and transformer [13], supports global-scale localization with the top 1 in street level (1km) photo geolocation estimation on Im2GPS3k benchmark [19].

In this study, we will evaluate wildlife image geolocation privacy by applying GeoCLIP to predict the geolocation of images in an open-source wildlife camera trap dataset and assess its accuracy by comparing the results with their known locations.

References

[1] Jorge A Ahumada, Eric Fegraus, Tanya Birch, Nicole Flores, Roland Kays, Timothy G O'Brien, Jonathan Palmer, Stephanie Schuttler, Jennifer Y Zhao, Walter Jetz, and 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):1–6, 2020.

[2] Kumar Ayush, Burak Uzkent, Chenlin Meng, Kumar Tanmay, Marshall Burke, David Lobell, and Stefano Ermon. Geography-aware self-supervised learning. *ICCV 2021*, 2022.

[3] Sara Beery, Arushi Agarwal, Elijah Cole, and Vighnesh Birodkar. The iwildcam 2021 competition dataset. *FGVC8 Workshop at CVPR*, 2021.

[4] Sara Beery and Elizabeth Bondi. Can poachers find animals from public camera trap images? *CV4Animals Workshop at CVPR*, 2021.

[5] Sara Beery, Elijah Cole, and Arvi Gjoka. The iwildcam 2020 competition dataset. *Fine-Grained Visual Categorization Workshop at CVPR*, 2020.

[6] Vicente Vivanco Cepeda, Gaurav Kumar Nayak, and Mubarak Shah. Geoclip: Clip-inspired alignment between locations and images for effective worldwide geo-localization. *NeurIPS*, 2023.

[7] Grace Chu, Brian Potetz, Weijun Wang, Andrew Howard, Yang Song, Fernando Brucher, Thomas Leung, and Hartwig Adam. Geo-aware networks for fine-grained recognition. *ICCVW 2019*, 2019.

[8] Lukas Haas, Michal Skreta, Silas Alberti, and Chelsea Finn. Pigeon: Predicting image geolocations. *Preprint*, 2023.

[9] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked autoencoders are scalable vision learners. *CVPR*, 2021.

[10] Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. The inaturalist species classification and detection dataset. In *CVPR*, 2018.

[11] Pang Wei Koh, Shiori Sagawa, Henrik Marklund, Sang Michael Xie, Marvin Zhang, Akshay Balsubramani, Weihua Hu, Michihiro Yasunaga, Richard Lanas Phillips, Irena Gao, Tony Lee, Etienne David, Ian Stavness, Wei Guo, Berton A. Earnshaw, Imran S. Haque, Sara Beery, Jure Leskovec, Anshul Kundaje, Emma Pierson, Sergey Levine, Chelsea Finn, and Percy Liang. Wilds: A benchmark of in-the-wild distribution shifts. 2021.

[12] Eric Müller-Budack, Kader Pustu-Iren, and Ralph Ewerth. Geolocation estimation of photos using a hierarchical model and scene classification. In Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss, editors, *Computer Vision – ECCV 2018*, pages 575–592, Cham, 2018. Springer International Publishing.

[13] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. 2021.

[14] Chaitanya Ryali, Yuan-Ting Hu, Daniel Bolya, Chen Wei, Haoqi Fan, Po-Yao Huang, Vaibhav Aggarwal, Arkabandhu Chowdhury, Omid Poursaeed, Judy Hoffman, Jitendra Malik, Yanghao Li, and Christoph Feichtenhofer. Hiera: A hierarchical vision transformer without the bells-and-whistles. *ICML*, 2023.

[15] Peter Soroye, Brandon P. M. Edwards, Rachel T. Buxton, Jeffrey P. Ethier, Acacia Frempong-Manso, Hannah E. Keefe, Albana Berberi, Maisy Roach-Krajewski, Allison D. Binley, Jaimie G. Vincent, Hsien-Yung Lin, Steven J. Cooke, and Joseph R. Bennett. The risks and rewards of community science for threatened species monitoring. *Conservation Science and Practice*, 4(9):e12788, 2022.

[16] Michael A. Tabak, Mohammad S. Norouzzadeh, David W. Wolfson, Steven J. Sweeney, Kurt C. Vercauteren, Nathan P. Snow, Joseph M. Halseth, Paul A. Di Salvo, Jesse S. Lewis, Michael D. White, Ben Teton, James C. Beasley, Peter E. Schlichting, Raoul K. Boughton, Bethany Wight, Eric S. Newkirk, Jacob S. Ivan, Eric A. Odell, Ryan K. Brook, Paul M. Lukacs, Anna K. Moeller, Elizabeth G. Mandeville, Jeff Clune, and Ryan S. Miller. Machine learning to classify animal species in camera trap images: Applications in ecology. *Methods in Ecology and Evolution*, 10(4):585–590, 2019.

[17] Devis Tuia, Benjamin Kellenberger, Sara Beery, Blair R. Costelloe, Silvia Zuffi, Benjamin Risse, Alexander Mathis, Mackenzie W. Mathis, Frank van Langevelde, Tilo Burghardt, Roland Kays, Holger Klinck, Martin Wikelski, Iain D. Couzin, Grant van Horn, Margaret C. Crofoot, Charles V. Stewart, and Tanya Berger-Wolf. Perspectives in machine learning for wildlife conservation. *Nature Communications*, 13(1), February 2022.

[18] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. 2017.

[19] Nam Vo, Nathan Jacobs, and James Hays. Revisiting im2gps in the deep learning era. *CVPR*, 2017.

[20] Adam Welz. Unnatural surveillance: How online data is putting species at risk. https://e360.yale.edu/features/unnatural-surveillance-how-online-data-is-putting-species-at-risk, 2023. Accessed: 2024-05-11.

[21] Tobias Weyand, Ilya Kostrikov, and James Philbin. *PlaNet - Photo Geolocation with Convolutional Neural Networks*, page 37–55. Springer International Publishing, 2016.