Darmstadt University of Applied Sciences – Data Science

Individual identification of patterned solitary species based on unlabeled video data

Master Thesis of Vanessa Süßle

Abstract

- Automatic processing and analysis of videos from camera traps without the need of manual interaction.
- Individual identification of animals based on visual characteristics.
- The program is tested on the leopard dataset from the PanAfrican Programme and achieves a rate of more than 85% correct matches between previously unknown individuals.
- Components of analytical computer vision as well as Deep Learning methods are applied.

Motivation

- Many species categorized as vulnerable on the IUCN Red List.
- Efficient and reliable monitoring of wild animals in their natural habitats is essential for wildlife conservatory, but a complex and time intense task for ecologists.
- Several steps from the filtering of false triggered footage to the identification and reidentification of individuals are required.

Research Question

The target of this thesis is to address the problem of the time consuming task to identify individual animals by developing a program that automatically processes the data taken from camera traps with computer vision methods, eliminating the manual tagging conducted by ecologists for the recognition of individuals based on unique fur patterns.

The system is implemented in Python and demonstrated on the leopard dataset from the PanAfrican Programme collected in Central and West Africa [1].

Methodology

For a proper functionality of the system the following assumptions have to apply to the investigated species and data at hand:

1. The inspected animal species has a solitary behavior.
2. Within one triggered video the same individual is seen throughout the frame sequences.
3. The inspected species is uniquely identifiable by coat or body marks.

Based on the solitary assumption all frames of the same video are initially automatically labeled with the same name ID. Due to the movement of the animal in the video footage different viewpoints might be captured leading to better matching chances. The patterns on the opposite flanks of the animal are independent.

Architecture

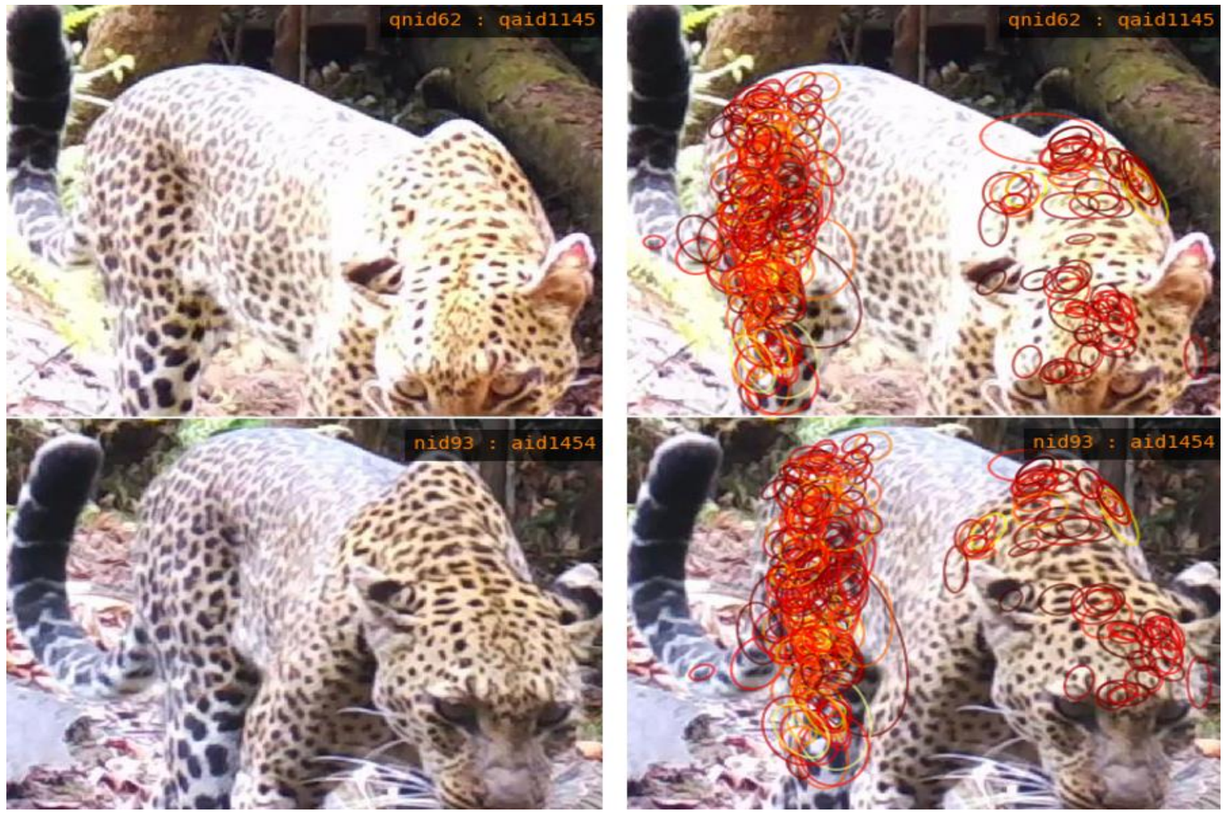
The raw video data from the camera traps passes through the following steps:

- **Detector:** a trained convolutional neural network locates the animal in each frame. Only the non-empty frames with a detection are further used [2].
- **Feature detection and matching:** the features of the fur pattern are detected and described with the SIFT algorithm [3].
- **Clustering:** based on the matching score from the SIFT algorithm frames that have a high probability of showing the same individual are merged into a cluster.

Results & Conclusion

The system is evaluated on the leopard dataset with 210 videos including 3183 non-empty frames. The videos were labeled by experts with the individual name ID of the leopards.

Different thresholds for the cluster edges are tested. A lower threshold leads to less mismatches, but more missed out matches. Decisions on the threshold must be adapted to the individual application area.



The images at the top and bottom show the same individual in frames from different videos. In the right images the prominent features are marked by red circles. The leopard was matched based on features mostly on the head and the right hind leg [1].

Reappearing issues causing mismatches are the matching of letters and digits in the automatic subtitle of the camera, which was subsequently fixed by cropping the images, or the matching of objects in the background captured due to the fixed camera scenery. Overall the system reached a correctness of 85.51% for the made matches after the above mentioned fixes.

Future Work

For future work techniques to cut out the background more detailed than rectangular could lead to a better result, but bear challenges due to noisy backgrounds in natural habitats.

Taking advantage of the video data optical flows could be used to assist that process.

For future data collections two opposite cameras could be placed to automatically collect footage from both sides of the animal. Furthermore the GPS data can be used as a criterion for the clustering.

References

[1] “Pan African Programme: The Cultured Chimpanzee,” 2016. http://panafrican.eva.mpg.de/english/where_we_work.php.

[2] S. Beery, D. Morris, and S. Yang, “Efficient Pipeline for Camera Trap Image Review,” 2019, [Online]. Available: http://arxiv.org/abs/1907.06772.

[3] J. P. Crall, “IBEIS - Image Based Ecological Information System - Repository.” [Online]. Available: https://github.com/Erotemic/ibeis