Improving an animal pose-estimation model for wildlife camera-trap imagery

Annotated pose of a male red deer in the Swiss National Park, from our 2023 data collection survey

Context:

Monitoring wildlife is important to understand their population dynamics, especially in a context of rapid climate change in the alpine regions. In recent years, camera traps have increasingly been used to monitor wildlife behavior and their interactions with the environment. While other methods such as telemetry remain widely used to measure behaviors related to energy acquisition and expenditure, camera-traps provide a new view on more fine-grained behaviors and interactions between individuals. This is supported by the advances in hardware capabilities and in automatic data processing. In 2023 and 2024, we collected around 15 hours of animal recordings in the Swiss National Park. The raw videos were used to train deep learning models for species and behavior recognition. To extend these methods towards the analysis of wildlife-environment interactions, we also aim to learn pose estimation models. These models take raw frames as input and predict animal keypoints (e.g., tail base, right ear, left front hoof). While initial results are promising, the performance of the learned models remain limited to infer animal pose on the complete dataset.

Project:

The objectives of this project are to train a model for wild mammals pose estimation from camera trap videos, to quantify its performance and to identify potential improvements. The project will be done using DeepLabCut and follows existing work on the topic (not published).

Specifically, the student will extend the training dataset by annotating more camera trap videos. The choice of videos to annotate will be based on the performance of the previously trained model. Then the student will work on fine-tuning a pose-estimation model to predict keypoints of interests that are shared across species. While the current model works well for

clear images, its performance is limited when the animal is partially occluded, too close to the camera or under low image quality.
A potential extension of this project is to then integrate predictions over time to obtain consistent predictions over a complete video.

Requirements:

- Interest in wildlife monitoring and machine learning
- Proficiency in Python and general image processing libraries
- Previous experience with DeepLabCut software appreciated but not required
- Passed an intro course to machine learning / deep learning successfully

Literature:

- Valentin Gabeff et al. 2025
- Shaokai Ye et al. 2023

Contact:

- valentin.gabeff(at)epfl.ch
- devis.tuia(at)epfl.ch & alexander.mathis(at)epfl.ch