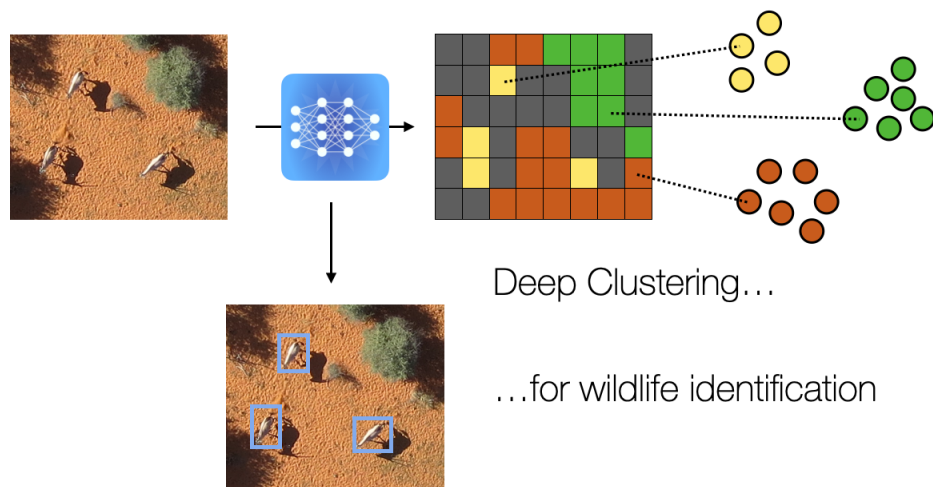


Master thesis subject

Detecting wildlife with powerful deep representations and few training data

Context

Deep Learning models like Convolutional Neural Networks (CNN) are ubiquitous across all major computer vision tasks, such as self-driving cars or object detection in images. In parts, this is due to CNNs learning a classifier *and* expressive features together on large amounts of training data. This generally fails for applications where only little annotated data is available. In this project, we intend to overcome this issue by employing techniques that pre-train CNNs for free.



Conventionally, CNNs are composed of a part that performs feature extraction (i.e., encoding each data point into a high-dimensional vector that summarises it appropriately), as well as a classifier that provides a prediction with reasoning based on said feature vectors. Due to the millions of parameters, such models require equally many training examples, which may not always be available. Although transfer learning (i.e., pre-training a CNN on another task where data availability is high) is often employed to address this limitation, it typically does not solve it.

A recent attempt to make CNNs ready for the final task is Self-Supervised Learning (SSL), where the model is pre-trained on the target data, but with labels that come for free. An example for this is the prediction of colour images from a greyscale input. The hope is that such models yield feature vectors that effectively summarise the image content appropriately, and thus only need little actual training data for the final task (e.g., object detection).

In this project, the student will investigate the feasibility of using SSL for wildlife detection in aerial images with CNNs. This task is particularly arduous in that labels are difficult and expensive to obtain. Instead, the student will first experiment with SSL techniques, such as deep clustering (cf. references) to pre-train the model. In a second step, the pre-trained model will be subjected to the task of wildlife detection, based on only few amounts of training data. If successful, the model will yield similar performances to counterparts that were trained on orders of magnitude more labels.

Objectives

- Familiarise and successfully set up a CNN-based object detector for animal localisation in UAV images
- Investigate the possibilities of using SSL tasks like deep clustering for model pre-training in object detection
- Provide an animal detector that, if pre-trained with SSL, yields detection performances similar to models trained with more labels

Requirements and practical info

- Background in machine/deep learning is welcome.
- Programming skills in Python.
- The thesis will be supervised from the Sion campus.
- Access to parallel computing resources is provided.

Literature

- Kellenberger, Benjamin, Diego Marcos, and Devis Tuia. "Detecting mammals in UAV images: Best practices to address a substantially imbalanced dataset with deep learning." *Remote Sensing of Environment* 216 (2018): 139-153.
- Caron, Mathilde, et al. "Deep clustering for unsupervised learning of visual features." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
- Doersch, Carl, and Andrew Zisserman. "Multi-task self-supervised visual learning." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.

Contact

Prof. Devis Tuia, devis.tuia@epfl.ch