

(3D) Object Discovery in Videos

Karteek Alahari: karteek.alahari@inria.fr

Cordelia Schmid: cordelia.schmid@inria.fr

<http://thoth.inriaples.fr>

Location of the internship: The internship will be in the Thoth team at Inria Grenoble, and will be co-supervised by Karteek Alahari (Inria researcher) and Cordelia Schmid (Inria Research Director). The team is specialized in computer vision, in particular visual recognition.

Topic: The traditional paradigm for learning representative models of data relies on manual annotation. For example, to learn a model for semantic segmentation of images, i.e., assigning object labels to pixels, the most successful methods require annotated training data [1], where every pixel in the entire training set is manually assigned a label. Such data is expensive to acquire reliably, and impossible to adapt to the context of large data collections. The challenge is to depart from this fully-supervised scenario. A promising research direction in this context is to automatically “discover” objects in an image collection. Object discovery involves localizing objects in a given image collection containing multiple dominant objects [2]. This problem: (i) is unsupervised, i.e., there are no image annotations to guide the discovery process, (ii) makes no assumption of a single dominant class or even a known number of object classes, and (iii) also needs to deal with noisy images in the collection, which do not contain any of the target objects.

The internship will focus on extending techniques such as [2–4] using more precise object representations (e.g., in object segments instead of bounding boxes), discover objects in video, and also to build a 3D interpretation of the scene in an unsupervised paradigm.

References

- [1] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *Trans. PAMI*, 40(4):834–848, 2018.
- [2] M. Cho, S. Kwak, C. Schmid, and J. Ponce. Unsupervised object discovery and localization in the wild: Part-based matching with bottom-up region proposals. In *CVPR*, 2015.
- [3] A. Karpathy, S. Miller, and L. Fei-Fei. Object discovery in 3d scenes via shape analysis. In *ICRA*, 2013.
- [4] S. Kwak, M. Cho, I. Laptev, J. Ponce, and C. Schmid. Unsupervised Object Discovery and Tracking in Video Collections. In *ICCV*, 2015.