Master's Thesis on Diverse Unsupervised Segmentation Proposals for Expert-driven Pixel-wise Annotation

Advisor: Simon Reiß (simon.reiss@kit.edu)

1   Motivation and relevance

Deployment of semantic segmentation systems is often slowed down by the time-consuming task of data annotation. Especially when expensive experts need to provide pixel-wise annotations as, e.g., in medical image segmentation tasks. Finding strategies to get more out of a few labeled examples can help immensely when training segmentation models and bringing them to real applications.

2   Coarse topic and starting point

Annotating images pixel-wise, as needed for training semantic segmentation algorithms is very costly, *i.e.* annotating a natural imagery dataset with 20 classes and circa 1.5 instances per class took around 240 seconds per image to annotate [2]. In domains where busy experts have to provide these annotations, getting hold of an expert to annotate images on a pixel-by-pixel basis is rather hard, as experts are occupied with important work (*e.g.* medical doctors are quite literally pre-occupied with saving lives).
To make the annotation process more streamlined, offering experts diverse proposal regions that can be selected with one click is a pathway to do this. Such regions have to be pre-computed without initial annotations. This boils down to the task of unsupervised image segmentation [3].
What would be interesting to investigate is a new method that not only does unsupervised image segmentation but has the property to produce multiple, diverse results [1, 4]. Thus, I am interested to answer this question with you: Can we train a model that produces diverse, sensible proposed regions and can they be used to train segmentation models?

3   Requirements

English language skills and good programming skills in Python are required. Some basic knowledge about the machine learning framework `https://pytorch.`

`org/` is recommended. You should have a solid understanding of deep learning models, e.g., by having attended lectures like {Deep Learning for Computer Vision, Computer Vision for Human Computer Interaction, Neural Networks, Machine Learning}. Take a quick glance at the referenced papers, if the topic sounds exciting, interesting and you feel like it's a big enough challenge for you, feel free to contact me.

References

[1] Tomer Amit, Eliya Nachmani, Tal Shaharbany, and Lior Wolf. Segdiff: Image segmentation with diffusion probabilistic models. *arXiv preprint arXiv:2112.00390*, 2021.

[2] Amy Bearman, Olga Russakovsky, Vittorio Ferrari, and Li Fei-Fei. What's the point: Semantic segmentation with point supervision. In *European conference on computer vision*, pages 549–565. Springer, 2016.

[3] Xu Ji, João F Henriques, and Andrea Vedaldi. Invariant information clustering for unsupervised image classification and segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9865–9874, 2019.

[4] Simon AA Kohl, Bernardino Romera-Paredes, Klaus H Maier-Hein, Danilo Jimenez Rezende, SM Eslami, Pushmeet Kohli, Andrew Zisserman, and Olaf Ronneberger. A hierarchical probabilistic u-net for modeling multi-scale ambiguities. *arXiv preprint arXiv:1905.13077*, 2019.