



Contribution ID: 28

Type: **Talk**

# Towards Universal Instance Segmentation Models in Biomedical Imaging

*Tuesday 10 September 2024 15:45 (15 minutes)*

Precise instance segmentation is a critical part of many fields of research in biomedical imaging. One key challenge is applying models to new data domains, typically involving pre-training a model on a larger corpus of data and fine-tuning it with new annotations for each specific domain. This process is labor-intensive and requires creating and maintaining multiple branched versions of the model. Working towards universal instance segmentation models in biomedical imaging, we propose to unify domain-adapted model branches into a single multi-expert model, following a foundation model paradigm. Our goal is to replace most existing fine-tuning scenarios with prompt-based user instructions, allowing the user to clearly state the task and object classes of interest. We hypothesize that such a combined approach improves generalization, as the base model can benefit from datasets that were previously only used for fine-tuning. A key challenge in the creation of such models is to resolve training conflicts and ambiguity in a pragmatic fashion when combining different segmentation tasks, datasets, and data domains. Such conflicts can occur if datasets focus on different classes in the same domain. For example, some datasets annotate all cells in microscopy images, while others focus on cells of a specific cell type. A naïve combination of such sets would create an ill-posed learning problem for most models, requiring them to infer their task from their input, which is undesirable in a universal setting. Models like SAM and MedSAM highlight the potential of prompting, but often require external detectors and fine-tuning. Here, we propose to leverage prompt-based task descriptions as a tool to manipulate general model behavior, such that user instructions yield domain expert models. We test our approach by training a Contour Proposal Network (CPN) on a multi-modal data collection, including the TissueNet dataset. Prompts, such as “cell segmentation” or simply “nuclei”, modulate underlying features, allowing the CPN to segment the respective object classes in TissueNet with a mean F1 score of 0.90 (0.88 for cells, 0.92 for nuclei), compared to 0.84 (0.81, 0.87) without prompting. Overall, the proposed approach introduces an interactive linguistic component that allows the conflict-free composition of various segmentation datasets, thus allowing to unify previously separated segmentation tasks. With that, we consider it an important step towards universal models.

**Primary author:** Mr UPSCHULTE, Eric (University Hospital Düsseldorf, Cécile & Oscar Vogt Institute for Brain Research)

**Co-authors:** Prof. HARMELING, Stefan (Technical University Dortmund, Department of Computer Science); DICKSCHEID, Timo (Forschungszentrum Jülich, Institute of Neurosciences and Medicine (INM-1)); Prof. AMUNTS, Katrin (Forschungszentrum Jülich, Institute of Neurosciences and Medicine (INM-1))

**Presenter:** Mr UPSCHULTE, Eric (University Hospital Düsseldorf, Cécile & Oscar Vogt Institute for Brain Research)

**Session Classification:** Contributed Talks - Segmentation and AI (co-Chairs: Claire Walsh, Jussi Tohka)