

Evaluating Foundation Model Representations for Cell Segmentation and Classification in Microscopy Images

Internship in the Biological Image Analysis Unit, Institut Pasteur, Paris, France

Introduction

Foundation models and large-scale pretraining have become central to visual representation learning, including in computational pathology. Models trained on natural images or large microscopy and histopathology datasets are increasingly reused as fixed feature extractors for downstream tasks such as diagnosis, disease sub-typing, biomarker quantification, treatment response estimation, and survival prediction [1, 2, 3]. While this paradigm offers strong performance and reduced training costs, recent work suggests that substantial performance gaps remain between general-purpose representations and domain-specific embeddings when applied to fine-grained cellular tasks [4]. In particular, cell segmentation and cell classification place strong demands on spatial precision and domain-specific visual cues [5]. This project is motivated by the need to better understand these gaps and to systematically analyse how patch-level embeddings transfer across datasets and tissue types for cell segmentation and classification in microscopy images, particularly in histopathology.

Objectives and Methodology

Students will first conduct a literature review on foundation models for microscopy images, with particular emphasis on histopathology. The review should cover pretraining strategies (supervised vs. self-supervised), dataset scale and domain specificity, and architectural design choices. Building on these insights, students will implement an experimental study in Python to systematically evaluate different foundation models as encoders within a UNETR architecture [6] for cell segmentation and cell classification. Models will be assessed both as frozen feature extractors and, where appropriate, under end-to-end fine-tuning, enabling direct comparison of these regimes. Experiments will analyse how factors such as pretraining domain, pretraining strategy, and representation design choices influence performance, with the goal of identifying limitations of current models and proposing directions for improving cellular-level modelling.

Contacts

Valentina Vadori, Biological Image Analysis Unit, Institut Pasteur, vvadori@pasteur.fr

References

- [1] Richard J Chen, Tong Ding, Ming Y Lu, Drew FK Williamson, Guillaume Jaume, Andrew H Song, Bowen Chen, Andrew Zhang, Daniel Shao, Muhammad Shaban, et al. Towards a general-purpose foundation model for computational pathology. *Nature medicine*, 30(3):850–862, 2024.
- [2] Ming Y Lu, Bowen Chen, Drew FK Williamson, Richard J Chen, Ivy Liang, Tong Ding, Guillaume Jaume, Igor Odintsov, Long Phi Le, Georg Gerber, et al. A visual-language foundation model for

computational pathology. *Nature medicine*, 30(3):863–874, 2024.

- [3] Eugene Vorontsov, Aican Bozkurt, Adam Casson, George Shaikovski, Michal Zelechowski, Kristen Severson, Eric Zimmermann, James Hall, Neil Tenenholtz, Nicolo Fusi, et al. A foundation model for clinical-grade computational pathology and rare cancers detection. *Nature medicine*, 30(10):2924–2935, 2024.
- [4] Valentina Vadori, Antonella Peruffo, Jean-Marie Graic, Livio Finos, and Enrico Grisan. Mind the gap: Evaluating patch embeddings from general-purpose and histopathology foundation models for cell segmentation and classification. In *Proceedings of the IEEE Engineering in Medicine and Biology Conference (EMBC)*, 2025. to appear.
- [5] Markus Marks, Uriah Israel, Rohit Dilip, Qilin Li, Changhua Yu, Emily Laubscher, Ahamed Iqbal, Elora Pradhan, Ada Ates, Martin Abt, et al. Cellsam: a foundation model for cell segmentation. *Nature Methods*, pages 1–9, 2025.
- [6] Ali Hatamizadeh, Yucheng Tang, Vishwesh Nath, Dong Yang, Andriy Myronenko, Bennett Landman, Holger R Roth, and Daguang Xu. Unetr: Transformers for 3d medical image segmentation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 574–584, 2022.