

Master/Bachelor Thesis Proposal: Foundation Model Empowered Computational Pathology Images Analysis

schuefflerlab.org

1 Background and methodology

Artificial intelligence (AI) can potentially transform cancer diagnosis and treatment by analyzing pathology images for precision medicine and decision support systems. Pathology’s clinical practice usually encompasses tasks like tumor classification, segmentation, subtyping, grading, staging, and whole slide matching. Although AI demonstrates promise in many pathological tasks, it still faces challenges in generalization and addressing rare diseases due to limited training data availability. [CDL+24, VBC+23]

Here, a foundation model may contribute to this challenge. A foundation Model refers to a general-purpose model pre-trained on typically unlabeled datasets, subsequently fine-tuned to apply to diverse downstream tasks [DFW+24]. To compare the proposed foundation model with previous state-of-the-art methods, we want to evaluate the performance of patch/slide level classification and segmentation tasks.

2 Tasks

- Paper reading and literature review
- Evaluate proprietary and public foundation models on public datasets
- Finetune the models to obtain the best performance
- Publicly available models: UNI [CDL+24], cTransPath [WYZ+22], Virchow [VBC+23], RudolfV [DFW+24], MSSM [CKF+23]
- Datasets for classification and detection: MHIST [WSR+21], PCAM [VLW+18], NCT-CRC [KHM18], CAMELYON16/17 [BGM+18]
- Datasets for segmentation: SegPath [KOS+23], PanNuke [GAKB+19]

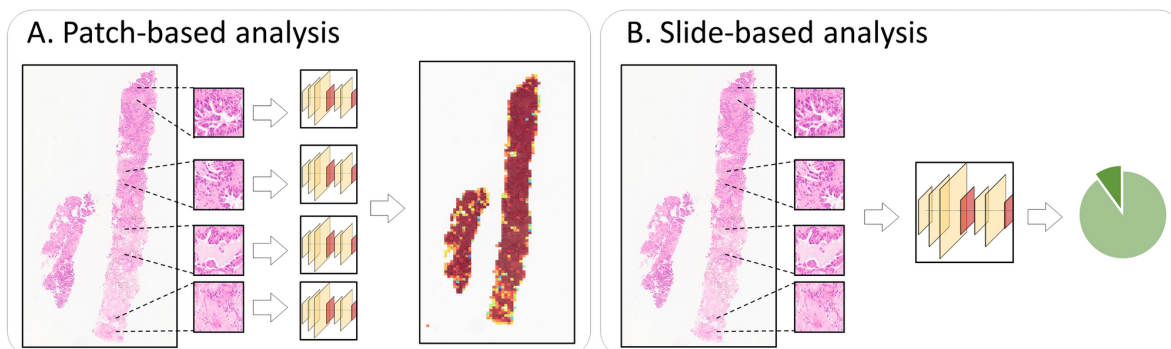


Figure 1: An example of patch-level and slide-level analysis of an HE whole slide image [KPL+22]

3 Requirements

Basic knowledge in at least one of the following areas:

- Pytorch and deep learning knowledge
- Medical image analysis

4 Supervision and Contact

Prof. Peter Schüffler and Jingsong Liu will be the supervisors. If you are interested, please briefly describe your prior experiences and attach your grade transcript, feel free to contact jingsong.liu@tum.de or peter.schueffler@tum.de.

References

- [BGM⁺18] Peter Bandi, Oscar Geessink, Quirine Manson, Marcory Van Dijk, Maschenka Balkenhol, Meyke Hermsen, Babak Ehteshami Bejnordi, Byungjae Lee, Kyunghyun Paeng, Aoxiao Zhong, et al. From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. *IEEE transactions on medical imaging*, 38(2):550–560, 2018.
- [CDL⁺24] 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.
- [CKF⁺23] Gabriele Campanella, Ricky Kwan, Eugene Fluder, Jennifer Zeng, Aryeh Stock, Brandon Veremis, Alexandros D. Polydorides, Cyrus Hedvat, Adam Schoenfeld, Chad Vanderbilt, Patricia Kovatch, Carlos Cordon-Cardo, and Thomas J. Fuchs. Computational Pathology at Health System Scale – Self-Supervised Foundation Models from Three Billion Images. October 2023. arXiv:2310.07033 [cs, eess].
- [DFW⁺24] Jonas Dippel, Barbara Feulner, Tobias Winterhoff, Simon Schallenberg, Gabriel Dernbach, Andreas Kunft, Stephan Tietz, Philipp Jurmeister, David Horst, Lukas Ruff, et al. Rudolfv: A foundation model by pathologists for pathologists. *arXiv preprint arXiv:2401.04079*, 2024.
- [GAKB⁺19] Jevgenij Gamper, Navid Alemi Koohbanani, Ksenija Benet, Ali Khuram, and Nasir Rajpoot. Pannuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. In *Digital Pathology: 15th European Congress, ECDP 2019, Warwick, UK, April 10–13, 2019, Proceedings 15*, pages 11–19. Springer, 2019.
- [KHM18] Jakob Nikolas Kather, Niels Halama, and Alexander Marx. 100,000 histological images of human colorectal cancer and healthy tissue. *Zenodo10*, 5281:6, 2018.
- [KOS⁺23] Daisuke Komura, Takumi Onoyama, Koki Shinbo, Hiroto Odaka, Minako Hayakawa, Mieko Ochi, Ranny Rahaningrum Herdiantoputri, Haruya Endo, Hiroto Katoh, Tohru Ikeda, et al. Restaining-based annotation for cancer histology segmentation to overcome annotation-related limitations among pathologists. *Patterns*, 4(2), 2023.
- [KPL⁺22] Sai Kosaraju, Jeongyeon Park, Hyun Lee, Jung Wook Yang, and Mingon Kang. Deep learning-based framework for slide-based histopathological image analysis. *Scientific Reports*, 12(1):19075, 2022.
- [VBC⁺23] Eugene Vorontsov, Alican Bozkurt, Adam Casson, George Shaikovski, Michal Zelechowski, Siqi Liu, Philippe Mathieu, Alexander van Eck, Donghun Lee, Julian Viret, et al. Virchow: A million-slide digital pathology foundation model. *arXiv preprint arXiv:2309.07778*, 2023.

- [VLW⁺18] Bastiaan S Veeling, Jasper Linmans, Jim Winkens, Taco Cohen, and Max Welling. Rotation equivariant CNNs for digital pathology. June 2018.
- [WSR⁺21] Jerry Wei, Arief Suriawinata, Bing Ren, Xiaoying Liu, Mikhail Lisovsky, Louis Vaickus, Charles Brown, Michael Baker, Naofumi Tomita, Lorenzo Torresani, et al. A petri dish for histopathology image analysis. In *International Conference on Artificial Intelligence in Medicine*, pages 11–24. Springer, 2021.
- [WYZ⁺22] Xiyue Wang, Sen Yang, Jun Zhang, Minghui Wang, Jing Zhang, Wei Yang, Junzhou Huang, and Xiao Han. Transformer-based unsupervised contrastive learning for histopathological image classification. *Medical Image Analysis*, 81:102559, October 2022.