Type: PhD Thesis / TV-L E13 (3y)

Title: Tumor Heterogeneity Quantification

Supervisor: Prof. Dr. Peter Schüffler (TUM), Computational Pathology

Collaborators: Saturn3 Consortium[1]

Keywords: Computational Pathology, Histology, Tumor heterogeneity, Quantification, Artificial Intelligence, Deep Learning

Objective: To quantify and analyze tumor heterogeneity in three hard-to-treat cancers based on pathology data. To correlate pathology heterogeneity with OMICs heterogeneity and molecular data.

Problem: Intratumor heterogeneity is a common problem for the treatment of cancer. It is believed to be an important factor for treatment success and recurrence / prognosis for the patient*(1–3)*. However, little is known about tumor heterogeneity, and causes, mechanisms and treatments remain unclear.

The 5y-project Saturn3 aims to comprehensively assess tumor heterogeneity in the three cancer types of the breast, the pancreas and the colon. Partners of 17 institutions develop new and apply a comprehensive list of established molecular, single cell, -omics, imaging and other methods on the studied data sets in order to gain new insights for the causes and mechanisms of heterogeneity.

Goal: In this project and as part of Saturn3, we aim to investigate the potential of artificial intelligence (AI) to detect, quantify and correlate intratumor heterogeneity in digital histology. The data and insights gained in this project shall be set in context with the other partners' approaches. Subgoals of this project are:

1. Automatically detect and segment tumor tissue in included longitudinal Saturn3 data sets.
2. Assess intratumor heterogeneity in the tumor regions.
    a. "known" heterogeneity (e.g. biomarker, IHC, molecular markers)
    b. "unknown" heterogeneity (e.g. patterns, clusters)
3. Quantify heterogeneity
4. Correlate quantitative heterogeneity with clinical data (recurrence, treatment, survival, …)

This project is not very well defined, as little is known about heterogeneity and its definitions. Therefore, the student is expected to work in the field of tumor heterogeneity, but it is possible to deviate from the above list of goals. However, it is expected to publish own progress and scientific results on national and international conferences and/or journals, and to participate actively in the Saturn3 project.

Data: Longitudinal datasets from the Saturn3 consortium, public data sets for heterogeneity (e.g. TCGA), TUM inhouse clinical data sets for Breast cancer, Colon cancer and Pancreatic cancer. Own data can additionally be generated/ collected as needed (e.g. required annotations, or more clinical data).

Computing: The student will use the LRZ high performance computing cluster for AI (GPU-based), and/or the group's computing infrastructure and/or a workstation.

[1] https://www.saturn3.org/

Teaching: For this PhD thesis, it is expected to mentor Bachelor or Master students and occasionally help in teaching classes (e.g. mentor seminar series). Active participation and shaping the group culture is expected, too.

Requirements

Knowledge in or interest in learning of: Python/R programming, deep learning, machine learning, Linux cluster usage, GPU programming, image processing, pathology, debugging, visualizations, analytical thinking, statistical analysis, presenting.

References

1. M. I. Jaber, B. Song, C. Taylor, C. J. Vaske, S. C. Benz, S. Rabizadeh, P. Soon-Shiong, C. W. Szeto, A deep learning image-based intrinsic molecular subtype classifier of breast tumors reveals tumor heterogeneity that may affect survival. *Breast Cancer Res.* **22**, 12 (2020).

2. G. Prokop, M. Örtl, M. Fotteler, P. Schüffler, J. Schobel, W. Swoboda, J. Schlegel, F. Liesche-Starnecker, Quantifying Heterogeneity in Tumors: Proposing a New Method Utilizing Convolutional Neuronal Networks. *Stud Health Technol Inform* **289**, 397–400 (2022).

3. Q. Zhong, J. H. Rüschoff, T. Guo, M. Gabrani, P. J. Schüffler, M. Rechsteiner, Y. Liu, T. J. Fuchs, N. J. Rupp, C. Fankhauser, J. M. Buhmann, S. Perner, C. Poyet, M. Blattner, D. Soldini, H. Moch, M. A. Rubin, A. Noske, J. Rüschoff, M. C. Haffner, W. Jochum, P. J. Wild, Image-Based Computational Quantification and Visualization of Genetic Alterations and Tumour Heterogeneity. *Scientific Reports* **6**, 24146 (2016).

Data: Longitudinal datasets from the Saturn3 consortium, public data sets for heterogeneity (e.g. TCGA), TUM inhouse clinical data sets for Breast cancer, Colon cancer and Pancreatic cancer. Own data can additionally be generated/ collected as needed (e.g. required annotations, or more clinical data).

Prof. Dr. Peter Schüffler · Trogerstraße 18 · 81675 Munich, Germany · https://schuefflerlab.org
TUM School of Computation (CIT) · TUM School of Medicine and Health · Munich Data Science Institute (MDSI)