Training a customized segmentation model for improved detection of Liver-Cells in high-dimensional spatial profiling data generated by imaging mass cytometry

Pius Martin[1], Laurenz Krimmel[1], Henrike Salié[1], Felix Röttele[1], Peter Hasselblatt[1] and Bertram Bengsch[1]

[1]Clinic for Internal Medicine II, Freiburg University Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany

1 Background

Segmentating cell masks from high-dimensional imaging data is required for downstream single-cell analysis. Current strategies include supervised machine learning pipelines that are however time-consuming and resource-intensive. Alternatively, deep learning models for automation of cell segmentation are being developed, such as Mesmer[1] used in DeepCell, and Cellpose[2]. However, the training data used for these models does not include a significant amount of liver cells. We thus set out to understand if the deep learning based models can reliably identify liver parenchymal and immune cell types in imaging mass cytometry data. We further set out to compare the performance to supervised machine learning-based segmentation and train a custom model using Cellpose for improved liver tissue cell segmentation.

2 Methods

We used 3 datasets from imaging mass cytometry experiments analyzing liver tissue from diseased (infection, cancer, autoimmune) tissue states (n=282) and the supervised machine learning – based segmentation masks. We performed a 80/20 Training-Test-Split and trained both a model based on the TN2-Model provided by Cellpose and a new model that also included the segmentation data. We then compared the performance of our new model against the Test split. We also compared Models trained on only one liver-segmented dataset against the ground-truth on others to exclude over-fitting

3 Results

We found that our model based on TN2 and additional liver-specific training masks outperforms the standard TN2 and Mesmer Model using standard segmentation quality metrics. We also investigated whether the Single-Cell-Data quality downstream is improved (e.g. less overlap between T- and B-Cells resulting in cleaner clusters).

4 Conclusions

Training a custom model for tissue types which were not included in the datasets used for training existing networks improves segmentation quality and allows improved automation of tissue-specific segmentation.

References

[1] Noah F. Greenwald et al. “Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning”. eng. In: *Nature Biotechnology* 40.4 (Apr. 2022), pp. 555–565. issn: 1546-1696. doi: 10.1038/s41587-021-01094-0.

[2] Marius Pachitariu and Carsen Stringer. “Cellpose 2.0: how to train your own model”. en. In: *Nature Methods* 19.12 (Dec. 2022). Number: 12 Publisher: Nature Publishing Group, pp. 1634–1641. issn: 1548-7105. doi: 10.1038/s41592-022-01663-4. url: https://www.nature.com/articles/s41592-022-01663-4 (visited on 06/28/2023).