A Masked Image Modelling Approach to Multiplex Tissue Imaging Panel Reduction

Zachary Sims, Young Hwan Chang

Department of Biomedical Engineering and Computational Biology Program,
Oregon Health & Science University (OHSU)

What if we could replace degraded marker channels from later CyCIF rounds with generated synthetic channels?

- A multiplex tissue imaging (MTI) panel size is limited due to technical artifacts, tissue loss, and long acquisition time.
- In our previous work [1], we demonstrated a proof-of-concept study where a deep learning model can be trained to predict the same information as a full dataset with fewer rounds of staining in CyCIF images.
- However, the computational intensiveness of selecting reduced panel sets prior to, and independent of, model training necessitates a more efficient and robust method of training a model versatile enough to be useful in practice.
- We use Masked Autoencoder (MAE) architectures [2][3] to reliably reconstruct full marker panels with fewer markers. Additionally, we exploit the transformer-based architecture and masked token reconstruction training objective to conduct panel selection in inference.
- This opens the door to creating an optimally reduced panel that captures the maximum amount of information with the fewest markers.

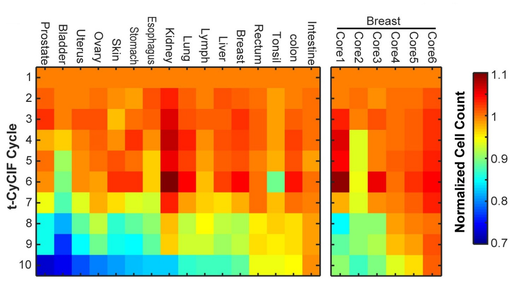


Figure 1. An increased number of CyCIF cycles results in cell loss. Heatmap shows that images from later CyCIF cycles contain significantly fewer cells than earlier cycles. This impacts the quality of markers applied in those later rounds (image source: [4])

Masked Autoencoder trained to reconstruct masked biomarker channels

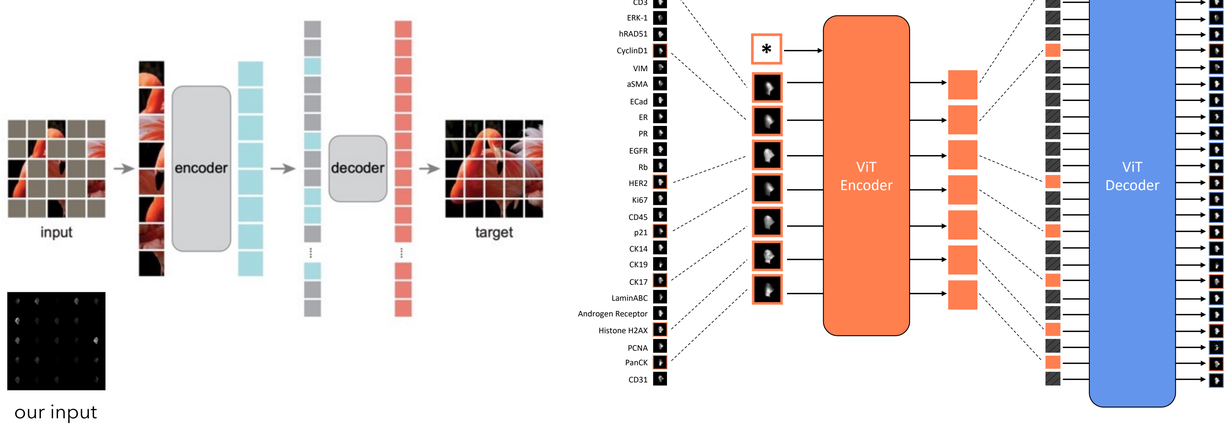


Figure 2. (Left) the original MAE architecture [2]. (Right) We adopt this method for MTI reconstruction by considering a channel-wise masking strategy with Vision Transformers (ViT) architecture.

MAE outperforms the previous approach [1] on the same panels as well as discovers new panels that further improve performance

In [1], the optimal reduced panels were found by grouping the markers that maximized the correlation to all the markers withheld from the panel. We show that MAE outperforms [1] on the same intensity correlation-based panels. Additionally, we find that our iteratively constructed panels further boost performance. In contrast to [1], MAE achieves these results without the need to retrain the model for a specific panel.

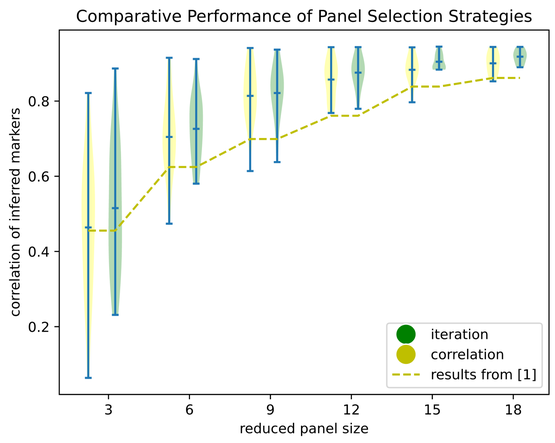


Figure 3. Spearman correlation was measured for each stain independently across the multiple reduced panels.

Highly correlated inferred staining intensities from multiple reduced panel sizes all from one model

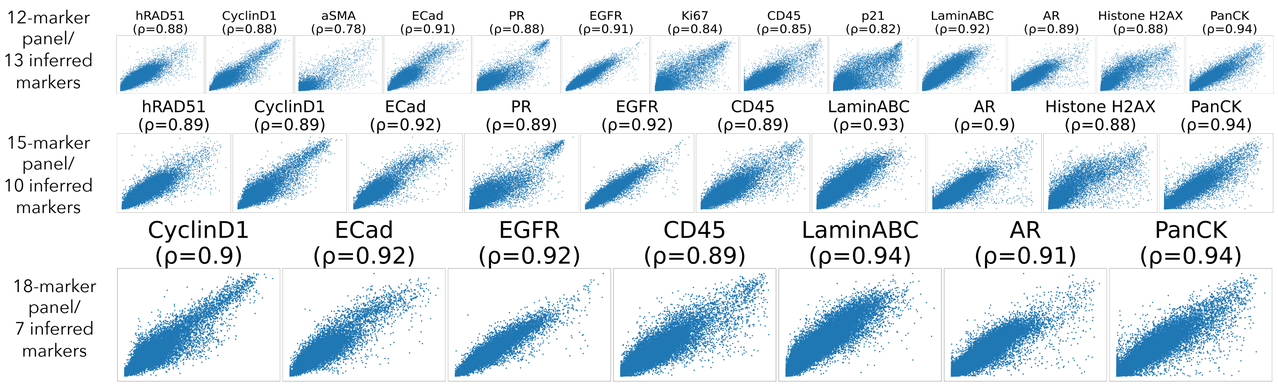


Figure 4. Predicted mean intensities from inferred marker channels versus ground truth along with their individual Spearman correlations. Each row is a different set of inferred markers generated from a different reduced panel size.

Iterative marker selection in inference reveals optimally reduced panel sets

$$Panel_1 = \{c_{DAPI}\}$$

$$Panel_n = Panel_{n-1} \cup \{\underset{c}{argmax}(MAE(X, Panel_{n-1} \cup \{c\}; \theta))\}$$

An iterative marker selection strategy constructs an optimal reduced panel of size *n* using MAE by iterating through prospective marker channels *c*, where *c* is a marker not already in the reduced panel of size *n-1*. θ refers to the trained model parameters, implying that selection is done in inference. Note that *argmax* here refers to the maximum average spearman correlation of inferred markers. We assign the DAPI channel as the first marker selected.

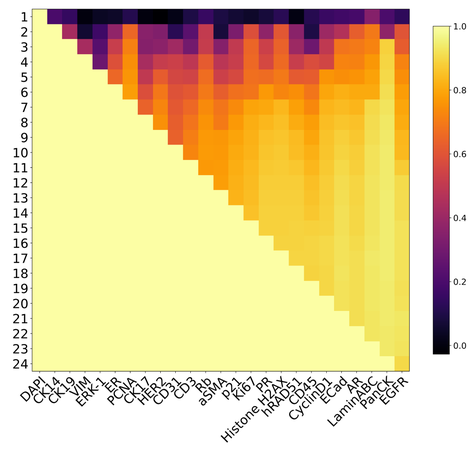
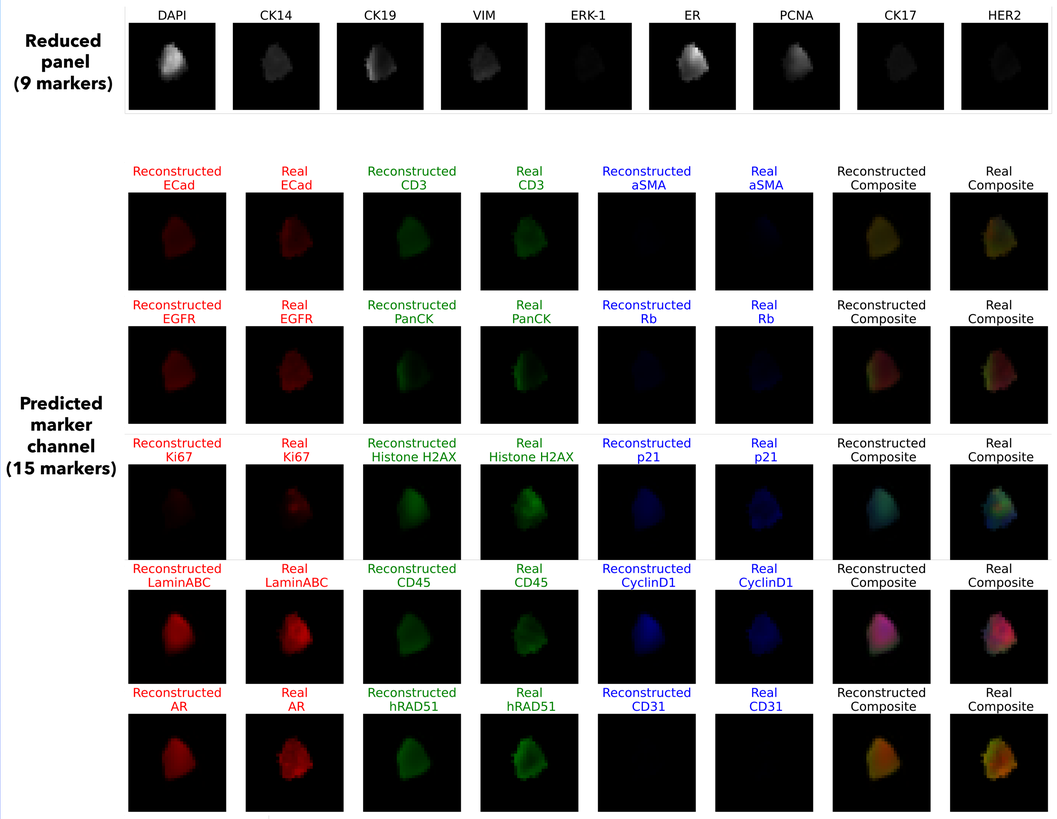


Figure 5. Optimally reduced panels and their inferred marker spearman correlations.

Inferred marker channels retain structural information



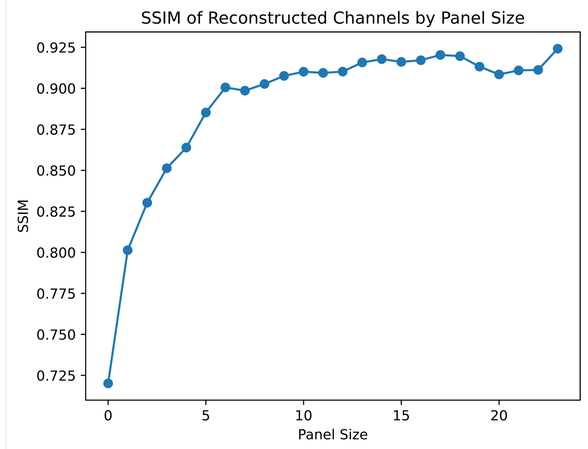


Figure 6. (top) Reconstructed markers generated from the best iteratively selected 9 marker panel. (bottom) the Structural Similarity Index Measure (SSIM) for each reduced panel size.

Conclusion: In this work we show that masked image modelling is a promising approach to discovering reduced MTI panels that can be used to infer additional biomarker information. This opens the door to the creation of larger MTI panels consisting of more informative markers by allowing less informative markers to be inferred by our model.

Dataset: The dataset used is a breast cancer (BC) tissue microarray (TMA) available on synapse from the Human Tumor Atlas Network (HTAN) TNP-TMA (https://www.synapse.org/#!Synapse:syn22041595). This BC TMA dataset is comprised of 88 cores and 6 different cancer subtypes: luminal A, luminal B, luminal B/HER+, Triple Negative, and Invasive Lobular Carcinoma.

Citations:

1. Ternes, L., Lin, J.-R., Chen, Y.-A., Gray, J. W., & Chang, Y. H. (2022). Computational multiplex panel reduction to maximize information retention in breast cancer tissue microarrays. *PLOS Computational Biology*, *18*(9). https://doi.org/10.1371/journal.pcbi.1010505
2. K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick. Masked autoencoders are scalable vision learners. arXiv preprint arXiv:2111.06377, 2021
3. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018
4. Jia-Ren Lin, Benjamin Izar, Shu Wang, Clarence Yapp, Shaolin Mei, Parin M Shah, Sandro Santagata, Peter K Sorger (2018) Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes eLife 7:e31657

Acknowledgment: This work was supported by NIH grants (U2CCA233280, R01 CA253860) and Kuni Foundation Imagination Grants.