Optimizing Channel Selection for U-Net Segmentation in MIBI-TOF of Triple-Negative Breast Cancer

Victor Gabriel Pereira de Sousa[1],**Daniel de Sousa Luz**[2],
Daniel Rodrigues de Sousa[3] , **Flávio Henrique Duarte de Araújo**[1,2,3]

[1]Programa de Pós Graduação em Engenharia Elétrica
Universidade Federal do Piauí - Teresina, Brasil

[2]Doutorado em Ciência da Computação Associação Maranhão Piauí
Universidade Federal do Piauí - Teresina, Brasil

[3]Departamento de Sistemas de Informação
Universidade Federal do Piauí - Picos, Brasil

`victorgabrielps204@gmail.com,daniel.luz@ifpi.edu.br`

`daniel.sousa@ufpi.edu.br, flavio86@ufpi.edu.br`

Abstract. *Multiplexed ion-beam imaging enables simultaneous quantification of dozens of protein markers at subcellular resolution. In this paper, we investigate how channel selection impacts U-Net–based segmentation performance on a public triple-negative breast cancer MIBI-TOF dataset with 41 images each with 48 channels. Our experiments show that compact, data-driven marker subsets achieve higher Dice than the recommended panel, suggested by the dataset paper, while using fewer channels. In summary, our experiments show that it is possible to reduce from seven channels to one without losing segmentation accuracy: our best single-channel model (dsDNA) achieves 89.36% Dice, practically matching the 89.41% Dice obtained by the full recommended panel.*

1. Introduction

Breast cancer remains one of the leading causes of cancer-related mortality among women[Bray et al. 2024], with triple-negative breast cancer (TNBC) representing one of its most aggressive forms due to the lack of estrogen receptor, progesterone receptor, and HER2 expression [Jie et al. 2025]. This absence limits the use of targeted therapies and is associated with poor clinical outcomes, making precise characterization of these tumors a continuing challenge.

Advances in high-dimensional imaging, particularly multiplexed ion beam imaging by time-of-flight (MIBI-TOF), have expanded the ability to study the tumor microenvironment in situ. MIBI-TOF uses metal-isotope–labeled antibodies to simultaneously quantify dozens of proteins at subcellular resolution, overcoming the limitations of conventional immunohistochemistry [Angelo et al. 2014]. Since its introduction in breast tumor analysis it has enabled detailed characterization of immune and tumor compartments, and studies such as that of [Keren et al. 2018] have established widely used benchmark datasets for TNBC.

Early and accurate detection of aggressive subtypes remains crucial, especially in metastatic progression, where mortality is highest. Imaging-based machine learning pipelines support this goal by enabling automated detection, stratification, and monitoring of tumors from high-dimensional images. Within these pipelines, segmentation is a fundamental step: errors in this stage propagate through downstream analyses and can compromise risk assessment or delay detection of disease spread. As a result, high-quality segmentation has become essential for computational workflows in breast cancer research [Petinrin et al. 2023].

Deep learning methods, particularly encoder–decoder architectures such as U-Net, now dominate biomedical image segmentation. The U-Net design integrates contextual information with precise spatial localization, making it particularly well suited for delineating nuclei and whole-cell boundaries in microscopy data. In breast cancer studies, U-Net-based models have been used extensively to quantify tumor morphology, cellular heterogeneity, and microenvironmental organization [Ronneberger et al. 2015]. Recent adaptations for multiplexed imaging, such as the deep-imcyto module within the TRACERx-PHLEX pipeline [Magness et al. 2024], have demonstrated that IMC-tailored segmentation workflows can outperform general-purpose approaches on dense and heterogeneous tissue images.

Despite these advances, typical MIBI-TOF experiments measure dozens of nuclear, membrane, and functional markers simultaneously, yet segmentation pipelines often rely on fixed subsets of these channels, selected by expert preference or inherited from prior work. This practice leaves open an important question: "how does segmentation quality depend on the specific choice and number of channels used as model input?". It remains unclear whether large marker panels are necessary or whether smaller, optimized subsets could yield comparable or even improved performance.

In this paper we first performs a systematic single-channel screening using U-Net to rank individual MIBI-TOF markers by Dice-based segmentation performance in TNBC. Second, it evaluates how segmentation quality scales with the number of input channels by training incremental multi-channel models built from the highest-ranked markers, characterizing performance saturation and identifying cases in which adding lower-ranked channels degrades accuracy. Finally, it compares data-driven marker subsets against a literature-recommended panel under a consistent training protocol, highlighting that compact, ranked panels can match or outperform broader conventional selections for this segmentation task.

The paper main contribution is to provide experimental evidence that explicit channel selection improves efficiency and can also improve performance, helping to design more economical marker panels for segmentation in spatial proteomics workflows.

2. Related Work

A natural starting point for this study is the seminal work of [Keren et al. 2018], which established one of the most influential MIBI-TOF benchmarks for triple-negative breast cancer (TNBC). In that study, the authors analyzed 41 TNBC cases with a 48 channels panel to characterize the tumor–immune microenvironment at subcellular resolution. Beyond its biological relevance, this dataset became an important reference for subsequent computational studies because it combines multiplexed imaging, spatial organization, and cell-level heterogeneity in a challenging breast cancer scenario.

Several later works reused this TNBC MIBI-TOF benchmark, although with objectives different from the one addressed in the present paper. For example, [Chen et al. 2020] employed the dataset to model spatially resolved tumor microenvironments using Spatial-LDA, aiming to identify recurrent microenvironmental programs and tumor organization patterns. Likewise, [Patwa et al. 2021] used the same family of data to derive tumor–immune predictors associated with clinical outcome, emphasizing prognostic interpretation rather than image segmentation itself. More recently, [Liu et al. 2023] used the TNBC MIBI-TOF dataset as one of the benchmarks for pixel-level clustering and robust phenotyping in highly multiplexed tissue imaging, focusing on reproducible annotation and phenotype discovery.

Although these studies demonstrate the broad utility of the dataset, they do not directly investigate a central issue for segmentation-oriented pipelines: how segmentation accuracy depends on the specific identity and number of input channels. In most cases, the multiplexed signal is treated as a rich substrate for downstream spatial analysis, phenotyping, or outcome prediction, while the channel configuration itself is inherited from prior biological conventions or fixed preprocessing choices.

A second line of related work is more directly connected to segmentation in multiplexed imaging. Recent pipelines such as TRACERx-PHLEX, introduced by [Magness et al. 2024], show that dedicated deep-learning workflows can substantially improve segmentation and downstream phenotyping in multiplexed imaging settings. However, even these more segmentation-centered approaches primarily emphasize pipeline robustness and analysis scalability, rather than systematically isolating the contribution of each marker or testing whether compact channel subsets can outperform broader expert-defined panels.

This gap is precisely the focus of the present study. In contrast to prior works, our paper treats channel selection itself as the experimental variable of interest. Using the public TNBC MIBI-TOF benchmark, we perform a systematic single-channel screening, construct incremental data-driven multi-channel panels, and compare them against the literature-recommended panel under a fixed U-Net protocol. In this way, the study moves beyond using multiplexed data as a given input and instead asks which channels are truly necessary for accurate segmentation.

Table 1. Comparison of related works using the TNBC MIBI-TOF benchmark.

Work	Objective	Use of TNBC MIBI-TOF	Difference from this work
[Keren et al. 2018]	Tumor–immune microenvironment characterization	Introduced the benchmark dataset and marker panel	Does not evaluate channel choice for U-Net segmentation
[Chen et al. 2020]	Spatial topic modeling of tumor microenvironments	Identifies recurrent spatial patterns	Focuses on spatial organization, not segmentation or channel ranking
[Patwa et al. 2021]	Outcome prediction from tumor–immune features	Extracts clinically relevant predictors	Focuses on prognosis, not segmentation efficiency
[Liu et al. 2023]	Pixel-level clustering and phenotyping	Uses the dataset for robust phenotype discovery	Focuses on phenotyping, not controlled U-Net channel subsets
[Magness et al. 2024]	Deep-learning segmentation in multiplexed imaging	Provides a segmentation-oriented reference pipeline	Does not isolate marker identity and number as variables
This work	Channel selection for U-Net segmentation	Compares single-channel, top-k, and recommended panels	Explicitly ranks channel informativeness under a fixed protocol

3. Methodology

The methodology follows a structured a workflow grounded in Dice metric support a clear comparison across single-channel models, incremental data-driven panels, and the recommended marker panel, enabling a direct assessment of how channel selection influences segmentation performance as show in the figure 1.

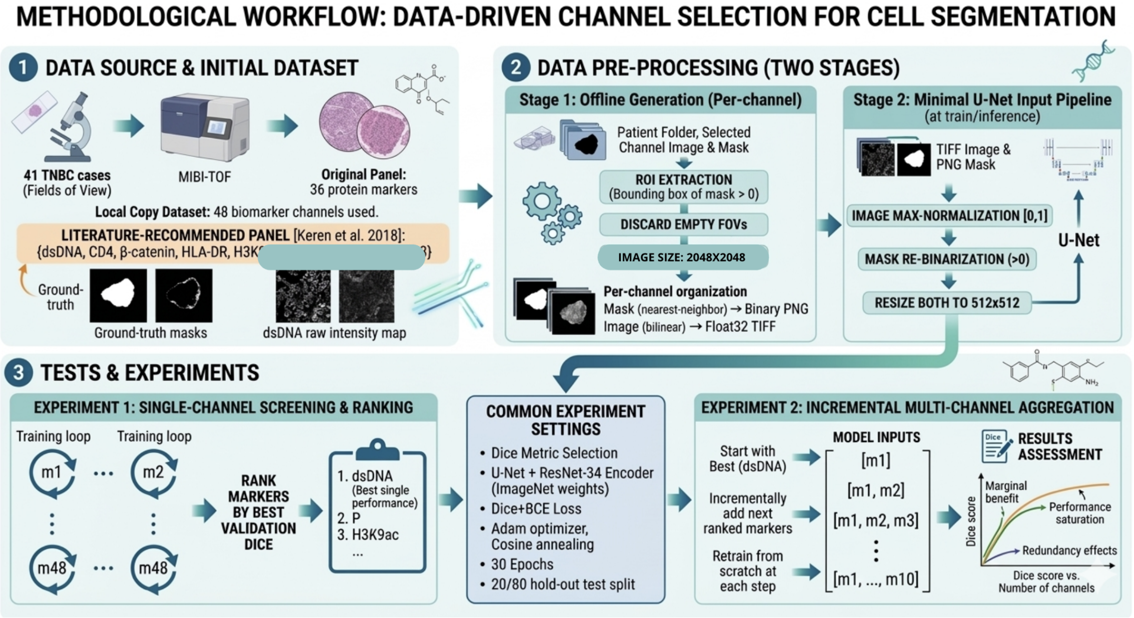

Figure 1. Workflow to investigate the best channels.

3.1. Dataset

We use the public MIBI-TOF dataset introduced by [Keren et al. 2018], which contains data from 41 TNBC cases (fields of view acquired per patient/sample).

the original study reports an antibody panel designed to measure 36 antigens protein markers in the tumor–immune microenvironment. In our local copy of the dataset (as used in our experiments), the image stack is provided with 48 channels, and we follow that channel set for all screening and multi-channel experiments.

Importantly, the dataset paper also describes a recommended marker subset commonly used to support segmentation by providing strong structural contrast for nuclei and cell boundaries. For nuclear signal, the authors build a robust nuclear image by combining dsDNA, H3K27me3, and H3K9ac, which are highly expressed in nuclei and therefore increase the signal-to-noise ratio for nuclear delineation. For membrane / whole-cell context, they further rely on membrane-enriched markers to resolve ambiguous cases, including immune- and epithelial-associated signals such as CD45 and HLA-DR (immune compartment) as well as Pan-Keratin and β-catenin (tumor/epithelial compartment), which provide complementary boundary cues when nuclear-only information is insufficient. In many reproductions of this benchmark, a compact recommended panel is operationalized as the set $\{$dsDNA, CD4, β-catenin, HLA-DR, H3K9ac, Pan-Keratin, H3K27me3$\}$, which we use as the literature-recommended baseline in our comparisons [Keren et al. 2018].

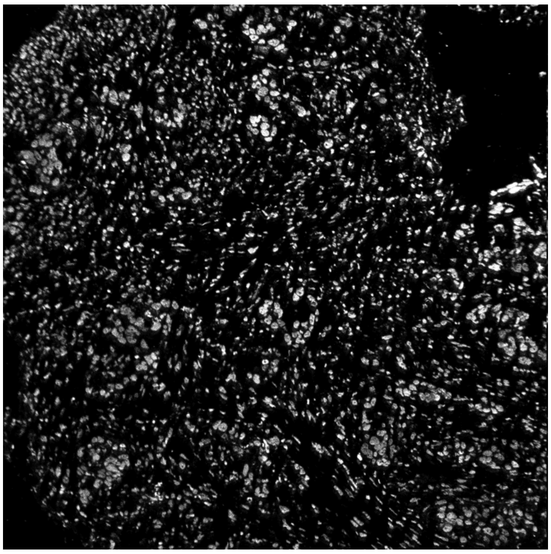

Figure 2. dsDNA channel.

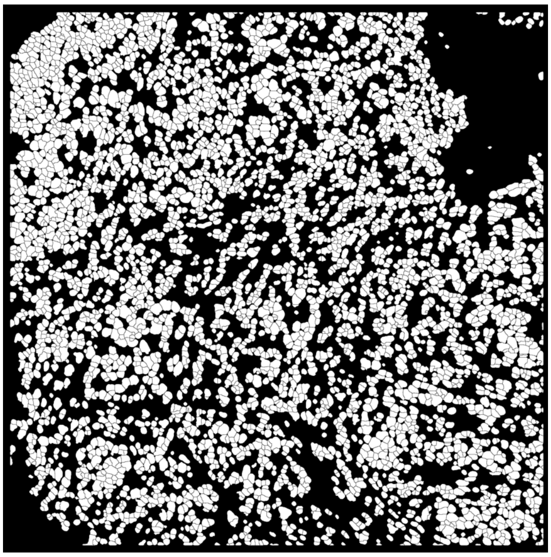

Figure 3. Mask Example.

Figure 3 shows an example of the binary/semantic ground-truth mask provided with the dataset, which we use as the supervision target in all pixel-wise segmentation experiments. Figure 2 visualizes the raw intensity map of the dsDNA channel for the same field of view. This marker provides strong nuclear contrast and, in our single-channel screening, it achieved the highest Dice score among all evaluated channels, motivating its role as a key input candidate for compact, data-driven panels.

3.2. Data Pre-processing

Data pre-processing was performed in two stages: (i) an offline, dataset-generation script that produces a per-channel dataset with paired images and masks, and (ii) the minimal input pipeline applied at training/inference time before feeding samples to the U-Net.

Stage 1 - Per-channel dataset generation: For each patient folder and for each selected channel, we loads the corresponding segmentation mask. The 2D image of the selected biomarker channel is then loaded and validated to have the same spatial shape as the mask. Next, a region of interest is extracted by trimming both the mask and the channel image to the exact bounding box of pixels with mask value greater than 0. Fields of view without any cell pixels are discarded. The mask uses nearest-neighbor interpolation to preserve discrete labels, whereas the biomarker image uses bilinear interpolation to preserve continuous intensity structure. Finally, the mask is stored as a binary image in PNG format, and the biomarker image is saved as a single-channel 32-bit floating TIFF. Outputs are organized in a per-channel structure, enabling systematic single-channel screening and controlled construction of multi-channel panels.

Stage 2 - Minimal U-Net input pipeline: During training/testing, each biomarker TIFF is read as a single-channel image, normalized by its per-image maximum so that intensities lie in $[0, 1]$. The corresponding mask PNG is read in grayscale and re-binarized using a threshold greater than 0. Both image and mask are then resized to 512×512 for network input. The preprocessing is limited to simple max-normalization, mask binarization, and resizing for computational efficiency and full reproducibility.

3.3. Tests

We conducted two complementary experiments to quantify how marker choice and the number of input channels affect U-Net segmentation performance on the dataset. In all experiments, we report and compare models using the validation Dice as the selection criterion.

3.3.1. Single-channel screening and ranking

First, we perform a single-channel screening over all available markers. For each channel m in the full MIBI-TOF panel, we train a U-Net whose input consists solely of the image corresponding to m (one model per marker). All models share the same architecture, loss function, and training schedule to ensure a fair comparison. Specifically, each model uses a U-Net decoder paired with a ResNet-34 encoder initialized with ImageNet weights.For optimization, we used Adam with a learning rate of 1×10^{-4}, combined with a cosine annealing learning-rate schedule over 30 epochs. The training objective was a composite Dice+BCE loss, defined as an equal-weight combination of binary cross-entropy with logits and soft Dice loss, encouraging both pixel-wise accuracy and overlap quality.After training, we record the best dice obtained during training and use it to rank markers according to their individual segmentation utility when used in isolation. All experiments were conducted under a fixed hyperparameter setting to ensure comparability across runs: we trained for 30 epochs, used a 20/80 hold-out split for test, and adopted a U-Net backbone with a ResNet-34 encoder.

3.3.2. Incremental multi-channel aggregation

Second, we evaluate how segmentation performance evolves as additional channels are added to the input. Starting from the best-ranked marker (dsDNA), we build multi-channel inputs by incrementally concatenating the next marker in the ranking and re-training the model from scratch at each step, again using the same fixed protocol as in Experiment 1. In other words, we first evaluate dsDNA+P, then dsDNA+P+H3K9ac, then dsDNA+P+H3K9ac+Ca, and so on until all nine channels are included. Each configuration is compared using dice, allowing us to measure (i) the marginal benefit of adding each marker, (ii) performance saturation as the number of channels increases, and (iii) whether adding lower-ranked channels introduces redundancy or noise that can reduce segmentation quality.

4. Results

This section reports the segmentation results from a systematic comparison of single-channel and multi-channel biomarker configurations. The findings show that markers encoding nuclear and structural information are the most effective for segmentation, whereas simply expanding the input panel does not consistently improve Dice performance. Taken together, these results indicate that segmentation quality is determined chiefly by marker informativeness and complementarity, rather than by the number of channels included.

Table 2. Top-20 single-channel results (U-Net, 30 epochs). Dice values correspond to the validation Dice (reported as Dice).

Channel	Dice (%)	Short description
dsDNA	89.36	Double-stranded DNA (nuclear signal; strong cue for nuclei localization).
P	85.24	Phosphorus (elemental; diffuse structural contrast from phosphate-rich biomolecules).
H3K9ac	81.63	Histone H3K9 acetylation (nuclear chromatin mark; reinforces nuclear morphology).
Ca	80.53	Calcium (elemental; broad structural/compartment contrast in tissue).
H3K27me3	79.07	Histone H3K27 trimethylation (nuclear chromatin mark; robust nuclear signal).
Na	75.47	Sodium (elemental; diffuse contrast that can help tissue/background separation).
phospho-S6	71.81	pS6 (mTOR/PI3K pathway activity; often highlights activated/proliferative cells).
HLA_Class_1	71.61	MHC-I (HLA-A/B/C; membrane-associated antigen presentation, useful for contours).
Fe	70.15	Iron (elemental; may correlate with blood/vascular regions or deposits; contrast cue).
C	70.07	Carbon (elemental; global tissue/background contrast, often diffuse).
PD-L1	70.04	Immune checkpoint ligand (tumor/APCs; immunosuppression marker).
CD163	69.85	TAM/macrophage marker (often M2-like; immunosuppressive myeloid signal).
OX40	69.80	T-cell activation/co-stimulation marker (activated T cells, mainly CD4).
B7H3	69.66	CD276 (B7 family; tumor/immune modulation marker in TME).
Lag3	69.61	LAG-3 checkpoint (T-cell dysfunction/exhaustion-associated marker).
CD209	69.59	DC-SIGN (dendritic cell / some myeloid populations; APC-related).
CD16	69.52	FCGR3A (NK/monocyte marker; cytotoxic/myeloid compartment cue).
FoxP3	69.34	Regulatory T-cell transcription factor (Treg; nuclear signal).
HLA-DR	69.31	MHC-II (APC activation/presentation; dendritic/macrophage/B-cell enriched).
CD45RO	69.29	Memory leukocyte marker (memory T-cell enriched; immune compartment cue).

Table 2 reports the validation Dice obtained when training the U-Net using a *single* MIBI-TOF channel as input. The highest Dice values are achieved by nuclear/structural channels like dsDNA and the chromatin marks H3K9ac and H3K27me3, consistent with prior TNBC MIBI-TOF processing where dsDNA and histone channels are combined to form a robust nuclear image for segmentation and reflect canonical nucleus-localized chromatin modifications [Keren et al. 2018, Kouzarides 2007].

A second group of strong performers comprises endogenous elemental channels (P, Ca, Na, Fe, and C), which are naturally captured by SIMS/ToF-based ion-beam imaging and can provide diffuse tissue/background contrast that helps delineate cellular structures [Angelo et al. 2014, Passarelli and Winograd 2011]. In contrast, functional and immunophenotyping markers such as phospho-S6 (mTOR pathway activity) [Saxton and Sabatini 2017], antigen-presentation markers (HLA class I and HLA-DR) [Neefjes et al. 2011], immune checkpoints/co-stimulation (PD-L1, LAG3, OX40, and B7-H3) [Pardoll 2012, Andrews et al. 2017, Croft 2010, Picarda et al. 2016], and lineage markers for macrophages/dendritic cells/NK or memory/Treg T cells (CD163, CD209, CD16, CD45RO, and FoxP3) [Etzerodt and Moestrup 2013, Geijtenbeek et al. 2000, Nimmerjahn and Ravetch 2008, Farber et al. 2014, Hori et al. 2003] tend to yield lower Dice because their expression is cell-type or state specific and does not consistently outline nuclei or full cell contours across all regions of the tissue.

Table 3 presents the top-10 single-channel models obtained from the screening experiment. Each row corresponds to a U-Net trained using only one biomarker channel as input and evaluated with Dice on the validation split. The results show that dsDNA is the most informative individual marker for pixel-wise segmentation, reaching, followed by P and H3K9ac.

Table 3. Incremental data-driven panels built from the single-channel ranking (top-k). Dice values are reported as validation Dice (val_dice).

k	Channels	Dice (%)
1	dsDNA	**89.36**
2	dsDNA + P	88.88
3	dsDNA + P + H3K9ac	88.24
4	dsDNA + P + H3K9ac + Ca	87.71
5	dsDNA + P + H3K9ac + Ca + H3K27me3	87.27
6	dsDNA + P + H3K9ac + Ca + H3K27me3 + Na	87.31
7	dsDNA + P + H3K9ac + Ca + H3K27me3 + Na + phospho-S6	87.99
8	dsDNA + P + H3K9ac + Ca + H3K27me3 + Na + phospho-S6 + HLA_Class_1	88.52
9	dsDNA + P + H3K9ac + Ca + H3K27me3 + Na + phospho-S6 + HLA_Class_1 + Fe	87.68
10	dsDNA + P + H3K9ac + Ca + H3K27me3 + Na + phospho-S6 + HLA_Class_1 + Fe + C	85.90

Table 3 reports the results of the incremental data-driven multi-channel experiment. Here, channels are added in descending order of the single-channel ranking (starting with dsDNA), and a new model is trained for each top-k panel. This table is intended to quantify how segmentation accuracy changes as additional high-ranked markers are concatenated to the input.

Table 4. Incremental combinations of the recommended panel (added in the order: dsDNA, CD4, β-catenin, HLA-DR, H3K9ac, Pan-Keratin, H3K27me3). Dice values are reported as Dice.

k	Channels	Dice (%)
2	dsDNA + CD4	89.38
3	dsDNA + CD4 + β-catenin	88.45
4	dsDNA + CD4 + β-catenin + HLA-DR	88.04
5	dsDNA + CD4 + β-catenin + HLA-DR + H3K9ac	87.08
6	dsDNA + CD4 + β-catenin + HLA-DR + H3K9ac + Pan-Keratin	86.58
7	dsDNA + CD4 + β-catenin + HLA-DR + H3K9ac + Pan-Keratin + H3K27me3	**89.41**

Table 4 evaluates the dataset-recommended marker panel by incrementally aggregating channels in the order specified by the dataset paper (dsDNA, CD4, β-catenin, HLA-DR, H3K9ac, Pan-Keratin, H3K27me3). This table provides a baseline that reflects a commonly adopted expert-driven channel subset for segmentation. The best result within this baseline occurs at $k = 7$ (dsDNA + CD4 + -catenin + HLA-DR + H3K9ac + Pan-Keratin + H3K27me3,89.41% Dice). Compared to the data-driven panels, these results indicate that compact subsets can match the recommended configuration and that larger fixed panels are not necessarily beneficial for pixel-wise segmentation in this dataset.

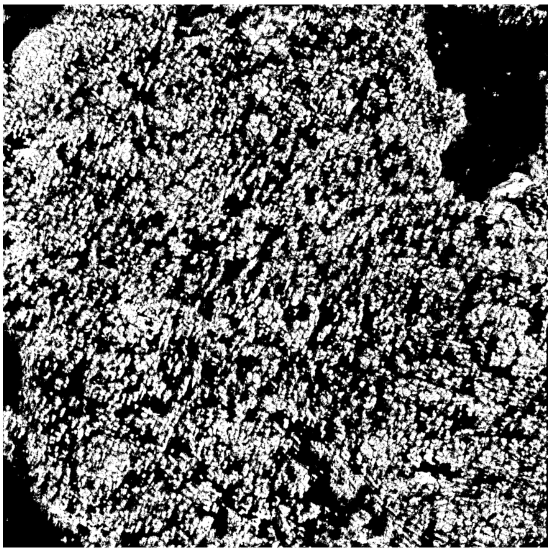

Figure 4. dsDNA predicted mask

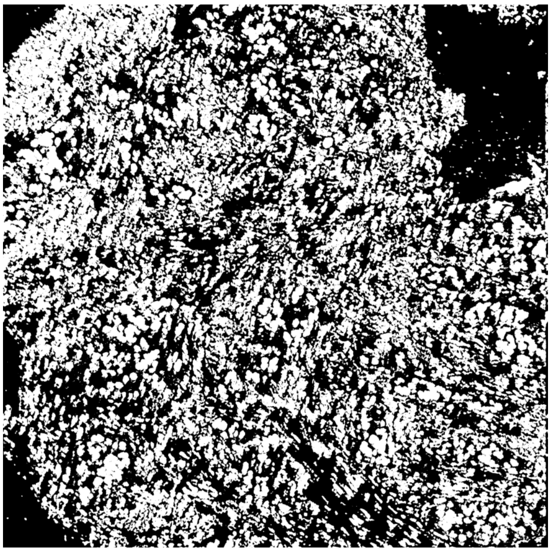

Figure 5. recommended panel predicted mask

5. Conclusion

This paper examined how marker selection affects U-Net semantic (pixel-wise) segmentation in MIBI-TOF TNBC by conducting a systematic single-channel screening under a fixed training protocol and then building incremental, data-driven multi-channel panels by aggregating top-ranked markers. We further contrasted these data-driven panels with the literature-recommended marker panel associated with the dataset. Overall, our results indicate that marker redundancy is substantial and that adding more channels does not necessarily improve segmentation; instead, extremely compact inputs can be sufficient and highly efficient for this benchmark. Data-driven channel selection can match the performance of broader expert-defined panels while using significantly fewer markers, as demonstrated by our single-channel dsDNA model (89.36% Dice) performing on par with the full 7-channel recommended panel (89.41% Dice). These findings highlight the need to explicitly validate marker choices and prioritize redundancy reduction when designing efficient segmentation workflows for spatial proteomics. As future work, we plan to (i) validate the stability of channel rankings across different train/validation splits and longer training schedules, (ii) extend the analysis to instance-level segmentation metrics and downstream per-cell quantification tasks, and (iii) investigate alternative normalization and fusion strategies that may better exploit complementary signals across channels.

6. Acknowledgments

The present work was carried out with the support of the Universidade Federal do Piauí (UFPI), the Instituto Federal do Piauí (IFPI), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Finance Code 001, the Fundação de Amparo à Pesquisa do Estado do Piauí (FAPEPI), and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

Andrews, L. P., Marciscano, A. E., Drake, C. G., and Vignali, D. A. A. (2017). LAG3 (CD223) as a cancer immunotherapy target. *Immunological Reviews*, 276(1):80–96.

Angelo, M., Bendall, S. C., Finck, R., Hale, M. B., Hitzman, C., Borowsky, A. D., Levenson, R. M., Lowe, J. B., Liu, S. D., Zhao, S., Natkunam, Y., and Nolan, G. P. (2014). Multiplexed ion beam imaging of human breast tumors. *Nature Medicine*, 20(4):436–442.

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., and Jemal, A. (2024). Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 74(3):229–263.

Chen, Z., Soifer, I., Hilton, H., Keren, L., and Jojic, V. (2020). Modeling multiplexed images with spatial-lda reveals novel tissue microenvironments. *Journal of Computational Biology*, 27(8):1204–1218.

Croft, M. (2010). Control of immunity by the TNFR-related molecule OX40 (CD134). *Annual Review of Immunology*, 28:57–78.

Etzerodt, A. and Moestrup, S. K. (2013). CD163 and inflammation: biological, diagnostic, and therapeutic aspects. *Antioxidants & Redox Signaling*, 18(17):2352–2363.

Farber, D. L., Yudanin, N. A., and Restifo, N. P. (2014). Human memory T cells: generation, compartmentalization and homeostasis. *Nature Reviews Immunology*, 14(1):24–35.

Geijtenbeek, T. B. H., Kwon, D.-S., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C. F., Middel, J., Cornelissen, I. L. M. H. A., Nottet, H. S. L. M., KewalRamani,

V. N., Littman, D. R., Figdor, C. G., and van Kooyk, Y. (2000). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell*, 100(5):587–597.

Hori, S., Nomura, T., and Sakaguchi, S. (2003). Control of regulatory T cell development by the transcription factor Foxp3. *Science*, 299(5609):1057–1061.

Jie, H., Ma, W., and Huang, C. (2025). Diagnosis, prognosis, and treatment of triple-negative breast cancer: A review. *Breast Cancer: Targets and Therapy*, pages 265–274.

Keren, L., Bosse, M., Marquez, D., Angoshtari, R., Jain, S., Varma, S., Yang, S.-R., Kurian, A., Van Valen, D., West, R., Bendall, S. C., and Angelo, M. (2018). A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell*, 174(6):1373–1387.e19.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, 128(4):693–705.

Liu, C. C., Greenwald, N. F., Kong, A., McCaffrey, E. F., Leow, K. X., Mrdjen, D., Cannon, B. J., Rumberger, J. L., Varra, S. R., and Angelo, M. (2023). Robust phenotyping of highly multiplexed tissue imaging data using pixel-level clustering. *Nature communications*, 14(1):4618.

Magness, A., Colliver, E., Enfield, K. S., Lee, C., Shimato, M., Daly, E., Moore, D. A., Sivakumar, M., Valand, K., Levi, D., et al. (2024). Deep cell phenotyping and spatial analysis of multiplexed imaging with tracerx-phlex. *Nature Communications*, 15(1):5135.

Neefjes, J., Jongsma, M. L. M. J., Paul, P., and Bakke, O. (2011). Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology*, 11(12):823–836.

Nimmerjahn, F. and Ravetch, J. V. (2008). Fcγ receptors as regulators of immune responses. *Nature Reviews Immunology*, 8(1):34–47.

Pardoll, D. M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer*, 12(4):252–264.

Passarelli, M. K. and Winograd, N. (2011). Lipid imaging with time-of-flight secondary ion mass spectrometry (ToF-SIMS). *Biochimica et Biophysica Acta*, 1811(11):976–990.

Patwa, A., Yamashita, R., Long, J., Risom, T., Angelo, M., Keren, L., and Rubin, D. L. (2021). Multiplexed imaging analysis of the tumor-immune microenvironment reveals predictors of outcome in triple-negative breast cancer. *Communications biology*, 4(1):852.

Petinrin, O. O., Saeed, F., Toseef, M., Liu, Z., Basurra, S., Muyide, I. O., Li, X., and Wong, K.-C. (2023). Machine learning in metastatic cancer research: Potentials, possibilities, and prospects. *Computational and Structural Biotechnology Journal*, 21:2454–2470.

Picarda, E., Ohaegbulam, K. C., and Zang, X. (2016). Molecular pathways: Targeting B7-H3 (CD276) for human cancer immunotherapy. *Clinical Cancer Research*, 22(14):3425–3431.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, volume 9351 of *Lecture Notes in Computer Science*, pages 234–241. Springer.

Saxton, R. A. and Sabatini, D. M. (2017). mTOR signaling in growth, metabolism, and disease. *Cell*, 168(6):960–976.