HExIF: An End-to-End Deep Learning Framework for Virtual Multiplex Immunofluorescence from H&E Slides

Rany Stephan[1,2*], Andrew Gentles[2*], Mohamed Kouchak[2] and Magdalena Matusiak[2]

[1*]Institute for Computational and Mathematical Engineering, Stanford University, California, USA.
[2]Department of Biomedical Data Science, Stanford University School of Medicine, California, USA.

*Corresponding author(s). E-mail(s): ranycs@stanford.edu; andrewg@stanford.edu;
Contributing authors: mohamedk@stanford.edu; mmatusiak@stanford.edu;

Abstract

Multiplexed immunofluorescence imaging has emerged as a powerful technology for characterizing the spatial organization of the tumor microenvironment at single-cell resolution. However, these technologies remain costly, time-intensive, and require specialized equipment, limiting their adoption in routine clinical practice. Here, we present HExIF (Hematoxylin and Eosin to Immunofluorescence), a computational framework for predicting spatially-resolved multiplex protein expression directly from standard hematoxylin and eosin (H&E)-stained tissue sections. Our approach integrates automated image registration, tissue segmentation, and a deep learning architecture based on hierarchical vision transformers to simultaneously predict expression patterns across a 20-marker immunofluorescence panel TODO: change this when done with CODEX as well. The framework employs a multi-component loss function designed to address the challenges of sparse marker expression and channel-specific intensity distributions. We demonstrate the methodology on an Orion multiplex imaging dataset with markers targeting macrophage subpopulations, T cell subtypes, and stromal components. This work establishes a foundation for virtual multiplexing that may accelerate spatial biomarker discovery and enable retrospective analysis of archival H&E specimens.

Keywords: Computational pathology Virtual staining
Multiplex immunofluorescence Deep learning Tumor microenvironment
Spatial transcriptomics

1 Introduction

Understanding the spatial organization of cells within the tumor microenvironment (TME) has become a central focus in cancer research and precision oncology Binnewies et al. (2018). The TME comprises a complex ecosystem of malignant cells, immune infiltrates, stromal components, and vasculature, whose spatial relationships profoundly influence disease progression, therapeutic response, and patient outcomes Schürch et al. (2020). While bulk profiling methods have provided valuable insights into the molecular landscape of tumors, they inherently obscure the spatial heterogeneity that underlies critical biological phenomena such as immune evasion, metabolic adaptation, and therapeutic resistance Marusyk et al. (2020).

Recent advances in multiplexed imaging technologies—including CO-Detection by indEXing (CODEX), Multiplexed Ion Beam Imaging (MIBI), Imaging Mass Cytometry (IMC), and cyclic immunofluorescence platforms such as Orion—have enabled simultaneous visualization of dozens to hundreds of protein markers at subcellular resolution Goltsev et al. (2018); Angelo et al. (2014); Giesen et al. (2014). These technologies have revealed previously unappreciated cellular neighborhoods, immune niches, and spatially-defined phenotypic states that carry prognostic and predictive significance across multiple cancer types Jackson et al. (2020). However, multiplexed imaging remains constrained by substantial cost, specialized instrumentation, extended acquisition times, and complex experimental workflows, limiting its scalability for large cohort studies and routine clinical deployment.

In contrast, hematoxylin and eosin (H&E) staining represents the cornerstone of diagnostic pathology, with billions of archived specimens available worldwide Zarella et al. (2019). H&E imaging is inexpensive, rapid, and universally accessible, yet it provides only morphological information without the molecular specificity of immunolabeling. This disparity has motivated growing interest in computational approaches that leverage deep learning to infer molecular features from H&E images, an approach broadly termed "virtual staining" or "in silico immunohistochemistry" Rivenson et al. (2020).

1.1 Spatial Single-Cell Imaging in Computational Pathology

The emergence of spatial single-cell imaging has catalyzed a paradigm shift in how we conceptualize tissue architecture and cellular interactions Lewis et al. (2021). Unlike traditional immunohistochemistry (IHC), which typically permits visualization of only one to three markers simultaneously, multiplexed platforms enable comprehensive phenotyping of individual cells within their native tissue context. This capability has proven particularly valuable for dissecting the immune microenvironment, where the

functional state of immune cells depends critically on their spatial positioning relative to tumor cells, vasculature, and other stromal elements Moldoveanu et al. (2022).

Orion, the multiplexed imaging platform employed in this study, utilizes cyclic immunofluorescence with spectrally-resolved detection to achieve simultaneous 20-channel imaging RareCyte, Inc. (2024). The marker panel used in this work was designed to capture key populations of the myeloid and lymphoid compartments, with particular emphasis on macrophage heterogeneity. Specifically, the panel includes markers for: nuclear identification (Hoechst); macrophage lineage and polarization (CD68, IBA1, CD163, FOLR2, SPP1); T cell subsets (CD3ε, CD8α, FOXP3); stromal components (FAP, SMA, Pan-CK); and additional markers relevant to immune function and tissue architecture (NLRP3, GFPT2, CD15, LYVE1, IL-4I1).

1.2 Immune Microenvironment Heterogeneity and Biological Motivation

The tumor immune microenvironment exhibits profound spatial heterogeneity that has emerged as a critical determinant of clinical outcomes and therapeutic response Thorsson et al. (2018). Among immune cell populations, tumor-associated macrophages (TAMs) have attracted particular attention due to their remarkable plasticity and context-dependent functions that can either promote or suppress tumor progression Cassetta and Pollard (2020).

Macrophages represent a phenotypically diverse population that defies simple classification into discrete subtypes. While the classical M1/M2 polarization paradigm provided an initial conceptual framework, single-cell and spatial analyses have revealed a continuum of macrophage states that vary with tissue context, disease stage, and spatial localization within the TME Mulder et al. (2021). Markers such as CD163 and FOLR2 have been associated with alternatively-activated or tissue-resident macrophage phenotypes, while SPP1 (osteopontin) expression marks a distinct macrophage population with pro-tumorigenic and pro-fibrotic properties Bill et al. (2023). The inflammasome component NLRP3 and the immunomodulatory enzyme IL-4I1 provide additional resolution for distinguishing macrophage functional states.

The spatial distribution of macrophage subpopulations carries prognostic significance across multiple tumor types. Studies have demonstrated that the proximity of specific macrophage phenotypes to tumor cells, vasculature, and lymphocyte populations correlates with clinical outcomes and response to immunotherapy Qi et al. (2022). However, comprehensive spatial characterization of macrophage heterogeneity currently requires multiplexed imaging, motivating the development of computational approaches to infer this information from more accessible H&E specimens.

1.3 Limitations of Current Spatial Phenotyping Pipelines

Existing approaches to virtual staining and computational immunohistochemistry face several methodological challenges that limit their utility for comprehensive spatial phenotyping:

Single-marker prediction. The majority of published methods focus on predicting individual immunohistochemical stains, such as Ki-67, HER2, or PD-L1 Xu

et al. (2017); Naik et al. (2020). While valuable, this approach fails to capture the correlated expression patterns and spatial relationships among multiple markers that define cellular phenotypes and tissue organization.

Limited marker panels. Methods that do address multi-marker prediction typically target small panels of 3–5 markers, insufficient for comprehensive immune phenotyping Fu et al. (2020). Scaling to larger panels introduces challenges related to class imbalance, heterogeneous expression distributions, and increased model complexity.

Sparsity and class imbalance. Many markers of biological interest—particularly those identifying rare immune populations—exhibit highly sparse expression patterns, with positive cells comprising less than 1% of the tissue area. Standard regression losses struggle to capture these sparse signals, often producing models that underestimate rare events Johnson and Khoshgoftaar (2019).

Registration and alignment. Generating paired H&E and multiplexed imaging data requires precise image registration, a non-trivial task given differences in tissue processing, sectioning, and imaging modalities. Registration errors propagate into training data noise, degrading model performance Borovec et al. (2018).

Intensity normalization. Multiplexed imaging data exhibits substantial variability in intensity distributions across channels, experimental batches, and tissue types. Appropriate normalization strategies are essential for training stable models and ensuring transferability Macenko et al. (2009).

1.4 Study Objectives and Contributions

This work presents HExIF (Hematoxylin and Eosin to Immunofluorescence), an end-to-end computational framework for predicting spatially-resolved multiplex immunofluorescence from standard H&E-stained tissue sections. Our primary contributions are:

1. **Comprehensive pipeline architecture.** We develop an integrated workflow encompassing automated image registration, tissue segmentation, training data generation, and deep learning-based prediction. The pipeline leverages the VALIS registration framework for robust multimodal alignment and employs morphological analysis for tissue microarray (TMA) core detection.
2. **20-channel simultaneous prediction.** Unlike prior work focused on single markers or limited panels, our approach predicts a comprehensive 20-marker Orion immunofluorescence panel from H&E input. The marker panel enables detailed characterization of macrophage subpopulations, T cell subtypes, and stromal architecture.
3. **Channel-aware training strategy.** We introduce adaptive sampling and loss weighting strategies that address the challenges of heterogeneous marker expression patterns. Channel-specific coverage statistics guide both training sample selection and loss computation, ensuring that rare markers receive adequate representation.
4. **Multi-component loss function.** The training objective combines center-window weighted reconstruction loss, per-channel coverage matching, structural

similarity constraints, total variation regularization, and an auxiliary marker presence classification head. This composite loss addresses multiple failure modes including intensity underestimation, spatial blurring, and hallucinated artifacts.

5. **Scalable distributed training.** The implementation supports multi-GPU distributed training via PyTorch DistributedDataParallel, enabling efficient processing of large-scale imaging datasets. Global intensity normalization parameters are computed on the training set and synchronized across processes.

The current work focuses on methodology development and validation using an Orion multiplex imaging dataset. We describe the computational approach in detail and present the framework as a foundation for subsequent quantitative evaluation and biological validation. Future extensions will incorporate additional multiplexed imaging modalities including CODEX, enabling cross-platform generalization studies.

2 Related Work

2.1 Multiplexed Imaging Technologies

The past decade has witnessed rapid development of multiplexed tissue imaging technologies that enable simultaneous detection of numerous protein markers at single-cell resolution. These platforms can be broadly categorized by their detection modality: mass spectrometry-based methods (IMC, MIBI), cyclic immunofluorescence (CODEX, CyCIF, Orion), and antibody-DNA barcoding approaches (CODEX) Tan et al. (2020).

CODEX (CO-Detection by indEXing) employs oligonucleotide-conjugated antibodies with iterative fluorescent reporter hybridization, enabling panels of 40–60 markers Goltsev et al. (2018). The platform has been extensively applied to characterize immune organization in lymphoid tissues, tumor microenvironments, and inflammatory diseases. Goltsev and colleagues demonstrated its utility for mapping spatially-defined immune niches in normal and diseased tissues Goltsev and Nolan (2023).

Imaging Mass Cytometry (IMC) and **MIBI** utilize metal-tagged antibodies with mass spectrometry detection, achieving panels of 40+ markers with subcellular resolution Giesen et al. (2014); Angelo et al. (2014). While offering excellent multiplexing capacity and minimal spectral overlap, these platforms have lower throughput and require specialized instrumentation.

Orion and related cyclic immunofluorescence platforms achieve multiplexing through iterative staining and imaging cycles with spectrally-resolved fluorophore detection RareCyte, Inc. (2024). The ArgoFluor series employed in this study provides 20-channel capacity with optimized spectral separation. These platforms balance multiplexing capacity with relatively accessible instrumentation and established immunofluorescence workflows.

2.2 Computational Pathology for Spatial Omics

Computational pathology encompasses a broad range of methods for extracting quantitative information from tissue images, with applications spanning diagnosis, prognosis, and biomarker discovery Echle et al. (2021). The integration of deep learning has

transformed the field, enabling extraction of features that correlate with molecular alterations, treatment response, and survival outcomes from standard histopathology images Bera et al. (2019).

Within the spatial omics context, computational methods address several interrelated tasks: cell segmentation and phenotyping, spatial neighborhood analysis, and integration with molecular profiling data Palla et al. (2022). Cell segmentation in multiplexed imaging typically employs nuclear detection followed by cytoplasmic boundary estimation, with recent approaches leveraging deep learning for improved accuracy Stringer et al. (2021). Following segmentation, cells are phenotyped based on marker expression patterns, often using clustering approaches or supervised classification.

2.3 Machine Learning for Cellular Phenotyping

Deep learning approaches for cellular phenotyping in tissue images have evolved from simple convolutional architectures to sophisticated multi-task and transformer-based models Campanella et al. (2019). Key methodological advances include:

U-Net and encoder-decoder architectures have become the dominant paradigm for dense prediction tasks in pathology, providing both localization and context through skip connections Ronneberger et al. (2015). Variants incorporating attention mechanisms, residual connections, and multi-scale feature aggregation have demonstrated improved performance on segmentation and regression tasks.

Vision transformers have recently emerged as powerful alternatives to convolutional networks for pathology applications Chen et al. (2021). The Swin Transformer architecture, which employs shifted windows for efficient self-attention computation, has shown particular promise for capturing long-range dependencies in tissue images Liu et al. (2021). We adopt a Swin Transformer encoder with a feature pyramid decoder in this work.

Virtual staining methods aim to computationally transform images from one staining modality to another. Rivenson and colleagues pioneered deep learning-based virtual staining, demonstrating that autofluorescence images could be transformed to appear as if H&E-stained Rivenson et al. (2019). Subsequent work has extended this concept to various stain transformations including H&E to IHC Burlingame et al. (2020), unstained to H&E de Haan et al. (2021), and autofluorescence to multiple special stains Zhang et al. (2020b).

HistoPlexer, developed by Andani and colleagues, represents a significant advance in multi-marker prediction from histopathology images Andani et al. (2024). Their approach employs a conditional generative model to predict multiplex immunofluorescence patterns, demonstrating feasibility of comprehensive virtual multiplexing. Our work builds on this foundation while introducing several methodological innovations including channel-aware sampling and multi-component loss design.

2.4 Spatial Statistics and Microenvironment Modeling

Quantitative analysis of spatial patterns in tissue images draws on methods from spatial statistics, graph theory, and computational geometry Baddeley et al. (2015). Common approaches include:

Spatial point patterns characterize the distribution of cells through summary statistics such as Ripley's K function, nearest neighbor distances, and spatial correlation functions Barua et al. (2018). These methods quantify clustering, dispersion, and co-localization among cell populations.

Graph-based representations model tissue architecture as networks where nodes represent cells and edges encode spatial relationships Pati et al. (2022). Community detection algorithms can identify cellular neighborhoods, while graph neural networks enable learning of spatial features for downstream prediction tasks.

Spatial transcriptomics integration combines imaging-based protein measurements with RNA profiling to provide complementary molecular and spatial information Marx (2021). While outside the scope of this work, such integration represents an important direction for comprehensive tissue characterization.

2.5 Macrophage Biology and Spatial Immunology

Tumor-associated macrophages represent one of the most abundant immune populations in the tumor microenvironment, comprising up to 50% of the tumor mass in some cancer types Qian and Pollard (2010). Their functional diversity and plasticity have made them both challenging to characterize and attractive as therapeutic targets.

Single-cell transcriptomic studies have identified multiple macrophage subpopulations within tumors, including $SPP1^+$ macrophages associated with tissue remodeling and angiogenesis, $FOLR2^+$ macrophages with tissue-resident phenotypes, and inflammatory macrophages expressing NLRP3 and other inflammasome components Cheng et al. (2021). Importantly, these subpopulations exhibit distinct spatial distributions, with $SPP1^+$ macrophages enriched at the tumor-stroma interface and $FOLR2^+$ macrophages in perivascular regions Zhang et al. (2020a).

The spatial organization of macrophages relative to T cells has emerged as a critical determinant of immunotherapy response. Studies have demonstrated that physical proximity between macrophages and $CD8^+$ T cells influences T cell activation, exhaustion, and cytotoxic function Peranzoni et al. (2018). Spatial metrics quantifying macrophage-T cell interactions have shown prognostic value in multiple cancer types and may predict response to immune checkpoint blockade.

Given the biological importance of macrophage spatial heterogeneity, we designed our marker panel to enable detailed characterization of macrophage subpopulations. The inclusion of CD68 (pan-macrophage), IBA1 (microglial/macrophage lineage), CD163 (alternative activation), FOLR2 (tissue-resident), SPP1 (pro-fibrotic), NLRP3 (inflammasome), and IL-4I1 (immunomodulatory) provides comprehensive coverage of macrophage diversity relevant to tumor immunity.

3 Data and Imaging Modalities

3.1 Orion Multiplex Imaging Platform

The Orion multiplex imaging platform enables simultaneous detection of 20 protein markers through cyclic immunofluorescence with spectrally-resolved ArgoFluor detection RareCyte, Inc. (2024). Unlike sequential staining approaches that require tissue stripping between cycles, Orion employs spectral unmixing algorithms to resolve overlapping fluorophore emissions, enabling single-round acquisition of the complete marker panel.

Image acquisition generates a multi-channel OME-TIFF stack with 20 fluorescence channels plus optional brightfield reference. Native resolution varies by objective but typically achieves subcellular detail sufficient for nuclear morphology and membrane marker localization. For this study, registered images were processed at full resolution prior to patch extraction for model training.

3.2 Marker Panel and Biological Annotation

The 20-channel Orion panel employed in this study was designed to comprehensively characterize the tumor immune microenvironment with emphasis on myeloid diversity. Table 1 provides the complete marker list with biological annotations.

Table 1 Orion 20-channel marker panel with biological annotations.

Channel	Marker	Fluorophore	Biological Function
01	Hoechst	–	Nuclear DNA stain
02	AF1	–	Autofluorescence control
03	SPP1	ArgoFluor520	Macrophage (pro-fibrotic)
04	CD68	ArgoFluor555L	Pan-macrophage/monocyte
05	CD3ε	ArgoFluor548	Pan-T cell
06	AF2	–	Autofluorescence control
07	IBA1	ArgoFluor660L	Microglia/macrophage
08	–	ArgoFluor572	Reserved
09	FAP	ArgoFluor602	Fibroblast activation
10	CD8α	ArgoFluor624	Cytotoxic T cells
11	CD163	ArgoFluor658	M2/alternative macrophage
12	FOLR2	ArgoFluor676	Tissue-resident macrophage
13	GFPT2	ArgoFluor698	Metabolic enzyme
14	NLRP3	ArgoFluor706	Inflammasome
15	FOXP3	ArgoFluor724	Regulatory T cells
16	CD15	ArgoFluor760	Granulocytes
17	LYVE1	ArgoFluor782	Lymphatic endothelium
18	SMA	ArgoFluor812	Smooth muscle/myofibroblast
19	Pan-CK	ArgoFluor845	Epithelial cells
20	IL-4I1	ArgoFluor874	Immunomodulatory enzyme

The panel enables identification of key immune populations including: (1) macrophage subsets distinguishable by CD68, IBA1, CD163, FOLR2, SPP1, and IL-4I1 expression patterns; (2) T cell populations marked by CD3ε with subset resolution via CD8α and FOXP3; (3) stromal components including FAP$^+$ cancer-associated fibroblasts and SMA$^+$ myofibroblasts; and (4) structural markers for epithelium (Pan-CK) and lymphatic vessels (LYVE1).

3.3 Image Acquisition and Raw Data Characteristics

Tissue microarray (TMA) specimens were processed following standard immunofluorescence protocols with antibody validation and spectral calibration per manufacturer recommendations. Whole-slide images were acquired at native Orion resolution with matched H&E staining performed on serial sections.

Raw Orion data exhibits several characteristics relevant to computational processing:

- **Dynamic range heterogeneity**: Signal intensity varies by several orders of magnitude across channels, with Hoechst nuclear staining exhibiting the highest signal and markers such as IL-4I1 showing sparse, low-intensity expression.
- **Spatial sparsity**: Many markers label rare cell populations, with positive pixels comprising less than 1–5% of tissue area. This extreme class imbalance necessitates specialized training strategies.
- **Background autofluorescence**: Channels AF1 and AF2 capture tissue autofluorescence for quality control and potential correction, though autofluorescence contribution varies by tissue type and processing.

3.4 Planned Extension to CODEX Imaging

While the current implementation focuses on Orion multiplex data, the framework architecture is designed for extension to additional multiplexed imaging modalities. CODEX imaging data from ccRCC (clear cell renal cell carcinoma) and ccOC (clear cell ovarian carcinoma) tissue microarrays will be incorporated in subsequent phases of this work. The CODEX panel includes 53 markers with substantial overlap to the Orion panel (CD3, CD8, CD68, CD163, FAP, CD15, SMA, Pan-CK), enabling cross-platform validation and transfer learning studies. Methodological adaptations for CODEX data will address differences in marker panel composition, intensity distributions, and spatial resolution.

4 Computational Pipeline Overview

4.1 System Architecture and Data Flow

The HExIF pipeline comprises four major processing stages: (1) multimodal image registration, (2) tissue segmentation and region extraction, (3) training data generation with intensity normalization, and (4) deep learning model training and inference. Figure **??** provides a schematic overview of the complete workflow.

The pipeline is implemented in Python with dependencies including PyVIPS for large-image handling, VALIS for registration, scikit-image for morphological processing, and PyTorch for deep learning components. Distributed training support enables efficient processing on multi-GPU systems.

4.2 Image Registration and Alignment

Accurate spatial alignment between H&E and Orion images is essential for generating paired training data. We employ the VALIS (Virtual Alignment of pathoLogy Image Series) registration framework, which provides robust multimodal alignment through a combination of rigid and non-rigid transformations Gatenbee et al. (2023).

The registration workflow proceeds as follows:

1. **Image loading**: H&E and Orion whole-slide images are loaded via Bio-Formats reader to ensure compatibility with vendor-specific OME-TIFF formats.
2. **Rigid alignment**: Initial global alignment is computed using intensity-based registration with mutual information similarity metric.
3. **Non-rigid refinement**: Local deformations are estimated using B-spline free-form deformation to accommodate tissue distortions from sectioning and processing.
4. **Warping**: The Orion multi-channel stack is warped to the H&E coordinate system, producing pixel-level correspondence between modalities.

Registration quality is assessed through overlay visualization and edge alignment metrics. Cores with substantial residual misalignment are flagged for manual review or exclusion from training.

4.3 Preprocessing and Artifact Handling

Following registration, several preprocessing steps prepare the data for model training:

Tissue segmentation distinguishes tissue regions from background using a combination of Laplacian edge detection and Otsu thresholding applied to H&E images. The Laplacian filter emphasizes tissue boundaries, while Otsu's method provides adaptive binarization. Morphological closing and opening operations remove small artifacts and fill holes.

TMA core detection identifies individual tissue cores within TMA images through connected component analysis of the tissue mask. Bounding boxes are computed for each core, and cores are extracted with consistent padding to ensure uniform patch dimensions.

Intensity normalization addresses the substantial dynamic range differences across Orion channels. A global quantile-based scaling procedure is applied:

$$\tilde{x}_c(u,v) = \frac{x_c(u,v) - q_c^{\text{low}}}{q_c^{\text{high}} - q_c^{\text{low}} + \epsilon} \tag{1}$$

where q_c^{low} and q_c^{high} denote the 1st and 99.5th percentile intensities for channel c computed across the training set. This normalization is followed by a log-transform:

$$z_c(u, v) = \log\left(1 + \max(\tilde{x}_c(u, v), 0)\right) \tag{2}$$

The log-transform compresses the dynamic range and stabilizes training for channels with Poisson-like noise characteristics. Quantile parameters are computed on the training set only and persisted for application to validation and test data.

4.4 Cell Segmentation and Nuclear Detection

While the current implementation focuses on pixel-level prediction rather than cell-level phenotyping, the Hoechst nuclear channel provides opportunity for downstream cell segmentation. Nuclear detection can be performed using standard approaches (e.g., Otsu thresholding, watershed) or deep learning-based methods. Predicted marker intensities can then be aggregated within nuclear or cellular regions for single-cell analysis.

4.5 Feature Extraction from Multiplex Channels

The 20-channel Orion output encodes rich spatial and molecular information. For each predicted marker channel, relevant features include:

- **Intensity statistics**: Mean, variance, and percentile values quantifying marker abundance.
- **Spatial coverage**: Fraction of pixels exceeding expression threshold, indicating population prevalence.
- **Texture features**: GLCM-based or learned features capturing spatial organization of marker expression.
- **Co-localization**: Correlation and overlap metrics between marker pairs identifying co-expression patterns.

5 Machine Learning for Cellular Phenotyping

5.1 Problem Formulation and Phenotype Definitions

We formulate virtual multiplexing as a dense regression problem: given an H&E image patch $\mathbf{X} \in \mathbb{R}^{H \times W \times 3}$, predict the corresponding multi-channel Orion intensity map $\mathbf{Y} \in \mathbb{R}^{H \times W \times C}$ where $C = 20$ denotes the number of marker channels. The prediction is performed in log-transformed, quantile-normalized space to stabilize training dynamics.

Let $f_\theta : \mathbb{R}^{H \times W \times 3} \to \mathbb{R}^{H \times W \times C}$ denote the neural network parameterized by θ. The model produces non-negative outputs $\hat{\mathbf{Y}} = f_\theta(\mathbf{X})$ via Softplus activation, ensuring predictions lie in the valid intensity range.

5.2 Feature Representation and Normalization

H&E input images are normalized using ImageNet statistics (mean $[0.485, 0.456, 0.406]$, standard deviation $[0.229, 0.224, 0.225]$), enabling effective transfer learning from pretrained vision encoders. Data augmentation during training includes random horizontal and vertical flips, rotation (± 10), and color jittering (brightness, contrast, saturation, hue).

Orion target channels undergo the global quantile scaling and log-transform described in Section 4.3. This normalization ensures consistent intensity distributions across channels despite their heterogeneous dynamic ranges, while the log-transform emphasizes low-intensity signals that may otherwise be dominated by bright markers during training.

5.3 Learning Framework and Model Architecture

We employ an encoder-decoder architecture combining a hierarchical vision transformer encoder with a feature pyramid decoder (Figure **??**).

Encoder: The Swin Transformer (swin_tiny_patch4_window7_224) provides the encoding backbone Liu et al. (2021). Swin Transformers partition the input image into non-overlapping patches and apply self-attention within local windows, achieving linear computational complexity with respect to image size. The hierarchical design produces multi-scale features at 1/4, 1/8, 1/16, and 1/32 of input resolution. We utilize pretrained ImageNet weights for initialization.

Decoder: A lightweight feature pyramid network (FPN) aggregates multi-scale encoder features. Lateral connections project encoder features to a common channel dimension (192 channels) via 1×1 convolution. Features are progressively upsampled using bilinear interpolation and combined with skip connections. The final decoder block reduces channels by half before the output projection.

Output: A 1×1 convolution projects decoder features to 20 channels, followed by Softplus activation ($\beta = 1.0$) to ensure non-negative predictions in the log-intensity domain.

The architecture is designed to capture both local morphological features (cell nuclei, tissue texture) and global contextual patterns (tissue architecture, spatial relationships) relevant to marker expression prediction.

5.4 Training Strategy and Validation Protocol

Data splitting: Paired H&E and Orion cores are randomly split into training and validation sets with a 80/20 ratio. Splitting is performed at the core level to prevent data leakage between spatially adjacent patches.

Patch sampling: Training patches of size 224×224 pixels are extracted using stratified sampling that favors regions containing positive marker expression. For each training iteration, a target channel is selected according to channel-specific sampling probabilities inversely proportional to marker coverage:

$$p_c \propto (\text{coverage}_c + \epsilon)^{-\alpha} \cdot \left(\bar{I}_c + \epsilon\right) \tag{3}$$

where coverage_c denotes the fraction of pixels exceeding threshold for channel c, $\bar{I}_c$ is the mean intensity, and $\alpha = 1.0$ controls sampling temperature. Patches are accepted if they contain sufficient positive pixels for the targeted channel, with multiple resampling attempts for rare markers.

Optimization: Models are trained using AdamW optimizer with learning rate 3×10^{-4}, weight decay 10^{-4}, and gradient clipping at norm 1.0. Learning rate scheduling employs cosine annealing with optional linear warmup. Mixed-precision training (FP16) is enabled via automatic mixed precision (AMP) for computational efficiency.

Distributed training: Multi-GPU training utilizes PyTorch DistributedDataParallel (DDP) with NCCL backend. Global quantile normalization parameters are computed on rank 0 and broadcast to all processes via filesystem synchronization.

5.5 Handling Rare and Ambiguous Cell Phenotypes

The extreme class imbalance inherent in multiplexed imaging—where rare immune populations may comprise less than 0.1% of tissue area—presents substantial challenges for model training. We address this through multiple complementary strategies:

Adaptive sampling: The channel-aware sampling procedure described above ensures that each training batch contains examples of rare marker expression. Channels with highest speckle scores (quantified by coefficient of variation) receive additional resampling budget.

Weighted loss: Pixel-level loss weights emphasize positive regions:

$$w_c(u, v) = 1 + \gamma \cdot \mathbb{1}\{y_c(u, v) > \tau\} \tag{4}$$

where $\gamma = 3.0$ provides boosted weight for pixels exceeding threshold $\tau = 0.10$. Channel-specific loss weights inversely proportional to coverage further emphasize rare markers.

Presence auxiliary task: An auxiliary classification head predicts binary marker presence (maximum intensity above threshold) for each channel, providing gradient signal even when pixel-level predictions are imprecise for very sparse markers.

6 Loss Function Design

The training objective comprises multiple complementary terms addressing different aspects of prediction quality:

$$\mathcal{L} = \lambda_{\text{rec}}\mathcal{L}_{\text{rec}} + \lambda_{\text{cov}}\mathcal{L}_{\text{cov}} + \lambda_{\text{ssim}}\mathcal{L}_{\text{ssim}} + \lambda_{\text{tv}}\mathcal{L}_{\text{tv}} + \lambda_{\text{pres}}\mathcal{L}_{\text{pres}} \tag{5}$$

Center-window reconstruction loss ($\mathcal{L}_{\text{rec}}$): Pixel-wise reconstruction error is computed within a central window to avoid boundary artifacts from convolutional padding:

$$\mathcal{L}_{\text{rec}} = \frac{1}{|\Omega|} \sum_{c=1}^{C} \sum_{(u,v)\in\Omega} w_c(u, v) \cdot |y_c(u, v) - \hat{y}_c(u, v)|^p \tag{6}$$

where Ω denotes the central 12×12 pixel window, $w_c(u, v)$ is the adaptive weight, and $p \in \{1, 2\}$ selects L1 or L2 loss. L1 loss provides robustness to outliers while L2 loss encourages smooth predictions.

Coverage loss ($\mathcal{L}_{\text{cov}}$): Per-channel mean intensity matching ensures global marker abundance is preserved:

$$\mathcal{L}_{\text{cov}} = \sum_{c=1}^{C} |\bar{y}_c - \bar{\hat{y}}_c| \tag{7}$$

where $\bar{y}_c$ and $\bar{\hat{y}}_c$ denote spatial means. Channel weights can be applied to emphasize rare markers.

Multi-scale structural similarity ($\mathcal{L}_{\text{ssim}}$): MS-SSIM loss encourages structural correspondence between predicted and target images:

$$\mathcal{L}_{\text{ssim}} = 1 - \text{MS-SSIM}(\mathbf{Y}, \hat{\mathbf{Y}}) \tag{8}$$

This perceptual loss complements pixel-wise terms by capturing spatial patterns at multiple scales.

Total variation regularization ($\mathcal{L}_{\text{tv}}$): TV regularization discourages hallucinated noise and speckle artifacts:

$$\mathcal{L}_{\text{tv}} = \sum_{c=1}^{C} (\|\nabla_h \hat{y}_c\|_1 + \|\nabla_v \hat{y}_c\|_1) \tag{9}$$

where ∇_h and ∇_v denote horizontal and vertical gradient operators.

Presence loss ($\mathcal{L}_{\text{pres}}$): Binary classification of marker presence:

$$\mathcal{L}_{\text{pres}} = \sum_{c=1}^{C} \text{BCE}\left(\sigma\left(\frac{\hat{y}_c^{\max} - \tau}{T}\right), \mathbb{1}\{y_c^{\max} > \tau\}\right) \tag{10}$$

where $\hat{y}_c^{\max}$ is the maximum predicted intensity for channel c, σ is the sigmoid function, and T is a temperature parameter. This auxiliary objective provides explicit supervision for rare marker detection.

Default loss weights are $\lambda_{\text{rec}} = 1.0$, $\lambda_{\text{cov}} = 0.1$, $\lambda_{\text{ssim}} = 0.15$, $\lambda_{\text{tv}} = 10^{-4}$, $\lambda_{\text{pres}} = 0.25$.

7 Experimental Design

This section will describe the experimental evaluation protocol including: dataset partitioning, evaluation metrics (per-channel PSNR, SSIM, correlation), baseline comparisons, cross-validation strategy, and ablation studies. Quantitative results will be added upon completion of systematic evaluation.

8 Results

This section will present quantitative and qualitative results including: (1) per-channel prediction accuracy metrics, (2) comparison of Swin-UNet vs. ConvNeXt encoder variants, (3) ablation studies on loss components and sampling strategies, (4) visualization of predicted vs. ground truth marker expression, and (5) analysis of failure modes and challenging cases. Results will be populated following systematic evaluation.

9 Discussion

9.1 Summary of Approach

We have presented HExIF, a comprehensive computational framework for predicting multiplex immunofluorescence from standard H&E histopathology images. The approach integrates automated image registration, tissue segmentation, channel-aware training strategies, and multi-component loss functions designed to address the unique challenges of sparse, heterogeneous marker expression patterns.

The 20-marker Orion panel employed in this study enables detailed characterization of the tumor immune microenvironment, with particular depth in macrophage subpopulation diversity. By simultaneously predicting all markers, the framework preserves spatial relationships and co-expression patterns that are essential for downstream phenotyping and spatial analysis.

9.2 Methodological Considerations

Several design choices warrant discussion:

Encoder architecture: The Swin Transformer backbone was selected for its ability to capture long-range spatial dependencies through hierarchical self-attention, which may be particularly relevant for learning relationships between distant tissue structures. The comparison with convolutional alternatives (ConvNeXt) will inform architecture selection for different marker types.

Loss function complexity: The multi-component loss incorporates five distinct terms, each addressing specific failure modes. While this complexity risks over-tuning, ablation studies will quantify the contribution of each component. The auxiliary presence loss is particularly relevant for rare markers where pixel-level accuracy may be limited.

Normalization strategy: Global quantile scaling with log-transform provides a practical solution for intensity heterogeneity but assumes consistent staining across the training set. Extension to multi-site or multi-batch data will require more sophisticated normalization approaches.

9.3 Limitations

Several limitations should be acknowledged:

- **Training data requirements**: The method requires spatially-registered H&E and multiplexed imaging pairs, limiting training to cohorts where both modalities are available on serial sections.

- **Marker panel specificity**: The trained model predicts the specific Orion panel used for training. Extension to different panels requires retraining with corresponding data.
- **Registration accuracy**: Prediction quality depends on registration accuracy. Residual misalignment introduces noise that may particularly affect sparse markers.
- **Generalization**: Performance on tissue types, disease contexts, or scanner configurations not represented in training data is unknown and will require prospective evaluation.

9.4 Biological Implications

Virtual multiplexing has potential applications across several areas:

Retrospective analysis: Millions of archival H&E specimens lack matched multiplexed imaging. Virtual staining could enable retrospective spatial biomarker analysis in large cohorts with clinical outcome data.

Prioritization for multiplexed imaging: Predicted marker patterns could guide selection of cases or regions for subsequent experimental validation with actual multiplexed staining.

Rapid screening: In clinical contexts where multiplexed imaging turnaround time is prohibitive, virtual predictions could provide preliminary spatial biomarker estimates to inform initial decision-making.

9.5 Future Directions

Several extensions are planned:

- Integration of CODEX data to enable cross-platform training and validation
- Uncertainty quantification to identify predictions requiring experimental confirmation
- Extension to whole-slide inference with efficient tiling strategies
- Development of downstream analysis pipelines for cell phenotyping and spatial statistics
- Evaluation on independent cohorts with clinical outcome correlation

10 Conclusion

We have developed HExIF, an end-to-end framework for predicting 20-channel multiplex immunofluorescence from hematoxylin and eosin-stained tissue images. The approach addresses key challenges in virtual multiplexing through automated registration, channel-aware training strategies, and multi-component loss functions designed for sparse marker expression. While quantitative evaluation is ongoing, the methodology establishes a foundation for virtual spatial biomarker analysis that may expand the accessibility and applicability of multiplexed tissue imaging.

Supplementary information. Supplementary materials will include: (1) detailed hyperparameter configurations, (2) extended marker panel characterization, (3) additional visualization examples, and (4) code availability information.

Acknowledgements. [Acknowledgements to be added]

Declarations

- **Funding**: [Funding information to be added]
- **Conflict of interest**: The authors declare no competing interests.
- **Ethics approval**: [IRB/ethics information for tissue specimens to be added]
- **Data availability**: [Data availability statement to be added]
- **Code availability**: Code will be made available upon publication at [repository URL to be added].
- **Author contributions**: R.S. developed the computational framework and performed experiments. A.G. supervised the project and provided guidance. All authors contributed to manuscript preparation.

References

Andani S, Yun T, Moyes A, et al (2024) HistoPlexer: Histopathology-based protein multiplex generation using deep learning. bioRxiv https://doi.org/10.1101/2024.01.09.574826, preprint

Angelo M, Bendall SC, Finck R, et al (2014) Multiplexed ion beam imaging of human breast tumors. Nature Medicine 20(4):436–442. https://doi.org/10.1038/nm.3488

Baddeley A, Rubak E, Turner R (2015) Spatial Point Patterns: Methodology and Applications with R. CRC Press

Barua S, Solis L, Parra ER, et al (2018) A functional spatial analysis platform for discovery of immunological interactions predictive of low-grade to high-grade transition of pancreatic intraductal papillary mucinous neoplasms. Cancer Informatics 17:1176935118782880. https://doi.org/10.1177/1176935118782880

Bera K, Schalper KA, Rimm DL, et al (2019) Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. Nature Reviews Clinical Oncology 16(11):703–715. https://doi.org/10.1038/s41571-019-0252-y

Bill R, Wirapati P, Messemaker M, et al (2023) CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control tumor immunity. Science 381(6657):515–524. https://doi.org/10.1126/science.ade2292

Binnewies M, Roberts EW, Kersten K, et al (2018) Understanding the tumor immune microenvironment (TIME) for effective therapy. Nature Medicine 24(5):541–550. https://doi.org/10.1038/s41591-018-0014-x

Borovec J, Munoz-Barrutia A, Kybic J (2018) Benchmarking of image registration methods for differently stained histological slides. In: 2018 IEEE International Conference on Image Processing (ICIP), IEEE, pp 3368–3372, https://doi.org/10.1109/ICIP.2018.8451318

Burlingame EA, Margolin AA, Gray JW, et al (2020) SHIFT: speedy histological-to-immunofluorescent translation of a tumor signature enabled by deep learning. Scientific Reports 10(1):17507. https://doi.org/10.1038/s41598-020-74500-3

Campanella G, Hanna MG, Geneslaw L, et al (2019) Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. Nature Medicine 25(8):1301–1309. https://doi.org/10.1038/s41591-019-0508-1

Cassetta L, Pollard JW (2020) Tumor-associated macrophages. Current Biology 30(15):R830–R838. https://doi.org/10.1016/j.cub.2020.06.075

Chen RJ, Lu MY, Weng WH, et al (2021) Multimodal co-attention transformer for survival prediction in gigapixel whole slide images. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp 4015–4025, https://doi.org/10.1109/ICCV48922.2021.00398

Cheng S, Li Z, Gao R, et al (2021) A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell 184(3):792–809.e23. https://doi.org/10.1016/j.cell.2021.01.010

Echle A, Rindtorff NT, Brinker TJ, et al (2021) Deep learning in cancer pathology: a new generation of clinical biomarkers. British Journal of Cancer 124(4):686–696. https://doi.org/10.1038/s41416-020-01122-x

Fu Y, Jung AW, Torne RV, et al (2020) Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. Nature Cancer 1(8):800–810. https://doi.org/10.1038/s43018-020-0085-8

Gatenbee CD, Baker AM, Prabhakaran S, et al (2023) Virtual alignment of pathology image series for multi-gigapixel whole slide images. Nature Communications 14(1):4502. https://doi.org/10.1038/s41467-023-40218-9

Giesen C, Wang HAO, Schapiro D, et al (2014) Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nature Methods 11(4):417–422. https://doi.org/10.1038/nmeth.2869

Goltsev Y, Nolan GP (2023) CODEX multiplexed tissue imaging. Nature Reviews Methods Primers 3:13. https://doi.org/10.1038/s43586-023-00196-w

Goltsev Y, Samusik N, Kennedy-Darling J, et al (2018) Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. Cell 174(4):968–981.e15. https://doi.org/10.1016/j.cell.2018.07.010

de Haan K, Zhang Y, Zuckerman JE, et al (2021) Deep learning-based transformation of H&E stained tissues into special stains. Nature Communications 12(1):4884. https://doi.org/10.1038/s41467-021-25221-2

Jackson HW, Fischer JR, Zanotelli VRT, et al (2020) The single-cell pathology landscape of breast cancer. Nature 578(7796):615–620. https://doi.org/10.1038/s41586-019-1876-x

Johnson JM, Khoshgoftaar TM (2019) Survey on deep learning with class imbalance. Journal of Big Data 6(1):27. https://doi.org/10.1186/s40537-019-0192-5

Lewis SM, Asselin-Labat ML, Nguyen Q, et al (2021) Spatial omics and multiplexed imaging to explore cancer biology. Nature Methods 18(9):997–1012. https://doi.org/10.1038/s41592-021-01203-6

Liu Z, Lin Y, Cao Y, et al (2021) Swin Transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp 10012–10022, https://doi.org/10.1109/ICCV48922.2021.00986

Macenko M, Niethammer M, Marron JS, et al (2009) A method for normalizing histology slides for quantitative analysis. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, pp 1107–1110, https://doi.org/10.1109/ISBI.2009.5193250

Marusyk A, Janiszewska M, Polyak K (2020) Intratumor heterogeneity: the Rosetta Stone of therapy resistance. Cancer Cell 37(4):471–484. https://doi.org/10.1016/j.ccell.2020.03.007

Marx V (2021) Method of the Year: spatially resolved transcriptomics. Nature Methods 18(1):9–14. https://doi.org/10.1038/s41592-020-01033-y

Moldoveanu D, Ramsay L, Labert M, et al (2022) Spatially mapping the immune landscape of melanoma using imaging mass cytometry. Science Immunology 7(70):eabi5072. https://doi.org/10.1126/sciimmunol.abi5072

Mulder K, Patel AA, Kong WT, et al (2021) Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. Immunity 54(8):1883–1900.e5. https://doi.org/10.1016/j.immuni.2021.07.007

Naik N, Madani A, Esteva A, et al (2020) Deep learning-enabled breast cancer hormonal receptor status determination from base-level H&E stains. Nature Communications 11(1):5727. https://doi.org/10.1038/s41467-020-19334-3

Palla G, Spitzer H, Klein M, et al (2022) Squidpy: a scalable framework for spatial omics analysis. Nature Methods 19(2):171–178. https://doi.org/10.1038/s41592-021-01358-2

Pati P, Jaume G, Foncubierta-Rodríguez A, et al (2022) Hierarchical graph representations in digital pathology. Medical Image Analysis 75:102264. https://doi.org/10.1016/j.media.2021.102264

Peranzoni E, Lemoine J, Vimeux L, et al (2018) Macrophages impede CD8 T cells from reaching tumor cells and limit the efficacy of anti–PD-1 treatment. Proceedings of the National Academy of Sciences 115(17):E4041–E4050. https://doi.org/10.1073/pnas.1720948115

Qi J, Sun H, Zhang Y, et al (2022) Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts and SPP1+ macrophages in colorectal cancer. Nature Communications 13(1):1742. https://doi.org/10.1038/s41467-022-29366-6

Qian BZ, Pollard JW (2010) Macrophage diversity enhances tumor progression and metastasis. Cell 141(1):39–51. https://doi.org/10.1016/j.cell.2010.03.014

RareCyte, Inc. (2024) Orion instrument platform: Technical documentation. https://rarecyte.com/orion/, accessed: 2024

Rivenson Y, Wang H, Wei Z, et al (2019) Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. Nature Biomedical Engineering 3(6):466–477. https://doi.org/10.1038/s41551-019-0362-y

Rivenson Y, de Haan K, Wallace WD, et al (2020) Emerging advances to transform histopathology using virtual staining. BME Frontiers 2020:9647163. https://doi.org/10.34133/2020/9647163

Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer, pp 234–241, https://doi.org/10.1007/978-3-319-24574-4_28

Schürch CM, Bhate SS, Barlow GL, et al (2020) Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. Cell 182(5):1341–1359.e19. https://doi.org/10.1016/j.cell.2020.07.005

Stringer C, Wang T, Michaelos M, et al (2021) Cellpose: a generalist algorithm for cellular segmentation. Nature Methods 18(1):100–106. https://doi.org/10.1038/s41592-020-01018-x

Tan WCC, Nerurkar SN, Cai HY, et al (2020) Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. Cancer Communications 40(4):135–153. https://doi.org/10.1002/cac2.12023

Thorsson V, Gibbs DL, Brown SD, et al (2018) The immune landscape of cancer. Immunity 48(4):812–830.e14. https://doi.org/10.1016/j.immuni.2018.03.023

Xu Y, Jia Z, Wang LB, et al (2017) Large scale tissue histopathology image classification, segmentation, and visualization via deep learning. BMC Bioinformatics 18(1):281. https://doi.org/10.1186/s12859-017-1685-x

Zarella MD, Bowman D, Aeffner F, et al (2019) A practical guide to whole slide imaging: a white paper from the Digital Pathology Association. Archives of Pathology & Laboratory Medicine 143(2):222–234. https://doi.org/10.5858/arpa.2018-0343-RA

Zhang L, Li Z, Skrber KM, et al (2020a) Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. Cell 181(2):442–459.e29. https://doi.org/10.1016/j.cell.2020.03.048

Zhang Y, de Haan K, Rivenson Y, et al (2020b) Digital synthesis of histological stains using micro-structured and multiplexed virtual staining of label-free tissue. Light: Science & Applications 9(1):78. https://doi.org/10.1038/s41377-020-0315-y