

Computational Anatomy And AI For Quantifying Tumor-Immune Spatial Geometry From Medical Images And Predicting Treatment Sensitivity

Ashok Kumar Dogra^{1*}, Disha Anoop Nagpure², Swapna Babu Budala³, Dr. Smriti Tiwari⁴,
Dr. Sana Fathima⁵, Dr Kumar Sambhav⁶, Avijit Kumar Chaudhuri⁷

^{1*}Assistant professor, Department of Biochemistry, Specialization in oncology,
Government Medical College, Srinagar, Jammu and Kashmir

²HOD(AIML department Alard college of Engineering, Pune), Department of AIML,
Specialization in Machine Learning, Alard college of Engineering, Pune, Pin code:
411057, India, Orcid Id: 0009-0008-95035367

³Assistant Professor, Department Of CSE-AIML, Aditya University, Surampalem,
Kakinada District, Andhra Pradesh, Pin Code: 533437, India, Orcid Id:0009-0007-8651-
8415

⁴Lecturer, Department of Pharmacy Practice, Teerthanker Mahaveer College of
Pharmacy, TMU University, Moradabad Orcid Id: 0009-0008-1818-0560

⁵Department of Histopathology, Specialization in MD PATHOLOGY, Bengaluru, Pin code :
560001, India, Orcid Id: 0000-0001-7614-9483

⁶Assistant Professor, Department of Anatomy, MD(Anatomy), All India Institute of
Medical Sciences, Bilaspur, Himachal Pradesh Orcid Id: 0000-0003-0012-3994

⁷Professor & HOD, Department of Computer Science & Engineering, Brainware
University, Barasat, Kolkata – 125, West Bengal, India, Pin code: 700125, Orcid Id: 0000-
0002-5310-3180

Abstract

The spatial organization of tumor and immune cells within the tumor microenvironment has emerged as a critical determinant of therapeutic response and clinical outcome. Advances in digital pathology, multiplex imaging, and radiological modalities now enable high-resolution acquisition of spatially resolved tumor-immune architecture. This review synthesizes recent progress in computational anatomy and artificial intelligence (AI) frameworks for quantifying tumor-immune spatial geometry from medical images and translating these representations into predictive models of treatment sensitivity. The image preprocessing and normalization strategies, segmentation and instance detection methods, and geometric modeling approaches including spatial point processes, graph-based representations, and topological descriptors. Modern AI architectures convolutional neural networks, graph neural networks, transformers, and multimodal foundation models for spatial feature learning and outcome prediction were further analyzed. Comparative evaluation highlights methodological trends, trade-offs in scalability and interpretability, and reproducibility challenges across studies. Persistent limitations include domain shift, annotation variability, computational burden of gigapixel whole-slide images, and the absence of standardized benchmarking protocols. Emerging directions such as self-supervised learning, federated training, and multiscale representation alignment are discussed as pathways toward robust and generalizable spatial analytics. Collectively, computational anatomy integrated with AI-driven spatial modeling provides a scalable framework for objective, quantitative assessment of tumor-immune interactions and precision prediction of treatment sensitivity from medical images.

Keywords: Computational anatomy; Tumor-immune spatial geometry; Medical image analytics; Graph neural networks; Treatment sensitivity prediction.

1. INTRODUCTION

Spatial arrangement of tumor and immune cells in the tumor microenvironment (TME) has become a key determinant of cancer progression, therapeutic response, and survival in patients. Although initially studies in oncology focused on tumor-intrinsic genomic changes, it is now common knowledge that cancer is a spatially organized ecosystem in which malignant cells, immune populations, stromal components and vascular networks interact within complex architectural niches. Functional information hidden in the relative positioning, patterns of clustering and neighborhood interactions between tumor and immune cells cannot be measured by bulk molecular measurements or simple cell-density parameters [1]. As a result, the tumor-immune spatial geometry has become the center of interest as a biological characterization and a computational object of quantitative modeling [2].

Although it is the gold standard of cancer diagnosis, traditional histopathological analysis is largely qualitative and semi-quantitative. Morphological features which can be analyzed using naked eye evaluation by trained pathologists, such as lymphocytic infiltration or immune exclusion, are subjective, with limited reproducibility and limited to the analysis of higher-order spatial interactions [3]. Geometric forms, interaction networks and spatial heterogeneity of tissue areas are not reflected in aggregate descriptions of measures such as immune cell counts, or tumor proportion scores. This mismatch of visual interpretation and quantitative characterization of space is what has given rise to the idea of computationalism in digital processes of pathology.

The latest advances in spatial profiling technology, including multiplex immunofluorescence, imaging mass cytometry and spatial transcriptomics, have drastically expanded the level according to which tumor-immune interactions can be inquired [4-5]. High dimensional image data is generated through these platforms in which every spatial coordinate has the capacity to encode phenotypic or molecular properties. As a result, tissue sections may be represented in the form of structured data object of pixel intensities, cell coordinates, molecular expression vectors and neighborhood graphs. Recent development of digital, high-throughput imaging of analog microscopy has offered a unique opportunity to apply computational anatomy, signal processing and artificial intelligence (AI) to provide a systematic quantification of spatial organization.

Computational anatomy provides a mathematical framework of structural variability and spatial relations of images data. It was originally developed to be applied to macroscopic anatomical experiments, but is now applied at the microarchitectural scale to characterize the topology of tissues, their deformation behaviour and their spatial heterogeneity. Together with image processing pipelines such as normalization, segmentation, and feature extraction computational anatomy can be used to study tumor-immune ecosystems using formal geometric measures, e.g. point processes, graphs and topological invariants. These models convert visual tissue structure to reproducible and measurable characteristics.

Artificial intelligence and, more specifically, deep learning have also contributed to this paradigm shift. The convolutional neural networks (CNNs), graph neural networks (GNNs), transformer networks, and weakly supervised learning models have demonstrated high capability in cellular structure segmentation, learning spatial features, and correlating features generated on images with survival or treatment response [6]. Unlike manual descriptors, AI-based models are able to learn hierarchical representations that experience local cellular morphology and distant contextual connections. This is particularly so concerning the whole-slide imaging (WSI), where data of gigapixels in size require scalable computing solutions to preserve the multi-scale spatial information.

The situation with methodology remains heterogeneous, despite the fast advancement. The imaging modalities differ in their spatial resolution, signal characteristics and dimensions, preprocessing pipelines are more or less robust to stain variations and batch effects, spatial modeling schemes are classical statistical point processes or deep learning graph structures. Moreover, the predictive models have different ways of synthesizing the spatial images and clinical outcomes. The computational methods must be organized in an organized manner so as to explain the trends in the methodology, highlight the strengths and the weaknesses therein, and envision the prospects of standardization and innovation. This review aims to:

1. Systematically review imaging modalities and computational image processing tools that we can use to quantify tumor-immune spatial organization of medical images.
2. Critically evaluate computational anatomy, geometric modeling strategies, and AI architectures for spatial feature extraction and representation learning.
3. Analyze predictive modeling frameworks that leverage spatial image-derived representations to estimate treatment sensitivity and clinical outcomes, while identifying technical challenges and future research directions in spatial image analytics.

2. IMAGING MODALITIES FOR TUMOR-IMMUNE SPATIAL ANALYSIS

The spatial geometry of tumor-immune interaction can be properly quantified based on the imaging modality that is employed to obtain the tissue information. The modalities vary in spatial resolution, dimensionality, signal properties, and artifact sensitivity, and have a direct impact on the preprocessing, segmentation and spatial modeling approaches. Whole-slide images in gigapixels, multiplex molecular imaging and radiological scans can all be used to capture tumor-immune interactions at a variety of spatial scales. To choose the right computational pipelines, one thus needs to know the strengths and weaknesses of each modality.

2.1 Digital Pathology (Whole-Slide Imaging)

Computational tumor-immune analysis is mostly based on whole-slide imaging (WSI). Scanned slides with magnifications of 20x-40x produce gigapixel images that maintain cellular morphology which can be used to assess tumor regions, immune infiltration, stromal compartments and vascular structures at the cell level. Nonetheless, WSIs are too large to process using patch-based processing and memory-efficient architectures, and require multi-resolution modeling. The standard CNNs use fixed-size inputs, which do not capture the global context and hence hierarchical aggregation and attention-based mechanisms are frequently employed to learn the long-range dependencies.

A significant problem is stain variability. Variations in staining procedures, scanner calibration, and laboratory procedures introduce inter-institutional variation which may bias segmentation and learned representations. Normalization and domain adaptation methods are thus essential towards reproducibility. Notably, after segmentation, WSIs give cellular coordinates that are explicit and allow graph construction and spatial point process modeling.

2.2 Multiplex and Spatially Resolved Imaging

Multiplex imaging Multiplex immunofluorescence (mIF) and imaging mass cytometry (IMC) are multimodal technologies that build upon the conventional histopathology by quantifying a range of protein markers in the same section. These images in multi-channels permit accurate phenotyping of immune subpopulations, with localization of space.

Making multiplex datasets Multiplex datasets can be considered high-dimensional image tensors, with each spatial location storing a vector of marker intensities. Although this makes biological resolution more accurate, it also makes it more difficult to calculate. Spectral unmixing and normalization are necessary when there is channel cross-talk, fluorescence decay, and signal noise. The large dimensions also require pre-predictive modeling feature reduction or multimodal fusion. Spatial transcriptomics introduces gene expression data to spatial coordinates, which connects molecular data with histological data and allows morpho-molecular simulations of tumor-immune interactions.

2.3 Radiological Imaging (MRI, CT, PET)

Non-invasive and three-dimensional representations of tumors on an organ and whole-body basis are available with radiological modalities like MRI, CT, and PET. They obtain macroscopic heterogeneity, vascularization, metabolic activity, and longitudinal treatment response [7]. Radiomics is a method that derives quantitative features such as intensity, texture, wavelet and shape descriptors of voxel representations. Radiological modeling is also regional and organ-scale as opposed to cell-level pathology analysis. Multi-scale analysis using radiology and histopathology is aided by integration of both techniques to be able to correlate the macroscopic tumor structure with the microscopic immune architecture.

2.4 Signal Characteristics and Image Artifacts

Signal variability and artifacts affect the quantitative spatial analysis across modalities. In fluorescence imaging, photon shot noise is a noise source, in pathology, scanner variability is a noise source and in radiology, motion or partial volume is a noise source. Noise modeling gives information to denoising strategies and robustness of segmentation.

Multi-center datasets have batch effects that cause systematic changes in intensities that may result in domain bias during AI training. Normalization and harmonization methods are hence necessary. Moreover, the heterogeneity of intensity that may be biological or technical makes the segmentation and feature extraction problematic and necessitates powerful preprocessing (Figure 1). Taken together, these modalities provide complementary perspectives of tumor-immune organization at spatial scales. Their specific computational limitations are direct sources of preprocessing design, segmentation plan, geometric modeling, and AI-based prediction frameworks described in the following sections.

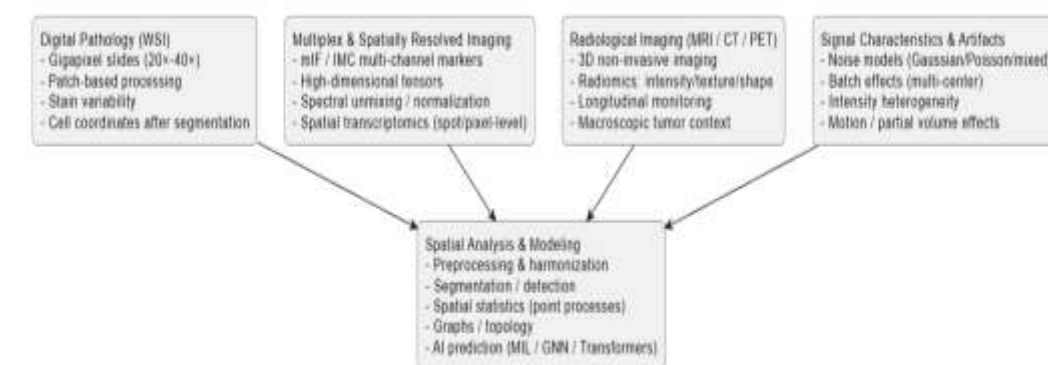


Figure 1: Imaging Modalities for Tumor-Immune Spatial Analysis

3. IMAGE PREPROCESSING AND ENHANCEMENT TECHNIQUES

Preprocessing of tumor-immune spatial geometry is crucial to ensure good quantification of the geometry, which involves stabilisation of intensity distributions, noise reduction, spatial misalignment correction and scalable computation on high-resolution images. Variability introduced during staining, scanning and acquisition may cause segmentation, spatial feature extraction and downstream predictive modeling bias. Preprocessing is thus a mathematically based step of the computational pipeline, which has a direct impact on feature fidelity, reproducibility and model generalization.

3.1 Color and Intensity Normalization

Normalization of color and intensity is designed to minimize inter-slide and inter-institution variability without compromising structural information. Stain variability in digital pathology distorts the intensity distributions of pixels and may impair the accuracy of segmentation and the robustness of deep learning models when not corrected. Normalization methods are usually divided into histogram-based, stain-deconvolution based, and deep learning based methods [8]. GAN-based approaches learn nonlinear relationships between domains of stains to minimize cross-site variance [9], whereas attention-directed residual networks focus on structure conservation in transformation.

In radiological imaging, statistical normalization like z-score scaling is commonly used to normalize the dynamic ranges of scans. Even though normalization enhances the cross-dataset generalization, over-transformation can either mask biologically relevant heterogeneity or artificial uniformity. There is thus a need to balance invariance and biological fidelity by using structure preserving strategies.

3.2 Denoising and Artifact Removal

Denoising and artifact elimination reduce the effects of degradation due to imaging systems and acquisition conditions. The classical methods, such as Gaussian filtering, median filtering, and non-local means, are still common in the medical image processing. Nonetheless, recent computational pathology work focuses on maintaining spatial relations in the process of enhancement to prevent the distortion of downstream geometric modeling.

The properties of noise are varied in various modalities, and they are often assumed to be of a Gaussian form in digital scanning and Poisson-type in fluorescence imaging. The balancing between signal amplification and edge preservation of the denoising should be effective since too much smoothing can result in loss of cell boundary delineation and loss of segmentation accuracy. State of the art preprocessing systems involve artifact correction in a system of spatial analysis pipeline, such that it is consistent with later modeling objectives.

3.3 Image Registration and Alignment

Image registration is applied to give spatial correspondence to images taken at dissimilar times, magnifications or modalities. Particularly, it is important in multiplex imaging and spatial transcriptomics where numerous channels or molecular layers must be correctly registered to maintain sound spatial coordinates [10]. Registration frameworks approximate rigid and affine or deformable transformations by optimizing measures of similarity e.g. mutual information.

The fact that the information is corrected to construct reliable cell interaction graphs and spatial neighborhood models is crucial especially when using multi-marker/ multi-section data [11]. Misregistration can introduce excess spatial offsets which propagate

errors to graph-based analysis or topological analysis which is why it is important to ensure that the alignment strategies are robust.

3.4 Patch Extraction and Multi-Resolution Processing

Whole-slide images are often larger than gigapixels and such large images inflexible from an end-to-end processing perspective. Patch-based strategies divide slides into manageable regions, which allows scalability in feature extraction and aggregation. Weakly supervised frameworks usually execute patch-wise representation learning then attention-based aggregation to generate slide-wise prediction. Transformer-based correlated multiple instances learning - further models inter-patch dependencies in order to preserve contextual relationships

Although the use of patch extraction enhances the scalability, it creates risks of sampling bias, contextual fragmentation, and multi-scale inconsistency (Table 1). Hierarchical and multi-resolution aggregation mechanisms have therefore become central to modern spatial image analysis pipelines in order to preserve the local cell morphology as well as the global tissue architecture.

Table 1. Image Preprocessing and Enhancement Techniques in Tumor-Immune Spatial Modeling

Category	Representative Methods	Core Objective	Strengths	Limitations	Impact on Spatial Modeling	Source
Color & Intensity Normalization	Histogram matching; Stain deconvolution; GAN-based normalization; Attention-guided residual learning; Z-score standardization (radiology)	Reduce inter-slide/inter-site variability while preserving morphology	Improves cross-cohort generalization; reduces stain bias; stabilizes intensity distributions	Risk of suppressing biologically meaningful heterogeneity; may introduce artificial uniformity	Directly affects segmentation robustness and feature reproducibility; critical for domain generalization	[8]
Denoising & Artifact Removal	Gaussian filtering; Median filtering; Non-local means; Modality-specific noise modeling	Suppress acquisition noise while preserving structural boundaries	Enhances signal quality; improves segmentation performance	Over-smoothing may distort cell boundaries; modality-dependent noise assumptions	Influences boundary accuracy and downstream spatial statistics	[7]

Image Registration & Alignment	Rigid, affine, deformable transformations; Mutual information optimization	Align multi-channel, multi-section, or multi-modal images	Enables accurate multi-marker integration; preserves spatial correspondence	Sensitive to misalignment errors; computationally intensive for deformable models	Essential for constructing valid spatial graphs and neighborhood relationships	[10]
Patch Extraction & Multi-Resolution Processing	Patch-based tiling; Attention-based aggregation; Correlated MIL; Hierarchical modeling	Enable scalable computation for gigapixel images	Memory-efficient; scalable; preserves local features	Risk of contextual fragmentation; sampling bias; multi-scale inconsistency	Critical for balancing local morphology and global spatial context	[30]
Batch Effect Correction & Harmonization	Domain adaptation; Cross-site normalization; Statistical harmonization	Mitigate systematic intensity shifts across institutions	Improves reproducibility and external validation	May not fully eliminate non-iid bias; complex to standardize	Central for robust cross-cohort predictive modeling	[9]

4. SEGMENTATION AND CELLULAR DETECTION METHODS

Accurate segmentation is fundamental for tumor-immune spatial modeling as downstream spatial statistics, graph construction and interaction analysis requires accurate delineation of nuclei, cells and tissue compartments. Errors at this stage are directly propagated into spatial features, making robustness across stains, scanners and tissue types essential.

4.1 Classical Image Processing Approaches

Classical methods of segmentation are based on preconceptual assumptions regarding intensity, edges and morphology. Thresholding methods are used to make cuts between tissue regions using either a global or adaptive cutoff, whereas watershed techniques are meant to represent intensity maps as topographic surfaces, and are used to partition touching nuclei. Morphological erosion, dilation, opening, and closing are morphological operations to refine binary masks and impose shape constraints on them. These methods are computationally inexpensive, interpretable and vulnerable to stain variations, non-homogenous textures and overlap of the nuclei. Tuning of parameters by hand further restricts scalability and cross dataset reproducibility.

4.2 Deep Learning-Based Segmentation

Deep learning is the new paradigm of biomedical segmentation because of its ability to learn hierarchical representations using complex tissue data. U-Net proposed an encoder-decoder architecture that has skip connections and maintains spatial resolution [12]. Variants of atrous convolution increased the receptive fields to improve contextual modelling [13].

nnU-Net high-level architecture and preprocessing adaptation, enhancing dataset reproducibility [14]. Mass annotation has made whole cell segmentation performance almost human with the data scale and label quality being vital [15].

Although deep models are better than classical models, they are still vulnerable to variations in domains with staining and scanner properties. Dilated architectures provide more contextual capture at a greater cost in terms of memory. Self-configuring frameworks enhance standardization at a high cost.

4.3 Instance Segmentation and Cell Phenotyping

To analyze tumor-immune interactions spatially, it is better to use instance segmentation rather than semantic segmentation since the former offers object-specific masks and centroids to use in the spatial statistics and modeling of cell-cell interactions.

Mask R-CNN supports prediction and separation of overlapping nuclei by the instance level [16]. HoVer-Net collaboratively carries out nucleus segmentation and classification of multi-tissue phenotyping [17]. Star-convex polygon methods simulate nuclei geometry in clustered separation [18], and Cellpose is a generalist methodology that has been shown to work effectively on a variety of morphologies [19].

The computational cost of instance models is heavier but they are more appropriate in quantification of spatial geometry. Generalist models are more portable and domain-specific models are more accurate when trained on similar pathology data.

4.4 Evaluation Metrics

Dice coefficient and Intersection-over-Union (IoU) to evaluate the performance of performance of segmentation in accuracy of overlap, and Hausdorff distance to evaluate performance of segmentation in accuracy of boundaries. Detection of instances is especially critical to precision and recall with missed or spurious nuclei directly distorting spatial statistics.

In the context of tumor immigration modeling, it is necessary to evaluate it not only in terms of mask overlap but also, centroid accuracy, cell count accuracy, computing cost and resistance to staining variation or independent validation (Table 2). It is necessary to report these metrics in a transparent way to enable the ability to replicate and translate these metrics to clinical practice at a larger scale.

Table 2. Segmentation and Cellular Detection Approaches for Tumor-Immune Spatial Modeling

Category	Representative Methods	Core Principle	Strengths	Limitations	Relevance to Tumor-Immune Spatial Modeling	Source
Classical Image Processing	Thresholding, Watershed,	Rule-based segmentation using	Computationally efficient; interpretable;	Sensitive to stain variability; limited	Useful as baseline or preprocessing	[18]

	Morphological operations	intensity, gradient, and morphological assumptions	low computational requirement	robustness to overlapping nuclei; manual parameter tuning	step; limited reliability for large-scale spatial quantification	
Semantic Deep Learning Segmentation	U-Net; DeepLab variants; nnU-Net	Encoder-decoder architectures with hierarchical feature learning	High segmentation accuracy; strong contextual modeling; adaptable pipelines	Sensitive to domain shift; requires annotated datasets; computationally demanding	Enables robust tissue compartment and region-level segmentation for downstream spatial modeling	[14]
Large-Scale Deep Segmentation Systems	Whole-cell human-level segmentation systems	Deep learning trained on extensive annotated datasets	Near human-level boundary precision; improved morphology extraction	Requires large-scale annotation; high infrastructure demand	Improves reliability of morphology-dependent spatial descriptors	[15]
Instance Segmentation Models	Mask R-CNN; HoVer-Net; Star-convex polygons; Cellpose	Object-level detection and mask prediction	Provides cell centroids and boundaries; essential for spatial statistics and graph construction	Higher memory and computational cost; sensitive to annotation noise	Critical for tumor-immune interaction modeling and cellular network construction	[16]
Generalist vs Domain-Specific Models	Cellpose (generalist); HoVer-Net (domain-specific)	Cross-domain training vs tissue-specific optimization	Generalist: portability; Domain-specific: higher task accuracy	Generalist: lower peak performance; Domain-specific: reduced transferability	Trade-off between scalability and precision in tumor-immune spatial quantification	[17]
Evaluation Metrics	Dice, IoU, Hausdorff distance, Precision, Recall	Overlap-based and boundary-sensitive evaluation	Standardized performance comparison	Overlap metrics may not reflect centroid accuracy or spatial reliability	For spatial modeling, evaluation should include centroid fidelity, cell counts, robustness, and computational cost	[15]

5. COMPUTATIONAL ANATOMY AND SPATIAL GEOMETRY MODELING

After segmentation and cell phenotyping, tumor-immune ecosystems can be represented as structured spatial objects such as point sets, graphs, and geometric complexes enabling quantitative analysis of how immune populations are arranged relative to tumor regions and to one another. Formally, let the tissue domain be defined as $\Omega \subset \mathbb{R}^2$, and let segmented cell centroids be represented as a marked point set:

$$\mathcal{X} = \{(x_i, m_i)\}_{i=1}^N, x_i \in \Omega$$

where x_i denotes spatial coordinates and m_i represents phenotypic marks (e.g., tumor, CD8+, macrophage). Computational anatomy offers a theoretical framework to study spatial variability and structural organization as studying tissue architecture as quantifiable geometry and not merely as visual patterning. This point of view is fundamental in the tumor-immune context since the clinically relevant behavior can be a consequence of the spatial interactions, like the immune exclusion at tumor peripheries, aggregation in stromal niches, or gradients of invasion which cannot be observed in the marginal cell densities [20]. Structured spatial modeling enables mapping from $\mathcal{X}$ to geometric descriptors $\Phi(\mathcal{X})$ that summarize interaction structure.

5.1 Spatial Point Process Modeling

Spatial point process models treat cell centroids as realizations of a stochastic process defined over Ω . The first-order intensity function describes spatial density:

$$\lambda(x) = \lim_{|B| \rightarrow 0} \frac{\mathbb{E}[N(B)]}{|B|}$$

where $N(B)$ denotes the number of points in region $B \subset \Omega$.

Ripley's K-function quantifies clustering across spatial scales:

$$K(r) = \frac{1}{\lambda} \mathbb{E}[\text{number of additional points within distance } r \text{ of a typical point}]$$

Under complete spatial randomness (CSR), $K(r) = \pi r^2$. Deviations from this baseline indicate aggregation or inhibition.

The pair-correlation function provides a scale-resolved interaction measure:

$$g(r) = \frac{1}{2\pi r} \frac{dK(r)}{dr}$$

Values $g(r) > 1$ indicate clustering, while $g(r) < 1$ indicates spatial inhibition. The functions are especially useful in the measurement of immune clustering and tumor-immune mixing. Nevertheless, spatial inhomogeneity and boundary corrections have to be considered in modeling in order to prevent biased inference of heterogeneous tissues.

5.2 Graph-Based Modeling of Cellular Interactions

Graph representations encode relational structure by defining a graph:

$$G = (V, E)$$

where nodes $V = \{v_i\}_{i=1}^N$ correspond to cells and edges $E \subseteq V \times V$ are constructed using k-nearest-neighbor or radius-based criteria:

$$E = \{(v_i, v_j) \mid \|x_i - x_j\| \leq r\}$$

The adjacency matrix $A \in \mathbb{R}^{N \times N}$ is defined as:

$$A_{ij} = \begin{cases} 1 & \text{if } (v_i, v_j) \in E \\ 0 & \text{otherwise} \end{cases}$$

The graph Laplacian is:

$$L = D - A$$

where D is the degree matrix with $D_{ii} = \sum_j A_{ij}$. Spectral properties of L characterize connectivity and diffusion-like behavior across tumor-immune networks. Centrality measures such as degree centrality:

$$C_D(v_i) = \sum_j A_{ij}$$

quantify cellular hubs, while betweenness centrality measures bridging roles in tumor-immune interaction pathways. Graph models allow integration of phenotypic marks and multi-scale relational information and directly support downstream graph learning. However, results depend strongly on edge-definition parameters and segmentation reliability.

5.3 Geometric and Topological Representations

Geometric approaches extend pairwise modeling by partitioning space and analyzing higher-order relationships. Voronoi tessellation defines regions:

$$V_i = \{x \in \Omega \mid \|x - x_i\| \leq \|x - x_j\|, \forall j \neq i\}$$

Each region V_i represents the spatial influence zone of cell i . Delaunay triangulation provides the dual graph of the Voronoi diagram and ensures stable neighborhood structure.

Topological data analysis employs a filtration function $f: \Omega \rightarrow \mathbb{R}$ to generate nested simplicial complexes $\mathcal{K}_\epsilon$, where connectivity evolves with scale parameter ϵ . Persistent homology tracks birth and death of topological features (connected components, loops):

$$\beta_k(\epsilon) = \text{rank of } H_k(\mathcal{K}_\epsilon)$$

where β_k are Betti numbers describing k -dimensional holes. These descriptors capture multi-scale structural invariants robust to small perturbations, but may introduce computational complexity in dense datasets.

5.4 Morphometric and Texture-Based Descriptors

Morphometric descriptors quantify geometric properties of segmented regions. For a region $R \subset \Omega$:

- Area: $A(R) = \int_R 1 \, dx$
- Perimeter: $P(R) = \int_{\partial R} 1 \, ds$
- Eccentricity derived from second-order central moments.

Fractal dimension D can be estimated using box-counting:

$$D = \lim_{\epsilon \rightarrow 0} \frac{\log N(\epsilon)}{\log (1/\epsilon)}$$

where $N(\epsilon)$ is the number of boxes of size ϵ needed to cover R .

Texture descriptors such as Haralick features are computed from the gray-level co-occurrence matrix $G(i, j, d)$, capturing intensity dependencies across displacement d . For example, contrast is defined as:

$$\text{Contrast} = \sum_{i,j} (i - j)^2 G(i, j)$$

While morphometric and texture descriptors are computationally efficient and interpretable, they require careful normalization to avoid redundancy and sensitivity to preprocessing variation (Table 3). When combined with spatial point process and graph-based features, they enrich the representation of tumor-immune geometry by capturing both relational structure and phenotypic morphology.

Table 3. Computational anatomy and spatial geometry modeling approaches for tumor-immune analysis.

Modeling Paradigm	Mathematical Representation	Captures	Strengths	Limitations	Typical Use Case	Source
Spatial Point Processes	$\lambda(x), K(r), g(r)$	Clustering, dispersion	Statistically grounded; scale-aware	Sensitive to spatial inhomogeneity	Immune clustering and exclusion analysis	[20]
Graph-Based Models	$G=(V,E), L=D-A$	Cell-cell interactions	Relational modeling; supports downstream GNN	Depends on edge construction rules	Tumor-immune interaction networks	[7]
Voronoi/Delaunay Geometry	Spatial partitions of Ω	Local dominance, neighborhood geometry	Interpretable spatial partitioning	Computational overhead in dense datasets	Microenvironment partitioning	[20]
Persistent Homology	Betti numbers β_k	Higher-order topology	Multiscale structural invariance	Complex parameter tuning	Global tissue architecture characterization	[20]
Morphometric Features	Area, perimeter, eccentricity	Cellular phenotype	Interpretable; low computational cost	Sensitive to segmentation errors	Tumor aggressiveness estimation	[15]
Texture Features (GLCM, Fractal)	GLCM statistics, box-counting dimension	Microarchitecture heterogeneity	Captures stromal and infiltration patterns	Redundant without normalization	Stromal organization and heterogeneity analysis	[8]

6. AI ARCHITECTURES FOR SPATIAL FEATURE LEARNING

AI architectures encode local morphology, relational structure, and long-range dependencies of tumor-immune organization to transform spatial representations of these structures into predictive embeddings. Spatial information is represented in different model families in different ways, including convolutional locality, graph-based interaction modeling, and global self-attention (Table 4). This is because their appropriateness is based on capacity to represent, scalability, interpretability and resistance to domain variability.

6.1 Convolutional Neural Networks

CNNs continue to be part of the core of extracting spatial features in medical images because of their translation-invariant receptive fields and hierarchical representation learning. CNNs are used in the tumor-immune analysis to extract multi-level features including texture, morphology and architectural context. The encoder-decoder systems like U-Net are able to produce dense spatial representations and atrous-based and attention-enhanced variants [21] are able to scale up to larger contextual receptive fields. In whole-slide imaging, CNNs are usually used on image patches, resulting in slide-level inferences as summarized embeddings.

CNNs are computationally compact, and they have developed optimization structures. But they are necessarily locally interacting and modeling of long-range interactions requires global aggregation schemes. Spatial continuity can also be broken by patch-based processing and multi-scale integration is required.

6.2 Graph Neural Networks

The graph neural networks (GNNs) are the extensions of deep learning to non-Euclidean spaces, which pass messages over graph-structured data. In tumor-immune modeling, nodes are cells, regions with phenotypic characteristics and edges are spatial closeness or forward hypothesized interactions. Graph Convolutional Networks generalize the Laplacian-based aggregation [22], Graph Attention Networks add an adaptive edge weighting [23] and inductive methods like GraphSAGE learn scalable representations [24]. Their expressive power is also further defined by theoretical analyses.

GNNs cannot help but learn the relational dependencies and heterogeneous cell types, making it possible to learn the interaction-based phenomenon, like immune clustering or tumor margin exclusion. They are more relationally interpretable than CNNs but are computationally intensive and dependent on the choices of building the graph.

6.3 Vision Transformers

ViTs are vision transformers that learn global interactions with the aid of self-attention, which calculates pairwise interactions between patch tokens. Ranked designs like Swin Transformer enhance efficiency by attention windows [25]. Whole-slide pathology involves the use of transformer-based correlated multiple instance learning models to model the spatial context and relationship of spatially remote regions.

Long-range tumor-immune interactions can be well modeled by transformers without the use of graphs, and attention maps can be interpreted to some degree. They however require a great deal of data and complexity of computing and attention may drive scalability down.

6.4 Geometric Deep Learning

Geometric deep learning represents Spatial topology as a direct part of neural architectures. In geometric deep learning, spatial topology is directly accommodated in graphs or structured domains. Graph based survival models in tumor immune modeling:

Graph-based survival models can be used both to instantiate spatial cell graphs by directly embedding them in predictive pipelines and interaction-based graph learning can explicitly use tumor immune relational structure [26].

Those methods combine the computational anatomy with the current AI through respect to spatial invariances and topology. The complexity of the model however grows with the size of the graph and large-scale datasets need effective sampling or pooling techniques.

6.5 Multi-Modal and Multi-Scale Fusion Models

Tumor immunoanalysis is becoming more and more heterogeneous, incorporating histopathology, spatial transcriptomics and radiology. Multimodal foundation models can use large-scale pretraining to learn task and resolution-rich transferable representations based on vector-quantization [27], and vector-quantization approaches can standardize and compress entire-slide embeddings [28]. Multi-scale fusion is a method which integrates patch-level and slide-level features by using hierarchical aggregation to enhance robustness, but at the cost of a more complex architecture. Such models improve the richness of the representation but interpretability in cases where such multiple modalities and scales are encoded simultaneously is difficult. Spatial encoding processes in architectures include local convolutional filters, as well as graph-based message passing and self-attention on a global scale. The model selection is made by trade-offs of representational power, computational feasibility, generalization capacity and interpretability to be used in clinical practice.

Table 4. AI architectures for spatial feature learning in tumor-immune modeling.

Architecture	Spatial Encoding Mechanism	Long-Range Context	Interpretability	Scalability	Representative References
CNN	Convolutional receptive fields	Limited to local neighborhoods	Low-Moderate	High	[17]
GNN	Message passing on graph G	Good (via relational edges)	Moderate-High	Moderate (graph size dependent)	[24]
Transformer	Self-attention $\text{softmax}(QK^T/\sqrt{d})V$	Strong global modeling	Moderate (attention maps)	Compute-intensive	[25]
MIL (Attention-based)	Attention-weighted aggregation of patch embeddings	Moderate	Moderate	High	[36]
Foundation Models	Large-scale pretraining &	Strong	Emerging	High training cost	[28]

	multimodal alignment				
--	-------------------------	--	--	--	--

7. PREDICTIVE MODELING OF TREATMENT SENSITIVITY

Tumor-immune spatial quantification becomes clinically relevant when spatial representations are translated into predictive models of treatment response, survival, or therapeutic sensitivity. In computational pathology, these frameworks map image-derived features—ranging from handcrafted spatial descriptors to learned embeddings—to patient-level outcomes. The primary focus lies in algorithmic robustness, representation learning, validation rigor, and cross-cohort generalization rather than drug-specific biology. The central challenge is constructing models that reliably associate high-dimensional spatial features with outcomes while minimizing overfitting and ensuring reproducibility.

7.1 Feature-Based Machine Learning Models

The feature-based methods work on descriptors explicitly extracted, like spatial clustering measures, graph measures, morphometric features and texture measures. They offer explainable levels of baseline and are particularly applicable in small datasets. SVMs build nonlinear decision boundaries with the help of the kernel functions, Random Forests make decisions stronger by aggregating trees together, and Gradient Boosting builds a stronger predictor by refining weak learners [29].

In tumor-immune modeling they are used to combine multi-scale spatial descriptors, e.g. immune clustering indices plus morphological variability to predict the response or survival end endpoint. They achieve high efficiency at computation time and do not consume as much data as deep networks, but their performance is highly sensitive to the quality of feature engineering. Handcrafted features can also lack fine nonlinear interaction space patterns that can be computed by learned representations.

7.2 Deep Learning Predictive Frameworks

Deep learning models are trained on direct image patch mappings, or mappings between spatial graphs, to clinical outcomes. Multiple instance learning (MIL) with weak-supervision uses slide patches as instances in a bag of labels, which allows prediction of outcomes without dense annotation. MIL models that are attention-guided have been shown to be effective at predicting survival given whole-slide images, and transformer-based correlated MIL is capable of improving the capacity to model context across spatially remote locations [30].

Graph based predictive models incorporate spatial cell interaction networks into neural networks. The initial graph convolutional survival models demonstrated that patch- or region-level graphs can represent spatial heterogeneity which is important to prognosis. Later methods represented whole-slide images as point clouds to encode contextual dependencies and interaction-directed graph learning emphasised the explicit use of tumor-immune relational structure to enhance survival prediction [31].

These end-to-end models eliminate the use of handcrafted features and describe the spatial patterns that are nonlinear and multi-scale. They however demand bigger data sets, more computing and cautious interpretability plans to prevent black box decision making.

7.3 Model Validation Strategies

Strong validation is of great importance in clinical translation. Cross-validation (e.g., k-fold or stratified sampling) measures internal generalization performance but does not consider domain shift with staining, scanner variation, or demographic variation. It is important that external validation on independent cohorts is used to measure real generalizability.

Regularization, dropout, early stopping, and data augmentation in deep learning and feature selection and ensemble in classical models are some examples of overfitting mitigation strategies. WSI frameworks with weak supervision are focused on sparse aggregation to avoid learning spurious correlations. The comparability and reproducibility are enhanced by the reporting of standardized measures, including indexes of concordance (C-index), AUC, calibration curves, and confidence interval. Altogether, predictive modeling involves a tradeoff between interpretability and computational efficiency (feature-based models) and expressive power and multi-scale interaction learning (deep networks). Strict validation and reporting is necessary to take into account that spatial predictors are modeling biologically relevant tumor-immune geometry and not artifacts of the dataset.

8. COMPARATIVE ANALYSIS OF EXISTING STUDIES

Strict comparative synthesis is necessary in order to raise tumor -immune spatial modeling beyond a set of independent methodological articles to a consistent computational system. The literature differs significantly in terms of imaging modality, the criteria used to segment the segmentation, spatial representation as well as learning architecture, scale of the dataset as well as design of validation; all of which directly impact the performance, interpretability, and reproducibility. Previous pipelines had a strong dependence on manually crafted spatial descriptors based on segmentation results as well as classical classifiers, but more recent systems prefer off-the-shelf deep learning designs, such as multiple instance learning (MIL), graph neural networks (GNNs), and transformer-based systems. There is also a difference in the amount of abstraction: some studies represent centroid of cells in terms of spatial statistics or adjacency graphs, whereas others consider entire slide images as sets of correlated patches which interact by attention processes.

The applicability of graph-based predictive modeling has become eminent because of its explicit representation of tumor-immune interactions. Initial graph convolutional survival models showed that it was possible to learn patch level graph based models [32], which were extended to contextual point-cloud models of entire slide images and interaction-guided graph learning systems focusing on microenvironment connectivity. At the same time, MIL methods with less supervision have made it possible to analyze gigapixel slides with no dense annotations, with attention-based survival models scaled to predict patient outcomes using patch embeddings [33]. Transformer-based correlated MIL also improves long-range contextual reasoning to a greater extent. Together, these changes indicate the transition of the engineered spatial characteristics to the learned multi-scale representations that can capture the complicated relational dependencies.

Even though the methods have improved, reproducibility is a significant issue. Lots of researches have high internal cross-validation but no external validation between institutions. The stain normalization variability and segmentation performance variability and graph construction strategy can significantly influence the downstream predictions. Also, the size of datasets is diverse and it makes it difficult to compare the metrics that are reported fairly. Besides, computational cost is also not consistently reported: CNN-based MIL models are often more memory-efficient, and graph-based and transformer models provide more relational modeling at a higher computational cost.

Table 5 summarizes representative studies according to imaging modality, segmentation strategy, spatial representation, AI architecture, dataset scale, performance metrics, and reported limitations, enabling structured comparison across methodological dimensions.

Table 5. Comparative Summary of Representative Tumor-Immune Spatial Modeling Studies

Imaging Modality	Segmentation Method	Spatial Modeling Approach	AI Architecture	Dataset Scale*	Performance Metric	Key Limitations	Study
WSI (H&E)	Patch-level (no dense cell segmentation)	MIL-based aggregation	Attention-based MIL	Moderate cohort	C-index (survival)	Weak supervision; limited external validation	[33]
WSI (H&E)	Patch-based	Weakly supervised patch aggregation	MIL framework	Large multi-task datasets	AUC / task-specific metrics	Domain shift sensitivity	[6]
WSI patches	Region segmentation	Patch graph modeling	Graph CNN	Small-moderate cohort	Survival metrics	Graph scalability constraints	[28]
WSI (point-cloud view)	Patch extraction	Graph-based contextual modeling	GCN	Moderate cohort	C-index	Graph construction dependency	[26]
WSI (H&E)	Patch extraction	Correlated patch modeling	Transformer-based MIL	Large-scale dataset	Classification metrics	High compute demand	[30]
WSI (H&E)	Cell/region segmentation	Tumor-immune interaction graph	Interaction-guided GNN	Moderate-large cohort	Survival metrics	Complex training pipeline	[31]
TME imaging	Cell detection	Spatial point process analysis	Statistical modeling	Variable datasets	Spatial statistics	Limited predictive modeling	[20]
TME imaging	Cell phenotyping	Systems-level spatial graph analysis	Hybrid pipeline	Moderate cohort	Phenotype association	Integration complexity	[7]

Based on these comparisons, there are a number of methodological trends. To begin with, self-evidently, it is the shift of local patch-based CNN embeddings into relational modeling schemes (graphs and transformers) that can encode long-range tumor-immune associations. Second, Downstream modeling is also impacted by the level of granularity in segmentation: instance-based segmentation can explicitly construct the graph at the cost of higher computational cost and increased risk of error propagation, whereas patch-based algorithms can be scaled at the cost of less cellular accuracy. Third, the rigor of validation is rather different; the research that includes multi-cohort analysis or cross-

site normalization methods offers more substantial information on the possibility of generalization as compared to one-centre studies.

In general, the comparative analysis shows that none of the modeling paradigms is dominant in all evaluation axes. Graph-based deep learning is more effective at relational encoding, feature-based and statistical spatial models are easier to interpret and have theoretical underpinnings, and transformer-based MIL architectures can learn global context effectively. Nevertheless, reproducibility, description of preprocessing procedures, and benchmarking are highly important issues to obtain consistent and transferable tumor-immune spatial prediction systems.

9. TECHNICAL CHALLENGES AND OPEN PROBLEMS

Despite the rapid progress of computational anatomy and AI to tumor-immune spatial modeling, it has several technical weaknesses to robustness and clinical use. Among them is domain shift where a model that has been trained using a single-institution dataset cannot be applied to other cohorts due to the variation of staining protocols and scanners, preprocessing pipelines, and demographics. This is particularly elevated in loosely supervised whole-slide learning, in which slide-level labels may augment non-diverse data spuriousness correlations. Although normalizations of stains reduce variance, they have a tradeoff between a biologically relevant structure and distributional consistency [34].

Change in annotation is also implications of performance of segmentation and phenotyping. Contrasting cell boundaries, Mobile immune cues and noise on labels all hamper even the high-quality deep learning models. Weak supervision reduces annotation cost; however, extra precaution of the dataset curation and split design is required to make biased spatial attribution. The other bottleneck is the computational scalability. The processing of gigapixel whole slide images and dense cellular graphs are memory intensive. Patch based pipelines are more practical, but less sensitive to long range dependencies, whereas transformer and graph-based architectures are more accurate, at a higher cost. Graph sampling and pooling techniques are more efficient although they can lead to loss in spatial fidelity. Imbalanced data also makes the prediction of these data more difficult, especially when the rate of survival or infrequent immune phenotypes is low in the data.

The problem of interpretability and multi-resolution integration is yet to be solved. Even though attention mechanisms and graph structure provide a partial transparency, it is challenging to attribute complex multi-hop spatial interactions. In addition, cellular- and slide-level information will need to be integrated whereby a strong cross-scale correspondence is needed. There are still no standardized benchmarking and validation protocols (Figure 2). Addressing domain shift, annotation noise, scalability, interpretability, and benchmarking is essential for developing generalizable and clinically reliable spatial AI systems.

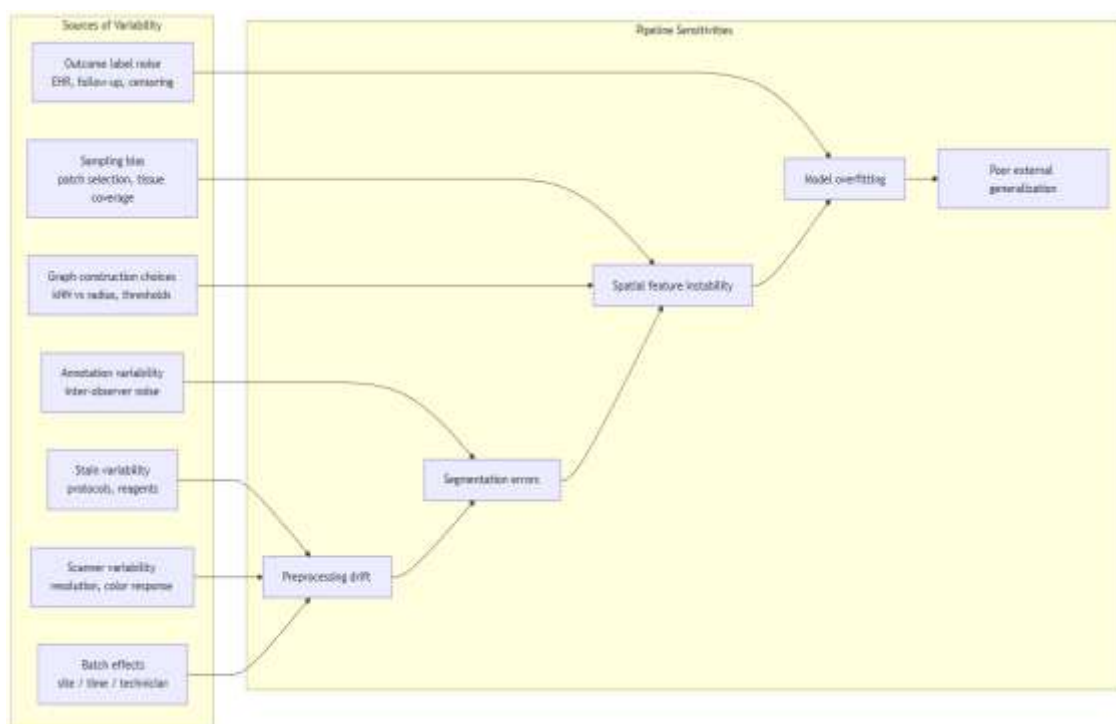


Figure 2. Sources of domain shift and variability in tumor-immune spatial modeling.

10. FUTURE DIRECTIONS IN SPATIAL IMAGE ANALYTICS

The next generation of tumor-immune spatial modeling will rely on the representation learning approaches that decrease the annotation pressure, improve the cross-cohort predictions, and combine the multiple scales and multimodal data. One of the key directions is the creation of foundation models in computational pathology. Massive pretraining with a variety of whole-slide datasets is expected to produce transferable embeddings, which can be used in segmentation, phenotyping, spatial interaction modeling, and outcome prediction. The use of multimodal foundation models shows that the matching of pathology features with other sources of complementary information enhances the generalization of tasks and institutions. Representations further aided by methods of vector-quantization are further standardized and compressed, improving scalability to large-scale application. An important agenda is to make sure that these models maintain biologically relevant tumor-immune geometry as opposed to revealing site-specific biases.

Another significant direction is the self-supervised learning. Contrastive and pretext-based goals allow learning representations using unlabeled data, eliminating the need to use large annotations. Self-supervised contrastive frameworks have demonstrated good performance in whole-slide representation learning in digital pathology [35]. Further studies can formulate the purpose of the research specific to the geometry of space, graph-based contrastive learning, multiscale consistency constraints, or topology-sensitive augmentations that do not alter interaction structure. Federated learning provides a system of multi-institution training that is not centralized and sensitive data but enhances the process of generalization, overcoming the limitations of privacy. Nevertheless, staining, scanner, and non-iid data distribution heterogeneity have significant technical challenges. To ensure the solid federated deployment, strong

normalization and domain adaptive optimization strategy will be necessary to curb the cross-site performance degradation.

A new frontier is physics-informed and biologically constrained modeling. Neural objectives may be enriched with spatial interaction priors or reaction-diffusion dynamics to make them more sample efficient and more plausible at a mechanistic scale. Such approaches would eventually enable the simulation of tumor-immune spatial evolution during therapy, in addition to outcome prediction. Finally, such improvements as inclusivity to computational efficiency, robustness, the quantification of uncertainty, and interpretation are required in clinical translation. Whole-slide inference must be scalable and reproducible as well as transparent on its decisions. Despite the fact that foundation and self-supervised models also improve the transferability, they come with new governance, auditing, and constant monitoring requirements. In sum, spatial image analytics of the future should be able to trade off expressive representation learning, scalability, reliability, and clinical trustworthiness in order to enable precision oncology.

11. CONCLUSION

The quantitative studies of tumor-immune spatial organization have been revolutionized by computational anatomy and artificial intelligence, which transforms the complex anatomy in tissues into a quantifiable geometry and relational forms of tissue architecture. The current technological progress in high-resolution imaging, preprocessing and segmentation allows dependable extraction of cells and region-level spatial features whereas statistical modeling, graph-based representation and topology-inspired approaches offer principled methods to describe patterns of clustering, exclusion and interaction within the microenvironment of the tumor. Recent models of learning such as weakly-supervised whole-slide models, transformers and graph neural networks have also allowed the connection of multiscale spatial embeddings to clinically-relevant outcomes. Spatial quantification enabled by AI has changed the paradigm of predictive modeling on density-based biomarkers to high-dimensional representation reflecting local morphology as well as spatial context. Nevertheless, strong translation to domain shift, transparent spatial attribution, and rigorous cross-independent cohort validation is essential to clinical translation. One of its critical weaknesses is that there is no standard benchmarking. The flexibility of imaging regimes, preprocessing plans, segmentation detail, spatial modeling options and evaluation plan makes it challenging to compare and reproducibly compare them fairly. It is critical to set common standards and align the reporting standards, metrics, and testing of cross-site generalization. The benchmarking frameworks should also measure the error propagation throughout the pipeline - through segmentation to spatial feature extraction and result forecasting. The foundation and self-supervised models, multimodal integration, and federated learning strategies will determine the future development. Finally, with the combination of computational anatomy and geometry-conscious AI and thorough validation, generalizable clinical-trustworthy systems predictive of treatment sensitivity based on tumor-immune spatial geometry can be created.

REFERENCES

1. Gong, D., Arbesfeld-Qiu, J.M., Perrault, E., Bae, J.W. and Hwang, W.L., 2024. Spatial oncology: Translating contextual biology to the clinic. *Cancer Cell*, 42(10), pp.1653-1675.
2. Janeiro, A.L., Wong, E.M., Jiménez-Sánchez, D., de Solorzano, C.O., Lozano, M.D., Teijeira, A., Schalper, K.A., Melero, I. and De Andrea, C.E., 2024. Spatially resolved tissue imaging

- to analyze the tumor immune microenvironment: beyond cell-type densities. *Journal for Immunotherapy of Cancer*, 12(5), p.e008589.
3. Heindl, A., Nawaz, S. and Yuan, Y., 2015. Mapping spatial heterogeneity in the tumor microenvironment: a new era for digital pathology. *Laboratory investigation*, 95(4), pp.377-384.
 4. Hu, C., Wang, F., Xu, H., Dong, X., Xiong, X., Zhang, Y., Zhao, T., He, Y., Deng, L. and Lu, X., 2025. Deep learning-based spatial analysis on tumor and immune cells of pathology images predicts MIBC prognosis. *Plos one*, 20(8), p.e0328816.
 5. Park, Y.M. and Lin, D.C., 2023. Moving closer towards a comprehensive view of tumor biology and microarchitecture using spatial transcriptomics. *nature communications*, 14(1), p.7017.
 6. Lu, M.Y., Williamson, D.F., Chen, T.Y., Chen, R.J., Barbieri, M. and Mahmood, F., 2021. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature biomedical engineering*, 5(6), pp.555-570.
 7. Hanson, G.F., Goundry, K.A., Wessel, R.E., Brown, M.G., Bullock, T.N. and Dolatshahi, S., 2025. A flexible systems analysis pipeline for elucidating spatial relationships in the tumor microenvironment linked with cellular phenotypes and patient-level features. *Frontiers in Immunology*, 16, p.1642527.
 8. Xu, C., Sun, Y., Zhang, Y., Liu, T., Wang, X., Hu, D., Huang, S., Li, J., Zhang, F. and Li, G., 2025. Stain Normalization of Histopathological Images Based on Deep Learning: A Review. *Diagnostics*, 15(8), p.1032.
 9. Breen, J., Zucker, K., Allen, K., Ravikumar, N. and Orsi, N.M., 2024. Generative adversarial networks for stain normalisation in histopathology. In *Applications of Generative AI* (pp. 227-247). Cham: Springer International Publishing.
 10. Hu, B., Sajid, M., Lv, R., Liu, L. and Sun, C., 2022. A review of spatial profiling technologies for characterizing the tumor microenvironment in immuno-oncology. *Frontiers in Immunology*, 13, p.996721.
 11. Elhanani, O., Ben-Uri, R. and Keren, L., 2023. Spatial profiling technologies illuminate the tumor microenvironment. *Cancer cell*, 41(3), pp.404-420.
 12. Ronneberger, O., Fischer, P. and Brox, T., 2015, October. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Cham: Springer international publishing.
 13. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818).
 14. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J. and Maier-Hein, K.H., 2021. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2), pp.203-211.
 15. Greenwald, N.F., Miller, G., Moen, E., Kong, A., Kagel, A., Dougherty, T., Fullaway, C.C., McIntosh, B.J., Leow, K.X., Schwartz, M.S. and Pavelchek, C., 2022. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nature biotechnology*, 40(4), pp.555-565.
 16. He, K., Gkioxari, G., Dollár, P. and Girshick, R., 2017. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).
 17. Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T. and Rajpoot, N., 2019. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical image analysis*, 58, p.101563.
 18. Schmidt, U., Weigert, M., Broaddus, C. and Myers, G., 2018, September. Cell detection with star-convex polygons. In *International conference on medical image*

- computing and computer-assisted intervention (pp. 265-273). Cham: Springer International Publishing.
19. Stringer, C., Wang, T., Michaelos, M. and Pachitariu, M., 2021. Cellpose: a generalist algorithm for cellular segmentation. *Nature methods*, 18(1), pp.100-106.
20. Benimam, M.M., Meas-Yedid, V., Mukherjee, S., Frafjord, A., Corthay, A., Lagache, T. and Olivo-Marin, J.C., 2025. Statistical analysis of spatial patterns in tumor microenvironment images. *Nature Communications*, 16(1), p.3090.
21. Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A.L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4), pp.834-848.
22. Kipf, T.N., 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.
23. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P. and Bengio, Y., 2017. Graph attention networks. *stat*, 1050(20), pp.10-48550.
24. Hamilton, W., Ying, Z. and Leskovec, J., 2017. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30.
25. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S. and Guo, B., 2021. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 10012-10022).
26. Chen, R.J., Lu, M.Y., Shaban, M., Chen, C., Chen, T.Y., Williamson, D.F. and Mahmood, F., 2021, September. Whole slide images are 2d point clouds: Context-aware survival prediction using patch-based graph convolutional networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 339-349). Cham: Springer International Publishing.
27. Ding, T., Wagner, S.J., Song, A.H., Chen, R.J., Lu, M.Y., Zhang, A., Vaidya, A.J., Jaume, G., Shaban, M., Kim, A. and Williamson, D.F., 2025. A multimodal whole-slide foundation model for pathology. *Nature medicine*, pp.1-13.
28. Li, H., Shui, Z., Zhang, Y., Zhu, C. and Yang, L., 2025. PathVQ: Reforming Computational Pathology Foundation Model for Whole Slide Image Analysis via Vector Quantization. *arXiv preprint arXiv:2503.06482*.
29. Xu, K., Hu, W., Leskovec, J. and Jegelka, S., 2018. How powerful are graph neural networks?. *arXiv preprint arXiv:1810.00826*.
30. Shao, Z., Bian, H., Chen, Y., Wang, Y., Zhang, J. and Ji, X., 2021. Transmil: Transformer based correlated multiple instance learning for whole slide image classification. *Advances in neural information processing systems*, 34, pp.2136-2147.
31. Shao, W., Shi, Y., Zhang, D., Zhou, J. and Wan, P., 2024. Tumor micro-environment interactions guided graph learning for survival analysis of human cancers from whole-slide pathological images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 11694-11703).
32. Li, R., Yao, J., Zhu, X., Li, Y. and Huang, J., 2018, September. Graph CNN for survival analysis on whole slide pathological images. In *International conference on medical image computing and computer-assisted intervention* (pp. 174-182). Cham: Springer International Publishing.
33. Yao, J., Zhu, X., Jonnagaddala, J., Hawkins, N. and Huang, J., 2020. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. *Medical image analysis*, 65, p.101789.
34. Madusanka, N., Padmanabha, P., Guruge, K. and Lee, B.I., 2025. Structure-Preserving Histopathological Stain Normalization via Attention-Guided Residual Learning. *Bioengineering*, 12(9), p.950.

35. Fashi, P.A., Hemati, S., Babaie, M., Gonzalez, R. and Tizhoosh, H.R., 2022. A self-supervised contrastive learning approach for whole slide image representation in digital pathology. *Journal of Pathology Informatics*, 13, p.100133.
36. Ilse, M., Tomczak, J. and Welling, M., 2018, July. Attention-based deep multiple instance learning. In *International conference on machine learning* (pp. 2127-2136). PMLR.