Preprints.org

Review

Not peer-reviewed version

Phenotyping of Histology Imaging Data with Histomics

Fnu Neha *, Deepshikha Bhati, Deepak Kumar Shukla

Posted Date: 21 April 2026

doi: 10.20944/preprints202604.1359.v1

Keywords: histomics; computational pathology; histopathology; whole slide images (WSI); artificial intelligence (AI); medical image processing; automated diagnosis; digital pathology

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Phenotyping of Histology Imaging Data with Histomics

Fnu Neha [1,*], Deepshikha Bhati [1] and Deepak Kumar Shukla [2]

[1] Department of Computer Science, Kent State University, Kent, OH 44242, USA
[2] Rutgers Business School, Rutgers University, Newark, NJ 07102, USA
* Correspondence: neha@kent.edu

Abstract

Whole-slide imaging has transformed histopathology into a data-intensive domain, with current approaches dominated by end-to-end deep learning that encode morphology implicitly within latent representations. This limits interpretability, reproducibility, and cross-dataset generalization. This review positions histomics as an intermediate phenotype representation layer that maps histological images to structured, multi-scale descriptors of tissue morphology, spatial organization, and architectural context. A unified taxonomy of histomic features across biological scales is presented, along with an analysis of artificial intelligence frameworks spanning classical machine learning, deep learning, weakly supervised learning, and multimodal integration. The review presents core failure modes in histomic pipelines, including segmentation dependence, feature instability, and domain shift, and examines their impact on robustness and generalization. Emerging trends in representation learning and multimodal modeling are analyzed in the context of phenotype-centric inference. Overall, this work reframes histomics as a representation-driven paradigm and outlines directions for developing stable, interpretable, and generalizable computational pathology systems.

Keywords: histomics; computational pathology; histopathology; whole slide images (WSI); artificial intelligence (AI); medical image processing; automated diagnosis; digital pathology

1. Introduction

Advances in whole-slide imaging and high-throughput digital microscopy have transformed histopathology from a qualitative, observer-driven practice into a data-intensive computational discipline. Early digital pathology focused on automating core diagnostic tasks, including tissue segmentation, tumor detection, and histologic grading. Subsequent developments expanded this scope through machine learning (ML) and deep learning (DL) models for classification, prognostic prediction, and outcome association. Despite these advances, current approaches increasingly rely on end-to-end pipelines in which whole-slide images (WSIs) are processed as raw inputs and morphological characteristics are encoded implicitly within latent representations. While effective for predictive performance, such works lack explicit and standardized representations of tissue phenotype, limiting interpretability, reproducibility, and cross-study comparability.

Histomics provides a framework for quantitative characterization of tissue architecture, cellular morphology, and spatial organization in histopathological images [1]. In contrast to prediction-centric pipelines, it emphasizes explicit feature extraction and phenotype-driven modeling, enabling direct association between measurable morphological patterns and clinical or biological outcomes. However, histomic pipelines remain sensitive to segmentation quality, staining variability, and acquisition differences, and lack standardized feature definitions, which constrains robustness and reproducibility [2].

Existing literature in digital and computational pathology confuses histomics with pathomics or DL–based image analysis, with limited distinction between feature-centric phenotyping and representation learning[1,3–5]. Methodological components, including feature taxonomy, spatial encoding, and

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

validation strategies, remain fragmented across application-specific studies [6]. Furthermore, feature instability, segmentation-induced variability, and dataset bias are insufficiently examined, limiting cumulative progress and reliable cross-study comparison [7].

This review reframes histomics as a representation-level abstraction rather than a feature extraction pipeline. Histomic descriptors are treated as a structured phenotype space bridging raw images and clinical predictions. From this view, segmentation dependence, feature instability, and domain shift emerge as consequences of misalignment between latent and phenotype representations. Formalizing this representation gap provides a unified basis to analyze learning paradigms and improve robustness, interpretability, and generalization in computational pathology.

The contributions of this review are as follows:

- Formalizes histomics as a phenotype-driven representation framework in computational pathology.
- Presents a unified taxonomy of histomic features.
- Analyzes learning paradigms under constraints of high dimensionality and weak supervision.
- Highlights challenges in standardization, validation, and multimodal integration.

The remainder of this paper is organized as follows: Section 2 presents the research methodology, Section 3 provides histomics background, Section 4 discusses ML and DL in histomics, Section 5 covers multimodal integration, Section 6 outlines emerging paradigms, Section 7 reviews related work, Section 9 discusses limitations and future directions, and Section 10 concludes this paper.

2. Research Methodology

This review adopts a reproducible methodology to identify, screen, and synthesize literature on histomics, with emphasis on feature-centric representation, methodological rigor, and translational relevance. The process follows database-driven retrieval, multi-stage screening, and qualitative synthesis. Table 1 summarizes the key components of the methodology.

Table 1. Summary of review methodology components and criteria.

Component	Description
Literature Sources	IEEE Xplore, ACM Digital Library, PubMed, Scopus, Web of Science, SPIE Digital Library, and Google Scholar; publications up to 2026.
Study Types	Peer-reviewed journal articles, conference proceedings, and selected preprints with methodological contributions.
Inclusion Criteria	Studies involving quantitative histopathological image analysis with explicit morphological, spatial, or texture-based features; contributions in feature design, representation, modeling, or validation using whole-slide or large-field microscopy data.
Exclusion Criteria	End-to-end ML/DL approaches without explicit feature representation; application-only studies lacking methodological detail; non-histopathological imaging; editorials, abstracts without full manuscripts, and non-English works.
Screening Process	Two-stage screening comprising title/abstract filtering followed by full-text review. Ambiguous cases were resolved based on relevance to feature-centric histomic modeling and interpretability.
Extracted Attributes	Feature categories, spatial modeling strategies, learning paradigms, datasets, validation protocols, and reported limitations.
Synthesis Approach	Qualitative comparative synthesis focusing on methodological patterns, feature stability, and validation practices rather than task-specific performance metrics.

The literature retrieval process was conducted using combinations of keywords such as *histomics*, *computational pathology*, *histopathological feature extraction*, *pathomics*, and *whole-slide image analysis*. The initial search yielded approximately 192 records across all databases. After duplicate removal and title/abstract screening, 132 studies were retained for full-text review. Following eligibility assessment

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

based on feature-centric representation and methodological rigor, a final set of 59 studies was included for qualitative synthesis.

2.1. Research Questions

This review is guided by the following research questions (RQs), which frame histomics as a phenotype representation problem rather than a task-specific modeling approach:

- **RQ1:** How is histomics defined as a phenotype representation framework in computational pathology?
- **RQ2:** What classes of histomic features and spatial descriptors are used within different learning paradigms?
- **RQ3:** How do different learning paradigms address challenges such as high dimensionality, weak supervision, and domain variability?

3. Background

Histopathology relies on microscopic examination of stained tissue sections to assess cellular morphology and tissue architecture. Diagnostic interpretation is based on visual assessment of structural patterns, including nuclear atypia, tissue organization, and staining characteristics [8]. While clinically effective, this process remains qualitative, subject to inter- and intra-observer variability, and difficult to standardize [9].

Whole-slide imaging enables digitization of histological specimens at subcellular resolution, transforming pathology into a computational imaging domain [10]. This shift has driven the development of ML and DL models for segmentation, classification, and prognostic prediction. These approaches optimize performance directly from image data but encode morphology implicitly within latent representations, limiting interpretability and cross-dataset reproducibility [11].

Histomics addresses this limitation by representing tissue as a structured phenotype space. Formally, it defines a mapping:

$$\mathcal{H} : I \mapsto z_p \tag{1}$$

where I denotes a WSI and $\mapsto z_p$ is a phenotype vector encoding morphological, spatial, and multi-scale descriptors.

Unlike latent embeddings, histomic representations decompose tissue into interpretable descriptors of cellular morphology, architectural organization, and spatial interactions. This enables phenotype-level analysis and integration with clinical and molecular data. However, these representations remain sensitive to segmentation quality, staining variation, and implementation choices, resulting in feature instability across studies [3].

Figure 1 illustrates a phenotype-centric histomics framework, wherein WSIs are decomposed into explicit multi-scale descriptors of cellular properties, tissue architecture, and spatial organization, enabling structured and interpretable representation learning.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

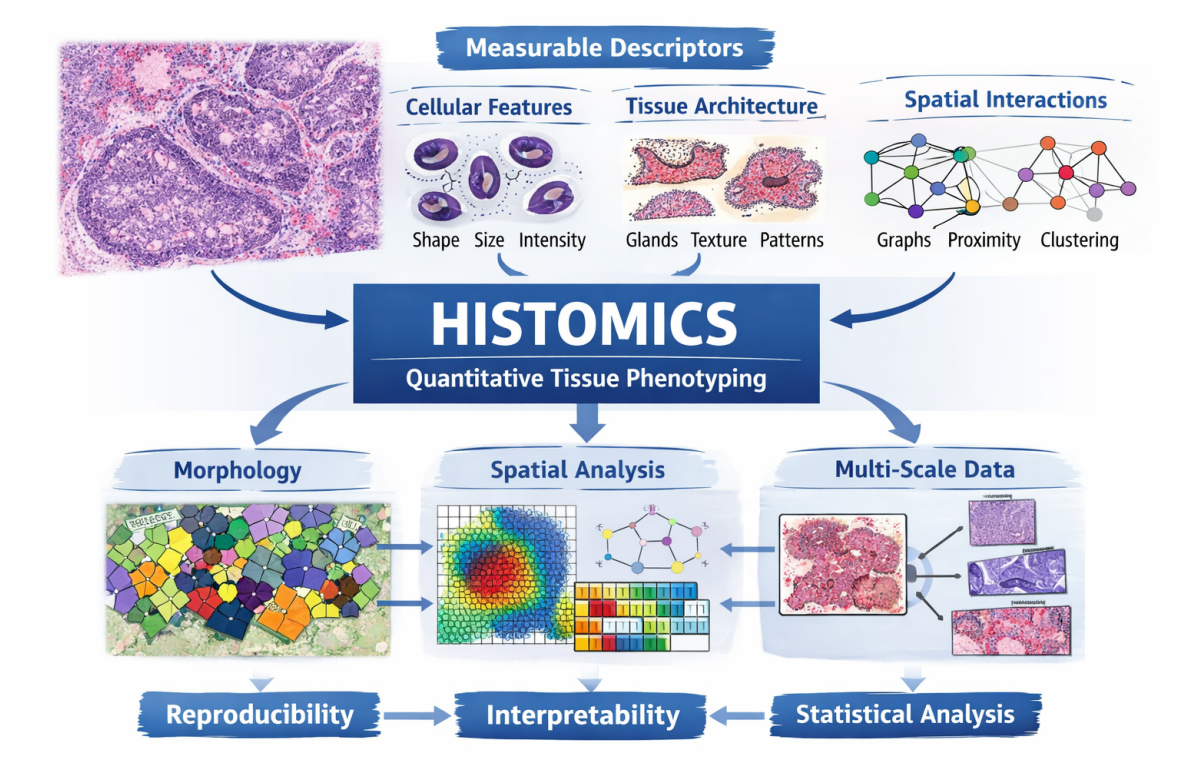

Figure 1. Phenotype-centric histomics framework for structured multi-scale tissue representation.

3.1. Key Components of Histomics

Histomic pipelines follow a structured sequence of preprocessing, segmentation, feature extraction, and aggregation [4,12]. These stages aim to preserve morphological fidelity while enabling quantitative representation of tissue phenotype. Figure 2 presents the overall workflow.

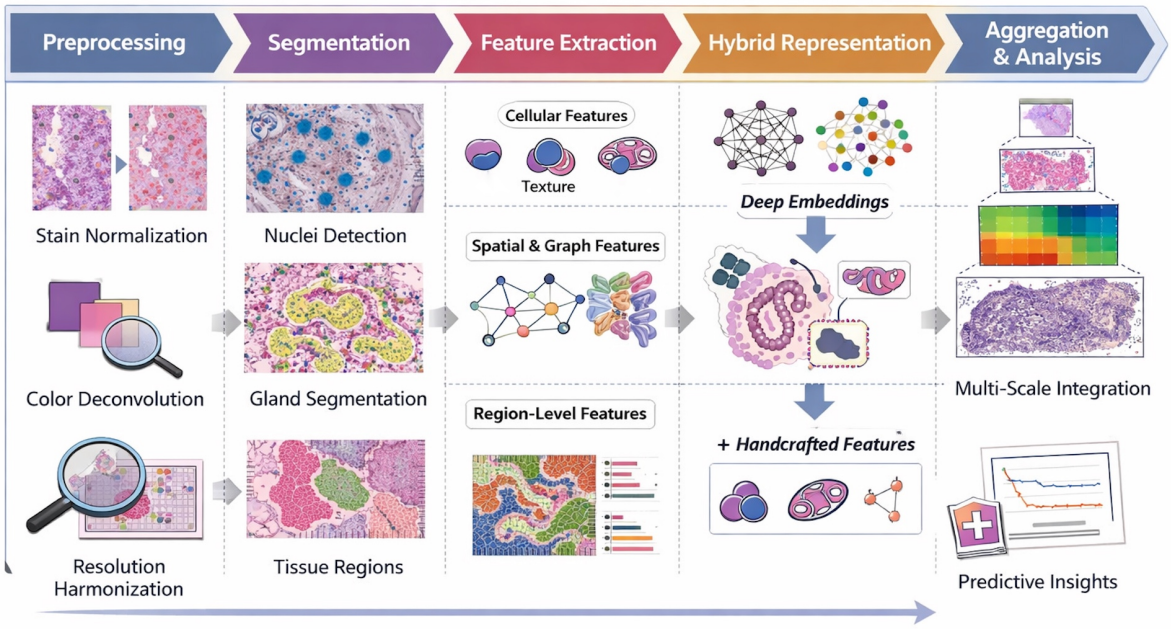

Figure 2. Feature-centric histomics pipeline illustrating preprocessing, biologically informed segmentation, multi-scale feature extraction, hybrid integration of handcrafted and deep representations, and hierarchical aggregation for interpretable analysis.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

3.1.1. Preprocessing

Preprocessing reduces non-biological variability introduced by staining protocols, scanner settings, and acquisition conditions [13]. Standard operations include stain normalization, color deconvolution, spatial resolution harmonization, and artifact suppression.

Stain normalization aligns color distributions across WSIs to a reference appearance, reducing inter-slide and inter-batch variability while preserving relative chromatic contrast. A common formulation is based on optical density (OD) transformation:

$$OD = -\log\left(\frac{I}{I_0}\right) \tag{2}$$

where I denotes the observed pixel intensity and I_0 represents the incident light intensity. This transformation linearizes stain absorption and enables normalization in stain space.

Color deconvolution separates composite RGB signals into stain-specific channels using a stain matrix S:

$$OD = S \cdot C \tag{3}$$

where OD is the optical density vector and C contains stain concentrations. This enables independent analysis of nuclear and cytoplasmic components and improves robustness of intensity- and texture-derived descriptors..

Spatial resolution harmonization enforces scale consistency across slides acquired at different magnifications through resampling:

$$I'(x, y) = I(\alpha x, \alpha y) \tag{4}$$

where α denotes the scaling factor, ensuring comparability of scale-dependent morphological and spatial features.

Additional steps such as tissue masking, background exclusion, and artifact filtering constrain feature extraction to biologically relevant regions while suppressing scanning noise and preparation artifacts. For example, a binary tissue mask $M(x, y)$ can be defined as:

$$M(x, y) = \begin{cases} 1, & \text{if } I(x, y) \in \text{tissue region} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

which restricts downstream analysis to relevant regions.

3.1.2. Segmentation

Segmentation delineates nuclei, cells, glands, stromal regions, and higher-order tissue compartments, providing the structural basis for localized phenotype quantification [14,15]. The choice of segmentation strategy is based on target scale, morphological complexity, and staining variability.

Classical image processing methods apply thresholding in optical density space, morphological operations, and watershed-based separation to segment well-contrasted nuclei and simple tissue regions.

ML-based approaches cast segmentation as pixel- or object-level classification, integrating handcrafted intensity, texture, and shape features to improve robustness under moderate variability. Framework such as *CellProfiler* is frequently used for feature extraction and classification-based segmentation pipelines.

DL-based methods model segmentation as dense prediction or instance separation tasks, leveraging convolutional architectures to capture contextual and multiscale dependencies, enabling accurate delineation of overlapping nuclei, irregular glandular structures, and heterogeneous tissue compartments.

Post-segmentation filtering based on geometric or intensity constraints is commonly applied to suppress implausible detections and segmentation artifacts.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

3.1.3. Feature Extraction

Feature extraction maps segmented histological structures to quantitative phenotype descriptors across cellular, regional, and spatial scales [4,6] as shown in Table 2. These descriptors encode morphology, intensity, texture, and spatial organization, forming a structured representation of tissue phenotype.

At the cellular scale, morphological features quantify geometric properties such as area, perimeter, eccentricity, and circularity, capturing deviations from regular structure associated with nuclear atypia and malignancy. Intensity features describe staining distribution through statistics such as mean optical density, variance, and entropy, reflecting chromatin density and intra-nuclear heterogeneity. Texture descriptors model local micro-patterns using representations such as Gray-Level Co-occurrence Matrix (GLCM) and Local Binary Patterns (LBP), capturing chromatic variation and tissue differentiation.

Spatial features extend this representation by modeling interactions between cellular components, including nearest-neighbor statistics, inter-nuclear distances, and density estimates, which characterize cellular organization, crowding, and architectural disruption. Graph-based representations further encode higher-order relationships by modeling tissue as a network, enabling topology-aware analysis of the tumor microenvironment.

At regional and tissue scales, structural descriptors summarize glandular architecture, lumen ratios, and tissue composition, while multi-scale aggregation integrates object-level features into slide-level representations through statistical pooling. Distributional descriptors further capture variability across tissue, enabling modeling of intra-tumoral heterogeneity.

The hierarchical organization of these representations across biological scales is summarized in Table 3.

Table 2. Interpretation of histomic features.

Feature Type	Representative Metrics	Formulation	Interpretation
Morphological	Area, perimeter, circularity, eccentricity	$A,\ P,\ \frac{4\pi A}{P^2}$	Encodes nuclear shape and boundary irregularity; deviations indicate atypia and malignancy
Intensity	Mean OD, variance, entropy	$\mu = \frac{1}{\lvert o\rvert}\sum I,\ \sigma^2 = \mathrm{Var}(I)$	Reflects chromatin density and staining heterogeneity linked to cellular activity
Texture	GLCM contrast, energy, homogeneity; LBP	$\sum(i-j)^2 P(i,j),\ \sum P(i,j)^2$	Captures local micro-patterns and chromatic variation associated with tissue differentiation
Spatial	Nearest-neighbor distance, density, clustering	$d_{ij} = \lVert x_i - x_j\rVert$	Models cellular organization, crowding, and architectural disruption
Graph-based	Node degree, adjacency, connectivity	$G = (V,E),\ A_{ij}$	Represents higher-order spatial interactions and tissue topology
Regional	Gland size, lumen ratio, composition	Area ratios, region statistics	Summarizes tissue architecture at mesoscopic scale
Aggregated	Mean, variance, pooling	$\bar{\mathbf{z}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_i$	Captures global phenotype distribution and heterogeneity
Distributional	Histograms, quantiles, skewness	$p(z),$ quantiles	Models feature variability and intra-tumoral heterogeneity

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

Table 3. Hierarchical organization of histomic representations across biological scales.

Scale	Representation Unit	Description	Modeling Role
Cellular	Individual objects (e.g., nuclei)	Local descriptors computed per segmented entity	Capture fine-grained phenotype and cellular variability
Spatial	Object neighborhoods	Relationships defined through proximity or connectivity	Encode organization, interaction, and structural arrangement
Regional	Tissue substructures	Aggregation within anatomically meaningful regions	Represent mesoscopic architecture and subtype patterns
Slide-level	Whole-slide representation	Statistical aggregation of lower-level descriptors	Capture global phenotype and heterogeneity
Multi-scale	Hierarchical integration	Joint modeling across cellular, spatial, and regional levels	Enable robust and context-aware representation learning

Feature normalization is applied prior to aggregation to ensure comparability across descriptors and to stabilize representations across datasets.

3.1.4. Hybrid Representation and Multi-Scale Aggregation

Recent approaches integrate handcrafted histomic features with learned deep embeddings [16–18]. Deep representations capture contextual and multiscale patterns, while handcrafted descriptors retain explicit morphological and spatial semantics. Fusion is typically performed at feature-, representation-, or decision-level, requiring normalization and alignment to mitigate scale mismatch and distributional bias.

Aggregation transforms instance-level descriptors (e.g., patches, nuclei, or regions) into slide-level representations under weak supervision. Multiple Instance Learning (MIL) frameworks treat slides as bags of instances, where attention-based pooling assigns adaptive weights:

$$z = \sum_{i=1}^{N} \alpha_i h_i, \quad \sum_{i=1}^{N} \alpha_i = 1 \tag{6}$$

where h_i denotes instance embeddings and α_i represents learned attention weights reflecting regional importance.

Transformer-based aggregation extends this formulation by modeling long-range dependencies across instances using self-attention, improving global context modeling in heterogeneous tissue regions. Graph-based aggregation further encodes spatial relationships by representing nuclei or regions as nodes with edges capturing proximity or morphological similarity, enabling topology-aware feature integration.

Hierarchical aggregation organizes features across multiple spatial scales (cell → region → slide), while statistical pooling (mean, variance, higher-order moments) provides compact summaries of feature distributions. Despite these advances, aggregation strategies introduce variability in representation stability and sensitivity to intra-tumoral heterogeneity, and their comparative impact remains underexplored across datasets and tasks.

3.2. Histomics in Context: Comparison with Existing Paradigms

Computational pathology approaches differ in how tissue morphology is represented. Pathomics relies on handcrafted descriptors, offering interpretability but limited scalability [19]. DL-based methods learn latent representations directly from images, achieving strong performance but reducing interpretability and reproducibility [20].

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

Histomics combines explicit descriptors with learned representations, enabling structured and interpretable phenotype modeling. However, challenges remain in standardization, feature stability, and integration across datasets. Table 4 presents a comparison of histomics with related paradigms.

Table 4. Comparison of histomics with related paradigms.

Paradigm	Representation	Interpretability	Strength	Limitation
Pathomics	Handcrafted features	High	Explicit descriptors	Limited scalability
DL-based pathology	Latent embeddings	Low	High performance	Limited interpretability
Histomics	Hybrid	Moderate–High	Structured phenotype modeling	Lack of standardization

4. Machine Learning and Deep Learning in Histomics

ML and DL in histomics formalize inference as a mapping $f_\theta : \mathcal{X} \to \mathcal{Y}$, where $\mathcal{X}$ denotes high-dimensional histomic representations derived from WSIs and $\mathcal{Y}$ corresponds to clinical or biological endpoints. as shown in Table 5, model design is constrained by high dimensionality, feature correlation, segmentation-dependent variability, sparse supervision, and domain shift, which affect stability and generalization.

Table 5. Mapping of modeling challenges to common solutions in histomics.

Modeling Challenge	Common Approach
Limited labeled data	Weakly supervised learning (e.g., MIL)
High feature dimensionality	Feature selection, regularization, dimensionality reduction
Feature redundancy	Sparse modeling, ensemble methods
Domain variability	Normalization, domain adaptation, augmentation
Latent spatial structure	Deep learning (CNNs, transformers)
Lack of supervision	Self- and semi-supervised learning

4.1. Supervised Modeling

In supervised modeling, histomic analysis is posed as a discriminative learning problem over fixed-length feature vectors or structured representations extracted at object, region, or slide scales [21]. Classical models such as support vector machines (SVMs) and ensemble methods operate on handcrafted histomic descriptors, optimizing convex or additive objectives under explicit regularization to control variance in high-dimensional, low-sample frameworks. For example, SVMs are used for tumor subtype classification based on nuclear morphology and texture features [22], while random forests and gradient boosting are applied for prognostic prediction and feature importance estimation. Feature selection and sparsity constraints, such as LASSO, are used to stabilize estimation and preserve interpretability [23].

Deep supervised models learn representations directly from image regions using convolutional or transformer-based architectures. CNNs such as ResNet and DenseNet are used for patch-level classification and grading, while vision transformers capture long-range spatial dependencies. Attention modules (e.g., CBAM) improve localization of discriminative regions. These models improve performance but require large labeled datasets and encode morphology implicitly, limiting interpretability and cross-dataset consistency.

4.2. Weakly Supervised, Semi-Supervised, and Self-Supervised Learning

Given limited annotations, learning is performed with slide-level labels. Weakly supervised approaches model slides as sets of instances using MIL, where attention-based pooling identifies

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

informative regions for tasks such as tumor detection and grading [24]. This reduces annotation requirements but introduces uncertainty and sensitivity to noise.

Semi-supervised models augment supervised objectives with unlabeled data using consistency regularization and pseudo-labeling. For example, teacher–student frameworks (e.g., Mean Teacher) enforce prediction consistency across perturbations, improving robustness under limited annotations [25].

Self-supervised learning removes external labels by defining pretext tasks such as contrastive learning (e.g., SimCLR, MoCo) or context prediction on image patches [26,27]. Encoders pretrained in this manner improve initialization and generalization for downstream histomic tasks.

Table 6 summarizes the learning techniques commonly adopted in histomics, highlighting how supervision level and input structure determine optimization objectives and modeling constraints.

Table 6. Learning techniques in histomics.

Paradigm	Input Structure	Objective	Constraints
Classical supervised ML	Fixed-length histomic feature vectors	Empirical risk minimization with margin-based or ensemble loss and explicit regularization	Limited samples, weak feature interactions, need for interpretability
Supervised DL	Image patches, tiles, or region embeddings	End-to-end minimization of task loss via hierarchical spatial encoders	Sufficient labeled data, high computational cost, implicit feature learning
Weakly supervised learning	Sets of local instances per slide	Slide-level loss with latent instance weighting or aggregation constraints	Instance relevance unobserved; global labels constrain local predictions
Semi-supervised learning	Labeled and unlabeled histomic representations	Supervised loss augmented with manifold, consistency, or entropy regularization	Data lie on smooth low-dimensional manifolds
Self-supervised learning	Unlabeled image regions or tiles	Pretext optimization based on spatial, contextual, or cross-scale consistency	Structural regularities encode transferable morphological priors

5. Integration with Multimodal Data

Recent advances in histomics increasingly position morphological phenotypes as one component of broader multimodal representations that integrate molecular, spatial, and clinical signals [28–30]. Multimodal integration involves joint inference over heterogeneous data types with mismatched resolutions, noise characteristics, and supervision regimes. Histomic features serve as spatially grounded, interpretable anchors that enable cross-modal alignment and biologically coherent fusion.

5.1. Histomics and Genomics Integration

Integration of histomic and genomic data links observable tissue morphology with underlying molecular alterations [31]. Histomic descriptors capture phenotypic manifestations of genomic dysregulation, while genomic profiles encode latent drivers not directly observable in tissue structure. Models typically adopt feature-level or representation-level fusion to predict clinical endpoints such as subtype, risk, or treatment response.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

Challenges arise from dimensionality imbalance, scale mismatch, and heterogeneous noise across modalities. Effective integration requires dimensionality reduction, normalization, and fusion strategies that preserve modality-specific signals while enabling interaction. Histomic features act as stabilizing priors; however, consistent alignment between morphological patterns and genomic signals is not guaranteed across datasets.

5.2. Histomics and Spatial Omics Integration

Spatial omics provides molecular measurements with spatial coordinates that can be aligned with histological structure [32]. Histomic features supply morphological context, enabling joint analysis of tissue architecture and spatial molecular heterogeneity.

Integration is achieved by mapping molecular measurements to histologically defined regions or cellular neighborhoods. This supports modeling of structure–function relationships, linking morphology with spatial gene or protein expression. However, alignment is sensitive to resolution differences, registration errors, and sampling sparsity, which can distort spatial correspondence.

5.3. Fusion Models and Multimodal Deep Networks

Multimodal fusion is implemented through early, intermediate, or late strategies [33,34]. Feature-level fusion concatenates modality-specific descriptors, while representation-level fusion integrates latent embeddings. Decision-level fusion combines modality-specific predictions.

Deep multimodal networks extend these approaches by learning cross-modal interactions using shared latent spaces, attention mechanisms, or gating functions. These architectures adaptively weight modalities based on relevance and data quality while preserving modality-specific structure. Despite flexibility, they require careful design to avoid overfitting, modality dominance, and misalignment between representations.

Table 7 summarizes fusion strategies and their associated modeling constraints.

Table 7. Fusion taxonomy for multimodal histomics aligned with learning models.

Fusion Strategy	Learning Model	Fusion Formulation	Constraints
Early (Feature-Level) Fusion	Classical supervised ML	Concatenation of modality-specific handcrafted features prior to modeling	Shared sample correspondence; limited tolerance to scale mismatch and noise heterogeneity
Representation-Level Fusion	Supervised DL	Joint or coordinated learning of latent embeddings across modalities	Sufficient labeled data; embeddings capture complementary but aligned semantics
Instance-Level Fusion	Weakly supervised learning	Aggregation of multimodal instances under slide-level supervision	Local correspondence latent; global labels constrain instance relevance
Consistency-Based Fusion	Semi-supervised learning	Alignment of modality-specific representations via consistency or smoothness constraints	Modalities share underlying phenotype manifold despite partial observation
Pretext-Aligned Fusion	Self-supervised learning	Cross-modal or cross-scale pretext objectives defining shared latent space	Structural correlations encode transferable multimodal priors
Decision-Level Fusion	Any supervision regime	Combination of modality-specific predictions under shared loss	Modalities conditionally independent given outcome

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

6. Emerging Paradigms

Recent developments in histomics reflect a shift from task-specific modeling toward scalable representation learning, interactive inference, and generative data augmentation. These frameworks address fundamental limitations of conventional pipelines, including annotation cost, domain sensitivity, and feature instability, while enabling more generalizable and adaptive histomic analysis.

6.1. Foundation Models and Representation Learning

Foundation models introduce a representation-centric framework in histomics, where large-scale self-supervised pretraining learns generic morphological embeddings transferable across tasks and datasets. These models are optimized over extensive collections of unlabeled WSIs, capturing multiscale tissue organization, cellular context, and architectural regularities without reliance on task-specific annotations.

From a modeling perspective, pretrained histology backbones define a shared latent space that decouples representation learning from downstream supervision, reducing sample complexity and improving robustness under distribution shift. Fine-tuning or linear probing over these embeddings enables efficient adaptation to classification, stratification, or prognostic tasks while preserving learned morphological priors. This paradigm reframes histomics as representation reuse rather than feature engineering or task-isolated optimization.

6.2. Interactive and Active Learning

Interactive and active learning addresses annotation bottlenecks by integrating human expertise directly into the model training loop [35]. In this, model uncertainty, disagreement, or expected information gain guides selective annotation of histomic instances, prioritizing regions or structures with maximal impact on decision boundaries.

Formally, learning proceeds under an adaptive sampling strategy, where annotation and model refinement are interleaved to minimize labeling effort while maximizing performance gains. Human-in-the-loop interaction further enables correction of systematic model errors, incorporation of domain knowledge, and iterative refinement of segmentation and classification outputs [36]. Such frameworks support efficient convergence under constrained annotation budgets while maintaining interpretability and expert oversight.

6.3. Virtual Staining and Synthetic Augmentation

Generative modeling introduces synthetic data generation as a mechanism to enhance histomic robustness and generalization. Virtual staining models learn mappings between unstained or alternative-stained tissue images and target histochemical appearances, enabling stain-invariant feature extraction and reducing dependency on physical staining protocols [37].

Beyond stain translation, generative augmentation synthesizes morphologically plausible variations in cellular density, texture, and tissue organization, enriching training distributions and mitigating overfitting. When incorporated into histomic pipelines, synthetic data act as regularizers on feature learning, improving stability across staining conditions and acquisition domains. These approaches expand the effective support of histomic feature spaces while preserving biologically consistent structure.

Table 8 presents emerging paradigms in histomics and the methodological limitations they address.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

Table 8. Emerging paradigms in histomics and the methodological limitations they address.

Emerging Paradigm	Primary Limitations Addressed	Technical Impact on Histomic Modeling
Foundation models and large-scale representation learning	Dependence on task-specific features, limited labeled data, poor cross-dataset transferability	Decouples representation learning from supervision; improves sample efficiency, robustness to domain shift, and reuse across tasks
Interactive and active learning	High annotation cost, inefficient labeling, slow model refinement	Introduces adaptive sampling and human-in-the-loop optimization, reducing labeling burden while stabilizing decision boundaries
Virtual staining	Stain variability, protocol dependence, scanner-induced color shift	Enables stain-invariant representations and harmonized feature extraction across acquisition conditions
Synthetic data augmentation	Overfitting, limited morphological diversity, sensitivity to rare phenotypes	Expands effective training distributions and regularizes feature learning through biologically plausible variation

7. Related Work

Histomics research spans feature-based modeling, hybrid DL, weakly supervised learning, multimodal integration, foundation models, and generative approaches as given in Table 9 and Table 10. Early work focused on handcrafted descriptors derived from segmentation, including morphology, texture, and spatial organization. Toolkits such as HistomicsTK [38] and CellProfiler [39]-based pipelines enabled systematic feature extraction, but performance remained highly dependent on segmentation quality and dataset-specific tuning.

Hybrid frameworks integrate handcrafted features with deep representations. Methods such as HistoEM [40], DeepCMorph [41], Histo-Miner [42], and CellViT [43] combine explicit morphological descriptors with CNN or transformer architectures. These approaches improve predictive performance while retaining partial interpretability, but remain constrained by segmentation errors, feature instability, and limited dataset diversity.

Weakly supervised learning has become central due to the lack of dense annotations. MIL frameworks, including DSMIL [44], and TransMIL [45], enable slide-level prediction for classification, survival analysis, and biomarker discovery. These models reduce annotation cost but introduce uncertainty in instance-level attribution and sensitivity to noisy or redundant features.

Recent work emphasizes multimodal integration. Models such as HE2RNA [46] link histological morphology with gene expression, while spatial omics approaches align histomic features with spatial transcriptomics data. These methods improve biological interpretability and predictive performance, but face challenges in modality alignment, resolution mismatch, and limited paired datasets.

A major shift in recent literature is toward foundation models and self-supervised learning. Methods such as CTransPath [47], RetCCL [48], Virchow [49], Virchow2, and Prov-GigaPath [50] learn transferable representations from large-scale WSI datasets. These models significantly improve downstream performance across tasks, reducing reliance on handcrafted features. However, studies show that performance varies under domain shift and remains sensitive to dataset composition, staining variability, and cohort bias.

Multimodal foundation models extend this paradigm by integrating visual, textual, and molecular data. Frameworks such as TITAN [51], CPath-Omni [52], and PathCLIP [53] enable cross-modal

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

learning, retrieval, and report generation. While promising, these models face challenges in aligning heterogeneous modalities and require large, well-annotated datasets.

Generative modeling introduces virtual staining and synthetic augmentation. Methods such as StainGAN [54], CycleGAN [55]-based stain transfer, diffusion-based models, and HistoXGAN [56] improve robustness to staining variability and expand training distributions. However, synthetic data may introduce artifacts and bias if not carefully validated.

Graph-based histomics models represent tissue as relational structures. Approaches such as cell-graph GNNs [57], and microenvironment-aware graph models capture higher-order spatial interactions. These models improve modeling of tissue architecture but depend heavily on segmentation accuracy and graph construction quality.

Vision-language models represent an emerging direction. Frameworks such as WsiCaption [58], PathCLIP, BioViL [59], and MedCLIP align histology images with textual descriptions or reports. These models enable weak supervision and multimodal reasoning, but alignment with domain-specific terminology remains limited.

Table 9. Comparative analysis of feature-based, hybrid, and weakly supervised histomics studies.

Study	Dataset	Features	Model	Task	Key Limitation
HistomicsTK	TCGA	Morphology + Texture	ML	Classification	Segmentation sensitivity
CellProfiler pipelines	Multi-dataset	Morphology	ML	Feature analysis	Limited scalability
HistoEM	Prostate WSI	Morphology + DL	CNN	Classification	Segmentation dependency
DeepCMorph	TCGA	Morphology-aware DL	CNN	Classification	Data-intensive
Histo-Miner	cSCC	Hybrid features	CNN + ViT	Therapy prediction	Limited dataset size
CellViT	Multi-cohort	Cell-level features	ViT	Segmentation	High computation
DSMIL	Multi-site	Instance embeddings	MIL	Classification	Limited interpretability
TransMIL	TCGA	Patch features	Transformer MIL	Classification	Noise sensitivity
HE2RNA	TCGA	Morphology → genomics	DL	Gene prediction	Modality mismatch
CTransPath	Multi-dataset	SSL embeddings	Transformer	Multi-task	Limited interpretability
RetCCL	Multi-dataset	Contrastive features	SSL	Representation learning	Weak phenotype alignment
UNI	Large-scale	Learned embeddings	SSL	Multi-task	Dataset bias
Virchow / Virchow2	Multi-cohort	Foundation embeddings	Transformer	Pan-cancer tasks	Generalization variability
Prov-GigaPath	Multi-cohort	Foundation features	Transformer	Multi-task	Data-intensive training

Table 10. Comparative analysis of foundation, generative, graph-based, and vision-language histomics models.

Study	Dataset	Features	Model	Task	Key Limitation
TITAN	WSI + reports	Image + text	Multimodal FM	Multi-task	Alignment complexity
CPath-Omni	Multi-modal	Image + molecular	Multimodal FM	Prognosis	Data scarcity
PathCLIP	Multi-modal	Image + text	Transformer	Retrieval	Weak domain alignment
StainGAN	Multi-stain	Image translation	GAN	Normalization	Artifact risk
HistoXGAN	Histology	Synthetic features	GAN	Reconstruction	Synthetic bias
BioViL / MedCLIP	Multi-modal	Visual-text embeddings	Transformer	Multimodal tasks	Limited pathology specificity
Cell-graph GNN	Multi-site	Cell graphs	GNN	Prognosis	Segmentation dependency

Across these studies, several consistent patterns emerge. Most approaches rely on limited datasets, predominantly TCGA [60], with insufficient external validation. Feature extraction remains tightly

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

coupled to segmentation quality, yet error propagation is rarely quantified. Feature definitions and implementations vary across studies, reducing reproducibility.

While foundation models improve representation learning, they introduce challenges related to domain bias, interpretability, and sensitivity to dataset composition. Weakly supervised and multimodal approaches improve data efficiency but introduce uncertainty in alignment and representation. Overall, the absence of standardized benchmarks, cross-dataset validation, and reproducible evaluation protocols remains a primary limitation in current histomics research.

8. Discussion

Computational pathology lacks a consistent representation of tissue phenotype. Histopathological diagnosis relies on structured, multi-scale features (morphology, texture, spatial organization), whereas computational models either encode these features implicitly (DL) or represent them explicitly without standardization. This creates a mismatch between computational representations and clinically meaningful features.

Latent embeddings do not preserve explicit correspondence to morphological variables, while explicit descriptors do not remain invariant across pipelines or datasets. As a result, phenotype representations are neither consistently interpretable nor reproducible.

Weak supervision and multimodal modeling further constrain representation fidelity. Slide-level labeling does not define instance-level relevance, and multimodal integration assumes correspondence across heterogeneous data sources. These factors limit preservation of coherent phenotype structure across scales.

Clinical interpretation is hierarchical and context-dependent, integrating spatial and multi-scale cues. Computational representations either isolate features or compress them through aggregation, reducing structural context and limiting correspondence with diagnostic reasoning.

The central issue is the absence of a stable, interpretable, and invariant phenotype representation, leading to models that capture dataset-specific correlations rather than consistent biological structure.

9. Limitations and Future Work

Histomics has evolved from handcrafted feature pipelines to representation learning and multimodal integration; however, limitations in histomic pipelines can be interpreted as structured perturbations within the phenotype representation space as shown in Table 11.

Let the total representation error be decomposed as:

$$\epsilon_{total} = \epsilon_{seg} + \epsilon_{feat} + \epsilon_{domain} + \epsilon_{agg} \tag{7}$$

where ϵ_{seg} arises from segmentation errors, ϵ_{feat} from feature instability, ϵ_{domain} from distributional shift across datasets, and ϵ_{agg} from aggregation inconsistencies in weakly supervised settings. This formulation reframes pipeline-level failures as representation-level inconsistencies that directly affect stability and generalization.

Table 11. Decomposition of representation-level errors in histomics.

Error Term	Source	Impact
ϵ_{seg}	Segmentation	Distorted morphology and spatial features
ϵ_{feat}	Feature extraction	Variability, redundancy, instability
ϵ_{domain}	Dataset shift	Reduced cross-dataset generalization
ϵ_{agg}	Aggregation (MIL, pooling)	Inconsistent slide-level representations

A central issue is the dependence on segmentation quality. Errors in nuclear detection or region delineation distort morphological and spatial descriptors and propagate through downstream models. Despite this dependency, segmentation uncertainty is rarely modeled, limiting confidence in extracted phenotypes.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

Feature instability further constrains reproducibility. Histomic descriptors vary with staining, preprocessing, and parameter choices, while strong correlations introduce redundancy and violate modeling assumptions. Classical models are sensitive to this instability, whereas deep models absorb it into latent representations, reducing interpretability without resolving underlying variability.

Domain shift remains a dominant bottleneck. Variations in staining protocols, scanners, and cohort composition alter feature distributions, causing models trained on datasets such as TCGA to degrade under external evaluation. Evidence from foundation model analyses indicates that large-scale representations often encode site-specific biases, suggesting that scale improves performance but does not ensure generalization.

Weak supervision and multimodal integration introduce additional uncertainty. Slide-level labeling obscures instance-level relevance, and multimodal learning assumes alignment across modalities with mismatched resolution and noise, often leading to inconsistent representations.

Generative augmentation expands data diversity but introduces risks of synthetic bias when realism is not validated.

A fundamental limitation across methods is the mismatch between feature properties and model assumptions. High-dimensional, correlated descriptors challenge classical approaches, while deep models encode morphology implicitly, limiting interpretability and stability.

Beyond modeling, current literature is constrained by evaluation practices, including reliance on single datasets, lack of external validation, and absence of uncertainty quantification.

These limitations can be summarized as core failure modes in histomics pipelines (see Table 12).

Table 12. Core failure modes in histomics pipelines and their implications.

Failure Mode	Source	Impact	Open Gap
Segmentation dependence	Preprocessing / detection	Distorted morphology and spatial features	Lack of uncertainty-aware segmentation
Feature instability	Staining, parameter choices	Poor reproducibility and variability	No standardized feature definitions
Domain shift	Scanner, cohort variation	Degraded cross-dataset performance	Limited invariant representations
Weak supervision	Slide-level labels	Uncertain instance attribution	Limited interpretability of attention weights
Multimodal misalignment	Resolution mismatch	Inconsistent feature correspondence	Lack of geometry-aware alignment

Thus, future histomic systems should model phenotype representations probabilistically to capture uncertainty arising from segmentation, feature extraction, and domain variability.

This includes: (i) uncertainty-aware histomic pipelines that explicitly propagate segmentation and feature-level uncertainty through downstream models, enabling confidence-aware predictions; (ii) invariant representation learning that disentangles morphology from acquisition-specific factors (e.g., stain, scanner, site) using causal or contrastive formulations; (iii) geometry-aware multimodal alignment that models partial and multi-scale correspondence between histology and spatial omics, rather than assuming direct alignment; (iv) hybrid interpretable–latent frameworks that enforce phenotype consistency constraints within deep representations to bridge explicit histomic descriptors and learned embeddings; (v) distributionally robust learning strategies that optimize performance under domain shift, including worst-case and out-of-distribution generalization; and (vi) standardized, large-scale benchmarking protocols with cross-cohort evaluation, uncertainty calibration, and failure-case analysis.

Advancing these directions is essential for developing histomic systems that are not only accurate but also stable, interpretable, and clinically reliable.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

10. Conclusions

Computational pathology has increasingly shifted toward end-to-end DL pipelines that achieve strong predictive performance but lack explicit and consistent representations of tissue phenotype. This limitation affects interpretability, reproducibility, and cross-dataset generalization, motivating the need for structured representation frameworks.

This work positions histomics as an intermediate phenotype representation layer that bridges raw image data and downstream prediction. By explicitly encoding morphological, spatial, and multi-scale tissue characteristics, histomic representations provide a structured alternative to purely latent embeddings and enable biologically grounded analysis.

A key insight developed in this review is that many persistent limitations in computational pathology arise from misalignment between latent representations learned by deep models and explicit phenotype descriptors. This representation gap provides a unified perspective for understanding failure modes such as segmentation dependence, feature instability, domain shift, and aggregation inconsistency, linking them to representation-level mismatch rather than isolated pipeline components.

While recent advances in DL and foundation models improve performance, they do not resolve inconsistencies in representation or guarantee robustness under distributional variation. Histomics addresses this by constraining representations to interpretable phenotype space and supporting more stable integration with clinical and molecular data.

Future progress requires shifting from task-specific optimization toward representation-centric design. This includes developing uncertainty-aware pipelines, learning invariant phenotype descriptors under domain variability, enabling geometry-aware multimodal alignment, and establishing standardized evaluation protocols for cross-dataset validation. Hybrid frameworks that explicitly couple latent and phenotype representations are likely to play a central role in reducing the representation gap.

This review has limitations. It emphasizes feature-centric and hybrid frameworks and does not fully capture purely end-to-end approaches. In addition, variability in feature definitions and reporting practices limits direct comparison across studies.

Overall, histomics provides a principled representation framework that addresses core limitations of current computational pathology systems and supports the development of models that are not only accurate, but also interpretable, stable, and clinically reliable.

References

1. Gurcan, M.N.; Boucheron, L.E.; Can, A.; Madabhushi, A.; Rajpoot, N.M.; Yener, B. Histopathological image analysis: A review. *IEEE reviews in biomedical engineering* **2009**, *2*, 147–171.
2. Pavone, A.M.; Giannone, A.G.; Cabibi, D.; D'Aprile, S.; Denaro, S.; Salvaggio, G.; Parenti, R.; Yezzi, A.; Comelli, A. Digital pathology: A comprehensive review of open-source histological segmentation software. *BioMedInformatics* **2024**, *4*, 173–196.
3. Li, X.; Li, C.; Rahaman, M.M.; Sun, H.; Li, X.; Wu, J.; Yao, Y.; Grzegorzek, M. A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches. *Artificial Intelligence Review* **2022**, *55*, 4809–4878.
4. Irshad, H.; Veillard, A.; Roux, L.; Racoceanu, D. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. *IEEE reviews in biomedical engineering* **2013**, *7*, 97–114.
5. Otálora, S.; Atzori, M.; Andrearczyk, V.; Khan, A.; Müller, H. Staining invariant features for improving generalization of deep convolutional neural networks in computational pathology. *Frontiers in bioengineering and biotechnology* **2019**, *7*, 474781.
6. Rundo, L.; Militello, C. Image biomarkers and explainable AI: handcrafted features versus deep learned features. *European Radiology Experimental* **2024**, *8*, 130.
7. Pan, L.; Feng, Z.; Peng, S. A review of machine learning approaches, challenges and prospects for computational tumor pathology. *arXiv preprint arXiv:2206.01728* **2022**.
8. Orchard, G.; Nation, B. *Histopathology*; OUP Oxford, 2012.
9. Choudhury, P.; Dhua, A.K.; Goel, P. Inter-observer and intra-observer reliability. In *Biostatistical Methods and Applications in Health Research*; Routledge, 2025; pp. 267–278.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

10. Ghaznavi, F.; Evans, A.; Madabhushi, A.; Feldman, M. Digital imaging in pathology: whole-slide imaging and beyond. *Annual Review of Pathology: Mechanisms of Disease* **2013**, *8*, 331–359.
11. Foroughi Pour, A.; White, B.S.; Park, J.; Sheridan, T.B.; Chuang, J.H. Deep learning features encode interpretable morphologies within histological images. *Scientific Reports* **2022**, *12*, 9428.
12. Ariotta, V.; Lehtonen, O.; Salloum, S.; Micoli, G.; Lavikka, K.; Rantanen, V.; Hynninen, J.; Virtanen, A.; Hautaniemi, S. H&E image analysis pipeline for quantifying morphological features. *Journal of pathology informatics* **2023**, *14*, 100339.
13. Madusanka, N.; Padmanabha, P.; Guruge, K.; Lee, B.i. Structure-Preserving Histopathological Stain Normalization via Attention-Guided Residual Learning. *Bioengineering* **2025**, *12*, 950.
14. Raman, S.; Singh, S.; Pécot, T.; Caserta, E.; Huang, K.; Rittscher, J.; Leone, G.; Machiraju, R. Capturing variations in nuclear phenotypes. *Journal of computational science* **2019**, *36*, 101014.
15. Krikid, F.; Rositi, H.; Vacavant, A. State-of-the-art deep learning methods for microscopic image segmentation: Applications to cells, nuclei, and tissues. *Journal of Imaging* **2024**, *10*, 311.
16. Huang, X.; Li, Z.; Zhang, M.; Gao, S. Fusing hand-crafted and deep-learning features in a convolutional neural network model to identify prostate cancer in pathology images. *Frontiers in Oncology* **2022**, *12*, 994950.
17. Fan, Z.; Lu, C.; Yue, J.; Anastasio, M.; Sun, L.; Wang, X.; Li, H. Enhancing deep learning interpretability for hand-crafted feature-guided histologic image classification via weak-to-strong generalization. *Journal of Medical Imaging* **2026**, *13*, 017502–017502.
18. Lu, C.; Fan, Z.; Wang, Z.; Anastasio, M.; Sun, L.; Wang, X.; Li, H. Hand-crafted feature-guided histologic image classification via weak-to-strong generalization. In Proceedings of the Medical Imaging 2025: Digital and Computational Pathology. SPIE, 2025, Vol. 13413, pp. 146–151.
19. Machiraju, G.B. *Going Beyond Salience: Deep Learning for the Discovery of Meaningful Pathomic Markers*; Stanford University, 2024.
20. Obeid, A.; Boumaraf, S.; Sohail, A.; Hassan, T.; Javed, S.; Dias, J.; Bennamoun, M.; Werghi, N. Advancing histopathology with deep learning under data scarcity: A decade in review. *arXiv arXiv:2410.19820* **2024**.
21. Jiang, C.; Hou, X.; Kondepudi, A.; Chowdury, A.; Freudiger, C.W.; Orringer, D.A.; Lee, H.; Hollon, T.C. Hierarchical discriminative learning improves visual representations of biomedical microscopy. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 19798–19808.
22. Sharma, S.; Mehra, R. Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight. *Journal of digital imaging* **2020**, *33*, 632–654.
23. Ma, S.; Huang, J. Penalized feature selection and classification in bioinformatics. *Briefings in bioinformatics* **2008**, *9*, 392–403.
24. Rony, J.; Belharbi, S.; Dolz, J.; Ayed, I.B.; McCaffrey, L.; Granger, E. Deep weakly-supervised learning methods for classification and localization in histology images: a survey. *arXiv preprint arXiv:1909.03354* **2019**.
25. Ran, L.; Li, Y.; Liang, G.; Zhang, Y. Pseudo labeling methods for semi-supervised semantic segmentation: A review and future perspectives. *IEEE Transactions on Circuits and Systems for Video Technology* **2024**, *35*, 3054–3080.
26. Kumar, P.; Rawat, P.; Chauhan, S. Contrastive self-supervised learning: review, progress, challenges and future research directions. *International Journal of Multimedia Information Retrieval* **2022**, *11*, 461–488.
27. Zhang, Y.; Zhang, X.; Qi, X.; Wu, X.; Chen, F.; Yang, G.; Fu, H. Content Generation Models in Computational Pathology: A Comprehensive Survey on Methods, Applications, and Challenges. *IEEE reviews in biomedical engineering* **2025**.
28. Tan, X.; Su, A.T.; Hajiabadi, H.; Tran, M.; Nguyen, Q. Applying machine learning for integration of multi-modal genomics data and imaging data to quantify heterogeneity in tumour tissues. *Artificial neural networks* **2020**, pp. 209–228.
29. Vincenzo, M.D. Review on multi-modal AI models to integrate imaging and omics data. Master's thesis, 2024.
30. Border, S.P. Increasing the Depth and Quality of Computational Pathology Research Through Interactive Tools and Multi-Modal Data. PhD thesis, State University of New York at Buffalo, 2025.
31. Chen, R.J.; Lu, M.Y.; Wang, J.; Williamson, D.F.; Rodig, S.J.; Lindeman, N.I.; Mahmood, F. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE transactions on medical imaging* **2020**, *41*, 757–770.
32. Lee, Y.; Lee, M.; Shin, Y.; Kim, K.; Kim, T. Spatial omics in clinical research: a comprehensive review of technologies and guidelines for applications. *International Journal of Molecular Sciences* **2025**, *26*, 3949.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

33. Haghighat, F.; Nemati, Z.; Rambodrad, A.; Negareshifard, P.; Jafari, E. Multimodal deep learning and data fusion in precision breast oncology: clinical Applications, fusion Strategies, and future directions. *InfoScience Trends* **2025**, *2*, 81–115.
34. Alshohoumi, F.; Al-Hamdani, A. Comparative Evaluation of Fusion Strategies Using Multi-Pretrained Deep Learning Fusion-Based (MPDLF) Model for Histopathology Image Classification. *Applied Sciences* **2026**, *16*, 1964.
35. Monarch, R.M. *Human-in-the-Loop Machine Learning: Active learning and annotation for human-centered AI*; Simon and Schuster, 2021.
36. Kumar, S.; Datta, S.; Singh, V.; Datta, D.; Singh, S.K.; Sharma, R. Applications, challenges, and future directions of human-in-the-loop learning. *IEEE Access* **2024**, *12*, 75735–75760.
37. Hussain, M.A.; Waris, M.A.; Akram, M.U.; Khan, M.J.; Asaf, M.Z.; Javaid, A.; Gilani, S.O.; Hazzazi, F. ViT-Stain: Vision transformer-driven virtual staining for skin histopathology via global contextual learning. *PloS one* **2026**, *21*, e0341311.
38. Pourakpour, F.; Szölgyén, Á.; Nateghi, R.; Gutman, D.A.; Manthey, D.; Cooper, L.A. HistomicsTK: A Python toolkit for pathology image analysis algorithms. *SoftwareX* **2025**, *31*, 102318.
39. Carpenter, A.E.; Jones, T.R.; Lamprecht, M.R.; Clarke, C.; Kang, I.H.; Friman, O.; Guertin, D.A.; Chang, J.H.; Lindquist, R.A.; Moffat, J.; et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome biology* **2006**, *7*, R100.
40. Ferrero, A.; Ghelichkhan, E.; Manoochehri, H.; Ho, M.M.; Albertson, D.J.; Brintz, B.J.; Tasdizen, T.; Whitaker, R.T.; Knudsen, B.S. HistoEM: A pathologist-guided and explainable workflow using histogram embedding for gland classification. *Modern Pathology* **2024**, *37*, 100447.
41. Ignatov, A.; Yates, J.; Boeva, V. Histopathological image classification with cell morphology aware deep neural networks. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 6913–6925.
42. Sancéré, L.; Lorenz, C.; Helbig, D.; Persa, O.D.; Dengler, S.; Kreuter, A.; Laimer, M.; Lang, R.; Fröhlich, A.; Landsberg, J.; et al. Histo-Miner: Deep learning based tissue features extraction pipeline from H&E whole slide images of cutaneous squamous cell carcinoma. *Plos Computational Biology* **2026**, *22*, e1013907.
43. Hörst, F.; Rempe, M.; Heine, L.; Seibold, C.; Keyl, J.; Baldini, G.; Ugurel, S.; Siveke, J.; Grünwald, B.; Egger, J.; et al. Cellvit: Vision transformers for precise cell segmentation and classification. *Medical image analysis* **2024**, *94*, 103143.
44. Li, B.; Li, Y.; Eliceiri, K.W. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 14318–14328.
45. Shao, Z.; Bian, H.; Chen, Y.; Wang, Y.; Zhang, J.; Ji, X.; et al. Transmil: Transformer based correlated multiple instance learning for whole slide image classification. *Advances in neural information processing systems* **2021**, *34*, 2136–2147.
46. Pronier, E.; Schmauch, B.; Romagnoni, A.; Saillard, C.; Maillé, P.; Calderaro, J.; Sefta, M.; Toldo, S.; Zaslavskiy, M.; Clozel, T.; et al. HE2RNA: a deep learning model for transcriptomic learning from digital pathology. *Cancer Res* **2020**, *80*, 2105–2105.
47. Wang, X.; Yang, S.; Zhang, J.; Wang, M.; Zhang, J.; Yang, W.; Huang, J.; Han, X. Transformer-based unsupervised contrastive learning for histopathological image classification. *Medical image analysis* **2022**, *81*, 102559.
48. Wang, X.; Du, Y.; Yang, S.; Zhang, J.; Wang, M.; Zhang, J.; Yang, W.; Huang, J.; Han, X. RetCCL: Clustering-guided contrastive learning for whole-slide image retrieval. *Medical image analysis* **2023**, *83*, 102645.
49. Vorontsov, E.; Bozkurt, A.; Casson, A.; Shaikovski, G.; Zelechowski, M.; Severson, K.; Zimmermann, E.; Hall, J.; Tenenholtz, N.; Fusi, N.; et al. A foundation model for clinical-grade computational pathology and rare cancers detection. *Nature medicine* **2024**, *30*, 2924–2935.
50. Xu, H.; Usuyama, N.; Bagga, J.; Zhang, S.; Rao, R.; Naumann, T.; Wong, C.; Gero, Z.; González, J.; Gu, Y.; et al. A whole-slide foundation model for digital pathology from real-world data. *Nature* **2024**, *630*, 181–188.
51. 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.; et al. A multimodal whole-slide foundation model for pathology. *Nature medicine* **2025**, pp. 1–13.
52. Sun, Y.; Si, Y.; Zhu, C.; Gong, X.; Zhang, K.; Chen, P.; Zhang, Y.; Shui, Z.; Lin, T.; Yang, L. Cpath-omni: A unified multimodal foundation model for patch and whole slide image analysis in computational pathology. In Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2025, pp. 10360–10371.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 April 2026 doi:10.20944/preprints202604.1359.v1

53. He, F.; Liu, K.; Yang, Z.; Chen, Y.; Hammer, R.D.; Xu, D.; Popescu, M. PathCLIP: Detection of genes and gene relations from biological pathway figures through image-text contrastive learning. *IEEE journal of biomedical and health informatics* **2024**, *28*, 5007–5019.
54. Shaban, M.T.; Baur, C.; Navab, N.; Albarqouni, S. Staingan: Stain style transfer for digital histological images. In Proceedings of the 2019 Ieee 16th international symposium on biomedical imaging (Isbi 2019). IEEE, 2019, pp. 953–956.
55. Nerrienet, N.; Peyret, R.; Sockeel, M.; Sockeel, S. Standardized cyclegan training for unsupervised stain adaptation in invasive carcinoma classification for breast histopathology. *Journal of Medical Imaging* **2023**, *10*, 067502–067502.
56. Howard, F.M.; Hieromnimon, H.M.; Ramesh, S.; Dolezal, J.; Kochanny, S.; Zhang, Q.; Feiger, B.; Peterson, J.; Fan, C.; Perou, C.M.; et al. Generative adversarial networks accurately reconstruct pan-cancer histology from pathologic, genomic, and radiographic latent features. *Science Advances* **2024**, *10*, eadq0856.
57. Yuan, Y.; Zhao, Z.; Fang, X.; Zhang, Q.; Zhong, W.; Xu, M.; Li, G.; Jiao, R.; Yu, H.; Wang, R.; et al. Cell graph analysis in hepatocellular carcinoma: predicting local recurrence and identifying spatial relationship biomarkers. *npj Precision Oncology* **2025**, *9*, 261.
58. Chen, P.; Li, H.; Zhu, C.; Zheng, S.; Shui, Z.; Yang, L. Wsicaption: Multiple instance generation of pathology reports for gigapixel whole-slide images. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2024, pp. 546–556.
59. Bannur, S.; Hyland, S.; Liu, Q.; Perez-Garcia, F.; Ilse, M.; Castro, D.C.; Boecking, B.; Sharma, H.; Bouzid, K.; Thieme, A.; et al. Learning to exploit temporal structure for biomedical vision-language processing. In Proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2023, pp. 15016–15027.
60. Tomczak, K.; Czerwińska, P.; Wiznerowicz, M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary Oncology/Współczesna Onkologia* **2015**, *2015*, 68–77.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.