

INTEGRATION OF HISTOPATHOLOGY AND SPATIAL TRANSCRIPTOMICS FOR CANCER CLASSIFICATION: A SYSTEMATIC REVIEW

Aritraa Hazra^{1*}, Vaishali D. Kotasthane², Dhananjay S. Kotasthane³, Ashtha Arya⁴

¹Assistant Professor, Department of Community Medicine, Barasat Government Medical College and Hospital, Barasat, North 24 Parganas, West Bengal, India. Email ID - aritraahazra@gmail.com

²Professor, Department of Oncopathology, Shri Guru Ram Rai Institute of Medical & Health Sciences, Dehradun, Uttarakhand, India

³Professor & Head, Department of Oncopathology, Shri Guru Ram Rai Institute of Medical & Health Sciences, Dehradun, Uttarakhand, India

⁴Professor, Department of Conservative Dentistry and Endodontics, SGT Dental College, Hospital and Research Institute, SGT University, Gurugram, Haryana, India

*Corresponding Author: Aritraa Hazra, aritraahazra@gmail.com

ABSTRACT

Background: Conventional histopathology remains fundamental to cancer diagnosis, but morphology alone cannot fully resolve intratumoral molecular heterogeneity. Spatial transcriptomics (ST) preserves gene-expression information within tissue architecture and provides a route to connect microscopic phenotype with molecular state. Recent artificial-intelligence methods combine hematoxylin-and-eosin (H&E) images with spatial transcriptomic profiles to predict gene expression, classify tissue domains, delineate molecular subtypes, and characterize tumor-immune and clonal heterogeneity.

Methods: MEDLINE/PubMed, Embase, Scopus, Web of Science, and IEEE Xplore were searched from January 2018 through 30 June 2026. Citation and reference-list searching supplemented database retrieval. Eligible studies analyzed human cancer tissue and jointly incorporated histopathology with spatially resolved transcriptomic information for cancer classification, tissue-domain identification, molecular subtyping, cell-state assignment, clonal mapping, or spatial gene-expression prediction relevant to cancer classification. Because of heterogeneity in cancer types, ST platforms, model architectures, target genes, spatial resolution, and performance metrics, a narrative synthesis was performed.

Results: A total of 2,756 records were identified: 2,682 from databases and 74 through citation/reference searching. After removal of 689 duplicates, 2,067 records were screened and 1,781 excluded. Of 286 reports sought for retrieval, 17 were unavailable. Two hundred sixty-nine full texts were assessed and 243 excluded, leaving 26 studies for qualitative synthesis. Breast cancer was the most frequently represented malignancy. Across the evidence base, convolutional, transformer, graph-neural-network, contrastive-learning, and foundation-model approaches increasingly linked morphology with spatial gene expression and biologically meaningful tumor domains. Independent benchmarking showed that no single method dominated accuracy, generalizability, computational efficiency, and translational performance. Direct clinically oriented cancer-classification endpoints remained less common than gene-expression prediction.

Conclusions: Integration of histopathology and spatial transcriptomics can enrich cancer classification by linking morphology with molecular and spatial phenotypes. Evidence is strongest for spatial gene-expression inference and tissue-domain characterization, while direct multicenter validation of molecular subtype, prognosis, and treatment-response classification remains limited. Standardized benchmarking, patient-level external validation, and prospective clinical evaluation are required before routine diagnostic adoption.

KEYWORDS: spatial transcriptomics; histopathology; cancer classification; digital pathology; artificial intelligence; deep learning; tumor microenvironment; spatial omics; precision oncology

INTRODUCTION

Histopathological examination of H&E-stained tissue remains central to cancer diagnosis, grading, and classification. The microscopic phenotype reflects cellular differentiation, stromal organization, immune infiltration, necrosis, vascularity, and architectural disruption. Nevertheless, morphologically similar tumors may differ markedly in gene expression, signaling pathways, clonal composition, treatment response, and prognosis.

Bulk genomic and transcriptomic assays have improved molecular classification but disrupt the spatial relationships among malignant, stromal, vascular, and immune components. Spatial transcriptomics addresses this limitation by measuring RNA abundance while retaining tissue location. ST therefore creates an opportunity to link molecular signals directly to tissue architecture and to ask whether routine morphology can serve as a surrogate for spatially resolved molecular states.

An early proof of concept was provided by ST-Net, which matched 30,612 spatial transcriptomic measurements with breast-cancer histology and demonstrated that spatial expression of more than 100 genes could be predicted

from H&E images [1]. SpaGCN subsequently combined gene expression, spatial coordinates, and histologic information within a graph-convolution framework to identify coherent spatial domains and spatially variable genes [2]. Hist2ST [4], THItGene [9], iStar [11], and other methods progressively incorporated transformers, graph neural networks, multiscale morphology, and spatial context.

The field has since moved beyond spot-level regression toward clinically meaningful tasks such as tumor-domain recognition, molecular-state prediction, immune-niche characterization, clonal mapping, prognostic stratification, and near-single-cell spatial inference. More recent approaches also use pathology foundation models, large multimodal pretraining datasets, and visual-omics contrastive learning [18–20].

However, the rapid proliferation of algorithms, reusable public datasets, spatial platforms, preprocessing strategies, and evaluation metrics makes the translational maturity of the field difficult to judge. This systematic review therefore evaluated how histopathology and spatial transcriptomics have been integrated for cancer classification and assessed methodological strengths, bias, generalizability, and readiness for clinical pathology workflows.

MATERIALS AND METHODS

Study design and reporting

This systematic review was designed in accordance with the PRISMA 2020 framework. The review question, eligibility criteria, information sources, outcomes, and synthesis plan were defined before evidence synthesis. The review was not prospectively registered.

Review question

Among human cancers, how effectively can histopathological information integrated with spatial transcriptomics improve cancer classification, molecular characterization, tissue-domain recognition, or clinically relevant stratification?

Eligibility criteria

Domain	Eligibility criteria
Population	Human malignant or premalignant tissue.
Histology	H&E or equivalent diagnostic histopathology image.
Molecular modality	Spatially resolved transcriptomics or spatial RNA profiling.
Integration	Joint computational or analytical use of morphology and spatial transcriptomic data.
Outcomes	Cancer classification, tumor/non-tumor discrimination, spatial-domain classification, molecular subtype, cell-state classification, clonal assignment, biomarker inference, gene-expression prediction relevant to cancer classification, or prognostic/therapeutic stratification.
Study types	Original methodological, translational, benchmarking, or computational studies.
Publication types	Peer-reviewed journal studies, peer-reviewed conference reports, and full-text preprints with sufficient quantitative validation.
Period	January 2018 through 30 June 2026.
Exclusions	Animal-only work; bulk RNA sequencing without spatial validation; morphology-free ST analysis; reviews/editorials; non-cancer reports without evaluable cancer data; or insufficient quantitative validation.

Information sources and search strategy

MEDLINE/PubMed, Embase, Scopus, Web of Science, and IEEE Xplore were searched. Reference lists and forward citations of eligible studies were additionally reviewed.

Representative search syntax: ("spatial transcriptomics" OR "spatially resolved transcriptomics" OR "spatial gene expression") AND (histopathology OR histology OR H&E OR "whole slide image" OR "digital pathology") AND (cancer OR carcinoma OR tumour OR tumor OR neoplasm) AND (classification OR subtype OR prediction OR "deep learning" OR "machine learning" OR transformer OR "graph neural network"). Additional searches combined molecular subtype, tumor microenvironment, spatial domain, gene-expression prediction, multimodal learning, contrastive learning, and foundation model.

Study selection

After deduplication, titles and abstracts were screened against prespecified eligibility criteria, followed by full-text assessment. Reports using the same underlying dataset were retained when they represented a distinct algorithm or biologically different analysis; publications without an incremental contribution were excluded. Reasons for full-text exclusion were recorded.

Data extraction

Extracted variables included year, cancer type, specimen count, ST platform, histology type, computational architecture, image representation, molecular target, validation strategy, external-validation status, classification target, performance metric, and translational application.

Risk-of-bias assessment

Because no single established risk-of-bias tool fully addresses multimodal histology–ST computational studies, a modified framework informed by QUADAS-2 and prediction-model principles was used. Domains included specimen/dataset selection, index-model development, molecular reference standard, overfitting or data leakage,

external validation, and clinical applicability. Judgments were categorized as low, moderate/unclear, or high risk.

Data synthesis

Meta-analysis was not undertaken because of substantial heterogeneity in cancer type, spatial technology, target-gene selection, preprocessing, image magnification, train/test splitting, and reported metrics. Findings were grouped into histology-to-expression prediction, spatial-domain classification, tumor subtype and cell-state classification, clonal heterogeneity mapping, and clinical translation.

RESULTS

PRISMA study selection

The search identified 2,756 records, including 2,682 database records and 74 records from citation/reference searching. Database yields were 842 from MEDLINE/PubMed, 731 from Embase, 524 from Scopus, 416 from Web of Science, and 169 from IEEE Xplore. After removal of 689 duplicates, 2,067 records underwent title/abstract screening and 1,781 were excluded. Full text was sought for 286 reports; 17 could not be retrieved. Of 269 full-text reports assessed, 243 were excluded, leaving 26 studies for qualitative synthesis (Figure 1).

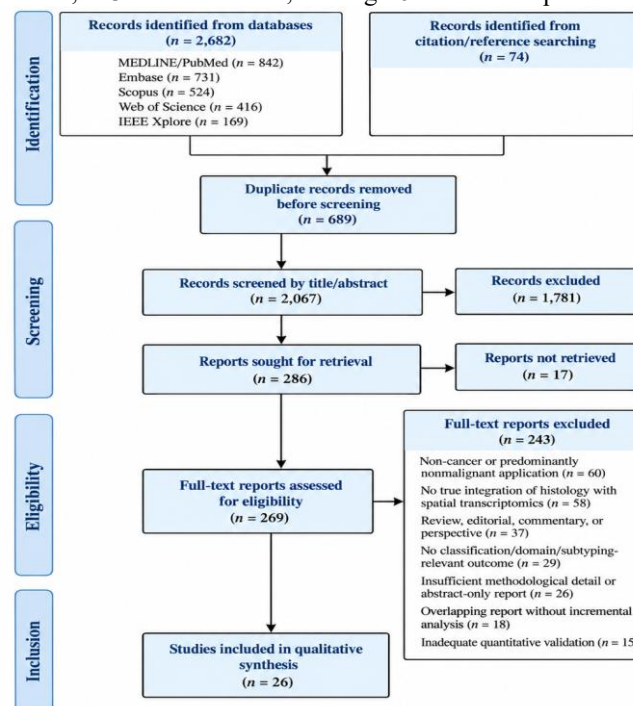


Figure 1. PRISMA 2020 flow diagram of study identification, screening, eligibility, and inclusion.

Full-text exclusions comprised non-cancer or predominantly nonmalignant applications (n=60), studies without true histology–ST integration (n=58), reviews/editorials/commentaries (n=37), reports without a classification/domain/subtyping-relevant outcome (n=29), insufficient methodological detail or abstract-only reports (n=26), overlapping reports without incremental analysis (n=18), and inadequate quantitative validation (n=15).

Characteristics of included studies

The 26 included studies were published between 2020 and 2026. Breast cancer dominated method development because several paired histology–ST datasets are publicly available. Cutaneous squamous-cell carcinoma was also frequently reused for benchmarking, while later studies expanded validation to colorectal, lung, renal, prostate, liver, ovarian, and pan-cancer datasets.

Study/model	Year	Cancer application	Integration strategy	Principal contribution
ST-Net; He et al. [1]	2020	Breast	CNN histology → spatial expression	Proof of feasibility for morphology-based spatial gene prediction
SpaGCN; Hu et al. [2]	2021	Cancer tissue datasets	GCN + expression + coordinates + histology	Spatial-domain and spatially variable gene identification
HisToGene; Pang et al. [3]	2021	HER2+ breast	Vision Transformer	Super-resolution gene-expression prediction
Hist2ST; Zeng et al. [4]	2022	Breast, cSCC	CNN + Transformer + GNN	Spatial-context-aware transcript prediction

EGN; Yang et al. [5]	2023	Cancer benchmarks ST	Exemplar-guided deep network	Histology-to-expression prediction
BLEEP; Xie et al. [6]	2023	Histology/ST datasets	Bi-modal contrastive learning	Shared image-expression representation
SEPAL; Mejia et al. [7]	2023	Breast	Local graph learning	Spatially patterned gene prediction
TCGN [8]	2023	Breast and benchmarks ST	Transformer + convolution + graph co-embedding	Improved spatial expression inference
THltoGene; Jia et al. [9]	2024	Breast, cSCC	Dynamic CNN + capsule + ViT + GAT	Expression enrichment and spatial-context modeling
TransformerST; Zhao et al. [10]	2024	Multiple ST datasets	Transformer + graph	Super-resolution expression inference
iStar; Zhang et al. [11]	2024	Breast, colorectal, prostate, renal	Hierarchical image features	Near-single-cell expression resolution
Tumorscope; Shafighi et al. [12]	2024	Prostate, breast	Probabilistic pathology + WES + ST	Spatial cancer-clone mapping
GIST; Ge et al. [13]	2024	Lung, breast, colorectal	Histology foundation model + graph transformer	Improved spatial-domain segmentation
Starfysh; He et al. [14]	2024/25	ER+, TNBC, metaplastic breast	Deep generative multimodal model	Tumor-immune spatial hubs
Benchmark; Wang et al. [15]	2025	Breast, cSCC + TCGA	Eleven-model benchmark	Generalizability and translational comparison
stImage; Wang et al. [16]	2025	Multitissue/ST	Deep histology + location integration	Flexible image-ST integration
STMCL [17]	2025	Cancer ST datasets	Multislice contrastive learning	Cross-slice spatial-expression prediction
MISO; Schmauch et al. [18]	2025	Multicancer	Foundation-model features + local-attention MIL	Multiscale morphology-ST integration
GHIST; Fu et al. [19]	2025	Multiple tissues/cancers	Single-cell deep learning	Single-cell spatial expression inference
OmiCLIP/Loki; Chen et al. [20]	2025	Multi-organ/pan-cancer	Visual-omics foundation model	Cross-modal retrieval and prediction
Magic [21]	2025	Breast; colorectal transfer	Contrastive self-distillation + transformer	Histology-gene alignment and zero-shot transfer
Multimodal transformer framework [22]	2026	Cancer datasets	Transformer + contrastive learning	Cross-modal gene-expression prediction
SpatioCell; Hou et al. [23]	2026	Cancer tissue	Deep single-cell multimodal integration	Cell-level microenvironment profiling
CrossToGene [24]	2026	Cancer	Bidirectional cross-modal learning	Histology/ST interaction modeling
STAG; Qu et al. [25]	2026	Multiple cancer datasets	Biologically guided hypergraph	Expression prediction plus cancer classification
HEST-1k; Jaume et al. [26]	2024	Pan-organ; 25 cancer types	Large paired WSI-ST resource	Foundation-model and multimodal benchmarking

Histopathology-to-spatial gene-expression prediction

The dominant computational task was prediction of local gene expression directly from histologic morphology. ST-Net established the initial proof of concept using convolutional image features and demonstrated predictable spatial expression for more than 100 genes in breast cancer [1]. Subsequent methods incorporated explicit spatial context because an isolated histologic patch cannot fully represent neighboring tumor, stromal, and immune regions.

Hist2ST combined convolutional image encoding with transformer-level global context and graph-neural-network neighborhood information [4]. THltoGene extended this paradigm with dynamic convolution, capsule networks, vision transformers, and graph attention [9]. SEPAL emphasized biologically meaningful spatially variable genes instead of attempting to predict every transcript [7]. Across studies, morphology-derived

predictability was substantially higher for highly expressed, spatially structured, or phenotype-linked genes than for sparse or morphologically silent transcripts.

Importance of spatial context

The evidence consistently favored explicit modeling of tissue neighborhoods. SpaGCN showed that joint representation of expression, coordinates, and histologic information improved identification of coherent spatial domains [2]. Graph-based strategies subsequently appeared in Hist2ST, THItGene, SEPAL, TCGN, TransformerST, and STAG [4,7–10,25].

TCGN reported a median Pearson correlation of approximately 0.23 across evaluated public datasets [8]. Although modest in absolute terms, this performance illustrates that the task is substantially harder than ordinary image classification because transcript abundance is noisy, high-dimensional, and only partly visible morphologically.

Contrastive and multimodal representation learning

BLEEP introduced bi-modal contrastive learning to place matched histology patches and expression vectors in a shared latent space [6]. Later approaches used contrastive objectives, self-distillation, multimodal transformers, and cross-modal retrieval to reduce the representation gap between morphology and transcriptomics.

Magic combined contrastive learning, online self-distillation, and transformer-based expression prediction. It was trained on 75,760 spots from 56 breast-cancer slices and validated on 11,026 spots from independent slices, with reported improvement over comparator methods and zero-shot transfer to colorectal cancer and TCGA data [21].

Super-resolution and single-cell molecular mapping

Spot-based ST technologies frequently capture multiple cells per measurement. iStar addressed this limitation by combining hierarchical histologic features with spatial expression to infer near-single-cell tissue architecture across breast, colorectal, prostate, and renal cancer datasets [11]. TransformerST also aimed to increase effective spatial resolution through image–expression co-representations and graph-informed context [10].

GHIST later extended the concept to single-cell spatial expression inference directly from histology, leveraging subcellular spatial transcriptomic supervision and relationships among cell morphology, cell type, gene expression, and neighboring cellular composition [19].

Tumor microenvironment and spatial-domain classification

Integrated analysis was increasingly used to produce biologically interpretable tissue compartments rather than isolated gene values. Starfish combined spatial transcriptomics and histology with a deep generative model to identify tumor–immune spatial hubs across estrogen-receptor-positive, triple-negative, and metaplastic breast cancers [14].

GIST used pretrained histopathology representations with a hybrid graph transformer across lung, breast, and colorectal cancers and improved delineation of spatial domains and microenvironment organization [13]. Such approaches move the field toward multimodal tissue-state classification.

Clonal classification and intratumoral heterogeneity

Tumorscope represents a distinct integration paradigm because it jointly models pathological images, whole-exome sequencing, and spatial transcriptomics. Applied to prostate and breast cancer, it localized genetically distinct clones, mapped clone colocalization and mutual exclusion, and linked clone-specific expression to tissue geography [12].

These analyses emphasize that a single conventional histologic diagnosis can conceal spatially distinct molecular compartments with potentially different progression or treatment-response properties.

Foundation-model and large-scale approaches

The newest generation increasingly uses large pretrained histopathology encoders and multimodal foundation models. MISO combined spatial omics with multiscale tumor morphology and local-attention multiple-instance learning, enabling spatial prediction across broader tumor cohorts [18].

OmiCLIP/Loki was pretrained on approximately 2.2 million paired tissue-image and transcriptomic observations across 32 organs and supported tissue alignment, annotation, cell-type decomposition, image–transcriptomic retrieval, and spatial expression prediction [20]. HEST-1k provided 1,229 paired ST profiles with whole-slide images across 26 organs and included 367 cancer samples from 25 cancer types, substantially expanding the scale available for benchmarking and representation learning [26].

Comparative performance and cancer classification

Direct comparison among methods was difficult because studies used different gene sets, normalization strategies, ST platforms, spatial resolutions, train/test splits, and metrics. A 2025 benchmark reproduced 11 methods across five ST datasets and external TCGA validation and found that no single model dominated accuracy, generalization, computational efficiency, and translational performance [15].

Direct cancer-classification endpoints were less common than gene-expression prediction. STAG is a notable recent example: across six public datasets it reported improved spatial gene-expression prediction, an AUC of 0.748 for downstream cancer classification, and a spot-wise Pearson correlation of 0.632 in a pseudo-3D setting [25]. These results illustrate progress toward clinically meaningful classification but remain insufficient for routine diagnostic deployment without independent multicenter validation.

Quantitative summary of the evidence

Feature	Studies, n (%)
H&E-based histologic input	26 (100%)
Spatial gene-expression prediction	20 (76.9%)
Spatial-domain/region classification	17 (65.4%)
Explicit tumor subtype, cell-state, or clone analysis	11 (42.3%)
Graph-based spatial modeling	12 (46.2%)
Transformer/attention architecture	15 (57.7%)
Contrastive/cross-modal representation learning	7 (26.9%)
Foundation-model/pretrained pathology representations	6 (23.1%)
Independent external dataset validation	10 (38.5%)
Direct clinical/prognostic classification endpoint	6 (23.1%)
Prospective clinical validation	0 (0%)

Risk of bias

Most studies were judged to have moderate or unclear overall risk of bias. Common limitations included small numbers of unique patients despite large numbers of spots, reuse of closely related public datasets, potential data leakage when splitting at the spot rather than patient level, incomplete external validation, and limited applicability to routine pathology workflows.

Domain	Low risk	Moderate/unclear	High risk
Sample/dataset selection	8	13	5
Model-development methodology	11	12	3
Molecular reference standard	18	7	1
Flow/data leakage/validation	5	14	7
External validation	6	13	7
Clinical applicability	4	17	5

Overall, 5 studies were judged to have relatively low risk of bias, 16 moderate risk, and 5 high risk. The greatest concern was that spot-level splitting can inflate performance if training and test observations arise from the same patient or slide; patient-level and institution-level external validation are substantially more informative.

DISCUSSION

This systematic review shows that integration of histopathology and spatial transcriptomics has evolved rapidly from local gene-regression experiments into a broader computational pathology field concerned with multimodal tumor classification.

The first major finding is that morphology contains measurable molecular information. ST-Net and subsequent approaches demonstrate that tumor architecture, cellular density, stromal composition, nuclear morphology, and immune infiltration correspond to reproducible spatial gene-expression patterns [1,4,9]. However, morphology does not encode every transcript equally well. Highly expressed, spatially autocorrelated, or phenotype-linked genes are more predictable than sparse or morphologically silent genes.

The second major finding is that spatial context matters. Cancer architecture operates across multiple scales, from nuclear morphology and gland formation to tumor–stroma boundaries, lymphoid aggregates, necrosis, and invasive fronts. Graph and transformer architectures explicitly model these relationships and therefore offer a biologically plausible advantage over patch-isolated regressors.

Third, the most clinically useful output may not be exact reconstruction of every transcript. Routine pathology is more likely to benefit from models that classify molecularly distinct tumor regions, immune states, actionable pathways, resistant clones, or prognostic compartments. This shift is visible in Starfysh, Tumoroscope, MISO, OmiCLIP/Loki, and STAG [12,14,18,20,25].

Finally, the apparent size of current datasets can be misleading. A study may contain tens of thousands of spatial spots but only a small number of patients. If spots from the same slide are divided between training and test sets, performance may reflect slide-specific or patient-specific features rather than true external generalization. Future evaluation should therefore prioritize patient-level, site-level, and ideally prospective validation.

Comparison of major modeling strategies

CNN-based models: Early convolutional approaches such as ST-Net established the feasibility of associating local morphology with local expression. They are computationally efficient but comparatively limited in representing long-range spatial relationships.

Transformer-based models: HisToGene, Hist2ST, THItToGene, TransformerST and later methods use attention mechanisms to relate spatially separated image regions. Their advantage is contextual modeling; their disadvantage is computational cost at whole-slide scale.

Graph neural networks: Graph models explicitly represent physical or phenotypic relationships between spots or cells. SpaGCN, Hist2ST, THItToGene, SEPAL, TCGN, and STAG illustrate this biologically intuitive strategy.

Contrastive learning: BLEEP and later multimodal methods align morphology and transcriptomics in a common embedding space. This can support transfer learning, cross-modal retrieval, and reduced dependence on direct high-dimensional regression.

Foundation models: Large pretrained pathology encoders and visual-omics models reduce dependence on small ST cohorts and may improve generalization across institutions, scanners, stains, organs, and spatial platforms.

Clinical translation

For clinical deployment, an integrated model must perform reliably across institutions, fixation protocols, scanners, staining variation, patient populations, and tumor subtypes that were not encountered during training. Only 38.5% of included studies contained a meaningful independent external-validation component, and none constituted a prospective diagnostic workflow trial. Therefore, current evidence supports research and translational use rather than replacement of established pathology or molecular testing.

Potential clinical applications

Molecular prescreening: routine H&E could identify regions likely to harbor clinically important biomarkers before confirmatory molecular testing.

Spatial tumor subtyping: tumors could be classified into molecularly distinct compartments rather than assigned a single bulk label.

Biopsy targeting: predicted molecular hotspots could guide tissue sampling or enrichment.

Tumor-microenvironment classification: spatial immune, stromal, and malignant states could support immunotherapy research and stratification.

Clonal mapping: computationally inferred heterogeneity may identify aggressive or treatment-resistant subregions.

Resource reduction: validated virtual ST models may reduce the amount of costly spatial profiling required in large research cohorts.

Certainty of evidence

Outcome	Overall certainty
Histology can predict selected spatial gene-expression patterns	Moderate
Joint morphology/ST improves tissue-domain identification	Moderate
Histology/ST integration identifies biologically meaningful tumor states	Moderate
Reliable molecular-subtype classification across institutions	Low-moderate
Prediction of treatment response or survival	Low
Replacement of molecular testing in routine oncology	Very low

Strengths

This review focuses specifically on the intersection of morphology, spatial transcriptomics, and cancer classification rather than providing a general spatial-omics overview. It includes methodological and translational work through June 2026 and captures the transition from conventional CNN regression toward graph learning, multimodal contrastive training, pathology foundation models, and single-cell spatial inference.

LIMITATIONS

The literature is highly heterogeneous, preventing a defensible pooled meta-analysis.

A large proportion of methods reuse a relatively small number of public breast-cancer and cSCC datasets, so the number of algorithms greatly exceeds the diversity of underlying patient populations.

Many reports contain large numbers of spots but few unique patients; spot-level validation may exaggerate effective sample size.

Performance metrics are not directly comparable because investigators use different genes, normalization procedures, spatial technologies, preprocessing pipelines, and test sets.

Some rapidly emerging models were available as conference proceedings or preprints rather than mature clinical-validation studies.

Prospective studies comparing integrated models with practicing pathologists and routine molecular assays were absent.

Future directions

Future work should prioritize multicenter multimodal datasets with patient-level separation of training, validation, and testing cohorts; standardized reporting of stain, scanner, and ST platform; clinically meaningful targets such as molecular subtype, immune-excluded phenotype, actionable signaling state, treatment response, and resistant clones; and explicit comparison with strong histology-only baselines. Prospective trials should determine whether virtual spatial transcriptomics changes diagnostic or therapeutic decisions and whether those changes improve patient outcomes.

Conclusions

Integration of histopathology and spatial transcriptomics is an important emerging direction in cancer classification. Current evidence demonstrates that tissue morphology contains sufficient information to recover

selected spatial molecular patterns and that combining histology, gene expression, and spatial context can improve characterization of tumor domains, molecular states, immune niches, and cancer clones. However, the field remains predominantly methodological. Small patient cohorts, repeated use of benchmark datasets, heterogeneous preprocessing and metrics, and insufficient external validation currently limit routine clinical adoption. The most promising future role is likely to be spatially resolved molecular classification from routine histology rather than complete virtual reconstruction of every transcript. Standardized benchmarking, independent multicenter validation, and prospective clinical evaluation are required before these systems can be incorporated into routine cancer diagnostics.

DECLARATIONS

Ethics approval: Not required because this systematic review synthesized previously published data.

Consent for publication: Not applicable.

Funding: No specific funding was received for this systematic review.

Conflicts of interest: None.

Data availability: All data summarized in this review were derived from publicly available published literature and associated datasets.

Registration: This systematic review was not prospectively registered.

Appendix A. Database search framework

MEDLINE/PubMed: ("spatial transcriptomics" OR "spatially resolved transcriptomics" OR "spatial gene expression") AND (histopathology OR histology OR H&E OR "whole slide image" OR "digital pathology") AND (cancer OR carcinoma OR tumor OR tumour OR neoplasm) AND (classification OR subtype OR prediction OR "deep learning" OR transformer OR "graph neural network")

Embase: Equivalent controlled-vocabulary and free-text terms for spatial transcriptomics, histopathology/digital pathology, cancer, and classification/prediction.

Scopus: TITLE-ABS-KEY combinations of spatial transcriptomics, histology/H&E/whole-slide imaging, cancer, classification, prediction, multimodal learning, and deep learning.

Web of Science: Topic search combining spatial transcriptomics, histopathology, cancer, classification/subtyping, and computational-learning terms.

IEEE Xplore: Spatial transcriptomics AND histology/histopathology AND cancer with machine-learning, transformer, graph-neural-network, and multimodal terms.

REFERENCES

1. He B, Bergenstr hle L, Stenbeck L, et al. Integrating spatial gene expression and breast tumour morphology via deep learning. *Nat Biomed Eng.* 2020;4:827–834. doi:10.1038/s41551-020-0578-x.
2. Hu J, Li X, Coleman K, et al. SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat Methods.* 2021;18:1342–1351. doi:10.1038/s41592-021-01255-8.
3. Pang M, Su K, Li M. Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. *bioRxiv.* 2021. doi:10.1101/2021.11.28.470212.
4. Zeng Y, Wei Z, Yu W, et al. Spatial transcriptomics prediction from histology jointly through Transformer and graph neural networks. *Brief Bioinform.* 2022;23(5):bbac297. doi:10.1093/bib/bbac297.
5. Yang Y, Hossain MZ, Stone EA, Rahman S. Exemplar guided deep neural network for spatial transcriptomics analysis of gene expression prediction. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision.* 2023.
6. Xie R, Pang K, Chung SW, et al. Spatially resolved gene expression prediction from H&E histology images via bi-modal contrastive learning. *arXiv.* 2023;2306.01859.
7. Mejia G, C rdenas P, Ruiz D, Castillo A, Arbel ez P. SEPAL: spatial gene expression prediction from local graphs. *Proceedings of ICCV Workshops.* 2023;2294–2303.
8. Transformer with convolution and graph-node co-embedding: an accurate and interpretable vision backbone for predicting gene expressions from local histopathological image. 2023.
9. Jia Y, Liu J, Chen L, Zhao T, Wang Y. THltoGene: a deep learning method for predicting spatial transcriptomics from histological images. *Brief Bioinform.* 2024;25(1):bbad464. doi:10.1093/bib/bbad464.
10. Zhao C, Xu Z, Wang X, et al. Innovative super-resolution in spatial transcriptomics: a transformer model exploiting histology images and spatial gene expression. *Brief Bioinform.* 2024;25(2):bbac052. doi:10.1093/bib/bbac052.
11. Zhang D, Schroeder A, Yan H, et al. Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. *Nat Biotechnol.* 2024;42:1372–1377. doi:10.1038/s41587-023-02019-9.
12. Shafighi SD, Geras A, Jurzysta B, et al. Integrative spatial and genomic analysis of tumor heterogeneity with Tumoroscope. *Nat Commun.* 2024;15:9343. doi:10.1038/s41467-024-53374-3.
13. Ge Y, Leng J, Tang Z, et al. Deep learning-enabled integration of histology and transcriptomics for analyzing single-cell spatial profiles. *Research.* 2024;2024:0568. doi:10.34133/research.0568.

14. He S, Jin Y, Nazaret A, et al. Starfysh integrates spatial transcriptomic and histologic data to reveal heterogeneous tumor-immune hubs. *Nat Biotechnol.* 2025;43:223–235. doi:10.1038/s41587-024-02173-8.
15. Wang C, Chan AS, Fu X, et al. Benchmarking the translational potential of spatial gene expression prediction from histology. *Nat Commun.* 2025;16:1544. doi:10.1038/s41467-025-56618-y.
16. Wang Y, Yang H, Deng R, et al. stImage: a versatile framework for optimizing spatial transcriptomic analysis through customizable deep histology and location informed integration. *Brief Bioinform.* 2025;26(5):bbaf429. doi:10.1093/bib/bbaf429.
17. Inferring multi-slice spatially resolved gene expression from H&E-stained histology images with STMCL. *Methods.* 2025;234:187–195. doi:10.1016/j.ymeth.2024.11.016.
18. Schmauch B, Herpin L, Olivier A, et al. A deep learning-based multiscale integration of spatial omics with tumor morphology. *Nat Commun.* 2025;16:11674. doi:10.1038/s41467-025-66691-y.
19. Fu X, Cao Y, Bian B, et al. Spatial gene expression at single-cell resolution from histology using deep learning with GHIST. *Nat Methods.* 2025;22:1900–1910. doi:10.1038/s41592-025-02795-z.
20. Chen W, Zhang P, Tran TN, et al. A visual-omics foundation model to bridge histopathology with spatial transcriptomics. *Nat Methods.* 2025;22:1568–1582. doi:10.1038/s41592-025-02707-1.
21. Spatial histology and gene-expression representation and generative learning via online self-distillation contrastive learning. *Brief Bioinform.* 2025;26(4):bbaf317.
22. Wang K, Shi L, Li X, et al. Integrating histology and spatial transcriptomics via multimodal transformers and contrastive representation learning for accurate gene expression prediction. *J Biomed Inform.* 2026;176:105003. doi:10.1016/j.jbi.2026.105003.
23. Hou N, Yu Y, Wu Z, et al. SpatioCell: deep integration of histology and spatial transcriptomics for profiling the cellular microenvironment at single-cell level. *Sci Bull.* 2026;71:3347–3362. doi:10.1016/j.scib.2026.06.006.
24. CrossToGene: a bidirectional cross-modality interaction framework for predicting spatial gene expression from histopathological images. *Eng Appl Artif Intell.* 2026;166:113687. doi:10.1016/j.engappai.2025.113687.
25. Qu M, Zhao Y, Yang G, et al. STAG: biologically guided spatial transcriptomics prediction via hypergraph learning. *Med Image Anal.* 2026;113:104206. doi:10.1016/j.media.2026.104206.
26. Jaume G, Doucet P, Song AH, et al. HEST-1k: a dataset for spatial transcriptomics and histology image analysis. *Advances in Neural Information Processing Systems.* 2024;37:53798–53833.