Proceedings of ICBioMed 2026 Symposium: Biomedical Engineering and Public Health: Real-Time Wearables for Disease Detection
DOI: 10.54254/2753-8818/2026.ES36957

Classification of Tumor Spatial Transcriptomics Data and Methods

Guanxiang Qiu

School of Integrated Traditional Chinese and Western Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, China
2253788898@qq.com

Abstract. Spatial transcriptomics technologies combine gene expression profiles with the spatial positions of tissues, providing an unprecedented perspective for dissecting the heterogeneity of the tumor microenvironment. However, tumor tissues are highly heterogeneous, and their discontinuous, sparse, and blurred-boundary features pose unique challenges to spatial domain identification. This paper systematically sorts out the distribution characteristics of current tumor spatial transcriptomics data in terms of acquisition platforms, cancer types, and data quality, and points out that data imbalance imposes substantial constraints on algorithm selection. Furthermore, according to the richness of input information, 21 mainstream computational methods are divided into three levels and six subclasses. The inherent assumptions and applicable boundaries of each type of method are explained, and their known applications in tumor research are summarized. On this basis, this paper constructs a decision framework based on data characteristics, and verifies the effectiveness of the framework through empirical analysis of seven representative methods on 10x Visium breast cancer data. The results show that the effectiveness of a method depends on the matching degree between its inherent assumptions and data characteristics, rather than the advancement of the method itself. This paper provides a systematic reference for tumor researchers to select appropriate methods according to their own data conditions.

Keywords: spatial transcriptomics, tumor microenvironment, spatial clustering, method classification, data characteristics

1. Introduction

Spatial transcriptomics technologies integrate gene expression profiles with tissue spatial localization, offering an unprecedented lens into the heterogeneity of the tumor microenvironment, intercellular communication, and spatial architecture [1, 2]. Tumor initiation, progression, metastasis, and drug resistance depend not only on the molecular characteristics of cells, but also on their spatial arrangement in tissues and the adjacent microenvironment. Tumor tissues exhibit high heterogeneity: diverse cell types, irregular structures, blurred boundaries, and frequent discontinuous small tumor foci or infiltrative regions. Cell types are mixed at the invasive tumor boundary, and conventional spatial smoothing assumptions easily lead to over-smoothed boundaries. Discontinuous

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.ES36957

small tumor foci, with few and scattered spots, render adjacency graph-based methods ineffective. For highly fibrotic tumors such as pancreatic cancer, degraded RNA quality and poor tissue adhesion further exacerbate algorithmic instability due to data sparsity. Tumor spatial domain identification is not merely an algorithmic problem, but a systemic challenge strongly constrained by data resources.

Existing reviews have systematically summarized the principles, application scenarios, and selected computational methods of spatial transcriptomics [3, 4], but most focus on the introduction of technical platforms or algorithm classification, lacking a systematic review from the perspective of the coupling between data constraints and methodological evolution. The data availability of different tumor types, the resolution and sensitivity of acquisition platforms, and tissue morphological features directly determine which algorithms can be effectively applied; conversely, algorithm design continuously advances the analytical capacity for scarce or challenging samples. This paper first analyzes the distribution of current tumor spatial transcriptomics data resources and their constraints on algorithm selection. It then sorts out the evolutionary trajectory of computational methods according to the richness of input information, explains the inherent assumptions and applicable boundaries of each type of method, and summarizes their known applications and findings in tumor research. On this basis, this paper proposes a decision framework based on data characteristics, verifies its effectiveness through empirical analysis, and finally discusses the joint limitations of data and methods and future directions.

2. Constraints of data resources on algorithm selection

2.1. Performance trade-offs of acquisition platforms

Spatial transcriptomics technologies are mainly divided into sequencing-based and imaging-based categories. Sequencing-based platforms (e.g., 10x Visium [5], Slide-seq [6], Stereo-seq [7]) capture RNA via spatial barcodes followed by sequencing, achieving whole-transcriptome coverage, but exhibit significant dropout effects. For 10x Visium, the transcript capture rate per spot is typically below 10%, with limited resolution (spot diameter usually greater than 10 μm). Imaging-based platforms (e.g., MERFISH [8], STARmap [9], CosMx [10], Xenium [11]) achieve single-molecule detection via fluorescence in situ hybridization and cyclic imaging, with high sensitivity and subcellular resolution. However, their gene panels are pre-designed (typically 2,000 to 6,000 genes), and they have zero sensitivity to genes outside the panel. The two types of technologies are complementary rather than substitutive. Detailed technical performance benchmarks can be found in the systematic evaluations by Ren et al. [12] and Yue et al. [13]. This paper focuses on four platforms with the highest degree of commercialization, highest throughput, and publicly available tumor datasets: the sequencing-based Stereo-seq v1.3 and Visium HD FFPE, and the imaging-based CosMx 6K and Xenium 5K.

Stereo-seq v1.3 [7] is based on DNA nanoball arrays, with 0.5 μm resolution and whole-transcriptome poly(A) capture capability, and can detect non-human RNA. However, it suffers from transcript diffusion issues, and its cell segmentation relies on nuclear swelling, resulting in low accuracy. In the benchmark test by Ren et al. [12], the median number of detected genes per 8 μm bin was approximately 700 to 800, lower than the 1,200 of Visium HD FFPE, with a higher proportion of invalid spatial barcodes. Visium HD FFPE [11] employs a probe hybridization capture system with 2 μm resolution, targeting approximately 18,085 human genes, and is specifically designed for FFPE samples. It has better diffusion control, and the normalized ratio of extra-tissue transcript signals is greatly reduced compared to Stereo-seq (approximately 1/3 to 1/2) [12]. However, it lacks an official cell segmentation algorithm and relies on bin-level analysis or external

DOI: 10.54254/2753-8818/2026.ES36957

tools. CosMx 6K [10] is based on multiplexed fluorescence in situ hybridization, covering 6,175 pre-designed genes (2,522 of which overlap with Xenium 5K), and supports three-channel whole-cell segmentation of nucleus, membrane, and cytoplasm. It has a high total transcript detection count (approximately 3,000 to 6,000 per cell), but the signal intensity of negative probes is significantly higher (median approximately 0.37 per cell), with high spatial autocorrelation and low inter-gene variability, limiting the detection of low-abundance targets [12]. Xenium 5K [11] also adopts multiplexed fluorescence in situ hybridization, detecting 5,001 pre-designed genes, with matched cell segmentation staining. The median number of transcripts per cell for shared genes is approximately 1,500 to 2,500, significantly higher than that of CosMx 6K. The negative probe signal is extremely low (median approximately 0.03 per cell), and the proportion of normalized negative control signals is only 1/10 to 1/5 that of CosMx. It has the highest gene-level correlation coefficient with scRNA-seq ($R \approx 0.88$), and ranks first in annotation consistency, immune cell discrimination ability, and consistency with CODEX protein data [12, 14, 15].

These platform characteristics directly determine the types of information that subsequent algorithms can utilize. Sequencing-based data provides whole-transcriptome breadth but lacks single-cell resolution, making it suitable for bin-based clustering or analyses requiring a wide dynamic range. Imaging-based data offers single-cell resolution but limited gene coverage, making it suitable for cell type annotation and fine spatial localization. Choosing between sequencing-based and imaging-based platforms is essentially a choice between "gene breadth" and "cellular precision", and this trade-off further influences algorithm selection.

2.2. Data distribution of major cancer types

According to published benchmark studies [12, 14, 15], the current publicly available tumor spatial transcriptomics data are highly unevenly distributed. Twelve slices of the human dorsolateral prefrontal cortex (DLPFC) from healthy individuals were generated using 10x Visium technology, with 3,400 to 4,800 spots per slice and complete annotations of 6 cortical layers and white matter [16]. Although not tumor samples, they have become a standard benchmark for algorithm development due to their regular structure. Breast cancer has the most abundant tumor data, including the official 10x Genomics dataset [17], Luminal B, HER2+, and triple-negative breast cancer samples provided by Wu et al. [18] (10 slices in total, 1,000 to 5,000 spots per slice), and breast cancer cohort data generated by Ren et al. [12] on the CosMx and Xenium platforms, totaling more than 20 public slices. Colorectal cancer and its liver metastasis samples have a moderate data volume. Wu et al. [19] provided primary tumors and paired liver metastasis samples on the 10x Visium platform (4 slices in total, approximately 2,000 to 4,000 spots per slice), but there are only 2 paired primary-metastasis datasets with 5 slices in total. Hepatocellular carcinoma and ovarian cancer have moderate data volumes. Ren et al. [12] systematically compared hepatocellular carcinoma (1 sample, approximately 600 to 2,000 spots) and ovarian cancer (1 sample, approximately 3,500 to 5,500 spots) using 10x Visium, CosMx 6K, and Xenium 5K, with fewer than 10 and 5 public slices respectively. Data for pancreatic ductal adenocarcinoma, HER2-positive small tumors, and lung cancer are markedly scarce. Pancreatic cancer samples have only 200 to 400 spots per slice [20]; HER2-positive small breast tumors (8 patients, 177 to 692 spots per sample) [21] have discontinuous structures; for lung cancer, there is currently only one CosMx dataset of a non-small cell lung cancer FFPE sample officially provided by NanoString (approximately 100,000 cells, no publicly available paired multi-slice data).

This data imbalance imposes substantial constraints on algorithm selection. For breast cancer with sufficient data, various algorithms can be fully trained and validated. For scarce data such as

DOI: 10.54254/2753-8818/2026.ES36957

pancreatic cancer, deep learning methods are prone to overfitting, while statistical models may incorrectly merge small structures due to dominant priors. Method selection cannot rely solely on benchmark rankings; the data characteristics of the target cancer type must also be considered.

2.3. Specific challenges of difficult samples for algorithms

The aforementioned "difficult" samples (small spot count, discontinuous domains, low cell density) pose specific challenges to computational methods. For graph neural network methods, the spatial adjacency graph degenerates into a sparse or even disconnected graph, message propagation fails, and embeddings are difficult to converge. For statistical models, spatial smoothing priors dominate the likelihood, causing small tumor domains to be incorrectly merged into the surrounding stroma, resulting in severe over-segmentation or under-segmentation. For deep learning methods, small samples easily lead to overfitting, and batch effects are difficult to correct. These challenges are not defects of the algorithms themselves, but conflicts between the inherent assumptions of the algorithms and the true distribution of tumor data. The reasons for the scarcity of data such as pancreatic cancer and lung cancer include low surgical resection rates, degraded RNA quality caused by fibrotic stroma, FFPE archival limitations on retrospective studies, and biopsy samples too small to meet chip size requirements. To address this, Visium HD FFPE with an enhanced permeabilization protocol should be prioritized for pancreatic cancer; FFPE-compatible panels of Xenium 5K or CosMx 6K should be used for HER2-positive small tumors; for non-small cell lung cancer, prospective cohorts should be established using Xenium 5K lung cancer custom panels, while matched FFPE and frozen tissues are collected for cross-validation. Public data repositories (e.g., SPATCH [12], SpatialDB [22]) should encourage the upload of raw images and segmentation results to facilitate the testing and optimization of computational methods on difficult samples.

3. Classification, evolution, and applicable boundaries of computational methods

To sort out existing methods, this paper divides them into three levels according to the input information modalities: gene expression only, gene expression plus spatial coordinates, and gene expression plus spatial coordinates plus histology images. This division reflects the progressive evolution of data availability and algorithmic capability: early methods only used expression data as a baseline, and with the addition of spatial coordinates and image information, algorithms have gradually become able to identify spatially continuous domains, characterize tumor boundaries, and integrate morphological features. Methods within each level are applicable to different tumor data scenarios due to differences in their inherent assumptions.

3.1. Level 1: gene expression only

Modularity optimization clustering (Louvain [23], Leiden [24]) takes only the gene expression matrix as input and groups cells with similar expression into the same community by optimizing modularity. This type of method does not utilize spatial information at all and cannot distinguish different functional regions within tumors; it is usually used as a baseline to demonstrate the value of spatial information. In tumor data, this type of method is only suitable for preliminary exploration of expression heterogeneity and should not be used for spatial domain identification.

DOI: 10.54254/2753-8818/2026.ES36957

3.2. Level 2: gene expression and spatial coordinates

3.2.1. Statistical modeling methods

Statistical modeling methods (BayesSpace [25], BASS [26], SpatialPCA [27]) explicitly model spatial adjacency relationships through Bayesian frameworks, hidden Markov random fields, or spatial principal component analysis. Their inherent assumption is that spatially adjacent cells or spots tend to belong to the same biological domain. This assumption is derived from observations of regular tissue structures, but often fails at invasive tumor boundaries: when tumor cells infiltrate the surrounding stroma, adjacent points may belong to different functional regions, causing spatial smoothing priors to incorrectly over-smooth the boundaries. The hierarchical prior of BASS is prone to overfitting on small-sample tumor data, incorrectly merging rare cell types. As an unsupervised dimensionality reduction method, the biological interpretability of the principal components of SpatialPCA decreases significantly when tumor heterogeneity is high. In tumor applications, BayesSpace has been used for spatial domain identification in breast cancer [25] but performs poorly in boundary regions; BASS achieved a high ARI (0.71) on MERFISH brain metastasis data [14], but its IoU on discontinuous small tumor domains is generally below 0.5.

3.2.2. Graph neural network methods

Graph neural network methods (SpaGCN [28], CCST [29], STAGATE [30], SpaceFlow [31], GraphST [32], SCAN-IT [33]) construct spatial adjacency relationships as graphs and propagate gene expression information on the graph using graph convolutional networks, graph attention networks, or graph mutual information maximization. CCST [29] adopts the Deep Graph Infomax [34] framework to maximize the mutual information between node representations and the global graph representation. Its inherent assumption is that the spatial adjacency graph can reliably reflect the connectivity of biologically relevant cells. When tumor regions are small, cells are sparse, or fibrosis is severe, the adjacency graph degenerates into a sparse or even disconnected graph, and message propagation fails. In cell-dense tumor regions (e.g., highly invasive breast cancer regions), graph neural networks can capture local structures well. In tumor data, GraphST achieved high clustering accuracy on 10x Visium breast cancer data [14]; SpaceFlow has the best spatial continuity index on MERFISH imaging-based data [32]; STAGATE has the best scalability on large Stereo-seq slices [31]. However, the IoU of all methods on discontinuous small tumor domains is generally below 0.3 [14]. The evolutionary trend of this type of method is from simple graph convolution to attention mechanisms and contrastive learning, gradually improving the ability to capture spatial heterogeneity. However, this trend is based on the assumption of regular data, and has limited improvement for sparse, discontinuous tumor samples.

3.2.3. Autoencoder methods

Autoencoder methods (SEDR [35], ADEPT [36], DeepST [37]) use deep autoencoders or variational autoencoders for dimensionality reduction and denoising of gene expression data, and then integrate spatial information through graph autoencoders. Their inherent assumption is that gene expression data can be compressed into a low-dimensional latent space, and spatial information can serve as an auxiliary reconstruction constraint. This type of method performs stably on small to medium-sized tumor datasets with high data quality, but variational inference has high computational cost on large-scale data; the imputation step of ADEPT may introduce spurious gene co-expression relationships.

DOI: 10.54254/2753-8818/2026.ES36957

DeepST adds a denoising module and a domain adversarial network on the basis of autoencoders, and can further integrate image information. In the evaluation by Yuan et al. [14], DeepST achieved an accuracy of 0.46 on tumor datasets, outperforming most pure autoencoder methods, but its memory consumption grows rapidly with the number of spots.

3.2.4. Spatial feature enhancement methods

Spatial feature enhancement methods (BANKSY [38]) use azimuthal Gabor filter kernels to capture the spatial variation trend of gene expression, calculate the neighborhood average expression and gradient features for each cell, and then directly concatenate the enhanced features into the original expression matrix. Their inherent assumption is that the expression gradient and mean difference between a cell and its neighborhood are sufficient to encode spatial microenvironment information without complex graph message passing. This method avoids complex graph neural networks and has high computational efficiency. In the benchmark test by Hu et al. [15], BANKSY took approximately 12 minutes to process an embryonic slice with about 30,000 spots, while STAGATE took about 6 minutes on the same task, but BANKSY's memory usage was only 60% of that of STAGATE. This method is mainly applicable to tissue structures with obvious spatial gradient changes; in regions with blurred tumor boundaries or uneven cell density, the directional sensitivity of Gabor filters may lead to feature enhancement distortion. On sparse pancreatic cancer data, BANKSY's performance degrades due to the lack of reliable neighborhood information.

3.3. Level 3: gene expression, spatial coordinates, and histology images

3.3.1. Contrastive learning methods

Contrastive learning methods (conST [39], ConGI [40]) align representations of different modalities (expression, spatial, image) in the latent space by designing multi-level contrastive losses. Their inherent assumption is that morphological features in histology images are consistent with gene expression, and complementary information can be extracted by aligning modalities. This type of method is highly dependent on high-quality H&E or immunofluorescence staining images. In necrotic regions, hemorrhagic regions, or areas with dense lymphocyte infiltration, image noise can cause alignment shifts and significantly degrade model performance. ConGI achieved an average ARI of 0.53 on the breast cancer dataset [15], but dropped to the middle range on non-tumor brain tissue data, indicating its high sensitivity to image quality. In tumor applications, ConGI successfully identified lesions and microvascular morphology in breast cancer H&E images, assisting in distinguishing invasive ductal carcinoma and ductal carcinoma in situ regions [40].

3.3.2. Multi-view fusion methods

Multi-view fusion methods (stMVC [41], stKeep [42], MCGAE [43]) actively construct graphs from multiple views (histology similarity graph, spatial location graph, expression similarity graph, etc.), fuse multi-view information through attention mechanisms or heterogeneous graph learning, and can incorporate manually annotated tumor regions or prior knowledge (e.g., PPI, GRN) as supervision signals. Their inherent assumption is that the tumor microenvironment can be described from multiple independent views, and fusing these views can yield more robust cell state representations. The cost of this type of method is the need for manually annotated tumor regions or high-quality prior knowledge, which may become a bottleneck in rare tumors lacking annotations or priors. In

DOI: 10.54254/2753-8818/2026.ES36957

tumor applications, stMVC detected more fine spatial domains in ovarian cancer than other methods (4 domains vs. 1 to 2), identified cell states and transition cells related to stemness, migration, and metastasis, and discovered gene modules associated with prognosis [41]. On this basis, stKeep [42] constructed a heterogeneous graph containing three types of nodes: cells, genes, and tissue regions. It learns cell modules, gene modules, and cell communication modules through three modules respectively, and uses contrastive learning to ensure comparable communication patterns between different cell states. In triple-negative breast cancer, stKeep discovered bipotent basal cells, tumor myoepithelial cells, and metastasis-related cell populations, identified invasion and metastasis communication pathways related to COMP, SPP1, etc., and key genes were significantly associated with patient survival [42]. MCGAE [43] constructs multiple augmented views for the gene expression matrix and spatial adjacency matrix respectively, learns view-specific representations through contrastive graph autoencoders, and then fuses them with attention mechanisms. It can also be extended to 3D multi-slice analysis. In colorectal cancer liver metastasis, MCGAE successfully identified tumor invasive regions and subdivided them into three functional subgroups; after combining images, the accuracy of tumor region identification improved by 20.75%, and it accurately distinguished IDC and DCIS regions in breast cancer [43].

3.3.3. Morphological normalization methods

Morphological normalization methods (stLearn [44]) extract morphological features of each spot using H&E images, calculate morphological distances, and perform spatial morphological normalization of gene expression based on these distances. Their inherent assumption is that histomorphological distance is positively correlated with gene expression distance. This method must rely on histology images and has certain advantages in the fine division of tumor boundaries and heterogeneous regions. However, necrotic regions, hemorrhagic regions, or regions with dense lymphocyte infiltration in H&E images may generate morphological noise unrelated to tumor parenchyma, causing the normalized expression signal to deviate from the true biological state. In addition, the normalization process relies on high-quality staining consistency of tissue sections, and staining differences between laboratories may affect the reproducibility of results. stLearn has been reported to be applied in the subtyping of cell types in breast cancer boundary regions [44].

4. Matching method selection with data conditions

The previous section constructed a classification system based on the input information modalities and inherent assumptions of methods, and proposed a decision framework accordingly. However, the practicality of this framework remains to be empirically verified. The performance differences and failure modes of different methods on the same data need to be observed and compared in real data to determine whether these differences can be explained by the logic that "data characteristics determine method effectiveness". To this end, this chapter selects seven representative methods to conduct spatial clustering experiments on the same real tumor data, and compares the observations with the decision framework in Chapter 4.

4.1. Experimental setup

The experiment uses the 10x Visium human breast cancer dataset [12] (3,798 spots, with H&E images), which is one of the most commonly used benchmarks in spatial transcriptomics tumor research. The seven selected models cover the three levels in the classification system of this paper:

DOI: 10.54254/2753-8818/2026.ES36957

Level 1 selects Louvain as the non-spatial baseline; Level 2 selects STAGATE (graph neural network), SEDR (autoencoder), and BANKSY (spatial feature enhancement); Level 3 selects stMVC (multi-view fusion), ConGI (contrastive learning), and stLearn (morphological normalization). All models run with default parameters, and the number of clusters is automatically determined by the algorithm.

4.2. Spatial continuity and boundary identification ability

The spatial clustering results of the seven models on breast cancer tissue slices are shown in Figure 1. As a non-spatial baseline, Louvain's clustering results show a random scatter distribution on the slice, with no spatial correlation between adjacent spots, and no continuous block of the same color can be found across the entire slice. This result confirms the fundamental role of spatial information.

Among Level 2 methods, STAGATE's clustering results form continuous large blocks of the same color, with clearly distinguishable boundaries between the tumor core, invasive boundary, and normal tissue. SEDR also produces continuous spatial partitions, but only identifies 5 clusters, with relatively coarse details; some subregions that can be distinguished in STAGATE are merged in SEDR. BANKSY goes to the other extreme, identifying 25 clusters with dense color changes and greatly amplified local differences, but the problem of over-segmentation is also obvious.

Among Level 3 methods, stMVC identifies more substructures in the tumor core region than STAGATE, and some partitions correspond to histological features on H&E images. ConGI's results are roughly comparable to STAGATE's, but with finer division at the tumor-stroma interface. stLearn only identifies 6 clusters, with large blocks and few details.

Figure 1 shows the clustering results of each method, with the actual number of clusters marked in the upper right corner of each subplot: 8 for Louvain, 7 for STAGATE, 5 for SEDR, 25 for BANKSY, 11 for stMVC, 7 for ConGI, and 6 for stLearn. Comparing these observations with the decision framework in Chapter 4: the framework states that graph neural networks perform better in regular dense tissues, which is confirmed by STAGATE's results; the framework notes that spatial feature enhancement is sensitive to local differences but requires careful parameter tuning, and BANKSY's over-segmentation precisely illustrates this risk; the framework indicates that multi-view fusion helps improve fineness, and the substructures identified by stMVC in the core region are exactly the embodiment of this gain.

4.3. Association between failure modes and data characteristics

The fundamental reason for method failure lies in the match between its inherent assumptions and the true distribution of data. STAGATE assumes that the spatial adjacency graph can reliably reflect biological connections between cells — this assumption is largely valid in cell-dense breast cancer tissue. BANKSY assumes that neighborhood gradient information is sufficient to encode microenvironment differences — in real data, this assumption amplifies local noise. stMVC assumes that image information can supplement the insufficiency of expression information — when H&E image quality is high, this assumption indeed brings an improvement in fineness. stLearn assumes that morphological distance is positively correlated with expression distance — but real H&E images contain morphological noise unrelated to tumor parenchyma, such as necrotic and hemorrhagic regions. Louvain completely fails in spatial domain identification tasks because it does not utilize spatial information at all.

These methods show huge performance differences on the same data, but the direction and magnitude of the differences can be precisely explained by the inherent assumptions of the methods.

DOI: 10.54254/2753-8818/2026.ES36957

The performance summary of each method is shown in Table 1.

Table 1. Performance summary of each method on real data

Method	Level	Number of clusters	Spatial continuity	Boundary clarity	Degree of match with assumptions
Louvain	Level 1	8	Poor (random scatter)	None	— (no spatial assumption)
STAGATE	Level 2	7	Excellent	Clear	High
SEDR	Level 2	5	Good	Relatively clear	Medium
BANKSY	Level 2	25	Medium (over-segmented)	Overly fine	Medium (assumption amplifies noise)
stMVC	Level 3	11	Excellent	Clearest	High
ConGI	Level 3	7	Excellent	Relatively clear	Relatively high
stLearn	Level 3	6	Medium	Relatively blurred	Medium-low

4.4. Retrospection and support for the decision framework

The above analysis provides empirical support for the decision framework in Chapter 4. The judgment in the framework that "graph neural networks are suitable for regular dense tissues" is verified by STAGATE's results; the reminder that "spatial feature enhancement is sensitive but requires careful parameter tuning" is confirmed by BANKSY's over-segmentation; the statement that "multi-view fusion improves fineness" is supported by stMVC's substructure identification.

Meanwhile, these observations also provide a basis for the refinement of the framework. BANKSY's performance indicates that parameter tuning of spatial feature enhancement is more difficult on real data than on simulated data; stLearn's performance shows that morphological normalization has stricter requirements on image quality than described in its original paper.

4.5. Summary

This chapter verifies the core logic of the decision framework in Chapter 4 through spatial clustering experiments of seven representative methods on the same real tumor data. The clustering results are shown in Figure 1. The effectiveness of a method is highly correlated with the rationality of its inherent assumptions, and the rationality of the assumptions depends on data characteristics. The fundamental reason for the different performance of the same method on different data is the change in the degree of match between data characteristics and assumptions. This conclusion provides empirical support for the decision framework in Chapter 4, and also points out a key direction for method development: rather than pursuing ranking improvement in benchmark tests, it is better to more clearly define the most applicable data conditions for each method.

DOI: 10.54254/2753-8818/2026.ES36957

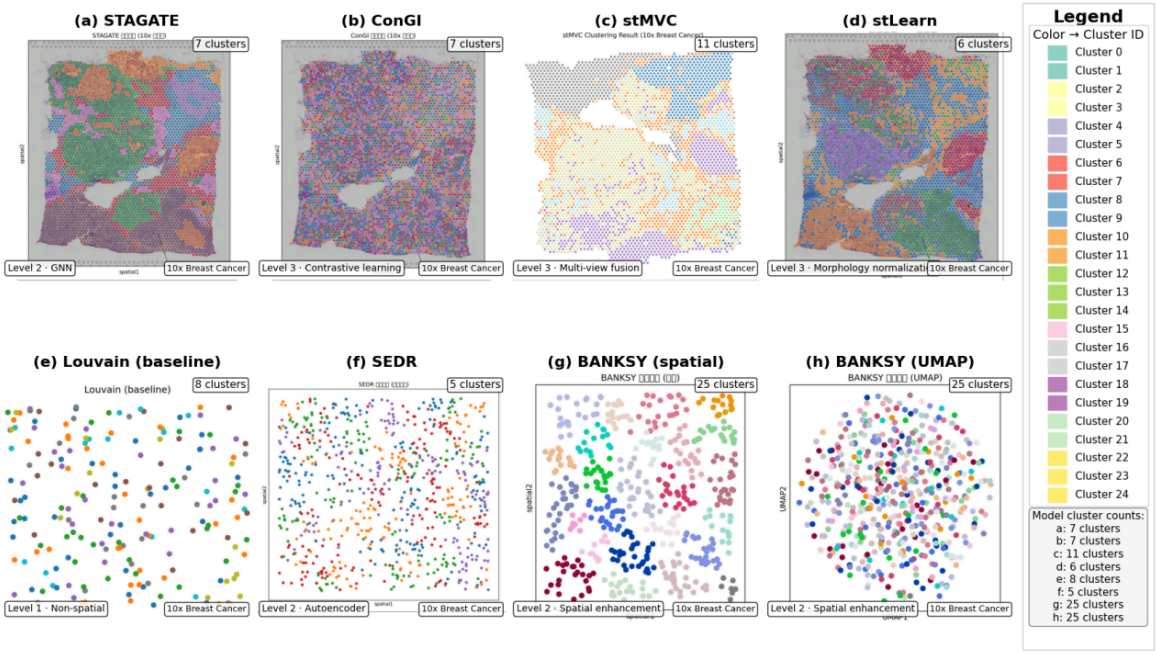


Figure 1. Comparison of spatial clustering results of representative methods at each level on the 10x Visium human breast cancer dataset

5. Coupling of data and methods and future directions

The difficulty of tumor spatial domain identification is essentially the conflict between data characteristics and method assumptions. Statistical modeling methods assume spatial smoothness but fail at invasive boundaries; graph neural network methods assume that adjacency graphs reflect biological similarity but degrade in sparse data; multi-view fusion methods introduce images and prior knowledge, but at the cost of relying on annotation quality and staining consistency. These conflicts need to be addressed collaboratively at both the data acquisition and method design levels.

At the data level, high-quality data for scarce cancer types such as pancreatic cancer and lung cancer should be prioritized for supplementation, especially publicly available datasets generated using FFPE-compatible platforms. Meanwhile, data repositories should encourage the upload of raw images and segmentation results. At the method level, it is necessary to develop graph models robust to sparse adjacency, self-supervised pre-training strategies that reduce the impact of image staining differences, active learning or weakly supervised methods that reduce reliance on manual annotation, and linear-complexity algorithms for large-scale data.

With the public availability of spatial transcriptomics data for more tumor types and advances in platform technology, the positive feedback between data and methods will continue to accelerate. In the future, spatial transcriptomics analysis will shift from the current "benchmark ranking-oriented" situation to a rational selection mode "oriented by data characteristics and biological problems". The decision framework and method classification system provided in this paper aim to provide a basis for this transition.

DOI: 10.54254/2753-8818/2026.ES36957

6. Conclusion

This paper systematically reviews the current status of spatial transcriptomics in tumor research from two dimensions: data resources and computational methods. At the data level, breast cancer data is the most abundant, while data for pancreatic cancer, HER2-positive small tumors, and lung cancer are severely insufficient, rooted in difficulties in sample acquisition, poor FFPE compatibility, and biopsy size limitations. At the method level, 21 mainstream computational methods are divided into three levels and six subclasses according to the richness of input information. Each type of method is based on specific inherent assumptions, which hold true in regular, dense tissue structures but fail in the discontinuous, sparse, and blurred boundary regions characteristic of tumors.

On this basis, this paper constructs a decision framework based on data characteristics and verifies its effectiveness through empirical analysis. Experimental results on 10x Visium breast cancer data show that the effectiveness of a method is highly correlated with the rationality of its inherent assumptions, and the rationality of the assumptions depends on data characteristics — methods with assumptions matching the data perform well; those with mismatched assumptions perform poorly. This conclusion supports the core view throughout the paper: the applicability of a method is determined by data characteristics, not by the complexity or advancement of the method itself.

Future work needs to develop graph models robust to sparse adjacency, self-supervised pre-training that reduces the impact of image staining differences, active learning strategies that reduce reliance on manual annotation, and linear-complexity algorithms for large-scale data. With the public availability of spatial transcriptomics data for more tumor types and advances in platform technology, the positive feedback between data and methods will continue to accelerate, ultimately driving spatial transcriptomics from technology-driven to biological discovery-driven.

References

[1] Rao, A., Barkley, D., Franca, G. S., et al. (2021). Exploring tissue architecture using spatial transcriptomics. *Nature, 596*, 211–220.
[2] Tian, L., Chen, F., & Macosko, E. Z. (2023). The expanding vistas of spatial transcriptomics. *Nature Biotechnology, 41*, 773–782.
[3] Moses, L., & Pachter, L. (2022). Museum of spatial transcriptomics. *Nature Methods, 19*, 534–546.
[4] Fang, S., Chen, B., Zhang, Y., et al. (2023). Computational approaches and challenges in spatial transcriptomics. *Genomics, Proteomics & Bioinformatics, 21*, 24–47.
[5] Ståhl, P. L., Salmén, F., Vickovic, S., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science, 353*, 78–82.
[6] Rodrigues, S. G., Stickels, R. R., Goeva, A., et al. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science, 363*, 1463–1467.
[7] Chen, A., Liao, S., Cheng, M., et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell, 185*, 1777–1792.e21.
[8] Moffitt, J. R., Bambah-Mukku, D., Eichhorn, S. W., et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science, 362*, eaau5324.
[9] Wang, X., Allen, W. E., Wright, M. A., et al. (2018). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science, 361*, eaat5691.
[10] He, S., Bhatt, R., Brown, C., et al. (2022). High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nature Biotechnology, 40*, 1794–1806.
[11] Janesick, A., Shelansky, R., Gottscho, A. D., et al. (2023). High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nature Communications, 14*, 8353.
[12] Ren, P., Zhang, R., Wang, Y., et al. (2025). Systematic benchmarking of high-throughput subcellular spatial transcriptomics platforms across human tumors. *Nature Communications, 16*, 212.

DOI: 10.54254/2753-8818/2026.ES36957

[13] Yue, Y., Fu, Y., Li, L., et al. (2024). Systematic comparison of sequencing-based spatial transcriptomic methods. *Nature Methods, 21*, 1743–1754.
[14] Xiong, C., Huang, S., Zhou, M., et al. (2025). A comprehensive comparison on clustering methods for multi-slice spatially resolved transcriptomics data analysis. *Briefings in Bioinformatics, 26*, bbaf471.
[15] Hu, Y., Xie, M., Li, Y., et al. (2024). Benchmarking clustering, alignment, and integration methods for spatial transcriptomics. *Genome Biology, 25*, 212.
[16] Maynard, K. R., Collado-Torres, L., Weber, L. M., et al. (2021). Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nature Neuroscience, 24*, 425–436.
[17] Xu, Z., Wang, W., Yang, T., et al. (2024). STOmicsDB: A comprehensive database for spatial transcriptomics data sharing, analysis and visualization. *Nucleic Acids Research, 52*, D1053–D1061.
[18] Wu, S. Z., Al-Eryani, G., Roden, D. L., et al. (2021). A single-cell and spatially resolved atlas of human breast cancers. *Nature Genetics, 53*, 1334–1347.
[19] Wu, Y., Yang, S., Ma, J., et al. (2022). Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer Discovery, 12*, 134–153.
[20] Li, Y., Du, Y., Li, R., et al. (2024). Spatial transcriptomics in pancreatic cancer: Advances, prospects and challenges. *Critical Reviews in Oncology/Hematology, 203*, 104430.
[21] Andersson, A., Larsson, L., Stenbeck, L., et al. (2021). Spatial deconvolution of HER2-positive breast tumors reveals novel intercellular relationships. *Nature Communications, 12*, 6012.
[22] Fan, Z., Chen, R., & Chen, X. (2020). SpatialDB: A database for spatially resolved transcriptomes. *Nucleic Acids Research, 48*, D233–D237.
[23] Blondel, V. D., Guillaume, J. L., Lambiotte, R., et al. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics, 2008*, P10008.
[24] Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports, 9*, 5233.
[25] Zhao, E., Stone, M. R., Ren, X., et al. (2021). Spatial transcriptomics at subspot resolution with BayesSpace. *Nature Biotechnology, 39*, 1375–1384.
[26] Li, Z., & Zhou, X. (2022). BASS: Multi-scale and multi-sample analysis enables accurate cell type clustering and spatial domain detection in spatial transcriptomic studies. *Genome Biology, 23*, 168.
[27] Shang, L., & Zhou, X. (2022). Spatially aware dimension reduction for spatial transcriptomics. *Nature Communications, 13*, 7203.
[28] Fu, Y., Nan, M., Ren, Q., et al. (2025). A multi-view graph convolutional network framework based on adaptive adjacency matrix and multi-strategy fusion mechanism for identifying spatial domains. *Bioinformatics, 41*, btaf172.
[29] Wang, C., & Yu, X. (2026). STransfer: A transfer learning-enhanced graph convolutional network for clustering spatial transcriptomics data. *Bioinformatics, 42*, btag049.
[30] Gui, Y., Tan, Z., Xu, Y., et al. (2025). Heterogeneous graph contrastive learning for integration and alignment of spatial transcriptomics data. *Briefings in Bioinformatics, 26*, bbaf497.
[31] Dong, K., & Zhang, S. (2022). Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention autoencoder. *Nature Communications, 13*, 1739.
[32] Ren, H., Walker, B. L., Cang, Z., et al. (2022). Identifying multicellular spatiotemporal organization of cells with SpaceFlow. *Nature Communications, 13*, 4076.
[33] Long, Y., Ang, K. S., Li, M., et al. (2023). Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nature Communications, 14*, 1155.
[34] Zixuan, C., Ning, X., Nie, A., et al. (2021). SCAN-IT: Domain segmentation of spatial transcriptomics images by graph neural network. *BMVC*.
[35] Xu, H., Fu, H., Long, Y., et al. (2024). Unsupervised spatially embedded deep representation of spatial transcriptomics. *Genome Medicine, 16*, 12.
[36] Hu, Y., Zhao, Y., Schunk, C. T., et al. (2023). ADEPT: Autoencoder with differentially expressed genes and imputation for robust spatial transcriptomics clustering. *iScience, 26*, 106792.
[37] Xu, C., Jin, X., Wei, S., et al. (2022). DeepST: Identifying spatial domains in spatial transcriptomics by deep learning. *Nucleic Acids Research, 50*, e131.
[38] Singhal, V., Chou, N., Lee, J., et al. (2024). BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nature Genetics, 56*, 431–441.
[39] Yang, Y., Cui, Y., Zeng, X., et al. (2025). STAIG: Spatial transcriptomics analysis via image-aided graph contrastive learning for domain exploration and alignment-free integration. *Nature Communications, 16*, 1067.
[40] Noh, M., Lee, S., Kim, S., et al. (2026). PathCLAST: Pathway-augmented contrastive learning with attention for interpretable spatial transcriptomics. *Briefings in Bioinformatics, 27*, bbag029.

DOI: 10.54254/2753-8818/2026.ES36957

[41] Zuo, C., Zhang, Y., Cao, C., et al. (2022). Elucidating tumor heterogeneity from spatially resolved transcriptomics data by multi-view graph collaborative learning. *Nature Communications, 13*, 5962.
[42] Zuo, C., Xia, J., & Chen, L. (2024). Dissecting tumor microenvironment from spatially resolved transcriptomics data by heterogeneous graph learning. *Nature Communications, 15*, 5057.
[43] Yang, Y., Zhang, C., Liu, Z., et al. (2024). MCGAE: Unraveling tumor invasion through integrated multimodal spatial transcriptomics. *Briefings in Bioinformatics, 26*, bbae608.
[44] Pham, D., Tan, X., Balderson, B., et al. (2023). Robust mapping of spatiotemporal trajectories and cell-cell interactions in healthy and diseased tissues. *Nature Communications, 14*, 7739.