

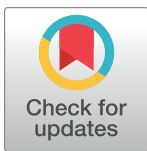
## RESEARCH ARTICLE

# MVST: Identifying spatial domains of spatial transcriptomes from multiple views using multi-view graph convolutional networks

Hao Duan<sup>1</sup>, Qingchen Zhang<sup>1</sup>, Feifei Cui<sup>1\*</sup>, Quan Zou<sup>2,3\*</sup>, Zilong Zhang<sup>1\*</sup>

**1** School of Computer Science and Technology, Hainan University, Haikou, China, **2** Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China, **3** Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou, China

\* [feifeicui@hainanu.edu.cn](mailto:feifeicui@hainanu.edu.cn) (FC); [zouquan@nclab.net](mailto:zouquan@nclab.net) (QZ); [zhangzilong@hainanu.edu.cn](mailto:zhangzilong@hainanu.edu.cn) (ZZ)



## OPEN ACCESS

**Citation:** Duan H, Zhang Q, Cui F, Zou Q, Zhang Z (2024) MVST: Identifying spatial domains of spatial transcriptomes from multiple views using multi-view graph convolutional networks. PLoS Comput Biol 20(9): e1012409. <https://doi.org/10.1371/journal.pcbi.1012409>

**Editor:** Jean Fan, Johns Hopkins University Whiting School of Engineering, UNITED STATES OF AMERICA

**Received:** April 12, 2024

**Accepted:** August 9, 2024

**Published:** September 5, 2024

**Copyright:** © 2024 Duan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The spatial transcriptomics datasets in this article are available here: (1) 10x Visium human breast cancer dataset: <https://support.10xgenomics.com/spatial-gene-expression/datasets>; (2) 10x Visium human dorsolateral prefrontal cortex dataset: <https://github.com/LieberInstitute/HumanPilot>; (3) The mouse brain anterior section dataset: <https://www.10xgenomics.com/resources/datasets>. The code implementation of the MVST algorithm has been

## Abstract

Spatial transcriptome technology can parse transcriptomic data at the spatial level to detect high-throughput gene expression and preserve information regarding the spatial structure of tissues. Identifying spatial domains, that is identifying regions with similarities in gene expression and histology, is the most basic and critical aspect of spatial transcriptome data analysis. Most current methods identify spatial domains only through a single view, which may obscure certain important information and thus fail to make full use of the information embedded in spatial transcriptome data. Therefore, we propose an unsupervised clustering framework based on multiview graph convolutional networks (MVST) to achieve accurate spatial domain recognition by the learning graph embedding features of neighborhood graphs constructed from gene expression information, spatial location information, and histopathological image information through multiview graph convolutional networks. By exploring spatial transcriptomes from multiple views, MVST enables data from all parts of the spatial transcriptome to be comprehensively and fully utilized to obtain more accurate spatial expression patterns. We verified the effectiveness of MVST on real spatial transcriptome datasets, the robustness of MVST on some simulated datasets, and the reasonableness of the framework structure of MVST in ablation experiments, and from the experimental results, it is clear that MVST can achieve a more accurate spatial domain identification compared with the current more advanced methods. In conclusion, MVST is a powerful tool for spatial transcriptome research with improved spatial domain recognition.

## Author summary

Spatial transcriptome sequencing can not only reveal the mechanisms of disease development, but can also be used to explore the structure of biological tissues, which are widely used in a variety of fields such as developmental biology and oncology. To utilize spatial transcriptome data to understand how different cells work together to perform complex functions, spatially relevant groups of cells must be identified, which leads to the task of

deposited at <https://github.com/dw666666/MVST.git>.

**Funding:** The work was supported by the National Natural Science Foundation of China (No. 62102064 to Z.L.Z., 62131004 to Q.Z., 62261018 to Z.L.Z.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

dissecting tissues into spatial domains. However, most of the currently available spatial domain identification tools have yet to fully utilize the information embedded in spatial transcriptomic data, and their identification remains to be improved for the fine-grained exploration of tissue structure. Here, we developed a tool for accurate spatial domain identification, and our tool makes full use of the various types of spatial transcriptome data from multiple perspectives, identifying the real grouping information in spatial transcriptome data as much as possible. Our results show that our spatial domain identification tool can identify spatial domains more accurately and provide effective help for biological organization structure exploration.

## Introduction

To explore the spatial structure of tissues and the spatial expression patterns of genes, numerous spatial transcriptome (ST) sequencing technologies have been developed to quantify the expression of a large number of genes in the spatial context of tissues and cells [1–3]. Transcript information obtained using spatial transcriptome technologies can correspond to tissue locations, enabling the restoration of gene expression status at various spatial locations within a tissue [4–6]. From the development history of spatial transcriptome sequencing technologies, they can be meticulously classified into five subcategories at present: i) technologies based on microdissected gene expression, e.g., LCM [7], ProximID [8]; ii) in situ hybridisation (ISH) technologies, e.g., seqFISH [9,10], MERFISH [11], osmFISH [12]; iii) in situ sequencing (ISS) technologies, such as STARmap [13]; iv) in situ capturing technologies, such as Slide-Seq [14], HDST [15], and the current mainstream spatial transcriptome sequencing technology-10X Visium (<https://www.10xgenomics.com/>), and v) in silico reconstruction of spatial data [16]. Spatial transcriptomics can not only help elucidate the spatial dynamics of normal organ development but also assist in the study of the mechanisms of disease onset and progression [17], and play a significant role in exploring the tumor microenvironment [18–21].

Identifying spatial domains refers to identifying regions with similarities in gene expression and histology. Identifying spatial domains is the most basic and critical aspect of spatial transcriptome data analysis, and is one of the major challenges in spatial transcriptome research. The identification of spatial domains can further demonstrate the spatial dependence of gene expression and provide a basis for further studies such as spatially variable gene identification. The identification of spatial domains can be achieved using traditional clustering methods and analytical tools developed for spatial domain identification.

Earlier methods used for spatial domain recognition, that is, traditional clustering methods such as k-means [22,23] and Louvain [24], which do not consider spatial location, are not ideal. Various spatial domain recognition methods have been developed that combine spatial location information, gene expression data, and histopathological images to better identify the spatial domains. For example, BayesSpace [25] achieves spatial domain recognition based on a spatial prior, that is, neighboring points have a high probability of belonging to the same cluster. Giotto [26] used a Hidden Markov Random Field model and spatial neighbors prior to identifying spatial domains. Currently, rapidly developing deep learning models such as graph neural networks have been applied to spatial domain recognition with good results. For example, SEDR [27] uses a variational graph self-encoder to embed spatial information into low-dimensional representations of gene expression to identify spatial domains. SpaGCN [28] identifies spatial structural domains by aggregating gene expressions from neighboring points using graph convolutional neural networks. CCST [29] employs an unsupervised graph

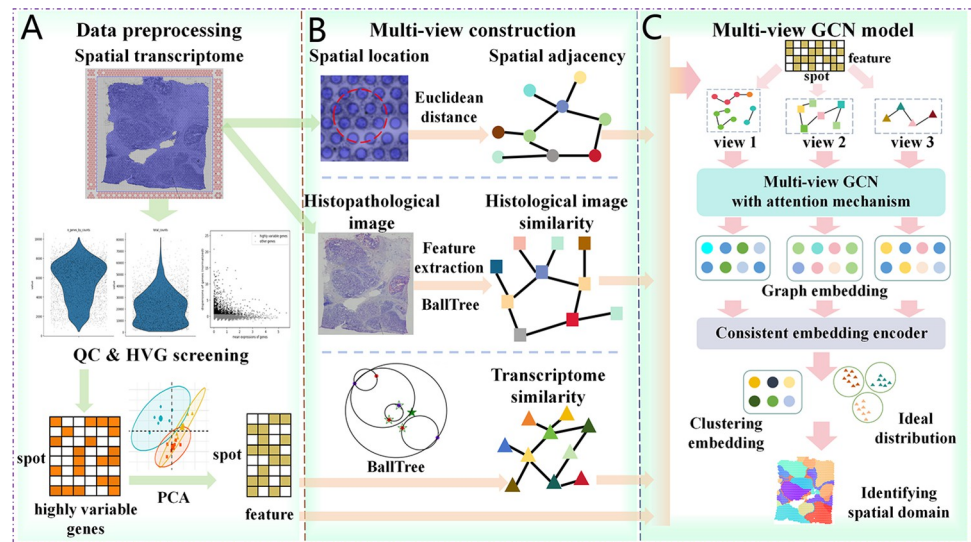
contrast learning model with mutual information maximization to learn node embeddings for identifying spatial domains. STAGATE [30] used a graph attention self-encoder to aggregate spatial and gene expression information to identify spatial domains and employed an attention mechanism to better explore spatial similarities at the boundaries of spatial domains. STMGCN [31] uses a graph convolutional network-based encoder to learn specific representations of multiple topological graphs to identify spatial domains. CONGI [32] uses contrast learning to adapt gene representations to image information and to accurately identify spatial domains. Most of the above methods construct adjacency matrices based on spatial locations and do not make full use of information, such as gene expression data and histopathological images, to construct graph structures. Currently, studies have shown that transcriptomics data and histopathology images have a common link [33], and more studies have achieved better results in spatial transcriptomics prediction, exploration of molecular features of spatial transcriptomes, prediction of clinical outcomes, and noise correction of spatial transcriptomics data by using histopathology image information [34–36]. Therefore, increasing the analysis of histopathology images can provide more comprehensive information for spatial transcriptomics-related studies, including spatial domain identification, to some extent. Although some methods also consider the use of information such as histopathological images, they do not analyze them in an all-round and multi-perspective manner, which may result in missing important information characterizing spatial expression patterns.

In this study, we propose a spatial domain recognition method, MVST, based on a multi-view graph convolutional network, which can combine gene expression data, spatial location information, and histopathological image information to learn low-dimensional embedding and achieve accurate recognition of spatial domains. MVST first constructs three undirected graphs based on gene expression data, spatial location, and histopathological images to represent the dependency of the data and features obtained by the dimensionality reduction of gene expression data as attributes of graph nodes. Second, low-dimensional graph embeddings are learned from each undirected graph by a multiview graph convolutional encoder with an attention mechanism and further fed into a coherent embedding encoder to obtain consistent cluster embeddings among multiple views, which are then used for spatial domain recognition. We tested the clustering performance of MVST using three real datasets from the current mainstream platform for spatial transcriptomes, 10x Visium, and some simulated datasets constructed based on the real datasets, as well as ablation experiments, and the results of the experiments show that compared with the current more advanced methods MVST has certain advantages in clustering analysis and can identify spatial domains more accurately.

## Results

### Overview of the MVST

MVST learns the embedded features of spots from three perspectives—gene expression, spatial location, and histopathological images—and thus identifies spatial domains. First, the spatial transcriptome gene expression data were preprocessed by quality control and normalization to obtain highly variable genes (HVGs) that expressed a high degree of variation among spots, and principal component analysis (PCA) was performed to reduce their dimensionality (Fig 1A). Next, the Euclidean distance between the spots was calculated based on the spatial coordinate position information of the spots to derive a spatial adjacency network. A gene expression similarity network was constructed through the nearest neighbor algorithm based on the gene expression feature information of the spots obtained from the preprocessing, and the histopathology image feature information corresponding to the spots was extracted and the histology image similarity network was constructed using the nearest neighbor algorithm.



**Fig 1. Workflow of MVST.** (A) MVST uses SCANPY to perform routine data preprocessing operations such as quality control and highly variable gene screening on spatial transcriptome gene expression data. PCA is used to downscale highly variable gene expression data to obtain a relatively low-dimensional feature representation of spot gene expression data. (B) The graph construction process of MVST reflects its multi-perspective feature, constructing graphs based on distance similarity, histopathological image similarity, and gene expression similarity. The spatial adjacency, histological image similarity, and gene expression similarity networks are obtained through this process. (C) The multi-view graph convolutional model of MVST includes a multi-view graph convolutional encoder with an attention mechanism and a coherent embedding encoder, in which the multi-view graph convolutional encoder learns the spot features from each of the three views, and the coherent embedding encoder integrates the features of the spot in each view to obtain the final clustered embedding.

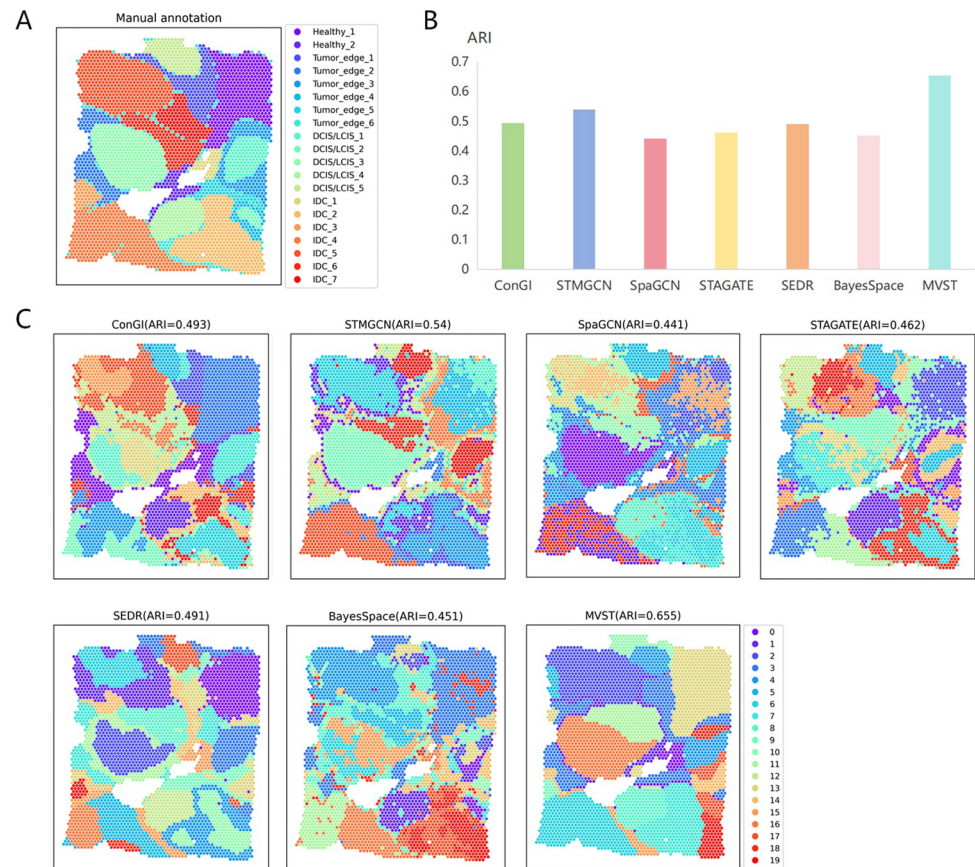
<https://doi.org/10.1371/journal.pcbi.1012409.g001>

This enabled the characterization of the interconnections between the spots (Fig 1B). Finally, the gene expression features and the constructed three-view networks were input into a multi-view graph convolutional encoder (the multi-view graph convolutional encoder contains three graph convolutional encoders, each of which corresponds to one view network) with an attention mechanism to learn the low-dimensional graph embeddings of the spots in each view. The resulting graph embeddings were further input into a coherent embedding encoder to obtain the final clustered embeddings and the ideal distributions based on geometric relationships and probability distributions, and then to identify the spatial domains (Fig 1C).

## MVST applied to human breast cancer tissue

To quantitatively evaluate the performance of MVST for the spatial domain recognition of human cancer tissues, we first applied it to a human breast cancer tissue dataset, a 10X Visium dataset. The human breast cancer tissue dataset, which has been manually annotated, consists of four main morphological types: ductal carcinoma in situ (DCIS/LCIS), healthy tissue, invasive ductal carcinoma (IDC), and tumor margins, and has been further subdivided into 20 regions (Fig 2A). We compared MVST with existing state-of-the-art methods on the benchmark dataset; specifically, MVST was compared with six more advanced spatial clustering methods, namely, ConGI, STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace, in terms of clustering accuracy in terms of the adjusted RAND index (ARI). The results showed a clear advantage over the other six spatial clustering algorithms, with MVST achieving the best clustering accuracy (ARI = 0.655), and its performance was 11.5% higher in terms of ARI than that of the STMGCN algorithm, which was the second-best performer (Fig 2B and 2C).





**Fig 2. MVST improves spatial domain recognition of human breast cancer tissues.** (A) Manual annotation of the spatial domain of the human breast cancer dataset by dividing the tissue slices into 20 regions. (B) Histogram of the performance of spatial domain recognition of human breast cancer dataset using MVST and existing state-of-the-art algorithms (ConGI, STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace), with the X-axis showing the names of the algorithms and the Y-axis showing the ARI value of the spatial domain recognition results of each algorithm. This is used to compare the results of the predicted spatial domains of each algorithm to the similarity of manually annotated regions. (C) ARI values and visualizations of spatial domain recognition results from MVST and other algorithms for the human breast cancer dataset.

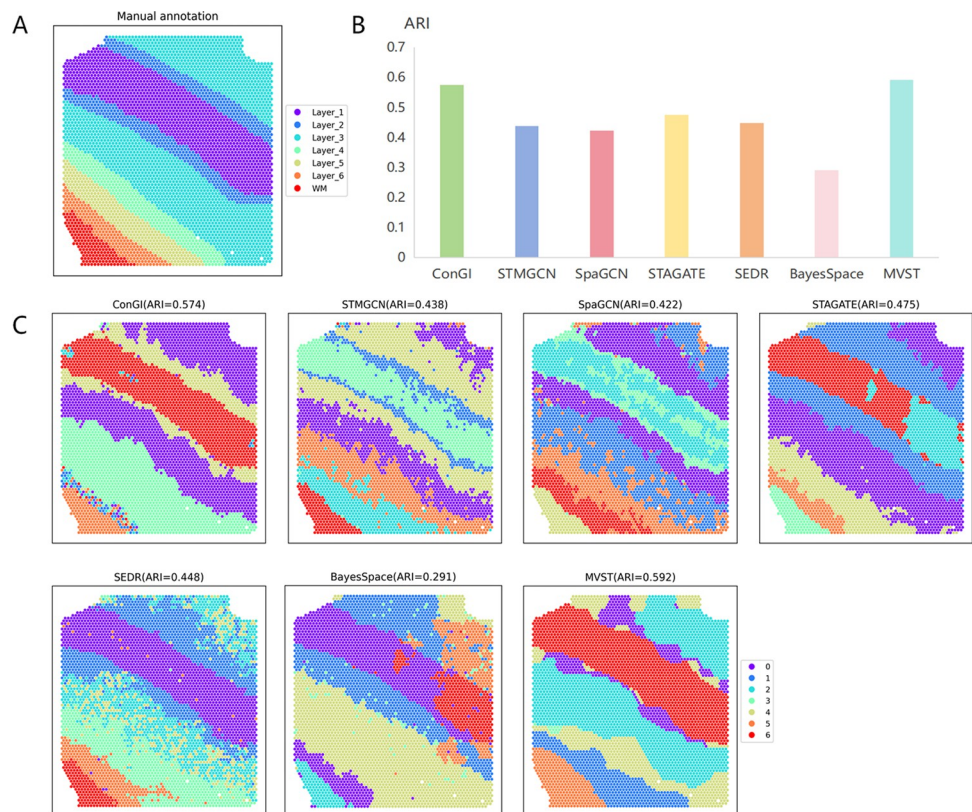
<https://doi.org/10.1371/journal.pcbi.1012409.g002>

By visualizing the spatial domain identification results of MVST with the other five spatial clustering algorithms on the human breast cancer tissue dataset, it is clear that MVST can identify the expected tissue structure more efficiently (Fig 2C). The clustering results of ConGI, STMGCN, SpaGCN, STAGATE, and BayesSpace roughly conformed to the tissue structure of human breast cancer tissue in the major partitions; their clustering boundaries were discontinuous, and there were more anomalies, which affected the accuracy of clustering. The clustering results of the SEDR, although with clearer boundaries, were still far from the actual human breast cancer tissue structure. The clustering results of STMGCN, SpaGCN, STAGATE, and SEDR clearly confused non-tumor regions and some tumor regions, which had a negative impact on the study of tumor tissues in patients with cancer. The clustering results of MVST have clearer boundaries, and at the same time, they are more consistent with the human breast cancer tissue structure partitioning; MVST has a particularly good recognition effect for Healthy\_1, and the IDC part of the region is particularly well recognized. These results illustrate the superiority of MVST for identifying the spatial domains of human tumor tissues.

## MVST applied to human dorsolateral prefrontal cortex tissue

Next, we applied MVST to a human dorsolateral prefrontal cortex (DLPFC) tissue dataset, a 10X Visium dataset, to further test the ability of MVST to identify the spatial domains of normal human tissues. The human dorsolateral prefrontal cortex (DLPFC) tissue dataset mentioned above contains 12 tissue section datasets that have been manually annotated based on genetic markers and cellular structures, and the human DLPFC tissue samples are divided into six cortical layers and one white matter layer. We selected slices 151509 and 151510 from this human dorsolateral prefrontal cortex tissue dataset as the benchmark dataset and compared our MVST method with six spatial clustering methods, namely ConGI, STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace, which are more advanced spatial clustering methods for spatial domain identification, thus verifying the good performance of our method.

The spatial domain division of the manual annotation of human dorsolateral prefrontal cortex tissue slices 151509 is shown in Fig 3A. For the human dorsolateral prefrontal cortex tissue section 151509 dataset, first, with respect to the evaluation criterion of the ARI used, the methods STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace performed less than optimally, and ConGI was similar to the ARI value of the result of spatial domain identification



**Fig 3. MVST enhances spatial domain recognition of human dorsolateral prefrontal cortex tissue slice 151509.** (A) Manual annotation of the spatial domain of human dorsolateral prefrontal cortex tissue slice 151509 dataset by dividing the tissue slice into seven layers. (B) Histogram of the performance of spatial domain recognition of human dorsolateral prefrontal cortex tissue slices 151509 dataset using MVST and existing state-of-the-art algorithms (ConGI, STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace), with the X-axis showing the names of the algorithms, and the Y-axis showing the ARI of the spatial domain recognition results value of each algorithm, which is used to compare the similarity between the spatial domains predicted by each algorithm and the manually annotated layers. (C) ARI values and visual presentation of the spatial domain recognition results of MVST and other algorithms for the human dorsolateral prefrontal cortex tissue slice 151509 dataset.

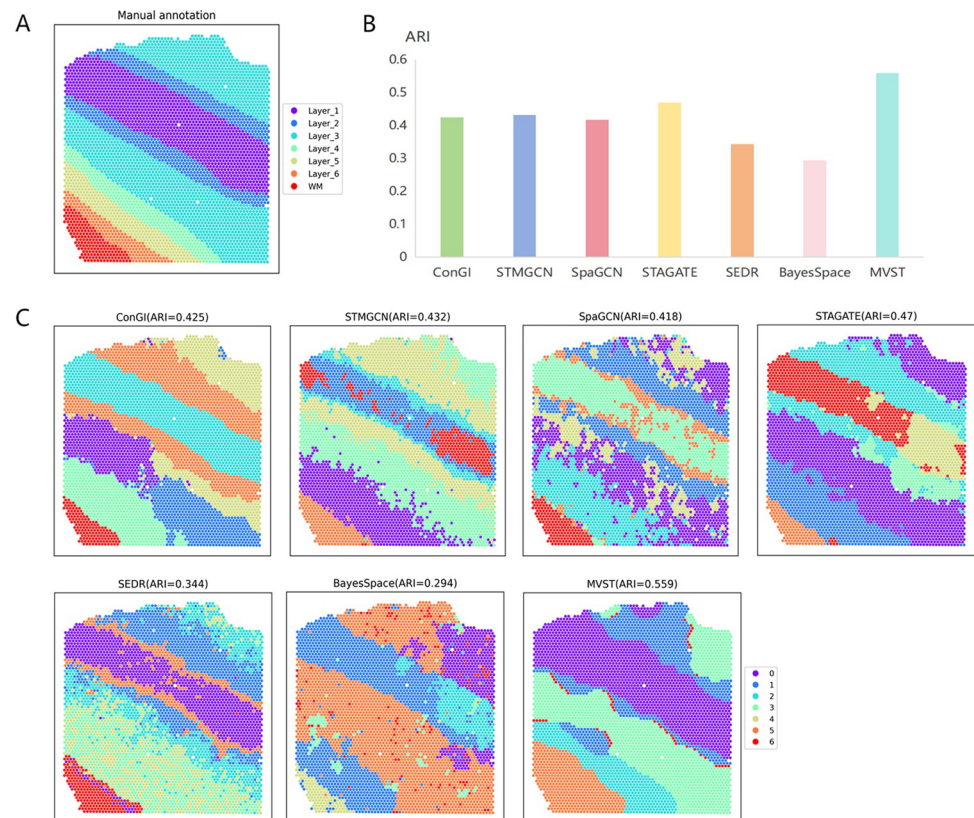
<https://doi.org/10.1371/journal.pcbi.1012409.g003>

obtained by our method. However, it was still slightly inferior, approximately 2% lower than the ARI values of the results obtained by our method (Fig 3B and 3C). In contrast, MVST achieved the highest clustering accuracy (ARI = 0.592) for the DLPFC151509 dataset (Fig 3C). Second, from the visualization of the spatial domain recognition results of all the methods used, all these methods roughly identified the hierarchical structure of the human dorsolateral prefrontal cortex tissue slices 151509; however, there were also large differences in the details (Fig 3C). Specifically, the tissue layering identified by BayesSpace was notably different from the true layered structure of human dorsolateral prefrontal cortex tissue section 151509. The three methods, STMGCN, SpaGCN, and SEDR, identified the overall organizational structure of the human dorsolateral prefrontal cortex tissue in the distribution; however, their spatial domain recognition results contained a large number of outliers, which could not correctly distinguish the various layers of tissues, and had a poor grasp of the spatial patterns of the spots. STAGATE's spatial domain identification results on the DLPFC151509 dataset, although containing only a relatively small number of anomalies, had a large error in the partitioning of each organizational layer, which may have occurred because of an overreliance on spatial relationships between spots during spatial domain identification. Although ConGI identifies the organizational layers Layer\_1, Layer\_2, Layer\_3, and VM more smoothly and clearly, it obfuscates Layer\_4, Layer\_5, and Layer\_6 and identifies numerous anomalies in Layer\_6. Our MVST method not only identifies the DLPFC151509 organizational hierarchy as correctly as possible in general, but also the identification of layers Layer\_4, Layer\_5 and Layer\_6 by MVST is significantly more detailed than by ConGI.

Next, the spatial domains of another tissue slice of the human dorsolateral prefrontal cortex, slice 151510, were identified using our MVST method, along with the six methods mentioned above, which are currently more advanced. Manually annotated spatial domain delineation of the human dorsolateral prefrontal cortex tissue slice 151510 is shown in Fig 4A, which is similar to the delineation of the DLPFC151509 slices. A plot of the spatial domain recognition results and a comparison of the clustering accuracy of each method for the human dorsolateral prefrontal cortex dataset slice 151510 are shown in Fig 4B and 4C, respectively. It can be seen that MVST is closest to the manually annotated stratification in terms of the ARI value, its spatial domain recognition results have the highest clustering accuracy compared to that of the other method (ARI = 0.559), and it outperforms the second-best-performing ConGI method by approximately 13%, which is a clear advantage. Second, in terms of the visualization of the spatial domain identification results, the spatial domains identified by MVST were relatively more consistent with manual annotation stratification. Although the boundaries of the spatial domains identified by ConGI and STAGATE were more distinct, some layers of the identified spatial domains were divided into two halves, which is inconsistent with the general rule of manual annotation layering. The spatial domains identified by the three methods, STMGCN, SpaGCN, and SEDR, were similar to their performances on DLPFC151509, with point-level noise, showing obvious fragmentation characteristics. The spatial domain identified by BayesSpace was similar to its performance on DLPFC151509, with the identified organizational hierarchies differing from the true hierarchical structure in a nontrivial manner.

From the above spatial domain recognition results and their analyses, compared with the six currently more advanced spatial domain recognition methods, namely ConGI, STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace, MVST has a stronger spatial domain recognition capability for human dorsolateral prefrontal cortex tissues, and the identified spatial domains are more in line with the true stratification of the tissues.





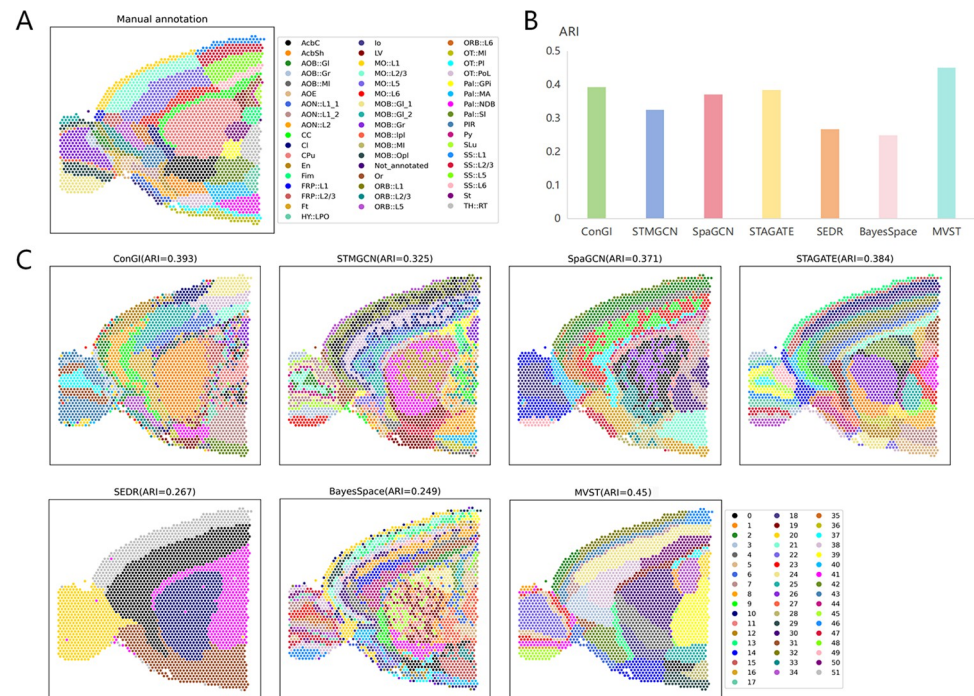
**Fig 4. MVST enhances spatial domain recognition of human dorsolateral prefrontal cortex tissue slice 151510.** (A) Manual annotation of the spatial domain of human dorsolateral prefrontal cortex tissue slice 151510 dataset by dividing the tissue slice into seven layers. (B) Histogram of the performance of spatial domain recognition of the human dorsolateral prefrontal cortex tissue slices 151510 dataset using MVST and existing state-of-the-art algorithms (ConGI, STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace), with the X-axis showing the name of each algorithm, and the Y-axis showing the ARI of the spatial domain recognition results of each algorithm value, which is used to compare the similarity between the spatial domains predicted by each algorithm and the manually annotated layers. (C) ARI values and visual presentation of spatial domain recognition results from MVST and other algorithms for the human dorsolateral prefrontal cortex tissue slice 151510 dataset.

<https://doi.org/10.1371/journal.pcbi.1012409.g004>

## MVST applied to mouse anterior brain tissue

In addition to applying MVST to a spatial transcriptome dataset of human tissues, we also applied MVST to a dataset of mouse anterior brain tissue slices to further test the spatial domain recognition capability of our method in tissues of species other than humans. The slices of the mouse anterior brain tissue slice dataset that we used were manually annotated, and they were more meticulously divided into 52 spatial domains (Fig 5A), which posed a considerable challenge to the spatial domain recognition algorithm. To evaluate the performance of MVST for spatial domain recognition on mouse anterior brain tissue slices, we compared MVST with the existing state-of-the-art methods mentioned above on the above benchmark dataset, and the results showed that our MVST method also achieved the best spatial domain recognition results in comparison with the mouse anterior brain tissue slice dataset. The clustering precision of the spatial domain recognition results, ARI, was up to 0.45, which is more than 5% higher than the ARI value obtained by the next best performing method, ConGI (Fig 5B and 5C). In terms of the delineation of the spatial domain, MVST has a better grasp of the general structure of the mouse anterior brain tissue slices; in particular, it better identifies





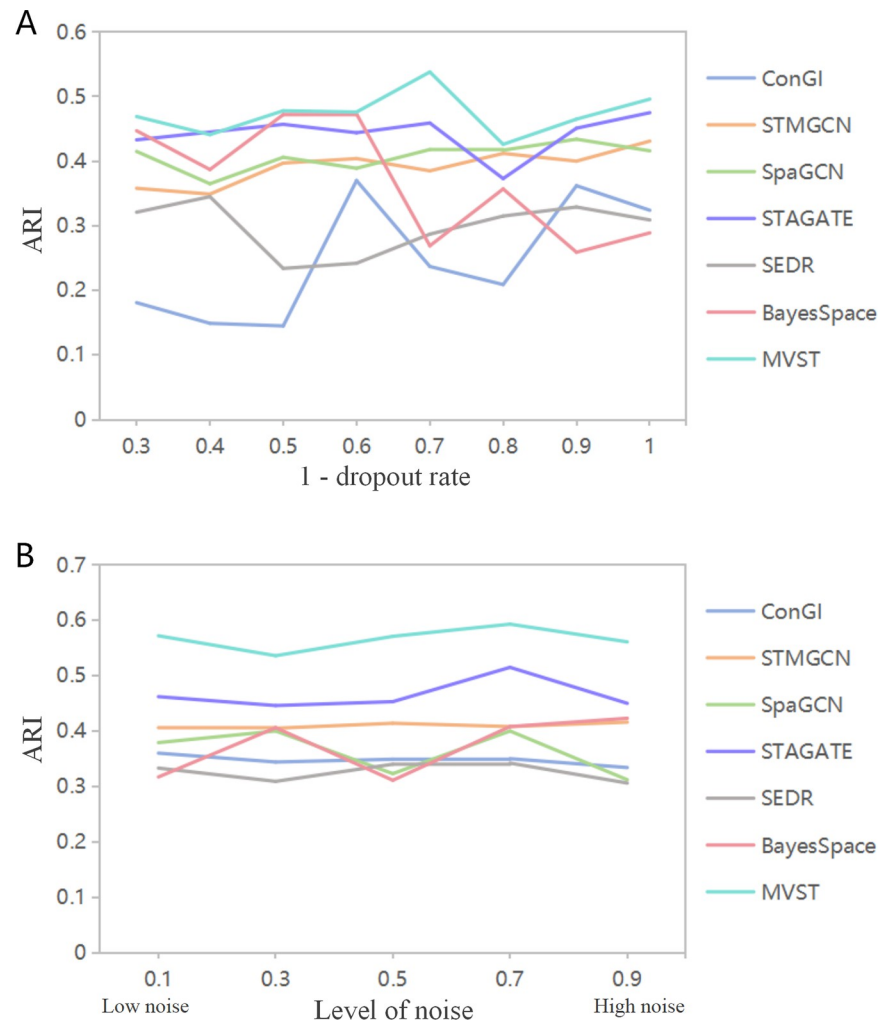
**Fig 5. MVST improves the spatial domain recognition of mouse anterior brain tissue slices.** (A) Manual annotation of the spatial domains of the mouse anterior brain tissue slice dataset by dividing the tissue slices into 52 regions. (B) Histogram of the performance of spatial domain recognition of the mouse anterior brain tissue slices dataset using MVST and existing state-of-the-art algorithms (ConGI, STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace), with the X-axis showing the names of the algorithms, and the Y-axis showing the ARI value of the spatial domain recognition results of each algorithm, which is used to compare the similarity between the spatial domains predicted by each algorithm and the manually annotated layers. (C) ARI values and visual presentation of spatial domain recognition results of MVST and other algorithms on the mouse anterior brain tissue slice dataset.

<https://doi.org/10.1371/journal.pcbi.1012409.g005>

the MOB region of the mouse anterior brain tissue. The boundaries of the spatial domain identification results of MVST are smoother, and at the same time there is almost no point-level noise. For the mouse anterior brain tissue slice dataset, SEDR was unable to identify the spatial domains in detail, and only a few rough delineations of the tissue slices were performed. Although ConGI, STMGCN, SpaGCN, and BayesSpace can perform a more detailed delineation of the spatial domains, the spatial domains identified by these methods appear slightly chaotic and more fragmented. The spatial domains delineated by STAGATE are more detailed, and the delineation boundaries are smoother; however, their similarity to the real configuration of the tissue in the anterior part of the mouse brain is lower, which can also be seen in the ARI values. In summary, MVST can achieve good results in the spatial domain recognition of mouse tissues and can also perform more detailed spatial domain recognition of tissues with complex structures.

### MVST applied to simulated data

As a result of experiments on simulated datasets that may better reflect the wide applicability and robustness of spatial domain recognition methods, we applied and compared the performance of MVST and six more advanced spatial domain recognition methods, such as ConGI, to simulated datasets constructed based on the human dorsolateral prefrontal cortex tissue slice 151509 dataset to further validate the robustness of MVST. The experimental results showed that the spatial domain recognition effect of ConGI and BayesSpace fluctuated greatly



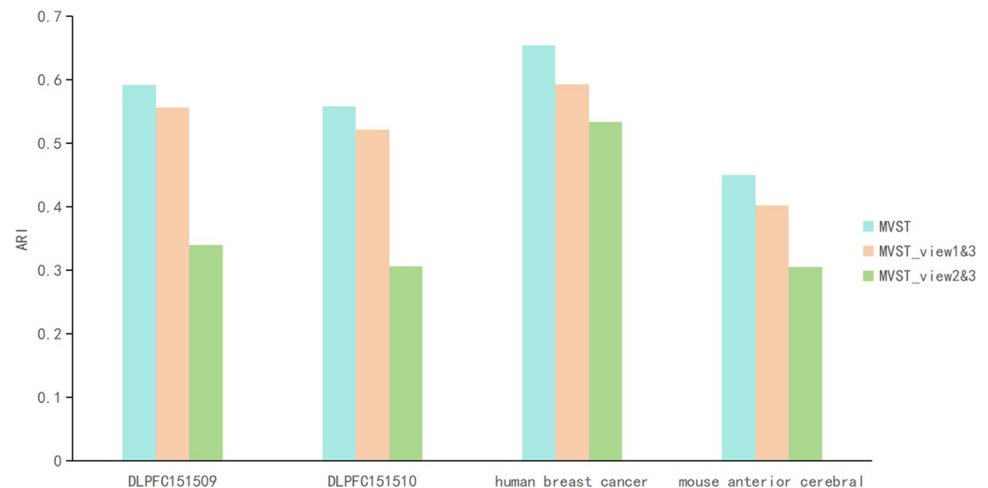
**Fig 6. Performance comparison between MVST and six comparison methods on simulated data.** (A) Performance comparison between MVST and six comparison methods when facing different sparsities of gene expression matrix; the X-axis represents different sparsities of gene expression matrix and the Y-axis represents ARI values. (B) Performance comparison between MVST and the six comparative methods when facing simulated data with different levels of added noise. The X-axis represents the level of added noise and the Y-axis represents the ARI value.

<https://doi.org/10.1371/journal.pcbi.1012409.g006>

in the face of changes in gene expression matrix sparsity, whereas the performance of STMGCN, SpaGCN, STAGATE, SEDR, and MVST was less affected by changes in gene expression matrix sparsity, and its performance was relatively stable (Fig 6A). Moreover, from the ARI values, the spatial domain recognition accuracy of our MVST method was mostly higher than that of the other methods when facing gene expression matrices with different sparsities. When facing simulated data with different levels of noise added, MVST and the remaining six spatial domain recognition methods show good robustness, and, from the ARI values, the spatial domain recognition accuracies of our method MVST are always higher than those of the other compared methods (Fig 6B). In summary, we conclude that the MVST is highly robust.

### Ablation study

In spatial transcriptomics, it is often assumed that cells adjacent to each other at spatial locations will exhibit similar gene expression profiles. The architecture of MVST uses both spatial



**Fig 7. Comparison of the performance of MVST with its two variants (MVST\_view1&3 and MVST\_view2&3) on the human dorsolateral prefrontal cortex, mouse anterior cerebral slice, and human breast cancer datasets, with the X-axis representing the different datasets and the Y-axis representing the ARI values.**

<https://doi.org/10.1371/journal.pcbi.1012409.g007>

neighborhood network maps and gene expression similarity network maps, which implies that MVST may capture redundant information, so we conducted ablation experiments by designing two variants of MVST to explore whether the possible capture of redundant information would have an impact on the spatial domain recognition effect of MVST. The first MVST variant uses only two views of the spatial adjacency network graph and histological image similarity network graph for spatial domain identification and is named MVST\_view1&3. The second MVST variant uses only two views of the gene expression similarity network graph and the histological image similarity network graph for spatial domain identification and is named MVST\_view2&3. We applied MVST and its two variants to the human dorsolateral prefrontal cortex dataset, mouse anterior brain slice dataset, and human breast cancer dataset used in the above ablation experiments. The results of the ablation experiments showed that the spatial domain recognition of MVST was better than that of the two variants of MVST on all three real spatial transcriptome datasets, including the human dorsolateral prefrontal cortex dataset, demonstrating that the framework structure of MVST is necessary and more reasonable (Fig 7). In addition, the results of the ablation experiments show that spatial and gene expression information contain cell grouping information that may not be identical; these two can complement each other in MVST to achieve better results, and there is no degradation of the spatial domain recognition performance owing to the possible redundancy of information. In the ablation experiments, MVST\_view1&3 outperformed MVST\_view2&3, suggesting that spatial location neighbourhood information may be more important than gene expression similarity information. This phenomenon corresponds to the fact that spatial information is crucial for spatial domain identification as described in the SpaGCN correspondence literature and others.

## Discussion

To fully integrate the spatial information, gene expression information, and histopathological image information of spatial transcriptomic data to explore the structure of organ tissues, we proposed MVST in this study. This process captures the similarity between spots in terms of gene expression, spatial location, and histopathological images from spatial transcriptomic data separately and learns the embedding of each spot using a multiview map convolutional

encoder with consistency encoder representation. MVST integrates histopathological images, gene expression, and spatial location information into spatial transcriptomic data, thus identifying spatial domains from a more comprehensive perspective. To validate the performance of our method, we tested MVST on three real spatial transcriptomics datasets. Our experimental results show that the MVST algorithm has a better competitive performance than existing methods.

In terms of processing histopathological images, MVST is different from existing methods such as SpaGCN and stLearn [37], which take histological images as input. It uses the features of this view of histopathological images to construct the graph structure individually and obtain embeddings in the graph convolutional encoder, which are then integrated with the embeddings obtained through the gene expression information view and the spatial location information view, respectively, to obtain the final clustering embedding. Methods such as SpaGCN and stLearn are more inclined to incorporate the features of histopathology images into the gene expression information and spatial location information of the spot, which are then uniformly fed into the algorithms or models for processing. Meanwhile, the model of MVST is flexible, allowing the user to adjust the weighting of three views: gene expression information, spatial location information, and histopathological image information. Moreover, based on the experiments we have conducted, the boundaries of the spatial domains identified by MVST on the datasets we used are relatively smooth, with no fragmented spatial domains. We hypothesize that this feature of MVST is more in line with exploring the organizational structure.

Irrespective of its good performance in spatial domain recognition, MVST has some limitations. One limitation of MVST is that it mainly supports spatial domain recognition on spatial transcriptome datasets generated by the currently dominant spatial transcriptome sequencing technology, 10X Visium. Therefore, the applicability of the method has to be further improved in future work, and we expect that the design idea of the MVST method can be extended to spatial transcriptome data with higher resolution in the near future. Another limitation is that the extraction, integration, and exploitation of histopathological image features of MVST are based on simpler ideas; therefore, the processing of histopathological images needs to be optimized as much as possible to make full use of these features in the future. In addition, the measure of similarity between spots of MVST in terms of gene expression and histopathological images is based on the BallTree algorithm, an optimized algorithm for KNN. Better results may be achieved if a more biologically meaningful approach is used in the next step. We expect that research in these areas will further improve the performance of MVST. From the perspective of research in the field of spatial domain recognition, with an increase in spatial transcriptome technology and the size of spatial transcriptome data, it is necessary to improve the efficiency and scalability of spatial domain recognition tools. Moreover, at present, large-language models such as BERT have achieved good results in the field of bioinformatics, and if large-language models are applied to spatial domain recognition, they may also achieve good results and breakthroughs.

## Materials and methods

### Data description

We applied MVST to three ST datasets generated by the currently dominant spatial transcriptome sequencing technology, 10X Visium, including the human dorsolateral prefrontal cortex [38], mouse anterior cerebral section, and human breast cancer datasets. The human dorsolateral prefrontal cortex dataset contained 12 samples of dorsolateral prefrontal cortex slices provided by the three subjects. The sequenced tissue comprised six neuronal layers and white



matter, with a different number of spots for each sample. The mouse anterior brain slice dataset contains 2695 spots, and the slices were manually annotated into 52 regions. The human breast cancer dataset contains 3798 spots, with 24,923 genes sequenced, and the slices were manually annotated into 20 regions. We applied MVST to simulated datasets constructed from the 151509 dataset of human dorsolateral prefrontal cortex tissue slices. Firstly, in terms of different expression matrix sparsities, we constructed simulated data with different sparsities by randomly deleting the count values of gene expression matrices of human dorsolateral prefrontal cortex tissue section 151509 according to different ratios (0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, and 0.7), and thus investigated the impact of gene expression matrix sparsity on our spatial domain recognition tool. When the ratio is 0.7, the probability that each value in the gene expression matrix is set to 0 is 0.7, and conversely, a ratio of 0 implies that the gene expression matrix remains unchanged. It should be noted that when the set ratio is larger than 0.7, the gene expression matrix is too sparse and the dimensions of the features of the spot cannot meet the requirements of the MVST model structure, so the maximum ratio is set to 0.7. Secondly, in terms of adding noise to the dataset, we used a Poisson distribution to add different levels of noise (10%, 30%, 50%, 70% and 90%) to the gene expression matrix of human dorsolateral prefrontal cortex tissue section 151509 to construct the simulated data with different levels of noise added to it, and thus to investigate the effect of different noise levels on our spatial domain recognition tool. Specifically, we used a Poisson distribution with the mean parameter set to the product of the noise level and the mean gene expression level.

## Data preprocessing

For all datasets, the SCANPY [39] tool was used for preprocessing. First, quality control filtering was performed on the gene expression data to filter genes expressed in fewer than five cells. Second, the gene expression data were normalized and legitimized. The top 2000 highly variable genes were then filtered, and the data were scaled. Finally, the preprocessed gene expression data were downsampled using PCA, and the first 1000 principal components were used as feature inputs for the spot. For datasets in which the number of spatial domain divisions is unknown, we use the Louvain algorithm implemented in SCANPY to select the number of spatial domain divisions during the data preprocessing stage.

## Construction of graphs

MVST identifies spatial domains from three perspectives. Therefore, we construct three different undirected graphs as inputs from each of the three perspectives: spatial adjacency network graphs, gene expression similarity network graphs, and histological image similarity network graphs. The details of the three undirected graphs are described below.

**Spatial adjacency network diagram.** To obtain the spatial adjacency between spots in the spatial transcriptome data, we calculated the Euclidean distance between the coordinates of the corresponding spatial locations of any two spots, selected the top 10 nearest spots as spatial neighbors according to the size of the distances between a spot and all other spots, and constructed the adjacency matrix, denoted as  $A$ . In adjacency matrix  $A$ , if spots  $i$  and  $j$  have a spatial neighborhood, then the matrix element  $a_{ij} = a_{ji} = 1$ ; otherwise,  $a_{ij} = a_{ji} = 0$ . At the time of data input, a spatial neighborhood network graph was constructed based on the adjacency matrix  $A$ .

**Gene expression similarity network diagram.** To obtain the gene expression similarity relationship between spots in the spatial transcriptome data, the first two thousand preprocessed high-variance gene expression data points were first downsampled to 50 dimensions

using PCA, and then the Ball Tree algorithm [40] was used to obtain the top 10 nearest neighbors of each spot in terms of gene expression similarity based on the downscaled gene expression features of the spot and construct the neighbor-joining matrix, denoted as  $B$ . In neighbor-joining matrix  $B$ , the matrix element  $b_{ij} = b_{ji} = 1$  if spot  $i$  has a gene expression similarity relationship with spot  $j$  otherwise,  $b_{ij} = b_{ji} = 0$ . At the time of data input, a gene expression similarity network graph was constructed based on the neighbor-joining matrix  $B$ .

**Histological image similarity network diagram.** To obtain the histological image similarity relationship between spots in spatial transcriptome data, first, the image features at the corresponding position of each spot were extracted from the H&E-stained histopathological images; then, the image features of the spots were normalized using the RobustScaler normalization method; finally, the Ball Tree algorithm was used to obtain the top 10 nearest neighbors for each spot's top 10 nearest neighbors in terms of histological image similarity and construct the adjacency matrix, denoted as  $C$ . In the adjacency matrix  $C$ , if spot  $i$  has a histological image similarity relationship with spot  $j$ , then the matrix element  $c_{ij} = c_{ji} = 1$ ; otherwise,  $c_{ij} = c_{ji} = 0$ . At the time of data input, the histological image similarity network graph was constructed based on adjacency matrix  $C$ .

Histopathology image feature extraction specifically involves extracting the texture and intensity features of the image corresponding to the position of each spot in the histopathology image through local (adaptive) histogram equalization and computation of gray-scale covariance matrices. Among them, the image texture feature is a global feature that can reflect important texture patterns and structural information in the image. We quantitatively describe the texture feature by calculating contrast, non-similarity, homogeneity, angular second-order moment (ASM), entropy, and correlation. The RGB three-channel intensity feature of the image corresponding to the position of each spot in the histopathological image was extracted as the intensity feature of the spot.

## MVST Architecture

MVST is based on graph neural networks to learn the embedding features of three views—spatial neighborhood network graph, gene expression similar network graph, and histological image similar network graph—and to learn the view consistency embedding. Specifically, the MVST has two modules: a multiview graph convolutional encoder and a consistent embedding encoder. The first module, the multi-view graph convolutional encoder, was used to learn the graph embedding features of the spot in the three views to reduce noise and redundancy. The second module, a Consistent Embedding Encoder, captures the consistency of the geometric relationships and probability distributions among different views to learn consistent cluster embeddings for spatial domain recognition.

**Multi-view graph convolutional encoder.** In a multi-view graph convolutional encoder, the  $m$ -th view maps the spot feature and the  $m$ -th view  $G_m$  to the  $d$ -dimensional graph embedding feature  $H_m$  and the output of the  $l$ -th multi-view graph convolutional encoder is shown by the following equation:

$$H_m^{(l)} = \sigma \left( D^{-\frac{1}{2}} G_m' D^{-\frac{1}{2}} H_m^{(l-1)} W^{(l)} \right) \quad (1)$$

where  $G_m' = G_m + IN$  is the relevance coefficient matrix with added self-connection,  $IN$  is the identity matrix,  $D_{ii} = \sum_j G_m'_{ij}$ ,  $W^{(l)}$  is the trainable parameter of the  $l$ -th multi-view auto-encoder layer and  $\sigma$  denotes the activation function.

After applying  $L$  multiview encoder layers, we obtain the graph embedding  $H_m$ , which retains all the information regarding the multiview graph structure  $G$  and node feature matrix  $X$ . Then, we decode the feature  $H_m$  containing almost all the information based on different

views. In this process, we used the same number of layers as the decoders, and each decoder layer tries to invert its corresponding encoder layer. In other words, the decoding process is the inverse of the encoding process.

In order to determine the correlation between nodes and neighboring points, we use an attention mechanism that shares parameters between nodes. The learnable correlation matrix  $U$  in each view is defined by the following equation:

$$U = \varphi(G_m \odot t_s^{(l)} H_m^{(l)} W^{(l)} + G_m \odot t_r^{(l)} H_m^{(l)} W^{(l)}) \quad (2)$$

where  $t_s^{(l)}$  and  $t_r^{(l)} \in \mathbb{R}^{1 \times d_l}$  represents the trainable parameters related to their own nodes and neighbor nodes, respectively;  $\odot$  refers to the element-wise multiplication with broadcasting capability; and  $\varphi$  denotes the activation function which is generally set as the sigmoid activation function.

The normalization matrix  $U$  was passed through to obtain the final correlation coefficient  $Y$ . Thus,  $Y_{ij}$  was calculated using the following equation:

$$Y_{ij} = \frac{\exp(U_{ij})}{\sum_{k \in N_i} \exp(U_{ik})} \quad (3)$$

where  $N_i$  is the set of all nodes adjacent to node  $i$ .

**Consistency embedded encoder.** In the consistent embedding encoder, for view  $m$ , we first employ a nonlinear feature extraction mapping embedding  $H_m$ , which is mapped to the low-dimensional space  $Z_m$ . We then learn a common clustering embedding  $Z$  that integrates all  $Z_m$ ,  $Z$  can be computed using the following equation:

$$Z = \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 \quad (4)$$

where  $\beta_1, \beta_2, \beta_3$  are hyperparameters and  $\beta_1 + \beta_2 + \beta_3 = 1$ .

We use the Student's t-distribution, which measures the similarity between the integrated node representation  $z_i$  and the center of mass  $\mu_j$ . Thus, the original probability distribution  $Q$  of  $Z$  is shown by this equation:

$$q_{ij} = \frac{(1 + \|z_i - \mu_j\|^2 / \alpha)^{-\frac{\alpha+1}{2}}}{\sum_{j'} (1 + \|z_i - \mu_{j'}\|^2 / \alpha)^{-\frac{\alpha+1}{2}}} \quad (5)$$

where  $\{\mu_j\}_{j=1}^k$  is the  $k$  initial cluster centroids,  $\alpha$  is the degree of freedom of the Student's t-distribution,  $q_{ij}$  is the probability of assigning node  $i$  to cluster  $j$ .

Next we compute the target probability distribution  $P$  of  $Z$ .  $p_{ij}$  is an element of  $P$  with  $0 \leq p_{ij} \leq 1$ . As shown in the equations below, we obtain  $p_{ij}$ :

$$p_{ij} = \frac{q_{ij}^2 / f_i}{\sum_{j'} q_{ij'}^2 / f_{j'}} \quad (6)$$

where  $f_i = \sum_j q_{ij}$  are soft cluster frequencies.

**Loss function.** The loss function consists of the reconstructed multiview node feature matrix loss  $L_e$ , geometric relationship consistency loss  $L_g$ , and probability distribution

consistency loss  $L_{pd}$ , which can be calculated using the following equations:

$$L_e = \min_{\theta} \sum_{i=1}^M \|X - X_i\|_F^2 \quad (7)$$

$$L_{gr} = \min_{\gamma} \sum_{i \neq j}^M \|Z_i - Z_j\|_F^2 \quad (8)$$

$$L_{pd} = \min_{\gamma} \sum_{m=1}^M \rho_m \|Q_m - P\|_F^2 \quad (9)$$

$$L = L_e + L_{gr} + L_{pd} \quad (10)$$

where  $M$  is the number of views (3 in the experiment),  $\theta$  and  $\gamma$  are the network parameter of the multi-view map convolutional encoder,  $Q_m$  is the soft distribution of  $Z_m$ ,  $P$  is the auxiliary distribution of  $Z$ , and  $\rho$  is the trade-off parameter.

**Spatial domain identification through clustering.** First, an auxiliary distribution  $P$  is learned from the clustered embedded features  $Z$  and then, the category label of each node is predicted from the auxiliary distribution  $P$ . For node  $i$ , the index with the highest probability value in  $p_i$  is its category label. The specific calculation process equation is as follows.

$$y_i = \underset{k}{\operatorname{argmax}}(p_{ik}) \quad (11)$$

## Evaluation criteria and baseline methodology

**Evaluation criteria.** We used the ARI [41], a commonly used clustering index, as an evaluation criterion for the clustering results. The equation for ARI is as follows:

$$ARI = \frac{\sum_i \sum_j \binom{n_{ij}}{2} - \frac{\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}}{\binom{n}{2}}}{\frac{\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}}{2} - \frac{\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}}{\binom{n}{2}}} \quad (12)$$

where  $n_{ij}$  is the number of data points within the same category  $u_i$  and cluster  $v_j$ ,  $a_i$  is the number of data points in category  $u_i$ ,  $b_j$  is the number of data points in cluster  $v_j$ ,  $n$  is the total number of data points.

**Baseline methodology.** To evaluate the performance of our method in identifying spatial domains, we compared MVST with ConGI, STMGCN, SpaGCN, STAGATE, SEDR, and BayesSpace. For all compared methods, we tested them on all datasets using default parameters and default data preprocessing processes and used the same number of real clusters in the clustering.

## Author Contributions

**Conceptualization:** Hao Duan.

**Data curation:** Hao Duan.

**Formal analysis:** Hao Duan.

**Funding acquisition:** Quan Zou, Zilong Zhang.

**Methodology:** Hao Duan.

**Project administration:** Hao Duan.



**Supervision:** Qingchen Zhang, Feifei Cui, Quan Zou, Zilong Zhang.

**Writing – original draft:** Hao Duan.

**Writing – review & editing:** Qingchen Zhang, Feifei Cui, Quan Zou, Zilong Zhang.

## References

1. Moses L, Pachter L: Museum of spatial transcriptomics. *Nature Methods* 2022, 19(5):534–546. <https://doi.org/10.1038/s41592-022-01409-2> PMID: 35273392
2. Liao J, Lu X, Shao X, Zhu L, Fan X: Uncovering an Organ's Molecular Architecture at Single-Cell Resolution by Spatially Resolved Transcriptomics. *Trends in Biotechnology* 2021, 39(1):43–58. <https://doi.org/10.1016/j.tibtech.2020.05.006> PMID: 32505359
3. Zhang Z, Cui F, Su W, Dou L, Xu A, Cao C, Zou Q: webSCST: an interactive web application for single-cell RNA-sequencing data and spatial transcriptomic data integration. *Bioinformatics* 2022, 38(13):3488–3489. <https://doi.org/10.1093/bioinformatics/btac350> PMID: 35604082
4. Marx V: Method of the Year: spatially resolved transcriptomics. *Nature Methods* 2021, 18(1):9–14. <https://doi.org/10.1038/s41592-020-01033-y> PMID: 33408395
5. Larsson L, Frisén J, Lundeberg J: Spatially resolved transcriptomics adds a new dimension to genomics. *Nature Methods* 2021, 18(1):15–18. <https://doi.org/10.1038/s41592-020-01038-7> PMID: 33408402
6. Longo SK, Guo MG, Ji AL, Khavari PA: Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nature Reviews Genetics* 2021, 22(10):627–644. <https://doi.org/10.1038/s41576-021-00370-8> PMID: 34145435
7. Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, Weiss RA, Liotta LA: Laser Capture Microdissection. *Science* 1996, 274(5289):998–1001. <https://doi.org/10.1126/science.274.5289.998> PMID: 8875945
8. Boisset J-C, Vivié J, Grün D, Muraro MJ, Lyubimova A, van Oudenaarden A: Mapping the physical network of cellular interactions. *Nature Methods* 2018, 15(7):547–553. <https://doi.org/10.1038/s41592-018-0009-z> PMID: 29786092
9. Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L: Single-cell in situ RNA profiling by sequential hybridization. *Nature Methods* 2014, 11(4):360–361. <https://doi.org/10.1038/nmeth.2892> PMID: 24681720
10. Shah S, Lubeck E, Zhou W, Cai L: In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. *Neuron* 2016, 92(2):342–357. <https://doi.org/10.1016/j.neuron.2016.10.001> PMID: 27764670
11. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X: Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 2015, 348(6233).
12. Codeluppi S, Borm LE, Zeisel A, La Manno G, van Lunteren JA, Svensson CI, Linnarsson S: Spatial organization of the somatosensory cortex revealed by osmFISH. *Nature Methods* 2018, 15(11):932–935. <https://doi.org/10.1038/s41592-018-0175-z> PMID: 30377364
13. Wang X, Allen WE, Wright MA, Sylwestrak EL, Samusik N, Vesuna S, Evans K, Liu C, Ramakrishnan C, Liu J et al: Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* 2018, 361(6400). <https://doi.org/10.1126/science.aat5691> PMID: 29930089
14. Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, Welch J, Chen LM, Chen F, Macosko EZ: Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 2019, 363(6434):1463–1467. <https://doi.org/10.1126/science.aaw1219> PMID: 30923225
15. Vickovic S, Eraslan G, Salmén F, Klughammer J, Stenbeck L, Schapiro D, Åijö T, Bonneau R, Bergensträhle L, Navarro JF et al: High-definition spatial transcriptomics for in situ tissue profiling. *Nature Methods* 2019, 16(10):987–990. <https://doi.org/10.1038/s41592-019-0548-y> PMID: 31501547
16. Asp M, Bergensträhle J, Lundeberg J: Spatially Resolved Transcriptomes—Next Generation Tools for Tissue Exploration. *BioEssays* 2020, 42(10). <https://doi.org/10.1002/bies.201900221> PMID: 32363691
17. Su Y, Cui F, Yan S, Zou Q, Cao C, Zhang Z: Human-Spa: An Online Platform Based on Spatial Transcriptome Data for Diseases of Human Systems. In: *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*: 2023. IEEE: 293–298.
18. Rao A, Barkley D, França GS, Yanai I: Exploring tissue architecture using spatial transcriptomics. *Nature* 2021, 596(7871):211–220. <https://doi.org/10.1038/s41586-021-03634-9> PMID: 34381231

19. Zhang L, Chen D, Song D, Liu X, Zhang Y, Xu X, Wang X: Clinical and translational values of spatial transcriptomics. *Signal Transduction and Targeted Therapy* 2022, 7(1). <https://doi.org/10.1038/s41392-022-00960-w> PMID: 35365599
20. Lewis SM, Asselin-Labat M-L, Nguyen Q, Berthelet J, Tan X, Wimmer VC, Merino D, Rogers KL, Naik SH: Spatial omics and multiplexed imaging to explore cancer biology. *Nature Methods* 2021, 18(9):997–1012. <https://doi.org/10.1038/s41592-021-01203-6> PMID: 34341583
21. Jung N, Kim T-K: Spatial transcriptomics in neuroscience. *Experimental & Molecular Medicine* 2023, 55(10):2105–2115. <https://doi.org/10.1038/s12276-023-01093-y> PMID: 37779145
22. Arthur D, Vassilvitskii S: K-Means++: The Advantages of Careful Seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, Louisiana, USA, January 7–9, 2007: 2007.
23. Wang Y, Zhai Y, Ding Y, Zou Q: SBSM-Pro: support bio-sequence machine for proteins. *arXiv preprint arXiv:230810275* 2023.
24. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E: Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008, 2008(10).
25. Zhao E, Stone MR, Ren X, Guenthoer J, Smythe KS, Pulliam T, Williams SR, Uyttingco CR, Taylor SEB, Nghiem P et al: Spatial transcriptomics at subspot resolution with BayesSpace. *Nature Biotechnology* 2021, 39(11):1375–1384. <https://doi.org/10.1038/s41587-021-00935-2> PMID: 34083791
26. Dries R, Zhu Q, Dong R, Eng C-HL, Li H, Liu K, Fu Y, Zhao T, Sarkar A, Bao F et al: Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biology* 2021, 22(1).
27. Fu H, Xu H, Chong K, Li M, Ang KS, Lee HK, Ling J, Chen A, Shao L, Liu L et al: Unsupervised Spatially Embedded Deep Representation of Spatial Transcriptomics. *bioRxiv* 2021.
28. Hu J, Li X, Coleman K, Schroeder A, Ma N, Irwin DJ, Lee EB, Shinohara RT, Li M: SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nature Methods* 2021, 18(11):1342–1351. <https://doi.org/10.1038/s41592-021-01255-8> PMID: 34711970
29. Li J, Chen S, Pan X, Yuan Y, Shen H-B: Cell clustering for spatial transcriptomics data with graph neural networks. *Nature Computational Science* 2022, 2(6):399–408. <https://doi.org/10.1038/s43588-022-00266-5> PMID: 38177586
30. Dong K, Zhang S: Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nature Communications* 2022, 13(1). <https://doi.org/10.1038/s41467-022-29439-6> PMID: 35365632
31. Shi X, Zhu J, Long Y, Liang C: Identifying spatial domains of spatially resolved transcriptomics via multi-view graph convolutional networks. *Briefings in Bioinformatics* 2023, 24(5). <https://doi.org/10.1093/bib/bbad278> PMID: 37544658
32. Zeng Y, Yin R, Luo M, Chen J, Pan Z, Lu Y, Yu W, Yang Y: Identifying spatial domain by adapting transcriptomics with histology through contrastive learning. *Briefings in Bioinformatics* 2023, 24(2). <https://doi.org/10.1093/bib/bbad048> PMID: 36781228
33. Cutiongco MFA, Jensen BS, Reynolds PM, Gadegaard N: Predicting gene expression using morphological cell responses to nanotopography. *Nature Communications* 2020, 11(1). <https://doi.org/10.1038/s41467-020-15114-1> PMID: 32170111
34. Wang S, Rong R, Zhou Q, Yang DM, Zhang X, Zhan X, Bishop J, Chi Z, Wilhelm CJ, Zhang S et al: Deep learning of cell spatial organizations identifies clinically relevant insights in tissue images. *Nature Communications* 2023, 14(1).
35. Zeng Y, Wei Z, Yu W, Yin R, Yuan Y, Li B, Tang Z, Lu Y, Yang Y: Spatial transcriptomics prediction from histology jointly through Transformer and graph neural networks. *Briefings in Bioinformatics* 2022, 23(5). <https://doi.org/10.1093/bib/bbac297> PMID: 35849101
36. Bae S, Choi H, Lee DS: Discovery of molecular features underlying the morphological landscape by integrating spatial transcriptomic data with deep features of tissue images. *Nucleic Acids Research* 2021, 49(10):e55–e55. <https://doi.org/10.1093/nar/gkab095> PMID: 33619564
37. Pham D, Tan X, Xu J, Grice LF, Lam PY, Raghubar A, Vukovic J, Ruitenberg MJ, Nguyen Q: stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues. *bioRxiv* 2020.
38. Maynard KR, Collado-Torres L, Weber LM, Uyttingco C, Barry BK, Williams SR, Catallini JL, Tran MN, Besich Z, Tippi M et al: Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nature Neuroscience* 2021, 24(3):425–436. <https://doi.org/10.1038/s41593-020-00787-0> PMID: 33558695
39. Wolf FA, Angerer P, Theis FJ: SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* 2018, 19(1). <https://doi.org/10.1186/s13059-017-1382-0> PMID: 29409532

40. Dolatshah M, Hadian A, Minaei-Bidgoli B: Ball\*-tree: Efficient spatial indexing for constrained nearest-neighbor search in metric spaces. *Computer Science* 2015.
41. Rand WM: Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association* 1971, 66(336):846–850.