ACCSCIENCE PUBLISHING

Tumor Discovery

ORIGINAL RESEARCH ARTICLE

Artificial intelligence-enabled spatially resolved transcriptomics reveal spatial tissue organization of multiple tumors

Teng Liu[1,2†], Jinxin Ye[3†], Chunnan Hu[1†], Zongbo Zhang[1], Zhuomiao Ye[2], Jiangnan Liao[2], and Mingzhu Yin[1,2*]

[1]Department of Clinical Research Center (CRC), Clinical Pathology Center (CPC), Cancer Early Detection and Treatment Center (CEDTC) and Translational Medicine Research Center (TMRC), Chongqing University Three Gorges Hospital, Chongqing University, Wanzhou, Chongqing, China

[2]Department of Chongqing Technical Innovation Center for Quality Evaluation and Identification of Authentic Medicinal Herbs, Chongqing University Three Gorges Hospital, Chongqing University, Wanzhou, Chongqing, China

[3]Physics and Technique Department of Radiation Oncology, Cancer treatment institute, Chongqing University Three Gorges Hospital, Chongqing University, Wanzhou, Chongqing, China

Abstract

Spatially resolved transcriptomics was honored as the Method of the Year 2020 by Nature Methods. This approach allows biologists to precisely discern mRNA expression at the cellular level within structurally preserved tissues. Leveraging artificial intelligence in spatial transcriptomic analysis enhances the understanding of cellular-level biological interactions and offers novel insights into intricate tissues, such as tumor microenvironments. Nevertheless, numerous existing clustering algorithms employing deep learning exhibit the potential for enhancement. In this paper, we focus on graph deep learning-based spatial domain identification for spatial transcriptomics (ST) data from multiple tumors. This identification enables the recognition of cell subpopulations in distinct spatial coordinates, aiding further studies on tumor progression, such as cell-cell communication, pseudo-time trajectory inference, and single-cell deconvolution. Initially, the gene expression profiles and spatial location information were transformed into a gene feature matrix and a cell adjacency matrix. A variational graph autoencoder was then applied to extract features and reduce the dimensions of these two matrices. Following training in the constructed graph neural networks, the latent embeddings of ST data were generated and could be leveraged for spatial domain identification. Through a comparison with established methods, our approach demonstrated superior clustering accuracy. The utilization of accurately segmented spatial regions enables downstream analyses of multiple tumors, encompassing the trajectory of tumor evolution, and facilitating differential gene expression analysis across various cell types.

Keywords: Spatial transcriptomics; Artificial intelligence; Graph neural network; Spatial domain identification; Tumor progression

†*These authors contributed equally to this work.*

***Corresponding author:**
Mingzhu Yin
(yinminzhu@cqu.edu.cn)

Citation: Liu T, Ye J, Hu C, *et al*. Artificial intelligence-enabled spatially resolved transcriptomics reveal spatial tissue organization of multiple tumors. *Tumor Discov*. 2024;3(1):2049.
https://doi.org/10.36922/td.2049

Received: October 16, 2023
Accepted: December 15, 2023
Published Online: March 6, 2024

Copyright: © 2024 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Spatiotemporal molecular medicine involves comprehending medicine across various dimensions, layers, perspectives, and dynamics through the integration of spatialization and temporalization of clinical phonemes with spatiotemporal molecular omics.[1] This emerging discipline strives to elucidate the pathogenesis, epidemiology, historical context, patient symptoms and signs, clinical measurements, and therapeutic approaches. Spatially resolved transcriptomics plays a crucial role in connecting and correlating information between histological sections and molecular profiles, such as cell-cell interactions, transcription factor distribution, spatial positioning, and mRNA expression within the cell through artificial intelligence, computerized programming, and visualization techniques.[2] Spatial transcriptomics (ST) technologies document transcriptomic expression data and their respective two-dimensional tissue locations at spatial resolution.[3] In the past decade, various ST technologies have been introduced to capture mRNA molecules and their locations, each differing in resolution, sensitivity, sequencing depth, and scale. Some notable technologies include seqFISH+,[4] MERFISH,[5] omsFISH,[6] 10× Visium,[7] Slide-seqV2,[8] 10× Xenium,[9] and Stereo-seq.[10]

Partitioning spatial domains in a tumor's spatial transcriptome data can shed light on the composition of the tumor microenvironment and the localized development of tumor cells. The primary goal of spatial clustering is to segment tumor tissues into distinct cell subpopulations, facilitating downstream analyses of differential gene expression, biological function grouping, and cell-cell communication.[11] Conventional methods classify cell types based on gene expression profiles using algorithms such as K-means and Louvain.[12] Giotto[13] identifies spatial domains through the construction of a hidden Markov random field model, leveraging both the feature matrix and spatial coordinates. stLearn[14] utilizes a standard Louvain clustering workflow to handle the cell adjacency matrix, subsequently using spatial positions to identify subgroups within broader clusters. BayesSpace[15] employs a fully Bayesian statistical model to enhance resolution and conduct clustering analysis, promoting the categorization of nearby spots into the same group based on a predefined spatial prior. Despite the potential of spatial location information to enhance clustering accuracy, the aforementioned spatial clustering algorithms have yet to achieve optimal performance.[16]

Graph deep learning (GDL) has emerged as a promising approach for integrating gene expression profiles and spatial location information to address spatial clustering tasks.[17] Numerous studies have proposed graph neural network (GNN) architectures aimed at reducing the dimensionality of ST data and identifying spatial regions. Notably, SpaGCN[18] presents a comprehensive framework that partitions data based on the aggregation of gene expression, histology, and spatial location. In the cell clustering for ST data technique, a sequence of graph convolutional networks is embedded within the deep graph infomax module to derive embeddings for cell nodes.[19] STAGATE[20] devises a graph attention autoencoder framework for spatial domain identification. This approach constructs a cell-type-aware spatial neighbor network to characterize the spatial similarity at domain boundaries. Similarly, GraphST[21] utilizes a graph self-supervised contrastive learning mechanism for the unified analysis of ST data, encompassing spatial clustering, multisample integration, and cell-type deconvolution. The contrastive learning framework within the spatial clustering module enables the acquisition of informative and discriminative spot representations by minimizing the embedding distance between spatially adjacent spots. All the algorithms mentioned above can be improved in terms of clustering performance.

Motivated by the successful application of diverse artificial intelligence algorithms in analyzing ST datasets, this paper introduces a variational graph autoencoder (VGAE) enriched with multiple layers of GNNs for identifying spatial domains in tumor tissues (Figure 1). Initially, gene expression profiles and spatial location information are transformed into a gene feature matrix and a cell adjacency matrix, respectively. The VGAE consists of two primary components: an encoder and a decoder. The encoder employs GNNs to extract matrix features, reduce the feature dimensions, and generate latent embeddings. In contrast, the decoder reconstructs two matrices based on these latent embeddings. Clustering methods were, then, applied to latent embeddings to detect spatial domains within the tumor tissue. Various loss functions are utilized to enhance the quality and performance of the latent embeddings by minimizing the overall errors, including the reconstruction error and Kullback-Leibler (KL) divergence error. The presented spatial clustering method is compared with existing approaches to demonstrate its effectiveness. Additional details are provided in the "Methods" section. In addition, we employ this clustering framework on different tumor tissues to investigate their respective tumor microenvironments.

2. Methods

2.1. VGAE

The VGAE is a machine learning model that is primarily utilized in the domains of graph representation learning

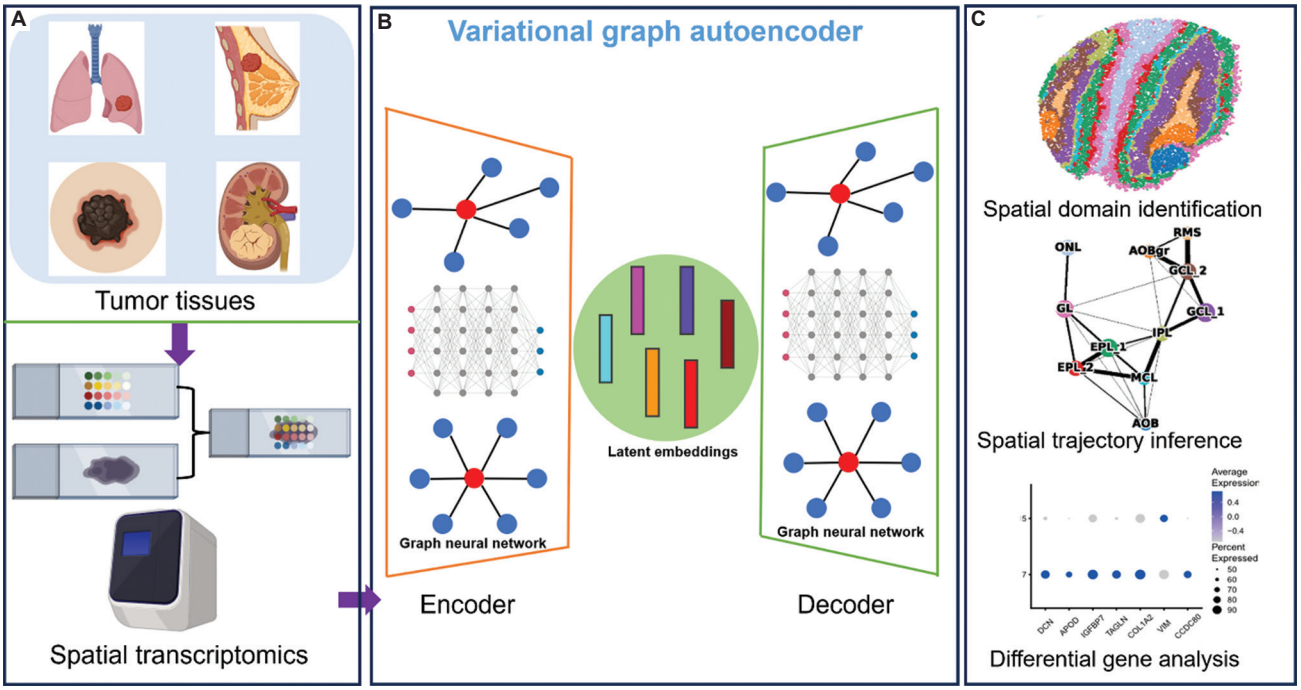


Figure 1. The proposed variational graph autoencoder-based spatial clustering framework. (A) Diverse spatial transcriptomics (ST) technologies are employed to generate distinct ST datasets of tumor tissues; (B) the ST data can be effectively processed using a graph deep learning framework to obtain latent embeddings that encapsulate meaningful and informative features; (C) these learned latent embeddings find application in various downstream analysis tasks, including spatial clustering, spatial trajectory analysis, and identification of differential genes.

and generative modeling. It amalgamates principles from variational autoencoders and GNNs, facilitating efficient acquisition and generation of representations for data organized in graph structures. The primary goal of VGAE lies in acquiring low-dimensional, continuous latent embeddings for spatial spots within the ST data. These latent embeddings can encapsulate meaningful and informative features, preserving both the structural and functional attributes inherent in spatial transcriptome datasets. In the encoding process, the initial GNN layer (Equation I) determines a lower-dimensional feature matrix, denoted as $\tilde{X}$, based on the given feature matrix X and the adjacency matrix A:[22]

$$\tilde{X} = GNN(X, A) = \mathrm{Re}LU(\tilde{A}XW_0)$$

$$\tilde{A} = D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \quad \text{(I)}$$

Where $\tilde{A}$ represents the symmetrically normalized adjacency matrix, signifies the rectified linear activation function, and W_0 stands for the weight parameter for this layer of the GNN. Subsequently, the second layer (Equation II) produces the mean and variance vectors of the feature matrix utilizing the weight parameters W_1:

$$\mu = GNN_{\mu}(X, A) = \tilde{A}\tilde{X}W_1$$

$$\log\sigma^2 = GNN_{\sigma}(X, A) = \tilde{A}\tilde{X}W_1 \quad \text{(II)}$$

The mean and variance vectors share the same weight parameters. Finally, the latent embeddings are determined by the reparameterization trick[22] according to these two values (Equation III):

$$Z = \mu + \sigma\varepsilon \quad \text{(III)}$$

where ε belongs to the standard normal distribution, ε ∈ Norm (0,1). Within the VGAE module, the incoming and predictive graphs are denoted by probability distributions, specifically q(Z|X, A) and p(A|Z). Since the decoder involves an inner product operation, the reconstruction of the adjacency matrix is formulated as an inner product (Equation IV):

$$p(A \mid Z) = \sigma(Z \cdot Z^T) \quad \text{(IV)}$$

The loss function encompasses two types of errors: reconstruction loss and a regularization term. The reconstruction loss encourages the decoded output to closely resemble the original input, while the regularization term promotes the learned latent distribution to align with the prior. The formulation of this loss function is as follows (Equation V):

$$L = L_{\text{recon}} + L_{KL}$$

$$L = \mathrm{E}_{q(Z|\mathrm{X},\mathrm{A})}[\log p(A \mid Z)] - KL(q(Z \mid X, A) \,||\, p(Z)) \qquad \text{(V)}$$

The term KL(.) denotes the Kullback-Leibler divergence between two probability distributions. Training VGAE to minimize this objective function enables the model to learn a probabilistic mapping of spots to a latent space, facilitating meaningful and informative representations of the structure and features of the ST data. The downstream clustering methods partition the latent embeddings to detect spatial domains. Subsequently, the resulting clustering labels are compared with the ground-truth to assess accuracy and performance.

2.2. GNN

In the VGAE model, the encoder utilizes two layers of GNNs to extract features and reduce dimensions. GNNs facilitate the processing of ST data by enabling spots to learn from and communicate with neighboring spots. Each spot aggregates information from its neighbors, subsequently updating its own representation based on this aggregated data. The Pytorch_pyG package[23] in Python offers multiple implementations of GNNs. Existing spatial clustering architectures typically incorporate only a single GNN. In this study, we opt for the simple graph convolution (SGC) ConvNet[24] to construct the encoder.

The SGC ConvNet highlights issues of model complexity and redundant computations within GDL. To address these defects, SGC ConvNet aims to minimize collapsing weight matrices and nonlinearities between successive layers. This streamlined linear model demonstrated comparable or even superior performance at both theoretical and experimental levels. Notably, the convolution kernel in SGC is redefined as a linear function (Equation VI):

$$\check{Y}_{SGC} = soft\max\left(S...SSX\Theta^{(1)}\Theta^{(2)}...\Theta^{(K)}\right) = soft\max\left(S^K X\Theta\right) \qquad \text{(VI)}$$

where S is the normalized adjacent matrix, X is the feature matrix, Θ is the weight matrix, and softmax indicates the normalized exponential function.

2.3. ST datasets

Various types of ST data from tumors were utilized to evaluate the proposed spatial clustering architecture. These datasets were generated using diverse ST technologies, resulting in variations in resolution, spot counts, and gene profiles. Specifically, the human dorsolateral prefrontal cortex (DLPFC) ST data was obtained from the 10× Visium platform, and the spatialLIBD project conducted an extensive study utilizing this data, encompassing spot-level information, layer-level data, and spatial marker genes.[25] The project comprises a total of 12 samples, each dissection covering six neuronal layers plus white matter. Consequently, eight samples were categorized into seven clusters, while the remaining four were grouped into five (ground-truth, cortical layers one to six, white matter). For validation of the GNN algorithm, sample 151673 was chosen as a representative due to its specificity. This sample entails 3,639 spots and 33,538 genes, with provided spot annotations.

The second ST dataset originates from human breast cancer tissue and is available through the 10× Visium dataset repository. This dataset holds significant value for the analysis of heterogeneous tumor and immune microenvironments, given its substantial intratumoral and intertumoral variations. To facilitate clustering estimation, the sample is divided into 20 regions using the SEDR[26] package, relying on pathological features and gene expression. These annotated regions provide the foundation for clustering evaluation. In total, this dataset encompasses 3798 spots and 36,601 genes.

The 10× Xenium technology represents a novel approach integrating single-cell, spatial, and *in situ* analysis of FFPE tissue. Notably, breast cancer tumor datasets associated with this technology were reprocessed and republished on December 6, 2022.[9] Leveraging its non-destructive workflow, Xenium spatially aligns RNA, protein, and histological data within a unified image. This feature empowers us to discern cell types and their corresponding gene expression profiles at a single-cell resolution. In the breast cancer tumor dataset, a remarkable seventeen distinct cell types have been identified, amounting to 164,079 cells and utilizing a 313-plex gene panel. To alleviate computational load, this manuscript employs a segmented version of this data for clustering comparison, comprising 15 cell types, 11,996 cells, and the 313 gene panel.

2.4. Data pre-processing and hyperparameters

In this article, data pre-processing and VGAE training are conducted within a Python virtual environment using PyTorch_pyG, Squidpy, and Scanpy toolkits. Initially, gene expression profiles undergo normalization and log transformation using Scanpy.[27] Users also have the option to select "SCTransform" for gene expression normalization. Three thousand highly variable genes are selected to construct the feature matrix. Subsequently, the scikit-learn toolkit[28] employs a nearest-neighbor search technique to calculate the adjacency matrix. The neighbors for each spot are determined using either the k-nearest neighbor or radius-nearest neighbor modes. Specifically, for the 10×

Visium technology, the radius-nearest neighbor mode is used to ensure each spot is in proximity to approximately five neighbors. Conversely, the k-nearest neighbor mode is employed to guarantee that each spot has precisely five neighbors for the 10× Xenium method.

On generating the feature and adjacency matrices, we utilize the VGAE module embedded with SGC ConvNets to learn latent embeddings through model training. This step is implemented in Python using PyTorch_pyG. The relevant hyperparameters are defined as follows: input channels are 3000 (representing the number of highly variable genes), hidden channels are 128, and output channels are 128. The learning rate is set to 1e-6, the number of epochs is 5000, the weight decay factor is 1e-4, gradient clipping is set at five, and the random seed is fixed at zero for all experiments. The model architecture comprises four hidden layers, and the activation function used outside each layer is the exponential linear unit function.

After generating the latent representations in the embedding space, we employ the K-means method as the downstream clustering approach for identifying spatial domains. The clustering process is facilitated using the scikit-learn package. The number of clusters in the K-means corresponds to the ground truth. The hardware utilized in this study includes an Intel (R) Core (TM) i9-12900F CPU at 2.40 GHz, 64 GB of memory, and a GeForce RTX 3090Ti GPU. To effectively run the GNN-based spatial clustering algorithms, the maximum number of spots should be below 20,000. Consequently, for the ST data from the 10× Xenium technology, we utilize a cropped subset of the data, ensuring it remains below this threshold. Given that this subset encompasses all genes and groups, it does not compromise the spatial clustering experiments.

2.5. Spatial clustering metrics

On acquiring the prediction labels using K-means, we employ the adjusted Rand index (ARI) to assess the similarity between these predicted labels and the ground truth. The ARI is a commonly utilized metric for evaluating clustering algorithms. It is calculated using the scikit-learn toolkit by comparing the two vectors. Assuming that $P = \{P_1, P_2, \ldots P_M\}$ and $G = \{G_1, G_2, \ldots G_M\}$ represent the predicted and ground-truth label sets, the ARI is defined as follows (Equation VII):

$$ARI = \frac{\sum_{i,j}\binom{n_{ij}}{2} - \left[\sum_i \binom{n_{i.}}{2}\sum_j \binom{n_{.j}}{2}\right] / \binom{n}{2}}{\frac{1}{2}\left[\sum_i \binom{n_{i.}}{2} + \sum_j \binom{n_{.j}}{2}\right] - \left[\sum_i \binom{n_{i.}}{2}\sum_j \binom{n_{.j}}{2}\right] / \binom{n}{2}} \quad \text{(VII)}$$

Where c is the number of cell types, n_i and n_j denote the number of spots belonging to P_i *and* G_j, and n_{ij} implies the number of spots located in P_i and G_j. A higher ARI indicates a greater similarity between the two groups, with all ARI values ranging from zero to one.

If the ground truth for the ST data is not available, the Silhouette Coefficient (SC) score and Davies-Bouldin (DB) score[29] are employed to assess the clustering performance. The SC score is computed based on the mean intra-cluster distance and mean nearest-cluster distance for the predicted labels, ranging from minus one to one. It signifies the dispersion level between clusters, and a higher SC score indicates better clustering accuracy. The DB score represents the average similarity measure of each cluster with the most similar cluster, ranging from zero to positive infinity. A lower DB index is preferred. These metrics are calculated using the scikit-learn package.

3. Results and discussion

3.1. Evaluating the clustering accuracy on human dorsolateral pre-frontal cortex data

We named our proposed spatial clustering method VGAE_SGC. To demonstrate its accuracy and effectiveness, we compared this method with six benchmark approaches: BayesSpace,[15] SpaGCN,[18] STAGATE,[20] SEDR,[26] Scanpy,[27] and DeepST.[30] The evaluation was performed on the ST data using sample 151673 from the human DLPFC dataset, and we calculated and compared the average ARI values across 12 samples. In the ground truth of sample 151,673, there were seven marked groups (six cortical layers and one white matter).

The VGAE_SGC approach exhibited the highest average ARI of 0.542 (Figure 2A). In addition, STAGATE and DeepST achieved ARI values exceeding 0.50. The ground truth of sample 151673 is illustrated in Figure 2B, with the spatial domain identification results closely aligned with this ground truth (ARI = 0.5253, Figure 2C). The spatial clustering results of the six compared methods are presented in Figure 2D, displaying ARI values lower than 0.50 for this sample. Through this comparison with the ground truth, we established the clustering accuracy of our proposed method. Subsequently, we validated this method using tumor ST datasets to assess its spatial clustering capabilities.

3.2. Deciphering multiple regions of human breast cancer from low-resolution technology

The analysis involved seven distinct clustering methods applied to the human breast cancer dataset obtained from 10× Visium. This ST dataset, manually annotated using the SEDR package, comprises 20 regions encompassing four

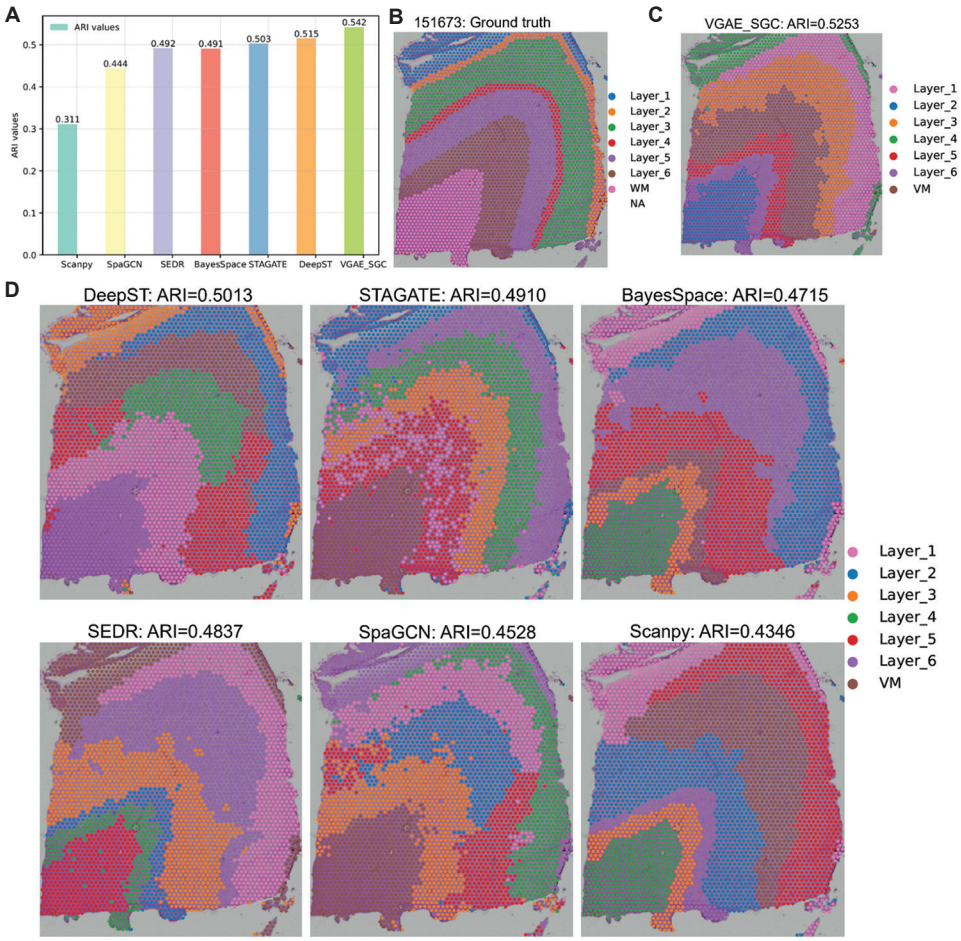


Figure 2. The spatial clustering results of seven compared methods on human DLPFC data; (A) the average adjusted Rand index (ARI) values in seven methods for twelve samples; (B) the ground truth of sample 15,1673 of this spatial transcriptomics (ST) data; (C) the spatial domain identification results of the proposed method VGAE_SGC; and (D) the spatial clustering results of six benchmarking methods for comparison.
Abbreviation: DLPFC: Dorsolateral pre-frontal cortex.

morphotypes: DCIS/LCIS, IDC, tumor edge, and healthy areas (Figure 3A). All seven approaches yielded ARI values exceeding 0.50, with VGAE_SGC achieving the highest ARI of 0.603 (Figure 3B). Figure 3C illustrates the spatial clustering outcomes of the top three methods. Notably, in the case of VGAE_SGC, several clusters (0, 10, 1, 13, 7, and 14) were closely aligned with the ground truth. These findings underscore the efficacy of VGAE_SGC in identifying cell subpopulations and effectively detecting the tumor microenvironment.

We performed a comprehensive downstream analysis of the spatial clustering results obtained from VGAE_SGC. Specifically, the IDC_5 region was divided into two distinct clusters, denoted as clusters 2 and 11, within VGAE_SGC. We, further, scrutinized the rationale behind the segmentation of the cell types. The differential gene expression profiles for clusters 2 and 11 are depicted in Figure 3D, revealing *COX6C* for cluster 2, and *ADIRF* and *HLA-B* for cluster 11. This analysis suggests disparate gene profiles for the two groups, which are indicative of distinct cell types. To elucidate the biological functions associated with these clusters, we conducted a gene set enrichment analysis of these differentially expressed genes. The enriched pathways for these clusters predominantly pertained to antigen processing and presentation (Figure 3E).

3.3. Partitioning tumor regions for the single-cell-resolution human breast cancer data

Finally, we validated our proposed spatial clustering framework, VGAE_SGC, using an alternative single-cell-resolution ST dataset. This dataset was produced utilizing the advanced 10x Xenium technology and pertains to breast cancer tissue, as shown in Figure 4A. The initial ST dataset boasts a high resolution, comprising 164,079 spots and 313 genes. To alleviate the computational demands, we segmented this primary dataset into a cropped subset,

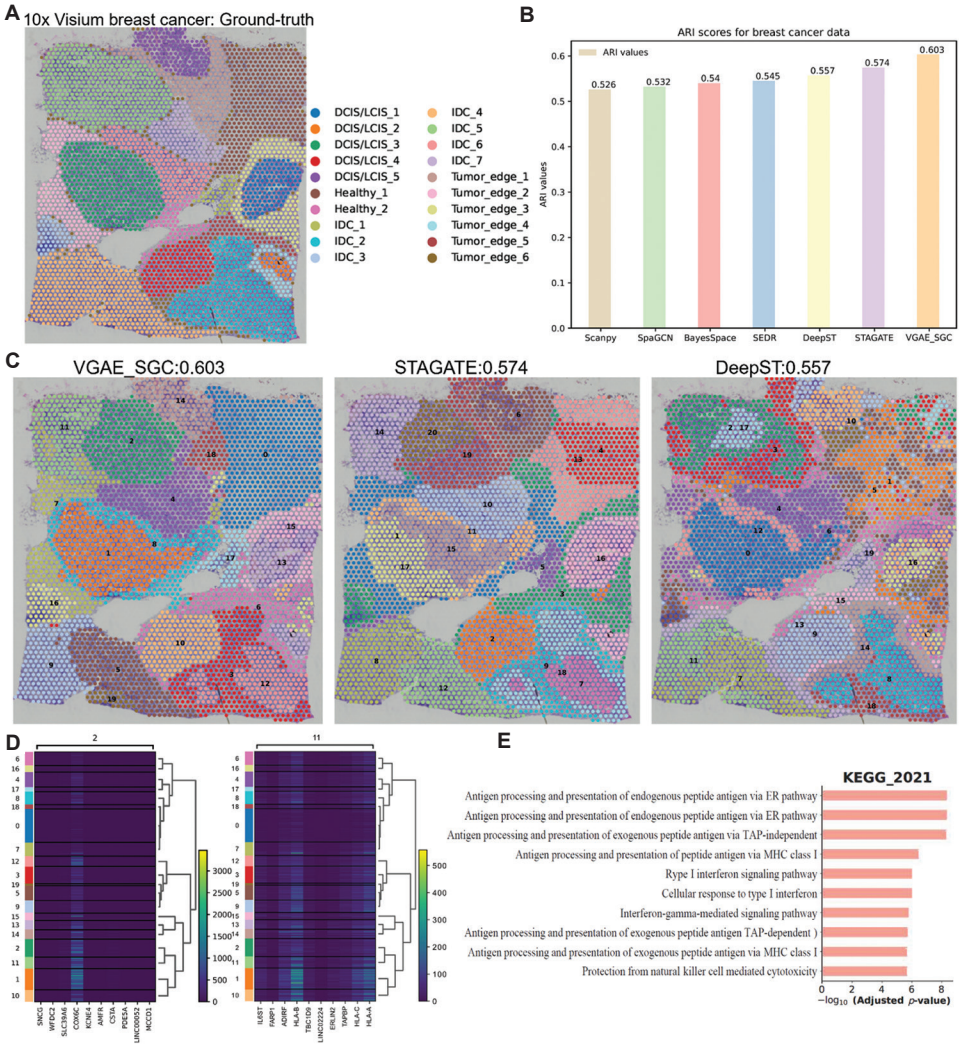


Figure 3. The spatial clustering and downstream analysis results on human breast cancer from 10× Visium platform: (A) The ground truth of breast cancer data labeled by SEDR package; (B) the adjusted Rand index (ARI) values of seven compared methods on this spatial transcriptomics (ST) data; (C) the spatial domain identification results of DeepST, STAGATE, and VGAE_SGC on the breast cancer data; (D) the differential gene analysis between group 2 or 11 and other clusters; and (E) the gene set enrichment analysis of differential genes in groups 2 and 11.

specifically encompassing 15 cell types with 11,996 cells and the same set of 313 genes (Figure 4A).

We conducted a comparative analysis of the clustering accuracy across the seven methods using the aforementioned cropped ST data. Notably, VGAE_SGC and STAGATE exhibit proximity in performance, outperforming the remaining methods, as shown in Figure 4B. Figure 4C portrays the spatial distribution of spots in the VGAE_SGC case, revealing that clusters 0, 1, 6, and 9 aligned with the annotated labels, as shown in Figure 4A. In addition to conducting differential gene and biological function analyses, we employed VGAE_SGC's clustering groups to infer the spatial development trajectory (Figure 4D). Initially, we present a uniform manifold approximation and projection (UMAP) plot illustrating VGAE_SGC's clustering results. Subsequently, we utilized the *scanpy.tl.paga* function within Scanpy to construct the spatial evolutionary trajectory, highlighting that group 0 represents regular cells and cluster 5 corresponds to tumor cells in breast cancer.

VGAE_SGC excelled in discerning cellular compartmentalization and diverse cellular subpopulations within spatial transcriptomic data. It also demonstrates proficiency in detecting the tumor microenvironment, which is crucial for uncovering neo-cells, elucidating carcinogenesis mechanisms, and advancing cancer treatments. However, it is essential to acknowledge the limitations of this approach. Notably, we did not leverage

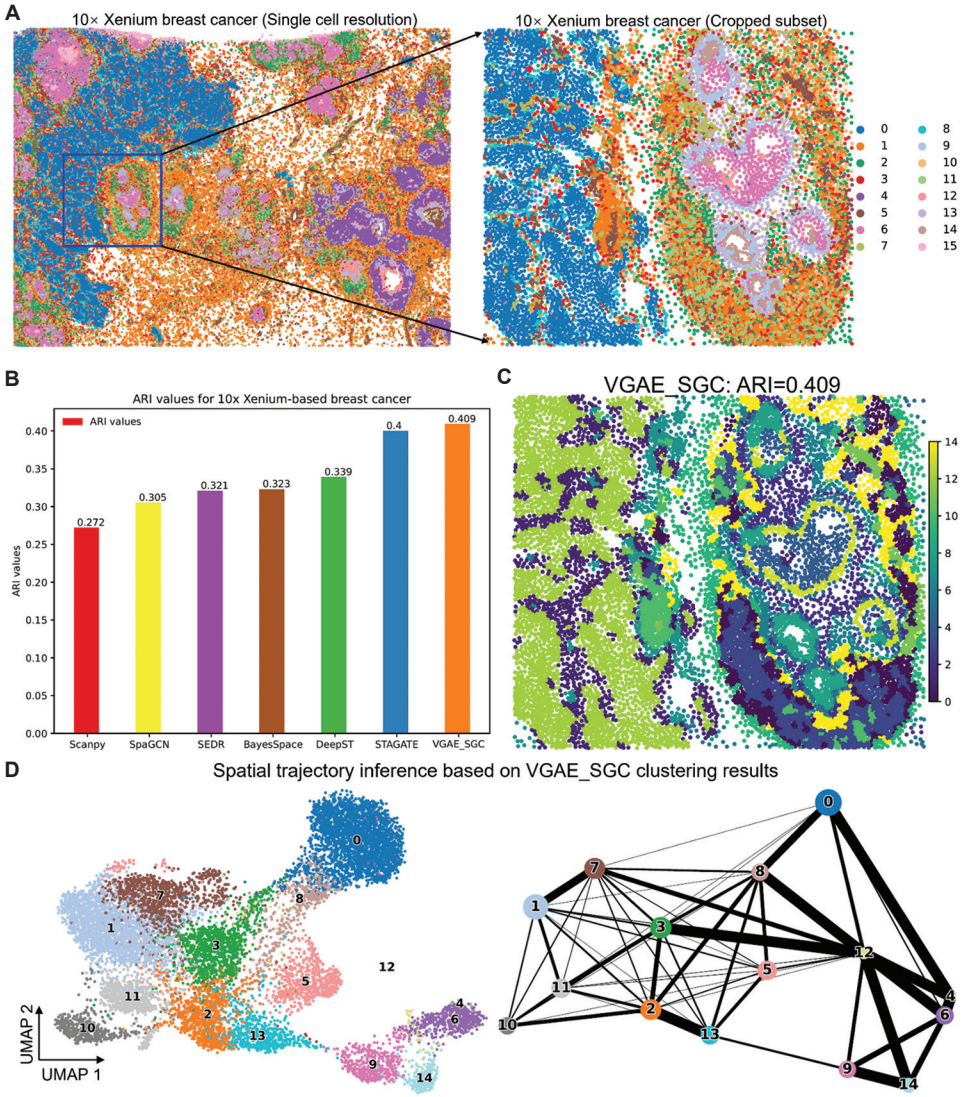


Figure 4. The proposed spatial clustering method, VGAE_SGC, undergoes validation using single-cell resolution breast cancer spatial transcriptomics (ST) data: (A) The primary breast cancer dataset is obtained through the 10× Xenium technology. To enhance computational efficiency, a cropped subset of this ST data is utilized in this study; (B) the adjusted Rand index (ARI) values for seven compared approaches are computed using the cropped ST data; (C) the spatial distribution of spots in the VGAE_SGC method yields an ARI value of 0.409; and (D) the reduced-dimension graph illustrates VGAE_SGC's spatial clustering outcomes and its associated inferred spatial evolutionary trajectory.

extensive single-cell data, which is a significant factor in optimizing the performance of our algorithm.

4. Conclusion

This article explores the utilization of artificial intelligence algorithms for spatial clustering tasks in spatially resolved transcriptomics. Specifically, we introduced a VGAE enhanced with SGC ConvNets to delineate spatial domains in ST data from various tumor tissues. The clustering accuracy of our architecture is demonstrated through comparisons with several existing clustering methodologies. We validated our proposed method, VGAE_SGC, using breast cancer datasets with varying resolutions. In addition, we delineated various downstream analysis tasks predicted for spatial clustering outcomes, including differential gene expression analysis, identification of biological functions, and inference of spatial developmental trajectories. Unsupervised learning modules face challenges in accurately identifying spatial regions within an organization. However, future research can focus on integrating methods to enhance the clustering results. Approaches such as contrastive and

semi-supervised methods, generative adversarial neural networks, and graph-based deep learning have shown promise for improving the accuracy and effectiveness of spatial clustering in organizational analysis.

Acknowledgments

We thank BioRender (biorender.com) for providing technical support in drawing the figures.

Funding

This work was supported in part by the National Key R&D Programs (NKPs) of China (Grant No. 2022YFC3601800) [M.Y.]; the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJZD-K202300105) [T.L.]; Chongqing medical scientific research project (Joint project of Chongqing Health Commission and Science and Technology Bureau) (Grant No. 2024QNXM004) [T.L.].

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Mingzhu Yin
Formal analysis: Jiangnan Liao
Investigation: Jinxin Ye, Chunlan Hu
Methodology: Teng Liu, Zongbo Zhang
Writing – original draft: Teng Liu, Zhuomiao Ye
Writing – review & editing: Jinxin Ye, Mingzhu Yin

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The data can be downloaded from the provided links. Human DLPFC: the primary source: https://www.nature.com/articles/s41593-020-00787-0; the pre-processed source: https://github.com/LieberInstitute/spatialLIBD); Human breast cancer: the primary source: https://www.10xgenomics.com/resources/datasets/human-breast-cancer-block-a-section-1-1-standard-1-1-0; the pre-processed source: https://github.com/JinmiaoChenLab/SEDR_analyses/); and Human breast cancer from 10x Xenium: the primary source: https://www.10xgenomics.com/products/xenium-in-situ/preview-dataset-human-breast; https://squidpy.readthedocs.io/en/stable/notebooks/tutorials/tutorial_xenium.html

References

1. Wang X, Fan J. Spatiotemporal molecular medicine: A new era of clinical and translational medicine. *Clin Transl Med*. 2021;11(1):e294.

   doi: 10.1002/ctm2.294

2. Wang X. Clinical trans-omics: An integration of clinical phenomes with molecular multiomics. *Cell Biol. Toxicol.* 2018;34:163-166.

   doi: 10.1007/s10565-018-9431-3

3. Rao A, Barkley D, França GS, *et al*. Exploring tissue architecture using spatial transcriptomics. *Nature*. 2021;596:211-220.

   doi: 10.1038/s41586-021-03634-9

4. Eng CHL, Lawson M, Zhu Q, *et al*. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature*. 2019;568:235-239.

   doi: 10.1038/s41586-019-1049-y

5. Chen KH, Boettiger AN, Moffitt JR, *et al*. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*. 2015;348:aaa6090.

   doi: 10.1016/B978-0-12-822824-1.00002-X

6. Codeluppi S, Borm LE, Zeisel A, *et al*. Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat Methods*. 2018;15:932-935.

   doi: 10.1038/s41592-018-0175-z

7. Ghorbani S, Jelinek E, Jain R, *et al*. Versican promotes T helper 17 cytotoxic inflammation and impedes oligodendrocyte precursor cell remyelination. *Nat Commun*. 2022;13:2445.

   doi: 10.1038/s41467-022-30032-0

8. Stickels RR, Murray E, Kumar P, *et al*. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat Biotechnol*. 2021;39:313-319.

   doi: 10.1038/s41587-020-0739-1

9. Janesick A, Shelansky R, Gottscho AD, *et al*. High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat Commun*. 2023;14:8353.

   doi: 10.1038/s41467-023-43458-x

10. Chen A, Liao S, Cheng M, *et al*. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*. 2022;185:1777-1792.e21.

    doi: 10.1016/j.cell.2022.04.003

11. Williams CG, Lee HJ, Asatsuma T, *et al*. An introduction to spatial transcriptomics for biomedical research. *Genome Med*. 2022;14:68.

    doi: 10.1186/s13073-022-01075-1

12. Bu Z, Li HJ, Zhang C, *et al*. Graph K-means based on leader identification, dynamic game, and opinion dynamics. *IEEE Trans Knowl Data Eng*. 2020;32:1348-1361.

doi: 10.1109/TKDE.2019.2903712

13. Dries R, Zhu Q, Dong R, *et al*. Giotto: A toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol*. 2021;22:78.

doi: 10.1186/s13059-021-02286-2

14. Pham D, Tan X, Balderson B, *et al*. Robust mapping of spatiotemporal trajectories and cell–cell interactions in healthy and diseased tissues. *Nat Commun*. 2023;14:7739.

doi: 10.1038/s41467-023-43120-6

15. Zhao E, Stone MR, Ren X, *et al*. Spatial transcriptomics at subspot resolution with BayesSpace. *Nat Biotechnol*. 2021;39:1375-1384.

doi: 10.1038/s41587-021-00935-2

16. Liu Y, Wang T, Duggan B, Sharpnack M, Huang K, Zhang J, *et al*. SPCS: A spatial and pattern combined smoothing method for spatial transcriptomic expression. *Brief Bioinform*. 2022;23:bbac116.

doi: 10.1093/bib/bbac116

17. Wu Z, Pan S, Chen F, *et al*. A comprehensive survey on graph neural networks. *IEEE Trans Neural Netw Learn Syst*. 2021;32:4-24.

doi: 10.1109/TNNLS.2020.2978386

18. Hu J, Li X, Coleman K, *et al*. 2021. SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat Methods*. 18:1342-1351.

doi: 10.1038/s41592-021-01255-8

19. Li J, Chen S, Pan X, *et al*. Cell clustering for spatial transcriptomics data with graph neural networks. *Nat Comput Sci*. 2022;2:399-408.

doi: 10.1038/s43588-022-00266-5

20. Dong K, Zhang S. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nat Commun*. 2022;13:1739.

doi: 10.1038/s41467-022-29439-6

21. Long Y, Ang KS, Li M, *et al*. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat Commun*. 2023;14:1155.

doi: 10.1038/s41467-023-36796-3

22. Thomas NK, Welling, M. Variational Graph Auto-Encoders. [arXiv Preprint] arXiv. 2016;2:1611.07308.

doi: 10.48550/arXiv.1611.07308

23. Fey M, Lenssen JE. Fast graph representation learning with pytorch geometric. [arXiv Preprint] arXiv. 2019;1903.02428.

doi: 10.48550/arXiv.1903.02428

24. Wu F, Souza A, Zhang T, *et al*. Simplifying Graph Convolutional Networks. In: *International Conference on Machine Learning*. Vol. 97. New York: PMLR; 2019. p. 6861-6871.

doi: 10.48550/arXiv.1902.07153

25. Pardo B, Spangler A, Weber LM, *et al*. SpatialLIBD: An R/Bioconductor package to visualize spatially-resolved transcriptomics data. *BMC Genomics*. 2022;23(1):1-5.

doi: 10.1186/s12864-022-08601-w

26. Xu H, Fu H, Long Y, *et al*. Unsupervised spatially embedded deep representation of spatial transcriptomics. *Genome Med*. 2024;16:12.

doi: 10.1186/s13073-024-01283-x

27. Wolf FA, Angerer P, Theis FJ. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol*. 2018;19:15.

doi: 10.1186/s13059-017-1382-0

28. Pedregosa F, Varoquaux G, Gramfort A, *et al*. Scikit-learn: Machine learning in python. *J Mach Learn Res*. 2011;12:2825-2830.

29. Addagarla SK, Amalanathan A. Probabilistic unsupervised machine learning approach for a similar image recommender system for e-commerce. *Symmetry*. 2020;12:1783.

doi: 10.3390/sym12111783

30. Xu C, Jin X, Wei S, *et al*. DeepST: Identifying spatial domains in spatial transcriptomics by deep learning. *Nucleic Acids Res*. 2022;50:e131.

doi: 10.1093/nar/gkac901