





**Please cite the Published Version**

Wang, Tianqi, Zhu, Huitong , Zhou, Yunlan, Ding, Weihong, Ding, Weichao , Han, Liangxiu  and Zhang, Xueqin  (2024) Graph attention automatic encoder based on contrastive learning for domain recognition of spatial transcriptomics. *Communications Biology*, 7 (1). 1351 ISSN 2399-3642

**DOI:** <https://doi.org/10.1038/s42003-024-07037-0>

**Publisher:** Springer

**Version:** Published Version

**Downloaded from:** <https://e-space.mmu.ac.uk/636364/>

**Usage rights:**  [Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

**Additional Information:** The version of record of this article, first published in *Communications Biology*, is available online at Publisher's website: <http://dx.doi.org/10.1038/s42003-024-07037-0>

**Data Access Statement:** All datasets analyzed in this paper can be downloaded in raw form from the original publication. Specifically, the first dataset is human dorsolateral prefrontal cortex data<sup>36</sup> captured using 10 Visium technology and can be downloaded from <http://research.libd.org/spatialLIBD/>. The second dataset is the mouse embryo E9.5 spatial transcriptomic data<sup>37</sup> obtained with Stereo-seq technology and can be downloaded from <https://db.cngb.org/stomics/mosta/>. The third dataset is human breast cancer data obtained from the public 10 Genomics database and can be downloaded from <https://www.10xgenomics.com/resources/datasets/humanbreast-cancer-block-a-section-1-1-standard-1-1-0>. The fourth dataset is mouse brain tissue data obtained from the public 10 Genomics database. The mouse brain tissue contains anterior and posterior regions. Here, we use mouse brain anterior data, which can be downloaded from <https://www.10xgenomics.com/resources/datasets>. The last dataset is de novo NEPC and ARPC data<sup>31</sup> obtained from the public 10 Genomics database, which can be downloaded from <https://db.cngb.org/stomics/datasets/STDS0000227>. The data used in this study have been uploaded to Zenodo and is freely available at: <https://zenodo.org/records/13731512>. All source data underlying the graphs and charts are presented in Supplementary Data. An open-source Python implementation of the GAAEST toolkit is accessible at <https://github.com/tqwang743/GAAEST-main>.

**Enquiries:**

If you have questions about this document, contact [openresearch@mmu.ac.uk](mailto:openresearch@mmu.ac.uk). Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

<https://doi.org/10.1038/s42003-024-07037-0>

# Graph attention automatic encoder based on contrastive learning for domain recognition of spatial transcriptomics

Check for updates

Tianqi Wang<sup>1,6</sup>, Huitong Zhu<sup>1,6</sup>, Yunlan Zhou<sup>2</sup>, Weihong Ding<sup>3</sup>, Weichao Ding<sup>1</sup>✉, Liangxiu Han<sup>4</sup> & Xueqin Zhang<sup>1,5</sup>✉

Spatial transcriptomics is an emerging technology that enables the profiling of gene expression in tissues while preserving spatial location information. This innovative approach is anticipated to provide a comprehensive understanding of the spatial distribution of different cells within tissues and facilitate in-depth analysis of tissue structure. To accurately recognize spatial domains from spatial transcriptomics, we have introduced a generalized deep learning method called GAAEST (Graph Attention-based Autoencoder for Spatial Transcriptomics). Our proposed approach effectively integrates both spatial location information and gene expression data from spatial transcriptomics. Specifically, it leverages spatial location details to construct a neighborhood graph and employs a graph attention network-based encoder to embed gene expression information into a spatially informed space. At the same time, to further optimize the learned potential embedding, self-supervised contrastive learning is introduced to capture spatial information at three levels: local, global and contextual feature of spots. Finally, the decoder reconstructs gene expressions, which are then clustered to identify spatial domains with similar expression patterns and spatial proximity. Based on our experiments conducted on multiple datasets, GAAEST consistently outperforms existing state-of-the-art methods. The proposed GAAEST demonstrates excellent capabilities in spatial domain recognition, positioning it as an ideal tool for advancing spatial transcriptomics research.

The organization of a multicellular organism is characterized by the presence of physically clustered cells that share similar characteristics and perform specific functions. These cells communicate with each other through intricate signaling mechanisms. They exhibit spatially coordinated patterns of gene expression and histology, giving rise to distinct spatial domains. By linking the gene expression profiling of cells to their spatial localization and visualizing the spatial distribution of different cell types within tissues, we can gain deeper insights into the spatial organization of cellular tissues and the progression of diseases. The recent advancements in spatial transcriptomics (ST) technology offer a new approach to integrating cellular spatial information with gene expression. This approach enables the characterization of gene expression patterns in relation to spatial information, which was previously not achievable with traditional non-spatial single-cell RNA sequencing (scRNA-seq) techniques. At present, the existing ST techniques are broadly divided into two categories: in situ hybridization

(ISH)-based and in situ capture sequencing-based<sup>1–3</sup>. ISH-based methods include MERFISH<sup>4</sup> and seqFISH<sup>5</sup>, in situ capture sequencing-based methods include Visium<sup>6</sup>, Slide-seq<sup>7</sup>, and Stereo-seq<sup>8</sup>. The mRNA transcripts captured with these methods will be referred to as spots in the following.

Deciphering spatial domains is a critical challenge in the field of ST. Currently, there are two main classes of methods used for spatial domain recognition: non-spatial clustering and spatial clustering. Representative classical non-spatial clustering methods, such as K-means, Louvain<sup>9</sup>, and Seurat<sup>10</sup>, usually utilize gene expression data as input to cluster spots into different regions. While these methods are capable of identifying segmented and hierarchical spatial domains for various types of ST data, the results they generate are often unstable and exhibit significant variability. To address this limitation, researchers have proposed incorporating spatial information to enhance spatial domain recognition. For instance, Giotto<sup>11</sup> simulates gene expression in neighboring cells by implementing a Hidden Markov Random

<sup>1</sup>School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China. <sup>2</sup>Department of Clinical Laboratory, Xinhua Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China. <sup>3</sup>Huashan Hospital Affiliated to Fudan University, Shanghai, China. <sup>4</sup>School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Manchester, UK. <sup>5</sup>Shanghai Key Laboratory of Computer Software Evaluating and Testing, Shanghai, China. <sup>6</sup>These authors contributed equally: Tianqi Wang, Huitong Zhu ✉e-mail: [weich@ecust.edu.cn](mailto:weich@ecust.edu.cn); [zxq@ecust.edu.cn](mailto:zxq@ecust.edu.cn)

Field (HMRF), allowing for the identification of spatial domains with expressive similarity. stLearn<sup>12</sup> homogenizes gene expression data using neighborhood information and morphological distances, and then performs unsupervised clustering of similar spots in tissues. SEDR<sup>13</sup> learns gene representations with a deep auto-encoder network and uses variogram auto-encoders to simultaneously embed spatial information. SpaGCN<sup>14</sup> takes into account both spatial location and histological information and identifies spatial domains by aggregating gene expression at neighboring spots via a graph convolutional network (GCN). RESEPT<sup>15</sup> defines spatial domain identification as an image segmentation problem, where 3D embeddings are mapped to RGB channels by a spatial retention map auto-encoder, and the embeddings are treated as 3D images. STAGATE<sup>16</sup> combines gene expression and spatial information to learn a low-dimensional embedding, and introduces a graph attention mechanism to detect spatial domains. BANKSY<sup>17</sup> takes into account both the cell's own transcriptome and its local microenvironment, and by balancing the contributions of the cell's intrinsic gene expression and the expression matrix of neighboring cells, it achieves the problem of domain segmentation within a machine learning framework. NeST<sup>18</sup> employs co-expression hotspots to mine nested structures in spatial data by computing single-gene hotspots, constructing hotspot networks, and extracting communities from the networks. All of the above methods have shown their superiority in spatial domain recognition compared to the previous baseline models. However, since these spatial domain recognition methods use unsupervised learning, the recognized domain boundaries are often uneven and do not match the annotations well.

Recently, several studies have adopted self-supervised learning technology to recognize spatial domains. For example, MUSE<sup>19</sup> reveals the heterogeneity of organization and structure in tissues by utilizing the heterogeneity of transcriptomic and morphological data, capturing and preserving information from each modality through self-reconstruction and self-supervision loss functions. However, self-supervised learning has limitations in coping with complex input data. In contrast, self-supervised contrastive learning has achieved breakthroughs in several tasks. It can further improve the richness and robustness of features by mining the internal information of data. Self-supervised contrastive learning frames the problem as a pretext task. In the context of spatial transcriptomics, recognizing spatial domains can also serve as pretext tasks. These tasks provide a structured learning framework for the model and help to learn meaningful feature representations by constructing positive and negative sample pairs, bringing similar samples closer and dissimilar samples farther apart. For instance, SpaceFlow<sup>20</sup> leverages a two-layer GCN to integrate gene expression data and spatial localization information, using contrastive learning and regularization techniques for optimization. GraphST<sup>21</sup> employs GCN and self-supervised contrastive learning to fully utilize ST data, enhancing the learning of latent representations and clustering of spatial information. These methods have obtained good clustering results on some datasets. However, GCN models exhibit limitations in dealing with graph structures characterized by long-distance dependencies and in dynamically learning the significance of inter-node connections. In addition, these methods have not comprehensively considered multi-level features in the contrastive learning process, resulting in them being proven insufficient in capturing the importance of spatial information.

To address the aforementioned challenges and achieve improved accuracy in spatial domain recognition, we propose an innovative generalized framework named GAAEST. The main contributions of this study are as follows:

In the construction of an autoencoder, considering that spatial transcriptome data usually contains multiple different cell types or has multiple gene expression features with significant local structural differences, we propose using GAT to construct the encoder. GAT has the ability to assign different weights to different adjacent nodes, overcome the interference of noise on gene expression features, and effectively capture long-distance dependency relationships between nodes. Specifically, we use two GATs that can extract multi-level feature representations of nodes, further improving the accuracy of gene expression features. The decoder uses two

FCNs, which help to reduce model resource consumption and training time while maintaining good performance.

In order to further enrich the representation extracted from the auto-encoder, and enhance the latent embeddings at different hierarchy levels, we design three contrastive learning mechanisms in GAAEST. Global feature-based contrastive learning improves the mutual information between the embedding of a single spot and the global embedding of the entire graph, thus endowing the network with the ability to learn global structures and have stronger robustness against neighboring noise. Context feature-based contrastive learning further promotes the network to learn the embedding of hierarchical structures by maximizing the mutual information between spot attributes and total cluster features. Especially, we introduce a local location-based supervised contrastive learning mechanism, which prompts the network to concentrate more on single-spot representations and emphasize spatial location-related properties in spot attributes. We combine these three types of contrastive learning to jointly promote the optimization of the feature extraction process.

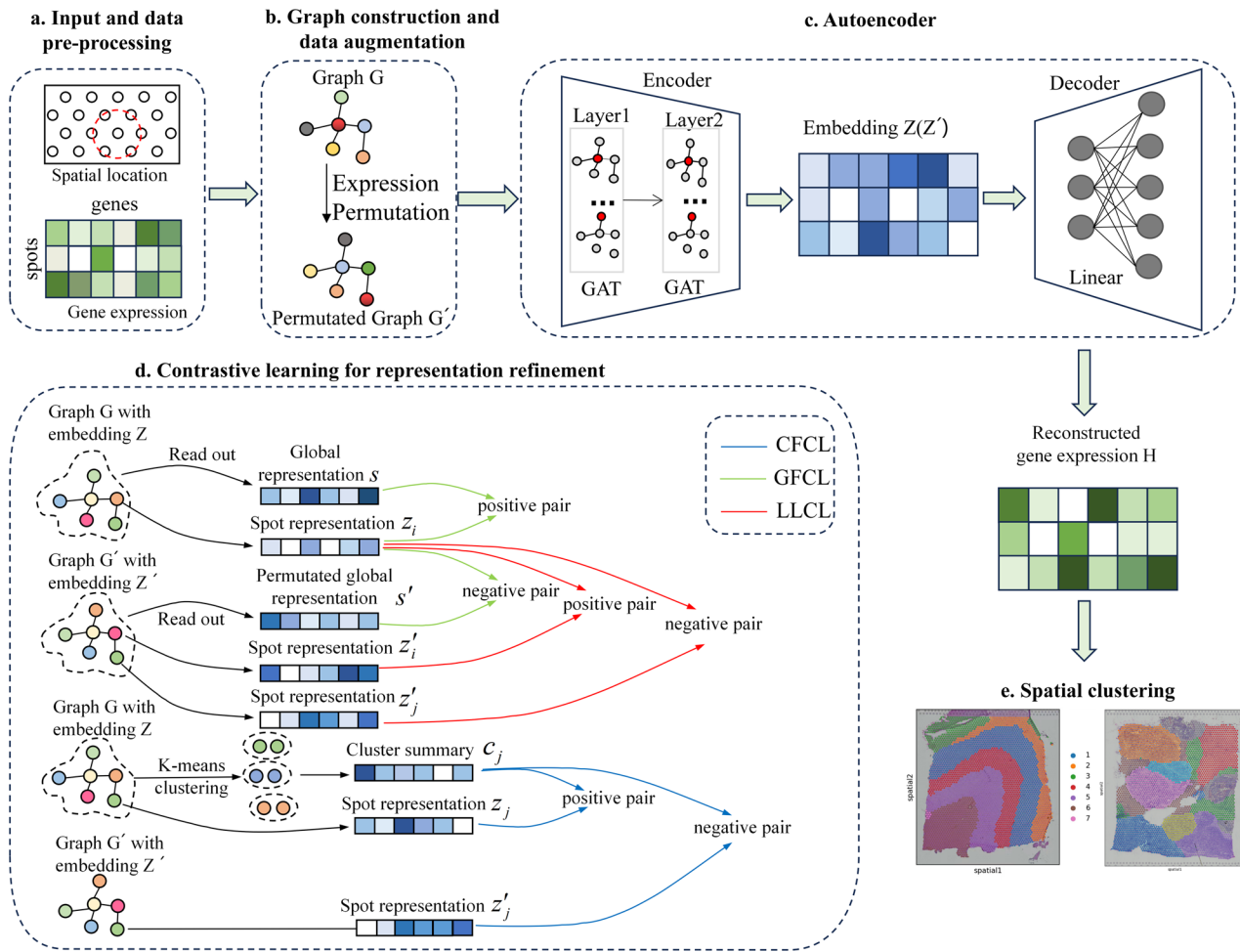
We design multiple experiments on five data sets, including spatial domain identification, gene expression pattern analysis, and functional analysis, across a variety of datasets, such as the human dorsolateral prefrontal cortex (DLPFC) dataset, mouse embryo E9.5 dataset, human breast cancer dataset, the mouse brain anterior dataset, and de novo neuroendocrine prostate cancer (NEPC) coexisting with androgen receptor pathway prostate cancer (ARPC) dataset. These experiments validate the effectiveness of the GAAEST. The results not only reveal potential biological principles and mechanisms to some extent, but also provide important references for defining disease-specific biomarkers, exploring therapeutic targets, and formulating personalized treatment strategies.

## Results

### Overview of GAAEST

For ST data, gene expression information plays a crucial role in determining biological characteristics, and spatial location information includes the spatial coordinates of spots within the tissue structure. To maximize the utilization of both types of information for spatial domain recognition, we propose a deep learning-based method named GAAEST to mine the hidden relationships between them. The established framework is shown in Fig. 1.

In GAAEST, there are five main components, namely data pre-processing, neighbor graph construction and data augmentation, auto-encoder, self-supervised contrastive learning for embedding refinement, and spatial clustering. Specifically, GAAEST takes ST data as input, which consists of a gene expression matrix and spatial location information of spots. Firstly, we apply a filtering process to eliminate genes with low variability from the raw gene expression matrix (Fig. 1a). Next, a neighbor graph is constructed to serve as the input for the auto-encoder. This graph captures the spatial neighborhood relationships through edges, while the spots' gene expression vectors are attached to the nodes. Additionally, a data augmentation technique is performed to construct a permuted neighbor graph (Fig. 1b). Then, a two-layer GAT is adapted to encode the input graph into a low-dimensional embedding, and the decoder with two linear layers is utilized to reconstruct the gene expression matrix (Fig. 1c). To further optimize the low-dimensional embedding, the self-supervised contrastive learning can be employed to capture the spatial context information (Fig. 1d). The local location-based contrastive learning is designed for making network place more emphasis on the representation of a single spot and enhancing the network's attention towards properties related to spatial location in spot attributes. Global feature-based contrastive learning improves the mutual information between the embedding of a single spot and the global embedding of the entire graph, thus endowing the network with the ability to learn global structures and have stronger robustness against neighboring noise. Context feature-based contrastive learning further promotes the network to learn the embedding of hierarchical structures by maximizing the mutual information between spot attributes and total cluster features. These three contrastive learnings form a complementary relationship with each other and jointly promote the optimization of representation. Finally,



**Fig. 1 | Overview of GAAEST. a** GAAEST takes in the gene expression matrix and spatial location information in ST data and performs data pre-processing. **b** A neighbor graph is constructed with spatial location information and data augmentation is employed to construct a permuted neighbor graph. **c** A two-layer GAT encodes the input into a low-dimensional embedding, and a linear layer-based decoder reconstructs the gene expression matrix. **d** The embedding is optimized by

using three-level self-supervised contrastive learning: local location-based contrastive learning (LLCL), global feature-based contrastive learning (GFCL), and context feature-based contrastive learning (CFCL). **e** After training, the reconstructed gene expression data are clustered to achieve the recognition of the spatial domain of ST data.

the reconstructed gene expression data are clustered by Mclust<sup>22</sup> method to achieve the recognition of spatial domain in ST data (Fig. 1e). We performed comprehensive testing and comparisons on different ST datasets, such as the human dorsolateral prefrontal cortex (DLPFC) dataset, mouse embryo E9.5 dataset, human breast cancer dataset, the mouse brain anterior dataset and de novo neuroendocrine prostate cancer (NEPC) coexisting with androgen receptor pathway prostate cancer (ARPC) dataset. The clustering results obtained from these experiments clearly demonstrate the advantages of GAAEST in spatial domain recognition.

**GAAEST improves the clustering accuracy in human dorsolateral prefrontal cortex dataset (DLPFC)**

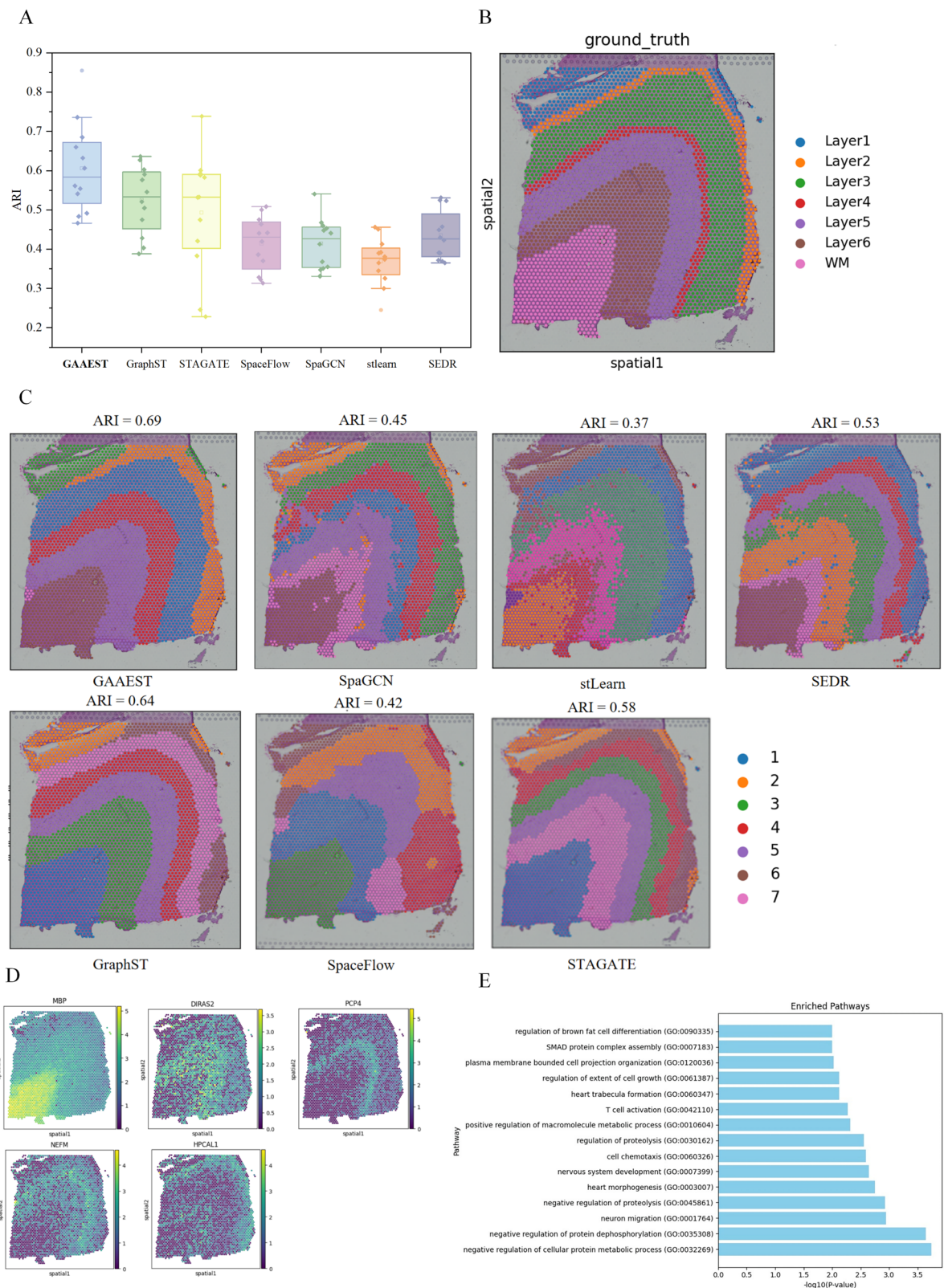
In the DLPFC dataset, Maynard et al. manually annotated the white matter (WM) and cortical layers (L1–L6) in 12 slices using gene markers and cellular structure<sup>23</sup>. To evaluate the performance of GAAEST in recognizing spatial domains, we compared it with six existing state-of-the-art methods, namely stLearn, SpaGCN, SEDR, STAGETE, SpaceFlow, and GraphST. Among these methods, stLearn and SpaGCN utilized histological images, while SEDR, STAGETE, SpaceFlow, and GraphST did not rely on histological images. The clustering results of the seven methods were illustrated in Fig. 2A. It was apparent from the values that across all 12 slices in the DLPFC dataset, GAAEST achieved the highest median ARI score among all methods, GraphST and STAGETE closely followed in performance, while

the remaining methods had lower ARI values. These findings demonstrated that the spatial domains identified by GAAEST exhibited the highest similarity with the manual annotation and aligned with the established understanding of cortical stratification in neuroscience.

Furthermore, we generated visualizations of the manual annotation and clustering results for each method on 151,673 slices, which were shown in Figs. 2B and 2C. Upon examination, it was evident that SpaceFlow struggled to adequately distinguish the seven layers. SpaGCN and stLearn only partially separated the hierarchical structure, falling short of accurately replicating the manual annotation. SEDR and STAGATE succeeded in clearly separating the seven layers, however, the boundaries appeared blurred, and there was noticeable noise interference. GraphST and GAAEST exhibited better layering results, but there were challenges with layer 3 and layer 4 in the manual annotation being appropriately represented by the pink and red areas in GraphST. Additionally, layer 2 was not well recognized by GraphST. In contrast, GAAEST successfully corresponded well with the manual annotation for all layers, with the exception of layer 4. Since layer 4 occupied a smaller area in the entire image, it had less influence on the ARI value.

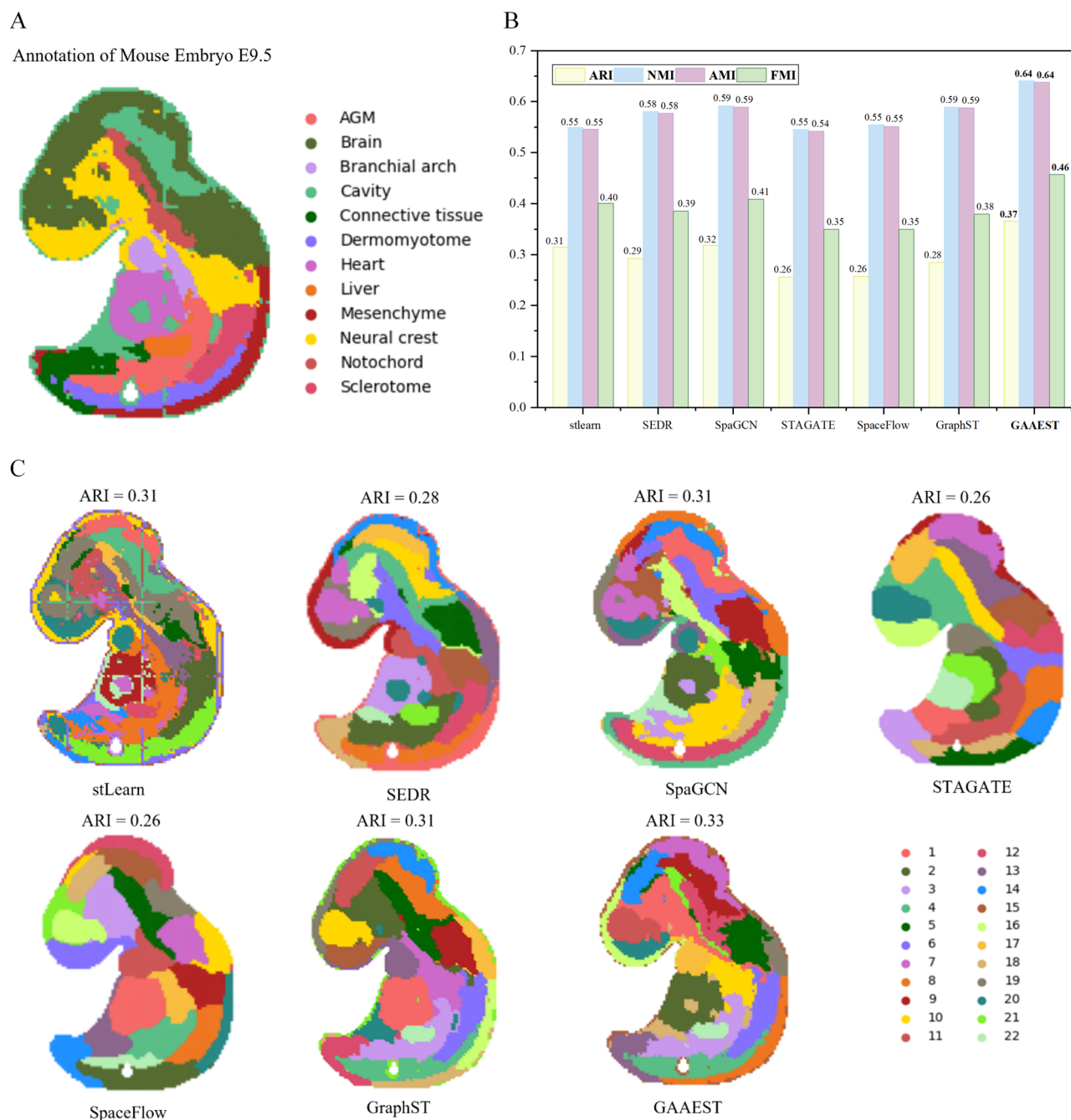
Additionally, we conducted a differential expression analysis to substantiate the biological significance of the spatial domains identified through GAAEST. In this analysis, we employed the identified clusters as a foundation for identifying genes exhibiting differential expression within each spatial domain. The detected differentially expressed genes (DEGs)





**Fig. 2 | Experimental results of GAAEST on the DLPC dataset. A** Box plots of ARI values for GAAEST and six comparative methods on 12 slices. In the boxplot, the center line represents the median quartile, the box limits represent the upper and lower quartiles, and the whiskers denote the 1.5× interquartile range. **B** Manual annotation results on 151,673 slices. **C** Visualization of the spatial clustering results

for GAAEST and six comparative methods on the 151,673 slice. **D** Visualization of gene expression for layer-specific genes in slide 151,673. **E** The enrichment analysis of GO for differentially expressed genes (25 genes) in domain 2 (orange layer in the result).



**Fig. 3 | Experimental results of GAAEST on mouse embryo E9.5 dataset.** **A** Tissue domain annotation of mouse embryo E9.5 data obtained from the original Stereo-seq study. **B** The value of ARI, NMI, AMI, and FMI for GAAEST and comparative methods. **C** Spatial clustering visualization of GAAEST and six comparative methods.

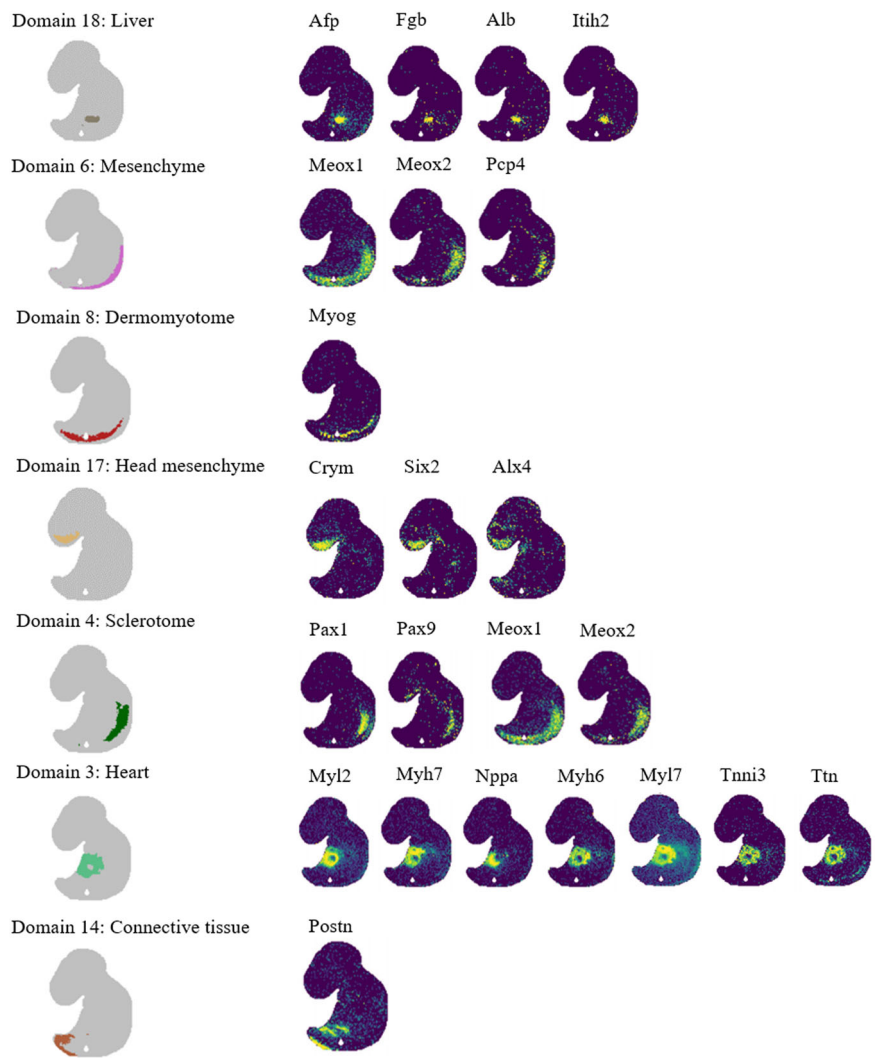
exhibited discernible expression patterns, as depicted in Fig. 2D. Notably, among these genes, PCP4 emerged as a well-established marker gene in the prefrontal cortex<sup>24</sup>. Furthermore, we performed Gene Ontology (GO) analysis on the detected 25 DEGs in region 2, characterized by a significant *p*-value of less than 0.01, and the results were presented in Fig. 2E. Based on it, several gene pathways associated with negative regulation of protein metabolism and protein dephosphorylation were identified in the examined region. Significant gene expression patterns related to neuronal migration, cardiac morphogenesis, and nervous system development were observed. Furthermore, cell chemotaxis, proteolytic regulation, and cell growth emerged as prominent factors influencing gene expression in this region. Notably, these findings aligned with the functional characteristics of the marker genes for cluster 2, namely *ENC1* and *HOPX*. *ENC1* and *HOPX* were

known to be involved in nervous system development and cell migration, which was consistent with the previously identified “Neuron migration” and “Nervous system development” gene pathways. Moreover, *HOPX* exhibited an association with “Heart Morphogenesis” as well.

**GAAEST refines the identification of known regions on the mouse embryo E9.5 dataset**

In this experiment, we utilized the mouse embryos E9.5 dataset, which included tissue domain annotations as shown in Fig. 3A. Initially, we set the number of clusters to 12, corresponding to the original annotation with 12 clusters. The results of all methods were presented in Supplementary Fig. 1, where GAAEST demonstrated the most favorable outcomes compared to other methods. To achieve a more detailed tissue segmentation with higher

**Fig. 4 | Visualization of specific spatial domain obtained from GAAEST clustering and domain-specific marker gene expression.** The gene expression heatmaps of domain-specific marker genes correspond well with the specific domains identified by GAAEST. The liver region is marked by *Afp*, *Fgb*, *Alb*, *Itih2*, mesenchyme by *Meox1*, *Meox2*, *Pcp4*, dermomyotome by *Myog*, head mesenchyme by *Crym*, *Six2*, *Alx4*, Sclerotome by *Pax1*, *Pax9*, *Meox1*, *Meox2*, heart by *Myl2*, *Myl7*, *Nppa*, *Myh6*, *Myh7*, *Tnni3*, *Ttn*, and connective tissue by *Postn*.



resolution, we incrementally increased the number of clusters. Supplementary Fig. 2 showcased the results of regional segmentation as the number of clusters progressively increased from 12 to 24. Through rigorous evaluation, we determined that a final number of 20 clusters yielded the highest evaluation index.

In the case of 20 clusters, we evaluated the spatial domain recognition performance of GAAEST and six comparative methods on the mouse embryo E9.5 dataset, as depicted in Fig. 3B. This analysis demonstrated that GAAEST reached the highest value of ARI, NMI, AMI, and FMI among all methods. The spatial clustering visualization results of each method are shown in Fig. 3C. Upon observation, it was apparent that the contours of the domains recognized by stLearn were not clear enough, and the continuity of the domains was poor. Also, STAGATE and SpaceFlow could not fully recognize the Mesenchyme region. SpaGCN and GraphST were unable to identify the liver region. STAGATE, SEDR, GraphST, and Spaceflow all struggled with Heart recognition. In contrast, the domain recognition results of GAAEST aligned better with the annotations, and the contour segmentation was clear.

To facilitate a more intuitive comparison between the domain recognition results of GAAEST and the annotation results, we drew the visualization diagram showcasing the specific spatial domain identified by GAAEST clustering alongside domain-specific marker gene expressions. As depicted in Fig. 4, the clusters of GAAEST could better match the annotated domains. More importantly, they exhibited a strong agreement with known marker genes associated with major organs. For instance, the liver region

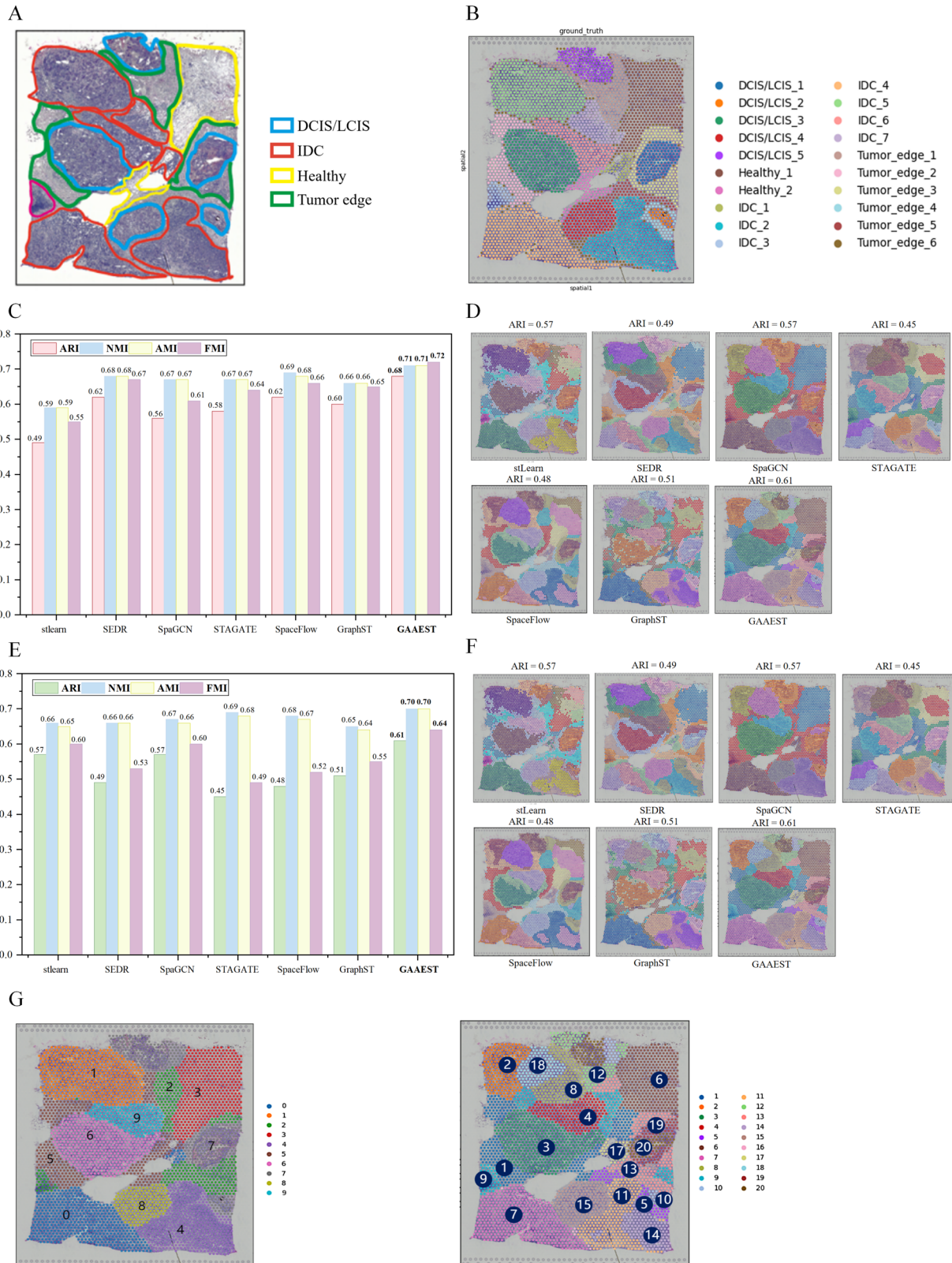
was marked by *Afp*, *Fgb*, *Alb*, *Itih2*, mesenchyme by *Meox1*, *Meox2*, *Pcp4*, dermomyotome by *Myog*, head mesenchyme by *Crym*, *Six2*, *Alx4*, Sclerotome by *Pax1*, *Pax9*, *Meox1*, *Meox2*, heart by *Myl2*, *Myl7*, *Nppa*, *Myh6*, *Myh7*, *Tnni3*, *Ttn*, and connective tissue by *Postn*. The gene expression heat maps of these marker genes corresponded well with the specific domains identified by GAAEST.

### GAAEST reveals the finer organization of the human breast cancer dataset

As shown in Fig. 5A, the human breast cancer data were mainly classified into four morphotypes: ductal carcinoma in situ/lobular carcinoma in situ (DCIS/LCIS), invasive ductal carcinoma (IDC), healthy tissue (Healthy), and tumor surrounding edge (Tumor edge). Clearly, the DCIS/LCIS domains were surrounded by IDC or Tumor edge. According to Fig. 5B, the annotated data were manually segmented into 20 domains, which consisted of 2 Healthy, 5 DCIS/LCIS, 7 IDC, and 6 Tumor edge domains.

Referring to literature<sup>25</sup>, we first designated the number of clusters to 10 and evaluated the spatial domain recognition performance of GAAEST and six comparative methods. The results were presented in Fig. 5C. It was observable that the ARI scores of stLearn, SpaGCN, and STAGATE ranged from 0.49 to 0.57, while the values of SEDR, SpaceFlow, GraphST, and GAAEST were all above 0.59. Particularly, GAAEST achieved the highest ARI score of 0.68. Also, in terms of NMI, AMI, and FMI, only GAAEST had values exceeding 0.7. These findings were reflected in the clustering visualization diagram, as depicted in Fig. 5D. The clustering results obtained by

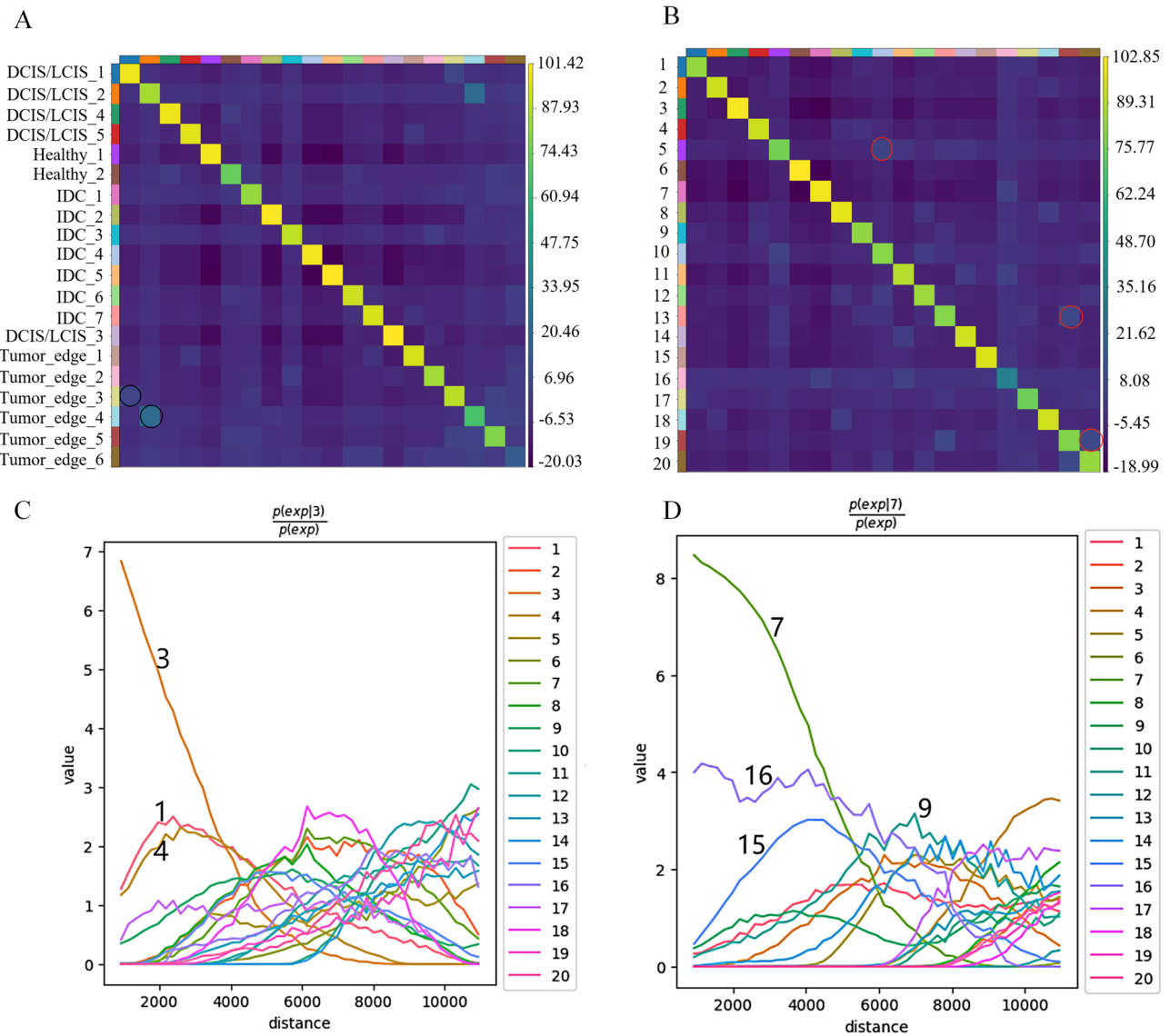




**Fig. 5 | Experimental results of GAAEST on human breast cancer.** **A** Spatial transcriptomic data of human breast cancer sample annotated by pathologists. **B** The manual annotation of 20 domains in human breast cancer data. **C** ARI, NMI, AMI, and FMI for GAAEST and comparative methods (cluster number = 10). **D** The clustering visualization of GAAEST and six comparative methods (cluster

number = 10). **E** ARI, NMI, AMI, and FMI for GAAEST and comparative methods (cluster number = 20). **F** The clustering visualization of GAAEST and six comparative methods (cluster number = 20). **G** The clustering visualization of GAAEST with cluster numbers ranging from 10 to 20.





**Fig. 6 | Statistical analysis of clustering results of human breast cancer data by GAAEST. A** Neighbor enrichment heatmap of ground truth. **B** Neighbor enrichment heatmap of GAAEST clustering results. **C** The co-occurrence score of domain 3 and all

other domains in GAAEST. **D** The co-occurrence score of domain 7 and all other domains in GAAEST.

stLearn, SpaGCN, and STAGATE appeared segmented and discontinuous, differing significantly from the manual annotation. Conversely, the clustered regions produced by SEDR, SpaceFlow, GraphST, and GAAEST exhibited continuity. Among them, the domains recognized by GAAEST had a high degree of overlap with the manually annotated domains, namely domains 0–7 corresponded to IDC\_4, IDC\_5, healthy\_1, IDC\_2, Tumor\_edge\_2, DCIS/LCIS\_3, DCIS/LCIS\_1, and DCIS/LCIS\_4, respectively. In order to achieve higher resolution in domain recognition, we increased the number of clusters to 20, and the results are presented in Fig. 5E. Overall, the evaluation values of all methods decreased compared to the previous experiment, but GAAEST still achieved the highest ARI, NMI, AMI, and FMI values. Obviously, the ARI values of SpaceFlow, STAGATE, and SEDR dropped below 0.5, which was quite a large decrease compared to previous results. The visualization results in Fig. 5F further supported this observation, as the clustering results of these three methods appeared less similar to the manual annotation displayed in Fig. 5B.

With the purpose of observing the impact of different cluster numbers, we drew the cluster visualization diagram of GAAEST, as shown in Fig. 5G. The number of clusters was 10 for the left image and 20 for the right. As observed from the figure, regardless of the clustering number being 10 or 20,

GAAEST outperformed other methods in terms of recognizing domains with higher regional continuity and less noise. As the number of clusters increased, the spatial domains with high heterogeneity, such as the tumor regions, became finer and thinner. Meanwhile, the spatial domains with low heterogeneity, such as healthy regions, still remained consistent. This observation indicated that GAAEST exhibited good robustness in capturing both fine-grained and homogeneous spatial domains.

So as to better compare with ground truth, we conducted the analysis below with a cluster number of 20. Utilizing the proximity information derived from the spatial connectivity graph of the clusters, we calculated the enrichment score for neighbor enrichment analysis<sup>26</sup>. The neighborhood enrichment heat maps of ground truth and GAAEST clustering results were depicted in Fig. 6A and Fig. 6B, respectively. In these heatmaps, each block represented the enrichment score between two domains, with a higher score indicating a stronger correlation between them. For example, in Fig. 6A, two black circles indicated that Tumor\_Edge\_3 was in close proximity to DCIS/LCIS\_1, and Tumor\_Edge\_4 surrounded DCIS/LCIS\_2. Similarly, in Fig. 6B, when GAAEST was applied for spatial domain recognition, three red circles highlighted the proximity relationships between Domain 5 and Domain 10, Domain 13 and Domain 19, and Domain 19 and Domain 20. These

observations were consistent with the annotated domains in Fig. 5B, namely, Tumor\_Edge\_4 surrounded DCIS/LCIS\_2, Tumor\_Edge\_5 neared Tumor\_Edge\_3, Tumor\_Edge\_3 being close to DCIS/LCIS\_1. Such analysis outcomes served as further evidence to support the clustering accuracy of the proposed GAAEST framework. For a specific domain, the co-occurrence score can reflect the degree of proximity between the two domains<sup>26</sup>. Figure 6C, D showed the co-occurrence scores for domains 3 and 7 in the GAAEST clustering results. From the figure, we observed that Domain 3 is the neighbor of Domains 1 and 4, while Domain 7 was adjacent to Domains 16, 15, and 9. The relationship of these domains could be confirmed in Fig. 5E, and it was also consistent with the distribution of annotated domains in Fig. 5B, namely DCIS/LCIS\_3 was the neighbor of Tumor\_Edge\_2 and IDC\_6, IDC\_4 neared Tumor\_Edge\_2, DCIS/LCIS\_4 and IDC\_3.

### GAAEST more precisely describes relevant tissue regions on the mouse brain anterior dataset

In this experiment, we evaluated the spatial domain identification performance of GAAEST on the mouse brain anterior dataset. The manual annotation provided 52 labeled regions, as illustrated in Fig. 7A. To align with this annotation, we set the number of clusters to 52. Figure 7B displayed histograms depicting the values of ARI, NMI, AMI, and FMI for GAAEST and six other comparative methods. As seen in the figure, it was evident that GAAEST achieved the highest clustering accuracy, while GraphST and stLearn exhibited minor differences in the three evaluation metrics but performed suboptimally compared to GAAEST. STAGATE and SpaceFlow demonstrated similar NMI and AMI values to GAAEST but showcased lower ARI and FMI scores. SEDR and SpaGCN exhibited the lowest clustering scores among the methods evaluated. Figure 7C illustrates the clustering visualization results of the seven methods. It was evident that SpaGCN and SEDR exhibited poor clustering performance, with different clusters mixing together and rough, indistinct boundaries between the clusters. SpaceFlow and STAGATE showcased clearer cluster boundaries, but many of the identified clusters did not align well with the manual annotations. The clusters generated by stlearn were closer to annotations, however, some clusters were mixed with others, and the cluster boundaries appeared relatively messy. In contrast, GraphST and GAAEST demonstrated the ability to capture the 52 clusters in the mouse brain anterior dataset more clearly, with distinct and well-defined boundaries.

Additionally, we further identified marker genes in the major regions of the mouse frontal brain. As shown in Fig. 7D, in our study, we detected the marker gene *GPR88* with the highest expression intensity in region 43 (striatum). This is consistent with the known literature<sup>27</sup>, stating that “*GPR88* is highly and almost exclusively expressed in the striatum”. Studies have shown that *GPR88* is involved in regulating various neurotransmission processes, such as dopaminergic neurotransmission, mental activity, and motor control<sup>28</sup>. It has the potential to serve as a beneficial target for the treatment of Parkinson’s disease, schizophrenia, and other neurological and mental disorders. Similarly, we detected the potential marker gene *Doc2b* (double C2, gamma) in the Juxtglomerular granule cells (GR) of region 4 of the main olfactory bulb (MOB). We validated this finding based on the widely accepted database Allen Brain Atlas (<https://portal.brain-map.org/>, experiment ID: 69174316). Related studies have shown that the *Doc2b* gene plays a crucial role in the insulin release process, potentially providing insights for the research and treatment of diabetes and related diseases<sup>29</sup>. In region 7 (primary olfactory cortex, MO), we detected the critical marker gene *Scn1b*, which has also been verified in the Allen Brain Atlas. Currently, there are therapeutic strategies that target *Scn1b* gene and its related channels, attempting to restore normal sodium channel function to inhibit epileptic seizures<sup>30</sup>.

### GAAEST performs domain recognition and spatial gene expression analysis on the de novo neuroendocrine prostate cancer coexisting with androgen receptor pathway prostate cancer dataset

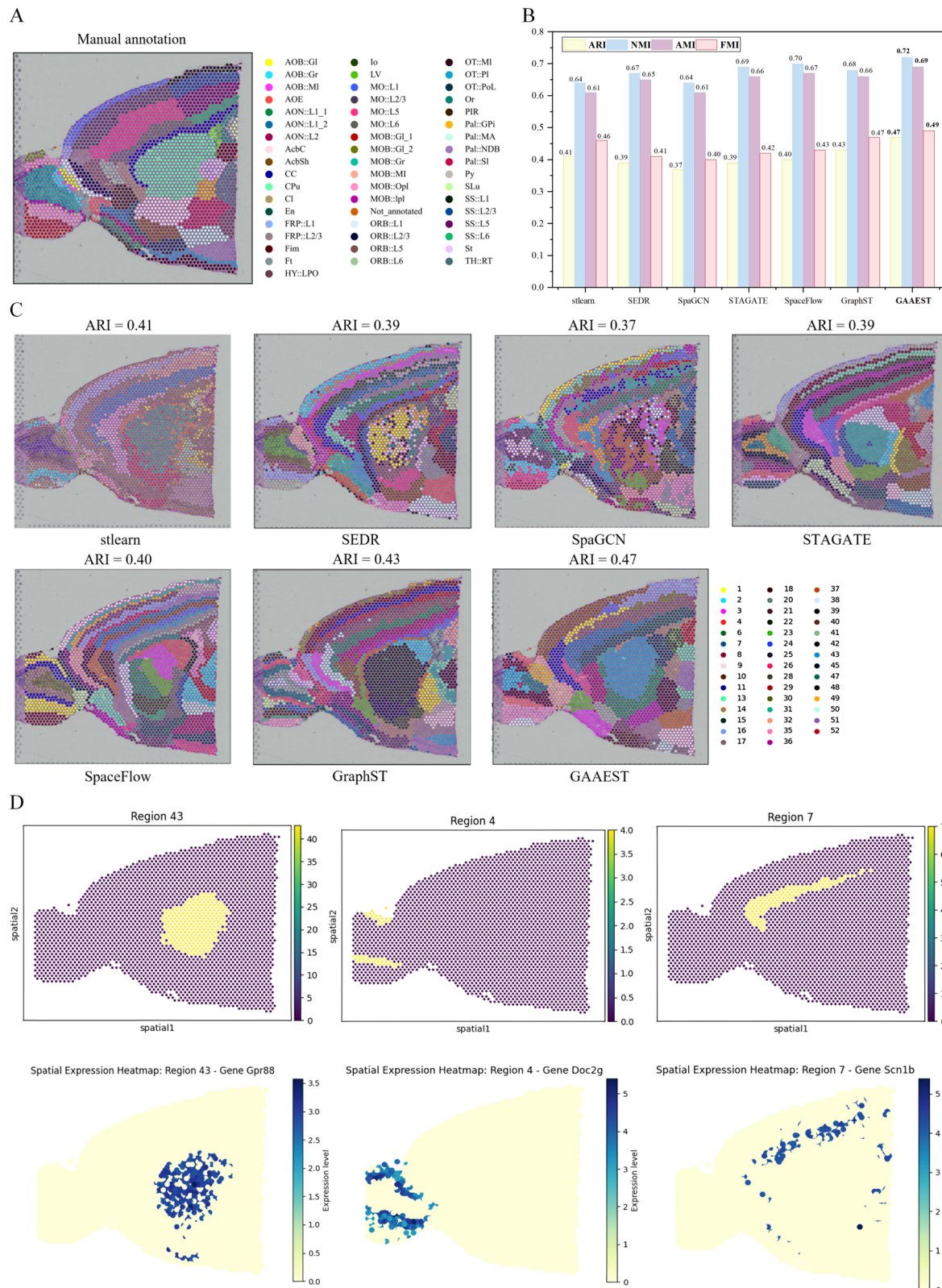
In this experiment, we evaluated the domain recognition performance of GAAEST on the human prostate cancer dataset. Usually, the occurrence of

simultaneous de novo neuroendocrine prostate cancer (NEPC) and androgen receptor pathway prostate cancer (ARPC) within the same tissue is exceedingly rare. However, in a recent report by Ryuta Watanabe et al.<sup>31</sup>, spatial gene expression analysis was performed on a 78-year-old man with coexisting metastatic de novo NEPC and ARPC. The Fig. 8A displayed a part of their findings, HE image revealed the separation of the NEPC region from the ARPC region, with intercalated noncancerous tissue between them, indicating the simultaneous and ectopic occurrence of ARPC and de novo NEPC within the same prostate. Spatial gene expression analysis utilizing CytAssist Visium classified the cells within the prostate tissue into 12 distinct clusters. Cluster 8 corresponded to the ARPC region, while Cluster 3 represented the NEPC region. Based on these findings, this dataset was further applied to GAAEST and other methods, with the number of clusters set to 12, yielding the region segmentation outcomes presented in Fig. 8B. Since a clear ground truth reference was unavailable for the dataset, we introduced two additional evaluation metrics, namely the Silhouette Coefficient score and the Davies-Bouldin score. Among all comparative methods, GAAEST demonstrated the most favorable evaluation index values, with its SC value about 11% higher than the second-best method and a DB value 0.8% lower. To facilitate a more distinct comparison, we generated separate visualizations for cluster 0 and cluster 4, which were obtained from the GAAEST clustering algorithm, as illustrated in Fig. 8C, D. Specifically, cluster 4 was found to correspond to the ARPC region, and cluster 0 represented the NEPC region. In comparison, although STAGATE, SpaGCN, and stLearn achieved some success in segmenting the NEPC area, they faced certain challenges in differentiating the ARPC area. The accuracy of other methods in dividing NEPC and ARPC areas was relatively low. These results further confirm GAAEST’s strong performance and discriminative ability.

Furthermore, to ascertain the biological relevance of the identified domains through GAAEST, a comprehensive analysis of differential gene expression was performed. Figure 8E, F displayed the 4 differentially expressed genes specific to cluster 0 and cluster 4 regions, respectively. Remarkably, differentially expressed gene maps of these two regions corresponded well with the NEPC and de novo ARPC regions depicted in Fig. 8A. In Fig. 8E, it was observed that the expression levels of NEPC signature genes, such as *TKTL1*, *CNN3*, and *LRMP*, were higher in the de novo NEPC region compared to the ARPC region. Studies showed that in metastatic prostate cancer, *TKTL1* average expression was significantly elevated. The high expression of *TKTL1* in the NEPC area was consistent with the fact that NEPC was a more invasive and aggressive subtype of prostate cancer. Moreover, the high expression of *CNN3* suggested that NEPC tissue might have a greater invasive capability and metastatic potential. The high expression of *LRMP* might impact immune system responses, leading to NEPC immune escape or immunosuppression, thereby increasing the risk of tumor progression. In Fig. 8F, the expression levels of ARPC signature genes, including *KLK2*, *KLK3*, and *ACPP*, were higher in the ARPC region than in the de novo NEPC region. The *KLK2*, *KLK3*, and *ACPP* genes were reported to be regulated by androgen receptors in ARPC, and their protein products played an important role in the development of prostate cancer. These observations provided cues for a deeper exploration of the mechanisms of prostate cancer development and helped provide more precise treatment strategies for specific subtypes of patients.

### GAAEST’s ablation and comparative experiments confirm the effectiveness of each module

To demonstrate the necessity and effectiveness of each module in GAAEST, we conducted an ablation experiment on the human breast cancer dataset. We used the same experimental parameters as mentioned earlier, with cluster numbers set to 10 and 20, respectively. The experimental results are shown in Supplementary Table 2. It could be seen that all three types of contrastive learning were favorable for feature embedding learning. Among them, the global feature-based contrastive learning had the most significant impact on performance improvement, and the ARI was reduced by 14%

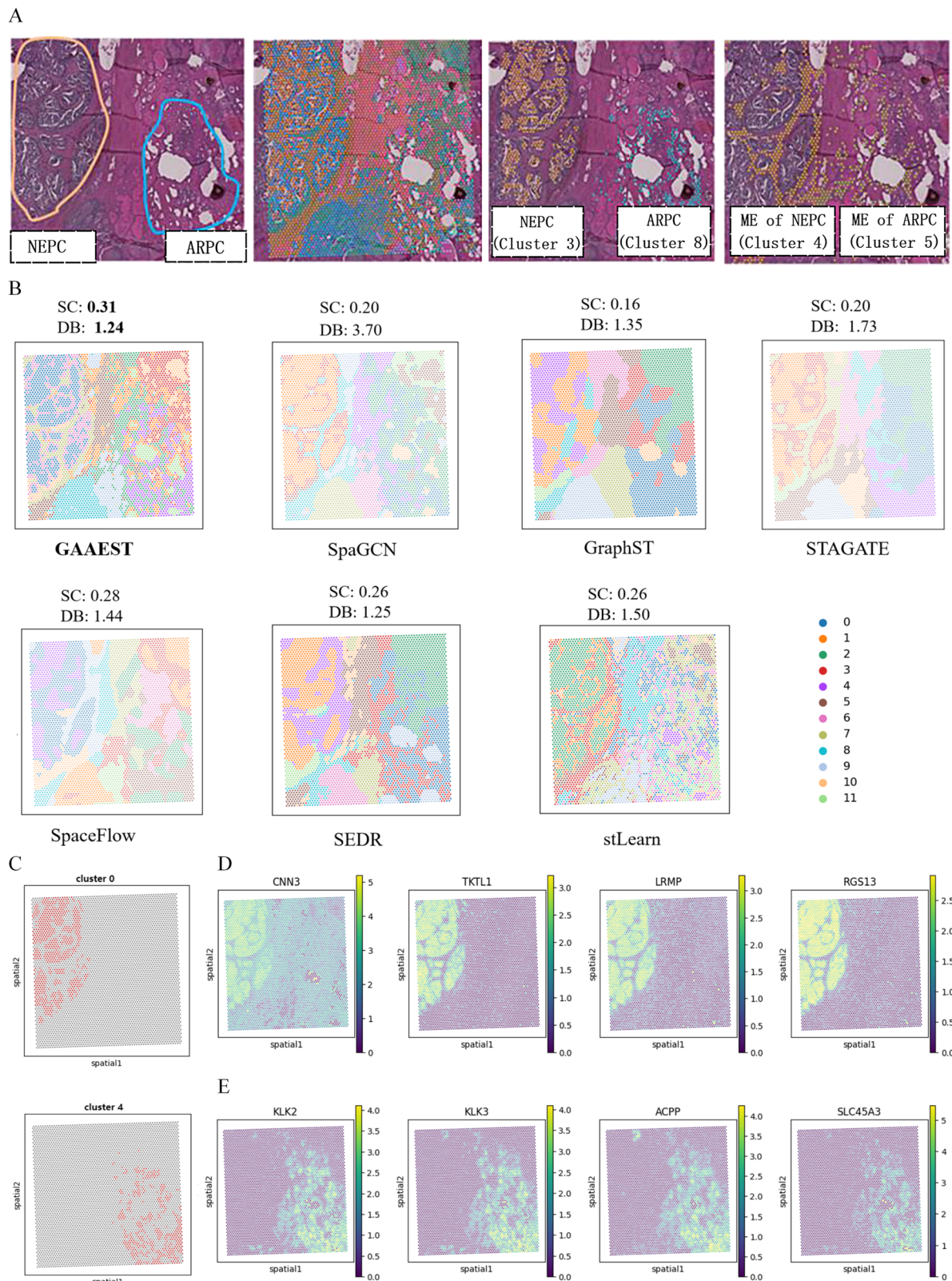


**Fig. 7 | GAAEST optimizes the identification of known tissue structures in the mouse brain anterior.** **A** Structural annotations of the brain anterior from the adult mouse. **B** The value of ARI, NMI, AMI, and FMI for GAAEST and comparative methods. **C** Clustering visualization of GAAEST and six comparative methods. **D** Visualization of gene expression for differentially expressed genes in regions 43, 4, and 7.

when the global feature was not considered. The reconstruction operation of the gene expression matrix also played a crucial role, as evidenced by a reduction in ARI of up to 65% when feature reconstruction was not incorporated. GAAEST achieved the best performance when all four

modules were used together. It was concluded that through the utilization of three-level contrastive learning and auto-encoder modules, GAAEST could effectively extract crucial features and achieve superior spatial clustering results. We visualized the state of the three modules in capturing spatial

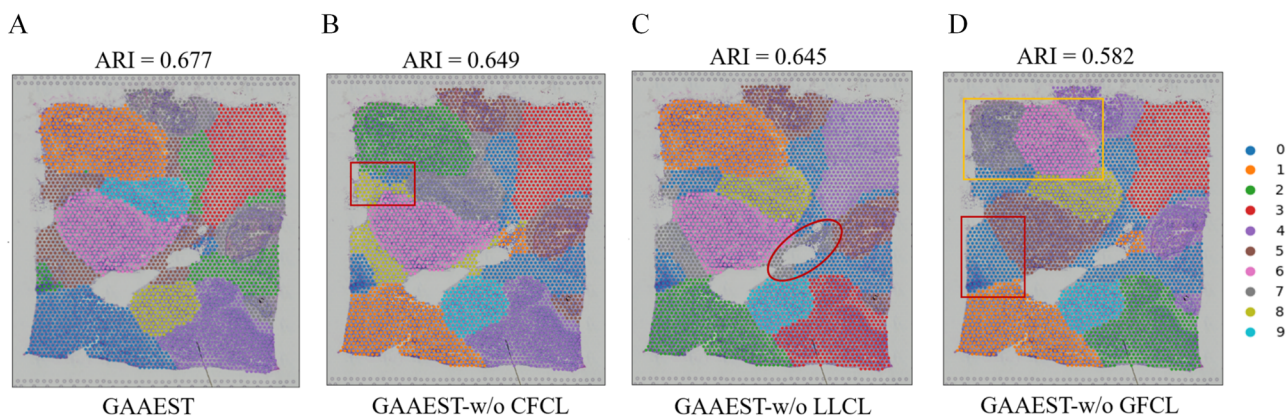




**Fig. 8 | GAAEST perform domain recognition and spatial gene expression analysis on de novo neuroendocrine prostate cancer coexisting with androgen receptor pathway prostate cancer dataset. A** Spatial gene expression analysis (CytAssist Visium) results provided by Ryuta Watanabe et al. **B** Clustering visualization of GAAEST and six comparative methods. **C** Visualization of specific cluster

0 obtained from GAAEST clustering. **D** Visualization of specific cluster 4 obtained from GAAEST clustering. **E** Spatial analysis of gene expression of four differentially expressed genes in NEPC region. **F** Spatial analysis of gene expression of 4 differentially expressed genes in ARPC region.





**Fig. 9 | Visualization results of GAAEST ablation experiment in human breast cancer dataset (Clustering number = 10).** **A** Clustering results of GAAEST on human breast cancer dataset. **B** Clustering results of GAAEST w/o CFCL on human

breast cancer dataset. **C** Clustering results of GAAEST w/o LLCL on human breast cancer dataset. **D** Clustering results of GAAEST w/o GFCL on human breast cancer dataset.

relationships when clustering is 10, as illustrated in Fig. 9. When removing CFCL, as shown in Fig. 9B, we can observe that the tumor boundary region within the red circle in the figure is identified as two parts, suggesting that the model may not accurately reflect the similarity between samples of the same tissue type. When removing LLCL, as shown in Fig. 9C, the gray region within the red circle, which should be divided into Healthy and IDC two regions, is instead classified as a single region, implying that the model may not have captured the local feature differences of the spots within the region. When removing the GFCL, as presented in Fig. 9D, the IDC region in the red box cannot be identified, and the IDC region in the yellow box is divided into two parts, indicating that the embedding may be less effective in preserving the global summary of the entire graph. By observing these results, we demonstrate that these three modules play unique and important roles in the model. Each module is indispensable for its contribution to representation learning and spatial relationship recognition.

We also conducted a comparative experiment to demonstrate the effectiveness of employing a GAT-based encoder and an FCN-based decoder in our GAAEST model. In this experiment, we replaced the auto-encoder structure in GAAEST with those used in STAGATE (2022) and GraphST (2023), while keeping the other structures unchanged. The dataset and parameters utilized were consistent with those employed in the ablation study, and the results were presented in Supplementary Table 3. As shown in the table, GAAEST's auto-encoder achieved the best performance, exhibiting an average improvement of 6% in terms of clustering accuracy compared to the sub-optimal method. This highlighted the superiority of GAAEST's auto-encoder in capturing and representing the underlying features in the gene expression matrix, leading to more effective spatial domain recognition.

## Discussion

Accurately identifying spatial domains is critical for comprehending tissue structure and biological functions. In order to uncover spatial domains with well-defined boundaries and minimal noise from spatial transcriptomic data, we propose a deep learning-based framework named GAAEST. This framework integrates the spatial location information and gene expression of spots to construct a neighborhood graph and uses a two-layer graph attention-based encoder to extract features and embed gene expression into the latent space. The local location-based, global featured-based, and context feature-based self-supervised contrastive learning are then employed to capture spatial information at three different levels. Meanwhile, by constructing these three loss functions, the learned potential embeddings are made to contain more informative and discriminative features. Finally, the gene expressions are reconstructed by the decoder, and a clustering algorithm is adopted to identify spatial domains with similar expressions. Based on extensive experiments conducted on multiple datasets, GAAEST

consistently outperforms existing state-of-the-art methods in spatial domain recognition tasks. These results highlight the unique advantages of GAAEST, showcasing its effectiveness in accurately identifying spatial domains and providing valuable insights into tissue organization and biological functions.

In the study, we conducted several experiments to evaluate the performance of GAAEST on different datasets. When applied to the human dorsolateral prefrontal cortex, GAAEST accurately revealed the laminar organization of the DLPFC, achieving an 8% improvement in clustering accuracy compared to the suboptimal method. Furthermore, GAAEST demonstrated superior performance in identifying known organizational structures in the mouse brain anterior dataset, outperforming the second-best solution by 9% in terms of accuracy. These results highlighted the excellence of GAAEST in depicting fine structures and its potential for uncovering intricate organizational patterns in complex brain datasets. In the mouse embryo dataset, the identified domains displayed a strong agreement with known marker genes associated with major organs, providing support for the effectiveness of GAAEST in capturing essential biological information. Moreover, in experiments conducted on the human breast cancer dataset, GAAEST consistently outperformed other methods across all evaluation metrics. Even when the number of clusters was increased to allow for more detailed segmentation, GAAEST still achieved an average 6% improvement compared to less effective techniques. This indicated that GAAEST offered a valuable approach to studying tumor evolution and investigating the interactions between tumor micro-environments. In the de novo NEPC coexisting with the ARPC dataset, GAAEST effectively identified the de novo NEPC region and ARPC region compared to other methods. The additional gene expression analysis was expected to help with the development of treatments for NEPC and improve the prognosis of patients with castration-resistant prostate cancer.

The primary feature of GAAEST lies in its utilization of graph attention networks and self-supervised contrastive learning to acquire latent embeddings. Similarly, STAGATE also employs graph attention networks to capture both gene expression and the spatial context of spots. However, the key distinction between them is that GAAEST incorporates graph self-supervised contrastive learning mechanisms to enhance features by leveraging spatial mutual information at three different levels. Furthermore, both SpaceFlow and GraphST also employ graph contrastive learning for spatial clustering. However, GAAEST demonstrates a significant performance advantage over these two methods due to its more rational definition of positive/negative pairs, objective function, and contrastive loss. These differences contribute to GAAEST's superior performance in spatial domain recognition tasks compared to SpaceFlow and GraphST. To validate the effectiveness of each of these differences, we conducted ablation experiments, which confirmed that the integration of gene expression and spatial

information with these improvements led to enhanced clustering performance in GAAEST.

Currently, spatial transcriptomics offers the possibility of studying gene expression and spatial information in the initial state of cells. However, there are also issues with low sensitivity or limited gene flux in transcriptional detection. To overcome these challenges and enhance spatial resolution, the integration of ST data with snRNA-seq data becomes essential to enhance spatial resolution and uncover the finer spatial structures and microenvironments of tissues. In our future work, we will focus on advancing these research directions to further improve the scalability of GAAEST. Additionally, we plan to investigate the behavior and associations of cellular tissues in a spatial context, enabling a deeper understanding of the intricate relationships within tissues.

## Methods

With the aim of fully utilizing the available information for spatial domain recognition, we propose a deep learning-based method named GAAEST, which aims to uncover the hidden relationship between gene expression and spatial location. An overview of our spatial domain recognition method is shown in Fig. 1. GAAEST consists of five main components: data pre-processing, neighbor graph construction, data augmentation, auto-encoder, self-supervised contrastive learning for embedding refinement, and spatial clustering. We provide a detailed description of each part below.

### Data pre-processing

Data pre-processing operation can provide high-quality data for subsequent GAAEST models. For GAAEST, the inputs are the gene expression matrix and spatial location information of spots. For the purpose of decreasing the difference between high and low expression genes and reducing the impact of noise data, we first perform a logarithmic transformation on the raw gene expression matrix and standardize it to library size. Then, the standardized gene expression matrix is scaled to unit variance and zero mean. At the same time, in order to focus more on genes that contribute significantly to cell changes, we select the first 3000 highly variable genes (HVGs) as input samples, which can effectively reduce dimensions and improve the efficiency of subsequent analysis.

After pre-processing, the gene expression matrix is set as  $X = \{x_1, x_2, \dots, x_N\} \subseteq R^{N \times F}$ , where  $N$  represents the total number of spots,  $F$  represents the feature dimension of a spot.

### Neighbor graph construction and data augmentation

The spatial location information within ST data is valuable for assessing the similarity between spots, which helps to recognize spatial domains. In order to effectively utilize this information, we first construct a neighbor graph.

To convert the spatial location information of spots into a neighbor graph  $G = (V, E)$ , the  $k$ -nearest neighbor algorithm is adopted. Here,  $V$  represents the set of spots,  $E$  represents the set of edges connecting the spots and  $k$  is the number of neighbors. Concretely, for a specific spot  $i \in V$ , we calculate the euclidean distance between this spot and all other spots, and choose the  $k$ -closest spots as its neighbors. By analogy, the adjacency relationships between all spots are obtained, and an adjacency matrix  $A \in R^{N \times N}$  is used to denote the adjacency relationships of all spots in the neighbor graph  $G$ . If spot  $j \in V$  is the  $k$ -closest neighbor of spot  $i$ , then  $A_{ij} = 1$ , otherwise 0. To address disparities in spot degrees and ensure that all spots have a more equal influence in calculations, we need to regularize the adjacency matrix  $A$ . Let  $\tilde{A} \in R^{N \times N}$  be the adjacency matrix after regularization, then

$$\tilde{A} = \bar{A} \times A \times \bar{A} \tag{1}$$

$$\bar{A} = \text{diag} \left( \text{power} \left( \sum_j A_j, -0.5 \right) \right) \tag{2}$$

where  $(\times)$  is the meaning of cross product,  $\text{diag}$  means creating or extracting the diagonal matrix, and  $\text{power}$  represents the exponentiation.

After constructing the neighborhood graph, we use a regularized adjacency matrix  $\tilde{A}$  and a preprocessed gene expression matrix  $X$  to represent the adjacency relationship and gene expression information of spots in the neighbor graph, respectively.

Meanwhile, we conduct data augmentation operations to generate positive and negative samples for the subsequent self-supervised contrastive learning task. Specifically, given the neighbor graph  $G = (V, E)$ , we perform a random permutation of the gene expression vectors between each spot while maintaining the original graph structure to obtain a permuted neighbor graph  $G' = (V', E')$ . At this time, for graph  $G'$ , the adjacency relationship of the spots remains unchanged, that is, the adjacency matrix still is  $\tilde{A}$ , but the gene expression matrix changes from  $X$  to  $X' = \{x'_1, x'_2, \dots, x'_N\} \subseteq R^{N \times F}$ .

### Autoencoder

In order to effectively integrate gene expression information and spatial location information in ST data, and to remove unnecessary noise from gene expression information, we use an auto-encoder framework to learn the feature embedding and reconstruct gene expression matrix. The designed auto-encoder includes two parts: an encoder  $E_w$  and a decoder  $D_w$ .

We employ a two-layer graph attention network (GAT)<sup>32</sup> as the encoder  $E_w$ . GAT network can utilize an attention mechanism to assign weights to the input data, which enables the model to emphasize key features and reduce attention to irrelevant ones. Encoder  $E_w$  takes the regularized adjacency matrix  $\tilde{A}$  and the preprocessed gene expression matrix  $X$  as inputs, and outputs the low-dimensional embedding  $Z = E_w(X, \tilde{A})$ ,  $Z = \{z_1, z_2, \dots, z_N\} \subseteq R^{N \times F'}$ , where  $F'$  represents the potential embedding dimension of spots. The specific formula of  $E_w$  is as follows:

$$X_i = \begin{cases} \text{ELU}(\text{BatchNorm}(\text{GAT}_i^{(K)}(X_{i-1}, \tilde{A}))) & \text{if } i = 1 \\ \text{ELU}(\text{BatchNorm}(\text{GAT}_i^{(K)}(\text{ELU}(X_{i-1}), \tilde{A}))) & \text{if } 1 < i < L \\ \text{GAT}_i^{(K)}(\text{ELU}(X_{i-1}), \tilde{A}) & \text{if } i = L \end{cases} \tag{3}$$

where  $X_i$  represents the  $i$ th layer's output,  $\text{GAT}_i^{(K)}$  denotes the  $i$ th GAT layer with  $K$  heads. Then  $Z = X_L$ ,  $L$  signifies the total number of GAT layers in encoder.

The decoder  $D_w$  takes  $Z$  as input and outputs the reconstructed gene expression matrix  $H = D_w(Z)$ ,  $H = \{h_1, h_2, \dots, h_N\} \subseteq R^{N \times F}$ . Since the primary objective of decoder  $D_w$  is to reconstruct the original data, it does not focus on feature selection and extraction in the data. Therefore, we directly use two fully connected layers in the decoder, deviating from the GAT structure used in the encoder structure. In  $D_w$ , the  $\text{ELU}$  activation function and batch normalization operation are applied to all the hidden layers except the bottleneck layer.

Finally, in the auto-encoder module, we adopt the mean square error loss function to calculate the reconstruction loss  $L_{\text{RECON}}$  of the gene expression matrix, which can be expressed as follows:

$$L_{\text{RECON}} = \sum_{i=1}^N \|x_i - h_i\| \tag{4}$$

where  $x_i$  and  $h_i$  are the original and reconstructed gene expression vectors of spot  $i$ , respectively.

### Self-supervised contrastive learning

Self-supervised contrastive learning is a training method that eliminates the need for annotated labels and is employed to learn meaningful feature representations from data. By comparing various views or transformations of the same sample, self-supervised contrastive learning enables the model to learn to group similar samples together and differentiate dissimilar

samples. To enhance the informativeness of embedding  $Z$  obtained from the encoder  $E_w$ , we further refine it through self-supervised contrastive learning across three dimensions: local location based, global feature based, and context feature based contrastive learning.

**Local location-based contrastive learning (LLCL).** Spatial location information in spatial transcriptome data can help to identify differences in gene expression between different tissue structures, thereby facilitating the recognition of spatial domain within spatial transcriptomics. Local location based contrastive learning is specifically designed to enhance the network's attention towards properties related to spatial location in spot attributes, so we use it to more accurately distinguish a specific spot from other spots in different locations.

For spots in the neighbor graph  $G$ , we designate the feature embedding derived from the encoder  $E_w$  to be  $Z = E_w(X, A)$ ,  $Z = \{z_1, z_2, \dots, z_N\} \subseteq R^{N \times F'}$ . Similarly, for spots in the permuted neighbor graph  $G'$ , we set it to be  $Z' = E_w(X', A)$ ,  $Z' = \{z'_1, z'_2, \dots, z'_N\} \subseteq R^{N \times F'}$ . Then, for a specific spot  $i$ , we set its feature embedding  $z_i$  as an anchor, its positive pair is  $(z_i, z'_i)$ , its negative pairs are  $(z_i, z'_j)_{i \neq j}$ . We utilize cosine similarity function to measure the similarity of feature expression between two spots as follows:

$$sim(a, b) = \frac{p(a) \cdot p(b)}{\|p(a)\| * \|p(b)\|} \tag{5}$$

where the function  $p(\cdot)$  represents a projection operation that is constructed using a two-layer multiple-layer perceptron (MLP). Thus, the loss function can be given as follows:

$$L_1(z_i, z'_i) = \log \left( \sum_{j=1}^N A_{[j \neq i]} \exp(sim(z_i, z'_j)) \right) - \log[\exp(sim(z_i, z'_i)/\tau)] \tag{6}$$

where  $\tau$  is the temperature hyperparameter,  $A_{[j \neq i]} \subseteq \{0, 1\}$  is the indicator function.

Considering the neighbor graph  $G$  and the permuted neighbor graph  $G'$  are equally important, we also need to set  $z'_i$  as anchor to calculate the loss  $L_1(z'_i, z_i)$ , so the total local location-based contrastive learning loss  $L_{LLCL}$  can be expressed as:

$$L_{LLCL} = \frac{1}{2N} \sum_{i=1}^N [L_1(z_i, z'_i), L_1(z'_i, z_i)] \tag{7}$$

**Global feature-based contrastive learning (GFCL).** In addition to considering local location information, we also need to establish a certain relationship between local and global features. This is because the local properties of individual spots are inherently consistent with the global properties when they are obtained from the identical part of the same organism. As a result, we improve the mutual information between the feature representations of every single spot and the global summary of the whole graph, thus allowing the network to cover both local and global-level features.

We employ a method resembling local location-based contrastive learning. Firstly, an average aggregation function  $R : Z(Z') \subseteq R^{N \times F'} \rightarrow s(s')$  is utilized to get the global summary  $s$  or  $s'$  of the graph:

$$s = R(E_w(X, \tilde{A})) \tag{8}$$

$$s' = R(E_w(X', \tilde{A})) \tag{9}$$

Also, for spot  $i$ , we set its feature embedding  $z_i$  as an anchor. Then its positive sample pair is  $(z_i, s)$ , and its negative sample pair is  $(z_i, s')$ . A bilinear transformation layer-based discriminator  $D_c : R^{F'} \times R^{F'} \rightarrow R$  is designed to give a higher probability score to positive pair than the negative pair. We use

the Jensen-Shannon-based binary cross entropy loss to construct the loss  $L_2(z_i, s, s')$ :

$$L_2(z_i, s, s') = \sum_{i=1}^N [\log D_C(z_i, s) + \log(1 - D_C(z_i, s'))] \tag{10}$$

Since the importance of neighbor graph  $G$  and the permuted neighbor graph  $G'$  are equal, we also need to set  $z'_i$  as anchor to calculate the loss  $L_2(z'_i, s', s)$ :

$$L_2(z'_i, s', s) = \sum_{i=1}^N [\log D_G(z'_i, s') + \log(1 - D_G(z'_i, s))] \tag{11}$$

Therefore, the total global feature-based contrastive learning loss  $L_{GFCL}$  is calculated as:

$$L_{GFCL} = \frac{1}{2N} (L_2(z_i, s, s') + L_2(z'_i, s', s)) \tag{12}$$

**Context feature-based contrastive learning (CFCL).** Inspired by literature<sup>33</sup>, we know that pots within the same tissue type tend to exhibit similarities in terms of marker genes and morphological structures. Building upon this knowledge, in addition to global and local level attributes, we further try to learn the embedding with context feature by maximizing the mutual information between the individual spot representation and the cluster-level summary.

We apply the  $K$ -means clustering algorithm on  $Z$  and acquire  $K$  initial clusters. Then, the centroid of each cluster  $\phi_k \subseteq R^{1 \times F'}$  is iteratively updated as follows:

$$\phi_k = \frac{\sum_i r_{ik} z_i}{\sum_i r_{ik}}, \text{ with } r_{ik} = \frac{\exp(-\gamma sim(\phi_k, z_i))}{\sum_k \exp(-\gamma sim(\phi_k, z_i))} \tag{13}$$

where  $k = 1, 2, \dots, K$ , and  $\gamma$  represent an inverse-temperature hyperparameter. For every individual spot in neighbor graph  $G$ , we aim to maximize the mutual information between its feature embedding  $z_i$  and corresponding cluster-level summary  $c_i$ , the value of  $c_i$  can be calculated as:

$$c_i = \sigma \left( \sum_{k=1}^K r_{ik} \phi_k \right) \tag{14}$$

where  $r_{ik}$  represents the likelihood of assigning spot  $i$  to cluster  $k$  and  $\sum_{k=1}^K r_{ik} = 1$ .

In line with the global feature-based contrastive learning, we design another bilinear transformation layer-based discriminator  $D_G : R^{F'} \times R^{F'} \rightarrow R$  to assign a higher score to the positive pair  $(z_i, c_i)$  than the negative pair  $(z'_i, c_i)$ . Then, we construct the context feature-based contrastive learning loss  $L_{CFCL}$  using the Jensen-Shannon-based binary cross-entropy loss, which is defined as follows:

$$L_{CFCL} = \sum_{i=1}^N E_{(X, \tilde{A})} [\log D_G(c_i, z_i)] + \sum_{i=1}^N E_{(X', \tilde{A})} [\log(1 - D_G(c_i, z'_i))] \tag{15}$$

weighted by  $\lambda_b$ , the total contrastive learning loss  $L_{CL}$  can be given as:

$$L_{CL} = \lambda_1 L_{LLCL} + \lambda_2 L_{GFCL} + \lambda_3 L_{CFCL} \tag{16}$$

**Total model loss**

With the  $L_{RECON}$  obtained in the Autoencoder section and  $L_{CL}$  obtained in the self-supervised contrastive learning section, we calculate the total loss



function  $L_{TOTAL}$  of the GAAEST model as follows:

$$L_{TOTAL} = \alpha L_{CL} + \beta L_{RECON} \quad (17)$$

where  $\alpha$  and  $\beta$  are both weight parameters. GAAEST obtains better-reconstructed gene expression  $H$  by minimizing the total loss.

Throughout the learning process in GAAEST, we construct the reconstruction loss for gene expression and the contrastive learning loss for embedding optimization. The former aims to capture informative features from both spatial location and gene expression, while the latter helps to refine the potential embeddings from three different levels and enable a more comprehensive and discriminative representation.

### Spatial clustering

After the training process, we obtain the optimized reconstructed gene expression matrix  $H$ . In order to distinguish domains that exhibit similar gene expression and spatial proximity, it is essential to perform spatial clustering operation. In this case, we use the non-spatial allocation algorithm Mclust to assign spots into different spatial domains.

Mclust is a clustering method based on the maximum expectation algorithm. It assumes that the data come from the mixed Gaussian distribution, and determines the best model and clustering result by the maximum likelihood function. Each cluster in the clustering results is considered as a spatial domain, encompassing spots that demonstrate similar gene expression profiles and spatial proximity.

To adjust the clustering results further, we design an optional optimization step. For a given spot  $i$ , we take the neighboring spots within a radius  $r$  as its neighbors, and then the labels assigned to these neighbor spots are queried. If more than 50% of the neighbors are assigned to label  $n$ , GAAEST will reassign the label  $n$  to spot  $i$ . The settings for  $r$  vary depending on the data type.

### GAAEST spatial domain identification algorithm

The GAAEST spatial domain identification algorithm is described as follows:

#### Algorithm. GAAEST

**Input:** Gene expression matrix, Spatial location matrix

Number of iteration  $T$ , Encoder  $E_w$ , Decoder  $D_w$

**Output:** Spatial domain recognition results

- 1: Get preprocessed gene expression  $X$
- 2: Construct neighbor graph  $G$  and permuted graph  $G'$
- 3: **For**  $t = 0$  to  $T-1$  **do**
- 4:   Get feature embeddings  $Z = E_w(G)$ ,  $Z' = E_w(G')$
- 5:   Get reconstructed feature  $H = D_w(Z)$
- 6:   Calculate reconstruction loss  $L_{RECON}$  by Eq. 4
- 7:   Calculate contrastive learning loss  $L_{CL}$  by Eq. 16
- 8:   Get total loss  $L_{TOTAL} = L_{RECON} + L_{CL}$
- 9:   Update the parameters of  $E_w$  and  $D_w$  by minimizing the total loss  $L_{TOTAL}$
- 10: **End for**
- 11:  $H = D_w(E_w(G))$
- 12: Apply Mclust to  $H$
- 13: **Return** the clustering results, in which each cluster is considered a spatial domain

### Experimental details

In experiments, we set  $k = 3$  for constructing the neighbor graph. The input, hidden, and output dimensions of the encoder are set to 3000, 256, and 64, respectively. Also, these parameters of the decoder are set to 64, 256, and 3000, respectively. In the contrastive learning module, the temperature hyperparameter is set to 0.5. In the spatial clustering module, the settings for optimization radius  $r$  vary depending on the datasets. For the DLPFC and Human breast cancer dataset, the value is 35. For the remaining datasets, the optional clustering optimization step is not needed as they all have fine-

grained domains, so the value is set to 0; During the training process, the total loss function is optimized by the Adam optimizer<sup>34</sup> with an initial learning rate of 0.001, weight decay of 0, and a total number of training rounds of 600.

GAAEST is trained on GeForce RTX A4000 Ti GPU with 16 GB of video memory, 6 x E5-2680 v4 CPU, 28 GB of RAM, and implemented by Python and PyTorch\_pyG<sup>35</sup>. The training time varies from 30 s to 3 min on different datasets.

### Evaluated metrics and criteria

For the spatial domain identification task of spatial transcriptome data, the adjusted rand index (ARI), normalized mutual information (NMI), adjusted mutual information (AMI), and Fowlkes–Mallows index (FMI) are employed as evaluation metrics if ground truth exists in the dataset. Silhouette coefficient (SC) score and Davies–Bouldin (DB) score are introduced when the ground truth doesn't exist. These metrics are widely utilized to assess the effectiveness of spatial domain identification models.

Adjusted rand index (ARI): the ARI measures the degree of agreement between the clustering results and the ground truth by comparing the consistency between all pairs of samples in their clusters. Assuming that  $U = \{U_1, U_2, \dots, U_C\}$  and  $V = \{V_1, V_2, \dots, V_C\}$  are two clustered label sets, the ARI value is calculated by the formula:

$$ARI = \frac{\sum_{i,j} \binom{n_{ij}}{2} - \left[ \sum_i \binom{n_i}{2} \sum_j \binom{n_j}{2} \right] / \binom{n}{2}}{\frac{1}{2} \left[ \sum_i \binom{n_i}{2} + \sum_j \binom{n_j}{2} \right] - \left[ \sum_i \binom{n_i}{2} \sum_j \binom{n_j}{2} \right] / \binom{n}{2}} \quad (18)$$

where  $n_i, n_j$  is the number of spots belonging to  $U_i$  and  $V_j$ , respectively,  $n_{ij}$  denotes the number of spots corresponding to  $U_i$  and  $V_j$ .

Normalized Mutual Information (NMI): NMI measures the mutual information between two sets. The NMI between  $U$  and  $V$  is calculated as:

$$NMI(U, V) = \frac{MI(U, V)}{\text{mean}(H(U), H(V))} \quad (19)$$

where MI is the mutual information between  $U$  and  $V$ , and  $H(U)$  and  $H(V)$  are the entropy of the clustering labels, the specific formulas are as follows:

$$MI(U, V) = \sum_{i=1}^{|U|} \sum_{j=1}^{|V|} P(i, j) \log\left(\frac{P(i, j)}{P(i)P(j)}\right) \quad (20)$$

$$H(U) = - \sum_{i=1}^{|U|} P(i) \log(P(i)) \quad (21)$$

$$H(V) = - \sum_{j=1}^{|V|} P(j) \log(P(j)) \quad (22)$$

where  $P(i) = |U_i|/N$  and  $P(j) = |V_j|/N$  are the probabilities of the selected spot falling into  $U_i$  or  $V_j$ .

Adjusted mutual information (AMI): AMI is a measure that incorporates the baseline of stochastic consistency and is an improvement over the ARI. The formula for AMI is as follows:

$$AMI(U, V) = \frac{MI(U, V) - E\{MI(U, V)\}}{\max\{H(U), H(V)\} - E\{MI(U, V)\}} \quad (23)$$

Fowlkes–Mallows Index (FMI): For the paired samples of the same category, the Fowlkes–Mallows index measures the degree of fit between the clustering results and the ground truth by calculating their similarity. The



formula for calculating the FMI is as follows:

$$\text{FMI} = \frac{\text{TP}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})}} \quad (24)$$

where TP and FP denote the number of samples that are correctly or incorrectly clustered in the ground truth, respectively. FN denotes the number of positively clustered samples that are not correctly identified.

Silhouette coefficient (SC) score: SC is computed by evaluating the mean intra-cluster distance ( $a$ ) and the mean nearest-cluster distance ( $b$ ) for each sample. The SC score for a sample is calculated as  $(b - a)/\max(a, b)$ , and the optimal SC score is 1, indicating well-separated and compact clusters, while the worst score is  $-1$ , indicating poor clustering.

Davies–Bouldin (DB) score: DB score provides an assessment of clustering quality by considering the average similarity between each cluster and its most similar cluster. Similarity is measured based on the ratio of within-cluster distances to between-cluster distances. A lower DB score indicates better clustering, with a minimum score of zero.

### Methods used for comparison

To demonstrate the effectiveness of GAAEST in spatial domain identification, we compare GAAEST with six state-of-the-art methods, including SEDR<sup>13</sup>, stLearn<sup>12</sup>, SpaGCN<sup>14</sup>, SpaceFlow<sup>20</sup>, STAGATE<sup>16</sup>, and GraphST<sup>21</sup>. To offer a comparative analysis of these methods, we present Supplementary Table 1, providing a summarized overview of some approaches, which are the state-of-the-art methods using feature extraction or contrastive learning. For each dataset, the number of target clusters set by the six comparison methods is the same. For SEDR, STAGATE, and GraphST, all parameters are kept at their default settings. For stLearn, in order to get better clustering results, we assign distinct weight values for each dataset within its SME\_normalize function. Specifically, in the case of DLPCF, mouse embryo and human breast cancer data, weights = “physical\_distance”. In the case of mouse brain anterior data, weights = “weights\_matrix\_pd\_md”. For SpaGCN, we set the parameter about “Histology” to “False” if no histology image exists in the dataset. For SpaceFlow, we adjust the resolution in clustering to get the target number of clusters, and the other parameters are default. The results of all comparison methods in the experiments are reproduced by us under the same experimental environment.

### Statistics and reproducibility

No statistical method was used to predetermine the sample size. All collected data were included in the analyses without exclusion. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

All datasets analyzed in this paper can be downloaded in raw form from the original publication. Specifically, the first dataset is human dorsolateral prefrontal cortex data<sup>36</sup> captured using 10× Visium technology and can be downloaded from <http://research.libd.org/spatialLIBD/>. The second dataset is the mouse embryo E9.5 spatial transcriptomic data<sup>37</sup> obtained with Stereo-seq technology and can be downloaded from <https://db.cngb.org/stomics/mosta/>. The third dataset is human breast cancer data obtained from the public 10× Genomics database and can be downloaded from <https://www.10xgenomics.com/resources/datasets/humanbreast-cancer-block-a-section-1-1-standard-1-1-0>. The fourth dataset is mouse brain tissue data obtained from the public 10× Genomics database. The mouse brain tissue contains anterior and posterior regions. Here, we use mouse brain anterior data, which can be downloaded from <https://www.10xgenomics.com/resources/datasets>. The last dataset is de novo NEPC and ARPC data<sup>31</sup> obtained from the public 10× Genomics database, which can be downloaded

from <https://db.cngb.org/stomics/datasets/STDS0000227>. The data used in this study have been uploaded to Zenodo and is freely available at: <https://zenodo.org/records/13731512>. All source data underlying the graphs and charts are presented in Supplementary Data.

### Code availability

An open-source Python implementation of the GAAEST toolkit is accessible at <https://github.com/tqwang743/GAAEST-main>.

Received: 6 January 2024; Accepted: 9 October 2024;

Published online: 18 October 2024

### References

- Asp, M., Bergenstråhle, J. & Lundeberg, J. Spatially resolved transcriptomes—next generation tools for tissue exploration. *BioEssays* **42**, 1900221 (2020).
- Rao, A., Barkley, D., França, G. S. & Yanai, I. Exploring tissue architecture using spatial transcriptomics. *Nature* **596**, 211–220 (2021).
- Palla, G., Fischer, D. S., Regev, A. & Theis, F. J. Spatial components of molecular tissue biology. *Nat. Biotechnol.* **40**, 308–318 (2022).
- Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
- Shah, S. et al. Dynamics and spatial genomics of the nascent transcriptome by intron seqFISH. *Cell* **174**, 363–376 (2018).
- Rao, N., Clark, S. & Habern, O. Bridging genomics and tissue pathology: 10x genomics explores new frontiers with the visium spatial gene expression solution. *Genet. Eng. Biotechnol. News* **40**, 50–51 (2020).
- Rodrigues, S. G. et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).
- Xia, K. et al. Single-cell Stereo-seq enables cell type-specific spatial transcriptome characterization in Arabidopsis leaves. *BioRxiv* <https://doi.org/10.1101/2021.10.20.465066> (2021).
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*. P10008 (2008).
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
- Dries, R. et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol.* **22**, 1–31 (2021).
- Pham, D. et al. stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues. *bioRxiv* <https://doi.org/10.1101/2020.05.31.125658> (2020).
- Xu, H. et al. Unsupervised spatially embedded deep representation of spatial transcriptomics. *Genome Med.* **16**, 12 (2024).
- Hu, J. et al. 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 (2021).
- Chang, Y. et al. Define and visualize pathological architectures of human tissues from spatially resolved transcriptomics using deep learning. *Computational Struct. Biotechnol. J.* **20**, 4600–4617 (2022).
- Dong, K. & Zhang, S. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention autoencoder. *Nat. Commun.* **13**, 1–12 (2022).
- Singhal, V. et al. BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nat. Gene.* **56**, 431–441 (2024).
- Walker, B. L. & Nie, Q. NeST: nested hierarchical structure identification in spatial transcriptomic data. *Nat. Commun.* **14**, 6554 (2023).

19. Bao, F. et al. Integrative spatial analysis of cell morphologies and transcriptional states with MUSE. *Nat. Biotechnol.* **40**, 1200–1209 (2022).
20. Ren, H., Walker, B. L., Cang, Z. & Nie, Q. Identifying multicellular spatiotemporal organization of cells with SpaceFlow. *Nat. Commun.* **13**, 4076 (2022).
21. Long, Y. et al. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with graphst. *Nat. Commun.* **14**, 1155 (2023).
22. Fraley, C., Raftery, A. E., Murphy, T. B., & Scrucca, L. mclust version 4 for R: normal mixture modeling for model-based clustering, classification, and density estimation. *Technical report*. **597** (2012).
23. Maynard, K. R. et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.* **24**, 425–436 (2021).
24. Watakabe, A. et al. Area-specific substratification of deep layer neurons in the rat cortex. *J. Comp. Neurol.* **520**, 3553–3573 (2012).
25. Xu, C. et al. DeepST: identifying spatial domains in spatial transcriptomics by deep learning. *Nucleic Acids Res.* **50**, e131 (2022).
26. Palla, G. et al. Squidpy: a scalable framework for spatial omics analysis. *Nat. Methods* **19**, 171–178 (2022).
27. Van Waes, V., Tseng, K. Y. & Steiner, H. GPR88: A putative signaling molecule predominantly expressed in the striatum: Cellular localization and developmental regulation. *Basal ganglia.* **1**, 83–89 (2011).
28. Laboute, T. et al. The orphan receptor GPR88 blunts the signaling of opioid receptors and multiple striatal GPCRs. *Elife* **9**, e50519 (2020).
29. Yu, H., Rathore, S. S., Davis, E. M., Ouyang, Y. & Shen, J. Doc2b promotes GLUT4 exocytosis by activating the SNARE-mediated fusion reaction in a calcium-and membrane bending-dependent manner. *Mol. Biol. cell* **24**, 1176–1184 (2013).
30. Chen, C. et al. Epilepsy and sudden unexpected death in epilepsy in a mouse model of human SCN1B-linked developmental and epileptic encephalopathy. *Brain Commun.* **5**, fcad283 (2023).
31. Watanabe, R. et al. Spatial Gene Expression Analysis Reveals Characteristic Gene Expression Patterns of De Novo Neuroendocrine Prostate Cancer Coexisting with Androgen Receptor Pathway Prostate Cancer. *Int. J. Mol. Sci.* **24**, 8955 (2023).
32. Veličković, P. et al. Graph attention networks. *International Conference on Learning Representations*. (ICLR, 2018).
33. Zong, Y. et al. conST: an interpretable multimodal contrastive learning framework for spatial transcriptomics. *BioRxiv* <https://doi.org/10.1101/2022.01.14.476408> (2022).
34. Bock, S. & Weiß, M. A proof of local convergence for the Adam optimizer. *IEEE Int. Joint Conf. Neural Netw.* (2019).
35. Fey, M. & Lenssen, J. E. Fast graph representation learning with PyTorch geometric. *ICLR 2019 Workshop on Representation Learning on Graphs and Manifolds*. (2019).
36. Pardo, B. et al. spatialLIBD: an R/Bioconductor package to visualize spatially-resolved transcriptomics data. *BMC Genomics* **23**, 434 (2022).
37. Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792 (2022).

## Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (NSFC No. 51975213).

## Author contributions

X. Z., T. W., and H. Z. envisaged the project and wrote the manuscript. T. W. and H. Z. implemented the model and code and performed the experiments. Y. Z. and W. D. analyzed and interpreted data of spatial transcriptomics, and validated the final method. X. Z. and L. H. provided model evaluation and critical manuscript revision. All authors read and approved the final paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-024-07037-0>.

**Correspondence** and requests for materials should be addressed to Weichao Ding or Xueqin Zhang.

**Peer review information** *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Adib Keikhosravi and Benjamin Bessieres.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024