INTERIOR-BOUNDARY AWARE GRAPH REPRESENTATION LEARNING FOR SPATIALLY RESOLVED TRANSCRIPTOMIC

Anonymous authors
Paper under double-blind review

ABSTRACT

Spatially resolved transcriptomics (SRT) can measure gene expression while preserving spatial information, providing an important tool for studying complex tissue structures and biological processes. In recent years, graph-based methods have been widely used to capture spatial relationships between spots for spatial domain identification. However, existing methods typically adopt the same strategy to model interior and boundary regions, without explicitly considering the structural differences between them. This may make local node representations overly similar and increase the mixing of information from heterogeneous neighbors. As a result, the distinctions between spatial domains can be weakened. Intuitively, interior spots and boundary spots exhibit distinct patterns in local neighborhood features and global relation structures, motivating us to distinguish these two types of spots and model their representations with different strategies. Based on this observation, we propose Boundary-aware Multi-level Relational Learning (BMRL), which estimates boundary characteristics by combining local expression differences with global semantic structure. Building on the identified interior and boundary spots, BMRL applies different representation learning constraints to them and further incorporates structural information at the region and slice levels. This allows structural information at different levels to complement each other and improves the integration of multi-slice data. Extensive experimental results show that BMRL achieves promising performance on various downstream tasks in both single-slice and multi-slice settings.

1 INTRODUCTION

Spatially resolved transcriptomics (SRT) can measure gene expression while preserving spatial location information (Ståhl et al., 2016; Rodriques et al., 2019; Stickels et al., 2021). It provides an important tool for studying tissue structures, cellular heterogeneity, and spatial molecular patterns (Rao et al., 2021; Longo et al., 2021). Spatial domain identification is one of the basic tasks in SRT analysis. It aims to divide spots with similar molecular features and spatial patterns into biologically meaningful tissue regions (Zhao et al., 2021; Hu et al., 2021). Therefore, learning high-quality spot representations that preserve both gene expression and spatial structure is important for accurately analyzing complex tissue structures (Dong & Zhang, 2022; Ren et al., 2022; Long et al., 2023).

In recent years, graph representation learning (Kipf & Welling, 2016; Veličković et al., 2017; Hamilton et al., 2017) has been widely used in SRT data analysis (Li et al., 2022; Xu et al., 2022; Xu et al.; Zong et al., 2022; Liu et al., 2023a). These methods usually build an adjacency graph based on the spatial locations of spots. Graph neural networks are then used to aggregate neighborhood information and encode gene expression and spatial structure into low-dimensional representations(Xu et al., 2022; Zong et al., 2022; Ren et al., 2022). However, interior and boundary spots exhibit different neighborhood characteristics. Interior spots are usually surrounded by relatively homogeneous neighbors. In contrast, boundary spots are more likely to have heterogeneous neighbors from different tissue regions. Directly aggregating information from these heterogeneous neighbors may make local spot representations overly similar and reduce their discriminability. Therefore, explicitly characterizing the structural differences between interior and boundary spots is important for spatial representation learning.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

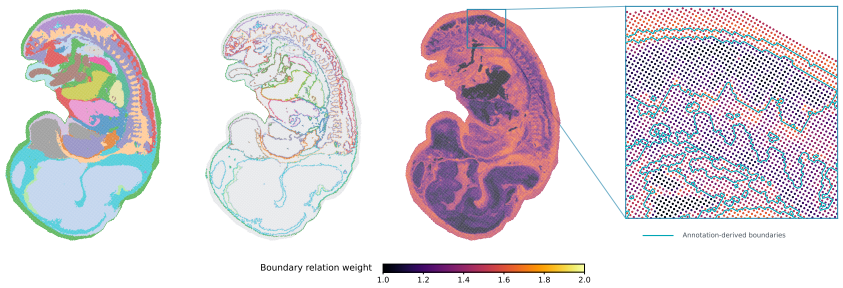


Figure 1: Tissue boundaries and boundary-aware weights in the E12.5 mouse embryo. From left to right are tissue annotations, reference boundary spots, fused boundary weights, and an enlarged view of the selected region. Different colors indicate different tissue categories. Brighter colors indicate higher boundary weights, and cyan lines denote reference interfaces derived from tissue annotations.

As shown in Fig. 1, we observe that interior and boundary spots exhibit distinct relational patterns in both local neighborhoods and global structures. Local expression differences capture fine-grained neighborhood changes but can be sensitive to noise (Walker et al., 2022; Shang & Zhou, 2022; Zhang et al., 2026). Global semantic information provides broader relational cues but may miss subtle local variations (Nonchev et al., 2026; Mason et al., 2023; Long et al., 2026; Cheng et al., 2026; Zhao & Min, 2025). Based on this observation, we design a boundary-aware mechanism to identify and exploit these differences. It combines local expression differences with global semantic information modeled by optimal transport to estimate a continuous boundary score for each spot. The estimated boundary score is then used to adaptively guide representation learning by assigning different learning constraints to spots with different boundary characteristics. To implement this adaptive learning strategy, we construct two augmented graph views and impose boundary-aware constraints within and across the views. Specifically, spots with stronger boundary characteristics place greater emphasis on cross-view semantic relation consistency, which helps preserve stable semantic structures under heterogeneous neighborhoods. In contrast, interior spots receive stronger spatial smoothing and local relation consistency constraints to exploit their relatively homogeneous neighborhoods and maintain intra-region coherence.

We further extend relational modeling from local spots to region and slice levels. Region-level relations provide higher-level semantic information shared by groups of spots. Slice-level relations preserve the overall structural information of each slice. Region-level relations provide higher-level semantic references that help stabilize the representations of spots belonging to similar tissue regions. Slice-level relations preserve the overall structural information of each slice and provide higher-level guidance for multi-slice representation learning. To realize these two levels of modeling, we use multi-granularity region prototypes and slice-level consistency constraints. This multi-level strategy improves representation consistency across slices and supports both single-slice analysis and multi-slice integration.

In summary, the main contributions of this work are as follows:

- We design a boundary-aware spatial representation learning mechanism that combines local expression differences with global semantic information for boundary estimation. The estimated boundary scores are used to adaptively guide representation learning. Interior spots are encouraged to maintain spatial coherence. Boundary-related spots are encouraged to preserve discriminative semantic relations.
- We develop a multi-level relational learning framework that models relational information at the local, region, and slice levels. Multi-granularity region prototypes and slice-level consistency provide complementary structural information, while shared prototype supervision and similarity scale coordination further support joint representation learning across multiple slices.
- We conduct spatial domain identification and multi-slice integration experiments on several SRT datasets. Ablation studies are also performed to analyze the effects of different learning constraints and verify the effectiveness of BMRL.

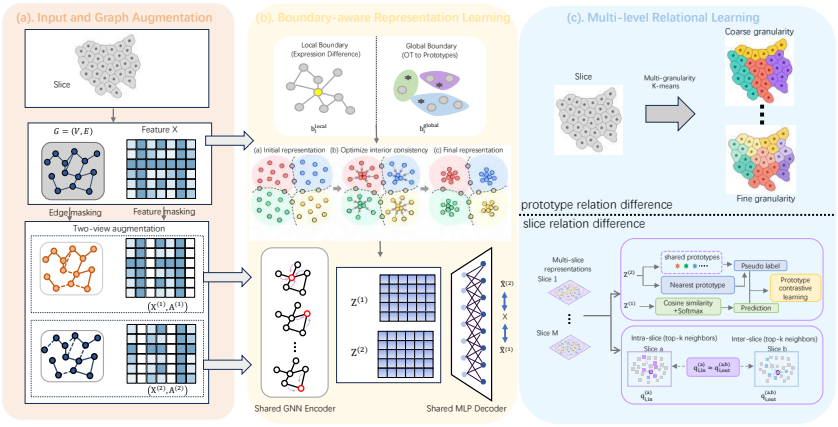


Figure 2: Overall framework of BMRL. The model first constructs two augmented graph views and performs boundary-aware representation learning by combining local expression differences with global prototype relations. It then models relations at the region and slice levels to capture multi-level spatial structures and support multi-slice integration.

2 Related Work

SRT representation learning aims to integrate gene expression and spatial information into low-dimensional representations, which can support downstream tasks such as spatial domain identification (SDI) (Zahedi et al., 2024). Recently, graph-based deep learning methods have obtained promising performance by treating spots as nodes and modeling tissue structures through spatial adjacency relations. SEDR jointly models gene expression and spatial information using an expression autoencoder and a variational graph autoencoder (Xu et al., 2024). SpaGCN integrates gene expression, spatial locations, and histology information through graph convolution, while STAGATE uses graph attention to adaptively aggregate neighboring spots (Hu et al., 2021; Dong & Zhang, 2022). SpaceFlow and GraphST learn spatially informative representations through graph-based self-supervised learning, while SpaCAE combines graph representation learning with contrastive learning (Ren et al., 2022; Long et al., 2023; Hu et al., 2024).

Joint analysis of multi-slice SRT data needs to address spatial misalignment and batch differences across slices (Liu & Yang, 2024). PASTE and PASTE2 use optimal transport to align spots across slices, with PASTE2 further supporting partial tissue overlap (Zeira et al., 2022; Liu et al., 2023b). STAligner learns batch-corrected representations through cross-slice matching, while GraphST extends graph self-supervised learning to multi-slice integration (Zhou et al., 2023; Long et al., 2023). SLAT and CAST further perform cross-slice alignment through graph-based matching and spatial registration, respectively (Xia et al., 2023; Tang et al., 2024). Spotscape introduces prototype contrastive learning and similarity scale matching to establish shared semantic references across slices and reduce batch effects (Oh et al., 2025). Our approach explicitly models the structural differences between interior and boundary regions. It further captures relational information at the spot, region, and slice levels to support both single-slice and multi-slice SRT analysis.

3 METHOD

This section introduces our proposed Boundary-aware Multi-level Relational Learning (BMRL). First, BMRL learns spot representations from the gene expression matrix and spatial graph. Then, boundary information is estimated from local expression variations and global semantic structures to guide representation learning. Finally, relational information is modeled at the spot, region, and slice levels to capture tissue structures at different semantic levels. The overall framework is shown in Figure 2.

3.1 Spatial Graph Representation Learning

Given spatial transcriptomics data containing N spots, the gene expression matrix is denoted as $X \in \mathbb{R}^{N\times F}$, where F is the number of gene features. We construct a spatial nearest neighbor graph based on the distances between spot coordinates. The spatial graph is represented as $G = (X, A)$, where $A \in \mathbb{R}^{N\times N}$ is the adjacency matrix. Specifically, $A_{ij} = 1$ if spots i and j are connected in the SNN graph while $A_{ij} = 0$ otherwise.

We generate two augmented graph views by randomly masking node features and spatial edges. The two augmented views are denoted as $G^{(1)} = (X^{(1)}, A^{(1)})$ and $G^{(2)} = (X^{(2)}, A^{(2)})$. A shared GNN encoder f_θ is used to encode the two augmented views:

$$Z^{(1)} = f_\theta\left(X^{(1)}, A^{(1)}\right), \qquad Z^{(2)} = f_\theta\left(X^{(2)}, A^{(2)}\right), \tag{1}$$

where $Z^{(1)}, Z^{(2)} \in \mathbb{R}^{N\times D}$, and D is the embedding dimension. To preserve gene expression information in the learned representations, we use a shared MLP decoder g_ϕ to reconstruct the original expression matrix X, and define the following reconstruction loss

$$\hat{X}^{(1)} = g_\phi(Z^{(1)}), \hat{X}^{(2)} = g_\phi(Z^{(2)}), \quad \mathcal{L}_{\mathrm{RE}} = \frac{1}{2}\left[\mathrm{MSE}(X, \hat{X}^{(1)}) + \mathrm{MSE}(X, \hat{X}^{(2)})\right]. \tag{2}$$

3.2 Boundary-aware Representation Learning

To distinguish spots inside spatial domains from boundary spots, we estimate boundary information at two levels: the local neighborhood level and the global semantic level. The local neighborhood level uses expression differences between a spot and its spatial neighbors. Formally, the local boundary score of spot i is defined as the average cosine distance between the spot and its spatial neighbors:

$$b_i^{\mathrm{local}} = \frac{1}{|\mathcal{N}(i)|}\sum_{j\in\mathcal{N}(i)}\left(1 - \frac{x_i^\top x_j}{\|x_i\|_2\|x_j\|_2}\right), \tag{3}$$

where $\mathcal{N}(i)$ denotes the spatial neighbors of spot i.

Local information alone may be insufficient to describe the broader semantic relations among spots. We therefore introduce optimal transport to model global semantic structure. Optimal transport provides a principled way to match two distributions by jointly considering all pairwise costs under global marginal constraints. In our setting, let $C \in \mathbb{R}^{N\times K}$ denote the assignment cost matrix between N spots and K region prototypes. We use K-means to initialize K region prototypes $\mu_k{}_{k=1}^{K}$. The assignment cost between spot i and prototype k is defined as

$$\bar{x}_i = \frac{x_i}{\|x_i\|_2}, \qquad C_{ik} = \|\bar{x}_i - \mu_k\|_2^2. \tag{4}$$

We use entropy-regularized optimal transport and solve it with the Sinkhorn algorithm (Cuturi, 2013) to obtain balanced assignments. Let $T \in \mathbb{R}^{N\times K}$ denote the transport plan. Each element T_{ik} represents the transport mass from spot i to prototype k and reflects their soft assignment relation. We solve the following optimal transport problem

$$\begin{aligned} T^* = \arg\min_{T\geq 0} \sum_{i=1}^{N}\sum_{k=1}^{K} T_{ik}C_{ik} + \epsilon\sum_{i=1}^{N}\sum_{k=1}^{K} T_{ik}\left(\log T_{ik} - 1\right), \\ \text{s.t.} \quad T\mathbf{1}_K = \frac{1}{N}\mathbf{1}_N, \qquad T^\top\mathbf{1}_N = \frac{1}{K}\mathbf{1}_K. \end{aligned} \tag{5}$$

Here, T^* denotes the optimal transport plan and ϵ controls the strength of entropy regularization. After obtaining the optimal transport solution, we convert the transport plan into a soft semantic assignment for each spot: $q_i = NT^*_{i,:}$. Based on these soft assignments, we extract global semantic cues for boundary estimation from two complementary aspects. We measure the assignment uncertainty using normalized entropy:

$$u_i = -\frac{1}{\log K}\sum_{k=1}^{K} q_{ik}\log(q_{ik}). \tag{6}$$

A larger u_i indicates a more uncertain semantic assignment. We further measure the discrepancy between the global semantic assignment of a spot and its local spatial context using the Jensen–Shannon divergence (Yang et al., 2024).

$$d_i = \mathrm{JS}\left(\mathbf{q}_i, \frac{1}{|\mathcal{N}(i)|}\sum_{j\in\mathcal{N}(i)}\mathbf{q}_j\right) \Big/ \max_m \mathrm{JS}\left(\mathbf{q}_m, \frac{1}{|\mathcal{N}(m)|}\sum_{j\in\mathcal{N}(m)}\mathbf{q}_j\right). \tag{7}$$

Here, $\mathcal{N}(i)$ is the set of spatial neighbors of spot i. $\mathrm{JS}(\cdot,\cdot)$ denotes the Jensen–Shannon divergence between two semantic assignment distributions. The index m runs over all spots and is used to obtain the maximum divergence for normalization. A larger d_i means that the semantic assignment of the spot is more different from its surrounding neighborhood. Therefore, the spot is more likely to be located in a potential semantic transition region. Finally, the global boundary score is defined as

$$b_i^{\mathrm{global}} = \frac{1}{2}\left(u_i + d_i\right). \tag{8}$$

It combines semantic assignment uncertainty and semantic context inconsistency to describe global boundary information. Then, we combine b_i^{global} and the local boundary score b_i^{local} by Rank fusion (Wang et al., 2024). Specifically, the two scores are first converted into normalized ranks to reduce the influence of different numerical scales. The ranked scores are then combined to obtain the final continuous boundary score $b_i \in [0,1]$. A larger b_i means that the spot has stronger boundary characteristics.

Using the boundary scores, we impose region-specific constraints on the spot representations. Boundary spots may aggregate heterogeneous neighbors, whereas interior spots benefit from local smoothness. We therefore define the boundary relation weight and interior spatial weight as

$$w_i^B = 1 + \alpha b_i, \qquad w_{ij}^I = \left(1 - b_i^{\mathrm{local}}\right)\left(1 - b_j^{\mathrm{local}}\right). \tag{9}$$

where $\alpha \geq 0$ is the boundary enhancement coefficient. w_i^B increases the contribution of boundary-related spots, while w_{ij}^I strengthens local consistency between interior spots.

To better preserve the semantic relations around boundary regions, we use w_i^B to reweight the relation consistency across the two augmented views. This encourages the relative semantic relations involving boundary-related spots to remain stable under random augmentations. Specifically, we model cross-view relations between spots. The boundary-aware relation consistency loss is defined as

$$S_{ij} = \cos\left(\mathbf{z}_i^{(1)}, \mathbf{z}_j^{(2)}\right), \qquad \mathcal{L}_{\mathrm{BRC}} = \frac{\sum_{i=1}^N \sum_{j=1}^N w_i^B w_j^B \left(S_{ij} - S_{ji}\right)^2}{\sum_{i=1}^N \sum_{j=1}^N w_i^B w_j^B + \epsilon}. \tag{10}$$

Larger $w_i^B w_j^B$ gives stronger constraints to unstable relations involving boundary-related spots.

For interior spots, their spatial neighborhoods are relatively homogeneous and can provide reliable local structural information. We encourage spatial smoothing within each augmented view to maintain local continuity among reliable neighboring spots. We also preserve local relation consistency across the two views to stabilize neighborhood structures under graph augmentation. For spatial smoothing, we adaptively adjust the contribution of each spatial edge according to the boundary characteristics of its connected spots.

$$\mathcal{L}_{\mathrm{IS}} = \frac{1}{2}\sum_{t=1}^{2}\left(\left[\sum_{(i,j)\in E} w_{ij}^I \left\|\mathbf{z}_i^{(t)} - \mathbf{z}_j^{(t)}\right\|_2^2\right] \Big/ \left[\sum_{(i,j)\in E} w_{ij}^I\right]\right). \tag{11}$$

Here, $t \in \{1,2\}$ denotes the two augmented representations. This objective encourages smooth representations along reliable spatial edges while reducing over-smoothing around potential boundaries. We further preserve local relation consistency across the two augmentations by comparing the cosine similarities of spatial edges:

$$\Delta s_{ij} = \cos\left(z_i^{(1)}, z_j^{(1)}\right) - \cos\left(z_i^{(2)}, z_j^{(2)}\right), \mathcal{L}_{\mathrm{local}} = \frac{\sum_{(i,j)\in E}(1-b_i)(1-b_j)\left(\Delta s_{ij}\right)^2}{\sum_{(i,j)\in E}(1-b_i)(1-b_j)}. \tag{12}$$

A lower boundary score gives a larger weight to the corresponding spatial edge. This helps keep local relations stable inside spatial domains.

3.3 MULTI-LEVEL RELATIONAL LEARNING

The section 3.2 estimates boundary information from expression variations and global semantic structures, and uses it to adaptively guide representation learning at different spatial positions. To further capture tissue organization at different semantic granularities, we model relational information at the spot, region, and slice levels. The spot level describes relations between spatially neighboring spots. The region level models semantic relations between spots and region prototypes. The slice level captures the overall semantic structure of each tissue slice.

At the region level, we perform multi-granularity K-means on the second augmented representation $\mathbf{Z}^{(2)}$. The index r denotes different granularity levels. At the r-th granularity, K_r denotes the number of region prototypes, which is determined by the base cluster number K and the granularity level r. These prototypes are used as shared region-level semantic references for the two augmented representations. For spot i, the prototype relation difference at the r-th granularity is defined as

$$\delta_i^{(r)} = \frac{1}{K_r} \sum_{k=1}^{K_r} \left[\cos\left(\mathbf{z}_i^{(1)}, \boldsymbol{\mu}_k^{(r)}\right) - \cos\left(\mathbf{z}_i^{(2)}, \boldsymbol{\mu}_k^{(r)}\right) \right]^2. \tag{13}$$

A smaller $\delta_i^{(r)}$ means that the semantic relations between the spot and the region prototypes are more stable across the two augmented representations. The region-level relation consistency loss is defined as

$$\mathcal{L}_{\text{proto}} = \frac{1}{R} \sum_{r=1}^{R} \left(\left[\sum_{i=1}^{N} w_i^B \delta_i^{(r)} \right] \Big/ \left[\sum_{i=1}^{N} w_i^B \right] \right). \tag{14}$$

Here, R is the number of prototype granularities and w_i^B is the boundary-adaptive weight defined in Section 3.2. This objective keeps the semantic relations between spots and region prototypes consistent across the two augmented representations.

To further keep the overall tissue representation stable under random changes, we apply average pooling to the spot representations in each slice. We then keep the overall representations of the same slice consistent across the two augmentations:

$$\mathcal{L}_{\text{slice}} = \frac{1}{M} \sum_{m=1}^{M} \left[1 - \cos\left(\text{AvgPool}\left(\{\mathbf{z}_i^{(1)}\}_{i \in \mathcal{V}_m} \right), \text{AvgPool}\left(\{\mathbf{z}_i^{(2)}\}_{i \in \mathcal{V}_m} \right) \right) \right]. \tag{15}$$

where $\mathcal{V}_m$ denotes the set of spots in the m-th slice, and $\text{AvgPool}(\cdot)$ denotes average pooling over all spot representations in the corresponding slice.

We further use the shared region prototypes to build semantic relations between different slices. Specifically, the prototype assignments obtained from the stop-gradient second augmented representation $\mathbf{Z}^{(2)}$ are used as pseudo-labels. The first augmented representation $\mathbf{Z}^{(1)}$ is then used to predict the corresponding prototype probabilities:

$$y_i^{(r)} = \arg\min_k \left\| \mathbf{z}_i^{(2)} - \boldsymbol{\mu}_k^{(r)} \right\|_2^2, \quad p_{ik}^{(r)} = \text{Softmax}_k\left(\cos\left(\mathbf{z}_i^{(1)}, \boldsymbol{\mu}_k^{(r)}\right) / \tau \right). \tag{16}$$

Here, $y_i^{(r)}$ is the prototype pseudo-label of spot i at the r-th granularity. $p_{ik}^{(r)}$ is the predicted probability that spot i belongs to the k-th prototype. τ is the temperature parameter. Then, we use $y_i^{(r)}$ as the supervision signal and apply a cross-entropy loss to the prototype prediction. The multi-scale prototype contrastive loss is defined as

$$\mathcal{L}_{\text{PCL}} = -\frac{1}{R} \sum_{r=1}^{R} \frac{1}{N} \sum_{i=1}^{N} \log p_{i, y_i^{(r)}}^{(r)}. \tag{17}$$

Here, R is the number of prototype scales. This objective encourages the first augmented representation to be consistent with the prototype assignments obtained from the second augmented representation. The shared prototypes therefore provide common semantic references across slices.

However, technical differences between slices may still make cross-slice similarities systematically lower than intra-slice similarities. To reduce this scale difference, we further apply cross-slice similarity scaling. For spot i in slice a, we calculate the average relation strength of the top-k similar

Table 1: Single-slice spatial domain identification performance on DLPFC, MTG, Mouse Embryo, and NSCLC. Results are reported as mean (standard deviation). The best and second-best results are shown in bold and underlined, respectively.

(a) DLPFC (Patient 1)

	Slice 151673			Slice 151674			Slice 151675			Slice 151676		
Method	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
SEDR	0.36 (0.08)	0.49 (0.08)	0.55 (0.06)	0.37 (0.08)	0.48 (0.07)	0.51 (0.07)	0.33 (0.06)	0.45 (0.05)	0.51 (0.03)	0.29 (0.03)	0.41 (0.04)	0.47 (0.02)
STAGATE	0.37 (0.04)	0.55 (0.03)	0.52 (0.04)	0.34 (0.03)	0.50 (0.02)	0.51 (0.03)	0.33 (0.03)	0.50 (0.03)	0.48 (0.03)	0.33 (0.00)	0.47 (0.01)	0.52 (0.01)
SpaCAE	0.21 (0.01)	0.37 (0.01)	0.43 (0.01)	0.25 (0.03)	0.38 (0.01)	0.44 (0.03)	0.23 (0.03)	0.41 (0.03)	0.42 (0.04)	0.23 (0.02)	0.34 (0.02)	0.43 (0.03)
SpaceFlow	0.42 (0.06)	0.57 (0.05)	0.57 (0.03)	0.37 (0.04)	0.51 (0.03)	0.53 (0.03)	0.38 (0.07)	0.55 (0.06)	0.53 (0.05)	0.38 (0.05)	0.51 (0.05)	0.53 (0.04)
GraphST	0.20 (0.02)	0.34 (0.03)	0.41 (0.02)	0.27 (0.02)	0.41 (0.01)	0.46 (0.01)	0.22 (0.02)	0.34 (0.01)	0.40 (0.02)	0.26 (0.05)	0.40 (0.05)	0.45 (0.04)
BMRL	**0.53** (0.02)	**0.68** (0.02)	**0.64** (0.01)	**0.52** (0.02)	**0.65** (0.01)	**0.64** (0.02)	**0.53** (0.02)	**0.66** (0.01)	**0.66** (0.02)	**0.45** (0.04)	**0.61** (0.03)	**0.62** (0.03)

DLPFC (Patient 2)

	Slice 151507			Slice 151508			Slice 151509			Slice 151510		
Method	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
SEDR	0.29 (0.06)	0.39 (0.07)	0.45 (0.06)	0.21 (0.02)	0.31 (0.02)	0.39 (0.02)	0.37 (0.04)	0.47 (0.04)	0.51 (0.05)	0.31 (0.05)	0.44 (0.04)	0.47 (0.04)
STAGATE	0.41 (0.01)	0.53 (0.01)	0.59 (0.00)	0.32 (0.01)	0.49 (0.00)	0.54 (0.01)	0.41 (0.02)	0.57 (0.02)	0.61 (0.04)	0.32 (0.03)	0.50 (0.02)	0.50 (0.02)
SpaCAE	0.28 (0.06)	0.41 (0.06)	0.46 (0.06)	0.20 (0.04)	0.31 (0.05)	0.40 (0.04)	0.31 (0.01)	0.44 (0.02)	0.50 (0.04)	0.27 (0.02)	0.42 (0.03)	0.45 (0.02)
SpaceFlow	0.55 (0.03)	0.68 (0.02)	0.71 (0.05)	0.44 (0.04)	0.57 (0.03)	0.58 (0.04)	0.53 (0.05)	0.66 (0.02)	0.65 (0.04)	0.50 (0.03)	0.64 (0.01)	0.61 (0.02)
GraphST	0.31 (0.01)	0.45 (0.01)	0.50 (0.01)	0.34 (0.01)	0.45 (0.02)	0.53 (0.02)	0.35 (0.01)	0.51 (0.01)	0.55 (0.02)	0.30 (0.02)	0.47 (0.01)	0.49 (0.03)
BMRL	**0.63** (0.04)	**0.72** (0.03)	**0.79** (0.03)	**0.53** (0.02)	**0.67** (0.01)	**0.67** (0.02)	**0.59** (0.01)	**0.72** (0.01)	**0.70** (0.01)	**0.55** (0.03)	**0.68** (0.01)	**0.66** (0.01)

DLPFC (Patient 3)

	Slice 151669			Slice 151670			Slice 151671			Slice 151672		
Method	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
SEDR	0.24 (0.07)	0.40 (0.07)	0.48 (0.06)	0.24 (0.06)	0.39 (0.05)	0.48 (0.05)	0.37 (0.10)	0.50 (0.09)	0.59 (0.07)	0.49 (0.09)	0.58 (0.06)	0.66 (0.07)
STAGATE	0.29 (0.05)	0.45 (0.07)	0.52 (0.04)	0.20 (0.01)	0.38 (0.01)	0.44 (0.01)	0.40 (0.07)	0.49 (0.03)	0.63 (0.06)	0.38 (0.02)	0.51 (0.04)	0.54 (0.01)
SpaCAE	0.21 (0.02)	0.28 (0.03)	0.43 (0.02)	0.21 (0.03)	0.28 (0.02)	0.43 (0.04)	0.38 (0.16)	0.29 (0.01)	0.49 (0.05)	0.25 (0.04)	0.35 (0.05)	0.50 (0.01)
SpaceFlow	0.30 (0.07)	0.48 (0.03)	0.51 (0.05)	0.34 (0.05)	0.50 (0.03)	0.56 (0.05)	0.54 (0.04)	0.67 (0.02)	0.67 (0.04)	0.60 (0.06)	0.70 (0.02)	0.73 (0.06)
GraphST	0.17 (0.04)	0.26 (0.04)	0.43 (0.02)	0.14 (0.01)	0.23 (0.00)	0.37 (0.01)	0.30 (0.05)	0.38 (0.03)	0.54 (0.03)	0.23 (0.01)	0.32 (0.02)	0.49 (0.01)
BMRL	**0.51** (0.01)	**0.62** (0.01)	**0.68** (0.01)	**0.50** (0.02)	**0.60** (0.01)	**0.68** (0.01)	**0.76** (0.03)	**0.76** (0.01)	**0.85** (0.03)	**0.80** (0.03)	**0.76** (0.02)	**0.87** (0.03)

	(b) MTG Control			**(b) MTG AD**			**(c) Mouse Embryo**			**(d) NSCLC**		
Method	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
SEDR	0.41 (0.02)	0.59 (0.02)	0.52 (0.02)	0.43 (0.08)	0.59 (0.07)	0.57 (0.07)	0.32 (0.02)	0.56 (0.01)	0.42 (0.02)	0.44 (0.08)	0.46 (0.06)	0.70 (0.08)
STAGATE	0.54 (0.00)	0.65 (0.00)	0.59 (0.00)	0.51 (0.01)	0.61 (0.01)	0.59 (0.01)	0.36 (0.01)	0.60 (0.01)	0.47 (0.01)	0.35 (0.05)	0.41 (0.04)	0.64 (0.02)
SpaCAE	0.37 (0.03)	0.52 (0.00)	0.44 (0.03)	0.22 (0.01)	0.40 (0.01)	0.40 (0.01)	0.34 (0.01)	0.60 (0.01)	0.48 (0.02)	0.32 (0.05)	0.38 (0.03)	0.62 (0.02)
SpaceFlow	0.66 (0.03)	0.74 (0.01)	0.70 (0.03)	0.54 (0.01)	0.71 (0.00)	0.65 (0.01)	0.42 (0.03)	0.60 (0.02)	0.49 (0.03)	0.53 (0.03)	0.52 (0.02)	0.75 (0.02)
GraphST	0.38 (0.00)	0.51 (0.00)	0.48 (0.00)	0.43 (0.06)	0.55 (0.05)	0.55 (0.04)	0.34 (0.01)	0.59 (0.02)	0.45 (0.01)	0.30 (0.00)	0.38 (0.00)	0.65 (0.00)
BMRL	**0.78** (0.04)	**0.80** (0.02)	**0.80** (0.04)	**0.69** (0.01)	**0.76** (0.01)	**0.83** (0.01)	**0.52** (0.01)	**0.67** (0.00)	**0.60** (0.01)	**0.62** (0.01)	**0.58** (0.01)	**0.79** (0.01)

spots in the same slice and in a target slice b:

$$q_{i,\text{in}}^{(a)} = \text{TopKMean}\left(\{S_{ij} \mid s_j = a\}, k\right), q_{i,\text{out}}^{(a,b)} = \text{TopKMean}\left(\{S_{ij} \mid s_j = b\}, k\right), \quad a \neq b. \tag{18}$$

Here, S_{ij} denotes the cross-augmentation similarity between spots i and j defined in Section 3.2, and s_j denotes the slice of spot j. $q_{i,\text{in}}^{(a)}$ measures the average relation strength between spot i and highly similar spots in the same slice. $q_{i,\text{out}}^{(a,b)}$ measures the corresponding relation strength between spot i and highly similar spots in slice b. The cross-slice similarity scaling loss is defined as

$$\mathcal{L}_{\text{SS}} = \frac{1}{|\mathcal{P}|} \sum_{(i,a,b)\in\mathcal{P}} \left(q_{i,\text{in}}^{(a)} - q_{i,\text{out}}^{(a,b)}\right)^2, \tag{19}$$

where $\mathcal{P}$ denotes the set of valid spot–slice comparison pairs. This objective encourages highly similar relations within and across slices to have a consistent scale.

4 Experiments

4.1 Experimental Settings

We evaluate BMRL on four spatial transcriptomics datasets, including DLPFC, MTG, Mouse Embryo, and NSCLC. We consider three tasks: single-slice spatial domain identification, multi-slice integration, and multi-slice alignment. Clustering performance is evaluated using ARI, NMI, and CA. For multi-slice integration, we additionally use batch silhouette score, kBET, graph connectivity, and PCR comparison to assess batch correction. LTARI is used for multi-slice alignment. We compare BMRL with representative spatial transcriptomics methods under the same evaluation protocol. Detailed dataset descriptions, baselines, evaluation protocols, and implementation settings are provided in Appendix A.

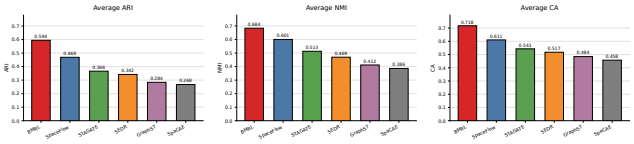


Figure 3: Average single-slice spatial domain identification performance across all datasets.

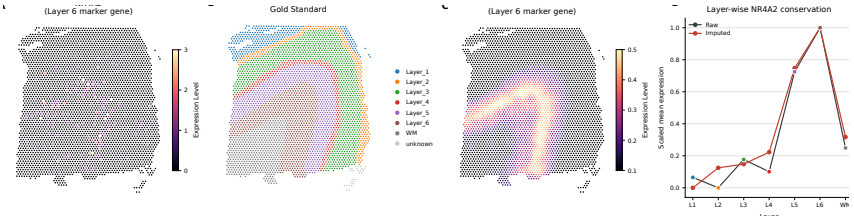


Figure 4: Gene expression reconstruction of the Layer 6 marker NR4A2 on DLPFC slice 151673.

4.2 Single-Slice Experimental Results

Table 1 reports the single-slice spatial domain identification results on DLPFC, MTG, Mouse Embryo, and NSCLC. We make the following observations: BMRL achieves the best or tied-best results in ARI, NMI, and CA across the 16 slice or sample settings. Fig. 3 further summarizes the average performance across all single-slice experiments, showing strong overall performance on all three metrics. On the 12 DLPFC slices, BMRL outperforms the compared methods on most metrics, showing its ability to preserve cortical spatial structures. BMRL also performs well on MTG, Mouse Embryo, and NSCLC, demonstrating consistent performance across different tissue types and sequencing platforms.

To evaluate whether BMRL can reconstruct clearer spatial gene expression patterns while preserving layer-specific enrichment, we analyze the reconstructed expression of the Layer 6 marker gene NR4A2 in DLPFC slice 151673. As shown in Fig. 4(A), the expression of NR4A2 is sparse in the raw data, and its spatial pattern is not clear. Fig. 4(B) shows the ground-truth cortical layer labels. After reconstruction by BMRL, NR4A2 shows a more continuous and clearer spatial expression pattern, as shown in Fig. 4(C). Furthermore, Fig. 4(D) compares the mean expression of NR4A2 across cortical layers and white matter. Both the raw and reconstructed expression reach the highest level in Layer 6. These results show that BMRL produces a clearer spatial expression pattern of NR4A2 while preserving its enrichment in Layer 6.

4.3 Multi-Slice Experimental Results

Homogeneous Integration. We first evaluate homogeneous multi-slice integration on the DLPFC dataset. For each patient, we jointly integrate the four tissue slices and use ARI, NMI, and CA to evaluate the spatial domain identification results after integration. Table 2 reports the homogeneous integration results of different methods on three DLPFC patients. BMRL achieves the best or competitive performance in ARI, NMI, and CA across the three patients, demonstrating its ability to integrate homogeneous slices while preserving spatial domain structure.

Heterogeneous Integration. We further evaluate heterogeneous multi-slice integration on the MTG dataset, which contains Control and AD samples with different biological states. In addition to ARI, NMI, and CA, we use Silhouette Batch, kBET, Graph Connectivity, and PCR comparison to evaluate batch effect correction. Table 4 reports the heterogeneous integration results of different methods on the MTG dataset. The results show that BMRL can integrate slices from different biological conditions while preserving spatial domain structures and domain connectivity. Although there are variations across different batch correction metrics, BMRL achieves a good balance between clustering quality and multi-slice integration performance.

Multi-slice Alignment. Finally, we evaluate the spatial alignment performance between slices from different developmental stages on the Mouse Embryo dataset. LTARI is used to measure the consis-

Table 2: Homogeneous multi-slice integration performance on DLPFC. Results are reported as mean (standard deviation). The best and second-best results are shown in bold and underlined, respectively.

Method	Patient 1			Patient 2			Patient 3		
	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
SEDR	0.38 (0.06)	0.49 (0.06)	0.56 (0.06)	0.32 (0.05)	0.44 (0.07)	0.48 (0.07)	0.43 (0.02)	0.51 (0.01)	0.56 (0.03)
STAGATE	0.31 (0.03)	0.46 (0.03)	0.49 (0.03)	0.30 (0.02)	0.46 (0.01)	0.48 (0.02)	0.31 (0.09)	0.43 (0.06)	0.54 (0.08)
SpaCAE	0.21 (0.03)	0.36 (0.02)	0.40 (0.02)	0.12 (0.06)	0.19 (0.07)	0.32 (0.05)	0.13 (0.05)	0.14 (0.05)	0.43 (0.06)
SpaceFlow	0.48 (0.03)	0.60 (0.02)	0.60 (0.02)	0.44 (0.05)	0.59 (0.02)	0.58 (0.04)	0.51 (0.02)	0.60 (0.01)	0.69 (0.05)
GraphST	0.18 (0.01)	0.32 (0.01)	0.38 (0.02)	0.25 (0.01)	0.39 (0.01)	0.42 (0.02)	0.25 (0.04)	0.30 (0.04)	0.50 (0.01)
PASTE	0.34 (0.00)	0.45 (0.00)	0.54 (0.00)	0.17 (0.00)	0.28 (0.00)	0.40 (0.00)	0.29 (0.00)	0.43 (0.00)	0.54 (0.00)
STAligner	0.38 (0.04)	0.52 (0.04)	0.55 (0.04)	0.29 (0.02)	0.45 (0.02)	0.48 (0.03)	0.37 (0.06)	0.47 (0.05)	0.59 (0.06)
CAST	0.26 (0.02)	0.37 (0.03)	0.42 (0.03)	0.30 (0.04)	0.43 (0.05)	0.47 (0.03)	0.38 (0.06)	0.40 (0.04)	0.56 (0.05)
Spotscape	0.57 (0.03)	0.70 (0.02)	0.67 (0.03)	0.53 (0.02)	**0.67** (0.01)	0.63 (0.02)	0.63 (0.09)	0.68 (0.03)	0.75 (0.09)
BMRL	**0.58** (0.02)	**0.71** (0.02)	**0.69** (0.02)	**0.56** (0.04)	**0.67** (0.01)	**0.73** (0.05)	**0.74** (0.05)	**0.69** (0.02)	**0.79** (0.05)

Table 3: Multi-slice spatial alignment performance on Mouse Embryo.

Method	LTARI
PASTE2	0.21 (0.02)
CAST	0.10 (0.00)
STAligner	0.46 (0.01)
SLAT	0.41 (0.11)
Spotscape	0.51 (0.01)
BMRL	**0.56** (0.01)

Table 4: Heterogeneous multi-slice integration performance on MTG across control and AD samples. Results are reported as mean (standard deviation). The best and second-best results are shown in bold and underlined, respectively.

Method	Clustering Metric			Batch Effect Correction Metric			
	ARI	NMI	CA	Silhouette batch	kBET	Graph connectivity	PCR comparison
GraphST	0.23 (0.02)	0.42 (0.00)	0.39 (0.01)	0.56 (0.00)	0.02 (0.01)	0.65 (0.02)	0.00 (0.00)
STAligner	0.38 (0.03)	0.54 (0.03)	0.49 (0.02)	0.62 (0.04)	**0.11** (0.08)	0.85 (0.04)	0.18 (0.10)
CAST	0.48 (0.07)	0.52 (0.06)	0.59 (0.06)	0.45 (0.02)	**0.11** (0.02)	0.81 (0.06)	**0.97** (0.03)
Spotscape	0.72 (0.04)	0.76 (0.01)	0.81 (0.05)	0.69 (0.01)	0.08 (0.02)	0.86 (0.03)	0.60 (0.08)
BMRL	**0.77** (0.01)	**0.78** (0.00)	**0.85** (0.01)	**0.73** (0.01)	0.08 (0.02)	**0.99** (0.01)	0.52 (0.06)

tency of the aligned labels. Table 3 reports the alignment results of different methods on the Mouse Embryo dataset. BMRL achieves the highest LTARI of 0.56. This result shows that the learned representations can better support the matching of tissue slices from different developmental stages.

4.4 Model Analysis

We conduct ablation studies on the two main components of BMRL: Boundary-aware Representation Learning and Multi-level Relational Learning. For Boundary-aware Representation Learning, removing $\mathcal{L}_{\mathrm{BRC}}$ causes the largest performance drop, while removing the boundary weighting, $\mathcal{L}_{\mathrm{IS}}$, or $\mathcal{L}_{\mathrm{local}}$ also degrades the performance. These results show that boundary-aware relation consistency, intra-domain smoothing, and local relation stability all contribute to spatial domain identification. For Multi-level Relational Learning, removing $\mathcal{L}_{\mathrm{SS}}$ leads to the largest performance decrease, while $\mathcal{L}_{\mathrm{proto}}$, $\mathcal{L}_{\mathrm{slice}}$, and $\mathcal{L}_{\mathrm{PCL}}$ also provide complementary improvements. These results verify the effectiveness of region- and slice-level relation modeling for multi-slice integration. Detailed results are provided in Appendix B.

We further analyze the effects of local and global information, the learned representations, and the sensitivity of BMRL to key hyperparameters. The local-global analysis shows that the two types of information provide complementary cues for boundary estimation. The similarity analysis shows that BMRL captures meaningful domain-level relational structures, while the sensitivity analysis demonstrates stable performance across different hyperparameter settings. Detailed results are provided in Appendix C.

5 Conclusion

In this work, we propose Boundary-aware Multi-level Relational Learning (BMRL) for SRT data analysis. BMRL considers the different structural characteristics of interior and boundary spots. It combines local expression differences with global relation information to estimate continuous boundary scores. The estimated boundary information is then used to guide spot representation learning. BMRL further models relations at the spot, region, and slice levels. This multi-level design captures complementary structural information and supports both single-slice and multi-slice analysis. Extensive experiments on multiple SRT datasets show that BMRL achieves strong performance on spatial domain identification, multi-slice integration, and spatial alignment. Ablation studies and model analyses further demonstrate the effectiveness of the proposed design.

AI USE STATEMENT

In this work, we used generative AI tools to improve the readability and language of the manuscript and to assist with literature retrieval and discovery. In particular, these tools were used to polish academic writing, search for relevant literature, and identify potentially related references. All AI-assisted text was reviewed and revised by the authors, and all suggested references were manually checked against their original sources. We have reviewed all AI-assisted work. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

Dharmesh D Bhuva, Chin Wee Tan, Agus Salim, Claire Marceaux, Marie A Pickering, Jinjin Chen, Malvika Kharbanda, Xinyi Jin, Ning Liu, Kristen Feher, et al. Library size confounds biology in spatial transcriptomics data. *Genome Biology*, 25(1):99, 2024.

Ao Chen, Sha Liao, Mengnan Cheng, Kailong Ma, Liang Wu, Yiwei Lai, Xiaojie Qiu, Jin Yang, Jiangshan Xu, Shijie Hao, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using dna nanoball-patterned arrays. *Cell*, 185(10):1777–1792, 2022a.

Shuo Chen, Yuzhou Chang, Liangping Li, Diana Acosta, Yang Li, Qi Guo, Cankun Wang, Emir Turkes, Cody Morrison, Dominic Julian, et al. Spatially resolved transcriptomics reveals genes associated with the vulnerability of middle temporal gyrus in alzheimer's disease. *Acta neuropathologica communications*, 10(1):188, 2022b.

Jinming Cheng, Pierce Kah Hoe Chow, and Nan Liu. Spneigh: spatial neighborhood and differential expression analysis for high-resolution spatial transcriptomics. *NAR Genomics and Bioinformatics*, 8(2):lqag039, 2026.

Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems*, 26, 2013.

Kangning Dong and Shihua Zhang. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nature communications*, 13(1):1739, 2022.

Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30, 2017.

Jian Hu, Xiangjie Li, Kyle Coleman, Amelia Schroeder, Nan Ma, David J Irwin, Edward B Lee, Russell T Shinohara, and Mingyao Li. Spagcn: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nature methods*, 18(11):1342–1351, 2021.

Yaofeng Hu, Kai Xiao, Hengyu Yang, Xiaoping Liu, Chuanchao Zhang, and Qianqian Shi. Spatially contrastive variational autoencoder for deciphering tissue heterogeneity from spatially resolved transcriptomics. *Briefings in Bioinformatics*, 25(2):bbae016, 2024.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Jiachen Li, Siheng Chen, Xiaoyong Pan, Ye Yuan, and Hong-Bin Shen. Cell clustering for spatial transcriptomics data with graph neural networks. *Nature Computational Science*, 2(6):399–408, 2022.

Teng Liu, Zhao-Yu Fang, Xin Li, Li-Ning Zhang, Dong-Sheng Cao, and Ming-Zhu Yin. Graph deep learning enabled spatial domains identification for spatial transcriptomics. *Briefings in Bioinformatics*, 24(3):bbad146, 2023a.

Xinhao Liu, Ron Zeira, and Benjamin J Raphael. Partial alignment of multislice spatially resolved transcriptomics data. *Genome research*, 33(7):1124–1132, 2023b.

Yuyao Liu and Can Yang. Computational methods for alignment and integration of spatially resolved transcriptomics data. 2024.

Jiangkai Long, Yanran Zhu, Xiao Zheng, Li Jin, and Chang Tang. A collaborative graph–transformer network for spatial transcriptomics data clustering with semantic graph induction. *Engineering Applications of Artificial Intelligence*, 181:115097, 2026.

Yahui Long, Kok Siong Ang, Mengwei Li, Kian Long Kelvin Chong, Raman Sethi, Chengwei Zhong, Hang Xu, Zhiwei Ong, Karishma Sachaphibulkij, Ao Chen, et al. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with graphst. *Nature communications*, 14(1):1155, 2023.

Sophia K Longo, Margaret G Guo, Andrew L Ji, and Paul A Khavari. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nature Reviews Genetics*, 22 (10):627–644, 2021.

Malte D Luecken, Maren Büttner, Kridsadakorn Chaichoompu, Anna Danese, Marta Interlandi, Michaela F Müller, Daniel C Strobl, Luke Zappia, Martin Dugas, Maria Colomé-Tatché, et al. Benchmarking atlas-level data integration in single-cell genomics. *Nature methods*, 19(1):41–50, 2022.

Kaishu Mason, Anuja Sathe, Paul Hess, Jiazhen Rong, Chi-Yun Wu, Emma Furth, Hanlee P Ji, and Nancy Zhang. Niche differential gene expression analysis in spatial transcriptomics data identifies context-dependent cell-cell interactions. *bioRxiv*, pp. 2023–01, 2023.

Kristen R Maynard, Leonardo Collado-Torres, Lukas M Weber, Cedric Uytingco, Brianna K Barry, Stephen R Williams, Joseph L Catallini, Matthew N Tran, Zachary Besich, Madhavi Tippani, et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nature neuroscience*, 24(3):425–436, 2021.

Kalin Nonchev, Sonali Andani, Joanna Ficek-Pascual, Marta Nowak, Bettina Sobottka, Viktor H Koelzer, and Gunnar Rätsch. Representation learning for multi-modal spatially resolved transcriptomics data. *Bioinformatics*, 42(7):btag316, 2026.

Yunhak Oh, Junseok Lee, Yeongmin Kim, Sangwoo Seo, Namkyeong Lee, and Chanyoung Park. Global context-aware representation learning for spatially resolved transcriptomics. *arXiv preprint arXiv:2506.15698*, 2025.

Anjali Rao, Dalia Barkley, Gustavo S França, and Itai Yanai. Exploring tissue architecture using spatial transcriptomics. *Nature*, 596(7871):211–220, 2021.

Honglei Ren, Benjamin L Walker, Zixuan Cang, and Qing Nie. Identifying multicellular spatiotemporal organization of cells with spaceflow. *Nature communications*, 13(1):4076, 2022.

Samuel G Rodriques, Robert R Stickels, Aleksandrina Goeva, Carly A Martin, Evan Murray, Charles R Vanderburg, Joshua Welch, Linlin M Chen, Fei Chen, and Evan Z Macosko. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, 363(6434):1463–1467, 2019.

Lulu Shang and Xiang Zhou. Spatially aware dimension reduction for spatial transcriptomics. *Nature communications*, 13(1):7203, 2022.

Patrik L Ståhl, Fredrik Salmén, Sanja Vickovic, Anna Lundmark, José Fernández Navarro, Jens Magnusson, Stefania Giacomello, Michaela Asp, Jakub O Westholm, Mikael Huss, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353 (6294):78–82, 2016.

Robert R Stickels, Evan Murray, Pawan Kumar, Jilong Li, Jamie L Marshall, Daniela J Di Bella, Paola Arlotta, Evan Z Macosko, and Fei Chen. Highly sensitive spatial transcriptomics at near-cellular resolution with slide-seqv2. *Nature biotechnology*, 39(3):313–319, 2021.

Zefang Tang, Shuchen Luo, Hu Zeng, Jiahao Huang, Xin Sui, Morgan Wu, and Xiao Wang. Search and match across spatial omics samples at single-cell resolution. *Nature methods*, 21(10):1818–1829, 2024.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Benjamin L Walker, Zixuan Cang, Honglei Ren, Eric Bourgain-Chang, and Qing Nie. Deciphering tissue structure and function using spatial transcriptomics. *Communications biology*, 5(1):220, 2022.

Siyi Wang, Qi Deng, Shiwei Feng, Hong Zhang, and Chao Liang. A survey on rank aggregation. In *IJCAI*, pp. 8281–8289, 2024.

Chen-Rui Xia, Zhi-Jie Cao, Xin-Ming Tu, and Ge Gao. Spatial-linked alignment tool (slat) for aligning heterogenous slices. *Nature Communications*, 14(1):7236, 2023.

Chang Xu, Xiyun Jin, Songren Wei, Pingping Wang, Meng Luo, Zhaochun Xu, Wenyi Yang, Yideng Cai, Lixing Xiao, Xiaoyu Lin, et al. Deepst: identifying spatial domains in spatial transcriptomics by deep learning. *Nucleic acids research*, 50(22):e131–e131, 2022.

H Xu, H Fu, Y Long, KS Ang, R Sethi, K Chong, et al. Unsupervised spatially embedded deep representation of spatial transcriptomics. genome med. 2024; 16 (1): 12.

Hang Xu, Huazhu Fu, Yahui Long, Kok Siong Ang, Raman Sethi, Kelvin Chong, Mengwei Li, Rom Uddamvathanak, Hong Kai Lee, Jingjing Ling, et al. Unsupervised spatially embedded deep representation of spatial transcriptomics. *Genome medicine*, 16(1):12, 2024.

Wenyi Yang, Pingping Wang, Shouping Xu, Tao Wang, Meng Luo, Yideng Cai, Chang Xu, Guangfu Xue, Jinhao Que, Qian Ding, et al. Deciphering cell–cell communication at single-cell resolution for spatial transcriptomics with subgraph-based graph attention network. *Nature Communications*, 15(1):7101, 2024.

Roxana Zahedi, Reza Ghamsari, Ahmadreza Argha, Callum Macphillamy, Amin Beheshti, Roohallah Alizadehsani, Nigel H Lovell, Mohammad Lotfollahi, and Hamid Alinejad-Rokny. Deep learning in spatially resolved transcriptomics: a comprehensive technical view. *Briefings in Bioinformatics*, 25(2):bbae082, 2024.

Ron Zeira, Max Land, Alexander Strzalkowski, and Benjamin J Raphael. Alignment and integration of spatial transcriptomics data. *Nature Methods*, 19(5):567–575, 2022.

Zhenhao Zhang, Antonio Jimeno Yepes, Jiang Bian, Fuyi Li, and Yuxi Liu. Gatorduo: Global-consistency dual-graph refinement with pseudo-label agreement for spatial transcriptomics. *Biorxiv: the Preprint Server for Biology*, 2026.

Edward Zhao, Matthew R Stone, Xing Ren, Jamie Guenthoer, Kimberly S Smythe, Thomas Pulliam, Stephen R Williams, Cedric R Uytingco, Sarah EB Taylor, Paul Nghiem, et al. Spatial transcriptomics at subspot resolution with bayesspace. *Nature biotechnology*, 39(11):1375–1384, 2021.

Jingcheng Zhao and Wenwen Min. Spaicl: image-guided curriculum strategy-based graph contrastive learning for spatial transcriptomics clustering. *Briefings in Bioinformatics*, 26(4):bbaf433, 2025.

Xiang Zhou, Kangning Dong, and Shihua Zhang. Integrating spatial transcriptomics data across different conditions, technologies and developmental stages. *Nature Computational Science*, 3 (10):894–906, 2023.

Yongshuo Zong, Tingyang Yu, Xuesong Wang, Yixuan Wang, Zhihang Hu, and Yu Li. const: an interpretable multi-modal contrastive learning framework for spatial transcriptomics. *BioRxiv*, pp. 2022–01, 2022.

A Experimental Settings

Datasets. - We evaluate BMRL on four spatial transcriptomics datasets: DLPFC, MTG, Mouse Embryo, and NSCLC. For single-slice spatial domain identification, all four datasets are used. DLPFC (Maynard et al., 2021) contains 12 slices from three patients, MTG (Chen et al., 2022b) contains control and Alzheimer's disease (AD) samples, Mouse Embryo contains slices (Chen et al., 2022a)

from different developmental stages, and NSCLC (Bhuva et al., 2024) is generated by the CosMx platform with subcellular spatial resolution. For multi-slice experiments, four slices from the same DLPFC patient are used for homogeneous integration, MTG is used for heterogeneous integration across control and AD samples, and Mouse Embryo is used for spatial alignment across developmental stages.

Baselines. To ensure a fair comparison, we select representative baseline methods for different experimental tasks. For the single-slice spatial domain identification task, we compare the proposed method with SEDR (Xu et al., 2024), STAGATE (Dong & Zhang, 2022), SpaCAE (Hu et al., 2024), SpaceFlow (Ren et al., 2022), and GraphST (Long et al., 2023). For the DLPFC homogeneous multi-slice integration task, we further include PASTE (Zeira et al., 2022), STAligner (Zhou et al., 2023), CAST (Tang et al., 2024), and Spotscape (Oh et al., 2025). For the MTG heterogeneous integration task, we use GraphST, Spotscape (Oh et al., 2025), STAligner, and CAST (Tang et al., 2024) as the compared methods. For the Mouse embryo multi-slice alignment task, we further compare our method with PASTE2 (Liu et al., 2023b), CAST, STAligner, and SLAT (Xia et al., 2023).
Evaluation Protocol. Since our method and the compared methods mainly aim to learn spot representations, we first obtain the representations generated by each method. Then, we perform different downstream evaluations for different experimental tasks. For the single-slice spatial domain identification task, we perform clustering on the learned spot embeddings following Long et al. (2023). We use Adjusted Rand Index (ARI), Normalized Mutual Information (NMI), and Clustering Accuracy (CA) to evaluate the clustering performance. For multi-slice integration, we evaluate clustering performance using the same metrics as in the single-slice experiments and assess batch correction using the batch silhouette score, kBET, graph connectivity, and PCR comparison, following Oh et al. (2025); Luecken et al. (2022). For the multi-slice alignment task, we use Label Transfer ARI (LTARI) to evaluate the alignment quality. LTARI measures the consistency between the aligned labels and the ground-truth labels (Xia et al., 2023). All experiments are repeated ten times with different random seeds, and we report the mean and standard deviation. In the experimental tables, bold values indicate the best results, while underlined values indicate the second-best results.

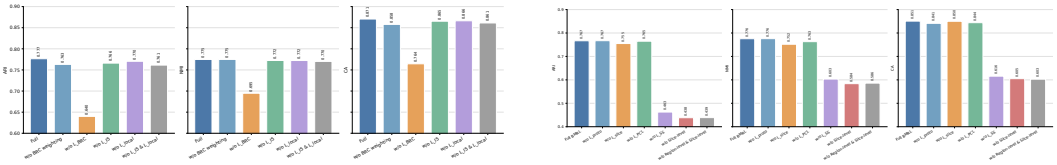
Implementation details. Unless otherwise specified, we use a two-layer GCN encoder with hidden and embedding dimensions of 256 and 64, respectively, together with an MLP decoder. The model is trained for 1,000 epochs using the Adam optimizer with a weight decay of 10^{-4}. The base learning rate is set to 10^{-4} for DLPFC, MTG, and NSCLC, and 10^{-5} for the mouse embryo dataset. Two augmented views are independently generated using feature masking and edge dropping, with both ratios set to 0.2. All experiments are conducted on NVIDIA GeForce RTX 4070 Laptop and RTX 4090 GPUs.

B Ablation Study

As shown in Fig. 5a, to evaluate the effects of different constraints in boundary-aware representation learning, we conduct ablation experiments on the spatial domain identification task. The model with all constraints is denoted as Full. Based on the full model, we remove the boundary weighting in $\mathcal{L}_{\mathrm{BRC}}$, remove the boundary-aware relation consistency loss $\mathcal{L}_{\mathrm{BRC}}$, and remove the two intra-domain constraints separately or together. These two constraints are the intra-view spatial smoothing loss $\mathcal{L}_{\mathrm{IS}}$ and the cross-view local relation consistency loss $\mathcal{L}_{\mathrm{local}}$.

The largest performance drop occurs when $\mathcal{L}_{\mathrm{BRC}}$ is removed. ARI decreases from 0.777 to 0.640, while CA decreases from 0.871 to 0.765. This result shows that cross-view global relation consistency is important for learning stable spatial representations. When $\mathcal{L}_{\mathrm{BRC}}$ is kept but its boundary weighting is removed, ARI and CA decrease to 0.763 and 0.858, respectively, while NMI remains almost unchanged. This confirms that explicitly distinguishing interior and boundary nodes is beneficial for controlling representation smoothness and preserving domain differences.

For low-boundary regions, $\mathcal{L}_{\mathrm{IS}}$ keeps spatial continuity by increasing representation smoothing on reliable spatial edges. In contrast, $\mathcal{L}_{\mathrm{local}}$ keeps the same spatial relation stable across the two augmented views. When $\mathcal{L}_{\mathrm{IS}}$ and $\mathcal{L}_{\mathrm{local}}$ are removed separately, ARI decreases to 0.766 and 0.771, respectively. When both constraints are removed, ARI further decreases to 0.761, while NMI and CA show similar trends. These results show that the two constraints improve intra-domain representations from two aspects: representation smoothing and relation stability. Using them together leads to better spatial domain identification performance.



(a) Boundary-aware representation learning. (b) Multi-slice relational learning.

Figure 5: Ablation studies of BMRL: (a) boundary-aware representation learning on DLPFC and (b) multi-slice relational learning on MTG.

Table 5: Analysis of local and global information for boundary estimation on DLPFC, MTG, Mouse Embryo, and NSCLC. Results are reported as mean (standard deviation).

(a) DLPFC (Patient 1)

	151673			151674			151675			151676		
Method	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
Local only	0.51 (0.01)	0.67 (0.02)	0.64 (0.01)	0.52 (0.02)	0.63 (0.02)	0.63 (0.02)	0.51 (0.02)	0.65 (0.01)	0.65 (0.01)	0.45 (0.05)	0.61 (0.03)	0.62 (0.04)
Global only	0.53 (0.01)	0.67 (0.02)	0.64 (0.02)	0.54 (0.03)	0.64 (0.03)	0.63 (0.02)	0.50 (0.01)	0.65 (0.01)	0.64 (0.01)	0.43 (0.03)	0.57 (0.02)	0.60 (0.02)
Local + Global	0.53 (0.02)	0.68 (0.02)	0.64 (0.01)	0.52 (0.02)	0.65 (0.01)	0.64 (0.02)	0.53 (0.02)	0.66 (0.01)	0.66 (0.02)	0.45 (0.04)	0.61 (0.03)	0.62 (0.03)

(b) DLPFC (Patient 2)

	151507			151508			151509			151510		
Method	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
Local only	0.52 (0.06)	0.69 (0.02)	0.63 (0.07)	0.52 (0.02)	0.67 (0.01)	0.67 (0.03)	0.59 (0.01)	0.72 (0.01)	0.70 (0.01)	0.58 (0.03)	0.66 (0.01)	0.65 (0.02)
Global only	0.51 (0.06)	0.64 (0.04)	0.66 (0.06)	0.53 (0.02)	0.67 (0.01)	0.65 (0.03)	0.56 (0.04)	0.70 (0.02)	0.68 (0.03)	0.53 (0.01)	0.67 (0.01)	0.63 (0.02)
Local + Global	0.63 (0.04)	0.72 (0.03)	0.79 (0.03)	0.53 (0.02)	0.67 (0.01)	0.67 (0.02)	0.59 (0.01)	0.72 (0.01)	0.70 (0.01)	0.55 (0.03)	0.68 (0.01)	0.66 (0.01)

(c) DLPFC (Patient 3)

	151669			151670			151671			151672		
Method	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
Local only	0.49 (0.01)	0.61 (0.01)	0.67 (0.01)	0.48 (0.02)	0.59 (0.01)	0.67 (0.01)	0.76 (0.02)	0.76 (0.01)	0.85 (0.02)	0.78 (0.04)	0.75 (0.03)	0.85 (0.05)
Global only	0.51 (0.01)	0.62 (0.01)	0.67 (0.01)	0.50 (0.02)	0.60 (0.01)	0.68 (0.01)	0.60 (0.13)	0.68 (0.05)	0.70 (0.10)	0.74 (0.05)	0.72 (0.03)	0.83 (0.05)
Local + Global	0.51 (0.01)	0.62 (0.01)	0.68 (0.01)	0.50 (0.02)	0.60 (0.01)	0.68 (0.01)	0.76 (0.03)	0.76 (0.01)	0.85 (0.03)	0.80 (0.03)	0.76 (0.02)	0.87 (0.03)

(d) MTG, Mouse Embryo, and NSCLC

	MTG Control			MTG AD			Mouse Embryo			NSCLC		
Method	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
Local only	0.77 (0.05)	0.79 (0.02)	0.79 (0.05)	0.68 (0.02)	0.75 (0.01)	0.82 (0.03)	0.52 (0.01)	0.67 (0.00)	0.59 (0.02)	0.62 (0.01)	0.58 (0.01)	0.78 (0.00)
Global only	0.77 (0.02)	0.79 (0.01)	0.76 (0.02)	0.68 (0.01)	0.75 (0.00)	0.82 (0.01)	0.52 (0.00)	0.66 (0.00)	0.59 (0.01)	0.61 (0.01)	0.57 (0.00)	0.77 (0.01)
Local + Global	0.78 (0.04)	0.80 (0.02)	0.80 (0.04)	0.69 (0.01)	0.76 (0.01)	0.83 (0.01)	0.52 (0.01)	0.67 (0.00)	0.60 (0.01)	0.62 (0.01)	0.58 (0.01)	0.79 (0.01)

As shown in Fig. 5b, to evaluate the effects of the region-level and slice-level learning objectives, we keep the spatial learning part fixed. We separately remove the prototype relation consistency loss $\mathcal{L}_{\text{proto}}$, the slice representation consistency loss $\mathcal{L}_{\text{slice}}$, the prototype pseudo-label supervision $\mathcal{L}_{\text{PCL}}$, and the cross-slice similarity scaling loss $\mathcal{L}_{\text{SS}}$. We also conduct joint ablation experiments.

After removing $\mathcal{L}_{\text{proto}}$, CA decreases from 0.851 to 0.841, while ARI and NMI remain almost unchanged. Removing $\mathcal{L}_{\text{slice}}$ or $\mathcal{L}_{\text{PCL}}$ also leads to different degrees of performance decrease. This shows that cross-view slice representation consistency and pseudo-label supervision from shared prototypes provide complementary benefits. Removing $\mathcal{L}_{\text{SS}}$ causes the largest performance drop among the individual ablations. ARI decreases from 0.767 to 0.463, showing that keeping intra-slice and inter-slice similarities at a consistent scale is important for the multi-slice task. When the three slice-level objectives are removed together, ARI further decreases to 0.439.

C METHOD ANALYSIS

Local-Global Boundary Analysis. As shown in Table 5, we further analyze the effects of local and global information on boundary estimation. Specifically, we compare three variants using only local expression differences, only global semantic information, and their combination. The local-only variant captures fine-grained changes in spatial neighborhoods, while the global-only variant provides broader semantic relations among spots. Combining the two types of information achieves

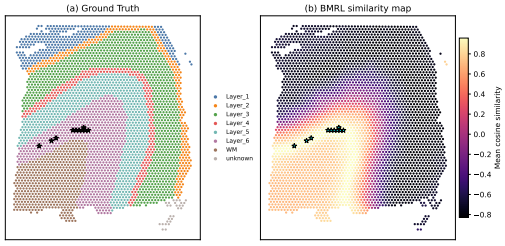


Figure 6: Similarity analysis of BMRL representations on DLPFC slice 151673 using 10 Layer 6 anchor spots.

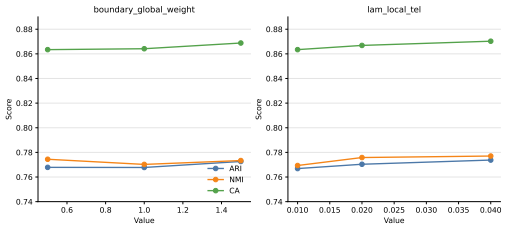


Figure 7: Sensitivity of BMRL to the global boundary weight and local relation loss weight on DLPFC slice 151671.

better overall performance than using either one alone. This result shows that local and global information provide complementary cues for distinguishing interior and boundary spots.

Similarity Analysis. Based on the experimental results, we further conduct a similarity analysis to evaluate whether BMRL can learn spot representations with meaningful domain-level relational structures. Specifically, we select 10 anchor spots from Layer 6 of DLPFC slice 151673 and calculate the average cosine similarity between these anchor spots and all other spots in the slice. As shown in Fig. 6, spots with high similarity are mainly located in the Layer 6 region containing the anchor spots and its nearby spatial regions. This spatial pattern is generally consistent with the ground-truth layer annotation. These results show that the representations learned by BMRL can preserve spatial domain structures and capture biologically meaningful spot-to-spot relations.

Sensitivity Analysis. We further analyze the sensitivity of BMRL to key hyperparameters. As shown in Fig. 7, when the global boundary weight and the local relation loss weight change within different ranges, ARI, NMI, and CA show only small changes. The overall performance remains stable. These results show that BMRL is robust to key hyperparameters.