ICML 2025

Global Context-aware Representation Learning for Spatially Resolved Transcriptomics

Yunhak Oh*, Junseok Lee*, Yeongmin Kim, Sangwoo Seo, Namkyeong Lee, & Chanyoung Park†

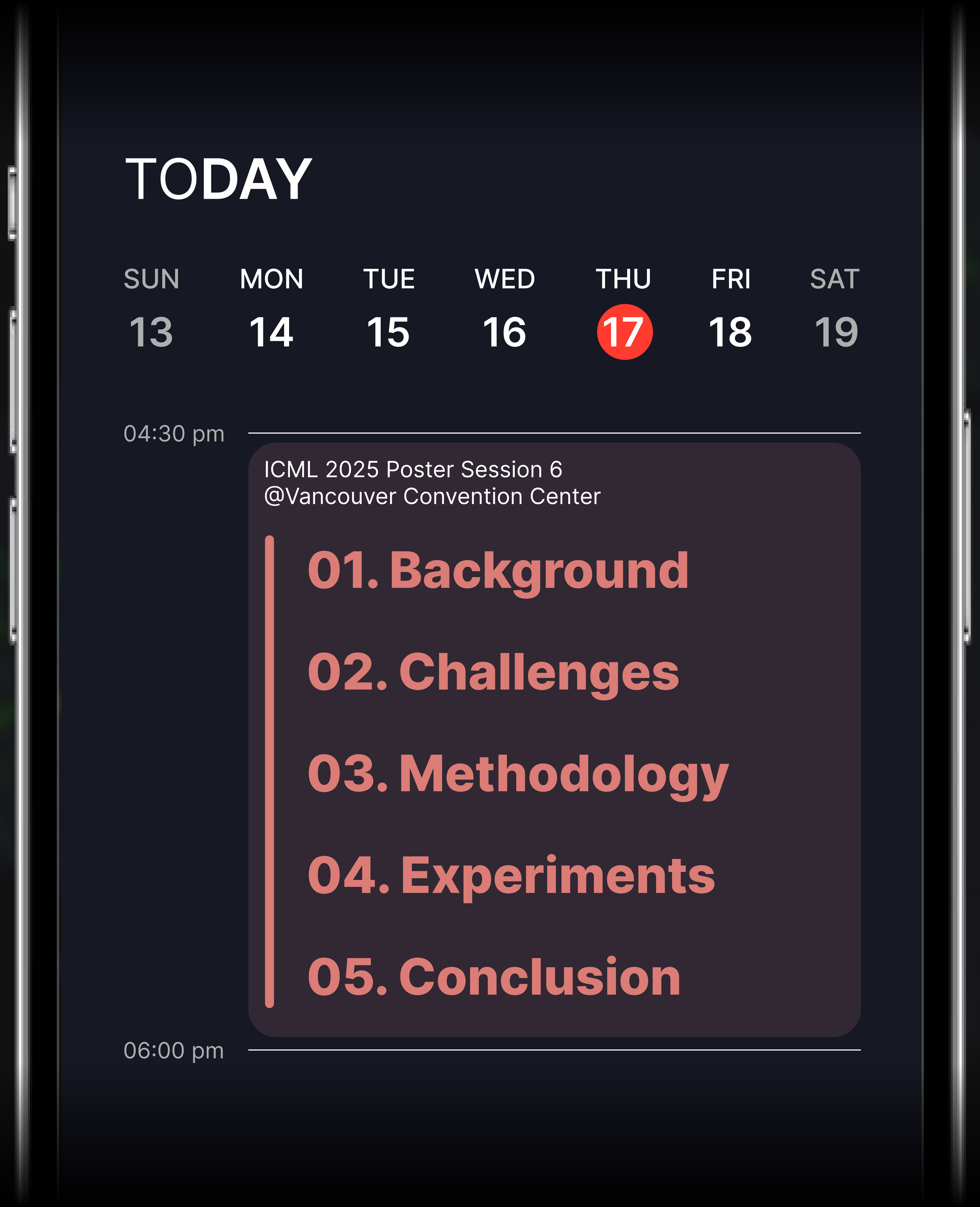
*: Equal Contribution
†: Corresponding Author

ICML
International Conference
On Machine Learning

KAIST

DSAIL Data Science &
Artificial Intelligence

Spotscape enhances spatial transcriptomics analysis by capturing **global cellular relationships** and **balancing the similarity scales** to more accurately map tissue structures across single and multiple slices



TODAY

SUN	MON	TUE	WED	THU	FRI	SAT
13	14	15	16	17	18	19

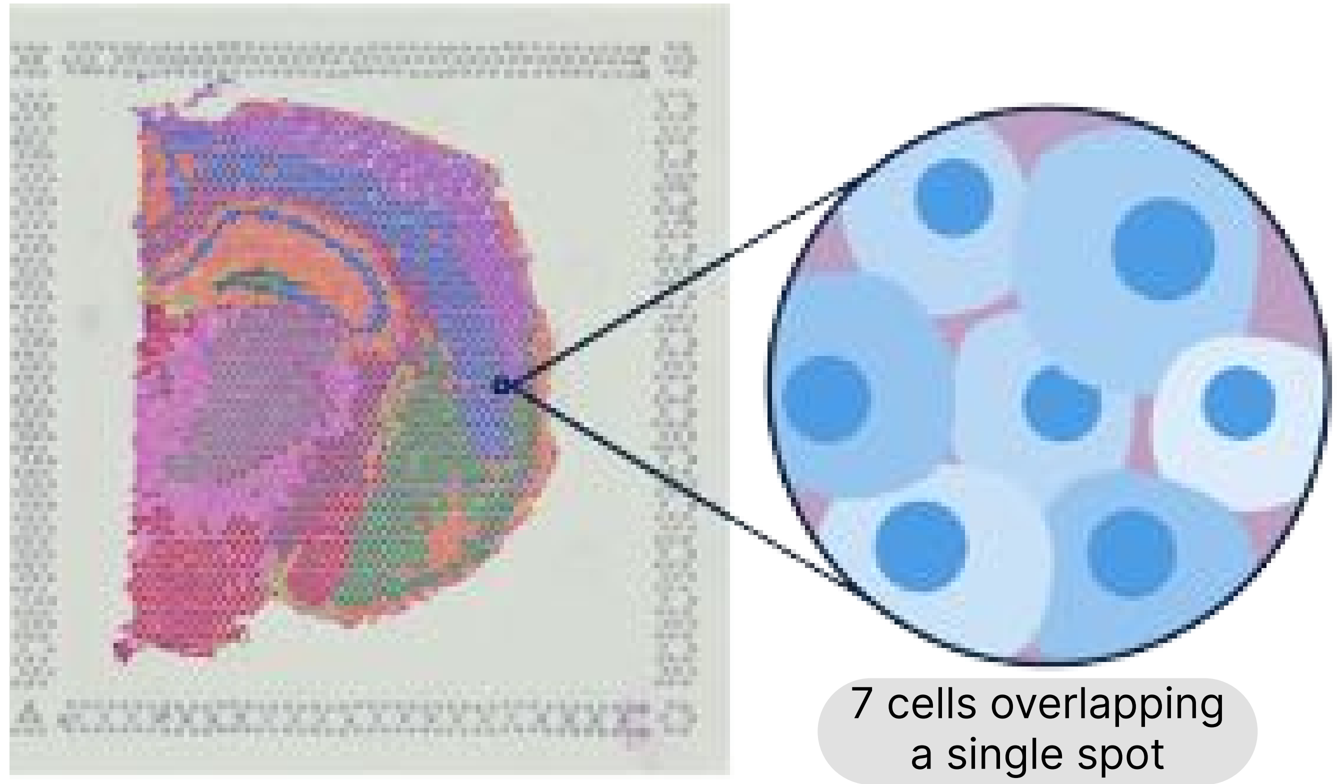
04:30 pm

ICML 2025 Poster Session 6
@Vancouver Convention Center

01. Background
02. Challenges
03. Methodology
04. Experiments
05. Conclusion

06:00 pm

Spatially Resolved Transcriptomics (SRT) BACKGROUND



7 cells overlapping a single spot

	Gene 1	Gene 2	Gene 3	Gene 4
Spot 1	18	1010	0	22
Spot 2	0	506	49	0
Spot 3	0	0	0	72

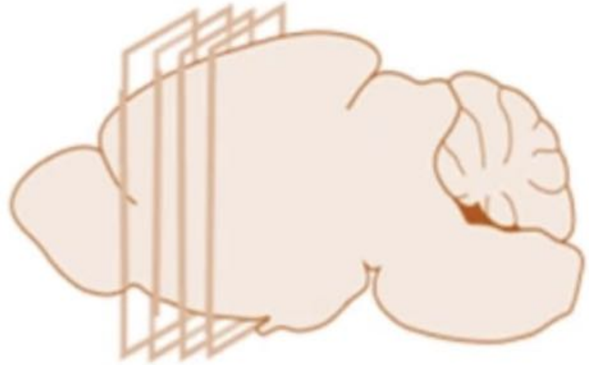
- **SRT** incorporate the **spatial context** and gene expression of cells
  - allowing for the identification of **where specific genes are expressed** within tissue
  - **enables the study of diseases**, such as cancer, by observing gene expression patterns in different tissue regions

Image credit: https://www.10xgenomics.com/analysis-guides/integrating-single-cell-and-visium-spatial-gene-expression-data

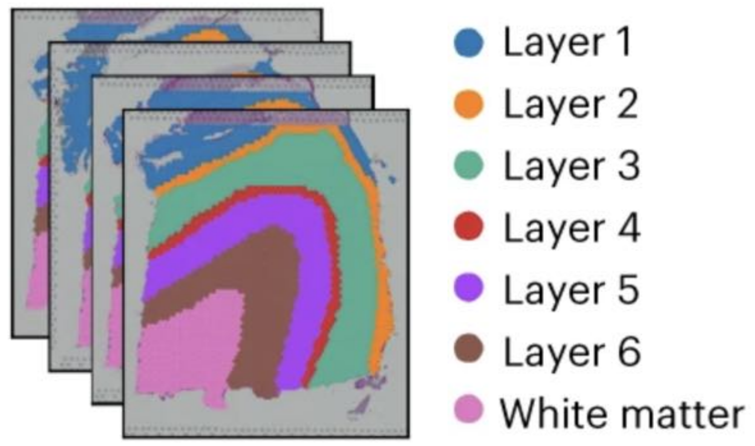
Spatially Resolved Transcriptomics (SRT) BACKGROUND

Representation Learning on SRT data for various downstream tasks

Homogeneous tasks



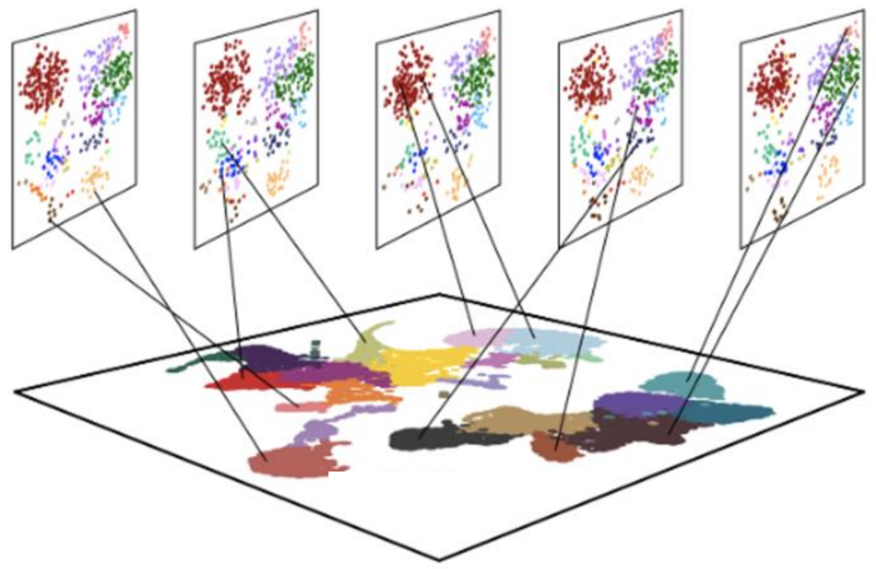
Single or Consecutive slices



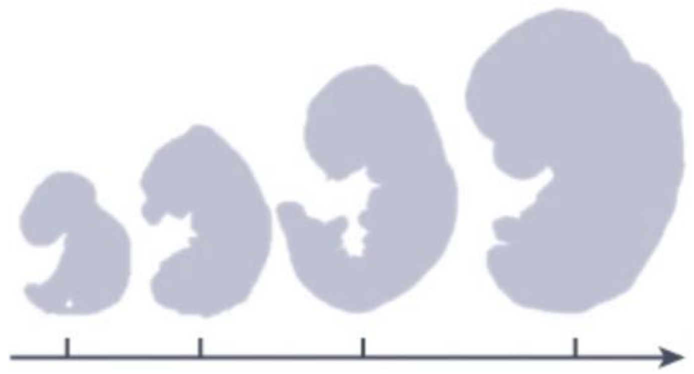
Spatial Domain Identification

Heterogeneous tasks

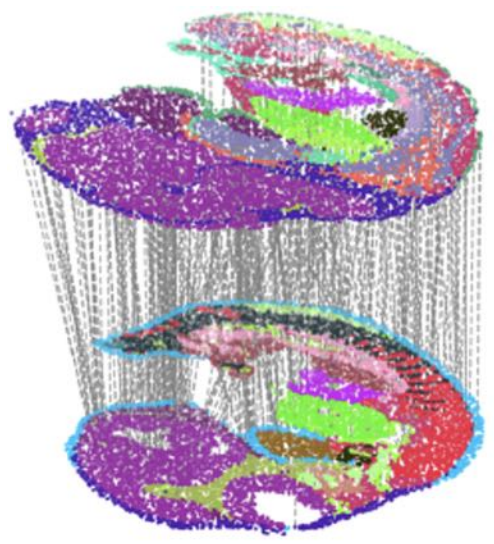
Normal vs. Disease



Integration



Development stages



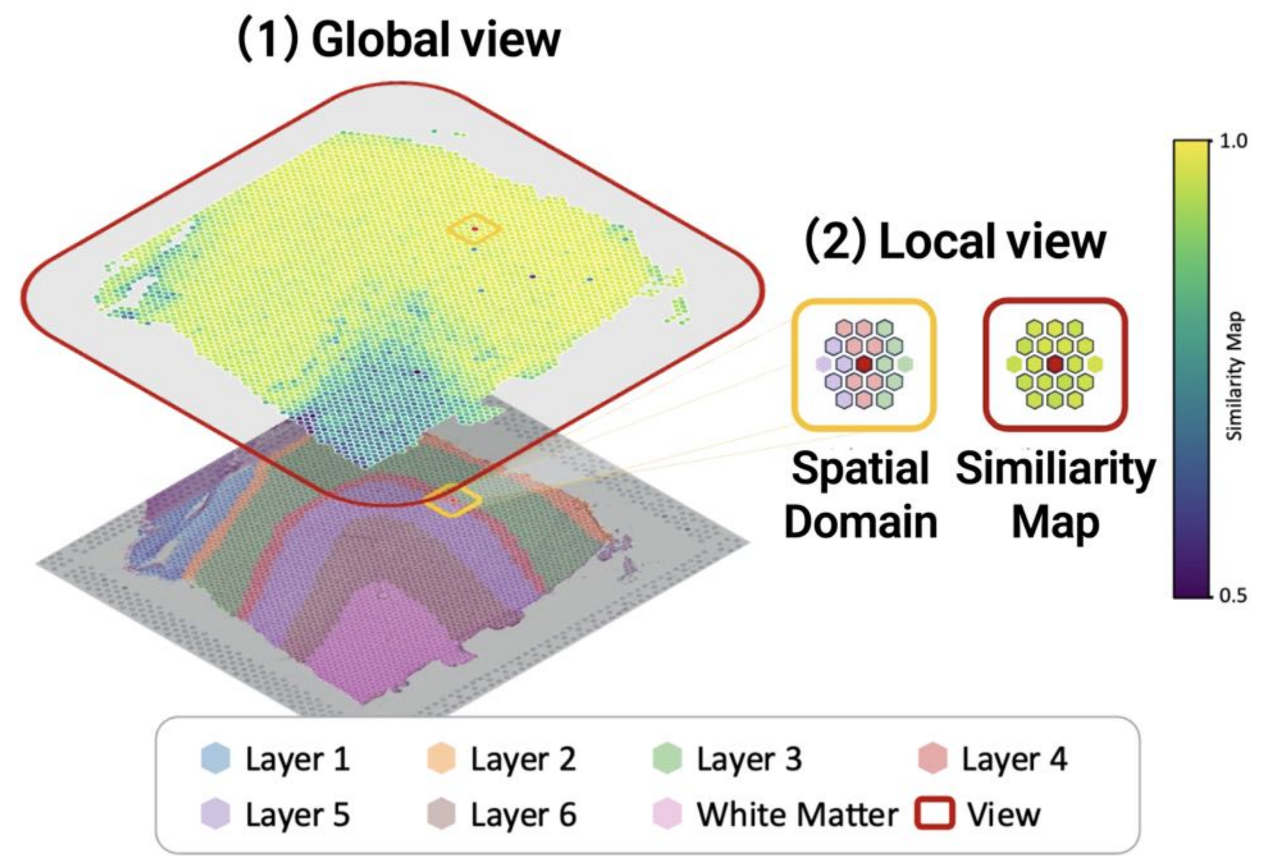
Alignment

Image credit: [Nature Computational Science 2023] Integrating spatial transcriptomics data across different conditions, technologies and developmental stages
https://www.archrproject.com/bookdown/defining-cluster-identity-with-scrna-seq.html

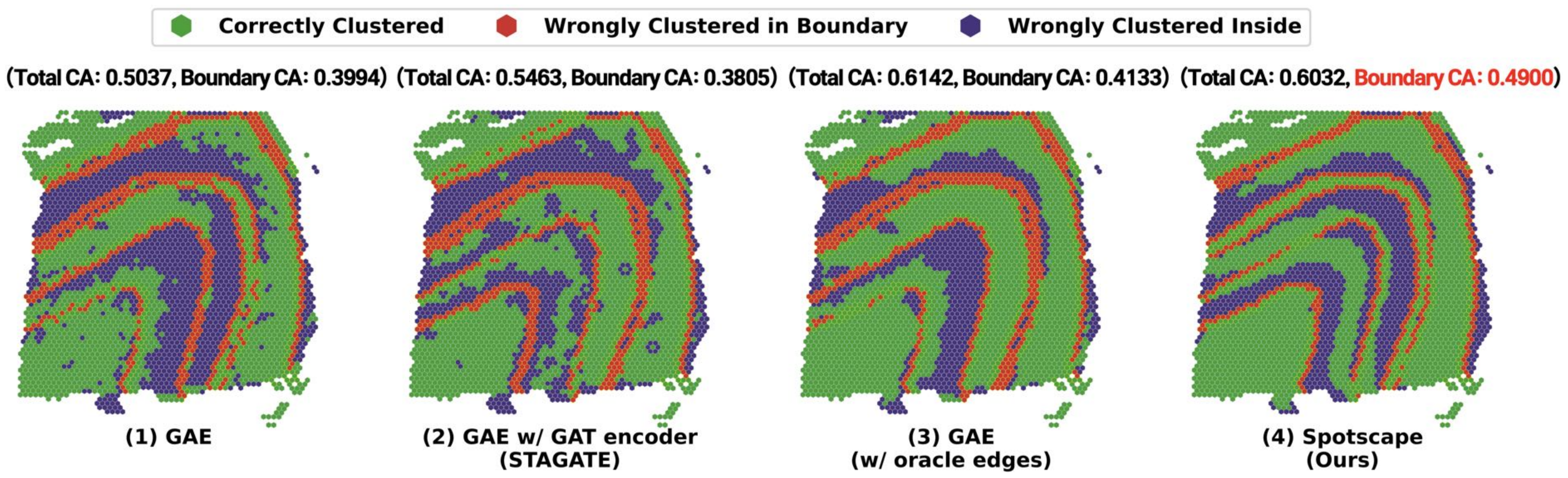
Continuous Nature CHALLENGE 1

Continuous nature of biological systems (SRT data)

→ Gene expression values vary smoothly along spatial coordinates



Feature similarity comparison



Clustering Performance comparison

Observations

- Local (spatially close) spots have high similarity, regardless of the spatial domain
    1. Results in a negative effect on GNNs by acquiring information from heterophilic nodes
    2. Difficult to learn the appropriate attention scores that gives high scores to homophilic nodes due to their low feature difference
    3. Even if it can learn appropriate edges (attentions), local view has insufficient information

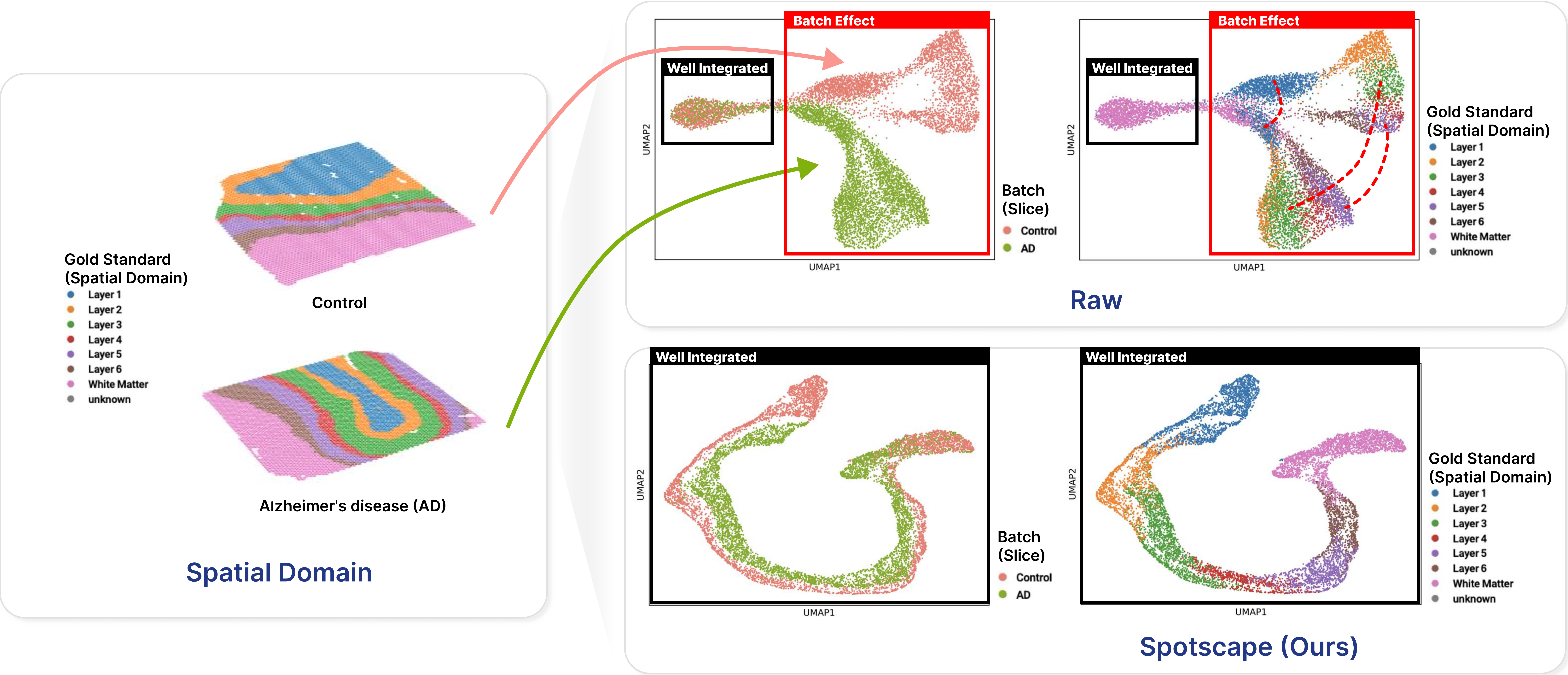
Should capture the relationships between spots in global context

Batch Effects

CHALLENGE 2

Batch effects in SRT data

→ Gene expression profiles from the same slice cluster together unexpectedly, regardless of their biological relevance

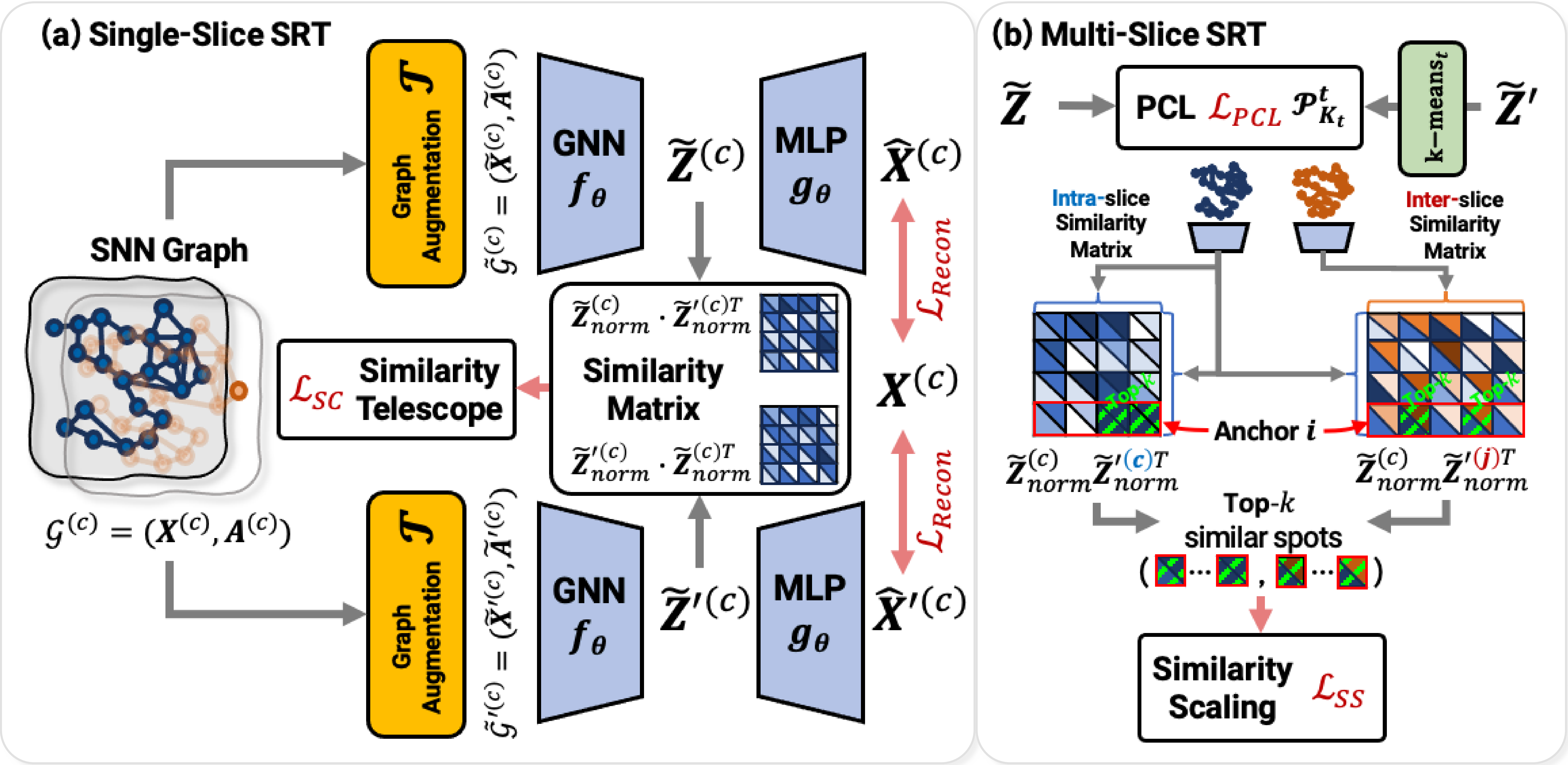


Should alleviate batch effects to extend multi-slices tasks

Key Ideas METHODOLOGY OVERVIEW

Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly **balances the similarity scales of inter- and intra-relationships** to mitigate batch effects
- **Grouping spots from the same spatial domain** while **distancing others** in latent space using PCL scheme to mitigate batch effects

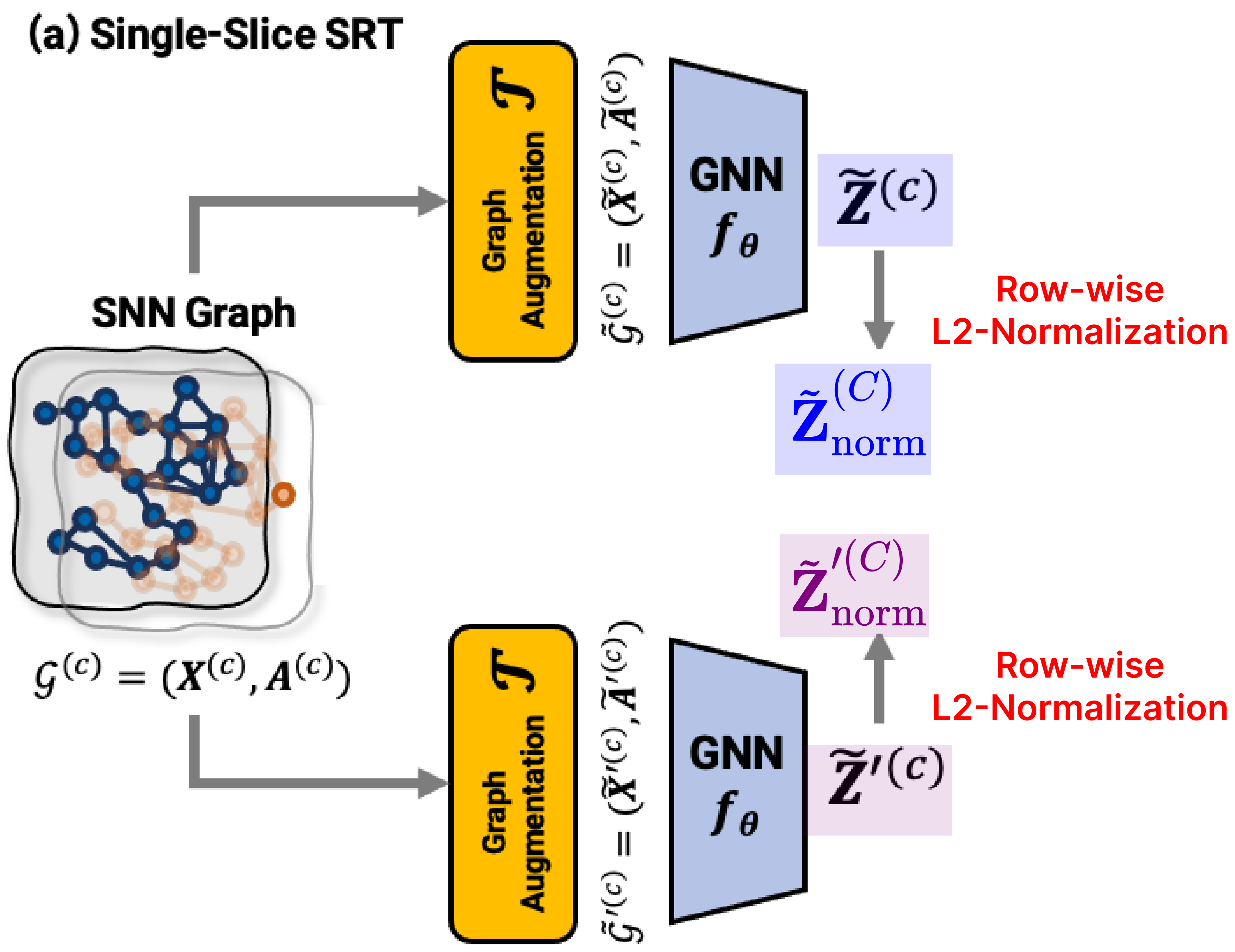


Key Ideas

METHODOLOGY 1

Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly **balances the similarity scales of inter- and intra-relationships** to mitigate batch effects
- **Grouping spots from the same spatial domain** while **distancing others** in latent space using PCL scheme to mitigate batch effects



(a) Single-Slice SRT

SNN Graph

$\mathcal{G}^{(c)} = (\boldsymbol{X}^{(c)}, \boldsymbol{A}^{(c)})$

Graph Augmentation $\mathcal{T}$

$\tilde{\mathcal{G}}^{(c)} = (\tilde{\boldsymbol{X}}^{(c)}, \tilde{\boldsymbol{A}}^{(c)})$

GNN f_θ

$\tilde{\boldsymbol{Z}}^{(c)}$

Row-wise L2-Normalization

$\tilde{\mathbf{Z}}^{(C)}_{\text{norm}}$

$\tilde{\mathbf{Z}}'^{(C)}_{\text{norm}}$

Row-wise L2-Normalization

$\tilde{\boldsymbol{Z}}'^{(c)}$

GNN f_θ

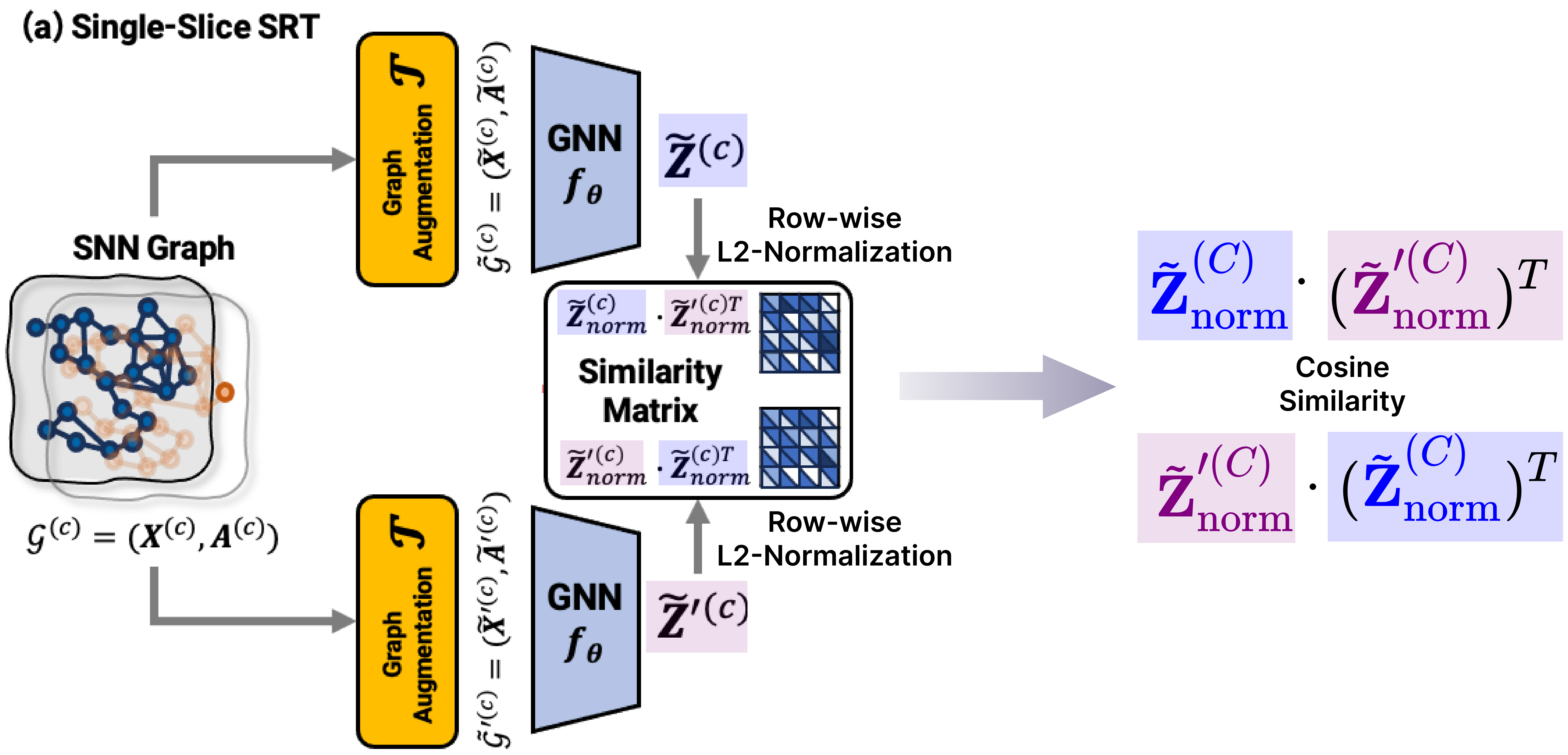
$\tilde{\mathcal{G}}'^{(c)} = (\tilde{\boldsymbol{X}}'^{(c)}, \tilde{\boldsymbol{A}}'^{(c)})$

Graph Augmentation $\mathcal{T}$

Key Ideas METHODOLOGY 1

Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly balances the similarity scales of inter- and intra-relationships to mitigate batch effects
- Grouping spots from the same spatial domain while distancing others in latent space using PCL scheme to mitigate batch effects

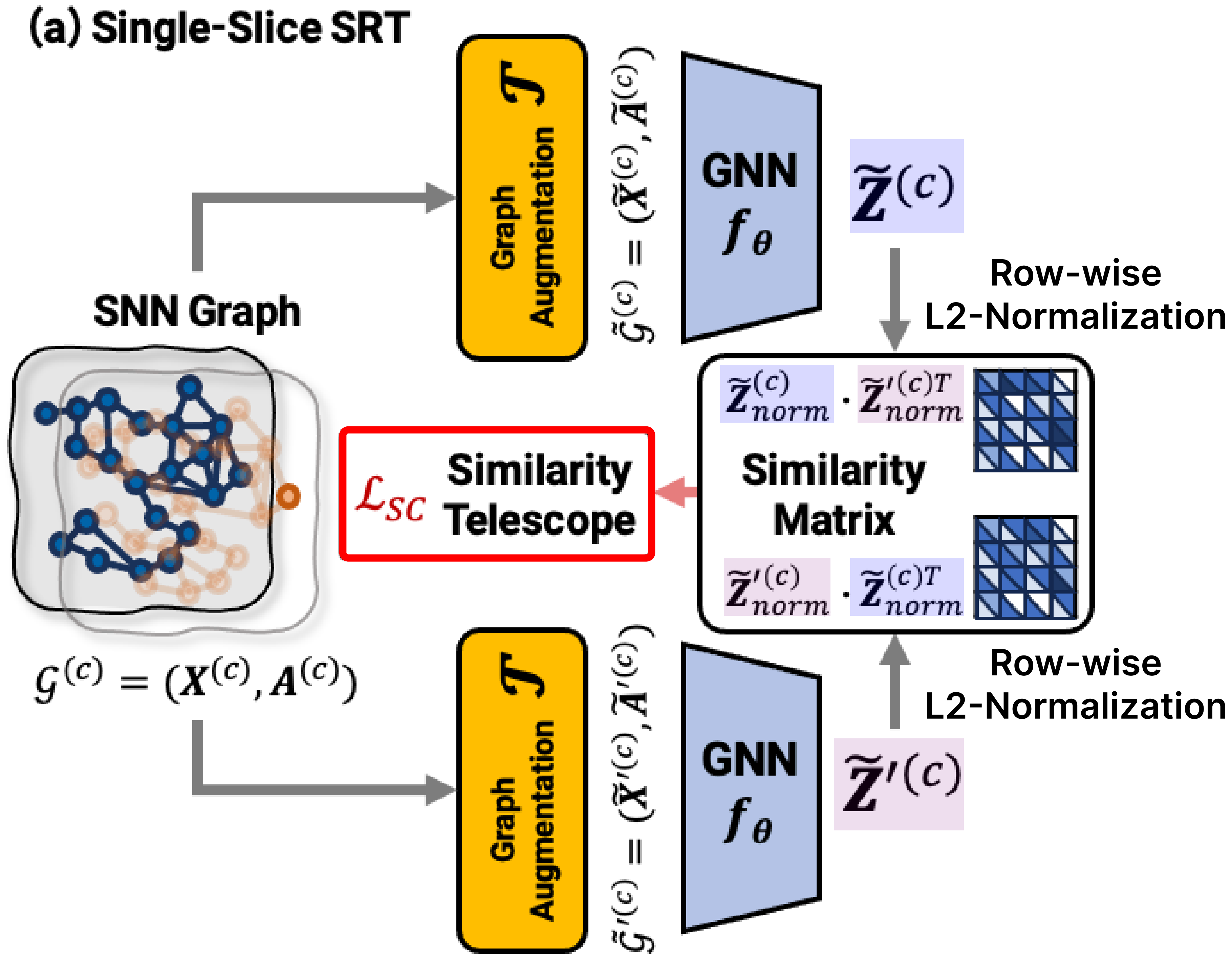


Key Ideas

METHODOLOGY 1

Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly balances the similarity scales of inter- and intra-relationships to mitigate batch effects
- Grouping spots from the same spatial domain while distancing others in latent space using PCL scheme to mitigate batch effects

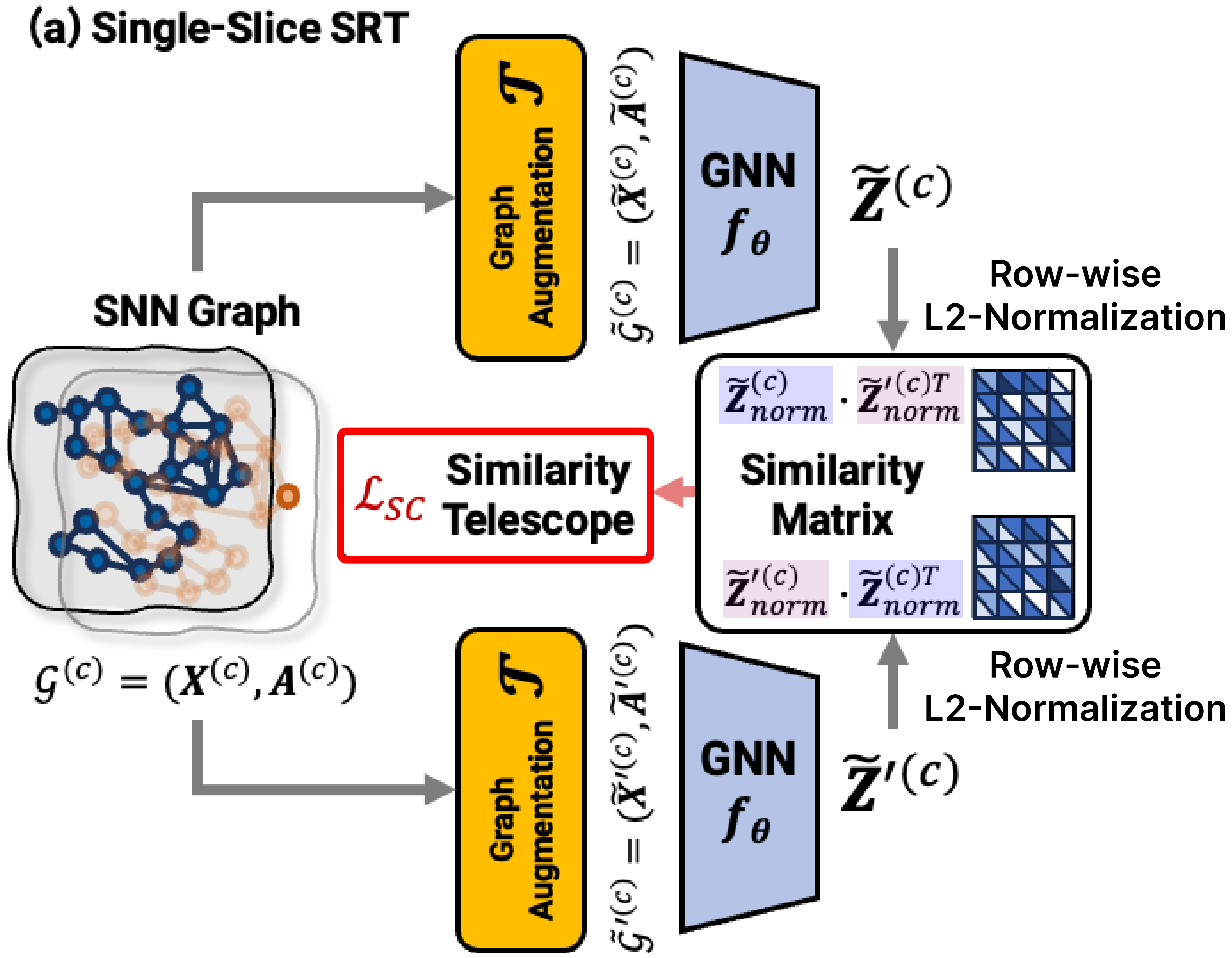

(a) Single-Slice SRT

$$\mathcal{L}_{\text{SC}}(\tilde{Z}^{(C)}, \tilde{Z}'^{(C)}) = \text{MSE}(\tilde{Z}^{(C)}_{\text{norm}}(\tilde{Z}'^{(C)}_{\text{norm}})^T, \tilde{Z}'^{(C)}_{\text{norm}}(\tilde{Z}^{(C)}_{\text{norm}})^T)$$

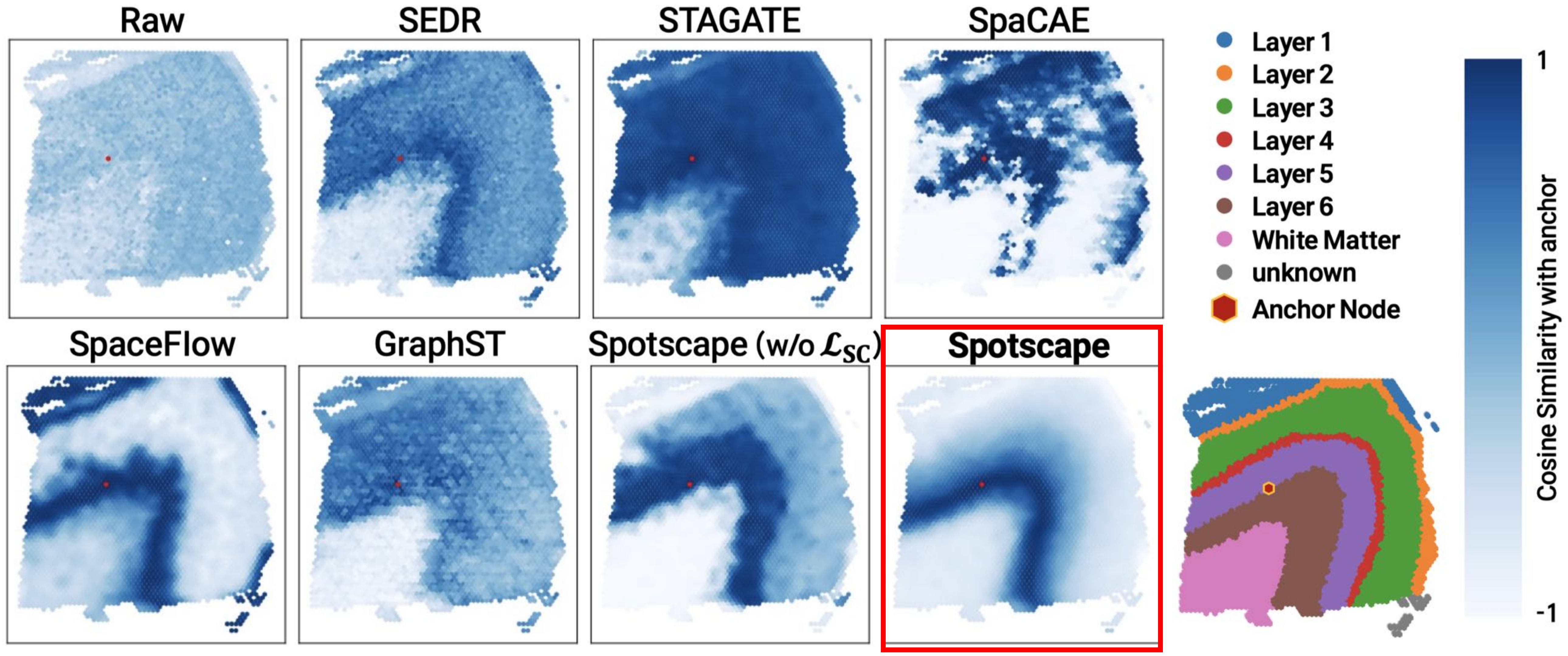
Key Ideas METHODOLOGY 1

Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly balances the similarity scales of inter- and intra-relationships to mitigate batch effects
- Grouping spots from the same spatial domain while distancing others in latent space using PCL scheme to mitigate batch effects



Similarity Consistency Loss



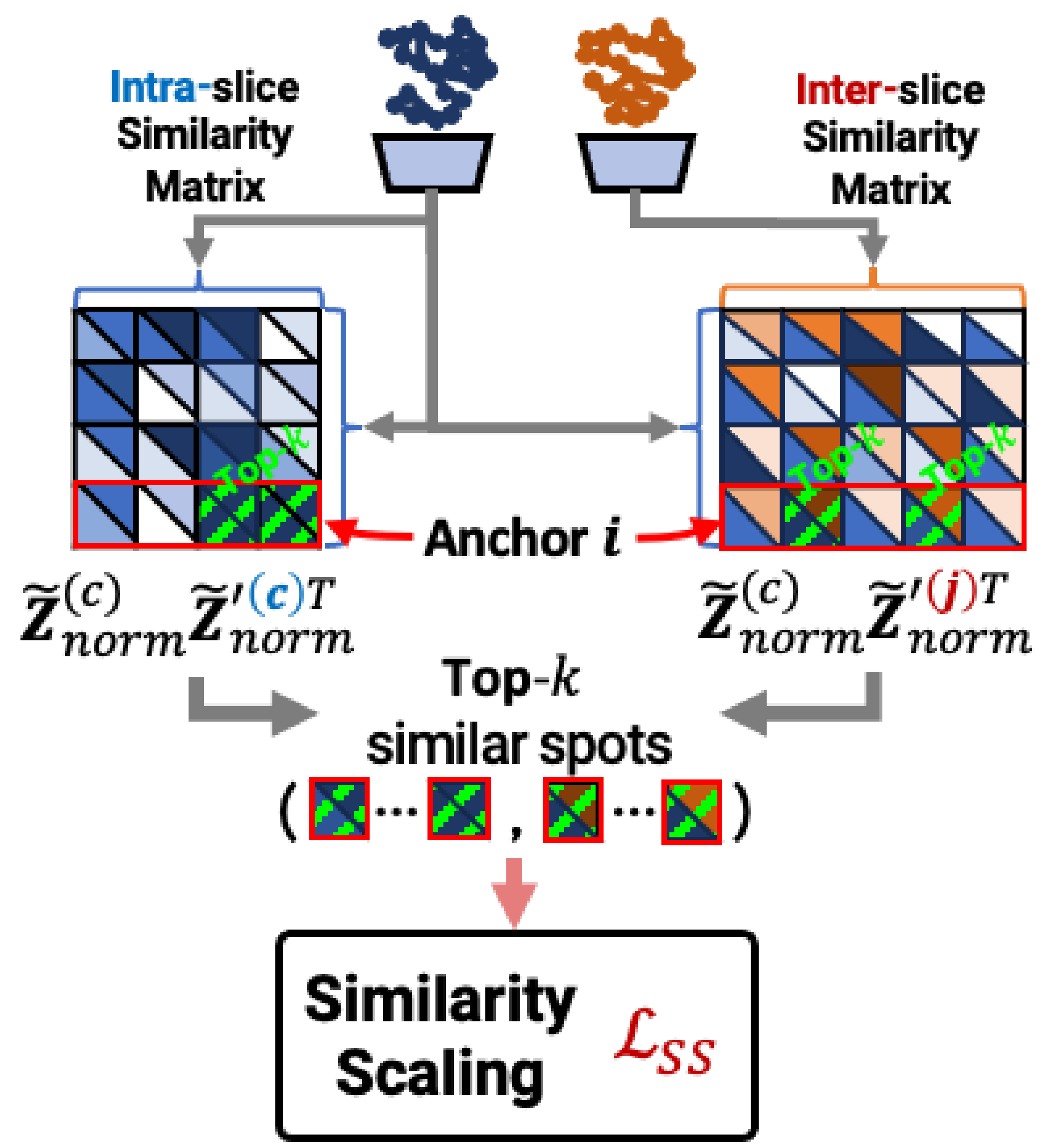
- **Spotscape** **captures relative similarities between spots**, aligning with the spatial dynamics of SRT data, unlike other baselines
- **Spotscape** exhibits varying similarity levels based on the true spatial domains, accurately reflecting spatial distance relationships

Key Ideas METHODOLOGY 2

Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly **balances the similarity scales of inter- and intra-relationships** to mitigate batch effects
- **Grouping spots from the same spatial domain** while **distancing others** in latent space using PCL scheme to mitigate batch effects

(b) Multi-Slice SRT

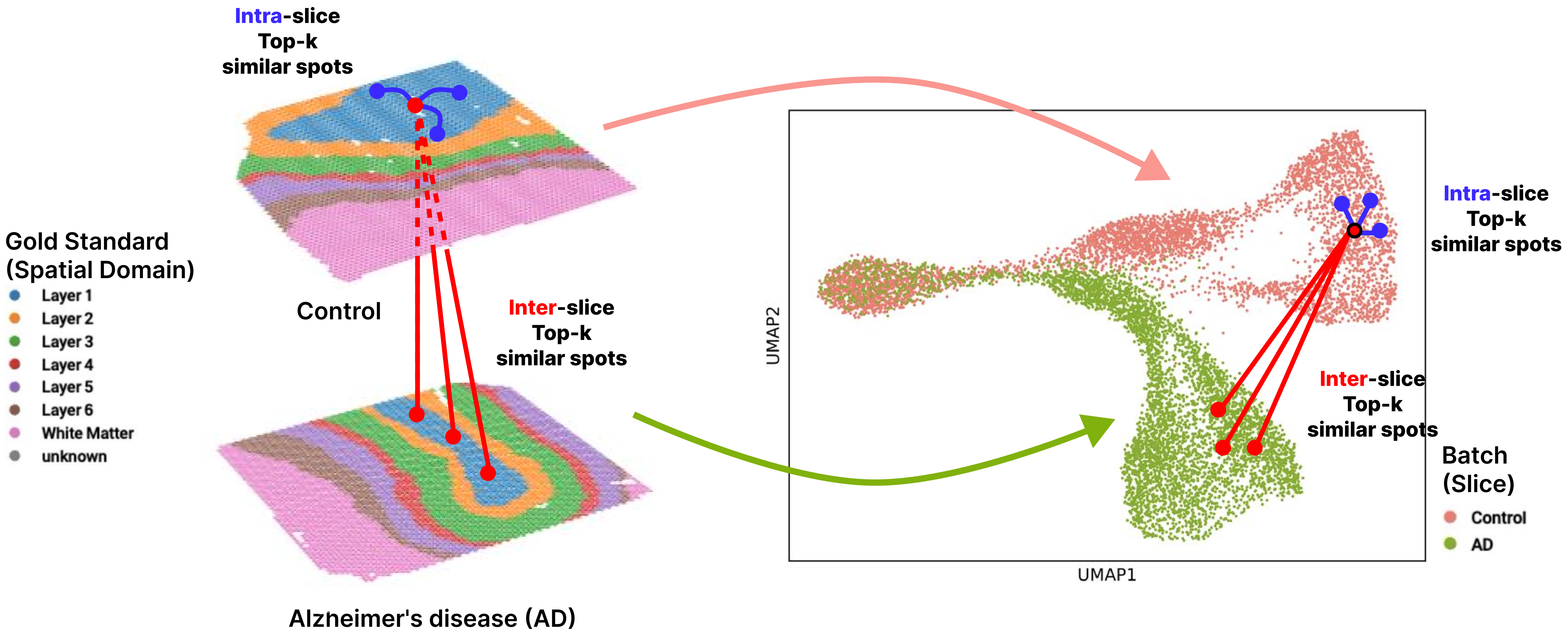


Key Ideas METHODOLOGY 2

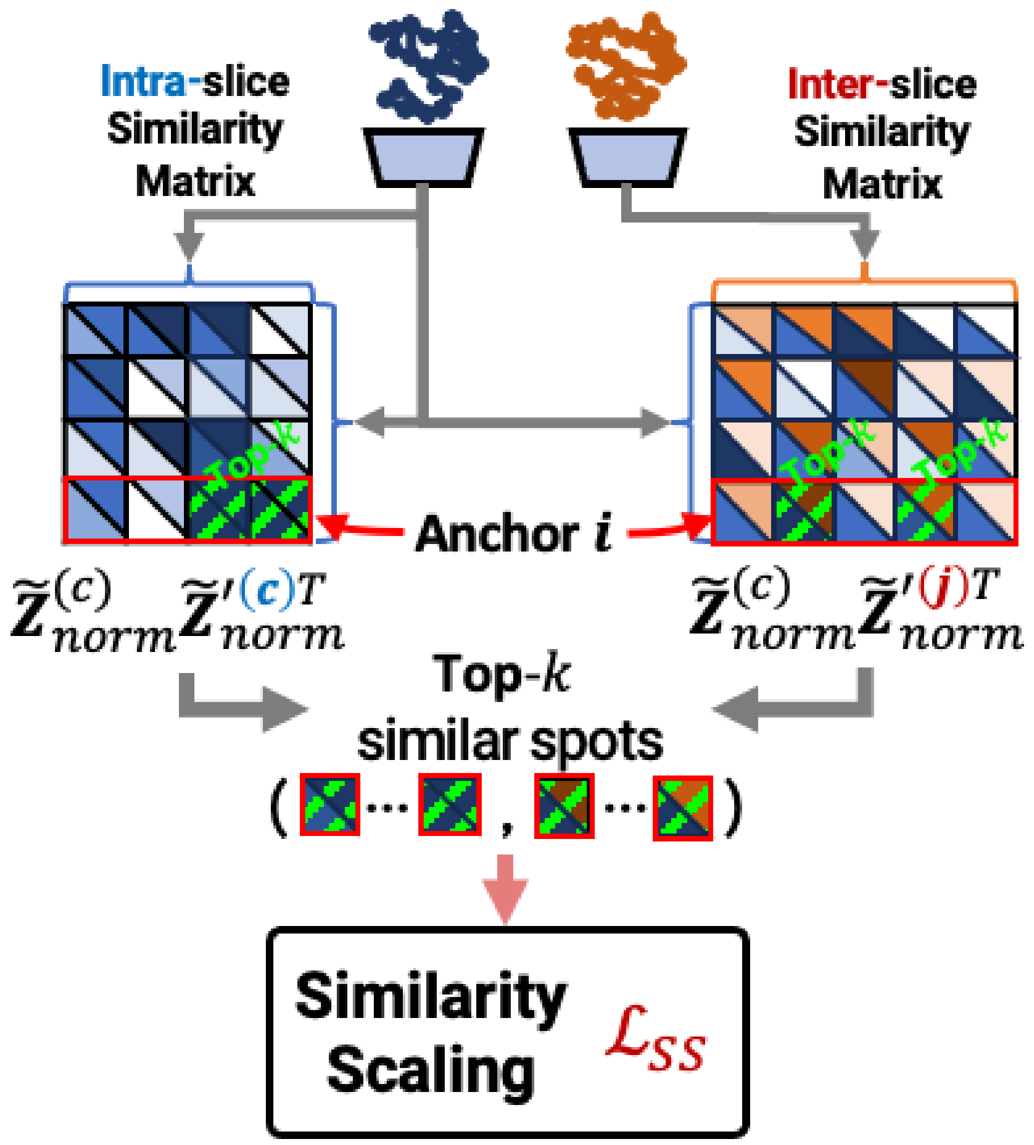
Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly **balances the similarity scales of inter- and intra-relationships** to mitigate batch effects
- **Grouping spots from the same spatial domain** while **distancing others** in latent space using PCL scheme to mitigate batch effects

Batch Effects: Experimental condition or noise ≫ Biological relevance



(b) Multi-Slice SRT

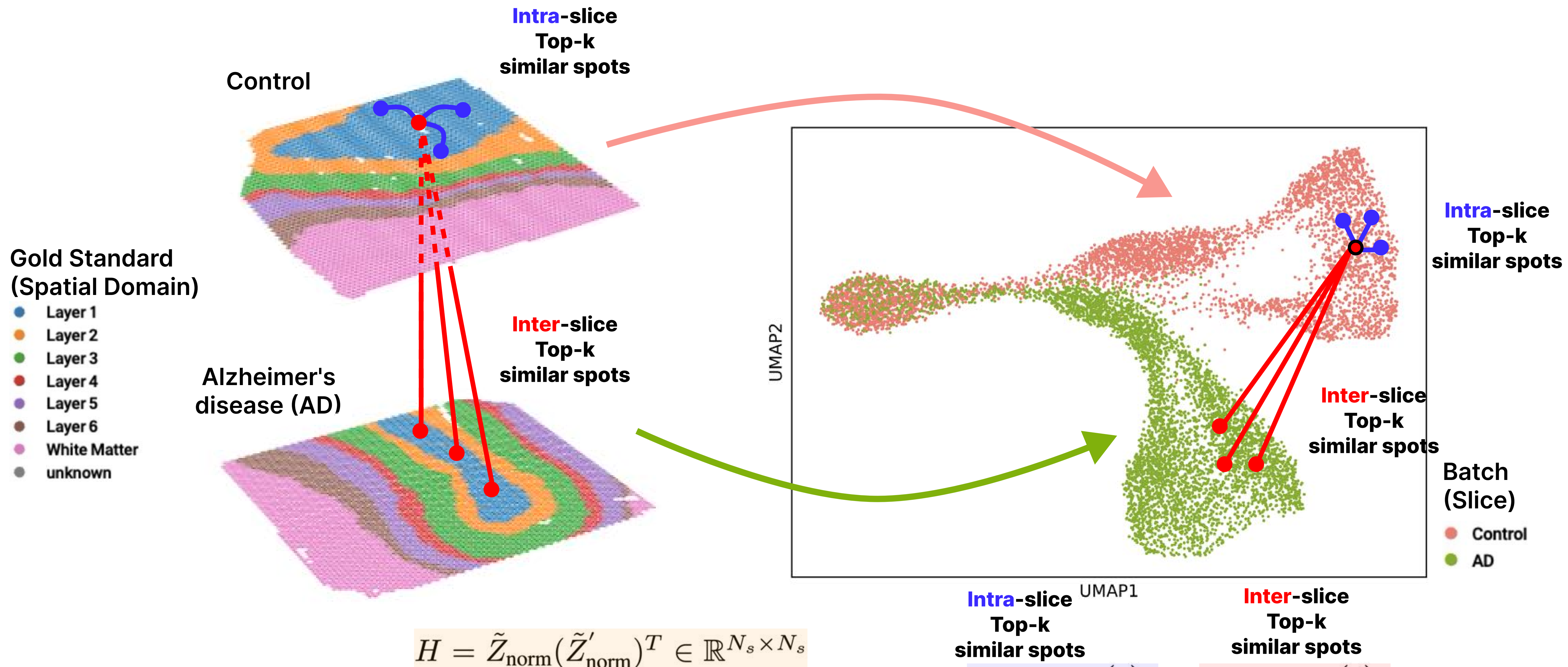


Key Ideas METHODOLOGY 2

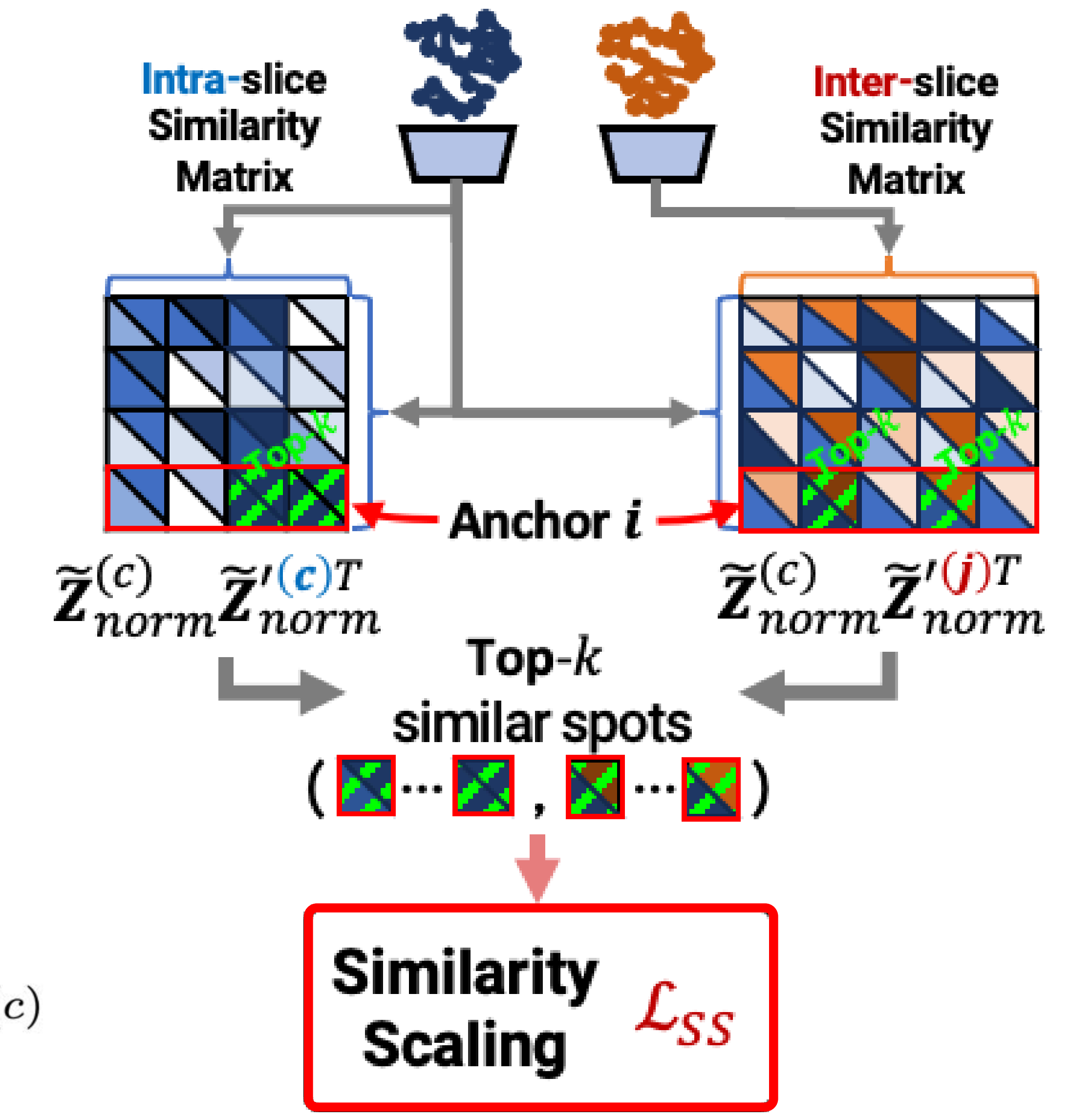
Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly **balances the similarity scales of inter- and intra-relationships** to mitigate batch effects
- **Grouping spots from the same spatial domain** while **distancing others** in latent space using PCL scheme to mitigate batch effects

Batch Effects: Experimental condition or noise ≫ Biological relevance



(b) Multi-Slice SRT



$$H = \tilde{Z}_{\text{norm}}(\tilde{Z}'_{\text{norm}})^T \in \mathbb{R}^{N_s \times N_s}$$

$$l_{\text{SS}}(H_i, \mathcal{G}^{(j)}) = (\text{Mean}(S^{(c)}_{\text{top}}) - \text{Mean}(S^{(j)}_{\text{top}}))^2, \text{for } i \in \mathcal{G}^{(c)}$$

$$\text{where} \quad S^{(c)}_{\text{top}} = \text{Top-}k_{l \in \mathcal{G}^{(c)}}(\underset{i\text{-th row}}{H_i}\underset{l\text{-th row}}{[l]}) = (a_1, a_2, \ldots, a_k),$$

$$S^{(j)}_{\text{top}} = \text{Top-}k_{l \in \mathcal{G}^{(j)}}(H_i[l]) = (b_1, b_2, \ldots, b_k)$$

$$\mathcal{L}_{\text{SS}} = \frac{1}{N_s(N_d - 1)} \sum_{i=1}^{N_s} \sum_{j=1}^{N_d} \mathbb{1}(i \notin \mathcal{G}^{(j)}) \cdot l_{\text{SS}}(H_i, \mathcal{G}^{(j)})$$

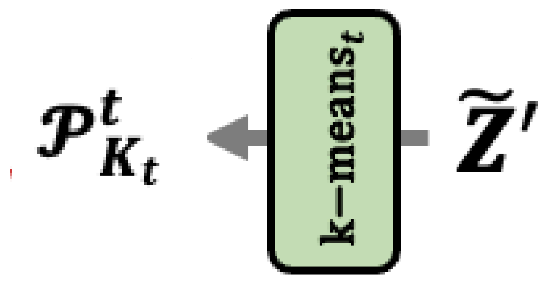
Key Ideas

METHODOLOGY 3

Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly **balances the similarity scales of inter- and intra-relationships** to mitigate batch effects
- **Grouping spots from the same spatial domain** while **distancing others** in latent space using PCL scheme to mitigate batch effects

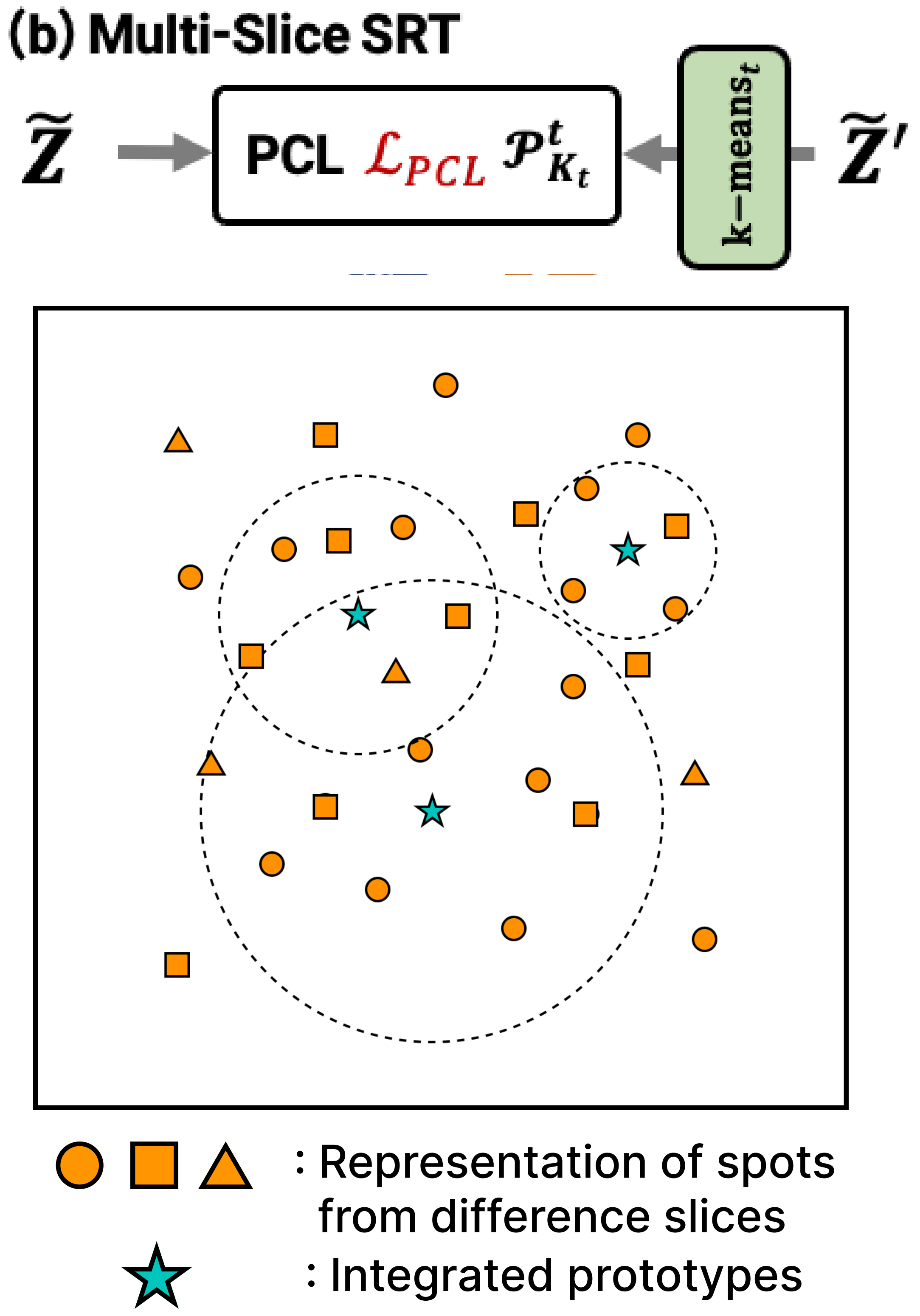
Prototype set of
t times k-means



$$l_{\mathrm{PCL}}(\tilde{Z}_i, P_{\mathrm{set}}) = \frac{1}{T}\sum_{t=1}^{T}\log\frac{e^{(\mathrm{sim}(\tilde{Z}_i, p^t_{\mathrm{map}_t(i)})/\tau)}}{\sum_{j=1}^{K_t} e^{(\mathrm{sim}(\tilde{Z}_i, p^t_j)/\tau)}},$$

$$\mathcal{L}_{\mathrm{PCL}} = -\frac{1}{N_s}\sum_{i=1}^{N_s} l_{\mathrm{PCL}}(\tilde{Z}_i, P_{\mathrm{set}}).$$

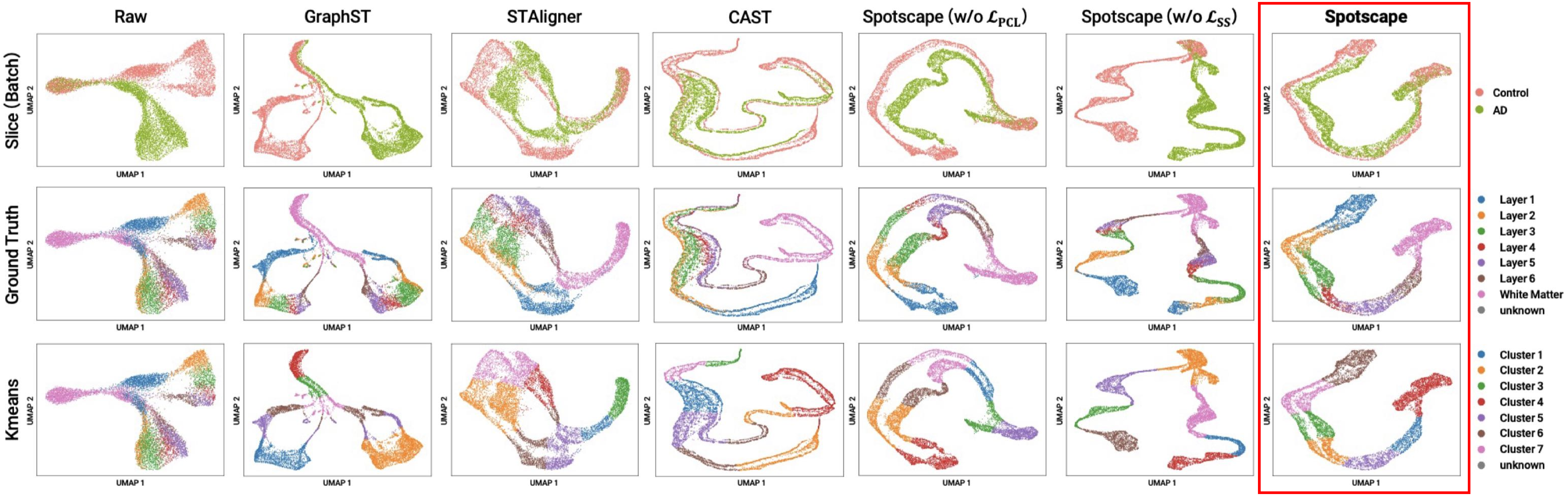
- To avoid the risk of obtaining inaccurate prototypes, the PCL loss gets involved to the training procedure after a **warm-up period**



Key Ideas METHODOLOGY 3

Key Ideas of Spotscape

- Capturing **global relationships** between cells by learning **robust similarities** with respect to different augmentations
- Explicitly **balances the similarity scales of inter- and intra-relationships** to mitigate batch effects
- **Grouping spots from the same spatial domain** while **distancing others** in latent space using PCL scheme to mitigate batch effects



- **Spotscape** effectively **integrates spot representations across slices** while preserving biological meaning

Various Downstream Tasks EXPERIMENTS

- **Extensive experiments on both single- and multi-slice tasks**

Single-Slice Tasks

- Spatial Domain Identification
- Trajectory Inference
- Denoising & Imputation

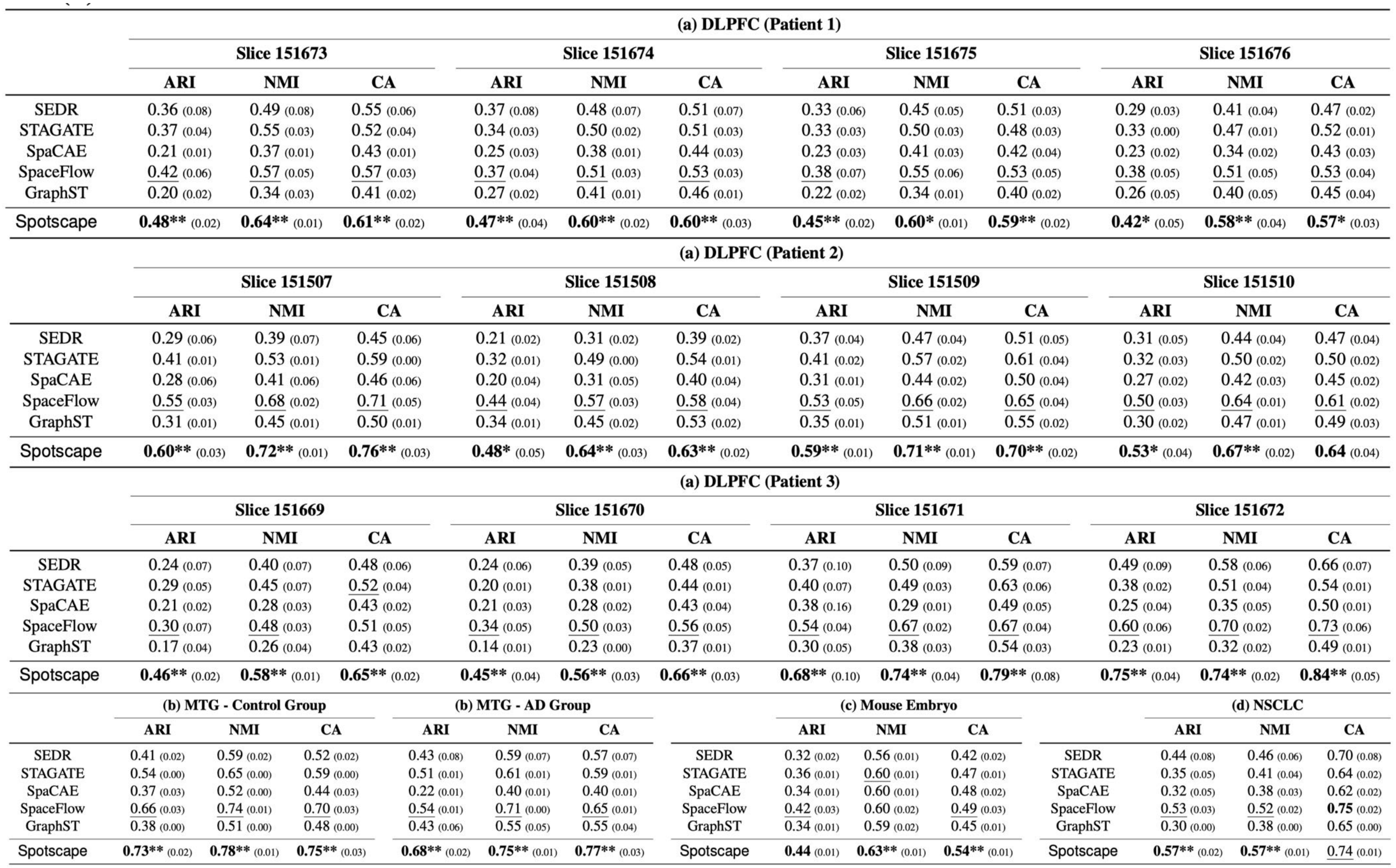
Multi-Slice Tasks

- Integration
- Alignment
- Differentially expressed gene (DEG) analysis

Various Downstream Tasks

EXPERIMENTS

Spatial Domain Identification



(a) DLPFC (Patient 1)

	Slice 151673			Slice 151674			Slice 151675			Slice 151676		
	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)
Spotscape	__0.48**__ (0.02)	__0.64**__ (0.01)	__0.61**__ (0.02)	__0.47**__ (0.04)	__0.60**__ (0.02)	__0.60**__ (0.03)	__0.45**__ (0.02)	__0.60*__ (0.01)	__0.59**__ (0.02)	__0.42*__ (0.05)	__0.58**__ (0.04)	__0.57*__ (0.03)

(a) DLPFC (Patient 2)

	Slice 151507			Slice 151508			Slice 151509			Slice 151510		
	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)
Spotscape	__0.60**__ (0.03)	__0.72**__ (0.01)	__0.76**__ (0.03)	__0.48*__ (0.05)	__0.64**__ (0.03)	__0.63**__ (0.02)	__0.59**__ (0.01)	__0.71**__ (0.01)	__0.70**__ (0.02)	__0.53*__ (0.04)	__0.67**__ (0.02)	__0.64__ (0.04)

(a) DLPFC (Patient 3)

	Slice 151669			Slice 151670			Slice 151671			Slice 151672		
	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)
Spotscape	__0.46**__ (0.02)	__0.58**__ (0.01)	__0.65**__ (0.02)	__0.45**__ (0.04)	__0.56**__ (0.03)	__0.66**__ (0.03)	__0.68**__ (0.10)	__0.74**__ (0.04)	__0.79**__ (0.08)	__0.75**__ (0.04)	__0.74**__ (0.02)	__0.84**__ (0.05)

	(b) MTG - Control Group			(b) MTG - AD Group		
	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)
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)
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)
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)
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)
Spotscape	__0.73**__ (0.02)	__0.78**__ (0.01)	__0.75**__ (0.03)	__0.68**__ (0.02)	__0.75**__ (0.01)	__0.77**__ (0.03)

	(c) Mouse Embryo		
	ARI	NMI	CA
SEDR	0.32 (0.02)	0.56 (0.01)	0.42 (0.02)
STAGATE	0.36 (0.01)	0.60 (0.01)	0.47 (0.01)
SpaCAE	0.34 (0.01)	0.60 (0.01)	0.48 (0.02)
SpaceFlow	0.42 (0.03)	0.60 (0.02)	0.49 (0.03)
GraphST	0.34 (0.01)	0.59 (0.02)	0.45 (0.01)
Spotscape	__0.44__ (0.01)	__0.63**__ (0.01)	__0.54**__ (0.01)

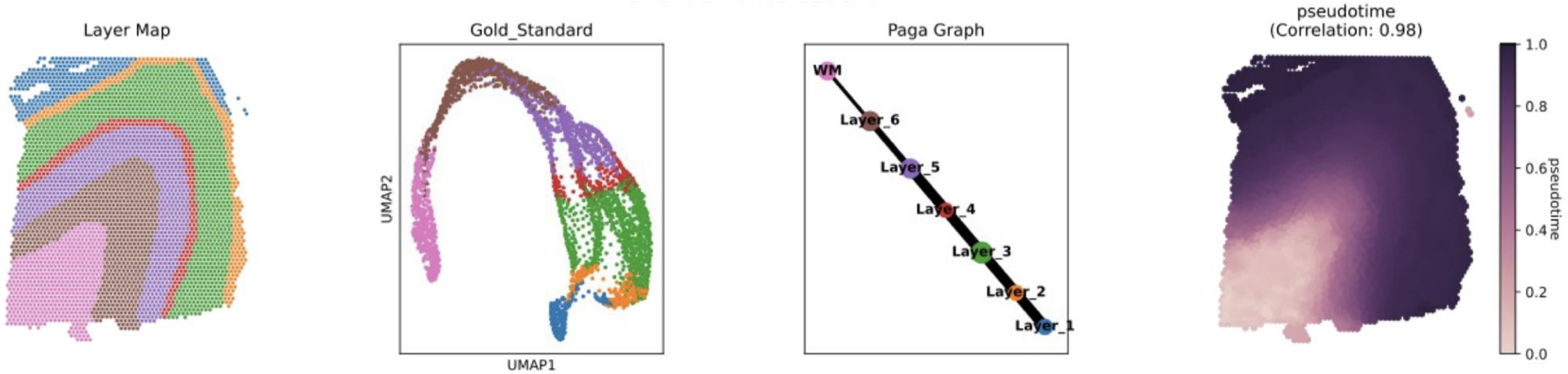
	(d) NSCLC		
	ARI	NMI	CA
SEDR	0.44 (0.08)	0.46 (0.06)	0.70 (0.08)
STAGATE	0.35 (0.05)	0.41 (0.04)	0.64 (0.02)
SpaCAE	0.32 (0.05)	0.38 (0.03)	0.62 (0.02)
SpaceFlow	0.53 (0.03)	0.52 (0.02)	__0.75__ (0.02)
GraphST	0.30 (0.00)	0.38 (0.00)	0.65 (0.00)
Spotscape	__0.57**__ (0.02)	__0.57**__ (0.01)	0.74 (0.01)

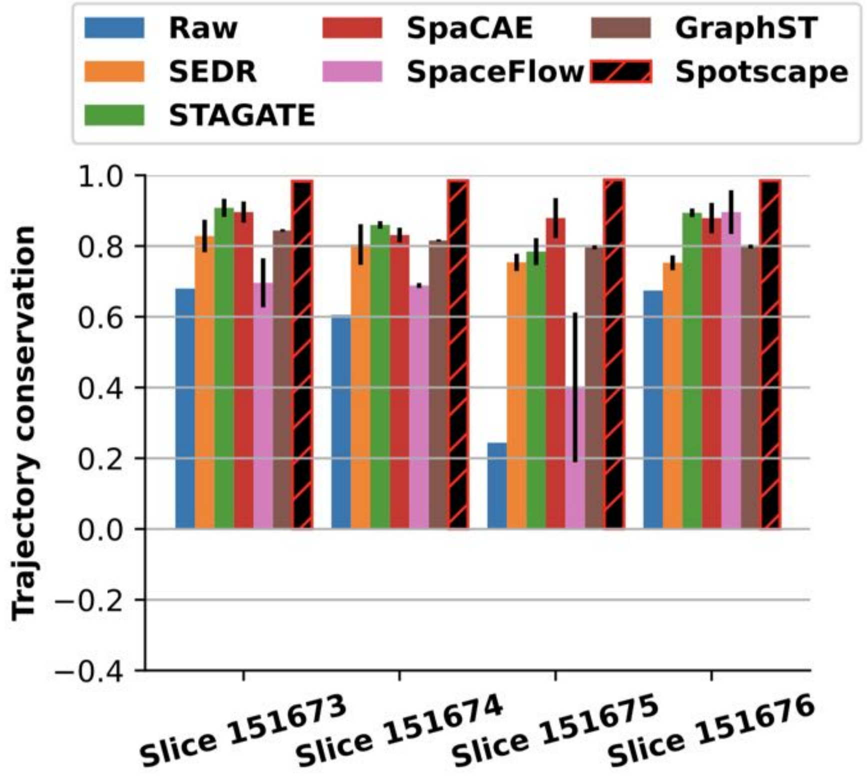
All experiments are repeated over 10 runs with different random seeds, and we report the mean and standard deviation of the results. For all experimental results, **Bold** indicates the best performance, underlining denotes the second-best, and an asterisk (*) marks statistically significant improvements of Spotscape over the top-performing baseline based on a paired t-test (**: $p < 0.01$, *: $p < 0.05$), with the numbers in parentheses representing the standard deviation.

- **Spotscape consistently outperforms** all baselines across 4 datasets and 16 slices
- **Moves beyond the limited insights** of local neighbor analysis by **capturing global contextual information**

Various Downstream Tasks **EXPERIMENTS**

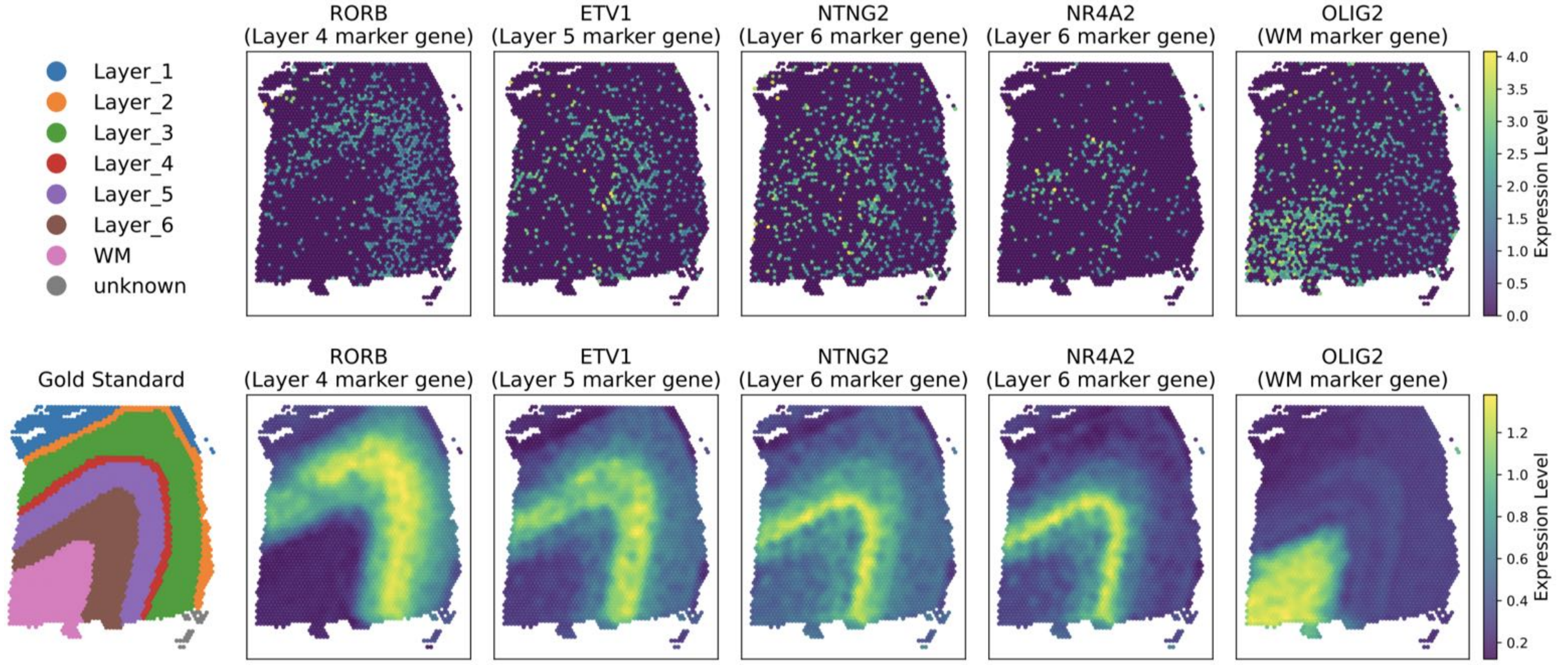
Trajectory Inference

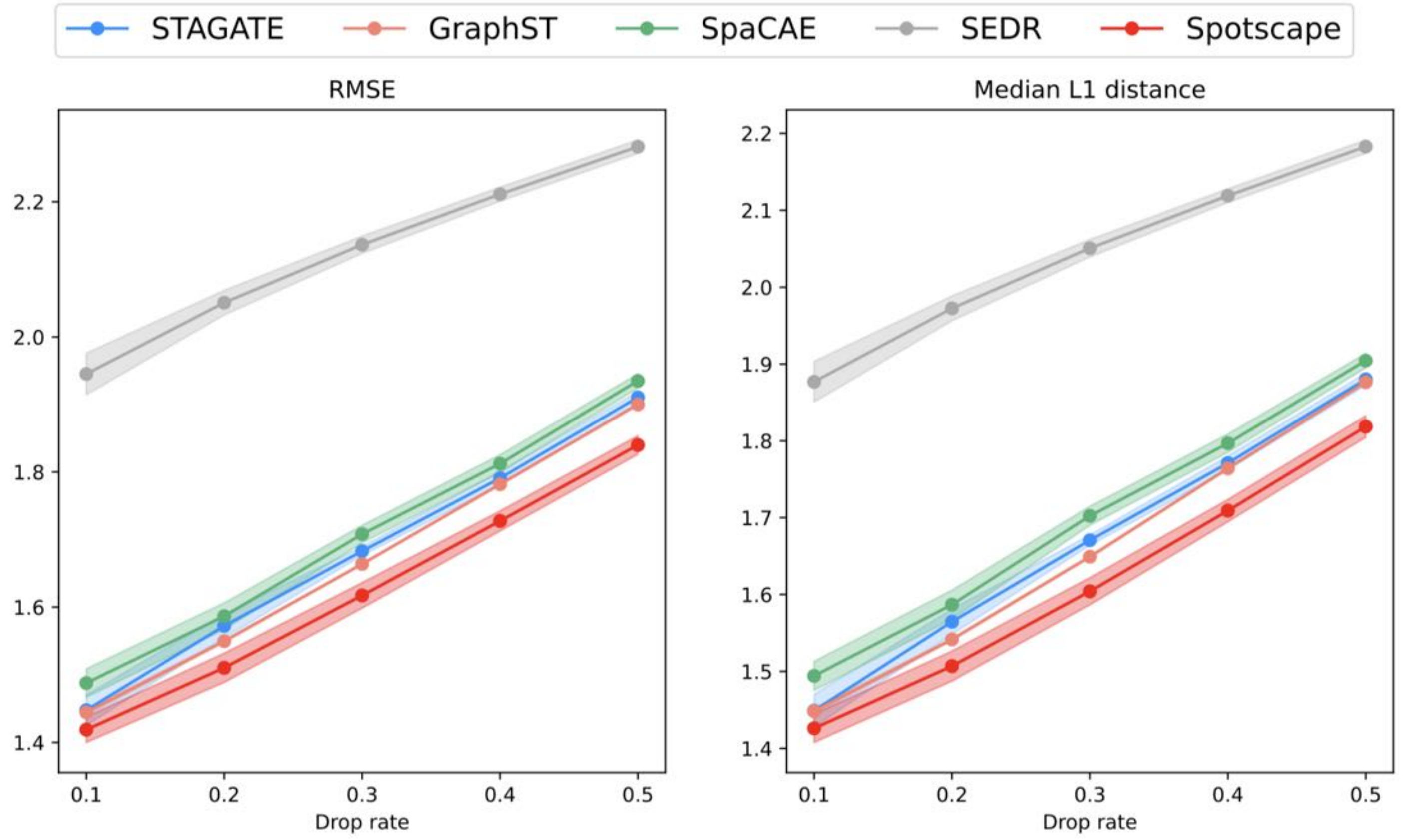




- **Spotscape** accurately captured the known developmental sequence, demonstrating its ability **to model biologically meaningful patterns**

Denoising & Imputation





- **Spotscape** clarifies **marker gene expression** for easier identification in noisy raw data in **denoising task**
- **Spotscape** achieves **best performance in imputation**, leading baselines on RMSE and L1-distance metrics

Various Downstream Tasks EXPERIMENTS

Integration

Table 2. Homogeneous integration performance on DLPFC data.

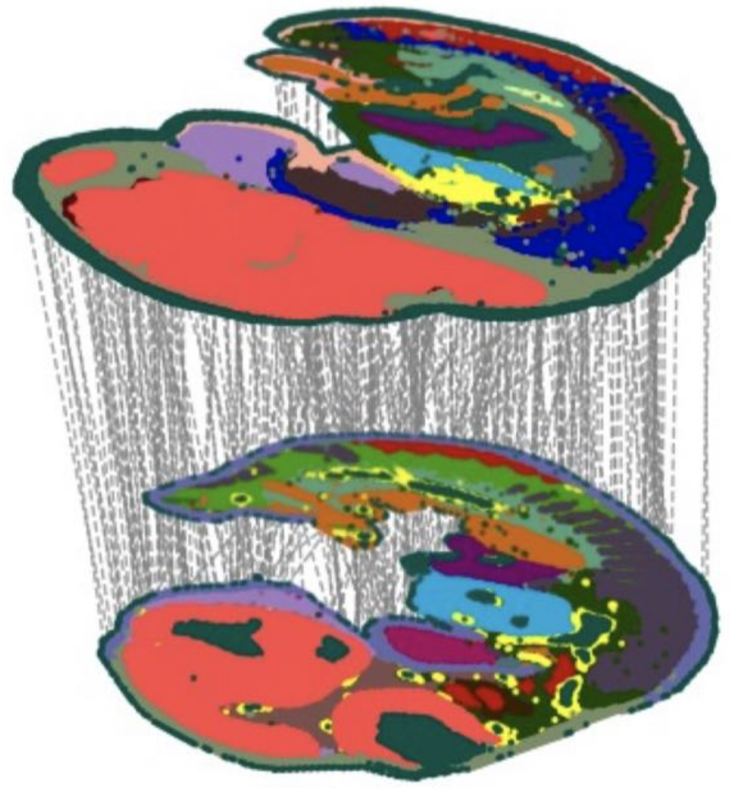
	Patient 1			Patient 2			Patient 3		
	ARI	NMI	CA	ARI	NMI	CA	ARI	NMI	CA
SEDR	0.38 (0.06)	0.49 (0.06)	0.56 (0.06)	0.32 (0.05)	0.44 (0.07)	0.48 (0.07)	0.43 (0.02)	0.51 (0.01)	0.56 (0.03)
STAGATE	0.31 (0.03)	0.46 (0.03)	0.49 (0.03)	0.30 (0.02)	0.46 (0.01)	0.48 (0.02)	0.31 (0.09)	0.43 (0.06)	0.54 (0.08)
SpaCAE	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)

Table 3. Heterogeneous integration performance on MTG data.

	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 (w/o $\mathcal{L}_{\text{PCL}}$)	0.61 (0.03)	0.71 (0.01)	0.70 (0.02)	0.67 (0.01)	0.03 (0.00)	0.79 (0.03)	0.50 (0.04)
Spotscape (w/o $\mathcal{L}_{\text{SS}}$)	0.47 (0.09)	0.60 (0.04)	0.59 (0.06)	0.24 (0.01)	0.00 (0.00)	0.63 (0.00)	0.00 (0.00)
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)

All experiments are repeated over 10 runs with different random seeds, and we report the mean and standard deviation of the results. For all experimental results, **Bold** indicates the best performance, underlining denotes the second-best, and an asterisk (*) marks statistically significant improvements of Spotscape over the top-performing baseline based on a paired t-test (**: $p < 0.01$, *: $p < 0.05$), with the numbers in parentheses representing the standard deviation.

Alignment



	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)

- **Homogeneous Integration:** Consistently outperforms all baseline methods when integrating multiple tissue slices from the same patient sample
- **Heterogeneous Integration**: Integrates diverse samples (e.g., Control vs. AD) by correcting batch effects, significantly outperforming competitors
- **Multi-slice Alignment**: Outperforms even specialized tools, successfully aligning slices across different developmental stages and technologies

Various Downstream Tasks **EXPERIMENTS**

DEG Analysis

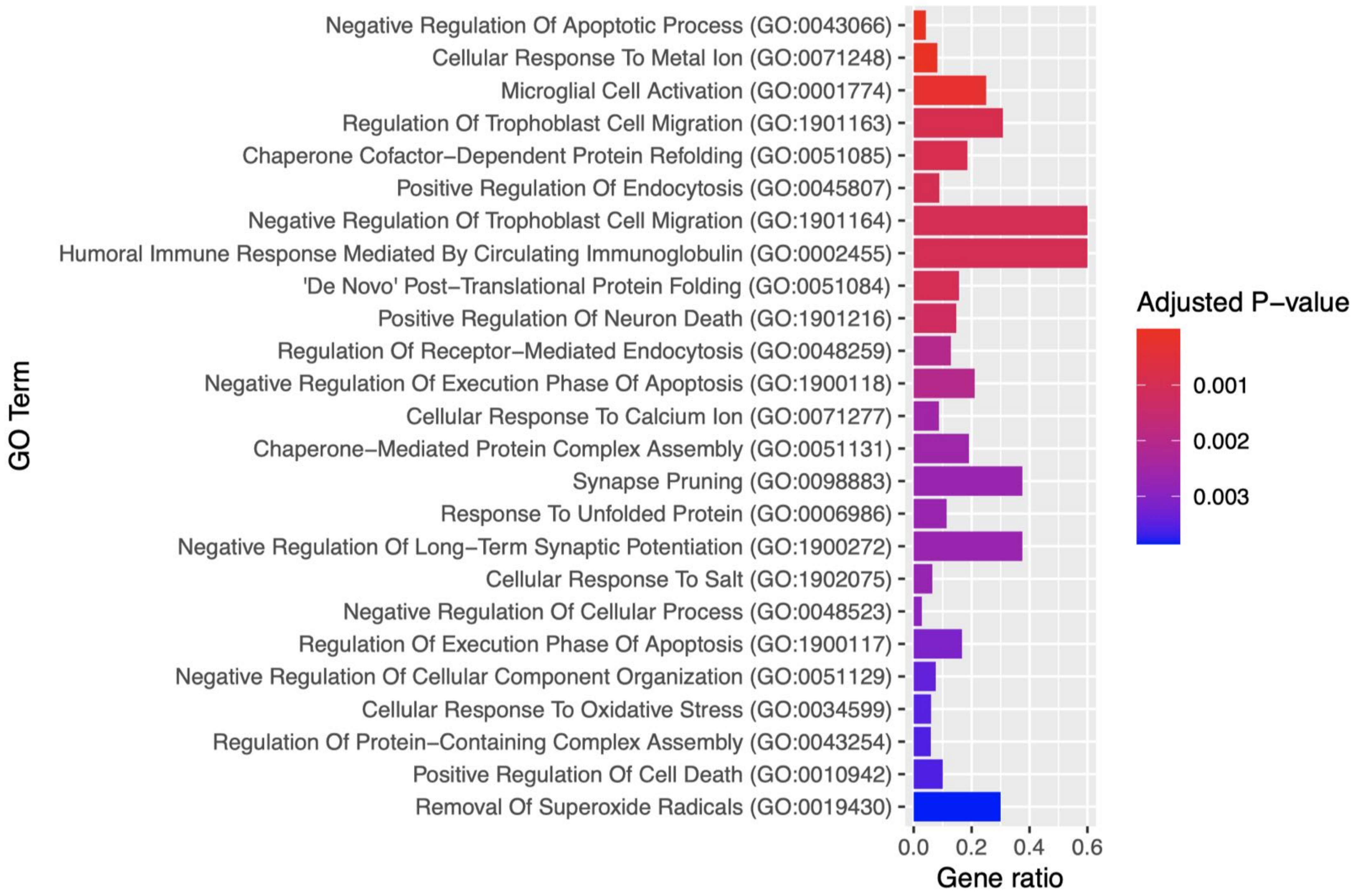


Figure 29. Top 25 biological process that DEGs between AD and Control enriched in a cluster assigned to layer 2.

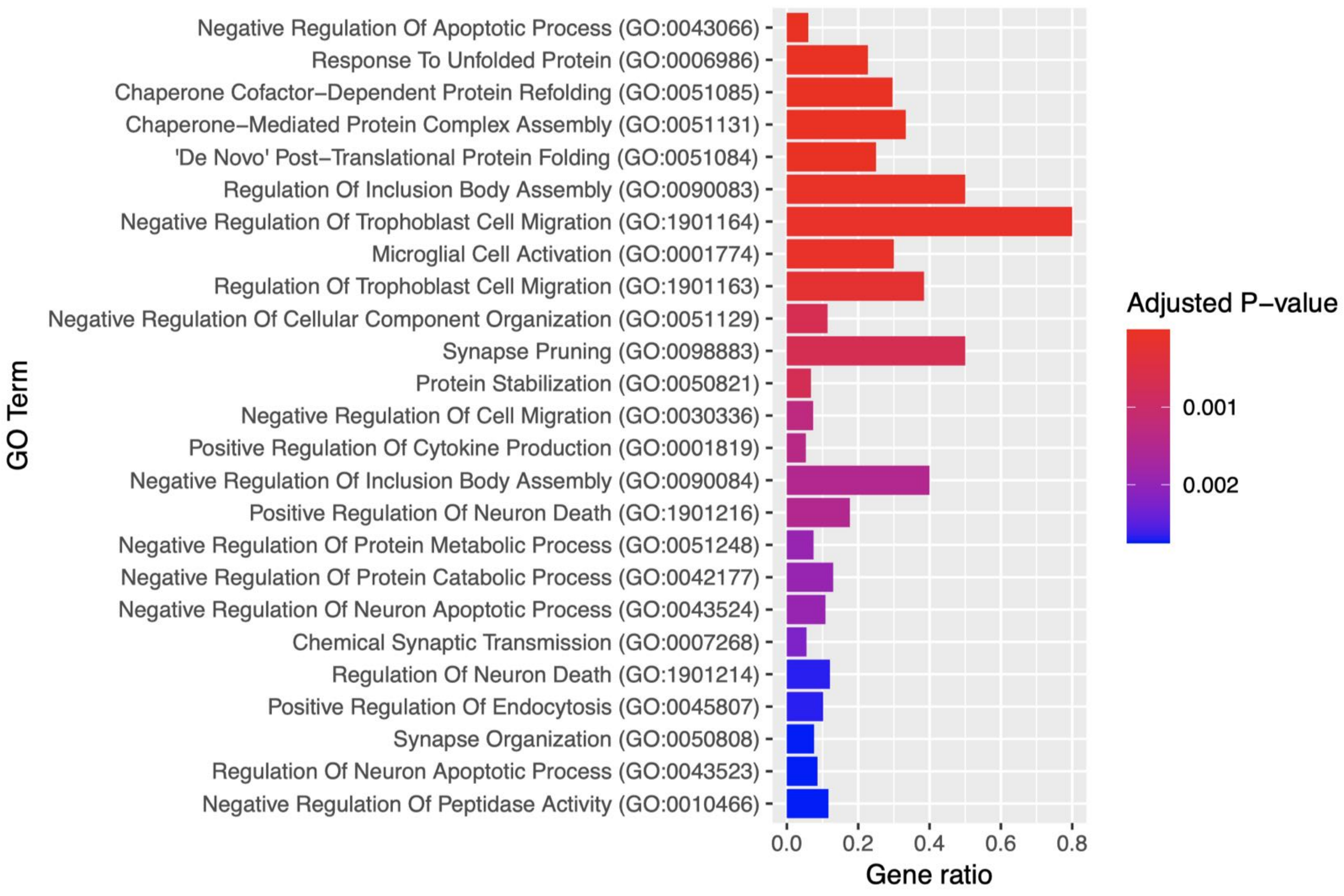
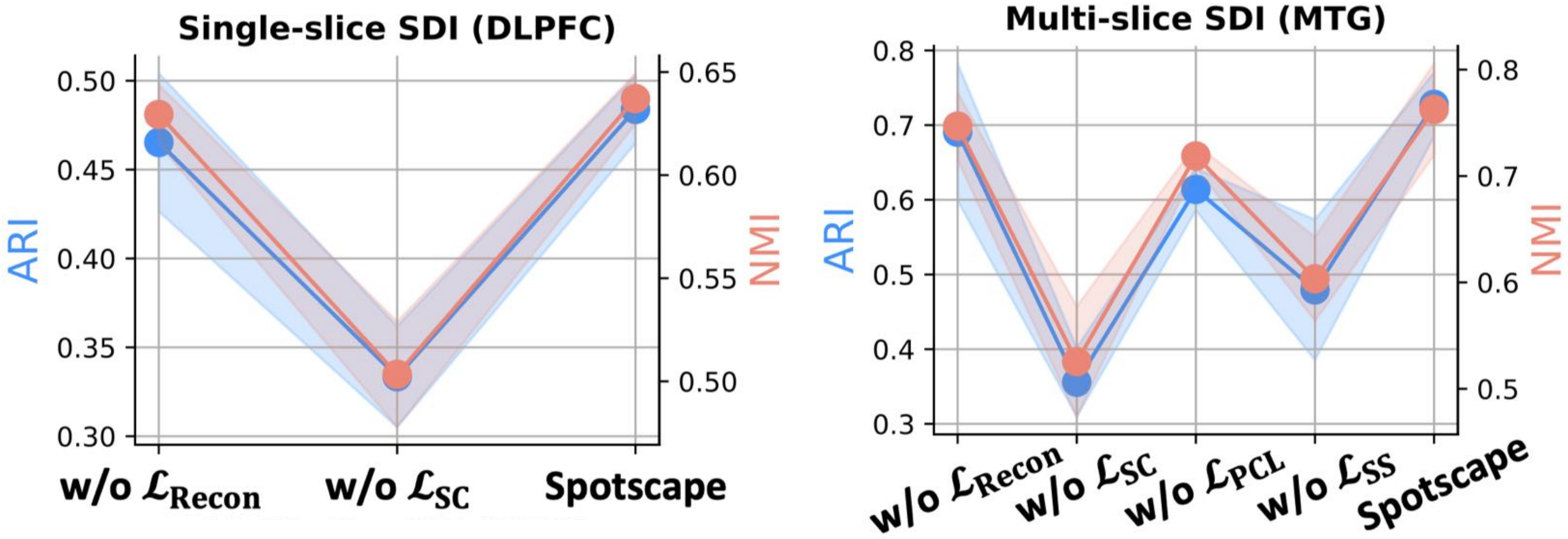


Figure 30. Top 25 biological process that DEGs between AD and Control enriched in a cluster assigned to layer 5.

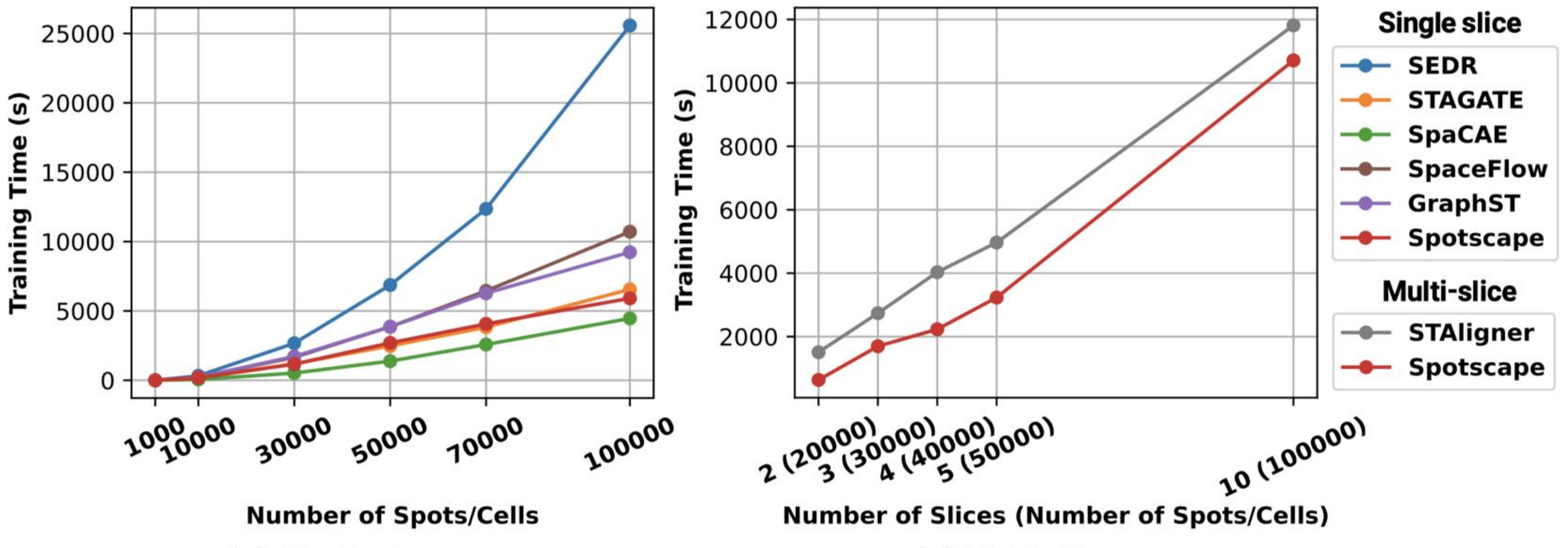
- **Biological Validation:** Gene analysis of Control vs. Alzheimer's samples validates **Spotscape**'s findings, as it correctly identified distinct early-stage (Layer 2) and late-stage (Layer 5) disease pathologies that align with known AD progression

Various Downstream Tasks EXPERIMENTS

Ablation study



Running-time analysis



(a) Single slice

(b) Multi-slice

- **Ablation Study: Removing key modules confirms their necessity;** the Similarity Telescope (global context) , Prototypical Contrastive Learning (grouping) , and Similarity Scaling (batch effect) are all critical for performance, while reconstruction loss primarily prevents degenerate solutions

- **Scalability**: **Spotscape** is fast and practical for high-throughput analysis , showing efficient training times on datasets up to 100,000 spots

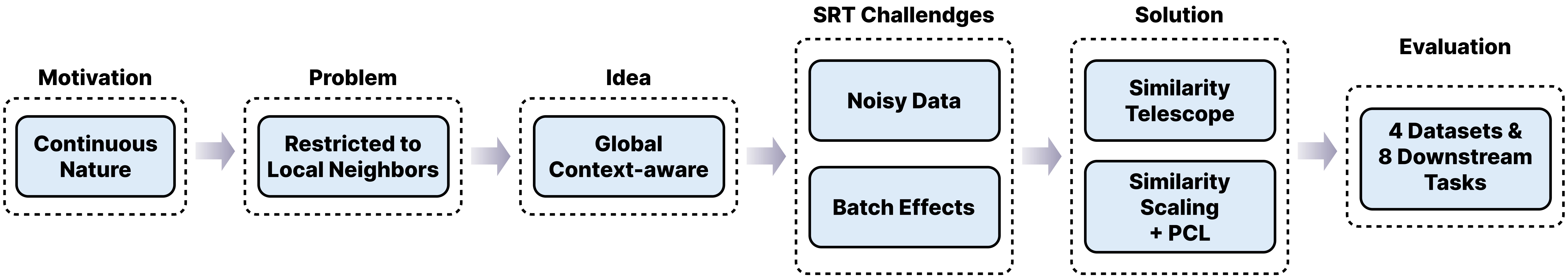
Conclusion

Limitations of previous works

- **Harnessing relationships between only spatially close (local) points** provides insufficient information to learn accurate representation
- Focusing on either spatial domain identification for a single slice or integration across multiple slices

Contributions of Spotscape

- **Reflecting global relationship information** by learning robust similarities w.r.t. augmentations
- **Alleviate batch effects** by explicitly balancing the intra- and inter- similarity



Global Context-aware Representation Learning for Spatially Resolved Transcriptomics

Paper: https://arxiv.org/abs/2506.15698

Code: https://github.com/yunhak0/Spotscape

Email: yunhak.oh@kaist.ac.kr

Paper

Code