ICML
International Conference
On Machine Learning

SpaEF: Spatially Resolved Transcriptomics Data Element-Wise Denoising Framework Powered by Large Models

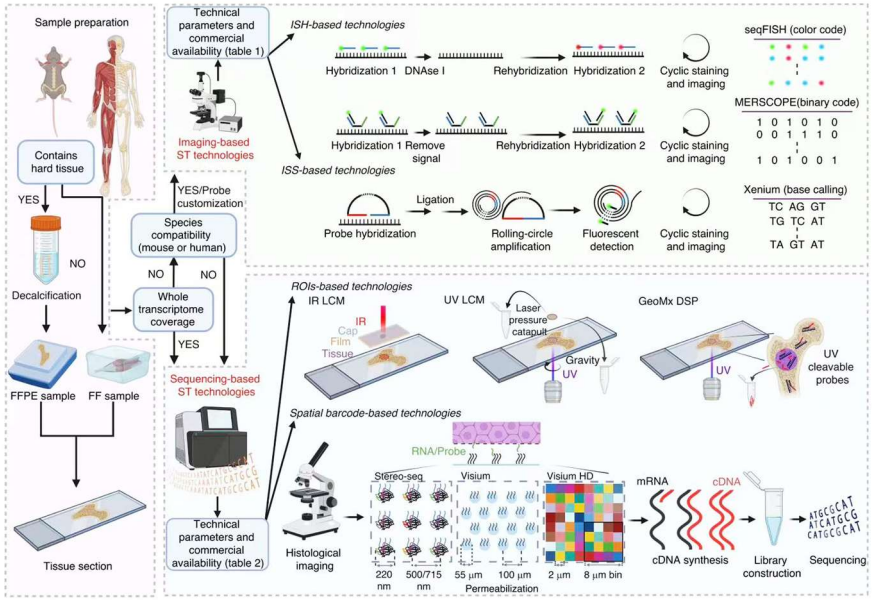
Zekuan Shang [1,2], Xiaosong Han [1,2], Liupu Wang [1,2], Wei Du [1,2],
Peng Zhao [1,2], Yuanshu Li [1,2], Yubin Xiao [1,2] & Xuan Wu [1,2] & You Zhou [1,2]
[1]College of Computer Science and Technology, Jilin University, Changchun, China
[2] Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Changchun, China.

Presenter: Zekuan Shang
PhD. Student

Background

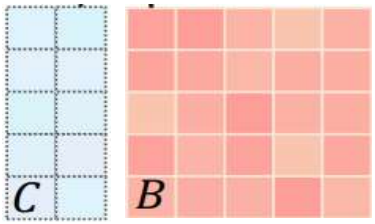
SRT tech and data

[1]



SRT Technology

SRT data



Spatially Resolved Transcriptomics (SRT) is a cutting-edge technique that provides biologists with rich insight into single-cell biology.

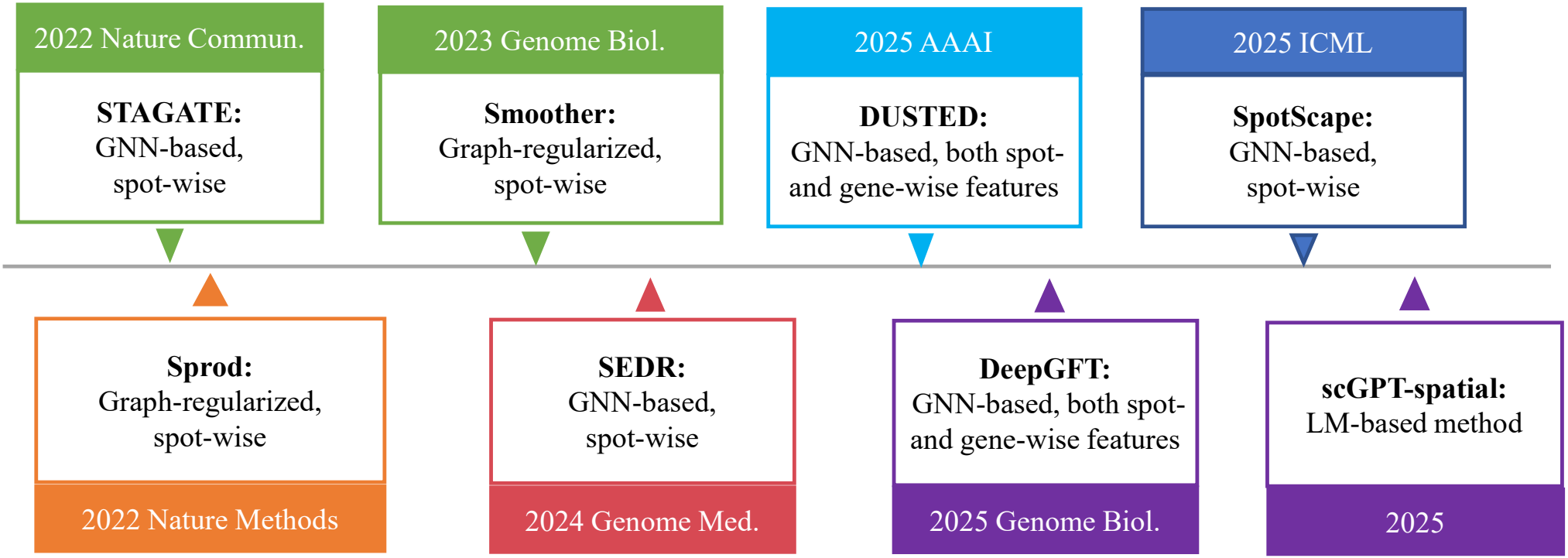
However, SRT data can be affected by biotechnological limitations (e.g., liquid phase diffusion) and operator-dependent variability, introducing significant noise that compromises downstream analyses.

SRT data consists of spot/cell spatial coordinates and an expression matrix (spots, genes), while the denoising of SRT data is to denoise the expression matrix.

1. Wang, H., Cheng, P., Wang, J. et al. Advances in spatial transcriptomics and its application in the musculoskeletal system. Bone Res 13, 54 (2025). https://doi.org/10.1038/s41413-025-00429-w

Background

Existing Methods



GNN or Graph-regularized method, constructing graphs of spots or genes.

Key Challenge

Challenges

1. Spurious similarity biases among spots;

Because of the narrow nature and staining-dependent variability of the histological images, incorporating histological images can amplify spurious similarity biases among spots; Despite this, directly discarding histological images would forfeit structural and morphological features that are often correlated with gene expression and that can provide useful priors for denoising

2. Insufficient capturing of gene relationships;

Gene relationships extend beyond simple co-expression, and often involve nonlinear mechanisms like path sharing and structural domain sharing, rendering co-expression analysis alone insufficient for gene graph construction

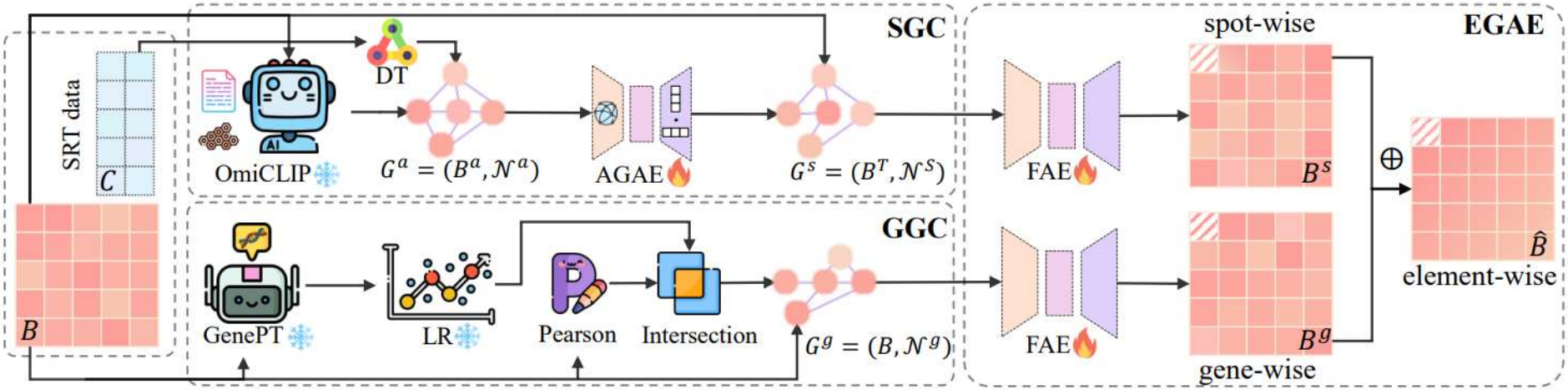
3. Rough denoising granularity.

During graph fusion, using matrix-wise weighting enforces a single global weight across all entries of the expression matrix, which limits the model's ability to adaptively denoise individual elements and thus potentially degrade denoising performance

Our Method

Framework: LMs-powered denoising

To address these three challenges, we propose the Large Models (LMs)-powered Spatially Resolved Transcriptomics data Element-wise Denoising Framework, named SpaEF, which comprises three modules, namely
Spot Graph Construction (SGC, with OmiCLIP LM to encode spots for challenge 1),
Gene Graph Construction (GGC, with GenePT LM to encode genes for challenge 2),
and Element-wise Graph Autoencoder (EGAE, with element-wise denoising for challenge 3).



Result

1. **Masking recovery** task on HOCWT

Table 1. RMSE values of various denoising methods across different mask proportions

Method	**Prop.=5%↓**	**Prop.=10%↓**	**Prop.=15%↓**	**Prop.=20%↓**	**Prop.=30%↓**	**Prop.=50%↓**
Sprod (Nat. Methods'22)	0.4103±0.0028	0.4155±0.0016	0.4232±0.0008	0.4327±0.0018	0.4509±0.0002	0.4949±0.0007
STAGATE (Nat. Commun.'22)	0.3739±0.0018	0.3759±0.0015	0.3817±0.0022	0.3882±0.0019	0.4074±0.0024	0.4643±0.0008
Smoother (Genome Biol.'23)	0.6600±0.0009	0.6600±0.0008	0.6599±0.0003	0.6597±0.0004	0.6596±0.0004	0.6598±0.0001
SEDR (Genome Med.'24)	1.0653±0.0021	1.0640±0.0019	1.0642±0.0026	1.0657±0.0034	1.0633±0.0048	0.9972±0.0125
DUSTED (AAAI'25)	0.3833±0.0012	0.3849±0.0028	0.3878±0.0032	0.3953±0.0043	0.4133±0.0051	0.4664±0.0051
DeepGFT (Genome Biol.'25)	0.3931±0.0014	0.3976±0.0011	0.4060±0.0012	0.4135±0.0017	0.4342±0.0020	0.4815±0.0044
SpaEF (ours)	0.3664±0.0004	0.3675±0.0005	0.3701±0.0011	0.3760±0.0042	0.3837±0.0007	0.4172±0.0034

Table 2. PCC values of various denoising methods across different mask proportions

Method	**Prop.=5%↑**	**Prop.=10%↑**	**Prop.=15%↑**	**Prop.=20%↑**	**Prop.=30%↑**	**Prop.=50%↑**
Sprod (Nat. Methods'22)	0.7016±0.0047	0.7021±0.0018	0.7015±0.0013	0.6983±0.0032	0.6991±0.0017	0.6953±0.0011
STAGATE (Nat. Commun.'22)	0.7602±0.0007	0.7595±0.0018	0.7595±0.0006	0.7591±0.0006	0.7575±0.0003	0.7528±0.0002
SEDR (Genome Med.'24)	0.2600±0.0014	0.2588±0.0006	0.2569±0.0011	0.2524±0.0016	0.2457±0.0018	0.2453±0.0016
DUSTED (AAAI'25)	0.7453±0.0008	0.7463±0.0006	0.7469±0.0006	0.7474±0.0008	0.7485±0.0006	0.7496±0.0004
DeepGFT (Genome Biol.'25)	0.7305±0.0007	0.7270±0.0006	0.7228±0.0001	0.7180±0.0003	0.7068±0.0008	0.6746±0.0004
SpaEF (ours)	0.7679±0.0011	0.7679±0.0016	0.7669±0.0016	0.7607±0.0041	0.7581±0.0030	0.7333±0.0059

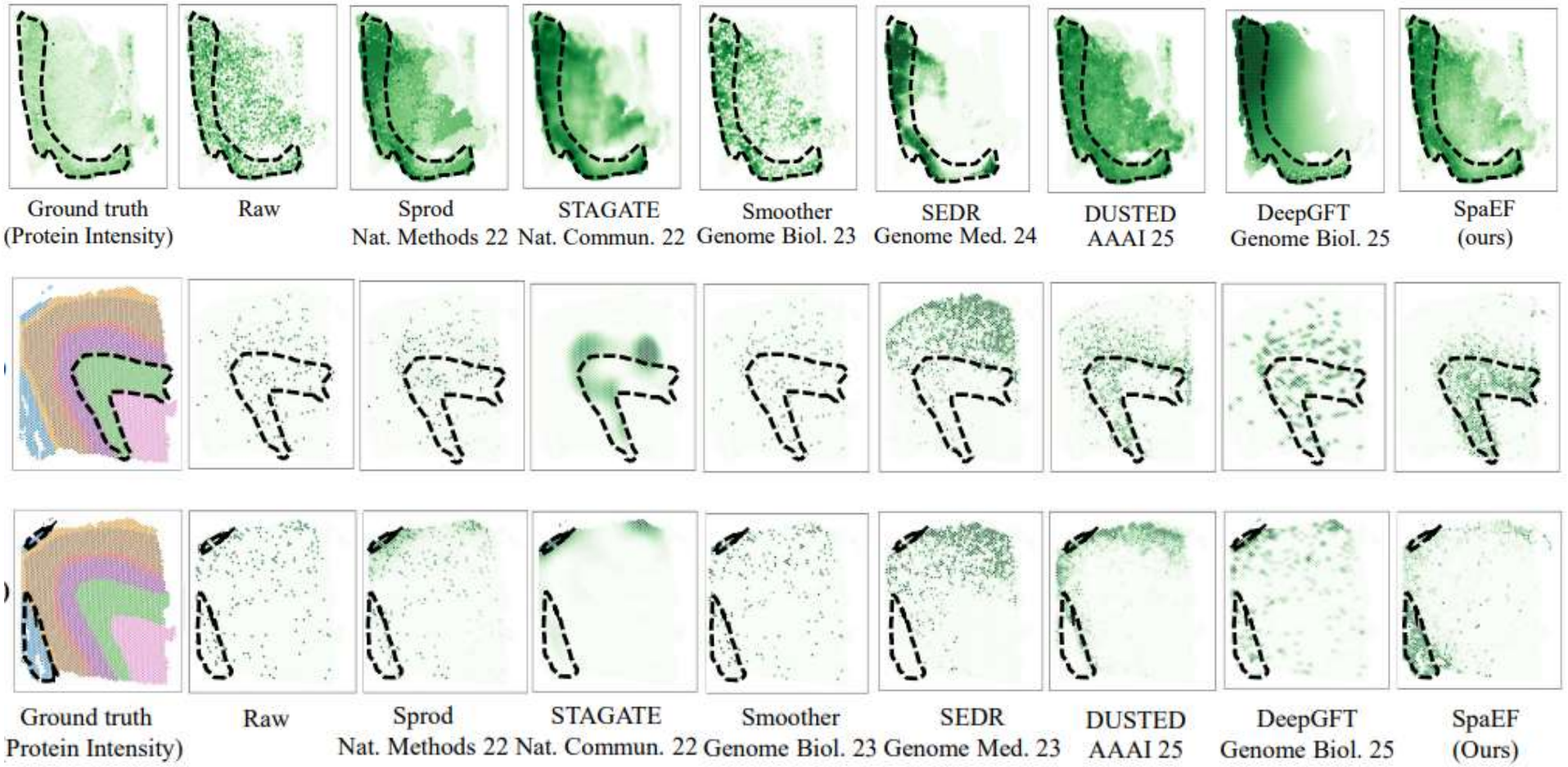
2. **Spatial Domain Identification and Clustering** on denoised DLPFC

Table 6. Results of Spatial Domain Identification and Clustering

Method	Domain.ARI	Domain.ACC	Clustering.ARI	Clustering.ACC
raw	0.262 ± 0.040	0.469 ± 0.034	0.040 ± 0.010	0.308 ± 0.005
Sprod (Wang et al., 2022)	0.274 ± 0.032	0.457 ± 0.003	0.270 ± 0.051	0.446 ± 0.038
STAGATE (Dong & Zhang, 2022)	0.338 ± 0.042	0.537 ± 0.024	0.282 ± 0.037	0.467 ± 0.033
Smother (Su et al., 2023)	0.360 ± 0.030	0.523 ± 0.037	0.199 ± 0.023	0.417 ± 0.018
DUSTED (Zhu et al., 2025a)	0.325 ± 0.045	0.511 ± 0.034	0.271 ± 0.031	0.488 ± 0.028
DeepGFT (Sun et al., 2025b)	0.306 ± 0.021	0.480 ± 0.021	0.295 ± 0.029	0.466 ± 0.026
SpaEF (ours)	0.353 ± 0.020	0.544 ± 0.019	0.297 ± 0.050	0.501 ± 0.052

Result

3. Spatial gene distribution on denoised HGBM and DLPFC



Result

4. Gene-protein & Gene-gene correlation on denoised HGBM and HDHBC

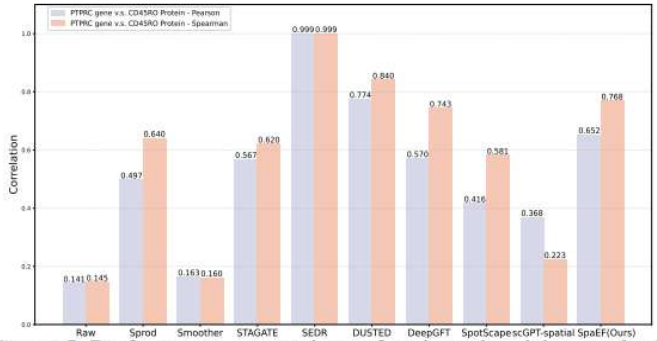


Figure 5. Performance comparison of various denoising methods on the HGBM dataset, quantified by PCC and SCC between CD44 and MYC gene expressions.

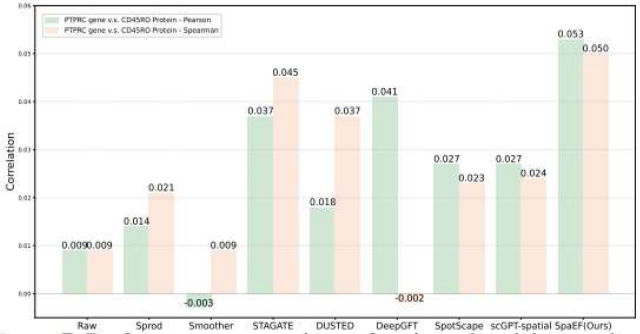


Figure 7. Performance comparison of various denoising methods on the HDHBC e, quantified by PCC and SCC between CD45RO protein intensity and PTPRC expression.

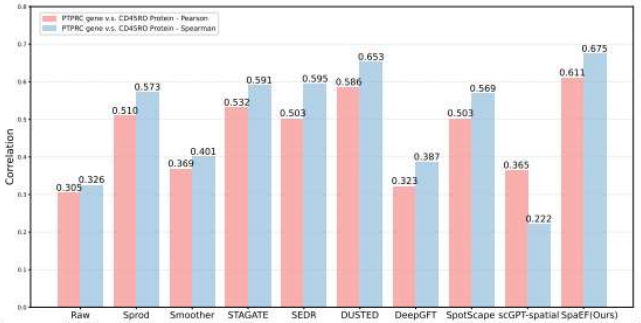


Figure 4. Performance comparison of various denoising methods on the HGBM dataset, quantified by PCC and SCC between PCNA protein intensity and PCNA expression.

5. Ablation Studies

Table 3. Ablation study on different SpaEF configurations

Method	PCC (Gap)	SCC (Gap)
w/o OmiCLIP (PCA)	0.531 (-13.09%)	0.605 (-10.37%)
w/o OmiCLIP (scGPT)	0.569 (-6.87%)	0.639 (-5.33%)
w/o OmiCLIP (Geneformer)	0.496 (-18.82%)	0.557 (-17.48%)
w/o GenePT (co-only)	0.540 (-11.62%)	0.614 (-9.04%)
w/o GenePT (scGPT)	0.548 (-10.31%)	0.620 (-8.15%)
w/o GenePT (Geneformer)	0.543 (-11.12%)	0.616 (-8.74%)
w/o AGAE	0.559 (-8.51%)	0.633 (-6.22%)
w/o Gene graph	0.539 (-11.78%)	0.615 (-8.89%)
w/o EWA (matrix-wise)	0.573 (-6.22%)	0.646 (-4.30%)
w/o EWA (cross-attention)	0.555 (-9.16%)	0.622 (-7.85%)
SpaEF	**0.611** (-)	**0.675** (-)

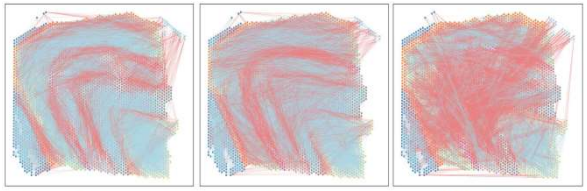


Figure 9. Visualization of spot graphs constructed using OmiCLIP, scGPT, and PCA on the DLPFC dataset.

Table 8. Modularity scores of different spot graphs

Ablation Method	Modularity Score
OmiCLIP	**0.543**
scGPT	0.491
PCA	0.405

Table 9. Functional precision of Gene Graph

Gene Graph	Precision
Co-expression only	0.94%
GenePT only	0.95%
GenePT&Co-expression (the adopted configuration)	1.46%

Table 10. Sensitivity analyses for different values of k_s and k_g

k_s	k_g	PCC	SCC	Clustering.ARI	Clustering.ACC
5	5	0.554	0.624	0.182	0.417
5	10	0.561	0.626	0.192	0.417
5	15	0.551	0.615	0.187	0.424
10	5	0.574	0.651	0.240	0.453
10	10	**0.611**	**0.675**	**0.297**	**0.501**
10	15	0.557	0.632	0.239	0.447
15	5	0.543	0.620	0.251	0.469
15	10	0.575	0.648	0.230	0.443
15	15	0.538	0.616	0.261	0.467

Table 11. Sensitivity analyses for different k_e values

k_e	PCC	SCC	Clustering.ARI	Clustering.ACC
15	0.569	0.644	0.253	0.478
25	0.532	0.609	0.234	0.465
50	**0.611**	**0.675**	**0.297**	**0.501**
65	0.548	0.625	0.268	0.480
75	0.565	0.635	0.240	0.465

Table 12. Sensitivity analyses for different threshold τ values

τ	PCC	SCC	Clustering.ARI	Clustering.ACC
0.4	0.545	0.614	0.251	0.462
0.5	**0.611**	**0.675**	**0.297**	**0.501**
0.6	0.540	0.611	0.274	0.485