



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Hyperbolic Contrastive Learning with Entailment for Spatial Transcriptomics

Daniela Vega^{*}[0009–0002–9731–7591], Paula Cárdenas^[0009–0005–1185–548X],
Hannah Ceballos^[0009–0005–5398–3770], Leonardo Manrique^[0009–0008–9428–6009],
and Pablo Arbeláez^[0000–0001–5244–2407]

Center for Research and Formation in Artificial Intelligence
Universidad de los Andes, Colombia

{d.vegaa,p.cardenasg,h.ceballos,dl.manrique,pa.arbelaez}@uniandes.edu.co

Abstract. Spatial Transcriptomics (ST) has transformed biomedical research by enabling the spatial mapping of gene expression across tissue sections. However, high operational costs, specialized equipment requirements, and sensitivity to experimental noise limit the accessibility and scalability of ST. Recent computer vision approaches aim to overcome these limitations by predicting spatial gene expression directly from histopathology images. While effective, current approaches often suffer from gene expression over-smoothing and overly uniform predictions across tissue regions, suggesting that further progress depends on learning representations that reflect the hierarchical and asymmetric structure of gene regulation and tissue morphology. To address these issues, we propose Hyperbolic Contrastive Learning with Entailment for Spatial Transcriptomics (HyCLoST), a hyperbolic contrastive learning model that captures the intrinsic hierarchical relationships within ST data. By leveraging hyperbolic geometry and a gene-to-image entailment loss, HyCLoST learns structured, biologically grounded representations that improve gene expression prediction accuracy, achieving a 6% reduction in MSE and an 8% increase in PCC across 26 ST datasets, over previous methods. Our source code is publicly available at <https://github.com/BCV-Uniandes/HyCLoST>.

Keywords: Spatial Transcriptomics · Hyperbolic Representation Learning · Gene Expression Prediction

1 Introduction

Spatial Transcriptomics (ST) enables the measurement of gene expression within the spatial architecture of tissues [14], effectively aligning histological appearance with local gene expression profiles and providing a rich view of tissue organization and disease mechanisms [21,13]. However, its high cost, specialized equipment

^{*} Corresponding author: d.vegaa@uniandes.edu.co

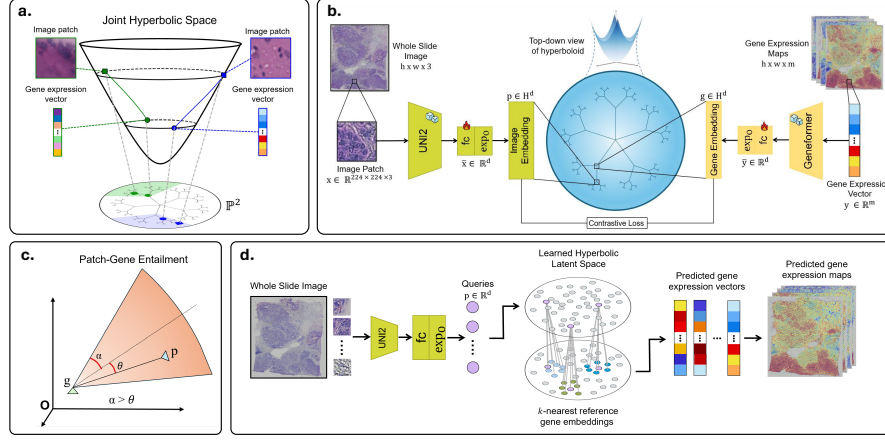


Fig. 1. (a) HyCLOST learns a joint hierarchical representation of histopathology patches and corresponding gene expression vectors via contrastive learning in hyperbolic space. (b) During training, image patches and paired gene expression vectors are encoded with UN2-h and Geneformer and projected into a shared hyperbolic space. (c) A patch–gene entailment objective models the asymmetric relationship between modalities. (d) In inference, a patch is encoded and projected into the learned hyperbolic space, and its prediction is obtained by averaging the k nearest gene embeddings.

requirements, and susceptibility to technical noise and missing measurements limit its scalability and the development of robust Artificial Intelligence (AI) models [10,25,12].

Numerous studies have explored computer vision methods to predict spatially resolved gene expression from histopathology images [9,29,23]. Existing approaches can be broadly categorized into architecture-focused and representation-focused methods. Architecture-focused models, such as ST-Net [9], HisToGene [18], EGN [28], EGN [27], and MERGE [7], refine network designs to capture local morphology and long-range dependencies, while BLEEP [26] introduces multimodal contrastive alignment in a shared Euclidean space. More recently, NH²ST [19] employs a dual-branch architecture combining cross-attention and contrastive learning: a query branch for paired histology patches and gene expression vectors, and a neighbor branch to capture multi-scale spatial context.

In these methods, performance improvements have largely plateaued, suggesting that further progress may depend more on representation geometry. Consequently, representation-focused methods such as TRIPLEX [4] and M2TGLGO [22] incorporate multi-resolution and biological functional priors, and HAGE [5] further injects gene co-expression relationships as prior knowledge into the image encoder via cross-attention, aligning gene-guided image and expression embeddings through contrastive learning, yet all still rely on Euclidean embeddings.

Biological systems are inherently hierarchical: gene expression constrains tissue organization through structured regulatory pathways that give rise to mor-

phology [8]. However, most existing approaches fail to model this hierarchy explicitly. Although DELST [3] introduces hyperbolic entailment, its hyperbolic component mainly serves as a regularizer, with multimodal alignment still performed in Euclidean space. To overcome these limitations, we propose Hyperbolic Contrastive Learning with Entailment for Spatial Transcriptomics (HyCLOST), a hierarchically structured contrastive framework that predicts gene expression from histopathology images using hyperbolic representations.

As shown in Fig. 1a, HyCLOST projects embeddings into a joint hyperbolic space, which naturally captures hierarchical relationships and enhances contrastive alignment. Inspired by [6], HyCLOST also incorporates an entailment loss to enforce a directional relationship between gene expression and image embeddings, reflecting the biological asymmetry in which molecular profiles give rise to tissue morphology. We build upon pretrained foundation models, UNI2-h [2] and Geneformer [24], to extract high-quality unimodal features. Furthermore, we address inconsistencies in prior evaluations by re-evaluating recent methods within the unified SpaRED benchmark [16], which standardizes comparison across 26 public datasets.

To summarize, our main contributions are: (1) We introduce HyCLOST, a hyperbolic contrastive learning framework that leverages hierarchical geometry to improve multimodal alignment and prediction accuracy in ST. (2) We propose a gene-to-image entailment loss that captures the directional relationships between gene expression profiles and tissue morphology. (3) We unify gene expression prediction evaluation under a standardized benchmark. We evaluate our model on the 26 datasets of the SpaRED benchmark [16], achieving an average 6% reduction in Mean Squared Error (MSE) and an 8% increase in Pearson Correlation Coefficient (PCC) compared to prior methods. To promote further research on ST, our source code is publicly available at <https://github.com/dvegaa00/HyCLOST>.

2 Method

The general workflow of HyCLOST is illustrated in Fig. 1b. We encode histopathology patches and gene expression profiles using pretrained foundation models (UNI2-h and Geneformer), project them to a shared latent space, and map them to a joint hyperbolic representation. Training optimizes two objectives: a symmetric contrastive loss using negative hyperbolic distance as logits, and an entailment loss that enforces directional gene-to-image consistency via semantic cones.

As shown in Fig. 1d, during inference, HyCLOST encodes test patches and projects them into the learned hyperbolic space. We obtain predictions via retrieval: for each test patch embedding, the model identifies the top- k nearest training gene embeddings under the Lorentzian inner product and averages their corresponding gene expression vectors. This non-parametric strategy is well suited to SpaRED’s heterogeneity across tissues and species, as retrieval adapts locally to nearby learned gene-expression contexts instead of forcing a

single parametric decoder to compress diverse output distributions. At scale, embeddings are precomputed once per dataset, so KNN cost grows linearly with dataset size and remains lightweight relative to feature encoding.

2.1 Data encoding and hyperbolic mapping

HyCLOST uses the Lorentz model [17] to represent hyperbolic geometry, where an n -dimensional hyperbolic space is embedded in the upper sheet of a two-sheeted hyperboloid in $\mathbb{R}^{n+1}$. Each vector $u \in \mathbb{R}^{n+1}$ is represented as $[u_{\text{space}}, u_{\text{time}}]$, with $u_{\text{space}} \in \mathbb{R}^n$ and $u_{\text{time}} \in \mathbb{R}$ denoting the spatial and time-like components [20]. The Lorentzian inner product between $u, v \in \mathbb{R}^{n+1}$ is defined as:

$$\langle u, v \rangle_{\mathbb{L}} = -u_{\text{time}}v_{\text{time}} + \sum_{j=1}^d u_j v_j, \quad (1)$$

where $\sum_{j=1}^d u_j v_j$, represents the Euclidean inner product $\langle u, v \rangle$ of u_{space} and v_{space} . Moreover, the hyperboloid curvature $-c$, is defined as:

$$\mathbb{L}^n = \left\{ u \in \mathbb{R}^{n+1} : \langle u, u \rangle_{\mathbb{L}} = -\frac{1}{c}, u_{\text{time}} = \sqrt{\frac{1}{c} + \|u_{\text{space}}\|^2} \right\}, \quad c > 0, \quad (2)$$

where $\|\cdot\|$ is the Euclidean norm. The curvature c is learned via $c = \exp(\tilde{c})$, with $\tilde{c}$ optimized jointly with the network. Thus, given a Euclidean embedding $z \in \mathbb{R}^n$, we map it from the tangent space at the origin to the hyperboloid via the exponential map (Eq. 3), with the time component determined by the hyperboloid constraint in Eq. 2:

$$z_{\text{space}} = \exp_{\mathbf{O}}(z) = \frac{\sinh(\sqrt{c} \|z\|)}{\sqrt{c} \|z\|} z \quad (3)$$

To construct the patch-gene hyperbolic space, each histopathology patch $x_i \in \mathbb{R}^{224 \times 224 \times 3}$ and its corresponding gene expression vector $y_i \in \mathbb{R}^m$ are encoded using pretrained models UNI2-h $f(\cdot)$ and Geneformer $h(\cdot)$, yielding embeddings $\hat{x}_i \in \mathbb{R}^{1536}$ and $\hat{y}_i \in \mathbb{R}^{768}$. These representations are linearly projected to a shared latent space of dimension $d = 512$, producing $\bar{x}_i, \bar{y}_i \in \mathbb{R}^{512}$. The projected embeddings are then mapped via Eq. 3 to the hyperbolic space, resulting in hyperbolic representations $p_i, g_i \in H^{512}$.

2.2 Contrastive learning with hyperbolic representations

After projecting image and gene embeddings into H^d , HyCLOST aligns both modalities using a symmetric contrastive objective based on hyperbolic geodesic distance. Following [6], the distance between embeddings p_i and g_i is computed as:

$$d_{\mathbb{L}}(p_i, g_i) = \frac{1}{\sqrt{c}} \cosh^{-1}(-c \langle p_i, g_i \rangle_{\mathbb{L}}). \quad (4)$$

We use the negative distance as similarity, defining logits for a given set of N paired samples as: $S_{ij}^{(P)} = -d_{\mathbb{L}}(p_i, g_j)$, $S_{ij}^{(G)} = -d_{\mathbb{L}}(g_i, p_j)$.

Each patch-gene pair (p_i, g_i) constitutes a positive match, while all other pairs in the set act as negatives. Consequently, the symmetric contrastive loss is:

$$\mathbb{L}_{con} = \frac{1}{2N} \sum_{i=1}^N \left[\ell(S_i^{(P)}, i) + \ell(S_i^{(G)}, i) \right], \quad (5)$$

where $\ell(S_i^{(P)}, i)$ denotes the cross-entropy loss with target i corresponding to the diagonal element. This objective pulls matched embeddings closer while pushing mismatched pairs apart, leveraging hyperbolic geometry to structure hierarchical relationships across modalities.

2.3 Gene-to-Image entailment

While the contrastive loss aligns paired image and gene embeddings symmetrically, the relationship between these modalities is inherently asymmetric: the visual appearance of a tissue arises from its underlying gene expression profile [8]. To model this directional dependency, we introduce an entailment loss inspired by [6], which enforces directional containment in hyperbolic space by encouraging image embeddings to lie within the entailment region induced by gene embeddings (gene to image). For each patch-gene pair (p_i, g_i) we compute the exterior angle at g_i in the hyperbolic triangle $(\mathbf{O}, p_i, g_i)$.

$$\text{ext}(g_i, p_i) = \cos^{-1} \left(\frac{p_{i,time} + g_{i,time} c \langle p_i, g_i \rangle_{\mathbb{L}}}{\|g_{i,space}\| \sqrt{(c \langle p_i, g_i \rangle_{\mathbb{L}})^2 - 1}} \right), \quad (6)$$

Intuitively, the closer p_i lies to this direction, the more consistent the histological appearance is with its underlying gene expression profile. As shown in Fig. 1c, each gene embedding g_i defines an entailment cone on the hyperboloid, parameterized by a half-aperture angle defined as:

$$\alpha_i = \sin^{-1} \left(\frac{2 r_{\min}}{\|g_i\| \sqrt{c}} \right), \quad (7)$$

where $r_{\min}$ denotes the minimum cone radius (set to 0.1) and c is the hyperbolic curvature. The entailment loss penalizes violations of this geometric constraint:

$$\mathcal{L}_{ent} = \frac{1}{2} \sum_{i=1}^N \max(0, \theta_i - \alpha_i), \quad (8)$$

where $\theta_i = \text{ext}(g_i, p_i)$.

Minimizing $\mathcal{L}_{ent}$ enforces a containment relation in which gene expression embeddings geometrically “entail” their patch counterparts. The overall training objective combines the contrastive and entailment components as $\mathcal{L} = \mathcal{L}_{con} + \mu \mathcal{L}_{ent}$, where μ controls the relative contribution of the entailment loss during training.

3 Results

3.1 Experimental Framework

Datasets. We use the 26 publicly available datasets from SpaRED [16], comprising 14 human and 12 mouse datasets across nine tissue types. These datasets are standardized for the gene expression prediction task of the top 128 or 32 most spatially autocorrelated genes. Following previous work [16,1], we pre-process the datasets using LGDiST [1]. We evaluate gene expression prediction using MSE and PCC computed only on genes directly measured by ST, excluding pre-completed values.

State-of-the-Art We compare HyCLOST against nine state-of-the-art methods: ST-Net [9], HisToGene [18], SEPAL [15], EGN [28], EGGN [27], BLEEP [26], TRIPLEX [4], MERGE [7], and M2TGLGO [22]. All methods were implemented following their official repositories.

For M2TGLGO, the original configuration limits training and evaluation to genes with available Gene2Vec embeddings, excluding on average 6 of 128 genes per dataset. We report results for the Gene2Vec-restricted variant (denoted by *) and for a Geneformer-based variant that covers the full gene set (reported without *). For TRIPLEX, 1-3 genes are predicted as constant across datasets, affecting evaluation. We report these results without *, and additionally provide metrics excluding constant gene predictions (denoted by *).

Implementation Details We train on a NVIDIA Quadro RTX 8000 using AdamW [11] (weight decay 0.2, batch size 128). The learning rate follows a linear warm-up (4,000 steps) to 5×10^{-4} with cosine decay over 120,000 total steps. To reduce computation, foundation model embeddings are precomputed and reused across experiments. For inference, we define k as 80.

3.2 Main Results

Tab. 1 compares HyCLOST with nine state-of-the-art methods across the 26 SpaRED datasets, resulting in eleven evaluated configurations due to starred variants. HyCLOST achieves the lowest average MSE and highest average PCC, with improvements of 6% and 8% over the second-best method, respectively. The improvement on both metrics indicates that HyCLOST accurately predicts absolute expression values and also captures spatially coherent expression patterns aligned with the ground truth. Moreover, statistical analysis confirms significant differences among methods (Friedman test, $\chi^2 = 117.02$, $p < 10^{-19}$). HyCLOST achieved the best average rank (2.79) and significantly outperformed 10 of 11 baseline configurations under paired one-sided Wilcoxon tests with Holm–Bonferroni correction ($\alpha = 0.05$).

Although recent approaches report competitive results, our evaluation in SpaRED shows that improvements over simpler models remain modest. In contrast, HyCLOST delivers consistent improvements in both MSE and PCC within an architecturally efficient, geometrically grounded framework. As shown in Tab. 1, it ranks first in 10 of 26 datasets, demonstrating robust performance across diverse tissues and acquisition settings.

Table 1. Average MSE and PCC over the 26 datasets in SpaRED. Symbol * denotes models missing a set of genes due to constraints (Section 3.1). The best results are highlighted in **bold**, and the second best are underlined. #Best indicates the number of datasets where each method achieves the best MSE. Avg. Rank denotes the average MSE-based rank across the 26 datasets, as used in the Friedman test. ✓ denotes statistically significant improvement of HyCLOST over the corresponding baseline in MSE under paired one-sided Wilcoxon tests with Holm–Bonferroni correction ($\alpha = 0.05$).

Model	MSE ($\downarrow$)	PCC ($\uparrow$)	#Best	Avg. Rank ($\downarrow$)	Sig.
TRIPLEX [4]	2.433 ± 1.712	0.388 ± 0.208	0	10.167	✓
HisToGene [18]	1.455 ± 0.581	0.166 ± 0.126	0	9.604	✓
TRIPLEX* [4]	1.126 ± 0.589	0.393 ± 0.211	1	7.833	✓
MERGE [7]	1.123 ± 0.736	0.339 ± 0.280	1	6.417	✓
BLEEP [26]	1.106 ± 0.588	0.317 ± 0.184	0	8.375	✓
M2TGLGO [22]	1.087 ± 0.417	0.174 ± 0.135	0	8.208	✓
SEPAL [15]	0.999 ± 0.841	0.393 ± 0.181	<u>4</u>	3.938	×
EGN [28]	0.982 ± 0.558	0.365 ± 0.154	2	5.021	✓
STNet [9]	0.974 ± 0.625	<u>0.400 ± 0.169</u>	<u>4</u>	<u>3.854</u>	✓
M2TGLGO* [22]	0.954 ± 0.380	0.294 ± 0.193	2	6.833	✓
EGGN [27]	<u>0.953 ± 0.442</u>	0.371 ± 0.173	2	4.958	✓
HyCLOST (Ours)	0.899 ± 0.436	0.431 ± 0.176	10	2.792	–

3.3 Qualitative Results

Fig. 2 compares the predictions of HyCLOST with those of the second-best performing model (EGGN, as observed in Sec. 3.2) and the two most recent methods (MERGE and M2TGLGO). The first row shows representative low-error cases. In this example, HyCLOST achieves the lowest MSE and better preserves the dynamic range and spatial patterns of the ground truth, whereas competing models tend to over-smooth or collapse predictions toward uniform values.

In the second row, all methods face more challenging cases and yield higher errors. Nevertheless, HyCLOST consistently achieves the lowest MSE and better captures extreme expression values, particularly in high-expression regions where competing models tend to flatten predictions. Overall, these results highlight HyCLOST’s ability to preserve spatial structure and dynamic range across diverse tissues and genes.

3.4 Ablation Experiments

We conduct ablation experiments to validate the geometric and biological principles underlying HyCLOST (Tab. 2). First, we assess the contribution of the entailment weight loss (Rows 1-3). The results confirm that enforcing a partial-order structure between image and gene embeddings produces a more informative hyperbolic space, thereby improving prediction accuracy.

Next, we evaluate the impact of using foundation models for both the gene and patch encoders. Replacing UNI2-h with a small Vision Transformer (ViT) increases MSE (Row 4 vs. Row 3), indicating less informative patch representations. Next, we keep UNI2-h as the patch encoder and add Geneformer as the

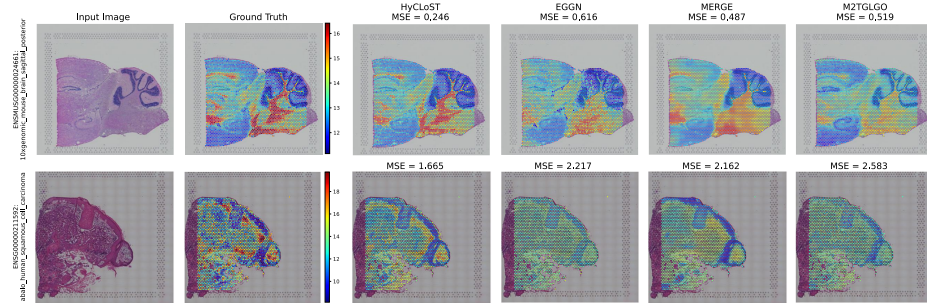


Fig. 2. Qualitative comparison of spatial gene expression predictions for two genes from two SpaRED datasets. For each gene, we show the input patch, gene expression ground truth, and predictions from HyCLOST, EGGN [27], MERGE [7], and M2TGLGO [22].

gene encoder. This change yields a significant performance improvement (Row 5 vs. Row 3).

Furthermore, we evaluate representation geometry. Replacing the hyperbolic space with a Euclidean one while keeping all other components unchanged substantially increases MSE (Row 6 vs. Row 5). This matched comparison isolates the contribution of the representation space and supports our hypothesis that hyperbolic geometry provides a better inductive bias than Euclidean embeddings for modeling the hierarchical structure induced by gene-regulatory organization.

Table 2. Average MSE over the 26 datasets from SpaRED in ablation experiments. The best results are highlighted in **bold**.

#	Entailment	Weight	Gene Encoder	Patch Encoder	Entailment Mode	Geometry	MSE ($\downarrow$)
1	0		None	UNI	None	Hyperbolic	0.958
2	0.2		None	UNI	Gene	Hyperbolic	0.942
3	0.5		None	UNI	Gene	Hyperbolic	0.933
4	0.5		None	ViT	Gene	Hyperbolic	1.206
5	0.5		Geneformer	UNI	Gene	Hyperbolic	0.899
6	0.5		Geneformer	UNI	Gene	Euclidean	1.043
7	0.5		Geneformer	UNI	Image	Hyperbolic	0.919

Finally, we analyze the directionality of the entailment relation. Comparing the configuration “gene entails image” (Row 5) with “image entails gene” (Row 7) shows that the former achieves a lower MSE. Although the average gap is moderate, the effect is consistent across datasets. Inspecting the per-dataset results underlying Tab. 2 reveals that the gene-to-image configuration attains lower MSE in 21 out of 26 datasets. This supports our analytical hypothesis that gene expression provides a more general molecular description that constrains plausible tissue morphologies, while histological appearance represents an observable outcome within a constrained set of possibilities [8].

4 Conclusions

In this paper, we present HyCLOST, a novel geometric learning framework for predicting gene expression from histopathology images. HyCLOST builds a joint hyperbolic space to model hierarchical relationships within ST data and guide contrastive alignment. A gene-to-image entailment loss captures the asymmetric dependence of tissue appearance on underlying gene expression, yielding predictions. Experiments on the 26 SpaRED datasets show average improvements of 6% in MSE and 8% in PCC over prior methods. Overall, our results indicate that refining the geometry and semantics of the representation space enables more expressive modeling of biological structure, highlighting a promising direction for spatial gene expression prediction.

Acknowledgments. This work was supported by Azure sponsorship credits granted by Microsoft’s AI for Good Research Lab.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Cárdenas, P., Manrique, L., Vega, D., Ruiz, D., Arbeláez, P.: Latent gene diffusion for spatial transcriptomics completion. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 901–910 (2025)
2. Chen, R.J., Ding, T., Lu, M.Y., Williamson, D.F., Jaume, G., Chen, B., Zhang, A., Shao, D., Song, A.H., Shaban, M., et al.: Towards a general-purpose foundation model for computational pathology. *Nature Medicine* (2024)
3. Chen, X., Huang, J.: Delst: Dual entailment learning for hyperbolic image-gene pre-training in spatial transcriptomics (2025), <https://arxiv.org/abs/2503.00804>
4. Chung, Y., Ha, J.H., Im, K.C., Lee, J.S.: Accurate spatial gene expression prediction by integrating multi-resolution features (2024), <https://arxiv.org/abs/2403.07592>
5. Dang, T.M., Li, H., Guo, Y., Ma, H., Jiang, F., Miao, Y., Zhou, Q., Gao, J., Huang, J.: HAGE: Hierarchical alignment gene-enhanced pathology representation learning with spatial transcriptomics. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer (2025)
6. Desai, K., Nickel, M., Rajpurohit, T., Johnson, J., Vedantam, R.: Hyperbolic image-text representations (2024), <https://arxiv.org/abs/2304.09172>
7. Ganguly, A., Chatterjee, D., Huang, W., Zhang, J., Yurovsky, A., Johnson, T.S., Chen, C.: Merge: Multi-faceted hierarchical graph-based gnn for gene expression prediction from whole slide histopathology images (2025), <https://arxiv.org/abs/2412.02601>
8. Gao, Y., Liang, J., Tian, M., Deng, W., Wang, L., Tao, S., Mou, T.: Histology image analysis of 13 healthy tissues reveals molecular-histological correlations. *Scientific Reports* **15**(1), 26812 (2025)
9. He, B., Bergenstråhle, L., Stenbeck, L., Abid, A., Andersson, A., Borg, Å., Maaskola, J., Lundeborg, J., Zou, J.: Integrating spatial gene expression and breast tumour morphology via deep learning. *Nature Biomedical Engineering* **4** (2020). <https://doi.org/10.1038/s41551-020-0578-x>

10. Jin, Y., Zuo, Y., Li, G., Liu, W., Pan, Y., Fan, T., Fu, X., Yao, X., Peng, Y.: Advances in spatial transcriptomics and its applications in cancer research. *Molecular Cancer* **23**(1), 129 (2024)
11. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014)
12. Li, H., Li, H., Zhou, J., Gao, X.: Sd2: spatially resolved transcriptomics deconvolution through integration of dropout and spatial information. *Bioinformatics* **38**(21), 4878–4884 (2022)
13. Liu, X., Peng, T., Xu, M., Lin, S., Hu, B., Chu, T., Liu, B., Xu, Y., Ding, W., Li, L., et al.: Spatial multi-omics: deciphering technological landscape of integration of multi-omics and its applications. *Journal of Hematology and Oncology* **17**(1), 72 (2024)
14. Marx, V.: Method of the year: spatially resolved transcriptomics. *Nature methods* **18**(1), 9–14 (2021)
15. Mejia, G., Cárdenas, P., Ruiz, D., Castillo, A., Arbeláez, P.: Sepal: spatial gene expression prediction from local graphs. In: *Proceedings of the IEEE/CVF International Conference on computer vision*. pp. 2294–2303 (2023)
16. Mejia, G., Ruiz, D., Cárdenas, P., Manrique, L., Vega, D., Arbeláez, P.: Enhancing gene expression prediction from histology images with spatial transcriptomics completion. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 91–101. Springer (2024)
17. Nickel, M., Kiela, D.: Learning continuous hierarchies in the lorentz model of hyperbolic geometry. In: *International conference on machine learning*. pp. 3779–3788. PMLR (2018)
18. Pang, M., Su, K., Li, M.: Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. *bioRxiv* (2021)
19. Qu, M., Wu, Y., Di, D., Gao, Y., Su, T., Song, Y., Fan, L.: Spatially gene expression prediction using dual-scale contrastive learning. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer (2025)
20. Ratcliffe, J.G.: *Foundations of hyperbolic manifolds*. Springer (2006)
21. Robles-Remacho, A., Sanchez-Martin, R.M., Diaz-Mochon, J.J.: Spatial transcriptomics: emerging technologies in tissue gene expression profiling. *Analytical Chemistry* **95**(42), 15450–15460 (2023)
22. Shi, H., Chi, C., Wan, P., Zhang, D., Shao, W.: Multi-modal topology-embedded graph learning for spatially resolved genes prediction from pathology images with prior gene similarity information. In: *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 20810–20819 (2025). <https://doi.org/10.1109/CVPR52734.2025.01938>
23. Song, T., Cosatto, E., Wang, G., Kuang, R., Gerstein, M., Min, M.R., Warrell, J.: Predicting spatially resolved gene expression via tissue morphology using adaptive spatial gnns. *Bioinformatics* **40**(Supplement_2), ii111–ii119 (2024)
24. Theodoris, C.V., Xiao, L., Chopra, A., Chaffin, M.D., Al Sayed, Z.R., Hill, M.C., Mantineo, H., Brydon, E.M., Zeng, Z., Liu, X.S., et al.: Transfer learning enables predictions in network biology. *Nature* **618**(7965), 616–624 (May 2023). <https://doi.org/10.1038/s41586-023-06139-9>
25. Wan, X., Xiao, J., Tam, S.S.T., Cai, M., Sugimura, R., Wang, Y., Wan, X., Lin, Z., Wu, A.R., Yang, C.: Integrating spatial and single-cell transcriptomics data using deep generative models with spatialscope. *Nature Communications* **14**(1), 7848 (2023)

26. Xie, R., Pang, K., Chung, S.W., Perciani, C.T., MacParland, S.A., Wang, B., Bader, G.D.: Spatially resolved gene expression prediction from h&e histology images via bi-modal contrastive learning (2023), <https://arxiv.org/abs/2306.01859>
27. Yang, Y., Hossain, M.Z., Stone, E., Rahman, S.: Spatial transcriptomics analysis of gene expression prediction using exemplar guided graph neural network. *Pattern Recognition* **145**, 109966 (Jan 2024). <https://doi.org/10.1016/j.patcog.2023.109966>
28. Yang, Y., Hossain, M.Z., Stone, E.A., Rahman, S.: Exemplar guided deep neural network for spatial transcriptomics analysis of gene expression prediction. In: *Proceedings - 2023 IEEE Winter Conference on Applications of Computer Vision, WACV 2023* (2023). <https://doi.org/10.1109/WACV56688.2023.00501>
29. Zeng, Y., Wei, Z., Yu, W., Yin, R., Yuan, Y., Li, B., Tang, Z., Lu, Y., Yang, Y.: Spatial transcriptomics prediction from histology jointly through transformer and graph neural networks. *Briefings in Bioinformatics* **23** (2022). <https://doi.org/10.1093/bib/bbac297>