Meta-Path based method for Cell Phenotyping on spatially resolved single-cell data

Varian Zhou, Yangyuan Zhang, John Cost

December 13th, 2024

1 Introduction

1.1 Spatial Omics Technologies

Spatial omics technologies have revolutionized the study of tissue architecture and cellular interactions by providing spatially resolved molecular data at single-cell resolution. Broadly, these technologies can be classified into two major categories: image-based and sequencing-based methods, each offering distinct advantages and having experienced significant advancements in recent years.

Image-based methods, such as CODEX [1] (Co-Detection by Indexing) and MERFISH [2] (Multiplexed Error-Robust Fluorescence In Situ Hybridization), employ multiplexed imaging to simultaneously detect dozens to hundreds of protein or RNA markers in situ. These approaches retain the spatial coordinates of individual cells, enabling the detailed mapping of cellular distributions and interactions within tissues. For example, CODEX utilizes fluorescently labeled antibodies to profile multiple proteins, while MERFISH encodes transcripts with combinatorial fluorescence barcodes, allowing for high-throughput spatial transcriptomic profiling.

In parallel, sequencing-based technologies, including Spatial Transcriptomics [3], Slide-seq [4], and High-Definition Spatial Transcriptomics (HDST) [5], combine next-generation sequencing with spatial barcoding to yield spatially resolved gene expression profiles. These sequencing-based methods complement image-based approaches by providing a broader and more unbiased overview of the transcriptome.

Both image-based and sequencing-based spatial omics technologies have now reached the critical capability of resolving individual cells within their spatial context. This high spatial resolution is transforming our understanding of cellular interactions, tissue architecture, and the molecular underpinnings of various biological processes and diseases.

1.2 GNNs in Spatial Omics Data and Cell Phenotyping

Graph Neural Networks (GNNs) have emerged as powerful tools for modeling complex relational data, making them particularly well-suited for spatial omics. In these settings, cells can be represented as nodes in a graph, and edges can encode spatial proximity or molecular similarity, capturing intricate relationships within the tissue. GNNs have been successfully applied to various tasks in spatial omics, including cell-cell interaction modeling [6], spatial clustering [7, 8], and tissue structure representation learning [9].

A key challenge in spatial omics is cell phenotyping, where existing but limited cell-type annotations are leveraged to annotate uncharacterized cells computationally. A representative work is STELLAR [10], which transfers cell-type labels from well-annotated reference datasets to unannotated datasets using GNNs. However, current approaches typically construct graphs based solely on spatial proximity, overlooking the potential benefits of integrating molecular similarity into the graph structure.

1.3 Bringing Spatial and Functional Similarity Together with Heterogeneous Graphs

To address this limitation, we propose incorporating both spatial and functional similarities into a heterogeneous graph model for cell phenotyping. Heterogeneous GNNs offer the flexibility to model nodes and edges of different types, naturally integrating spatial and molecular relationships. By leveraging metapaths—sequences of edges that capture composite relationships—we can achieve more interpretable message passing across heterogeneous neighborhoods.

In this work, we focus on Heterogeneous Graph Attention Networks (HAN) [11] as a means to fuse spatial and functional information. By exploiting multiple meta-paths and learning their relative importance, HAN can potentially enhance the accuracy of cell phenotyping tasks. The ability to incorporate and weight distinct relational contexts both improves interpretability and opens new avenues for understanding the cellular and molecular organization of tissues.

2 Methods

2.1 Heterogeneous Graph Construction

Given the spatial transcriptomic data and the count matrix $\mathbf{X}$, which describes the cell expression profile, we construct a heterogeneous graph as follows. First, we compute the similarity matrix:

$$\mathbf{A}_{sim} = \text{normalize}(\mathbf{X}) \cdot \text{normalize}(\mathbf{X})^{\top}. \tag{1}$$

For each cell, we retain only the top-k similarity values in its corresponding row, masking all others. This ensures that each node is connected to its k most

similar neighbors, based on gene expression profiles. Next, we construct the spatial proximity graph $\mathbf{A}_{spa}$ by calculating pairwise 2-D distances among cells and retaining only edges between cells whose distance is less than a predefined threshold *distance_threshold*.

Finally, we integrate the normalized expression data and the two adjacency matrices (functional similarity and spatial proximity) into a heterogeneous graph:

$$\mathbf{G} = (\mathbf{X}, \mathbf{E}, \mathcal{T}), \tag{2}$$

where $\mathcal{T}_e \in \{\textit{spatially\_close\_to}, \textit{functionally\_similar\_to}\}$ characterizes the edge types.

2.2 Metapath-based Methods

2.2.1 Metapath

Given a heterogeneous graph $\mathcal{G} = (\mathbf{V}, \mathbf{E}, \mathcal{T})$, a meta-path $\mathcal{P}$ is defined as a sequence of relations that connect nodes of potentially different types. If $R_1, R_2, \ldots, R_n \in \mathcal{T}^E$ are distinct edge types with adjacency matrices $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n$, then the meta-path neighborhood structure represented by the composite relation $R_1 \circ R_2 \circ \cdots \circ R_n$ is given by:

$$\mathbf{A}_{\mathcal{P}} = \mathbf{A}_n \cdots \mathbf{A}_2 \mathbf{A}_1. \tag{3}$$

Each meta-path has a specific semantic meaning derived from its constitutive relations. In our context, the two fundamental relations are *spatially_close_to* and *functionally_similar_to*, enabling meta-paths that capture, for example, cells with similar gene expression profiles as those of a cell's spatial neighbors, or cells that are spatially close to cells functionally similar to the reference cell.

2.2.2 Heterogeneous Graph Attention (HAN) Networks

Heterogeneous Graph Attention Network (HAN) [11] generalizes the graph attention mechanism to heterogeneous graphs. It incorporates both node-level attention, which focuses on aggregating information from meta-path-specific neighbors, and semantic-level attention, which identifies the most informative meta-paths.

Formally, given a set of meta-paths $\{\mathcal{P}_1, \mathcal{P}_2, \ldots\}$ and their corresponding adjacency matrices $\{\mathbf{A}_{\mathcal{P}_1}, \mathbf{A}_{\mathcal{P}_2}, \ldots\}$, HAN first learns meta-path-specific embeddings. For a meta-path $\mathcal{P}_i$, node-level attention scores are computed to select and aggregate features from relevant neighbors:

$$\mathbf{h}_{v,\mathcal{P}_i} = \sum_{u \in N_v^{\mathcal{P}_i}} \alpha_{vu}^{\mathcal{P}_i} \mathbf{h}'_u, \tag{4}$$

where $\alpha_{vu}^{\mathcal{P}_i}$ is the learned attention weight for neighbor u under meta-path $\mathcal{P}_i$, and $\mathbf{h}'_u$ is the transformed feature embedding of node u.

Next, a semantic-level attention mechanism is applied to integrate the meta-path-specific embeddings $\{\mathbf{h}_{v,\mathcal{P}_1}, \mathbf{h}_{v,\mathcal{P}_2}, \dots\}$ into a final embedding:

$$\mathbf{h}_v = \sum_i \beta_i \mathbf{h}_{v,\mathcal{P}_i}, \tag{5}$$

where β_i denotes the importance of meta-path $\mathcal{P}_i$.

By dynamically weighting both neighbors and meta-paths, HAN unifies spatial and functional contexts into a comprehensive node representation. Importantly, this approach can adaptively learn the relevance of different meta-paths rather than relying on manual selection. As a result, HAN not only provides interpretable embeddings grounded in meaningful relational structures but also simplifies the process of leveraging heterogeneous relational information.

3 Dataset

3.1 CODEX Dataset

The data relevant to these studies typically consists of spatially resolved single-cell datasets, which include high-dimensional molecular profiles such as gene expression levels, combined with the spatial coordinates of each cell within a tissue sample. Our project leverages a spatial proteomics dataset [10] using CODEX technique [1]. CODEX enables the simultaneous detection of multiple protein markers in tissue samples, offering detailed spatial and molecular insights into cellular composition. In this case, the CODEX dataset was generated from multiplexed imaging of 24 sections of the human intestine from three donors (B004, B005, B006), using a panel of 47 antibodies. Additionally, human tonsil and Barrett's Esophagus (BE) tissues were imaged using 57 antibodies. Following imaging, processes like image-stitching and single-cell segmentation were applied. The final dataset includes information on 870,000 cells from the intestine and 220,000 cells from the tonsil and BE, with fluorescence values from each marker. The *B004_training_dryad.csv* dataset contains approximately 250,000 cells from donor B004's small intestine and colon, with expertly annotated cell types, providing sufficient instances for training and testing our models. The *BE_Tonsil_l3_dryad.csv* dataset includes about 130,000 annotated cells from a donor's human tonsil and Barrett's Esophagus.

3.2 Batch Normalization

In the original dataset descriptions, the data underwent standard preprocessing, including normalization to address batch effect [10]. This was further supported by our review of the STELLAR repository, where we did not observe any explicit application of batch normalization techniques. Based on this, we assumed that batch correction had already been performed on the dataset.

However, to ensure robustness and verify this assumption, we applied a batch correction method, Scanorama [12], which is specifically designed for integrating and correcting batch effects in scRNA-seq datasets. Scanorama leverages

techniques such as mutual nearest neighbor matching and low-dimensional embeddings to identify and adjust for batch effects efficiently.

After running Scanorama, we observed minimal differences in cross-tissue annotation accuracy before and after batch correction. This strongly suggests that batch normalization was likely incorporated during the preprocessing of the datasets, aligning with the dataset's original descriptions. Despite this, applying Scanorama allowed us to confirm the robustness of our downstream analysis, ensuring consistency across tissue types.

Method	**Cross-Tissue Accuracy**
Before Batch Correction	81.1%
After Batch Correction	81.2%

Table 1: Comparison of cross-tissue annotation accuracy before and after batch correction. The hyperparameters are default settings in following cross-tissue experiments.

3.3 Cell Type Distribution

3.3.1 Colon and Small Intestine

We examined and visualized cell type compositions in two different tissue regions: the colon and the small intestine. This analysis aimed to highlight the heterogeneity of cell types across tissues while identifying similarities and discrepancies in their distributions.

In both tissues, we observed a high degree of similarity in the relative abundance of major cell types, such as Enterocytes, Smooth Muscle, and TA cells. This indicates a conserved cellular structure and function between these regions. However, a single rare cell type mismatch was noted, which underscores subtle differences in tissue-specific biology.

These findings provide an excellent basis for testing annotation models, particularly when addressing tissue heterogeneity. The similarity in cell type composition, combined with slight regional differences, presents a robust scenario to evaluate the model's ability to accurately annotate cell types while considering the inherent diversity of tissues.

3.3.2 Barrett's Esophagus (BE) and Tonsil

For the second pair of tissue regions—Barrett's esophagus (BE) and tonsil—our analysis revealed striking differences in cell type composition. In Barrett's esophagus, the dominant cell type is Glandular Epithelium, accounting for 32.0% of the composition, while the tonsil is predominantly composed of Innate immune cells, representing 35.6%. Furthermore, the tonsil shows a significant presence of T cells (28.0%) and B cells (18.3%), whereas these populations are less prominent in Barrett's esophagus. Conversely, Barrett's esophagus ex-

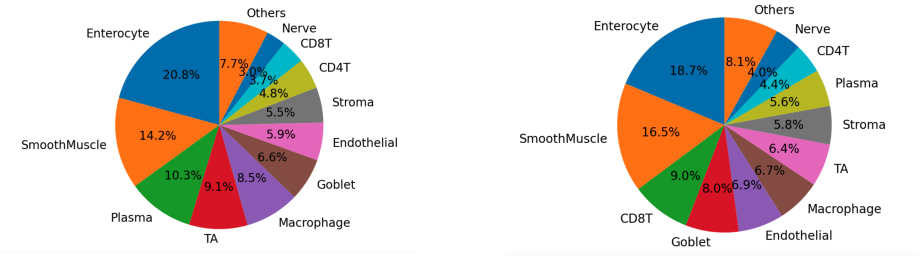

Figure 1: The cell type compositions for colon and small intestine individually.

hibits a notable abundance of Smooth Muscle cells (19.6%) and Endothelial cells (13.4%), which are far less prevalent in the tonsil.

These stark differences emphasize the challenges of cross-tissue annotation in such scenarios. Given the non-overlapping profiles of key cell types, conventional annotation models are likely to struggle in accurately predicting cell identities across these regions. To address this, we initially used STELLAR, which is specifically designed to generalize across datasets and discover novel cell types. The model's claimed ability to handle significant tissue heterogeneity makes it an ideal choice for this pair, where traditional approaches might fail to capture the full diversity of cellular compositions.

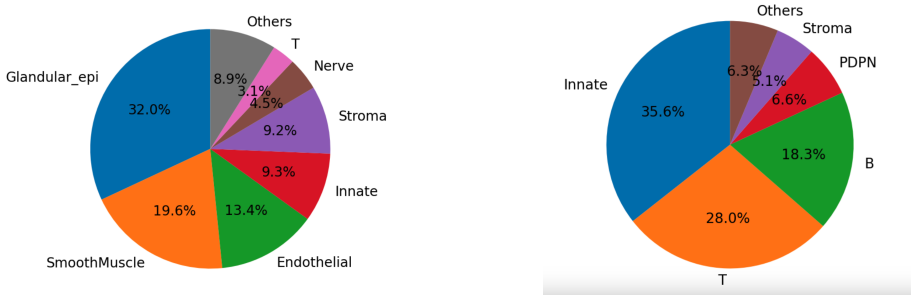

Figure 2: The cell type compositions for tonsil and BE individually.

4 Experiments

4.1 Experimental Setup

Our experiments were designed to evaluate cell type annotation performance under two scenarios: intra-region and cross-tissue annotation.

Intra-region annotation involves annotating cell types within the same tissue type. For this, we used the colon dataset, following a traditional node classification approach. The dataset was split into training (70%), validation (10%), and testing (20%) subsets to evaluate model performance.

Cross-tissue annotation involves using an annotated reference dataset from one tissue region to predict cell types in a different, unlabeled tissue region. Specifically, we trained models on the colon dataset and tested their performance in annotating cell types in the small intestine dataset.

For both intra-region and cross-tissue annotation, we first employed a fixed distance threshold d and a set number of most similar neighbors (top K) in molecular profiles ($d = 30$, $k = 5$). These parameters were used to evaluate the performance of four spatially related methods—HAN (Heterogeneous Attention Network), RGCN (Relational Graph Convolutional Network), GCN-Multi, and GCN-Spatial—alongside a multi-layer perceptron (MLP), which served as a non-spatial baseline for comparison.

RGCN (Relational Graph Convolutional Network): A graph convolutional network that explicitly models multiple types of relationships (edges) by associating each edge type with a unique weight matrix.

GCN-Multi: A simple GCN that integrates both molecular and spatial relationships into a single edge list but does not distinguish between the two types of relationships. This method treats all edges equally, irrespective of their origin (molecular or spatial).

GCN-Spatial: A GCN that considers only the spatial relationships between cells, ignoring molecular similarities. This method models adjacency based purely on spatial proximity.

Following the initial evaluation, the two most promising methods—HAN and RGCN—were further tested under varying distance thresholds and top K values. This was done to assess the robustness of these methods and uncover any underlying patterns or rules in their behavior.

4.2 Intra-region Annotation

We evaluated the performance of five different methods—HAN, RGCN, GCN-Multi, GCN-Spatial, and MLP—on the colon dataset in the intra-region annotation task. The results are summarized in terms of accuracy, as shown in the table below.

Among the methods, HAN achieved the highest accuracy (92.5%), demonstrating its effectiveness in leveraging spatial relationships and heterogeneous information. The MLP, despite being a non-spatial baseline, also performed well, with an accuracy of 91.1%, suggesting that molecular profiles alone carry substantial predictive power for intra-region annotation.

RGCN followed with an accuracy of 89.9%, showing promise as a relational model for handling spatially structured data. However, GCN-Multi (82.5%) and GCN-Spatial (74.0%) performed less effectively, possibly due to their limited ability to handle complex relationships or reliance solely on spatial relationships.

The results highlight the importance of modeling both spatial and molecular relationships effectively for intra-region annotation tasks.

To evaluate the robustness of the HAN model, we tested it using different distance thresholds (30, 40, 50) and varying top K values (1, 3, 5, 10). The accuracy results indicate that HAN consistently performs well across a wide

Methods	HAN	RGCN	GCN-Multi	GCN-Spatial	MLP
Accuracy (%)	92.5	89.9	82.5	74.0	91.1

Table 2: Performance of different methods in intra-region annotation on the colon dataset.

range of configurations, with slight variations depending on the choice of parameters. The combination of a threshold of 30 and top K=1 yielded the highest accuracy (93.2%), demonstrating the model's ability to adapt to finer spatial relationships.

Accuracy (%)	**Top** $K = 1$	**Top** $K = 3$	**Top** $K = 5$	**Top** $K = 10$
$d = 30$	**93.2**	92.3	92.5	92.6
$d = 40$	93.1	92.4	92.5	92.8
$d = 50$	92.8	92.4	92.7	92.5

Table 3: Accuracy of HAN across different distance thresholds (d) and top K values.

Despite the minimal differences in accuracy across the configurations, the observation is that smaller distance thresholds (d=30) and lower TopK values (K=1) tend to yield the highest performance. This suggests that HAN benefits from focusing on fine-grained spatial relationships and a selective set of neighbors rather than incorporating broader spatial contexts, which might introduce irrelevant information or noise.

These results also demonstrate the robustness of HAN, as it consistently achieves high accuracy (above 92%) across all tested configurations. However, the subtle variations indicate that careful tuning of distance thresholds and TopK values can further optimize performance depending on the dataset characteristics.

4.3 Cross-tissue Annotation

In the cross-tissue annotation task, we trained and fine-tuned the models on the colon dataset and applied them to annotate cell types in the small intestine dataset. This scenario represents a more challenging case due to the differences in cell type composition and spatial relationships between the two tissue regions.

The results, summarized in the table below, show that HAN and MLP achieved comparable performance, with accuracies of 81.1% and 81.5%, respectively. This indicates that molecular profiles alone (captured by MLP) are quite informative, while HAN benefits from integrating spatial and heterogeneous information to achieve similar performance.

RGCN performed slightly worse than HAN, achieving an accuracy of 79.9%, demonstrating its ability to generalize across datasets, albeit less effectively than HAN in this context. GCN-Multi and GCN-Spatial exhibited significantly

lower accuracies (72.6% and 63.2%, respectively), highlighting their limitations in handling cross-tissue differences.

Despite the minimal differences between HAN and MLP, the observation is that HAN shows strong generalizability when incorporating both spatial and heterogeneous relationships, making it better suited for cross-tissue annotation tasks than purely spatial approaches like GCN-Spatial.

Methods	**HAN**	**RGCN**	**GCN-Multi**	**GCN-Spatial**	**MLP**
Accuracy (%)	81.1	79.9	72.6	63.2	81.5

Table 4: Performance of different methods in cross-tissue annotation (trained on colon, tested on small intestine).

The table below shows the accuracy of HAN under different configurations of distance thresholds (d) and topK values in cross-tissue annotation tasks. The performance trends reveal important insights into HAN's behavior and its potential to outperform the non-spatial MLP model. With better parameter choices, such as d=50 and TopK=1, HAN can outperform the MLP model, demonstrating its superior ability to integrate spatial and molecular relationships.

As we can see, the accuracy improves slightly as the distance threshold increases, with d=50 yielding the best accuracy across most topK values; This indicates that a larger distance threshold allows HAN to capture more relevant spatial relationships across tissues, which is especially beneficial for cross-tissue tasks.

On the other hand, lower TopK values (e.g., K=1 and K=3) consistently yield better performance compared to higher K values. This suggests that focusing on fewer, most similar neighbors is more effective than considering a broader range, which may introduce less relevant or noisy connections.

We ran similar experiment on the RGCN model. RGCN's behavior, however, seems don't exhibit the same trend as HAN 6. The RGCN model's performance is oscillating between 77.9-79.1 across different TopKs and d's, and doesn't show significant differences.

This suggests that

Accuracy (%)	**Top** $K = 1$	**Top** $K = 3$	**Top** $K = 5$	**Top** $K = 10$
$d = 30$	82.3	81.6	81.1	80.0
$d = 40$	82.6	81.1	80.9	79.6
$d = 50$	**82.7**	80.8	80.4	79.6

Table 5: Accuracy of HAN across different distance thresholds (d) and top K values in cross-tissue evaluation.

The confusion matrix shows that some cell types are more frequently misclassified than others, particularly in both HAN and MLP models. Key misclassification patterns include: 1) Enterocytes misclassified as Goblets: Enterocytes and Goblet cells are both epithelial cells, which may share overlapping features

Accuracy (%)	**Top** $K = 1$	**Top** $K = 3$	**Top** $K = 5$	**Top** $K = 10$
$d = 30$	78.4	77.9	78.8	78.5
$d = 40$	78.4	78.9	78.0	**79.1**
$d = 50$	77.9	78.0	78.4	77.8

Table 6: Accuracy of RGCN across different distance thresholds (d) and top K values in cross-tissue evaluation.

in their molecular profiles, leading to confusion; 2) Goblet cells misclassified as TA cells: Transit-amplifying (TA) cells are progenitors and may share markers with Goblet cells; 3) Enterocytes misclassified as TA cells: Similar to the Goblet misclassification, this suggests overlapping molecular signatures between epithelial cells and their progenitor counterparts.

The misclassifications align with biological characteristics of the involved cell types. For example, Enterocytes and Goblet cells both belong to the epithelial lineage. Their shared structural and functional roles in maintaining tissue homeostasis could result in overlapping molecular features that complicate classification. Besides, TA cells are progenitors that give rise to more differentiated cell types like Enterocytes and Goblets. This shared lineage may explain why these cells are often misclassified, as their molecular profiles might reflect intermediate states.

4.4 Runs of STELLAR on Tonsil-BE cross annotation and Other Issues

We attempted to integrate our heterogeneous GNN approaches into the STELLAR workflow and evaluated their performance on the tonsil-BE dataset pair. However, several critical issues emerged during these experiments, particularly with prediction consistency and compatibility of objective functions.

To study and evaluate the embeddings of cells in STELLAR, we ran the default script provided in the repo, which trains and evaluates STELLAR on the TonsilBE dataset. We found that even though the original paper claims to have robust novel cell type discovery ability, it still produces false clustering annotations.

We observed that the cell-type assignment by STELLAR lacked robustness, particularly when handling scarce cell types. For example, smooth muscle cells, which are underrepresented in the training set, were often misclassified as innate immune cells.

This misclassification highlights a key limitation of STELLAR's reliance on having sufficiently similar cell pairs within a mini-batch. When the training data is imbalanced, or when certain cell types are rare, the model struggles to adequately represent and distinguish these rare types, leading to significant misclassifications.

This issue becomes more pronounced in the prediction output, which relies heavily on a predefined number of new cell types. As shown in the figures, the

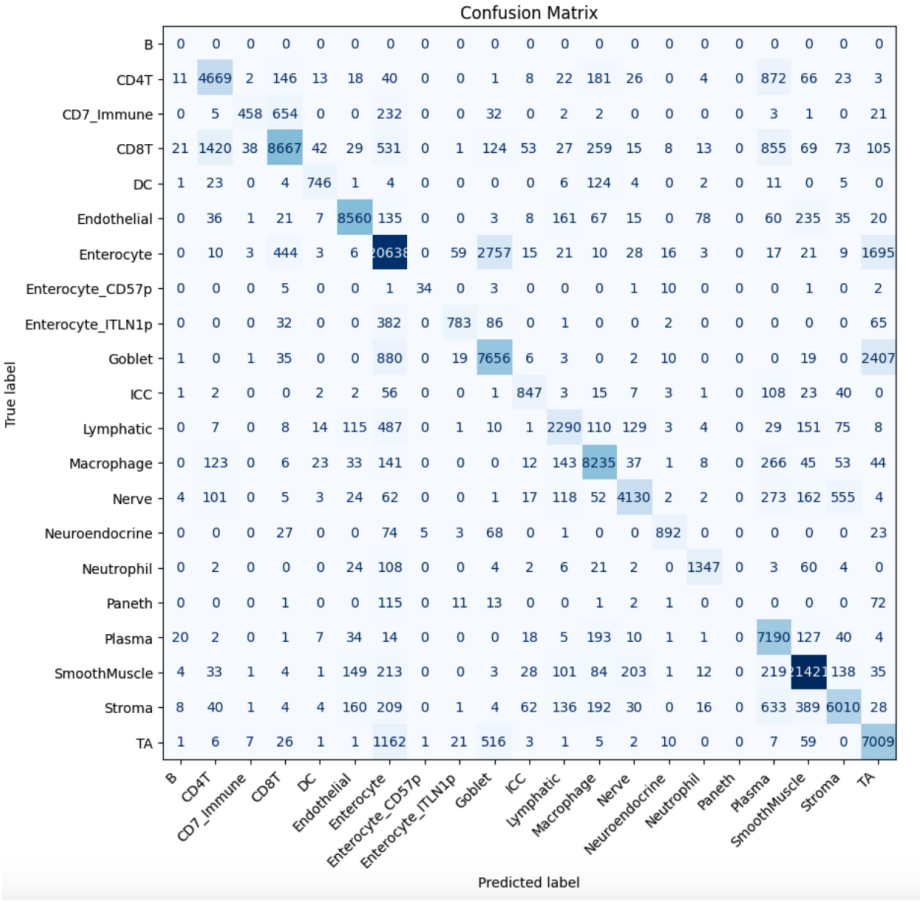

Figure 3: Confusion Matrix for Cross-tissue Evaluation Using Default HAN Model for Prediction

output mappings are inconsistent and fail to adapt to novel cell types in a biologically meaningful way. Such reliance on predefined settings undermines the model's ability to generalize effectively across datasets with distinct compositions.

Besides, applying STELLAR to our dataset settings caused errors in the calculation of the cross-entropy loss function. This issue appears to stem from a dimension mismatch introduced during the integration of novel cell types.

Therefore, temporarily, we are hampered by these obstacles and cannot acheive the integration of our approaches with STELLAR. However, we still believe jointly training in a somewhat semi-supervised way like STELLAR will closer the gap in cell label transferring in different tissues, so we summarize STELLAR's design here for future exploreation. Its objective function is designed to balance cell type similarity and novelty detection. The key components

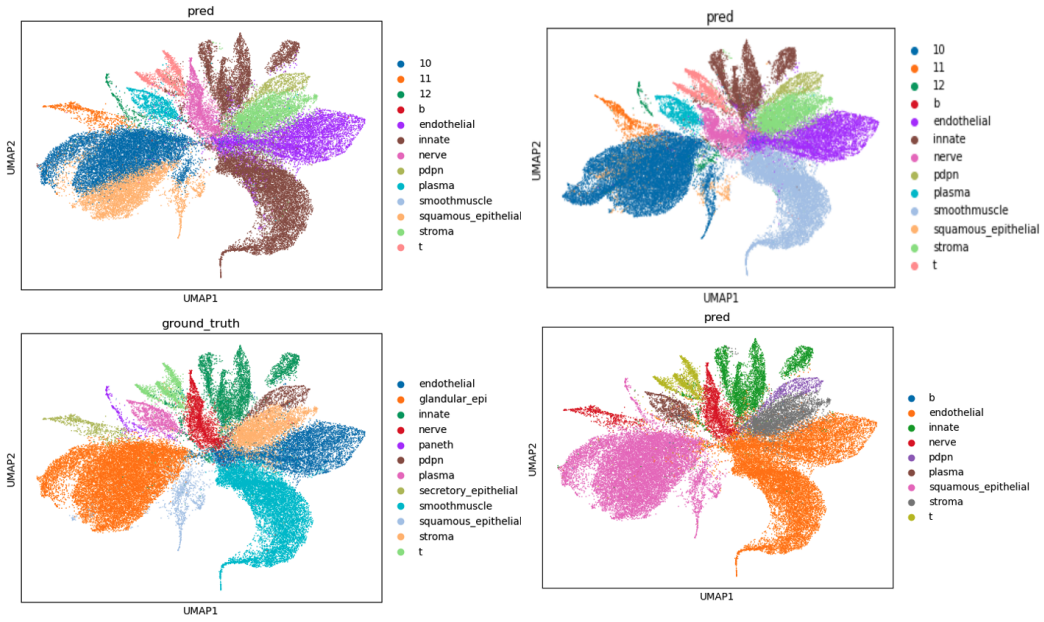

Figure 4: Top-left: novel classes set to be 3 (ours); Top-right: novel classes set to be 3 (original repo); Bottom-left: ground truths; Bottom-right: novel classes set to be 0 (ours)

include: 1) Within-batch similarity loss: The model aims to align similar cell types across mini-batches by minimizing the distance between embeddings of matching cells in the reference and query datasets. 2) Novelty detection loss: To identify new cell types, STELLAR incorporates a regularization term that encourages the network to separate novel cell types from known ones in the embedding space. 3)Cross-entropy loss: This term is used for classification tasks, aligning predicted cell types with ground truth annotations.

It is also worth mentioning that, this objective function relies on the assumption that new cell types are represented uniformly and sufficiently across mini-batches, which becomes problematic in scenarios with imbalanced or scarce data.

5 Discussion

5.1 Combined Approach Performance

Our proposed combined approach, which integrates molecular and spatial information with heterogeneous graph, outperforms pure spatial models. This aligns with the intuition that cell type identification is primarily driven by

molecular profiles. Spatial information, when used as a complementary feature, enhances performance and robustness, especially in heterogeneous tissue scenarios. This allows the combined approach to surpass MLP in cross-tissue annotation, demonstrating its adaptability to varying tissue compositions.

5.2 Parameter Sensitivity

The selection of the distance threshold (d) and the number of nearest neighbors (top K) impacts model performance, particularly in cross-tissue annotation tasks. Consistently, smaller K values yield better results in both intra-region classification and cross-tissue annotation. This suggests that focusing on fewer, most similar neighbors helps capture more relevant local relationships, avoiding noise introduced by distant or less related neighbors. This highlights the importance of parameter tuning for optimal results in diverse tissue settings.

5.3 High Accuracy of Non-spatial Methods

Non-spatial methods, such as MLP, perform well because expert-annotated cell types are largely derived from molecular-level clustering. In these annotations, molecular profiles provide the foundational information for cell type classification, while spatial relationships are often used as a secondary validation step during the ground-truth annotation process. This dependence on molecular clustering explains why MLP, despite lacking spatial features, can achieve high accuracy.

5.4 Future Work

A promising direction is to implement a joint training scheme that leverages both labeled and unlabeled datasets. By integrating unlabeled data, the model can learn broader patterns and relationships, improving generalizability to novel datasets.

Given that having some ground truth annotations is always beneficial for out-of-training-set data, a future approach might focus on constructing a larger and more comprehensive model. This model would enable training on a diverse set of annotated and unannotated tissue slides, incorporating tissue or region type as a contextual embedding. Such a model could adaptively handle tissue heterogeneity, improving cross-tissue annotation accuracy and robustness.

References

[1] Yury Goltsev et al. “Deep profiling of mouse splenic architecture with CODEX multiplexed imaging”. In: *Cell* 174.4 (2018), pp. 968–981.

[2] Kok Hao Chen et al. “Spatially resolved, highly multiplexed RNA profiling in single cells”. In: *Science* 348.6233 (2015), aaa6090.

[3] Patrik L Ståhl et al. "Visualization and analysis of gene expression in tissue sections by spatial transcriptomics". In: *Science* 353.6294 (2016), pp. 78–82.

[4] Samuel G Rodriques et al. "Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution". In: *Science* 363.6434 (2019), pp. 1463–1467.

[5] Sanja Vickovic et al. "High-definition spatial transcriptomics for in situ tissue profiling". In: *Nature methods* 16.10 (2019), pp. 987–990.

[6] Ziyang Tang et al. "spaCI: deciphering spatial cellular communications through adaptive graph model". In: *Briefings in Bioinformatics* 24.1 (2023), bbac563.

[7] Jian Hu et al. "SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network". In: *Nature methods* 18.11 (2021), pp. 1342–1351.

[8] Kangning Dong and Shihua Zhang. "Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder". In: *Nature communications* 13.1 (2022), p. 1739.

[9] Yahui Long et al. "Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST". In: *Nature Communications* 14.1 (2023), p. 1155.

[10] Maria Brbić et al. "Annotation of spatially resolved single-cell data with STELLAR". In: *Nature Methods* 19.11 (2022), pp. 1411–1418.

[11] Xiao Wang et al. "Heterogeneous graph attention network". In: *The world wide web conference*. 2019, pp. 2022–2032.

[12] Brian Hie, Bryan Bryson, and Bonnie Berger. "Efficient integration of heterogeneous single-cell transcriptomes using Scanorama". In: *Nature biotechnology* 37.6 (2019), pp. 685–691.