MIMYR: GENERATIVE MODELING OF MISSING TISSUE IN SPATIAL TRANSCRIPTOMICS

Ajinkya Deshpande
Machine Learning Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, USA

Zhilei Bei
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, MA, USA

Jian Ma*
Ray and Stephanie Lane
Computational Biology Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, USA

Spencer Krieger†
Ray and Stephanie Lane
Computational Biology Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, USA

ABSTRACT

Spatial transcriptomics enables the study of how gene expression is organized across tissues, revealing cell interactions within their native microenvironments in health and disease. However, tissue damage during sectioning and the allocation of intermediate slices to other assays often result in regions or entire planes missing from the data, limiting downstream analysis. Here, we introduce MIMYR, a generative framework for reconstructing realistic spatial transcriptomics data in unmeasured tissue regions. MIMYR addresses this challenge through three coupled components: predicting cell locations via guided diffusion, assigning cell types through supervised classification, and generating gene expression profiles with a transformer conditioned on spatial and cellular context. MIMYR accurately reconstructs held-out regions in mouse brain data and generalizes across experimental conditions, including variations in gene panels and slicing orientations. After finetuning on limited Alzheimer's disease data, MIMYR captures disease-associated transcriptional changes in unmeasured brain regions. By enabling high-fidelity spatial imputation from limited training data, MIMYR extends the utility of spatial transcriptomics, allowing researchers to recover unmeasured tissue states and deepen the study of tissue spatial organization and dynamics.

1 INTRODUCTION

Spatial transcriptomics (ST) enables direct measurement of gene expression within intact tissue architecture, making it possible to investigate how cells organize, communicate, and adapt in their native environments (Bressan et al., 2023; Moses & Pachter, 2022; Liu et al., 2024; Maynard et al., 2021). Spatial structure is tightly linked to biological function (Rao et al., 2021), and understanding these patterns requires continuous spatial context rather than isolated snapshots (Chidester et al., 2023; Krieger et al., 2025). Despite advances in spatial transcriptomics, physical sectioning remains a major source of technical artifacts (Kummerfeld et al., 2025). Slicing can introduce tears, folds, and geometric distortions, and in some cases entire regions of tissue are lost. In a MERFISH mouse brain atlas (Zhang et al., 2023), for example, many slices exhibit sizeable missing patches (Fig. S1). These gaps limit downstream analyses and compromise computational methods that rely on intact tissue sections. Additionally, in many workflows, intermediate sections are allocated to complementary assays (e.g., histology, single-cell sequencing), creating missing planes that interrupt spatial continuity (Liu et al., 2023). These gaps obscure depth-dependent transitions, blur anatomical

*Correspondence: jianma@cs.cmu.edu
†Correspondence: skrieger@andrew.cmu.edu

boundaries, and impede reconstruction of cellular neighborhoods. Consequently, spatial transcriptomics datasets often capture an incomplete view of tissue architecture. Because the detection of spatial features – rare cell types, layer-specific structures, and enriched cell–cell adjacencies – relies on contiguous sampling (Baker et al., 2023), removing sections sharply lowers the probability of recovering these patterns, especially in structured organs. Missing planes therefore do more than reduce data volume: they eliminate spatial signals that are central to biological interpretation and weaken downstream analyses.

Reconstructing missing tissue sections is therefore essential for enabling comprehensive spatial analysis. Recent methods have begun to address incomplete or under-sampled spatial transcriptomics data, but each tackles only part of the reconstruction problem. LUNA (Yu et al., 2025) learns atlas-derived spatial priors to reassemble dissociated cells into tissue structures, yet it cannot generate new cell types or transcriptomes for unmeasured regions. stDiff (Li et al., 2024) denoises or imputes low-quality expression by transferring abundance patterns from scRNA-seq, improving observed spots but not filling missing tissue. Other approaches aim to enhance or interpolate existing measurements without generating new cellular layouts: GNTD (Song et al., 2023) densifies sparse sequencing-based ST data via graph-guided tensor decomposition but operates only on existing spots; C2-STi (Que et al., 2025) interpolates intermediate histology sections using spatial transcriptomics as auxiliary input rather than reconstructing new cells or full transcriptomes; and SpatialZ (Lin et al., 2024) seeks 3D reconstruction from planar ST slices but assigns gene expression using a lookup-based procedure rather than a generative model. Among generative approaches, STDIFFUSION (Khan et al., 2025) extends diffusion models to spatial transcriptomics but relies on blending heuristics for in-between slices and cannot robustly handle missing planes or interior gaps. Collectively, existing methods do not provide a continuous generative model capable of reconstructing truly missing tissue regions under heterogeneous panels, modalities, and conditions.

Here, we introduce MIMYR, a generative framework that tackles the full spatial-transcriptomic reconstruction problem by explicitly modeling the three components required to rebuild missing tissue: cell locations, cell identities, and complete gene-expression profiles. MIMYR comprises three coordinated modules: (1) a plane-conditioned, backward-guided diffusion model that synthesizes realistic spatial layouts and reconstructs cellular positions for entire intermediate slices; (2) a supervised classifier that assigns cell types consistent with tissue architecture; and (3) a transformer-based generator that produces spatially coherent transcriptomes conditioned on the inferred cell locations and identities. Across spatial and functional metrics, MIMYR reconstructs realistic spatial transcriptomics data, generalizes to sparse data regimes and heterogeneous gene panels, and handles alternative slicing orientations such as sagittal brain sections. MIMYR can also extend transcriptomes that were measured on a limited gene panel to include all genes available from the pretraining data and generate wild-type control slices for disease tissues while preserving biologically interpretable differences. Together, these capabilities close a key methodological gap, enabling cross-condition, cross-panel, and cross-slice analyses in datasets with missing tissue, limited samples, or disease-specific perturbations.

2 METHOD

2.1 MIMYR FRAMEWORK OVERVIEW

MIMYR frames spatial transcriptomics reconstruction as three linked generative tasks: predicting cellular locations, assigning cell identities, and inferring full gene-expression profiles (**Fig.** 1).

1. A plane-conditioned diffusion model learns tissue-level cell density and uses guided inference to generate cell coordinates for missing regions.
2. A multi-layer perceptron (MLP) assigns cell identities by mapping spatial coordinates to cluster labels learned from annotated slices in a reference atlas.
3. A transformer generates full gene expression profiles conditioned on predicted spatial position, inferred cell state, and optional metadata such as species, disease state, gene panel, or technology.

The final output is a complete spatially resolved transcriptome for every generated cell in the reconstructed region. We adopt a staged formulation because, in practice, reconstruction is often a conditional completion problem rather than generation from scratch. Many experiments already

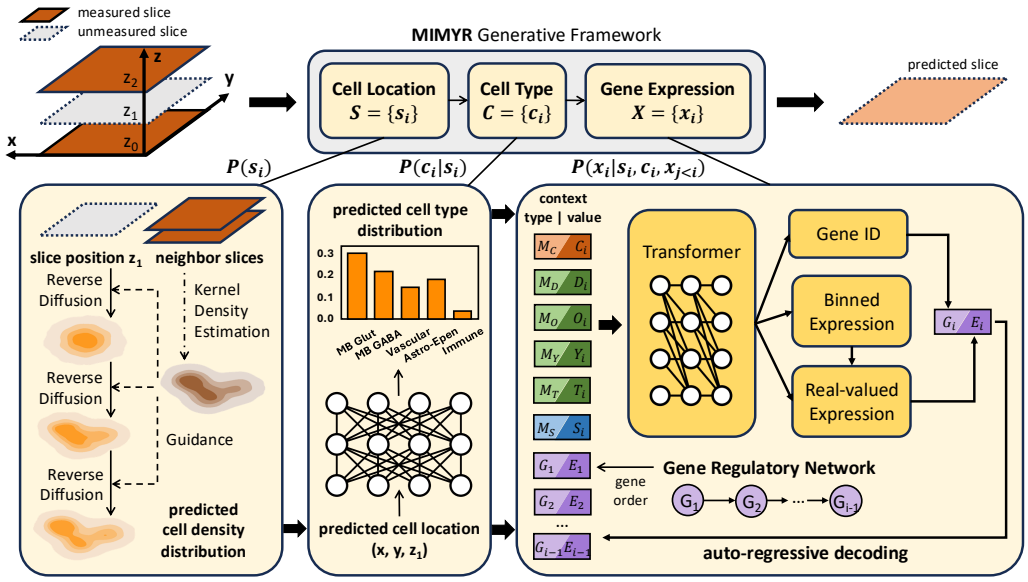

Figure 1: **Overview of the MIMYR generative framework**. The MIMYR framework reconstructs unmeasured spatial transcriptomic regions through a multi-stage generative pipeline. **(1)** Guided by cell density distributions from neighboring slices, a diffusion model predicts plausible cell locations on the target plane $S = \{s_i\}$. **(2)** Based on these predicted coordinates, an MLP estimates cell-type probabilities $P(c_i \mid s_i)$, from which cell identities are sampled. **(3)** Conditioned on the predicted cell location M_S and cell type M_C, as well as metadata variables M_D, M_O, M_Y, M_T representing species, organ, disease state, and technology, a transformer autoregressively generates gene identifiers G and corresponding expression values E according to an order derived from a gene regulatory network. Together, these modules produce spatially coherent and biologically consistent gene-expression reconstructions across the tissue.

provide cell coordinates from imaging, while others supply cell identities but measure only a limited gene panel. By separating locations, types, and expression, any available component can be held constant and the remainder generated conditionally. When all information is missing, the full pipeline is used; when partial observations exist, the model reduces to only the necessary stages. This modular structure allows the framework to interface naturally with heterogeneous assays and avoids retraining for each new setting.

MIMYR introduces several technical innovations that advance spatial-omics reconstruction. First, whereas existing approaches struggle to generalize to unseen intermediate slices and rely on simple interpolation schemes such as slice blending, MIMYR learns an explicit plane-conditioned generative model that predicts a continuous family of intermediate sections and generates an arbitrary number of cells based on the target slice thickness. Second, the diffusion component incorporates backward-guidance, leveraging information from neighboring slices when available, producing structurally consistent reconstructions without paired supervision. Third, the expression module introduces a new, biologically informed tokenization strategy that orders gene tokens using a gene regulatory network, imposing biologically meaningful structure on the generation process. Finally, the expression module conditions on metadata such as disease state, enabling synthesis of gene-expression profiles under varying biological conditions rather than producing a single, condition-agnostic distribution. Full architectural details, training hyperparameters, and implementation specifics are provided in Appendix B.

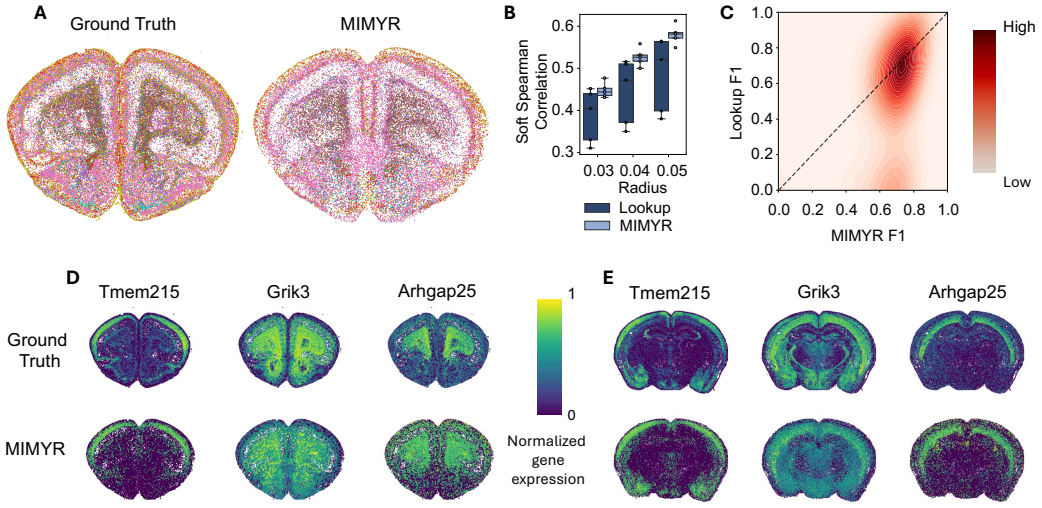

Figure 2: **Performance evaluation using a multi-sample atlas. A.** Spatial maps of test slice 1 comparing ground truth cell types (left) to MIMYR's predictions (right). **B.** Soft Spearman correlation at neighborhood radii $r \in \{0.03, 0.04, 0.05\}$ comparing MIMYR to a rule-based lookup baseline. **C.** KDE plot of F1 scores per spot across test slice 1, comparing MIMYR (x-axis) to the lookup baseline (y-axis), where points below the $y{=}x$ reference line indicate superior performance by MIMYR. **D.** Spatial expression for three example genes (*Tmem215*, *Grik3*, *Arhgap25*) on test slice 1 comparing ground truth (top) and MIMYR's prediction (bottom). **E.** Spatial expression for three example genes (*Tmem215*, *Grik3*, *Arhgap25*) on test slice 3 comparing ground truth (top) and MIMYR's prediction (bottom).

3 RESULTS

3.1 EVALUATION SETUP

We evaluated MIMYR across diverse large-scale spatial transcriptomics datasets – including a whole-brain MERSCOPE atlas (Yao et al., 2023) and a companion MERFISH atlas (Zhang et al., 2023), as well as an Alzheimer's disease MERFISH dataset (Johnston et al., 2025) – to assess reconstruction performance across samples, gene panels, slicing orientations, and disease contexts using standardized train/validation/test splits. We compared MIMYR against two baselines that follow the same three-stage decomposition of location generation, cell-type assignment, and gene-expression synthesis. The first baseline is a simple *rule-based lookup* method that assigns cell locations from the spatially nearest reference slice (or uniformly in zero-shot settings), infers cell types via local majority voting among nearby reference cells, and copies gene-expression profiles from the closest reference cell of the predicted type. The second baseline, *SpatialZ* (Lin et al., 2024), is a more advanced non-generative lookup-based framework for constructing dense 3D atlases from sparsely sampled planar slices. It refines initial cell coordinates sampled from reference slices via sliced Wasserstein distance optimization, assigns cell identities through distance-weighted k-nearest-neighbor voting, and synthesizes gene expression by similarity-weighted probabilistic sampling from similar reference cells. These baselines allow us to assess the benefit of MIMYR's atlas-scale pretraining and generative modeling beyond both simple and more advanced lookup-based reference-matching approaches. Additional details on datasets, baseline implementations, and evaluation metrics are provided in Appendix C.

3.2 RECONSTRUCTION OF MISSING TISSUE REGIONS IN AN ATLAS-SCALE MOUSE BRAIN DATASET

Missing tissue regions in spatial transcriptomics, caused by tearing, deformation, or incomplete measurements, create gaps ranging from small irregular holes within a slice to entire unmeasured planes that obscure biological structure and complicate downstream analysis. To evaluate how well MIMYR

Table 1: Soft Spearman Correlation (ρ) at Different Radii Over 5 Held-Out Slices From the MERSCOPE Mouse Brain Atlas (mean $\pm$ std).

Method	Soft ρ@0.03	Soft ρ@0.04	Soft ρ@0.05	Soft ρ@0.07	Soft ρ@0.1
Rule-based	0.39 ± 0.06	0.44 ± 0.08	0.49 ± 0.09	0.54 ± 0.11	0.60 ± 0.12
SpatialZ	0.44 ± 0.12	0.50 ± 0.14	0.54 ± 0.15	0.60 ± 0.17	0.66 ± 0.19
MIMYR	$\mathbf{0.45 \pm 0.02}$	$\mathbf{0.53 \pm 0.02}$	$\mathbf{0.58 \pm 0.02}$	$\mathbf{0.65 \pm 0.02}$	$\mathbf{0.70 \pm 0.02}$

reconstructs such missing areas in an atlas-style setting, where many tissue slices are available for training, we withheld individual slices from a MERSCOPE mouse brain atlas (Yao et al., 2023) and trained the model on the remaining sections. Each withheld slice serves as a large, biologically realistic missing region, enabling direct assessment of reconstruction fidelity under matched morphology and gene panels. Notably, because the reconstruction pipeline does not rely on observable cells within the target slice, reconstruction of irregular in-slice regions is conceptually analogous to full-slice reconstruction.

Visual inspection of predicted cell types shows that MIMYR accurately restores tissue organization and major anatomical boundaries (**Fig.** 2A). Laminar structures and regional divisions are preserved, although thin layers and sharp transitions appear slightly smoothed, and a small number of fine-scale domains are mislabeled. Predicted expression for genes such as *Tmem215*, *Grik3*, and *Arhgap25* recovers both dominant spatial gradients and local expression hotspots, indicating that the model captures large-scale architecture as well as fine-grained transcriptional structure (**Fig.** 2D–E).

We next compared MIMYR's reconstruction accuracy against the rule-based lookup baseline and SpatialZ that both rely on reference-matching rather than learning a global generative distribution. We evaluated reconstruction quality using two neighborhood-aware metrics – soft Spearman correlation ρ and soft F1 – calculated at different neighborhood radii, which together quantify how well local gene-expression patterns and spatial activity hotspots are preserved (full descriptions in Appendix C.3). The results were averaged across 5 held-out test slices. SpatialZ improves over the simple lookup across all radii, indicating that more sophisticated spatial alignment and neighbor weighting are beneficial when sufficient reference slices are available (**Tab.** 1). However, MIMYR consistently achieves higher soft Spearman scores with substantially lower variance across slices (**Tab.** 1). This reduced variance suggests greater robustness to local anatomical irregularities and slice-specific noise, whereas lookup-based methods remain sensitive to the quality and density of nearby reference sections. The performance gap widened at larger radii, reflecting MIMYR's ability to recover coherent gene-expression organization beyond immediate local matches. Although larger radii naturally tolerate positional offsets and therefore raise overall correlations, the increasing margin emphasizes that MIMYR reconstructs local expression structure more faithfully than the baselines, rather than merely reproducing global intensity patterns.

To evaluate gene-level accuracy beyond correlation-based metrics, we binarized ground-truth and predicted expression and computed a per-spot soft F1 score. Across slices, the rule-based lookup baseline reached an average soft F1 of 0.65, whereas MIMYR achieved 0.70 (absolute improvement +0.05; 9% relative gain). A two-dimensional kernel density estimate of per-spot F1 for slice 1 showed the highest-density region concentrated below the $y{=}x$ diagonal, indicating systematically higher F1 values for MIMYR (**Fig.** 2C). Similar distributions were observed for other slices, with modest shifts in peak density across samples. Together, these analyses demonstrate that MIMYR captures biologically meaningful spatial gene expression patterns with greater fidelity than the rule-based approach.

With abundant training data from the same dataset, MIMYR generalizes effectively to held-out slices, outperforming the baselines across spatial radii. The model preserves large-scale anatomical organization and reproduces fine-grained expression trends, with remaining discrepancies largely confined to boundary sharpness and minor dynamic-range compression at the smallest spatial scales. Collectively, these results show that MIMYR can reliably reconstruct missing regions in spatial transcriptomics datasets, enabling more complete and biologically accurate analysis of partially observed tissues.

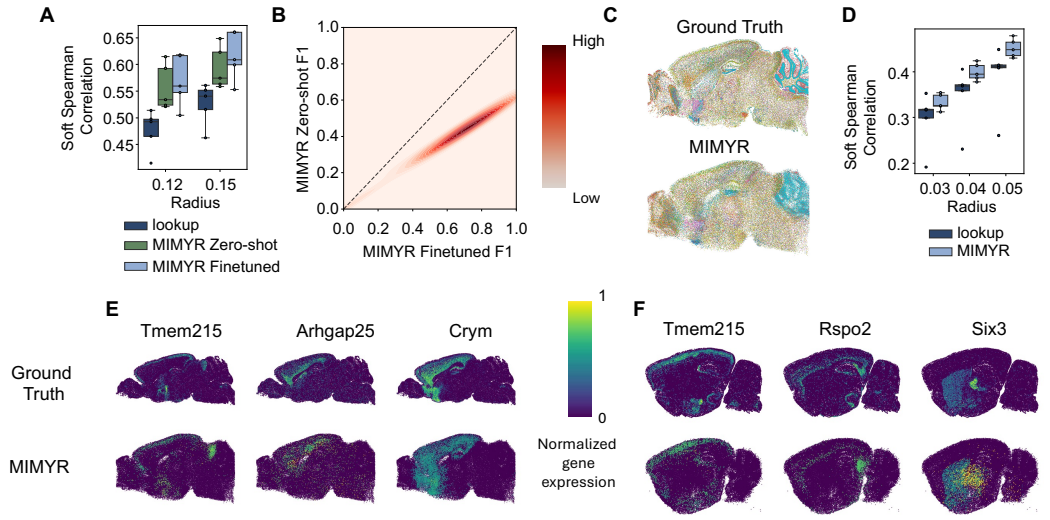

Figure 3: **Cross-setting performance and qualitative comparisons. A.** Soft Spearman correlation at radii $r \in \{0.12, 0.15\}$ for MIMYR Zero-shot, MIMYR Finetuned, and the rule-based lookup baseline in the cross-gene-panel setting. **B.** KDE plot of F1 scores per spot for test slice 1, comparing MIMYR Finetuned (x-axis) to MIMYR Zero-shot (y-axis), where points below the $y{=}x$ reference line indicate superior performance by MIMYR Finetuned. **C.** Spatial maps for a sagittal test slice comparing ground truth cell types (top) to MIMYR's predictions (bottom). **D.** Soft Spearman correlation at radii $r \in \{0.03, 0.04, 0.05\}$ for MIMYR Finetuned and the lookup baseline over sagittal test slices. **E.** Spatial expression for three example genes (*Tmem215*, *Arhgap25*, *Crym*) on sagittal test slice 1, comparing ground truth (top) and MIMYR's prediction (bottom). **F.** Spatial expression for three example genes (*Tmem215*, *Rspo2*, *Six3*) on sagittal test slice 3, comparing ground truth (top) and MIMYR's prediction (bottom).

3.3 Transfer Learning Enables Robust Reconstruction in Sparse Data Regimes

In many practical settings, only a few tissue slices are available, creating a sparse data regime that challenges model generalization. While earlier experiments assumed abundant training data from the same atlas dataset, we next examined how MIMYR performs when only limited measurements are available. To do so, we adopted a transfer-learning framework in which a model trained on one dataset is evaluated on a related target dataset. We considered two regimes. In the *zero-shot setting*, the model is applied directly to the target dataset without fine-tuning, requiring complete generalization from the source domain. In the *few-shot setting*, a small set of target slices is provided for fine-tuning, and performance is re-evaluated on held-out slices to quantify how much limited in-domain supervision improves reconstruction. The target dataset uses a different gene panel, adding variability akin to real experimental conditions (Haber et al., 2025).

For zero-shot evaluation, we applied MIMYR, trained exclusively on the coronal slices of the MERSCOPE brain atlas (Yao et al., 2023), directly to an independent MERFISH atlas (Zhang et al., 2023) with no adaptation (**Fig.** 3A). Across slices, MIMYR achieved mean soft-Spearman correlations of 0.56 and 0.59 at radii 0.12 and 0.15, outperforming the rule-based lookup baseline (0.48, 0.53) by +0.08 and +0.06, and SpatialZ (0.52, 0.56) by +0.04 and +0.03 (**Tab.** 2).

In the few-shot setting, the model was fine-tuned on a small subset of target-domain slices, and the rule-based lookup baseline used the same slices for neighborhood voting and gene-expression transfer. Fine-tuning increased performance to 0.57 and 0.62 at the same radii, corresponding to improvements over lookup of +0.09 and +0.09, and improvements over SpatialZ of +0.05 and +0.06 (**Tab.** 2). The distribution of per-spot soft F1 scores showed the highest-density region below the $y{=}x$ diagonal, indicating consistent benefits from fine-tuning (**Fig.** 3B). Even limited supervision allowed the model to adjust to measurement- and panel-specific batch effects, improving expression prediction for both in-panel and out-of-panel genes. Remaining errors were largely localized to low-coverage regions and rare cell types.

Table 2: Soft Spearman Correlation (ρ) at Different Radii Over Cross-Gene-Panel Test Slices (mean $\pm$ std).

Method	Soft ρ@0.05	Soft ρ@0.09	Soft ρ@0.12	Soft ρ@0.15
Rule-based	0.30 ± 0.04	0.41 ± 0.04	0.48 ± 0.04	0.53 ± 0.04
SpatialZ	0.33 ± 0.11	0.46 ± 0.17	0.52 ± 0.19	0.56 ± 0.21
MIMYR Zero-shot	$\mathbf{0.38 \pm 0.05}$	$\mathbf{0.51 \pm 0.05}$	0.56 ± 0.04	0.59 ± 0.04
MIMYR Finetuned	0.35 ± 0.05	0.50 ± 0.05	$\mathbf{0.57 \pm 0.05}$	$\mathbf{0.62 \pm 0.05}$

Table 3: Soft Spearman Correlation (ρ) at Different Radii Over Sagittal Test Slices (mean $\pm$ std).

Method	Soft ρ@0.03	Soft ρ@0.04	Soft ρ@0.05	Soft ρ@0.07	Soft ρ@0.1
Rule-based	0.30 ± 0.06	0.35 ± 0.07	0.39 ± 0.07	0.46 ± 0.09	0.53 ± 0.10
SpatialZ	0.18 ± 0.17	0.21 ± 0.20	0.23 ± 0.22	0.27 ± 0.26	0.32 ± 0.31
MIMYR Finetuned	$\mathbf{0.33 \pm 0.02}$	$\mathbf{0.40 \pm 0.02}$	$\mathbf{0.45 \pm 0.02}$	$\mathbf{0.53 \pm 0.02}$	$\mathbf{0.61 \pm 0.02}$

To assess orientation transfer, we next evaluated MIMYR on sagittal slices from the MERFISH atlas (Zhang et al., 2023), introducing both rotational and molecular domain shifts. After fine-tuning on a small number of sagittal slices, MIMYR continued to outperform the lookup baselines (**Fig.** 3D; **Tab.** 3), achieving mean soft Spearman gains of +0.04, +0.05, and +0.06 at radii 0.03, 0.04, and 0.05 over the rule-based baseline, with SpatialZ performing substantially worse. Qualitative agreement is visible in the cell-type maps (**Fig.** 3C), and representative gene-expression reconstructions for sagittal test slices 1 and 3 show preservation of large-scale anatomical organization and major expression gradients, with remaining discrepancies primarily confined to thin boundaries and other fine spatial features (**Fig.** 3E-F).

Overall, MIMYR demonstrates strong cross-dataset generalization and efficient adaptation under limited supervision. Zero-shot transfer already surpasses strong lookup baselines on a new dataset, and light fine-tuning closes most of the remaining gap to in-domain performance. The model's gains are largest at broader spatial radii, indicating robust recovery of global structure under strong domain shifts.

3.4 RECONSTRUCTING DISEASE-ASSOCIATED SPATIAL TRANSCRIPTOMES

We next evaluated whether MIMYR retains biological fidelity in a pathological setting using an Alzheimer's disease MERFISH dataset (Johnston et al., 2025), which includes wild-type, Trem2^{R47H}, 5xFAD, and Trem2^{R47H};5xFAD brain sections. We fine-tuned the model on three slices from each non-wild-type genotype, reserving two slices for validation and testing. The rule-based lookup baseline method uses the validation slice for lookup-based prediction. Across genotypes, MIMYR consistently outperformed the baseline, producing spatial and transcriptional reconstructions that more faithfully capture the underlying pathological organization.

Because the gene-expression module can generate spatial transcriptomes conditioned on metadata tokens, we examined two relevant scenarios for downstream analysis. In the first scenario, we held out 20 of the 300 genes during fine-tuning and prompted the model to output the full transcriptome (2,000 pretraining genes) using scRNA-seq technology tokens. MIMYR successfully generated expression for 1,811 of the 2,000 genes and accurately predicted expression for both in-panel genes and the held-out genes (**Fig.** 4A). Predicted spatial expression closely matched ground truth (**Fig.** 4B) for these held-out genes. Notably, several Alzheimer's-associated genes absent from the MERFISH panel – including *Tyrobp* (microglia) (Haure-Mirande et al., 2022) and *Cux2* (Layer 2/3 glutamatergic neurons) (Otero-Garcia et al., 2022) – displayed spatial patterns consistent with known cell-type distributions (**Fig.** 4C), demonstrating the model's capacity to extend transcriptomic coverage in biologically meaningful ways.

In the second scenario, we used metadata conditioning to generate expression for the same test slice under both Alzheimer's and wild-type conditions and analyzed the resulting differential expression patterns. Importantly, the model was not fine-tuned on any wild-type slices from this

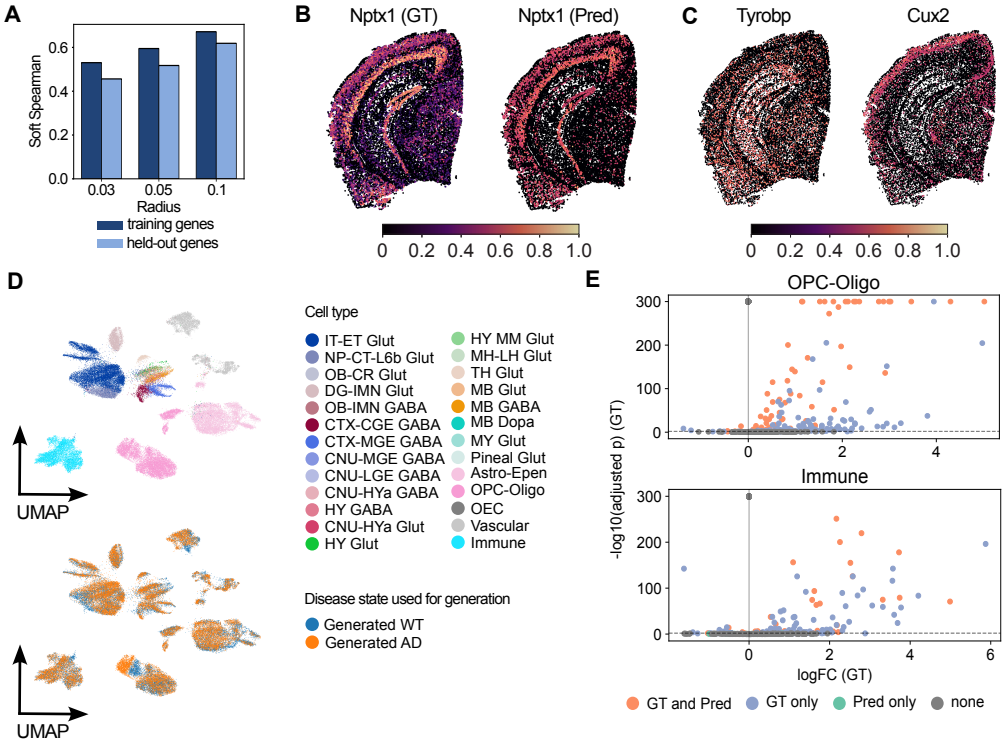

Figure 4: **Applying MIMYR to an Alzheimer's disease dataset. A.** Soft Spearman correlation at radii $r \in \{0.03, 0.05, 0.10\}$ for the finetuned MIMYR model evaluated on 280 training genes and 20 held-out genes not seen during training. Here, MIMYR predicts expression for all 2,000 genes in its vocabulary. **B.** Spatial plots comparing ground-truth and MIMYR-predicted expression of a held-out gene, *Nptx1*. **C.** Spatial plots showing MIMYR-predicted expression of Alzheimer's marker genes that were not measured in the original dataset. **D.** UMAP visualization of MIMYR-generated gene expression conditioned on wild-type or disease state, colored by cell type (top) and disease state (bottom). **E.** Volcano plots for oligodendrocytes (top) and immune cells (bottom) showing differentially expressed genes between disease states in ground-truth and generated data, colored by whether they are differential in ground truth, generated, both, or neither.

dataset. Nevertheless, we observed subtle but consistent shifts in expression profiles between cells of the same type when conditioned on different disease states (**Fig.** 4D). The largest shifts occurred in oligodendrocyte-lineage cells (OPC-Oligo), immune cells (including microglia), and astrocyte-ependymal cells (Astro-Epen). To quantify correspondence with true disease biology, we computed differentially expressed genes (DEGs) between wild-type and Alzheimer's states in the generated data and compared them to DEGs from the real dataset (wild-type versus Trem2^{R47H};5xFAD). For these three major cell types, all DEGs identified in the generated data were contained within the ground-truth DEGs (**Fig.** 4E), although a subset of experimentally observed DEGs were not recovered by the model.

Overall, these results demonstrate that MIMYR can robustly extend the transcriptome to nearly all genes available from pretraining and can generate synthetic wild-type controls for disease datasets while still recapitulating biologically meaningful gene-expression shifts across conditions. These capabilities support downstream analyses even with limited samples, sparse panels, or missing spatial regions.

4 DISCUSSION

In this work we introduced MIMYR, a unified framework for regenerating missing or damaged regions in spatial transcriptomics tissues. The method decomposes tissue reconstruction into three sequential stages: (1) generating plausible cell locations given a specified tissue region, optionally

guided by neighboring slices; (2) predicting cell identity conditioned solely on spatial position; and (3) producing realistic gene-expression profiles for each generated cell using the predicted locations and cell types.

Across multiple analyses, we showed that MIMYR yields coherent and biologically plausible completions. The reconstructed tissues exhibit spatially consistent cell-type distributions and gene-expression patterns that align well with ground truth. We also demonstrated robustness in data-sparse settings and showed that with minimal fine-tuning, MIMYR transfers effectively across samples, gene panels, and slice orientations (e.g., sagittal). These results indicate that the model does not rely on dataset-specific heuristics, but rather captures transferable spatial and transcriptional structure that generalizes across gene panels and technologies.

Beyond reconstruction, MIMYR provides practical utility for biological inference. We showed that the model can augment limited gene panels by predicting spatial patterns of unmeasured genes in diseased tissues. By manipulating disease-state metadata during generation, the framework can also simulate corresponding healthy or control tissues, enabling counterfactual analyses that are otherwise inaccessible. This ability to generate inferred "what-if" states opens opportunities for exploring disease-associated perturbations, estimating unmeasured gene programs, and supporting downstream workflows such as DEG analysis or spatial neighborhood characterization.

MIMYR also suggests several promising directions for future work. Its generative nature naturally extends to interpolating in-between slices, enabling denser 3D reconstructions and providing powerful data augmentation for spatial analysis methods. While the architectural framework is fundamentally general and separate models could be trained for different tissues and species, its practical effectiveness likely depends on the availability of large-scale, atlas-level spatial transcriptomics data to provide a robust training prior. At present, these comprehensive, high-resolution references are primarily concentrated in the mouse brain; consequently, while our approach is broadly applicable, its current deployment is tied to the maturing landscape of spatial atlases in other domains. Furthermore, a significant practical constraint is the reliance on aligned slices achieved through the Allen CCF. Existing CCF-based registration requires manual steps and can be time-consuming; developing alignment-free methods or more automated registration pipelines would further broaden applicability. Another opportunity for improvement lies in integrating additional spatial modalities such as histology, protein imaging, or chromatin accessibility, which could further strengthen conditioning signals for reconstruction.

Overall, MIMYR fills a gap in an under-explored problem space and establishes a flexible foundation for future generative models in spatial biology. It enables new biological analyses, supports inference in low-data regimes, and expands the computational toolkit available for spatial transcriptomics research.

REFERENCES

Ethan AG Baker, Denis Schapiro, Bianca Dumitrascu, Sanja Vickovic, and Aviv Regev. In silico tissue generation and power analysis for spatial omics. *Nature Methods*, 20(3):424–431, 2023.

Arpit Bansal, Hong-Min Chu, Avi Schwarzschild, Soumyadip Sengupta, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Universal guidance for diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 843–852, 2023.

Haiyang Bian, Yixin Chen, Xiaomin Dong, Chen Li, Minsheng Hao, Sijie Chen, Jinyi Hu, Maosong Sun, Lei Wei, and Xuegong Zhang. scMulan: a multitask generative pre-trained language model for single-cell analysis. In *International Conference on Research in Computational Molecular Biology*, pp. 479–482. Springer, 2024.

Dario Bressan, Giorgia Battistoni, and Gregory J Hannon. The dawn of spatial omics. *Science*, 381 (6657), 2023.

Benjamin Chidester, Tianming Zhou, Shahul Alam, and Jian Ma. SPICEMIX enables integrative single-cell spatial modeling of cell identity. *Nature Genetics*, 55(1):78–88, 2023.

Darren A Cusanovich, Andrew J Hill, Delasa Aghamirzaie, Riza M Daza, Hannah A Pliner, Joel B Berletch, Galina N Filippova, Xingfan Huang, Lena Christiansen, William S DeWitt, et al. A single-cell atlas of in vivo mammalian chromatin accessibility. *Cell*, 174(5):1309–1324, 2018.

Ellie Haber, Ajinkya Deshpande, Jian Ma, and Spencer Krieger. Unified integration of spatial transcriptomics across platforms with LLOKI. *Genome Research*, 2025.

Jean-Vianney Haure-Mirande, Mickael Audrain, Michelle E Ehrlich, and Sam Gandy. Microglial TYROBP/DAP12 in Alzheimer's disease: Transduction of physiological and pathological signals across trem2. *Molecular Neurodegeneration*, 17(1):55, 2022.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.

Kevin G Johnston, Bereket T Berackey, Kristine M Tran, Alon Gelber, Zhaoxia Yu, Grant R MacGregor, Eran A Mukamel, Zhiqun Tan, Kim N Green, and Xiangmin Xu. Single-cell spatial transcriptomics reveals distinct patterns of dysregulation in non-neuronal and neuronal cells induced by the Trem2 R47H Alzheimer's risk gene mutation. *Molecular Psychiatry*, 30(2):461–477, 2025.

Sumeer Ahmad Khan, Xabier Martínez de Morentin, Vincenzo Lagani, Robert Lehmann, Narsis A Kiani, David Gomez-Cabrero, and Jesper Tegnér. stDiffusion: A diffusion based model for generative spatial transcriptomics. In *ICLR 2025 Workshop on Machine Learning for Genomics Explorations*, 2025.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv*, 2014.

Spencer Krieger, Ellie Haber, and Jian Ma. EYKTHYR reveals transcriptional regulators of spatial gene programs. *bioRxiv*, pp. 2025–05, 2025.

Erich Kummerfeld, Leland Williams, Yinzhao Wang, Samuel T Peters, Elizabeth Schmidt, Mickayla DuFresne-To, David Bernlohr, Paul Robbins, Sayeed Ikramuddin, Oyedele Adeyi, et al. Artifacts in spatial transcriptomics data: their detection, importance, prevalence, and prevention. *Briefings in Bioinformatics*, 26(4):bbaf306, 2025.

Kongming Li, Jiahao Li, Yuhao Tao, and Fei Wang. stDiff: a diffusion model for imputing spatial transcriptomics through single-cell transcriptomics. *Briefings in Bioinformatics*, 25(3), 2024.

Senlin Lin, Zhikang Wang, Yan Cui, Qi Zou, Chuangyi Han, Rui Yan, Zhidong Yang, Wei Zhang, Rui Gao, Jiangning Song, et al. Bridging the dimensional gap from planar spatial transcriptomics to 3D cell atlases. *bioRxiv*, pp. 2024–12, 2024.

Longqi Liu, Ao Chen, Yuxiang Li, Jan Mulder, Holger Heyn, and Xun Xu. Spatiotemporal omics for biology and medicine. *Cell*, 187(17):4488–4519, 2024.

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

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

Lambda Moses and Lior Pachter. Museum of spatial transcriptomics. *Nature Methods*, 19(5):534–546, 2022.

Marcos Otero-Garcia, Sameehan U Mahajani, Debia Wakhloo, Weijing Tang, Yue-Qiang Xue, Samuel Morabito, Jie Pan, Jane Oberhauser, Angela E Madira, Tamara Shakouri, et al. Molecular signatures underlying neurofibrillary tangle susceptibility in Alzheimer's disease. *Neuron*, 110 (18):2929–2948, 2022.

Ningfeng Que, Xiaofei Wang, Jingjing Chen, Yixuan Jiang, and Chao Li. Adaptive spatial transcriptomics interpolation via cross-modal cross-slice modeling. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 45–54. Springer, 2025.

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

David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 1986.

Tianci Song, Charles Broadbent, and Rui Kuang. GNTD: reconstructing spatial transcriptomes with graph-guided neural tensor decomposition informed by spatial and functional relations. *Nature Communications*, 14(1):8276, 2023.

Quanxin Wang, Song-Lin Ding, Yang Li, Josh Royall, David Feng, Phil Lesnar, Nile Graddis, Maitham Naeemi, Benjamin Facer, Anh Ho, et al. The Allen mouse brain common coordinate framework: a 3D reference atlas. *Cell*, 181(4):936–953, 2020.

Zizhen Yao, Cindy TJ Van Velthoven, Michael Kunst, Meng Zhang, Delissa McMillen, Changkyu Lee, Won Jung, Jeff Goldy, Aliya Abdelhak, Matthew Aitken, et al. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature*, 624(7991):317–332, 2023.

Tingyang Yu, Chanakya Ekbote, Nikita Morozov, Jiashuo Fan, Pascal Frossard, Stéphane d'Ascoli, and Maria Brbić. Tissue reassembly with generative AI. *bioRxiv*, pp. 2025–02, 2025.

Meng Zhang, Xingjie Pan, Won Jung, Aaron R Halpern, Stephen W Eichhorn, Zhiyun Lei, Limor Cohen, Kimberly A Smith, Bosiljka Tasic, Zizhen Yao, et al. Molecularly defined and spatially resolved cell atlas of the whole mouse brain. *Nature*, 624(7991):343–354, 2023.

A Motivating Examples

Spatial transcriptomics data frequently contain large missing regions caused by tissue tearing, folding, or sectioning artifacts (Fig. S1), motivating the need for methods that can reliably reconstruct unmeasured tissue in a biologically consistent manner.

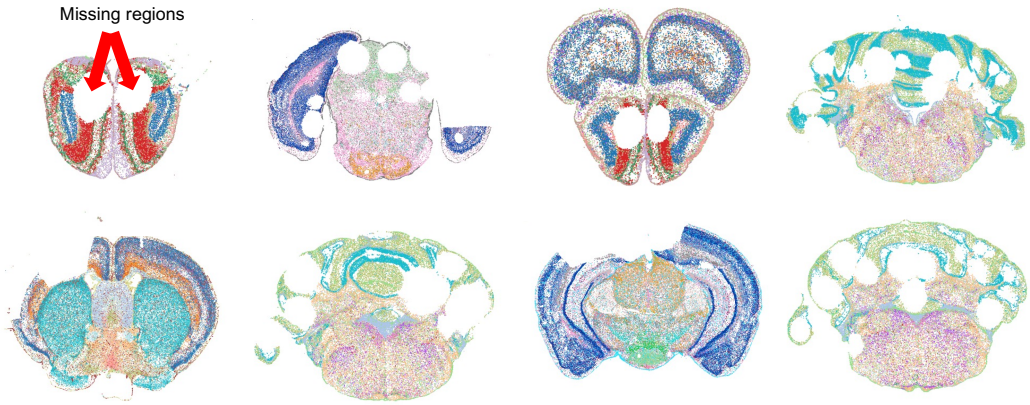

Figure S1: Tissue tearing, folding, and other artifacts frequently create missing regions in spatial transcriptomics datasets. In a dataset used in this study (Yao et al., 2023), several slices exhibit substantial holes. Some examples are shown to illustrate the scale of these missing sections.

B Mimyr Architecture

B.1 Pipeline Design

Reconstructing a missing tissue region requires reasoning about where cells should be located, what identities they should take on, and how their gene-expression programs should manifest within their spatial context. To make this problem tractable, we decompose it into three sequential generative stages that reflect this hierarchy (**Fig.** 1). First, we infer a plausible spatial layout of cells in the unobserved slice using a diffusion model that learns the global anatomical density and adapts it to arbitrary slicing planes. Second, given these reconstructed coordinates, we assign cell-type identities using a lightweight classifier that captures how cell types are spatially organized across the tissue and transfers this structure to unlabeled regions. Finally, conditioned on both spatial position and predicted identity, as well as additional sample-level attributes, we generate full gene-expression profiles using a transformer-based model that captures local transcriptional neighborhoods and broader regulatory structure. Together, these components form an integrated framework that reconstructs realistic and spatially coherent cellular and molecular landscapes in missing tissue sections.

B.2 Predicting Cell Locations with Guided Diffusion

To infer plausible cell coordinates within unobserved regions of a tissue, we use a diffusion-based generative model that learns the underlying spatial density of cells while conditioning on an explicit representation of the slicing plane. Trained on reference slices from samples that collectively span the full anatomical extent of the organ, the model captures a continuous probability distribution across the three-dimensional anatomical space (x, y, z) and learns how spatial density varies across arbitrary section orientations.

Formally, we model the conditional density $p(\mathbf{s} \mid \mathbf{c})$ over spatial coordinates $\mathbf{s} \in \mathbb{R}^3$ given a plane descriptor $\mathbf{c} \in \mathbb{R}^6$, where $\mathbf{c} = (p_x, p_y, p_z, n_x, n_y, n_z)$ encodes a point on the plane and the unit normal vector of that plane. The model is implemented as a denoising diffusion probabilistic model (DDPM) (Ho et al., 2020). During training, Gaussian noise is progressively added to the true cell coordinates according to the forward process:

$$q(\mathbf{s}_t \mid \mathbf{s}_{t-1}) = \mathcal{N}(\sqrt{1-\beta_t}\,\mathbf{s}_{t-1},\, \beta_t \mathbf{I}), \tag{1}$$

and a neural network θ learns the conditional reverse process by predicting the injected noise $\epsilon_\theta(\mathbf{s}_t, \mathbf{c}, t)$ at each timestep t, given both the noisy coordinates $\mathbf{s}_t$ and the plane condition $\mathbf{c}$. At inference time, generating the spatial layout of a slice reduces to sampling from the reverse diffusion process while supplying the desired plane descriptor $\mathbf{c}$. This conditioning mechanism allows the model to generalize to arbitrary intermediate slice positions and orientations, and it naturally supports variable slice thickness by scaling the number of sampled points to match the expected density for the specified plane.

When neighboring slices are available, we optionally refine the generated coordinates using information from the closest observed section. Instead of relying solely on the plane-conditioned density, we incorporate backward universal guidance (Bansal et al., 2023) to bias samples toward structures that are known to appear adjacent to the target slice.

$$\widetilde{\epsilon}(\mathbf{s}_t, \mathbf{c}, t) = \epsilon_\theta(\mathbf{s}_t, \mathbf{c}, t) + \eta\sqrt{\alpha_t/(1-\alpha_t)}\,\nabla_{\widehat{\mathbf{s}}_0^t} \log p_{\mathrm{KDE}}(\widehat{\mathbf{s}}_0^t) \tag{2}$$

where α_t denotes the noise-schedule coefficient at timestep t, η is a guidance-strength hyperparameter, and $\widehat{\mathbf{s}}_0^t$ denotes the model's estimate of the original clean coordinates at timestep t. Specifically, we compute a KDE on the nearest real slice and use its gradient as a weak anatomical prior during sampling. At each reverse diffusion step, the samples are nudged toward regions of higher KDE density. We choose backward guidance here due to its computational simplicity and stronger adherence to the energy function. This preserves the model's global distribution while using neighboring sections to enforce local anatomical consistency when available.

B.3 Predicting Cell Identity with a Multi-layer Perceptron

We infer cluster identities using a lightweight multi-layer perceptron (MLP) (Rumelhart et al., 1986) trained to map each cell's spatial coordinates to its cell-type label, using annotated slices for supervision and applying the model to predict labels in unannotated regions.

Formally, let the labeled dataset be

$$\mathcal{D}_L = \{(\mathbf{s}_i, c_i)\}_{i=1}^{N_L}, \tag{3}$$

where $\mathbf{s}_i \in \mathbb{R}^3$ denotes spatial coordinates and $c_i \in \{1, \ldots, C\}$ the corresponding cell-type label. We learn a function to approximate the posterior for each cell i

$$f_\theta : \mathbb{R}^3 \to \Delta^{C-1}, \qquad f_\theta(\mathbf{s}_i) = \mathbf{p}_i = (p_{i1}, \ldots, p_{iC}), \tag{4}$$

where Δ^{C-1} denotes the C-class probability simplex, $p_{ic} \geq 0$ and $\sum_{c=1}^{C} p_{ic} = 1$.

The MLP consists of multiple fully connected layers with ReLU activations and a softmax output layer. We optimize the categorical cross-entropy loss using the Adam optimizer (Kingma & Ba, 2014):

$$\mathcal{L} = -\sum_{i=1}^{N_L} \log \left[f_\theta(\mathbf{s}_i)\right]_{c_i}. \tag{5}$$

When sufficient labeled data are available, the MLP is trained directly on the labeled dataset $\mathcal{D}_L$. Otherwise, we first pretrain the model on a fully annotated reference sample to learn global correspondences between spatial coordinates and cell-type structure. We then finetune the weights on the available labeled slices from the target sample, enabling the model to adapt to local spatial variations while retaining global structural priors. This transfer learning setup is applicable to the gene-expression prediction module as well.

After training, the model estimates cell-type probabilities $\mathbf{p}_j = f_\theta(\mathbf{s}_j)$ for each cell j in the unlabeled slices $\mathcal{D}_U$. Instead of taking the most likely class, we sample cell-type assignments from this predictive distribution $\widehat{c}_j \sim \mathbf{p}_j$ which yields more diverse cell-type assignments.

B.4 Using Location and Cell Identity to Predict Gene Expression

After inferring each cell's spatial position and identity, we generate its full gene-expression profile. Let x_i denote the expression profile of cell i with type c_i and spatial location $\mathbf{s}_i$. The baseline problem is to learn a conditional generator of the form:

$$x_i \sim p_\theta(x_i \mid c_i, \mathbf{s}_i), \tag{6}$$

where θ represents the learnable parameters of the model. We extend this formulation to incorporate additional sample-level attributes such as species, disease state, gene panel, technology, and organ as conditioning variables.

To learn θ, we employ a transformer-based architecture that builds on the generative framework introduced in scMulan (Bian et al., 2024). Specifically, we use the same dual-embedding token representation, where each gene or metadata attribute is encoded as a 2-tuple: an identity embedding and a value (expression) embedding, which are summed to form the input. We also maintain the autoregressive next-token prediction strategy, however we predict only expressed genes to manage data sparsity.

Building upon this foundation, we introduce four novel extensions designed for spatial robustness and biological causality: tokenization of spatial metadata, three-head decoder, stagnation-aware pruning, and a gene regulatory network-based sequence ordering.

We represent spatial locations as discrete tokens by binning x, y, and z coordinates into 100 bins each. During training, we inject random integer noise – uniformly sampled from $\{-2, -1, 0, 1, 2\}$ – into these coordinates to enhance spatial robustness. Gene expression levels are similarly discretized into 100 bins to support the binned expression prediction head.

The transformer outputs feed into three decoder heads: (1) a gene identity head for reconstructing gene tokens, (2) a binned expression head for predicting discretized expression levels, and (3) a real-valued expression head for regressing continuous expression values. The gene identity and binned expression heads receive the direct output of the transformer, while the real-valued expression head operates on the output of the binned expression prediction.

Our training objective minimizes a weighted combination of losses: cross-entropy for gene identity and binned expression, and mean squared error (MSE) for real-valued expression. A scaling factor α controls the relative weighting between gene identity and expression terms, while the binned and real-valued expression losses are balanced equally. We found that $\alpha = 1$ achieves the optimal trade-off between classification accuracy (identity) and regression precision (expression).

At inference, the model autoregressively synthesizes complete expression profiles by recursively feeding predicted identities and binned values back into the sequence. Generation continues until an end-of-sequence token is reached or the model detects stagnation via a predefined number of "non-descending tokens," at which point we apply end-of-sequence pruning. While we tested both 0.25M and 23M-parameter variants, the medium 23M-parameter model is used for all evaluations.

To ground this autoregressive process in biological causality, we address the inherent lack of gene ordering. An arbitrary sequence could introduce non-biological bias; therefore, we impose a grounded order by constructing a directed gene regulatory network (GRN). Using a mouse brain ATAC-seq atlas (Cusanovich et al., 2018) processed via Signac, we restrict the GRN to our gene vocabulary and apply a condensation and tie-breaking procedure to linearize the graph. This produces a total order where, for every directed edge (g_1, g_2), the source precedes its target ($g_1 < g_2$). For strongly connected components (cyclic motifs), we break edges minimally and prioritize nodes with higher out-degree, resulting in a reproducible, topology-consistent sequence for modeling.

B.5 Hyperparameters

For the location module, we trained a conditional diffusion model with a learning rate of 2×10^{-4}, a batch size of 2.05×10^5, and 1,000 training epochs. After this initial phase, we selected "good" slices – sections without holes or major artifacts – and continued training for an additional 500 epochs. This two-stage procedure enables the model to learn the global distribution of cell locations from all slices while refining local spatial structure using high-quality slices. The diffusion process used 70 timesteps and a cosine noise schedule.

For the cell-type module, we trained the MLP classifier using the Adam optimizer with a learning rate of 10^{-3}, a batch size of 32, and standard cross-entropy loss. Model selection was performed using validation performance.

For the gene-expression module, we used a context length of 500, a learning rate of 5×10^{-4}, $\alpha = 1$, and an effective batch size of 512. Our gene identity vocabulary was constructed from the union of all genes present across our training and evaluation datasets, supplemented with highly variable

genes from the single-cell reference to reach a total of 2,000 genes. Pretraining was performed on slices from the MERFISH atlas (Zhang et al., 2023) and on accompanying scRNA-seq data, where each scRNA-seq cell was assigned the spatial coordinates of a randomly selected MERFISH cell of the same cell type. To improve representation of underrepresented cell types and those with strong spatial variation, we balanced the dataset by upsampling rarer populations and spatially variable cell types. The initial model was pretrained for 4 epochs.

B.6 Scalability

All models were trained using a mixture of NVIDIA A6000, A100, and H100 GPUs. The diffusion model and MLP were trained on single GPUs each, with training time for the diffusion model approximately 1.5 days, and around 3 hours for the MLP. The gene expression module was trained using distributed training on 4 GPUs and training the medium model with 23M parameters took 2 hours.

C Evaluation Setup

C.1 Datasets

We evaluated Mimyr across multiple large-scale spatial transcriptomic datasets spanning healthy and diseased mouse brains. Our primary dataset is a whole-brain MERSCOPE atlas comprising approximately ten million cells and profiling more than 500 genes (Yao et al., 2023). These data are integrated with scRNA-seq to define over 5,000 molecularly distinct clusters, all registered to the Allen Mouse Brain Common Coordinate Framework (CCF) (Wang et al., 2020). Four randomly selected slices were held out for validation and five for testing, with all remaining sections used for training. This dataset provides a foundation for assessing reconstruction performance in a regime with abundant training data.

To assess generalization across samples, gene panels, and spatial orientations, we additionally used a companion MERFISH dataset (Zhang et al., 2023) that offers comparable whole-brain coverage aligned to the same CCF but measured over a distinct panel of more than 1,100 genes. For the panel-shift and cross sample experiments – where the target dataset contains a non-overlapping gene panel – we fine-tuned the pretrained model on four slices, used one slice for validation, and evaluated on five held-out slices. We also report a zero-shot version of this experiment without finetuning. To evaluate robustness to orientation, we introduced a sagittal setting: models were fine-tuned on sagittal slices and evaluated on independent sagittal test slices using a 4/1/5 split for fine-tuning, validation, and testing.

To examine Mimyr's ability to generalize to a disease context, we used an Alzheimer's disease MERFISH dataset (Johnston et al., 2025), which includes 19 brain sections spanning four genotypes: wild-type, Trem2^{R47H}, 5xFAD, and Trem2^{R47H};5xFAD. Because CCF registration was not available for this dataset, we applied a preprocessing step to align all slices to the Allen CCF, adapting the registration workflow from (Yao et al., 2023) for consistent coordinate mapping. We also mapped the cell type labels to the single-cell reference dataset via label transfer in a shared latent space. Specifically, following the integration protocol of our primary dataset (Yao et al., 2023), we co-embedded the cells into a 100-dimensional latent space based on the normalized and log1p-transformed counts of the shared genes, and hierarchically assigned cell type labels by nearest neighbor voting with a confidence threshold.

C.2 Baseline Methods

Rule-based Lookup. We compared Mimyr against a rule-based baseline designed to mirror its three-stage prediction process – location, cell type, and gene expression – while relying solely on spatial proximity and local voting. All baseline predictions are drawn from a reference set composed of the available training or fine-tuning slices. For location prediction, the baseline assigns cell coordinates from the spatially nearest slice in the reference set; when no fine-tuning data are available (zero-shot transfer), locations are instead sampled uniformly along a circle centered on the target region. Given these locations, cell types are assigned by majority voting among the 20 nearest neighbors in the reference slice. Gene expression profiles are then copied from the closest refer-

ence cell that shares the same predicted cell type. This sequential, rule-based procedure provides a consistent and interpretable baseline across all evaluation settings.

SpatialZ. SpatialZ (Lin et al., 2024) synthesizes "virtual" slices by identifying spatially similar cellular neighborhoods in nearby reference slices and transferring gene-expression profiles via a lookup-based procedure rather than learning a generative distribution. Given two adjacent experimentally measured slices, SpatialZ first initializes virtual cell coordinates by sampling spatial positions from the neighboring real sections. These sampled coordinates are then refined by minimizing a sliced Wasserstein distance objective between the virtual slice and its adjacent real slices via stochastic gradient descent. Cell types are then inferred using a distance-weighted voting strategy within each cell's k nearest neighbors in Euclidean space. For each virtual cell, nearest neighbors are retrieved from both adjacent real slices, and their labels are aggregated using inverse-distance weighting. Finally, to generate expression profiles, SpatialZ constructs a spatial niche representation for each cell using multi-scale neighborhood statistics that encode the distribution of surrounding cell types. Gene expression values are sampled probabilistically from the expression profiles of reference cells within similar niches.

C.3 Evaluation Metrics

We evaluate reconstruction quality using two neighborhood-aware metrics: *soft Spearman correlation* and *soft F1*, which jointly measure local spatial and transcriptional coherence between the generated and ground-truth tissues.

Soft Spearman Correlation For each ground-truth cell located at position $\mathbf{x}_i$, we define a local neighborhood (or "spot") as all cells within a fixed spatial radius r:

$$\mathcal{N}_{\text{gt}}(i) = \{\, j \mid \|\mathbf{x}_j^{\text{gt}} - \mathbf{x}_i\|_2 \le r \,\}, \qquad \mathcal{N}_{\text{pred}}(i) = \{\, j \mid \|\mathbf{x}_j^{\text{pred}} - \mathbf{x}_i\|_2 \le r \,\}. \tag{7}$$

Within each spot, we aggregate the gene expression vectors over all neighboring cells:

$$\mathbf{g}_i = \sum_{j \in \mathcal{N}_{\text{gt}}(i)} \mathbf{X}_j^{\text{gt}}, \qquad \mathbf{p}_i = \sum_{j \in \mathcal{N}_{\text{pred}}(i)} \mathbf{X}_j^{\text{pred}}, \tag{8}$$

where $\mathbf{X}_j^{\text{gt}}$ and $\mathbf{X}_j^{\text{pred}}$ denote the gene expression vectors of cell j in the ground-truth and predicted tissues, respectively. We then compute the Spearman correlation $\rho_i = \text{corr}_S(\mathbf{g}_i, \mathbf{p}_i)$ between the aggregated expression profiles of the two corresponding spots. The final score is the mean correlation across all ground-truth cells:

$$\text{SoftSpearman} = \frac{1}{N} \sum_{i=1}^{N} \rho_i. \tag{9}$$

This measures how well local spatial expression gradients are preserved in the generated tissue.

Soft F1 Score To quantify local agreement in gene activation, we compute a radius-based *soft F1* score using the same neighborhood definition. For each cell i, after summing expressions within the corresponding spots to obtain $\mathbf{g}_i$ and $\mathbf{p}_i$, we binarize each gene as active if its aggregated expression is nonzero:

$$\mathbf{g}_i^{\text{bin}} = \mathbb{1}[\mathbf{g}_i > 0], \qquad \mathbf{p}_i^{\text{bin}} = \mathbb{1}[\mathbf{p}_i > 0]. \tag{10}$$

We then compute local precision and recall as

$$\text{Precision}_i = \frac{|\mathbf{p}_i^{\text{bin}} \cap \mathbf{g}_i^{\text{bin}}|}{|\mathbf{p}_i^{\text{bin}}|}, \qquad \text{Recall}_i = \frac{|\mathbf{p}_i^{\text{bin}} \cap \mathbf{g}_i^{\text{bin}}|}{|\mathbf{g}_i^{\text{bin}}|}, \tag{11}$$

and their harmonic mean

$$\text{F1}_i = \frac{2\,\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}. \tag{12}$$

The global score is averaged across all ground-truth spots:

$$\text{SoftF1} = \frac{1}{N} \sum_{i=1}^{N} \text{F1}_i. \tag{13}$$

This metric evaluates whether predicted expression hotspots overlap with true active regions within each radius-defined neighborhood, providing a spatially tolerant measure of local transcriptional agreement.

Table S4: Location Ablation at Radius $r = 0.5$ Over 5 Held-Out Slices (mean $\pm$ std).

Method	Gridized L1@0.5↓	Gridized KL@0.5↓
Closest slice	0.0041 ± 0.0023	0.096 ± 0.179
MIMYR	0.0027 ± 0.0011	0.001 ± 0.001

Table S5: Cluster Ablation at Radius $r = 0.5$ Over 5 Held-Out Slices (mean $\pm$ std).

Method	Soft Accuracy@0.05↑	Delaunay Colocalization↓
Majority baseline	0.401 ± 0.076	5.45 ± 1.06
MIMYR	0.472 ± 0.090	4.59 ± 0.31

D ABLATIONS: DISENTANGLING THE THREE STAGES OF RECONSTRUCTION

To determine which components of MIMYR contribute to its performance, we ablate each stage of the pipeline – cell locations, cell identities, and gene expression – and replace the learned component with a rule-based alternative while keeping the remaining stages fixed. This isolates the contribution of MIMYR at different levels of the reconstruction hierarchy.

Location generation. We tested whether MIMYR's location prediction module using guided diffusion is necessary for realistic spatial layouts or if it is sufficient to simply copy cell coordinates from a nearby slice. We evaluated the reconstructed cell locations using two density-based metrics. *Gridized L1* measures the absolute difference between predicted and ground-truth cell densities evaluated over a spatial grid. *Gridized KL* compares the full probability distributions, penalizing misplaced mass more strongly and highlighting structural mismatches.

Using the closest slice as a proxy produces reasonable cell locations but fails to match the true density landscape (**Tab.** S4). MIMYR's location module reduces Gridized L1 from 0.0041 to 0.0027 (a $\sim$34% reduction) and lowers KL from 0.096 to 0.001. This shows that MIMYR's location module learns a continuous anatomical prior rather than reusing discrete arrangements from neighboring tissue slices.

Cell identity assignment. We next replaced MIMYR's cell identity module with majority voting among nearby reference cells. We report *Soft Accuracy*, which compares predictions to ground truth using the cosine similarity of the distribution of cell types within a local spatial neighborhood, rewarding agreement in regional composition rather than exact one-to-one matches. We also compute *Delaunay colocalization*, which builds adjacency graphs from the Delaunay triangulation of cells within the tissue and measures how well predicted cell-type contacts match those computed from the ground truth.

MIMYR improves over the baseline with the soft accuracy increasing from 0.401 to 0.472 and Delaunay error reducing from 5.45 to 4.59. These gains indicate that global spatial correspondences matter: local voting propagates annotation noise and boundary artifacts, whereas the learned mapping recovers smoother, anatomically coherent organization.

Gene expression generation. Lastly, we compared MIMYR's gene expression module against a strong non-parametric rule that copies the transcriptome of the nearest reference cell with the same cell type. As with previous evaluations, we evaluate predictions using *Soft Spearman*, which measures correlation of summed gene expression within local spots, measuring preservation of spatial expression structure even when cells are not in one-to-one correspondence. MIMYR's gene expression module achieves slightly higher agreement than the lookup baseline (0.516 vs. 0.512) while also exhibiting lower variability across slices.

This ablation in particular is performed in a data-sparse setting designed to mimic realistic deployment, where only limited neighboring information is available. This is because lookup from a neighboring slice can be thought of as an upper-bound, since the same cell types when adjacent have

Table S6: Expression Ablation at Radius $r = 0.5$ Over 5 Held-Out Slices (mean $\pm$ std).

Method	Soft Spearman↑
Lookup	0.512 ± 0.051
MIMYR	0.516 ± 0.041

very similar expression. The main advantage of MIMYR is that it is able to extend the gene panel, something lookup is inherently incapable of doing.

Each module of MIMYR improves over its rules-based counterpart. We note that one downside of the three-step generation process can be the propagation of errors from one step to the next, however these ablations show that the earlier steps of MIMYR's generation tend to show the highest improvements over the baseline. Rule-based copying can perform well when close references are available, but it cannot extrapolate outside the observed data distribution.

E DATA AVAILABILITY

All datasets used in this study are publicly available. The primary MERSCOPE whole-brain atlas and companion MERFISH dataset are available through the published resources of Yao et al. (Yao et al., 2023) and Zhang et al. (Zhang et al., 2023), respectively, including aligned coordinates and cell-type annotations. The Alzheimer's disease MERFISH dataset from Johnston et al. (Johnston et al., 2025) is similarly accessible through its associated publication.

F CODE AVAILABILITY

The source code for MIMYR is available on GitHub: https://github.com/gkrieg/mimyr.