SUICA: Learning Super-high Dimensional Sparse Implicit Neural Representations for Spatial Transcriptomics

Qingtian Zhu†, **Yumin Zheng**†, Yuling Sang, Yifan Zhan, Ziyan Zhu, Jun Ding*, Yinqiang Zheng*

UTokyo · McGill · ICML International Conference On Machine Learning

Background

- Spatial Transcriptomics (ST) is a spatially resolved and high-dimensional measurement of gene expression.
- Whole Slice Imaging (WSI) vs Transcriptomics vs ST:
  - **WSI:** Shows a static view of the structure, shape, and organization of cells. (Morphological features)
  - **Transcriptomics:** Profiles cellular states in a sample based on gene expression but loses spatial information.
  - **ST:** Provides a functional view of gene expression profiles across a tissue section with matched images.

Challenges

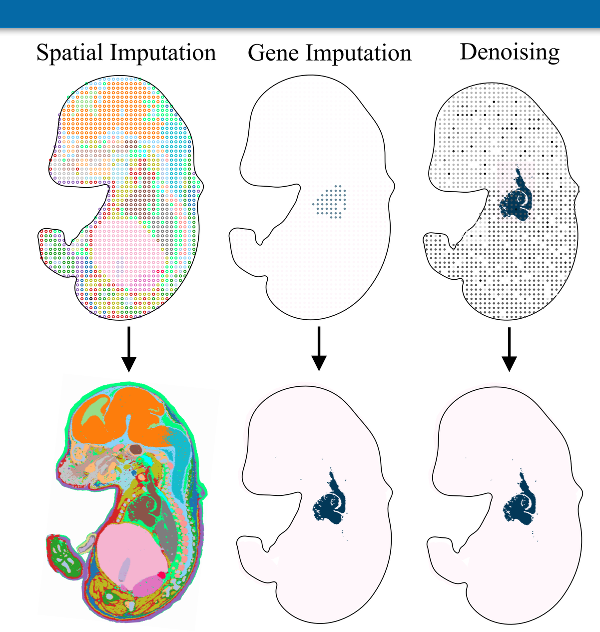
- ST data is high-dimensional, noisy and very sparse.
  - The high sparsity and noisy nature of ST weakens the bio-signatures for analysis
- Trade-offs in ST: high resolution vs high cost
  - No existing ST platform is both affordable and capable of providing high resolution.

Contributions

- Introduced SUICA to model ST data as a continuous, compact representation while preserving authenticity.
- Enabled Implicit Neural Representations (INRs) to process high-dimensional gene expression through a Graph Autoencoder and a classification loss.
- Demonstrated SUICA's strong imputation and denoising performance across various ST datasets, facilitating subsequent analyses.

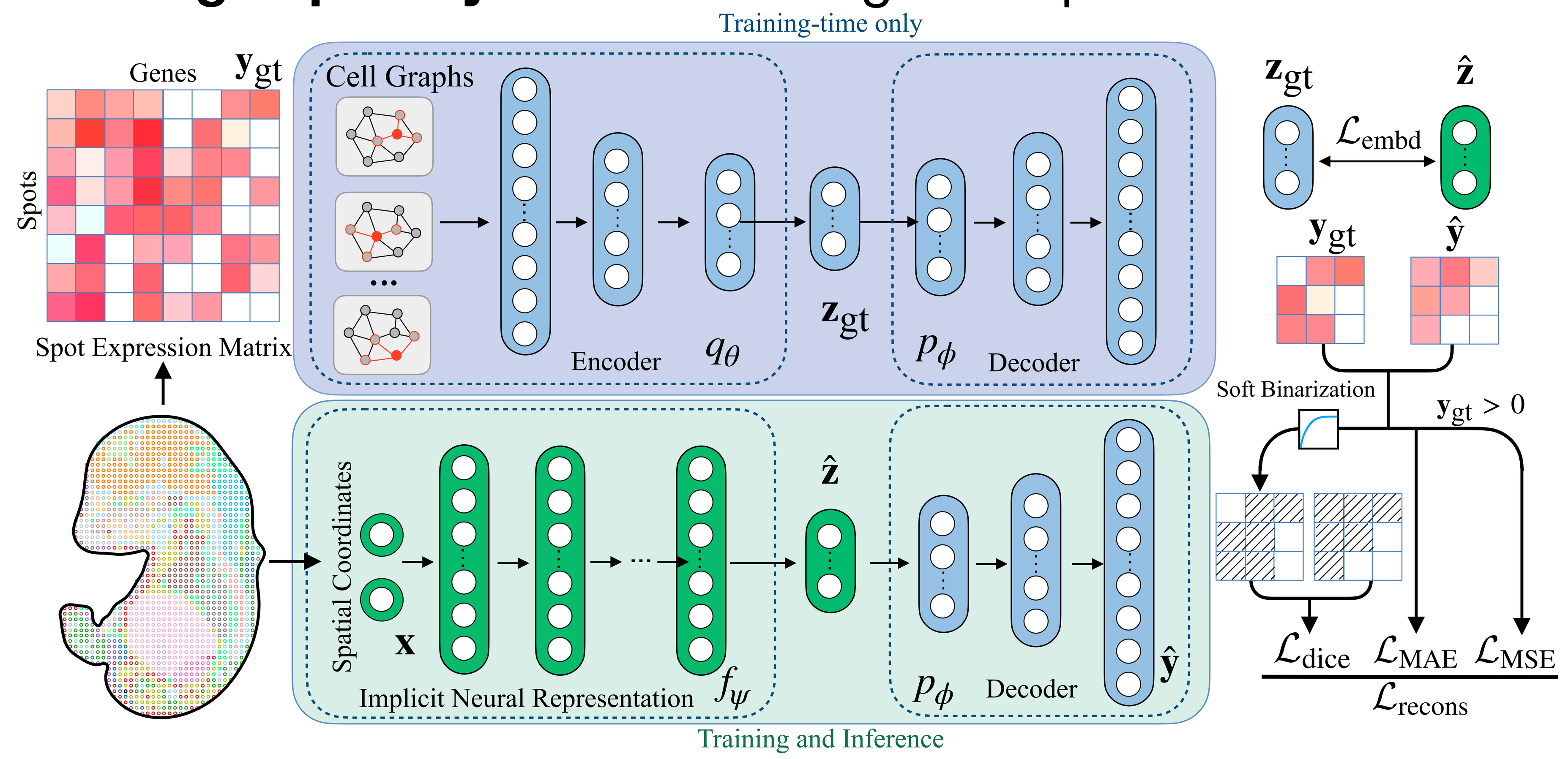
Value

- Applicable to all ST platforms
- Produce ST data with
  - Higher resolution
  - Lower noise level
  - Stronger bio-signatures
- No extra cost
- High data efficiency



Methodology

- **Challenges in applying INRs to ST data:**
  - **High dimensionality:** 1k~30k genes per spatial location
  - **High sparsity:** Zero-inflated gene expression values



- **SUICA overall framework:**
  - Train a graph autoencoder (gae) to learn low-dimensional representations of ST data

$$\mathcal{L}_{gae} = \frac{1}{|\mathbf{M}_\mathbf{y}|}\sum_{\mathbf{M}_\mathbf{y}}(\hat{\mathbf{y}} - \mathbf{y}_{gt})^2$$

  - Learn the coordinate-to-representation mapping using INRs

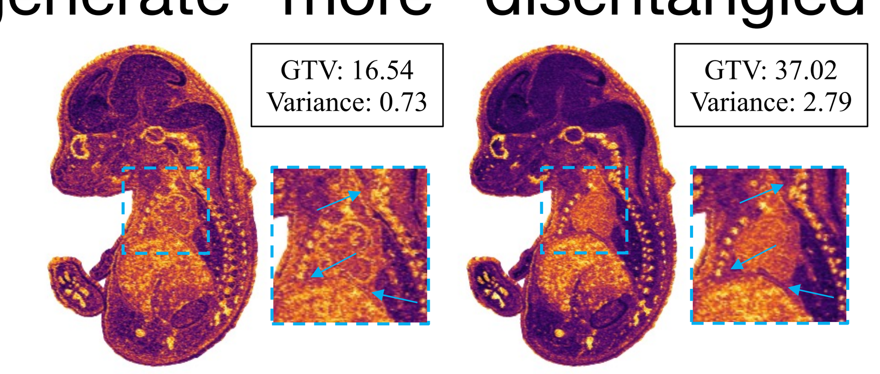
$$\mathcal{L}_{embd} = \frac{1}{|\mathbf{M}_\mathbf{z}|}\sum_{\mathbf{M}_\mathbf{z}}(\hat{\mathbf{z}} - \mathbf{z}_{gt})^2$$

  - Reconstruct gene expression profiles from the low dimensional representations with DICE loss to address zero inflation issue

$$\mathcal{L}_{recons} = \frac{1}{|\mathbf{M}_\mathbf{y}^+|}\sum_{\mathbf{M}_\mathbf{y}^+}(\hat{\mathbf{y}}-\mathbf{y}_{gt})^2 + \frac{1}{|\mathbf{M}_\mathbf{y}|}\sum_{\mathbf{M}_\mathbf{y}}|\hat{\mathbf{y}}-\mathbf{y}_{gt}| + \lambda\mathcal{L}_{dice}$$

$$\mathcal{L}_{dice} = 1 - \frac{2\sum(\tanh(\hat{\mathbf{y}}) \circ \mathrm{sgn}(\mathbf{y}_{gt})) + \epsilon}{\sum\tanh(\hat{\mathbf{y}}) + \sum\mathrm{sgn}(\mathbf{y}_{gt}) + \epsilon}$$

- The graph encoder can generate more disentangled representations.
  - GTV: graph total variance

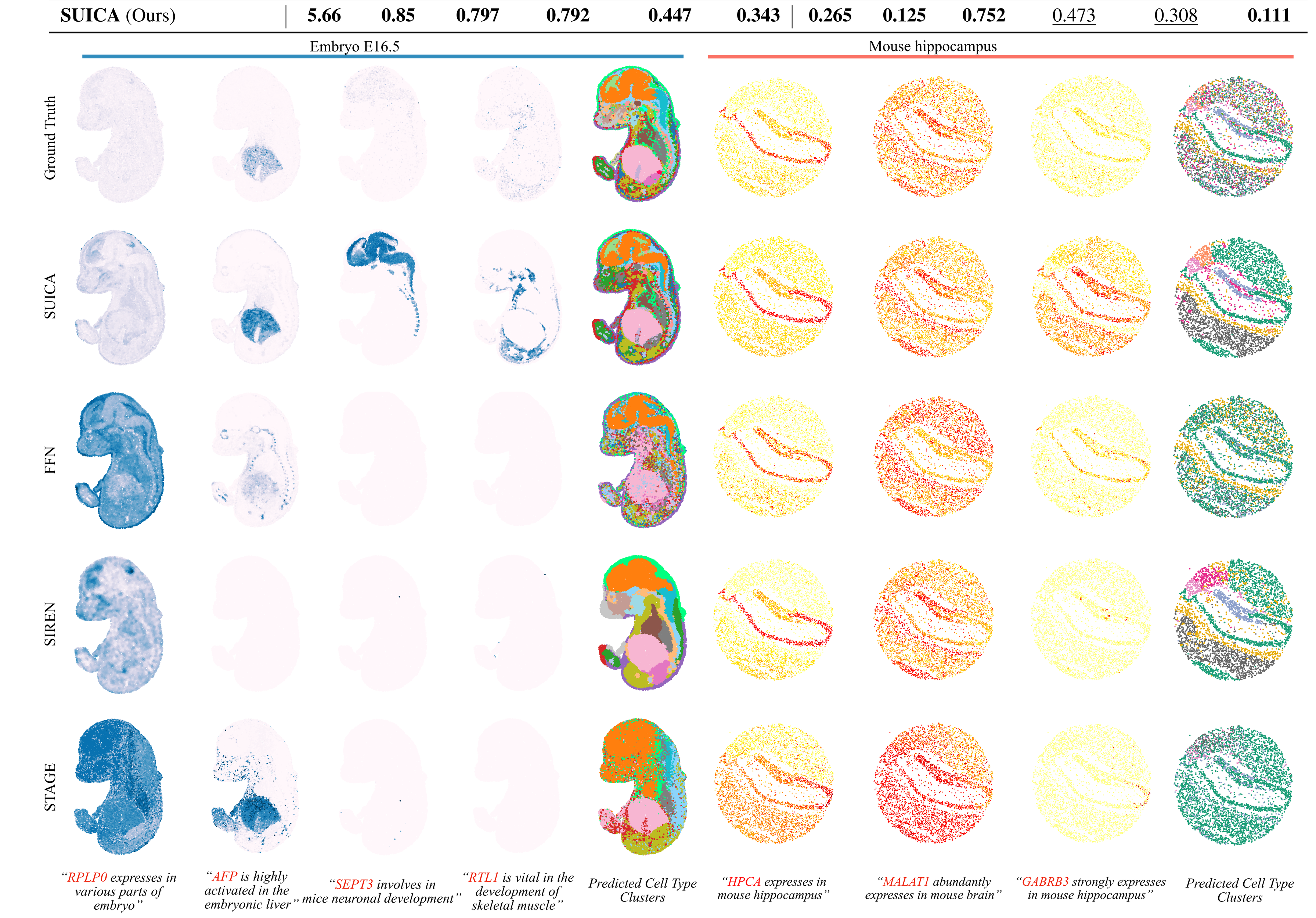


(a) AE (b) GAE

Results

- Quantitative and visual comparisons of spatial imputation (super-resolution) performance on various ST platforms.

Methods	Stereo-seq MOSTA MAE↓	MSE↓	Cosine↑	Pearson↑	Spearman↑	ARI↑	Mouse hippocampus Slide-seqV2 MAE↓	MSE↓	Cosine↑	Pearson↑	Spearman↑	ARI↑
FFN (Tancik et al., 2020)	<u>6.51</u>	1.20	0.706	0.718	<u>0.400</u>	0.143	0.378	0.215	0.499	0.442	0.274	0.0523
SIREN (Sitzmann et al., 2020)	7.21	1.31	0.661	0.678	0.247	<u>0.289</u>	0.383	0.216	0.494	0.452	0.248	<u>0.110</u>
STAGE (Li et al., 2024)	6.52	<u>1.11</u>	<u>0.732</u>	<u>0.747</u>	0.365	0.139	<u>0.351</u>	<u>0.198</u>	<u>0.587</u>	**0.483**	**0.314**	0.0361
SUICA (Ours)	**5.66**	**0.85**	**0.797**	**0.792**	**0.447**	**0.343**	**0.265**	**0.125**	**0.752**	<u>0.473</u>	<u>0.308</u>	**0.111**



Methods	Visium-Human Brain MAE↓	MSE↓	Cosine↑	Pearson↑	Spearman↑	Visium-Mouse Brain MAE↓	MSE↓	Cosine↑	Pearson↑	Spearman↑	ARI↑
FFN (Tancik et al., 2020)	5.76	0.881	0.772	0.786	<u>0.402</u>	5.95	5.85	0.832	0.741	0.581	0.000587
SIREN (Sitzmann et al., 2020)	6.58	0.933	0.756	0.747	0.196	5.35	4.29	0.878	<u>0.804</u>	0.647	<u>0.359</u>
STAGE (Li et al., 2024)	6.19	0.805	0.795	0.772	0.223	<u>4.55</u>	<u>3.20</u>	<u>0.918</u>	**0.825**	**0.666**	0.140
TRIPLEX (Chung et al., 2024)	**4.75**	**0.560**	**0.881**	**0.850**	0.319	9.35	14.0	0.00	-0.00682	-0.00715	0.358
UNIv2 (Chen et al., 2024)	7.30	1.41	0.723	0.633	0.129	6.94	7.88	0.790	0.631	0.425	0.228
SUICA (Ours)	<u>4.99</u>	<u>0.567</u>	<u>0.860</u>	<u>0.846</u>	**0.445**	**3.68**	**2.45**	**0.932**	0.800	<u>0.660</u>	**0.393**

- Quantitative comparisons of the gene imputation and denoising performance on mouse embryo E16.5 stereo-seq data.

Methods	Gene Imputation MAE↓	MSE↓	Cosine↑	Pearson↑	Spearman↑	Denoising MAE↓	MSE↓	Cosine↑	Pearson↑	Spearman↑
FFN (Tancik et al., 2020)	4.88	0.963	0.731	0.610	0.251	7.90	1.95	0.266	0.285	0.0523
SIREN (Sitzmann et al., 2020)	6.44	1.12	0.675	0.652	0.124	7.91	1.97	0.112	0.103	0.0166
STAGE (Li et al., 2024)	<u>4.69</u>	<u>0.738</u>	**0.802**	<u>0.705</u>	<u>0.264</u>	<u>7.60</u>	<u>1.66</u>	<u>0.606</u>	<u>0.630</u>	<u>0.182</u>
SUICA (Ours)	**4.30**	**0.724**	<u>0.798</u>	**0.714**	**0.317**	**6.03**	**0.934**	**0.733**	**0.737**	**0.379**

- Ablation studies:

Settings	Embryo E16.5 MSE↓	Cosine↑	Pearson↑	Human Brain MSE↓	Cosine↑	Pearson↑
Vanilla INR	2.35	0.668	0.653	9.33	0.756	0.747
+AE	1.60	0.789	0.751	11.27	0.695	0.691
+Dice	1.48	0.806	0.747	7.05	0.826	0.800
+Graph	1.47	0.807	0.761	5.67	0.860	0.846

- Data efficiency:

%	MAE↓	MSE↓	Cosine↑	Pearson↑
80%	8.01	1.47	0.807	0.761
60%	7.96	1.52	0.801	0.752
40%	8.00	1.59	0.790	0.739
20%	8.14	1.62	0.786	0.738