HoloGene

Learning to Lift Sliced Spatial Transcriptomics to Holistic 3D Gene Fields

Yumin Zheng[1,2,*], **Qingtian Zhu**[3,*], Ziyan Zhu[4], Kailu Song[1,2], Yinqiang Zheng[3,†], Jun Ding[1,2,5,†]

[1]McGill University [2]MUHC Research Institute [3]The University of Tokyo [4]Carnegie Mellon University [5]Mila – Quebec AI Institute

Code

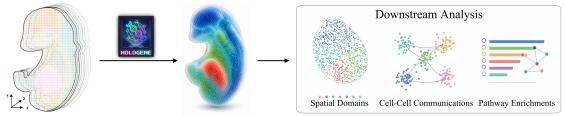
From a few 2D sections to **gene expression at any *(x, y, z)***: HoloGene lets **depth *z* modulate** a dense in-plane *(x, y)* field, instead of treating *z* as a third coordinate.

1 Why 3D gene fields?

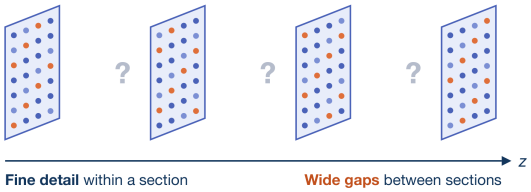
Biology is 3D. Organogenesis and tissue organization cannot be read from one slice, so spatial transcriptomics (ST) is moving from 2D sections to 3D atlases.

But 3D ST is still measured slice by slice. The tissue between sections is never sequenced.

Goal: a continuous gene field, queryable at any *(x, y, z)*, for virtual sections and volumetric analysis.



2 What makes it hard?



Anisotropy. Standard INRs encode *(x, y, z)* symmetrically and ignore that depth is sampled far more sparsely. Sections are also processed independently, which adds slice-specific variation.

> **Our answer.** Feed the network only the dense *(x, y)* plane and let depth *z* modulate it.

TAKEAWAYS

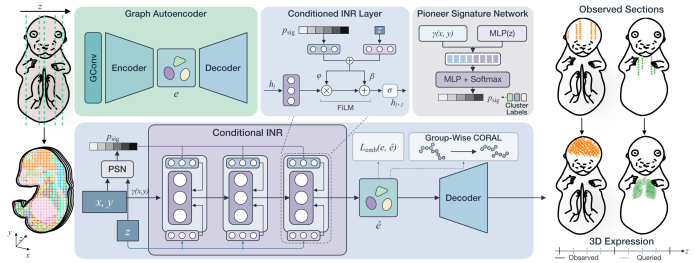
Acquisition geometry matters: moving *z* from input to condition alone already improves reconstruction.

Virtual sections are usable for biology: pathways, cell–cell communication and DEGs are preserved.

Scope: needs pre-registered sections; queries between sections are model-based estimates.

3 Idea: depth *z* modulates an in-plane *(x, y)* field

$$\hat{g}(x, y, z) = \mathrm{Dec}(S_\theta(\underbrace{x, y}_{\text{dense: input}} \mid \underbrace{z, p_{\text{sig}}}_{\text{sparse: condition}}))$$

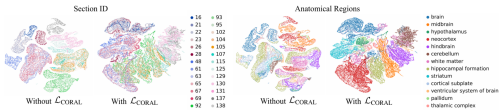


1 Compress
A graph autoencoder turns expression into low-dimensional embeddings e, the regression target.

2 Guide
The Pioneer Signature Network predicts tissue-state probabilities p_{sig} from Leiden pseudo-labels.

3 Synthesize
The INR sees only *(x, y)*; *z* and p_{sig} scale and shift every layer (FiLM).

4 Align
Signature-stratified CORAL matches latent covariances across slices, so slice identity is not learned as biology.



With CORAL, sections mix better while anatomical regions stay separated (whole mouse brain).

4 Does each part matter?

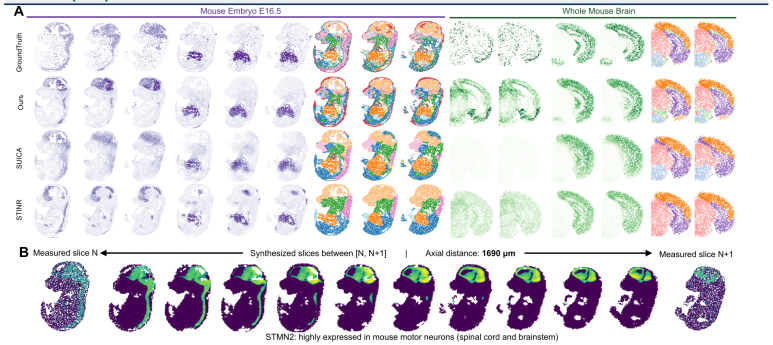
Conditioning on *z* (FiLM-only) beats isotropic XYZ input; PSN and CORAL add complementary gains.

Method	MOSTA E16.5			Whole Mouse Brain		
	MSE_m↓	Pearson↑	ARI↑	MSE_m↓	Pearson↑	ARI↑
Isotropic XYZ	0.2413	0.6869	0.1310	2.5460	0.4446	0.4155
FiLM-only	0.2354	0.7132	0.1454	2.5041	0.4636	0.4231
FiLM + PSN	0.2308	0.7134	0.1440	2.4899	0.4636	0.4398
FiLM + CORAL	**0.2235**	**0.7154**	0.1569	**2.4715**	0.4644	0.4354
Full HoloGene	0.2303	0.7146	**0.1637**	2.4835	**0.4648**	**0.4670**

5 Does it work on unseen sections?

Most accurate reconstruction (80/20 split along *z*; test sections never seen): best on 9 of 10 metrics; whole-brain ARI rises from 0.428 to 0.467.

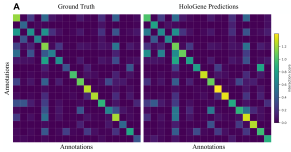
Method	MOSTA E16.5					Whole Mouse Brain				
	MSE_m↓	Cos_m↑	Pearson↑	$Spear_m$↑	ARI↑	MSE_m↓	Cos_m↑	Pearson↑	$Spear_m$↑	ARI↑
FFN	0.2588	0.7630	0.6525	0.5034	0.1009	2.5616	0.7443	0.4414	0.4009	0.4275
SIREN	0.2835	0.7373	0.6165	0.4695	0.0187	2.5778	0.7717	0.4624	0.4262	0.3837
STAGE	0.2987	0.7234	0.6655	0.5237	0.1213	2.9103	0.7356	0.4020	0.3662	0.3313
SpatialZ	0.3145	0.7029	0.5315	0.4722	0.1188	3.3409	0.5279	0.2335	0.2070	0.0221
STINR	0.2332	0.7971	0.6986	0.5308	0.1422	2.7411	0.7782	0.4289	0.3895	0.2586
SUICA-3D	0.2972	0.7233	0.6772	0.5412	0.1393	2.7776	**0.7803**	0.4130	0.3839	0.3618
HoloGene (Ours)	**0.2303**	**0.8054**	**0.7146**	**0.5503**	**0.1637**	**2.4835**	0.7748	**0.4648**	**0.4315**	**0.4670**



Downstream biology is preserved, not just expression values: best on all 8 pathway, cell–cell communication (CCC) and DEG metrics.

Method	MOSTA E16.5				Whole Mouse Brain			
	Path. F1↑	CCC Pearson↑	CCC Top-F1↑	DEG F1↑	Path. F1↑	CCC Pearson↑	CCC Top-F1↑	DEG F1↑
FFN	0.5348	0.8654	0.8892	0.4389	0.3462	0.9708	0.8800	0.8346
SIREN	0.3419	0.8444	0.8851	0.4056	0.5538	0.9767	0.8956	0.8615
STAGE	0.4068	0.8986	0.8880	0.2861	0.4242	0.8977	0.5444	0.6269
SpatialZ	0.6490	0.8806	0.8642	0.5421	0.4935	0.9752	0.8991	0.8731
STINR	0.5674	0.8786	0.8823	0.4306	0.4545	0.9618	0.8489	0.7192
SUICA-3D	0.5163	0.9122	0.8509	0.4167	0.3529	0.9245	0.8222	0.6692
HoloGene (Ours)	**0.7239**	**0.9342**	**0.9138**	**0.6278**	**0.5714**	**0.9804**	**0.9022**	**0.8808**

Cell-Cell Communications on MOSTA E16.5



ACKNOWLEDGMENTS

Fonds de recherche Santé Québec
CIHR IRSC
NSERC CRSNG
AMED
JST-Mirai Program
AIRIS
科研費 KAKENHI