STINR: Deciphering Spatial Transcriptomics via Implicit Neural Representation

Supplementary Material

6. Implementation Details

STINR is optimized by using Adam optimizer with learning rate 0.001 for all datasets. We have consistently used a four-layer INR $f_\theta(\cdot)$ for all datasets in this work. The number of hidden neurons of the INR network is set to 400 for the human lymph node data and 200 for other ST datasets. The depth of the encoder-decoder is set to 4 (two for encoder and two for decoder) for all datasets, and the number of hidden neurons of the encoder-decoder is set to 200 for all datasets. For gene imputation, the loss functions (5), (7), and (8) are refined through introducing a mask to avoid calculating loss for those zero-count spots. All experiments on conducted on a PC of i5-12400F CPU and RTX 3090 GPU, with 24 GB memory. Running time is reported using the Time package in Python. All quantitative results are rounded to three decimal places.

For gene imputation and denoising, we additionally added the weighted TV regularization into the loss function, which is defined as

$$\mathcal{L}_{\mathrm{TV}} := \alpha_{\mathrm{TV}} \left| \mathbf{w}|_{(i)} \sum_{g=1}^{G} \frac{\partial f_\theta(\boldsymbol{V})|_{(i,g)}}{\partial \boldsymbol{V}|_{(i,g)}} \right|, \qquad (10)$$

where α_{TV} is a trade-off parameter, $\boldsymbol{V}$ is the input coordinate matrix of INR, and $f_\theta(\cdot)$ is the INR. We set the parameter α_{TV} to $\alpha_{\mathrm{TV}} = 8 \times 10^{-4}$ for mouse brain datasets and $\alpha_{\mathrm{TV}} = 1 \times 10^{-4}$ for mouse embryo datasets when conducting gene imputation and denoising. The TV regularization uses the gradients of INR outputs w.r.t. the input coordinates to enhance local smoothness of the output $\boldsymbol{X}_{\mathrm{rec}} = f_\theta(\boldsymbol{V})$ by penalizing such gradients. Here, $\mathbf{w} \in \mathbb{R}^N$ is a weighted parameter that assigns different local smooth weight to different spots. Specifically, we update $\mathbf{w}$ every 100 iterations by using the following paradigm:

$$\mathbf{w}|_{(i)} = \left(\sum_{g=1}^{G} \frac{\partial f_\theta(\boldsymbol{V})|_{(i,g)}}{\partial \boldsymbol{V}|_{(i,g)}} + \epsilon \right)^{-1}. \qquad (11)$$

After updating the weight parameter $\mathbf{w}$, we fix it in the next 100 iterations until we update it again. The philosophy of updating the weight parameter $\mathbf{w}$ by using the reciprocal of the gradients is to assign larger weights to smoother regions with lower gradients, while assigning lower weights to sharper regions with larger gradients. The ϵ is a small constant to avoid dividing by zero.

7. Dataset Details

The mouse brain coronal datasets were sourced from 10x Genomics[11]. We selected three "Mouse Brain Section (Coronal)" files and used the filtered barcode matrix to conduct experiments. The mouse brain single-cell data was obtained from Mousebrain.org[12] by selecting the "L5_All.loom" file, which contains gene expressions with one cell per column and one gene per row. The mouse embryo organogenesis datasets were sourced from STomicsDB[13], and we selected the E9.5, E10.5, and E11.5 growing stages under configuration E1S1 to conduct experiments. The DLPFC datasets were downloaded from SpatialLIBD[14] by selecting the "h5_filtered" files under raw data section. Additional metadata files of DLPFC were downloaded from HumanPilot[15]. The single-cell reference data of DLPFC was downloaded from GEO[16] by selecting the file "GSE144136_GeneBarcodeMatrix_Annotated.mtx.gz". The human SCC dataset was sourced from GEO[17] by selecting the "GSE144240_RAW.tar" file. The human lymph node multi-omics data was sourced from GEO[18] by selecting the "GSM8195494 A1 Lymph Node" file. We remark that the imputation method GNTD [26] utilizes additional protein-protein interaction network datasets as auxiliary information for each ST dataset to perform imputation. For easier implementation on different ST datasets and fair comparisons, we have not considered such auxiliary information in order to test the stand-alone reconstruction ability of each method. Future work on combining STINR with additional biological structure priors such as protein interactions would be interesting.

8. More Results

We provide some real-world gene imputation examples by using raw sparse gene expressions in Fig. 9. STINR performs well for real gene imputation. We have also provided more qualitative gene imputation results on mouse

---

[11] https://www.10xgenomics.com/datasets
[12] http://mousebrain.org/adolescent/
[13] https://db.cngb.org/stomics/mosta/download/
[14] http://spatial.libd.org/spatialLIBD
[15] https : / / github . com / LieberInstitute / HumanPilot/tree/master/10X
[16] https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144136
[17] https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144240
[18] https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE263617

brain and mouse embryo datasets in Fig. 11 and Fig. 12. STINR more accurately recovers the spatial maps of gene expressions. We also present more qualitative cell-type deconvolution results on DLPFC in Fig. 10. The cell proportion maps of STINR are visually more pleasant and smooth, more robustly suggesting the cortex layer specificity of each cell-type within the human brain. Quantitatively, STINR obtains consistently better deconvolution accuracy by evaluating the layer specificity of each deconvoluted cell-type through AUC. We have also provided more spatial domain detection results on DLPFC in Fig. 13. STINR more accurately identifies cortical layers of the human brain.

9. More Discussions

The proposed INR reconstruction is effective for gene imputation/denoising because INR can implicitly capture the smoothness between spots, and hence can recover the gene expression at arbitrary input coordinate by leveraging neighboring spots. We further test the robustness of STINR's cell-type deconvolution results w.r.t. various degrees of gene expression sparsity (Table 6). STINR is robust to such sparsity. We further include more cell-type deconv. results (Table 7 and Fig. 8) to show the superiority of STINR. Besides, we further test a sparse Lasso penalty loss for cell types, which could further enhance the performance, i.e., AUC 0.887 vs. 0.880 (w/ vs. w/o Lasso) on DLPFC. While STINR is designed for ST, it can be easily extended to multi-omics by using multiple INRs. For consistency, we have used a simple encoder-decoder structure (same as ST) for multi-omics, and STINR already shows competitive performances and potential under such simple setting. We can consider more advanced multi-modal network structures (e.g., cross-modality attention network or multi-modal representation learning) to enhance STINR for multi-omics in future research.

Our method depends on the single-cell reference matrix. We perform a robustness test of our method w.r.t. an inaccurate reference matrix by adding noise to the reference matrix (Table 6). STINR is robust to an inaccurate reference matrix. We perform a principal ablation study of INR modeling by directly inputting the observed data X_s into the MLP (i.e., discrete representation) instead of coordinates (i.e., continuous representation). The results for gene imputation are MSE 0.185 vs. 0.156 (INR), which reveals the benefits of continuous representation using INR. Since we have used an encoder-decoder, the gene imputation result of the reconstructed ST data X_{rec} is similar to that of the reconstructed normalized ST data $q_\kappa(Z)$, e.g., MSE 0.156 vs. 0.158 (X_{rec} vs. $q_\kappa(Z)$). Finally, using the reconstructed ST data X_{rec} (the output of INR) for spatial domain detection is better than using the original ST data X_s, e.g., ARI 0.604 vs. 0.377 (X_{rec} vs. X_s). We further compare STINR with more SOTA methods DestVI (Nat. Bio., 2022), CARD

Table 6. Robustness tests of STINR's cell-type deconvolution results w.r.t. sparsity of gene expressions and noise perturbation of single-cell references. Average AUC scores across eight excitatory (Ex_) cell types in DLPFC are reported.

Original	Sparsity degree				Noise deviation			
	10%	20%	30%	40%	0.01	0.02	0.03	0.04
0.880	0.879	0.878	0.876	0.872	0.878	0.873	0.868	0.861

Table 7. Quantitative cell-type deconvolution results on different slices of an adult mouse brain ST dataset (GSE147747). Average AUC scores of four hippocampal cell types are reported.

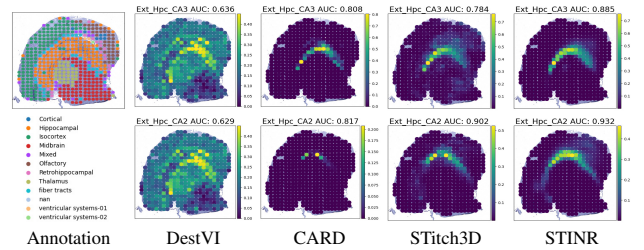| Slice ID	21A	22A	23A	24A	25A	26A	27A
DestVI	0.671	0.633	0.619	0.601	0.630	0.694	0.646
CARD	0.822	0.825	0.813	0.840	0.828	0.803	0.823
STitch3D	0.878	0.859	0.836	0.855	0.789	0.835	0.848
STINR	**0.892**	**0.863**	**0.864**	**0.859**	**0.860**	**0.858**	**0.884**

Figure 8. Cell deconvolution on mouse brain Slice 25A. Compared to DestVI (Nat. Bio., 2022), CARD (Nat. Bio., 2022), and STitch3D (Nat. MI, 2023), STINR more accurately recovers hippocampal (orange annotation) regional cell types (Ext_Hpc_CA).

(Nat. Bio., 2022), and STitch3D (Nat. MI, 2023) for cell-type deconvolution (Table 7 and Fig. 8). STINR achieves better quantitative and visual performances.
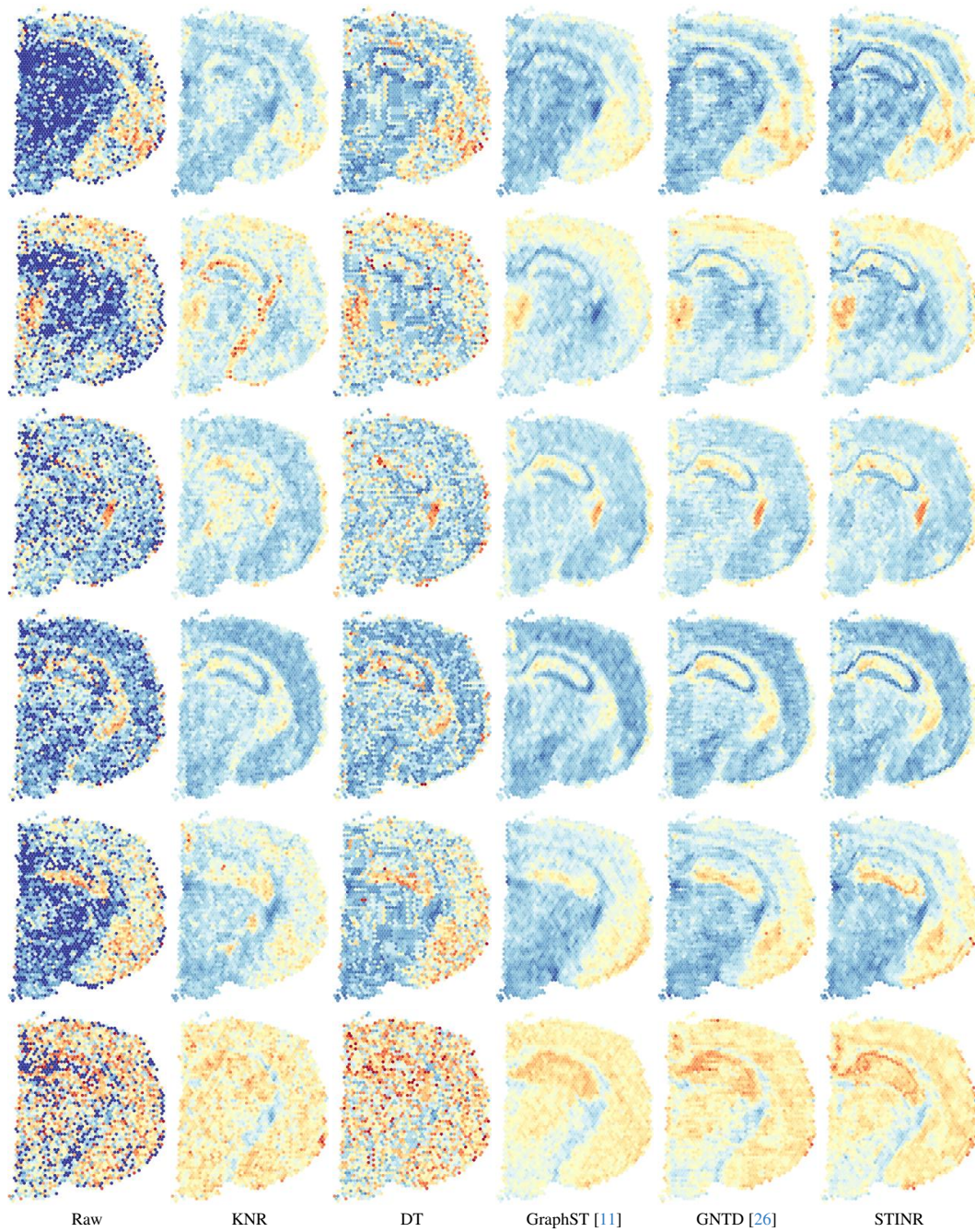
Figure 9. Some authentic examples of real-world gene imputation for raw sparse ST data (some relatively low expression genes in the mouse brain coronal section 1) by different methods. STINR more clearly recovers gene expressions revealing known tissue profiles such as hippocampus. Top to down list the genes "atp2b4", "olfm2", "aldh2", "id4", "ak5", and "ola1".
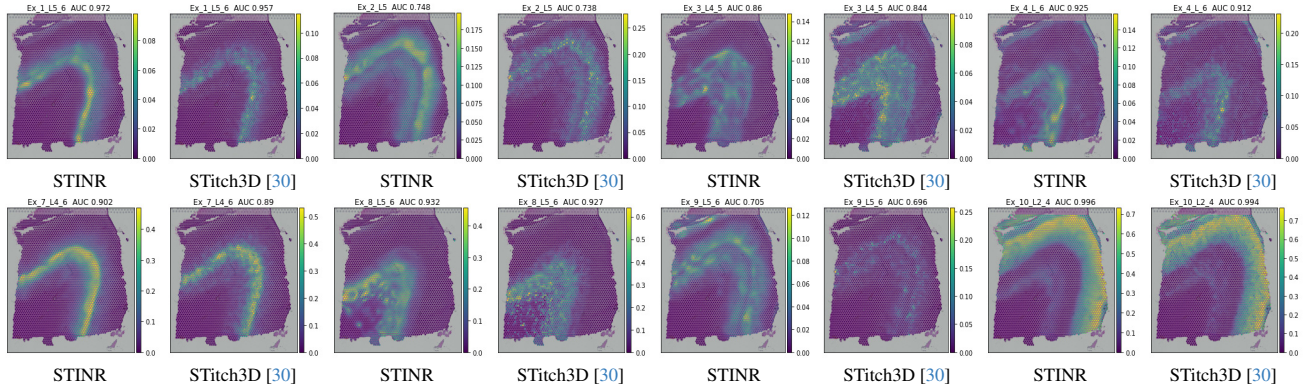
Figure 10. Cell-type deconvolution results for different cell-types on DLPFC section #151673, along with the area under curve (AUC) evaluations on top of the figures.
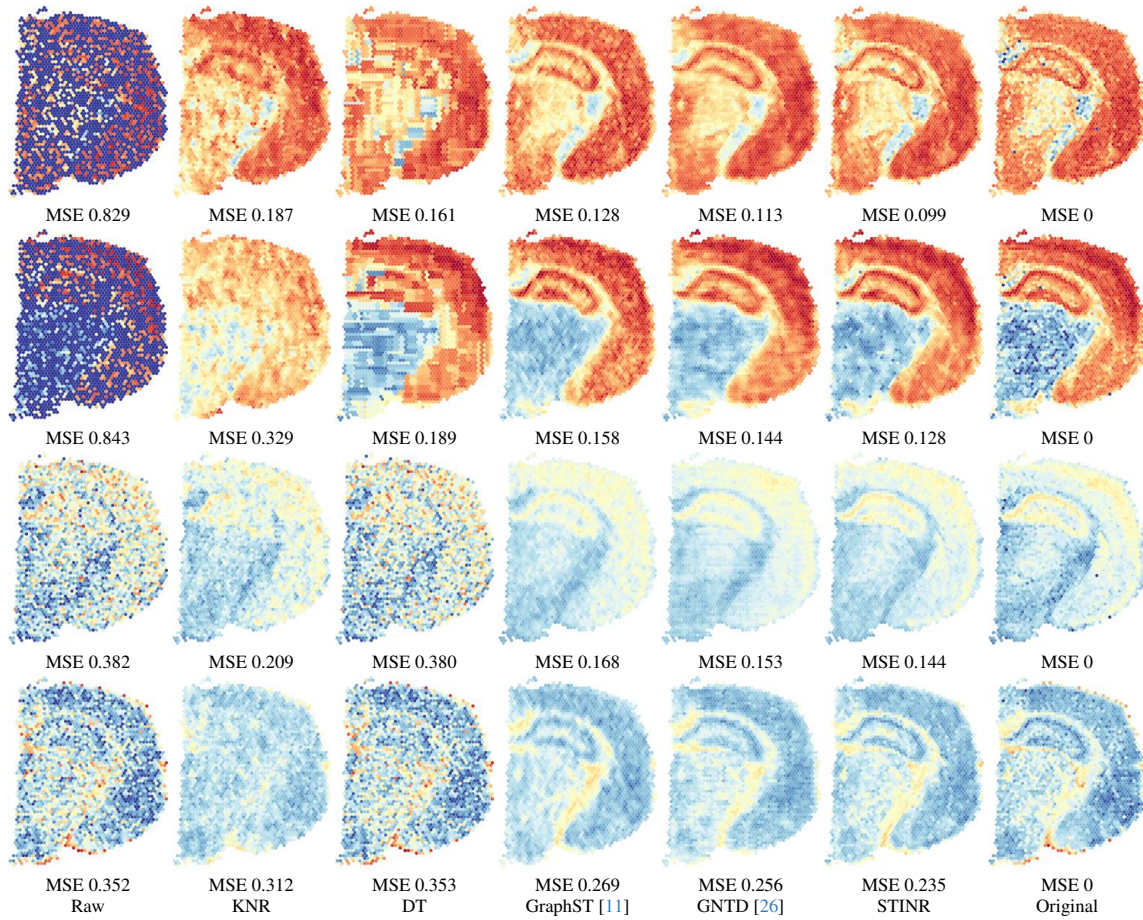

Figure 11. Qualitative results for simulated ST imputation and denoising on mouse brain section 1. From top to down list the results for imputation on gene "rtn1" with SR 0.7, imputation on gene "nrgn" with SR 0.7, denoising on gene "ppp3ca" with noise deviation 1, and denoising on gene "ptgds" with noise deviation 1.
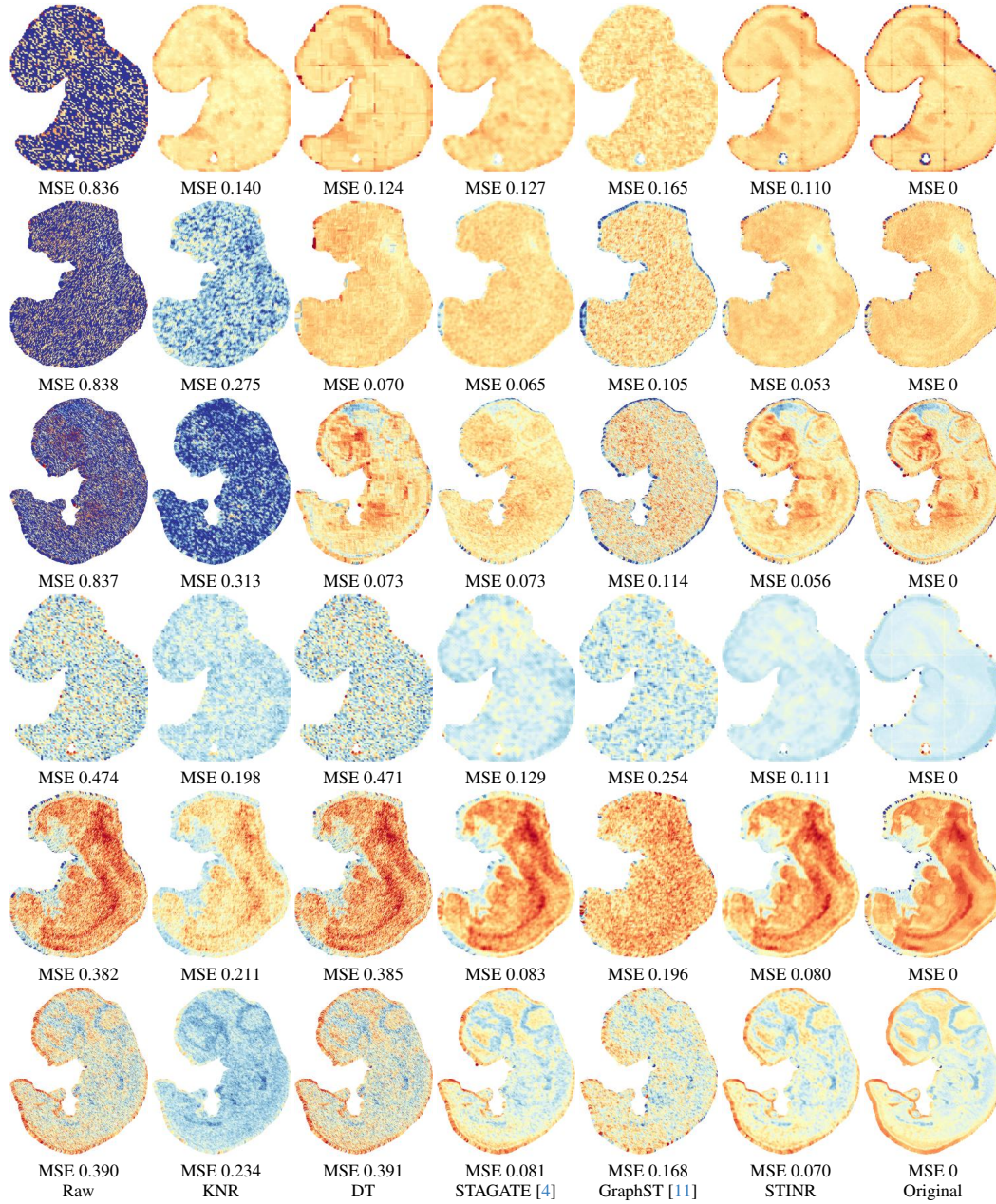
Figure 12. Qualitative results for simulated ST imputation and denoising on mouse embryo organogenesis under growing stages E9.5, E10.5, and E11.5. From top to down respectively list the results for imputation on gene "Rpl41" with SR 0.7 of three growing stages and the results for denoising on gene "Gm42418" with noise deviation 1 of three growing stages.
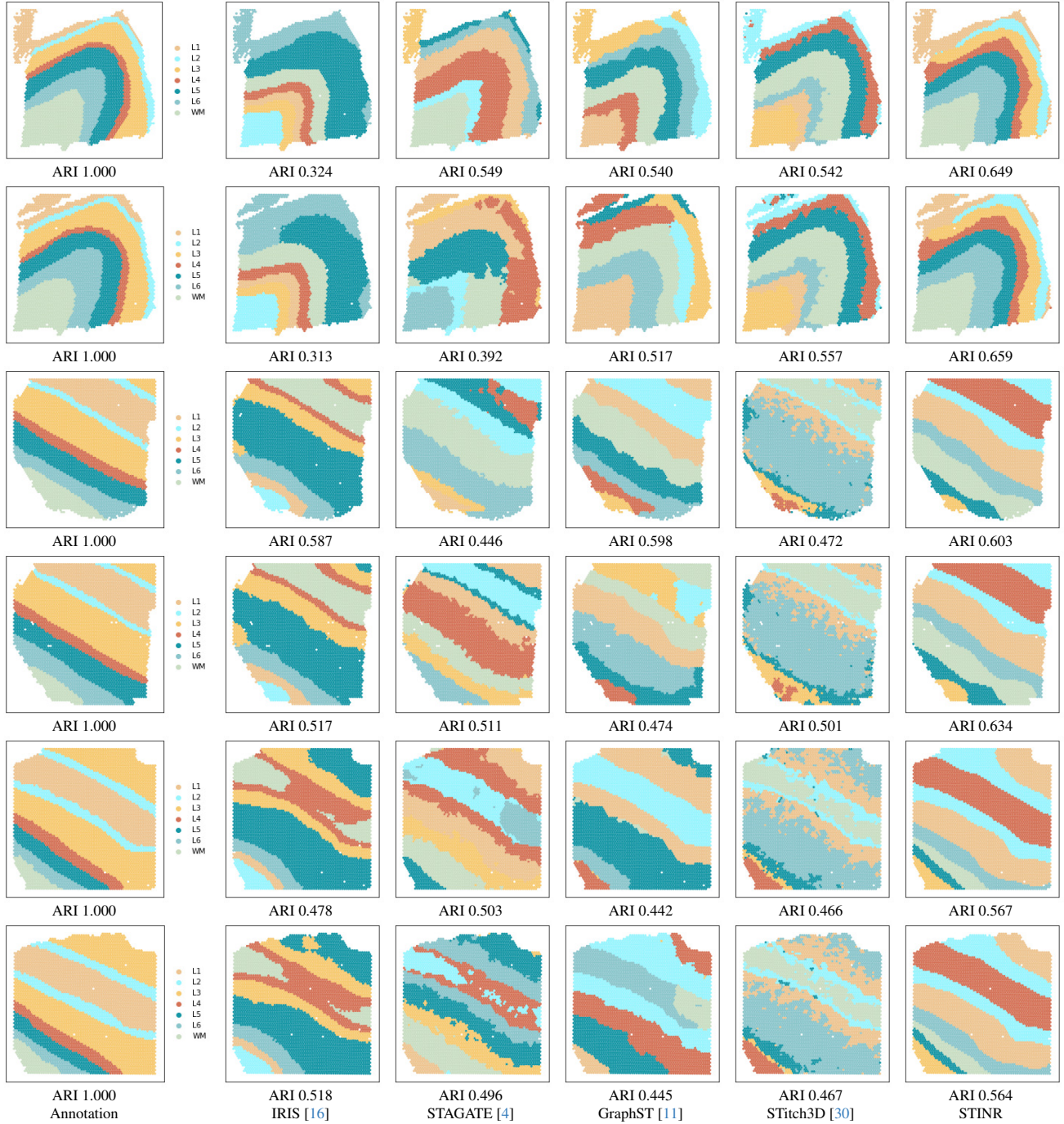
Figure 13. Spatial domain detection results by different methods on the ST data DLPFC sections #151675, #151676, and #151507-#151510. Note that the colors of each domain do not necessarily correspond to the same detected tissues across different methods.