Nucleic Acids Research, 2025, **53**, gkaf158
https://doi.org/10.1093/nar/gkaf158
Methods

stAI: a deep learning-based model for missing gene imputation and cell-type annotation of spatial transcriptomics

Guangsheng Zou[1,†], **Qunlun Shen**[1,†], **Limin Li**[2], **Shuqin Zhang**[1,3,*]

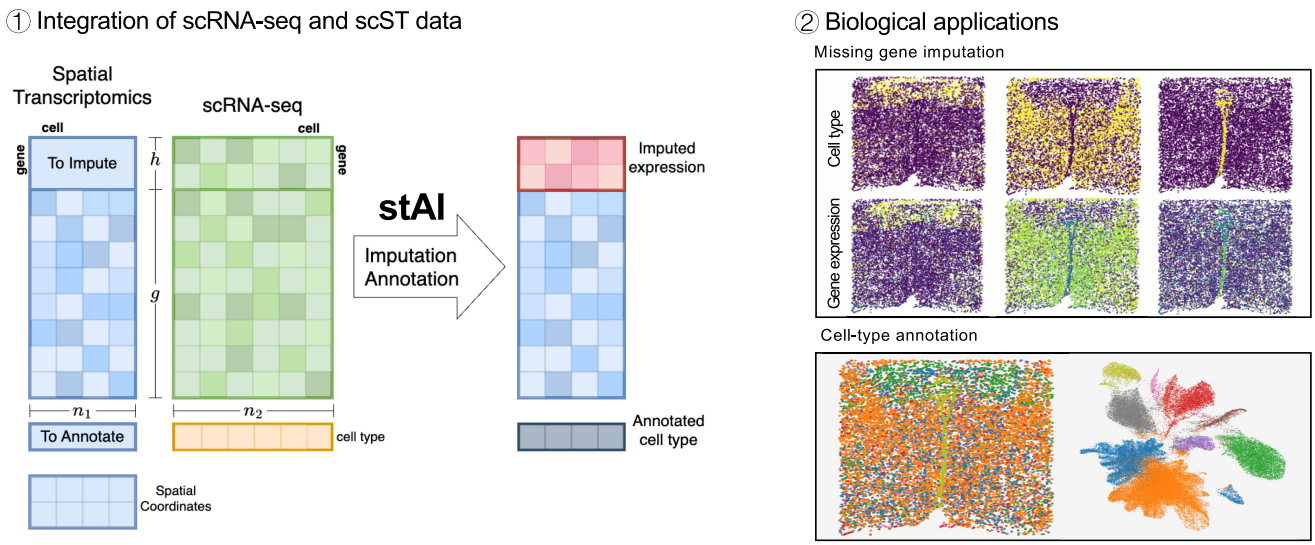
[1]School of Mathematical Sciences, Fudan University, Shanghai 200433, China
[2]School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China
[3]Center for Applied Mathematics, Research Institute of Intelligent Complex Systems, and Shanghai Key Laboratory for Contemporary Applied Mathematics, Fudan University, Shanghai 200433, China

[*]To whom correspondence should be addressed. Email: zhangs@fudan.edu.cn
[†]These authors contributed equally.

Abstract

Spatial transcriptomics technology has revolutionized our understanding of cellular systems by capturing RNA transcript levels in their original spatial context. Single-cell spatial transcriptomics (scST) offers single-cell resolution expression level and precise spatial information of RNA transcripts, while it has a limited capacity for simultaneously detecting a wide range of RNA transcripts, hindering its broader applications. Characterizing the whole transcriptome level and comprehensively annotating cell types represent two significant challenges in scST applications. Despite several proposed methods for one or both tasks, their performance remains inadequate. In this work, we introduce stAI, a deep learning-based model designed to address both missing gene imputation and cell-type annotation for scST data. stAI leverages a joint embedding for the scST and the reference scRNA-seq data with two separate encoder-decoder modules. Both the imputation and annotation are performed within the latent space in a supervised manner, utilizing scRNA-seq data to guide the processes. Experiments for datasets generated from diverse platforms with varying numbers of measured genes were conducted and compared with the updated methods. The results demonstrate that stAI can predict the unmeasured genes, especially the marker genes, with much higher accuracy, and annotate the cell types, including those of small size, with high precision.

Graphical abstract



Received: June 23, 2024. **Revised:** December 29, 2024. **Editorial Decision:** February 11, 2025. **Accepted:** February 18, 2025

Introduction

The emergence of single-cell RNA sequencing (scRNA-seq) technology [1] has revolutionized our understanding of cellular heterogeneity by measuring gene expression at the single-cell level. However, scRNA-seq data loses the spatial information of cells on tissues during the cell dissociation process, preventing the association of gene expression profiles with the spatial behavior of cells. Spatial transcriptomics (ST) techniques [2] allow for investigation of the spatial distribution of gene expression within tissues and organs by generating the RNA transcripts' level in their original spatial context. By capturing the spatial information of gene expression patterns, researchers can gain valuable insights into the organization and function of complex biological systems [2, 3]. Single-cell spatial transcriptomics (scST) technologies that utilize fluorescence in situ hybridization or in situ sequencing, such as seqFISH [4], osmFISH [5], MERFISH [6], and STARmap [7] can offer sing-cell resolution expression and precise spatial information of RNA transcripts. Nonetheless, these methods have a relatively limited capacity for simultaneously detecting a wide range of RNA transcripts, which restricts their widespread applications.

One urgent need for the application of scST is to predict the spatial distribution of undetected transcripts. Imputing the undetected transcripts provides an opportunity to create a more comprehensive map of gene expression within tissues, addressing the limitations of current scST technologies related to the number of detectable genes. Another critical challenge in the utilization of scST data is cell-type annotation. Accurate identifying and localizing the cell types is crucial for understanding tissue architecture, function, and disease pathogenesis. Despite the availability of numerous cell-type annotation methods for scRNA-seq data [8–11], their efficacy is limited by the small number of measured genes in scST data.

Improvements in computational techniques for spatially resolved transcriptomics are expected to enable precise prediction of the expression level of the unmeasured genes from other measured genes by integrating the reference scRNA-seq datasets. One class of the existing gene imputation methods is joint embedding-based, including SpaGE [12], gimVI [13], stPlus [14], and so on. Such methods typically first perform joint embedding to establish a shared latent space for both scST and scRNA-seq reference data. Subsequently, for each cell in scST data, imputation of certain genes is obtained by using a weighted average of the expression of corresponding genes in certain cells (usually the k-nearest neighbors) from scRNA-seq data (e.g., SpaGE, stPlus) or through generative methods such as gimVI. Mapping-based methods such as Seurat [15] and Tangram [16] can also be applied to gene imputation. Seurat [15] as a well-known package applies canonical correlation analysis (CCA) to project ST data and scRNA-seq data into a shared latent space, maps the scRNA-seq data to the ST data by the learned archors, and imputes the undetected transcripts of scST with the expression level of corresponding genes in scRNA-seq data. Tangram [16] learns the mapping from scRNA-seq data to scST data using a linear combination of cells in scRNA-seq data to reconstruct the scST data, and imputes the unmeasured genes using the learned probabilistic mapping. Some other methods are also dedicated to this area. ENGEP [17] develops an ensemble learning framework to fuse results of different scRNA-seq datasets and different imputation models, while TISSUE [18] presents a general framework for estimating the uncertainty in spatial gene expression predictions.

Considering the complementary information in scRNA-seq data and scST data, integrating both of them would yield a richer understanding of cellular identities and their interactions within intricate tissue structures. Many methods have focused on transferring the well-annotated cell types from reference scRNA-seq data to ST data. Seurat [15] transfers the labels by mapping the annotated cells from scRNA-seq data to scST data in the shared latent space. Tangram [16] transfers the annotations by the learned linear mapping. scANVI [11] proposes a variant of its former probabilistic generative model scVI [19] by introducing a semi-supervised module that leverages the information of existing annotations. Spatial-ID [20] first generates a series of pseudo labels for scST data and then constructs a GNN-based model, which is trained and supervised with the pseudo labels.

Although several methods have been proposed for missing gene imputation and cell-type annotation in scST data, only a few are capable of performing both tasks, and most demonstrate insufficient performance in either one of the tasks. Integrating both tasks would be advantageous as cells from the same type should exhibit similar gene expression patterns and vice versa. Moreover, the spatial information provided by STs is rich and valuable, yet most existing methods overlook this information in constructing the model. Additionally, previous methods typically adopt an unsupervised framework for imputation, which fails to fully utilize the relationships between different genes to capture the complex patterns of gene expression.

Inspired by the limitations of existing methods, we propose a deep learning-based model called stAI for spatial transcriptomics cell-type Annotation and gene Imputation. This model can both impute missing genes and assign cell-type labels for scST data. It first performs a joint embedding for scST and reference scRNA-seq data with two separate encoder-decoder modules and then incorporates a maximum mean discrepancy (MMD) loss into the shared latent space to align the distribution of the latent codes. Gene imputation is conducted in the latent space by a weighted combination of the neighboring single cells from scRNA-seq in a supervised manner by setting some measured genes as calibration, while cell-type annotation is also performed in the latent space by a non-linear classifier trained with labels of the scRNA-seq data. The model is trained in an end-to-end manner. We conducted experiments for different datasets with varying numbers of measured genes and compared them with the updated methods. The results demonstrate that stAI can give better gene imputation results and cell-type annotations.

Materials and methods

Cell-type annotation and gene imputation method for scST data

Overall framework

stAI employs an encoder-decoder-based backbone framework to effectively capture a shared embedding space for both scST and scRNA-seq data (Fig. 1). Within the shared latent space, each sample from the scST data can identify its closest neighbors in the scRNA-seq data using predefined metrics. Imputation is then performed by generating a weighted linear combination of the corresponding gene expressions from

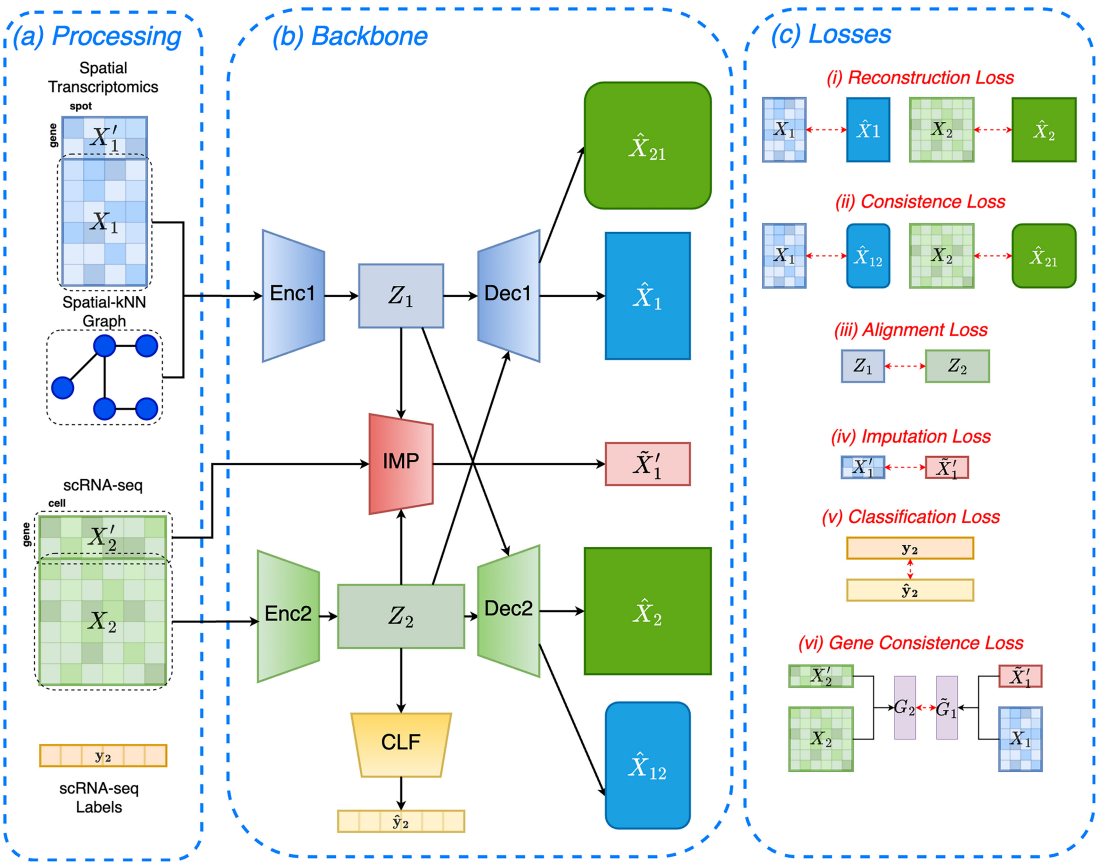


Figure 1. Overview of the model. (**A**) Data processing. Shared genes from scRNA-seq data and scST data are filtered. Spatial-*k*NN graph for scST data is constructed. (**B**) The backbone module of the stAI model. *Enc*1 is a GAT encoder that takes the gene expression profile of training genes as well as the spatial-*k*NN graph as inputs. *Dec*1 is an MLP decoder that attempts to reconstruct the input gene profile. *Enc*2 and *Dec*2 are MLP-based encoders and decoders applied to the scRNA-seq dataset. *CLF* is a classifier trained in a supervised manner on the scRNA-seq dataset. The *IMP* module performs the imputation for the calibration genes. The *IMP* module initially learns a similarity score between two datasets. It then utilizes a weighted sum of measured expression values from calibration genes in scRNA-seq data to impute the corresponding genes in scST data, resulting in the imputed result $\tilde{X}_1$. (**C**) Training losses of the overall model. (i) Reconstruction loss: the MSE between the input gene expression matrices and the reconstructed matrices; (ii) consistence loss: the cosine similarity between the input expression matrices and the expression matrix reconstructed by the other dataset's decoder; (iii) alignment loss: the max-mean discrepancy (MMD) between the latent embedding of two datasets; (iv) imputation loss: the PCC between the measured expression and the imputed expression of calibration genes in stAI. Calibration genes are a subset of measured genes in scST data. During the training process, the model attempts to predict the expression of these calibration genes as they are not measured, and the imputation loss is defined as the PCC between the ground truth expression and predicted expression of calibration genes; (v) classification loss: the cross entropy between the label of scRNA-seq data and the corresponding outputs of the classifier; (vi) gene consistence loss: the differences of the PCCs between training genes and calibration genes in the two datasets.

scRNA-seq data. Additionally, a classifier is trained using the latent embeddings of scRNA-seq data, which not only restricts a more consistent embedding but also transfers the labels in scRNA-seq data to scST data. To train the model in a supervised manner for gene imputation, we split the measured genes into k folds and leave one fold as calibration genes and others as training genes. The forward step only uses training genes while calibration genes are employed in the backward process to train a submodel, i.e. the loss function calculation and gradient backpropagation. Our final model is the ensemble of k submodels by averaging the predicted gene expression from the k submodels.

Model inputs

Denote $X_1 \in \mathbb{R}^{g \times n_1}$, $X_2 \in \mathbb{R}^{g \times n_2}$ the expression matrix of scST and scRNA-seq data of shared genes, where n_1, n_2 denote the number of samples (cells) in scST and scRNA-seq data, respectively and g is the number of shared genes. The nearest neighbor graph is calculated based on the spatial coordinates of scST, denoted by G. Denote the expression of genes to impute in the scST data by $X_1' \in \mathbb{R}^{h \times n_1}$, and the corresponding genes in scRNA-seq data by $X_2' \in \mathbb{R}^{h \times n_2}$, where h is the number of genes to impute. Denote the labels of scRNA-seq data by y_2. The inputs of the forward process include the gene expression matrix of both scST data and scRNA-seq data, i.e. X_1 and X_2, the graph G, and the known expression matrix of genes to impute in scRNA-seq data X_2', as well as cell labels of scRNA-seq data y_2.

Network modules

Encoder-decoder backbone

We utilize two encoder-decoder pairs to project the two datasets into a shared latent space. For the scRNA-seq data, we employ a multi-layer perception (MLP) for both the encoder and decoder. In the case of scST, which includes additional spatial information, we leverage a graph-based network as the encoder to effectively incorporate the spatial context.

Specifically, we employ a graph attention module (GAT) [21] as the encoder for the STs, while using an MLP decoder instead of a graph-based decoder. We consciously opt not to employ a graph-based decoder for two main reasons: first, an MLP-based decoder adequately reconstructs the original expressions; secondly, using an MLP-based decoder ensures compatibility with the decoder for scRNA-seq data, enabling us to compute the consistence loss.

Denote $\mathcal{E}_1$, $\mathcal{E}_2$ the encoders for scST and scRNA-seq data, and $\mathcal{D}_1$, $\mathcal{D}_2$ the decoders. Then the latent embeddings are:

$$Z_1 = \mathcal{E}_1(X_1, G) \in \mathbb{R}^{d\times n_1}, \quad Z_2 = \mathcal{E}_2(X_2) \in \mathbb{R}^{d\times n_2},$$

where d is the dimension of the latent space. The reconstruction results of the decoders are:

$$\hat{X}_1 = \mathcal{D}_1(Z_1) \in \mathbb{R}^{g\times n_1}, \quad \hat{X}_2 = \mathcal{D}_2(Z_2) \in \mathbb{R}^{g\times n_2}.$$

Additionally, stAI also attempts to reconstruct the latent embeddings using coupling decoders, that is, using the decoder of the other dataset:

$$\hat{X}_{1,2} = \mathcal{D}_2(Z_1) \in \mathbb{R}^{g\times n_1}, \quad \hat{X}_{2,1} = \mathcal{D}_1(Z_2) \in \mathbb{R}^{g\times n_2},$$

where $\hat{X}_{1,2}$ is the expression matrix reconstructed by scRNA-seq's decoder from latent embeddings of scST and $\hat{X}_{2,1}$ the opposite. This coupling decoder benefits the batch correction between the two datasets.

Imputation module

After obtaining the latent embeddings of the scST and scRNA-seq data, we proceed to compute the pairwise similarities between each sample in the scST data and that in the scRNA-seq data. For each gene that requires imputation in scST, we use a weighted combination of the gene expression levels from the corresponding gene in the scRNA-seq dataset to generate the imputed value. The weights are determined by normalizing the calculated similarity scores.

Under previous notations, the imputation process is listed as follows:

$$S = NormSim(Z_1, Z_2) \in \mathbb{R}^{n_1\times n_2}, \quad \tilde{X}_1' = X_2' \times S^\top \in \mathbb{R}^{b\times n_1},$$

where *NormSim* can be any user-specified normalized similarity score, and we use the Euclidean distance as default. $\tilde{X}_1'$ represents the calibration genes, which guide the training of S and is essential for the training of Z_1, Z_2. Specifically, S is defined as follows:

$$W_{i,j} = \left\|Z_1[:,i] - Z_2[:,j]\right\|_2, \quad S_{i,j} = \frac{\exp(-W_{i,j})}{\sum_k \exp(-W_{i,k})}.$$

Alternatively, we can also use a learnable similarity score. Inspired from the attention mechanism [22], we parameterize two matrices and use the scaled dot product to represent the similarity matrix: denote $Q \in \mathbb{R}^{d\times d}$ and $K \in \mathbb{R}^{d\times d}$ as two parameterized matrices, which will be updated through the training process, and keep constant during the inference process. Then the weight matrix W is defined as $W = (QZ_1)^\top(KZ_2)$, and S can be computed similarly. In such a case, the imputation process could be considered as a cross-attention module with the value matrix being identity.

Annotation module

Another task is to transfer the labels in scRNA-seq data to scST data. stAI incorporates an additional classifier into the latent space, which is trained using supervision from the scRNA-seq labels. Subsequently, we can employ this classifier on the latent embeddings of scST data to infer their cell types. Denote the estimated classifier by $\hat{C}$, and the annotation results of STs are calculated as:

$$\mathbf{y}_2 = \hat{C}(Z_2), \quad \hat{\mathbf{y}}_1 = \hat{C}(Z_1),$$

where Z_1 and Z_2 are the latent code of scST data and scRNA-seq data. $\hat{\mathbf{y}}_1$ and $\mathbf{y}_2$ denote the estimated labels of scST data and the given labels of scRNA-seq data, respectively.

Loss functions

We design six loss functions to cover all the possible relations in the model. Details of the function definition are put in the supplementary materials.

Reconstruction loss

The reconstruction loss is commonly used in auto-encoders and is calculated as:

$$\mathcal{L}_{\text{recon}} = \text{MSE}(X_1, \hat{X}_1) + \text{MSE}(X_2, \hat{X}_2),$$

where X_1, X_2 are the original data and $\hat{X}_1$, $\hat{X}_2$ are reconstructed data.

Alignment loss

After projecting the two datasets into a latent space separately, we have to align their distributions in the latent space to conduct the subsequent tasks. In the context of domain adaptation and generative modeling, the MMD [23] is a popular loss function used to measure the distance between two probability distributions. We apply the MMD loss between the latent embeddings of two datasets, which is calculated by a mixture of K RBF kernels. We fix $K = 5$ to calculate the MMD loss:

$$\mathcal{L}_{\text{MMD}} = \text{MMD}(Z_1, Z_2).$$

Consistence loss

To keep the cells of the same type from the two datasets consistent in the latent space, inspired by the single-cell integration method Portal [24], we apply another consistence loss, which measures the cosine distance between the original data and the data reconstructed from the coupling decoder:

$$\mathcal{L}_{\text{consist}} = \frac{1}{n_1}\sum_{i=1}^{n_1}\left(1 - \cos(X_1[:,i], \hat{X}_{1,2}[:,i])\right) + \frac{1}{n_2}\sum_{i=1}^{n_2}\left(1 - \cos(X_2[:,i], \hat{X}_{2,1}[:,i])\right),$$

where X_1, X_2 are the original expression matrix of two datasets, $\hat{X}_{1,2}$, $\hat{X}_{2,1}$ are the matrix reconstructed by the decoder of the other dataset. cos is the cosine similarity.

Imputation loss

Given the imputation results on calibration genes, we can evaluate the imputation performance by calculating the imputation score. It could be chosen the same as the evaluation metric or not. The default imputation score is set as the gene-wise Pearson correlation coefficient (PCC), which is defined as

$$\mathcal{L}_{\text{impute}} = \frac{1}{|\text{Calib}|}\sum_{j\in\text{Calib}}\left(1 - \text{corr}(X_1[j,:], \tilde{X}_1'[j,:])\right),$$

where $X_1[j,:]$ is the measured ground truth expression of calibration gene j, $\tilde{X}_1'[j,:]$ is the imputed expression for the

corresponding gene j, 'Calib' is the set of calibration genes, and corr is PCC.

Classification loss

The classification loss is the cross entropy calculated on the annotated scRNA-seq data:

$$\mathcal{L}_{\text{classify}} = -\sum_{i=1}^{n_2}\sum_{c=1}^{C} y_{ic}\log(p_{ic}),$$

where y_{ic} is a binary indicator of whether class c is the correct classification for observation i and p_{ic} is the predicted probability that observation i is of class c.

Gene consistence loss

The expression of different genes is not always independent. We assume that the relationship between genes in scST is similar to that in scRNA-seq data, and then we can align the relationship between the imputed genes and the measured genes in scST with the relationship between the corresponding genes in scRNA-seq data. We first calculate the PCC between genes to impute and the remaining genes both in the scRNA-seq data $A_2 \in \mathbb{R}^{g\times h}$ and in scST data $A_1 \in \mathbb{R}^{g\times h}$:

$$A_2[i,\ j] = \text{corr}(X_2[i,:], X_2[j,:]),$$

$$A_1[i,\ j] = \text{corr}(X_1[i,:], \tilde{X}_1'[j,:]).$$

Then the gene consistence loss is calculated by:

$$\mathcal{L}_{\text{gene}} = \|A_1 - A_2\|_F.$$

The total loss function is a weighted sum of the above losses:

$$\mathcal{L} = \lambda_{\text{recon}}\mathcal{L}_{\text{recon}} + \lambda_{\text{MMD}}\mathcal{L}_{\text{MMD}} + \lambda_{\text{consist}}\mathcal{L}_{\text{consist}}$$
$$+\lambda_{\text{impute}}\mathcal{L}_{\text{impute}} + \lambda_{\text{classify}}\mathcal{L}_{\text{classify}} + \lambda_{\text{gene}}\mathcal{L}_{\text{gene}},$$

and the weights for each loss term are tuned by leaving a validation consisting of a subset of randomly chosen genes and cells.

Training processes

One of the novelties of our model is that we formulate a framework with supervised imputation loss while previous methods are all trained unsupervisedly to the best of our knowledge. We split the measured genes in scST data into k folds, and train k independent submodels. For each submodel, we leave one fold as calibration genes and construct a model to impute this fold using the remaining measured genes, which we call training genes. Hence, during the training process, we can teach the model to impute the calibration genes more accurately and hence stAI can update its parameters in a supervised manner. The final predicted gene expression is the average of those predicted from each of the k submodels.

In the loss function calculation, i.e. the backward process, in addition to the inputs of the forward process, we need two more data to calculate the loss functions: the cell labels in scRNA-seq data, denoted by $\mathbf{y}_2$, and the ground truth expression of calibration genes in scST. The ground truth expression of calibration genes is only available during the training process with the setting of k folds split.

Datasets

We applied the proposed method to three spatial single-cell transcriptomics datasets with annotation paired with one scRNA-seq for each. The first dataset is a cell atlas of the mouse somatosensory cortex generated using osmFISH [5] paired with Zeisel [25]. Only 33 marker genes were selected in the scST dataset. The second dataset is for mouse hypothalamic preoptic region generated using MERFISH [6] paired with Moffit [6]. 153 genes were included in the scST dataset. The third consists of one dataset consisting of brain tissues of a mouse model of Alzheimer's disease at 8 and 13 months, and one of healthy status generated by STARmap PLUS including 2611 genes [26]. The corresponding scRNA-seq dataset was downloaded from GEO. The first two datasets are examples used in SpaGE [12], and the third dataset represents a newer addition to the datasets utilized in this research area.

Data preprocessing

We first filtered the shared genes between the paired datasets. Then for all the scRNA-seq datasets, we normalized each cell by dividing the counts by the total number of transcripts within that cell, scaling by 10^4 and $\log(1 + x)$ transformed. Then the expression with each cell was scaled to zero mean and unit variance. The preprocessing steps were implemented by the Python package 'scanpy'. For scST, we scaled the data into zero mean and unit variance. It has to be noted that for models that do not need normalization, i.e. gimVI for imputation, scANVI for annotation, and Tangram for both tasks, we used the expression before normalized as input, which exhibits a better performance than after normalized. Besides, for the annotation task, we retained the shared cell types between two datasets to benchmark the macro F1 score fairly.

Evaluation

For the imputation task, we performed a cross-validation for each dataset. As the osmFISH and MERFISH datasets have a small number of genes, we used leave-one-out cross-validation to evaluate the performance of the model: for each gene, we dropped its expression in scST and trained the model using the remaining genes to predict it. For the STARmap dataset, which has a total of 2611 shared genes, we performed a 10-fold validation: we attempted to impute the expression level of one of these folds using the model trained from the remaining genes. After deriving the imputation results for each single gene, we compared the imputed result with the ground truth expression of scST. We used PCC, rooted mean squared error (after normalization) (RMSE), and JS-divergence (JSD) as three evaluation metrics to depict the model's performance from different aspects. Furthermore, since we are always more interested in marker genes or high-variable genes than others, we also evaluated the imputation performances of marker genes of each cell type. To further demonstrate the ability of our model to predict unmeasured genes, we also selected several marker genes that were not measured in scST data and visualized their spatial distribution compared to the corresponding cell types. For the annotation task, we filtered the two datasets by retaining shared cell types and evaluated the annotation results on scST by accuracy, macro F1 score, and weighted F1 score against the ground truth. Definitions of these evaluation metrics are put in the supplementary materials.

Benchmarked methods

We adopted several state-of-the-art baseline methods for each task. For the imputation task, we adopted Tangram [16], SpaGE [12], and gimVI [13], which have shown good per-

formances in the previous studies. Annotation models were chosen as scANVI [11], Tangram [16], and Seurat [15].

Results

stAI performs accurate imputation and annotation on osmFISH dataset

We applied stAI to the dataset pair of osmFISH [5] and Zeisel [25] for imputation and annotation tasks. To evaluate the imputation performance, we alternately held one gene as the test and trained the model using all the remaining genes for all four methods including stAI, Tangram [16], SpaGE [12], and gimVI [13]. The results are shown in Fig. 2A–D. Fig. 2A shows box plots of the three benchmark scores PCC, MSE, and JSD for all 33 genes. stAI achieved a median PCC of 0.395, much higher than SpaGE (0.281), Tangram (0.290), and gimVI (0.179). It showed a 36% improvement over the second-best model in terms of median PCC, and also outperformed other methods in terms of MSE and JSD scores. These three evaluation metrics consistently demonstrate that stAI outperforms the compared state-of-the-art methods. Notably, the variance of stAI over different genes is larger, indicating that it can discriminate genes that have the potential for accurate imputation. We also conducted one-sided Mann–Whitney U test to assess the significance of our improvement versus other methods, and the corresponding *P*-values are shown in Fig. 2A. Additionally, the imputation results for each individual gene were compared. Fig. 2B illustrates the PCC between the predicted expression and the ground truth for each gene using stAI compared with other methods, revealing superior performance of stAI over most genes, particularly against SpaGE and gimVI.

Furthermore, to visualize the imputation performance of each method, we plotted the spatial distribution of the imputed gene expression compared with the ground truth. Fig. 2C shows that stAI produced the most consistent spatial expression pattern with that of the measured genes. We selected three genes *Sox10*, *Anln*, and *Ttr* to demonstrate the different performance of the methods. *Sox10* is a gene on which most models performed well and stAI performed the best. All models but gimVI resulted in a consistent imputed expression with the measured. *Anln* is a gene on which stAI did not perform the best according to the PCC. However, the visualization results of stAI are still comparable. Gene *Ttr* had the lowest PCC using all models, but the spatial distribution image of stAI and SpaGE still demonstrates consistency. To verify the ability of stAI to impute the real unmeasured genes, we chose two cell types with distinct spatial patterns: Pyramidal cells and Oligodendrocyte cells. We selected the top two marker genes of each cell type using the scRNA-seq data and imputed these four genes with different methods. The PCC between the imputed expression and the indicator of the corresponding cell type is annotated below each subfigure (Fig. 2D). stAI can predict these markers more consistently with the cell types and achieved a higher PCC. SpaGE also exhibited a consistent spatial distribution though slightly less than stAI, while Tangram and gimVI performed much worse.

We then evaluated the performance of stAI for the annotation task. Compared with three baseline methods scANVI [11], Tangram [16], and Seurat [15], stAI could accurately identify the cell types and achieved higher evaluation scores as shown in Fig. 2E. stAI achieved the highest accuracy of 0.774, compared to 0.747 for scANVI, 0.602 for Tangram, and 0.430 for Seurat. stAI also demonstrated superior performance on macro F1 score and weighted F1 score, which reflects not only the overall performance but also the effectiveness in handling minor classes. We summarized the macro F1 score of each cell type in Fig. 2F, of which stAI has the best performance on all cell types but Interneurons, which consists of 779 cells from a total of 4310 cells. Notably, stAI exhibits significantly better performance on Microglia, the smallest cell type, consisting of 55 cells only. It achieved a precision of 0.36 and a recall of 0.82 for this cell type, while the second-best model scANVI only achieved a precision of 0.16 and a recall of 0.29. Seurat, on the other hand, failed to successfully identify this small cell type.

Then to visualize the annotation results, we plotted the spatial distribution of different cell types in Fig. 2G. A higher consistency between stAI and ground truth is shown in the figure. scANVI only exhibits comparable performance on larger cell types such as Pyramidal. Tangram and Seurat show noisier cell type distributions compared to stAI and scANVI, indicating the effectiveness of neural networks even if the scRNA-seq data comprises only one-third of the cell amount found in STs. The river plot of all four models is shown in Fig. 2H. scANVI misclassified some cells as Pyramidal cells, the largest cell type, while Seurat misclassified many cells as Interneurons. Neither Tangram nor Seurat successfully classified the Pyramidal cells. Compared with the baseline methods, as shown in the river plot, stAI exhibits much better annotation performance.

In addition, we evaluated the performance of stAI in annotating sub-class cells. The two predominant cell types, Pyramidal cells and Oligodendrocyte cells, were chosen for visualizing the annotation results, and the UMAP plot is presented in Supplementary Fig. S1. The annotated sub-class cell types from each main cell type show a high degree of consistency with the ground truth, and the cells belonging to the same sub-class are positioned closely together.

stAI outperforms state-of-the-art methods on imputation by a large margin and also better on annotation on MERFISH dataset

We conducted experiments by leaving one of the 153 genes as the test and training the model using the remaining 152 genes shared in the MERFISH scST dataset and Moffit scRNA-seq dataset to evaluate the imputation performance. The box plots illustrating the performance on all genes are shown in Fig. 3A. stAI outperformed other baseline methods by a large margin on the PCC. It achieved a median PCC of 0.423, compared to 0.308 for SpaGE, 0.249 for Tangram, and 0.170 for gimVI. It surpassed the second-best model SpaGE by over 37% and showed statistically significant improvements over other baseline methods with *P*-values of 1.38e-5 for SpaGE, 3.08e-9 for Tangram, and 1.02e-17 for gimVI. stAI also outperformed other models in terms of MSE and JSD. The scatter plots for the paired PCCs between the measured expression and the predicted of each gene using stAI and the compared methods are shown in Fig. 3B. Of all the 153 genes, stAI outperformed SpaGE on 150 genes except for *Galr2*, *Igf2r*, and *Serpine1*, Tangram on all but seven genes: *Cckar*, *Cspg5*, *Galr2*, *Igf2r*, *Kiss1r*, *Mki67*, *Nup62cl*, and *Serpine1*, and gimVI on all but three genes: *Galr2*, *Igf2r*, and *Nup62cl*.

Since we are always more interested in marker genes or highly variable genes, we evaluated the performance of

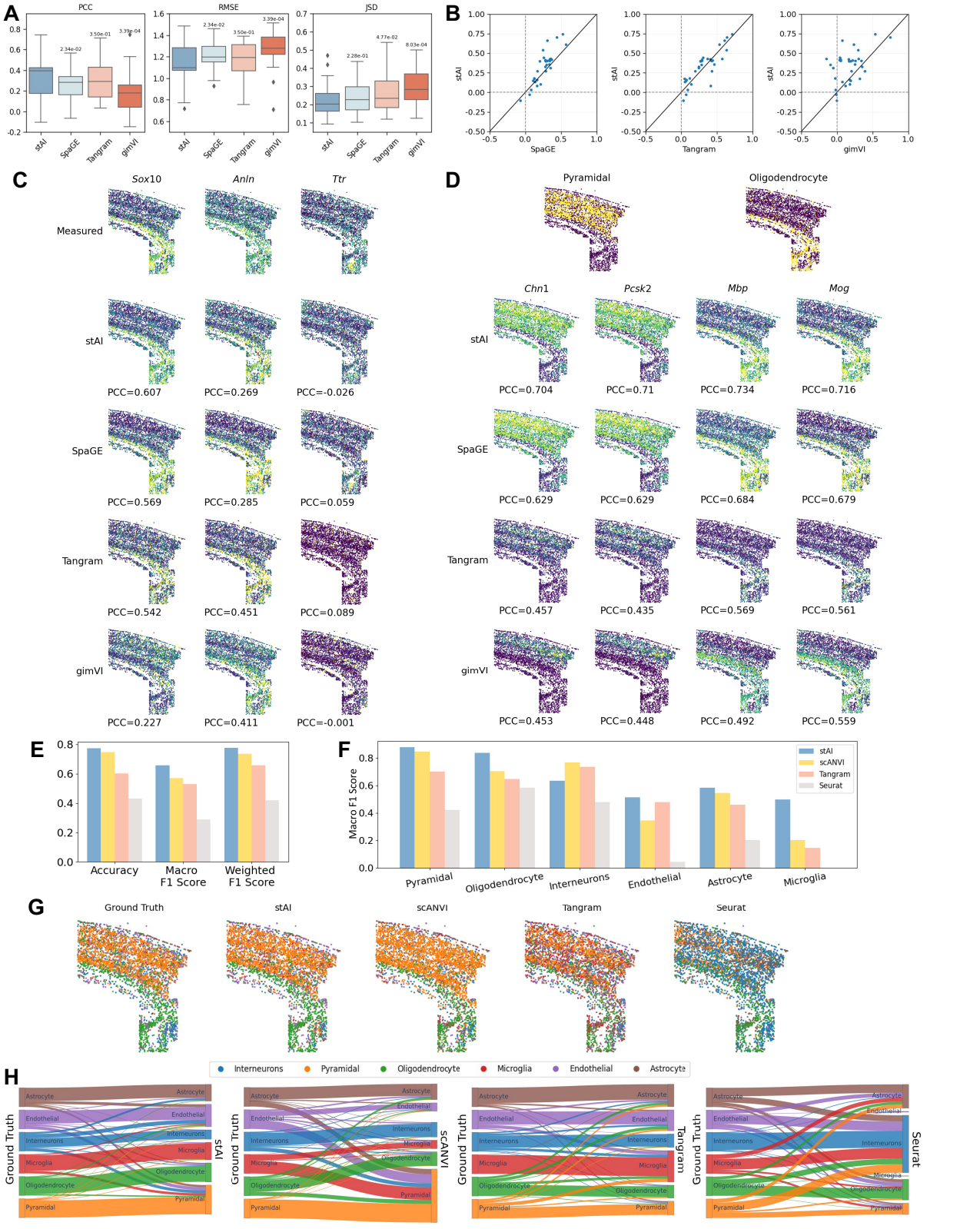


Figure 2. Imputation and annotation results on osmFISH dataset. (**A**) Overall performance measured by PCC, RMSE, and JSD. The annotated *P*-value corresponds to the significance level of the one-side Mann–Whitney U test. (**B**) The scatter plot of paired PCC between stAI and compared method for all single genes. The *x*-axis and *y*-axis represent the PCC of the baseline method and stAI, respectively. Each point lying above the $y = x$ line represents a gene on which stAI outperforms the corresponding baseline. (**C**) Visualization of the predicted expression for example measured genes. The PCC between the imputed expression and the ground truth is annotated. (**D**) Visualization of the predicted expression for unmeasured genes. The PCC between the imputed expression and the indicator of corresponding cell types is annotated. (**E**) The overall performance of annotation results measured by accuracy, macro F1 score, and weighted F1 score. (**F**) Macro F1 score for each cell type. (**G**) Spatial distribution of the original cell types and the annotated cell types. (**H**) River plot for the annotation results.

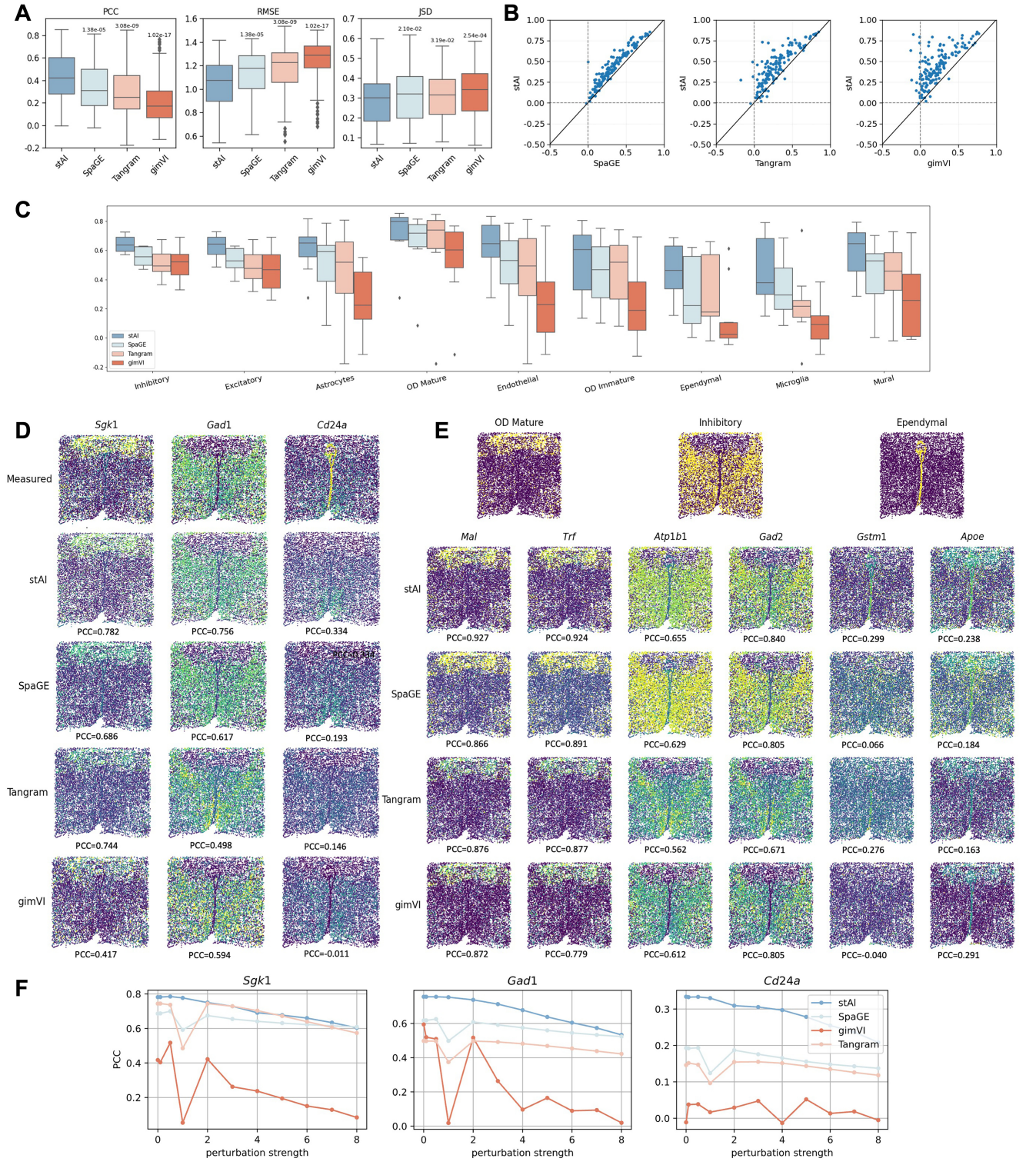


Figure 3. Imputation results on MERFISH dataset. (**A**) Overall performance measured by PCC, rooted MSE (RMSE), and JSD. The annotated *P*-value corresponds to the significance level of the one-side Mann–Whitney U test. (**B**) The scatter plot for paired PCC between stAI and the compared method for all single genes. The *x*-axis and y-axis represent the PCC of the baseline method and stAI, respectively. Each point lying above the $y = x$ line represents stAI outperforms the corresponding baseline in that gene. (**C**) The PCC between the ground truth and the imputed expression for marker genes of each cell type. (**D**) Visualization of the predicted expression for example measured genes. The PCC between the imputed expression and the ground truth is annotated. (**E**) Visualization of the predicted expression for unmeasured genes. The PCC between the imputed expression and the indicator of corresponding cell types is annotated. (**F**) Performance comparison when the scST data are perturbed by different noise levels. The noise follows normal distribution with zero mean and standard deviation being *k* times the mean expression of each corresponding gene, where *k* is the perturbation strength.

imputation on marker genes of different cell types. By filtering the marker genes of each cell type using Python package 'scanpy', we chose the top 10 markers and calculated the PCC of these genes between the given gene expression and the predicted. Across different cell types, stAI consistently exhibited significantly better performance than baseline methods as shown in Fig. 3C. The performance on OD Mature marker genes achieved a PCC of nearly 0.8 with the ground truth. The imputation results on marker genes are generally better than those on other genes, providing us more confidence for downstream analyses on the imputed results generated by stAI.

To visualize the imputation results, we selected several marker genes of OD Mature, Inhibitory, and Ependymal cells. For each of the three selected cell types, we first chose one marker gene that was measured in scST to evaluate the imputed expression against the ground truth as shown in Fig. 3D. stAI performed much better than the baseline methods, especially on *Cd24a*, where the PCC between the imputed value and the ground truth using stAI (0.334) is much higher than the second-best method SpaGE (0.193). Even though all these methods performed well on *Sgk1*, stAI still predicted slightly more consistently with measured expression. We also chose the top two marker genes for each selected cell type that are not measured in scST data and plotted the imputed expression of these genes as well as the spatial distribution of corresponding cell type in Fig. 3E. For the marker genes of Ependymal cells, though the PCC produced by stAI is not much higher, the spatial pattern is much clearer than that produced by other methods. stAI outperformed other baselines both on the visualization and on the PCC between imputed expression and indicator of corresponding cell types.

As during the data generation processes, noise often occurs due to various reasons, we further conducted a perturbation experiment to demonstrate that as long as the perturbation is not destructive, stAI can still reconstruct unmeasured genes accurately. We modeled the noise as the normal distribution with zero mean. To account for the expression level of different genes, the standard deviation was set as k times the mean expression of each corresponding gene, with k representing the perturbation strength. The expression less than zero was set to zero. We considered $k = 0, 0.5, 1, 2, 3, 4, 5, 6, 7$, and 8, where the case $k = 0$ means no perturbation, whereas the case $k = 1$ means perturbing the expression by a normal distribution with standard deviation equal to its average expression level. We also opted for the three representative genes: *Sgk1*, *Gad1*, and *Cd24a*. Fig. 3F shows the comparison results evaluated by PCC. It is clear that stAI outperformed the compared methods. For stAI, the perturbation with a strength less than 1 does not have much influence on the performance and the PCC decreases with a linear trend as the perturbation goes larger. Supplementary Fig. S2 shows the visualization of the imputation with stAI. The imputation performance does not have a significant decrease both from the spatial visualization and the PCC. Even under a strong perturbation, the PCC score is still significant and there also exists a slight boundary similar to measured transcriptomics.

stAI can also transfer the annotation of the Moffit scRNA-seq dataset to the MERFISH dataset accurately. We chose the shared cell types of both datasets. stAI achieved the highest cell-type annotation scores in terms of accuracy, macro F1 score, and weighted F1 score compared with baseline models as shown in Fig. 4A. It excelled not only in overall performance, measured by accuracy, but also in the performance of small cell types, as indicated by the macro F1-score. We then calculated the macro F1 score of each cell type to evaluate the performance as in Fig. 4B. Results indicate that except for the OD Immature cells, stAI achieved the highest macro F1 score. Particularly impressive is the performance of the cell type Ependymal, which accounts for 1961 cells out of 64,373 in total. stAI achieved a macro F1 score of 0.92, while scANVI neglected this cell type despite its presence in both the scST dataset and the scRNA-seq dataset, and the score of other methods is much lower.

We further visualized the annotation results by the spatial distribution of different cell types as well as the UMAP plot in Fig. 4C and D. It is evident that stAI achieved better annotation results than other models. The Ependymal cells consist of less than 1% of scRNA-seq data, and stAI is the only method to successfully distinguish this cell type, even if it exhibited a clear cluster in both the UMAP plot and the spatial distribution plot as shown in Fig. 4C and D. scANVI misclassified this cell type to Astrocytes, while Tangram and Seurat classified part of Ependymal cells to Astrocytes. The capability to capture spatial information of stAI is probably the key feature contributing to its best performance on Ependymal cells and other cell types. The river plot of classification results is shown in Fig. 4E. It is clear that stAI consistently delivered near-perfect classification results, whereas other methods exhibited varying degrees of distortion across certain cell types.

Additionally, we evaluated the annotation performance of all the compared methods when the perturbation is imposed on the scST data using the same settings as those used for gene imputation. stAI consistently delivered the best annotation across all three evaluation criteria as the level of perturbation varied (Fig. 4F). With more noise was added, the annotation accuracy of stAI, Tangram and Seurat decreased, while scANVI showed some fluctuations in its performance.

stAI achieves better imputation score and annotation score on STARmap dataset despite the difficulty of imputation

Compared to the two previous datasets, this scST dataset contains a larger number of measured genes (2611 genes). Similarly, we first evaluated the performance of stAI on gene imputation. We divided the genes into ten folds with each fold serving as the test and training the model using the remaining. Fig. 5A exhibits the overall imputation performance on 2611 genes, evaluated by PCC, rooted MSE, and JSD. Even though all methods have relatively low scores, stAI still outperformed other baseline methods, and the Mann–Whitney U tests indicate a higher level of significance. Fig. 5B gives the scatter plot comparing the PCC of each individual gene between the predicted and the ground truth with baseline methods. stAI demonstrated better performance on most genes, especially when compared with SpaGE and gimVI. We further evaluated the imputation performance on the marker genes in Fig. 5C. The performance of all the compared methods on these marker genes is much better and stAI outperformed other baseline methods.

Although stAI outperformed the compared methods, its performance is not as good as expected from all the evaluation metrics. To explore the reason for this phenomenon, we evaluated the imputation results of stAI with all 2611 genes alternatively as missing. As shown in Supplementary Fig. S3A, the overall PCCs on all genes are much less than that on marker

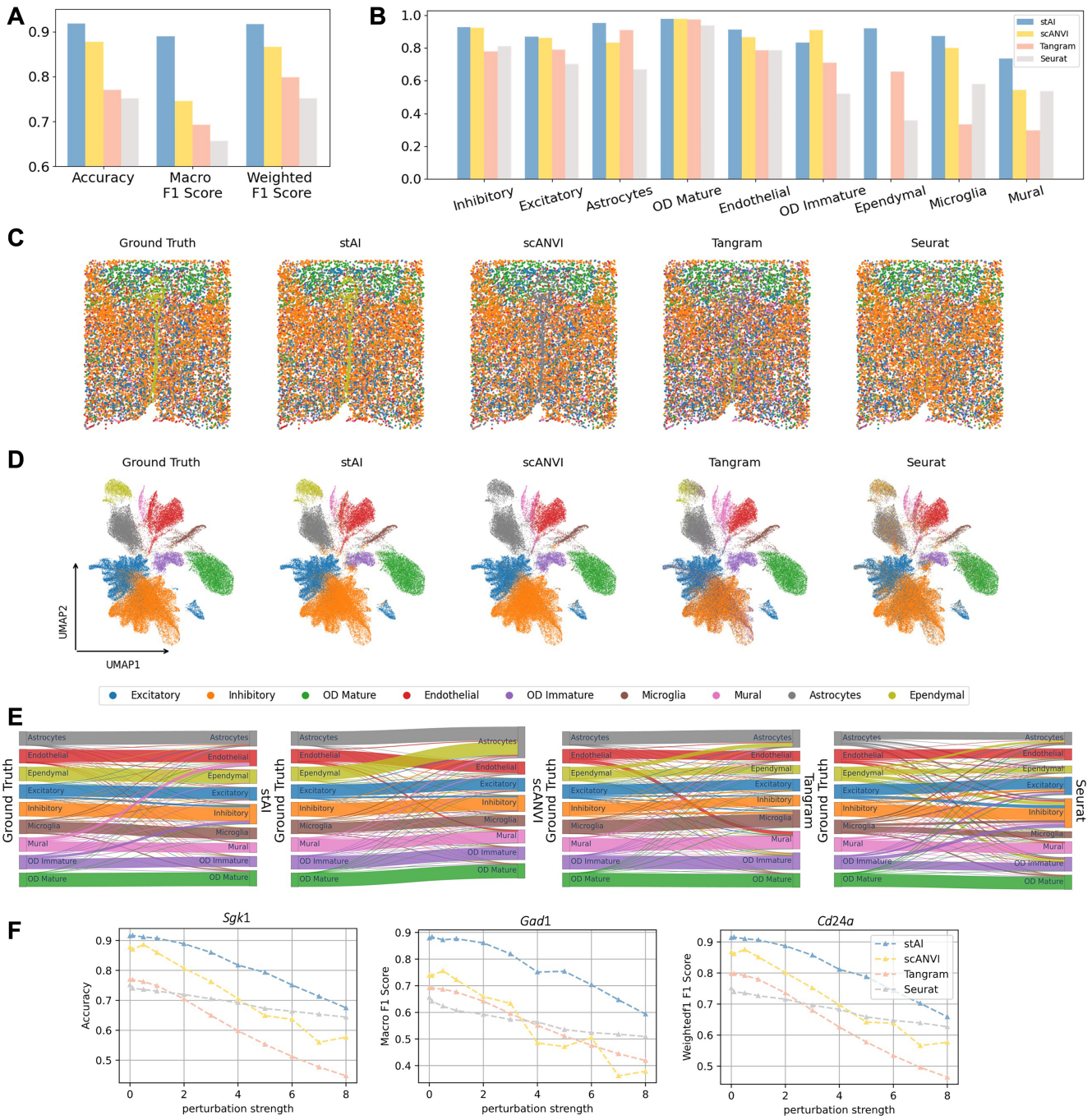


Figure 4. Annotation results on MERFISH dataset. (**A**) The overall performance of annotation results measured by accuracy, macro F1 score, and weighted F1 score. (**B**) Macro F1 score for each cell type. (**C**) Spatial distribution of different cell types for the ground truth and prediction results of different models. (**D**) UMAP plot colored by ground truth cell types and annotation results. (**E**) River plot for the annotation results. (**F**) Performance comparison when the scST data are perturbed by different noise levels. The noise follows normal distribution with zero mean and standard deviation being k times the mean expression of each corresponding gene, where k is the perturbation strength.

genes. This discrepancy may stem from the co-expression patterns of the genes, as those with strong relationships to other genes (such as marker genes) tend to be imputed with better accuracy. Furthermore, we also found that highly variable genes are more likely to be imputed more accurately than others as shown in Supplementary Fig. S3B. The higher the genes are variably expressed, the higher the PCC correlations are.

The annotation results on this dataset also show the superiority of stAI. The overall annotation scores are shown in Fig. 5D, where stAI performed the best, scANVI showed comparable performance and Tangram performed worst. The two reasons that Tangram did not perform well on this dataset could be summarized as: first, all models except Tangram perform data alignment prior to classification, whereas Tangram learns a mapping between unaligned datasets; second, given that the dataset contains the expression of 2611 genes, a linear mapping may be too simplistic to accurately capture the dataset structure. The performance of all methods on each

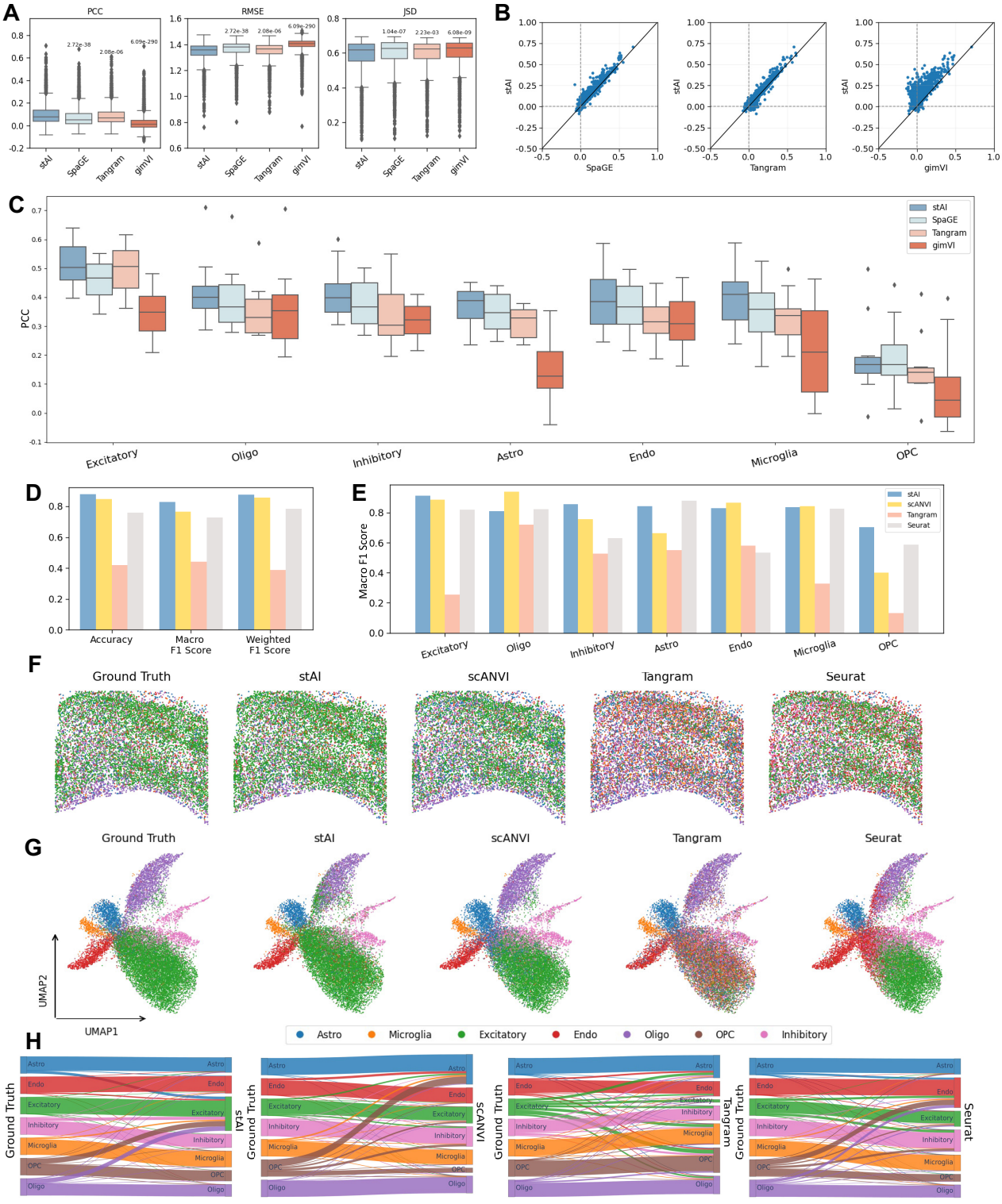


Figure 5. Imputation and annotation results on STARmap dataset. (**A**) Overall performance measured by PCC, rooted MSE (RMSE), and JSD. The annotated *P*-value corresponds to the significance level of the one-side Mann–Whitney U test. (**B**) The scatter plot for paired PCC between stAI and the compared method for all single genes. The *x*-axis and *y*-axis represent the PCC of the baseline method and stAI, respectively. Each point lying above the $y = x$ line represents a gene on which stAI outperforms the corresponding baseline. (**C**) PCC between the ground truth and the imputed expression for all single marker genes of each cell type. (**D**) The overall performance of annotation results measured by accuracy, macro F1 score and weighted F1 score. (**E**) Macro F1 score for each cell type. (**F**) Spatial distribution of different cell types for the ground truth and prediction results of different models. (**G**) UMAP plot colored by ground truth cell types and annotation results. (**H**) River plot for the annotation results.

single cell type measured by the macro F1 score is shown in Fig. 5E. Of all seven cell types, stAI performed the best on Excitatory, Inhibitory, and Oligodendrocyte Precursor Cell (OPC), while scANVI performed the best on Microglia, Endothelial (Endo), and Oligodendrocyte (Oligo), and Seurat performed best on Astrocyte (Astro). The UMAP plot and the spatial distribution of different cell types are shown in Fig. 5F and G. The UMAP plot indicates a relatively clear boundary for each cell type, enabling most methods to achieve good performance on this dataset. stAI and scANVI have comparable performance according to the UMAP plot. stAI misclassified some Oligodendrocyte cells into Excitatory Neuron cells while scANVI misclassified some Excitatory Neuron cells into Inhibitory Neuron cells. Tangram yielded more confused results, especially for Excitatory Neuron cells, while Seurat classified more cells into Endothelial cells. The river plot of classification results is shown in Fig. 5H, from which stAI and scANVI both have better performance compared to other methods.

To investigate the robustness of stAI, we conducted an additional annotation experiment on a mismatched dataset pair: the normal mouse cortex scRNA-seq data as well as the ST data from mice with Alzheimer's disease. We again compared stAI's performance against scANVI, Tangram, and Seurat (Supplementary Fig. S4A). The performance of stAI still surpassed all other algorithms, which demonstrates the robustness and annotation capability of stAI. The visualization highlights the consistency between the predicted and actual cell types (Supplementary Fig. S4B), and the precision and recall for each cell type are also provided (Supplementary Fig. S4C), along with the corresponding results for the matched dataset for comparison. Even though the scRNA-seq data and scST data are generated under different conditions, stAI still delivered reliable annotations.

Discussion

We have demonstrated the ability of the proposed model stAI in both imputating the unmeasured genes and annotating the cell types on scST data with paired scRNA-seq data as a reference. stAI employs an encoder-decoder framework to align the scST data and scRNA-seq data. It employs a *k*NN based method to perform gene imputation in a supervised manner. During the training pipeline, all genes are split into training and calibration sets by cross-validation, and the training process is supervised by the calibration genes. From this pipeline, the pattern of gene expression can be captured more accurately, which can also be extended to other methods. For the annotation step, having integrated the scST data and scRNA-seq data, a classifier trained on scRNA-seq data is directly used to predict the cell type on scST data.

We benchmarked stAI against several state-of-the-art methods for both gene imputation and cell-type annotation. For the imputation task, stAI accurately predicts missing genes unmeasured in scSTs, and results on three datasets exhibit that stAI outperforms other baseline methods significantly. For the annotation task, stAI not only performs best on accuracy metric to measure overall performance but also best on macro F1 score, which demonstrates its superior ability to distinguish small cell types better than other methods.

In this study, we performed experiments on scST data from brain samples generated by three platforms. Since different tissues may exhibit distinct cellular compositions and gene profiles, conducting experiments on diverse tissues is very important, as it enables the validation and refinement of stAI and related methods, ensuring the methods' robustness and applicability across different biological contexts. This will be left as one of our subsequent tasks. With the availability of scST data from a broader range of tissues, we expect significant advancements in development of methods and an improved understanding of tissue-specific biological processes and interactions, as well as insights into the unique spatial organization and functional roles of different cell types.

stAI still has some limitations. For one thing, the *k*NN-based imputation method neglects the heterogeneity between two datasets, leading to the imputed results following the distribution of scRNA-seq data rather than scST data even though the PCC coefficients are much higher. For another, the annotation step does not consider new cell types and cannot discover novel cell types that do not exist in scRNA-seq data. Addressing these issues will be considered as our next tasks.

Acknowledgements

This work is supported in part by the National Key Research and Development Program of China [No. 2021YFC2701601 to S.Z.], Science and Technology Commission of Shanghai Municipality [No. 23JC1401000 to S.Z.], and National Natural Science Foundation of China [No. 12222115 to L.L., No. 12471350 to S.Z.].

Author contributions: G.Z. and S.Z. conceived the project. G.Z. and Q.S. developed and implemented the package. G.Z., Q.S., L.L., and S.Z. validated the methods and wrote the manuscript. All authors read and approved the final manuscript.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

Science and Technology Innovation Plan of Shanghai Science and Technology Commission (Grant Number: 23JC1401000); National Key Research and Development Program of China (Grant Number: 2021YFC2701601), National Natural Science Foundation of China (Grant Number: 12222115, 12471350). Funding to pay the Open Access publication charges for this article was provided by Science and Technology Commission of Shanghai Municipality [No. 23JC1401000].

Data availability

The osmFISH dataset was downloaded in loom format from http://linnarssonlab.org/osmFISH/, and the paired Zeisel dataset (GSE60361) was downloaded from http://linnarssonlab.org/cortex/. The MERFISH dataset was downloaded from the Dryad repository (https://doi.org/10.5061/dryad.8t8s248), and we used the data of the first mouse (Animal_ID == 1). The STARmap PLUS datasets of the mouse with normal

healty status and Alzheimer's disease were downloaded from https://singlecell.broadinstitute.org/single_cell/study/SCP1375/integrative-in-situ-mapping-of-single-cell-transcriptional-states-and-tissue-histopathology-in-an-alzheimer-disease-model#study-download. Its paired scRNA-seq dataset was downloaded from GEO (GSE124952).

Code availability

stAI is implemented with Python and is available at: https://github.com/gszou99/stAI and https://doi.org/10.5281/zenodo.14866342.

References

1. Tang F, Barbacioru C, Wang Y *et al.* mrna-seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009;**6**:377–82. https://doi.org/10.1038/nmeth.1315
2. Ståhl PL, Salmén F, Vickovic S *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 2016;**353**:78–82. https://doi.org/10.1126/science.aaf2403
3. Rao A, Barkley D, França GS *et al.* Exploring tissue architecture using spatial transcriptomics. *Nature* 2021;**596**:211–20. https://doi.org/10.1038/s41586-021-03634-9
4. Eng CHL, Lawson M, Zhu Q *et al.* Transcriptome-scale super-resolved imaging in tissues by rna seqfish+. *Nature* 2019;**568**:235–9. https://doi.org/10.1038/s41586-019-1049-y
5. Codeluppi S, Borm LE, Zeisel A *et al.* Spatial organization of the somatosensory cortex revealed by osmfish. *Nat Methods* 2018;**15**:932–5. https://doi.org/10.1038/s41592-018-0175-z
6. Moffitt JR, Bambah-Mukku D, Eichhorn SW *et al.* Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* 2018;**362**:eaau5324. https://doi.org/10.1126/science.aau5324
7. Wang X, Allen WE, Wright MA *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* 2018;**361**:eaat5691. https://doi.org/10.1126/science.aat5691
8. Aran D, Looney AP, Liu L *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* 2019;**20**:163–72. https://doi.org/10.1038/s41590-018-0276-y
9. Li C, Liu B, Kang B *et al.* Scibet as a portable and fast single cell type identifier. *Nat Commun* 2020;**11**:1818. https://doi.org/10.1038/s41467-020-15523-2
10. Cortal A, Martignetti L, Six E *et al.* Gene signature extraction and cell identity recognition at the single-cell level with cell-id. *Nat Biotechnol* 2021;**39**:1095–102. https://doi.org/10.1038/s41587-021-00896-6
11. Xu C, Lopez R, Mehlman E *et al.* Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Mol Syst Biol* 2021;**17**:e9620. https://doi.org/10.15252/msb.20209620
12. Abdelaal T, Mourragui S, Mahfouz A *et al.* Spage: spatial gene enhancement using scrna-seq. *Nucleic Acids Res* 2020;**48**:e107. https://doi.org/10.1093/nar/gkaa740
13. Lopez R, Nazaret A, Langevin M *et al.* A joint model of unpaired data from scrna-seq and spatial transcriptomics for imputing missing gene expression measurements. arXiv, https://arxiv.org/abs/1905.02269, 6 May 2019, preprint: not peer reviewed.
14. Chen S, Zhang B, Chen X *et al.* stplus: a reference-based method for the accurate enhancement of spatial transcriptomics. *Bioinformatics* 2021;**37**:i299–307. https://doi.org/10.1093/bioinformatics/btab298
15. Stuart T, Butler A, Hoffman P *et al.* Comprehensive integration of single-cell data. *Cell* 2019;**177**:1888–902. https://doi.org/10.1016/j.cell.2019.05.031
16. Biancalani T, Scalia G, Buffoni L *et al.* Deep learning and alignment of spatially resolved single-cell transcriptomes with tangram. *Nat Methods* 2021;**18**:1352–62. https://doi.org/10.1038/s41592-021-01264-7
17. Yang ST, Zhang XF . Engep: advancing spatial transcriptomics with accurate unmeasured gene expression prediction. *Genome Biol* 2023;**24**:293. https://doi.org/10.1186/s13059-023-03139-w
18. Sun ED, Ma R, Navarro Negredo P *et al.* Tissue: uncertainty-calibrated prediction of single-cell spatial transcriptomics improves downstream analyses. *Nat Methods* 2024;**21**:444–54. https://doi.org/10.1038/s41592-024-02184-y
19. Lopez R, Regier J, Cole MB *et al.* Deep generative modeling for single-cell transcriptomics. *Nat Methods* 2018;**15**:1053–8. https://doi.org/10.1038/s41592-018-0229-2
20. Shen R, Liu L, Wu Z *et al.* Spatial-id: a cell typing method for spatially resolved transcriptomics via transfer learning and spatial embedding. *Nat Commun* 2022;**13**:7640. https://doi.org/10.1038/s41467-022-35288-0
21. Veličković P, Cucurull G, Casanova A *et al.* Graph attention networks. *Stat* 2017;**1050**:10–48550.
22. Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. *Neurocomputing* 2021;**452**:48–62. https://doi.org/10.1016/j.neucom.2021.03.091
23. Li CL, Chang WC, Cheng Y *et al.* Mmd gan: Towards deeper understanding of moment matching network. *Adv Neu Inf Proc Syst*. Vol. **30**, 2017.
24. Zhao J, Wang G, Ming J *et al.* Adversarial domain translation networks for integrating large-scale atlas-level single-cell datasets. *Nat Comput Sci* 2022;**2**:317–30. https://doi.org/10.1038/s43588-022-00251-y
25. Zeisel A, Muñoz-Manchado AB, Codeluppi S *et al.* Cell types in the mouse cortex and hippocampus revealed by single-cell rna-seq. *Science* 2015;**347**:1138–42. https://doi.org/10.1126/science.aaa1934
26. Zeng H, Huang J, Zhou H *et al.* Integrative in situ mapping of single-cell transcriptional states and tissue histopathology in a mouse model of alzheimer's disease. *Nat Neurosci* 2023;**26**:430–46. https://doi.org/10.1038/s41593-022-01251-x

Received: June 23, 2024. **Revised:** December 29, 2024. **Editorial Decision:** February 11, 2025. **Accepted:** February 18, 2025