

METHOD

Open Access



Unsupervised spatially embedded deep representation of spatial transcriptomics

Hang Xu^{1†}, Huazhu Fu^{2†}, Yahui Long¹, Kok Siong Ang¹, Raman Sethi¹, Kelvin Chong¹, Mengwei Li¹, Rom Uddamvathanak¹, Hong Kai Lee¹, Jingjing Ling¹, Ao Chen^{3,4}, Ling Shao⁵, Longqi Liu⁶ and Jinmiao Chen^{1,7*}

Abstract

Optimal integration of transcriptomics data and associated spatial information is essential towards fully exploiting spatial transcriptomics to dissect tissue heterogeneity and map out inter-cellular communications. We present SEDR, which uses a deep autoencoder coupled with a masked self-supervised learning mechanism to construct a low-dimensional latent representation of gene expression, which is then simultaneously embedded with the corresponding spatial information through a variational graph autoencoder. SEDR achieved higher clustering performance on manually annotated $10 \times$ Visium datasets and better scalability on high-resolution spatial transcriptomics datasets than existing methods. Additionally, we show SEDR's ability to impute and denoise gene expression (URL: <https://github.com/JinmiaoChenLab/SEDR/>).

Keywords Spatial transcriptomics, Spatial clustering, Variational graph auto-encoder, Batch integration, Trajectory inference, Gene imputation

Background

Single-cell omics technologies enable measurements at single-cell resolution, and this has led to discoveries of new subpopulations across various tissues, in

both healthy and diseased states. However, the dissociation of tissue into single cells prior to high-throughput omics data acquisition leads to cellular spatial information being lost, hindering our ability to dissect the spatial organization and intercellular interactions of individual cells. While computational tools have been developed to predict cell–cell interactions from ligand and receptor expression, they require validation using immunohistochemistry (IHC) or immunofluorescence (IF) experiments. Emerging spatial omics technologies overcome these limitations by retaining the spatial location of gene/protein expression measurements. Such spatially resolved transcriptomes of histological tissues enable the reconstruction of tissue architecture and cell–cell interactions [1–9]. This approach has proven valuable in many applications, including studies on brain disorders [2, 10], tumor microenvironments [3, 11], and embryonic development [12].

Among the currently available spatial transcriptomics approaches, *in situ* capturing-based technologies

*Hang Xu and Huazhu Fu contributed equally to this work.

†Correspondence:

Jinmiao Chen

chen_jinmiao@immunola-star.edu.sg

¹ Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A*STAR), Singapore 138648, Singapore

² Institute of High-Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore 138632, Singapore

³ BGI Research-Southwest, BGI, Chongqing 401329, China

⁴ JFL-BGI STOmics Center, Jinfeng Laboratory, Chongqing 401329, China

⁵ UCAS-Terminus AI Lab, University of Chinese Academy of Sciences, Beijing, China

⁶ BGI-ShenZhen, Shenzhen, 518103, China

⁷ Immunology Translational Research Program, Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore (NUS), 5 Science Drive 2, BlkMD4, Level 3, Singapore 117545, Singapore



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

such as 10 \times Genomics Visium and Nanostring GeoMX DSP are highly popular owing to their accessibility and ability to profile large numbers of mRNA targets within each spot. In principle, a histological section from a tissue sample is permeabilized, and the released mRNA is captured either by spatially arrayed oligos on the slide surface or by pre-hybridized RNA-target barcodes in manually defined regions of interest (ROIs). However, both technologies suffer from mRNA capture spot size limitations, with each spot covering multiple cells. To overcome this, several computational methods have been developed to deconvolve the cell mixtures of spatial spots [13–20]. Recent improvements in mRNA capture methods have led to smaller capture spots that are not greater than 10 μm in diameter. These high-resolution spatial transcriptomics methods can obtain spatially resolved transcriptomes with increased spatial fidelity without compromising the number of genes captured. They include Slide-seq [4, 7], DBiT-seq [8], Stereo-seq [5], PIXEL-seq [6], and Seq-Scope [9]. Some of these methods offer sub-cellular resolution and usually voxel binning or cell segmentation is performed to produce a gene-by-cell expression matrix for downstream analysis. Capture area sizes have also improved and increased the overall cell count throughput, necessitating new computational methods that can handle big spatial data.

Another class of methods relies on fluorescence imaging where DNA probes with attached fluorophores cyclically hybridize to cellular mRNAs, coupled with sequencing-by-ligation or sequencing-by-synthesis techniques to determine the mRNA sequence bases. Early examples of fluorescence *in situ* hybridization (FISH) methods had low gene coverage, but MERFISH [21] and seqFISH+ [22] can achieve high gene coverage with sub-cellular resolution. Commercial vendors like 10 \times Genomics, Vizgen, and Nanostring are debuting their imaging-based methods, Xenium [23], MERSCOPE, and CosMX [24], respectively, which also offer sub-cellular resolution. Similar to other *in situ* sequencing or barcoding-based methods with sub-cellular resolution, data preprocessing is needed to generate the desired gene-by-cell expression matrices.

When analyzing spatial transcriptomics data, combining both gene expression and spatial information to learn a discriminative representation for each cell or spot is crucial. However, established workflows such as Seurat [25] still employ pipelines designed for single-cell RNA-seq (scRNA-seq) analysis, which primarily focuses on the gene expression data and ignores the spatial arrangement of cells. Recently, several methods have been developed for spatial transcriptomics to overcome this limitation. For example, Giotto [26] and

BayesSpace [27] utilize Markov random field models to detect domains with spatially coherent gene expression. stLearn [28], SpaGCN [29], DeepST [30], and STAGATE [31] adopt deep learning approaches to identify spatial domains. The latter three are based on graph neural networks, and they construct neighborhood graphs of cells or spots based on their spatial adjacency. This process can be time-consuming for datasets with large numbers of cells or spots. Some other methods combine gene expression and spatial information into a new feature matrix or latent representation that can be used for clustering and other follow-up analyses. For instance, UTAG [32] calculates an inner product between the spatial adjacency matrix and the expression matrix, returning a new matrix of spatially aggregated expression values. SpatialLDA [33] introduces a spatial regularization term to the latent Dirichlet allocation model to encourage agreement between neighboring cells and outputs spatially aware topics. SpaGene [34] detects spatially variable genes on which non-negative matrix factorization is applied to generate latent factors. Furthermore, another key challenge in spatial transcriptomics analysis is drop-out events and data sparsity, which is not yet taken into account by these existing methods.

In this work, we developed an unsupervised spatially embedded deep representation (SEDR) method for learning a low-dimensional latent representation of gene expression embedded with spatial information. The SEDR model consists of two main components, a deep masked autoencoder network for learning latent representation and a variational graph autoencoder network for embedding spatial information. These two components are optimized jointly to generate a latent representation suited for spatial transcriptomics data analysis. We applied SEDR to 10 \times Genomics Visium, Slide-seq, and Stereo-seq datasets, demonstrating its ability to achieve better representations for various follow-up analysis tasks, namely clustering, visualization, trajectory inference, batch effects correction, and gene expression imputation.

Methods

Model structure

SEDR implements a variational graph autoencoder [35] (VGAE) coupled with a masked self-supervised learning framework to learn a latent representation from gene expression profiles and spatial information. As shown in Fig. 1, the SEDR framework contains two major components, i.e., data masking and latent representation learning. Next, we will elaborate on each component of the framework.

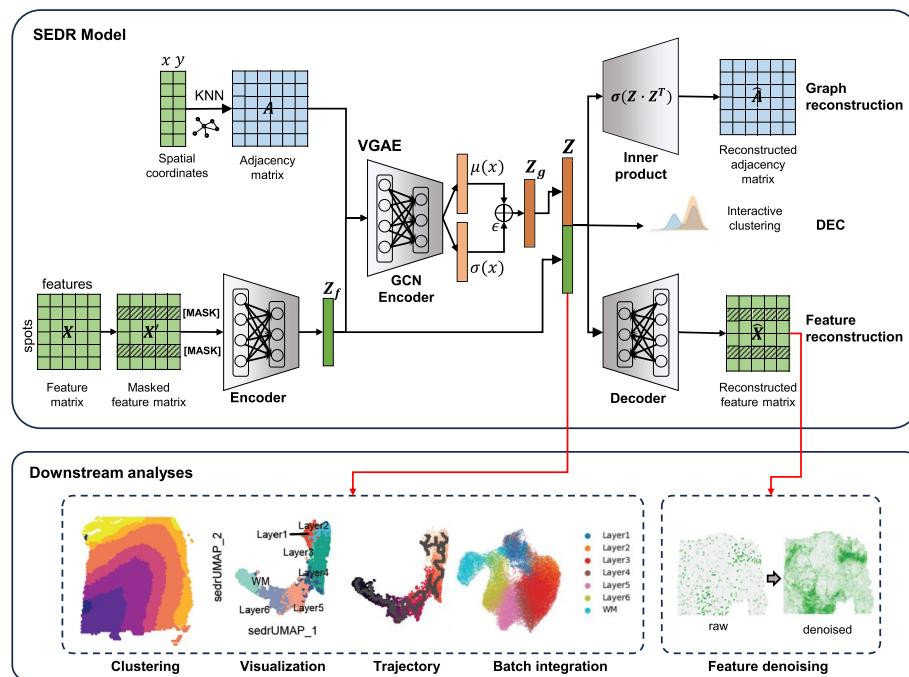


Fig. 1 Overview of SEDR. SEDR learns a low-dimensional latent representation of gene expression embedded with spatial information by jointly training a masked self-supervised deep autoencoder and a variational graph convolutional autoencoder. The low-dimensional embedding produced by SEDR can be used for downstream visualization, spot clustering, trajectory inference, and batch effect correction. The reconstructed feature matrix can be used to impute the raw gene expression with dropouts

Data masking

Before the latent representation learning, we first generate a masked gene expression matrix as input. The masked matrix is then fed into the deep autoencoder and the VGAE, respectively. Specifically, let $X \in \mathbb{R}^{n \times m}$ be a spot-by-gene expression matrix with n spots and m genes. We first randomly sample a subset of spots $SP_{\text{sub}} \subset SP$ and mask each of their gene expression vector $x_i \in \mathbb{R}^m (i \in SP_{\text{sub}})$ with a learnable vector $\tilde{x}_i \in \mathbb{R}^m$. Thus, the input expression matrix X is re-defined as the masked expression matrix X' as follows:

$$x'_i = \begin{cases} \tilde{x}_i, & i \in SP_{\text{sub}} \\ x_i, & i \notin SP_{\text{sub}} \end{cases}$$

One of the main objectives of this masked self-supervised framework is to reconstruct the masked gene expressions of spots in SP_{sub} given the remaining gene expressions X' and spatial adjacency matrix A . Formally, the masked reconstruction loss is denoted as follows:

$$\mathcal{L}_{\text{mask}} = \sum_{i=1}^{|SP_{\text{sub}}|} \|x_i - \tilde{x}_i\|_F^2.$$

Graph construction for spatial transcriptomics data

To create a graph representing the spot-spot spatial relationships in spatial transcriptomics data, we calculate the Euclidean distances between spots using the spatial coordinates. We then use the K -nearest neighbors for each spot to construct an adjacency matrix. The adjacency matrix, denoted by A , is a symmetric matrix, where $A_{ij} = A_{ji} = 1$ if i and j are neighbors, and 0 otherwise. To save CPU memory and running time, the adjacency matrix is stored as a sparse matrix.

Latent representation learning

The latent representation of gene expression is learned using a deep autoencoder. The encoder part consists of two fully connected stacked layers and generates a low-dimensional representation $Z_f \in \mathbb{R}^{n \times d_f}$ from the masked gene expression matrix $X' \in \mathbb{R}^{n \times m}$. Meanwhile, the decoder part with one fully connected layer reconstructs the expression matrix $\hat{X} \in \mathbb{R}^{n \times m}$ from the latent representation $Z \in \mathbb{R}^{n \times d}$, which is obtained by concatenating the low-dimensional representation Z_f and spatial embedding $Z_g \in \mathbb{R}^{n \times d_g}$. Here, d_f, d_g , and d are the dimensions of the low-dimensional expression

representation learned by the encoder, the spatial embedding learned by the graph convolutional neural network (GCN), and the final latent representation of SEDR, respectively, with $d = d_f + d_g$. For the decoder, SEDR has two modes, clustering and gene imputation. For the clustering mode, the decoder is a GCN-based decoder which can capture more spatial information, while a plain linear decoder is used for gene imputation since it helps to avoid over-smoothing caused by a GCN-based decoder. The objective function of the deep autoencoder maximizes the similarity between the input gene and reconstructed expressions measured by the mean squared error (MSE) loss function $\sum (X - \hat{X})^2$.

With the adjacency matrix A and its degree matrix D , the VGAE learns a graph embedding Z_g with the following format: $g : (A, Z_f) \rightarrow Z_g$, where Z_f is the feature representation from the deep autoencoder. The inference part of the VGAE is parameterized by a two-layer GCN [36]:

$$g(Z_g|A, Z_f) = \prod g(z_i|A, Z_f), \text{ with } g(z_i|A, Z_f) = \mathcal{N}(z_i|\mu_i, \text{diag}(\sigma_i^2))$$

where μ is the matrix of mean vectors, and σ is the matrix of standard deviation. μ and $\log(\sigma)$ are obtained by two-layer GCN which is defined as follows:

$$\text{GCN}_\mu(A, Z_f) = \tilde{A} \text{ReLU}(\tilde{A} Z_f W_0) W_\mu,$$

$$\text{GCN}_{\log(\sigma)}(A, Z_f) = \tilde{A} \text{ReLU}(\tilde{A} Z_f W_0) W_\sigma,$$

with weight matrices W_0, W_μ, W_σ , and symmetrically normalized adjacency matrix $\tilde{A} = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$. In computational programming, the distribution of the spatial embedding Z_g cannot be fully depicted. Instead, it is obtained by reparametrization:

$$z_g = \mu + \sigma \odot \epsilon$$

where $\epsilon \sim \text{Normal}(0, 1)$. After obtaining Z_g , merged latent representation Z is obtained by concatenating Z_g and Z_f , and the reconstructed adjacency matrix \hat{A} are generated as follows:

$$\hat{A} = \sigma(Z \cdot Z^T)$$

The objective of the VGAE is to minimize the cross-entropy (CE) loss between the input adjacency matrix A and the reconstructed adjacency matrix \hat{A} , while simultaneously minimizing the Kullback–Leibler (KL) divergence between $g(Z_g|A, Z_f)$ and the Gaussian prior:

$$p(Z_g) = \prod_i \mathcal{N}(z_i|0, I)$$

Batch effect correction for spatial transcriptomics

Spatial relationships only exist within a single spatial transcriptomic measurement; spots from different transcriptomic measurements have no direct spatial relations. Let A^k and Z_f^k denote the adjacency matrix and deep gene representation of spatial omics k , respectively; we then create a block-diagonal adjacency matrix A^k and concatenate the deep gene representation in the spot dimension as follows:

$$A = \begin{bmatrix} A^1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A^K \end{bmatrix} Z_f = \begin{bmatrix} Z_f^1 \\ \vdots \\ Z_f^K \end{bmatrix}$$

where K is the number of spatial omics. Based on this formulation, the different spatial omics datasets (of

potentially different sizes) are transformed into multiple graph instances in the form of one block-diagonal adjacency matrices as inputs to SEDR.

To remove batch effects and enhance the compactness of its latent representation, SEDR employs an unsupervised deep embedded clustering (DEC) method [37] to iteratively group the spots into different clusters. To initialize the cluster centers, we employ the KMeans function in scikit-learn on the learned latent representations. The number of clusters is pre-defined as a hyperparameter. With this initialization, DEC improves the clustering using an unsupervised iterative method of two steps. In the first step, a soft assignment q_{ij} of latent point z_i to cluster center μ_j is calculated using Student's t -distribution:

$$q_{ij} = \frac{\left(1 + \|z_i - \mu_j\|^2\right)^{-1}}{\sum_{j'} \left(1 + \|z_i - \mu_{j'}\|^2\right)^{-1}}$$

In the second step, we iteratively refine the clusters by learning from their high-confidence assignments with the help of an auxiliary target distribution p based on q_{ij} :

$$p_{ij} = \frac{q_{ij}^2 / \sum_i q_{ij}}{\sum_{j'} (q_{ij'}^2 / \sum_i q_{ij'})}$$

Based on the soft assignment q_{ij} and auxiliary target distribution p_{ij} , an objective function is defined using the KL divergence:

$$KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}$$

The SEDR parameters and cluster centers are then simultaneously optimized using stochastic gradient descent (SGD) with momentum.

Clustering

After obtaining latent representation, the results can be used to cluster for the spatial data. SEDR accepted various clustering methods. In this manuscript, mcluster (R package mclust v6.0.0) was selected because the clustering performance is high, and the number of clusters can be determined. We used Python to create a wrapped mcluster function in SEDR code.

Data cohorts

For benchmarking SEDR with other competing methods, we utilized a human dorsolateral prefrontal cortex (DLPFC) dataset described by Kristen et al. [2]. This dataset contains 12 sections, each with 3460–4789 spots, which had been manually annotated into 7 cortical layers, namely layers 1–6 and white matter (WM). To support the ability of SEDR to denoise spatial transcriptomics data, a human ovarian cancer dataset (3493 spots), and a human lymph node dataset (4035 spots) were downloaded from the 10× database. In addition, to demonstrate that SEDR can be applied to high-resolution spatial transcriptomics data, we used two datasets of similar sample size that were generated with Stereo-seq and Slide-seqV2, respectively [5, 7]. The Stereo-seq dataset contains 19,109 spots and 27,106 genes, and the Slide-seqV2 dataset contains 21,724 spots and 21,220 genes. In the case study of tumor heterogeneity, we downloaded a Visium dataset for human breast cancer from the 10× database, which consists of 3798 spots and 36,601 genes.

Results

Overview of SEDR

SEDR learns a gene expression representation in a low-dimensional latent space with jointly embedded spatial information (Fig. 1). Given a set of spatial transcriptomics data, SEDR first learns a non-linear mapping from the gene expression space to a low-dimensional feature space using a deep autoencoder network. A masked self-supervised learning mechanism is used to enforce the encoder to capture more gene expression information through pretext tasks. Simultaneously, a variational graph autoencoder is utilized to aggregate the gene representation with the corresponding spatial neighborhood relationships to produce a spatial embedding. Next, the gene representation and spatial embedding are concatenated to

form the final latent representation used to reconstruct the gene expression. Thereafter, an unsupervised deep clustering method [37] is employed to enhance the compactness of the learned latent representation. This iterative deep clustering generates a form of soft clustering that assigns cluster-specific probabilities to each spot, leveraging the inferences between cluster-specific and spot-specific representation learning. Finally, the learned latent representation can be applied towards various downstream analysis tasks, including clustering, data visualization, trajectory inference, and batch integration. Using the scaled expression as input, the reconstructed matrix can be used to impute and denoise the raw data with dropouts.

Quantitative assessment of SEDR on human dorsolateral prefrontal cortex (DLPFC) dataset

To perform a quantitative comparison of SEDR with competing methods, we downloaded the LIBD human DLPFC data with manual annotation, acquired using the 10× Genomics Visium spatial transcriptomics platform [2]. We chose this dataset because the human DLPFC has clear and established morphological boundaries which can serve as the ground truth [38]. We first applied the standard Seurat pipeline [25] to process and cluster spots using only gene expression profiles and set the result as the benchmarking baseline to investigate the extent to which spatial information improves spot clustering. Next, current methods that integrate gene expression with associated spatial information were used to benchmark SEDR, namely SpatialLDA [33], Giotto [26], stLearn [28], SpaGene [34], SpaGCN [29], BayesSpace [27], UTAG [32], DeepST [30], and STAGATE [31] (Additional file 1: Supplementary methods). The computational features of those methods are summarized in Table 1.

First, we considered one cortex slice for illustration, #151,673 (Fig. 2A) with 3639 spots and 33,538 genes. We found that SEDR achieved the best performance in terms of both layer borders and adjusted rand index (ARI), followed by STAGATE (Fig. 2A). When comparing the results on all 12 DLPFC samples, we employed 6 quantitative measures, ARI, adjusted mutual information (AMI), purity score, homogeneity, completeness, and v measure. These quantitative measures reflect the goodness of matching between ground truth and predictions from different aspects (Additional file 1: Supplementary methods). For all 6 metrics, SEDR had statistically significant higher scores than competing methods, except for DeepST and STAGATE (Mann–Whitney *U* test, *p*-value < 0.05, Fig. 2B). Compared to DeepST and STAGATE, the median scores were still higher for SEDR though the difference was not statistically significant. In addition, the Silhouette score was used to measure

Table 1 Summary of features of the methods for detecting spatial domains. Compared to other methods, SEDR allows the implementation of more types of data and provides more information for downstream analyses, including latent representation and de-noised feature values. In addition, it uses GPU to accelerate calculations

Methods	Model	Resolution	Latent representation	De-noising	Batch integration	Programming	GPU
Seurat	Principal component analysis	Spot or single cell	✓	✗	✓	R	✗
SpatialLDA	Latent Dirichlet allocation	Single cell	✓	✗	✗	Python	✗
Giotto (HMRF)	Hidden Markov random field	Spot or single cell	✗	✗	✗	R	✗
stLearn	Spaital morphological gene expression normalization	Spot or single cell	✓	✗	✓	Python	✓
SpaGene	Spatial network (KNN)	Single cell	✓	✗	✗	R	✗
SpaGCN	Graph convolutional network	Spot or single cell	✗	✗	✓	Python	✗
BayesSpace	Bayesian model with a Markov random field	Spot	✗	✗	✗	R	✗
DeepST	Variational graph autoencoder	Spot or single cell	✓	✗	✓	Python	✓
STAGATE	Graph attention autoencoder	Spot or single cell	✓	✗	✓	Python	✓
UTAG	Graph+clustering	Single cell	✗	✗	✓	Python	✗
SEDR	Variational graph autoencoder+masked self-supervised	Spot or single cell	✓	✓	✓	Python	✓

the matching between the latent representation and the ground truth labels (Additional file 1: Supplementary methods). The Silhouette score for SEDR was the highest and only not statistically significant with respect to DeepST (Additional file 1: Fig. S1). The lack of statistical significance when compared to DeepST can be attributed to the high variance in DeepST's Silhouette scores.

As the baseline method, the Seurat pipeline processes spatial transcriptomic data without considering the spatial information, obtaining the poorest results. Among the spatially informed methods, methods that use graph convolutional frameworks (GCN) including SpaGCN, DeepST, STAGATE, and SEDR, generally achieved better performance than others, indicating the superior ability of GCNs to integrate spatial information. When it comes to the usage of histological information, stLearn, SpaGCN, and DeepST employed the H&E image of the spots, but their performance was not outstanding, suggesting the limited utility of image information, which is potentially due to the low quality of the images. Additionally, the image processing step is inefficient in terms of running time and memory usage, especially for high-resolution images. This can make such methods unsuitable for larger spatial transcriptomics data as such data are growing in availability and scale.

To explore the robustness of SEDR and competing methods with respect to hyperparameters, we tested the top 3 methods, namely SEDR, STAGATE, and DeepST (Fig. 2) with a varying number of nearest neighbors (K), which determines the level of local spatial smoothing. There was a slight fluctuation of ARI scores for all 3 methods on the 12 DLPFC sections with different K values (Additional file 1: Fig. S2). But for each K , SEDR

achieved the best ARI score over the other methods. Furthermore, even the lowest median ARI score (0.532 when $K=14$) of SEDR was higher than the best score of STAGATE (0.50 when $k=18$).

We next tested SEDR's low-dimensional representation features in trajectory inference [39]. Monocle3 [40] was employed to perform trajectory inference on the same DLPFC slice (#151,673) with the low-dimensional representations from Seurat, stLearn, DeepST, STAGATE, and SEDR (Additional file 1: Supplementary methods). In the UMAP plot, the RNA-only representation (Seurat) clearly segregated the WM as a separate cluster from the other layers and the cortical layers showed mixing with no clear ordering according to their developmental trajectory (Fig. 2C, top). In contrast, the spatially informed methods' latent embeddings mostly recapitulated the ordering from the WM to layer 6 and sequentially to layer 1. Setting WM as the root, the pseudo-time of trajectory was computed with Monocle. With DeepST and SEDR's latent representations, Monocle could accurately infer the "inside out" developmental trajectory. By plotting trajectories on the original spatial positions, we observed a strong correlation between the ordering of pseudotime and the physical location of cortical layers (Fig. 2C, middle and bottom). Comparatively, the trajectories from stLearn and STAGATE's embeddings, while improving over the RNA-only inference, still showed incorrect layer locations based on the inferred orderings. This demonstrated that incorporating spatial information into pseudo-time trajectory inference can be beneficial and SEDR's latent representation is well-suited for this application [38, 41, 42].

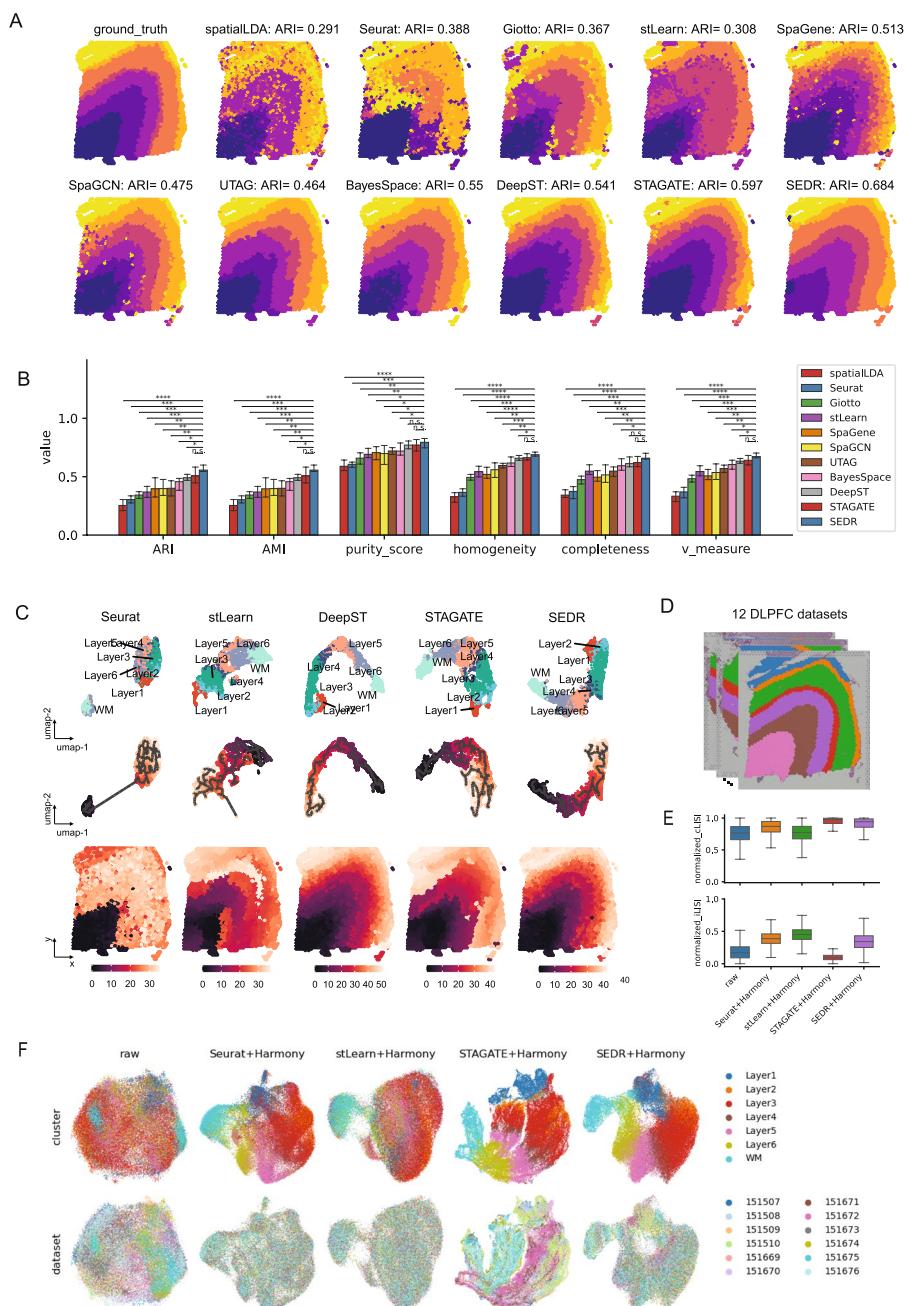


Fig. 2 Quantitative assessment of SEDR on the human dorsolateral prefrontal cortex (DLPFC) dataset. **A** Manual annotation for the DLPFC #151673 section and clustering results of eleven methods (SpatialLDA, Seurat, Giotto, stLearn, SpaGene, SpaGCN, BayesSpace, UTAG, STAGATE, DeepST, and SEDR). **B** Barplot for the 6 performance metrics (ARI, AMI, purity score, homogeneity, completeness, v measure) on the clustering results of the DLPFC 12 sections. Notation for statistical significance testing: $n.s$ p -value > 0.05 , * p -value < 0.05 , ** p -value < 0.005 , *** p -value < 0.0005 , **** p -value < 0.00005 . **C** UMAP and spatial visualization of Monocle 3 pseudo-time trajectories inferred with the latent representation by the tested methods of DLPFC slice #151673. UMAP plots with ground-truth labels (above), UMAP plots overlaid with Monocle 3 pseudo-time trajectories (middle), and Monocle 3 pseudo-time ordering on spatial coordinates (bottom). **D** The 12 human DLPFC sections with manual annotation. **E** Normalized cLISI and iLISI scores for DLPFC section integration results using the latent representations of four methods (Seurat, Harmony, STAGATE, SEDR). **F** UMAPs of integration results for four methods

SEDR corrects for batch effects

The proliferation of spatial omics applications is generating ever-increasing volumes of spatially resolved omics data across different labs. However, differences in protocols and technologies complicate comparisons and data integration when trying to achieve consensus on spatially resolved tissue atlases. As with scRNA-Seq, removing batch effects in spatial omics datasets is a significant challenge. The deep embedded clustering (DEC) loss function employed in the SEDR model helps retain biological variations while reducing technical variations. Here, we tested the use of joint embeddings across multiple batches and projected them into a shared latent space.

We evaluated three methods, SEDR, stLearn, and STAGATE, using their latent representations of the DLPFC dataset in batch integration (Fig. 2D, E). To establish a baseline, we used Seurat to calculate the PCA embeddings for raw data (Fig. 2F). For batch integration, we employed Harmony due to its superior performance in scRNA-seq data integration [43]. With Seurat's PCA embeddings, Harmony mixed the batches evenly but also mixed the cortical layers together (e.g., layers 1 and 2, Seurat UMAP, Fig. 2F). With stLearn's embeddings, the cortical layers were also well-mixed. For STAGATE and SEDR's embeddings, the different cortical layers' cells showed clear separation with the developmental ordering visible. However, STAGATE's integrated output showed the batches to be integrated into two major clusters. In contrast, the batch mixing was much more even in SEDR's output. We also assessed the batch integration results with the integration LISI (iLISI) and cell type LISI (cLISI) metrics for all 12 slices (Fig. 2E, Additional file 1: Supplementary methods). In terms of cLISI, stLearn was the poorest with a score similar to the uncorrected data, and this corroborated the visual inspection of the UMAP. In terms of iLISI, the SEDR-derived integration was third (Mann–Whitney U test <0.05), again matching the observations made on the plotted UMAPs. Taking both cell type separation and batch integration into consideration, we consider SEDR to be the overall best-performing method.

The integration results also suggested that the inclusion of spatial information within embeddings can have variable results. In the case of SEDR's embeddings, the spatial information was integrated such that it improved the cortical layer cell label separation over the typical PCA embedding. On the other hand, it may have interfered with the batch integration as in STAGATE's case. Overall, this example demonstrated that the combination of SEDR and Harmony can be effective for batch integration of spatial transcriptomics data.

Effective noise removal and imputation of spatial transcriptomics with SEDR

Spatial transcriptomics offers unprecedented opportunities in dissecting tissue heterogeneity but suffers from measurement noise including dropouts. Consequently, effective imputation with noise removal can help reveal spatially resolved features within the data. With masked self-supervised learning, SEDR can construct a denoised and imputed gene expression matrix, which cannot be achieved by other competing methods for learning latent representations. Instead of using principal components (PCs) as input for clustering tasks, SEDR uses the processed gene expression matrix as input and generates the reconstructed expression matrix with the decoder module.

On the $10\times$ Genomics Visium human ovarian cancer dataset (downloaded from the $10\times$ website) [44] (Fig. 3A, B), the PTPRC (CD45) gene expression showed poor correlation (0.193) with the protein expression from immunofluorescent (IF) staining. This is primarily explained by the high levels of dropouts for PTPRC expression in the data. Using SEDR, we successfully imputed the PTPRC expression to attain a higher correlation of 0.499. Visually, we can observe the imputed region in the top left which was captured by IF but not found in the original data, indicating that such denoising is not simply spatial smoothing.

To further investigate SEDR's denoising and imputation capabilities, we next applied SEDR to a healthy human lymph node dataset (Fig. 3C) [44]. The lymph node contains substructures called germinal centers (GCs), which can be identified through H&E staining. We first outlined three GCs and plotted the denoised values of three gene markers for substructures within the GC. As shown in Fig. 3D, the denoised BCL6, FCER2, and EGR1 expressions better delineated the germinal center structures compared to the raw expressions. The gene BCL6 marks the mature B cells found in the center of the GC, while FCER2 is a marker for naïve B cells located at the marginal zone of the follicle containing the GC, and EGR1 marks activated B cells found outside the follicles [45]. We also assessed the gene–gene correlations before and after the denoising. IGHD is known to positively correlate with MS4A1 and CD1C and negatively with CD3D. We found SEDR's denoised values to show stronger expected gene–gene correlations than the raw and the Sprod [44] denoised values (Fig. 3E, F).

SEDR can handle high-resolution spatial transcriptomics

Newly emerging methods such as Stereo-seq [5], PIXEL-Seq [6], and Seq-Scope [9] can achieve sub-micrometer and thus subcellular resolution. With

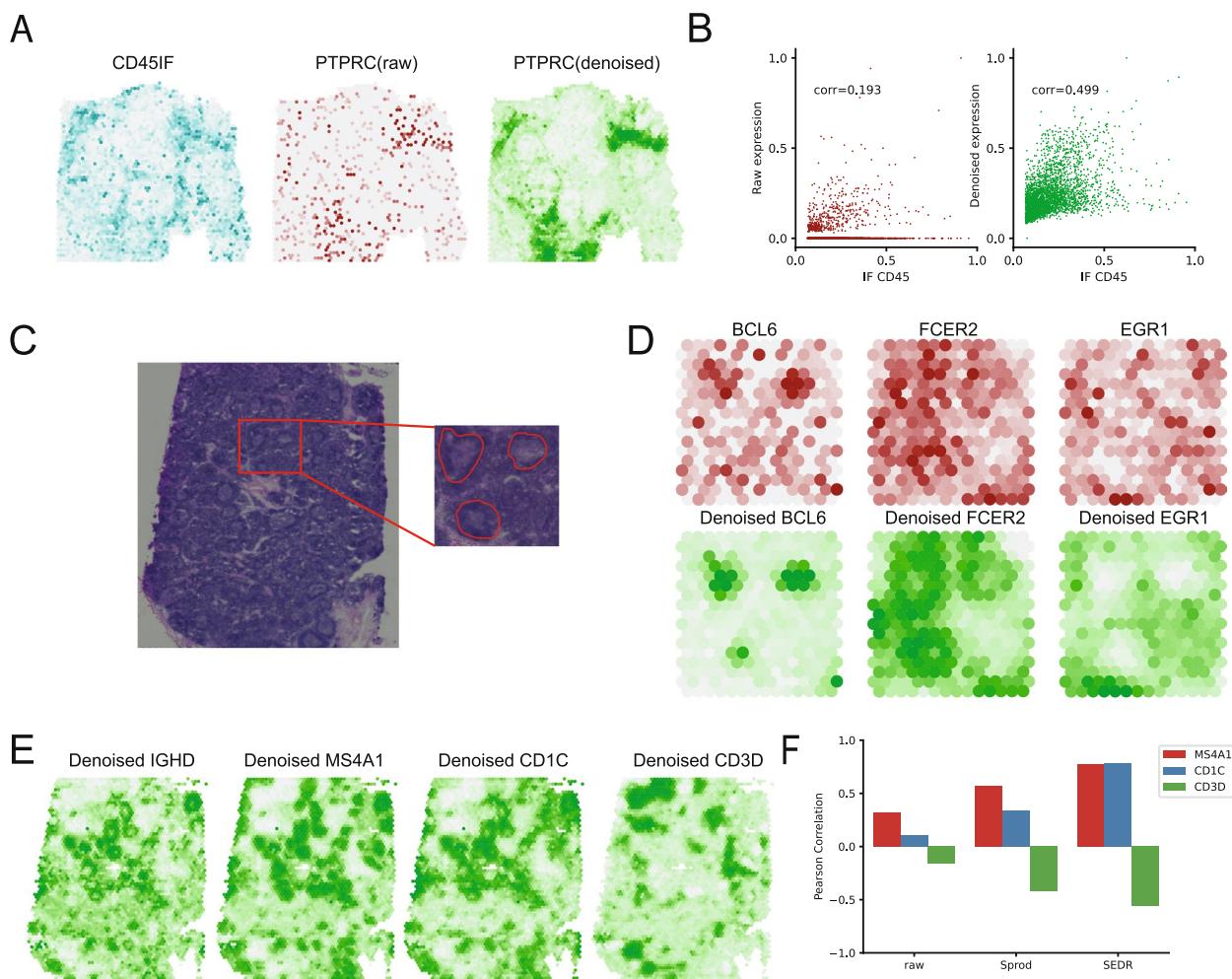


Fig. 3 Denoised gene expression with SEDR. **A** Immunofluorescence (IF) values, raw RNA-seq counts, and denoised gene expression by SEDR for PTPRC (CD45) on the ovarian cancer dataset. **B** Scatter plot of scaled raw counts and de-noised PTPRC expression by SEDR. **C** H&E image with identified germinal centers (GCs) in the lymph node tissue sample. **D** Denoised expression of three marker genes within the selected GCs. BCL6, FCER2, and EGR1 are the markers for different regions of GCs. **E** Denoised expression of IGHD and three genes that are correlated with IGHD: MS4A1 and CD1C are positively correlated with IGHD while CD3D is negatively correlated with IGHD. **F** Pearson correlation between IGHD and the correlated genes (MS4A1, CD1C, CD3D) as measured by raw counts, de-noised value by Sprod, and de-noised value by SEDR

continued technology advancement, the spatial resolution and number of cells detected per tissue will significantly improve, producing large datasets with high throughput. As such, we evaluated SEDR's performance on Stereo-seq and Slide-seq data of mouse olfactory bulb tissues. We first consider the Stereo-seq data with 19,109 spots and 27,106 genes. The coronal section of a mouse olfactory bulb can be divided into 7 layers, the olfactory nerve layer (ONL), glomerular layer (GL), external plexiform layer (EPL), mitral cell layer (MCL), internal plexiform layer (IPL), granule cell layer (GCL), and rostral migratory stream (RMS) (Fig. 4A). We performed dimension reduction and unsupervised clustering using 10 methods to computationally reconstruct

the spatial distribution of tissues within the olfactory bulb. With this data, stLearn, SpaGCN, UTAG, DeepST, STAGATE, and SEDR all produced distinct clusters that match the layers in the olfactory bulb, consistent with the annotated staining. (Fig. 4A, B). We further plotted the individual clusters per method to visualize their differences (Fig. 4C). For stLearn, SpaGCN, and STAGATE, they were unable to fully remove the technical noise from Stereo-seq seen near the core of the olfactory bulb, as visible in their cluster 9. Overall, we found DeepST, STAGATE, and SEDR to be competitive with the most similar clusters which also well matched the markers of the different tissue domains within the olfactory bulb (Fig. 4D). Namely,

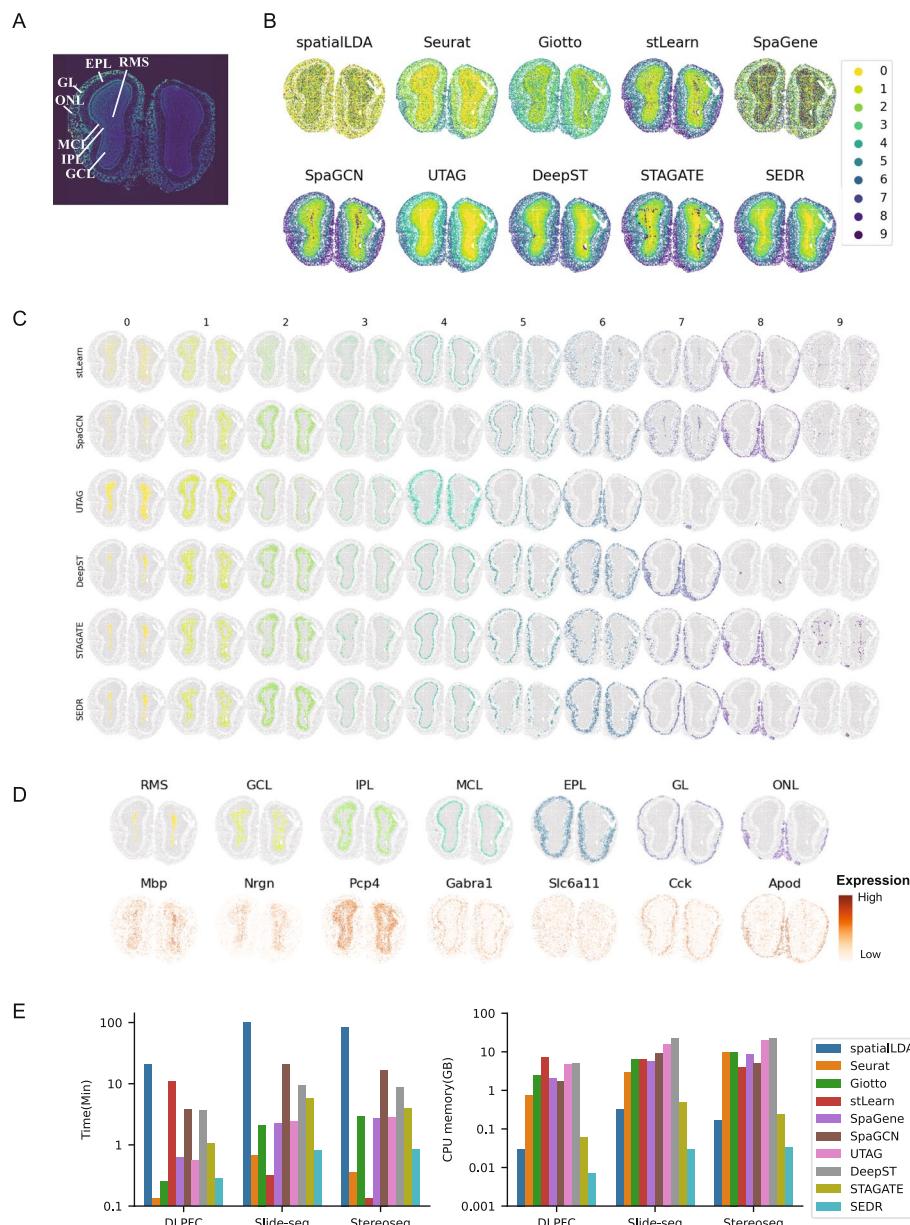


Fig. 4 Application of SEDR on Stereo-seq dataset. **A** Laminar organization of a DAPI-stained mouse olfactory bulb. **B** Unsupervised clustering results from SEDR and nine competing methods on the olfactory bulb Stereo-seq data. **C** Visualization of individual identified clusters for selected methods (stLearn, SpaGCN, UTAG, DeepST, STAGATE, and SEDR). **D** Predicted olfactory bulb layers and corresponding marker genes. **E** Computational efficiency of ten methods tested on slice #151673 of the DLPFC set, and Slide-seq data and Stereo-seq data of mouse olfactory bulb. Left: run time of methods. Right: CPU memory usage of methods

cluster 0 corresponded to the RMS, cluster 1 to GCL, cluster 2 to IPL, clusters 3 and 4 to MCL, clusters 5 and 6 to EPL, cluster 7 to GL, and cluster 8 to ONL. We also applied the methods to the Slide-seq data acquired from the mouse olfactory bulb with 21,724 spots and 21,220 genes. The clusters obtained are provided in Additional file 1: Fig. S3. Similar to the Stereo-seq data,

stLearn, SpaGCN, UTAG, DeepST, STAGATE, and SEDR showed good performance.

With the increasing sample size of high-resolution spatial transcriptomics data, the scalability of processing methods becomes correspondingly more and more important. Here, we tested the run time and memory usage of the ten methods. As shown in Fig. 4E, SEDR

required less time than other methods except for Seurat and stLearn. It is reasonable for Seurat to require less run time as the Seurat pipeline only performed the baseline PCA dimension reduction of the gene expression matrix and ignored spatial information. For stLearn, it required much more time on the $10 \times$ Visium DLPFC but less time on Slide-seq and Stereo-seq data, because its pipeline for $10 \times$ Genomics Visium data contains a time-consuming image processing step, while the pipeline for the other two datasets does not contain such step. For CPU memory usage, SEDR required less CPU memory than all other methods, which can be explained by the use of a sparse matrix in its graph construction and computational steps involved in graph multiplication, greatly reducing the CPU memory requirement. We also compared SEDR with DeepST and STAGATE for GPU memory usage. SEDR also required less GPU memory than the other two methods (Additional file 1: Fig. S4A). To further demonstrate SEDR's capacity in handling large datasets, we created a series of augmented datasets based on the mouse olfactory bulb Slide-seq data, which contained 1/2, 1, 2, and 4 times the number of spots compared to the original data (20,000 spots). The results demonstrated that the cost of time, CPU, and GPU memory were still acceptable even in the case of $4 \times$ data size (80,000 spots) (Additional file 1: Fig. S4B).

Dissecting tumor heterogeneity and immune microenvironments using SEDR

Intratumoral heterogeneity in cancer complicates effective treatment formulations and is associated with poor survival prospects [46]. Spatial transcriptomics offers advantages over scRNA-seq at dissecting and characterizing intratumoral heterogeneity and tumor-immune crosstalk by retaining the spatial information that can help reconstruct spatially distributed domains and distance-dependent interactions. In this example, we tested SEDR on the $10 \times$ Genomics Visium spatial transcriptomics acquired data for human breast cancer, which is known for its high intratumoral and intertumoral differences [47]. To aid in interpreting the SEDR's results, we performed manual pathology labeling based on the H&E staining. It should be noted that, unlike the cerebral cortex which has clear and established morphological boundaries, tumor tissues are highly heterogeneous and encompass complex microenvironments; manual labeling solely based on tumor morphology is inadequate for characterizing such complexity. Based on the pathological features, we first manually segmented the histology image into 20 regions, which we then grouped into 4 main morphotypes: ductal carcinoma *in situ*/lobular carcinoma *in situ* (DCIS/LCIS), healthy tissue (Healthy), invasive ductal carcinoma (IDC), and tumor

surrounding regions with low features of malignancy (Tumor edge) (Fig. 5A, Additional file 1: Fig. S5A).

To detect spatial domains, eleven methods, namely SpatialLDA, Seurat, Giotto, stLearn, SpaGene, SpaGCN, UTAG, BayesSpace, DeepST, STAGATE, and SEDR, were used to produce the same number of clusters (20). Visually, all methods agreed with the manual annotations at the macroscopic level (Fig. 5A). Giotto's clusters showed the highest levels of fragmentation, more than Seurat's, a non-spatially aware method. Notably, SEDR, UTAG, and DeepST divided the tumor region DCIS/LCIS_3 into an outer "ring" and a tumor core, while BayesSpace, SpaGene, and STAGATE divided the region into two halves.

To explore the cell types in SEDR cluster 11 (tumor core) and cluster 13 (tumor edge) in SEDR clustering results (Fig. 5B), Seurat 3 was employed to deconvolute spatial data against the scRNA-seq reference data for human breast [48] (Additional file 1: Supplementary methods). For each spot, we obtained a composition vector for cell types (Additional file 1: Fig. S5B). Interestingly, we found that cluster 13 contains significantly more tumor-associated macrophages (TAMs) than cluster 11 (Fig. 5C). TAM infiltration is known to be strongly associated with poor survival rate in solid tumor patients due to its promotion of tumor angiogenesis and induction of tumor migration, invasion, and metastasis [49, 50].

We also performed differential expression analysis followed by pathway enrichment analysis (Fig. 5D, Additional file 1: Supplementary methods). In cluster 11, we observed the upregulation of interferon signaling pathways (IFIT1, IFITM1, IFITM3, and TAP1) and NK or neutrophil activities (FCGR3B and TNFSF10) (Fig. 5D, Additional file 1: Fig. S6). In addition, RHOB was upregulated in this region, pointing towards reduced metastatic potential [51]. Therefore, cluster 11 represented a region where cancer growth was limited by pro-inflammatory immune responses. On the other hand, in cluster 13, we observed the presence of TAMs (Additional file 1: Fig. S5B), memory B cells (IGHG1, IGHG3, IGHG4, IGLC2, and IGLC3), and fibroblasts (COL1A1, COL1A2, COL3A1, COL5A1, COL6A1, COL6A2, and FN1) (Additional file 1: Fig. S6). Upregulated cathepsin activity (CTSB, CTSD, and CTSZ) and complement pathway (C1QA, C1S) indicated pro-tumor activity by the TAMs in this region [52–54]. Overall, cluster 13 represented a region with an immune-suppressed pro-tumor microenvironment and a high potential for cancer metastasis.

In summary, SEDR analysis dissected intratumoral heterogeneity within visually homogeneous tumor regions and revealed the tumor outer ring (cluster 13) with TAM infiltration and cancer-associated fibroblasts (CAFs), both of which have been reported to facilitate tumor spread [55, 56].

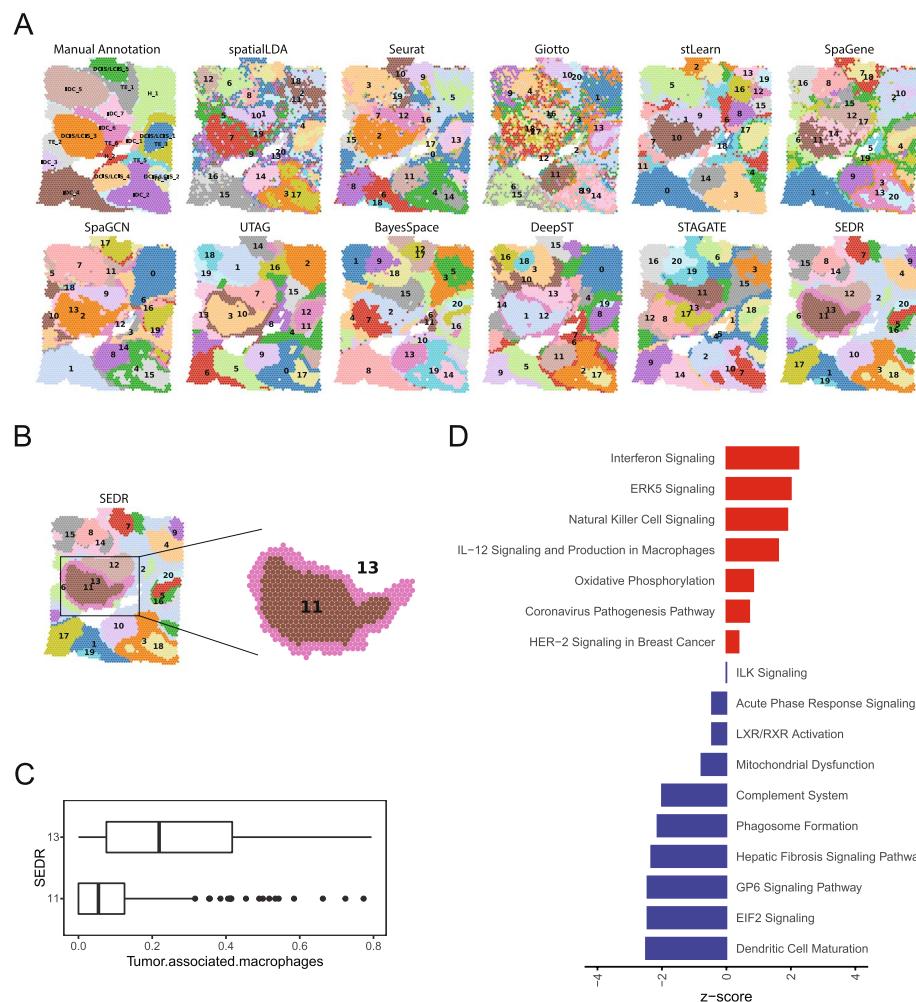


Fig. 5 Application of SEDR on 10×Visium spatial transcriptomics data of human breast cancer. **A** Manual pathology labeling based on H&E staining (annotation) and clustering results of eleven methods. **B** SEDR clusters 11 (core) and 13 (out ring) that captured the annotated DCIS/LCIS_3 region. **C** Percentage of tumor-associated macrophages (TAMs) in clusters 11 and 13. **D** Enriched pathways for differentially expressed genes identified between clusters 11 and 13

Discussion

Cell type heterogeneity is a feature of both healthy and diseased tissue. Capturing this heterogeneity, coupled with its spatial arrangement in the tissue, is crucial when studying the roles of cells and their crosstalk. Spatial omics technologies represent the state-of-the-art approaches for capturing omics data with corresponding spatial information from tissue samples. In this paper, we introduced SEDR, which leverages cutting-edge graph neural network techniques to achieve a better representation of spatial omics data that can be used for clustering and further downstream analyses. SEDR first learns a low-dimensional latent space representation of the transcriptome information with a deep autoencoder

network coupled with masked self-supervised learning, which is then aggregated with spatial neighborhood information by a variational graph autoencoder to create a spatial embedding. This spatial embedding is then concatenated with the encoded gene expression to reconstruct the final gene expression for further analyses. We first demonstrated SEDR's efficacy in delineating the different cerebral cortex layers with higher clarity than competing methods and recapitulating the associated development order by using the joint latent representation with Monocle 3. We also demonstrated SEDR's ability in data imputation with human ovarian cancer and health lymph node tissue data acquired with the 10×Genomics Visium technology. SEDR is not

restricted to Visium-acquired data, as we also showed its efficacy on high-resolution Stereo-seq and Slide-seq data, being able to more accurately recover tissue domains compared to competing methods.

To enhance the analytical power and resolution of spatial omics, we need to integrate multiple datasets from the same tissue. Similar to single-cell transcriptomic data, spatial omics datasets generated in different batches also contain batch-specific systematic variations that present a challenge to batch-effect removal and data integration. In our study, we demonstrated that by combining SEDR and Harmony, we were able to effectively remove batch effects present.

Spatial omics technologies such as Stereo-seq are able to measure a large number of spots in a single experiment through high spatial resolutions and large tissue sizes. Based on current trends, we expect to see ever-increasing throughput from spatial omics experiments, which will result in spatial omics big data that poses significant challenges to data analysis and integration. Computational methods that employ GCNs require the entire graph to be loaded into GPU memory, which inhibits their application to very large datasets. Currently, the current version of SEDR is the most CPU and GPU memory efficient and second fastest among methods that do not employ image processing. To enable SEDR to scale with larger datasets, we will further improve the memory efficiency of SEDR using a GCN mini-batch or parallel techniques to construct large-scale graphs for spatial omics data of high throughput and resolution. Furthermore, technologies with a capture spot size smaller than the diameter of a cell will also require new computational methods that can accurately delineate cells based on capture spots. We plan to integrate cell segmentation based on H&E or DAPI staining images into the SEDR workflow to handle such data.

The current SEDR methodology employs gene expression and spatial information and does not make use of histology images. Contemporary methods such as SpaGCN and stLearn use histological images as input, but in a suboptimal fashion, as demonstrated in our study. Specifically, SpaGCN utilizes histology image pixels as features by calculating the mean color values from the RGB channels directly. However, the pixel values are easily affected by noise and cannot provide semantic features for cell analysis. A potentially more effective approach is to adopt a deep CNN model which can learn high-level representations of histology images. stLearn introduces a deep learning model to extract the image features of spots and integrates them with the spatial location and gene expression. However,

stLearn employs a model pre-trained on natural images which is not fine-tuned for histology images. In the future, we will incorporate histology images as an additional modality into the SEDR model. We will employ an image autoencoder network to first learn image features, followed by joint learning of the latent representation by integrating gene expression, image morphology, and spatial information.

In summary, SEDR is a promising new approach that builds an integrated representation of spots using both transcriptomic data and spatial coordinates. SEDR-derived low-dimensional embedding enables more accurate clustering, trajectory inference, batch effect correction, gene expression imputation, and denoising. Our model is also able to handle spatial transcriptomics with capture spot sizes ranging from 50 μm to less than 1 μm . Furthermore, we applied SEDR on a human breast cancer sample to reveal heterogeneous sub-regions within the seemingly homogeneous tumor region and shed light on the role of immune microenvironments on tumor invasiveness.

Conclusions

In this work, we introduce SEDR as a novel method that processes spatial transcriptomics data to derive deep representation, which benefits various downstream analyses, including spatial clustering, batch integration, trajectory inference, gene expression imputation, and denoising. To obtain more accurate and informative low-dimensional representations from spatial transcriptomics data, SEDR integrates spatial information with RNA-seq data using a variational graph autoencoder, which improves the embedding results significantly compared to the one purely based on RNA-seq data.

We conclude by discussing the future development of SEDR. Although SEDR works well on data generated by 10 \times Genomics Visium, Slide-seq, and Stereo-seq, the ability for it to integrate other types of spatial transcriptomics data is not fully investigated, especially for the technologies based on multiplexed imaging, such as SeqFISH [22] and MERFISH [57], which are different from spot-based ones in terms of data format, resolution, and data quality. We plan to upgrade SEDR to enable it to analyze more data types. We also plan to employ a mini-batch strategy to train models on part of the graph instead of the whole dataset to further enhance its scalability. Furthermore, we will also upgrade SEDR to include more downstream analyses and visualization methods to make it more user-friendly.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-024-01283-x>.

Additional file 1: Supplementary methods. **Fig. S1.** Silhouette score for 5 methods on 12 DLPFC sections. n.s.: p-value > 0.05, *: p-value < 0.05, **: p-value < 0.005, ***: p-value < 0.0005, ****: p-value < 0.00005. **Fig. S2.** ARI boxplot for 3 methods on 12 DLPFC datasets with different K (number of nearest neighbors). **Fig. S3.** Unsupervised clustering results for SEDR and competing methods on olfactory bulb Slide-seq data. **Fig. S4.** Computational requirements of SEDR. A) GPU memory usage for DeepST, STAGATE and SEDR when processing DLPFC, mouse olfactory bulb Slide-seq and Stereo-seq data. B) Time, CPU memory and GPU memory costed by SEDR on simulated data. Simulated data is generated with Slide-seq data (20000 spots) by randomly selecting $\frac{1}{2}$, 1, 2, 4 times of the spots as the raw data. **Fig. S5.** Human breast cancer histology and cell type mixtures of spatial spots. A) H&E staining. B) Probability of cell types for spots that was predicted by Seurat. **Fig. S6.** Differentially expressed genes (DEGs) between SEDR cluster 11 and cluster 13 in human breast cancer data.

Acknowledgements

Not applicable.

Authors' contributions

HF designed the model framework of SEDR. HX and YL enhanced the model design. HX performed most of the data analyses. The remaining analyses were done by RS, YL, KC, ML, HKL, and JL. Figures were plotted by HX, KSA, JC, HF, and HX drafted the manuscript. KC, JL, ML, KSA, and LS helped to revise the manuscript. ML and RU also provided help for maintaining the computational server and managing systems. AC and LL provided the mouse olfactory bulb Stereo-seq data. JC conceptualized and supervised the study. All authors read and approved the final manuscript.

Funding

This research was supported by funding from A*STAR, NRF, NMRC, Singapore (grant id: OFIRG18nov-0103; NRF-CRP19-2017-04; FY21_CF_HTPO_SEED_ID_SlgN_C211418002; FY21_CF_HTPO_SEED_AI3_SlgN_C211118015; A*STAR UIBR; NRF-CRP26-2021-0001).

Availability of data and materials

(1) The LIBD human dorsolateral prefrontal cortex (DLPFC) dataset was downloaded from https://research.libd.org/globus/jhpce_HumanPilot10x/index.html [2]. (2) The 10 \times Genomics Visium spatial transcriptomics and Stereo-seq of mouse olfactory bulb datasets are downloadable from https://github.com/JinmiaoChenLab/SEDR_analyses/tree/master/data [58]. (3) The Slide-seqV2 data can be downloaded from https://singlecell.broadinstitute.org/single_cell/study/SCP815/ [7]. (4) Scripts to reproduce the benchmarking in Fig. 2 are available at https://github.com/JinmiaoChenLab/SEDR_analyses/ [59]. (5) SEDR was written in Python using the PyTorch library. An open-source implementation of SEDR has been released at <https://github.com/JinmiaoChe/nLab/SEDR/> [60].

Declarations

Ethics approval and consent to participate

All datasets used in this study have been previously published. The research conformed to the principles of the Helsinki Declaration.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 29 September 2022 Accepted: 2 January 2024

Published online: 12 January 2024

References

- Stahl PL, Salmen F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, Giacometto S, Asp M, Westholm JO, Huss M, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*. 2016;353:78–82.
- Maynard KR, Collado-Torres L, Weber LM, Uytingco C, Barry BK, Williams SR, Catallini JL 2nd, Tran MN, Besich Z, Tippiani M, et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat Neurosci*. 2021;24:425–36 (https://research.libd.org/globus/jhpce_HumanPilot10x/index.html).
- Ji AL, Rubin AJ, Thrane K, Jiang S, Reynolds DL, Meyers RM, Guo MG, George BM, Mollbrink A, Bergenstrahle J, et al. Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell*. 2020;182(497–514):e422.
- Rodrigues SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, Welch J, Chen LM, Chen F, Macosko EZ. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science*. 2019;363:1463–7.
- Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, Qiu X, Yang J, Xu J, Hao S, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*. 2022;185(1777–1792):e1721.
- Fu X, Sun L, Chen JY, Dong R, Lin Y, Palmiter RD, Lin S, Gu L. Continuous polony gels for tissue mapping with high resolution and RNA capture efficiency. *bioRxiv*. 2021;2021.2003.2017.435795.
- Stickels RR, Murray E, Kumar P, Li J, Marshall JL, Di Bella DJ, Arlotta P, Macosko EZ, Chen F. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat Biotechnol*. 2021;39:313–9 (https://singlecell.broadinstitute.org/single_cell/study/SCP815).
- Liu Y, Yang M, Deng Y, Su G, Enninfoul A, Guo CC, Tebaldi T, Zhang D, Kim D, Bai Z, et al. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell*. 2020;183(1665–1681):e1618.
- Cho CS, Xi J, Si Y, Park SR, Hsu JE, Kim M, Jun G, Kang HM, Lee JH. Microscopic examination of spatial transcriptome using Seq-Scope. *Cell*. 2021;184(3559–3572):e3522.
- Lein E, Borm LE, Linnarsson S. The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science*. 2017;358:64–9.
- Yoosuf N, Navarro JF, Salmen F, Stahl PL, Daub CO. Identification and transfer of spatial transcriptomics signatures for cancer diagnosis. *Breast Cancer Res*. 2020;22:6.
- van den Brink SC, Alemany A, van Batenburg V, Moris N, Blotenberg M, Vivie J, Baillie-Johnson P, Nichols J, Sonnen KF, Martinez Arias A, van Oudenaarden A. Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature*. 2020;582:405–9.
- Dong R, Yuan GC. SpatialDWLS: accurate deconvolution of spatial transcriptomic data. *Genome Biol*. 2021;22:145.
- Andersson A, Bergenstrahle J, Asp M, Bergenstrahle L, Jurek A, Fernandez Navarro J, Lundeberg J. Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Commun Biol*. 2020;3:565.
- Biancalani T, Scalia G, Buffoni L, Avasthi R, Lu Z, Sanger A, Tokcan N, Vanderburg CR, Segerstolpe A, Zhang M, et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat Methods*. 2021;18:1352–62.
- Gayoso A, Lopez R, Xing G, Boyeau P, Valiollah Pour Amiri V, Hong J, Wu K, Jayasuriya M, Mehlan E, Langevin M, et al. A Python library for probabilistic analysis of single-cell omics data. *Nat Biotechnol*. 2022;40:163–6.
- Lopez R, Li B, Keren-Shaul H, Boyeau P, Kedmi M, Pilzer D, Jelinski A, David E, Wagner A, Addad Y, et al. Multi-resolution deconvolution of spatial transcriptomics data reveals continuous patterns of inflammation. *bioRxiv*. 2021;2021.2010.443517.
- Elosua-Bayes M, Nieto P, Mereu E, Gut I, Heyn H. SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res*. 2021;49:e50.
- Danaher P, Kim Y, Nelson B, Griswold M, Yang Z, Piazza E, Beechem JM. Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. *Nat Commun*. 2022;13:385.
- Song Q, Su J. DSTG: deconvoluting spatial transcriptomics data through graph-based artificial intelligence. *Brief Bioinform*. 2021;22:bbaa414.
- Zhang M, Eichhorn SW, Zingg B, Yao Z, Cotter K, Zeng H, Dong H, Zhuang X. Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature*. 2021;598:137–43.

22. Eng CL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, Yun J, Cronin C, Karp C, Yuan GC, Cai L. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature*. 2019;568:235–9.
23. Janesick A, Shelansky R, Gottsch AD, Wagner F, Rouault M, Beliakoff G, de Oliveira MF, Kohlway A, Abousoud J, Morrison CA. High resolution mapping of the breast cancer tumor microenvironment using integrated single cell, spatial and *in situ* analysis of FFPE tissue. *bioRxiv*. 2022;2022.2010.2006.510405.
24. He S, Bhatt R, Brown C, Brown EA, Buhr DL, Chantranuvatana K, Danaher P, Dunaway D, Garrison RG, Geiss G, et al. High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat Biotechnol*. 2022;40:1794–806.
25. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(3573–3587):e3529.
26. Dries R, Zhu Q, Dong R, Eng CL, Li H, Liu K, Fu Y, Zhao T, Sarkar A, Bao F, et al. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol*. 2021;22:78.
27. Zhao E, Stone MR, Ren X, Guenther J, Smythe KS, Pulliam T, Williams SR, Uytengco CR, Taylor SEB, Nghiem P, et al. Spatial transcriptomics at subspot resolution with BayesSpace. *Nat Biotechnol*. 2021;39:1375–84.
28. Pham D, Tan X, Xu J, Grice LF, Lam PY, Raghubar A, Vukovic J, Ruitenberg MJ, Nguyen Q. stLearn: integrating spatial location, tissue morphology and gene expression to find cell types, cell-cell interactions and spatial trajectories within undissociated tissues. *bioRxiv*. 2020;2020.2005.125658.
29. Hu J, Li X, Coleman K, Schroeder A, Ma N, Irwin DJ, Lee EB, Shinohara RT, Li M. SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat Methods*. 2021;18:1342–51.
30. Xu C, Jin X, Wei S, Wang P, Luo M, Xu Z, Yang W, Cai Y, Xiao L, Lin X, et al. DeepST: identifying spatial domains in spatial transcriptomics by deep learning. *Nucleic Acids Res*. 2022;50:e131.
31. Dong K, Zhang S. Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nat Commun*. 2022;13:1739.
32. Kim J, Rustam S, Mosquera JM, Randell SH, Shaykhiev R, Rendeiro AF, Elemento O. Unsupervised discovery of tissue architecture in multiplexed imaging. *Nat Methods*. 2022;19:1653–61.
33. Chen Z, Soifer I, Hilton H, Keren L, Jojic V. Modeling multiplexed images with spatial-LDA reveals novel tissue microenvironments. *J Comput Biol*. 2020;27:1204–18.
34. Liu Q, Hsu CY, Shyr Y. Scalable and model-free detection of spatial patterns and colocalization. *Genome Res*. 2022;32:1736–45.
35. Kipf TN, Welling M. Variational graph auto-encoders. pp. [arXiv:1611.07308](https://arxiv.org/abs/1611.07308); 2016;[arXiv:1611.07308](https://arxiv.org/abs/1611.07308).
36. Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:160902907*. 2016.
37. Xie J, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. In: Proceedings of the 33rd International Conference on Machine Learning (Maria Florina B, Kilian QW eds.), vol. 48. pp. 478–487. Proceedings of Machine Learning Research: PMLR; 2016:478–487.
38. Gilmore EC, Herrup K. Cortical development: layers of complexity. *Curr Biol*. 1997;7:R231–234.
39. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014;32:381–6.
40. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566:496–502.
41. Chini M, Hanganu-Opatz IL. Prefrontal cortex development in health and disease: lessons from rodents and humans. *Trends Neurosci*. 2021;44:227–40.
42. Nadarajah B, Parnavelas JG. Modes of neuronal migration in the developing cerebral cortex. *Nat Rev Neurosci*. 2002;3:423–32.
43. Tran HTN, Ang KS, Chevrier M, Zhang X, Lee NYS, Goh M, Chen J. A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol*. 2020;21:12.
44. Wang Y, Song B, Wang S, Chen M, Xie Y, Xiao G, Wang L, Wang T. Sprod for de-noising spatially resolved transcriptomics data based on position and image information. *Nat Methods*. 2022;19:950–8.
45. King HW, Orban N, Riches JC, Clear AJ, Warnes G, Teichmann SA, James LK. Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Sci Immunol*. 2021;6:eabe6291.
46. Nguyen PHD, Ma S, Phua CZJ, Kaya NA, Lai HLH, Lim CJ, Lim JQ, Wasser M, Lai L, Tam WL, et al. Intratumoural immune heterogeneity as a hallmark of tumour evolution and progression in hepatocellular carcinoma. *Nat Commun*. 2021;12:227.
47. Poljak K. Heterogeneity in breast cancer. *J Clin Invest*. 2011;121:3786–8.
48. Pal B, Chen Y, Vaillant F, Capaldo BD, Joyce R, Song X, Bryant VL, Penington JS, Di Stefano L, Tubau Ribera N, et al. A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *EMBO J*. 2021;40:e107333.
49. Asiry S, Kim G, Filippou PS, Sanchez LR, Entenberg D, Marks DK, Oktay MH, Karagiannis GS. The cancer cell dissemination machinery as an immunosuppressive niche: a new obstacle towards the era of cancer immunotherapy. *Front Immunol*. 2021;12:654877.
50. Kuroda H, Jamiyan T, Yamaguchi R, Kakumoto A, Abe A, Harada O, Masunaga A. Tumor microenvironment in triple-negative breast cancer: the correlation of tumor-associated macrophages and tumor-infiltrating lymphocytes. *Clin Transl Oncol*. 2021;23:2513–25.
51. Ju JA, Godet I, DiGiacomo JW, Gilkes DM. RhoB is regulated by hypoxia and modulates metastasis in breast cancer. *Cancer Reports*. 2020;3:e1164.
52. Olson OC, Joyce JA. Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. *Nat Rev Cancer*. 2015;15:712–29.
53. Roumenina LT, Daugan MV, Noe R, Petitprez F, Vano YA, Sanchez-Salas R, Becht E, Meilleroux J, Clec'h BL, Giraldo NA, et al. Tumor cells hijack macrophage-produced complement C1q to promote tumor growth. *Cancer Immunol Res*. 2019;7:1091–105.
54. Fraser D, Melzer E, Camacho A, Gomez M. Macrophage production of innate immune protein C1q is associated with M2 polarization (lINM1P_434). *Am Assoc Immunol*. 2015;194:56.11.
55. Monteran L, Erez N. The dark side of fibroblasts: cancer-associated fibroblasts as mediators of immunosuppression in the tumor microenvironment. *Front Immunol*. 2019;10:1835.
56. Lin Y, Xu J, Lan H. Tumor-associated macrophages in tumor metastasis: biological roles and clinical therapeutic applications. *J Hematol Oncol*. 2019;12:76.
57. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*. 2015;348:aaa6090.
58. Xu H. Data for SEDR analysis. Github. 2023; https://github.com/JinmiaoChenLab/SEDR_analyses/tree/master/data.
59. Xu H. Code for benchmarking SEDR with other methods. Github. 2023; https://github.com/JinmiaoChenLab/SEDR_analyses.
60. Xu H. Code for SEDR. GitHub. 2023; <https://github.com/JinmiaoChenLab/SEDR>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.