LETSmix: a spatially informed and learning-based domain adaptation method for cell-type deconvolution in spatial transcriptomics

AUTHORS

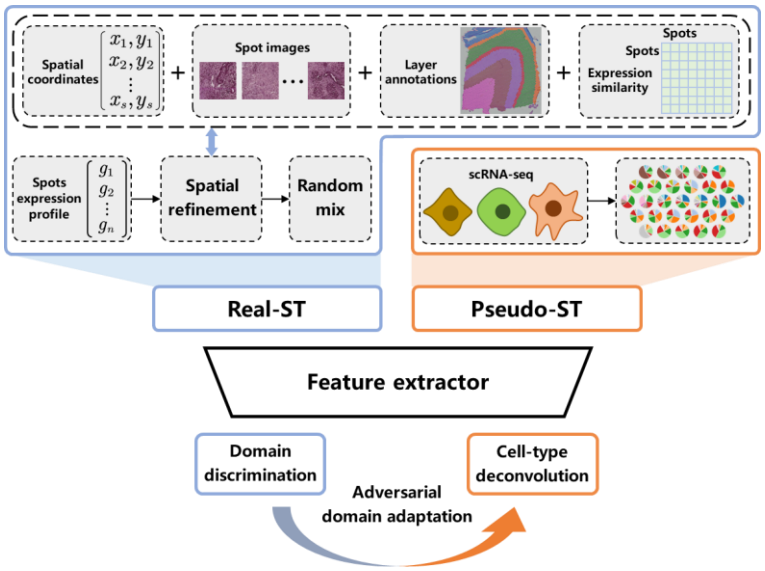
Yangen Zhan[1,†], Yongbing Zhang[2,†,*], Zheqi Hu[2,†], Yifeng Wang[2], Zirui Zhu[2], Sijing Du[2] and Xiu Li[1,*]

[1] Division of Information Science and Technology, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518052, China
[2] School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China

* To whom correspondence should be addressed. Email: ybzhang08@hit.edu.cn
Correspondence may also be addressed to Xiu Li. Email: li.xiu@sz.tsinghua.edu.cn

† These authors contributed equally to this work.

GRAPHICAL ABSTRACT



ABSTRACT

Spatial transcriptomics (ST) has revolutionized our understanding of gene expression patterns by incorporating spatial context. However, many ST technologies operate on heterogeneous cell mixtures due to limited spatial resolution. Current methods for cell-type deconvolution often underutilize spatial context information inherent in ST and the paired histopathological images, meanwhile neglect domain variances between ST and the reference single-cell RNA sequencing (scRNA-seq) data. To address these issues, we present LETSmix, a deep learning-based domain adaptation method trained on labelled pseudo-spots generated from scRNA-seq data, and mixed real-spots that are refined by a designed LETS filter leveraging correlations among neighboring spots with similar morphological features. The performance of LETSmix is demonstrated across three public ST datasets through comprehensive assessments, setting a new record among current state-of-the-art models. Our findings indicate that LETSmix accurately estimates the proportions of various cell types, and effectively maps them to the expected areas. The utilization of domain adaptation techniques enables LETSmix to achieve highly stable results when trained with different reference scRNA-seq datasets. Applications of LETSmix to diverse tissues, including the human dorsolateral prefrontal cortex, human pancreatic ductal adenocarcinoma, and mouse liver, showcase its robust performance and generalizability across different biological contexts.

INTRODUCTION

In the last decade, the continuous advancement of single-cell RNA sequencing (scRNA-seq) technology facilitates high-throughput sequencing of individual cells, revealing comprehensive gene expression profiles within them (1,2). This technological leap has unearthed cellular heterogeneity, enabling the identification of diverse cell types, subpopulations, and transcriptional state alterations within complex cell populations (3). However, the intrinsic nature of scRNA-seq—sequencing individual cells—necessitates the isolation of cells from their native tissue prior to sequencing, hampering the integration of spatial information for analysis (4). To unravel the spatial distribution of gene transcriptomes and the tissue microenvironment, spatial transcriptomics (ST) technology has emerged. This innovative approach enables the detection of gene expression profiles across numerous locations within tissue regions while retaining their positional information (5,6). Currently, ST techniques are broadly categorized into image-based and sequence-based methodologies, which complement each other in a manner (7). Image-based methods like MERFISH (8) and seqFISH+ (9) utilize fluorescent probes to target specific genes, offering high resolution even at the subcellular level (10). However, due to their reliance on targeted fluorescent probes, these techniques are confined in detecting a limited number of genes, typically a few hundred (11). On the other hand, sequence-based techniques, such as Spatial Transcriptomics (ST) (12) and SLIDE-seq (13), may capture expression profiles of a complete gene repertoire, but their sequencing resolution is lower, often encompassing multiple cells within each detection site, referred to as a "spot" (7). Notably, these sequence-based techniques usually not only provide the positional coordinates of each detection site but also H&E stained histological images of the sequenced tissue. Despite ongoing advancements in improving their resolution, accurately assigning individual cells to each spot remains a challenge (14).

In sequence-based ST data, to discern the spatial distribution of diverse cell types and gain deeper insights into their compositional structure, prevailing methods integrate scRNA-seq with ST data (15,16). scRNA-seq furnishes gene expression profiles of individual cells alongside their corresponding cell-type information, facilitating the inference of cell type-specific gene expression traits. Integration of this information with expression patterns observed in each spot from ST allows the analysis of contribution from each cell type at each

spot (17,18,19). Current mainstream cell-type deconvolution methodologies can be broadly categorized into three types: statistical probabilistic models, matrix factorization-based models, and deep learning-based models (16,20,21). Probabilistic models such as RCTD (19), Stereoscope (22), Cell2location (23) and SONAR (24) assume that gene expressions in scRNA-seq and ST data follow certain distributions, such as the Negative Binomial or Poisson distribution. Although various influencing factors, such as technology sensitivity, batch effects, and per-location shifts, can be specifically parameterized within these methods, the regressed parameters may deviate from their original design. Similarly, these factors can also be transformed into parameter matrices in models based on non-negative matrix factorization (NMF), exemplified by SPOTlight (17), SpatialDWLS (18), and CARD (25). The learned cell-type gene signatures from scRNA-seq are used to disentangle each spot in ST by the aforementioned two modelling strategies. On the other hand, deep learning approaches are gaining traction as promising solutions because of their capacity to apprehend and leverage intricate patterns and correlations within data. Techniques such as graph networks (26,27) capitalize on spatial relationships, while domain adaptation methods (28,29) bridge technical variances between scRNA-seq and ST. DSTG (30) constructs a graph network, generating synthetic data from scRNA-seq to approximate true cell-type proportions in ST. Furthermore, self-supervised training with variational graph autoencoders applied in Spatial-ID (31), tissue histological image integration applied in SpaDecon (32), and domain-adversarial learning applied in CellDART (33) are innovative deep learning strategies, improving cell-type deconvolution accuracy in ST data.

However, these existing methods either overlook spatial correlations among individual sequencing spots or inadequately leverage ancillary information available in ST datasets such as the paired high-resolution histopathological images. Additionally, the inherent lower sensitivity of ST sequencing technology in detecting genes compared to scRNA-seq results in generally diminished gene expression levels (31). This discrepancy, termed the “platform effect”, poses significant challenges for current cell-type deconvolution methodologies, demanding a high level of coherence between the reference scRNA-seq and ST data (34). Here, we introduce a novel deconvolution method, termed LETSmix, which comprehensively leverages spatial context information and employs domain adaptation to effectively tackle these challenges. LETSmix firstly constructs a LETS filter to refine raw ST data by capturing

spatial correlations, which utilizes information from Layer annotations, spots gene Expressions, histological image Texture features, and Spatial positions of spots. LETSmix employs a domain-adversarial training paradigm, and consists of a shared feature extractor and two separate classifiers. The source classifier estimates cell-type proportions and is trained on labelled pseudo-spots synthesized from the reference scRNA-seq dataset, while the domain classifier contributes to the learning of domain-invariant gene expression features using data from both real-spots and pseudo-spots. Importantly, to address the issue of unbalanced samples between the two domains, where infinite pseudo-spots can be generated in the source domain while only limited real-spots are available in the target domain, a mixup procedure is performed on the spatially refined real-spot data before feeding them into the feature extractor. These mixed real-spots aid in alleviating the data scarcity issue in the target domain and enhancing the robustness and generalization capabilities of the proposed LETSmix model.

Rigorous benchmarking analyses on three publicly available datasets substantiate the superior performance of LETSmix over existing state-of-the-art methods, both qualitatively and quantitatively. Comparative evaluations encompass the statistical probabilistic model Cell2location, NMF-based model CARD, and learning-based models like SpaDecon and CellDART. Evaluation across diverse tissues—human brain (DLPFC), human pancreatic ductal adenocarcinoma (PDAC), and mouse liver (Liver)—demonstrates the precision of LETSmix in mapping layer-specific cell types within real-ST datasets and inferring their proportions in individual spots. Ablation analyses underscore the significance of distinct functional components within the model structure. Furthermore, a series of controlled experiments assess the performance of LETSmix across varied hyperparameter choices.

MATERIAL AND METHODS

Public dataset collection

Human dorsolateral prefrontal cortex (DLPFC) data. The 10X Visium DLPFC dataset is derived from a postmortem 30-year-old neurotypical subject (35). Our experiments include all twelve ST samples, each containing 3000-5000 spots with 33,538 common genes. Layer annotations

are provided by assigning each spot to a specific cortical layer, ranging from L1 to L6 and WM (white matter), based on layer-specific marker genes and expert inspections (36). A few spots without layer annotations are discarded in our experiments. The reference single-nucleus dataset (also referred to as scRNA-seq) is obtained from the DLPFC tissues of different postmortem individuals without neurological disorders (37). This scRNA-seq dataset consists of 56,561 cells and 30,062 genes. We focus on 10 layer-specific excitatory neurons to evaluate the controlled models using AUC and ER metrics.

Human pancreatic ductal adenocarcinoma (PDAC) data. The PDAC dataset is acquired from two tumorous tissue sections of two patients, respectively, denoted as PDAC-A and PDAC-B (38). The ST datasets are collected using the Spatial Transcriptomics technology while their matched scRNA-seq datasets are obtained through inDrop. The paired ST and scRNA-seq data are both derived from the same tissues of the same patients. Additionally, an external scRNA-seq dataset named PDAC-Peng (39) obtained through 10x Chromium is used to evaluate model performance under unmatched situation. Cell types and their compositions are different in the three scRNA-seq datasets. Although the true cell-type composition of each spot in the ST data of PDAC-A and PDAC-B remains unknown, the overall composition of the whole tissue, i.e. the average cell-type proportions in all spots, should be close to that of the matched reference scRNA-seq dataset. Thus, a JSD value can be calculated to measure the consistency between the estimated overall cell-type composition and the expected ground truth. Moreover, a few annotated cell types in the reference scRNA-seq dataset are assumed to be enriched within specific tissue regions (Supplementary Table S1). Based on this prior knowledge, the ER metric is implemented to compare the performance of different deconvolution methods in predicting their regional distribution patterns.

Mouse liver (Liver) data. The Liver dataset is obtained from the paper (40). Three consecutive Visium slices of healthy mouse liver tissues are analyzed in this study, each delineated into five distinct regions. These ST samples are primarily composed of hepatocytes, making it challenging for deconvolution tools to accurately identify other cell types. The reference scRNA-seq datasets include cells obtained from three different digestion protocols: ex vivo digestion, in vivo liver perfusion, and frozen liver single-nucleus RNA-seq (referred to as "exVivo scRNA-seq", "inVivo scRNA-seq", and "nuclei scRNA-seq", respectively) (21). All three scRNA-seq datasets contain the same nine cell types. Notably, portal vein and central vein

endothelial cells (ECs) are anticipated to be exclusively present in the portal and central regions, respectively. Moreover, previous researches based on confocal microscopy suggests that the average proportion of each cell type in all spots within the ST sample is equivalent to that in the nuclei scRNA-seq dataset (21,40). Therefore, JSD values were computed to evaluate the estimated proportions in ST, and the cell-type proportions in the ex vivo and in vivo scRNA-seq datasets were adjusted to match those in the nuclei scRNA-seq dataset.

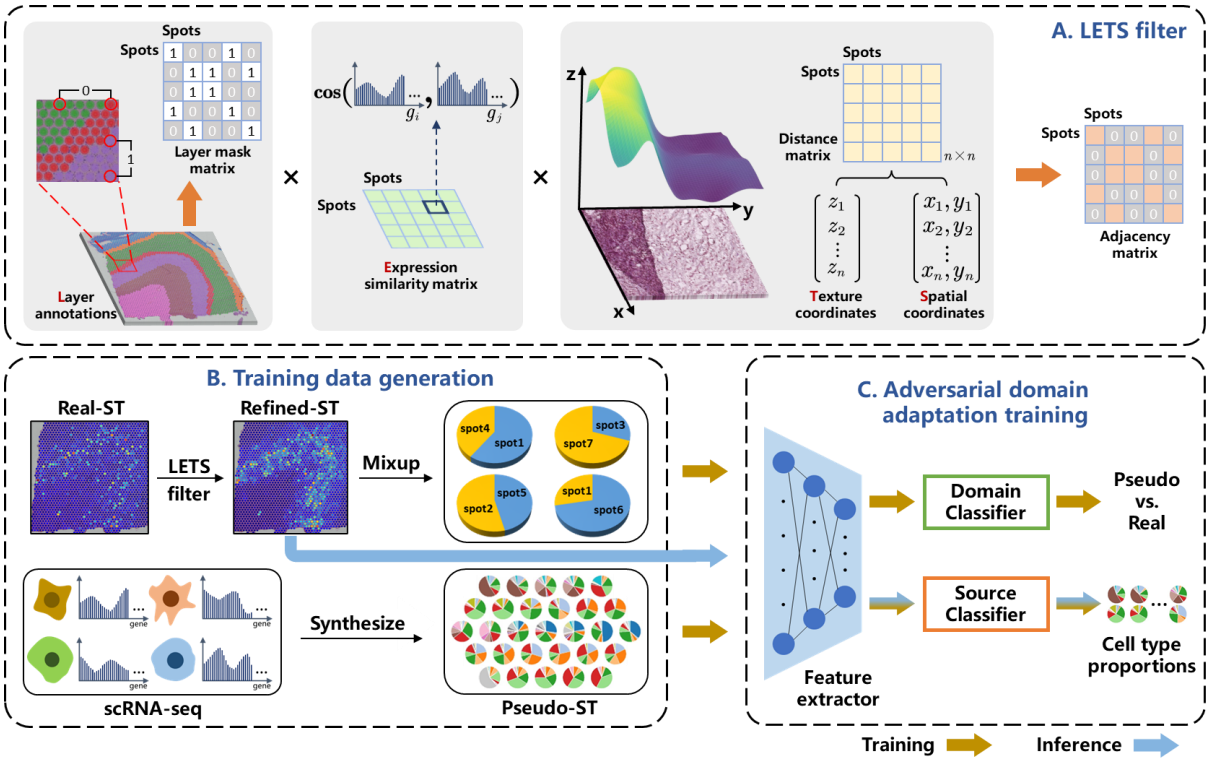


Figure 1. Schematic overview of the LETSmix workflow. LETSmix is designed to perform cell-type deconvolution for spatial transcriptomics based on the labelled reference scRNA-seq data. (**A**) An adjacency matrix reflecting internal correlations between spots in ST is constructed leveraging information from layer annotations, spots gene expression, histological image texture features and spots spatial coordinates. (**B**) Pseudo-ST data with known cell type compositions are synthesized by randomly selecting cells from the reference scRNA-seq dataset. Real-ST data are refined by the LETS filter constructed from (**A**), and then undergo a mixup procedure for data augmentation. (**C**) The network structure comprises a shared feature extractor and two classifiers: the source classifier estimates cell type proportions, and the domain classifier identifies spots as real or pseudo. The two branches are trained in an adversarial manner, aiming to eliminate domain shifts in the extracted

features. After finishing training the model, the spatially refined real-ST data are directly fed into the feature extractor and the source classifier for cell-type deconvolution.

Implementation of LETSmix

A schematic diagram illustrating the LETSmix method is presented in Fig. 1. The overall network framework comprises three main components. Firstly, leveraging region annotations, gene expression data, pathological images, and spot coordinates from the ST dataset, an adjacency matrix is constructed, accurately capturing the spatial correlations among different spots. This matrix is subsequently employed to perform local smoothing on the ST dataset, emphasizing spatial relationships between neighboring spots with similar morphological features. Secondly, a fixed number of cells are randomly selected from the scRNA-seq dataset to synthesize gene expression data for each pseudo-spot. Meanwhile, the spatially refined real-ST data are also randomly mixed to generate more samples in the target domain. In the final stage, the synthesized pseudo-ST data and mixed real-ST data are simultaneously fed into a deep learning network for adversarial domain adaptation training. Here, the domain classifier aims to differentiate the domain of the input spot expression data, while the source classifier estimates cell-type compositions within each spot in both pseudo-ST and real-ST datasets.

Spatial refinement for ST data. Expression data could be noisy due to limited cell numbers in each spot. To address this, we construct an adjacency matrix that aggregates information from neighboring spots. Initially, the distance $d_{i,j}$ between spots i and j is calculated incorporating spatial, morphological and expression features. Following the method adopted in SpaDecon, the coordinate z_i for spot i, representing its histological image texture features, is defined as

$$z_i = \frac{\mu_r\sigma_r^2+\mu_g\sigma_g^2+\mu_b\sigma_b^2}{\sigma_r^2+\sigma_g^2+\sigma_b^2}, \quad (1)$$

where μ_r, μ_g, μ_b and σ_r, σ_g, σ_b are the means and standard deviations, respectively, of the RGB values across all pixels in the $2r \times 2r$ image patch of spot i, and r is the radius of each spot. z_i is further scaled as

$$z_i^* = \frac{z_i-\mu_z}{\sigma_z} \times \max(\sigma_x, \sigma_y). \quad (2)$$

Here, μ_z and σ_x, σ_y, σ_z are the means and standard deviations, respectively, of (x, y, z) coordinates of all spots in the ST data. This scaling ensures that the weight of histological features approximately matches that of spatial location features. The Euclidean distance between two spots i and j using the three coordinates is given by

$$d_{i,j} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 + \left(z_i^* - z_j^*\right)^2}. \quad (3)$$

The total number of spots in the real-ST data is denoted by n, and the adjacency matrix $\mathrm{A} = \left[a_{i,j}\right]_{n\times n}$ is constructed based on $d_{i,j}$ combined with expression data and layer annotation information, as

$$a_{i,j} = \exp\left(-\frac{d_{i,j}^2}{2l_i^2}\right) \times c_{i,j} \times m_{i,j}, \quad (4)$$

where l_i is a scaling factor controlling the degree of local smoothing, $c_{i,j}$ is the cosine similarity between the expression data of two spots, and $m_{i,j}$ is an element from a mask matrix $\mathrm{M} = \left[m_{i,j}\right]_{n\times n}$ constructed using layer annotations, as

$$m_{i,j} = \begin{cases} 1, \text{ if spots } i \text{ and } j \text{ are in the same layer} \\ 0, \text{ otherwise} \end{cases}. \quad (5)$$

Note that SpaDecon only utilizes spatial coordinates and histological images to construct the adjacency matrix, whereas we leverage additional information from layer annotations and gene expression data. This improvement brought by our designed adjacency matrix accurately reflects the intrinsic spatial correlations within ST data, thereby enhancing the accuracy of cell-type deconvolution results. Next, we describe the selection strategy of an appropriate value for l_i. Define s_i by

$$s_i = \sum_{j=1}^{n} a_{i,j} - 1, \quad (6)$$

which represents the summed weights of other spots with respect to the target smoothing spot i. The value of s_i is determined by l_i. Although the relationship between them cannot be definitively ascertained through numerical methods, it can be established that s_i increases monotonically with respect to l_i. Thus, we start by specifying the desired value of s_i and then iteratively approach the value of l_i, as outlined in Supplementary Algorithm S1. In our experiments, s_i of different spots where $i \in \{1,2,\ldots,n\}$ are set to be the same value $\tilde{s}$. Note

that SpaDecon employs a similar approximation technique for determining the value of l. However, the scaling factor l in SpaDecon is a scalar, ensuring only the average summed weights of other spots to be a specific value, resulting in varying degrees of local smoothing for different spots. Despite that n different values in the vector l are required to be approximated in our improved method, the designed algorithm maintains a computational complexity similar to the original. It requires only a single round of approximation to directly compute each value in the vector l, thereby not compromising the operational speed of the model. With the constructed adjacency matrix A, we perform matrix multiplication to obtain smoothed gene expression data G_r^* of real-ST data, as

$$\mathrm{G_r^* = AG_r}. \tag{7}$$

Mixup for ST and scRNA-seq data. Expression data of each pseudo-spot are synthesized from randomly selected k cells in the reference scRNA-seq dataset. We use $[g_{c1}, g_{c2}, \dots, g_{ck}]$ to represent the gene expression profile of each selected cell. Meanwhile, random weights $[w_{c1}, w_{c2}, \dots, w_{ck}]$ are assigned to each cell, representing its proportion in the generated pseudo-spot, and the sum of k weights is constrained to be 1. Thus, the gene expression profile of the pseudo-spot is the weighted sum of k cells, denoted as:

$$g_p = \sum_{i=1}^{k} w_{ci} g_{ci}\,, \text{ where } \sum_{i=1}^{k} w_{ci} = 1. \tag{8}$$

Similarly, the ground truth cell-type proportions in this pseudo-spot are also the weighted sum of the cell types of the selected k cells. In this way, the total amount of gene expression in a pseudo-spot is equivalent to that in a cell from scRNA-seq data.

Through various combinations, an infinite number of pseudo-spots can be synthesized, whereas the quantity of real-spots is typically limited. To address the discrepancy in data volume between the target domain and the source domain during the subsequent domain adaptation process, we also apply mixup to real-ST data refined by the LETS filter, thereby significantly increasing the amount of data in the target domain. Specifically, a mixup ratio λ is randomly sampled from a Beta distribution $Beta(\alpha, \alpha)$ for $\alpha > 0$. The resulting mixed-spot is generated by:

$$g_m = \lambda g_{ri}^* + (1 - \lambda) g_{rj}^*, \tag{9}$$

where g_{ri}^{*} and g_{rj}^{*} are two randomly selected refined real-spots. In fact, mixup is a commonly used data augmentation strategy in the field of computer vision. We find it particularly suitable for ST data, where concerns about its potential negative impacts such as the loss of target images or confusing merged objects are not as relevant.

Adversarial domain adaptation training for the deep learning network. The training process of the LETSmix network, consisting of non-linearly activated fully connected layers (detailed in Supplementary Table S2), is conducted using the pytorch-gpu package (version 1.12.0). Initially, a shared feature extractor F is utilized to derive low-dimensional features from both pseudo-ST and locally smoothed real-ST data. The subsequent source classifier S and domain classifier D are trained in an adversarial manner. Specifically, we first pretrain the feature extractor and the source classifier using only the pseudo-ST data for cell-type deconvolution, leaving the domain classifier aside. Then, the two branches are trained alternatively. In the domain classifier training procedure, learnable weight parameters in the feature extractor and source classifier are frozen. This process aims to effectively distinguish whether the extracted features originate from pseudo-ST or real-ST. At this stage, the loss function is defined as the binary cross-entropy between the predictions of the domain classifier and the assigned domain label, formalized as

$$\mathcal{L}_d = y_d \cdot \log\left(D(F(x))\right) + (1 - y_d) \cdot \log\left(1 - D(F(x))\right), \qquad (10)$$

where x and y_d denote the input data and the domain label (0 for pseudo-ST and 1 for real-ST), respectively. In the source classifier training procedure, we freeze the trained domain classifier and use it to classify the domain of input data. At this stage, the feature extractor aims to deliberately confuse the domain discriminator, preventing it from accurately distinguishing the domain of the extracted features. This serves the purpose of eradicating domain variances existed in the feature space. Simultaneously, the source classifier endeavours to accurately estimate the proportions of different cell types within each input pseudo-spot. Thus, there are two components in the loss function. The first component is defined as the Kullback-Leibler divergence (KLD) between the estimations of the source classifier and the cell-type proportion labels, while the second component is the binary cross-entropy between the predictions of the frozen domain classifier and the inverted assigned domain label, which can be formalized as

$$\mathcal{L}_s = \sum_{i=1}^{z} y_{i,s} \log\left(\frac{S(F(x))_i}{y_{i,s}}\right) + (1 - y_d) \cdot \log\left(D(F(x))\right) + y_d \cdot \log\left(1 - D(F(x))\right), \qquad (11)$$

Here, z is the total number of cell types, while $y_{i,s}$ and $S\big(F(x)\big)_i$ stand for the label and estimation for the ith cell type, respectively. Note that the KLD loss is only calculated for input data from the source domain (pseudo-ST), while the binary cross-entropy loss is calculated for data from both domains. Maintaining the same batch size for data from pseudo-ST and real-ST, we find that it is crucial for the training iterations of the domain classifier to be more than that of the source classifier within a cycle of alternation, where the ratio between the two branches is d: 1. This observation aligns with logical expectations because the ability of the domain discriminator to guide the feature extractor in learning domain-invariant features hinges upon its accurate domain differentiation.

Preprocessing spatial and single-cell datasets

The preprocessing of ST and scRNA-seq data is conducted in Python using Scanpy packages (version 1.9.3). Initially, m highly expressed marker genes for each cell type in scRNA-seq data are selected through Wilcoxon rank-sum test. Note that expression data in scRNA-seq undergo normalization and log1p-transformation before the marker gene selection process, ensuring an equivalent total amount of gene expression for each cell. The subsequent procedures utilize the intersection between the extracted marker genes and all genes in the ST data. Following this, expression data of the selected marker genes in both scRNA-seq and ST are once again normalized and log1p-transformed. Importantly, the normalized total amount of expressions in each cell from the scRNA-seq data should be $(1 + \tilde{s})$ times larger than that in each spot from the ST data. This adjustment accounts for the real-ST refinement process, ensuring that the total counts in a generated pseudo-spot and a real spot are equal. Finally, the expression data in pseudo-ST and real-ST are min-max scaled, respectively, to ensure that the values lie between 0 and 1. This comprehensive preprocessing pipeline sets the stage for subsequent analyses in LETSmix, enhancing the comparability and accuracy of the deconvolution results across different datasets.

Evaluation metrics

In the experiments presented in this paper, three quantitative metrics are employed to evaluate the performance of various cell-type deconvolution methods: Area Under the Curve

(AUC), EnRichment (ER), and Jensen-Shannon Divergence (JSD). AUC and ER gauge the agreement between predicted spatial distribution patterns of regionally restricted cell types and layer annotations, while JSD assesses the alignment of the estimated general cell-type composition in the ST dataset with prior knowledge from the matched scRNA-seq dataset. In the context of receiver operating characteristic (ROC) analysis, a spot is labelled as 1 if it resides in the target region of the layer-specific cell type under examination; otherwise, it is labelled as 0. The AUC value is then calculated between the estimated proportions of this cell type in all spots and the assigned labels. Despite the common use of the AUC metric in other cell-type deconvolution studies (33,41), it is acknowledged to have certain limitations. Notably, not all spots within the target region necessarily contain the layer-specific cell type under investigation in reality. Additionally, setting the label as 1 does not realistically represent the actual proportion of this cell type in a spot within the target region. Therefore, even if the model accurately predicts cell-type compositions in each spot, achieving an AUC value of 1 is improbable. To address these limitations, an additional metric called EnRichment (ER) is introduced in this study. ER calculates the estimated total counts of a layer-specific cell type within the target region divided by the total counts of that cell type in the entire tissue section. Due to the unknown total number of cells in each spot, this value is approximated using the estimated proportions of that cell type as a substitute. The formalization of ER is given by:

$$ER_j = \frac{\sum_{i \in \Omega'} p_{i,j}}{\sum_{i \in \Omega} p_{i,j}}, \qquad (12)$$

where Ω' and Ω denote spots in the target region and the entire tissue section, respectively. $p_{i,j}$ is the estimated proportion of the jth cell type in the ith spot. In cases where multiple cell types share a common target region, such as cancer clone A and cancer clone B cells in the PDAC-A dataset, their estimated proportions in each spot are firstly summed together as a broader cell type, and an ER value is calculated based on this merged cell type. For the JSD metric, the overall proportions of various cell types are computed in both scRNA-seq and ST data. The calculation for the latter involves averaging the cell-type compositions across all estimated spots, as illustrated by Formula (13).

$$P_{st} = (\sum_{i=1}^{n} p_i)/n,$$
$$JSD(P_{sc}||P_{st}) = \frac{1}{2}\left(D_{KL}(P_{sc}||M) + D_{KL}(P_{st}||M)\right), \quad (13)$$
where $M = (P_{sc} + P_{st})/2, D_{KL}(P||Q) = \sum_i P_i \log(P_i/Q_i)$.

All quantitative results presented in this paper are obtained from five repeated experiments. The models under evaluation were trained from scratch five times with different random seeds. In each repeated experiment, the learned parameters in LETSmix are saved at the training epoch where the best performance is achieved. The distribution patterns of 10 layer-specific cell types in the DLPFC datasets estimated by different models were examined using the AUC and ER metrics. The AUC metric was also utilized in the ablation and hyperparameter analysis for the DLPFC dataset. For the PDAC dataset, ER and JSD values are calculated for different models under matched conditions, and only ER values are computed in unmatched conditions. Similarly, the ER and JSD metrics were applied in the Liver dataset to analyze the estimated distribution patterns of two regionally restricted cell types and the general cell-type compositions, respectively.

Method comparison

The deconvolution performance of LETSmix is benchmarked against other four state-of-the-art computational methods: CellDART (33), SpaDecon (32), Cell2location (23), and CARD (25). CellDART and SpaDecon are two deep learning-based models, and can be seen as different simplified versions of LETSmix. CellDART also follows a source-domain dual-branch framework, leveraging the adversarial domain adaptation strategy to bridge technical variances between ST and the scRNA-seq data, but completely ignores the inherent correlations among spots in the ST dataset and the significant imbalance between the number of samples in the source and the target domain. SpaDecon simply utilizes information from spot coordinates and histological images to locally smooth the ST dataset. The model is trained on scRNA-seq data with a supervised clustering algorithm, and is directly employed to infer cell-type proportions in ST. Cell2location assumes that gene expression data in scRNA-seq and ST data follow a negative binomial distribution. Cell type signatures are inferred from the reference scRNA-seq dataset and used to estimate cell type proportions decomposed from ST data. The model incorporates parameters capturing various sources of data variability. In CARD, the count matrics of scRNA-seq and ST data are decomposed into different parameter matrices. It infers cell type signatures from the

reference scRNA-seq dataset and estimates the cell type composition matrix from ST data using NMF regression. The model introduces a conditional autoregressive assumption, leveraging position coordinates information to consider spatial correlations among spots. For all methods, the default parameters are applied for analyses unless otherwise specified.

Parameter selection

The key parameters of the proposed LETSmix method include the degree of the used spatial context information ($\tilde{s}$), the ratio of training iterations between the domain classifier and the source classifier (d), the number of marker genes per cell type (m), and the number of cells per pseudo-spot (k). To search optimal values for these parameters, we investigated the performance variation of LETSmix on the DLPFC dataset across a spectrum of conditions as a reference (Supplementary Figure S1).

Different choices for $\tilde{s}$ revealed that optimal performance is achieved when its value ranges from 0.5 to 1. Extremes in the utilization of spatial background information, either too low or too high, impeded the cell-type deconvolution process. However, considering that a parameter similar to $\tilde{s}$ is incorporated in SpaDecon, we set the value of $\tilde{s}$ to be 0.5, which is the default value applied in SpaDecon, for a fair comparison.

Effective discrimination by the domain discriminator between the two domains is crucial in guiding the feature extractor to eliminate domain differences, which is substantiated by the fact that more training iterations for the domain discriminator lead to enhanced performance. However, improvements in deconvolution performance stop when the value of d reaches 10, as the discriminative capacity of the domain discriminator itself has inherent limitations. Notably, the official implementation in CellDART did not prioritize additional training iterations for the domain classifier. In our reproduced experimental outcomes, adhering to the training methodology outlined in the public code of CellDART yields notably lower accuracy in deconvolution results compared to those reported in the original paper. To maintain fairness, all experiments in this paper maintain a proportional ratio of 10:1 for training iterations between the domain and source classifiers within both LETSmix and CellDART.

Increasing the number of marker genes per cell type during the data preprocessing stage consistently yields improved results, providing more comprehensive information for the model but at the expense of higher computational cost. Empirically, the value of m is set to be 50 for all three public datasets used in this study.

Similar to m, increasing the value of k consistently leads to enhanced deconvolution performance. With sufficient data, the model is able to learn more diverse features and adapt to complex application scenarios by including more cells in each pseudo-spot. In our experiments, k is set to be 8 for all datasets as it is the default value applied in CellDART for a fair comparison.

RESULTS

Overview of LETSmix

The proposed cell-type deconvolution method integrates spatial correlations among spots, meanwhile eliminates domain variances between ST and the reference scRNA-seq data. Illustrated in Figure 1, both scRNA-seq and ST data undergo preprocessing before entering the deep learning network. The synthesis of pseudo-ST from annotated scRNA-seq data involves randomly selecting cells to generate pseudo-spots, with their total gene expression counts determined by a weighted sum of the chosen cells. This procedure accumulates a substantial corpus of pseudo-ST data with known cell-type compositions, facilitating supervised training of the feature extractor and source classifier. Real-ST data are locally smoothed with ancillary spatial context information. Considering potential noise in expression data due to limited cell numbers per spot and sequencing techniques, it is rational to aggregate information from neighboring spots with similar histological and gene expression features (32). To this end, an adjacency matrix, termed the LETS filter, is designed to signify spatial, histological, and gene expression similarities among real-ST spots. Further optimization of this matrix incorporates layer annotations as masks, restricting information sharing to spots within the same layer. Multiplying the constructed adjacency matrix with the original real-ST expression count matrix effectively enhances data quality. Due to inherent technical variances between ST and scRNA-seq, the refined real-ST and the synthetic

pseudo-ST data undergo adversarial training within the deep learning network, employing a label inverting technique commonly used in domain adaptation methods. This technique ensures that the domain classifier effectively discerns real-spots from pseudo-spots during training, while the extracted features deceive the trained domain classifier during the training of the source classifier. Notably, augmented training data are generated by mixing spots randomly selected from the spatially refined real-ST data to amend for the significant disparity in data volume between real-spots and pseudo-spots. Thus, the source classifier, trained exclusively with labelled pseudo-ST, demonstrates superior performance in estimating cell-type compositions from real-ST data.

To assess the efficacy of our proposed deconvolution method, LETSmix is applied to three representative public real ST datasets (Supplementary Table S3). These datasets encompass 12 ST samples of human brain cortex (DLPFC) data, 2 ST samples of human pancreatic ductal adenocarcinoma (PDAC) data, and 3 ST samples of mouse liver (Liver) data. All ST samples in the DLPFC dataset share a common reference scRNA-seq dataset, while each PDAC ST sample is paired with a matched reference scRNA-seq dataset, supplemented by an external scRNA-seq dataset. The Liver dataset comprises 3 distinct scRNA-seq datasets with different sequencing protocols. The performance of LETSmix is benchmarked against other state-of-the-art methods (Supplementary Table S4), including CellDART, SpaDecon, Cell2location, and CARD, through qualitative heuristic inspection and quantitative evaluation using metrics such as Area Under the Curve (AUC), EnRichment (ER), and Jensen-Shannon Divergence (JSD). Among these metrics, AUC and ER evaluate the enrichment of regionally restricted cell types within target areas, while JSD measures the consistency of the overall cell-type proportions between the entire ST tissue region and the matched reference scRNA-seq dataset.

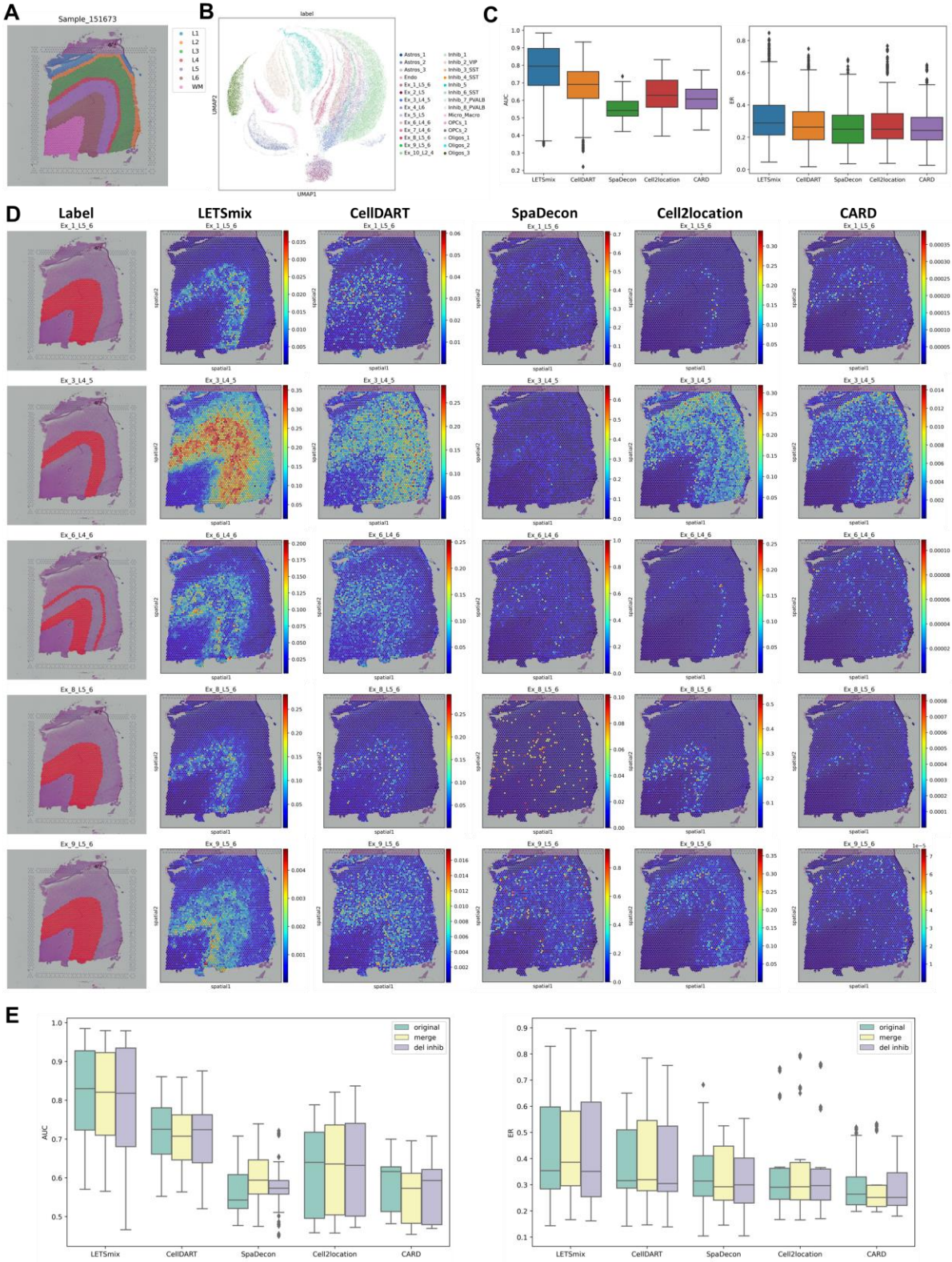


Figure 2. Application to the human brain cortex 10X Visium dataset. (**A**) Layer annotations of the ST sample named "151673" in the DLPFC dataset. (**B**) UMAP representation of the reference scRNA-seq dataset. (**C**) Box plots display the calculated AUC and ER for estimated cell type distribution in all 12 ST samples. (**D**) Estimated proportion heatmaps of 5 layer-

specific excitatory neurons by 5 models and the ground truth layer annotations. (**E**) Box plots show the calculated AUC and ER for the estimated cell type distribution in the "151673" ST sample. "original" includes all 28 cell types in the scRNA-seq dataset. "merge" means cell subtypes are merged before model training except the 10 excitatory neurons. "del Inhib" means inhibitory neurons are deleted from the scRNA-seq dataset. Each box plot in (**C**) and (**E**) comprises 50 datapoints, representing scores for 10 layer-specific cell types in 5 repeated experiments. Each box plot ranges from the third and first quartiles with the median as the horizontal line, while whiskers represent 1.5 times the interquartile range from the lower and upper bounds of the box.

Application to the human brain cortex 10X Visium dataset

We assess the performance of LETSmix using a 10X Visium dataset derived from postmortem neurotypical human DLPFC tissues (35). In this dataset, 12 ST samples exhibit clear structural stratification, ranging from L1 to L6 and WM (white matter) annotated by pathologists (Figure 2A, Supplementary Figure S2). A shared reference scRNA-seq dataset (37) is applied to train each deconvolution method for benchmarking, as visualized by the UMAP representation in Figure 2B. Among the 28 annotated cell types, the spatial mapping results of 10 layer-specific excitatory neurons are used to quantify the deconvolution performance of different tools.

Figure 2C presents the box plots of AUC and ER values achieved by different models on all 12 ST samples. The ranks of all these models remain consistent across the two metrics, with LETSmix consistently holding the top position. Due to the higher sensitivity of the AUC metric compared to ER, the evaluation results based on the AUC metric exhibited more pronounced differences in cell-type deconvolution performances across different models. In Figure 2D, we visualize the spatial distribution heatmaps of 5 layer-specific cell types estimated by each model trained exclusively on the "151673" ST sample. Compared to layer annotations, estimations of excitatory neurons from LETSmix are more accurate and coherent than those of other models. All excitatory neuron cell types demonstrate reasonable regionally restricted patterns. For example, the distinct gap between layer 4 and 6 can be clearly observed in the predictions of LETSmix for Ex_6_L4_6 cells. In contrast, the estimation results from other models are either excessively sparse or entirely fail to identify this cell

type. Additionally, only LETSmix demonstrates the ability to correctly predict significantly more Ex_3_L4_5 cells within layer 4 and 5. Compared to CellDART, LETSmix exhibits a more continuous spatial distribution and fewer false positive results, crediting to the utilization of information from spatial context. SpaDecon struggles to accurately predict the distribution of these cell types, mainly due to its neglection of the domain differences between scRNA-seq and ST data during the modelling process. The same issue is also present in the CARD model. Although Cell2location also considers the domain shifts between data from the two sequencing technologies through traditional statistical probabilistic approaches, its performance falls short compared to the deep learning-based domain adaptation method employed in this study. Supplementary Figure S3 provides results of all cell types estimated by LETSmix. Regarding other non-neuronal cells, LETSmix predicts that astrocytes are primarily distributed in layer 1 and 6, while oligodendrocytes are mainly located in the white matter region, which is consistent with findings from other biomedical studies (33), demonstrating the credibility and reliability of predictions made by LETSmix.

Furthermore, LETSmix and other models are tested under three different conditions on the "151673" ST sample. The AUC and ER values are calculated for each excitatory neuron type in 5 repeated experiments. As shown in Figure 2E, the scRNA-seq dataset is used with the original 28 cell types for model training in the "original" condition. These 28 cell types contain several cell subtypes. In the "merge" condition, these subtypes are merged into a broader category before training the models. For example, Astros_1, Astros_2, and Astros_3 are merged into the Astros cell type. Only the 10 layer-specific excitatory neuronal cell subtypes used for metric calculations are not merged, resulting in 16 cell types in total. In the "del Inhib" condition, all cells belonging to the Inhibitory subtype are removed from the scRNA-seq data before model training, which accounts for approximately 20% of the entire dataset. For both metrics and three conditions, LETSmix always achieves the highest scores, evidently outperforms other models, in accord with the visual inspection in Figure 2D. Additionally, note that there may exist certain domain variances among the sequenced data even when originating from the same tissue and using the same technology for spatial transcriptomics, which is commonly referred to as "batch effects" in other related works (19, 23). This phenomenon can also be observed in our experiment results. Specifically, we find that the performance of LETSmix trained exclusively on the "151673" sample is 0.01 higher

than that shown in Fig. 2c, which is achieved by using all 12 ST samples as the target domain for model training. Therefore, to obtain more accurate cell-type deconvolution results, we recommend training the model separately for each ST sample. Nonetheless, even when using all the ST samples simultaneously, LETSmix still exhibits significantly better deconvolution performance compared to the other four models.

To validate the effectiveness of LETSmix, we evaluate its performance on the "151673" DLPFC sample in the absence of certain designed components within its workflow (Supplementary Table S4). The key determining factor affecting model performance is the application of the domain adaptation training strategy. Without such a technique, the average AUC value achieved by LETSmix decreases drastically by 0.22. This underscores the essential consideration of domain differences between scRNA-seq and ST data. With the aid of information from layer annotations and spot expression similarity to build the LETS filter, the average AUC value increases by 0.008, indicating that the constructed adjacency matrix in LETSmix is able to more accurately depict inherent correlations among spots in ST. The scaling factor l used in constructing the adjacency matrix for SpaDecon is a scalar, controlling only the general degree of local smoothing for ST data, which introduces variations in the utilization of spatial contextual information among different spots. By transforming it to a vectorized factor, our method effectively eliminates such variations and achieves more precise and tailored scaling for each spot, leading to a more stable spatial refinement and a notable 0.017 increasement in the average AUC value of the deconvolution results. When completely neglecting interactions among spots, the average AUC value significantly decreases by approximately 0.1. This once again underscores the critical role of spatial context information. Note that since AUC and ER both measure whether the estimated results are concentrated in the target regions, and the former is more sensitive and exhibits greater variability based on previous findings, only the AUC metric was applied in this sub-section of the experiment. If the ER metric is used instead, it is challenging to clearly reflect the effects of different components in the LETSmix model.

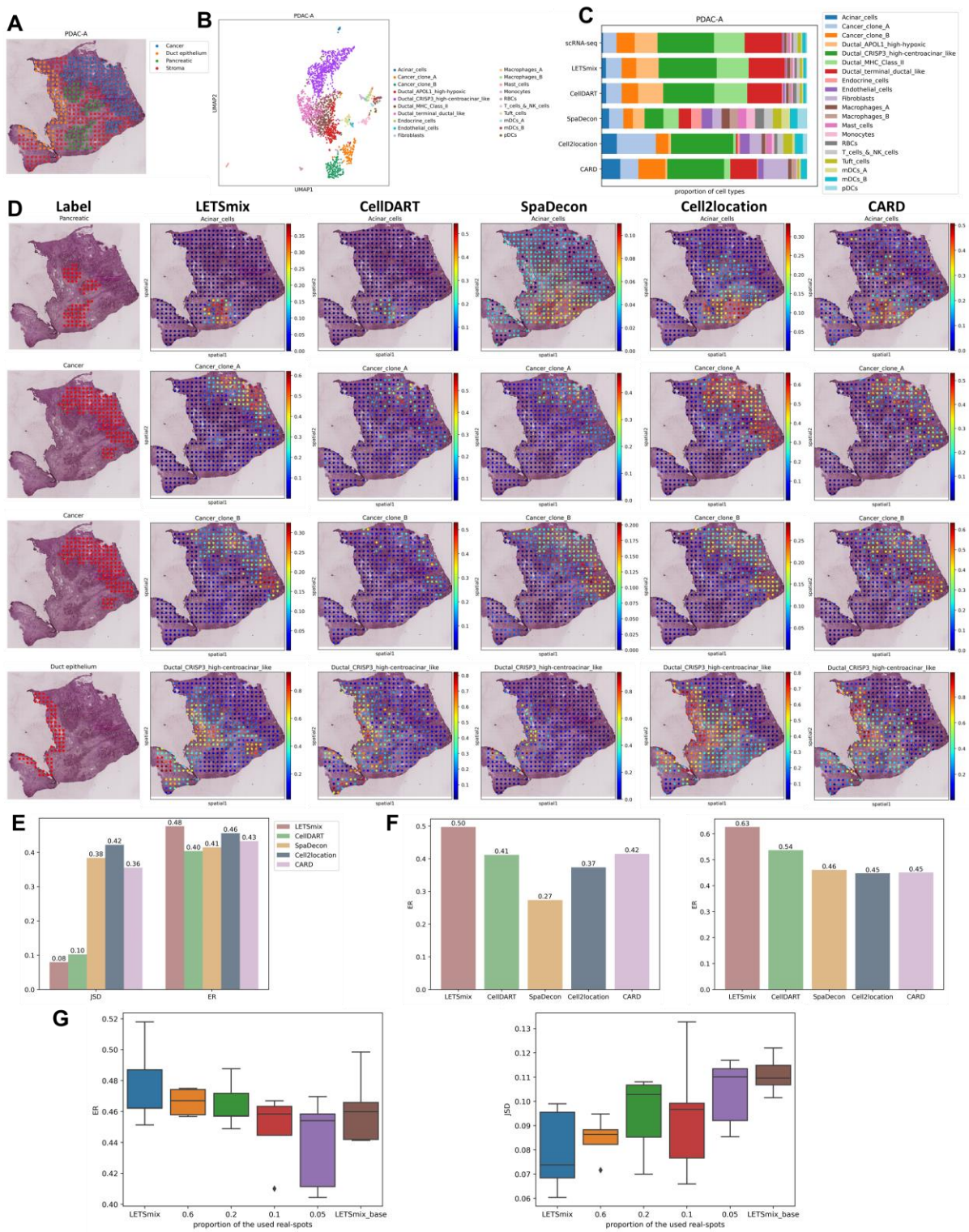


Figure 3. Application to the pancreatic ductal adenocarcinoma ST dataset. (**A**) Region annotations of the PDAC-A ST sample. (**B**) UMAP representation of the reference PDAC-A scRNA-seq dataset. (**C**) Stacked bar plots show the overall cell-type compositions in the PDAC-A ST sample estimated by each model using the paired PDAC-A scRNA-seq dataset. Ground truth is shown in the first row (denoted as "scRNA-seq"). The predicted proportion of

each cell type is the average value in 5 repeated experiments. (**D**) Estimated proportion heatmaps of 4 regionally restricted cell types by each model trained with matched PDAC-A ST and scRNA-seq data, and the ground truth layer annotations. (**E**) Model comparisons with matched ST and scRNA-seq data from PDAC-A. JSD and ER metrics are calculated using prior knowledge of cell-type compositions and localizations, respectively, in the PDAC-A tissue. (**F**) Model comparisons through the ER metric evaluated in the PDAC-A tissue. Left: models are trained with unmatched scRNA-seq data from PDAC-B. Right: models are trained with unmatched scRNA-seq data from PDAC-Peng. (**G**) The performance of LETSmix using varying proportions of available ST data for the mixup-augmented domain adaptation training. The model is tested on the PDAC-A ST sample through ER and JSD metrics, and is trained with matched scRNA-seq data. "LETSmix_base" represents scores achieved by the proposed method without the mixup procedure. Each box plot ranges from the third and first quartiles with the median as the horizontal line, while whiskers represent 1.5 times the interquartile range from the lower and upper bounds of the box.

Application to the pancreatic ductal adenocarcinoma ST dataset

The second dataset used for evaluation originates from cancerous tissues in human PDAC patients. Here, ST and the paired scRNA-seq data are collected following different protocols compared to those in the previously analyzed DLPFC dataset. We first apply LETSmix to a ST sample denoted as PDAC-A using the paired scRNA-seq data for model training. The ST sample is delineated into four distinct regions by pathological experts (Figure 3A). Among annotated cell types in scRNA-seq (Figure 3B), acinar, cancer, and ductal cells are expected to be mainly located within specific regions. Since both ST and scRNA-seq data in PDAC-A are collected from the same individual and the same tissue region, the average proportion of each cell type of all spots from ST is mirrored by the cell-type composition in scRNA-seq data. As visualized in Figure 3C, LETSmix and CellDART most accurately estimate the overall cell-type proportions in ST. Due to the complete neglection of domain differences, the input spot data exhibits significant distances from every clustering center in the feature space of SpaDecon trained on scRNA-seq data, leading to a similar estimated proportion for each cell type. It is worth noting that attempts were made to employ a clustering algorithm in LETSmix similar to the one used in SpaDecon. However, evaluation results obtained from the DLPFC dataset were notably inferior to the LETSmix model with currently applied classifier. Hence, it

is believed that the source classifier structure composed of fully connected layers used by LETSmix and CellDART is more suitable for cell-type deconvolution tasks compared to the supervised clustering training strategy employed by SpaDecon. This preference may stem from the capacity of the classifier to capture intricate data relationships, facilitating more effective differentiation among cell types, while clustering may not sufficiently capture the nuanced patterns required for precise cell-type identification.

Fig 3D presents the spatial distribution pattern of cell types inferred by each model. Credited to the effective utilization of comprehensive spatial context information, acinar cells estimated by LETSmix mainly distribute in the lower half of the pancreatic area, closely aligning with the manual annotation. In contrast, other models exhibited some problems in inferring this cell type: some were excessively sparse (such as CellDART), while others generated numerous false-positive results, extending the predictions beyond the pancreatic area into stroma and cancer regions (as seen in SpaDecon and Cell2location). Cancer clone A and cancer clone B cell types are expected to primarily distribute within the cancer region. The distinguishing marker genes for these two cell types are TM4SF1 and S100A4, respectively (Supplementary Figure S4A). TM4SF1 is significantly associated with tumor migration and invasion (42), indicating that areas with high TM4SF1 expression may represent late-stage cancer regions with metastatic potential. On the contrary, S100A4 serves as an early prognostic marker for pancreatic cancer (43). Upon examination, the upper half of the cancer region in PDAC-A tissue shows higher TM4SF1 expression, while the lower half exhibits higher S100A4 expression (Supplementary Figure S4B). Thus, it is inferred that the upper region corresponds to late-stage cancer likely populated by cancer clone A cells, while the lower region represents early-stage cancer, possibly populated by cancer clone B cells. Among results obtained from the five deconvolution models, LETSmix is able to most accurately identify such nuanced differences in the spatial distribution between the two cancer cell types. Although Cell2location generally delineated the cancer regions accurately, it failed to capture the relationship between these two cell types, incorrectly predicting a significant presence of Cancer clone A cells in the lower region. LETSmix also reasonably inferred distributions of other cell types in the PDAC-A tissue, consistent with the corresponding marker genes (Supplementary Figure S4B and S5). For example, spots enriched with TFF3, VIM, and CD74 marker genes are also estimated to have high

proportions of ductal terminal, endothelial, and mDCs cells, respectively. Based on the test results under two quantification metrics in Figure 3E, both the lowest JSD value and the highest ER value are achieved by LETSmix, showing that our proposed method is not only capable of mapping cells to their expected locations in ST, but also precisely estimating the proportions of various cell types. Similar to LETSmix, the CellDART model, also employing domain adaptation techniques, evidently outperforms other methods in the JSD metric evaluation, demonstrating the effectiveness of the applied adversarial training strategy in mitigating domain shifts between the two datasets. Despite that Cell2location performs well in predicting the enrichment of various cell types within specific regions, it ranks the poorest in the JSD metric, indicating its inability to accurately estimate the specific proportions of various cell types within each spot. This limitation was also reflected in the previous visualization about the prediction results of two cancer subtypes.

Next, the performance of different models is evaluated on scRNA-seq and ST data both collected from another tissue region, denoted as PDAC-B (Supplementary Figure S6). Distributions of cancer, ductal centroacinar, and RBCs cells predicted by each method are compared in Supplementary Figure S6C. Proportion heatmaps of the remaining cell types predicted by LETSmix are shown in Supplementary Figure S7, and Supplementary Figure S8 displays the distribution patterns of their corresponding marker genes for reference. In comparison to the region annotations of the PDAC-B tissue shown in Supplementary Figure S6A, all models primarily predicted the distribution of cancer clone A cells within the cancer region. However, CellDART, Cell2location, and CARD incorrectly predicted considerable cancer cells in the interstitium area, meanwhile SpaDecon estimated a very low content of cancer cells within the target region. Furthermore, the distribution of ductal centroacinar cells predicted by the SpaDecon model diffused from the ductal area to almost the entire tissue. The presence of RBCs should be minimal or absent in the pancreatic tissue, as PDAC tumors often compress and disrupt blood vessels, leading to reduced blood flow and impaired vascular function. This is also reflected in the stacked barplot of scRNA-seq shown in Supplementary Figure S6D. However, except for LETSmix and CellDART, all other models estimated a substantial presence of RBCs in this tissue, which reaffirms the high accuracy of LETSmix in predicting the proportions of various cell types. As shown in Supplementary Figure S6E, similar to previous results on PDAC-A, although Cell2location performed well

under the ER marker, its JSD value was significantly higher than that of other models. On the contrary, LETSmix achieves satisfying performance under both metrics.

To further investigate the role of domain adaptation in the LETSmix model, we conducted additional tests in two scenarios of data mismatch. Specifically, when assessing cell-type deconvolution performance on the ST data from PDAC-A, models are trained using reference scRNA-seq data from PDAC-B (Supplementary Figure S9) and an external dataset denoted as PDAC-Peng (Supplementary Figure S10), respectively. The PDAC-Peng dataset is collected from 24 primary PDAC tumors and 11 normal pancreas. Between the two ductal subtypes shown in Supplementary Figure S10A, ductal cell type 1 was identified as non-malignant while the other as malignant by previous researches (39), suggesting that ductal cell type 2 may infiltrate into the cancer region. In results obtained using PDAC-B scRNA-seq data (Supplementary Figure S9C), it can be observed that predictions made by LETSmix maintain high consistency with previously achieved results using the matched PDAC-A scRNA-seq dataset. However, other models exhibit significant changes in the predicted distributions compared to the previous outcomes. Specifically, acinar and ductal centroacinar cells estimated by other models exhibited severe diffusion into non-target regions. Since scRNA-seq data from PDAC-B lack cells of the cancer clone B type, predictions for cancer clone A cells inferred by LETSmix appeared to be a fusion of both cancer cell types. Furthermore, except for LETSmix and Cell2location, other models mistakenly inferred a high presence of RBCs in this tissue. Visual inspections in Supplementary Figure S10C using data from PDAC-Peng further underscore that consistent prediction for acinar cells is made by LETSmix. Our proposed method also correctly estimates distribution patterns of the two ductal subtypes, where non-malignant ductal cell type 1 mainly enriched in the ductal region and the malignant ductal cell type 2 distributed across ductal and cancer regions. In evaluations based on the ER metric shown in Figure 3F, the superiority in the performance of LETSmix against other methods becomes even more pronounced than results achieved in Figure 3E, which substantiates its stability in cell-type deconvolution when faced with considerable domain shifts between scRNA-seq and ST data, showcasing the ability to mitigate the need for alignment between these two data sources.

Given the pronounced sample imbalance between the source and target domains in the PDAC dataset, with only 428 spots available in ST samples from PDAC-A and 224 spots from

PDAC-B, experiments are conducted to evaluate the specific impact of the mixup process introduced in LETSmix using ST and scRNA-seq data both from PDAC-A. As the number of available spots in the ST data gradually decreased, the performance of LETSmix under ER and JSD metrics is illustrated in Figure 3G. The model without performing mixup operations on the spatially refined real-ST data is referred to as LETSmix_base. Combining the evaluation results of both metrics, LETSmix achieves comparable deconvolution performance to LETSmix_base when using only 5% of the spots from PDAC-A, i.e., 21 real-spots data, for the domain adaptation training. This strongly validates the critical role of mixup technique introduced in our designed LETSmix workflow, enabling the model to effectively address scenarios with limited sample availability.

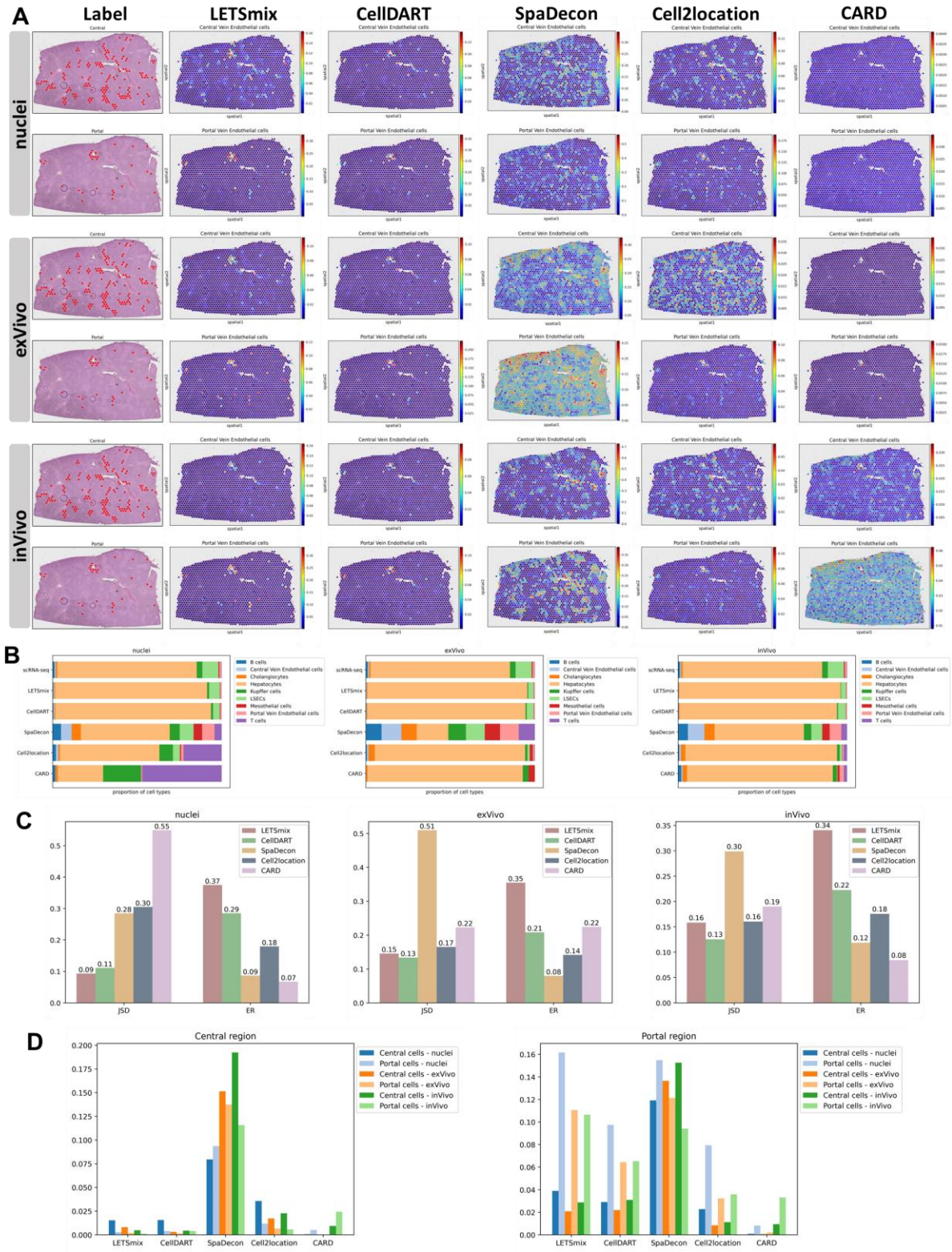


Figure 4. Application to the 10x Visium healthy mouse liver dataset. (**A**) Estimated proportion heatmaps of central vein and portal vein endothelial cells in the JBO001 Visium slice by 5 models and the ground truth layer annotations, using three scRNA-seq, respectively, as the reference datasets. (**B**) Stacked bar plots show the overall cell type compositions estimated

by each model and the ground truth. The predicted proportion of each cell type is the average value in three Visium slices in 5 repeated experiments. (**C**) Model comparisons through JSD and ER metrics calculated using prior knowledge of cell-type compositions and localizations, respectively, in the mouse liver tissues. Each bar represents the average value of the involved cell types in three Visium slices in 5 repeated experiments. (**D**) Comparisons of the average proportions of two regionally restricted cell types in the target area estimated by each model. Each bar represents the average value of the involved cell type in three Visium slices and in 5 repeated experiments.

Application to the healthy mouse liver 10x Visium dataset

Lastly, we employed the five deconvolution models to analyze three Visium slices of healthy mouse liver tissues (Supplementary Figure S11A). Three scRNA-seq datasets obtained with different experimental protocols are provided for the joint analysis, denoted as nuclei, exVivo, and inVivo, respectively (Supplementary Figure S11B). Similar to experiments carried on the PDAC dataset, the ER metric is implemented to measure the enrichment of portal vein and central vein ECs within portal and central regions, respectively. Meanwhile, JSD values are calculated for the estimated overall cell-type composition in ST samples based on that in the matched nuclei scRNA-seq dataset.

Figure 4A illustrates the estimation results of two regionally restricted cell types on the JBO001 Visium slice using three scRNA-seq datasets, respectively, to train each model. From the shown proportion heatmaps, it can be observed that LETSmix consistently provides the most accurate predictions among the compared methods for the spatial distribution patterns of these two cell types. However, it is also acknowledged that the spatial distribution of central vein ECs estimated by LETSmix shows certain discrepancies compared to the annotated central regions, and there is a tendency to misidentify substantial central vein ECs in the portal area. Similarly, predictions for this cell type made by other methods also exhibit similar issues, potentially with more pronounced discrepancies. This difficulty may be attributed to the scarcity of central vein ECs within the applied scRNA-seq datasets (Figure 4B), limiting the ability of each model to sufficiently capture the characteristics of this cell type. Nevertheless, upon closer inspection of predicted central vein ECs by LETSmix (Supplementary Figure S10C), it can be observed that they still maintain a high similarity to

the manually annotated central region, albeit with lower estimated proportions. Moreover, estimations for portal vein ECs by LETSmix are almost in perfect agreement with region annotations. LETSmix consistently maintains high consistency in its estimation results when trained with the three different scRNA-seq datasets. Cell2location also produces relatively decent estimation results for central vein ECs when trained with nuclei and inVivo scRNA-seq datasets, showing certain correlation with the annotated central region. However, it predicts numerous false positive results for both central vein ECs and portal vein ECs. Additionally, when using exVivo scRNA-seq as the reference dataset, Cell2location exhibits a significant decrease in deconvolution performance, indicating inferior stability compared to LETSmix. Although CellDART also produces stable prediction results with different scRNA-seq datasets, it estimates a relatively lower content for both types of ECs and does not accurately identify their locations compared to LETSmix. SpaDecon falsely predicts the occurrence of the two endothelial cell types across almost the entire tissue region, especially when trained with exVivo scRNA-seq data. On the contrary, the results of CARD hardly show the presence of these two cell types. Despite that estimation results for the two cell types made by CARD trained on the inVivo dataset also distribute throughout the entire tissue region, closer inspection reveals that the predicted two cell types only account for very small proportions, with the upper limit of the colorbar much lower than 0.1. Furthermore, results given by SpaDecon and CARD exhibit significant differences when trained with different scRNA-seq datasets.

Quantitative Evaluations further confirm the exceptional performance of LETSmix compared to other controlled deconvolution methods. Figure 4B visualizes the overall cell-type proportions across three ST Visium slices estimated by different models, and Figure 4C presents their JSD and ER scores. Under assessments of the JSD metric, LETSmix and CellDART are the most accurate predictors for the general proportions of different cell types, where CellDART achieves slightly lower JSD values using exVivo and inVivo scRNA-seq data. However, LETSmix significantly outperforms CellDART in terms of the ER metric and maintains the highest accuracy in identifying the expected locations of each cell type across three scRNA-seq datasets. SpaDecon provides relatively uniform estimations for all cell types as shown in Figure 4B, with only hepatocytes being noticeably more abundant than other cell types, which aligns with experiment results observed in the PDAC dataset. On the other

hand, when trained with the nuclei scRNA-seq dataset, Cell2location and CARD tend to predict an excessive amount of T cells. Although this issue is alleviated in the exVivo and inVivo results, where Cell2location and CARD achieve improved deconvolution performances, they are still significantly behind LETSmix. In fact, their performance on the PDAC dataset also surpasses that on the DLPFC dataset, with the former utilizing single-cell RNA-seq data and the latter utilizing single-nucleus RNA-seq data. This suggests a more pronounced domain shift between single-nucleus RNA-seq and ST data compared to single-cell RNA-seq data. Conversely, the performance of LETSmix using nuclei scRNA-seq dataset was even slightly better than using exVivo and inVivo datasets, as indicated by the lower JSD value and the higher ER value. A similar trend can be observed in the performance of CellDART, which also applies domain adaptation techniques. This implies that confounding information unrelated to the platform effect between ST and scRNA-seq data may be inadvertently introduced into features learned by the domain classifier when the degree of domain shift is inconspicuous. This, in turn, could impede the learning process of the source classifier.

Additionally, we investigate the differences between estimated proportions of the two endothelial cell types within central and portal regions, respectively, as shown in Figure 4D. In the central region, irrespective of the scRNA-seq dataset utilized, only LETSmix and Cell2location are able to accurately identify the quantity relationship between the two cell types, with central vein ECs significantly outnumbering portal vein ECs. However, Cell2location also incorrectly estimates substantial portal vein ECs. CellDART achieves desirable results only when utilizing the nuclei scRNA-seq dataset. Trained with other two scRNA-seq datasets, proportions of both cell types predicted by CellDART are too scarce, and thus their quantity differences are less distinct. SpaDecon and CARD exhibit suboptimal performance, producing unreasonable results where the number of portal vein ECs exceeds that of central vein ECs, meanwhile SpaDecon excessively estimates proportions of the two cell types. In the portal region, LETSmix consistently predicts more portal vein ECs when trained with each scRNA-seq dataset. CellDART achieves similar results, but the predicted proportions of portal vein ECs are significantly lower than that of LETSmix, while proportions of central vein ECs remain the same. Although SpaDecon predicts a large number of portal vein ECs in this region, it also inaccurately predicts a high proportion of central vein ECs. Cell2location correctly identifies a higher content of portal vein ECs in this area, but the

disparity between portal and central vein ECs is less pronounced than in LETSmix. CARD predicts excessively low content for both cell types when trained with the exVivo scRNA-seq dataset, aligning with observations in Figure 4A.

To sum up, LETSmix predicts the spatial distribution of different cell types with higher accuracy when compared to CellDART, benefiting from the ability to utilize additional spatial context information in ST data. Although CARD also considers spatial correlations among spots in ST using their positional coordinates, the scattered distribution patterns of different regions within the Liver dataset make it challenging to accurately capture inherent correlations based solely on coordinate information. LETSmix, on the other hand, overcomes this limitation by leveraging region annotations to guide the construction of the LETS filter. Credited to the use of domain adaptation techniques, only LETSmix and CellDART maintain high consistency in their estimation results when trained with the three different scRNA-seq datasets. This confirms that the proposed LETSmix model is more versatile and effectively alleviates the requirement for a high degree of matching between ST and scRNA-seq data.

DISCUSSION

The identification of spatial distribution patterns for specific cell types plays a pivotal role in elucidating their positions, densities, and interactions within tissue structures, facilitating a comprehensive understanding of tissue complexity and pathological changes. Seq-based ST technologies measure average gene expressions within cell mixtures. Through cell-type deconvolution, the positions and relative proportions of different cell types can be delineated on a spatial level, contributing to a more nuanced comprehension of tissue structure and cellular interactions. Additionally, spatial positions of spots and histological image information provide visual cues for tissue structure and cell distribution, enabling researchers to correlate ST data with specific cell types or structural features. This correlation aids in identifying differences and heterogeneity in cell types across different tissue regions, providing crucial insights for in-depth analysis. Furthermore, due to variations in data processing, detection sensitivity, technical specificity, cell handling, and sample preparation, certain domain shifts exist between ST and reference scRNA-seq data, which may impede the joint analysis of cell-type deconvolution. In this study, we introduce LETSmix, a deep

learning-based method trained on pseudo-spots synthesized from reference scRNA-seq data, and real-spots augmented by mixing spots from ST data. LETSmix effectively utilizes spatial context information to construct a LEST filter, which enhances the continuity of the spatial distribution in deconvolution results and reduces noise in raw ST data. Meanwhile, LETSmix employs adversarial domain adaptation techniques to facilitate a seamless transition of robust deconvolution capabilities trained on simulated pseudo-ST to real ST data, enhancing the generalizability of LETSmix across different domains.

LETSmix excels in constructing precise spatial maps of cell-type composition for ST samples. Evaluated across three datasets from distinct tissues, LETSmix notably outperforms other advanced deconvolution methods (CellDART, SpaDecon, Cell2location, and CARD) through both visual inspections and quantitative analyses leveraging prior knowledge of general cell-type locations and compositions. Beyond its effective deconvolution capability, LETSmix also boasts fast computational speed. Prior to model training, LETSmix conducts highly variable gene selection, significantly reducing the dimensionality of input expression data. In the DLPFC dataset, a shared reference scRNA-seq dataset is used to deconvolve 12 ST samples. Marked domain differences exist between scRNA-seq and ST data, and a high density of spots in ST indicates considerable structural features and noticeable stratification. These characteristics allow advantages of LETSmix to be fully manifested. In visualizations of layer-specific cell-type proportion heatmaps, predictions made by LETSmix closely align with layer annotations. Ablation and hyperparameter analyses conducted on this dataset strongly demonstrate the benefits of the employed domain adaptation technique and the utilized spatial context information. In the PDAC dataset, performances of different models are evaluated using matched and unmatched scRNA-seq data as references to deconvolve two ST samples. Although fewer spots are available in this dataset and the domain shifts are less significant in the matched scenario, LETSmix still exhibits a nonnegligible performance advantage by mixing spots to effectively augment data samples in the real-ST domain. As the domain shift intensifies with unmatched reference scRNA-seq and ST data, LETSmix continues to reliably estimate cell-type distribution patterns, showcasing robustness. Finally, the performance of LETSmix was assessed on mouse liver tissues using scRNA-seq data with different digestion protocols. Despite a weaker hierarchical structure among different functional regions, the incorporation of region annotation information enables LETSmix to

accurately capture inherent spatial correlations in ST data. Regardless of the scRNA-seq dataset used for model training, LETSmix consistently demonstrates the ability to accurately determine the spatial distribution patterns of different cell types.

As a versatile plug-in module aiming at optimizing the quality of ST data by capturing inherent spatial correlations, the designed LETS filter in this study can be easily transferred to other deconvolution models. Through local smoothing on adjacent spots with similar morphological features, their expression profiles exhibit higher similarity, ensuring the continuity of deconvolution results in space. With the ongoing advancement in ST clustering research (44, 45, 46), even in the absence of manually annotated regions for the ST dataset, we anticipate that advanced clustering models can be employed to guide the construction of the adjacency matrix in LETSmix. In order to address the lack of actual cell-type proportion labels in real-ST, LETSmix adopts the same pseudo-ST generation approach as in CellDART, where a fixed number of cells with random weights is selected from scRNA-seq data to synthesize the gene expression profile for each pseudo-spot. Although theoretically generating a sufficient number of pseudo-spots may simulate various combinations of cell types, there is room for improvement in this synthesis method. For instance, future work could explore alternative enhancement strategies, addressing issues such as class-unbalanced sample sizes, where characteristics of rare cell types might be overshadowed by the dominating cell types. Additionally, although the domain adaptation technique implemented in this work is augmented by mixup and thus is able to address the issue of unbalanced samples between source and target domains, it still relies on relatively conventional adversarial training strategies, which may exhibit limitations in certain application scenarios. When the degree of domain shifts between ST and the reference scRNA-seq data is minimal, posing a challenge for the domain discriminator in LETSmix to accurately distinguish between the source and target domains, the learned confusing features may negatively impact the source classifier, leading to a performance inferior to that observed in the absence of the domain adaptation technique. We look forward to discovering more suitable domain adaptation algorithms for ST analysis in the future.

In conclusion, LETSmix emerges as a valuable tool in the realm of spatial transcriptomics, providing enhanced capabilities for cell-type deconvolution. Its incorporation of spatial context information and effective domain adaptation techniques contribute to its ability to

accurately delineate spatial distribution patterns. Even though LETSmix has already demonstrated superior performance compared to other state-of-the-art models across multiple datasets, there is still room for further improvements. We anticipate that the suggested method has broad applications in comprehensively mapping tissue architecture across diverse biological contexts, aiding biomedical researchers in understanding cellular interactions, developmental processes, and pathological mechanisms within complex biological systems.

DATA AVAILABILITY

Three publicly available datasets were analysized in this study. Raw data can be obtained through the following websites or GEO accession numbers: (1) Human DLPFC (ST: http://research.libd.org/spatialLIBD/; scRNA-seq: GSE144136); (2) Pancreatic Ductal Adenocarcinoma (ST: GSM3036911, GSM3405534; scRNA-seq: GSE111672, https://download.cncb.ac.cn/gsa/CRA001160/) (3) Mouse Liver (ST: GSM5764 414, GSM5764415, GSM5764416; scRNA-seq: GSE192740). Additionally, we have curated and made available all the aforementioned datasets, which can be directly downloaded from the following website: https://zenodo.org/records/10687718.

An open-source implementation of the LETSmix algorithm in Python is available at https://github.com/ZhanYangen/LETSmix/, including codes for data preprocessing, network construction, model training, and results analysis.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online, including Supplementary Figures S1-11, Supplementary Tables S1-4, and Supplementary Algorithm S1.

ACKNOWLEDGEMENTS

Author contributions. Yangen Zhan: Conceptualization, Investigation, Data curation, Methodology, Software, Formal Analysis, Visualization, Writing – original draft. Yongbing

Zhang: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. Zheqi Hu: Data curation, Software, Validation. Yifeng Wang: Conceptualization, Writing – review & editing. Zirui Zhu: Data curation, Software, Validation. Sijing Du: Data curation. Xiu Li: Funding acquisition, Project administration, Supervision, Writing – review & editing.

FUNDING

This work was supported in part by the STI 2030-Major Projects [2021ZD0201404], in part by the National Natural Science Foundation of China [62031023, 62331011], in part by the Shenzhen Science and Technology Project [GXWD20220818170353009], and in part by the Fundamental Research Funds for the Central Universities [HIT.OCEF.2023050].

CONFLICT OF INTEREST

The authors declare no competing interests.

REFERENCES

1. Kolodziejczyk,A.A., Kim,J.K., Svensson,V., Marioni,J.C. and Teichmann,S.A. (2015) The technology and biology of single-cell RNA sequencing. *Mol. Cell*, **58**, 610-620.
2. Svensson,V., Vento-Tormo,R. and Teichmann,S.A. (2018) Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.*, **13**, 599-604.
3. Zhu,L., Yang,P., Zhao,Y., Zhuang,Z., Wang,Z., Song,R., Zhang,J., Liu,C., Gao,Q., Xu,Q. *et al.* (2020) Single-cell sequencing of peripheral mononuclear cells reveals distinct immune response landscapes of COVID-19 and influenza patients. *Immunity*, **53**, 685-696. e683.
4. Kharchenko,P.V. (2021) The triumphs and limitations of computational methods for scRNA-seq. *Nat. Methods*, **18**, 723-732.

5. Longo,S.K., Guo,M.G., Ji,A.L. and Khavari,P.A. (2021) Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.*, **22**, 627-644.
6. Rao,A., Barkley,D., França,G.S. and Yanai,I. (2021) Exploring tissue architecture using spatial transcriptomics. *Nature*, **596**, 211-220.
7. Asp,M., Bergenstråhle,J. and Lundeberg,J. (2020) Spatially resolved transcriptomes—next generation tools for tissue exploration. *Bioessays*, **42**, 1900221.
8. Chen,K.H., Boettiger,A.N., Moffitt,J.R., Wang,S. and Zhuang,X. (2015) Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*, **348**, aaa6090.
9. Eng,C.H.L., Lawson,M., Zhu,Q., Dries,R., Koulena,N., Takei,Y., Yun,J., Cronin,C., Karp,C., Yuan,G.C. *et al.* (2019) Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature*, **568**, 235-239.
10. Cho,C.S., Xi,J., Si,Y., Park,S.R., Hsu,J.E., Kim,M., Jun,G., Kang,H.M. and Lee,J.H. (2021) Microscopic examination of spatial transcriptome using Seq-Scope. *Cell*, **184**, 3559-3572. e3522.
11. Xia,C., Fan,J., Emanuel,G., Hao,J. and Zhuang,X. (2019) Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc. Natl. Acad. Sci.*, **116**, 19490-19499.
12. Ståhl,P.L., Salmén,F., Vickovic,S., Lundmark,A., Navarro,J.F., Magnusson,J., Giacomello,S., Asp,M., Westholm,J.O., Huss,M. *et al.* (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, **353**, 78-82.
13. Rodriques,S.G., Stickels,R.R., Goeva,A., Martin,C.A., Murray,E., Vanderburg,C.R., Welch,J., Chen,L.M., Chen,F. and Macosko,E.Z. (2019) Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, **363**, 1463-1467
14. Chen,A., Liao,S., Cheng,M., Ma,K., Wu,L., Lai,Y., Qiu,X., Yang,J., Xu,J., Hao,S. *et al.* (2022) Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, **185**, 1777-1792. e1721.
15. Liao,J., Lu,X., Shao,X., Zhu,L. and Fan,X. (2021) Uncovering an organ's molecular architecture at single-cell resolution by spatially resolved transcriptomics. *Trends Biotechnol.*, **39**, 43-58.

16. Li,B., Zhang,W., Guo,C., Xu,H., Li,L., Fang,M., Hu,Y., Zhang,X., Yao,X., Tang,M. *et al.* (2022) Benchmarking spatial and single-cell transcriptomics integration methods for transcript distribution prediction and cell type deconvolution. *Nat. Methods*, **19**, 662-670.
17. Elosua-Bayes,M., Nieto,P., Mereu,E., Gut,I. and Heyn,H. (2021) SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res.*, **49**, e50-e50.
18. Dong,R. and Yuan,G.C. (2021) SpatialDWLS: accurate deconvolution of spatial transcriptomic data. *Genome Biol.*, **22**, 145.
19. Cable,D.M., Murray,E., Zou,L.S., Goeva,A., Macosko,E.Z., Chen,F. and Irizarry,R.A. (2022) Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.*, **40**, 517-526.
20. Chen,J., Liu,W., Luo,T., Yu,Z., Jiang,M., Wen,J., Gupta,G.P., Giusti,P., Zhu,H., Yang,Y. *et al.* (2022) A comprehensive comparison on cell-type composition inference for spatial transcriptomics data. *Brief Bioinform.*, **23**, bbac245.
21. Sang-aram,C., Browaeys,R., Seurinck,R. and Saeys,Y. (2023) Spotless: a reproducible pipeline for benchmarking cell type deconvolution in spatial transcriptomics. eLife 12. *RP88431*.
22. Andersson,A., Bergenstråhle,J., Asp,M., Bergenstråhle,L., Jurek,A., Fernández Navarro,J. and Lundeberg,J. (2020) Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Commun. Biol.*, **3**, 565.
23. Kleshchevnikov,V., Shmatko,A., Dann,E., Aivazidis,A., King,H.W., Li,T., Elmentaite,R., Lomakin,A., Kedlian,V., Gayoso,A. (2022) *et al.* Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.*, **40**, 661-671.
24. Liu,Z., Wu,D., Zhai,W. and Ma,L. (2023) SONAR enables cell type deconvolution with spatially weighted Poisson-Gamma model for spatial transcriptomics. *Nat. Commun.*, **14**, 4727.
25. Ma,Y. and Zhou,X. (2022) Spatially informed cell-type deconvolution for spatial transcriptomics. *Nat. Biotechnol.*, **40**, 1349-1359.
26. Hu,J., Li,X., Coleman,K., Schroeder,A., Ma,N., Irwin,D.J., Lee,E.B., Shinohara,R.T. and Li,M. (2021) SpaGCN: Integrating gene expression, spatial location and histology to

identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods*, **18**, 1342-1351.

27. Xu,H., Fu,H., Long,Y., Ang,K.S., Sethi,R., Chong,K., Li,M., Uddamvathanak,R., Lee,H.K., Ling,J. *et al.* (2024) Unsupervised spatially embedded deep representation of spatial transcriptomics. *Genome Med.*, 16, 12.
28. Lopez,R., Nazaret,A., Langevin,M., Samaran,J., Regier,J., Jordan,M.I. and Yosef,N. (2019) A joint model of unpaired data from scRNA-seq and spatial transcriptomics for imputing missing gene expression measurements. arXiv doi: https://doi.org/10.48550/arXiv.1905.02269, 6 May 2019, preprint: not peer reviewed.
29. Zhou,X., Chai,H., Zeng,Y., Zhao,H. and Yang,Y. (2021) scAdapt: virtual adversarial domain adaptation network for single cell RNA-seq data classification across platforms and species. *Brief Bioinform.*, **22**, bbab281.
30. Song,Q. and Su,J. (2021) DSTG: deconvoluting spatial transcriptomics data through graph-based artificial intelligence. *Brief Bioinform.*, **22**, bbaa414.
31. Shen,R., Liu,L., Wu,Z., Zhang,Y., Yuan,Z., Guo,J., Yang,F., Zhang,C., Chen,B., Feng,W. *et al.* (2022) Spatial-ID: a cell typing method for spatially resolved transcriptomics via transfer learning and spatial embedding. *Nat. Commun.*, **13**, 7640.
32. Coleman,K., Hu,J., Schroeder,A., Lee,E.B. and Li,M. (2023) SpaDecon: cell-type deconvolution in spatial transcriptomics with semi-supervised learning. *Commun. Biol.*, **6**, 378.
33. Bae,S., Na,K.J., Koh,J., Lee,D.S., Choi,H. and Kim,Y.T. (2022) CellDART: cell type inference by domain adaptation of single-cell and spatial transcriptomic data. *Nucleic Acids Res.*, **50**, e57-e57.
34. Chen,J., Luo,T., Jiang,M., Liu,J., Gupta,G.P. and Li,Y. (2023) Cell composition inference and identification of layer-specific spatial transcriptional profiles with POLARIS. *Sci. Adv.*, **9**, eadd9818.
35. Maynard,K.R., Collado-Torres,L., Weber,L.M., Uytingco,C., Barry,B.K., Williams,S.R., Catallini,J.L., Tran,M.N., Besich,Z., Tippani,M. *et al.* (2021) Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.*, **24**, 425-436.

36. Hawrylycz,M.J., Lein,E.S., Guillozet-Bongaarts,A.L., Shen,E.H., Ng,L., Miller,J.A., Van De Lagemaat,L.N., Smith,K.A., Ebbert,A., Riley,Z.L. *et al.* (2012) An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, **489**, 391-399.
37. Nagy,C., Maitra,M., Tanti,A., Suderman,M., Théroux,J.F., Davoli,M.A., Perlman,K., Yerko,V., Wang,Y.C., Tripathy,S.J. *et al.* (2020) Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons. *Nat. Neurosci.*, **23**, 771-781.
38. Moncada,R., Barkley,D., Wagner,F., Chiodin,M., Devlin,J.C., Baron,M., Hajdu,C.H., Simeone,D.M. and Yanai,I. (2020) Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat. Biotechnol.*, **38**, 333-342.
39. Peng,J., Sun,B.F., Chen,C.Y., Zhou,J.Y., Chen,Y.S., Chen,H., Liu,L., Huang,D., Jiang,J., Cui,G.S. *et al.* (2019) Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res.*, **29**, 725-738.
40. Guilliams,M., Bonnardel,J., Haest,B., Vanderborght,B., Wagner,C., Remmerie,A., Bujko,A., Martens,L., Thoné,T., Browaeys,R. *et al.* (2022) Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches. *Cell*, **185**, 379-396. e338.
41. Bae,S., Choi,H. and Lee,D.S. (2023) spSeudoMap: cell type mapping of spatial transcriptomics using unmatched single-cell RNA-seq data. *Genome Med.*, **15**, 19.
42. Zheng,B., Ohuchida,K., Cui,L., Zhao,M., Shindo,K., Fujiwara,K., Manabe,T., Torata,N., Moriyama,T., Miyasaka,Y. *et al.* (2015) TM4SF1 as a prognostic marker of pancreatic ductal adenocarcinoma is involved in migration and invasion of cancer cells. *Int. J. Oncol.*, **47**, 490-498.
43. Fei,F., Qu,J., Zhang,M., Li,Y. and Zhang,S. (2017) S100A4 in cancer progression and metastasis: A systematic review. *Oncotarget*, **8**, 73219.
44. Hu,Y., Zhao,Y., Schunk,C.T., Ma,Y., Derr,T. and Zhou,X.M. (2023) ADEPT: Autoencoder with differentially expressed genes and imputation for robust spatial transcriptomics clustering. *Iscience*, **26**.
45. Li,Z., Chen,X., Zhang,X., Jiang,R. and Chen,S. (2023) Latent feature extraction with a prior-based self-attention framework for spatial transcriptomics. *Genome Res.*, **33**, 1757-1773.

46. Long,Y., Ang,K.S., Li,M., Chong,K.L.K., Sethi,R., Zhong,C., Xu,H., Ong,Z., Sachaphibulkij,K., Chen,A. *et al.* (2023) Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat. Commun.*, **14**, 1155.