communications biology

A Nature Portfolio journal

Article

https://doi.org/10.1038/s42003-026-09659-y

DANST enables cell-type deconvolution in spatial transcriptomics using deep domain adversarial neural networks

Xueqin Zhang[1] ✉, Zhichao Wu[1], Tianqi Wang[1] ✉, Yunlan Zhou[2], Weihong Ding[3], Huitong Zhu[1] & Qing Zhang[4]

Spatial transcriptomics is an emerging technology that can analyze gene expression profiles of tissues while preserving spatial location information. To restore cell type proportions from mixed gene expression data, here we present DANST, a deconvolution framework based on deep domain adversarial neural networks. By integrating single-cell RNA sequencing (scRNA-seq) with inferred spatial coordinates, we construct pseudo-spatial data. DANST utilizes a variational autoencoder to learn refined feature representations and introduces a domain adversarial architecture to align feature distributions between pseudo and real data, enabling accurate label transfer. Benchmarking on human and mouse datasets shows that DANST achieves superior deconvolution accuracy compared with existing methods. These findings highlight its effectiveness for tumor microenvironment analysis and potential clinical utility.

Recent advances in spatial transcriptomics (ST) have enabled the integration of cellular spatial information and gene expression, which is not possible with traditional non-spatial single-cell RNA sequencing technology (scRNA-seq). However, current technical limitations also prevent ST from achieving gene coverage comparable to scRNA-seq while achieving single-cell resolution. Specifically, the widely used 10x Visium platform can capture transcriptome data with a throughput comparable to scRNA-seq[1], but because it uses 55 μm capture spots, which are larger than typical cells (5–10 μm), the spatial transcriptome data obtained is at multi-cellular resolution[2,3]. The latest sequencing-based technologies, including Slide-seq[4], DBiT-seq[5], Stereo-seq[6], PIXEL-seq[7], and Seq-Scope[8], provide sub-cellular spatial resolution, but there are high deletion events, resulting in a very sparse gene expression matrix. At the same time, methods based on fluorescence in situ hybridization (FISH) can achieve subcellular resolution, but lack genome-wide gene coverage, and the latest seq-Fish technology is still limited to 10,000 genes[9]. Therefore, in order to analyze ST data from low-resolution capture technologies and infer the relative proportions of different cell types from mixed cell spots, conducting studies on cell type deconvolution is important for improving the biological interpretability of the data.

Currently, many cell type deconvolution algorithms have been developed to decompose expression into percentage contributions from different cells or cell types. RCTD[10] assumes that the count of each gene in each ST spot satisfies the Poisson distribution and uses maximum likelihood estimation to predict the proportion of cell types. SPOTlight[11] has developed a regression framework based on non-negative matrix factorization (NMF) to reveal the underlying structure and characteristics of the data and infer the contribution of each cell type by decomposing the original data matrix into a non-negative low-rank approximate matrix. Stereoscope[12] estimates the proportion of cell types using the maximum likelihood method based on the assumption that the gene counts in ST data follow a negative binomial distribution. Cell2location[13] constructs a hierarchical Bayesian-based framework to model and infer ST data by establishing a hierarchical model, providing an interpretable tool for revealing the spatial distribution and regulatory mechanisms of cell types. However, while many existing methods utilize scRNA-seq data with known cell type labels, they often do not fully exploit this information[14], which may limit their predictive accuracy.

Also, several recent approaches use scRNA-seq data with known cell type labels to construct pseudo ST data, thereby obtaining labeled reference data and converting unsupervised learning into semi-supervised or fully supervised learning. SPACEL[15] uses scRNA-seq data to simulate pseudo spots and combines a probabilistic Variational Autoencoder (VAE), a Multi-Layer Perceptron Feature Extractor (MLPF) and a Multi-Layer Perceptron Proportion Quantifier (MLPQ) to provide a more robust and accurate framework for estimating the proportion of cell types in real ST data. SD2[16] randomly combines cells from scRNA-seq data to synthesize

[1]School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China. [2]Department of Clinical Laboratory, Xinhua Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China. [3]Huashan Hospital Affiliated to Fudan University, Shanghai, China. [4]Shanghai Institute of Technology, Shanghai, China. ✉e-mail: zxq@ecust.edu.cn; tqwang743@163.com

pseudo ST spots, applies an auto-encoder to obtain low-dimensional nonlinear representations of real and pseudo ST spots, and constructs a neighbor graph to connect real and pseudo spots base on their feature representations. Then, inputs it into graph convolutional neural network to predict the cell type composition of real ST spots. Another component of the method involves constructing a mapping matrix to transfer cell type information from scRNA-seq data to ST data. GraphST[17] uses a contrastive learning mechanism without enhancement to train a mapping matrix from cells to spots, and maps scRNA-seq data to the ST space based on its learned features. SpatialcoGCN[18] obtains the feature representation corresponding to each cell type by summing the gene expression profiles of the same cell type in the scRNA-seq data, and uses the K nearest neighbor algorithm to establish a link graph between cell types and spots. A graph convolutional network is utilized to obtain a mapping matrix from spots to cell types, which is used to estimate the different cell type composition of each spot in the ST data. spaDecon[19] clusters scRNA-seq data to obtain the cluster center of each cell type, and calculates the distance from each spot in the ST data to the cluster center, using this as the weight to obtain a mapping matrix from spots to cell types. However, a common limitation of these methods is that they only partially address the distributional differences between scRNA-seq data and ST data[10], which may reduce the robustness of the models when applied across different data sources. To solve this problem, Tangram[20] introduces a batch integration method to narrow the technical differences between scRNA-seq and ST data. CellDART[21] and spSeudoMap[22] propose domain adaptation methods to improve the accuracy of overall prediction. Although these three cell type deconvolution methods take into account the differences in feature distribution, they make limited use of the spatial information in ST data, resulting inaccuracy of cell type identification and abundance estimation still insufficient[23]. Thus, effective integration of spatial information remains a key challenge in cell type deconvolution.

To overcome the above problems, here we show a general framework based on deep domain adversarial neural network (DANST). The main contributions of this study are as follows:

To enhance the accuracy of cell type deconvolution in ST data, we propose an effective pseudo ST data construction strategy that enables robust spatial-level integration between scRNA-seq and ST data. Specifically, scRNA-seq data with clear cell type labels are used to generate pseudo ST data, while spatial clustering is performed on real ST data to calculate the distance between each pseudo spot and the centers of clusters. These distances serve as weights to construct pseudo spatial coordinates. Compared with traditional approaches, this strategy introduces a distance-weighted mechanism that better exploits spatial information, allowing the constructed neighborhood graph to more effectively capture structural correspondences between the two data modalities and thereby improve overall deconvolution performance.

In order to solve the problem of the difference in feature distribution between scRNA-seq data and ST data, DANST introduces a domain adversarial architecture. The designed domain adversarial network contains a feature extractor, a domain discriminator, and a proportion predictor. The feature extractor and proportion predictor are trained to maximize the performance of the proportion prediction task, while the domain discriminator is trained to distinguish features from different data types. In this way, the network is able to learn features that are both useful for the task and generalizable to different data types, thereby effectively enhancing the domain adaptability of the overall model.

By benchmarking DANST on multiple synthetic and real datasets from mouse heart, mouse brain, and human breast cancer, our results demonstrate that DANST significantly outperforms existing methods, achieving up to an 18.3% improvement in average ranking scores. Furthermore, DANST accurately identifies the distribution of luminal cells within the invasive ductal carcinoma region of human breast cancer samples, providing biologically interpretable insights into the tumor microenvironment. These findings suggest that DANST is a robust tool for clinical spatial biology and cross-platform data integration, offering new insights into complex biological systems and the improvement of medical treatments.

Results

Overview of DANST

To address the challenge of cross-platform domain shifts in ST deconvolution, we developed DANST, a deep domain adversarial neural network. The workflow of DANST comprises five core components as illustrated in Fig. 1: (1) Data pre-processing and pseudo-spot construction; (2) Generation of pseudo coordinates; (3) Neighborhood graph construction and feature representation refinement; (4) Domain adversarial training, which aligns the feature distributions of reference and target datasets; and (5) Cell type proportion prediction. By integrating a variational auto-encoder (VAE) with an adversarial learning strategy, the model aims to minimize platform-specific noise and enhance the accuracy of spatial cell type mapping.

Benchmark experiment

This experiment is the benchmark experiment for cell type deconvolution task, which evaluates the accuracy of predicting cell type proportions. The experiment compares DANST with GraphST[17], CellDART[21], Tangram[20], SpaDecon[19], spSeudoMap[22], and SpatialcoGCN[18]. Here, GraphST, CellDART are graph convolution-based methods, Tangram, SpaDecon are mapping-based methods and spSeudoMap, SpatialcoGCN are recently proposed methods.

Refer to SpatialcoGCN, we leverage the artificial ST dataset with spatial location information, which is generated by integrating eight scRNA-seq datasets and nine real ST datasets from the same region or tissue type. This experiment selects the most representative datasets 5, 8, and 12 from different tissues. Among them, Dataset 5 is a mouse heart dataset generated using the 10x Visium platform. It is constructed from (i) an scRNA-seq dataset containing 10,146 cells across eight cell types (10x Visium, mouse heart), and (ii) a real ST dataset containing 1,835 spots (10x Visium, mouse heart). Dataset 8 is a mouse cortex dataset generated using seqFISH+ and Smart-seq. It contains 953 spots, each spot including 4–15 cells. It is constructed from (i) an scRNA-seq dataset containing 14,249 cells across 15 cell types (Smart-seq, mouse cortex), and (ii) a real ST dataset containing 524 spots (seqFISH +, mouse cortex). Dataset 12 is a mouse visual cortex dataset generated using seqFISH+ and Smart-seq. It is constructed from (i) an scRNA-seq dataset containing 13,823 cells across 15 cell types (Smart-seq, mouse cortex), and (ii) a real ST dataset containing 1,154 spots (seqFISH +, mouse cortex).

The experimental results of seven methods on three datasets are shown in Fig. 2A, C.

From the figures, it can be seen that compared with other methods, DANST achieves the highest average ranking scores on all datasets, which are 18.3%, 18.1%, and 13.4% higher than the sub-optimal method, respectively. Additionally, DANST records the highest average values in PCC, SSIM, and COSSIM, exceeding those of the inferior method by 11.3%, 20.8%, and 10.8%, respectively. It also demonstrates the lowest average values in RMSE and JSD, which are 28.2 and 22.7% lower than those of the second best method.

In general, DANST has a relatively good cell type deconvolution result on the artificially synthesized ST dataset. At the same time, for different datasets, the best method is always DANST, while the sub-optimal and worst methods are constantly changing, which also confirms that for different data sources, DANST is more robust than other methods.

The experiment on mouse brain anterior dataset

This experiment is conducted using a 10X Visium ST real dataset from the anterior part of the mouse brain. We first verify the ability of the DANST in elucidating the microscopic anatomical structures of complex tissues. By analyzing scRNA-seq data from the anterior end of the mouse brain, it can be concluded that this tissue contains 28 different cell types, each with a unique gene expression pattern. Using these scRNA-seq data, the DANST method performs cell type deconvolution on the ST data and reconstruct the proportion of cell types at each spot in the ST data. In order to more intuitively demonstrate the results of DANST cell type deconvolution, we

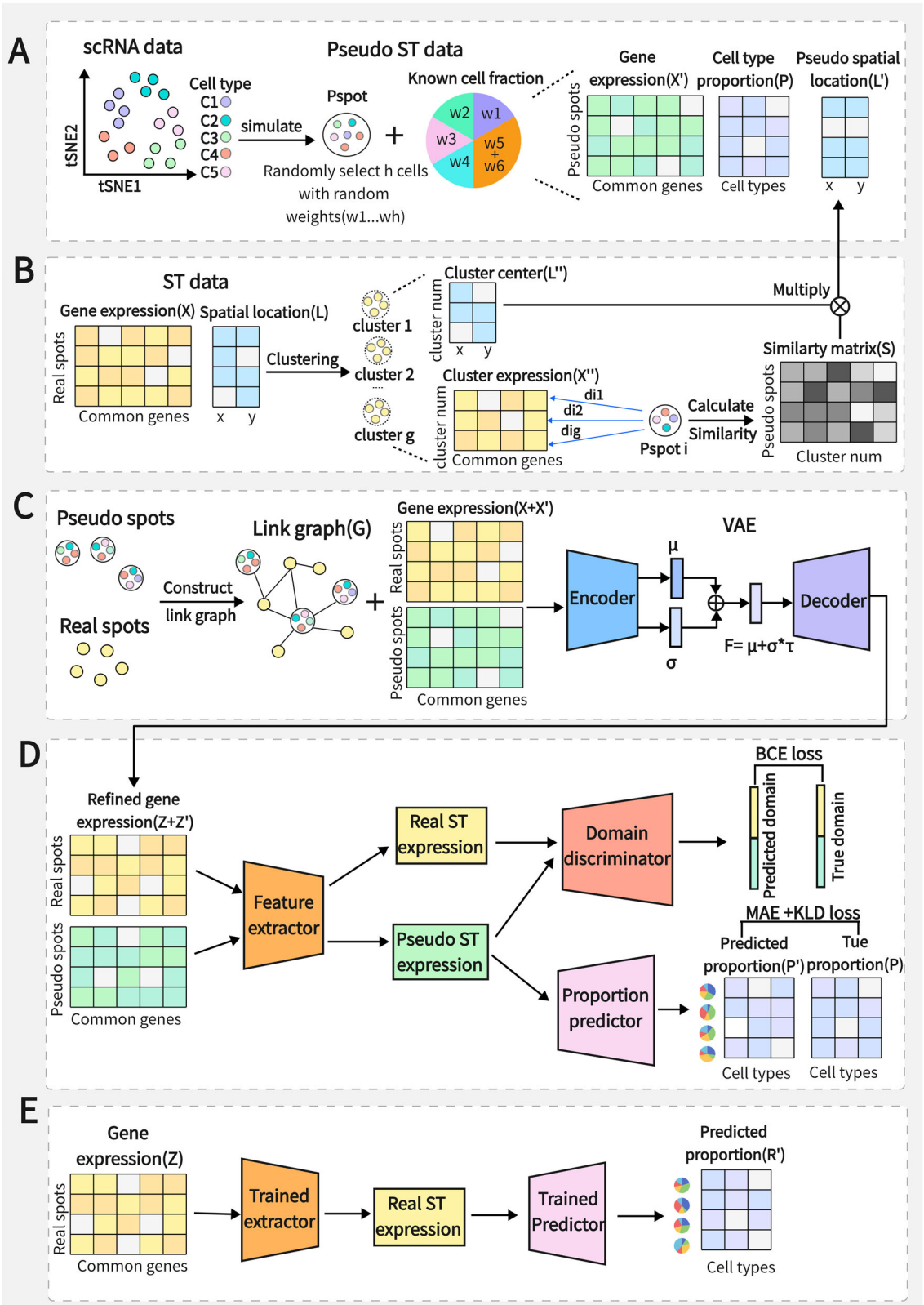

Fig. 1 | Overall framework of DANST. A Use scRNA-seq data to construct pseudo ST data with known cell type proportions, and perform data pre-processing. **B** Cluster the real ST data, use the distance between the pseudo spot and the center of each cluster as the weight, and calculate the pseudo coordinates of the pseudo ST data. **C** Design a variational auto-encoder module to denoise and optimize the features of the real ST data and pseudo ST data. **D** Design a domain adversarial module, introduce a feature extractor, a domain discriminator, and a proportion predictor. Then, construct a two-stage alternating training process to train the generative adversarial ability of the model. **E** Use the optimized feature extractor and proportion predictor to predict the cell type proportions of real ST data.

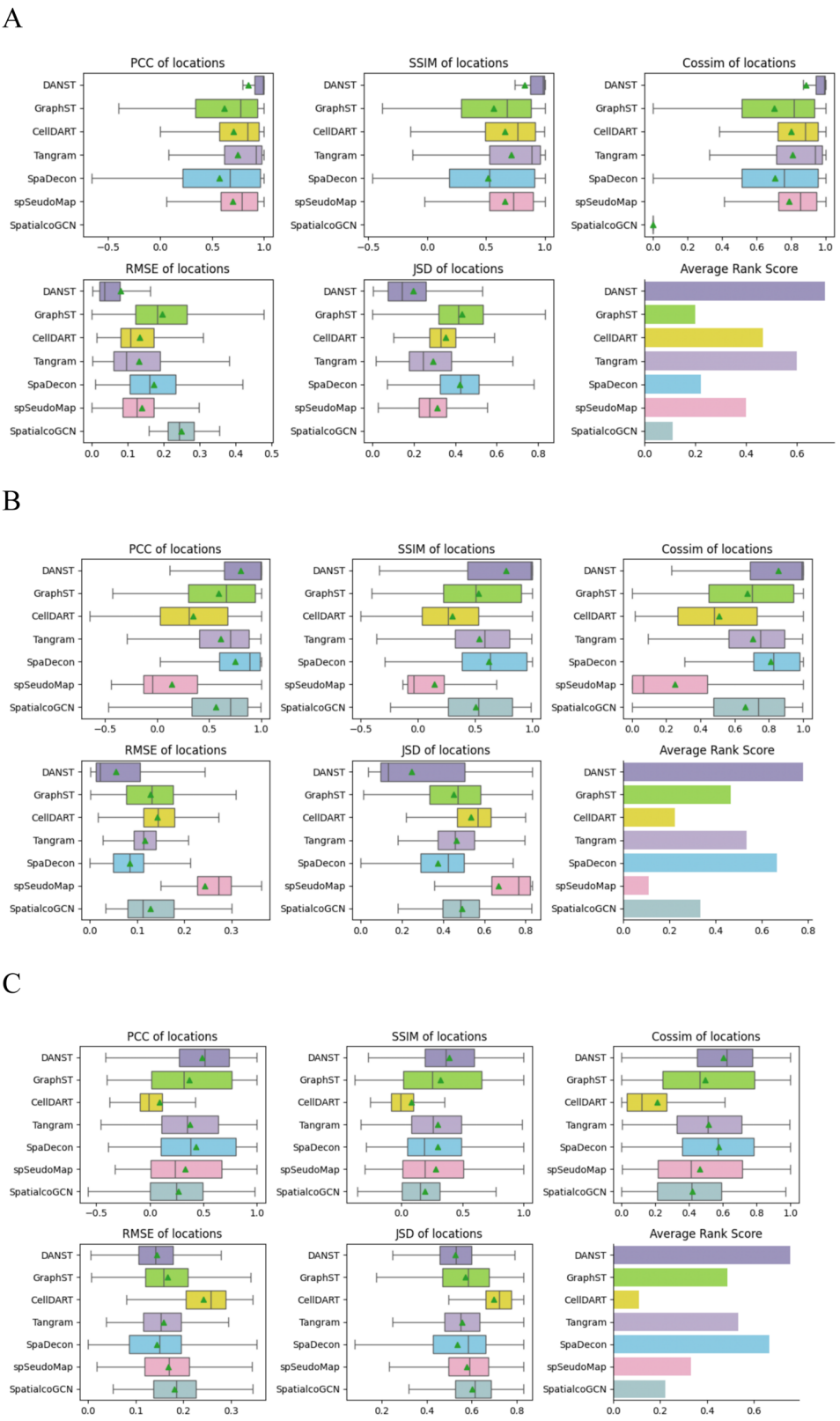

Fig. 2 | Evaluation of cell type deconvolution performance on the synthetic ST datasets. The cell type deconvolution capabilities of DANST and the other six methods are compared, and the results are presented in the form of box plots and bar graphs. Box plot: the middle line represents the median, the upper and lower limits of the box represent the upper and lower quartiles, and the triangle represents the mean. **A** Comparison of the ability of various methods to predict cell type proportions on the synthetic ST dataset 5. The box plot shows the results of PCC, SSIM, COSSIM, RMSE, and JSD, and the bar graph shows the results of ARS (an index aggregated from PCC, SSIM, COSSIM, RMSE, and JSD). **B** Comparison of the ability of various methods to predict cell type proportions on the synthetic ST dataset 8. **C** Comparison of the ability of various methods to predict cell type proportions on the synthetic ST dataset 12. For box plots, each ST spot is treated as one biologically independent sample; $n = 831$ for dataset 5, $n = 953$ for dataset 8, and $n = 1179$ for dataset 12.

visualize the distribution heat map of each cell type on the tissue image, as demonstrated in Fig. 3A. From the figure, it can be observed that the proportions of neuronal subtypes corresponding to cortical layers L2–3, L4, L5, L6, as well as oligodendrocytes, are relatively high in their respective regions, while those in other areas are relatively low, indicating that the predicted results of DANST clearly reveal the cortical structure of the mouse brain's anterior end.

To further quantify the cell-type deconvolution result of DANST, as shown in Fig. 3B, 52 different domains artificially labeled by domain experts based on histological features and spatial context, are employed as the ground truth. The PCC values between each cell type distribution and corresponding label are calculated, and the specific results are presented in Fig. 3C. It is evident that the similarities between L2/3, L5, L6 and their corresponding labels are all greater than 0.6, indicating that the distributions

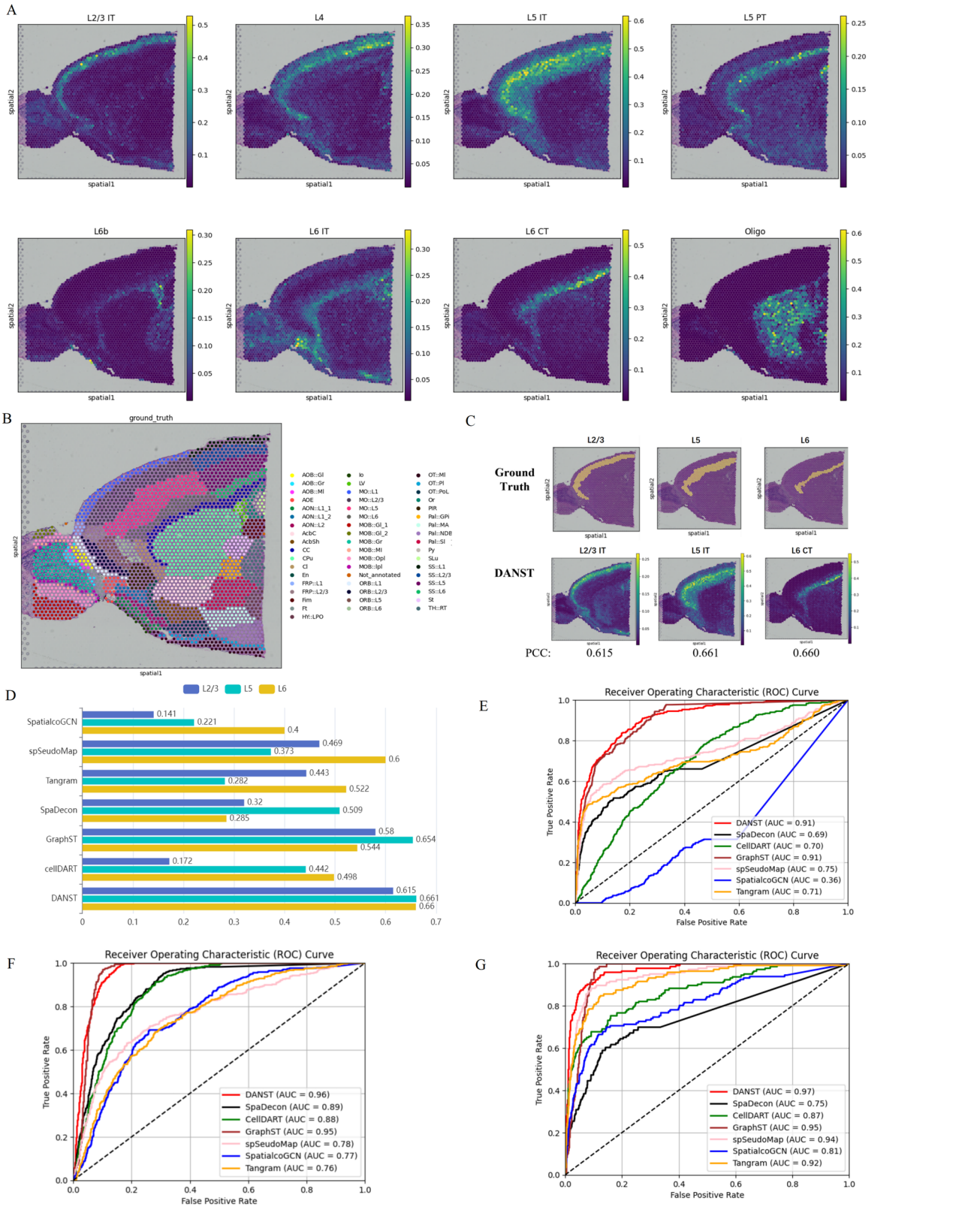

Fig. 3 | Cell type deconvolution experiment and result analysis on the mouse anterior brain dataset. A Based on the cell type proportion of each spot predicted by DANST, the distribution heat map of each cell type on the tissue image is visualized. **B** The domain division result of the mouse brain anterior data. **C** The similarity between the distribution of L2/3, L5, L6 cell and the ground truth, with the PCC value as the similarity index. Annotation correspondence: DANST('L2/3 IT', 'L5 IT', 'L6 IT') vs. Traditional ('L2/3', 'L5', 'L6'). **D** Accuracy comparison of DANST and other six cell type deconvolution methods in predicting the distribution of L2/3, L5, and L6 cells, with the accuracy index being PCC. **E–G** The accuracy of the L2/3, L5, and L6 cell distribution results predicted by DANST and other six cell type deconvolution methods are evaluated by ROC curve. On the ROC curve, the X-axis represents the false positive rate (FPR) and the Y-axis represents the true positive rate (TPR). The area under the curve (AUC) is an important indicator for evaluating model performance. AUC values range from 0 to 1: AUC = 1: perfect model; AUC = 0.5: random model; AUC < 0.5: poor model.

of L2/3, L5 and L6 in tissue image obtained by DANST are generally consistent with the ground truth.

In addition, in order to compare DANST with the GraphST, CellDART, Tangram, SpaDecon, spSeudoMap and SpatialcoGCN, the experiment performs the same operation on these six comparison methods, and the results are depicted in Fig. 3D and Supplementary Fig. 1. Through comparative analysis, it can be found that for L2/3, the method with the highest PCC value is DANST (PCC = 0.615), which is 6% higher than the second best method GraphST (PCC = 0.580). This phenomenon can also be seen from Supplementary Fig. 1 that the results of DANST and GraphST both relatively accurately show the location of the L2/3 cell in the tissue image, but GraphST performs slightly worse than DANST. For L5, the methods with the highest and second highest PCC values are still DANST (PCC = 0.661) and GraphST (PCC = 0.654), both of which show good stratification results in Supplementary Fig. 1. For L6, DANST (PCC = 0.661) is 10% higher than the sub-optimal method spSeudoMap (PCC = 0.600). This is also well demonstrated in Supplementary Fig. 1 that DANST clearly shows the area where L6 cells are located, while spSeudoMap has partially blurred contour boundaries.

In order to more accurately and comprehensively evaluate the cell type deconvolution results of DANST and the comparison methods, the experiment further introduces the receiver operating characteristic (ROC) curve[24], and the specific results are illustrated in Fig. 3E–G. As can be seen from the figure, for cell types L2/3, L5, and L6, the curves where DANST are located are closest to the upper left corner and have achieved the highest AUC value, further confirming the high prediction accuracy and sensitivity of the DANST method.

The experiment on mouse brain coronal section dataset

This experiment uses the coronal section of mouse brain[25] to more accurately observe the spatial patterns of different cell types in specific brain regions, thereby gaining a more comprehensive understanding of the complex structure and function of the mouse brain. The representative cell type deconvolution method spSeudoMap[22] is selected for comparison since it uses the same dataset in the experiment and shows superior performance in comparison with CellDART[21] and Cell2location[13]. The experimental results are shown in Fig. 4A. It is apparent that both methods can obtain obvious hierarchical structures (L2-L6), but only DANST can map Hpc-CA1 and Hpc-CA3 more clearly.

Since there is no manually annotated domain division for mouse brain coronal section, the ST data of mouse brain coronal section is spatially clustered with reference to spSeudoMap, providing an indirect validation of the deconvolution results. Each cluster is renamed according to its anatomical location and visualized on the tissue image, thus achieving domain division[26]. The specific results are presented in Fig. 4B. As can be seen from the figure, the divided domains include: Amy (amygdala), Amy_Pir (amygdala or olfactory cortex), Cor_out (outer cortex), Cor_mid (middle cortex), Cor_in (inner cortex), CP (caudate nucleus), EP (ventricular membrane), Hippo (hippocampus), Hypo (hypothalamus), Men (meninges), Thal_lat (lateral thalamus), Thal_med (medial thalamus) and WM (white matter). Then, in order to further confirm the accuracy of DANST's cell type deconvolution results on mouse brain coronal section dataset, the distribution ratio of each cell type in each domain is calculated experimentally, as shown in Fig. 4C.

From the figure, it can be observed that Ext_L25 is mainly located in the middle cortex, Ext_23 is positioned within the outer cortex, whereas Ext_L56 resides in the inner cortex. Ext_Hpc_CA1, Ext_Hpc_CA3, and Ext_Hpc_DG1 are situated in the hippocampus. Ext_Thal_1 is found in the thalamus, while Ext_Pir occupies areas of the amygdala or the olfactory cortex. In all, these cell types are all distributed in the expected anatomical domains, which proves that DANST can accurately predict the spatial composition of the main cell types that make up the cell subpopulation.

The experiment on human breast cancer dataset

This experiment uses more clinically relevant human breast cancer tissue dataset to perform cell type deconvolution experiment[27], thereby further revealing the diversity of cellular components within the tumor and helping to identify different cell sub-populations and their roles in tumor progression[28]. The ST data contains 3798 spots and 36,601 genes. Using its corresponding scRNA-seq data, the DANST method predicts the cell type proportion of each spatial spot in the ST data. Then, we visualize the distribution of each cell type on the tissue image, as shown in Fig. 5A. At the same time, pathologists manually annotated the ST data based on H&E images and spatial expression data of known breast cancer marker genes. It is divided into 20 regions in total, including four morphological types: Healthy, Tumor edge, Ductal Carcinoma In Situ/Lobular Carcinoma In Situ (DCIS/LCIS) and Invasive Ductal Carcinoma (IDC), as illustrated in Fig. 5B.

Combining Fig. 5A, B, it can be seen that luminal progenitor and luminal cell mainly locate in the IDC and DCLS/LCIS region, fibroblast mostly lies in the Tumor_edge region, and muscle cell are mainly positioned in the healthy region[29]. Further focusing on immune cells, it can be found that there are multiple immune subgroups in the IDC region[30], with T cells and macrophage/DC/monocytes (myeloid cells) present in IDC_5, 6, 7 and 8 regions. Meanwhile, there are myeloid and plasma cells reside in the IDC 4 region. In the DCIS/LCIS region, only myeloid cells are placed in DCIS/LCIS_2, 4 and 5, while there are no immune cells exist in DCIS/LCIS_1. B cells are mainly situated in IDC_3, Tumor_edge_3 and a part of the Healthy region. Small concentrations of plasma cells are found in the marginal areas of Tumor_edge region. It is worth noting that macrophage cell can be observed in tumor tissues, which is a clinical concern as these cells are associated with tumor progression[31]. At the same time, it can be observed that there is a significant deficiency of B cells, as it is also associated with poor prognosis in patients[32]. Therefore, we can infer that the patient's survival rate may be lower.

The observation from Fig. 5B also reveals that a large number of luminal cells exist in the IDC region. This experiment conducts an in-depth analysis of this phenomenon. Firstly, the literature shows that luminal cells are very sensitive to estrogen and other sex hormones, which play a key role in the development and function of breast tissue. At the same time, it can promote tumor growth and metastasis by interacting with other cells such as stromal cells and immune cells. They affect the behavior of surrounding cells in the tumor micro-environment by secreting cytokines and chemokines. In addition, luminal cells have a strong proliferation ability and can proliferate rapidly in the tumor micro-environment[33,34]. Secondly, the gene ontology enrichment analysis[35] of the top 30 differentially expressed genes (DEGs) in the IDC region is performed, and the results are demonstrated in Fig. 5C. As can be seen from the figure, the top four main functions of this region are regulation of steroid biosynthesis regulation (GO:0050810), O-glycosylation processing (GO:0016266), viral process (GO:0016032) and positive regulation of growth (GO:0045927). Among them, the regulation of steroid biosynthesis involves regulating the synthesis process of steroid hormones. O-glycosylation processing affects cell-to-cell interactions and signal transduction; viral process indicates the presence of virus-related signals in the tumor micro-environment; positive regulation of growth promotes cell proliferation and growth and promotes tumor progression. Through comparative analysis, it can be concluded that the various functions of luminal cells given in the literature correspond to the main functions of the IDC region obtained through gene ontology enrichment analysis. Therefore, the phenomenon that a large number of luminal cells are present in the IDC region is reasonable and accurate, which further proves the accuracy of the DANST cell type deconvolution results. At the same time, the large number of luminal cells in IDC region may indicate the tumor's dependence on hormones, especially its sensitivity to estrogen. Such tumors are usually classified as estrogen receptor positive (ER +), and patients may respond well to hormonal therapies such as tamoxifen or aromatase inhibitors. The metabolic reprogramming of luminal cells may be associated with aggressiveness and metastatic potential of the tumor. If these cells exhibit

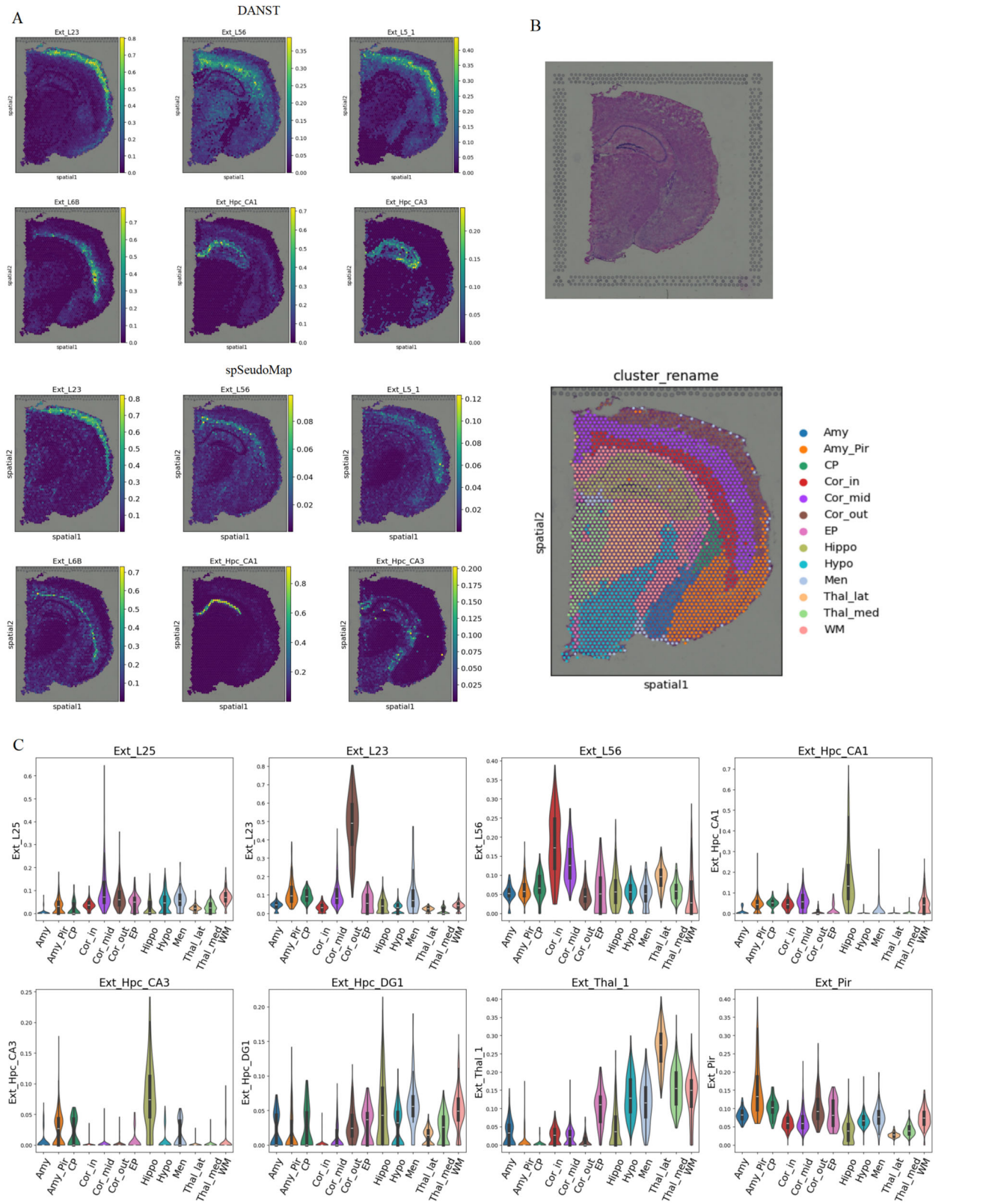

Fig. 4 | Cell type deconvolution experiment and result analysis on the mouse brain coronal section dataset. A The distribution heat map of specific cell types in mouse brain coronal section predicted by DANST and spSeudoMap. **B** Spatial spots are clustered according to gene expression profiles, and each cluster is named according to the anatomical location. **C** The distribution of each cell type at each anatomical location is quantitatively visualized in the form of a violin plot. The violin plot contains a box plot section showing the median, upper and lower quartiles of the data.(i.e., one-quarter and three-quarter values). For the violin and box plots, each ST spot is treated as one biologically independent sample ($n = 2702$).

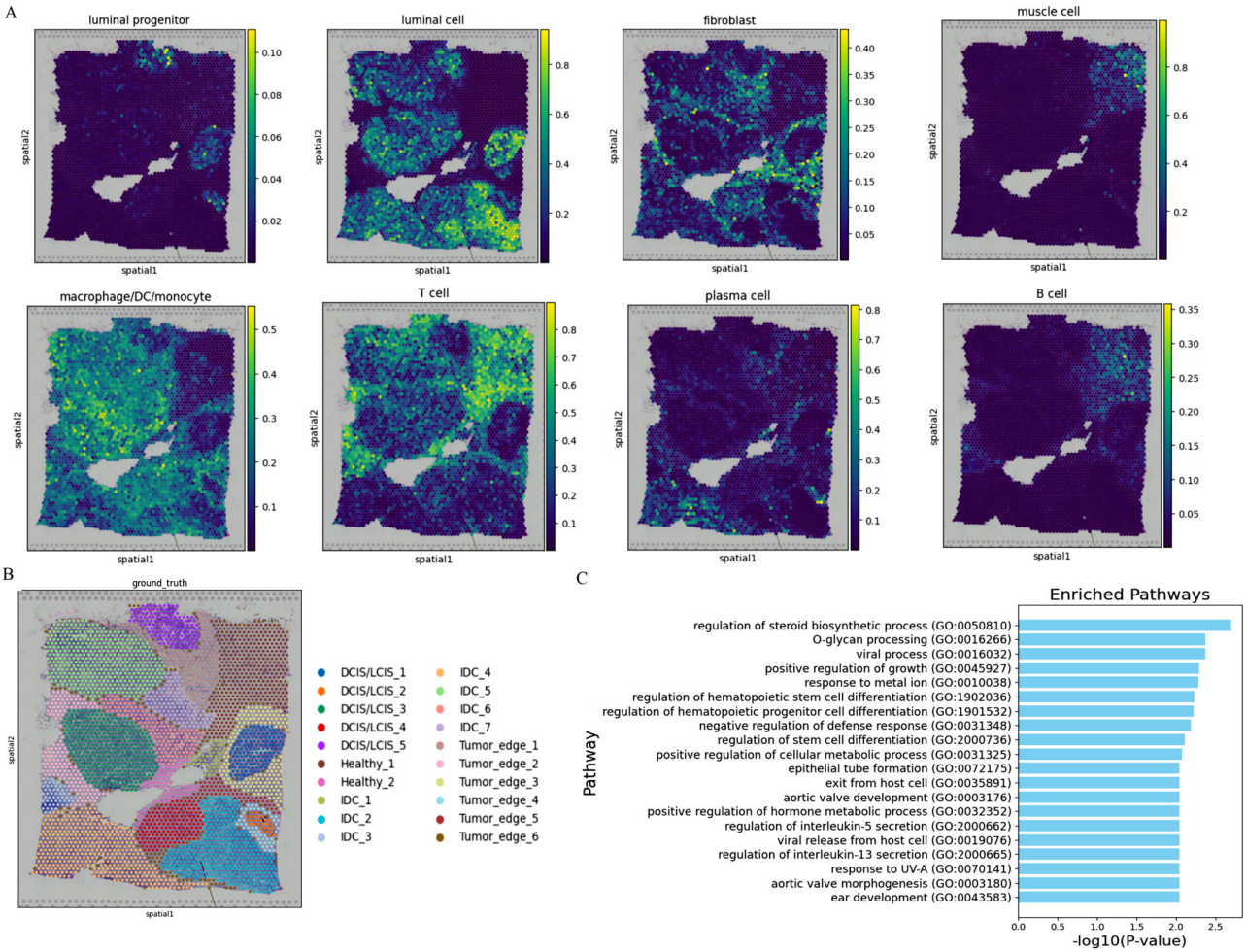

Fig. 5 | Cell type deconvolution experiment and result analysis on the human breast cancer dataset. A The distribution heat map of specific cell types in human breast cancer predicted by DANST. **B** ST data of human breast cancer sample annotated by pathologists, which includes Invasive ductal carcinoma(IDC), Ductal carcinoma in situ(DCIS), Lobular carcinoma in situ(LCIS), Tumor edge and Healthy region. **C** Gene ontology enrichment analysis of the top 30 differentially expressed genes in the IDC region.

metabolic adaptation in the IDC region, this may indicate a worse prognosis, thereby affecting patient's survival rate.

Discussion

Cell type deconvolution is a basic task in ST analysis. This task enables the discovery of new cell types and sub-populations, revealing their spatial distribution and functional characteristics within tissues. To this end, we propose a domain adversarial network-based framework named DANST. The framework constructs pseudo ST data using scRNA-seq data with clear cell type labels, and calculates pseudo coordinates of pseudo ST data with the assistance of the spatial location information from real ST data. The two types of data are connected by constructing a neighbor graph, and feature optimization is performed through a variational auto-encoder structure. By designing a domain adversarial module, the cell type label is transferred from pseudo data to real data.

The main features of DANST are the construction of pseudo spots and domain adversarial training. Existing methods such as CellDART also use scRNA-seq data to construct pseudo ST data. The main difference between DANST and them is that pseudo coordinates are additionally constructed for pseudo spots, and a neighbor graph of pseudo spots and real spots is constructed based on the location information to connect the two types of data. In addition, spSeudoMap also uses domain adversarial training. However, DANST uses gene representations that have been feature-optimized and integrated with spatial location information. During the training process, DANST uses more reasonable objective functions and contrast losses. It also designs the two-stage training strategy. These differences make DANST superior to spSeudoMap in cell type deconvolution tasks.

Benchmark experiments on artificial synthetic datasets show that DANST outperforms six comparison methods in the cell type deconvolution task, and the average optimization ratio can reach 16.6%. In addition, experiments on a real mouse brain anterior dataset show that the DANST method has the ability to accurately reveal the cell types and hierarchical structure of the cerebral cortex. Experiments on a mouse brain coronal slice dataset illustrate that DANST can more accurately observe the spatial patterns of different cell types in specific brain regions. Experiments on human breast cancer datasets have shown that DANST can accurately predict the distribution of various types of cells in breast cancer tissue images, thereby further revealing the diversity of cellular components within tumors and providing tools for studying tumor evolution and the interaction between tumors and the tumor micro-environment. Furthermore, DANST can process data obtained from different experimental platforms and has been verified on 10xVisium, seqFISH + , and Smart-seq data.

In conclusion, DANST can efficiently perform cell type deconvolution tasks, which helps better understand the cellular and functional diversity of tissues. With the availability of more ST data and the public release of scRNA-seq data, DANST will become feasible for cell type deconvolution tasks in a broader context. Based on this, we can apply DANST in clinical

research and medical practice, exploring its specific applications in tumor micro-environments, immune responses, drug reactions, and regenerative medicine. This will enable us to assess its impact on disease diagnosis and treatment planning. Additionally, discovering previously unidentified cell types or sub-populations through deconvolution techniques represents a promising future research direction. By utilizing high-throughput single-cell sequencing technologies in conjunction with deconvolution methods, we can advance our understanding of biological mechanisms and reveal cellular diversity and functionality. These research avenues not only contribute to the advancement of cell type deconvolution technologies but also offer new insights into the understanding of complex biological systems and the improvement of medical treatments.

Methods

Data pre-processing and pseudo spots construction

Gene expression data are often highly skewed, making raw counts unsuitable for direct analysis. Without proper normalization, pseudo spots may contain excessive noise and biases, thereby reducing representativeness and impairing accuracy in spatial mapping and clustering. Therefore, appropriate preprocessing is necessary to ensure reliable pseudo spot construction.

Data pre-processing

There are two types of data input to DANST, namely ST data with spatial location information and gene expression information, and scRNA-seq data with cell type labels and gene expression information. In order to reduce unnecessary noise and impact in the data, data pre-processing is required before constructing pseudo spots.

For scRNA-seq data, the raw gene expression information is first normalized to a target sum per cell and subsequently log-transformed via the SCANPY[36] package. Finally, this processed gene expression information is scaled to unit variance and zero mean. Differential expression (DE) analysis is performed on normalized expression profiles, including both cell-level and gene-level normalization. For each cell type, the top d genes with the highest log-fold change and statistically significant p-values ($p < 0.05$, adjusted for multiple testing) are selected, resulting in a total of q marker genes. For ST Data, we use the similar logarithmic transformation and scaling operation. Then, the common genes shared by scRNA-seq data and ST data are screened out to achieve gene expression alignment for subsequent model input.

pseudo spots construction

After data pre-processing, we utilize preprocessed scRNA-seq data to construct pseudo ST data, which helps stabilize gene expression variance, reduce noise, and ensure comparability with the ST dataset. A fixed number of cells h are randomly sampled from the scRNA-seq data, and each cell is given a random weight$(w_1, w_2, \ldots, w_h)$. The resulting cell mixture is called pseudo spot. Among the cells that make up the pseudo spot, the weighted sum of the gene expression of each cell is calculated to obtain the gene expression of the pseudo spot $x' \in R^{1\times t}$, which is specifically expressed as:

$$x' = w_1 * cell_1 + \ldots + w_h * cell_h, \tag{1}$$

Among them, $cell_i, i \in \{1, 2, \ldots, h\}$ represents the gene expression of cell i in scRNA-seq data. At the same time, by adding the weights assigned to each cell type, the proportion of each cell type $p \in R^{1\times c}$ in pseudo spot is obtained, which is specifically expressed as:

$$W_j = \sum_{i\in T_j} w_i, \tag{2}$$

$$p_j = \frac{W_j}{\sum_{k=1}^{c} W_k}, p \in R^{l\times c}, \tag{3}$$

Where c is the total number of cell types, where T_j denotes the index set of cells belonging to cell type j, c is the total number of cell types, and p represents the cell type proportion vector of the pseudo spot. W_j represents the weight assigned to the cell type j.

Repeat this process to generate n pseudo spots, which forms pseudo ST data. Then, the mean and standard deviation of gene expression at each spot are estimated from the real ST data, and each pseudo spot is down-sampled using the down-sample Matrix function provided by Scuttle[37], thus obtaining pseudo ST data with gene expression distribution similar to the real ST data.

By stacking the previously defined pseudo spot vectors x' and p, we obtain the gene expression matrix $X' = \{x'_1, x'_2, \ldots, x'_n\} \in R^{n\times t}$ and the cell type proportion matrix $P = \{p_1, p_2, \ldots, p_n\} \in R^{n\times c}$, the real ST data can be represented by a gene expression matrix $X = \{x_1, x_2, \ldots, x_m\} \in R^{m\times t}$ and a spatial location matrix $L = \{l_1, l_2, \ldots, l_m\} \in R^{m\times 2}$. Among them, n and m are the number of pseudo spots and real spots respectively, and t is the number of common genes.

The generated pseudo spots are mainly used for the later model training, providing mixed expression data with known cell compositions and reliable cell type annotations. The downstream spatial analyses and biological interpretations are still based on real ST data.

Generation of pseudo coordinates

After constructing pseudo ST data, it is necessary to construct pseudo coordinates for each pseudo spot, which are used for subsequent neighbor graph construction.

Mclust[38] is a spatial clustering method based on Gaussian mixture models, which can effectively partition spatial expression patterns into interpretable clusters. For real ST data, we use Mclust for spatial clustering to obtain g clusters. For each cluster, the spatial coordinates of each spot are averaged as the cluster center $\widetilde{L} = \{\widetilde{l}_1, \widetilde{l}_2, \ldots \widetilde{l}_g\} \in R^{g\times 2}$, and the feature representation of each spot is averaged to obtain the feature representation of the cluster $\widetilde{X} = \{\widetilde{x}_1, \widetilde{x}_2, \ldots \widetilde{x}_g\} \in R^{g\times t}$. Based on the feature representation of the pseudo spot and the clusters, the feature similarity between the two is calculated, and the pseudo coordinates of the pseudo spots are calculated with this similarity as the weight. Specifically, for the pseudo spot i, the similarity value S_{ij} between its feature representation x_i and the feature representation of the cluster j can be expressed as:

$$S_{ij} = \frac{1}{d_{ij} + \alpha}, \tag{4}$$

$$d_{ij} = cdist(x_i, \widetilde{x}_i), \tag{5}$$

Among them, α is an offset used to avoid zero errors when taking the reciprocal, $cdist$ represents the calculation of Euclidean distance. The calculation formula for the pseudo coordinate l'_i of the pseudo spot P spot i is:

$$l'_i = \frac{\sum_{j=1}^{k} S_{ij}\widetilde{l}_j}{\sum_{j=1}^{k} S_{ij}}, \tag{6}$$

By using the above method, pseudo coordinate can be constructed for each pseudo spot, and the pseudo spatial location matrix of the pseudo ST data $L' = \{l'_1, l'_2, \ldots, l'_n\} \in R^{n\times 2}$ is obtained.

Neighborhood graph construction and feature optimization

(1) Neighborhood graph construction

In order to effectively connect real ST data and pseudo ST data, it is necessary to construct neighbor graph before feature optimization. We apply k-nearest neighbor algorithm to convert the spatial location

information of all pseudo spots and real spots into a neighbor graph $G = (V, E)$, where V represents the spot set and E represents the edge set connecting the spots. For a specific spot $i \in V$, the Euclidean distance between this spot and all other spots is calculated, and the k spots closest to it are selected as neighbors. And so on, the adjacency relationship between all spots is obtained. Here, we introduce the adjacency matrix $A \in R^{(n+m)\times(n+m)}$ to represent the adjacency relationship of all spots in the neighbor graph$G$, where $n + m$ is the total number of spots. If the spot $j \in V$ is the neighbor of spot $i \in V$, then $A_{ij} = 1$, otherwise $A_{ij} = 0$. In order to normalize the degree difference of the nodes, the adjacency matrix is needed to be regularized. Let $\widetilde{A} \in R^{(n+m)\times(n+m)}$ is the adjacency matrix after regularization, and its specific calculation formula is:

$$\widetilde{A} = \bar{A} \cdot A \cdot \bar{A}, \tag{7}$$

$$\bar{A} = diag\left(\left(\sum_j^N A_j\right)^{-\frac{1}{2}}\right), \tag{8}$$

Among them, $(\cdot)$ is dot multiplication, *diag* represents the operation of creating or extracting diagonal matrices, and *power* is the power operation of numbers.

(2) feature optimization

After the neighbor graph is constructed, we perform the feature optimization operation. In order to effectively integrate the gene expression information and spatial location information in the data, also removing unnecessary noise in the gene expression information, such as low capture efficiency, PCR amplification bias, and low-abundance genes, a variational auto-encoder is utilized for feature extraction and feature reconstruction.

The encoder architecture is designed as follows: both the mean encoder $Encoder_\mu$ and the variance encoder $Encoder_\sigma$ consist of two stacked graph convolution layers, a batch normalization layer, an ELU layer (non-linear activation function) and a dropout layer.

The feature representation of the real ST data X and pseudo ST data X' as well as the adjacency matrix $\widetilde{A}$ are input into the encoder to obtain feature embedding F:

$$\mu = Encoder_\mu\left([X, X'], \widetilde{A}\right), \tag{9}$$

$$\sigma = Encoder_\sigma\left([X, X'], \widetilde{A}\right), \tag{10}$$

$$F = \mu + \sigma * \tau \tag{11}$$

where τ is a random variable whose value is sampled from $N(0, 1)$ and $[,]$ Indicates vertical splicing. The decoder is constructed by two fully connected layers, and the feature embedding generated by the encoder is input into the decoder for feature reconstruction. The loss between the original and reconstructed input is calculated with the mean squared error (MSE) loss, and the specific formula is as follows:

$$[Z, Z'] = Decoder(F), \tag{12}$$

$$Loss_{reconstruction} = MSE(X, Z) + MSE(X', Z'), \tag{13}$$

Among them, Z and Z' represent the feature inputs reconstructed from real ST data and pseudo ST data respectively, which are used as the input of the next module as the data after feature optimization.

Domain adversarial training

Since there are certain distribution differences between pseudo ST data and real ST data, it is necessary to make the features of the two types of data tend to be consistent in the representation space. Domain adversarial network can effectively solve the problem of feature distribution differences between the source domain and the target domain through domain training and feature alignment mechanism[39]. Therefore, DANST introduces domain adversarial training, so that the model can better adapt to the target domain after training on the source domain, thereby realizing cross-domain knowledge transfer.

The process of domain adversarial training is shown in Fig. 1D. The domain adversarial module consists of three parts: feature extractor, domain discriminator, and proportion predictor. The feature extractor consists of two fully connected layers, each with batch normalization and ELU activation functions. Its main function is to extract the domain-inseparable features that are useful for the target task from the two different regional distributions of pseudo ST data and real ST data. The domain discriminator consists of two fully connected layers. The output of the first layer is input to the second layer after batch normalization, ELU activation, and dropout. The output of the second layer is then passed through the Sigmoid function to obtain the result of domain discrimination. The structure of the proportion predictor is similar to that of a feature extractor, and is mainly used to predict the proportion of cell types. The module adopts two-stage alternating training.

In the first stage, the feature extractor is used to extract the feature representation of the pseudo ST data and the real ST data, and input it into the domain discriminator to obtain the result of domain discrimination. Then, we flip the domain labels and construct the loss of domain discrimination through binary cross entropy (BCE) loss. At the same time, the feature representation of the pseudo ST data is input into the proportion predictor to predict the cell type proportion, and the mean absolute error (MAE) loss and kullback leibler divergence (KLD) are constructed with the known cell type proportion of the pseudo ST data. During training, the domain discriminator is fixed, and the model parameters of the feature extractor and proportion predictor are optimized. In the second stage, the feature representations of ST data and real ST data are input into the domain discriminator to obtain the results of domain discrimination. Then, we calculate the loss of domain discrimination through BCE. During training, the proportion predictor and feature extractor are fixed, the model parameters of the domain discriminator are optimized. The specific formula is as follows:

$$V = Feature_{Extractor(Z)}, \tag{14}$$

$$V' = Feature_{Extractor(Z')}, \tag{15}$$

$$D = Domain_{discriminator(V)}, \tag{16}$$

$$D' = Domain_{discriminator(V')}, \tag{17}$$

$$Loss_{adv1} = BCE(D, D_p) + BCE\left(D', D_r\right), \tag{18}$$

$$P' = Proportion_{Predictor(V')}, \tag{19}$$

$$Loss_{pre1} = MAE(P, P') + KLD(P, P'), \tag{20}$$

$$Loss_{DAN1} = Loss_{adv1} + Loss_{pre1}, \tag{21}$$

$$Loss_{DAN2} = Loss_{adv2} = BCE\left(D, D_r\right) + BCE(D', D_p), \tag{22}$$

Among them, V and V' represent the feature embedding of the real ST data and the pseudo ST data after passing through the feature extractor, D and D' represent the predicted domain labels of the real ST data and the pseudo ST data, D_r andD_p represent the true domain label of real ST data and pseudo ST data, P and P' represent the real and predicted cell type proportion of the pseudo ST data, $Loss_{adv1}$ and $Loss_{pre1}$ represent the domain discrimination loss and the cell type prediction loss, respectively,

$Loss_{DAN1}$ and $Loss_{DAN2}$ represents the total loss of the first and second stage training.

The two stages are trained alternately. The first stage of training prompts the main network to learn features that are insensitive to domain classification and beneficial for cell type deconvolution task, while the second stage of training prompts the domain discriminator to distinguish which domain the features come from. This two stage of training forms an adversarial relationship. Finally, the training ends when the loss change is less than the set threshold or reaches the pre-set training round. Through domain adversarial training, the model can learn features that are both useful for the specific task and have generalization capabilities for different domains, thereby enhancing the model's robustness to various data sources, further reducing the misclassification rate across different datasets and improving overall prediction accuracy.

Cell type proportion prediction

Based on the above model training, the optimized feature extractor and proportion predictor is used to realize the cell type proportion of real ST data. The specific formula is as follows:

$$R = Refined\_Proportion\_Predictor(Refined\_Feature\_Extractor(Z)), \quad (23)$$

Among them, R is the predicted result of the cell type proportion of each spot in real ST data,*Refined_Feature_Extractor* and *Refined_Proportion_Predictor* represent the optimized feature extractor and proportion predictor, respectively.

Experimental details

The proposed DANST model is implemented with python and PyTorch_pyG[40]. In the pseudo spot construction and data preprocessing module, the number of differentially expressed genes d is equal to 30, and the number of cells contained in each pseudo spot h is set to 10. In the pseudo coordinate generation module, the number of clusters b equals 25. In the feature optimization module, the number of neighbors $k = 10$, and the dimensions of the encoder's hidden layer and output layer are 256 and 128, respectively, while the dimensions of the decoder's layers are opposite. In the domain adversarial training module, the dimensions of the two layers in feature extractor are 256 and 128, respectively, and the dimensions of the two layers in domain discriminator are 64 and 2, respectively. The hidden layer dimension in proportion predictor is 64, and the output layer dimension is the total number of cell types. During the training process, the total loss function is optimized by the Adam optimizer[41], with an initial learning rate of 0.001 and a weight decay of 0.

DANST is trained on a GeForce RTX A4000 Ti GPU with 16 GB video memory, and a 6 x E5-2680 v4 CPU with 28 GB memory. The training time ranges from 1 minute to 4 minutes depending on different datasets.

Evaluated metrics and criteria

For the cell type deconvolution task of spatial transcriptome data, we employ the root mean square error (RMSE), person correlation coefficient (PCC), Jensen-Shannon divergence (JSD), structural similarity (SSIM), cosine similarity (COSSIM) and average ranking score (ARS) as evaluation indicators. The specific calculation formulas are as follows.

(1) Root mean square error (*RMSE*):

$$RMSE = \sqrt{\frac{1}{c}\sum_{j=1}^{c}(r'_{ij} - r_{ij})^2}, \quad (24)$$

Among them, r_{ij} represents the proportion of cell type j in the spot i in the true result, r'_{ij} is the predicted result, and c is the total number of cell types. For each spot in the ST data, a lower value indicates a higher deconvolution accuracy.

(2) Pearson correlation coefficient (*PCC*):

$$PCC = \frac{E\left[\left(r'_i - \mu'_i\right)\left(r_i - \mu_i\right)\right]}{\sigma'_i \sigma_i}, \quad (25)$$

Among them, r'_i and r_i are the predicted and true results of cell type proportion in spot i, μ'_i and μ_i are the mean value of the predicted and true results of cell type proportion in spot i, σ'_i and σ_i are the variance of the predicted and true results of cell type proportion in spot i. The PCC value ranges from -1 to 1, and the closer to 1, the higher the similarity. For each spot in ST data, a higher PCC value indicates better prediction accuracy.

(3) Jensen-Shannon divergence (*JSD*):

$$JSD = \frac{1}{2}KL\left(r'_i, |, \frac{r'_i + r_i}{2}\right) + \frac{1}{2}KL\left(r_i, |, \frac{r'_i + r_i}{2}\right) \quad (26)$$

$$KL(a_i||b_i) = \sum_{j=0}^{c}\left(a_{ij} * \log\frac{a_{ij}}{b_{ij}}\right) \quad (27)$$

Where, r'_i and r_i are defined similarly to PCC, $KL(a_i||b_i)$ are the KL divergences between the two probability distributions a_i and b_i. For each spot, a lower value indicates a higher deconvolution accuracy.

(4) Structural similarity (*SSIM*):

$$SSIM = \frac{(2\mu'_i\mu_i + \alpha^2)(2\mathrm{cov}(r_i, r'_i) + \beta^2)}{\left(\mu_i'^2 + \mu_i^2 + \alpha^2\right)\left(\sigma_i'^2 + \sigma_i^2 + \beta^2\right)} \quad (28)$$

Among them, r'_i, r_i, μ'_i, μ_i, σ'_i and σ_i are defined similarly to those in PCC, α and β are set to 0.01 and 0.03, respectively. $\mathrm{cov}(r_i, r'_i)$ is the covariance. For each spot in ST data, a higher value indicates a higher deconvolution accuracy.

(5) Cosine similarity (*COSSIM*):

$$COSSIM = \frac{\sum_{j=1}^{c} r'_{ij} r_{ij}}{||r'_i|| \times ||r_i||} \quad (29)$$

Among them, the definitions of r_{ij} and r'_{ij} are similar to those in RMSE, and the definitions of r_i and r_i are similar to those in PCC, $||\bullet||$ represents L2 norm normalization. For each spot in ST data, a higher value indicates better deconvolution accuracy.

(6) Average ranking score (*ARS*):

We uses ARS to comprehensively evaluate the relative accuracy of the deconvolution results of various methods. After calculating the five indicators of RMSE, PCC, JSD, SSIM and COOSSIM, the PCC, COSSIM and SSIM values of each method are sorted in ascending order, and the RMSE and JSD values of each method are sorted in descending order. The value of ARS is calculated by the following formula:

$$ARS = \frac{1}{5 \times M}\left(R_{RMSE} + R_{PCC} + R_{JSD} + R_{SSIM} + R_{COSSIM}\right), \quad (30)$$

Among them, M is the number of comparison methods, R indicates the ranking of the method on a certain evaluation indicator. For a certain dataset, the method with the highest value performs the best.

Statistics and reproducibility

No statistical method is used to predetermine the sample size. All collected data are included in the analyses without exclusion. The experiments are not randomized. The investigators are not blinded to allocation during experiments and outcome assessment.

Data availability

We use six sets of ST data and the corresponding scRNA-seq data. Three of them are artificially synthesized data, which are the datasets used by Wang et al.16 to conduct benchmark experiments and can be downloaded at https://figshare.com/articles/dataset/SpatialcoGCN_data/22682611. The other three sets are real data, namely the mouse brain anterior dataset, the mouse brain coronal slice dataset, and the human breast cancer dataset. The ST data of the mouse brain anterior dataset is obtained from the 10x Genomics dataset[2] (https://www.10xgenomics.com/resources/datasets), and the corresponding scRNA-seq data is acquired from the whole cortex and hippocampus of the mouse (https://portal.brain-map.org/atlases-and-data/rnaseq/mouse-whole-cortex-and-hippocampus-10x). The ST data and corresponding scRNA-seq data of the mouse brain coronal section dataset are collected from the 10x Genomics dataset and E-MTAB-11115[42] (https://www.ebi.ac.uk/biostudies/arrayexpress/studies/EMTAB-11115), respectively. The ST data of the human breast cancer dataset is retrieved from https://www.10xgenomics.com/resources/datasets/human-breast-cancer-block-asection-1-1-standard-1-1-0, and the scRNA-seq data is downloaded from the database DISCO[43] (https://www.immunesinglecell.org/). The data used in this study have been uploaded to Zenodo and is freely available at: https://doi.org/10.5281/zenodo.18213061[44]. All source data underlying the graphs and charts are provided as Supplementary Data 1, Supplementary Data 2, Supplementary Data 3, and Supplementary Data 4. All other data supporting the findings of this study are available from the corresponding author upon reasonable request.

Code availability

An open-source Python implementation of the DANST toolkit is accessible at https://github.com/ZhichaoWu7/DANST.The version of the code described in this paper has been deposited in Zenodo with the identifier https://doi.org/10.5281/zenodo.18212150[45].

Received: 7 March 2025; Accepted: 27 January 2026;
Published online: 09 February 2026

References

1. Smith, K. D., Prince, D. K., MacDonald, J. W., Bammler, T. K. & Akilesh, S. Challenges and opportunities for the clinical translation of spatial transcriptomics technologies. *Glomerular Dis.* **4**, 49–63 (2024).
2. Asp, M., Bergenstråhle, J. & Lundeberg, J. Spatially resolved transcriptomes—next generation tools for tissue exploration. *Bioessays* **42**, e1900221 (2020).
3. 10x Genomics. https://www.10xgenomics.com/resources/datasets/ (2023).
4. Rodriques, S. G. et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).
5. Liu, Y. et al. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell* **183**, 1665–1681 (2020).
6. Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792 (2022).
7. Fu, X. et al. Polony gels enable amplifiable DNA stamping and spatial transcriptomics of chronic pain. *Cell* **185**, 4621–4633 (2022).
8. Cho, C.-S. et al. Microscopic examination of spatial transcriptome using seq-scope. *Cell* **184**, 3559–3572.e22 (2021).
9. Eng, C.-H. L. et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235–239 (2019).
10. Cable, D. M. et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* **40**, 517–526 (2022).
11. 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.* **49**, e50 (2021).
12. Andersson, A. et al. Spatial mapping of cell types by integration of transcriptomics data. *Commun. Biol.* **3**, 77 (2020).
13. Kleshchevnikov, V. et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).
14. Garmire, L. X. et al. Challenges and perspectives in computational deconvolution of genomics data. *Nat. Methods* **21**, 391–400 (2024).
15. Xu, H. et al. SPACEL: deep learning-based characterization of spatial transcriptome architectures. *Nat. Commun.* **14**, 7603 (2023).
16. Li, H., Li, H., Zhou, J. & Gao, X. SD2: spatially resolved transcriptomics deconvolution through integration of dropout and spatial information. *Bioinformatics.* **38**, 4878–4884 (2022).
17. Long, Y. et al. Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nat. Commun.* **14**, 1155 (2023).
18. Wang, Y., Wan, Y. & Zhou, Y. SpatialcoGCN: deconvolution and spatial information–aware simulation of spatial transcriptomics data via deep graph co-embedding. *Brief. Bioinform.* **25**, bbae130 (2024).
19. Coleman, K. et al. SpaDecon: cell-type deconvolution in spatial transcriptomics with semi-supervised learning. *Commun. Biol.* **6**, 378 (2023).
20. Biancalani, T. et al. Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat Methods* **18**, 1352–1362 (2021).
21. Bae, S. et al. CellDART: cell type inference by domain adaptation of single-cell and spatial transcriptomic data. *Nucleic Acids Res.* **50**, e57 (2022).
22. Bae, S., Choi, H. & Lee, D. S. spSeudoMap: cell type mapping of spatial transcriptomics using unmatched single-cell RNA-seq data. *Genome Med.* **15**, 19 (2023).
23. Liu, Z. et al. SONAR enables cell type deconvolution with spatially weighted Poisson-Gamma model for spatial transcriptomics. *Nat. Commun.* **14**, 4727 (2023).
24. Davis, J. & Goadrich, M. An introduction to ROC analysis. *Pattern Recognit. Lett.* **27**, 861–874 (2006).
25. Ståhl, P. L. et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
26. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature.* **563**, 72–78 (2018).
27. Wu, S. Z. et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat. Genet.* **53**, 1334–1347 (2021).
28. Yang, W. et al. Single-cell RNA reveals a tumorigenic microenvironment in the interface zone of human breast tumors. *Breast Cancer Res.* **25**, 100 (2023).
29. Kumar, T. et al. A spatially resolved single-cell genomic atlas of the adult human breast. *Nature* **620**, 181–191 (2023).
30. Keren, L. et al. A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell* **174**, 1373–1387 (2018).
31. Carron, E. C. et al. Macrophages promote the progression of premalignant mammary lesions to invasive cancer. *Oncotarget* **8**, 50731–50746 (2017).
32. Hu, Q. et al. Atlas of breast cancer infiltrated B-lymphocytes revealed by paired single-cell RNA-sequencing and antigen receptor profiling. *Nat. Commun.* **12**, 2186 (2021).
33. Roy, M., Fowler, A. M., Ulaner, G. A. & Mahajan, A. Molecular Classification of Breast Cancer. *PET Clin.* **18**, 441–458 (2023).
34. Song, Y. et al. DDHD2 is involved in the malignant progression of early luminal A breast cancer by changing cell membrane proteins and immune responses functionality. *Oncol. Transl. Med.* **10**, 231–244 (2024).
35. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
36. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
37. McCarthy, D. J., Campbell, K. R., Lun, A. T. L. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 (2017).

38. Fraley, C., Raftery, A. E., Murphy, T. B. & Scrucca, L. mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. Report No. 597 (University of Washington, 2012).
39. Ganin, Y. et al. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* **17**, 1–35 (2016).
40. Fey, M. & Lenssen, J. E. Fast graph representation learning with PyTorch geometric. *ICLR 2019 Workshop on Representation Learning on Graphs and Manifolds*. (ICLR, 2019).
41. Bock, S., & Weiß, M. A proof of local convergence for the Adam optimizer. *IEEE International Joint Conference on Neural Networks* (IEEE, 2019).
42. Kleshchevnikov V., et al. Single-nucleus RNA-seq from adult mouse brain sections paired to 10X Visium spatial RNA-seq. E-MTAB-11115, (EMBL-EBI, 2022).
43. Li, M. et al. DISCO: a database of deeply integrated human singlecell omics data. *Nucleic Acids Res.* **50**, D596–D602 (2022).
44. Wu, Z. Processed datasets for DANST enables cell-type deconvolution in spatial transcriptomics using deep domain adversarial neural networks. *Zenodo* https://doi.org/10.5281/zenodo.18213061 (2026).
45. Wu, Z. ZhichaoWu7/DANST: First version for publication (v1.0.0). *Zenodo* https://doi.org/10.5281/zenodo.18212150 (2026).

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (NSFC No.51975213) and Clinical Project of Shanghai Municipal Health Commission (No.202340054).

Author contributions

X.Z. and T.W. envisaged the project and wrote the manuscript. T.W. and Z.W. implemented the model and code, performed the experiments. H.Z., Y.Z., and W.D. analyzed and interpreted data of S.T., and validated the final method. X.Z. and Q.Z. provided model evaluation and critical manuscript revision. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-026-09659-y.

Correspondence and requests for materials should be addressed to Xueqin Zhang or Tianqi Wang.

Peer review information *Communications Biology* thanks Yanghong Guo, Nam D. Nguyen and Chitrasen Mohanty for their contribution to the peer review of this work. Primary Handling Editors: Laura Rodríguez Pérez. A peer review file is available.