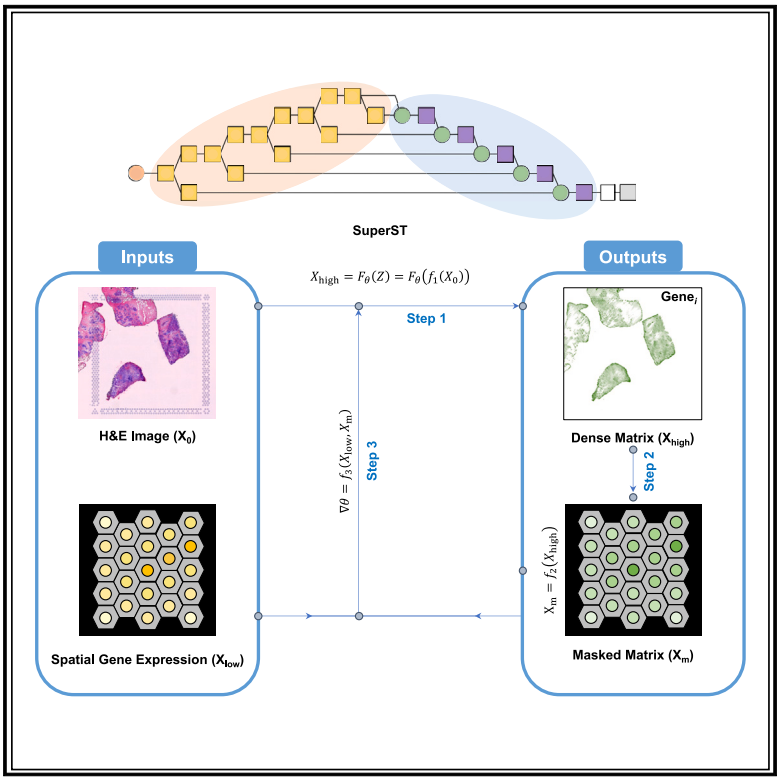
Cell Reports
Methods

Report

Generation of super-resolution images from barcode-based spatial transcriptomics by deep image prior

Graphical abstract



Highlights

- SuperST enhances spatial transcriptomics resolution using deep image prior
- It outperforms other methods in generating dense gene expression matrices
- Reduced parameter sensitivity enables reliable performance across datasets

Authors

Jeongbin Park, Seungho Cook, Dongjoo Lee, ..., Hyung-Jun Im, Daeseung Lee, Hongyoon Choi

Correspondence

chy1000@snu.ac.kr

In brief

Park et al. present SuperST, an algorithm that enhances spatial transcriptomics by reconstructing high-resolution, dense gene expression matrices from low-resolution raw data. Using deep image prior, SuperST produces spatial patterns that closely resemble immunofluorescence images. The method overcomes limitations in resolution and zero inflation, enabling applications based on computer vision.

Park et al., 2025, Cell Reports Methods 5, 100937
January 27, 2025
https://doi.org/10.1016/j.crmeth.2024.100937

CellPress

Cell Reports Methods

CellPress
OPEN ACCESS

Report

Generation of super-resolution images from barcode-based spatial transcriptomics by deep image prior

Jeongbin Park,[1] Seungho Cook,[1] Dongjoo Lee,[1] Jinyeong Choi,[1] Seongjin Yoo,[1] Sungwoo Bae,[1] Hyung-Jun Im,[1,2,3,4,5] Daeseung Lee,[1] and Hongyoon Choi[1,6,7,8,*]

[1]Portrai, Inc., Dongsullagil, 78-18 Jongrogu, Seoul, Republic of Korea
[2]Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, Seoul National University, Seoul 08826, Republic of Korea
[3]Department of Applied Bioengineering, Graduate School of Convergence Science and Technology, Seoul National University, Seoul 08826, Republic of Korea
[4]Cancer Research Institute, Seoul National University, 03080 Seoul, Republic of Korea
[5]Research Institute for Convergence Science, Seoul National University, 08826 Seoul, Republic of Korea
[6]Department of Nuclear Medicine, Seoul National University Hospital, 03080 Seoul, Republic of Korea
[7]Department of Nuclear Medicine, Seoul National University College of Medicine, 03080 Seoul, Republic of Korea
[8]Lead contact
*Correspondence: chy1000@snu.ac.kr
https://doi.org/10.1016/j.crmeth.2024.100937

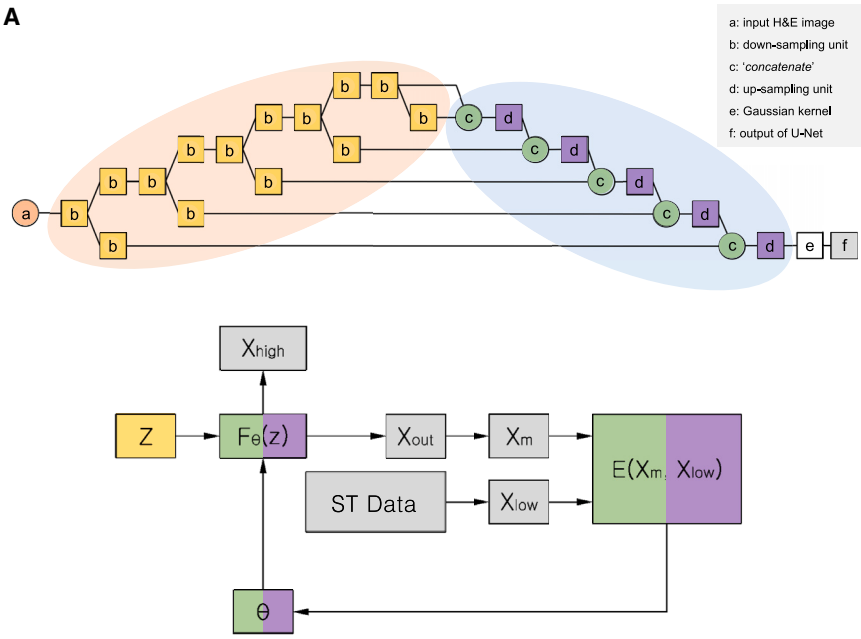
MOTIVATION Enhancing the resolution of barcode-based spatial transcriptomics can deepen our understanding of biological phenomena using both existing and newly generated data. We propose a method, SuperST, that aims to overcome the limitations of spatially resolved transcriptomics (ST) methods that rely on segmentation and/or exhibit parameter-dependent performance. SuperST utilizes an end-to-end convolutional neural network and is robust to parameters such as image size, learning epochs, or learning rate. To enhance usability, we integrate Visium-specific diffusion features and offer user-friendly Python-based implementations.

SUMMARY

Spatially resolved transcriptomics (ST) has revolutionized the field of biology by providing a powerful tool for analyzing gene expression *in situ*. However, current ST methods, particularly barcode-based methods, have limitations in reconstructing high-resolution images from barcodes sparsely distributed in slides. Here, we present SuperST, an algorithm that enables the reconstruction of dense matrices (higher-resolution and non-zero-inflated matrices) from low-resolution ST libraries. SuperST is based on deep image prior, which reconstructs spatial gene expression patterns as image matrices. Compared with previous methods, SuperST generated output images that more closely resembled immunofluorescence images for given gene expression maps. Furthermore, we demonstrated how one can combine images created by SuperST with computer vision algorithms. In this context, we proposed a method for extracting features from the images, which can aid in spatial clustering of genes. By providing a dense matrix for each gene *in situ*, SuperST can successfully address the resolution and zero-inflation issue.

INTRODUCTION

Spatially resolved transcriptomics (ST) has become an emerging technology in biology. There are currently two types of ST methods: barcode-based ST and image-based ST.[1] Barcode-based ST, also known as spot-based ST, such as 10x Visium,[2] differs from image-based ST[1] in that it provides whole gene expression data and typically involves thousands of spots. However, there are tens of cells in a spot in the barcode-based ST and empty spaces outside the barcodes, making it difficult to interpret the underlying biology and cellular-level phenomena. Additionally, RNA diffusion in tissues during experiments limits the resolution of ST, particularly in the case of barcode-based ST. The reported velocities of moving RNAs or RNA-protein complexes range from 0.65 to 1.5 μm/s,[3–6] depending on the tissues, conditions, and molecules of interest.



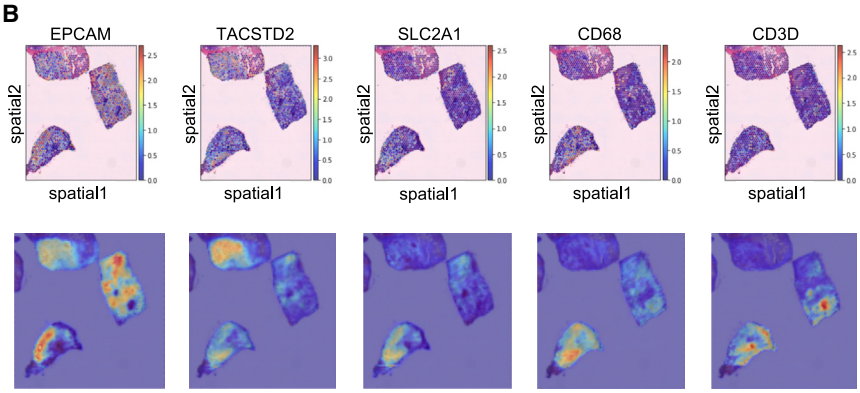


Figure 1. The design of SuperST

(A) The schematic representation of SuperST. Here, a, b, c, d, e, and f represent *input_1*: an input H&E image, a conceptual down-sampling unit, *concatenate*, a conceptual up-sampling unit, a Gaussian smoothing kernel, and the output of U-Net, respectively. Note that while DIP is a method for image restoration using neural networks, U-Net refers to a specific type of neural network architecture that is often utilized to implement the DIP method. The algorithm runs with $F_\theta(z) \rightarrow X_{out}$ by *num_iter* times of execution at first for updating U-Net and once with $F_\theta(z) \rightarrow X_{high}$ for predicting high-resolution images. "F" denotes a conceptual function that links the CNN architectures (b) to the output (f). The matching parts in the top and bottom images are shown in the same color. The detailed information can be found in the STAR Methods section.

(B) The comparison of conventional spatial feature plots and high-resolution images made by SuperST from a publicly available breast cancer dataset. The pixels in high-resolution images darker than 95 percentiles of each image are not shown.

To date, several algorithms have been developed to improve the resolution of ST. BayesSpace used the Bayesian approach to predict sub-spot gene expression, considering the structure of neighboring spots.[7] The spatially aware dimension reduction method (SpatialPCA) extracted a low-dimensional representation using probabilistic principal component analysis (PCA) and predicted gene expression in unmeasured spots in a spatially aware imputation manner.[8] Deep spatial data fusion, also known as XFuse, was a histology-dependent super-resolution image generation algorithm using a deep-generative model for Visium ST.[9] A recent algorithm, tumor edge structure and lymphocyte multi-level annotation (TESLA), involves determining neighborhood relationships from H&E images and imputing super-pixels to generate super-resolution gene expression images.[10] Another recent algorithm, iStar, uses a pre-trained hierarchical vision transformer (HViT) to extract tissue features at multiple scales from histology images, which are then used by a feedforward neural network to predict super-pixel-level gene expressions.[11]

There have also been various attempts to assume gene expression in unknown regions by introducing well-known probabilistic models, such as Gaussian distribution, negative binomial distribution, and Poisson distribution.

We defined a dense matrix as a matrix with a resolution as high as a regular image and appropriately improved from zero-inflated data. A complete dense matrix would be a stack of two-dimensional (2D) images, each representing a different gene map. Due to the limited performance of the previous methods for enhancing ST resolution, reconstructing dense image matrices from barcode-based data is still challenging. However, by treating the ST data as dense image matrices, it is possible to apply image processing methods such as segmentation,[12] registration, contouring, transformation, and integration[13,14] for gene expression data. This approach allows us to effectively employ ST data beyond just obtaining high-resolution data, advancing the field of ST and enhancing our understanding of gene expression *in situ*. In this regard, we proposed and validated an algorithm, SuperST, to generate dense matrices from low-resolution Visium ST libraries by applying deep image prior (DIP).

RESULTS

Dense image matrices of gene expression from ST data by DIP

SuperST was based on an exemplary DIP algorithm[15] (Figures 1A and S1). Briefly, DIP is a deep learning approach that reconstructs

high-quality images from low-quality inputs, such as noisy or blurred images, without the need for pre-labeled training data. However, SuperST can generate the output image even when a random matrix is used as input, although this will impact the performance of SuperST (Figure S2). It uses an untrained neural network, in this case, U-Net,[16] to constrain the image reconstruction process and utilize the intrinsic capability of the network to capture the underlying latent structure of the image data, creating high-resolution images. For DIP, there were two modules in the algorithm: down-sampling (colored in orange) and up-sampling (colored in blue), shown in Figure 1A. The input for this process was the histology image instead of a noisy image, and the output was generated through U-Net. Specifically, while the image size can be easily modified in SuperST, we designated the size of the gene expression map to be a 256 × 256 matrix for each gene. We resized the histology image, the input for U-Net, to 256 × 256 and fed it into the U-Net to predict a set of gene expression matrices of specific genes. We then applied a convolution term of Gaussian smoothing to account for the diffusion of mRNA. This term considered the diffusion of mRNA expression in the periphery for each position, which was counted as gene expression for each position of the spot. In Visium, the extent of mRNA diffusion can differ, so SuperST allows for adjusting the size of the Gaussian kernel or even omitting this step altogether. In our code, *demask_image_t* indicates SuperST images assuming zero diffusion, while *demask_image_t_g* means SuperST images with a modifiable Gaussian kernel size specified by *kernel_size*. The output 256 × 256 image was then masked only with pixels corresponding to the central position of each spot, and a loss function was created with the difference between the value of each masked pixel and the corresponding gene expression value for each spot. To minimize this loss function, the weight of the neural network model was iteratively updated. As a result, the dense matrices for the gene expression were generated, followed by the masking process. The example of the human breast cancer samples[17] is presented in Figure 1B.

Optimal parameters for SuperST

To evaluate the performance of SuperST in predicting the image matrix of specific gene expression data, we compared the output images obtained from SuperST with immunofluorescence (IF) images. For the comparisons, we selected genes reported in a study for their notable correlation between RNA and protein levels: PTPRC (Spearman = 0.79, p = 2.08E−06), CD3G (Spearman = 0.74, p = 0.01), and Pecam1 (Spearman = 0.73, p = 1.29E−05).[18] Rbfox3, well known as a mature neuron marker,[19] was also considered in our analysis, as it is expected to show a strong correlation between its RNA and protein abundance.

We assessed the correspondence between the outputs of SuperST generated from gene expression images and the IF images obtained from tissues (Figure 2). This was achieved by evaluating the pixel-wise correlation between the two images. When exploring image similarity metrics including Pearson correlation coefficient (*pearson*) and mutual information (*mutual*), the variation of overall coefficients decreased when the number of iterations (*num_iter*) increased, but the similarity tended to decrease further beyond a certain number of iterations (Figure S3). This could be due to the overfitting of SuperST, as can be seen in the case where *num_iter* equals 1,024 (Figure S4). However, with a small *num_iter*, the output image did not accurately represent the input gene expression. For example, *S100a9*, a well-known marker gene for neutrophils, was expressed in the inner necrotic core of the 4T1 tumor,[20] as revealed by the outcomes of a cell type decomposition algorithm[21] (Figure S5A). When *num_iter* was less than 128, the output images did not show the predominant *S100a9* expression in the core region, compared with the conventional spatial mapping (Figure S5B). Based on these findings, the *num_iter* parameter was chosen within a range of 128–512. Notably, choosing the *num_iter* parameter is crucial. Setting it too high makes images overly reflect gene expression sparsity, while setting it too low underestimates the gene expression. An optimal range of the *num_iter* ensures that SuperST performs best, and determining the appropriate range is not challenging in a practical sense. The performance of SuperST was further examined across varying output dimensions in multiples of two required for the U-Net model. Analogous to the observations made with a 256 × 256 output size, the 128 × 128 and 512 × 512 dimensions exhibited a comparable overfitting tendency as the *num_iter* approached 1,024 (Figures S3E–S3L). Additionally, consistent with previous findings, a local maximum in Pearson correlation coefficients was identified within the *num_iter* range of 128–512, aligning with the patterns previously delineated.

Comparison with other methods

As previous methods to provide high-resolution image matrices from ST data, XFuse, TESLA, and iStar were conducted for four different datasets, along with SuperST and conventional visualization (Figure 2B). It was noteworthy that the output images of the super-resolution algorithms better reflected the gene expression profiles than sparse Visium data, which poorly exhibited continuous gene expression due to the nature of zero-inflated data. SuperST showed the highest Pearson correlation coefficients between the output images and the IF images. XFuse and TESLA exhibited inappropriately contoured outputs in some cases (Figure 2B, *Pecam1* for XFuse and *Rbfox3* for TESLA). Also, there were no available outputs for certain genes (e.g., *CD3G*) when running XFuse in a default setting, while SuperST can reliably generate dense matrices for all genes without being greatly constrained by parameters such as image size, learning epochs, or learning rate. Moreover, discrepancies were identified between the output images of XFuse, TESLA, or iStar and the IF images. For example, in the *Mouse 4T1* sample, the inner necrotic section and the outer stromal compartment in the anti-Pecam1 antibody IF image were more accurately depicted in SuperST when compared with XFuse, TESLA, and iStar.

Application of computer vision algorithms with SuperST

One advantage of the dense matrix resulting from applying SuperST is that it can be utilized by various computer vision algorithms that are difficult to use on traditional ST data. Typically, these algorithms are designed to process pixel-based images,

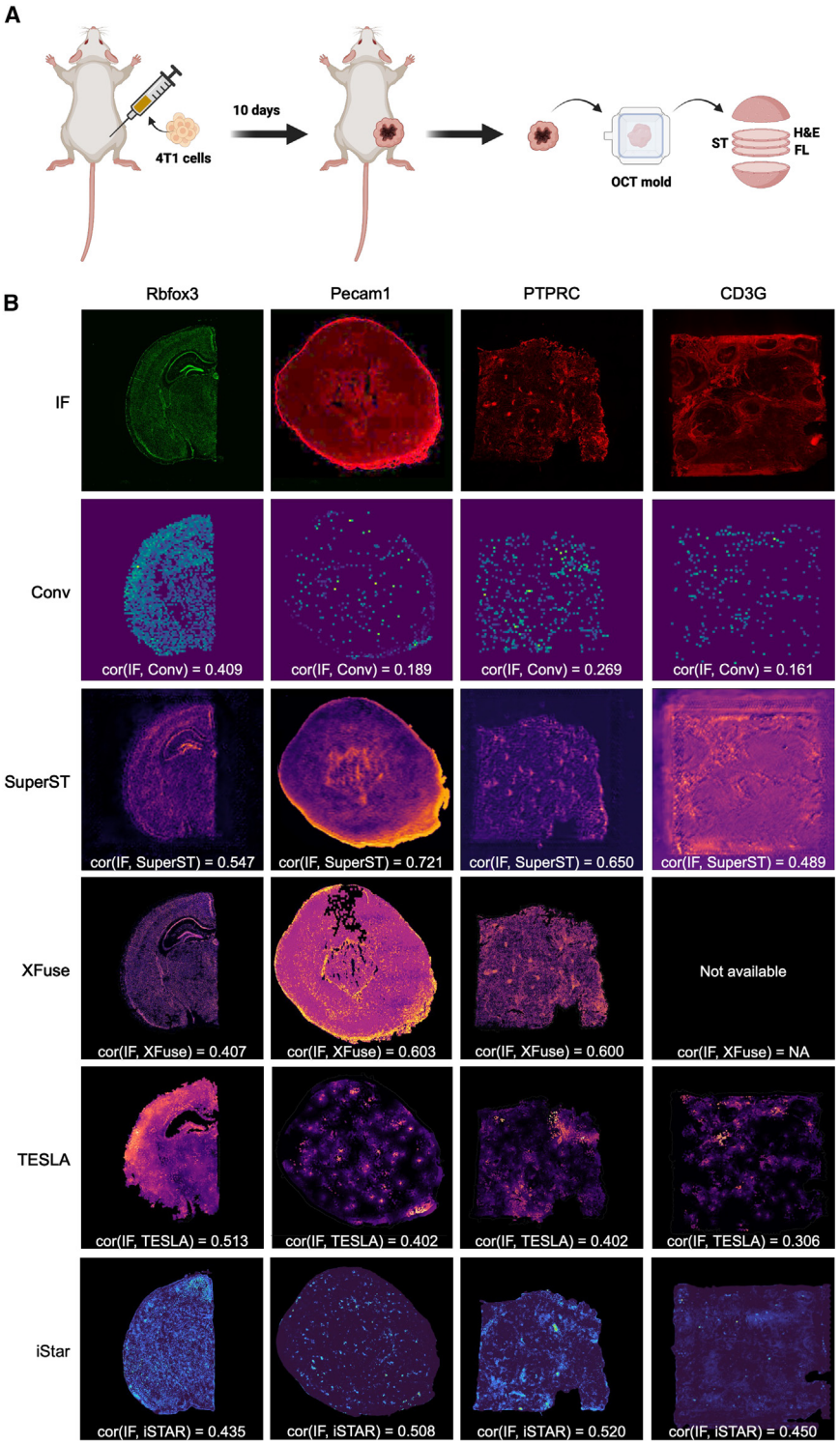


Figure 2. Comparison between conventional visualization, SuperST, XFuse, TESLA, and iStar with the IF image

(A) The acquisition of Visium ST library for *Mouse 4T1* with a recommended tissue preparation protocol, followed by the primary and the fluorescent secondary antibody treatment. H&E, ST, and FL refer to the H&E-stained tissue, the tissue used for acquiring the Visium ST library, and the tissue utilized for fluorescence imaging, respectively.

(B) Each Pearson correlation coefficient between IF and each super-resolution algorithm (SuperST, XFuse, TESLA, and iStar) output was represented for *Mouse Brain* (*Rbfox3*), *Mouse 4T1* (*Pecam1*), *Human Ovarian* (*PTPRC*), and *Human Ductal* (*CD3G*). *Conv*, indicating the conventional visualization of ST data (negative control), was also compared with IF. Super-resolution algorithms outperformed the conventional visualization in most cases. Also, the correlation coefficient was found to be the highest in SuperST compared with the other algorithms.

which differ from the nature of ST spots. We combined VGG16, one of the pre-trained convolutional neural network (CNN) models, with SuperST to extract image features from gene expression data. As a result, this approach enabled the extraction of 512D feature vectors from each super-resolution gene expression image, the output of SuperST. Utilizing these feature vectors, distinct clusters of genes exhibiting similar spatial expression patterns were identified through the application of K-means clustering (K = 3). This method effectively grouped genes based on their expression profiles, highlighting the spatial distribution and co-expression patterns within the tissue samples (Figure 3). In other words, the output of SuperST and the pre-trained CNN model can be utilized to obtain gene sets with similar spatial expression characteristics based on the pattern of spatial distribution of each gene (i.e., spatially variable gene sets). The score maps created by three gene clusters were associated with cancer cell division, fibrous tissue, and neutrophil chemotaxis, respectively (Figure 3B). This observation aligns with three distinct areas within the 4T1 solid tumor: regions densely enriched with active cancer cells, peripheral regions of the tumor with abundant fibroblasts, and necrotic tissues, where neutrophils mainly accumulate (Figure 3C).[20] Similarly, the results of spot clustering obtained based on gene expression showed similarity to the score maps in terms of the number of clusters and their spatial distribution (Figures S5C and S5D). Furthermore, the clustering based on pixels was shown, demonstrating the reliance of SuperST on *num_iter* as demonstrated earlier. (Figure S5E).

CellPress
OPEN ACCESS

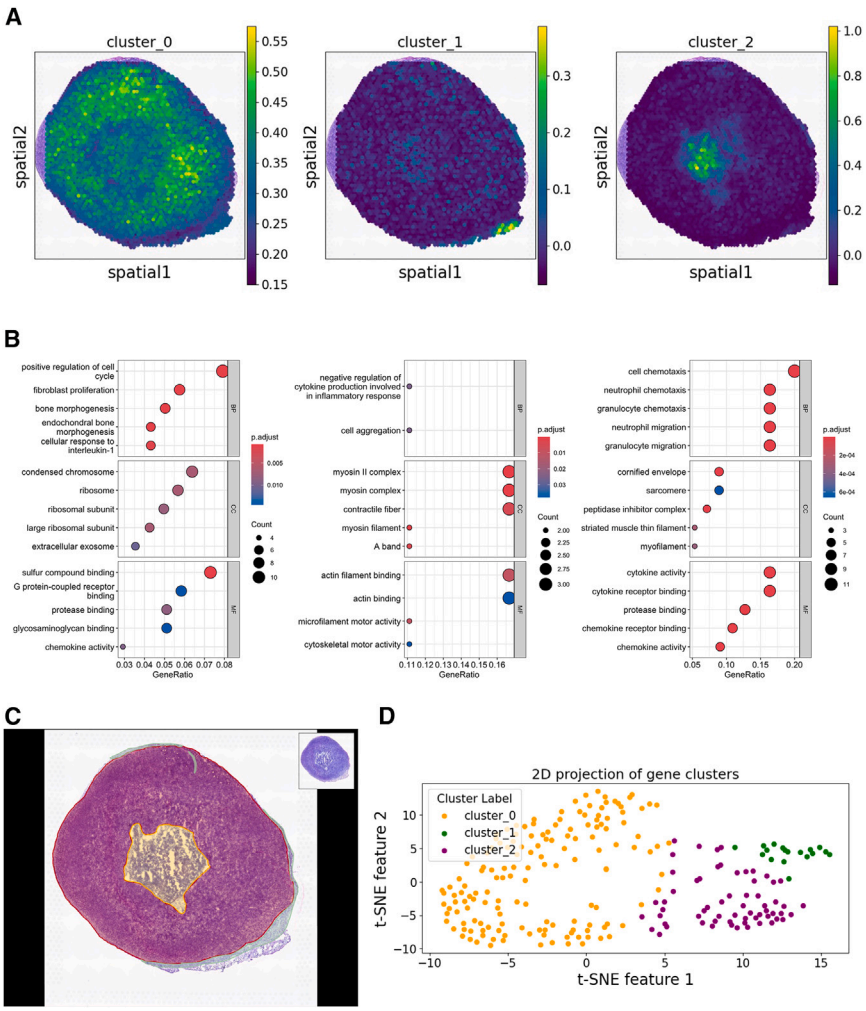


Figure 3. The spatial gene clustering with SuperST and pre-trained CNN algorithm

(A) The resultant three gene clusters (i.e., cluster_0, cluster_1, and cluster_2) derived from SuperST and VGG16, a representative pre-trained CNN algorithm. Firstly, SuperST was performed for spatially variable genes (SVGs) derived from highly variable genes (HVGs). Secondly, features were extracted from the images generated by SuperST using VGG16. Then, the 512D features were clustered using K means clustering (K = 3). Following this, each gene list from a cluster was scored. Consequently, three specific zones were identified, corresponding to the cancer-rich region, the fibrotic tissue, and the necrotic core.

(B) Gene Ontology (GO) plots corresponding to the spatial gene clusters. Each GO plot corresponds to the spatial feature plot directly above it in (A).

(C) The histological annotation made by a pathologist. The tumor (red) is relatively homogeneous and solidly distributed, with a thin layer of stroma (green) observed around it. The central part of the tumor (yellow) differs from the surrounding area because the tissue has fallen off the slide, and it is presumed to be a part of the tumor that has undergone necrosis.

(D) The distribution of SVGs in a 2D projection, displaying distinct borders between gene clusters.

DISCUSSION

The SuperST algorithm proposed in this study aims to address the challenge of reconstructing dense image matrices from barcode-based ST data by predicting gene expression data as images using the DIP algorithm. The study demonstrated that DIP produced high-resolution image matrices of gene expression from low-resolution ST data. The study evaluated the performance of SuperST by comparing the output images obtained from SuperST with IF images and assessing the correspondence between both images using image similarity metrics. The results showed that SuperST generated output images that were highly correlated with IF images.

While interpolation assumes a specific spatial distribution for gene expression, super-resolution algorithms, including SuperST, do not rely on such assumptions. SuperST can generate varying output images in a probabilistic way, and it can also provide clues for optimal parameters through semi-parameter independence (i.e., existence of a range of parameters that result in constantly optimal outputs). After searching for the best parameters, we conducted a comparison between the SuperST results and IF images to examine the molecular expression in the tissue. It is noteworthy that IF detects proteins that are not identical to mRNA transcripts of ST.[22,23] Therefore, we concentrated on genes that are recognized for their strong correlations between RNA and protein levels. As a result, SuperST exhibited a stronger correlation with the spatial patterns observed in IF, surpassing other methods (i.e., XFuse, TESLA, and iStar).

We have developed SuperST as an algorithm designed to provide high resolution and broad gene coverage simultaneously. It provides a way to overcome the limitations of barcode-based ST, which can only collect data from specified positions and struggles to interpret empty spaces. Image-based methods such as Xenium and multiplexed error-robust FISH (MERFISH) inherently offer high resolution, thus diminishing the need for super-resolution algorithms. However, these techniques do not adequately encompass the entire transcriptome with respect to gene coverage. Even as the coverage of gene panels expands, challenges persist regarding the efficiency of probe hybridization, which can vary by gene, inaccurate cell segmentation algorithm, and batch-to-batch consistency.[24] For methods like Stereo-seq and Seq-scope, although they provide higher resolution compared with Visium, the diffusion issue inherent to the approaches significantly impacts their performance, potentially necessitating the use of super-resolution imaging. The use of SuperST proposes several significant implications for ST. For example, various spatial omics modalities,

such as IF images, can be aligned with ST data by effortlessly improving the ST data's resolution using SuperST to closely match the corresponding omics data. Also, SuperST showed the potential to be applied in imaging analysis workflows, as can be seen in the feature extraction for spatial gene clustering by using a pre-trained CNN architecture (i.e., VGG16). In addition, the major differences compared to previously proposed algorithms include the capability to address the diffusion problems in ST, the ability to produce an appropriate mixture of input image and gene expression by adjusting *num_iter*, semi-parameter independence, and the design as an end-to-end AI algorithm, which facilitates ease of use and robust performance.

Although we validated SuperST using spatially matched IF data, further validation with large-scale, multi-molecular expression studies will enhance the reliability and applicability of this algorithm. Given these considerations, to validate the performance of SuperST transcriptome-wide, we conducted an additional study using paired Visium-Xenium data from a human patient with breast cancer (Figure S6). Following the benchmarking guidelines outlined in the iStar literature, we quantified the correspondence with Xenium using the structural similarity index measure (SSIM) instead of Pearson correlation and root-mean-square error (RMSE) and reconstructed pseudo-Visium gene expression based on Xenium. The results demonstrated that SuperST outperforms conventional visualization, XFuse, TESLA, and iStar. Notably, when XFuse was executed, there was no result of *APOBEC3B*, and the performance of XFuse and iStar showed SSIM values at the level of H&E images, which are unrelated to gene expression. The difficulty of these methods seems to stem from the challenge of applying appropriate hyperparameters that can adapt to the specific characteristics of each dataset. In this aspect, the advantage of SuperST, which can achieve robust performance through semi-parameter independence, is highlighted.

In conclusion, our results demonstrate that SuperST could generate dense image matrices with high-resolution information from barcode-based ST data, outperforming existing methods. Furthermore, the image data generated by SuperST could be easily integrated with other spatial omics technologies, such as proteomics and metabolomics, which could lead to a more comprehensive understanding of complex biological systems. Overall, SuperST displays great potential for advancing the field of ST and expanding our understanding of gene expression *in situ*.

Limitations of the study

The study also has a limitation that need to be considered. The proposed algorithm requires key parameters for DIP, which need to be optimized for generalization. It was discovered that the number of iterations (*num_iter*) is crucial and needs to fall within a specific range. However, it also needs to be customized according to the characteristics of the data provided. Also, this study focuses solely on *num_iter* for optimal SuperST performance, but other factors like image size, kernel size, and sigma value should be tuned for better performance. Despite these limitations, the SuperST algorithm provides an easy-to-use application with relatively robust performance.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Hongyoon Choi (chy1000@snu.ac.kr).

Materials availability

The study did not produce any unique reagents.

Data and code availability

- The ST dataset generated for this study, *Mouse 4T1*, is available upon a reasonable request from the lead contact. Otherwise, all the other datasets used in this study are publicly available and are listed in the key resources table.
- The scripts produced for this study are available in our GitHub repository (https://github.com/portrai-io/SuperST). An archival DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This research was supported by the National Research Foundation of Korea (NRF-2020M3A9B6038086, NRF-2023R1A2C2006636, and NRF-2022M3A9D3016848) and Korea Medical Device Development Fund grant funded by the Korean government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (project nos. 1711137868 and RS-2020-KD000006), as well as a grant from the Korea Technology and Information Promotion Agency (RS-2023-00304101).

AUTHOR CONTRIBUTIONS

Conceptualization, H.C.; methodology, H.C. and J.P.; software, J.P. and S.Y.; validation, J.P. and Dongjoo Lee; formal analysis, J.P.; investigation, J.P. and J.C.; resources, S.Y.; data curation, J.P.; writing – original draft, J.P., H.-J.I., Daeseung Lee, and H.C.; writing – review & editing, all authors; visualization, J.P. and S.B.; supervision, S.C.; project administration, H.C.; funding acquisition, H.C. and H.-J.I.

DECLARATION OF INTERESTS

H.-J.I., Daeseung Lee, and H.C. are the co-founders of Portrai, Inc. H.C., H.-J.I., J.P., and Daeseung Lee were listed as inventors in patent applications for the technology presented in this manuscript (Korean patent application no. 10-2023-0060674).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2024.100937.

Received: January 10, 2024
Revised: June 17, 2024
Accepted: December 4, 2024
Published: December 26, 2024

REFERENCES

1. Moffitt, J.R., Lundberg, E., and Heyn, H. (2022). The emerging landscape of spatial profiling technologies. Nat. Rev. Genet. *23*, 741–759.
2. 10x Genomics. Spatial Gene Expression. https://www.10xgenomics.com/products/spatial-gene-expression.
3. Mor, A., Suliman, S., Ben-Yishay, R., Yunger, S., Brody, Y., and Shav-Tal, Y. (2010). Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. Nat. Cell Biol. *12*, 543–552.
4. Shav-Tal, Y., Darzacq, X., Shenoy, S.M., Fusco, D., Janicki, S.M., Spector, D.L., and Singer, R.H. (2004). Dynamics of single mRNPs in nuclei of living cells. Sci. Technol. Humanit. *304*, 1797–1800.
5. Park, H.Y., Lim, H., Yoon, Y.J., Follenzi, A., Nwokafor, C., Lopez-Jones, M., Meng, X., and Singer, R.H. (2014). Visualization of dynamics of single endogenous mRNA labeled in live mouse. Science *343*, 422–424. https://doi.org/10.1126/science.1239200.
6. Fusco, D., Accornero, N., Lavoie, B., Shenoy, S.M., Blanchard, J.M., Singer, R.H., and Bertrand, E. (2003). Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. Curr. Biol. *13*, 161–167. https://doi.org/10.1016/s0960-9822(02)01436-7.
7. Zhao, E., Stone, M.R., Ren, X., Guenthoer, J., Smythe, K.S., Pulliam, T., Williams, S.R., Uytingco, C.R., Taylor, S.E.B., Nghiem, P., et al. (2021). Spatial transcriptomics at subspot resolution with BayesSpace. Nat. Biotechnol. *39*, 1375–1384.
8. Shang, L., and Zhou, X. (2022). Spatially aware dimension reduction for spatial transcriptomics. Nat. Commun. *13*, 7203. https://doi.org/10.1038/s41467-022-34879-1.
9. Bergenstrahle, L., He, B., Bergenstrahle, J., Abalo, X., Mirzazadeh, R., Thrane, K., Ji, A.L., Andersson, A., Larsson, L., Stakenborg, N., et al. (2022). Super-resolved spatial transcriptomics by deep data fusion. Nat. Biotechnol. *40*, 476–479. https://doi.org/10.1038/s41587-021-01075-3.
10. Hu, J., Coleman, K., Zhang, D., Lee, E.B., Kadara, H., Wang, L., and Li, M. (2023). Deciphering tumor ecosystems at super resolution from spatial transcriptomics with TESLA. Cell Syst. *14*, 404–417.e4.
11. Zhang, D., Schroeder, A., Yan, H., Yang, H., Hu, J., Lee, M.Y.Y., Cho, K.S., Susztak, K., Xu, G.X., Feldman, M.D., et al. (2024). Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. Nat. Biotechnol. *42*, 1372–1377. https://doi.org/10.1038/s41587-023-02019-9.
12. Lee, D., Park, J., Cook, S., Yoo, s.-j., Lee, D., and Choi, H. (2023). IAM-SAM: Image-based Analysis of Molecular signatures using the Segment-Anything Model. Preprint at bioRxiv. https://doi.org/10.1186/s13059-024-03380-x.
13. Bae, S., Choi, H., and Lee, D.S. (2021). Discovery of molecular features underlying the morphological landscape by integrating spatial transcriptomic data with deep features of tissue images. Nucleic Acids Res. *49*, e55. https://doi.org/10.1093/nar/gkab095.
14. Bergenstråhle, J., Larsson, L., and Lundeberg, J. (2020). Seamless integration of image and molecular analysis for spatial transcriptomics workflows. BMC Genom. *21*, 482. https://doi.org/10.1186/s12864-020-06832-3.
15. Ulyanov, D., Vedaldi, A., and Lempitsky, V. (2018). Deep Image Prior. pp. 9446-9454.
16. Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation (Springer), pp. 234–241.
17. Wu, S.Z., Al-Eryani, G., Roden, D.L., Junankar, S., Harvey, K., Andersson, A., Thennavan, A., Wang, C., Torpy, J.R., Bartonicek, N., et al. (2021). A single-cell and spatially resolved atlas of human breast cancers. Nat. Genet. *53*, 1334–1347. https://doi.org/10.1038/s41588-021-00911-1.
18. Wang, D., Eraslan, B., Wieland, T., Hallström, B., Hopf, T., Zolg, D.P., Zecha, J., Asplund, A., Li, L.H., Meng, C., et al. (2019). A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol. Syst. Biol. *15*, e8503. https://doi.org/10.15252/msb.20188503.
19. Ianevski, A., Giri, A.K., and Aittokallio, T. (2022). Fully-automated and ultra-fast cell-type identification using specific marker combinations from single-cell transcriptomic data. Nat. Commun. *13*, 1246. https://doi.org/10.1038/s41467-022-28803-w.
20. Park, J., Choi, J., Lee, J.E., Choi, H., and Im, H.J. (2022). Spatial Transcriptomics-Based Identification of Molecular Markers for Nanomedicine Distribution in Tumor Tissue. Small Methods *6*, e2201091. https://doi.org/10.1002/smtd.202201091.
21. 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. https://doi.org/10.1093/nar/gkac084.
22. Maier, T., Güell, M., and Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. FEBS Lett. *583*, 3966–3973.
23. Koussounadis, A., Langdon, S.P., Um, I.H., Harrison, D.J., and Smith, V.A. (2015). Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. Sci. Rep. *5*, 10775.
24. Williams, C.G., Lee, H.J., Asatsuma, T., Vento-Tormo, R., and Haque, A. (2022). An introduction to spatial transcriptomics for biomedical research. Genome Med. *14*, 68. https://doi.org/10.1186/s13073-022-01075-1.
25. 10x Genomics. Datasets. https://www.10xgenomics.com/resources/datasets.
26. 10x Genomics. Human Ovarian Cancer: Targeted, Immunology Panel. Stains: DAPI, Anti-PanCK, Anti-CD45. Retrieved from https://www.10xgenomics.com/resources/datasets/human-ovarian-cancer-targeted-immunology-panel-stains-dapi-anti-pan-ck-anti-cd-45-1-standard-1-2-0.
27. 10x Genomics. Invasive Ductal Carcinoma Stained With Fluorescent CD3 Antibody. Retrieved from https://www.10xgenomics.com/resources/datasets/human-ovarian-cancer-targeted-immunology-panel-stains-dapi-anti-pan-ck-anti-cd-45-1-standard-1-2-0.
28. Janesick, A., Shelansky, R., Gottscho, A.D., Wagner, F., Williams, S.R., Rouault, M., Beliakoff, G., Morrison, C.A., Oliveira, M.F., Sicherman, J.T., et al. (2023). High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. Nat. Commun. *14*, 8353. https://doi.org/10.1038/s41467-023-43458-x.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rat anti-PECAM1 primary antibody	Invitrogen	14-0311-82, RRID: AB_467201
Alexa Fluor 647 conjugated goat anti-rat IgG secondary antibody	Invitrogen	A21247, RRID: AB_141778
Deposited data		
Public human breast cancer Visium data[17]	Wu et al.[17]	https://doi.org/10.5281/zenodo.3957257
Mouse Brain Visium[25]	10x Genomics	https://www.10xgenomics.com/resources/datasets
Mouse 4T1 Visium	This paper	Available upon a reasonable request from the lead contact
Human Ovarian Visium[26]	10x Genomics	https://www.10xgenomics.com/resources/datasets
Human Ductal Visium[27]	10x Genomics	https://www.10xgenomics.com/resources/datasets
Public human breast cancer paired Visium-Xenium data[28]	Janesick et al.[28]	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE243280
Experimental models: Cell lines		
4T1 tumor cells	ATCC	N/A
Experimental models: Organisms/strains		
A 6-week-old, female BALB/c mouse	SNU, College of Medicine	N/A
Software and algorithms		
SuperST (GitHub)	This paper	https://github.com/portrai-io/SuperST
SuperST (Zenodo)	This paper	https://zenodo.org/records/14207705
XFuse	Bergenstråhle et al.[9]	https://github.com/ludvb/xfuse
TESLA	Hu et al.[10]	https://github.com/jianhuupenn/TESLA
iSTAR	Zhang et al.[11]	https://github.com/daviddaiweizhang/istar

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Tissue slice preparation

The study used both public datasets and generated a new ST dataset (*Mouse 4T1*) from a 6-week-old, female BALB/c mouse bearing a 4T1 tumor. The mouse was prepared by injecting 4T1 tumor cells, which had a passage number of 30–40 and were not genetically modified, into its right thigh region. The tumor sample was collected and frozen 10 days after injection. Three tissue slices were acquired from the tumor tissue.

Data acquisition

One slice was treated with rat anti-PECAM1 primary antibody (14-0311-82, Invitrogen) and Alexa Fluor 647 conjugated goat anti-rat IgG secondary antibody (A21247, Invitrogen). The IF image was then obtained by observing it with confocal microscope (STELLARIS 5, Leica microsystems). Another slice was used for the H&E staining. The last slice was treated with methanol, frozen, cryo-sectioned, and used to generate a block embedded in optical cutting temperature (OCT) compound (25608-930, VWR, USA) and the ST library. The preparation of the ST library followed a recommended Visium protocol of 10x Genomics.

Ethics declaration

Animal experiments for this study were approved by the Woojung Bio IACUC of the Republic of Korea with the approval code of IACUC2001-003.

METHOD DETAILS

Datasets

A public breast cancer dataset was used to compare the results with and without SuperST.[17] Three public datasets were additionally utilized for the validation of SuperST. One dataset (*Mouse Brain*) was 'Adult Mouse Brain Section 2 (Coronal)' data publicly available from 10x Genomics.[25] The tissue in the dataset was treated with anti-GFAP antibody (glial fibrillary acidic protein, red) and anti-NeuN

antibody (mature neuron marker, green) along with DAPI staining (blue). Also, two datasets provided by 10x Genomics including the human ovarian cancer sample (*Human Ovarian*; CD45: Cy5 - red)[26] and the human invasive ductal carcinoma sample (*Human Ductal*; CD3: Alexa Fluor 647 - red)[27] were used.

SuperST algorithm construction

SuperST generates a high-resolution image X_{high} from a low-resolution expression matrix X_{low} by employing U-Net to realize DIP. It applies down-sampling and up-sampling modules, with an adjustable Gaussian smoothing kernel to account for mRNA diffusion in the ST data. The following equation illustrates the mechanism of SuperST.

$$X_{out} \text{ or } X_{high} = F_{\theta}(Z)$$

Here, $Z \in \mathbb{R}^{C' \times H' \times W'}$ is C′ numbers of CNN architectures of height H′ and width W′ learned from X_0. X_{out} or $X_{high} \in \mathbb{R}^{3 \times H \times W} = \mathbb{R}^{3 \times 256 \times 256}$ is the image derived from Z. θ is a network parameter and F is a function made up of deep learning network. Also, the input image X_0 is a histology image, either 'spatial/tissue_hires_image.png' or 'spatial/tissue_lowres_image.png,' resized to 256 × 256 for U-Net processing. Alternatively, random images may also be used. After generating an output high-resolution image X_{out} from F, it is converted to low-resolution matrix X_m to be compared with the low-resolution ST gene expression matrix of a gene, X_{low}. Here, θ is determined to minimize the cost function of E between X_m and X_{low}.

Parameters

DIP, originally utilizing overfitting, does not exhibit stable performance with an increase in the number of epochs ('*num_iter*') compared with conventional deep learning. Instead, it produces output images that converge to the input gene expression. Therefore, adjusting the *num_iter* in the SuperST algorithm using DIP is quite crucial, and the rate at which SuperST converges to the input gene expression can vary depending on the genes and tissue types, inevitably leading to different optimal *num_iter* values. However, it's important to note that during the transition from the input image to the input gene expression, SuperST demonstrates an interval (128–512) of stable and optimal performance, which is relatively consistent across different samples and can even be easily determined through visual inspection.

Computer vision analysis

SuperST was used to define distinct tissue regions by combining with a pre-trained deep learning algorithm, VGG16, which recognizes textural patterns from images. The *Mouse 4T1* data was used to apply SuperST. Firstly, 1,000 highly variable genes (HVGs) were identified using *scanpy.pp.highly_variable_genes* Python package. Then, 233 spatially variable genes (SVGs) were isolated from HVGs by using *squidpy.gr.sepal* Python package. In the following step, SuperST leverages DIP to generate dense image matrices from the low-resolution spatial expression of the 233 SVGs with *num_iter* of 400. The final high-resolution images were used to extract 512D feature vectors for each gene via a pre-trained convolutional neural network model (VGG16). Subsequently, we apply K-means clustering (K = 3) to these features to identify spatially and transcriptionally distinct tissue regions. Lastly, a gene score for each spatial gene cluster was calculated using *scanpy.tl.score_genes* Python package.

QUANTIFICATION AND STATISTICAL ANALYSIS

Benchmarking

When performing SuperST for *Mouse Brain* (Rbfox3), *Human Ovarian* (PTPRC), and *Human Ductal* (CD3G), the input images were derived by subtracting the fluorescent antibody channel from the given 3-channel images. Also, for *Mouse 4T1* (Pecam1), the H&E staining image was used as the input, representing the general use cases for SuperST. The inputs for XFuse and TESLA were the same as those for SuperST. Also, XFuse was prepared in a default setting according to the literature,[9] and the *invcv+* image was selected as the representative output image among the XFuse outputs (i.e., *stdv*, *mean*, and *invcv+* images). In addition, TESLA was run with *apertureSize* of 5, *res* of 3, and the first type of contouring algorithm of three. Lastly, the conventional visualization method of Visium data was achieved by enlarging spot-wise expressions in a square form.

Evaluation

To validate the results, Pearson correlation coefficients and mutual information metrics were computed between SuperST outputs and IF images, with each sample being assessed 20 times. Based on the correlation patterns across varying iteration counts, optimal parameters were determined. Similar procedures were applied to other methods, analogous to SuperST.

Validation via xenium

After acquiring the paired Visium-Xenium public data for human breast cancer,[28] we ran SuperST on 307 common genes between the two spatial transcriptomics datasets and obtained 307 dense images of 256 × 256 dimension. Subsequently, we rotated the SuperST results by about 90° and cropped them to map into the Xenium AnnData's obs. We then generated the outputs of SuperST as well as the conventional visualization, XFuse, TESLA, and iStar outcomes for the common genes, mapped them onto Xenium, and calculated the SSIM values with Xenium-based pseudo-Visium gene expression image. Also, we compared the SSIM values between each pair of methods, using Wilcoxon rank-sum test.