GCNLA: Inferring Cell-cell Interactions from Spatial Transcriptomics with Long Short-term Memory and Graph Convolutional Networks

Chao Yang, Xiuhao Fu, Zhenjie Luo, Leyi Wei, Jingbing Li, Feifei Cui, Quan Zou, Qingchen Zhang, and Zilong Zhang

***Abstract*— Spatial transcriptomics analysis methods offer an opportunity to investigate highly diverse biological tissues. Cell-cell communication is fundamental for maintaining physiological homeostasis in organisms and coordinating complex biological processes. Identifying cell-cell interactions is critical for understanding cellular activities. The interaction of a cell with other cells depends on several factors, and most of the existing methods that consider only gene expression information of neighbouring cells and spatial location information are somewhat limited. In this paper, we propose a network architecture based on graph convolution network and long short-term memory attention module-GCNLA, which contains graph convolution layer, long short-term memory network, attention module, and residual connections. GCNLA not only learns the spatial structure of cells but also captures interaction information between distal cells, the attention module further extracting and enhancing features related to cell-cell interactions. Finally, the inner product decoding calculates the cosine similarity, which is used to infer cell-cell interactions. In addition, GCNLA is capable of reconstructing the complete cell-cell interaction network. The experimental results on seqFISH and MERFISH demonstrate that the GCNLA network structure has better robustness and noise immunity. The potential features learned by GCNLA enable other downstream analyses, including single-cell resolution cell clustering based on spatial information resolving cell heterogeneity.**

***Index Terms*— Cell-cell interactions, Long short-term memory, Graph convolution network, Attention.**

The work was supported by the National Natural Science Foundation of China (No. 62131004, 62261018, 62262015) and Hainan Provincial Natural Science Foundation of China (324MS009). (*Corresponding author: Zilong Zhang.*)

Chao Yang, Xiuhao Fu, Zhenjie Luo, Feifei Cui, Qingchen Zhang, and Zilong Zhang are with the School of Computer Science and Technology, Hainan University, Haikou, 570228, China.(e-mail: zhangzilong@hainanu.edu.cn)

Leyi Wei is with the Centre for Artificial Intelligence driven Drug Discovery, Faculty of Applied Science, Macao Polytechnic University, Macao SAR, China, and also with the School of Informatics, Xiamen University, Xiamen, China.

Jingbing Li is with the School of Information and Communication Engineering, Hainan University, Haikou 570228, China.

Quan Zou is with Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, 610054, China, and also with the Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou, 324000, China.

The GCNLA is implemented by Python using the PyTorch framework. All source codes are available at https://github.com/sharonycc/GCNLA.

I. Introduction

IN biological systems, cells communicate information with each other by secreting signaling molecules directly or indirectly [1], [2]. The physical connection and exchange of information between cells by direct or indirect means is called cell-cell interactions (CCIs) [3]. CCIs can be short-range contact-dependent interactions or can be achieved by secreting signaling molecules. CCIs co-ordinate the development, homeostasis and single-cell functions of organisms [4]. Moreover, CCIs also influence physiological functions among cellular tissues to some extent [4]. The regulation of CCIs is essential for maintaining the normal function of tissues and organs, while their dysregulation may lead to a variety of diseases, including cancers, autoimmune diseases, and infections [5]–[9]. The intracellular information of individual cells is inextricably linked to interactions with the multicellular environment [10]–[12]. Therefore, understanding the function of individual cells in preserving tissue homeostasis and reacting to the surroundings requires the identification and inference of these connections [13]. However, due to the complexity and heterogeneity of CCIs, comprehensively resolving these interactions remains a challenge.

There have been a number of studies exploring CCIs. For example, Browaeys et al. [14] developed NicheNet, a method for predicting CCIs mediated by ligand-target linkages by combining expression data from interacting cells with a priori knowledge of signal transduction and gene regulatory networks. Cheng et al. [15] proposed scMLnet, which models functional intercellular communication and intracellular gene regulatory networks, to explore the microenvironmental regulation of SARS-CoV-2 receptor ACE2 expression. Wang et al. [16] proposed SoptSC, an optimization method based on similarity matrices, to infer cell-cell communication through structured inter-cell similarity matrices. Sun et al. [17] discovered the involvement of macrophages and LAMP3+ DCs in the regulation of T cell activity and interactions with TASCs based on single-cell RNA sequencing (scRNA-seq). Jin et al. [18] developed CellChat, which is able to infer and analyses CCIs networks from scRNA-seq and perform comparative analyses of CCIs under different conditions. Ma et al. [19] proposed CellCommuNet, which can provide information on the strength of CCIs and related signaling pathways, thereby

enabling the exploration of differences in CCIs between normal and pathological conditions. These methods described above are basically based on scRNA-seq. Although scRNA-seq can analyze differences in gene expression at the single-cell level, revealing cell type heterogeneity [20]–[23]. However,the spatial location of cells in tissues and the analysis of the spatial dependency of CCIs cannot be obtained using standard scRNA-seq [24].

In order to solve the above problems, Spatial Transcriptomics (ST) technology has emerged, which is able to perform gene expression analysis while maintaining spatial information, thereby providing a more effective means to explore CCIs [25]–[30]. Unfortunately, traditional ST data have a low resolution of gene expression [31]. Currently, many studies integrate single-cell transcriptomics with spatial information, as single-cell transcriptomics provides high-resolution gene expression profiles, while ST provides information on the spatial distribution of gene expression, making the two approaches complementary [32]–[34]. Cang et al. [35] proposed SpaOTsc, a structure-optimized transmission-dependent method that uses spatial measurements of fewer genes to recover the spatial properties of scRNA-seq, which in turn captures intercellular communication in Drosophila embryos. Dries et al. [36] presented Giotto, which explores CCIs by combining RNA-seq and ST. Recent advances in spatially resolved transcriptomics have made it feasible to quantify gene expression profiles at single-cell or subcellular precision, while maintaining information about the spatial location of cells [37]–[40].

CCIs networks often have complex topologies, which contain multiple types of cells and their interaction relationships. Graph Convolutional Networks (GCNs) are well-suited for graph-structured data, capturing complex relationships between nodes and edges [41], [42]. In CCIs tasks, GCNs can reveal connectivity among cells by leveraging their topological and feature-based associations [41], [42]. Meanwhile, GCNs are able to aggregate node features at different scales, including local neighborhood features and global graph structure information, which is crucial for studying CCIs [42]. Currently, there have been many studies using GCNs to handle CCIs tasks. Yuan et al. [43] developed GCNG, which encodes spatial information into graphs and combines it with expression data to infer interactions between genes via convolutional networks. Li et al. [44] developed DeepLinc, a graph neural network-based reconstruction of cellular interaction networks from single-cell spatial transcriptomics. Bafna et al. [11] proposed CLARIFY, a multilevel graph autoencoder that uses single-cell spatial transcriptome data to refine intracellular and extracellular interactions. Yang et al. [45] proposed DeepTalk, which resolves spot-based ST data and infer CCIs with graph-attention network. The studies presented above all used graph-related networks to construct interaction models, but it is extremely important to learn the dependencies between cells and other cells in CCIs in order to reconstruct a more complete graph of cellular communication [1].

Based on these considerations, we propose a graph autoencoder and Long Short-Term Memory (LSTM) attention framework-GCNLA, which combines GCNs and LSTM networks for inferring CCIs. CCIs are constructed as graph structures, where nodes represent cells and edges represent interactions. GCNs can effectively capture the spatial dependencies in the graph structure, LSTM modeling of potential dependency chains associated with CCIs messaging to further capture relevant features of cellular communication [46]. The combination of GCNs and LSTM enhances the prediction of CCIs and improves noise immunity and robustness of the overall model. At the same time, LSTM introduces the attention module, which can make the model focus more on effective information and effectively enhance the computational efficiency. Therefore, GCNLA is developed to refine CCIs by analyzing spatial transcriptome data to generate cell-level representations of potential features, which combine spatial location information and single-cell gene expression. To evaluate the performance, we compare GCNLA with existing methods in two datasets, and the results show that GCNLA is able to enhance the robustness and noise immunity of the model while improving the prediction of CCIs. The contribution of this paper is as follows:

1. We introduce a LSTM attentional unit that is capable of extracting distal dependencies of CCIs, and the attention module further enhances features associated with CCIs. The combination of GCNs handles the spatial topological relationships of cells efficiently, which makes the potential features more context-dependent and learns more complex cellular communication.

2. GCNLA is able to refine the reconstructed CCIs profile and explore distal CCIs. The reconstructed CCIs network shows high physiological relevance.

3. We compare GCNLA with other methods aimed at reconstructing CCIs profiles in seqFISH and MERFISH, confirming the advantages of GCNLA in inferring CCIs. GCNLA shows high robustness and noise immunity in inferring CCIs.

The structure of this work is organized as follows. Section II describes the research methodology, including the GCNLA architecture, and experimental details. Section III shows the experimental setup and results. Section IV discusses the implications of the results and the limitations. Finally, Section V summarizes the study.

II. Materials and Methods

We describe a framework based on graph autoencoders and LSTM attention networks. The method first constructs the graph structure, gets the potential feature representation by GCNs and long short-term attention networks, and finally trains the model according to the task requirements. The overall GCNLA network framework is shown in **Fig.** 1.

A. Cell Graph Construction

In this work, we hypothesized that neighbouring cells in biological tissues are more likely to interact in certain ways than distant, randomly chosen non-neighbouring cells. It is difficult to represent the full range of CCIs by relying only on distance relationships. We further hypothesize that the gene expression profile of each single cell contains features associated with CCIs.

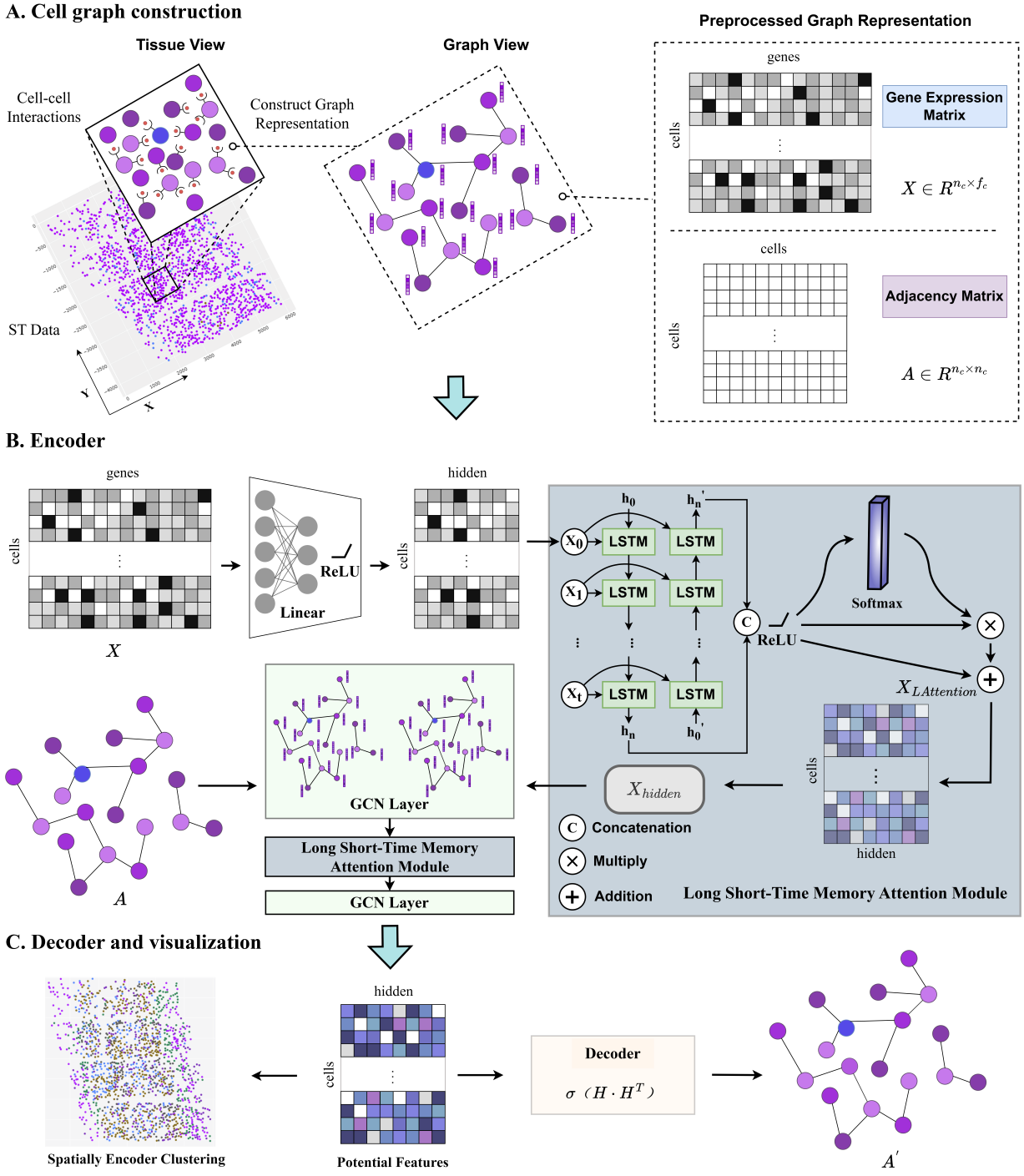


Fig. 1. The overall framework of the GCNLA network. **A** Construction of the cell graph. The first column represents the cell-cell interactions. The second column shows the modeling cell graph for the interactions. The third column represents the preprocessed graph representation matrix. **B** GCNLA encoder architecture. Consists of a long short-term memory attention module and a graph convolutional network. **C** The GCNLA decoder architecture.

In order to exploit the spatial component of the data, we determined CCIs based on spatial proximity. Therefore, we constructed graph structures at cellular level. We consider each cell to be a vertex of the graph structure, and construct edges of the graph structure based on a network of spatial connections, edges connecting cell vertices, i.e., CCIs. We represent the adjacency matrix used to describe cells and edges as $A \in \mathbb{R}^{n_c \times n_c}$, where n_c represents the number of cells. If there is an interaction between cell i and cell j, then $A_{i,j} = 1$. Conversely, if there is no interaction between cell i and cell j, then $A_{i,j} = 0$.

In order to understand the ability of these genes within a cell to interact with genes in adjacent cells, we identified genes that build cell graph using standard ligand-receptor database

[47]. We used the matrix X to represent the node features of the graph, which is the gene expression features of each cell in the graph (i.e., the vertex of the graph). Essentially, matrix $X \in \mathbb{R}^{n_c \times f_c}$ is essentially gene expression matrix, where n_c represents the number of cells and f_c represents the number of genes per cell.

Ultimately, the adjacency matrix A and the feature matrix X together construct the original cellular neighbourhood graph G, which was used as input data to train our graph network.

B. GCNLA Framework

The adjacency matrix A and the feature matrix X are fed to GCNLA. GCNLA is a method that combines GCNs and LSTM attention networks. The first stage of GCNLA is composed of linear layer, LSTM attention module, and GCNs. To avoid missing data and retain information from the original input data, we used skip connection after the graph convolutional network in the first stage. The second stage of training consists mainly of LSTM attention modules and GCNs. The feature matrix X are trained through two stages to obtain the latent representation. GCNs capture features of individual cells themselves and their neighbours, while LSTM capture features of cells' implicit interrelationships with other cells. In addition, Attention Modules help identify and emphasize key features, thereby improving the relevance and accuracy of feature representations. Hence, the latent representation integrates spatial information and intercellular dependencies. Subsequently, the decoder performs inner product on the latent representation to generate a reconstructed CCIs. The model is optimized according to the reconstruction capability of the cell adjacency matrix.

On the other hand, the latent representation acquired through training can represent potential information of cell interaction landscapes and gene expression profiles, making it useful for single-cell visualization and clustering tasks.

1) Long Short-Term Memory Attention Module: The LSTM Attention Module is composed of the LSTM Network and the Attention Mechanism.

LSTM is a particular kind of recurrent neural network (RNN) structure [46]. LSTM is improved on traditional RNNs. LSTM can effectively capture long and short-term dependencies and overcome the limitations of traditional RNNs, such as the gradient vanishing problem. We feed the linear layer-processed feature matrix into a two-layer bidirectional LSTM network to capture the distal dependencies between cells. Each LSTM layer is composed of multiple LSTM cells, each LSTM cells is composed of Cell State cs_t, Hidden State hs_t, and Gates. The cs_t is the memory part of the LSTM network that is able to maintain dependencies over long periods of time. The Hidden State hs_t is the output of LSTM at each time step. There are three key gates in the Gates, which are the Input Gate ig_t, the Output Gate og_t, and the Forget Gate fg_t. We assume that the input features at time t is x_t, which is computed as follows:

$$fg_t = \sigma\left(W_{if}x_t + b_{if} + W_{hf}hs_{t-1} + b_{hf}\right). \quad (1)$$

$$ig_t = \sigma\left(W_{ii}x_t + b_{ii} + W_{hi}hs_{(t-1)} + b_{hi}\right). \quad (2)$$

$$cg_t = \tanh\left(W_{ig}x_t + b_{ig} + W_{hg}hs_{(t-1)} + b_{hi}\right). \quad (3)$$

$$og_t = \sigma\left(W_{io}x_t + b_{io} + W_{ho}hs_{(t-1)} + b_{ho}\right). \quad (4)$$

$$cs_t = fg_t \odot cs_{t-1} + ig_t \odot cg_t. \quad (5)$$

$$hs_t = og_t \odot \tanh(cs_t). \quad (6)$$

Where cg_t represents the Cell Gate, $\odot$ is the Hadamard product, $\sigma(\cdot)$ is the sigmoid activation function, and W represents the weight. Here, we employ a bidirectional LSTM. To acquire the hidden state $\overleftrightarrow{hs_t}$ at time t, the forward LSTM hidden state $\overrightarrow{hs_t}$ and the backward LSTM hidden state $\overleftarrow{hs_t}$ are horizontally concatenated. The formula is as follows:

$$\overleftrightarrow{hs_t} = \left[\overrightarrow{hs_t}, \overleftarrow{hs_t}\right]. \quad (7)$$

The hidden states of all the cells are concatenated to obtain the output features of the LSTM. The LSTM processed features are fed into the attention module following the Rectified Linear Unit (ReLU) activation function. The softmax function is applied to the feature tensor to generate attention weights, and these weights are used to multiply the input features element by element to finally obtain the weighted features. The features of the ReLU activation function are skip-connected to the weighted features to obtain the output features of the LSTM attention module. The formula to calculate the attention module is as follows:

$$\begin{aligned} X_{\text{LAttention}} = {} & \text{ReLU}(X_{\text{LSTM}}) + \\ & \text{ReLU}(X_{\text{LSTM}}) \times \left(\text{softmax}\left(\text{ReLU}(X_{\text{LSTM}})^T\right)\right)^T \end{aligned} \quad (8)$$

Where X_{LSTM} represents the output features of the LSTM, $X_{LAttention}$ represents the output features of the attention module.

2) GCNs layer: GCNs are a neural network model for graph-structured data that can learn representations of nodes, edges, and entire graphs. Therefore, we use GCNs in the encoding layers of GCNLA.

GCNs are capable of a range of message passing and integration. The whole can be viewed as a function $Z = f(X, A)$ with respect to the node features and the adjacency matrix. GCNs use edges to transfer information between neighboring vertices after obtaining the vertex features and adjacency matrix of the graph, which in turn embeds the vertex features into a more efficient representation Z. It is worth noting that GCNs can be stacked similarly to convolutional neural networks. In our model, the message-passing rule for GCNs is as follows [42]:

$$X' = \sigma\left(\hat{D}^{-1/2}\hat{A}\hat{D}^{-1/2}X_{\text{LAttention}}W\right). \quad (9)$$

Where $X_{LAttention}$ represents the final output features of LSTM Attention Module, $\sigma(\cdot)$ represents an activation function. Notably, $\hat{A} = A + I_n$ indicates that a self-looping adjacency matrix has been inserted, $\hat{D}$ is the degree matrix of $\hat{A}$, and W is a trainable weight matrix.

3) Decoder: We use inner product decoders for the graph reconstruction task. The inner product decoder is defined as follows:

$$A' = \sigma(HH^T). \tag{10}$$

Where Z represents the embedding matrix of the node of the encoder, $\sigma(\cdot)$ represents an activation function. The inner product decoder calculates the cosine similarity score for each pair of embedding features. The cosine similarity score represents the likelihood that an edge exists between two candidate vertices. Next, a sigmoid function is used to convert the cosine similarity matrix into a probability representing the likelihood of the existence of an edge. The adjacency matrix is reconstructed by capturing the relationships between nodes in the graph.

C. Training Strategy

In GCNLA, we employ binary cross entropy (BCE) to construct the loss function. BCE computes the loss between the predicted and target values, which is a measure of the difference between the predicted probability distribution and the true distribution. The BCE loss formula is defined as follows:

$$\begin{aligned} L(A, A') = -\frac{1}{n^2}[&\sum_{i=1}^{n}\sum_{j=1}^{n}\left[A_{(i,j)}\log(A'_{(i,j)})\right] \\ &+\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(1-A_{(i,j)})\log(1-A'_{(i,j)})\right]] \end{aligned} \tag{11}$$

Where $A_{(i,j)}$ represents the ground truth label of the neighboring edges between cell i and j, $A'_{(i,j)}$ represents the predicted probability score of GCNLA for the edge between cell i and j.

D. Method's Validation

GCNLA infers the interaction probability scores of all cell pairs and reconstructs CCIs network from pairs with high probability scores. The cell-adjacency graph is reconstructed with strong reliability depending on the optimal threshold determined during training.

We used the initially defined cell adjacency graph as a positive sample of the CCIs network, and randomly selected an equal number of non-adjacent cell pairs as a negative sample. We chose genes labeled as ligand or receptor as feature genes for used in GCNLA. The positive and negative samples were randomly divided into 7:3 as training and test sets.

In the primal cell-adjacency graph, edges of different proportions were randomly selected and discarded. The processed edges were used to train the GCNLA to evaluate the effectiveness of edge in the reconstructed CCIs network. Meanwhile, randomly generated different numbers of fake edges added to the primal adjacency graph were fed into the GCNLA. The tolerance of GCNLA to noise was evaluated by using the real existing edges as positive samples and the mixed fake edges as negative samples.

Previous methods have been used to evaluate the enrichment or depletion of interactions between or within cell type [44], [48]. Given the above strategy, we first used reconstructed CCIs networks to generate distance distributions and determine thresholds to categorize distal interactions. Subsequently, we randomly shuffled the reconstructed network 2,000 times to produce a null distribution of distal interaction profiles based on the number of different types of CCIs in 2,000 networks. We calculated P-values by comparing the observed number of interactions with the null distribution to indicate the frequency with which random values are higher or lower than the observed values, i.e., specific cell types are enriched or depleted.

We performed cell clustering using K-Means [49] and Leiden [50], [51] algorithms. The optimal clustering number was identified with the Silhouette Score and the Calinski-Harabasz Score. Furthermore, Uniform Manifold Approximation and Projection (UMAP) and t-Distributed Stochastic Neighbor Embedding (t-SNE) were employed for visualization and dimension reduction.

In this study, we employed accuracy (ACC), precision, recall, specificity, Matthews correlation coefficient (MCC), and F1-score as evaluation metrics to evaluate the performance of GCNLA in inferring CCIs [53], which are defined as follows:

$$\text{ACC} = \frac{TP+TN}{TP+FP+FN+TN}. \tag{12}$$

$$\text{Precision} = \frac{TP}{TP+FP}. \tag{13}$$

$$\text{Recall} = \frac{TP}{TP+FN}. \tag{14}$$

$$\text{Specificity} = \frac{TN}{FP+TN}. \tag{15}$$

$$\begin{aligned} \text{MCC} = &\frac{TP\times TN - FP\times FN}{\sqrt{(TP+FP)(TP+FN)}} \\ &\times \frac{1}{\sqrt{(TN+FP)(TN+FN)}}. \end{aligned} \tag{16}$$

$$\text{F1-score} = \frac{2\times \text{Recall}\times \text{Precision}}{\text{Recall}+\text{Precision}}. \tag{17}$$

Where TP stands for true positive, TN for true negative, FP for false positive, and FN for false negative. The existence of a cell-cell edge is positive, whereas the absence of one is negative.

In addition, we used Average Precision (AP) and Area Under the Receiver Operating Characteristic Curve (AUROC) to further evaluate the model performance. The AP measures the performance of the model under all possible thresholds, the precision-recall curve is a weighted average of the precision achieved under each threshold, and the AP is the area under the precision-recall curve. AUROC evaluates the overall performance of the classification model, which is assessed by calculating the area under the ROC curve. The horizontal axis of the ROC curve is 1-specificity, whereas the vertical axis is recall.

III. Result

GCNLA combines GCNs, LSTM networks and attention mechanisms. Briefly, GCNLA learns from single-cell spatial transcriptomics to generate potential feature representation, which in turn capture the inherent relationship between gene expression and CCIs.

A. Datasets

The datasets used in our study were two published datasets. One was the mouse visual cortex using the seqFISH technology [54], and the other was from the mouse hypothalamic preoptic area using the MERFISH technology [55]. Both datasets are ST data at the resolution of a single cell. The specifics of each dataset are shown in **Table** I.

The seqFISH dataset used in our study was derived from Zhu et al.'s analysis of scRNA-seq data from the mouse visual cortex, providing high-precision quantification of 125 genes in 1,597 cells [40], [54], [56]. The MERFISH dataset of the mouse hypothalamic preoptic area was selected from the section at Bregma+0.11 mm of animal No.18, which contained the highest number of single cells. After removing the fuzzy cells, the dataset contained 160 genes in 4975 cells [44], [55].

B. Experimental setup

Graph data typically consists of numerous nodes and edges with a complex structure. GPUs have significant advantages in handling parallel computations, significantly accelerating the computational process and improving efficiency. Therefore, our proposed GCNLA was implemented on the PyTorch framework on the NVIDIA GeForce RTX 4050 Laptop GPU. The network employs the Adam optimizer with a learning rate of 0.001 and an epoch set to 1500.

C. Performance of the GCNLA on the test set

Although there has been some work focused on studying CCIs such as DeepLinc [44] and CLARIFY [11], which we have briefly described in the INTRODUCTION section. Both works also used graph structures. DeepLinc was a variational graph autoencoder for CCIs reconstruction [44]. CLARIFY employed a graph autoencoder to enable the reconstruction and refinement of CCIs [11]. Therefore, our work focuses on comparisons with DeepLinc and CLARIFY.

We defined the neighborhood graph of cells by based on adjacent cells at spatial locations. Specifically, each cell with its five adjacent nearest neighbor cells served as neighboring cell pairs with potential direct interactions, defined as direct contacts. In general, it is biologically reasonable to assume that most cells can be in direct communication with multiple other cells in real biological tissues [44]. This follows the same assumptions as in DeepLinc and CLARIFY. Meanwhile, the Cell Graph for our work was constructed following the same process as DeepLinc and CLARIFY. We also used spatial coordinates to calculate the Euclidean distance between cells to find the K nearest cells based on the Euclidean matrix. Therefore, we constructed adjacency matrices using the K nearest neighbour algorithm, which constitutes the adjacency matrix. We utilized neighboring cells with direct contact as positive samples to discover transcriptome features associated with CCIs.

We used the edges of cells with interactions as positive samples. and negative samples were randomly sampled in a 1:1 ratio to the positive samples to enable learning of transcriptomic features associated with CCIs. In the process of training the model, we selected 70% of the edges to train GCNLA, while the remaining 30% were used as a test set to assess GCNLA. To evaluate the performance of the reconstruction, we used AP and AUROC to assess the ability to reconstruct the edges of the test set during the training period. The results on seqFISH and MERFISH datasets are shown in **Table** II, while the performance was compared with that of DeepLinc and CLARIFY.

We evaluated the robustness of GCNLA with different edge partitions by varying the number of test edges. We evaluated our model using AP and AUROC on seqFISH and MERFISH datasets. The test edges were divided into 10%, 30%, 50%, 70%, and 90%, and each group repeated the experiment five times to generate box plots. The results are shown in **Fig.** 2. GCNLA outperforms DeepLinc and CLARIFY in all division test set ratios on the seqFISH and MERFISH datasets. Our model GCNLA is more robust than DeepLinc and CLARIFY, according to the results in **Fig.** 2. Our model is more stable with different test set division ratios and maintains stable performance even when trained on less data.

D. Performance of GCNLA in cell-cell interactions and noise experiments

We introduced false-positive edges and false-negative edges to the input training graph, in order to evaluate the tolerance of GCNLA to noise. False positive edges are false edges that we add at a rate of 0.1 to 0.9 times the input training graph with known CCIs edges. Meanwhile, false negative edges are real edges that we remove from the input training graph at a rate of 0.1 to 0.9 times the known CCIs edges. We trained GCNLA using these noisy inputs and evaluated the overall performance on the edge test set in each case. The results are shown in **Fig.** 3.

E. Ablation experiments

We performed ablation experiments on the seqFISH and the MERFISH datasets to verify the efficacy of each module in GCNLA.

The LSTM network was selected to better handle the complex high-dimensional data by capturing and learning the dependencies between cells. To be able to make the LSTM network more focused on task-relevant features, we introduced an attention module in LSTM. To evaluate the contribution of the LSTM and the attention module, we compared the GCNLA network with a degraded version without the LSTM and the attention module. In addition, spatial location information is crucial in CCIs studies, and we used GCNs to construct graph structures related to cellular spatial neighborhoods, and extracted inter-cell local and global structural features by iteratively aggregating neighbor node information. To evaluate

TABLE I
STATISTICS OF THE DATASETS USED FOR GCNLA.

Method	Cell Number	Gene Number	Tissue	Reference
seqFISH	1597	125	Mouse visual cortex	Zhu, et al., Nature Biotechnology, 2018
MERFISH	4975	160	Mouse hypothalamic preoptic region	Moffitt, et al., Science, 2018

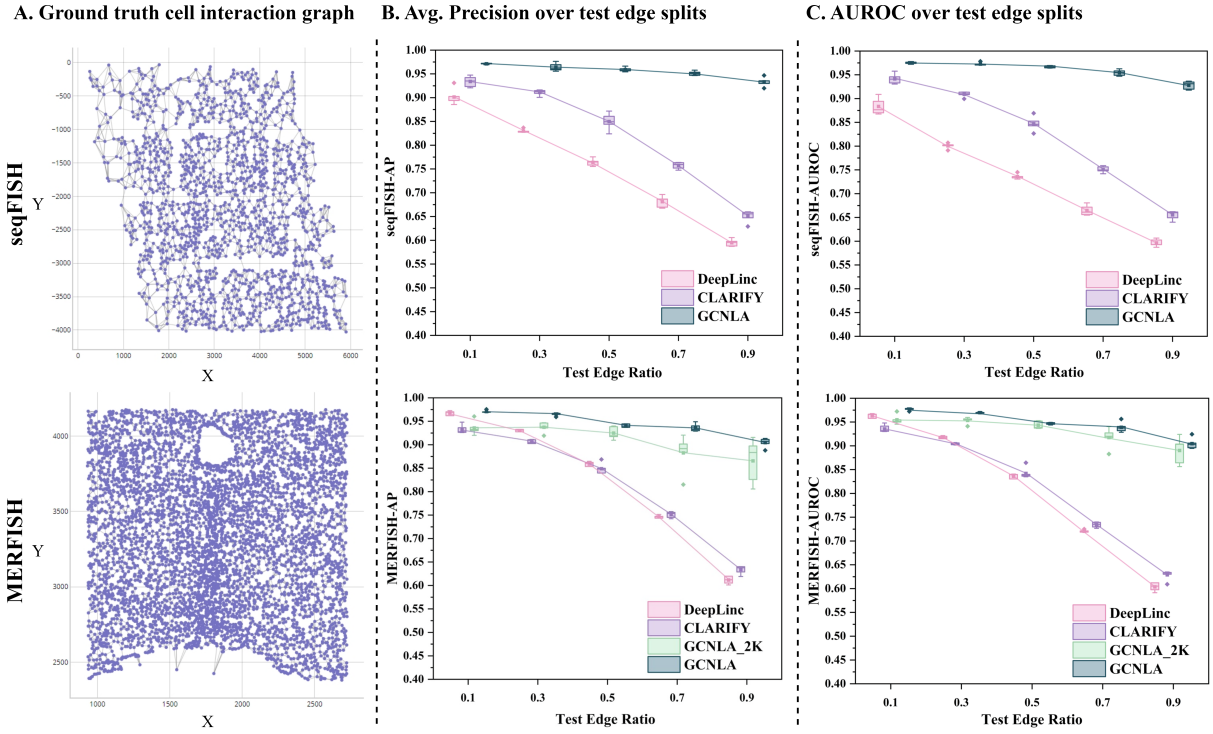


Fig. 2. Experimental performance of GCNLA in reconstructing cell-cell interactions. **A** The spatial transcriptomics dataset with realistic cellular interaction edges. **B** Comparison of the AP performance of GCNLA with DeepLinc and CLARIFY on different training/test splits. **C** Comparison of the AUROC performance of GCNLA with DeepLinc and CLARIFY on different training/test splits. The horizontal axis represents the test edge percentage.

TABLE II
AP, AUROC, AND AUPRC ARE COMPUTED IN THE TEST SET. THE BEST PERFORMANCES IN EACH DATASET ARE SHOWN IN BOLD.

DataSet	Model	Test Set	
		AP	AUROC
seqFISH	DeepLinc [44]	0.8293	0.8028
	Clarify [11]	0.9130	0.9106
	GCNLA	**0.9557**	**0.9724**
MERFISH	DeepLinc [44]	0.8981	0.8763
	Clarify [11]	0.9063	0.9064
	GCNLA_2k	0.9256	0.9405
	GCNLA	**0.9681**	**0.9720**

the contribution of the GCNs network, we compared the GCNLA network with a degraded version without GCNs.

The contribution and importance of each sub-module in GCNLA were evaluated with the training and test sets divided in 7:3 ratio. **Table** III presents the findings from the ablation experiments.

F. Reconstructing cell-cell interaction graphs

We compared the reconstructed CCIs network with the original CCIs network graph. The CCIs graph reconstructed by GCNLA consists of a large difference from the original cell-neighboring network consisting of proximal cells in direct contact. GCNLA provides a more comprehensive reconstruction of the cellular interaction graph through learning a cellular neighborhood graph simply defined by the spatial location information of cells, estimating missing interactions and filtering potential misconnections.

We visualized reconstructed CCIs network graphs for localized regions of different tissues, as shown in **Fig.** 4**A**. It can be noticed from the figure that GCNLA removed a small number of initially defined direct CCIs and introduced new CCIs relationships. Meanwhile, we revealed the capability of GCNLA to reconstruct distal interactions. As shown in **Fig.** 4**B** and **Supplementary file: Fig. S1**, the distal interactions between neurons in mouse visual cortex seqFISH are demonstrated. The enrichment or depletion of certain cell type interactions was assessed by comparison with randomly generated networks. The reconstructed distal CCIs network has different cell type patterns. In the mouse visual cortex seqFISH and

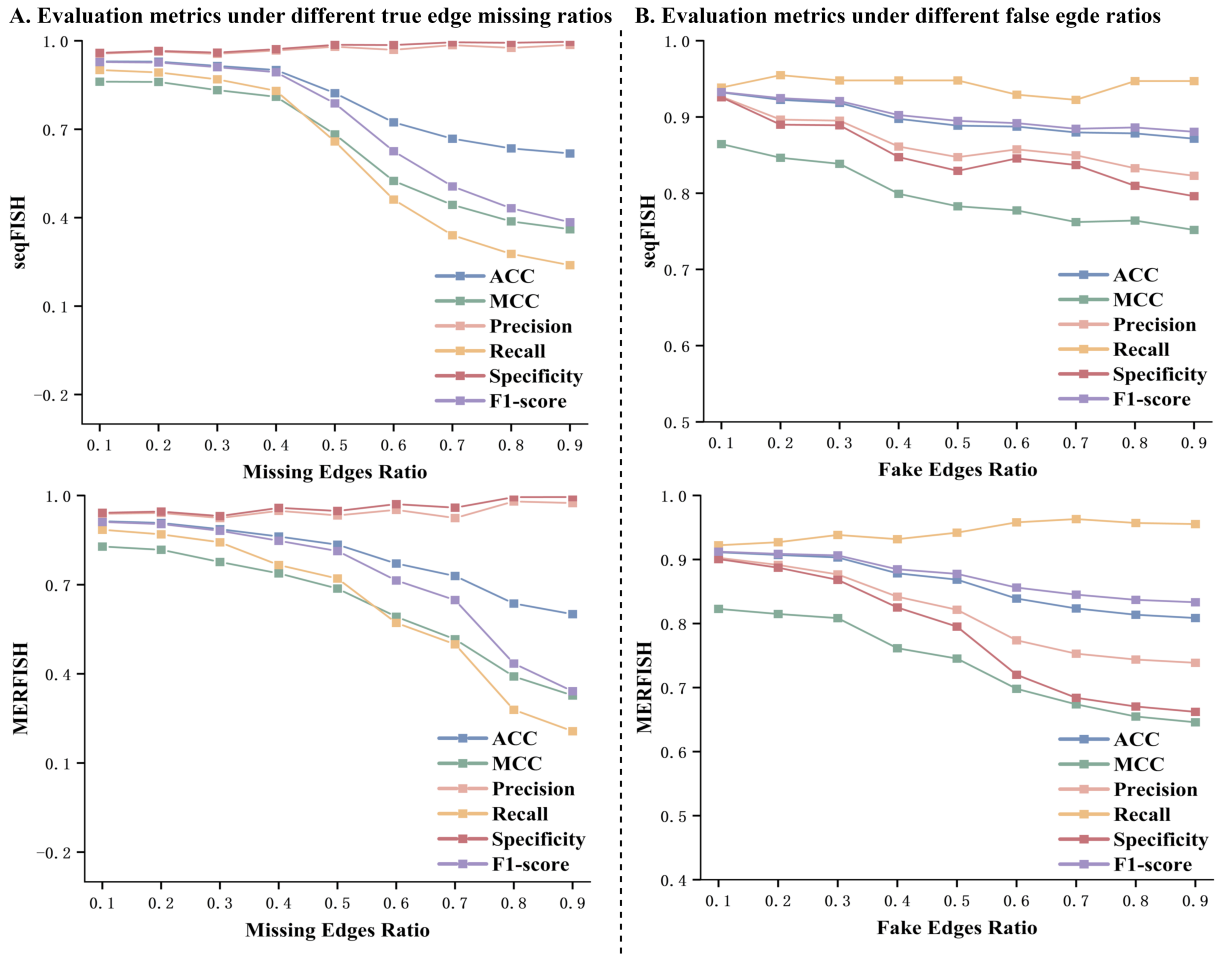


Fig. 3. Experimental performance of GCNLA in the presence of perturbed edge training sets. **A** Evaluation metrics under different proportions of missing true edges in the training data. **B** Evaluation metrics under different proportions of false edges added to the training data.

TABLE III
EVALUATION RESULTS OF NETWORK STRUCTURE ABLATION EXPERIMENTS ON TWO DATASETS. THE BEST PERFORMANCES IN EACH DATASET ARE SHOWN IN BOLD.

Dataset	Module			Test Set							
	LSTM	Attention	GCNs	AP	AUROC	ACC	MCC	Precision	Recall	Specificity	F1-score
seqFISH	✓			0.8583	0.8874	0.8524	0.7100	0.8143	0.9130	0.7918	0.8608
	✓	✓		0.9165	0.9271	0.8784	0.7600	0.8468	0.9241	0.8328	0.8837
			✓	0.9498	0.9508	0.7965	0.6386	**0.9715**	0.6109	**0.9821**	0.7501
	✓	✓	✓	**0.9557**	**0.9724**	**0.9373**	**0.8747**	0.9449	**0.9292**	0.9454	**0.9368**
MERFISH	✓			0.8606	0.8624	0.8244	0.6515	0.7971	0.8704	0.7784	0.8321
	✓	✓		0.8791	0.9040	0.8538	0.7094	0.8298	0.8901	0.8174	0.8589
			✓	0.9232	0.9153	0.7688	0.5878	**0.9515**	0.5664	**0.9711**	0.7101
	✓	✓	✓	**0.9681**	**0.9720**	**0.9281**	**0.8571**	0.9480	**0.9056**	0.9506	**0.9264**

mouse hypothalamic preoptic area MERFISH datasets, distal interactions between neurons were significantly enhanced (**Supplementary file: Fig. S3**). In contrast, endothelial cells showed a reduction in distal connections. This supports the view that endothelial cells form the endothelial layer in blood vessels, which in effect influences endothelial interactions between cells to some extent. In the mouse hypothalamic preoptic area MERFISH, distal interactions of pericytes reconstituted by GCNLA were depleted (**Supplementary file: Fig. 3**).

G. GCNLA potential cell embedding representation space refinement

In this section, we evaluated the ability of GCNLA to the embed input normalized count cell features into potential feature representations. The cell potential feature representations obtained from GCNLA were able to better capture the spatial distribution of cells, which confirming that the cell potential feature representations from GCNLA were more spatially relevant.

First, we visualized the pairwise Euclidean distances be-

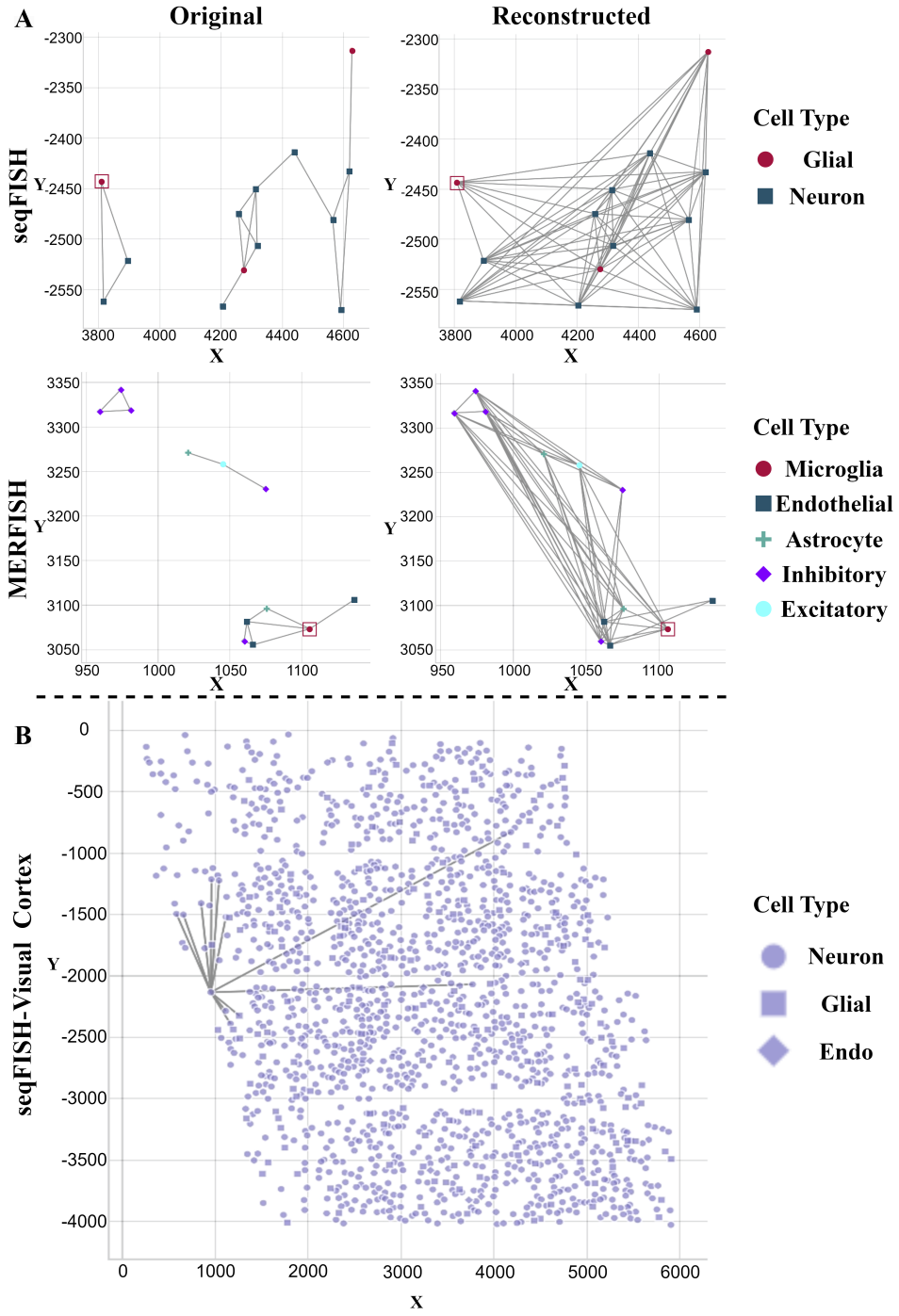


Fig. 4. GCNLA reconstructed cell-cell interaction graphs. **A** Representative tissue sections in the seqFISH and MERFISH datasets of the original CCIs network and the reconstructed CCIs network. The red box indicates the interaction of a microglia with other cells. **B** Example of neuronal distal interactions in the seqFISH dataset.

tween cells. The heatmap of Euclidean distances is shown in 5**A**, which represents an $n_c \times n_c$ matrix with the values in row i and column j representing the Euclidean distance between cell i and cell j. This Euclidean distance was calculated based on the coordinates of the ST data. As shown in **Figs.** 5**B** and 5**C**, we generated cell-cell similarity matrices utilizing latent features from GCNLA and initial normalized count cell features. In the two situations mentioned above, the matrix heatmaps are displayeds in **Figs.** 5**B** and 5**C**. The values in row i and column j in **Fig.** 5**B** represent the Euclidean distance between the initial feature vectors of cell i and cell j, while in **Fig.** 5**C** the values represent the Euclidean distances between the potential feature representations of cell i and cell j obtained from GCNLA. In contrast, the distribution of initial features is closer to a uniform distribution, while the latent feature representation obtained from GCNLA reveals an underlying structure that does not identical to the spatial distribution.

To quantify the results of the spatial refinement further, we

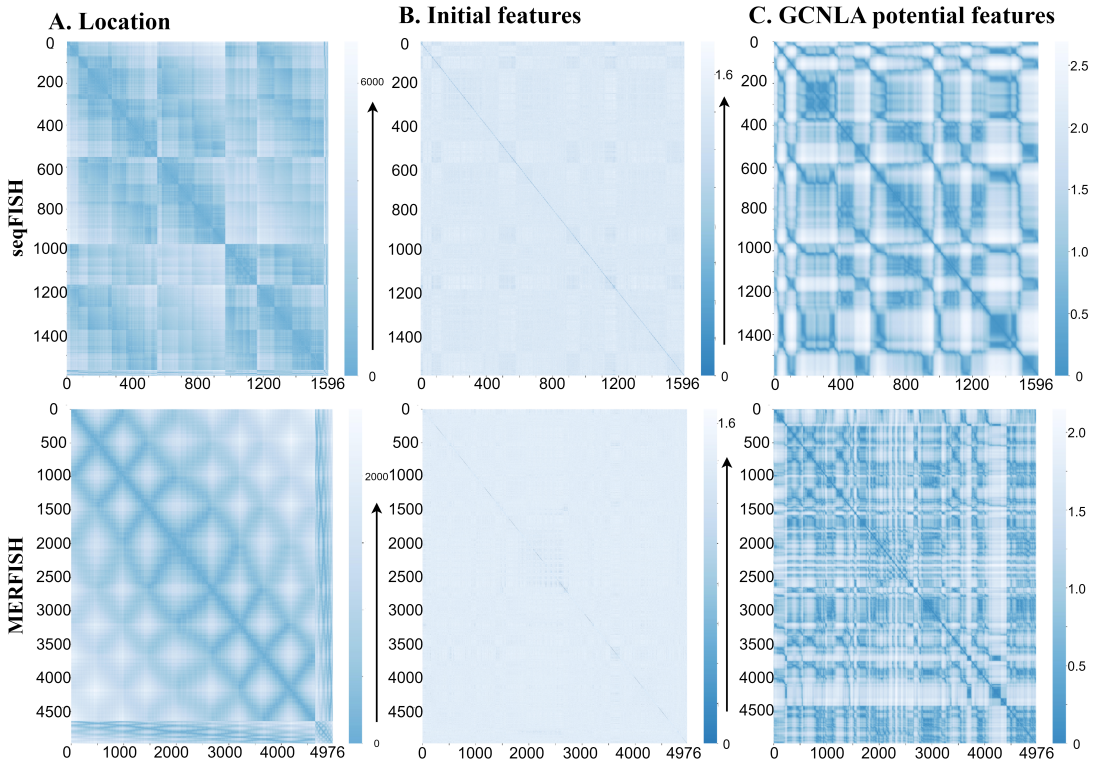


Fig. 5. Refinement of spatial information visualization. **A** Cell-cell heatmap in pairs based on location information. **B** Heatmap using initial cell features. **C** Heatmap using GCNLA potential features.

calculated the Spearman correlation between the cell location and the initial feature heatmap, as well as the Spearman correlation between the cell location and the heatmap of the cell potential features obtained from GCNLA. The results are presented in **Table** IV.

TABLE IV
THE SPEARMAN CORRELATION BETWEEN THE INITIAL FEATURE MATRIX, CELL EMBEDDING AND SPATIAL POSITION. THE BEST PERFORMANCES IN EACH DATASET ARE SHOWN IN BOLD.

Dataset	Region	Initial Cell features	GCNLA Cell embedding
seqFISH	Entire matrix	0.1053	**0.2388**
	Block diagonal	0.2522	**0.4474**
MERFISH	Entire matrix	0.0348	**0.1664**
	Block diagonal	0.1971	**0.2658**

H. GCNLA identifies spatial domains

For both seqFISH and MERFISH datasets, we used K-Means [49] and Leiden [50], [51] algorithms to cluster the potential feature representations obtained from GCNLA. Each cluster was defined as a spatial domain, mapped to each single cell. The results are visualized in **Fig.** 6**A**. In addition, UMAP and t-SNE were employed to minimize dimensionality and visualize potential features of single cells [52]. The dimension reduction and visualization results of UMAP and t-SNE in initial features and potential features of GCNLA are shown in **Figs.** 6**B**, 6**C**, 6**D**.

The findings demonstrate the ability of GCNLA to spatially refine single-cell features by demonstrating the spatial correlation of GCNLA potential feature representations.

IV. DISCUSSION

ST analysis techniques offer unprecedented opportunities to study highly heterogeneous cellular tissue. Recent developments in ST have made it possible to explore CCIs utilizing gene expression data. However, it has always been difficult to infer CCIs from ST data with comprehensiveness at single-cell resolution. Currently, ST data is subject to a number of technical limitations, including low gene coverage, data sparsity, etc. Thus, there is a need to develop data mining techniques that are extremely resilient to these technological constraints.

We introduce GCNLA, a network framework that integrates graph autoencoder with LSTM attention network. GCNLA is capable of identify CCIs from ST data at single-cell resolution. Meanwhile, further downstream analysis could be performed for features extracted by GCNLA.

In particular, we used the GCNLA model that encodes intrinsic features of ST data and decodes them according to the CCIs task. We used two different modes of data integrated with the cellular neighbor-adjacency matrix of the graph structure and high-throughput transcriptomic data to construct the cellular graph structure. First, processing gene expression matrices through LSTM attention networks can capture intercellular dependencies while reinforcing key gene expression features. Next, the enhanced gene expression information, together with

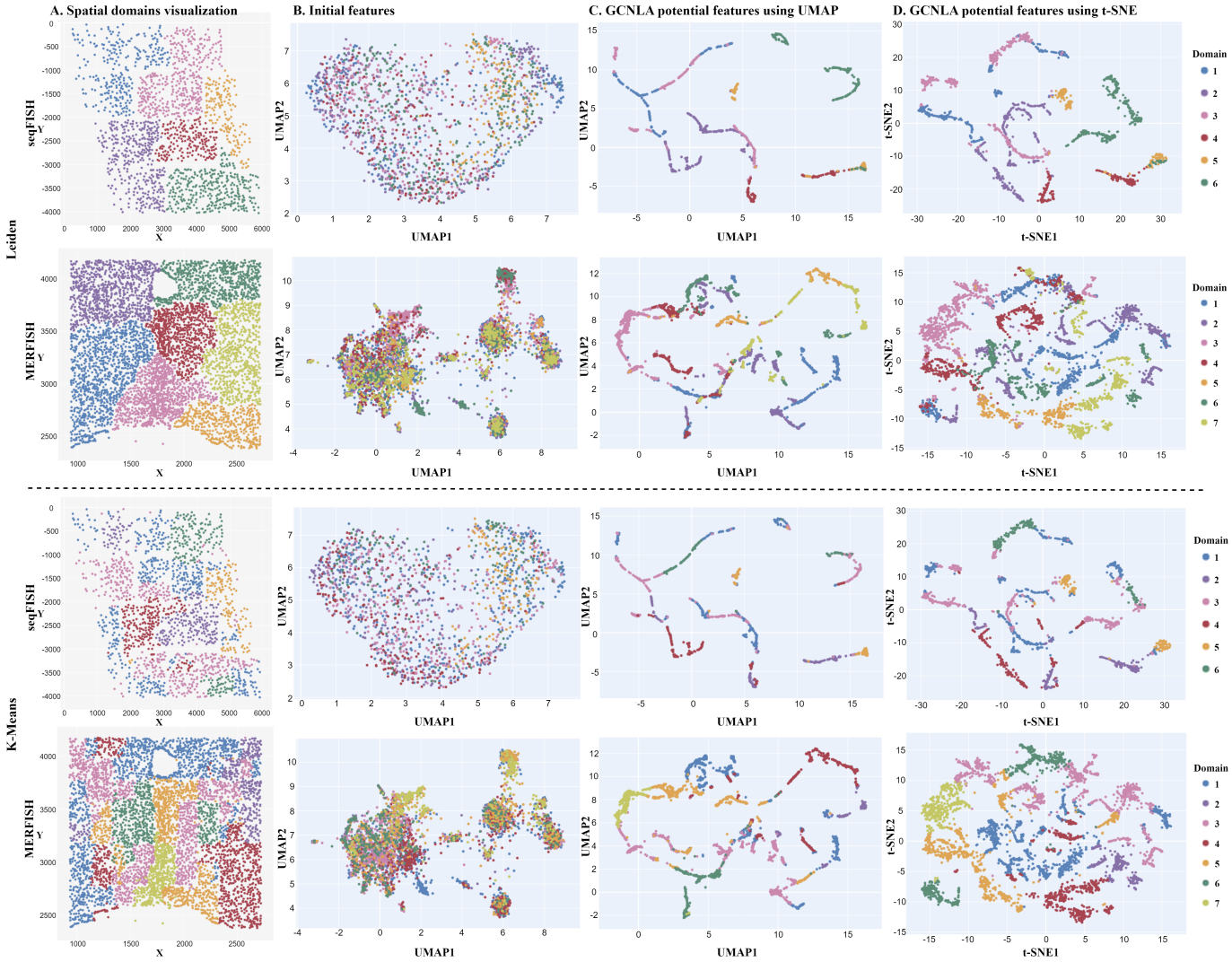


Fig. 6. Cell re-clustering based on GCNLA potential features defines the spatial domain. **A** Leiden and K-Means re-clustering using cells with GCNLA potential features. **B** The UMAP graph using cells with initial features. **C** The UMAP graph using cells with GCNLA potential features. **D** The t-SNE graph using cells with GCNLA potential features.

the neighbor matrix, is fed into the graph coding layer to capture the global cell interaction graph structural features and learn the spatial location information of the cells. Through extensive testing in real data, GCNLA has shown high efficiency and better performance in inferring CCIs. Thereby, the effectiveness and reliability of LSTM networks and GCNs in CCIs tasks were demonstrated.

We evaluated the reconstruction capability of CCIs using GCNLA by testing the model across different proportions of the test set. The experimental results with different proportions of edges in the test set (**Fig.** 2) show the stabilization of the model performance by GCNLA even with more edges in the test set. The anti-noise experiment findings (**Fig.** 3) demonstrate that our model GCNLA is more stable with the addition of false edges than with the loss of true edges. This further proves that GCNLA is more tolerant to noisy data in the presence of false edges.

In the local reconstructed CCIs network graphs from seqFISH in mouse visual cortex and MERFISH in mouse hypothalamus preoptic area, microglia acted as a hub cell demonstrating dense connections with other cells (**Fig.** 4**A**). This finding recapitulates the well-recognized interactions of microglia with other cell types in the Central Nervous System [44], [57]. The analysis of distal interactions in the MERFISH dataset of the mouse hypothalamus preoptic area indicated a depletion of endothelial cell connections (**Supplementary file: Fig. S3**). This supports the view that endothelial cells form the endothelial layer in blood vessels, which in effect influences endothelial interactions between cells to a certain degree. In addition distal interactions of pericytes are depleted, considering that pericytes embedded in the capillaries basement membranes interact directly with endothelial cells [58]. This affirms with the findings of Li et al. [44] and laterally further confirms the potential of GCNLA in reconstructing CCIs profiles.

In the Euclidean heatmap of potential feature representations versus location distributions obtained through GCNLA (**Fig.** 5), we observe that the initial feature distribution is

essentially uniformly distributed (**Fig.** 5**B**). However, the potential feature representation obtained through the GCNLA model reveals an underlying structure (**Fig.** 5**C**), but is does not identical to the spatial distribution. The reason is that the location information is not the only information of the potential feature representation, which integrates gene expression, CCIs information, etc. To further analyze the potential features obtained from the coding part of our model, we clustered the potential feature representations of the cells using the K-Means and Leiden algorithm, which define each cluster as a spatial domain, with the results mapped back to each single cell mapping. The clustering results (**Fig.** 5**A**) provide additional visual confirmation that the potential features of the GCNLA are clearly spatially clustered into distinct domains. The dimension reduction and visualization results of UMAP and t-SNE demonstrate that the potential feature representations learned by GCNLA have a clearer feature space structure than the initial features (**Figs.** 6**B**, 6**C**, 6**D**). This further demonstrates that the potential features encoded by GCNLA have stronger intrinsic structure, with promising applications to other downstream tasks. All of the experimental results above effectively demonstrate that GCNLA effectively integrates spatial information with single-cell gene expression, showing robust capability in reconstructing CCIs.

Unfortunately, GCNLA is unable to infer the direction as well as the strength of CCIs due to the ST data lacks more specific information. Similar to most models based on GCNs, GCNLA struggle to effectively interpret the encoded significant features. The continuous development of ST technology and breakthroughs in the interpretability of GCNs will contribute to further improvements of GCNLA in the future. If more information can be provided in the future, GCNLA can be further developed by incorporating features including cell morphology and metabolic profile.

V. Conclusion

In this study, we present GCNLA, a graph autoencoder and LSTM attention network approach, which is a deep data mining tool designed to infer CCIs from ST data. GCNLA combines GCNs and LSTM networks to infer CCIs. Our developed GCNLA model does not rely on prior knowledge of cell or gene functions. The experimental results in seqFISH and MERFISH datasets show that GCNLA achieves high efficiency in inferring CCIs, while also providing high robustness and high noise immunity. In addition, the representation of potential features of CCIs learned by GCNLA enables cell clustering, identification of spatial domains, which in turn informs the understanding of tissue structure.

References

[1] T. J. Bechtel, T. Reyes-Robles, O. O. Fadeyi, and R. C. Oslund, "Strategies for monitoring cell-cell interactions," *Nat Chem Biol*, vol. 17, no. 6, pp. 641-652, Jun 2021, doi: 10.1038/s41589-021-00790-x.

[2] E. Armingol, H. M. Baghdassarian, and N. E. Lewis, "The diversification of methods for studying cell-cell interactions and communication," *Nat Rev Genet*, vol. 25, no. 6, pp. 381-400, Jun 2024, doi: 10.1038/s41576-023-00685-8.

[3] X. Wang, A. A. Almet, and Q. Nie, "The promising application of cell-cell interaction analysis in cancer from single-cell and spatial transcriptomics," *Semin Cancer Biol*, vol. 95, pp. 42-51, Oct 2023, doi: 10.1016/j.semcancer.2023.07.001.

[4] E. Armingol, A. Officer, O. Harismendy, and N. E. Lewis, "Deciphering cell-cell interactions and communication from gene expression," *Nat Rev Genet*, vol. 22, no. 2, pp. 71-88, Feb 2021, doi: 10.1038/s41576-020-00292-x.

[5] C. Bonnans, J. Chou, and Z. Werb, "Remodelling the extracellular matrix in development and disease," *Nat Rev Mol Cell Biol*, vol. 15, no. 12, pp. 786-801, Dec 2014, doi: 10.1038/nrm3904.

[6] B. Bakhshandeh *et al.*, "Mechanotransduction in tissue engineering: Insights into the interaction of stem cells with biomechanical cues," *Exp Cell Res*, vol. 431, no. 2, pp. 113766, Oct 15 2023, doi: 10.1016/j.yexcr.2023.113766.

[7] D. Pascut, M. Y. Pratama, N. V. T. Vo, R. Masadah, and C. Tiribelli, "The Crosstalk between Tumor Cells and the Microenvironment in Hepatocellular Carcinoma: The Role of Exosomal microRNAs and their Clinical Implications," *Cancers (Basel)*, vol. 12, no. 4, pp. 823, Mar 29 2020, doi: 10.3390/cancers12040823.

[8] J. X. Zhou, R. Taramelli, E. Pedrini, T. Knijnenburg, and S. Huang, "Extracting Intercellular Signaling Network of Cancer Tissues using Ligand-Receptor Expression Patterns from Whole-tumor and Single-cell Transcriptomes," *Sci Rep*, vol. 7, no. 1, pp. 8815, Aug 18 2017, doi: 10.1038/s41598-017-09307-w.

[9] R. S. Sealfon, A. K. Wong, and O. G. Troyanskaya, "Machine learning methods to model multicellular complexity and tissue specificity," *Nature Reviews Materials*, vol. 6, no. 8, pp. 717-729, 2021, doi: 10.1038/s41578-021-00339-3.

[10] L. Bich, T. Pradeu, and J. F. Moreau, "Understanding Multicellularity: The Functional Organization of the Intercellular Space," *Front Physiol*, vol. 10, pp. 1170, 2019, doi: 10.3389/fphys.2019.01170.

[11] M. Bafna, H. Li, and X. Zhang, "CLARIFY: cell-cell interaction and gene regulatory network refinement from spatially resolved transcriptomics," *Bioinformatics*, vol. 39, no. 39 Suppl 1, pp. i484-i493, Jun 30 2023, doi: 10.1093/bioinformatics/btad269.

[12] Z. Zhang, F. Cui, C. Cao, Q. Wang, and Q. Zou, "Single-cell RNA analysis reveals the potential risk of organ-specific cell types vulnerable to SARS-CoV-2 infections," *Computers in biology and medicine*, vol. 140, pp. 105092, 2022, doi: 10.1016/j.compbiomed.2021.105092.

[13] X. Zhou *et al.*, "Circuit Design Features of a Stable Two-Cell System," *Cell*, vol. 172, no. 4, pp. 744-757 e17, Feb 8 2018, doi: 10.1016/j.cell.2018.01.015.

[14] R. Browaeys, W. Saelens, and Y. Saeys, "NicheNet: modeling intercellular communication by linking ligands to target genes," *Nat Methods*, vol. 17, no. 2, pp. 159-162, Feb 2020, doi: 10.1038/s41592-019-0667-5.

[15] J. Cheng, J. Zhang, Z. Wu, and X. Sun, "Inferring microenvironmental regulation of gene expression from single-cell RNA sequencing data using scMLnet with an application to COVID-19," *Brief Bioinform*, vol. 22, no. 2, pp. 988-1005, Mar 22 2021, doi: 10.1093/bib/bbaa327.

[16] S. Wang, M. Karikomi, A. L. MacLean, and Q. Nie, "Cell lineage and communication network inference via optimization for single-cell transcriptomics," *Nucleic Acids Res*, vol. 47, no. 11, pp. e66, Jun 20 2019, doi: 10.1093/nar/gkz204.

[17] K. Sun et al., "scRNA-seq of gastric tumor shows complex intercellular interaction with an alternative T cell exhaustion trajectory," *Nat Commun*, vol. 13, no. 1, pp. 4943, Aug 23 2022, doi: 10.1038/s41467-022-32627-z.

[18] S. Jin, M. V. Plikus, and Q. Nie, "CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics," *Nat Protoc*, vol. 20, no. 1, pp. 180-219, Jan 2025, doi: 10.1038/s41596-024-01045-4.

[19] Q. Ma, Q. Li, X. Zheng, and J. Pan, "CellCommuNet: an atlas of cell-cell communication networks from single-cell RNA sequencing of human and mouse tissues in normal and disease states," *Nucleic Acids Res*, vol. 52, no. D1, pp. D597-D606, Jan 5 2024, doi: 10.1093/nar/gkad906.

[20] R. K. Gupta and J. Kuznicki, "Biological and Medical Importance of Cellular Heterogeneity Deciphered by Single-Cell RNA Sequencing," *Cells*, vol. 9, no. 8, Jul 22 2020, doi: 10.3390/cells9081751.

[21] T. Stuart and R. Satija, "Integrative single-cell analysis," *Nat Rev Genet*, vol. 20, no. 5, pp. 257-272, May 2019, doi: 10.1038/s41576-019-0093-7.

[22] F. Tang *et al.*,, "mRNA-Seq whole-transcriptome analysis of a single cell," *Nat Methods*, vol. 6, no. 5, pp. 377-82, May 2009, doi: 10.1038/nmeth.1315.

[23] V. Svensson, R. Vento-Tormo, and S. A. Teichmann, "Exponential scaling of single-cell RNA-seq in the past decade," *Nat Protoc*, vol. 13, no. 4, pp. 599-604, Apr 2018, doi: 10.1038/nprot.2017.149.

[24] X. Shen, Y. Zhao, Z. Wang, and Q. Shi, "Recent advances in high-throughput single-cell transcriptomics and spatial transcriptomics," *Lab Chip*, vol. 22, no. 24, pp. 4774-4791, Dec 6 2022, doi: 10.1039/d2lc00633b.

[25] P. L. Stahl *et al.*,, "Visualization and analysis of gene expression in tissue sections by spatial transcriptomics," *Science*, vol. 353, no. 6294, pp. 78-82, Jul 1 2016, doi: 10.1126/science.aaf2403.

[26] S. G. Rodriques *et al.*,, "Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution," *Science*, vol. 363, no. 6434, pp. 1463-1467, Mar 29 2019, doi: 10.1126/science.aaw1219.

[27] Z. Zhang *et al.*,, "webSCST: an interactive web application for single-cell RNA-sequencing data and spatial transcriptomic data integration," *Bioinformatics*, vol. 38, no. 13, pp. 3488-3489, Jun 27 2022, doi: 10.1093/bioinformatics/btac350.

[28] Y. Sun *et al.*,, "A comprehensive survey of dimensionality reduction and clustering methods for single-cell and spatial transcriptomics data," *Brief Funct Genomics*, pp. elae023, Jun 11 2024, doi: 10.1093/bfgp/elae023.

[29] C. Yan *et al.*,, "Integration tools for scRNA-seq data and spatial transcriptomics sequencing data," *Briefings in Functional Genomics*, pp. elae002, 2024, doi: 10.1093/bfgp/elae002.

[30] H. Duan, Q. Zhang, F. Cui, Q. Zou, and Z. Zhang, "MVST: Identifying spatial domains of spatial transcriptomes from multiple views using multi-view graph convolutional networks," *PLOS Computational Biology*, vol. 20, no. 9, pp. e1012409, 2024, doi: 10.1371/journal.pcbi.1012409.

[31] Y. Su, S. Yan, C. Cao, F. Cui, Q. Zou, and Z. Zhang, "HumanSpa: An Online Platform Based on Spatial Transcriptome Data for Diseases of Human Systems," *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2023, pp. 293-298, doi: 10.1109/BIBM58861.2023.10385302.

[32] Q. Chen *et al.*,, "Integrating single-cell and spatial transcriptomics to elucidate the crosstalk between cancer-associated fibroblasts and cancer cells in hepatocellular carcinoma with spleen-deficiency syndrome," *J Tradit Complement Med*, vol. 14, no. 3, pp. 321-334, May 2024, doi: 10.1016/j.jtcme.2023.11.008.

[33] C. Zhang, L. Wang, and Q. Shi, "Computational modeling for deciphering tissue microenvironment heterogeneity from spatially resolved transcriptomics," *Comput Struct Biotechnol J*, vol. 23, pp. 2109-2115, Dec 2024, doi: 10.1016/j.csbj.2024.05.028.

[34] R. Moncada *et al.*,, "Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas," *Nat Biotechnol*, vol. 38, no. 3, pp. 333-342, Mar 2020, doi: 10.1038/s41587-019-0392-8.

[35] Z. Cang and Q. Nie, "Inferring spatial and signaling relationships between cells from single cell transcriptomic data," *Nat Commun*, vol. 11, no. 1, pp. 2084, Apr 29 2020, doi: 10.1038/s41467-020-15968-5.

[36] R. Dries *et al.*,, "Giotto: a toolbox for integrative analysis and visualization of spatial expression data," *Genome Biol*, vol. 22, no. 1, pp. 78, Mar 8 2021, doi: 10.1186/s13059-021-02286-2.

[37] J. H. Lee *et al.*,, "Highly multiplexed subcellular RNA sequencing in situ," *Science*, vol. 343, no. 6177, pp. 1360-3, Mar 21 2014, doi: 10.1126/science.1250212.

[38] X. Wang *et al.*,, "Three-dimensional intact-tissue sequencing of single-cell transcriptional states," *Science*, vol. 361, no. 6400, Jul 27 2018, doi: 10.1126/science.aat5691.

[39] K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, and X. Zhuang, "Spatially resolved, highly multiplexed RNA profiling in single cells," *Science*, vol. 348, no. 6233, pp. aaa6090, 2015, doi: 10.1126/science.aaa6090.

[40] S. Shah, E. Lubeck, W. Zhou, and L. Cai, "In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus," *Neuron*, vol. 92, no. 2, pp. 342-357, Oct 19 2016, doi: 10.1016/j.neuron.2016.10.001.

[41] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Advances in neural information processing systems*, vol. 29, 2016.

[42] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:.1609.02907*, 2016, doi: 10.48550/arXiv.1609.02907.

[43] Y. Yuan and Z. Bar-Joseph, "GCNG: graph convolutional networks for inferring gene interaction from spatial transcriptomics data," *Genome Biol*, vol. 21, no. 1, pp. 300, Dec 10 2020, doi: 10.1186/s13059-020-02214-w.

[44] R. Li and X. Yang, "De novo reconstruction of cell interaction landscapes from single-cell spatial transcriptome data with DeepLinc," *Genome Biol*, vol. 23, no. 1, pp. 124, Jun 3 2022, doi: 10.1186/s13059-022-02692-0.

[45] W. Yang et al., "at single-cell resolution for spatial transcriptomics with subgraph-based graph attention network," *Nat Commun*, vol. 15, no. 1, pp. 7101, Aug 18 2024, doi: 10.1038/s41467-024-51329-2.

[46] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition," *arXiv preprint arXiv:1402.1128*, 2014, doi: 10.48550/arXiv.1402.1128.

[47] X. Shao, J. Liao, C. Li, X. Lu, J. Cheng, and X. Fan, "CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice," *Brief Bioinform*, vol. 22, no. 4, Jul 20 2021, doi: 10.1093/bib/bbaa269.

[48] D. Schapiro et al., "histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data," *Nat Methods*, vol. 14, no. 9, pp. 873-876, Sep 2017, doi: 10.1038/nmeth.4391.

[49] S. Vassilvitskii and D. Arthur, "k-means++: The advantages of careful seeding," *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 2006, pp. 1027-1035.

[50] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, pp. P10008, 2008, doi: 10.1088/1742-5468/2008/10/P10008.

[51] V. A. Traag, L. Waltman, and N. J. van Eck, "From Louvain to Leiden: guaranteeing well-connected communities," *Sci Rep*, vol. 9, no. 1, pp. 5233, Mar 26 2019, doi: 10.1038/s41598-019-41695-z.

[52] L. McInnes, J. Healy, and J. Melville, "Umap: Uniform manifold approximation and projection for dimension reduction," *arXiv preprint arXiv:.03426*, 2018, doi: 10.48550/arXiv.1802.03426.

[53] Y. Wang, Y. Zhai, Y. Ding, and Q. Zou, "SBSM-Pro: Support Bio-sequence Machine for Proteins," *arXiv e-prints*, pp. arXiv:2308.10275, 2023, doi: 10.48550/arXiv.2308.10275.

[54] Q. Zhu, S. Shah, R. Dries, L. Cai, and G. C. Yuan, "Identification of spatially associated subpopulations by combining scRNAseq and sequential fluorescence in situ hybridization data," *Nat Biotechnol*, vol. 36, no. 12, pp. 1183-1190, Oct 29 2018, doi: 10.1038/nbt.4260.

[55] J. R. Moffitt *et al.*,, "Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region," *Science*, vol. 362, no. 6416, Nov 16 2018, doi: 10.1126/science.aau5324.

[56] B. Tasic *et al.*,, "Adult mouse cortical cell taxonomy revealed by single cell transcriptomics," *Nat Neurosci*, vol. 19, no. 2, pp. 335-46, Feb 2016, doi: 10.1038/nn.4216.

[57] M. Prinz, S. Jung, and J. Priller, "Microglia biology: one century of evolving concepts," *Cell*, vol. 179, no. 2, pp. 292-311, 2019, doi: 10.1016/j.cell.2019.08.053.

[58] M. D. Sweeney, S. Ayyadurai, and B. V. Zlokovic, "Pericytes of the neurovascular unit: key functions and signaling pathways," *Nat Neurosci*, vol. 19, no. 6, pp. 771-83, May 26 2016, doi: 10.1038/nn.4288.

Chao Yang is currently pursuing a Professional Doctorate degree in electronic information at Hainan University, Haikou, China. Her research interests include bioinformatics, machine learning and artificial intelligence.

Xiuhao Fu is currently working toward the Master degree in electronic information, Hainan University, Haikou, China. His research interests include bioinformatics, machine learning and AI.

Zhenjie Luo is currently working toward the academic master degree in computer science and technology, Hainan University, Haikou, China. His research interests include diffusion model and drug molecule generation.

Leyi Wei is currently a Professor at School of Information,Xiamen University, Xiamen, China. He received Ph.D. degrees in Computer Science and Technology from Xiamen University in 2016. He worked as a project researcher in the lab of Functional analysis in silico, Institute of Medical Science, University of Tokyo during 2018 - 2019. His research interests include Bioinformatics, Machine Learning, Deep Learning, and Computational Biology. He has published 90+ peer-reviewed papers in top-tier journals such as Bioinformatics, Briefings in Bioinformatics, and BMC Genomics, etc. He has got around 8000 citations in Google Scholar.

Jingbing Li (Member, IEEE) was born in 1966. He received the B.S. degree in electronic information engineering from the Wuhan University of Technology, Wuhan, China, in 1989, the M.S. degree in automation from the Beijing Institute of Technology, in 1996, and the Ph.D. degree in control theory and control engineering from Chongqing University, in 2007. In 2006, he joined Zurich University, Switzerland, as a Visiting Scholar. In 2011, he went to the Intelligent Laboratory of Ritsumeikan, Japan, as a Visiting Scholar. From 2013 to 2014, he was also a Visiting Scholar with Troy University, USA. He is currently a Full Professor with Hainan University. His research interests include artificial intelligence, network security, image processing, and medical image watermarking. He received two Second Class Prizes of the Science and Technology Progress Award of Hainan Province, in 2007 and 2012, and the Teaching Masters Award in Hainan Province, in 2010.

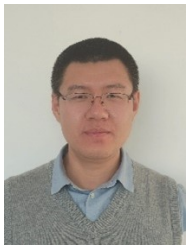
Feifei Cui is currently an associate professor in the School of Computer Science and Technology, Hainan University. She received the M.S. degree in computer application technology from Shandong University, Jinan, China, in 2012 and the Ph.D. degree in bioinformatics from the University of Tokyo, Japan in 2020. She worked as a postdoctoral researcher in the University of Electronic Science and Technology of China. Her research interests include bioinformatics, deep learning and biological data mining.

Dr. Cui is a winner of the Japanese Government (Monbukagakusho: MEXT) Scholarship Program from 2016 to 2019, and the Postdoctoral International Exchange Program from 2020 until 2022.

Quan Zou (M'13-SM'17) is currently a Professor in the Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China. He received the BSc, MSc and the PhD degrees in computer science from Harbin Institute of Technology, China, in 2004, 2007 and 2009, respectively. He worked in Xiamen University and Tianjin University from 2009 to 2018 as an Assistant Professor, Associate Professor and Professor. His research is in the areas of bioinformatics, machine learning and parallel computing.

Several related works have been published by Science, Briefings in Bioinformatics, Bioinformatics, IEEE/ACM Transactions on Computational Biology and Bioinformatics, etc. Google scholar showed that his more than 100 papers have been cited more than 5000 times.

Dr. Zou is the editor-in-chief of Current Bioinformatics, associate editor of IEEE Access, and the editor board member of Computers in Biology and Medicine, Genes, Scientific Reports, etc. He was selected as one of the Clarivate Analytics Highly Cited Researchers in 2018 and 2022. He is also a reviewer for many impacted journals and NSFC (National Natural Science Foundation of China).



Qingchen Zhang is currently a professor at the School of Computer Science and Technology, Hainan University. He obtained his Ph.D. in Software Engineering from Dalian University of Technology, China, in 2015. He has conducted postdoctoral research and served as an assistant professor at St. Francis Xavier University in Canada. His main research areas include machine learning, medical big data, and blockchain.

Zilong Zhang is currently an associate professor in the School of Computer Science and Technology, Hainan University. He received the Ph.D. degree in bioinformatics from the University of Tokyo, Japan in 2020. He worked as a postdoctoral researcher in the University of Electronic Science and Technology of China. His research interests include bioinformatics, machine learning and graph neural network.