# Investigation of pair-wise single-cell interactions by statistically interpreting spatial cell state correlation learned by self-supervised graph inductive bias transformer

Xiao Xiao[1]

[1]Department of Biostatistics, Yale School of Public Health, New Haven, CT

**Abstract**

Image-based spatial transcriptomics (ST) offers spatial gene expression profile at the single-cell resolution and provides information to understand intercellular communication that is critical for maintaining tissue development and organ function. Disruption of normal cell-cell interactions (CCI) can lead to disease onset and progression. Current CCI analysis methods face several limitations, including subjection to the number of measured ligand-receptor genes in image-based spatial transcriptomics, limited graph encoding power, inadequate use of spatial information, and low interpretability. Here, we present GITIII, an interpretable self-supervised graph transformer-based language model that treats cells as words (nodes) and their cell neighborhood as a sentence to explore the communications among cells. Enhanced by multilayer perceptron-based distance scaler, physics-informed attention mechanism, and a state-of-the-art, expressive, and lightweight graph transformer model, GITIII infers CCI by investigating how the state of a cell is influenced by the spatial organization, ligand expression, cell types and states of neighboring cells. With its interpretable architecture, GITIII can be used to understand how the sender cell influences target genes in the receiver cell, visualize the spatial pattern and utility of CCI, identify significant CCI networks, perform CCI-informed cell subtyping, and compare CCI strength between disease groups. Applications to four ST datasets from several species, organs, and platforms, GITIII effectively identified and quantitatively interpreted key CCI patterns driving within-sample heterogeneity and disease progression, thus improving our understanding of brain structures, tumor microenvironments, and the interplay among different cell types responding to neighboring CCIs.

## Introduction

In multicellular organisms, cell-cell interactions (CCI) are pivotal in maintaining biological function and organization. These interactions, which can involve direct physical contacts mediated by a diverse array of cell adhesion molecules among adjacent cells or more distant communications through paracrine and autocrine signaling by secreted factors, trigger intracellular cascades that profoundly influence gene expression and cell behavior. This intricate process is shaped by several factors, including the spatial proximity between the sender and receiver cells, the microenvironment surrounding the receiver cell, the strength of the ligand-receptor (LR) signaling pathway, and the expression of other genes that modulate downstream CCI effects. Together, these elements play a decisive role in development, metabolism, and disease progression.

Considering the importance of single-cell resolution measurements for investigating CCI, recent studies have primarily utilized single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics data. To analyze CCI, previous methods typically begin with constructing a ligand-receptor (LR) pair signaling database[1-3]. Subsequently, the communication strength for each signaling pathway is calculated using methods such as the production of LR expression[3], co-expression correlation[4], thresholding[5], differential expression[6], and tensor decomposition[7]. These approaches facilitate the construction of CCI networks between cell types, analysis of downstream effects of CCIs[8], comparison between different contexts[3, 7], and clustering of signal pathways[2], which can be used to better understand the key factors in the development of embryos and the progression of diseases[9]. However, these methods often aggregate expression data within the same cell type, overlooking intra-cell type heterogeneity. This means two cells of the same type might interact under one niche but not another. Additionally, these methods typically identify only the most extreme communication strengths in permutation tests as statistically

significant CCIs, disregarding less pronounced interactions that do exist and make impacts. Most importantly, CCI functions through LR pairs influenced by intercellular distance, yet spatial information is absent in scRNA-seq data.

Spatial transcriptomics (ST) offers invaluable insights into CCI by providing gene expression levels with spatial coordinates. Currently, various spatial transcriptomics techniques are available to measure spatial gene expression. These techniques fall into two categories: barcode-based methods (e.g., 10x visium and Slide-seq[10]) and imaging-based methods (e.g. seqFISH[11], MERFISH[12], 10x xenium[13], and CosMX[14]). Barcode-based methods enable whole transcriptome analysis but provide a lower resolution that encompasses multiple cells in one spot. In contrast, imaging-based methods provide single-cell resolution, making it better suited for investigating intercellular interactions, albeit limited to measuring only a few hundred genes.

Currently, various methods have been developed to integrate spatial information for more detailed analysis of intercellular interactions in barcode-based methods[15]. To utilize spatial information, previous approaches either limit the analysis to proximal cells and spots, as seen in SpaTalk[16], COMMOT[17], CellPhoneDB V3[18], and SCVA[19], or weigh communication scores by diffusion formulae, such as SpatialDM[20], BATCOM[21], GraphComm[22], and SpaCeNet[23]. However, the influence of distance on CCI is neither binary nor uniform across different signaling pathways; some pathways require direct cell-cell contact, while others function through ligands with varying diffusion coefficients. Most critically, in imaging-based spatial transcriptomics, which offers single-cell resolution for precise CCI analysis, the reliance on the LR database is limited due to the often low or missing inclusion of LR pairs in measured gene sets, restricting the application of these database-dependent methods.

Graph deep learning has become increasingly prevalent in transcriptomic data analysis, covering tasks such as spatial domain identification[24, 25], identifying joint biomarkers[26], correlating multi-modal data[27], and inferring gene interaction[28]. Given the intricate nature of CCI and the existing knowledge gap, the renowned capability of deep learning to model complex patterns for graph-structured data could be significantly beneficial. Several studies have employed rudimentary forms of Graph Neural Networks (GNNs) to explore cell-cell interaction. For example, NCEM[29] and GCNG[28] have utilized graph convolutional neural networks (GCNs)[30] to encode CCI, treating cells as nodes and spatial relationships as edges. stKeep[31], DeepTalk[32] and SPACE[33] explored CCI-informed subclusters using embeddings from graph attention network[34] (GAT)-based model, whereas CGCom[35] interpreted attention scores as communication strength indicators. And HoloNet[36] integrates ligand-receptor (LR) pair information, multi-view GCN[37], dense spatial graphs, and the distance scaler to encode and decode CCI. However, besides limited usage in single-cell-resolution spatial transcriptomics with few or no inclusion of LR pairs, these methods have several limitations, including high reconstruction (prediction) loss, potential misinterpretations of attention scores, limited graph encoding capabilities, inadequate integration of distance information, and the restrictive receptive field of GNN models. Furthermore, previous deep-learning models may not effectively learn CCI during training, often defaulting to decoding ligand-receptor expression from the input LR product or relying on the principle that proximity implies similarity, as commonly observed in ST. Moreover, previous studies have not simultaneously addressed the specific direction, strength, or target gene of CCI, nor have they accurately measured the precise influence of one sender cell on a receiver cell for each cell on a slide. To enhance the understanding and encoding of CCI, the implementation of a more powerful graph encoding model, a model architecture that facilitates

easier interpretation, and as much LR knowledge as possible in the given single-cell-resolution spatial transcriptomics dataset are essential for a comprehensive depiction of CCI's nature.

In this study, we introduce GITIII (Graph Inductive Transformer for Intercellular Interaction Investigation), a self-supervised deep learning model designed to analyze intercellular interactions. GITIII addresses the complexity of CCI by treating the transcriptomics profile and cell types of each cell as node features, with ligand expression levels adjusted by learnable distance embeddings as edge features. The encoding of these node and edge features is achieved using a modified version of the GRaph Inductive bias Transformer (GRIT)[38], which encodes input tensors in a language model manner. The unique architecture of our model allows the encoder's output to be interpreted as the influence of neighboring cells on the central cell for each targeted gene. These influences are then aggregated and compared with the observed gene expression to update the model parameters.

With the ability to decode the functions of CCI between single cells, a novel and unprecedented comprehensive understanding of the interactions between cells is available using our method. We dissect the three fundamental aspects of CCI - directionality, strength, and functional role - providing a detailed analytical and visual framework. This enables the identification of distinct communication patterns within and across patients at various disease stages. Then, by employing partial linear regression-based statistical approaches, our methods identify specific target genes significantly influenced by CCIs between certain cell types, and how these interactions vary across different patient stages. While boasting a wide range of functions, our research also establishes that our graph neural network model significantly outperforms existing GNN methodologies used in earlier studies in terms of self-supervised learning loss. We apply our method to 4 single-cell-resolution spatial transcriptomics datasets of different techniques

and unravel key CCI patterns in the mouse brain, Alzheimer's disease, lung cancer, and breast cancer.

## Results

### Overview of GITIII

GITIII is a deep-learning model for CCI inference with high interpretability to characterize intercellular communication landscapes at single-cell resolution. GITIII is described in Methods, with its technical details provided in Supplementary Note. Briefly, GITIII takes the gene expression, cell type annotation, and spatial coordinates of each cell as inputs (Fig. 1a). The model decomposes the gene expression levels of each cell into two parts: “cell-type expression,” representing the average gene expression across cells of the same type, and “cell-state expression,” representing the deviation of gene expression from the average expression across cells of the same type. For each cell, GITIII identifies its 49 nearest neighbors and leverages their information to predict the cell-state expression of the center cell. The model consists of two components: embedding and encoding. In the embedding component, the node features are constructed by cell-type embedding and gene expression profile processed by multi-layer perceptron (MLP), whereas the edge features are derived from ligand expressions scaled by a distance scaler, which is calculated based on the distance between the sender and receiver cells through learnable neural networks (Fig. 1b). In the encoder component, the model employs a modified one-layer Graph Inductive bias Transformer (GRIT) block to better adapt to the complexity of CCI and enhance interpretability. Here, each cell is treated as a word in a transformer-based language model, with a neighborhood of 50 cells representing a sentence (Fig. 1c). The encoder outputs the influence tensor that quantifies the impact of each neighboring cell on the cell-state expression of the center cells. By summing the influence from all neighboring cells, we obtain the reconstructed (predicted)

cell-state expression of the center cell. Mean squared error is used as the loss function to train the model parameters.

The influence tensor, characterizing how receiver cells' target genes are influenced by its neighboring sender cells, can be used to visualize spatial CCI patterns, investigate within-sample CCI heterogeneity, identify significant CCIs, and compare CCI strength between different samples, thus improving our understanding of CCI patterns across spatial regions. We can also construct CCI-informed interpretable features to assess the impact of neighboring cell types on the cell-state expression of a particular gene. Based on these features, we further cluster cells into subtypes with distinct CCI patterns and examine how CCIs affect the cell state of an individual cell (Fig. 1d). Furthermore, we use the Wald test to examine the partial dependence of one cell's state on the influence of other cell types and identify significant CCI networks. Additionally, we calculate the neighborhood permutation-based strength of CCIs targeting various genes and compare them across different samples using a cell neighborhood-based permutation approach (Fig. 1e).

**GITIII reveals CCI network and within-cell-type CCI heterogeneity across mouse brain regions**

The core features of GITIII include assessing the influence of sender cells on their corresponding receiver cells, visualizing the spatial distribution and effects of these communications, detecting significant CCI networks, and identifying cell subtypes. We first illustrated these features using a MERFISH dataset from the mouse primary motor cortex (PMC) [39]. This ST dataset contains 254 genes measured in 284,098 cells in 64 slices from two mice, with annotations among 24 cell types. We used slice 201 from mouse 1 with 6,137 cells for demonstration (Fig. 2a). Among the 254 genes, 47 were found to be ligand genes and no LR pairs were found in the ligand-receptor databases NeuronChatDB[3] and CellChatDB[2].

We trained GITIII to estimate the influence tensor, where the strength of CCI naturally decreases with the distance between two communicating cells (Fig. 2b). For each cell on the slide, we considered it as a receiver cell and selected the top 5 sender cells with the strongest interactions to form CCI pairs. Among these selected pairs, we further identified the top 10 cell-type interactions by ranking them based on sender and receiver cell types. Fig. 2c shows the Uniform Manifold Approximation and Projection (UMAP) of these CCI pairs[40], revealing clear separations among the 10 cell-type interactions. This demonstrates GITIII's ability to capture distinct CCI patterns among interactions between different cell types, consistent with the patterns observed in our UMAP analysis.

We used CCI-informed interpretable features to identify subtypes of astrocytes and L2/3 IT. Astrocytes are reported to be functionally diverse in different brain regions[41]. For example, the original MERFISH study [39] clustered them into three subtypes, with one subtype distributed across all layers and two other subtypes enriched mostly in the white matter (Fig. S1b). GITIII identified five distinct subtypes in this slice (Fig. 2d, upper panel), each spatially distributed in different cortical layers (Fig. 2e). Briefly, astrocyte subtype 1 is located in layer 1, subtype 2 is located in layers 2 and 3, subtype 3 is located in layers 4 and 5, subtype 4 is located in layer 6, and subtype 5 is located in the white matter (Fig. S1a). For L2/3 IT, the MERFISH study [39] clustered them into five subtypes without clear spatial boundaries (Fig. S1d). In contrast, GITIII detected three distinct subtypes with clear separations (Fig. 2d, 2e), where subtype 1 is located in the upper part of layers 2 and 3, closer to layer 1, while subtypes 2 and 3 are located in the deeper part of layers 2 and 3, closer to layer 4 (Fig. S1c).

Next, we investigated differentially expressed genes among the subtypes of astrocytes and L2/3 IT and assessed the influence of their adjacent cell types on these cells. Fig. 2g shows that

astrocyte subtypes 1 and 2 were strongly influenced by L2/3 IT neurons, subtype 3 was mostly influenced by L4/5 IT and L5 IT neurons, subtype 4 was strongly influenced by L6 CT and L6 IT neurons, and subtype 5 was predominantly influenced by oligodendrocytes and L6b neurons. The inferred interactions between the different types of excitatory neurons and astrocyte subtypes, as well as that between oligodendrocytes and astrocyte subtype 5, confirmed that intercellular communications were most likely to happen between physically close cells (Fig. S1a, S1c). Astrocyte subtypes 1 and 2 had increased expression of *Cux2* and *Lama3*, while subtype 5 showed increased expression of *Igfbp5* and *Gfap* (Fig. 2f). For L2/3 IT neurons, subtype 1 was mostly influenced by astrocytes, Lamp5, Pvalb, and Vip neurons, while subtypes 2 and 3 were strongly influenced by L4/5 IT neurons as they were physically close to layer 4 than subtype 1, and subtype 3 was also influenced by Pvalb. Under these influences, L2/3 IT subtype 1 had upregulated expression levels of *Camk2d*, *Trp53i11*, *Trpc6,* and *Otof*, subtype 2 showed increased expression of *Rorb* and *Unc5d*, while subtype 3 showed increased expression of *Sulf2* and *Adra1b* (Fig. 2f).

We further identified significant CCI networks among cell types and visualized inferred CCI patterns and their spatial distributions. We selected three genes *Rorb*, *Gfap*, and *Cux2* for visualization due to their high predictive accuracy and biological significance. Figure 2h, shows the observed (left) and predicted (right) expression by GITIII of these genes across all cells, where the Pearson correlation coefficients between the observed and predicted expression exceed 0.4 and the adjusted p-values less than $1 \times 10^{-24}$.

We then visualized the spatial distribution of the top 5% of the strongest CCI pairs influencing the expression of these three genes (Fig. 2i) and identified the CCI network using partial linear regression. Figure 2j illustrates that L4/5 IT significantly increased the expression of *Rorb* in L2/3 IT, where we observed more CCI pairs between L4/5 IT and L2/3 IT subtype 2 (Fig

2i, Fig. S1f), agreeing with the upregulated expression of *Rorb* in L2/3 IT subtype 2 (Fig. 2f). The interactions between oligodendrocytes and astrocytes resulted in the increased expression of *Gfap* (Fig. 2i, 2j, Fig. S1g), which confirmed the influence we inferred of oligodendrocytes on astrocyte subtype 5 in the white matter (Fig. S1c) and the upregulated expression of *Gfap* in astrocyte subtype 5 (Fig. 2f). *Gfap* is a marker of astrocytes during development and the astrocyte-oligodendrocyte interaction is important for myelin regeneration[42]. *Cux2*, a transcription factor controlling neuronal proliferation and differentiation[43], showed increased expression in astrocytes from the L2/3 IT region (Fig. 2i, 2j). This aligns with the upregulated expression of *Cux2* in astrocyte subtype 2, also located in layers 2 and 3 (Fig. S1a, S1h), and supports previous findings on the influence of neurons on astrocyte morphology and proliferation[44].

**GITIII identifies various intercellular interaction patterns within and across patients with Alzheimer's Disease**

Beyond basic CCI network inference, our model can be extended to the analysis of multiple samples, allowing investigation into the similarities and differences of CCI networks both within and across samples. To illustrate these functionalities, we applied GITIII to the SEA-AD dataset, which contains data from 366,272 cells across 69 images of the human middle temporal gyrus from 27 donors, obtained using MERFISH. The dataset includes the measurements of 140 genes, and we utilized the 24 "subclass" cell type annotation from the original study[45]. For demonstration, we used the sample 'H20.33.001.CX28.MTG.02.007.1.02.03' (Fig. 3a) to illustrate within-sample heterogeneity. Additionally, while the dataset measures 28 ligand genes, only 2 ligand-receptor (LR) pairs from neuronchat and cellchat were identified among these genes.

After training the model, we visualized the functionality of the 10 most frequent CCI pairs (Fig. 3b) and calculated how interaction strength decays with the distance between two cells,

aligning with physical laws. To illustrate within-sample heterogeneity, we selected microglia and L2/3 IT due to their abnormal metabolism in AD as reported in previous studies[46-48]. The UMAP visualization of CCI-informed interpretable features reveals that microglia can be subdivided into three subtypes, each localized to a distinct cortical layer (Fig. 3e, left panel, and Fig. S2a). Specifically, in subtype 3 of microglia, we observed significant up-regulation of the *CUX2* gene, and down-regulation of *NTNG2*, *MOG*, and *FEZF2* (Figure 3f), indicating the influence of L2/3 IT (Fig. 3g, 3h) on microglial functions including cell proliferation, synapse formation[49], microglial accumulation along axons[50, 51], myelin growth and integrity[52], and the determination of microglial states[53] in AD condition.

Regarding L2/3 IT neurons, as depicted in Figure 3d, right panel, we identified two main subtypes with distinct spatial patterns with more clear spatial separation than the expression-based annotation in the original paper (Fig. S2b). Subtype 2 is predominantly influenced by Layer 4 Intratelencephalic neurons (L4 IT) with lesser influence from L2/3 IT, while subtype 1 is relatively more influenced by microglia and vasoactive intestinal peptide (VIP)-expressing neurons. In subtype 1, there was a notable down-regulation of the *RORB* and *MOG* genes, both of which are linked to Alzheimer's disease and indicate neuronal vulnerability[46]. Given that this subtype is predominantly influenced by microglia, it is plausible to hypothesize that microglia may contribute to the downregulation of the *RORB* gene in L2/3 IT neurons in AD, which aligns with our inferred interactions from microglia to L2/3 IT targeting *RORB* (Fig. 3h, right panel).

After analyzing the Cell-Cell Interaction (CCI) patterns within a single sample, we assessed our model's prediction accuracy and calculated the CCI networks for all samples in the dataset. We aggregated these networks using Stouffer's Z-score method combined with Fisher's combined test, while also controlling for Type II error that is absent in previous studies. To demonstrate, we

visualized the prediction accuracy (Fig. 3i) and the CCI network for genes *RORB*, *MOG*, *CD74*, and *ROBO2* (Fig. 3j), which are identified as potential risk genes for AD.

As illustrated in Figure 3h, the PCC between measured and predicted expressions for these five genes are all above 0.14 (with adjusted p-values <1e-24), demonstrating our model's robust capability in decoding the impact of CCIs on cell state. For *RORB*, statistical analysis shows that Layer 4 Intratelencephalic neurons (L4 IT) predominantly enhance RORB expression in other cells, while L2/3 IT, the primary receiver, sees *RORB* levels negatively influenced by microglia and VIP neurons.

Regarding the myelin oligodendrocyte glycoprotein (*MOG*) gene, we noted significant up-regulation in cells interacting with oligodendrocytes, with down-regulation in cells interacting with L2/3 IT and immune-related glial cells. This pattern underscores the role of oligodendrocytes in myelin generation[54] and highlights potential disruptions caused by degenerating neurons[55] and immune-related glial cells[56, 57], which may serve as a new therapeutic target.

*CD74*, known for its role in immune responses[58], was found to be significantly upregulated in cells interacting with microglia, consistent with previous research[59]. This finding highlights the influence of inflammatory microglia on surrounding cells[60].

For *ROBO2*, despite the interactions between neuron cells, a significant change in CCI between microglia and oligodendrocytes was detected. Although the *TREM2* pathway is recognized for primarily mediating microglial responses to amyloid-beta, demyelination, and the maintenance of oligodendrocyte health[61-64], the specific underlying mechanisms remain unclear. However, recent research has identified TREM1, a protein within the TREM family, as a newly recognized ligand for the ROBO2 receptor[65]. This discovery suggests a potential indirect

link between the ROBO2 and TREM2 pathways, which may contribute to microglial-induced damage in oligodendrocytes.

After identifying the significant CCI network, we quantified the communication strengths and compared them between two Alzheimer's Disease (AD) progression groups in the dataset. Figure 3k presents a UMAP visualization of CCI strengths of 69 slides from patients with varying cognitive statuses, revealing distinct differences between the two progression groups. As depicted in Figure 3l, patients in the dementia stage of AD show enhanced communication between microglia, oligodendrocytes, and excitatory neurons in the deeper cortical layers, compared to those in the early stages of the disease. This indicates a higher level of oligodendrocyte and microglia dysfunction that leads to the impairment of neuron functions, which further influences cognitive status.

**GITIII quantifies and interprets how tumor microenvironments influence gene expression heterogeneity in specific cell types in NSCLC**

Quantifying and interpreting the tumor microenvironment (TME), as well as exploring its impact on cellular states, has long presented challenges. With the introduction of GITIII, we can now accurately measure the TME where a cell resides and deduce how this microenvironment influences the cell's state. To illustrate this, we applied GITIII on the cosmx NSCLC dataset, which comprises 980 genes analyzed across 769,114 cells from eight formalin-fixed paraffin-embedded (FFPE) slides derived from five non-small cell lung cancer (NSCLC) samples[14]. In our study, we utilized the "CellType" cell type annotation from the original study, which contains 22 cell types. For demonstration, we choose 'Lung6' (Fig. 4a) for CCI-informed subtyping and CCI network analysis.

After training and evaluating our model, we first examined how GITIII encodes CCI across different cell types (Fig. 4b) and distances (Fig. 4c). The results showed a clear distinction in CCI between different cell types and a strong tendency for closer distances to exhibit stronger CCI. Next, we constructed interpretable CCI-informed features and conducted subtyping analysis for three cell types.

Figure 4d-g, upper panel, illustrates that macrophages can be categorized into three subtypes. The cells in subtype 1 are predominantly influenced by other macrophages, subtype 2 by structural cells, and subtype 3 by tumor cells. Subtype 3 cells express higher levels of *CD68*, *PSAP*, *GPNMB*, and HLA-family genes, indicating enhanced phagocytosis[66], inflammation[67], antigen processing, and presentation[68] due to increased interactions with neighboring macrophages. In subtype 3, the influence from tumor cells significantly up-regulates *SPP1* expression in macrophages, which plays a crucial role in modulating cancer progression[14, 69, 70]. In endothelial cells, a distinct differentiation is observed between subtype 4 and others, where the influence from tumor cells significantly up-regulates the expression of *KRT6A* and *EGFR*, suggesting that tumor cells modulate endothelial's behavior via *EGFR* signaling, promoting proliferation and angiogenesis[71]. Furthermore, we categorized tumor cells into two groups: those predominantly influenced by other tumor cells and those at the tumor boundary influenced by diverse cell types. Differential expression analysis revealed that boundary tumor cells exhibit reduced expression of genes associated with cancer proliferation[72], immunosuppression[73], and metastasis[74], highlighting the suppressive impact of immune cells at the tumor boundary.

We then determined the significant CCI networks targeting *SPP1*, *EGFR*, *KRT6A*, and *GPNMB* (Fig. 4h). Our observations revealed that networks targeting *SPP1* predominantly involve macrophages, where macrophages increase the expression of *SPP1* in their neighboring cells and

tumor cells increase the expression of *SPP1* in macrophages. In contrast, the CCI network for *EGFR* is primarily composed of tumor cells as sender types, suggesting that tumor cells overexpress *EGFR* and its ligands[75, 76], thereby influencing neighboring cells and promoting tumor proliferation and metastasis[77]. Regarding *KRT6A*, our analysis indicates that tumor cells significantly up-regulate the expression of *KRT6A* in their neighboring cells, while fibroblasts enhance *KRT6A* expression in tumor cells. This interaction underscores the role of fibroblasts in promoting tumor migration[78] and the capacity of tumor cells to modify the microenvironment, facilitating cell migration and proliferation[79]. For *GPNMB*, we identified an up-regulation in structural cells induced by tumor cells. Given that overexpression of *GPNMB* in the mammary epithelium accelerates tumor onset and growth[80], this suggests a feedback network between tumor and structural cells that enhances tumor cell proliferation.

After identifying CCI networks, we sought to understand the broad-scale influences of one cell type on another. To this end, we conducted a Gene Ontology (GO) enrichment analysis using ClusterProfiler[81]. This analysis examined genes significantly influenced by interactions between tumor cells and endothelial cells, as well as between tumor cells and macrophages, with genes measured in this dataset being used as background genes. As illustrated in Figure 4 i-l, fibroblasts influence tumor cells in remodeling the extracellular matrix[82], while tumor cells influence endothelial cells by promoting cytoskeleton and fiber growth[83]. These interactions collectively enhance tumor proliferation and metastasis. Regarding macrophages, we noted that they impact tumor cells in cell cycle regulation[84], chromosome organization, and p53 signal pathway[85]. Conversely, tumor cells enhance the role of macrophages in apoptotic processes[86] while modifying their extracellular organelles[87]. These interactions highlight potential therapeutic targets and reveal the complex role of the microenvironment on each cell type[88].

**GITIII uncovers key drivers of within-cell type heterogeneity in breast cancer**

To further illustrate the capacity of our model to quantify and interpret the tumor microenvironment (TME) and analyze single-cell-resolution spatial transcriptomics data across diverse platforms, we applied GITIII to a breast cancer dataset generated by Xenium, encompassing 313 genes across three slides[13]. Our model was trained and evaluated using data from 'Sample1_rep1' and 'Sample1_rep2' slides, which included 19 distinct, supervised cell type annotations spanning 274,492 cells. Of the 313 genes measured, 33 were identified as ligand genes. The 'Sample1_rep1' slide was specifically used for demonstration purposes (Figure 5a).

We first examined how our model encodes CCI patterns across different cell types by generating a UMAP visualization of significant CCI pairs and evaluating the influence of neighboring cells at varying distances on the final prediction. As shown in Figure 5b, CCI pairs from different cell type combinations have distinct impacts on the state of the corresponding receiver cell. Figure 5c further demonstrates that our model accurately captures natural CCI patterns, with spatial proximity enhancing the influence of the sender cell on the receiver cell.

We then systematically examined how immune cells, tumor cells, and structural cells within the tumor microenvironment (TME) are influenced by their neighboring cells. Starting with B cells from the immune cell category, Figure 5d-g, leftmost panel, demonstrates that B cells can be divided into two subtypes with distinct cell state expression patterns. Specifically, the up-regulation of *MS4A1*, *BANK1*, and *PTPRC*, along with the down-regulation of *TENT5C*, *SLAMF7*, *ITM2C*, and *NZB1*, was observed in the subtype 1, which is significantly more influenced by CD4+ T cells. This finding is consistent with previous research[89] where down-regulation of certain B-cell surface markers, including immune checkpoints, is a common immune escape mechanism indicative of an immunosuppressive environment produced by regulatory T cells influenced by

tumor-derived factors. For macrophages, we identified subtypes exhibiting intense interactions with invasive tumor cells and ductal carcinoma in situ (DCIS), characterized by the up-regulation of *SCD*, *CD9*, and *FOXA1*. This suggests that macrophages, under the influence of tumor cells, exhibit diminished ability to induce programmed cell death in tumor cells and antigen presentation, as evidenced by the abnormal expression patterns of *SCD*[90] and FOXA1[91].

Next, we examined the cell state heterogeneity of tumor cells, specifically DCIS. As illustrated in Figure d-g, the third row, DCIS can be categorized into three subtypes, each corresponding to a distinct distance range from the tumor boundary. For subtype 1, located on the tumor boundary, we observed increased expression of *MYLK*, *ACTA2*, and *KRT14*. This observation aligns with previous findings and suggests a progressive role for DCIS cells at the tumor boundary[92].

Finally, we analyzed structural cells, including endothelial and stromal cells, and observed that subtypes interacting with invasive tumor cells exhibit distinct gene expression patterns compared to other subtypes. In endothelial, subtype 2 associated with invasive tumors exhibited up-regulation of *CCND1*, *KDR*, *TCM*, and *KRT7*, suggesting that tumor cells prompt endothelial cells to support angiogenesis[93] and tumor growth[94]. For stromal cells, subtype 3, which interacts more intensively with invasive tumors, showed increased expression of *KRT7*, *FOXA1*, *FASN*, and *KRT8*. This indicates that tumor-resident stromal cells contribute to breast cancer invasion by promoting the expression of basal epithelial markers such as *KRT14*[95].

**Benchmarking and sensitivity test**

In addition to the unprecedented comprehensive deciphering of CCI, our model also demonstrates its effectiveness in encoding the complex nature of CCI compared with GNN methods used in previous studies. To demonstrate it, our model is benchmarked against four widely

used and sophisticated deep learning model backbones in the domain of spatial transcriptomics data analysis on the four datasets we used in this study. This comparison specifically focuses on each model's ability to reconstruct and predict the state of a central cell based on the information from its neighboring cells. To be more specific, with the same random seed and training hyperparameters, we compared GCN used in GCNG[28], SpaGCN[24], GraphST[25]; multi-view GCN (mv-GCN) used in Spatial-MGCN[96], HoloNet[36]; GAT used in CGCom[35], spaCI[97], SPACE[33], DeepTalk[32] and stKeep[31]; And original graph transformer (GT) used in SiGra[98]. To demonstrate the superiority of the distance scaler fitted by a neural network, we also compared the original mv-GCN to mv-GCN with our distance scaler (mv-GCN1).

The evaluation metric is the mean square error (MSE) between the measured cell state expression and the predicted cell state expression on the validation set. Furthermore, to demonstrate statistical significance, we computed the Pearson Correlation Coefficient (PCC) between the ground truth and the predictions for each measured gene. We then conducted a non-parametric paired Wilcoxon test. The p-values obtained from this test were adjusted using the Benjamini-Hochberg correction method to compare our model with other models, ensuring a robust assessment of statistical differences.

As shown in Figure 6 a-d, our model achieves the lowest MSE in decoding CCI, significantly surpassing other GNN models, with adjusted p-values below 1e-9 in comparison with all models across all datasets. Additionally, we confirmed that incorporating our MLP-based distance scaler into the mv-GCN framework markedly reduces MSE across four datasets, underscoring the efficacy of our distance scaler.

We then conducted an ablation study and a sensitivity test to evaluate whether the introduction of each component in our model enhances its ability to encode CCI and assess its

response to various settings. As shown in Figure 6e, the introduction of modified attention mechanisms, MLP-based distance-scaler, and distance embedding would improve our model's ability to encode CCI. Furthermore, we examined our model's performance in the absence of ligand gene information and cell type annotations. We discovered that omitting this information does not diminish our model's encoding capabilities. This finding suggests that the architecture of our attention calculation and feature embedding is capable of autonomously learning this information from the input data to encode CCI. Consequently, our model remains robust even with incomplete information, indicating its applicability in scenarios where cell type annotations or ligand information may be incomplete.

In the field of spatial transcriptomics data analysis, most GNN-based models employ multiple GNN layers. However, our findings indicate that a single layer of our encoder outperforms other models. As shown in Figure 6f, adding extra encoder layers to our model only leads to marginal improvement but would cause great damage to the interpretability. As a result, we employ a single-layer graph transformer encoder to maintain interpretability.

Finally, we evaluated the sensitivity of our model to training hyperparameters, specifically the learning rate, and examined how the number of neighbors in the cell neighborhood used for prediction affects the final prediction accuracy of our self-supervised learning. As depicted in Figure 6g, a learning rate between 0.0001 and 0.0001 yields the optimal results. Furthermore, Figure 6h illustrates that prediction loss decreases as the number of neighbors increases. However, since the decrease in prediction loss becomes insignificant when the number of neighbors exceeds 50, we have opted to set the number of neighbors at 50 to reduce computational resource demands.

**Discussion**

We introduced GITIII, a state-of-the-art graph neural network that achieves unprecedented precision in encoding and decoding cell-cell interactions (CCI) within single-cell-resolution spatial transcriptomics. Equipped with an MLP-based distance scaler, distance-mediated attention calculation, and the integration of node and edge features, GITIII outperforms existing GNN models in capturing the complex nature of CCI. The model's interpretable architecture allows for the precise determination of how each CCI pair influences its target gene in the receiver cell. These advancements enable the analysis and visualization of CCI pair's function, CCI's spatial pattern, significant CCI network, CCI-informed cell subtyping analysis, and comparison of CCI strength between groups. Besides its high performance, GITIII operates without the limitations related to the low or absent inclusion of measured ligand-receptor genes, establishing it as the optimal tool for comprehensive analysis of CCI in single-cell-resolution spatial transcriptomics.

We applied GITIII to four datasets from various species, organs, and platforms, demonstrating its capability to decode CCI influences and its superiority over other GNN models. This application has enriched our understanding of brain structure, tumor microenvironments, and CCI functions among diverse cell types. In the mouse brain dataset, GITIII reveals the heterogeneity of within-cell-type CCIs across different brain regions, pinpointing key gene markers with unique expression patterns in the same cell type across various brain layers. In the Alzheimer's disease dataset, GITIII was used to elucidate how glial cells affect excitatory neurons, which results in neuronal vulnerability. GITIII also identified cell subtypes with varying expressions of AD-risk genes and uncovered differences in CCI strengths between AD patients with distinct cognitive status. For the non-small cell lung cancer dataset, we use GITIII to quantitatively describe the tumor microenvironment and determine how each cell types respond to CCI from different sender cells and different kinds of tumor microenvironments. Lastly, in the

breast cancer dataset, we systematically examined how immune cells, tumor cells, and structural cells within the tumor microenvironment interact with and influence their neighboring cells, identifying how tumors manage to proliferate and migrate by altering the cell states of their neighboring cells.

While boosting a wide range of functionality and significantly outperforming other GNN models, there are still some limitations in GITIII. Firstly, although we can ascertain how cell state expression varies with the microenvironment, it remains challenging to discern whether these changes result from cell-cell interactions or programmed cell differentiation. Secondly, due to the lack of signaling pathways in the provided data, we can determine the influence of the sender cell on the receiver cell, but we are unable to identify the specific pathways through which this influence is exerted. To overcome these challenges, we plan to integrate single-cell RNA sequencing (scRNA-seq) data and accordingly modify the architecture of GITIII, which would enable us to identify the specific pathways through which sender cells influence each gene's states in receiver cells.

## Methods

### The algorithm of GITIII

The GITIII pipeline comprises four main components: the deep learning model, which includes the embedding module and a modified graph inductive bias transformer for estimating the influence of each sender cell on the target genes in each receiver cell; the CCI patterns visualization module; the cell subtyping analysis module; and the network analysis module. For convenience, a list of notations used below is included in Supplementary Note 1.

### Data preprocessing

Consider an imaging-based single-cell resolution spatial transcriptomics dataset, which includes the expression matrix, spatial coordinates, and cell type annotations for each cell from one or more tissue sections. We first normalize the expression matrix using log transformation to scale the values. Next, we calculate the average expression for each cell type to derive the cell type expression, and subtract the expression profile of each cell by their corresponding cell type expression to obtain their cell state expression. Subsequently, we identify the top k (default to 49) nearest neighbor for each cell to construct a cell neighborhood around this cell. Lastly, we exclude cells whose distance to their nearest neighbor exceeds a certain threshold (default to 80 pixels) as this may indicate that the cell was displaced from its original location during the slide preparation.

**Deep-learning model architecture**

Our deep learning model utilizes the features of each cell's neighborhood to predict its cell state expression, thereby learning how cells with varying states and types influence the central receiver cell (Figure 1a). The deep learning model consists of two components: the embedding module (Figure 1b) and a single-layer modified graph inductive bias transformer encoder (Figure 1c). In the model, each cell neighborhood is conceptualized as either a sentence, with cells representing individual words, or as a graph, with cells corresponding to nodes. After embedding and encoding, the model produces an influence tensor that quantifies the influences of each neighboring cell on the target genes of the central receiver cell. By summing these influences, the model generates a predicted cell state expression, which is then compared with the measured cell state expression for backward propagation.

**Model input**

The input to the deep learning model consists of three components, each characterizing a distinct aspect of a cellular neighborhood containing 50 cells. The first matrix, denoted as Exp,

encodes the expression profiles of all cells within this neighborhood. The second matrix, $D$, quantifies the distances between each pair of cells. The third component, cell type vector $T$, indicates the cell types of each individual cell. Together, these matrices furnish a comprehensive representation of the cellular microenvironment, facilitating our analysis. The dimensions of each matrix are as follows:

- The expression matrix $\text{Exp} \in \mathbb{R}^{50\times c}$, where 50 represents the number of cells in the neighborhood and $c$ is the number of measured genes in this dataset.
- The distance matrix $\text{distance} \in \mathbb{R}^{50\times 50}$, describing the pairwise distances between the cells.
- The cell type vector $T \in \mathbb{R}^{50}$, where each entry corresponds to one of the $t$ cell types identified in this dataset.

**Node feature embedding**

The node feature embedding consists of three components: expression masking, expression embedding, and cell type embedding. First, we substitute the actual expression of cells with the same cell type as the central cell (the cell to be predicted) with a close approximation of its cell-type expression. This substitution aims to prevent the model from relying on the common "the closer, the more similar" rule observed in spatial transcriptomics, which can act as a confounder. Instead, it encourages the model to learn from the influences of CCI. Next, we apply a multi-layer perceptron (MLP) to encode the expression features, yielding the expression embedding. Finally, we enhance the node features by adding the cell type embedding, which effectively captures the characteristics of the cell type categories. Consequently, the final node embedding $X \in \mathbb{R}^{50\times D_V}$ is defined as:

$$X = \text{MLP}\big(\text{mask}(\text{Exp})\big) + \text{embed}(T)$$

where $D_V$ denotes the node embedding dimension, MLP refers to an MLP transformation along the last dimension of the input tensor (see Supplementary Note 1 for more information), mask refers to the expression masking operation described earlier, and embed refers to the PyTorch token embedding layer.

**Distance embeddings**

To model the effects of distance, we employ five transformations of the original distance between each cell in the cell neighborhood, including a transformation based on the diffusion equation, and apply multi-layer perceptrons (MLPs) to accurately capture the complex effects of distance on CCI. Specifically, we generate three distance embeddings for edge embeddings and the graph transformer encoder, with two of them scaled to the range of 0-1 using the sigmoid function. These embeddings are denoted as Distance_embedding1 , and scaled Distance_embedding2 and Distance_embedding3 (see Supplementary Note 2 for details).

We diverge from the commonly used exponential decay diffusion model, which fails to adequately represent the dynamics in biological environments, for several reasons. First, certain ligand-receptor (LR) pairs are effective only when cells are in direct contact. In such cases, Fick's Second Law of Diffusion is not applicable, as there is no diffusion of the ligand. Second, the assumptions underlying the diffusion equation—specifically, that ligand concentration is initially zero except at the source and that space is infinite—are often not met in biological contexts. For example, tumor microenvironments may exhibit high local concentrations of growth factors secreted by cancer cells. Additionally, in the brain-blood barrier, diffusion is constrained not only by the selective permeability of the brain environment but also by the limited spaces within the brain's structural compartments. Third, the diffusion equation assumes that no reactive processes affect the particles during diffusion. However, in environments such as the brain, neurotransmitters

crossing synaptic gaps are rapidly broken down by enzymes like acetylcholinesterase, significantly reducing their concentration and effectiveness.

**Edge feature embedding**

Edge features are generated based on the distances between cells and the ligand expression levels of each sender cell, which capture the spatial organization and the measurable CCI patterns as reflected by the measured genes in the ligand-receptor (LR) database.

In brief, we first multiply the ligand expression by the scaled Distance_embedding2, which is then further encoded by a linear transformation. Subsequently, we concatenate this tensor with the unscaled Distance_embedding1 and apply another linear transformation to obtain the final edge embedding $E \in \mathbb{R}^{50\times50\times D_E}$, where $D_E$ represents the embedding dimension of the edges. By integrating the complex effects of distance on the ligands' impact, the edge embedding $E$ accurately characterizes the interactions between every pair of cells within the 50-cell neighborhood. For more details on the edge embedding, see Supplementary note 3.

**Graph attention calculation**

We adapted the Graph Inductive Bias Transformer (GRIT) model to more effectively encode both the graph structure and expression profiles within cellular neighborhoods. GRIT is recognized as a state-of-the-art model in deep learning for graph-structured data, distinguished by its exceptional ability to differentiate between heterogeneous graphs. To tailor the GRIT model for encoding features of cellular neighborhoods, we introduced three principal modifications. First, we replaced the original random walk graph structure encoding, which is useful in sparse graphs, with Distance_embedding3 to better capture the two-dimensional and dense spatial organization of cellular neighborhoods. Second, we refined the computation within the attention mechanism to better simulate the "AND logic" (interaction occurs only if both ligands and receptors are present)

and the distance-scaled nature of CCI. Lastly, we removed the degree scaler, which adjusts the data's scale based on node connectivity, recognizing that in dense graphs like cellular neighborhoods, each cell (node) generally maintains a consistent connectivity degree.

Similar to the original attention mechanism, we first generate the Key, Query, and Value matrices through three linear transformations of the node features. The Key and Query matrices are then matrix multiplied to form the attention matrix. This attention matrix is subsequently integrated with the edge features and distance embeddings to capture the distance-scaled nature of CCI. It is then normalized using the softmax function to obtain the final attention scores. Next, we combine the Value matrix with the edge features whose destination is the corresponding node. This operation not only enhances the graph encoding capacity but also reflects the nature of CCI, where the influences represented by edge features are aggregated across all neighboring cells to impact the receiver cell.

Now that we have the attention scores $\alpha \in \mathbb{R}^{50\times 50}$ and the Value matrix $V \in \mathbb{R}^{50\times D_V}$. We then select the attention scores from other cells in the cell neighborhood to the center receiver cell to predict and remove the row of this cell in the Value matrix by $\alpha = \alpha[0,1:] \in \mathbb{R}^{1\times 49}$ and $V = V[1:,:] \in \mathbb{R}^{49\times D_E}$. We then multiply the value matrix with the attention scores and use an MLP to transform the last dimension from $D_V$ to the number of measured genes $c$, forming the influence matrix $i \in \mathbb{R}^{49\times c}$ that represents the influence of each 49 neighboring cells on the $c$ target genes in the center receiver cell. Summing over the first dimension of $i$ we get the predicted cell state expression $\hat{y} \in \mathbb{R}^{c}$ which can be used for comparison with the true cell state expression to help the model learn how each cell with varying type and state influence the receiver cells. For more details, please refer to Supplementary Note 4.

**Interpretation of the influence tensor**

Concatenating the influence matrix for all cells within one slide results in the influence tensor $I \in \mathbb{R}^{N\times 49\times c}$, where $N$ indicates the number of cells within a slide. This tensor quantifies, for each of the $N$ cells on the slide, how their top 49 nearest neighbors influence the expression of the $c$ measured genes.

All non-attention transformations in our deep learning model are executed along the last dimension of the input tensor. Moreover, the attention mechanism maintains the order and representation of each token (cell), ensuring that the identity of each cell (token) is preserved throughout the encoding process. Recognizing this, the influence tensor $I$ can be interpreted as the impact of each of the 49 neighboring cells on the center cell targeted each $c$ measured gene.

Compared with interpreting the attention score $\alpha$ as the strength or scaler of CCI, the influence tensor offers enhanced accuracy. This improvement is because the influence of each token in the calculation of attention in our model or GAT is represented by the product of its embedding vector and the corresponding attention score. Consequently and apparently, relying solely on the attention score to represent influence would introduce bias.

**Normalization and aggregation of the influence tensor**

After obtaining the influence tensor, we aim to quantify how much, on average, each cell is influenced by other cell types. To achieve this, we normalize the influence tensor $I$ as follows:

$$I_p = \frac{|I|}{\text{sum}(|I|, \text{dim=-2, keepdim=True})}$$
$$I_s = \text{mean}\left(I_p, \text{dim=-1}\right)$$

This calculation yields two tensors: the normalized influence tensor $I_p \in \mathbb{R}^{N\times 49\times c}$ and $I_s \in \mathbb{R}^{N\times 49}$. Here, $I_p$ represents the proportional influence, ensuring that for each target gene in each cell, the corresponding influence from the 49 neighboring cells sums to one. $I_s$ estimates how each

of the $N$ cells is proportionally influenced by its top 49 nearest neighbors, which is averaged across all genes. $I_s$ will serve as a criterion for CCI pair selection in the subsequent visualization step.

In addition, users may set the values in $I$ with the corresponding $I_p$ less than a certain threshold (default to 0.02) and set them to 0 and re-calculate $I_p$ and $I_s$, since these small values are likely to be random noise.

We also aim to understand how target genes in each receiver cell are influenced by different cell types. To achieve this, we aggregate the influence tensor by sender cell type, resulting in a matrix $\mathbb{I} \in \mathbb{R}^{N \times t \times c}$. This matrix quantifies the influence of $t$ neighboring sender cell types on each of the $c$ genes expressed by $N$ receiver cells within the tissue sample.

**Visualization of CCI patterns**

GITIII enables three types of visualization to explore the spatial patterns and heterogeneity of CCI, including visualization of the effect of distance, visualization of CCI pairs (one sender cell influencing one receiver cell is called a CCI pair), and visualization of the spatial directions and strengths of CCI.

For the visualization of the distance effect, we flatten the influence tensor, take its absolute values, and plot them against the distance between the corresponding sender cell and receiver cell. We then fit a distance scaler line using Locally Weighted Scatterplot Smoothing (LOWESS). In the visualization of CCI pairs, we select the top 5 strongest CCI pairs for each receiver cell based on the values in $I_s$, treat the influence on the receiver cell targeting each gene as features, and use UMAP to visualize them. Here, one point in the UMAP represents one CCI pair, not one cell. For the visualization of the spatial directions and strengths of CCI, we also select the top 5 strongest CCI pairs for each cell and further select the top 5% strongest CCI pairs, plotting the strength and the positions of their corresponding sender and receiver cells on the slide.

**Cell subtyping analysis**

The CCI-informed interpretable microenvironment features $F$ for the $N$ cells in the tissue section are constructed by flattening the aggregated influence tensor $\mathbb{I} \in \mathbb{R}^{N \times t \times c}$ into $F \in \mathbb{R}^{N \times (t \times c)}$. Alternatively, one may construct $F$ by using the absolute values or by normalizing $\mathbb{I}$ along the second dimension before flattening to improve subtyping performance.

Based on $F$, UMAP can be used to visualize the microenvironment in which each cell resides. Leiden clustering [99] can then be performed to identify subgroups. Additionally, the corresponding aggregated influence tensor for cells in each subgroup can be visualized using heatmaps.

**Network analysis**

**Inference of z-scores for statistical test of significant CCI**

We developed an algorithm based on the Wald test of a partial linear model to determine whether interactions between sender and receiver cell types significantly influence the expression of each target gene in the receiver cells. Unlike previous methods that only controlled for type I error (p-value), our approach is also capable of controlling for type II error (1 - statistical power).

For a significant CCI between two cell types, a higher proportional influence should consistently lead to a higher or lower expression of the target gene in the receiver cell. This relationship posits that a significant correlation should exist between the proportional influence and cell state expression if there is a significant CCI from one sender cell type to one receiver cell type. To statistically test it, we construct a simple linear model defined as $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, where $Y_i$ represents the cell state expression of gene i of type B cells, $X$ denotes the proportional influence of type A cells on type B targeting gene i extracted from the aggregated influence tensor $\mathbb{I}$, and the error term $\epsilon_i$ is assumed to be independently and identically distributed (i.i.d.) following

a normal distribution with zero mean and unknown variance $\sigma^2$. The analysis focuses on testing the null hypothesis $\beta_1 = 0$ (CCI from ) using the Wald test. Under this framework, the z-score test statistic $z = \frac{\hat{\beta}_1}{\sqrt{\mathrm{var}(\hat{\beta}_1)}}$ is assumed to follow a Student-t distribution with $n_B - 2$ degrees of freedom. Given a large $n_B$, the number of type B cells per slide, the Student-t distribution approximates a standard normal distribution. The test statistic is computed as $z = \frac{r\sqrt{n_B-2}}{\sqrt{1-r^2}}$, where $r$ is the Pearson correlation coefficient $r = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}}$. The corresponding two-sided test p-value is calculated by:

$$p_{iAB} = 2\big(1 - \Phi(|z_{iAB}|)\big)$$

where $\Phi$ is the cumulative distribution function of the normal distribution. Then, p-values are adjusted for multiple comparisons using the Benjamini–Hochberg procedure, with interactions deemed significant at a threshold below 0.05.

**Calculation of statistical power for statistical test of significant CCI**

After calculating the p-values, we first find the largest p-values $\alpha$ that produce a significant adjusted p-value after the p-value adjustment. Its corresponding z-score is $z_{1-\alpha/2} = -\Phi^{-1}(\alpha/2)$, then the corresponding type II error $\beta$ can be expressed as:

$$\beta = \Phi\big(-|z_{iAB}| + z_{1-\alpha/2}\big) - \Phi\big(-|z_{iAB}| - z_{1-\alpha/2}\big)$$

Then, the statistical power of this test can be calculated by $1 - \beta$.

**Aggregation of CCI network across samples**

We employ Stouffer's Z-score method and Fisher's combined probability test to integrate patient-level inferred Cell-Cell Interaction (CCI) scores into a group-level CCI based on the metadata for each patient. For instance, consider the inferred CCI from cell type A to cell type B targeting gene i, for all the patients, we get a z-value vector $Z_{iAB} \in \mathbb{R}^S$, where $S$ denotes the

number of patients in the group. Then the aggregated z-value and its corresponding p-value for Stouffer's Z-score method are calculated using the following formulas:

$$\bar{Z_{iAB}} = \frac{\sum_{s=1}^{S} Z_{siAB}}{\sqrt{S}}$$
$$\bar{p_{iAB}} = 2(1 - \Phi(|\bar{Z_{iAB}}|)$$

where $Z_{siAB}$ donates the z-score of CCI from cell type A to cell type B targeting gene i in s-th sample.

To ensure the statistical significance for group-aggregated CCI network, we further applied Fisher's combined probability test to determine the significance of aggregated network:

$$p_{siAB} = 2\big(1 - \Phi(|Z_{siAB}|)\big)$$
$$\chi^2 = -2\sum_{s=1}^{S} \ln(p_{siAB})$$
$$\bar{p_{iAB}} = 1 - \text{chi2cdf}(\chi^2, 2S)$$

where chi2cdf donates the cumulative distribution function of the chi-squared distribution.

**Computation of CCI strength**

Previous studies on calculating CCI strength have predominantly utilized two methods: multiplying Ligand-Receptor (LR) pairs and spatially permuting all cells within a slide. However, the multiplication of LR pairs is susceptible to batch effects, and spatial permutation across entire slides yields inconsistent results when applied to subsets of the original slides. To address these issues, we developed a novel method for computing CCI strength based on permutations within cell neighborhoods.

We begin by defining the cell neighborhood. This involves counting the number of different cell types that exert a significant influence, previously discussed as not being filtered out and set to zero. Considering a specific CCI targeting gene $i$ within a single slide, suppose $n_a$ cells of type A significantly influence $n_b$ cells of type B, out of a total of $N_b$ cells sending non-zero influence

to cells of type B. The actual proportional influence, $I_A$, is calculated as the sum of influences from all $n_a$ cells of type A that have non-zero influences on $n_b$ cells of type B. The inferred communication strength is then defined as $\frac{I_A}{n_b} - \frac{n_a}{N_b}$, quantifying the extent to which the observed influence from cell type A on cell type B exceeds what is expected in the specific neighborhood.

**Implementation details**

To optimize our model's performance without requiring extensive GPU resources, we carefully selected a modest hyperparameter setting for the number of parameters in the model. Specifically, for all models across the four datasets, the number of cells in a cell neighborhood is set to 50 and the embedding dimension of the node features is fixed at 256. The dimensionality of the edge features varies according to the number of measured ligand genes in each dataset, resulting in the edge dimensions of 48, 48, 128, and 128 for the Mouse brain, AD, NSCLC, and IBD datasets, respectively. For the optimization process, we chose a learning rate of 0.0001, utilizing the AdamW optimizer (citation needed here). Regarding the batch size, we set it at 90 for both MERFISH datasets and 30 for the two Cosmx datasets, aligning with the 24 GB GPU memory capacity of the NVIDIA RTX A5000 Graphics Card. This configuration ensures that our model is both efficient and accessible, even on less powerful hardware setups. We used Pytorch 2.0.1 to build our deep learning model and Scanpy for data preprocessing and the conversion of raw data into CSV format, which is compatible with our model's dataloader.

During the training process, we randomly (with the fixed random seed) divided the entire dataset into training and validation subsets using a 4:1 ratio with a fixed random seed. We trained the model for 50 epochs and the model exhibiting the lowest Mean Squared Error (MSE) loss on the validation set was selected for downstream analysis.

**Code availability**

GITIII is implemented in the open-source Python using PyTorch. The source code and a detailed tutorial are publicly available at (https://github.com/lugia-xiao/GITIII).

**Data availability**

The mouse brain MERFISH dataset is available at ftp://download.brainimagelibrary.org:8811/02/26/02265ddb0dae51de/. The SEA-AD dataset is available at https://sea-ad-spatial-transcriptomics.s3.amazonaws.com/middle-temporal-gyrus/all_donors-h5ad/SEAAD_MTG_MERFISH.2024-02-13.h5ad. The CosMx human NSCLC dataset is available at https://nanostring.com/products/cosmx-spatial-molecular-imager/ffpe-dataset/nsclc-ffpe-dataset/. The breast cancer dataset is available at https://www.10xgenomics.com/products/xenium-in-situ/preview-dataset-human-breast. The ligand-receptor database of CellChat, NeuronChat, and MultiNicheNet is publicly available, and an integrated version is available on our GitHub page.

**Competing Interests**

No competing interests.

**References**


1. Efremova, M., et al., *CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes.* Nat Protoc, 2020. **15**(4): p. 1484-1506.
2. Jin, S., et al., *Inference and analysis of cell-cell communication using CellChat.* Nat Commun, 2021. **12**(1): p. 1088.
3. Zhao, W., et al., *Inferring neuron-neuron communications from single-cell transcriptomics through NeuronChat.* Nat Commun, 2023. **14**(1): p. 1128.
4. Graeber, T.G. and D. Eisenberg, *Bioinformatic identification of potential autocrine signaling loops in cancers from gene expression profiles.* Nat Genet, 2001. **29**(3): p. 295-300.
5. Choi, H., et al., *Transcriptome analysis of individual stromal cell populations identifies stroma-tumor crosstalk in mouse lung cancer model.* Cell Rep, 2015. **10**(7): p. 1187-201.

6. Wang, T., et al., *Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data.* BMC Bioinformatics, 2019. **20**(1): p. 40.

7. Tsuyuzaki, K., M. Ishii, and I. Nikaido, *Sctensor detects many-to-many cell-cell interactions from single cell RNA-sequencing data.* BMC Bioinformatics, 2023. **24**(1): p. 420.

8. Browaeys, R., W. Saelens, and Y. Saeys, *NicheNet: modeling intercellular communication by linking ligands to target genes.* Nat Methods, 2020. **17**(2): p. 159-162.

9. Li, L.X., et al., *Single-Cell and CellChat Resolution Identifies Collecting Duct Cell Subsets and Their Communications with Adjacent Cells in PKD Kidneys.* Cells, 2022. **12**(1).

10. Stickels, R.R., et al., *Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2.* Nat Biotechnol, 2021. **39**(3): p. 313-319.

11. Eng, C.L., et al., *Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH.* Nature, 2019. **568**(7751): p. 235-239.

12. Chen, K.H., et al., *RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells.* Science, 2015. **348**(6233): p. aaa6090.

13. Janesick, A., et al., *High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis.* Nat Commun, 2023. **14**(1): p. 8353.

14. He, S., et al., *High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging.* Nat Biotechnol, 2022. **40**(12): p. 1794-1806.

15. Armingol, E., H.M. Baghdassarian, and N.E. Lewis, *The diversification of methods for studying cell-cell interactions and communication.* Nat Rev Genet, 2024. **25**(6): p. 381-400.

16. Shao, X., et al., *Knowledge-graph-based cell-cell communication inference for spatially resolved transcriptomic data with SpaTalk.* Nat Commun, 2022. **13**(1): p. 4429.

17. Cang, Z., et al., *Screening cell-cell communication in spatial transcriptomics via collective optimal transport.* Nat Methods, 2023. **20**(2): p. 218-228.

18. Garcia-Alonso, L., et al., *Mapping the temporal and spatial dynamics of the human endometrium in vivo and in vitro.* Nat Genet, 2021. **53**(12): p. 1698-1711.

19. Arnol, D., et al., *Modeling Cell-Cell Interactions from Spatial Molecular Data with Spatial Variance Component Analysis.* Cell Rep, 2019. **29**(1): p. 202-211 e6.

20. Li, Z., et al., *SpatialDM for rapid identification of spatially co-expressed ligand-receptor and revealing cell-cell communication patterns.* Nat Commun, 2023. **14**(1): p. 3995.

21. Wu, D., et al., *Inferring Cell-Cell Communications from Spatially Resolved Transcriptomics Data Using a Bayesian Tweedie Model.* Genes (Basel), 2023. **14**(7).

22. So, E., et al., *GraphComm: A Graph-based Deep Learning Method to Predict Cell-Cell Communication in single-cell RNAseq data.* bioRxiv, 2023: p. 2023.04.26.538432.

23. Lück, N., et al., *SpaCeNet: Spatial Cellular Networks from omics data.* bioRxiv, 2022: p. 2022.09.01.506219.

24. Hu, J., et al., *SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network.* Nat Methods, 2021. **18**(11): p. 1342-1351.

25. Long, Y., et al., *Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST.* Nat Commun, 2023. **14**(1): p. 1155.

26. Zhang, X., et al., *Graph-based autoencoder integrates spatial transcriptomics with chromatin images and identifies joint biomarkers for Alzheimer's disease.* Nat Commun, 2022. **13**(1): p. 7480.

27. Xiao, X., et al., *Transformer with convolution and graph-node co-embedding: An accurate and interpretable vision backbone for predicting gene expressions from local histopathological image.* Med Image Anal, 2024. **91**: p. 103040.

28. Yuan, Y. and Z. Bar-Joseph, *GCNG: graph convolutional networks for inferring gene interaction from spatial transcriptomics data.* Genome Biol, 2020. **21**(1): p. 300.

29. Fischer, D.S., A.C. Schaar, and F.J. Theis, *Modeling intercellular communication in tissues using spatial graphs of cells.* Nat Biotechnol, 2023. **41**(3): p. 332-336.

30. Kipf, T.N. and M. Welling *Semi-Supervised Classification with Graph Convolutional Networks*. 2016. arXiv:1609.02907 DOI: 10.48550/arXiv.1609.02907.

31. Zuo, C., J. Xia, and L. Chen, *Dissecting tumor microenvironment from spatially resolved transcriptomics data by heterogeneous graph learning.* Nat Commun, 2024. **15**(1): p. 5057.

32. Yang, W., et al., *Deciphering cell-cell communication at single-cell resolution for spatial transcriptomics with subgraph-based graph attention network.* Nat Commun, 2024. **15**(1): p. 7101.

33. Li, Y., et al., *Tissue module discovery in single-cell-resolution spatial transcriptomics data via cell-cell interaction-aware cell embedding.* Cell Syst, 2024. **15**(6): p. 578-592 e7.

34. Veličković, P., et al. *Graph Attention Networks*. 2017. arXiv:1710.10903 DOI: 10.48550/arXiv.1710.10903.

35. Wang, H., et al., *CGCom: a framework for inferring Cell-cell Communication based on Graph Neural Network.* bioRxiv, 2023.

36. Li, H., et al., *Decoding functional cell-cell communication events by multi-view graph learning on spatial transcriptomics.* Brief Bioinform, 2023. **24**(6).

37. Khan, M.R. and J.E. Blumenstock *Multi-GCN: Graph Convolutional Networks for Multi-View Networks, with Applications to Global Poverty*. 2019. arXiv:1901.11213 DOI: 10.48550/arXiv.1901.11213.

38. Ma, L., et al. *Graph Inductive Biases in Transformers without Message Passing*. 2023. arXiv:2305.17589 DOI: 10.48550/arXiv.2305.17589.

39. Zhang, M., et al., *Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH.* Nature, 2021. **598**(7879): p. 137-143.

40. McInnes, L., J. Healy, and J. Melville *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. 2018. arXiv:1802.03426 DOI: 10.48550/arXiv.1802.03426.

41. Ben Haim, L. and D.H. Rowitch, *Functional diversity of astrocytes in neural circuit regulation.* Nat Rev Neurosci, 2017. **18**(1): p. 31-41.

42. Molina-Gonzalez, I., et al., *Astrocyte-oligodendrocyte interaction regulates central nervous system regeneration.* Nat Commun, 2023. **14**(1): p. 3372.

43. Cubelos, B., et al., *Cux-2 controls the proliferation of neuronal intermediate precursors of the cortical subventricular zone.* Cereb Cortex, 2008. **18**(8): p. 1758-70.

44. Nagata, I., G. Keilhauer, and M. Schachner, *Neuronal influence on antigenic marker profile, cell shape and proliferation of cultured astrocytes obtained by microdissection of distinct layers from the early postnatal mouse cerebellum.* Brain Res, 1986. **389**(1-2): p. 217-32.

45. Gabitto, M.I., et al., *Integrated multimodal cell atlas of Alzheimer's disease.* Res Sq, 2023.

46. Leng, K., et al., *Molecular characterization of selectively vulnerable neurons in Alzheimer's disease.* Nat Neurosci, 2021. **24**(2): p. 276-287.

47. Leng, F. and P. Edison, *Neuroinflammation and microglial activation in Alzheimer disease: where do we go from here?* Nat Rev Neurol, 2021. **17**(3): p. 157-172.

48. Gao, C., et al., *Microglia in neurodegenerative diseases: mechanism and potential therapeutic targets.* Signal Transduct Target Ther, 2023. **8**(1): p. 359.

49. Suzuki, T., et al., *CUX2 deficiency causes facilitation of excitatory synaptic transmission onto hippocampus and increased seizure susceptibility to kainate.* Sci Rep, 2022. **12**(1): p. 6505.

50. Fujita, Y., et al., *Netrin-G1 Regulates Microglial Accumulation along Axons and Supports the Survival of Layer V Neurons in the Postnatal Mouse Brain.* Cell Rep, 2020. **31**(4): p. 107580.

51. Yaguchi, K., et al., *Identification of transcriptional regulatory elements for Ntng1 and Ntng2 genes in mice.* Mol Brain, 2014. **7**: p. 19.

52. McNamara, N.B., et al., *Microglia regulate central nervous system myelin growth and integrity.* Nature, 2023. **613**(7942): p. 120-129.

53. Whalley, K., *Neuronal neighbours tune microglial identity.* Nat Rev Neurosci, 2022. **23**(10): p. 582-583.

54. Kuhn, S., et al., *Oligodendrocytes in Development, Myelin Generation and Beyond.* Cells, 2019. **8**(11).

55. Desai, M.K., et al., *Early oligodendrocyte/myelin pathology in Alzheimer's disease mice constitutes a novel therapeutic target.* Am J Pathol, 2010. **177**(3): p. 1422-35.

56. Depp, C., et al., *Myelin dysfunction drives amyloid-beta deposition in models of Alzheimer's disease.* Nature, 2023. **618**(7964): p. 349-357.

57. Stadelmann, C., et al., *Myelin in the Central Nervous System: Structure, Function, and Pathology.* Physiol Rev, 2019. **99**(3): p. 1381-1431.

58. Su, H., et al., *The biological function and significance of CD74 in immune diseases.* Inflamm Res, 2017. **66**(3): p. 209-216.

59. Bryan, K.J., et al., *Expression of CD74 is increased in neurofibrillary tangles in Alzheimer's disease.* Molecular Neurodegeneration, 2008. **3**.

60. Olah, M., et al., *Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease.* Nat Commun, 2020. **11**(1): p. 6129.

61. Colonna, M. and Y. Wang, *TREM2 variants: new keys to decipher Alzheimer disease pathogenesis.* Nat Rev Neurosci, 2016. **17**(4): p. 201-7.

62. Ulland, T.K. and M. Colonna, *TREM2 - a key player in microglial biology and Alzheimer disease.* Nat Rev Neurol, 2018. **14**(11): p. 667-675.

63. McQuade, A., et al., *Gene expression and functional deficits underlie TREM2-knockout microglia responses in human models of Alzheimer's disease.* Nat Commun, 2020. **11**(1): p. 5370.

64. Poliani, P.L., et al., *TREM2 sustains microglial expansion during aging and response to demyelination.* J Clin Invest, 2015. **125**(5): p. 2161-70.

65. Liu, T., et al., *Soluble TREM-1, as a new ligand for the membrane receptor Robo2, promotes hepatic stellate cells activation and liver fibrosis.* Journal of Cellular and Molecular Medicine, 2021. **25**(24): p. 11113-11127.

66. Zhang, J., et al., *Role of CD68 in tumor immunity and prognosis prediction in pan-cancer.* Sci Rep, 2022. **12**(1): p. 7844.

67. van Leent, M.M.T., et al., *Prosaposin mediates inflammation in atherosclerosis.* Sci Transl Med, 2021. **13**(584).

68. Jin, W., et al., *Single-cell RNA-Seq reveals transcriptional heterogeneity and immune subtypes associated with disease activity in human myasthenia gravis.* Cell Discov, 2021. **7**(1): p. 85.

69. Bill, R., et al., *macrophage polarity identifies a network of cellular programs that control human cancers.* Science, 2023. **381**(6657): p. 515-524.

70. Yang, Q., et al., *Single-Cell RNA Sequencing Reveals the Heterogeneity of Tumor-Associated Macrophage in Non-Small Cell Lung Cancer and Differences Between Sexes.* Frontiers in Immunology, 2021. **12**.

71. Lichtenberger, B.M., et al., *Autocrine VEGF signaling synergizes with EGFR in tumor cells to promote epithelial cancer development.* Cell, 2010. **140**(2): p. 268-79.

72. Sharma, S.V., et al., *Epidermal growth factor receptor mutations in lung cancer.* Nat Rev Cancer, 2007. **7**(3): p. 169-81.

73. Lazaratos, A.M., M.G. Annis, and P.M. Siegel, *GPNMB: a potent inducer of immunosuppression in cancer.* Oncogene, 2022. **41**(41): p. 4573-4590.

74. Saha, S.K., et al., *KRT19 directly interacts with beta-catenin/RAC1 complex to regulate NUMB-dependent NOTCH signaling pathway and breast cancer properties.* Oncogene, 2017. **36**(3): p. 332-349.

75. Arteaga, C.L., *Overview of epidermal growth factor receptor biology and its role as a therapeutic target in human neoplasia.* Seminars in Oncology, 2002. **29**(5): p. 3-9.

76. Werner, S.L., et al., *Abstract A144: Therapeutically targeting high-affinity ligand activation of EGFR with MM-151, an oligoclonal therapeutic.* Molecular Cancer Therapeutics, 2011. **10**(11_Supplement): p. A144-A144.

77. Astsaturov, I., et al., *Synthetic lethal screen of an EGFR-centered network to improve targeted therapies.* Sci Signal, 2010. **3**(140): p. ra67.

78. Li, Y.Y., et al., *Cancer-associated fibroblasts contribute to oral cancer cells proliferation and metastasis via exosome-mediated paracrine miR-34a-5p.* EBioMedicine, 2018. **36**: p. 209-220.

79. Che, D., et al., *KRT6A Promotes Lung Cancer Cell Growth and Invasion Through MYC-Regulated Pentose Phosphate Pathway.* Frontiers in Cell and Developmental Biology, 2021. **9**.

80. Maric, G., et al., *GPNMB augments Wnt-1 mediated breast tumor initiation and growth by enhancing PI3K/AKT/mTOR pathway signaling and beta-catenin activity.* Oncogene, 2019. **38**(26): p. 5294-5307.

81. Wu, T., et al., *clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.* Innovation (Camb), 2021. **2**(3): p. 100141.

82. DeLeon-Pennell, K.Y., T.H. Barker, and M.L. Lindsey, *Fibroblasts: The arbiters of extracellular matrix remodeling.* Matrix Biol, 2020. **91-92**: p. 1-7.

83. Lai, Y., et al., *Interleukin-8 Induces the Endothelial Cell Migration through the Activation of Phosphoinositide 3-Kinase-Rac1/RhoA Pathway.* International Journal of Biological Sciences, 2011. **7**(6): p. 782-791.

84. Denisenko, T.V., A.D. Pivnyuk, and B. Zhivotovsky, *p53-Autophagy-Metastasis Link.* Cancers (Basel), 2018. **10**(5).

85. Jung, H., H.A. Seong, and H. Ha, *Critical role of cysteine residue 81 of macrophage migration inhibitory factor (MIF) in MIF-induced inhibition of p53 activity.* J Biol Chem, 2008. **283**(29): p. 20383-96.

86. Pavlyukov, M.S., et al., *Apoptotic Cell-Derived Extracellular Vesicles Promote Malignancy of Glioblastoma Via Intercellular Transfer of Splicing Factors.* Cancer Cell, 2018. **34**(1): p. 119-135 e10.

87. Chen, F., et al., *Extracellular vesicle-packaged HIF-1alpha-stabilizing lncRNA from tumour-associated macrophages regulates aerobic glycolysis of breast cancer cells.* Nat Cell Biol, 2019. **21**(4): p. 498-510.

88. Horrevorts, S.K., et al., *Glycan-Modified Apoptotic Melanoma-Derived Extracellular Vesicles as Antigen Source for Anti-Tumor Vaccination.* Cancers, 2019. **11**(9).

89. Gates, J.D., et al., *Circulating regulatory T cells (CD4+CD25+FOXP3+) decrease in breast cancer patients after vaccination with a modified MHC class II HER2/neu (AE37) peptide.* Vaccine, 2010. **28**(47): p. 7476-82.

90. Lonne, G.K., et al., *Protein Kinase Cδ Supports Survival of MDA-MB-231 Breast Cancer Cells by Suppressing the ERK1/2 Pathway.* Journal of Biological Chemistry, 2009. **284**(48): p. 33456-33465.

91. He, Y., et al., *FOXA1 overexpression suppresses interferon signaling and immune response in cancer.* J Clin Invest, 2021. **131**(14).

92. Schultz, S., et al., *Progression-specific genes identified in microdissected formalin-fixed and paraffin-embedded tissue containing matched ductal carcinoma in situ and invasive ductal breast cancers.* BMC Med Genomics, 2018. **11**(1): p. 80.

93. Takahashi, Y., et al., *Expression of vascular endothelial growth factor and its receptor, KDR, correlates with vascularity, metastasis, and proliferation of human colon cancer.* Cancer Res, 1995. **55**(18): p. 3964-8.

94. Cao, H.H., et al., *In leukemia, knock-down of the death inducer-obliterator gene would inhibit the proliferation of endothelial cells by inhibiting the expression of*

*and.* Peerj, 2022. **10**.

95. Hanley, C.J., et al., *Tumor-Resident Stromal Cells Promote Breast Cancer Invasion through Regulation of the Basal Phenotype.* Molecular Cancer Research, 2020. **18**(11): p. 1615-1622.

96. Wang, B., et al., *Spatial-MGCN: a novel multi-view graph convolutional network for identifying spatial domains with attention mechanism.* Brief Bioinform, 2023. **24**(5).

97. Tang, Z., et al., *spaCI: deciphering spatial cellular communications through adaptive graph model.* Brief Bioinform, 2023. **24**(1).

98. Tang, Z., et al., *SiGra: single-cell spatial elucidation through an image-augmented graph transformer.* Nat Commun, 2023. **14**(1): p. 5618.

99. Traag, V.A., L. Waltman, and N.J. van Eck, *From Louvain to Leiden: guaranteeing well-connected communities.* Scientific Reports, 2019. **9**: p. 5233.

**Figures**

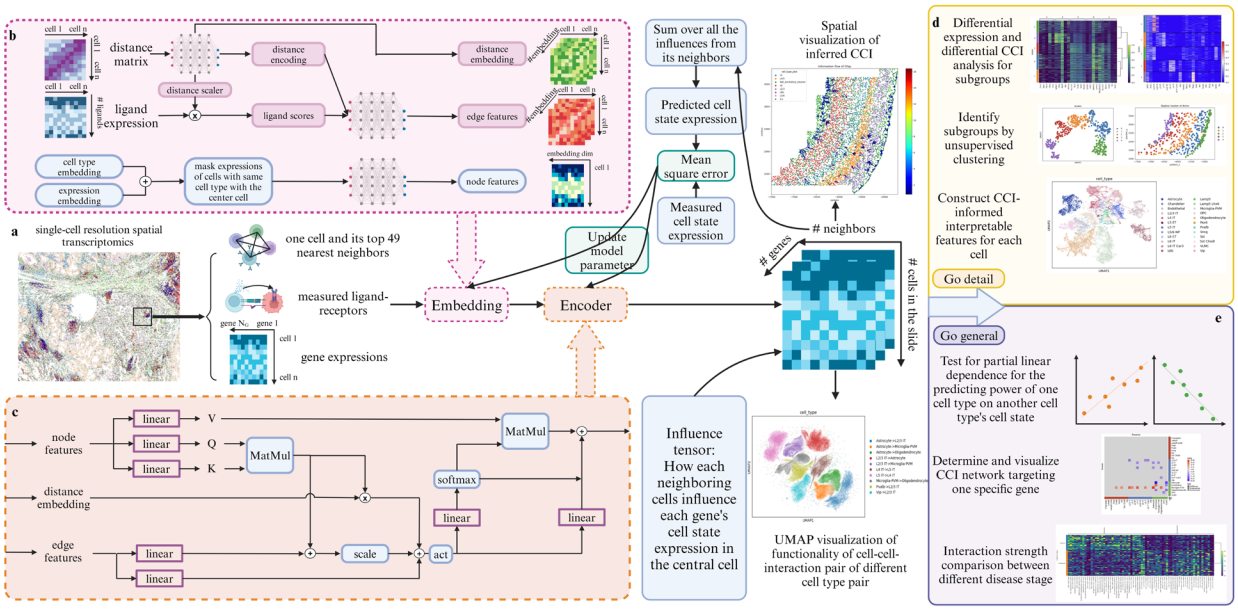


**Figure 1. Overview of GITIII. a.** GITIII Workflow. The model employs self-supervised learning to decipher the influence of cell-cell interaction (CCI), predicting cell state expression from ligand data, neighboring cell type and cell expression profiles, and spatial organization. Subsequently, we evaluate our model's interpretability by analyzing how each cell communication pair affects gene expression in the receiver cell, facilitating visualization and further analyses. **b.** Detailed view of the input data embeddings. The craft of node features incorporates cell type embeddings and expression encodings, while the edge features between two nodes (cells) contain the ligand gene expression scaled by distance and the embedding of the distance between two cells. **c.** Detailed view of the GITIII encoder layer. A specially designed graph attention mechanism was used to better capture the CCI patterns in the cell neighborhood. The edge features are updated through linear transformations combined with the attention matrix from the node features and the embedding of distance relationships. Node features are aggregated in a manner akin to traditional attention mechanisms, augmented by edge features weighted by attention scores **d.** Functionalities of GITIII in analyzing within-sample CCI heterogeneity. CCI-informed interpretable features were first constructed by analyzing how much this cell's cell state expression for each gene is influenced

by its neighboring cell types, it is then visualized by UMAP and further divided into different sub-groups. Differential expression and differential CCI analysis are then performed to interpret the differences between sub-groups. **e.** Functionalities of GITIII in identifying CCI network. Significant CCI networks that illustrate how each sender cell types influence one specific gene's cell state expression in the receiver cell types are first determined by the Wald test on partial linear regression. Then the strength of CCI is calculated by neighborhood permutation, which enables comparison of CCI network between patient groups.

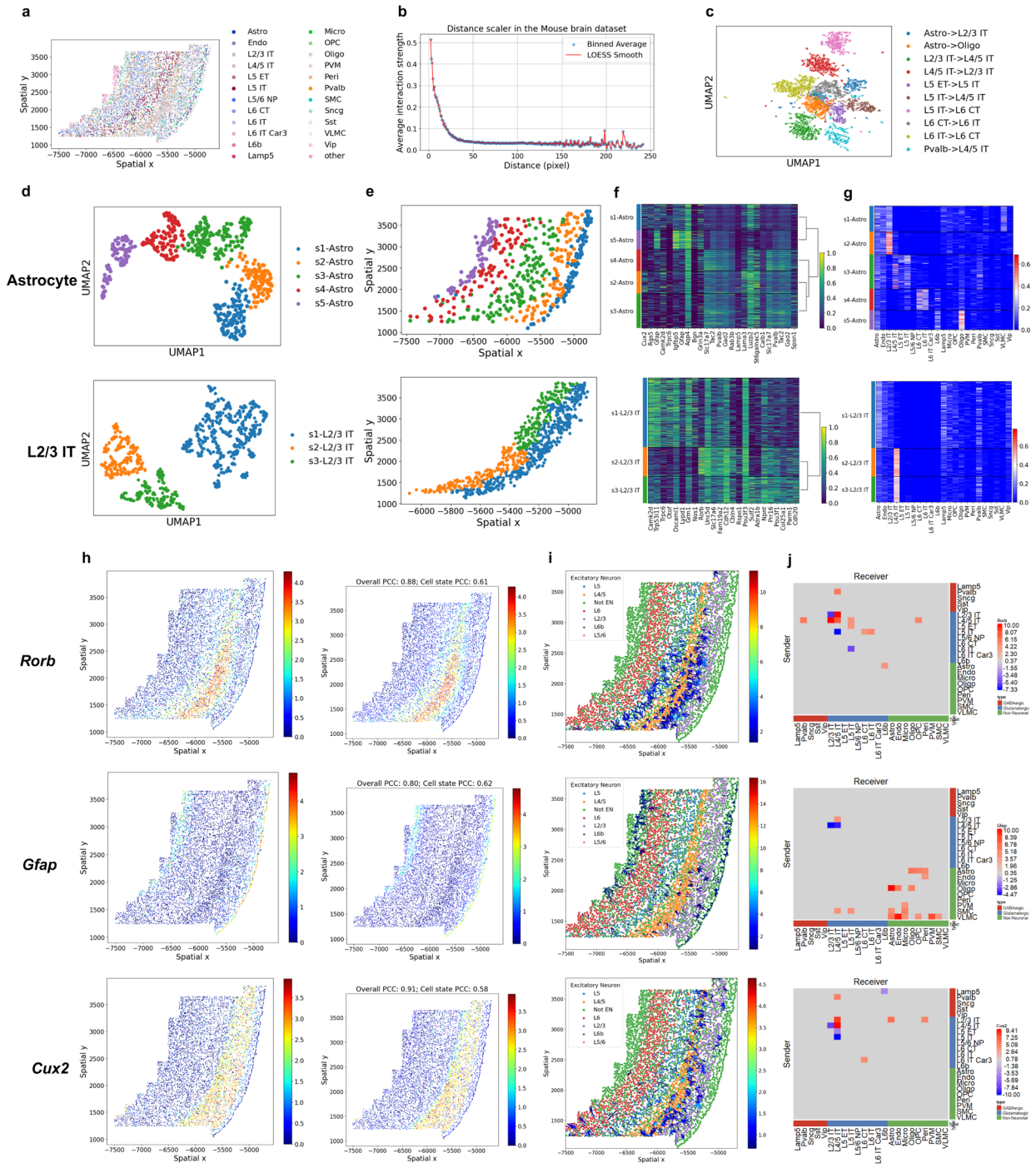


**Figure 2. Visualization of intercellular communication networks and CCI-based subtyping for mouse primary motor cortex. a.** Spatial distribution of different cell types in the "mouse1_slice201" slide. **b.** The relationship between the average CCI strength and the distance

between two cells. **c.** UMAP visualization of the top 5 strongest CCI pairs for each cell in the "mouse1_slice201" slide, colored by the cell types combination of sender and receiver cells. The features of one point (CCI pair) indicate how much each target gene in a receiver cell is influenced by a sender cell. **d.** The UMAP visualization of the CCI-informed interpretable features of astrocytes (upper) and L2/3 IT neurons (lower), with color indicating the identified subtypes by the unsupervised Leiden clustering method. **e.** The spatial distribution of each identified subtype in astrocytes (upper) and L2/3 IT neurons (lower). **f.** Differential expression analysis for each identified subtype using the Wilcoxon rank-sum test. **g.** Proportional influence on each cell by their corresponding neighboring cell types averaged across all genes. **h.** Visualization of measured gene expression (left) and predicted gene expression (right) of *Rorb*, *Gfap*, and *Cux2*. **i.** Visualization of predicted spatial CCI distributions**.** The figure shows the top 5% strongest CCIs targeting the corresponding genes, with the arrow indicating the influences from one sender cell to one receiver cell and the color of the arrow indicating the magnitude of the influence. **j.** The inferred CCI networks between different cell types targeting genes of Rorb, Gfap, and Cux2, where values indicating the corresponding z-scores of the probability of one sender cell type significantly influence the given gene in the receiver cell type, with negative values indicating down-regulation and positive values indicate up-regulation.

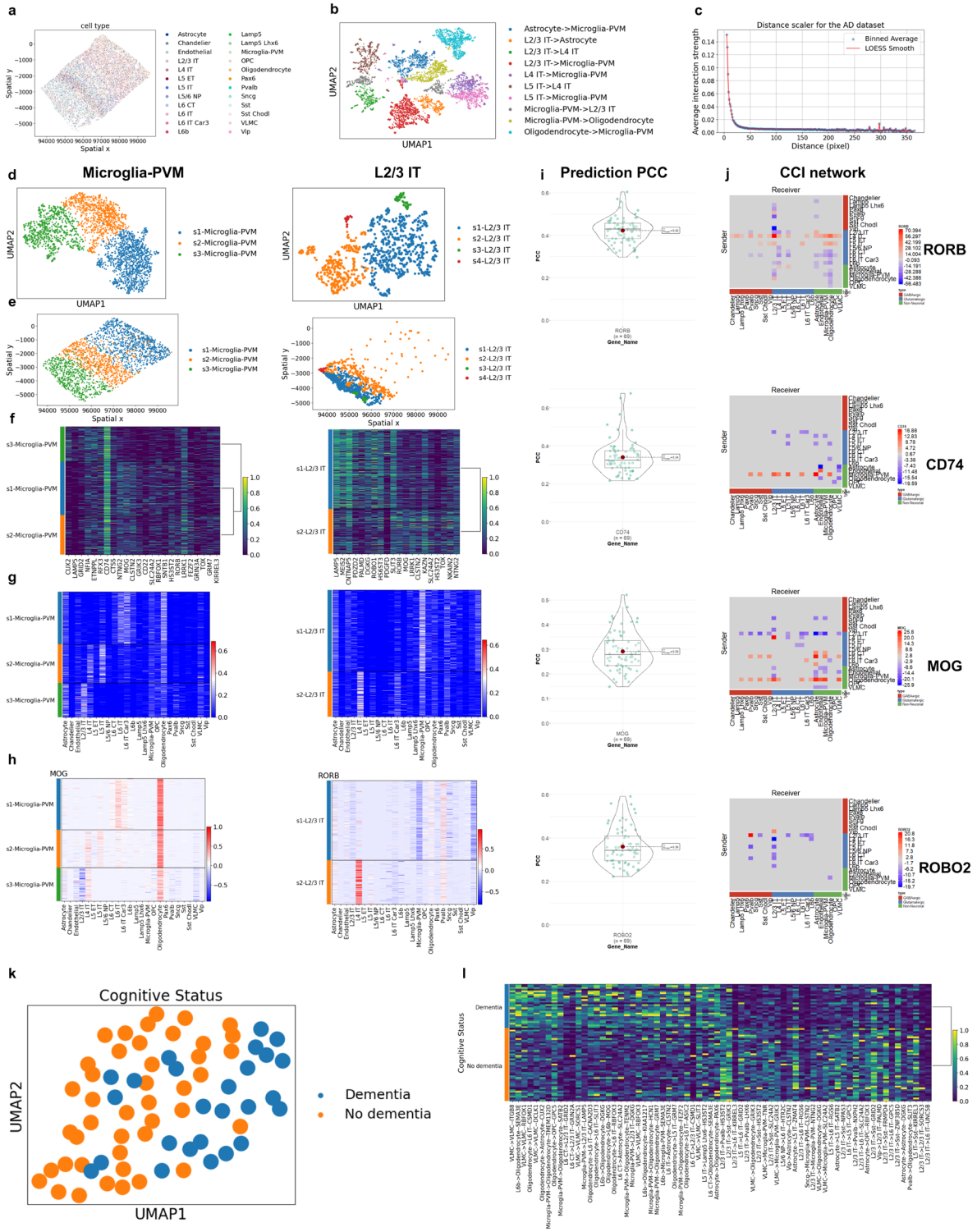


**Figure 3. Various CCI patterns within and across patients in the SEA-AD dataset. a.** Spatial distribution of 24 cell types on the 'H20.33.001.CX28.MTG.02.007.1.02.03' slide. **b.** UMAP visualization demonstrating the function of the top 10 most frequently occurring CCI cell type

combinations, with the features of one point indicating how much each target gene in a receiver cell is influenced by a sender cell. **c.** The relationship between the average CCI strength and the distance between two cells. **d.** The UMAP visualization of the CCI-informed interpretable features of microglia (upper) and L2/3 IT neurons (lower), with color indicating the identified subtypes by the unsupervised Leiden clustering method. **e.** The spatial distribution of each identified subtype in microglia (upper) and L2/3 IT neurons (lower). **f.** Differential expression analysis for each identified subtype using the Wilcoxon rank-sum test. **g.** Proportional influence on each cell by their corresponding neighboring cell types averaged across all genes. **h.** The proportional influence of other cell types on the *MOG* gene of microglia targeting and on the *RORB* gene of L2/3 IT, with positive values indicating up-regulation and negative values indicating down-regulation. **i.** Visualization of our model's prediction accuracy for genes *RORB*, *MOG*, *CD74*, and ROBO2. **j.** Aggregated CCI networks across all samples in the SEA-AD dataset targeting genes of RORB, MOG, CD74, and ROBO2, where values indicating the corresponding z-scores of the probability of one sender cell type significantly communicate with one receiver cell type targeting the given gene, with negative values indicating down-regulation and positive values indicating up-regulation. **k.** UMAP visualization of each slide using the inferred CCI strength as features, with color indicating which cognitive status group this slide belongs to. **l.** Differential CCI strength analysis between the two cognitive status groups, with features indicating the CCI from the sender cell type to the receiver cell type targeting one gene.

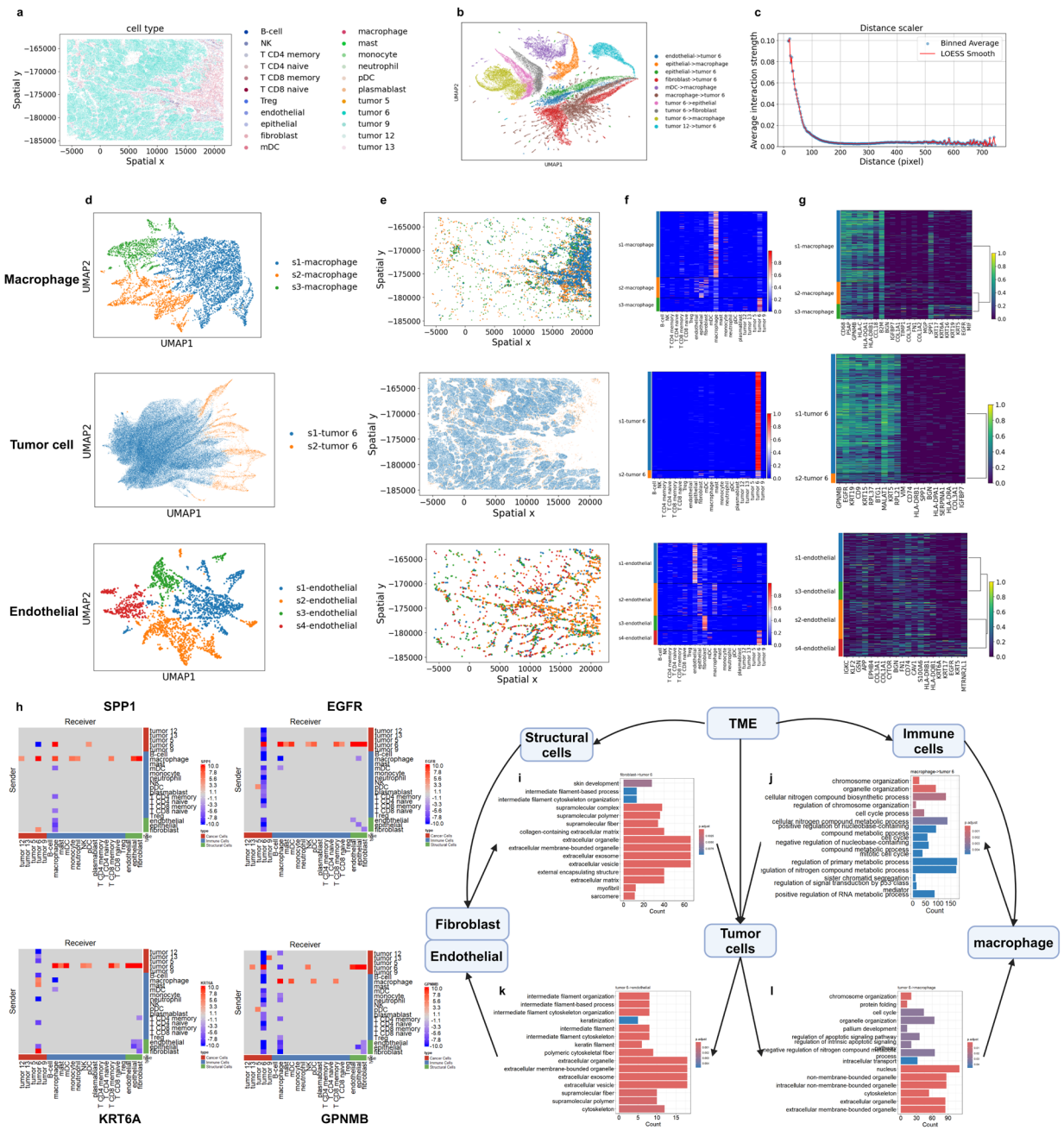


**Figure 4. Quantitative analysis of tumor microenvironment in NSCLC and its impact on different cell types. a.** Spatial distribution of 22 cell types within the 'Lung6' slide. **b.** UMAP visualization demonstrating the function of the top 10 most frequently occurring CCI cell type combinations, with the features of one point indicating how much each target gene in a receiver cell is influenced by a sender cell. **c.** The relationship between the average CCI strength and the

distance between two cells. **d-g.** Subtyping analysis of macrophage (upper), tumor cell (middle), and endothelial cells (lower) by **(d)** UMAP visualization of interpretable CCI-informed features of each cell, **(e)** Spatial distribution cells in each sub-group, **(f)** Proportional influence on each cell by their corresponding neighboring cell types averaged across all genes, and **(g)** Differential expression analysis for each sub-group. **h.** Visualization of CCI networks targeting SPP1, EGFR, KRT6A, and GPNMB, where values indicate the corresponding z-scores of the probability of one sender cell type significantly communicating with one receiver cell type targeting the given gene, and negative values indicate down-regulation and positive values indicate up-regulation. **i-l.** GO enrichment analysis for genes significantly influenced by CCI **(i)** from fibroblast to tumor cell, **(j)** from tumor cell to macrophage, **(k)** from tumor cell to endothelial cell, and **(l)** from tumor cell to macrophage.

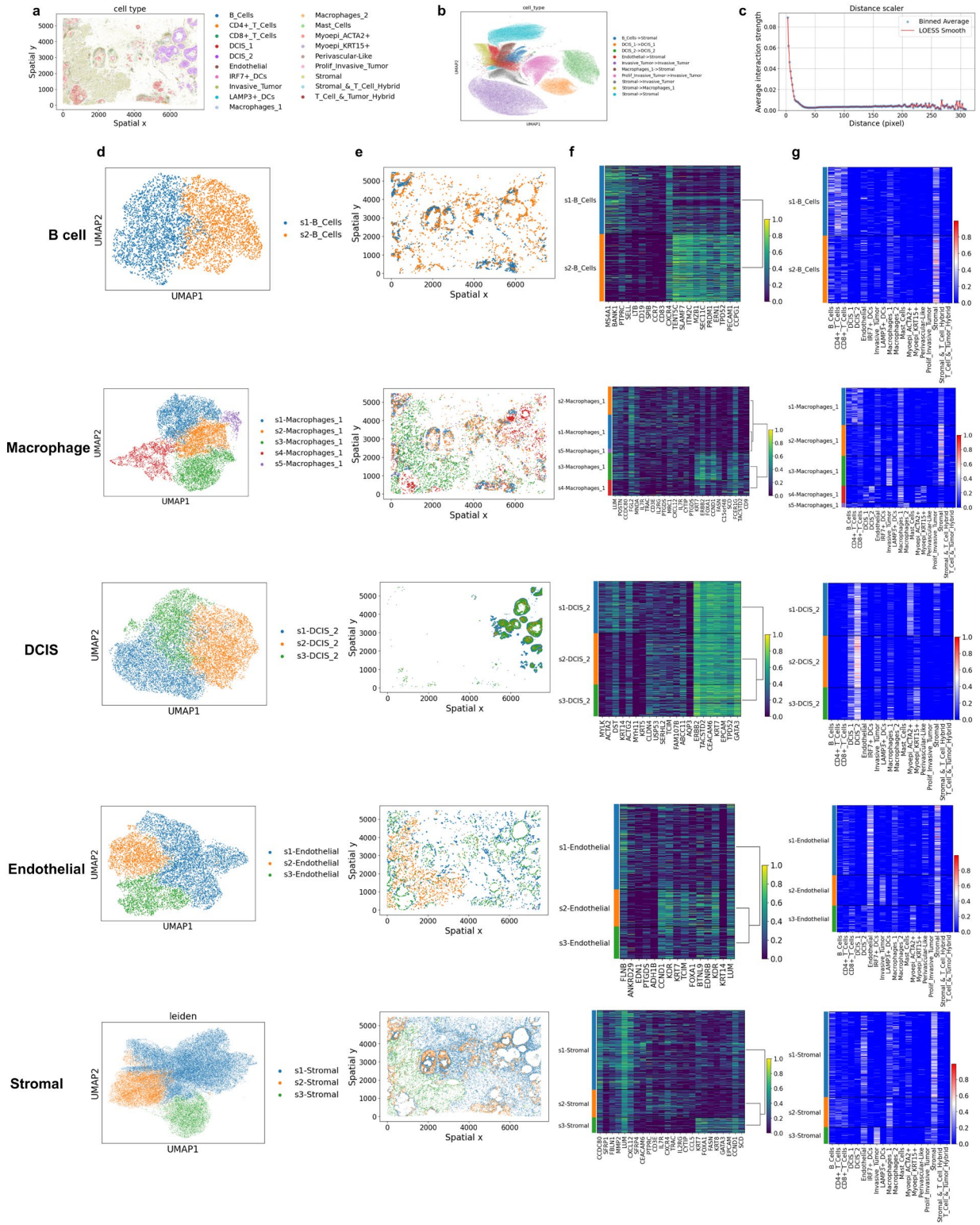


**Figure 5. Analysis of CCI-driven spatial gene expression heterogeneity in breast cancer. a.** Spatial distribution of 19 cell types within the 'Sample1_rep1' slide. **b.** UMAP visualization

demonstrating the function of the top 10 most frequently occurring CCI cell type combinations, with the features of one point indicating how much each target gene in a receiver cell is influenced by a sender cell. **c.** The relationship between the average CCI strength and the distance between two cells. **d-g.** Subtyping analysis of B cells, macrophage, ductal carcinoma in situ (DCIS), endothelial, and stromal cells by **(d)** UMAP visualization of interpretable CCI-informed features of each cell, **(e)** Spatial distribution cells in each sub-group, **(f)** Proportional influence on each cell by their corresponding neighboring cell types averaged across all genes, and **(g)** Differential expression analysis for each sub-group.

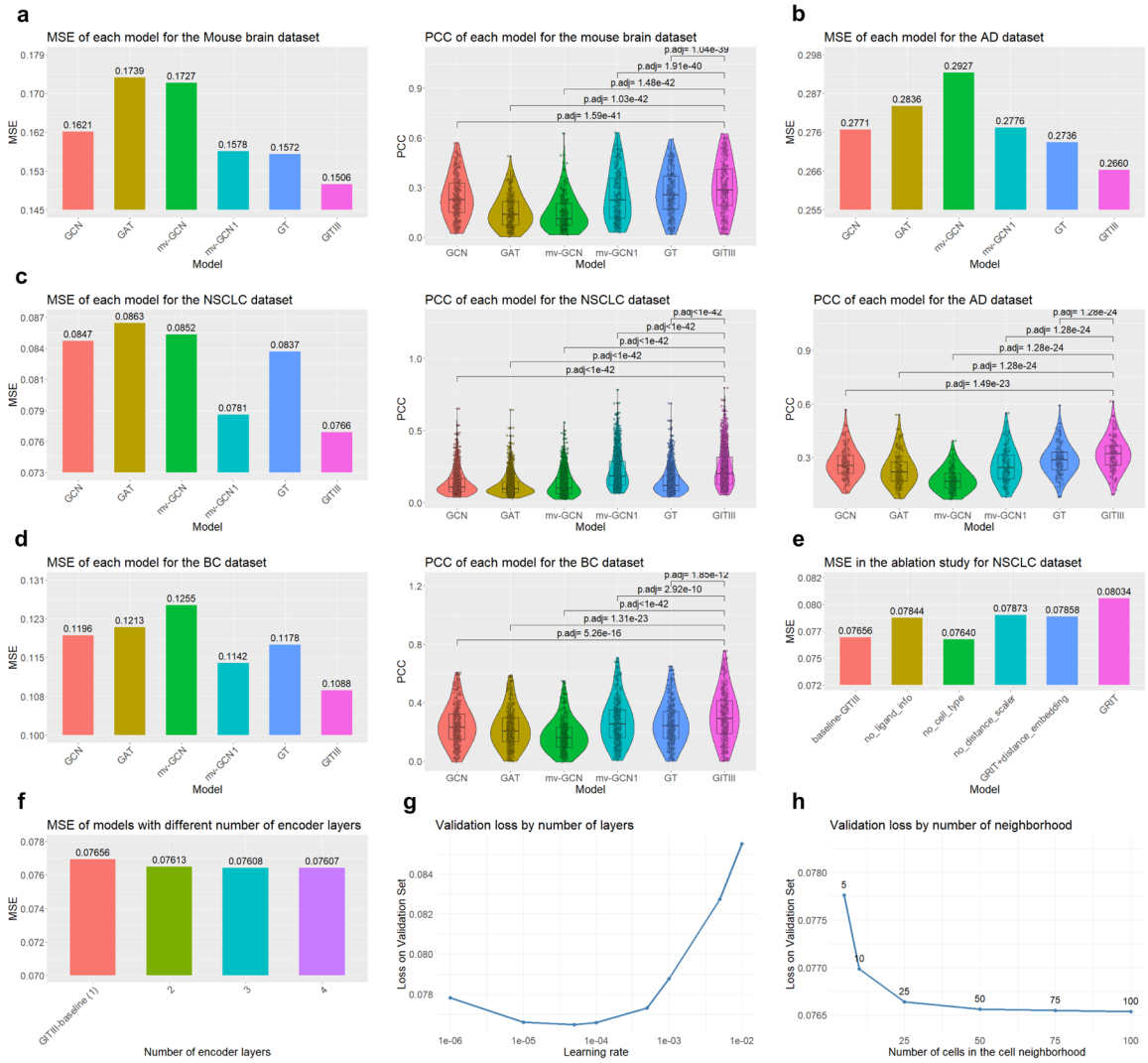


**Figure 6. Benchmarking results, ablation study, and sensitivity test. a-d.** The comparison of mean square error (MSE) and Pearson Correlation Coefficient (PCC) between the predicted cell state expression and the measured cell state expression on **a.** the Mouse brain dataset, **b.** the AD dataset, **c.** the NSCLC dataset, and **d.** the breast cancer dataset between GCN, GAT, mv-GCN, mv-GCN with MLP-based distance scaler (mv-GCN1), GT, and our model. Adjusted p-values indicate the significance of the Wilcoxon rank-sum test comparison between each model's prediction PCCs of each gene. **e.** Comparison of MSE in the ablation study. **f.** Comparison of MSE for models with

different numbers of encoder layers. **g.** Comparison of MSE for models trained by different learning rates. **h.** Comparison of MSE for models utilizing different numbers of neighbors in the cell neighborhood to predict the cell state expression of the central cell.