DISCOVERING SPATIAL DIFFERENTIAL EXPRESSION WITH GRAPH CONVOLUTIONAL NETWORKS

Roman Kouznetsov, Jackson Loper, and Jeffrey Regier

University of Michigan, Department of Statistics

Single-cell spatial transcriptomics measures gene expression for individual cells and pairs these measurements with the positions of cells within a tissue sample. With spatial transcriptomic data, how neighboring cells interact can be investigated by posing a pair of prediction problems. The first problem is to predict a gene's expression in each cell using the attributes of the cell. The second problem is to predict the same responses using the attributes of the cell *and its neighbors*. Genes may then be ranked according to difference in predictive performance: genes for which the predictions from the second problem are more accurate than the first may carry important spatial correlates. Highly ranked genes may warrant further investigation through follow-up experiments.

Mixture of Experts for Spatial Signaling genes Identification (MESSI) is one method that adopts this general approach [1]. However, MESSI is limited to using fixed-dimensional encoding of the attributes of neighboring cells as predictors. MESSI encodes neighboring information as the sum of gene expressions over all neighboring cells, with neighborhood structure determined by Delaunay triangulation. These encodings ignore the number of neighbors and the interactions among them. We propose instead to encode neighboring information as a graph, thus avoiding the limitations of fixed-dimensional encodings. Figure 1 illustrates the benefit of using graph-based encoding instead of a fixed-dimensional encoding.

Methods We transform spatial transcriptomics data into a graph in which nodes represent cells and pairs of nearby cells (within 60 μm) are connected by an edge. We augment this graph with node attributes representing each cell's type and the expression levels for certain genes (the "covariate genes"). We also augment the graph with edge attributes representing the distances between nearby cells. Then, we fit a network to predict expression of other genes (the "response genes") using the full graph. We refer to this network as DeepST. DeepST is a graph convolutional network (GCN), which is a neural network that takes a graph as input, along with node attributes and edge attributes, and outputs node-level predictions [2].

We rank genes based on the contrast between DeepST's predictions and those of a baseline prediction method that bases its predictions only on the attributes of the cell (and not its neighbors). Our baseline method is in fact a special case of DeepST, where no cells are connected by edges. In this special case, DeepST is equivalent to a feed-forward neural network that takes as input the covariate gene expression levels.

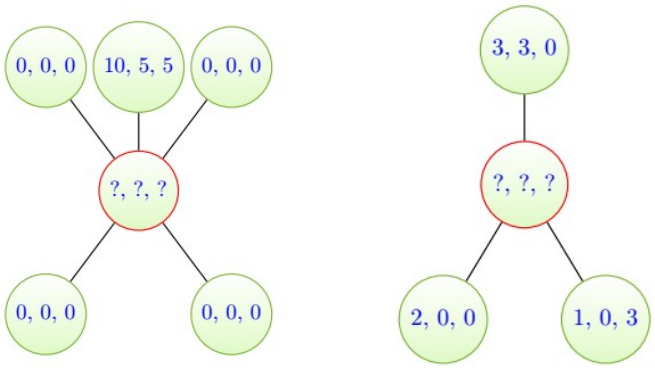

Figure 1: Two graphs, each representing the neighborhood of the cell (node) in the center. The node attributes indicate neighbors' expression levels for three genes. These neighborhoods are quite different in structure, yet both have the same element-wise mean gene expression levels: (2, 1, 1).

Results We applied our method and alternative methods to the mouse hypothalamus spatial transcriptomics dataset introduced in [3]. We used the expression levels of ligand and receptor genes as covariates; these are known to facilitate intercellular communication. We trained each method to predict the expression levels of all other genes. DeepST's predictions achieved 0.327 mean average error (MAE), outperforming a boosted regression tree prediction rule that relies on fixed-dimensional neighborhood encodings and achieved 0.367 MAE. The boosted regression trees themselves outperformed MESSI, which achieved 0.442 MAE using the same fixed-dimensional encodings. All reported MAE values are for held-out data, and hyperparameters were tuned using a separate validation set. Note that the MESSI regression algorithm was specifically designed for this problem and this dataset, whereas the boosted regression tree method is an off-the-shelf machine learning algorithm. In our experiments, we used the LightGBM [4] implementation of boosted regression trees and configured it to target the MAE loss rather than the more common mean squared error (MSE) loss. MESSI has previously been reported to outperform boosted regression trees in terms of MAE. We speculate that in [4], boosted regression trees may have been trained to target MSE loss rather than MAE, the metric used to report relative performance.

Our baseline method (i.e., DeepST without edges) achieved 0.350 MAE; unsurprisingly, this spatially ignorant baseline method had greater error than the spatially aware DeepST method. On the other hand, this MAE is actually less than the MAE found in the two alternative spatially-aware methods (MESSI and gradient boosting). This is concerning given that the purpose of estimating predictive models is to rank genes based on the difference between a spatially aware model and a baseline. When spatially ignorant methods outperform spatially aware methods, it becomes difficult to draw conclusions from the contrast. Any such conclusions may say as much about the restrictive parametric form of the regressors as it does about the ways cells interact. This highlights the necessity of using the best possible predictors when ranking genes by comparing predictor accuracies.

Next, we explored DeepST's ability to reveal genes that are spatially differentially expressed. For each response gene, we compared DeepST's predictive performance with the performance of the baseline estimator. Figure 2 shows the differences. EBF3, CPNE5 and ERMN stand out, with DeepST achieving 3.2%, 3.8% and 5.0% reduction in mean absolute error, respectively. CPNE5 is already known to guide the spatial organization of the murine brain [5]. EBF3 and ERMN may be genes to investigate further in validation experiments.

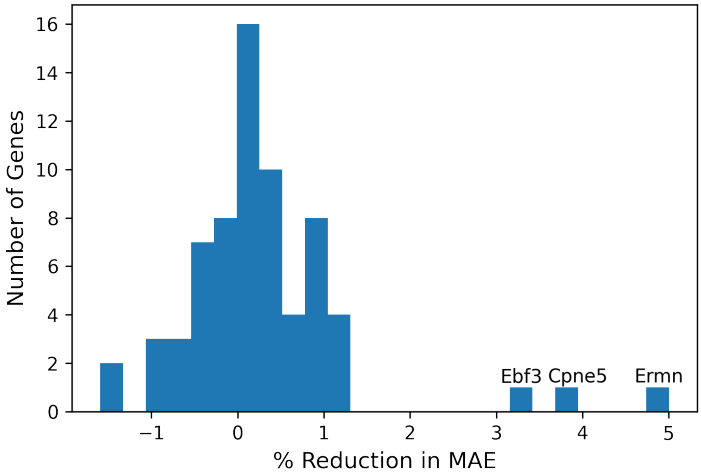

Figure 2: A histogram showing how mean absolute error (MAE) changes for each of 84 response genes when neighborhood information is included in the predictors. The MAE changes by less than 1.0% for most genes, but for EBF3, CPNE5, and ERMN, the MAE is reduced by 3.2%, 3.8%, and 5.0%, respectively.

Conclusion Spatial transcriptomics provides a new lens for studying cell-cell communication, but there are many potential analysis pitfalls. Posing paired prediction problems offers one way to interpret spatial transcriptomics data, but this approach can fail when prediction algorithms are insufficiently flexible, either due to a reliance on fixed-dimensional neighborhood encodings or due to the limited expressivity of the prediction rule that maps neighborhood encodings to predictions. To obtain sufficiently flexible prediction model, we developed DeepST, a graph convolutional network that learns on graphs defined from spatial transcriptomics datasets. The contrast between DeepST's predictions and the predictions from a baseline regressor lacking access to cell neighborhood information provides insights into how cells interact.

References

[1] Dongshunyi Li, Jun Ding, and Ziv Bar-Joseph. Identifying signaling genes in spatial single cell expression data. *Bioinformatics*, 37(7):968–975, 2021.

[2] Si Zhang, Hanghang Tong, Jiejun Xu, and Ross Maciejewski. Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1):1–23, 2019.

[3] Jeffrey R Moffitt, Dhananjay Bambah-Mukku, Stephen W Eichhorn, Eric Vaughn, Karthik Shekhar, Julio D Perez, Nimrod D Rubinstein, Junjie Hao, Aviv Regev, Catherine Dulac, et al. Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science*, 362(6416):eaau5324, 2018.

[4] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 2017.

[5] Xuefeng Ding, Yanbing Jin, Yan Wu, Yanrui Wu, Haitao Wu, Lei Xiong, Xiaoguo Song, Shuhong Liu, Wenhong Fan, and Ming Fan. Localization and cellular distribution of CPNE5 in embryonic mouse brain. *Brain Research*, 1224:20–28, 2008.