deepST: A Graph Convolutional Autoencoder for Spatial Transcriptomics

Roman Kouznetsov, Jackson Loper, Jeffrey Regier

Abstract

Spatial transcriptomics (ST) measures gene expression for individual cells and pairs these measurements with the positions of cells within a tissue sample. This opens the door for statistical methods to explore how neighboring cells interact. The statistical structure of these interactions can be investigated by posing prediction problems. For example, we can see which subsets of genes in neighboring cells are most predictive of gene expression in target cells. We can infer conditional independence structures by comparing prediction accuracy obtained from different subsets. Existing methods pursuing this vision use fixed-dimensional summaries of the attributes of neighboring cells, ignoring the number of neighbors and the interactions among them. We here propose deepST, a denoising graph convolutional autoencoder that accounts for these subtleties. For a large MERFISH hypothalamus dataset, deepST imputes missing expression levels for response genes more accurately than other state-of-the-art methods including gradient boosting, attaining an 8.7% reduction in absolute error. We also find that gradient boosting itself outperforms existing methods in this domain such as *"Mixture of Experts for Spatial Signaling genes Identification"*, attaining a 7.2% reduction in absolute error. This error reduction is critical because we are using differences in predictive accuracy to uncover biological structure, and these differences in prediction accuracy due to biological causes are often on the order of 1%.

Introduction

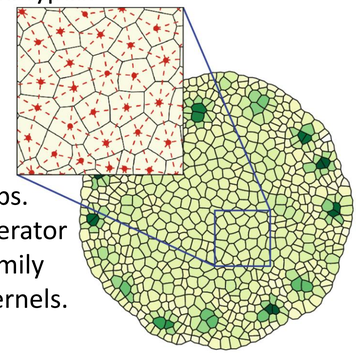
- Given that cells communicate with one another, it could be useful to use neighboring expressions as inputs for predicting gene expression.
- Recent advancements in graph convolutional networks (GCNs) have allowed for richer prediction results on graph-structured data.
- We propose deepST, a graph convolutional autoencoder (GCAE).

Methods

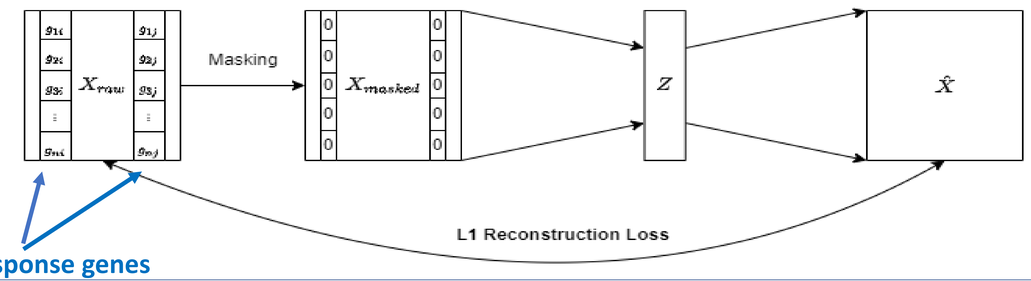
The MERFISH dataset contains expressions of the following gene types:

- **Control**: "genes" associated with blank barcodes
- **Ligand**: genes responsible for creating ligands
- **Receptor**: genes responsible for creating receptors
- **Response**: genes whose expressions we want to predict

- Each tissue sample represented a graph, with nodes representing cells and edges representing cell-cell relationships.
- For convolutional layers, we use GMMConv, a convolution operator with a parameterized patch that allows us to study a wider family of functions that can be modeled via a mixture of Gaussian kernels.



A typical tissue of cells. The dashed red lines represent edges based on cell proximity (D. Draelants et al., 2015).

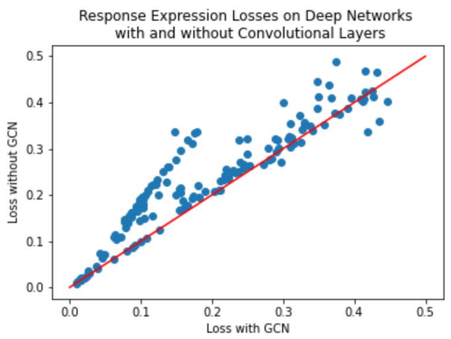


Results

- deepST better predicts response expressions than previous work.
- Use of convolution layers typically lowers loss for response expressions.
- The model utilizes a deep architecture with a kernel size of 10 for pseudo-coordinate learning.

radius	training loss	validation loss	test loss
0 μm	0.3110	0.3209	0.3367
1 μm	0.3038	0.3313	0.3363
2 μm	0.3013	0.3100	0.3339
4 μm	0.3162	0.3177	0.3423
8 μm	0.3256	0.3222	0.3338
16 μm	0.2952	0.3066	0.3409
32 μm	0.2872	0.3108	0.3350

deepST (ours)	LightGBM L_1	MESSI	XGBoost L_2
0.327	0.367	0.442	0.477



Discussion

- deepST is a model that has a significantly larger amount of learnable parameters, leading to a better predictive performance.
- deepST requires only ≈25 minutes more of training time than its gradient boosting competitors and is ≈25x faster than MESSI.
- If deepST is used to only predict the expressions of a single celltype, the training time does not change, while for competing methods, the training time triples.
- Surprisingly, increasing the radius for cell neighbor consideration does not significantly change model performance.
- Masking 50% of response gene expressions at random yields an imputation problem for which our model performs well.

Conclusion

- We demonstrate an increased performance in response gene prediction using GCAEs.
- Previous methods have been limited in neighborhood expression and learnable parameters. deepST is a state-of-the-art model that avoids these bottlenecks.
- While we demonstrate this method is effective, the fact that neighboring information does not significantly add to predictive performance is surprising.

Future Directions

Currently, we have shown that using a deterministic graph decoder-encoder structure manages to increase response gene prediction accuracy. A natural next step is to learn a graph *variational* autoencoder that learns latent distributions rather than embeddings.