Diffusion on Spatiotemporal Gene Expressions

Rohola Zandie

April 2025

1 Data

The dataset comprises gene expression data from cells on slides (with spatial information) at various time points. Each time point includes multiple slices, and each slice is represented as a tensor of shape $G \times W \times L$, where G denotes the number of genes, W is the width, and L is the length of the scanned slice. Each element in the tensor reflects the expression level of a specific gene at a given spatial location (x, y).

There are many genes in one slide (in the order of 25000). Genes are basically like RGB colors for an image in this context. An example visualization of the gene **Col3a1** is shown in Figure 1. Each gene's image is saved separately. Genes also come with expression level (basically magnitude), although the expression values lie within the range $[0, 9]$, they can be normalized to a desired range (e.g., $[-1, 1]$) for downstream image processing tasks.

The dataset presents two main challenges. First, the number of genes (i.e., channels) is quite large—approximately 25,000—but remains consistent across all datasets. Second, the gene expression tensors are highly sparse, with most positions and time points exhibiting zero expression values.

2 Goal

The primary objective of this project is to develop a conditional diffusion model capable of generating realistic gene expression patterns conditioned on both the gene identity and time point. This involves selecting an appropriate model architecture—particularly one that can effectively handle the high dimensionality and sparsity of the dataset—and training it on the large-scale dataset with efficiency and getting good accuracy of reconstruction (you can think about it as video generation basically, but we want to conditions the generation based on genes)

We have been thinking about different modeling, and it seems like treating each gene expression image (where the conditions are actually the gene names) at each time step is the best way to do this, but we are open to other modeling. With this modeling, after training, we can also get the embedding of each gene (condition) and also analyze those which is interesting for us.

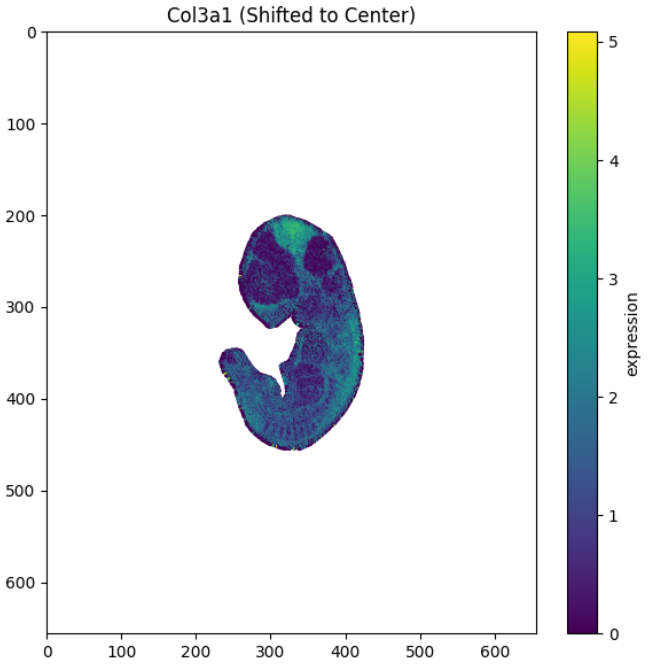

Figure 1: The expression values of the Gene **Col3a1** on one slice and one time point