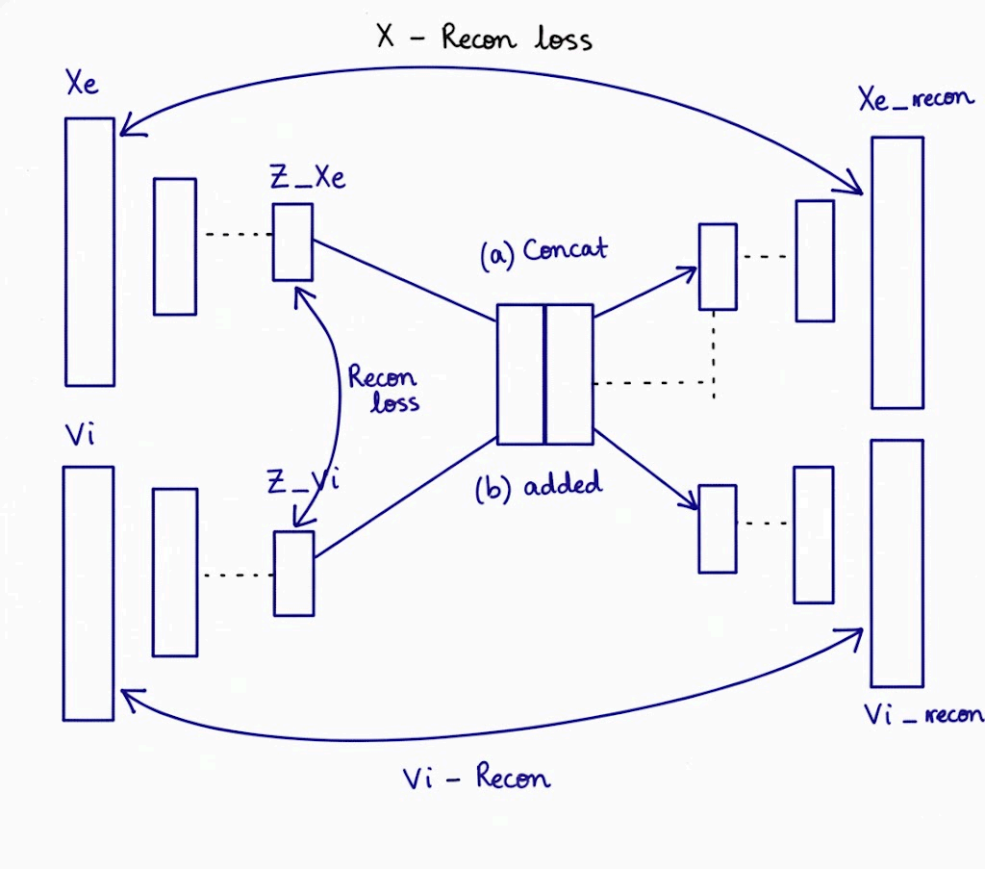
In spatial transcriptomics, different technologies capture complementary aspects of gene expression. For example, Xenium provides high-resolution, single-cell measurements but is limited to a subset of genes, while Visium offers genome-wide coverage but at lower spatial resolution, typically spanning multiple cells.

In our UGP project, we were tasked with the problem of creating a joint representation of xenium and visium data to solve two tasks: imputation and deconvolution.

Approaches:

1. In our initial approach, the architecture was like this:

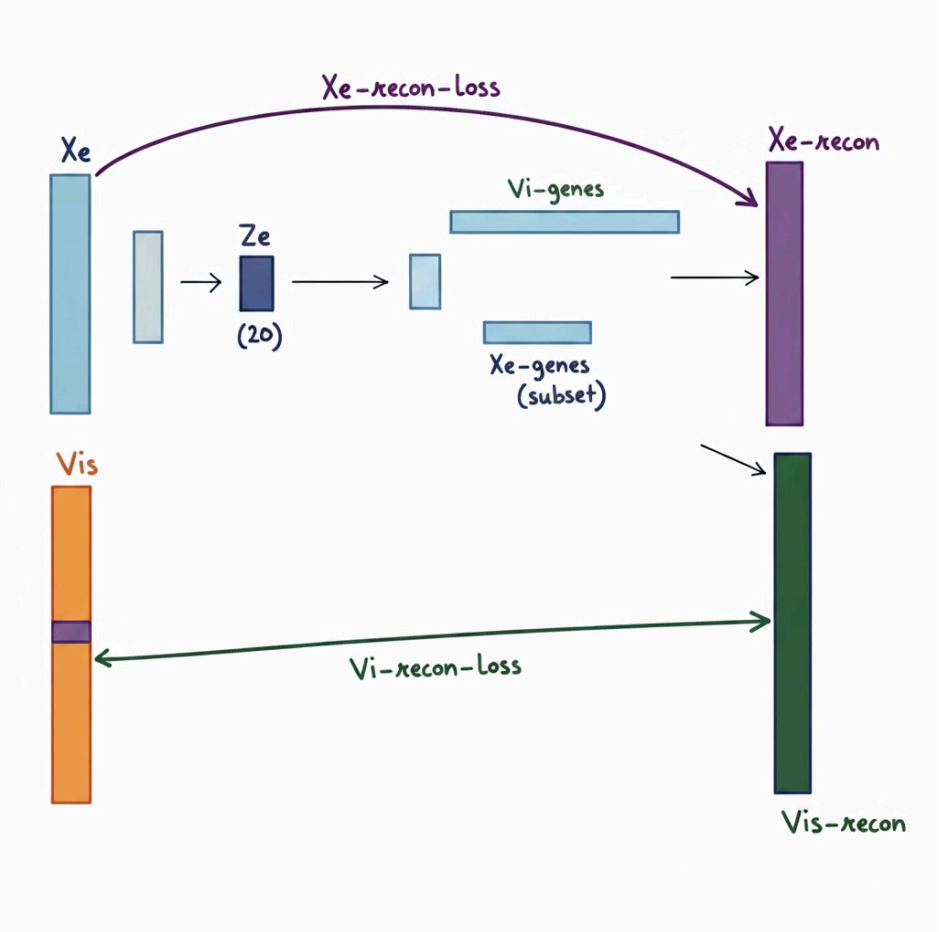


While this gave very excellent reconstruction results, a R^2 of more than 0.9 on both xenium and visium, this failed to solve the imputation problem as the model failed to infer xenium from visium only.

2. To solve that, we masked a large amount of xenium and just gave input as visium, while giving a huge improvement from our original model , this gave us inspiration for the next move.

3. We just gave xenium , encoded into a latent embedding (of size = no. of cell types), applied non negativity constraints on it, then used a decoder to decode both visium

and xenium from it. Visium was only used to do loss calculations (loss = loss_vi +loss_xe + a regulariser that promoted a cell to be in a particular cell type only)



This gave performance on par with the SIID paper on the initial runs, in the future we are planning to optimize it much further through more hyperparameter optimizations.