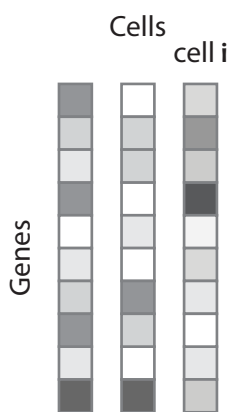


Single-cell Atlas



Tokenizer

Transcriptomic Foundation Model

Transformer Block
(layer I)

Activations

x_i^l

x N layers

Input Setup

Activations, x_i^l

Sparse Autoencoders

Reconstruction

x_{ij}^l

$f(x_{ij}^l)$

x_{ij}^r

Linear Probes

Directions

W

Classification

$\text{softmax}(Wx_{ij}^l)$

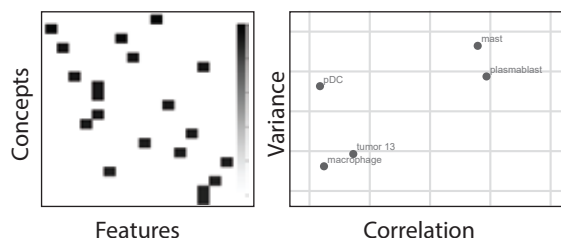
Gene Ontology Labels

g:Profiler



Decomposition

Interpretation of Model Knowledge via SAE

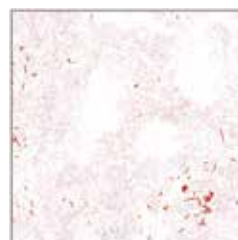


Microenvironment Discovery with TIS



Interpretability analyses

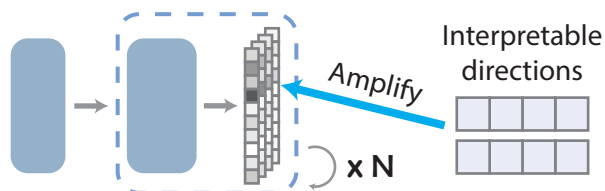
Interpretable Cell-cell Communication



Crosstalk

- cross-compartment
- within-compartment

Steering-guided Perturbation Analysis



Application