THE UNIVERSITY OF CHICAGO | DEPARTMENT OF STATISTICS

Master's Thesis Presentation

Yuhan Wang
Department of Statistics
The University of Chicago

"Ablation and Architectural Extensions of STAMP for Spatial Transcriptomics"

May 7, 2026, at 1:00 PM
Jones 111, 5747 S. Ellis Avenue

Abstract

Spatial Transcriptomics Analysis with Topic Modeling to Uncover Spatial Patterns (STAMP) is a spatially aware topic modeling framework that uses gene expression counts and spatial neighborhood information to infer interpretable spatial topics and topic-associated gene patterns. Although the original model incorporates adjacency-based propagation in the inference network and a background component in the topic-gene parameterization, the empirical importance of these design choices remains unclear.

This thesis studies these components through replication, ablation, and architectural extension of STAMP. This thesis compares eight encoder architectures, including the original adjacency-based encoder, non-spatial ablations, residual graph-feature variants, and multiscale filter-bank encoders. Experiments are conducted on two spatial transcriptomics datasets, SMI Lung and Visium Mouse Brain, across multiple topic numbers, random seeds, and spatial split strategies. Architectures are evaluated using topic consistency, downstream gene-level signatures, marker-gene recovery, and biological annotation.

The results do not support a single universal ranking of encoder architectures. In SMI Lung, most architectures show broadly similar topic consistency. In Visium Mouse Brain, the background-free non-spatial ablation shows the clearest loss of consistency, especially at larger topic numbers. GoM DE, marker-gene, and annotation analyses further show that architectures generally recover comparable broad biological structures, while differing in topic granularity, spatial smoothness, and allocation of topic components across biological labels.

Overall, adjacency-based propagation and multiscale graph features do not provide a uniform advantage across datasets and evaluation criteria. Instead, stable and interpretable topic recovery depends on the interaction between dataset structure, background modeling, topic number, and encoder construction.

For information about building access for persons with disabilities, please contact Keisha Prowoznik at 773.702-0541 or send an email to kprowoznik@uchicago.edu. If you wish to subscribe to our email list, please visit the following website: https://lists.uchicago.edu/web/info/statseminars.