SPELL: Spatial Prompting with Chain-of-Thought for Zero-Shot Learning in Spatial Transcriptomics

Sumeer A. Khan[1], Xabier M. de Morentin[1], Abdel Rahman Alsabbagh[1], Mahmoud Zahran[1], Vincenzo Lagani[1], Robert Lehmann[1], Narsis A. Kiani[2], David Gomez-Cabrero[1,3], Jesper Tegnér[1,2]*

[1]King Abdullah University of Science & Technology (KAUST), Saudi Arabia; [2]Karolinska Institute, Sweden; [3]Navarrabiomed, Spain; *Corresponding author

Introduction

- **Zero-Shot Learning (ZSL) for Cell-Type Classification**
  - Limited exploration of ZSL in spatially resolved transcriptomics, particularly in integrating spatial context with marker gene semantics for accurate cell-type classification
- **Spatial Prompt-Enhanced Zero-Shot Learning**
  - Combining graph autoencoder (GAE)-derived spatial embeddings with chain-of-thought (CoT) prompting to create human-interpretable classification prompts for zero-shot learning
- **Overcoming Limitations of Traditional ZSL Approaches**
  - SPELL surpasses conventional ZSL methods by explicitly encoding local cellular neighborhoods via spatial k-nearest neighbor graphs and generating interpretable CoT prompts that integrate marker gene expression and spatial embedding norms

Data

- **MERFISH, Stereo-Seq, and MIBI-TOF Datasets for SPELL Evaluation**
  - **MERFISH Dataset**
    - Comprises 12 consecutive slices from the mouse hypothalamic preoptic region, (Moffitt et al., 2018; curated in Palla et al., 2022)
  - **Stereo-seq Drosophila Dataset**
    - Provides a 3D high- resolution gene expression map capturing spatial and temporal dynamics during Drosophila embryonic and larval development (Qiu et al., 2024)
  - **MIBI-TOF Dataset**
    - Offers single-cell metabolic profiles, phenotypes, and spatial organization of CD8+ T cells and colorectal carcinoma, enabling analysis of cellular interactions (Hartmann et al., 2021; curated in Palla et al., 2022)

Conclusion

- **Effective Spatial Prompting with SPELL**
  - SPELL integrates spatial and molecular data into CoT prompts, leveraging LLMs for zero-shot cell-type classification, achieving strong accuracy and interpretability across MERFISH, MIBI-TOF, and Stereo-seq datasets
- **Future Directions**
  - Enhancing SPELL with additional spatial modalities, optimized marker selection, advanced spatial metrics, and validation across diverse tissues and platforms will further improve its robustness and applicability

References

- J. R. Moffitt, et.al., *Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region.* Science, 362(6416):eaau5324, 2018
- Felix J Hartmann et al. *Single-cell metabolic profiling of human cytotoxic t cells.* Nature biotechnology, 39(2):186–197, 2021

جامعة الملك عبدالله
للعلوم والتقنية

King Abdullah University of
Science and Technology

Karolinska Institutet

NAVARRABIOMED
CENTRO DE INVESTIGACIÓN BIOMÉDICA

Methods

- **Integrated Workflow for Zero-Shot Classification in Spatial Transcriptomics**
  - SPELL comprises three primary components: (i) extraction of marker genes, (ii) construction and embedding of a spatial graph using a Graph AutoEncoder (GAE), and (iii) generation of chain-of-thought prompts for zero-shot classification
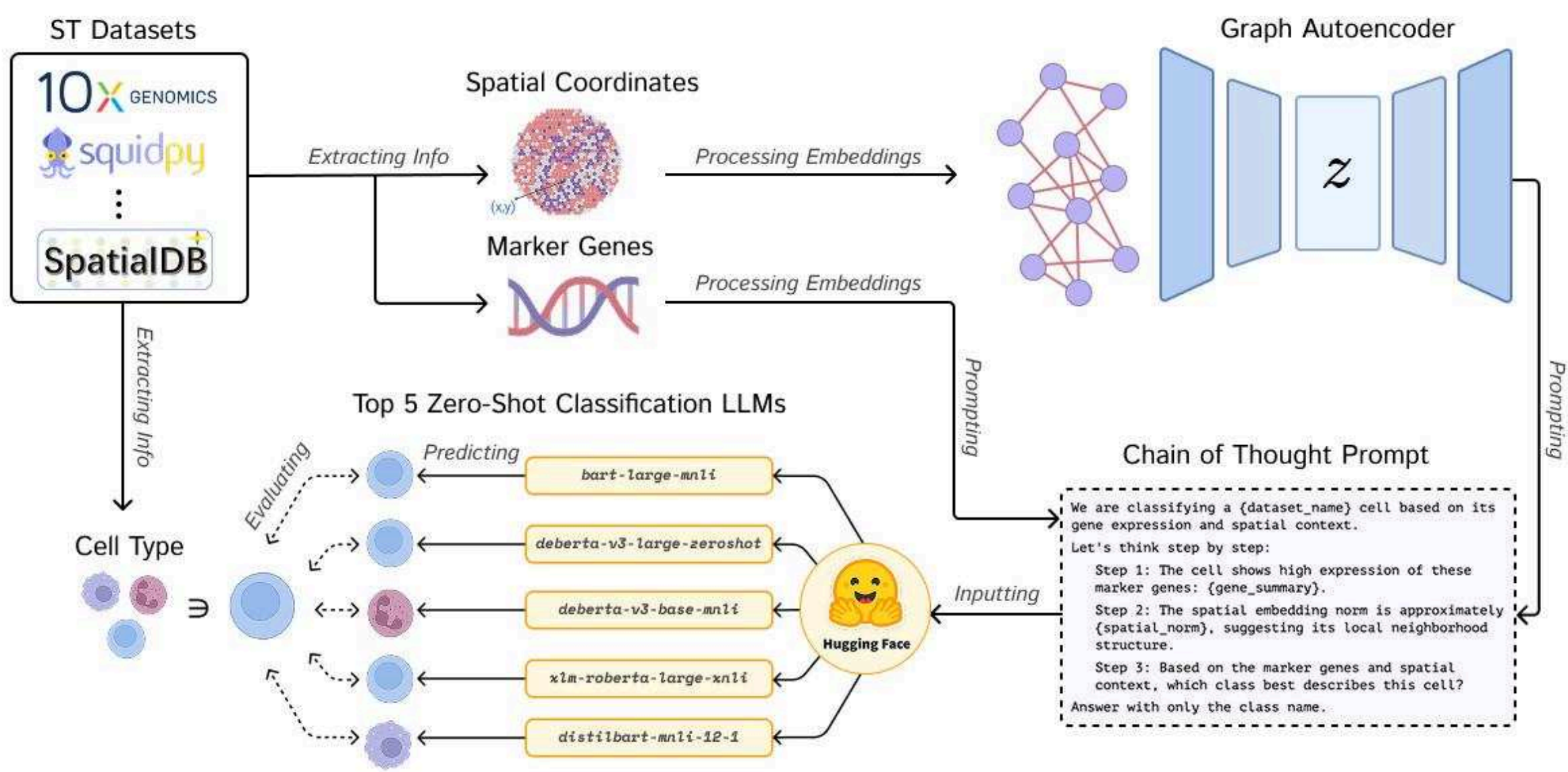

Results

- **High Accuracy on MERFISH and MIBI-TOF**
  - The SPELL framework evaluated five models, with distilbart-mnli-12-1i and bart-large-mnli achieving top accuracies on MERFISH (64%) and MIBI-TOF (52%) datasets, leveraging marker gene data and spatial context via Chain-of-Thought (CoT) prompting
- **Challenges with Stereo-seq Drosophila**
  - Both BART models showed reduced performance on the Stereo-seq dataset due to platform differences and added temporal complexity, highlighting challenges in generalizing zero-shot classification to non-mammalian systems

Model	MERFISH		MIBI-TOF		Stereo-Seq	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
distilbart-mnli-12-1	**0.640**	**0.558**	0.390	0.269	**0.210**	**0.080**
bart-large-mnli	0.010	0.002	**0.520**	**0.429**	0.120	0.026
deberta-v3-base-mnli-fever-anli-ling-wanli-binary	0.150	0.039	0.510	0.362	0.080	0.012
deberta-v3-large-zeroshot-v1.1-all-33	0.040	0.003	0.350	0.253	0.020	0.001
xlm-roberta-large-xnli	0.160	0.075	0.030	0.013	0.070	0.009
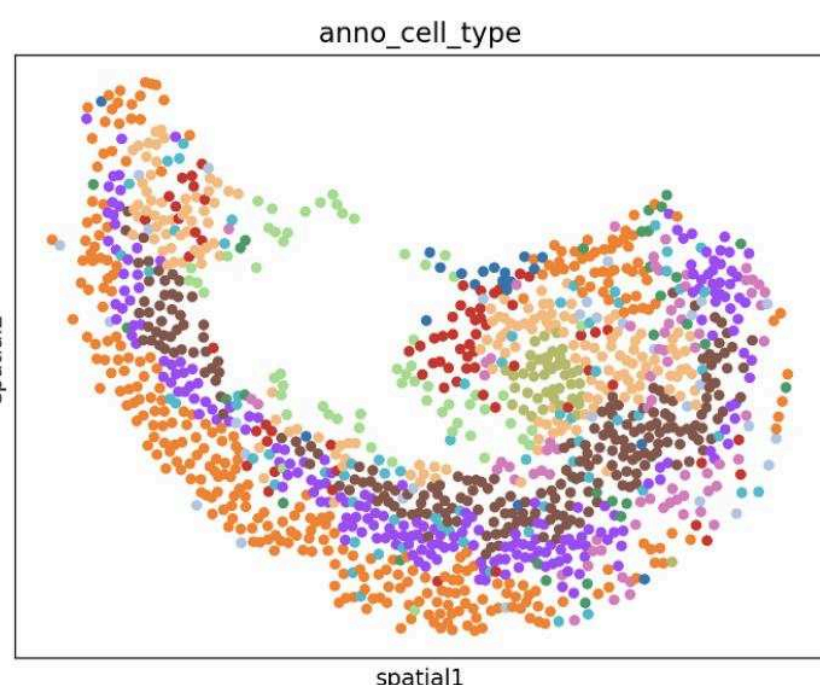
Drosophila Stereo-seq Slice
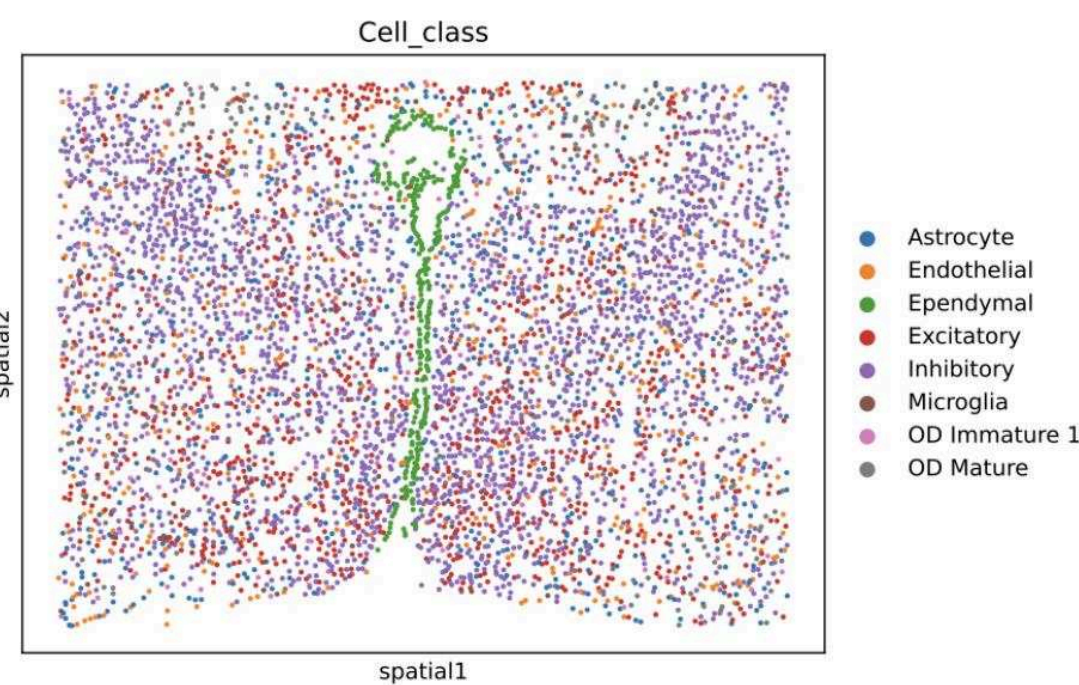
MERFISH Slice

- **Ablation Study**
  - **Spatial Context Removal**
    - Omitting spatial embedding norm from CoT prompt, using only marker gene summary, caused performance drops across models
  - **MERFISH Performance Decline**
    - Distilbart-mili-12-1i accuracy fell to 49% (F1-score: 0.338) and bart-large-mnli to 4% (F1-score: 0.003), compared to 64% and 52% with full model

Model Identifier	Accuracy	F1-Score
distilbart-mnli-12-1	0.490	0.338
xlm-roberta-large-xnli	0.150	0.039
deberta-v3-large-zeroshot-v1.1-all	0.020	0.003
deberta-v3-base-mnli-fever-anli-ling-wanli-binary	0.130	0.037
Bart-large-mnli	0.040	0.003