

Nathaly Dongo Mendoza – Computational Modeling and Simulation TU Dresden, Kather Lab EKFZ

A picture is worth a thousand genes

Predicting Spatial Gene Expression from Liver Histology with Scalable Interpolant Transformers

BACKGROUND

The Hidden Molecular Layer

Hospital routine H&E tissue slides reveal **structure**, but the underlying **gene activity** driving disease remains invisible.

Spatial Transcriptomics

Maps gene expression onto tissue spatial coordinates, but costs €500-10000 per slide, limiting clinical translation.

RESEARCH QUESTION

Can we recover this molecular information directly from tissue images that hospitals already produce every day?

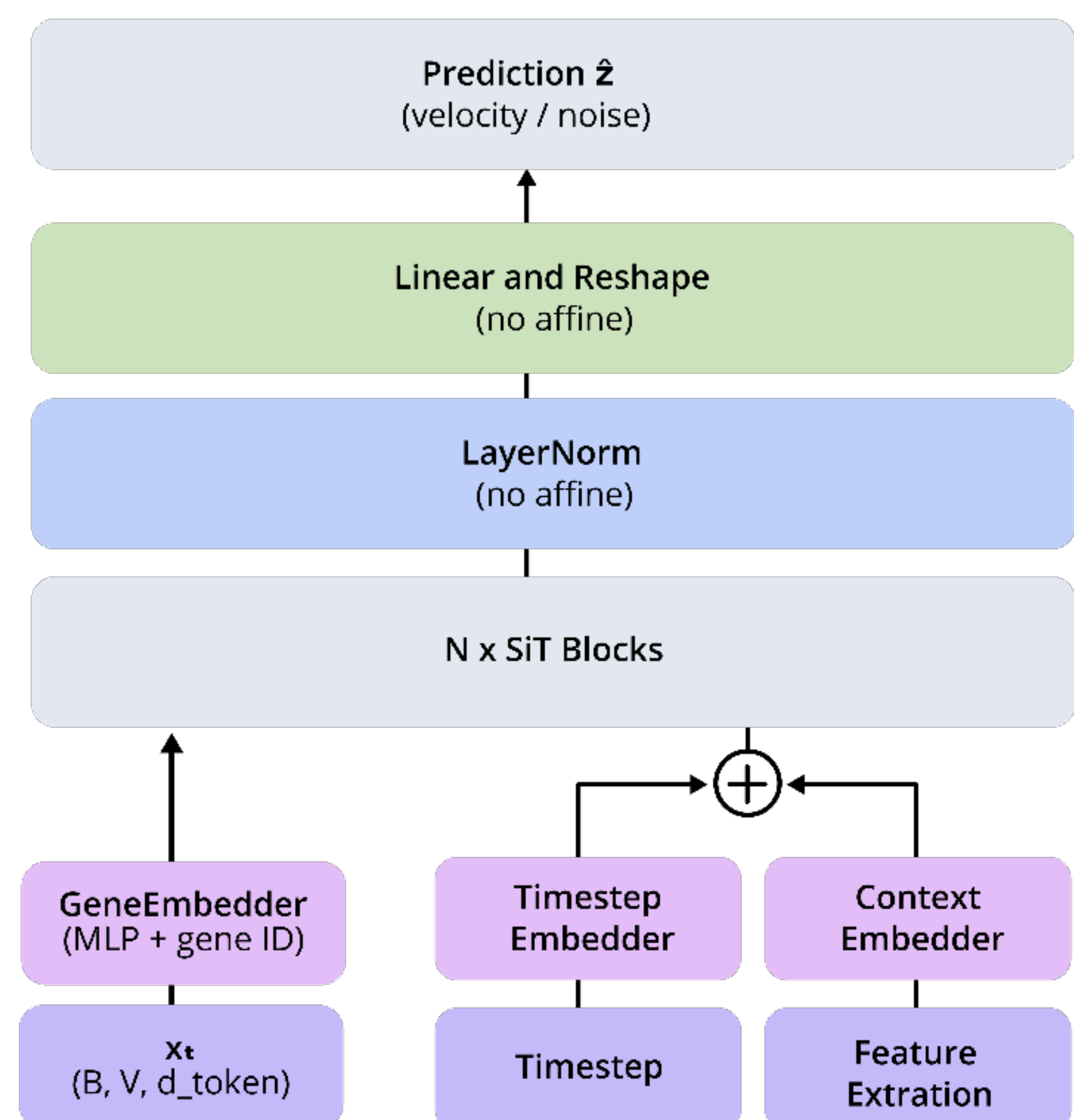
METHODOLOGY

Datasets & Feature Extraction

99 liver slides from HEST-1K + ST-Image

Modeling SiT

Interpolant transformers x Sampler x Target

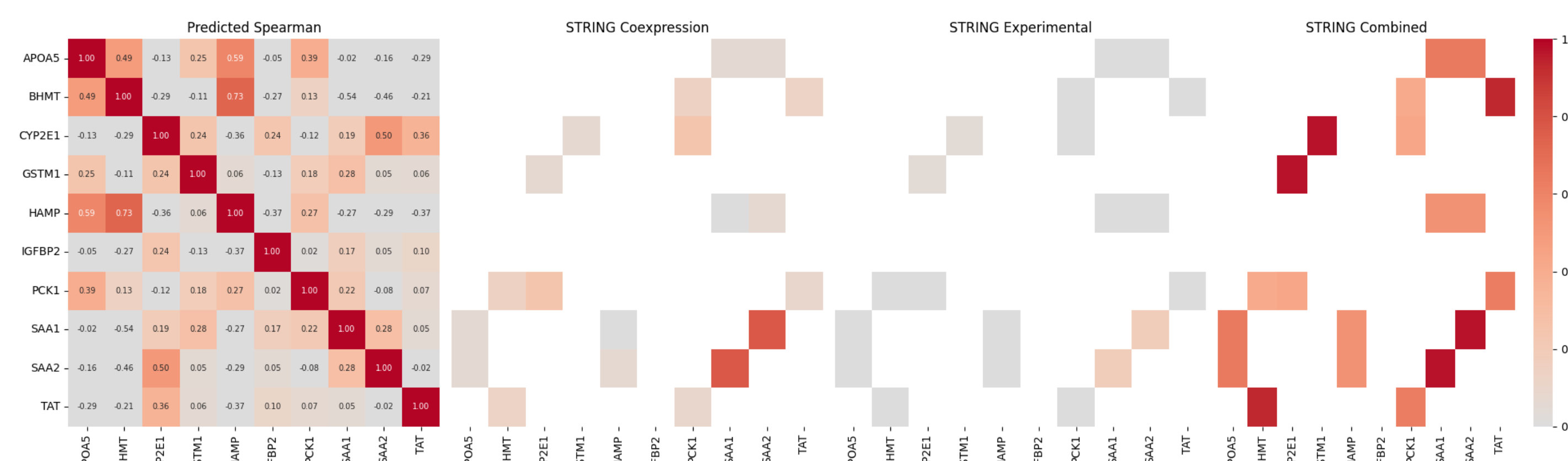


Input: H&E stained whole-slide image and the spatial transcriptomics model.

INTERPRETABILITY

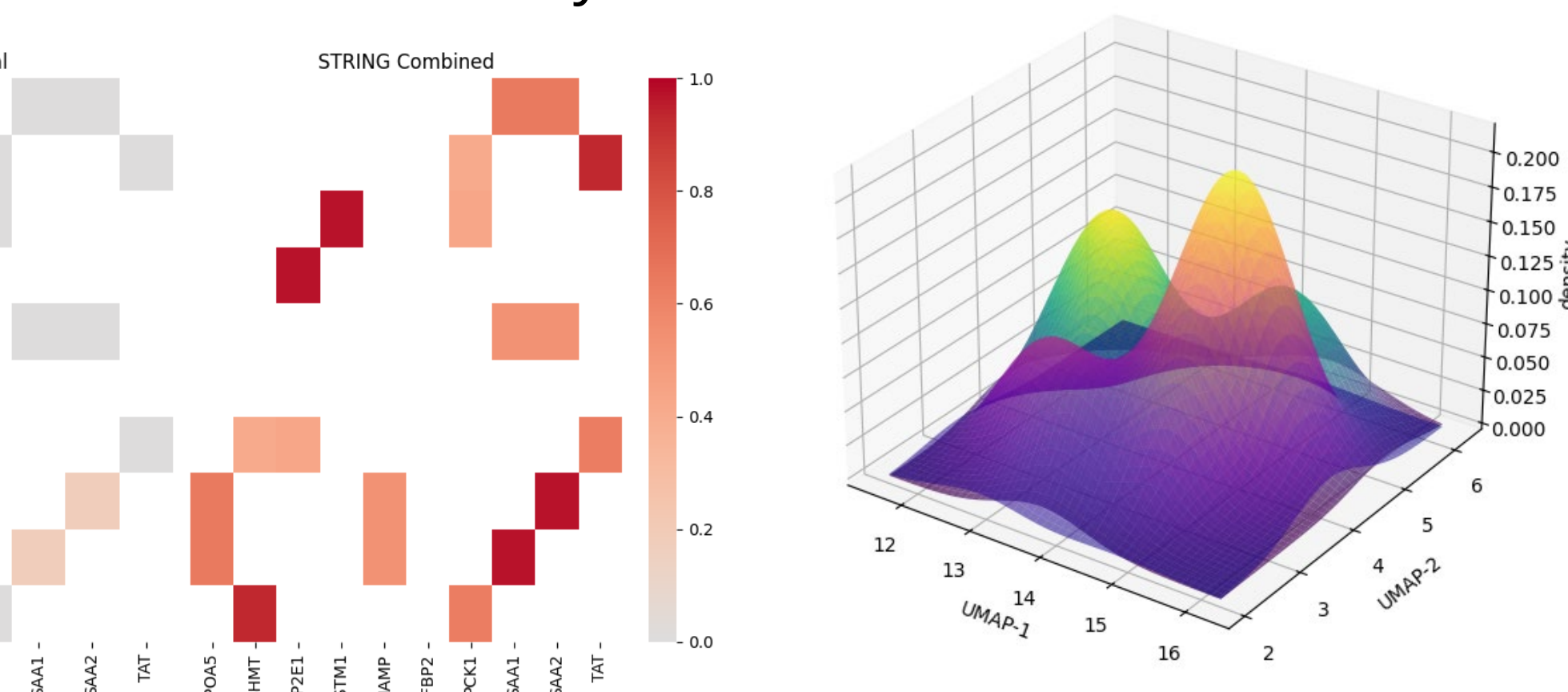
Biologically Meaningful Gene-Gene Relationships

Spearman correlation matrix of predicted expressions partially aligns with STRING database interactions.



Generative Trajectories from Healthy to Cancer

Embedding interpolation (α : 0→1) shifts latent context from healthy to cancer slides.



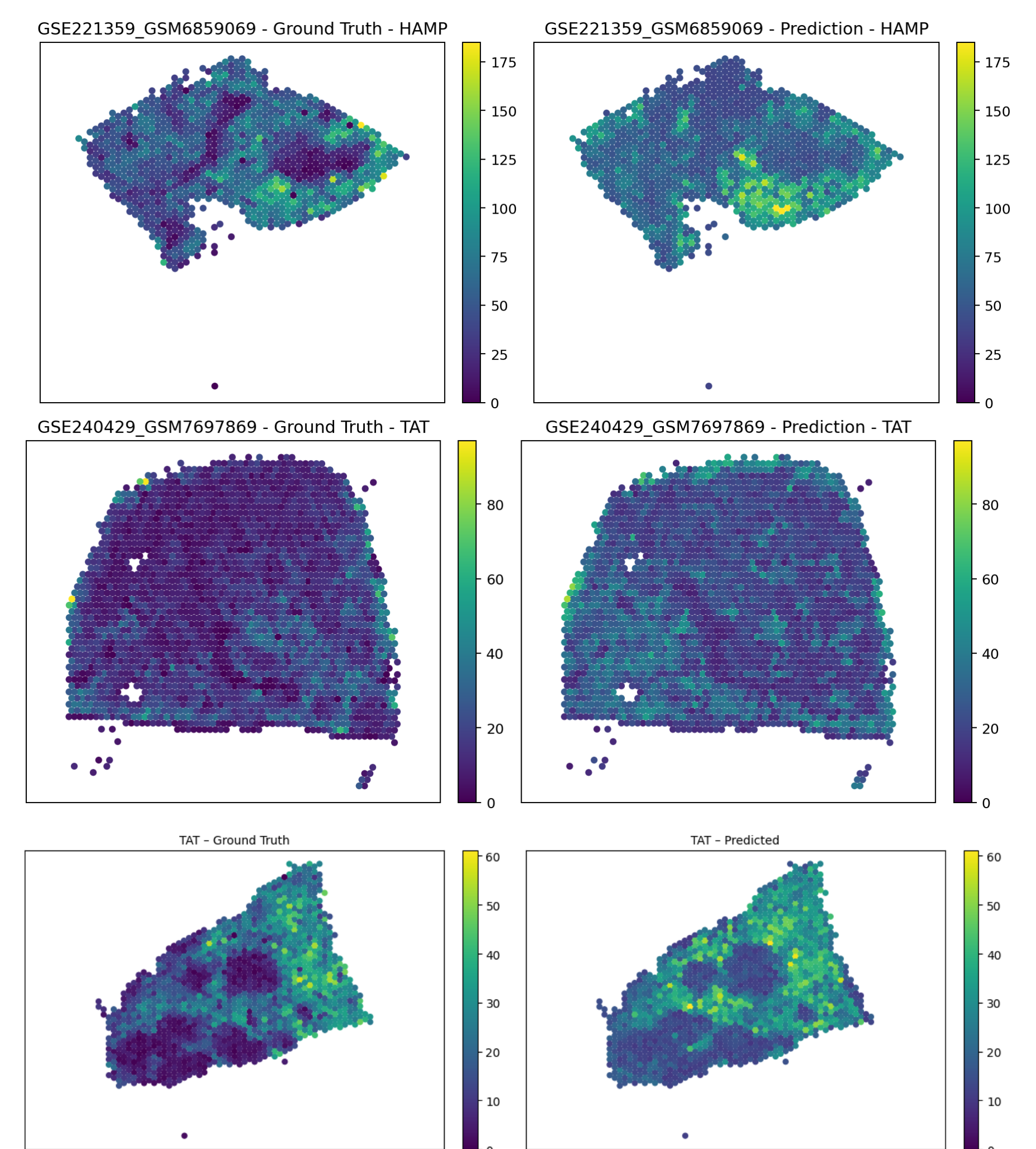
References

N. Ma, M. Goldstein, M. S. Albergo, N. M. Boffi, E. Vanden-Eijnden, and S. Xie, "Sit: Exploring flow and diffusion-based generative models with scalable interpolant transformers," 2024. [Online]. Available: <https://arxiv.org/abs/2401.08740>
G. Jaume, P. Doucet, A. H. Song, M. Y. Lu, C. Almagro-Pérez, S. J. Wagner, A. J. Vaidya, R. J. Chen, D. F. K. Williamson, A. Kim, and F. Mahmood, "Hest-1k: A dataset for spatial transcriptomics and histology image analysis," arXiv preprint arXiv:2406.06393, 2024. [Online]. Available: <https://arxiv.org/abs/2406.06393>
J. N. Kather, L. R. Heij, H. I. Grabsch, C. Loeffler, A. Echle, H. S. Muti, J. Krause, J. M. Niehues, K. Sommer, P. Bankhead, L. F. Kooreman, J. Schulte, C.-A. Weis, D. Hirsch, K. Aldape, I. Zlobec, V. H. Koelzer, P. Boor, J. Galon, R. Nawroth, B. Markl, P. Schirmacher, H. Brenner, J. Chang-Claude, M. Hoffmeister, C. Trautwein, A. T. Pearson, and N. Halama, "Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer," Nature Medicine, vol. 25, pp. 1054–1056, 2019.

RESULTS

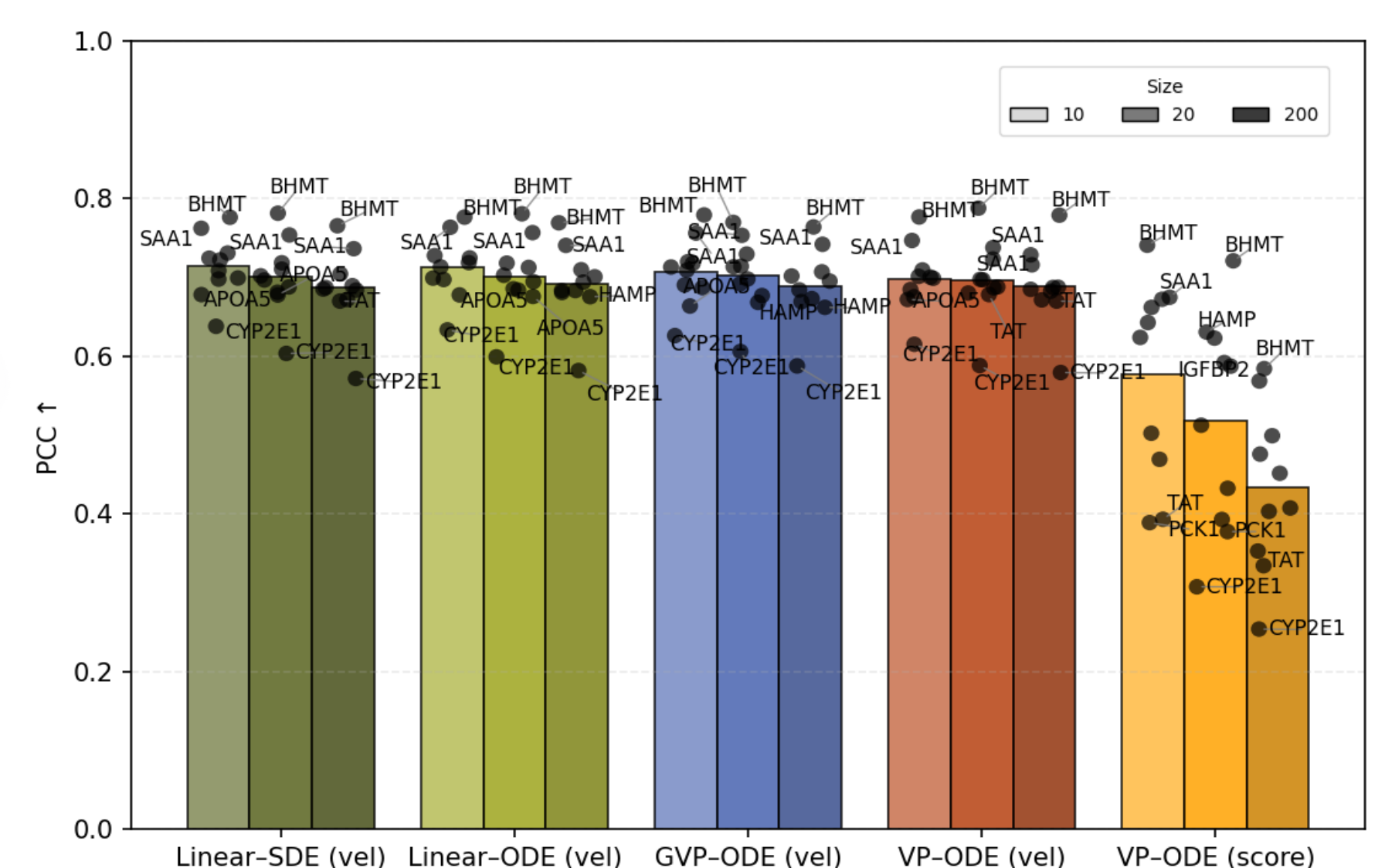
Predictions Recover Tissue-Scale Spatial Structure

HVG markers show broad spatial support with smooth gradients. Model predictions closely track ground-truth across 222-gene panels.



Performance: Velocity Targets Outperform Score Consistently

All velocity-based formulations produced consistent results. PCC values typically ranged from 0.65 to 0.78. SDE sampler had the highest pearson correlation.



CONCLUSION

- Model recovers biologically meaningful gene-gene associations and cancer-directed expression shifts.
- H&E slides carry sufficient phenotypic information to approximate spatial transcriptomics at scale.
- Velocity-based SiT (Linear, GVP, VP) consistently outperforms score-based approaches (PCC 0.65–0.78 vs. <0.50).
- Linear-SDE achieves 1.6× speedup (189s vs. 300s) without accuracy loss.