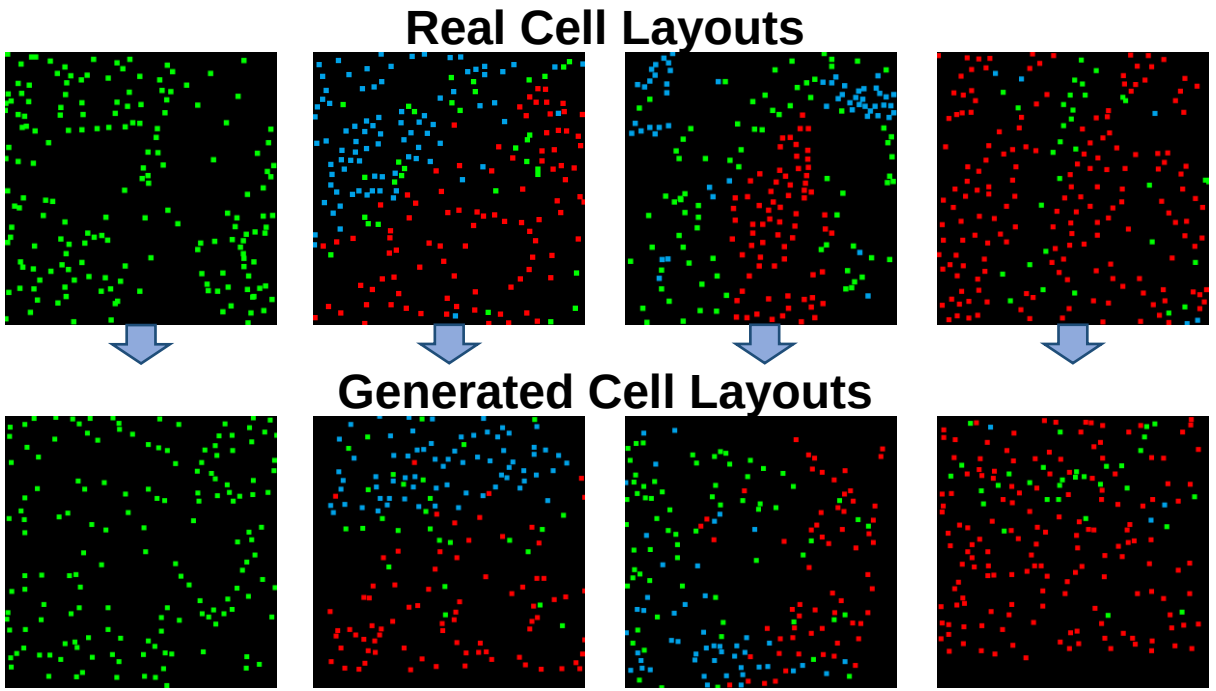
Topology-Guided Multi-Class Cell Context Generation for Digital Pathology

Shahira Abousamra, Rajarsi Gupta, Tahsin Kurc, Dimitris Samaras,
Joel Saltz and Chao Chen

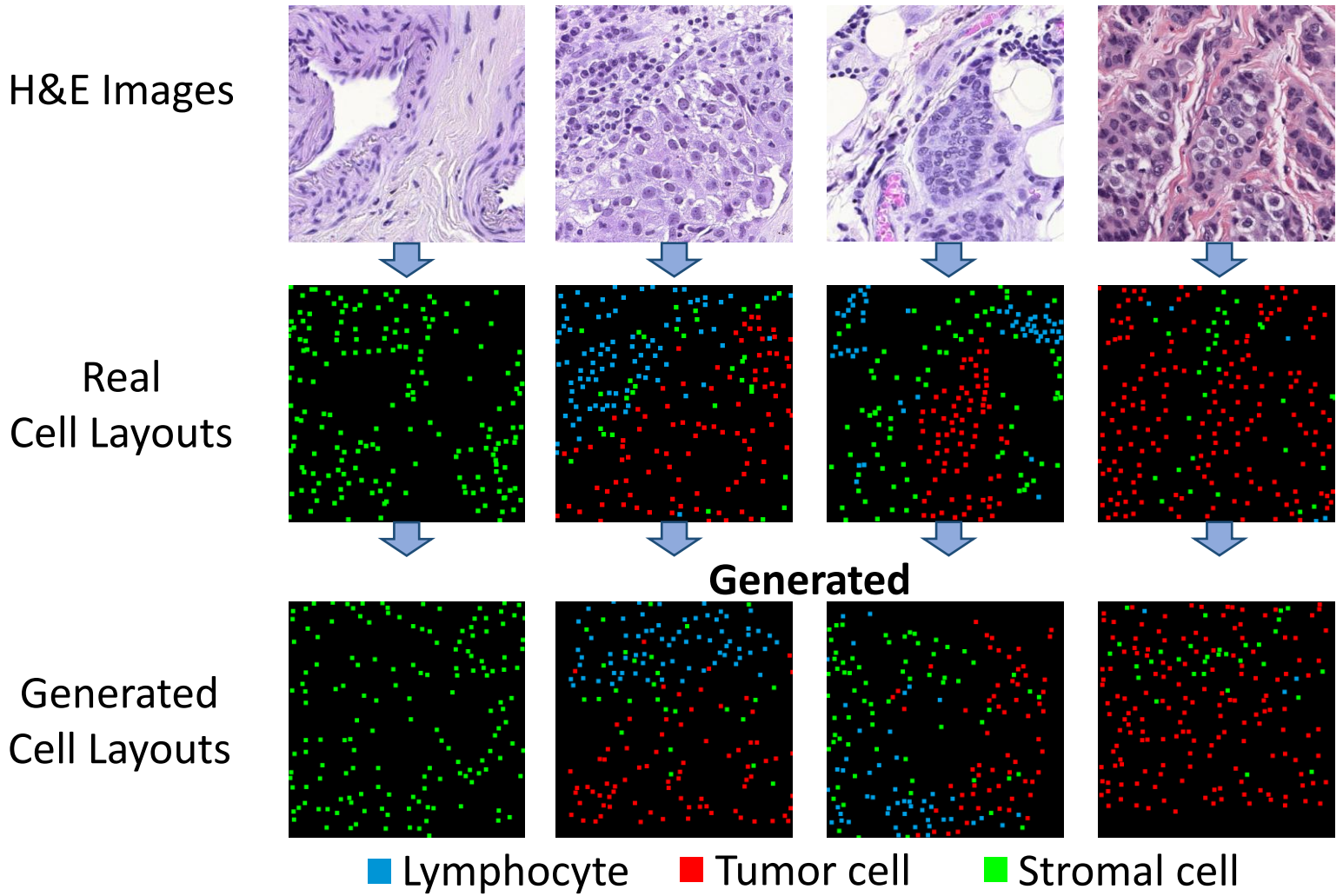
TUE-AM-316



Stony Brook University

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

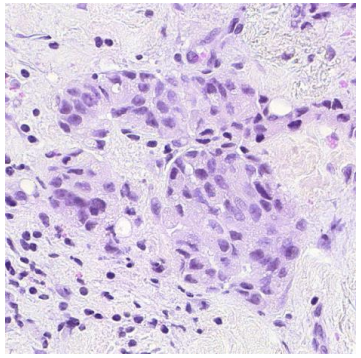
Topology-Guided Multi-Class Cell Context Generation for Digital Pathology



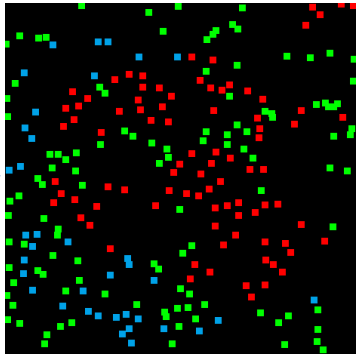
- Synthetic Data
- Data Augmentation
- Discriminative Distribution of Cells
- Realistic Cell Layouts
- Conditioned on Spatial and Topological Patterns

Topology-Guided Multi-Class Cell Context Generation for Digital Pathology

Real Image

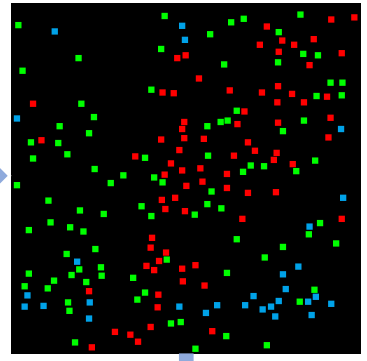


Real Cell Layout

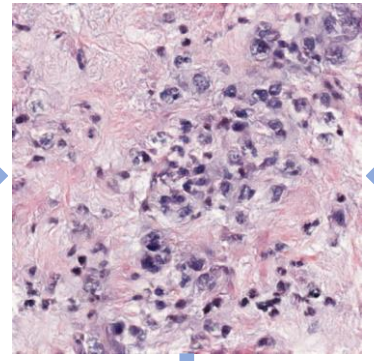


Lym. Tumor Stromal

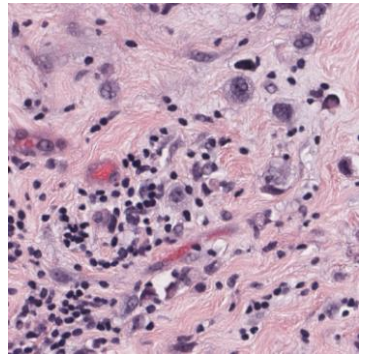
Fake Cell Layout



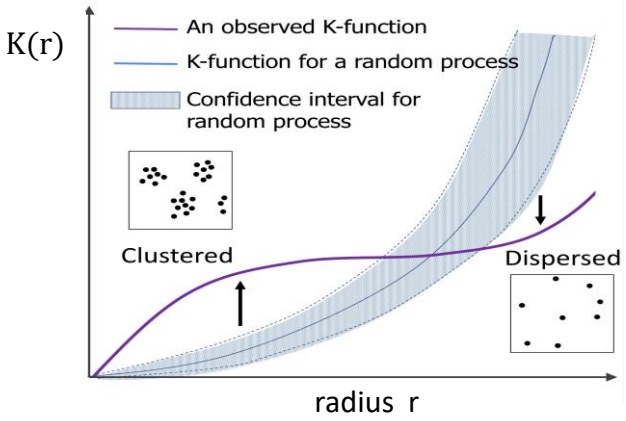
Fake Image



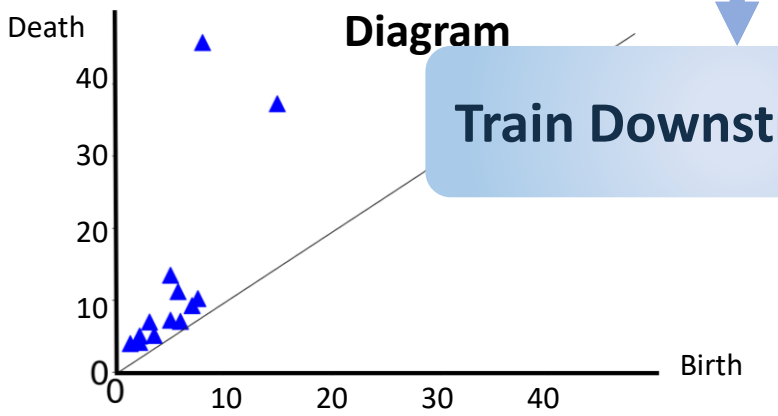
Real Texture



Ripley's K



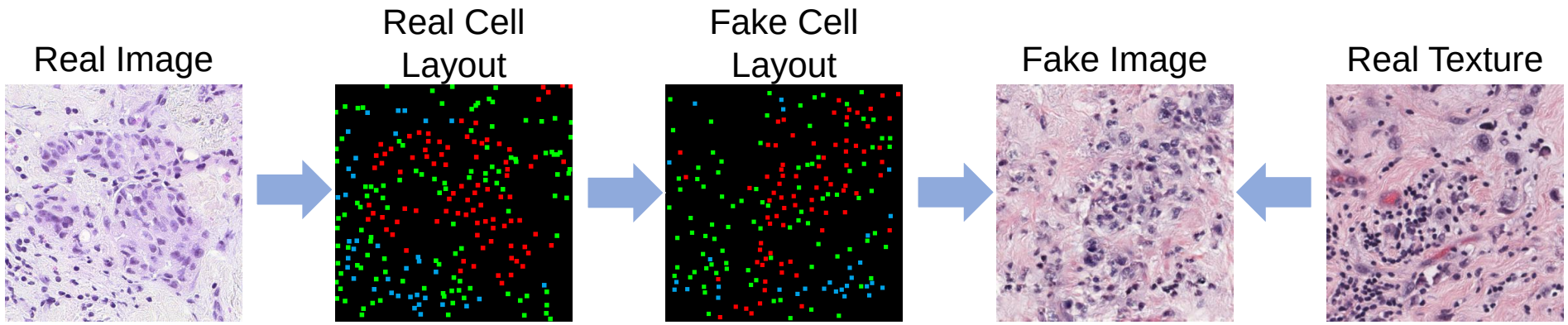
Persistence Diagram



Train Downstream Task

Introduction

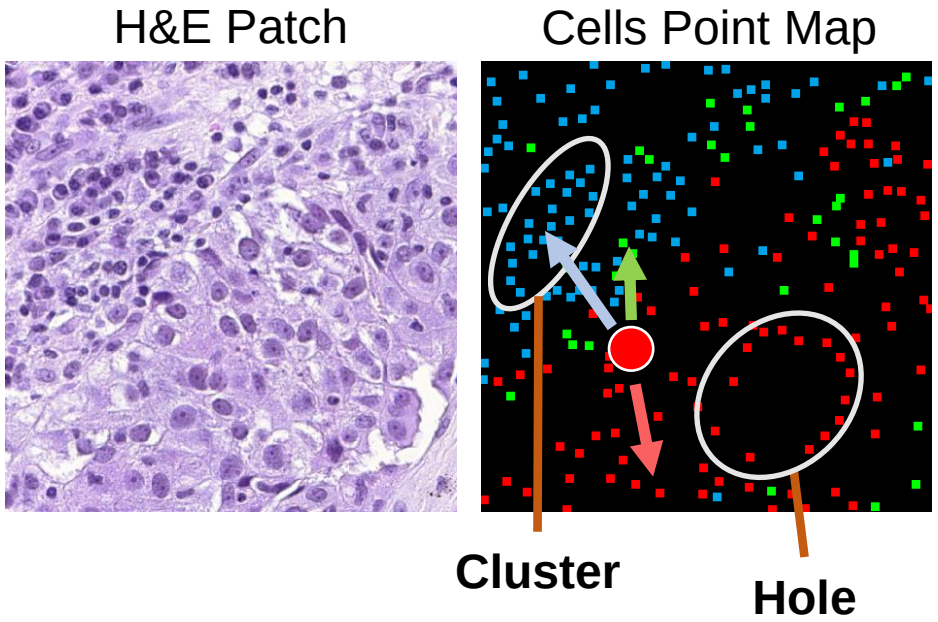
- Pathology image analysis suffers from **limited annotations**.
- **Augment** labeled data with **synthetic labeled** data.
- **Generating pathology images** usually involves **two steps**:
  1. *Generating spatial layout* of cells .
  2. *Filling in stains and textures.*
- **Cell Context**
  - **Important** for pathology data analysis.
  - The **arrangement** of cells.
  - Their **spatial co-localization.**



Cell Configuration Descriptors

Challenges:

Hard for models to learn the underlying distribution.

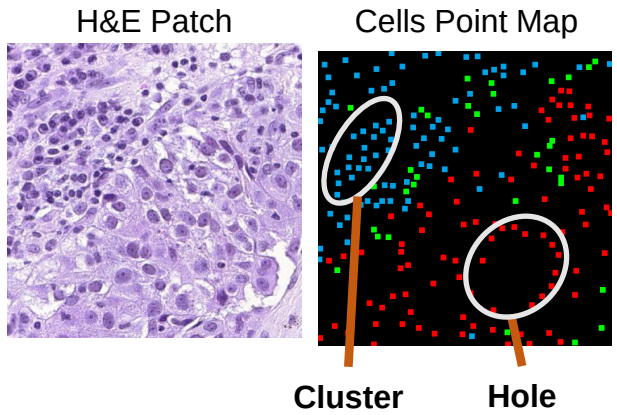


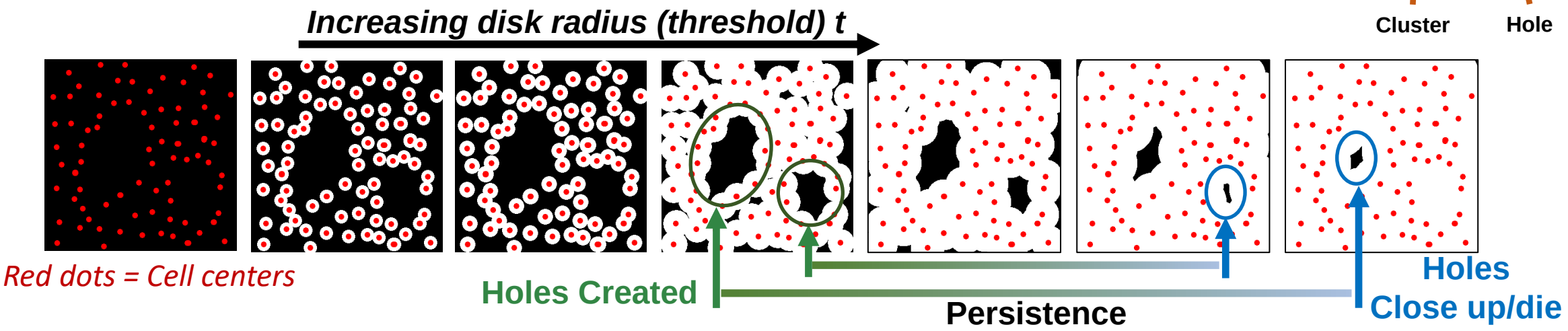
Effective cell configuration descriptors needs to **capture**:

1. *Structural patterns* such as clusters and holes of a reference cell layout. Topological Features
2. S*patial co-localization* of different types of cells. Spatial Statistics

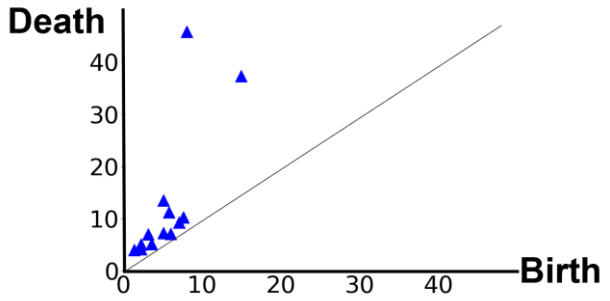
1. Topological Features

- Topological features to represent clusters and holes.
- Persistent Homology Filtration of Point Set[1]





Encode into a Persistence Diagram:



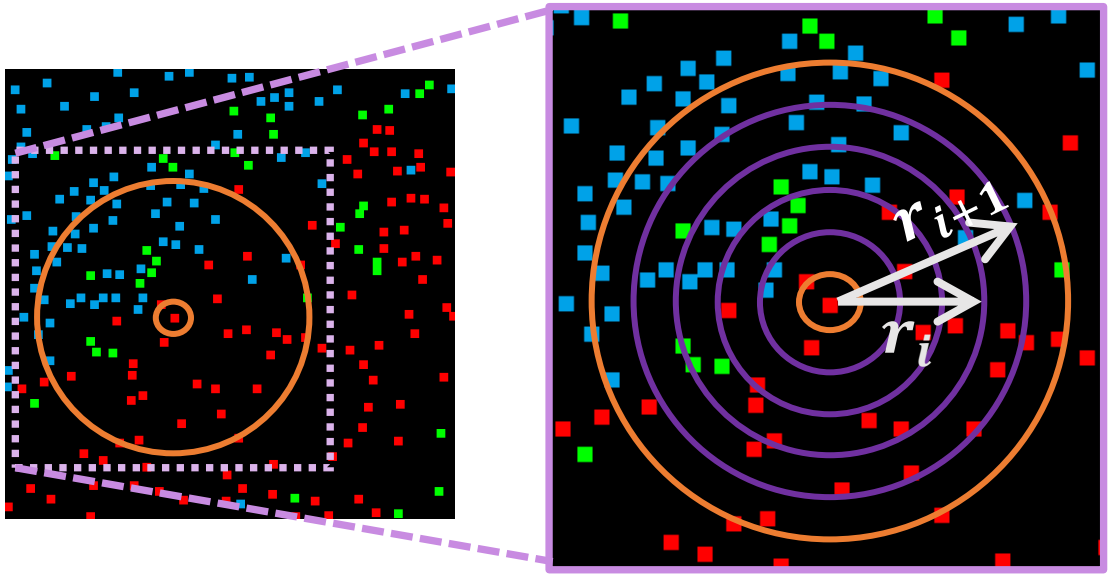
Persistence Diagram

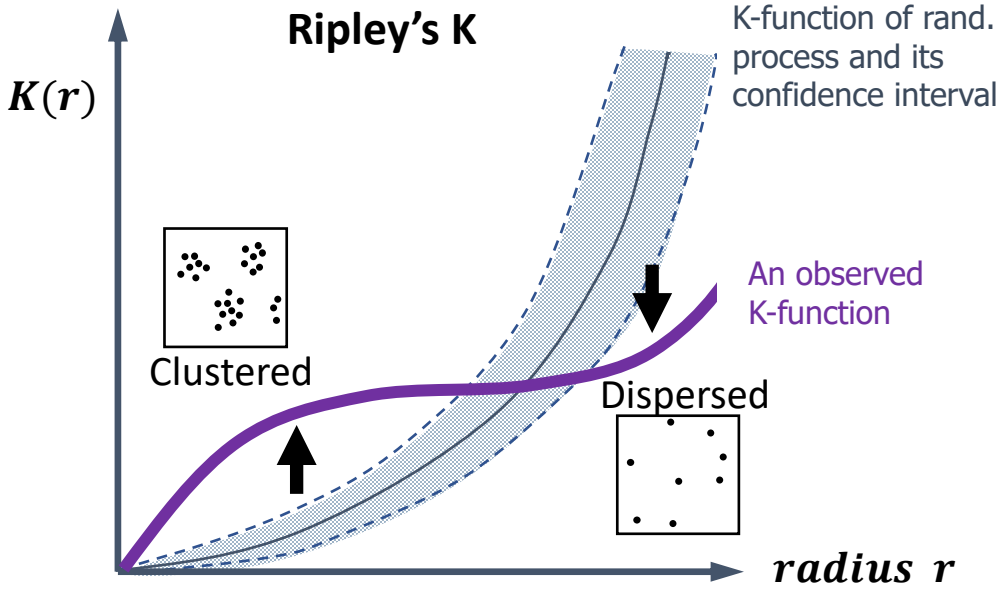
[1] Edelsbrunner et al. Computational Topology an Introduction. American Mathematical Society (AMS) 2010.

2. Spatial Statistics Features

Cross K-functions (Ripley's K)[2]

- Describes the distribution of target class of points surrounding a source.
- **Cells types co-localization**
- **Characterize holes**

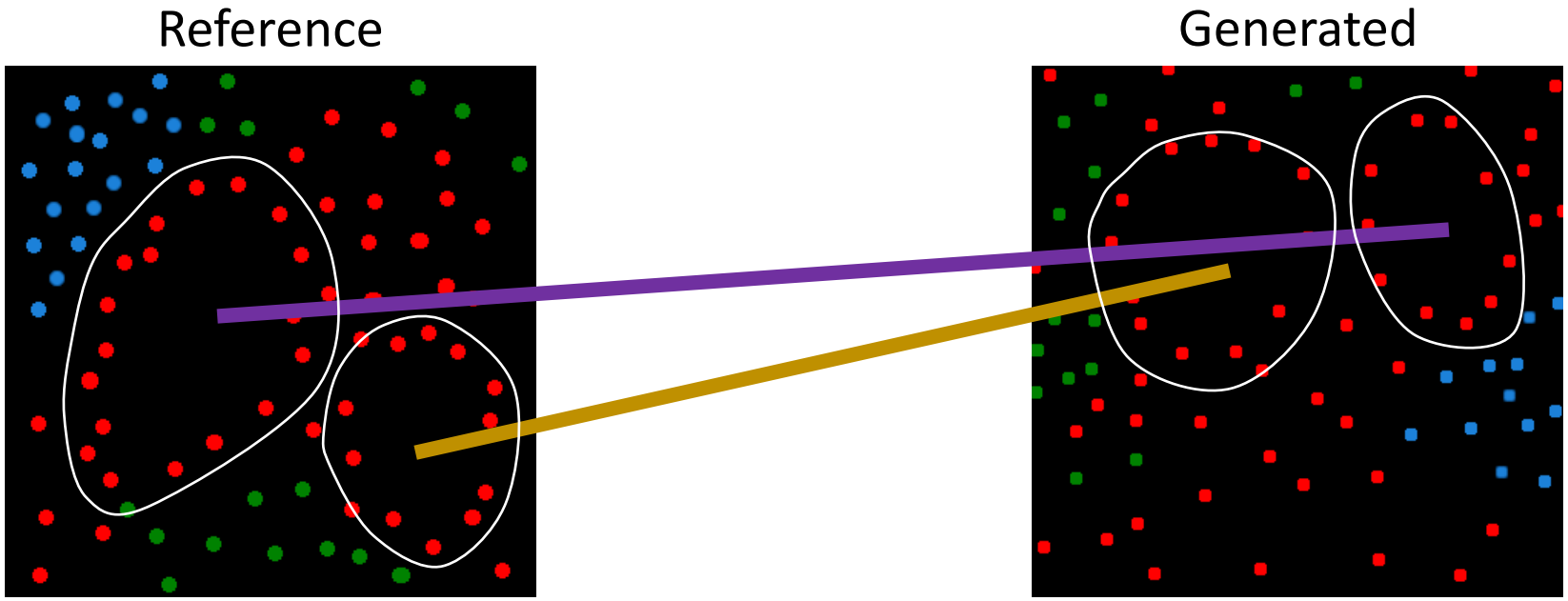




[2] Ripley, B.D. The second-order analysis of stationary point processes, Journal of Applied Probability 13 (2), 255–266, 1976.

Matching Structures (Gaps/Holes)

- Goal: Gen. and Ref. layouts have similar spatial distribution patterns.
- Find correspondence between holes in the generated and the reference layouts.

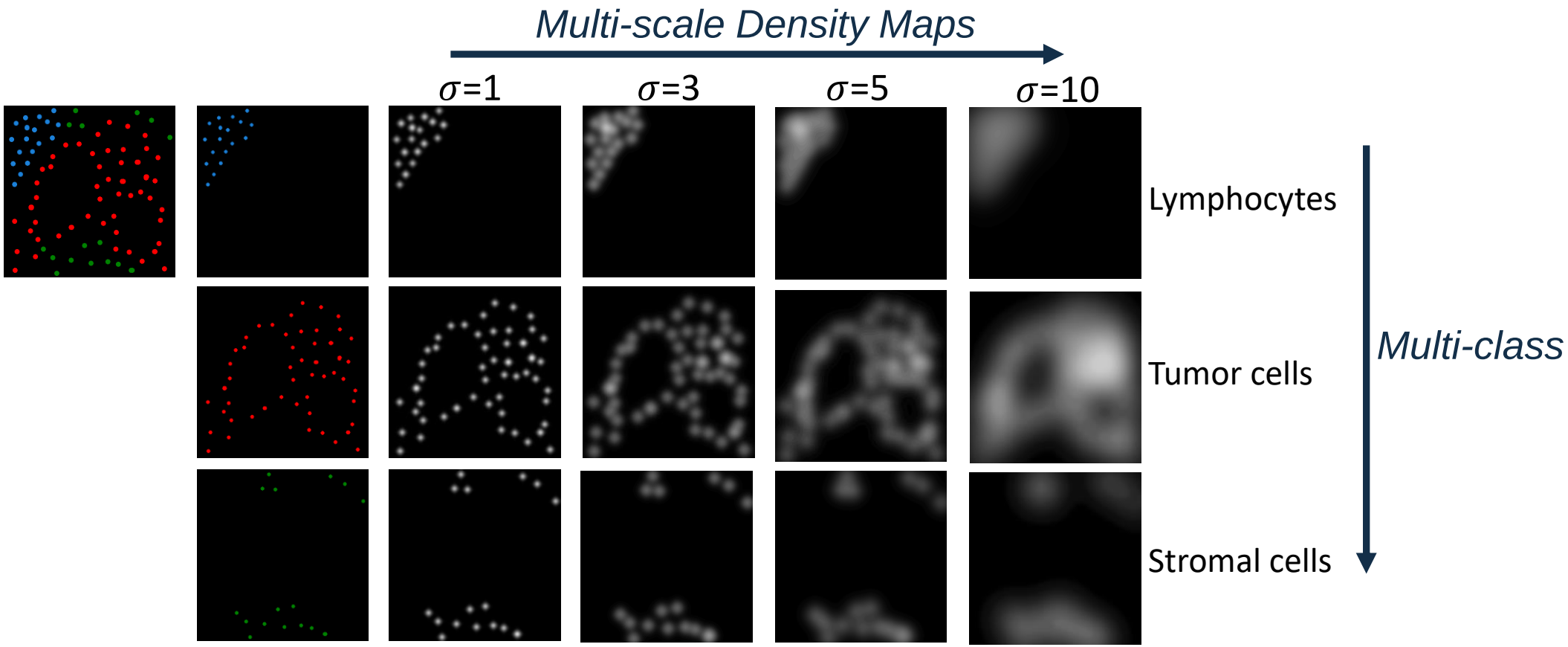


Match based on size (persistence) and spatial context (cross K functions)

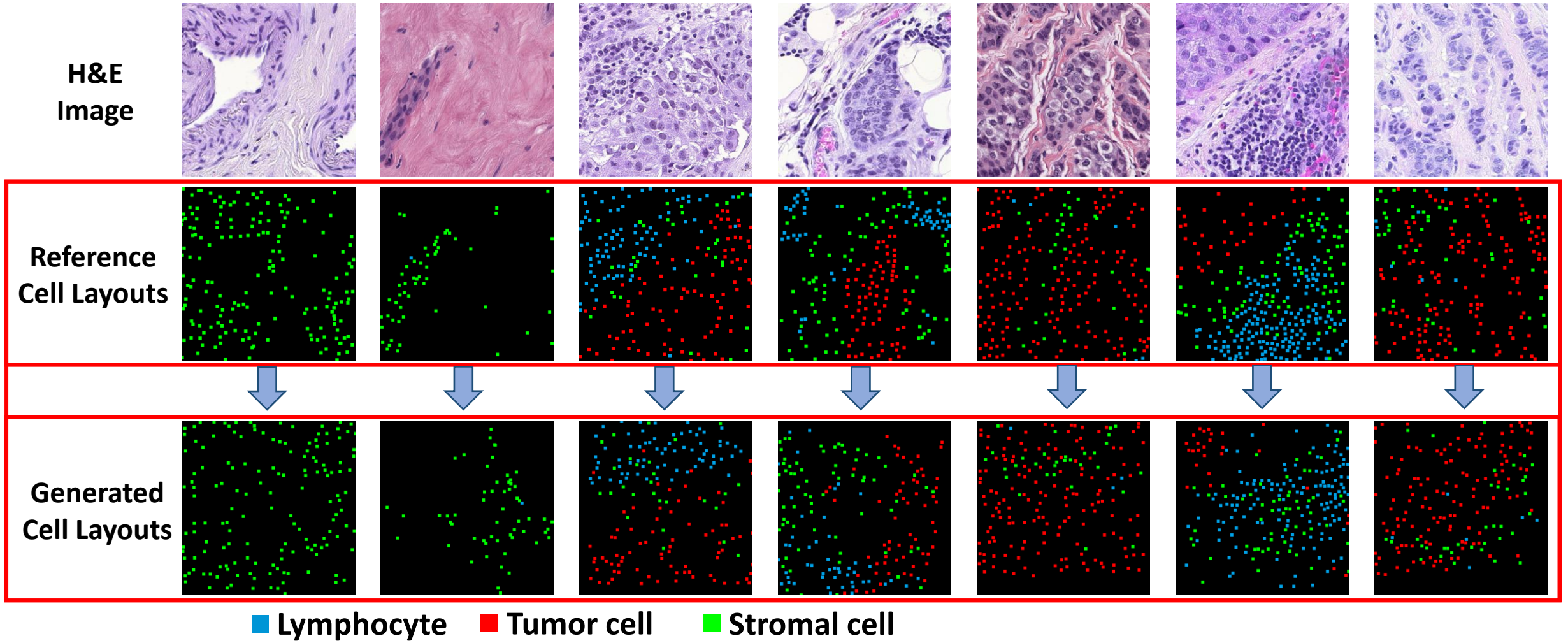
Cell Configuration Loss $\mathcal{L}_{cc}$

Matched locations should have neighborhoods w/ similar spatial context.

i.e., have similar values in the multi-class, multi-scale density maps.



Cells Layout Generation: Sample Results



Cells Layout Evaluation

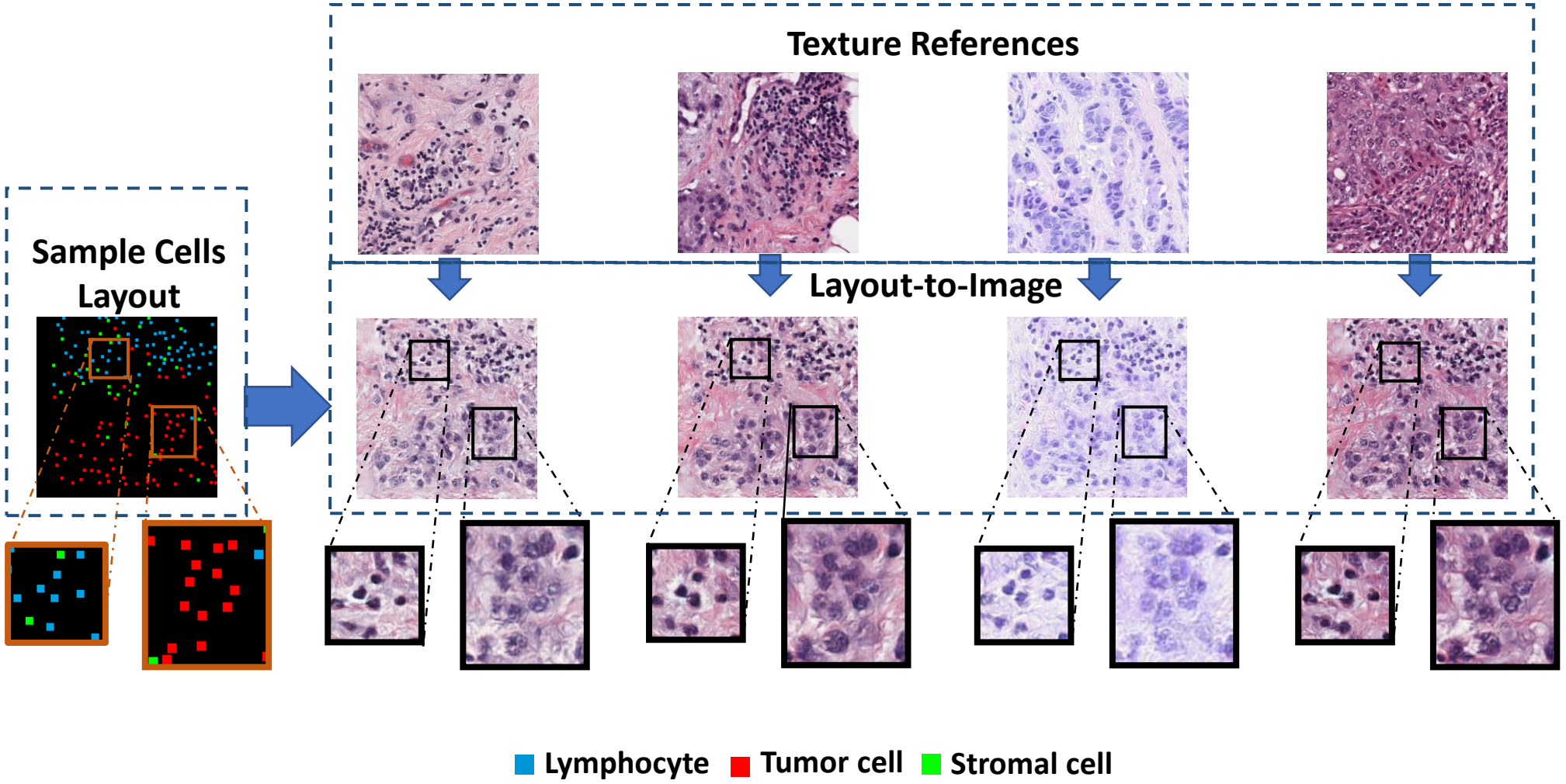
- Cross K-function:

	Cross K-function - MAE				Cross K-function - RMSE			
Method	Lym.	Tumor	Stro.	Mean	Lym.	Tumor	Stro.	Mean
w/o Spatial Descriptors + w/o $\mathcal{L}_{cc}$	0.555	0.096	0.424	0.359	0.829	0.127	0.666	0.541
w/o $\mathcal{L}_{cc}$	0.592	0.126	0.402	0.373	0.861	0.176	0.683	0.573
w/o Cross K-function Descriptor	0.417	0.154	0.431	0.334	0.602	0.226	0.583	0.470
Ours	0.413	0.146	0.357	**0.306**	0.611	0.201	0.509	**0.440**

- Persistence Diagram (PD):

	PD –EMD				PD - Cell Configuration Matching			
Method	Lym.	Tumor	Stro.	Mean	Lym.	Tumor	Stro.	Mean
w/o Spatial Descriptors + w/o $\mathcal{L}_{cc}$	0.28	0.082	0.19	**0.184**	0.8	1.74	1.66	1.4
w/o $\mathcal{L}_{cc}$	0.249	0.203	0.156	0.202	0.9	1.69	1.79	1.46
w/o Cross K-function Descriptor	0.237	0.167	0.17	0.191	0.75	1.74	1.77	1.42
Ours	0.246	0.141	0.165	**0.184**	0.74	1.64	1.71	**1.36**

Texture Generation: Sample Results



Multi-Class Cell Classification

- Train with augmentation data generated by our method.

	F-Score			
Method	Lym.	Tumor	Stro.	Mean
U-Net	0.498	0.744	0.476	0.572
U-Net + Aug. (Rand.)	0.625	0.735	0.472	0.611
U-Net + Aug. (Ours)	**0.65**	**0.768**	**0.511**	**0.644**
MCSpatNet	0.635	0.785	**0.553**	0.658
MCSpatNet + Aug. (Rand.)	0.652	0.772	0.506	0.644
MCSpatNet + Aug. (Ours)	**0.678**	**0.8**	0.522	**0.667**

Conclusion

- **First time to explicitly model and generate realistic multi-class** cell layouts with desirable spatial configuration.
- Propose novel **cell configuration loss** that uses **persistent homology** and **spatial statistics** to **model the cell context.**
- Qualitative and quantitative results show that that our method generates cell layouts with **realistic spatial and structural patterns**.
- Improve performance in **downstream tasks** such as cell classification.
- **Future Work**:
  - **Modeling more complex structures**.
  - Applying to **other downstream tasks**.

Conclusion

- **First time to explicitly model and generate realistic multi-class** cell layouts with desirable spatial configuration.
- Propose novel **cell configuration loss** that uses **persistent homology** and **spatial statistics** to **model the cell context**.
- Qualitative and quantitative results show that our method generates cell layouts with **realistic spatial and structural patterns**.
- Improv fication.
- **Future Work**:
  - **Modeling more complex structures**.
  - Applying to **other downstream tasks**.

Thank You!

Poster Session: TUE-AM-316