Morphological Prototyping for Unsupervised Slide Representation Learning for Computational Pathology

Mass General Brigham
HARVARD MEDICAL SCHOOL
Mahmood Lab AI for Pathology

Andrew H. Song*, Richard J. Chen*, Tong Ding, Drew F.K. Williamson, Guillaume Jaume, Faisal Mahmood
asong@bwh.harvard.edu, richardchen@g.harvard.edu, faisal.mahmood@bwh.harvard.edu

CVPR SEATTLE, WA JUNE 17-21, 2024

Motivation

Computational Pathology

- Patient-level outcome prediction of **digitized tissue sections** (whole-slide images, WSIs), of up to ***100,000 x 100,000*** pixels (at 0.5μm/pixel) [1]
- ***Multiple Instance Learning*** (MIL) (1) tokenizes WSI into a set of image patches encoded using a pretrained vision encoder and (2) aggregates patch embeddings into a slide embedding for patient-level task.

Limitations of MIL

- Resulting slide embeddings are ***specific*** to the downstream task
- Due to **large-p** (# of parameters) and **small-n** (# of patients), unstable training for supervised models

Can we create a *task-agnostic*, *unsupervised* slide embedding?

Representations summarizing slide statistics

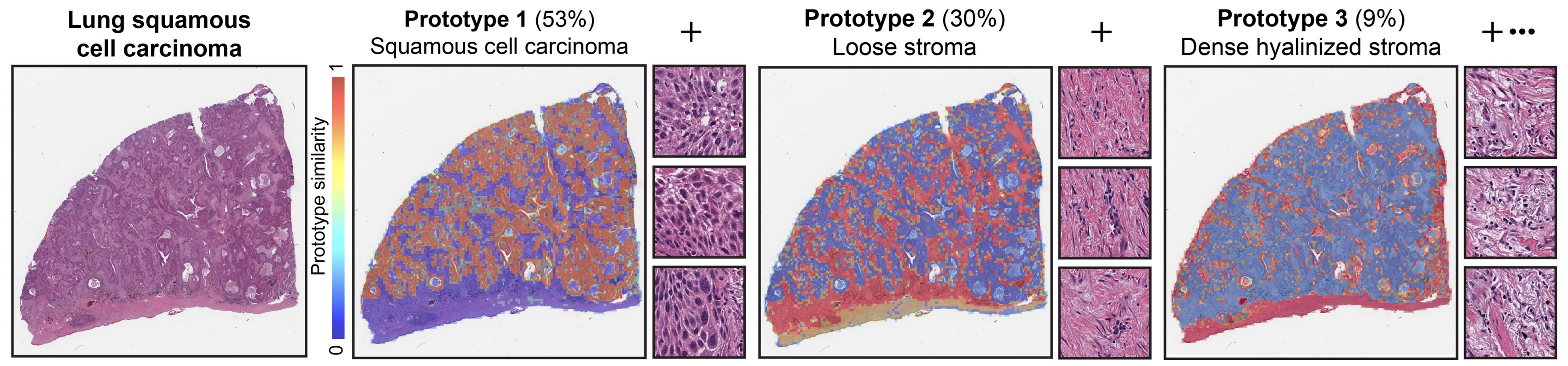
Redundant morphological information in WSI

- Handful of morphological patterns repeated throughout the tissue (e.g., cancer cells, stroma, fat)

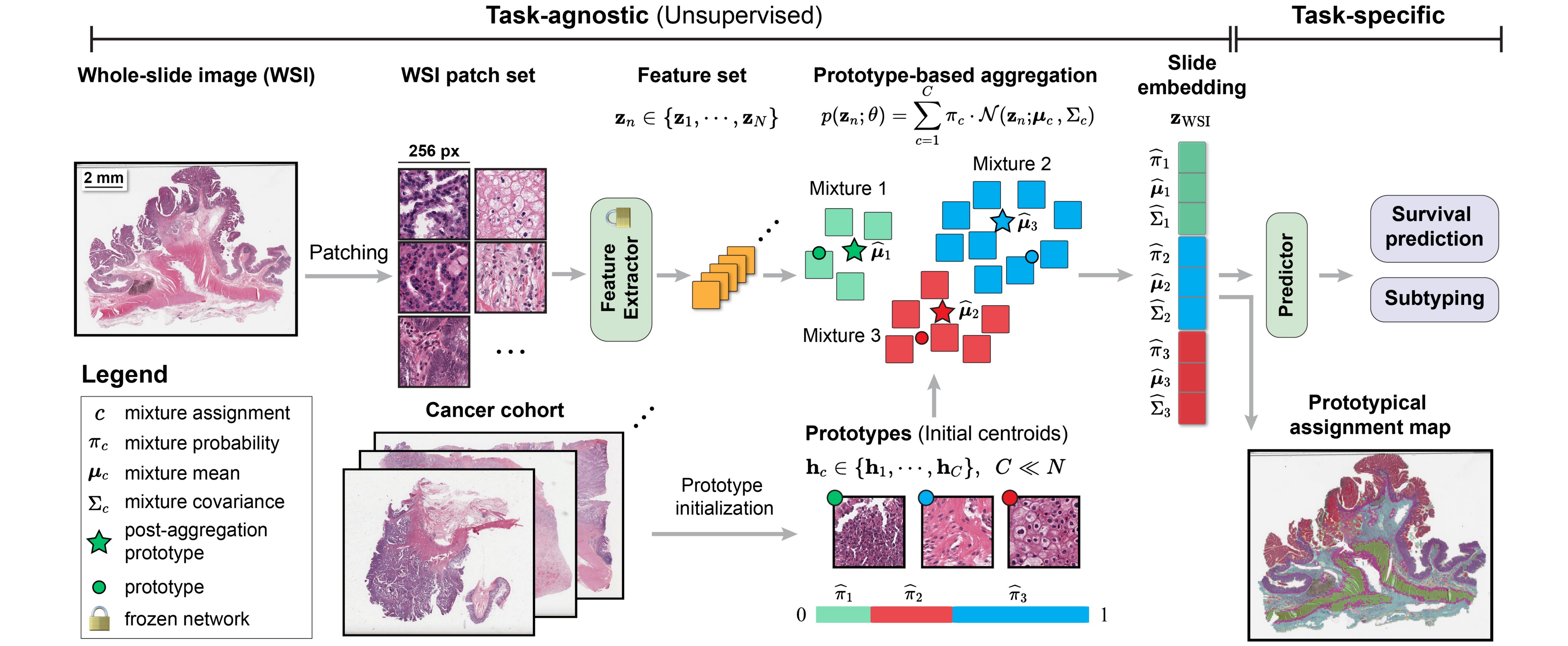
Prototype-based summarization of WSI

- **WSI** ≅ Distribution of ***morphological concepts*** (a.k.a. ***prototypes***)
- Summarization of WSI based on two important conditions
  - Feature representation of each concept
  - Cardinality (proportion) of each concept in WSI
- Huge compression
  - Prototypes ($C = 8\sim32$) ≪ patches per WSI ($N \sim 10^4$)
- Optimal transport [3], Gaussian mixture models [4] are good candidates

⟹ Prototype **A**ggregatio**N**-based framework for compact **HE**terogenous slide set **R**epresentation (**PANTHER**)



PANTHER for Slide Representation Learning



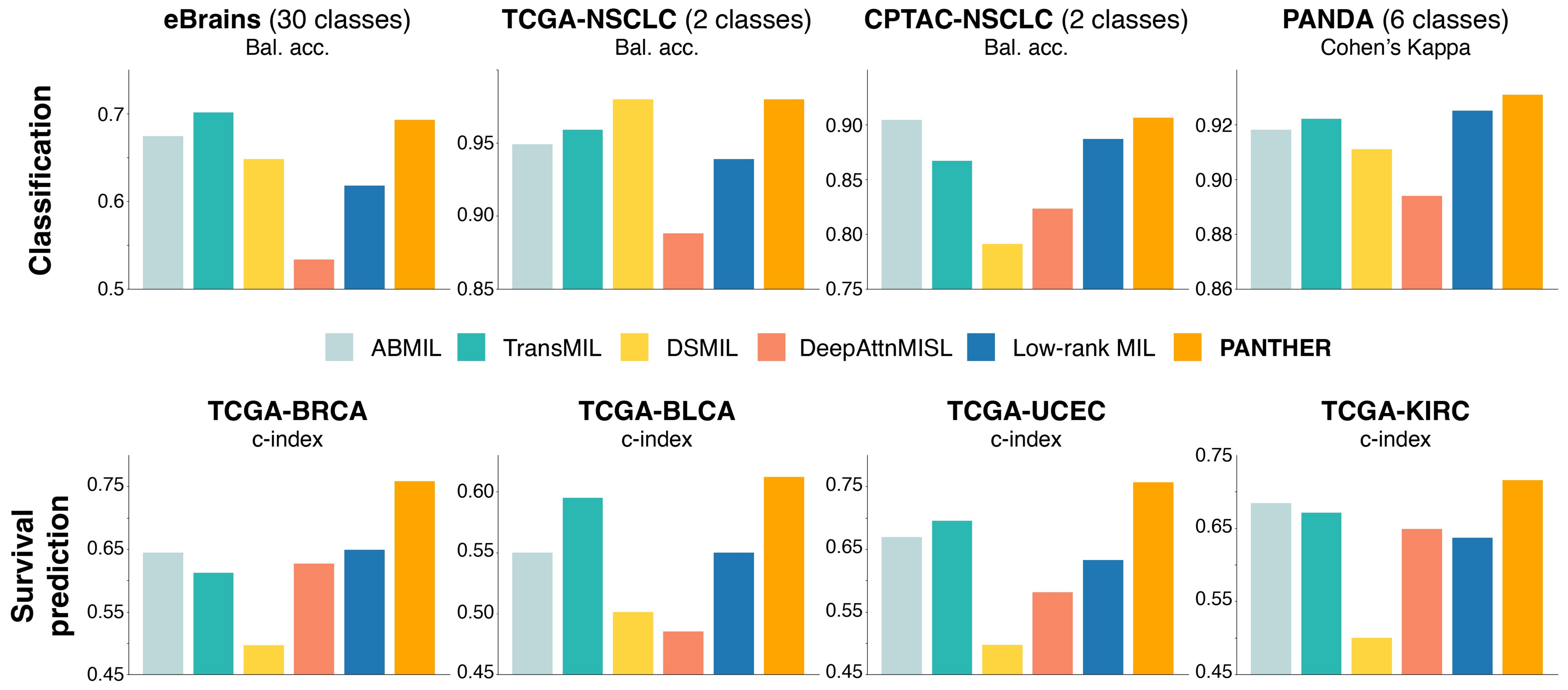
Generative model for patch embedding (Gaussian Mixture Model)

- $p(z_n;\theta) = \sum_{c=1}^{C} \pi_c \cdot N(z_n;\mu_c,\Sigma_c)$ ⟹ Each component: a prototype and its distribution
- **Unsupervised** $z_{\mathrm{WSI}} = \left[\hat{\pi}_1, \hat{\mu}_1, \hat{\Sigma}_1, \ldots, \hat{\pi}_C, \hat{\mu}_C, \hat{\Sigma}_C\right] \in \mathbb{R}^{c\cdot(2d+1)}$ ⟹ EM algorithm for param. Estimation
- Per-prototype feed-forward network (linear or MLP) for downstream task

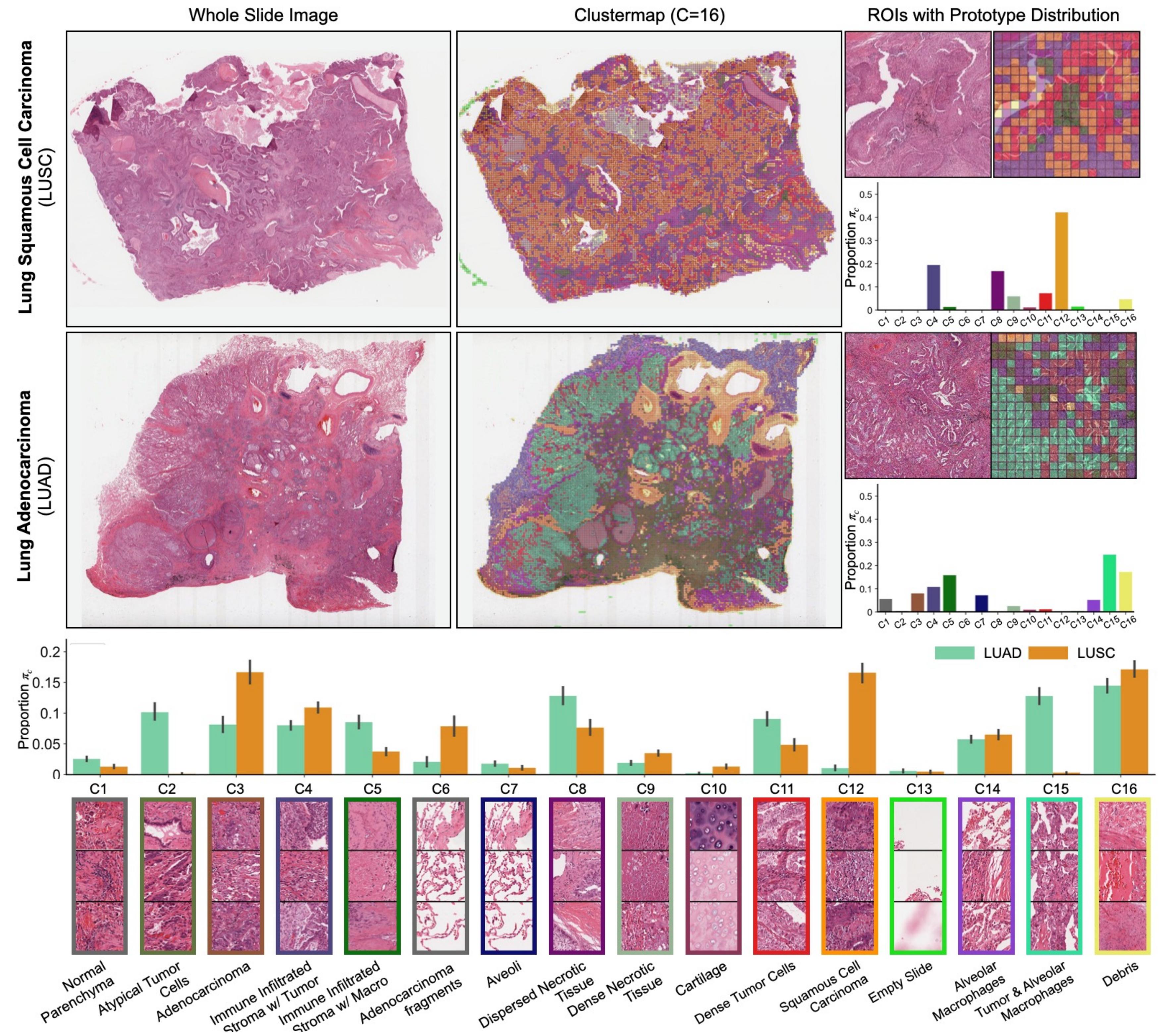
PANTHER for slide-level evaluation

Unsupervised slide embedding for downstream task

- Extensive evaluation on 4 cancer classification and 9 cancer survival datasets
- Competitive performance against other **supervised MIL** baselines



PANTHER for Interpretability



Prototype-oriented interpretability

- Visualization of the most similar prototype on WSI (Clustermap)
- Quantification of prototype distribution per WSI

References

- [1] Song et al., Artificial intelligence for digital and computational pathology. *Nature Reviews Bioengineering*, 2023
- [2] Ilse et al., Attention-based deep multiple instance learning. *ICML*, 2018
- [3] Mialon et al., A Trainable Optimal Transport Embedding for Feature Aggregation and its Relationship to Attention. *ICLR*, 2021
- [4] Kim M., Differentiable Expectation-Maximization for Set Representation Learning. *ICLR*, 2022

PAPER