

Histoscope

Expert-grounded SAE inspection for computational pathology



How well can an automatic metric identify *coherent morphology* in SAE features?
Two pathologists tested 100 feature panels.

THE ANSWER

Precise, but with a **blind spot**.

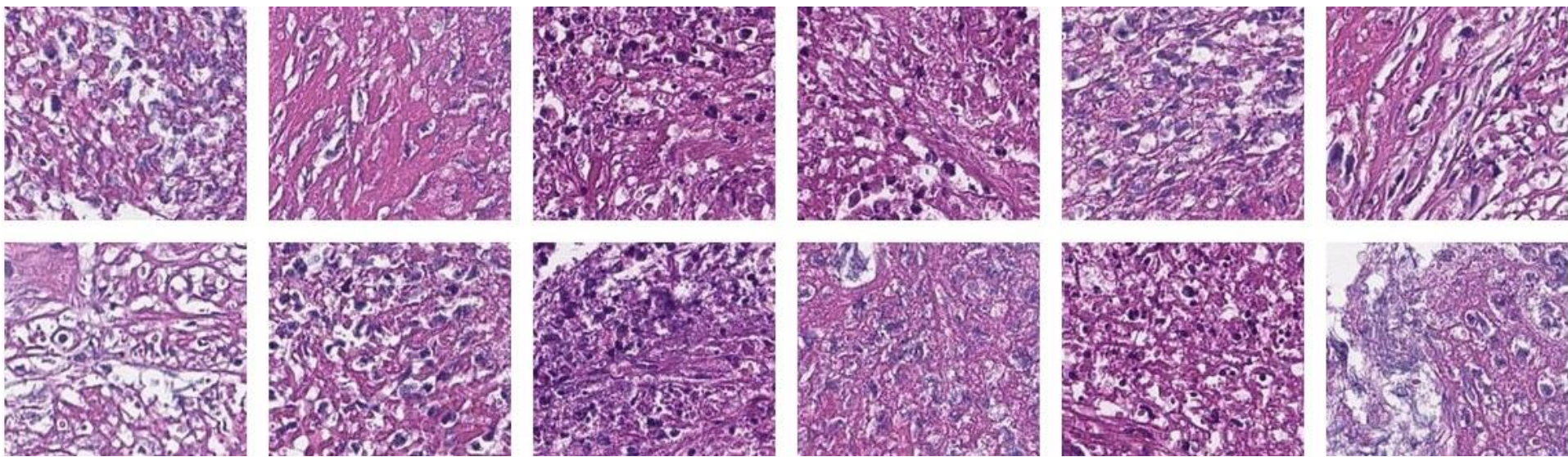
Class-selectivity works for class-concentrated morphology but misses coherent cross-class patterns.
Histoscope makes SAE features inspectable for expert evaluation.

✓ WHEN CLASS-SELECTIVITY WORKS · class-concentrated morphology

Necrosis

AUPRC 0.96

Expert consensus: **MONOSEMANTIC**



Feature #14964 · Concentrated in a single diagnostic class

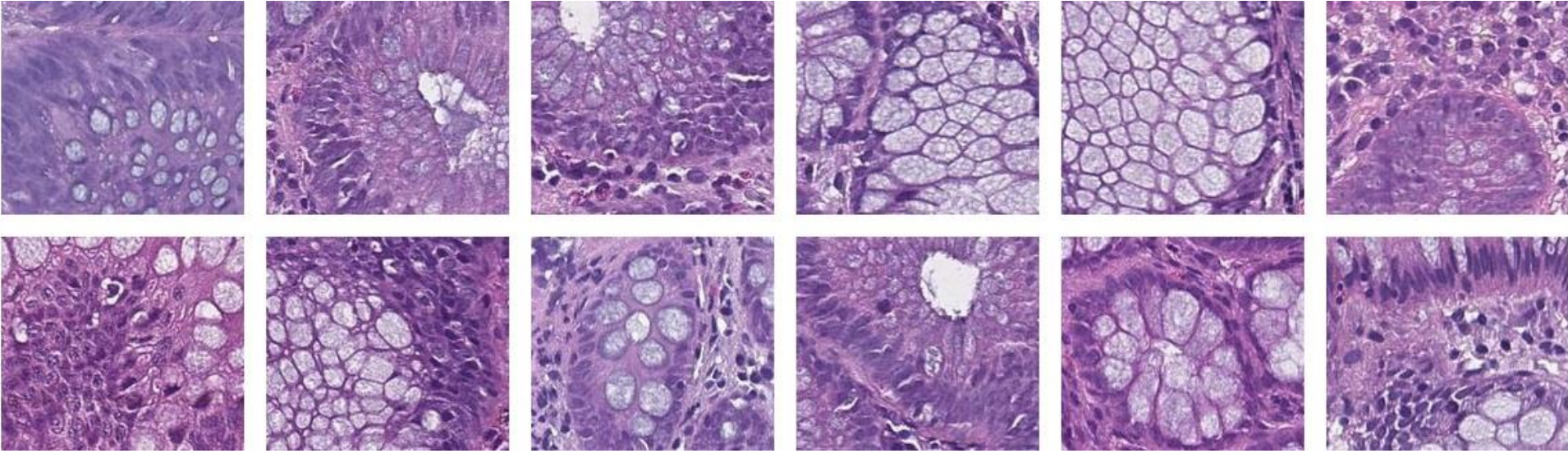
VS

✗ WHEN CLASS-SELECTIVITY MISSES · cross-class coherent morphology

Goblet-cell mucus

AUPRC 0.20

Expert consensus: **MONOSEMANTIC**



Feature #25534 · Coherent morphology across multiple classes

01 Tool labels vs. expert consensus

resolved set, n = 88 · 12 unresolved excluded

		HISTOSCOPE LABEL (AUPRC RULE)	
		AUPRC-mono monosemantic	AUPRC-poly polysemantic
Expert monosemantic		50 TP	32 FN
Expert polysemantic		0 FP	6 TN

AUPRC selectivity = how consistently a feature fires for a single diagnostic class.

✓ 50 / 50

AUPRC-mono features confirmed by pathologists — perfect precision.

✗ 32 / 38*

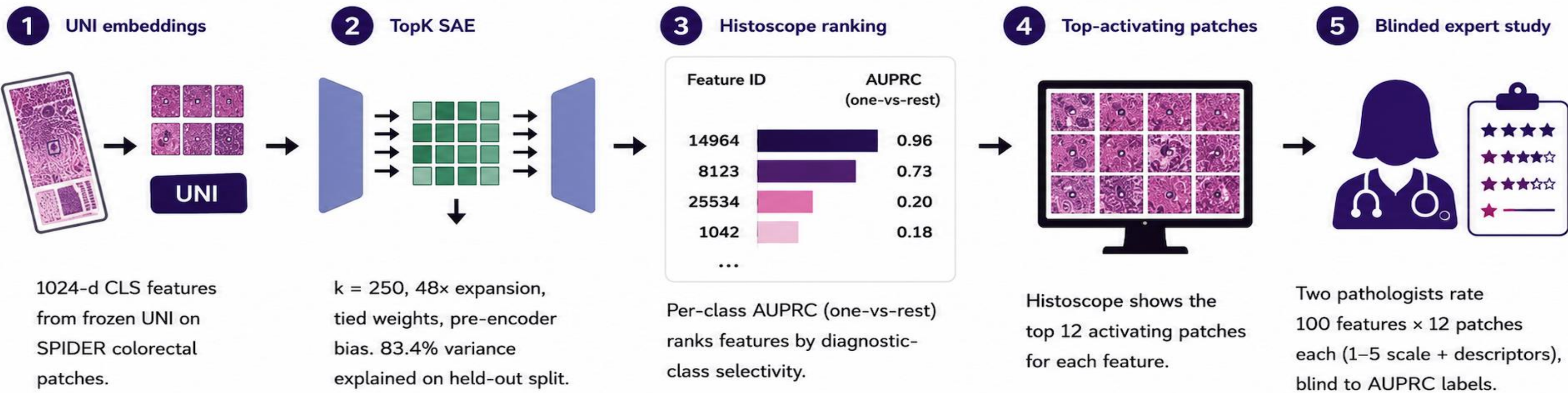
AUPRC-poly features also judged monosemantic — the metric misses coherent cross-class morphology.

★ Most misses were not noise

29 of 32 rescued AUPRC-poly features spanned ≥ 2 diagnostic classes, vs. 45 of 50 confirmed AUPRC-mono features being single-class.

02 From foundation model to expert judgement

UNI → SAE → ranking → blinded study



03 The Histoscope interface



Fig. 1 Histoscope interface — feature #25534. The feature × diagnostic-class heatmap and top-activating patches (left); selecting a cell opens the feature summary, margins & purity, AUPRC selectivity, and per-class activation distribution (right).

04 Why it matters

In computational pathology, **monosemanticity** is not the same as **diagnostic-class selectivity**.

A feature can be morphologically coherent while spanning multiple diagnostic labels.

Use class-selectivity as a **filter**, not a verdict:
Histoscope exposes the tissue evidence for expert review.



Paper



Code & tool

Mirza Nasir Hossain¹ · Sarah Bell² · Gareth Bryson² · David Harris-Birtill¹

¹ University of St Andrews · ² NHS Greater Glasgow and Clyde · Email: mnh3@st-andrews.ac.uk