



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Supplementary Material

Thumbnail-Based Tile Selection for Efficient Whole-Slide Image Classification

S1 Patient-grouped bootstrap

Cohort intervals resample *patients* with replacement, carrying all of a patient's slides together, so correlated slides cannot be split across a bootstrap draw. Grouping changes the interval only where a patient contributes more than one slide. In the breast test split (254 slides from 119 patients, 63% of them with more than one) it widens the interval. In TCGA-RCC (937 slides from 895 patients) the grouped and ungrouped intervals agree to three decimals (95% CI [0.912, 0.954] against [0.912, 0.954] for *u-net*), and in UBC-OCEAN, where each case contributes exactly one slide, grouping is the identity. We therefore report the grouped interval throughout: it is the conservative choice on breast and costs nothing on the other two.

S2 Training-free selectors and end-to-end cost

Table S1. Training-free thumbnail selectors on breast ER status, mean $\pm$ s.d. test AUROC over the ten folds at fixed tile counts (the median slide holds 8,018 tiles, so 8 tiles is about 0.1%). The full bag scores 0.776. A regular *lattice* matches *random* everywhere – uniform coverage alone carries no signal – and the Macenko *stain* map is below it. Both learned selectors separate from all three in the tail.

Selector	4096	1024	256	64	16	8	4
<i>random</i>	0.776 $\pm$ 0.040	0.774 $\pm$ 0.040	0.768 $\pm$ 0.040	0.754 $\pm$ 0.043	0.722 $\pm$ 0.042	0.698 $\pm$ 0.047	0.668 $\pm$ 0.046
<i>lattice</i>	0.775 $\pm$ 0.041	0.771 $\pm$ 0.039	0.767 $\pm$ 0.037	0.753 $\pm$ 0.037	0.719 $\pm$ 0.046	0.699 $\pm$ 0.039	0.659 $\pm$ 0.054
<i>stain</i>	0.771 $\pm$ 0.044	0.768 $\pm$ 0.050	0.754 $\pm$ 0.058	0.717 $\pm$ 0.057	0.678 $\pm$ 0.056	0.661 $\pm$ 0.057	0.646 $\pm$ 0.068
<i>attn</i>	0.776 $\pm$ 0.039	0.774 $\pm$ 0.041	0.769 $\pm$ 0.044	0.757 $\pm$ 0.053	0.741 $\pm$ 0.060	0.728 $\pm$ 0.063	0.710 $\pm$ 0.068
<i>u-net</i>	0.774 $\pm$ 0.042	0.772 $\pm$ 0.045	0.769 $\pm$ 0.050	0.758 $\pm$ 0.048	0.734 $\pm$ 0.056	0.726 $\pm$ 0.053	0.712 $\pm$ 0.054

Table S2. End-to-end per-slide cost, median over 20 breast slides on the deployment filesystem. The main text prices the neural components only; reading a tile off the slide costs an order of magnitude more than encoding it, and both routes pay it per tile, so the saving survives while the absolute latencies do not.

Stage	Paid	Median
Open the slide	once per slide	211 ms
Thumbnail + resize to 224px	once per slide	14 ms
U-Net forward	once per slide	3.2 ms
Read one 256 px tile at 20 $\times$	per tile	81 ms
CONCH forward on one tile	per tile	8.1 ms
Budget path ($k = 46$ tiles)		4.3 s
Full bag (8018 tiles)		715 s
Speed-up		165$\times$