

Molecular Subtype Prediction for Breast Cancer Using H&E Specialized Backbone

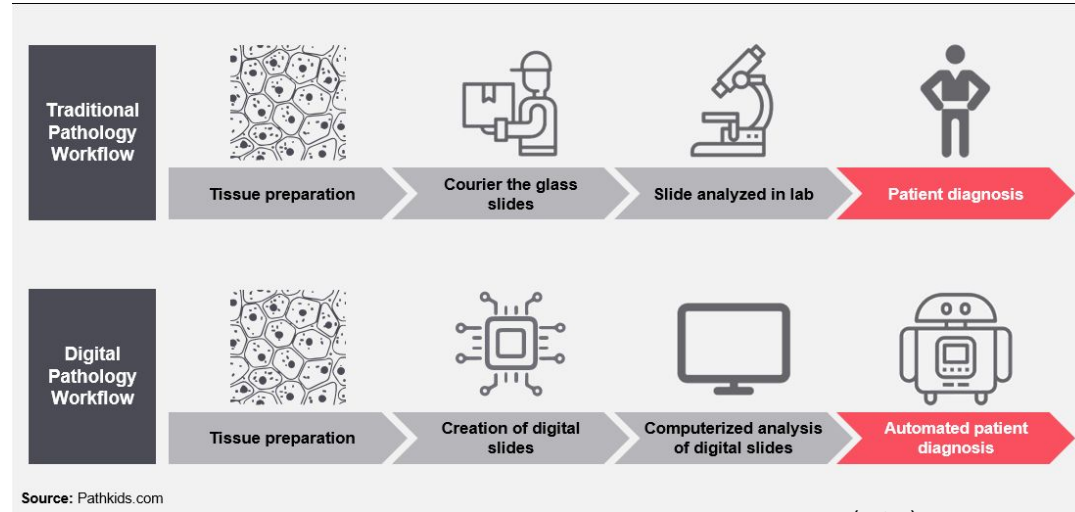
Samaneh Abbasi-Sureshjani, Anıl Yüce, Simon Schönenberger, Maris Skujevskis, Uwe Schalles, Fabien Gaire, Konstanty Gorski

Proceedings of the MICCAI Workshop on Computational Pathology, PMLR
156:1-9, 2021 [\[link\]](#)

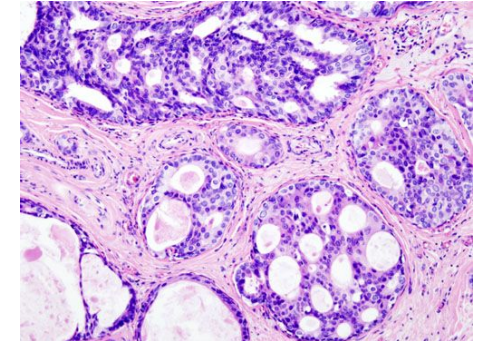
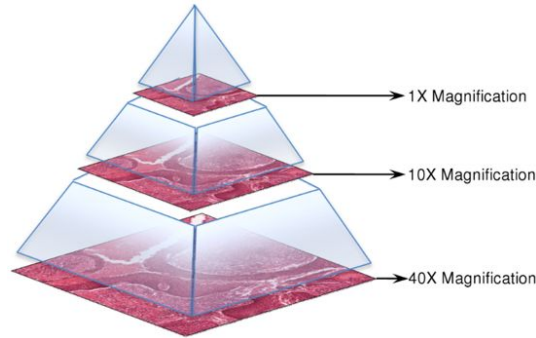
Digital (Histo)Pathology

DP

- Histo: the Greek word for 'Tissue' i.e., thin slices of an organ/animal/plant that is removed from the host to be examined.
- Steps: Tissue preparation \Rightarrow fixation \Rightarrow staining \Rightarrow scanning \Rightarrow Read to process
- H&E: cellular morphological features
- IHC: cellular surface antigens



Hematoxylin and eosin (H&E)

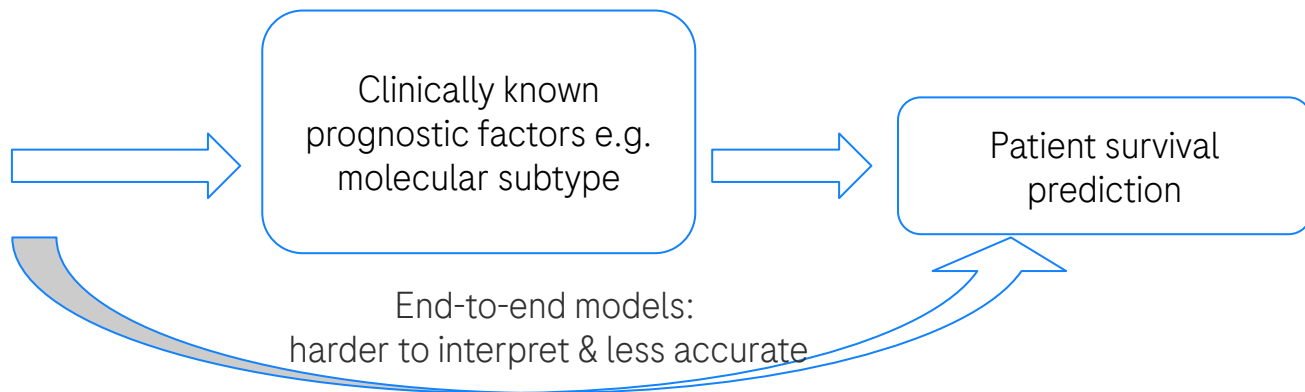
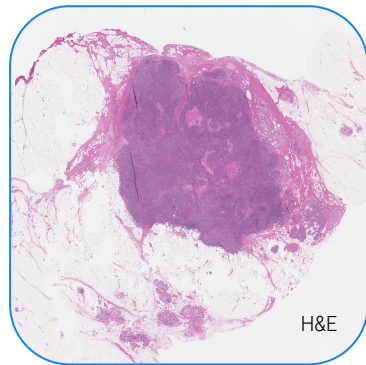


H&E based prognostic algorithms

Personalize treatment options for early breast cancer patients



AI powered - DP prognostic model for better stratification of HER2 eBC patients



Impact:



eBC patients' stratification beyond standard prognostic biomarkers



Future clinical trials design support



Optimization of treatment options and increase patients' access to tailored therapies

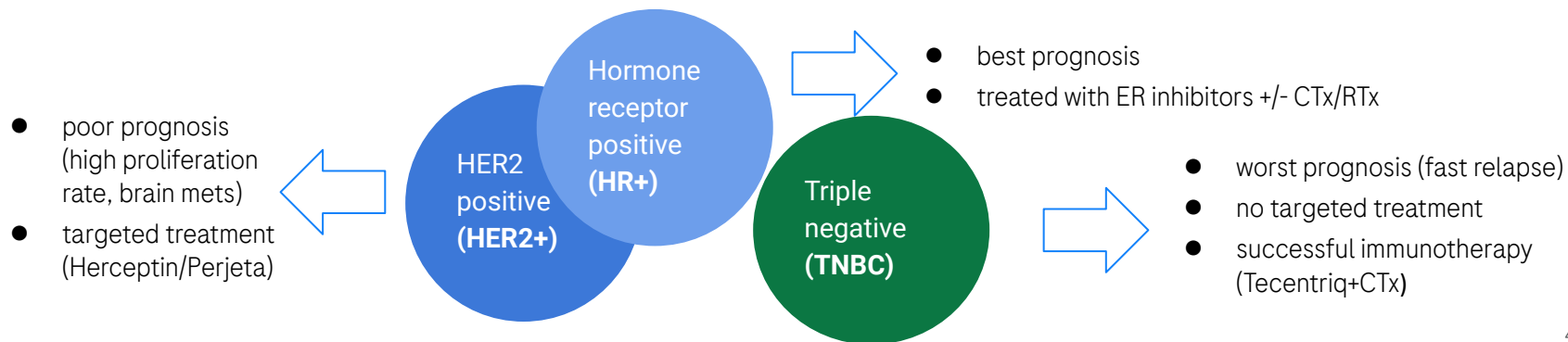


Providing additional tools supporting clinical decision making process

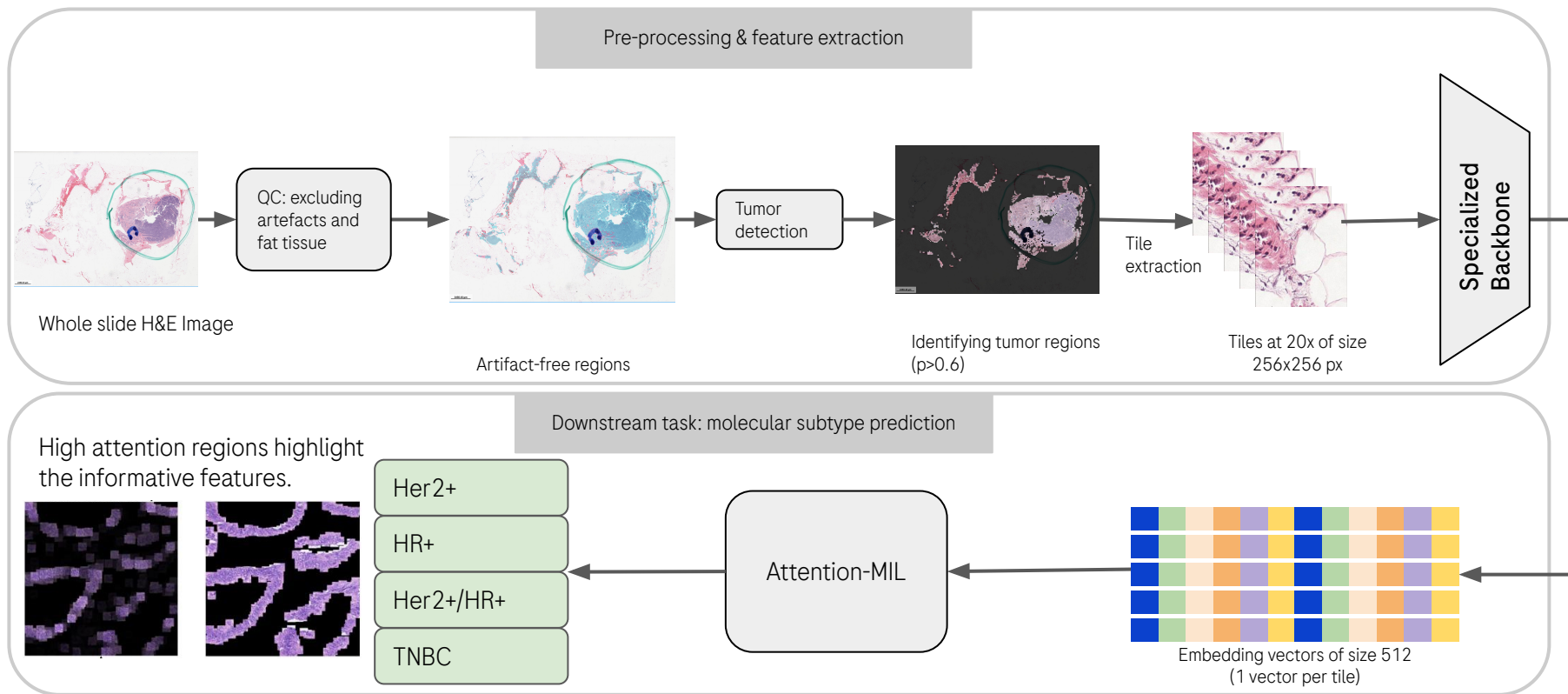
Molecular subtypes of invasive breast cancer

Controlling how the cells behave

- The molecular subtypes are based on the genes the cancer cells express
- Presence of specific proteins (receptors) for hormones determines the categories
- The standard clinical practice:
 - Analysis based on multiple immunohistochemistry (IHC) stainings for each biomarker
 - Expensive and inconsistent when lacking resources.
- One step prediction based on H&E slides?

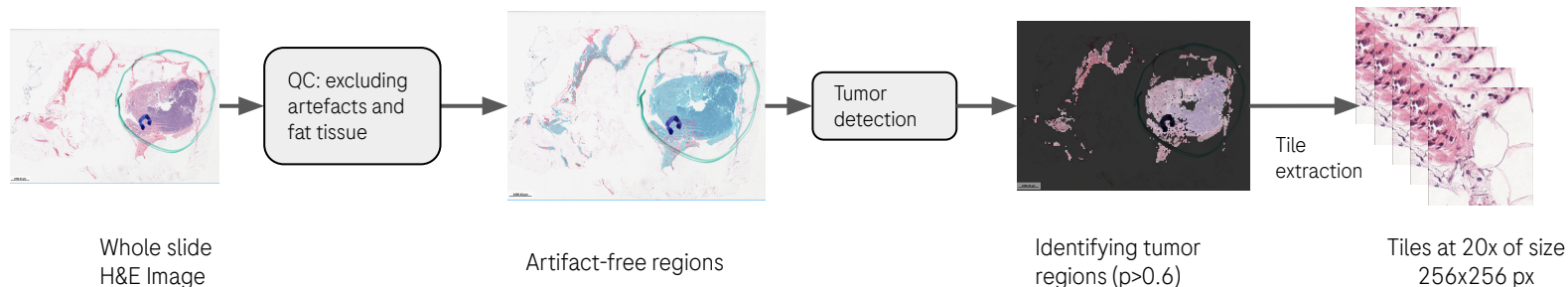


Algorithm Overview



Preprocessing

QC, Tumor detection & tile extraction

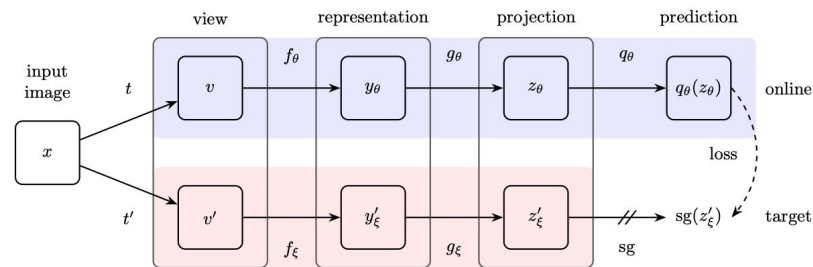


- Identifying good quality tissue regions
- Excluding background, artefacts such as blurred tiles and pen markers
- Detecting the tumor region (excluding the healthy tissue)
- Tile extraction at 20x magnification with 256x256 resolution
 - Good combination of cellular and global context

Feature Extraction

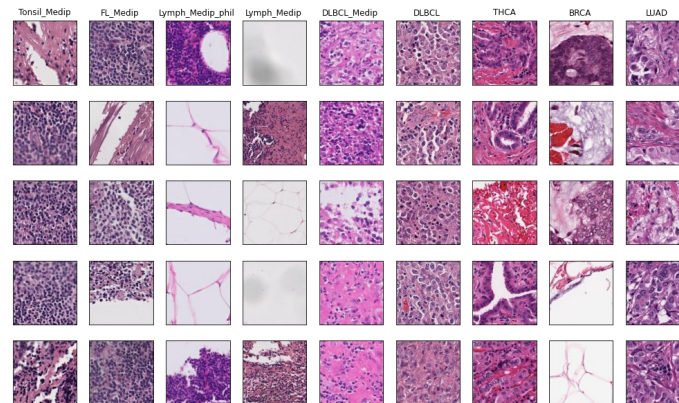
Specialized Backbone

- Pre-extraction of features to help speeding up the training.
- Imagenet-based backbones are sub-optimal for DP.
- Self-supervised learning: a set of auxiliary tasks are created to supervise the model.
 - No expensive annotations needed
 - Various augmentations applied
 - Encoders: Resnet architectures
- Training set: ~1m tiles (256×256 tiles at 0.5 μm/pixel) of different tissue types
- Backbone has been evaluated on different benchmark tasks.



$$\mathcal{L}_{\theta, \xi} \triangleq \|\overline{q_{\theta}}(z_{\theta}) - \overline{z'_{\xi}}\|_2^2 = 2 - 2 \cdot \frac{\langle q_{\theta}(z_{\theta}), z'_{\xi} \rangle}{\|q_{\theta}(z_{\theta})\|_2 \cdot \|z'_{\xi}\|_2}.$$

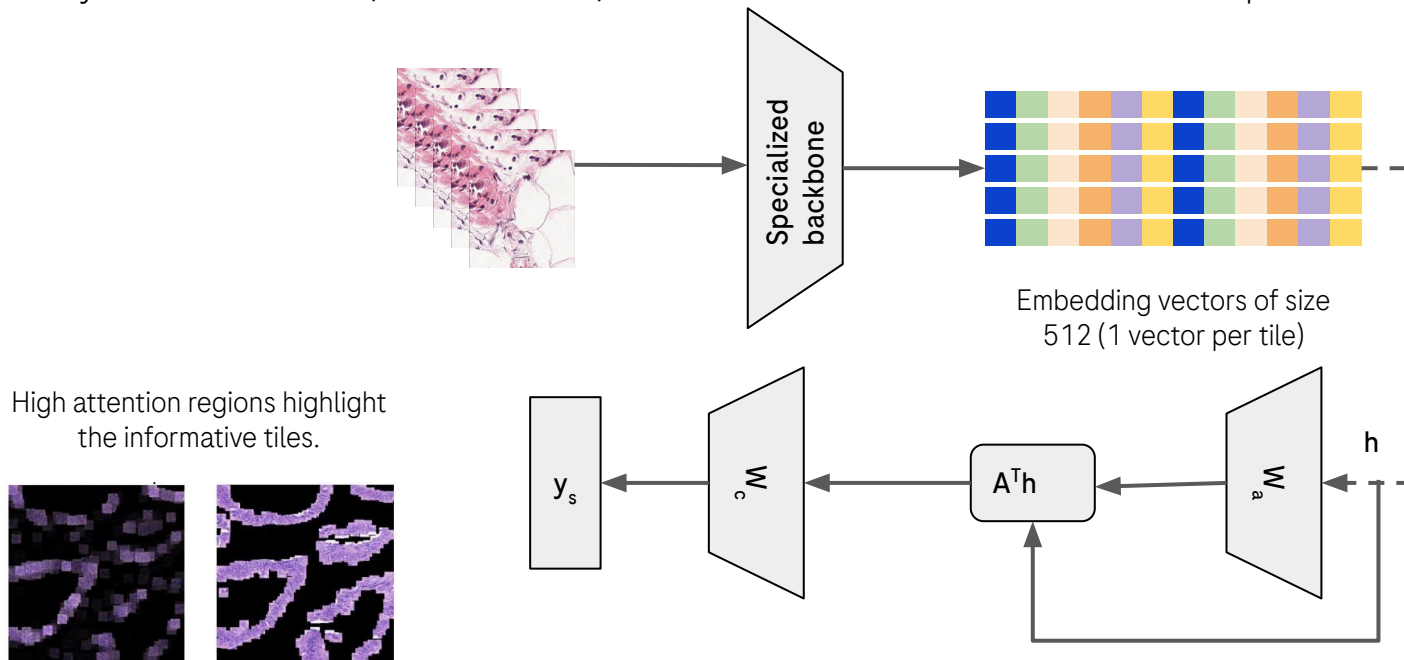
Different tissue types



End-to-end deep learning model

Embedding-based approach

Only slide-level labels (based on IHCs) are available thus an attention-based deep MIL framework is used.



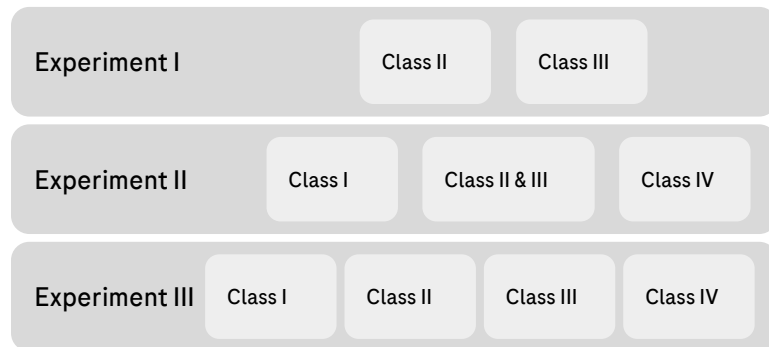
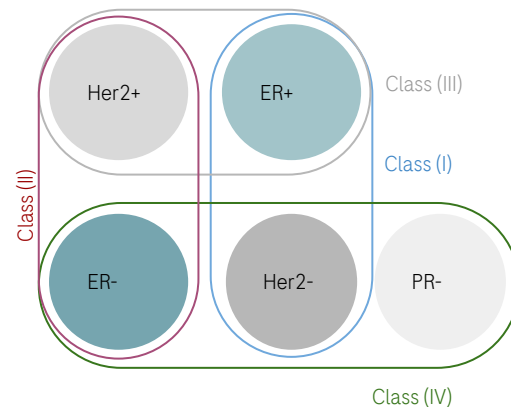
Experiment Design

Datasets & experiments

- A diverse set of public and private datasets with different scanners
- Samples may have different antigens or hormonal receptor status
- Created four main categories
 - Some classes may have overlaps.
- Test 1: 20% from all cohorts
- Test 2: unseen data sources in the test set (all CVD data at test time)
- Different experiments to control the complexity and samples size.

An overview of the number of slides per cohort used in train/test sets

Cohort	class (I)	class (II)	class (III)	class (IV)	Total
TCGA-BRCA	493/121	32/11	123/29	121/33	769/194
CVD	150/36	50/12	82/26	160/40	442/114
ID1	71/19	9/2	24/4	6/1	110/26
ID2	939/248	118/26	168/44	513/128	1738/446
Test 1	1653/424	209/51	397/103	800/202	3059/780
Test 2	1503/574	159/101	315/185	640/362	2617/1055



Results

Overall & per dataset

- Our specialized backbone archives comparable results to the backbone pre-trained on imagenet
- Better performance in terms of generalization to unseen data
- Comparative results to the SOTA techniques

Test AUCs obtained per cohort and experiment. Num: Number of test slides

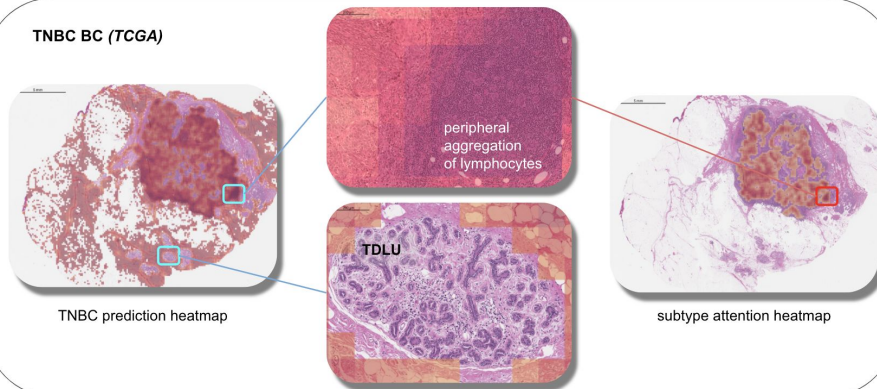
Setting	Backbone	TCGA-BRCA		CVD - DP200		CVD - Philips		ID1		ID2	
		Num	AUC	Num	AUC	Num	AUC	Num	AUC	Num	AUC
Exp.1 & Test 1	ImageNet H&E	40	0.639 0.649	19	0.885 0.731	19	0.782 0.769	6	1.00 0.25	70	0.736 0.783
Exp.1 & Test 2	ImageNet H&E	40	0.680 0.803	85	0.703 0.653	84	0.712 0.649	6	0.750 0.375	70	0.758 0.782
Exp.2 & Test 1	ImageNet H&E	193	0.758 0.760	56	0.918 0.951	56	0.929 0.923	26	0.877 0.750	446	0.83 0.823
Exp.2 & Test 2	ImageNet H&E	193	0.755 0.735	277	0.654 0.559	272	0.529 0.69	26	0.920 0.879	446	0.828 0.833
Exp.3 & Test 1	ImageNet H&E	193	0.722 0.739	56	0.898 0.897	56	0.865 0.869	26	0.675 0.787	446	0.802 0.783
Exp.3 & Test 2	ImageNet H&E	193	0.721 0.715	277	0.601 0.553	272	0.538 0.639	26	0.759 0.735	446	0.811 0.801

Reminder notes:

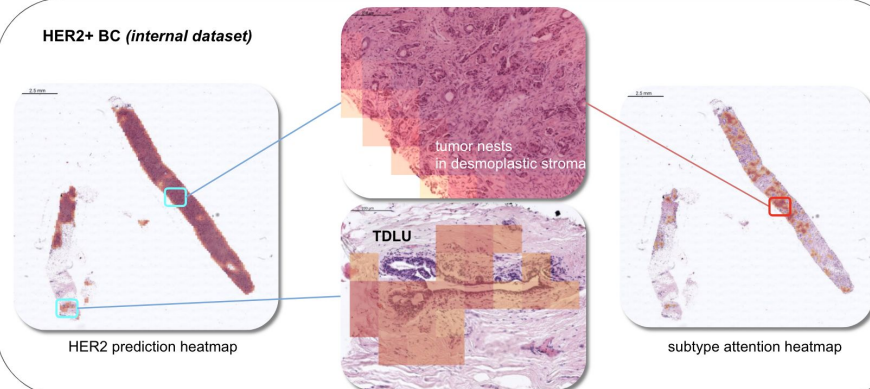
- Test 1 is with training images from all datasets
- Test 2 contains images from unseen datasets
- Experiment I: (her2+, hr+) vs. (her2+, hr-),
- Experiment II: (hr+, her2-) vs. (her2+, hr+/-) vs. tnbc,
- Experiment III: (hr+, her2-) vs. (her2+, hr+) vs. (her2+, hr-) vs. tnbc

Explainability

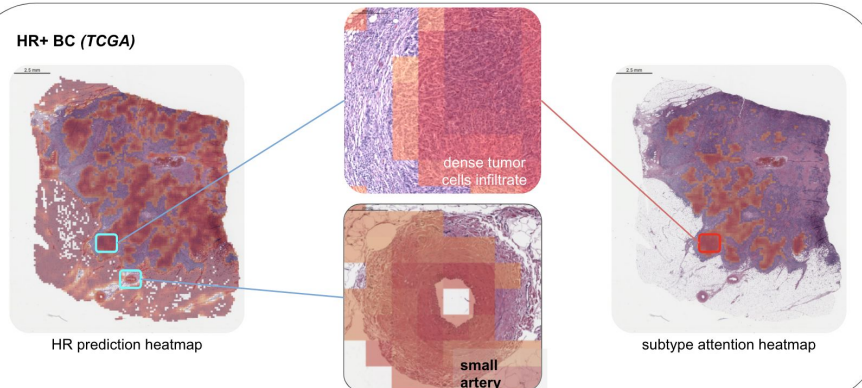
TNBC BC (TCGA)



HER2+ BC (internal dataset)



HR+ BC (TCGA)



- Attention heatmaps are created to get better insights to the model
- Pathologist's review suggests that the network focuses on biologically meaningful regions such as tumor epithelial cells, collagen-rich stroma and TILs aggregates
- The lymphocytic infiltration in some tumor nests and necrotic areas are informative for differentiating the ER- samples from ER+ ones within her2+ population (experiment I).

Conclusions

- H&E images contain biologically interpretable morphological patterns that are predictive of molecular subtypes
- Faster and less expensive prediction based on H&E compared to multiple IHCs
- Advantages of pre-training a specialized backbone on H&E WSI:
 - Transfer learning when lacking large and precisely annotated dataset
 - Optimal backbone for each dataset compared to pre-training on natural images
- The discriminatory morphological patterns for HR+ are harder to identify when both ER and Her2 are present:
 - Her2 overexpression is a dominant transformation mechanism in tumors.
- Combining with other risk factors and using for patient survival prediction

Doing now what patients need next