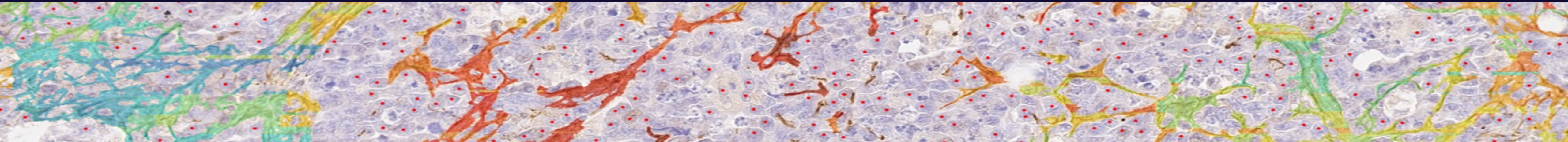
PS3: A Multimodal Transformer Integrating Pathology Reports with Histology Images and Biological Pathways for Cancer Survival Prediction

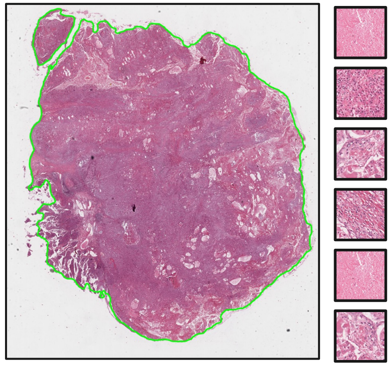
Manahil Raza, Ayesha Azam, Talha Qaiser and Nasir Rajpoot
Tissue Image Analytics Centre, University of Warwick, UK



Introduction

Survival Prediction | Introduction

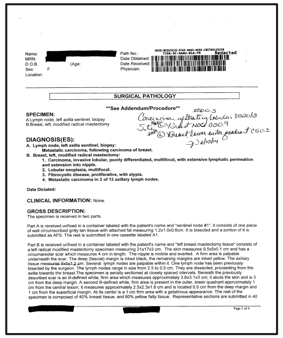
Whole Slide Images



WSI Patches > 1000

Genomic Data

Pathology Reports

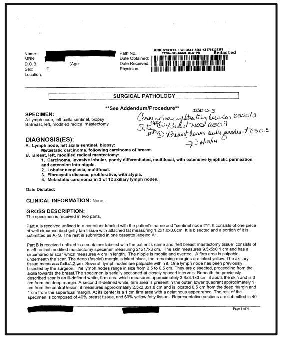


Text Segments < 100

Survival Prediction | Introduction

- Readily available – routine clinical workflows
- Expert assessments from pathologists
- Clinical and diagnostic details

Pathology Reports



Name: [redacted]
MRN:
D.O.B. (Age:
Sex: F
Location:

Path No.: UUID:0C023E1B-3FA3-4A43-AE6E-C8E7851252F8
TCGA-3C-AAAU-01A-PR Redacted
Date Obtained:
Date Received:
Physician:

SURGICAL PATHOLOGY

****See Addendum/Procedure****

ICD-O-3
Carcinoma, infiltrating lobular 8520/3
Site NOS (L) Breast NOS C50.9
path (L) Breast, lower outer quadrant C50.5
6/13/14

SPECIMEN:
A:Lymph node, left axilla sentinel, biopsy
B:Breast, left, modified radical mastectomy

DIAGNOSIS(ES):

A. Lymph node, left axilla sentinel, biopsy:
Metastatic carcinoma, following carcinoma of breast.

B. Breast, left, modified radical mastectomy:
1. **Carcinoma, invasive lobular, poorly differentiated, multifocal, with extensive lymphatic permeation and extension into nipple.**
2. **Lobular neoplasia, multifocal.**
3. **Fibrocystic disease, proliferative, with atypia.**
4. **Metastatic carcinoma in 3 of 12 axillary lymph nodes.**

Date Dictated:

CLINICAL INFORMATION: None.

GROSS DESCRIPTION:
The specimen is received in two parts.

Part A is received unfixed in a container labeled with the patient's name and "sentinel node #1". It consists of one piece of well circumscribed gritty tan tissue with attached fat measuring 1.2x1.0x0.6cm. It is bisected and a portion of it is submitted as AFS. The rest is submitted in one cassette labeled A1.

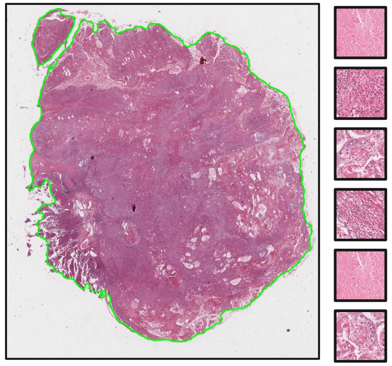
Part B is received unfixed in a container labeled with the patient's name and "left breast mastectomy tissue" consists of a left radical modified mastectomy specimen measuring 21x17x3 cm. The skin measures 9.5x5x0.1 cm and has a circumareolar scar which measures 4 cm in length. The nipple is mobile and everted. A firm area is palpable underneath the scar. The deep (fascial) margin is inked black, the remaining margins are inked yellow. The axillary tissue measures 9x5x1.2 cm. Several lymph nodes are palpable within it. One lymph node has been previously bisected by the surgeon. The lymph nodes range in size from 2.5 to 0.5 cm. They are dissected, proceeding from the axilla towards the breast.The specimen is serially sectioned at closely spaced intervals. Beneath the previously described scar is an ill-defined white, firm area which measures approximately 3.8x3.1x3 cm; it abuts the skin and is 3 cm from the deep margin. A second ill-defined white, firm area is present in the outer, lower quadrant approximately 1 cm from the central lesion; it measures approximately 2.5x2.3x1.8 cm and is located 0.5 cm from the deep margin and 1 cm from the superficial margin. At its center is a 1 cm firm area with a gelatinous appearance. The rest of the specimen is composed of 40% breast tissue, and 60% yellow fatty tissue. Representative sections are submitted in 40

[redacted] [redacted] Page 1 of 4

Text Segments < 100

Survival Prediction | Introduction

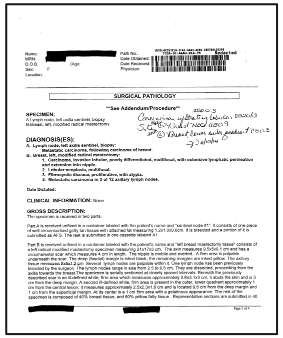
Whole Slide Images



WSI Patches > 1000

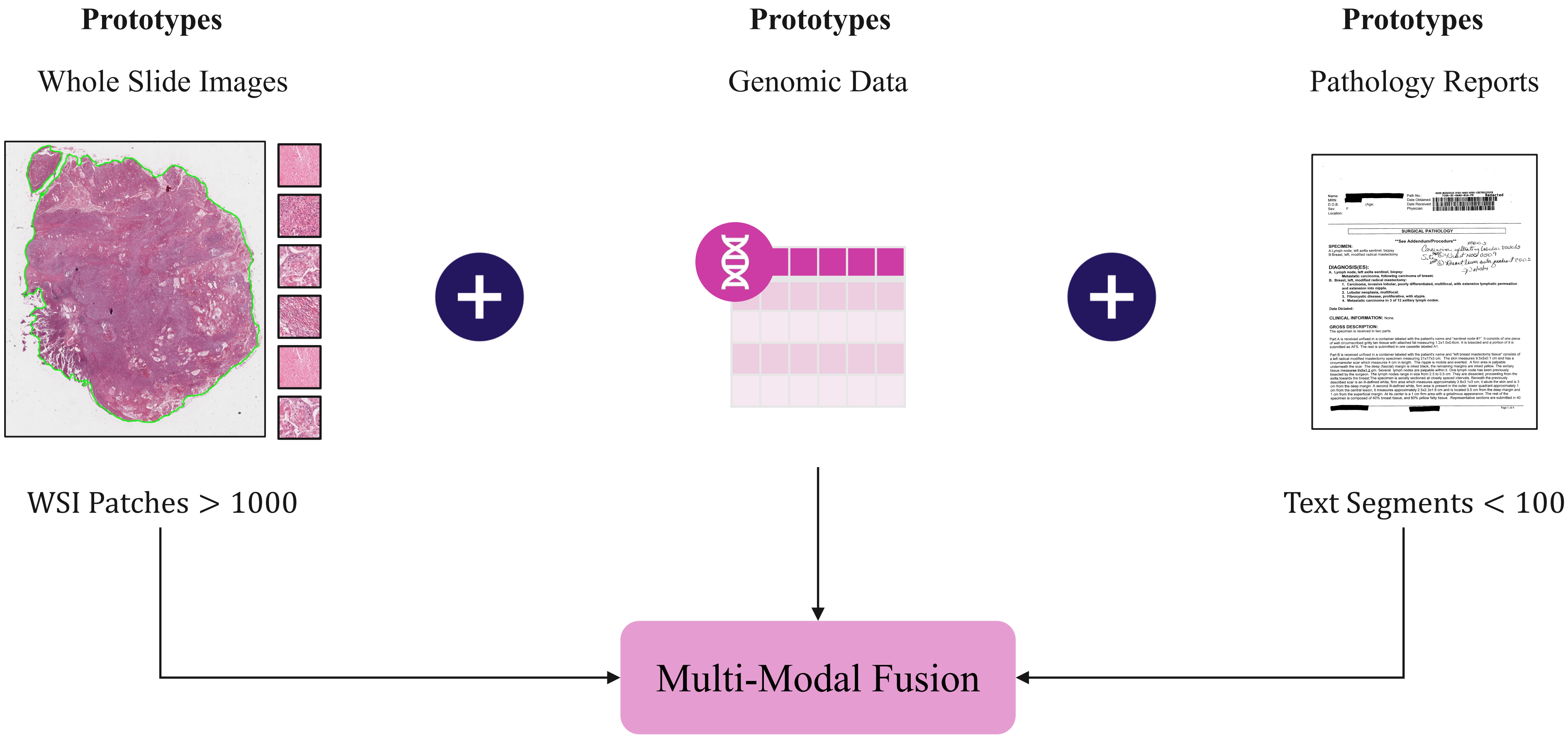
Genomic Data

Pathology Reports



Text Segments < 100

Survival Prediction | Introduction



Prototypes

Whole Slide Images

WSI Patches > 1000

Prototypes

Genomic Data

Prototypes

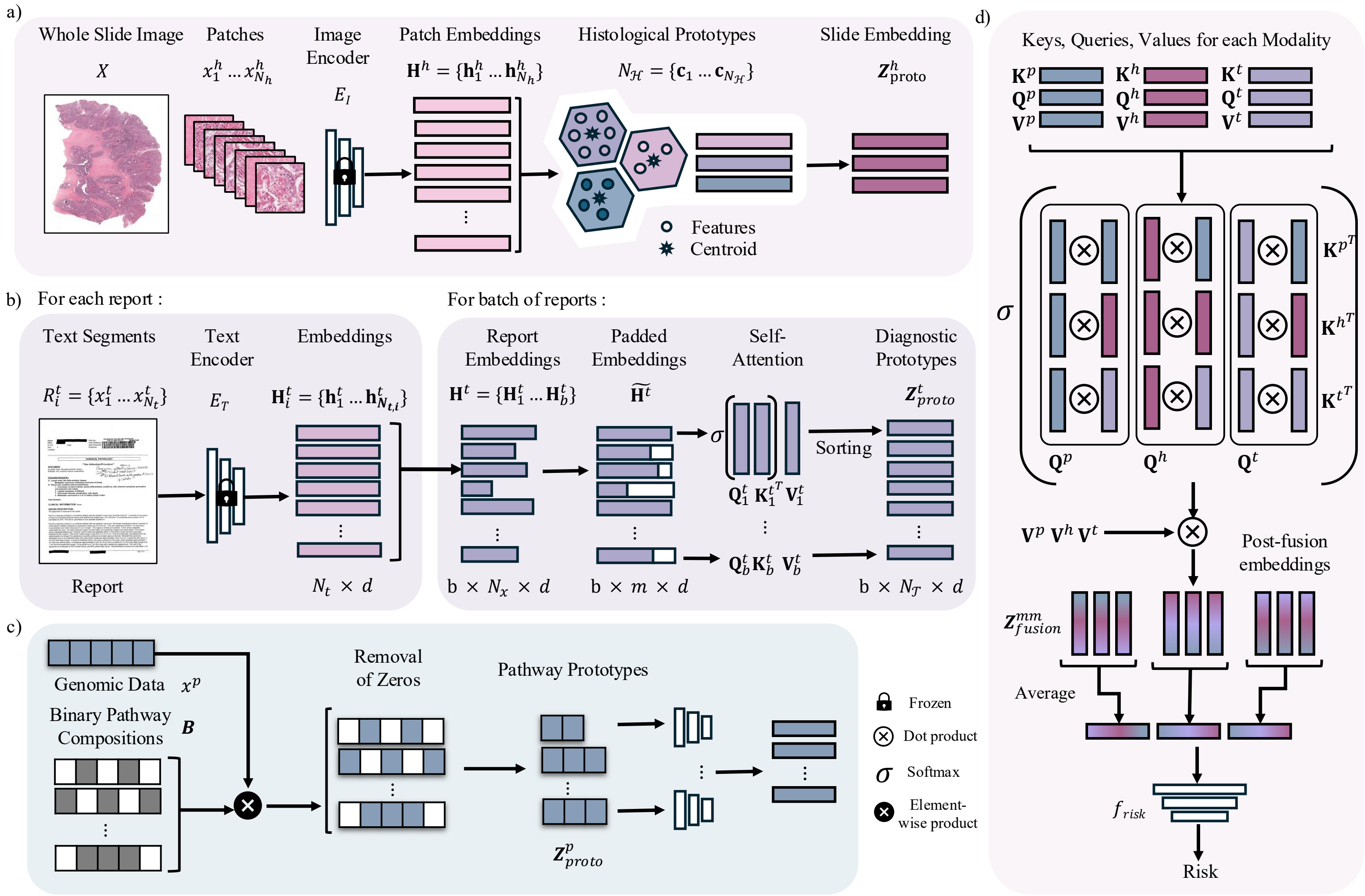
Pathology Reports

Text Segments < 100

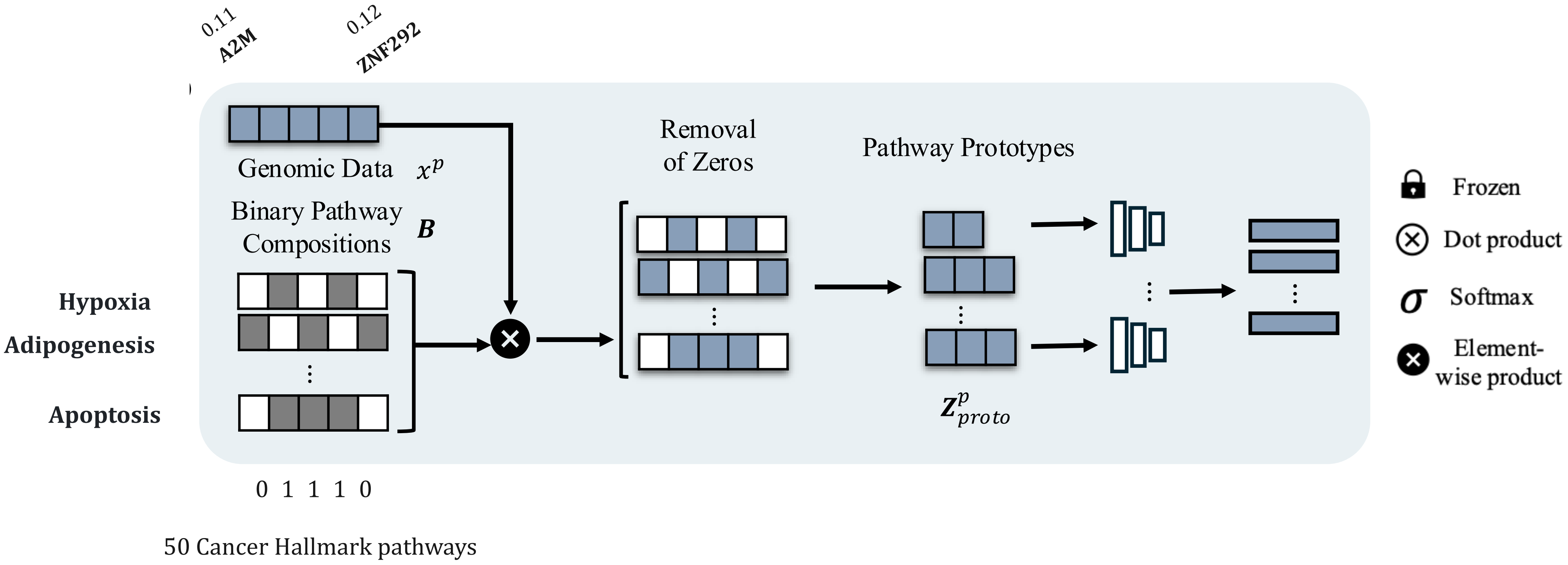
Multi-Modal Fusion

Methodology

Survival Prediction | Model Overview



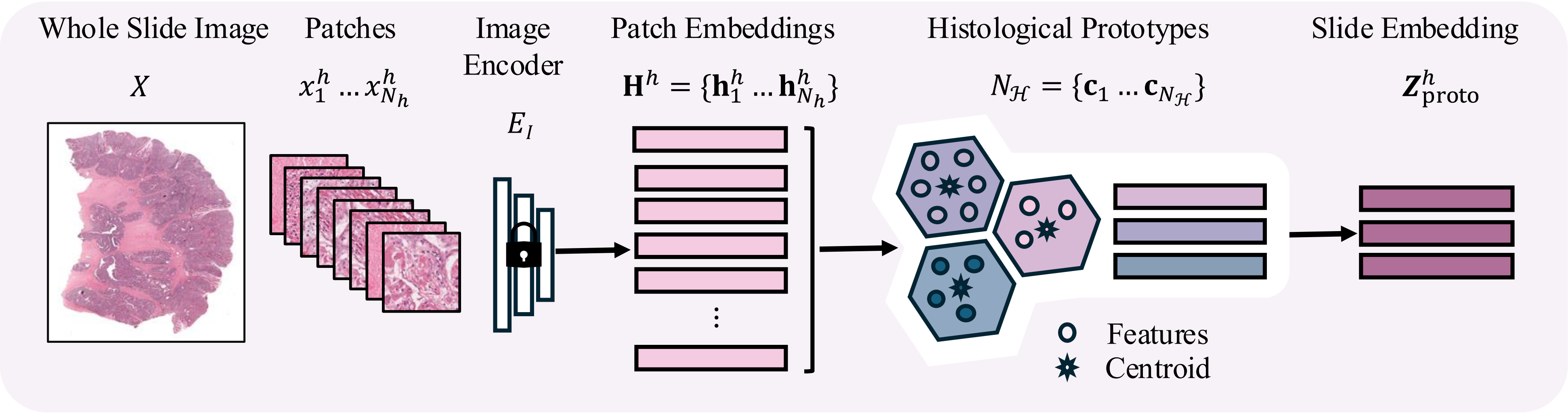
Survival Prediction | Pathway Prototypes



Liberzon, Arthur, et al. "The molecular signatures database hallmark gene set collection." *Cell systems* 1.6 (2015): 417-425.

Song, A. H., Chen, R. J., Jaume, G., Vaidya, A. J., Baras, A., & Mahmood, F. Multimodal Prototyping for cancer survival prediction. In *Forty-first International Conference on Machine Learning*.

Survival Prediction | Histological Prototypes



Whole Slide Image

X

Patches

$x_1^h \dots x_{N_h}^h$

Image Encoder

E_I

Patch Embeddings

$\mathbf{H}^h = \{\mathbf{h}_1^h \dots \mathbf{h}_{N_h}^h\}$

Histological Prototypes

$N_{\mathcal{H}} = \{\mathbf{c}_1 \dots \mathbf{c}_{N_{\mathcal{H}}}\}$

○ Features

✷ Centroid

Slide Embedding

$\boldsymbol{Z}_{\text{proto}}^h$

🔒 Frozen

⊗ Dot product

σ Softmax

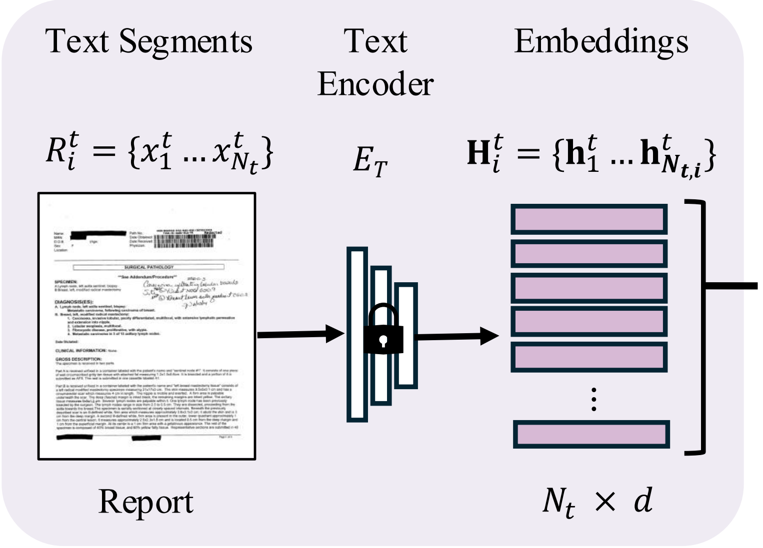
✖ Element-wise product

Song, Andrew H., et al. "Morphological prototyping for unsupervised slide representation learning in computational pathology." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.

Song, A. H., Chen, R. J., Jaume, G., Vaidya, A. J., Baras, A., & Mahmood, F. Multimodal Prototyping for cancer survival prediction. In *Forty-first International Conference on Machine Learning*.

Survival Prediction | Diagnostic Prototypes

For each report :



Text Segments

Text Encoder

Embeddings

$R_i^t = \{x_1^t \ldots x_{N_t}^t\}$

E_T

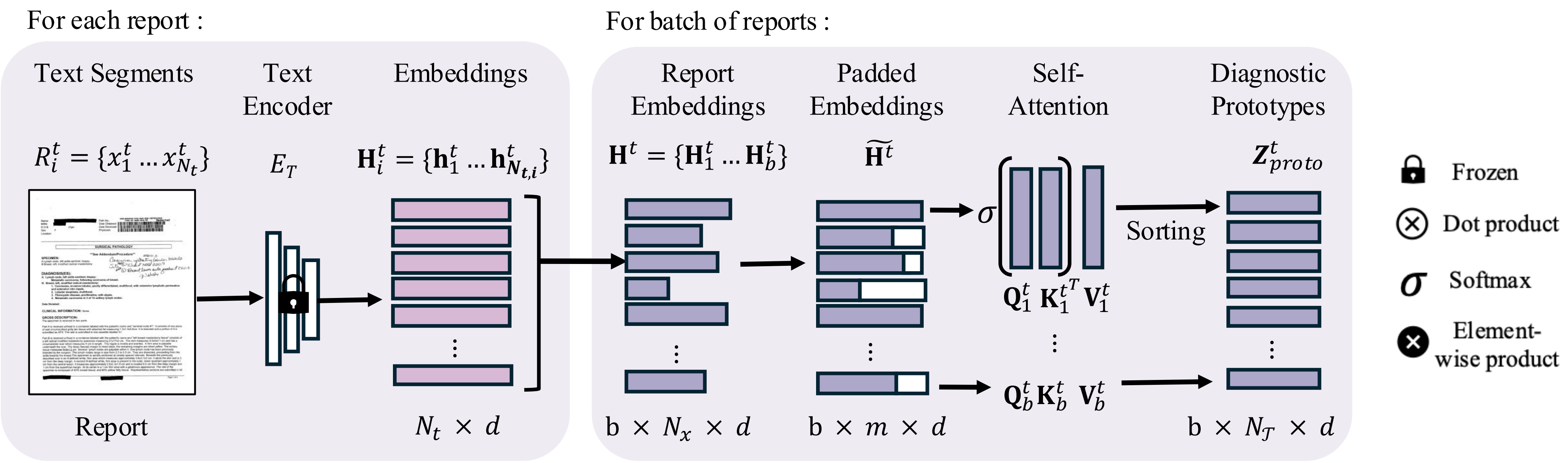
$\mathbf{H}_i^t = \{\mathbf{h}_1^t \ldots \mathbf{h}_{N_{t,i}}^t\}$

Report

$N_t \times d$

The number of text chunks = number of text embeddings, varies for each report

Survival Prediction | Diagnostic Prototypes

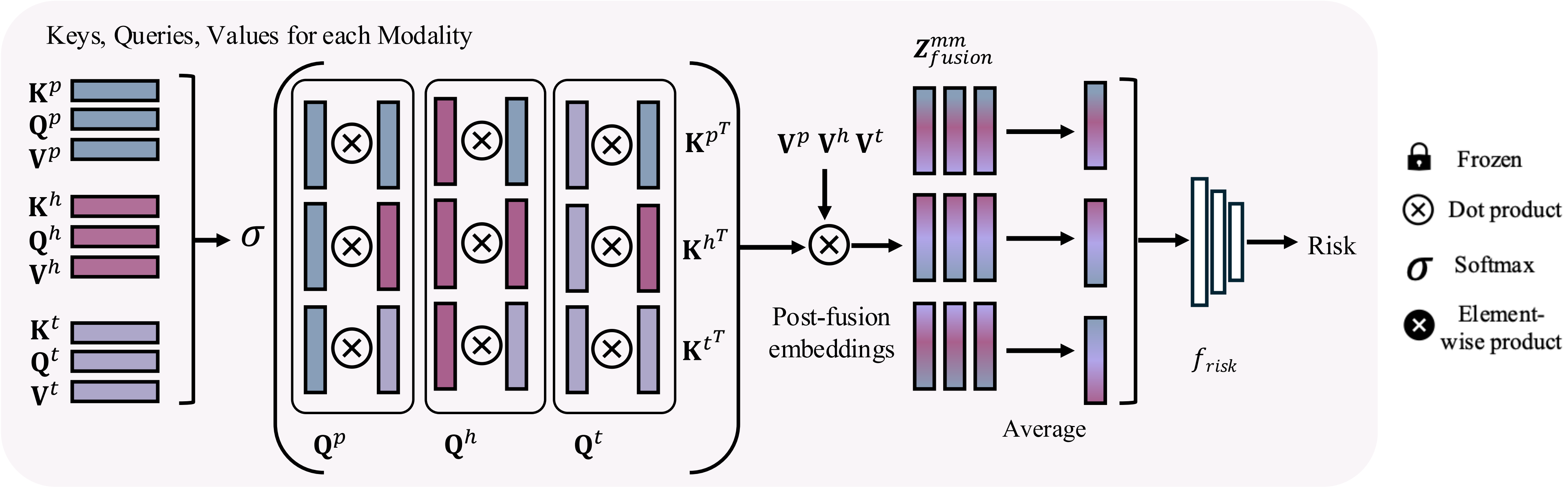


The number of text chunks = number of text embeddings, varies for each report

For each fold find:
m = max length
$N_{\mathcal{T}}$ = target length

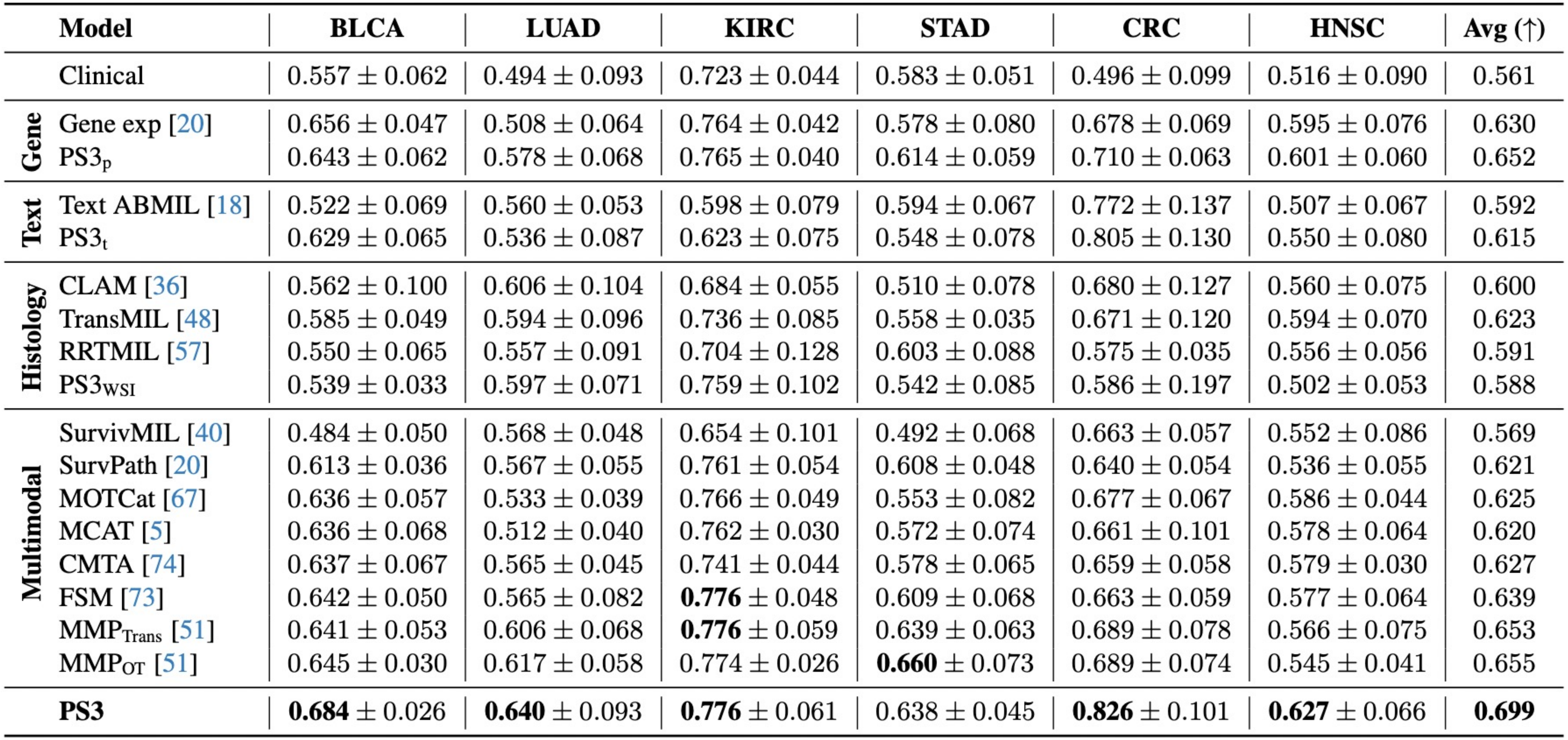
Based on the attention scores select the top $N_{\mathcal{T}}$ embeddings

Survival Prediction | Multimodal Transformer



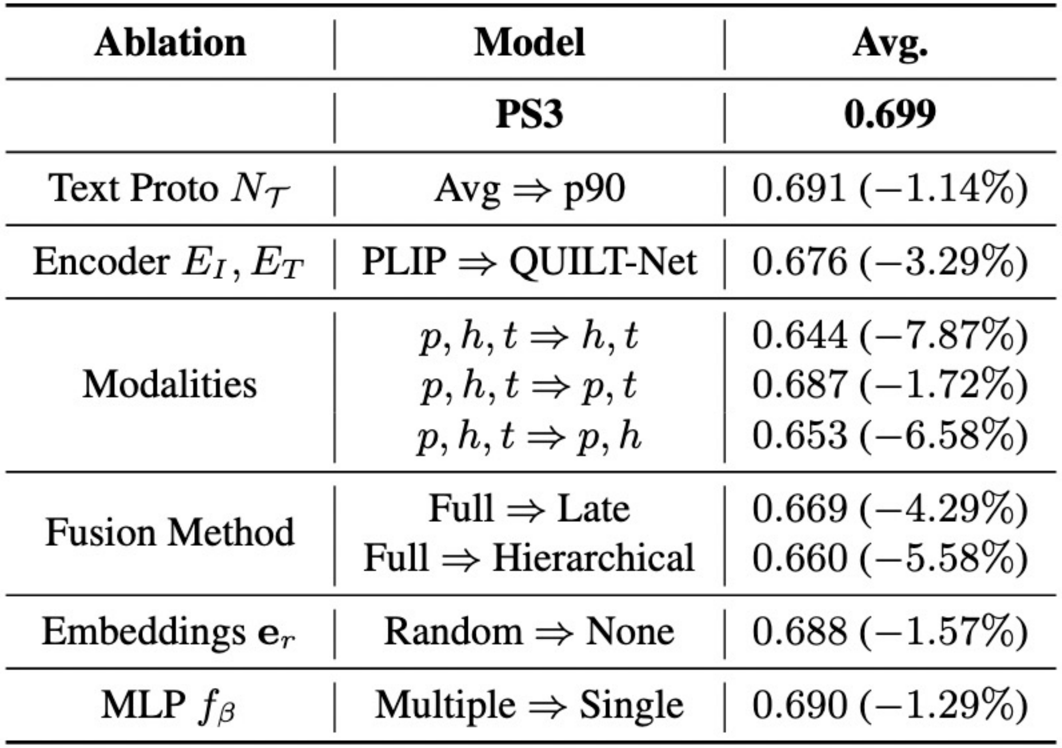
Results

Survival Prediction | C-Index Results

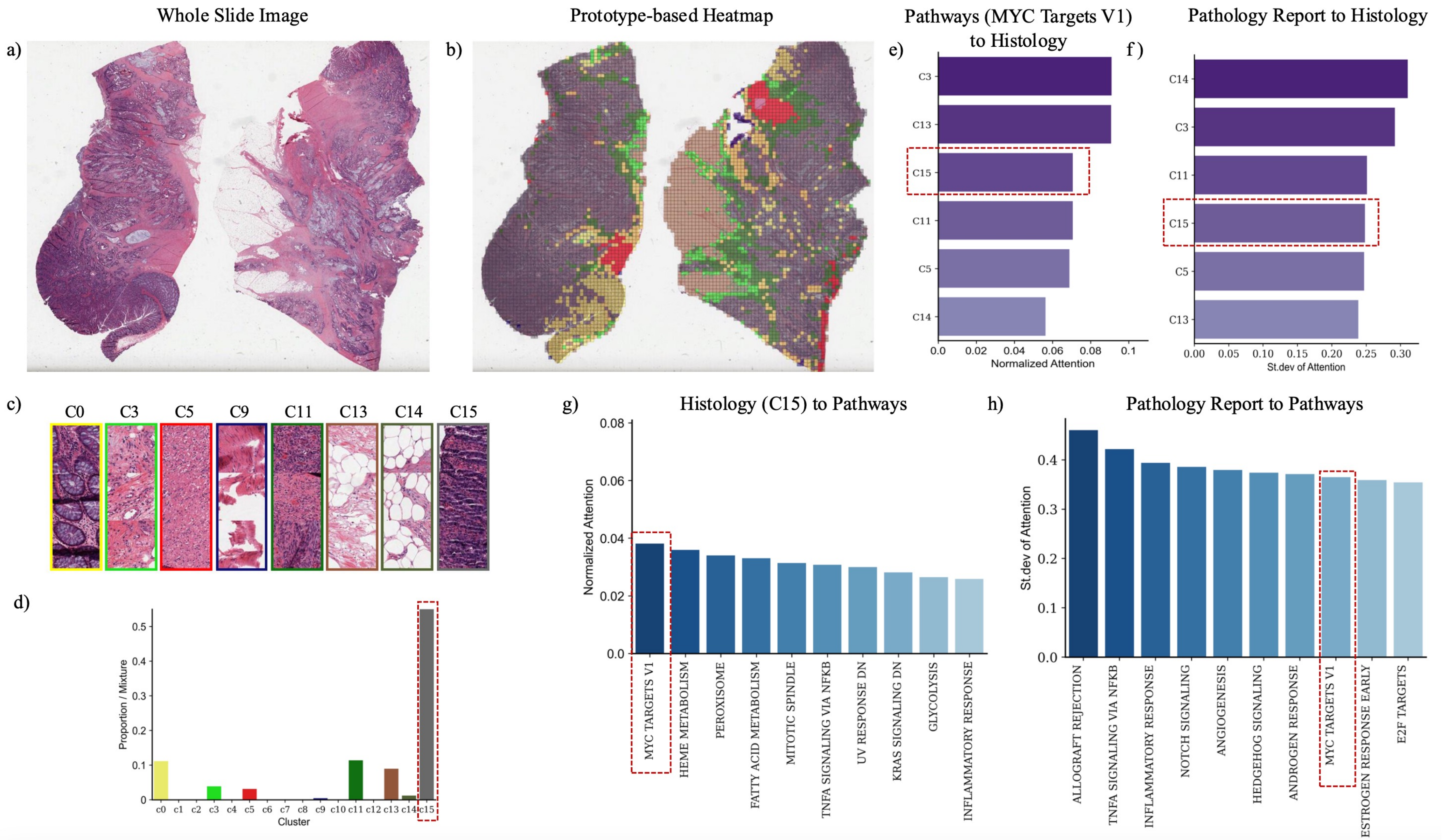


	Model	BLCA	LUAD	KIRC	STAD	CRC	HNSC	Avg (↑)
	Clinical	0.557 ± 0.062	0.494 ± 0.093	0.723 ± 0.044	0.583 ± 0.051	0.496 ± 0.099	0.516 ± 0.090	0.561
Gene	Gene exp [20]	0.656 ± 0.047	0.508 ± 0.064	0.764 ± 0.042	0.578 ± 0.080	0.678 ± 0.069	0.595 ± 0.076	0.630
	PS3$_p$	0.643 ± 0.062	0.578 ± 0.068	0.765 ± 0.040	0.614 ± 0.059	0.710 ± 0.063	0.601 ± 0.060	0.652
Text	Text ABMIL [18]	0.522 ± 0.069	0.560 ± 0.053	0.598 ± 0.079	0.594 ± 0.067	0.772 ± 0.137	0.507 ± 0.067	0.592
	PS3$_t$	0.629 ± 0.065	0.536 ± 0.087	0.623 ± 0.075	0.548 ± 0.078	0.805 ± 0.130	0.550 ± 0.080	0.615
Histology	CLAM [36]	0.562 ± 0.100	0.606 ± 0.104	0.684 ± 0.055	0.510 ± 0.078	0.680 ± 0.127	0.560 ± 0.075	0.600
	TransMIL [48]	0.585 ± 0.049	0.594 ± 0.096	0.736 ± 0.085	0.558 ± 0.035	0.671 ± 0.120	0.594 ± 0.070	0.623
	RRTMIL [57]	0.550 ± 0.065	0.557 ± 0.091	0.704 ± 0.128	0.603 ± 0.088	0.575 ± 0.035	0.556 ± 0.056	0.591
	PS3$_{WSI}$	0.539 ± 0.033	0.597 ± 0.071	0.759 ± 0.102	0.542 ± 0.085	0.586 ± 0.197	0.502 ± 0.053	0.588
Multimodal	SurvivMIL [40]	0.484 ± 0.050	0.568 ± 0.048	0.654 ± 0.101	0.492 ± 0.068	0.663 ± 0.057	0.552 ± 0.086	0.569
	SurvPath [20]	0.613 ± 0.036	0.567 ± 0.055	0.761 ± 0.054	0.608 ± 0.048	0.640 ± 0.054	0.536 ± 0.055	0.621
	MOTCat [67]	0.636 ± 0.057	0.533 ± 0.039	0.766 ± 0.049	0.553 ± 0.082	0.677 ± 0.067	0.586 ± 0.044	0.625
	MCAT [5]	0.636 ± 0.068	0.512 ± 0.040	0.762 ± 0.030	0.572 ± 0.074	0.661 ± 0.101	0.578 ± 0.064	0.620
	CMTA [74]	0.637 ± 0.067	0.565 ± 0.045	0.741 ± 0.044	0.578 ± 0.065	0.659 ± 0.058	0.579 ± 0.030	0.627
	FSM [73]	0.642 ± 0.050	0.565 ± 0.082	**0.776** ± 0.048	0.609 ± 0.068	0.663 ± 0.059	0.577 ± 0.064	0.639
	MMP$_{Trans}$ [51]	0.641 ± 0.053	0.606 ± 0.068	**0.776** ± 0.059	0.639 ± 0.063	0.689 ± 0.078	0.566 ± 0.075	0.653
	MMP$_{OT}$ [51]	0.645 ± 0.030	0.617 ± 0.058	0.774 ± 0.026	**0.660** ± 0.073	0.689 ± 0.074	0.545 ± 0.041	0.655
	PS3	**0.684** ± 0.026	**0.640** ± 0.093	**0.776** ± 0.061	0.638 ± 0.045	**0.826** ± 0.101	**0.627** ± 0.066	**0.699**

Survival Prediction | Ablation Studies

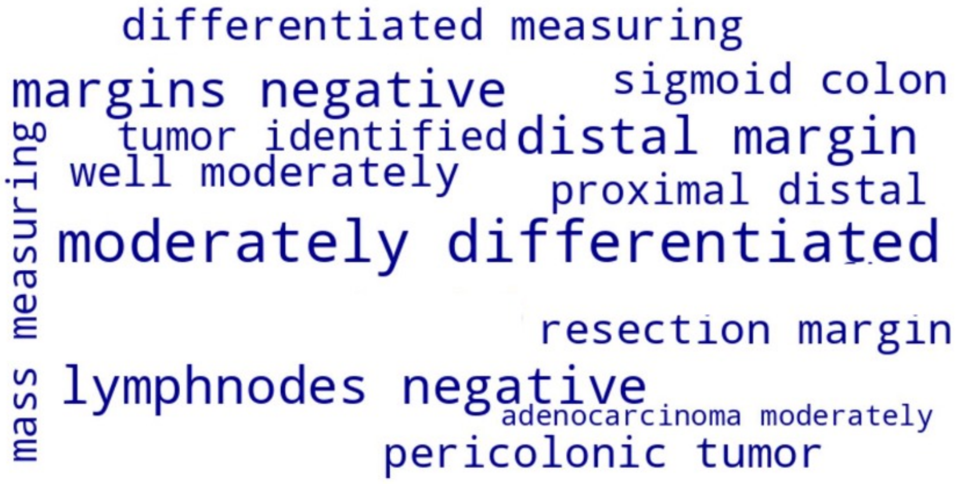


Ablation	**Model**	**Avg.**
	PS3	**0.699**
Text Proto $N_{\mathcal{T}}$	Avg $\Rightarrow$ p90	0.691 ($-$1.14%)
Encoder E_I, E_T	PLIP $\Rightarrow$ QUILT-Net	0.676 ($-$3.29%)
Modalities	$p, h, t \Rightarrow h, t$ $p, h, t \Rightarrow p, t$ $p, h, t \Rightarrow p, h$	0.644 ($-$7.87%) 0.687 ($-$1.72%) 0.653 ($-$6.58%)
Fusion Method	Full $\Rightarrow$ Late Full $\Rightarrow$ Hierarchical	0.669 ($-$4.29%) 0.660 ($-$5.58%)
Embeddings $\mathbf{e}_r$	Random $\Rightarrow$ None	0.688 ($-$1.57%)
MLP f_β	Multiple $\Rightarrow$ Single	0.690 ($-$1.29%)

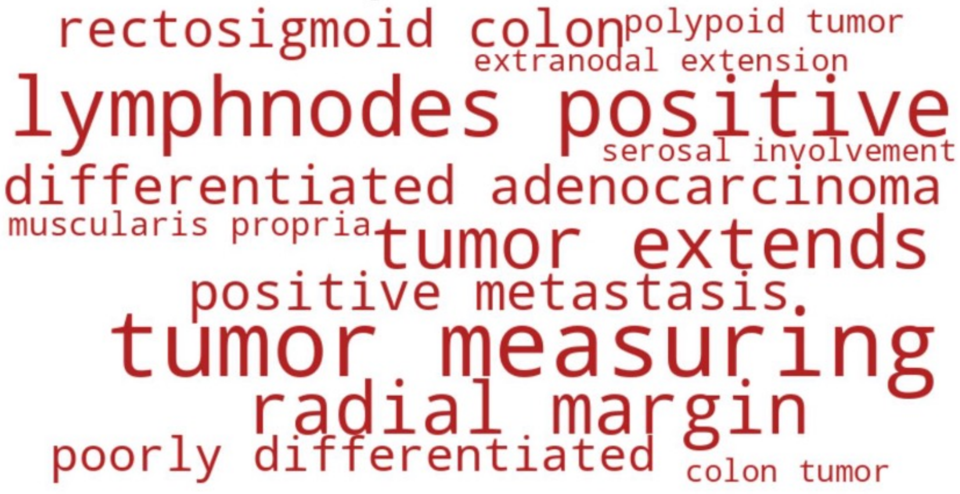

Whole Slide Image
a)
Prototype-based Heatmap
b)
Pathways (MYC Targets V1) to Histology
e)
C3
C13
C15
C11
C5
C14
0.0
0.02
0.04
0.06
0.08
0.1
Normalized Attention
Pathology Report to Histology
f)
C14
C3
C11
C15
C5
C13
0.00
0.05
0.10
0.15
0.20
0.25
0.30
St.dev of Attention
c)
C0
C3
C5
C9
C11
C13
C14
C15
d)
Proportion / Mixture
0
0.1
0.2
0.3
0.4
0.5
c0
c1
c2
c3
c4
c5
c6
c7
c8
c9
c10
c11
c12
c13
c14
c15
Cluster
g)
Histology (C15) to Pathways
Normalized Attention
0.0
0.02
0.04
0.06
0.08
MYC TARGETS V1
HEME METABOLISM
PEROXISOME
FATTY ACID METABOLISM
MITOTIC SPINDLE
TNFA SIGNALING VIA NFKB
UV RESPONSE DN
KRAS SIGNALING DN
GLYCOLYSIS
INFLAMMATORY RESPONSE
h)
Pathology Report to Pathways
St.dev of Attention
0.0
0.1
0.2
0.3
0.4
ALLOGRAFT REJECTION
TNFA SIGNALING VIA NFKB
INFLAMMATORY RESPONSE
NOTCH SIGNALING
ANGIOGENESIS
HEDGEHOG SIGNALING
ANDROGEN RESPONSE
MYC TARGETS V1
ESTROGEN RESPONSE EARLY
E2F TARGETS

Survival Prediction | Word Clouds

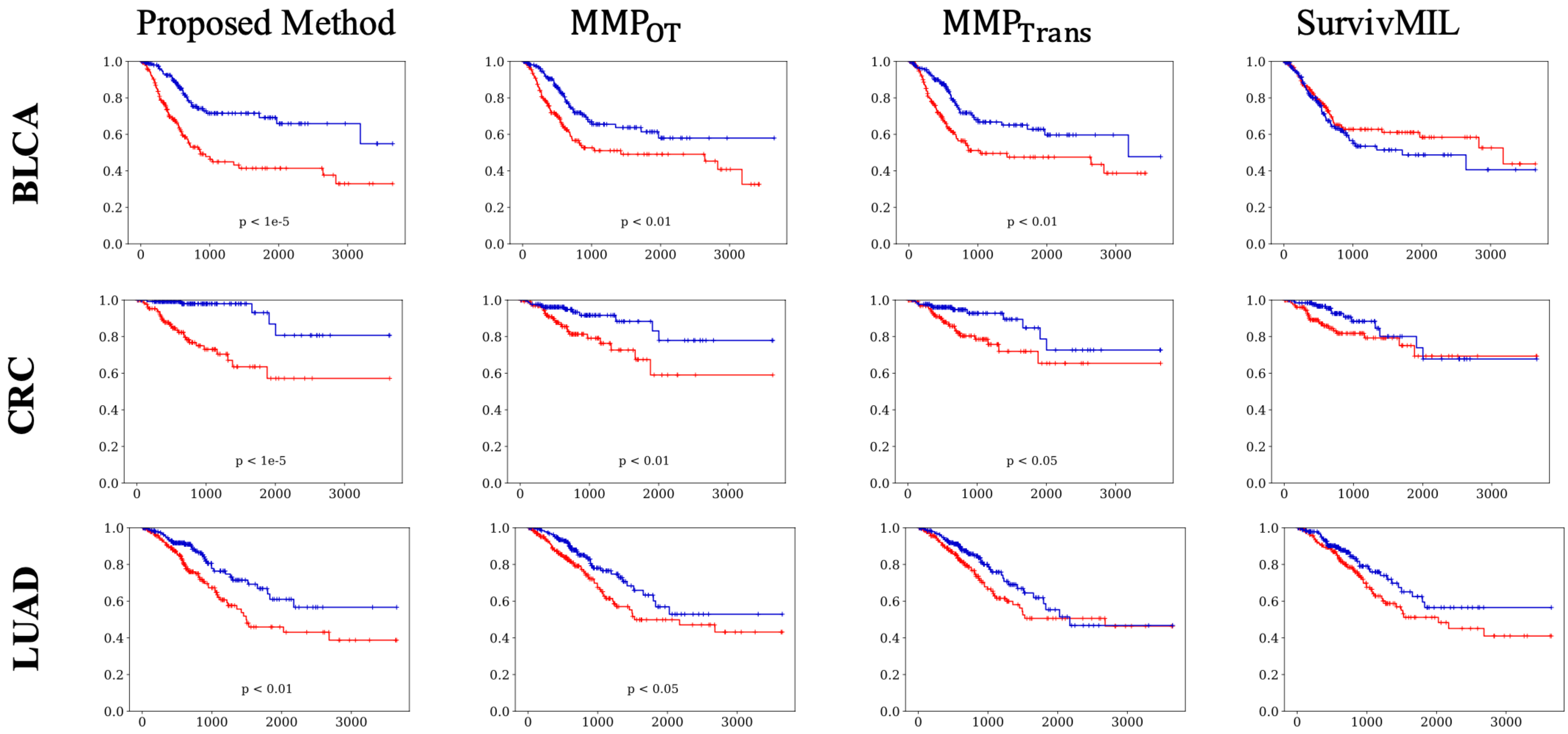
Low-risk word clouds



High-risk word clouds



Survival Prediction | Survival Curves



Thank You

TIA Centre

TISSUE IMAGE ANALYTICS