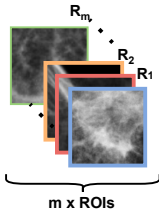


Detection
Module (θ)



top k

ViT (DINO)

ROBERTa

E_T

E_V

MAXPOOL

e_v^1
 e_v^2
 e_v^3
 \vdots
 e_v^k

K

Q

Cross-Attention

Projection

Projection

\otimes

SoftMax

Projection

V

E_A



\hat{y}

Clinical History

<Patient is K/C/of carcinoma
left breast status post CT>

Masking

