Mammo-CLIP: A Vision Language Foundation Model to Enhance Data Efficiency and Robustness in Mammography

Shantanu Ghosh[1], Clare B. Poynton[2], Shyam Visweswaran[3], Kayhan Batmanghelich[1]

[1]Dept. Of Electrical and Computer Engineering, Boston University

[2]Boston University Chobanian & Avedisian School of Medicine

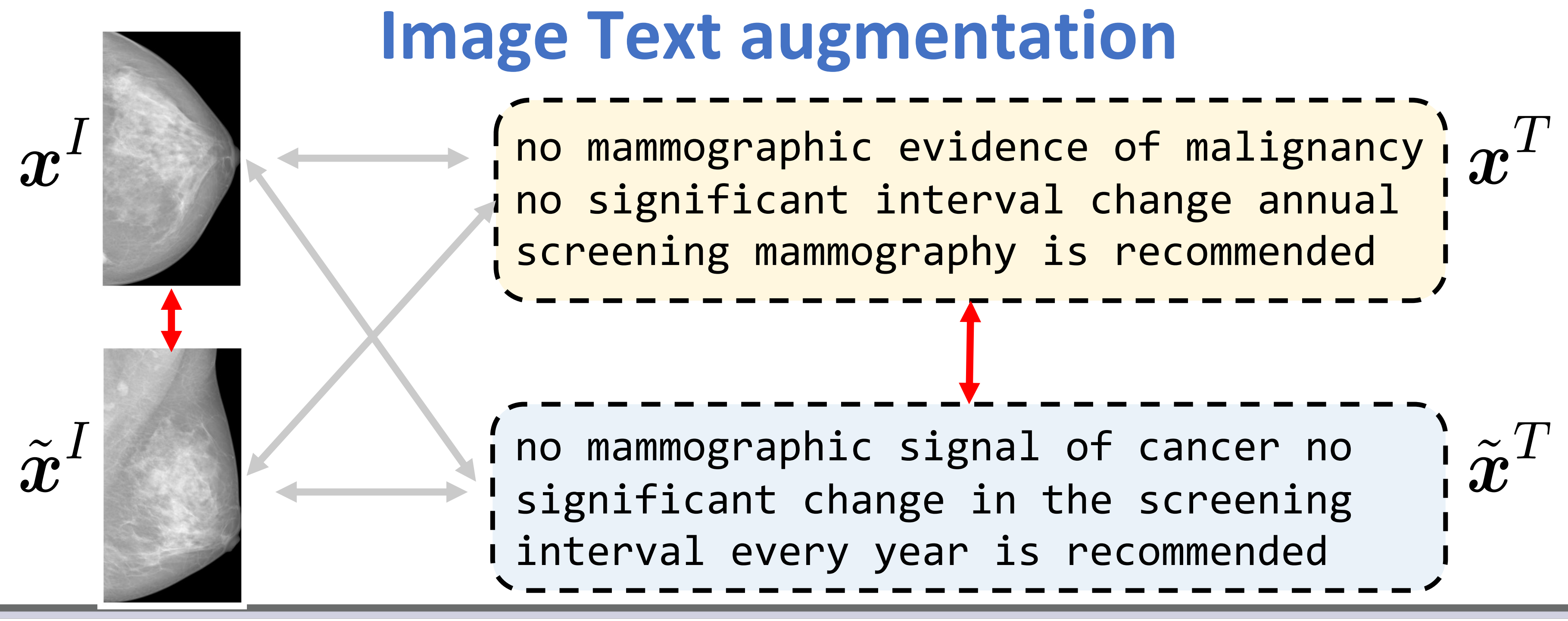
[3]Intelligent Systems Program (ISP), University of Pittsburgh

TLDR: A vision language model trained on both mammogram-report pairs and mammogram-attribute datasets, enhancing data efficiency, robustness, and interpretability

Motivation

- Scarcity of diverse, annotated mammogram datasets for effective CAD training.
- Vision-Language Models enhance robustness and data efficiency for medical imaging..
- Existing models lose critical diagnostic details due to reduced image resolution.
- Improving AI transparency with feature alignment between images and reports.

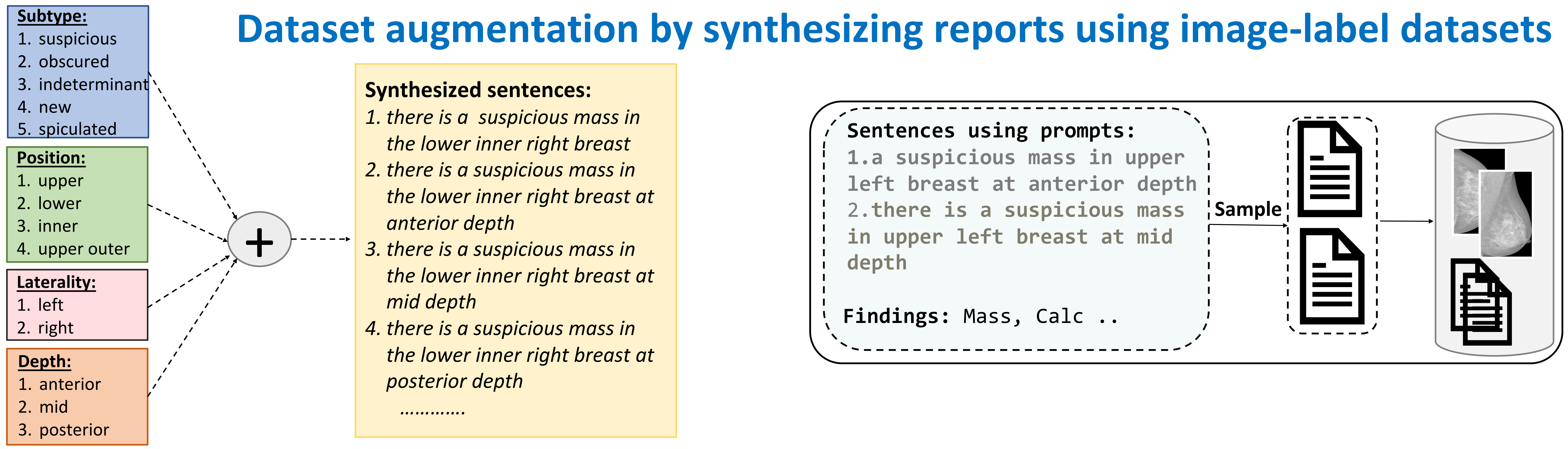
Image Text augmentation



no mammographic evidence of malignancy no significant interval change annual screening mammography is recommended

no mammographic signal of cancer no significant change in the screening interval every year is recommended

Dataset augmentation by synthesizing reports using image-label datasets



Subtype:
1. suspicious
2. obscured
3. indeterminant
4. new
5. spiculated

Position:
1. upper
2. lower
3. inner
4. upper outer

Laterality:
1. left
2. right

Depth:
1. anterior
2. mid
3. posterior

Synthesized sentences:
1. *there is a suspicious mass in the lower inner right breast*
2. *there is a suspicious mass in the lower inner right breast at anterior depth*
3. *there is a suspicious mass in the lower inner right breast at mid depth*
4. *there is a suspicious mass in the lower inner right breast at posterior depth*

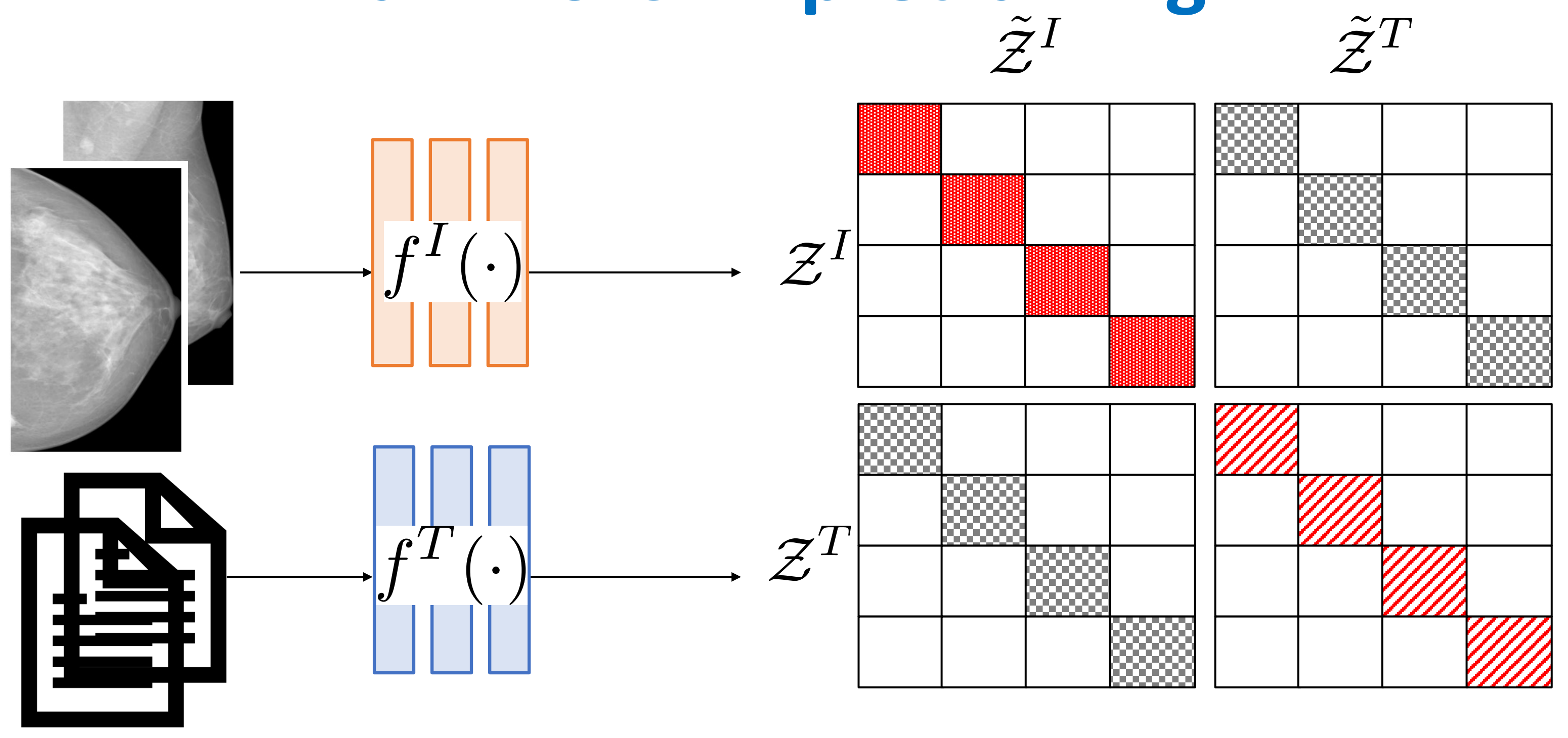
   ....

Sentences using prompts:
1.a suspicious mass in upper left breast at anterior depth
2.there is a suspicious mass in upper left breast at mid depth

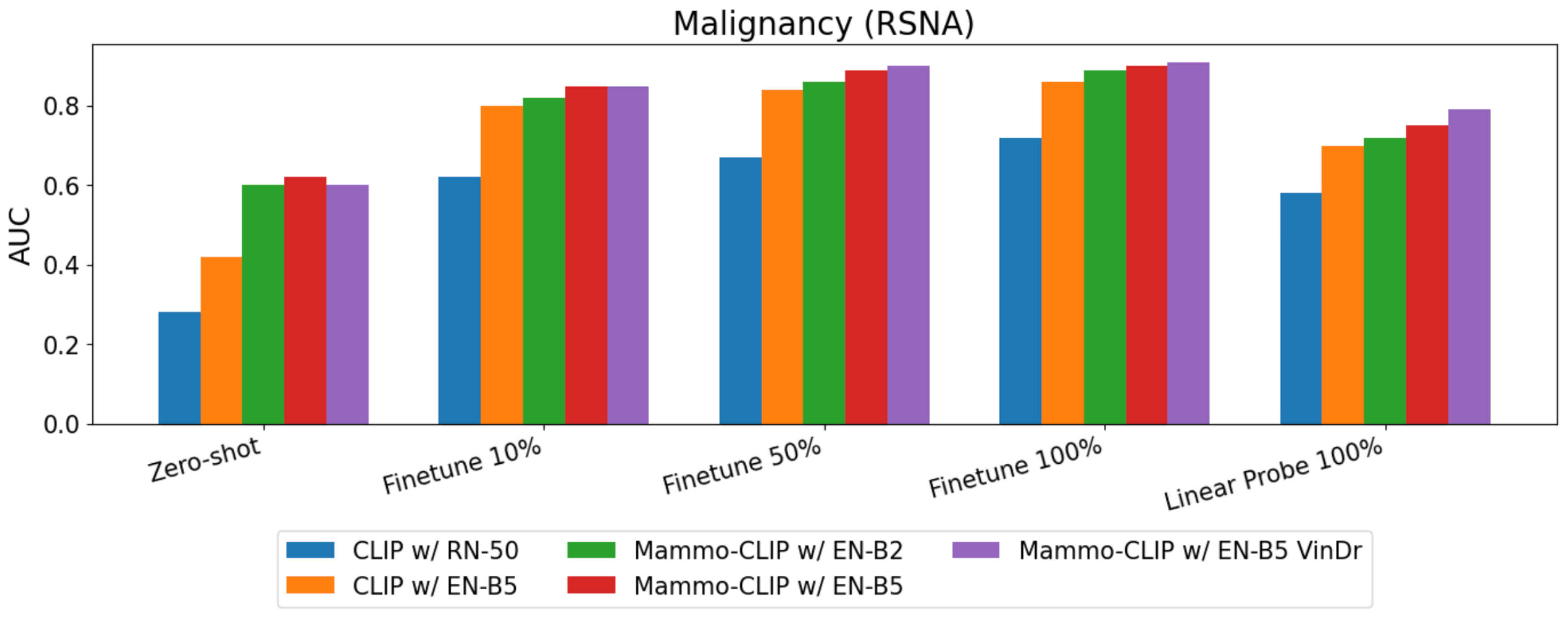
Findings: Mass, Calc ..

Sample

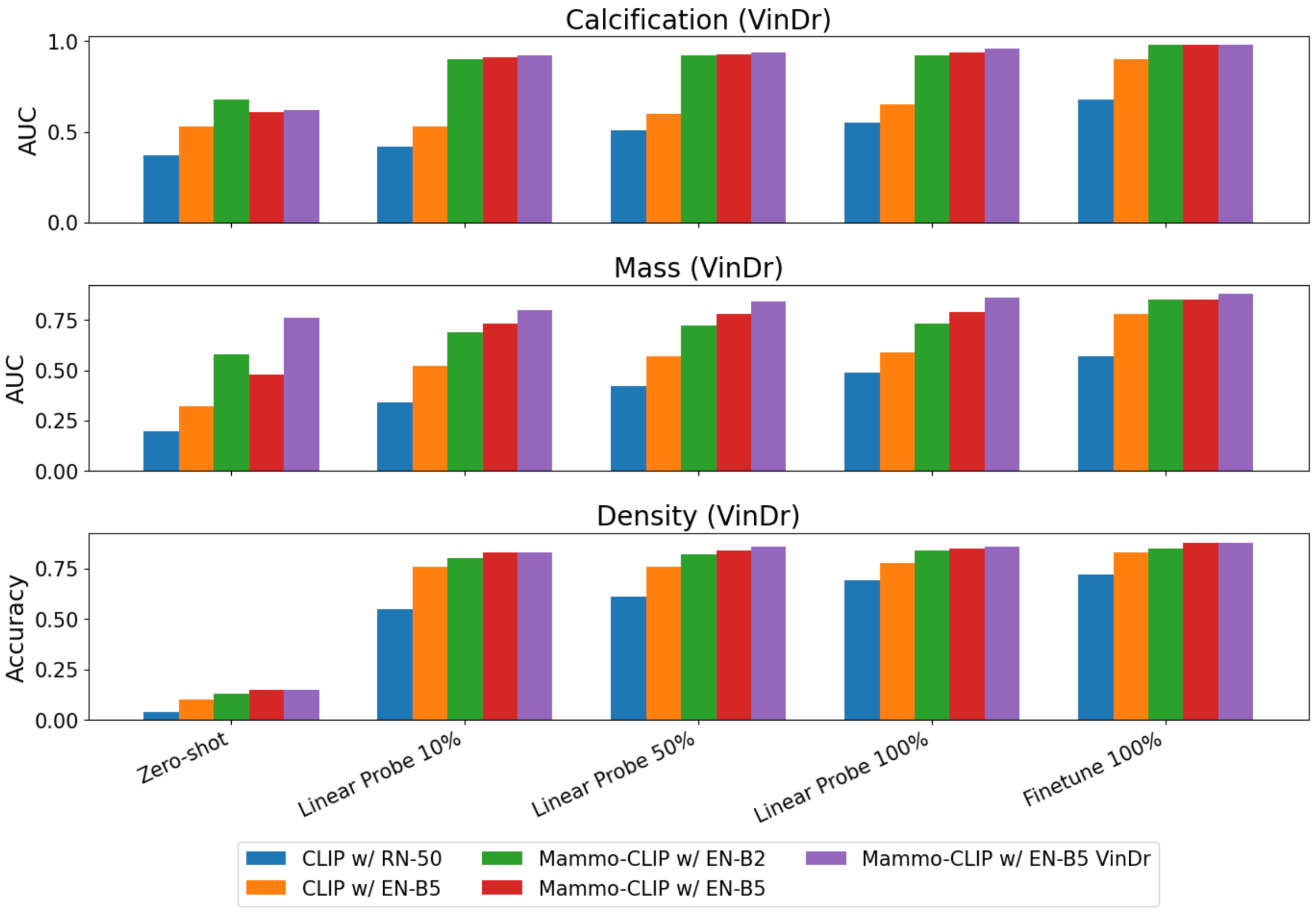
Mammo-CLIP pretraining



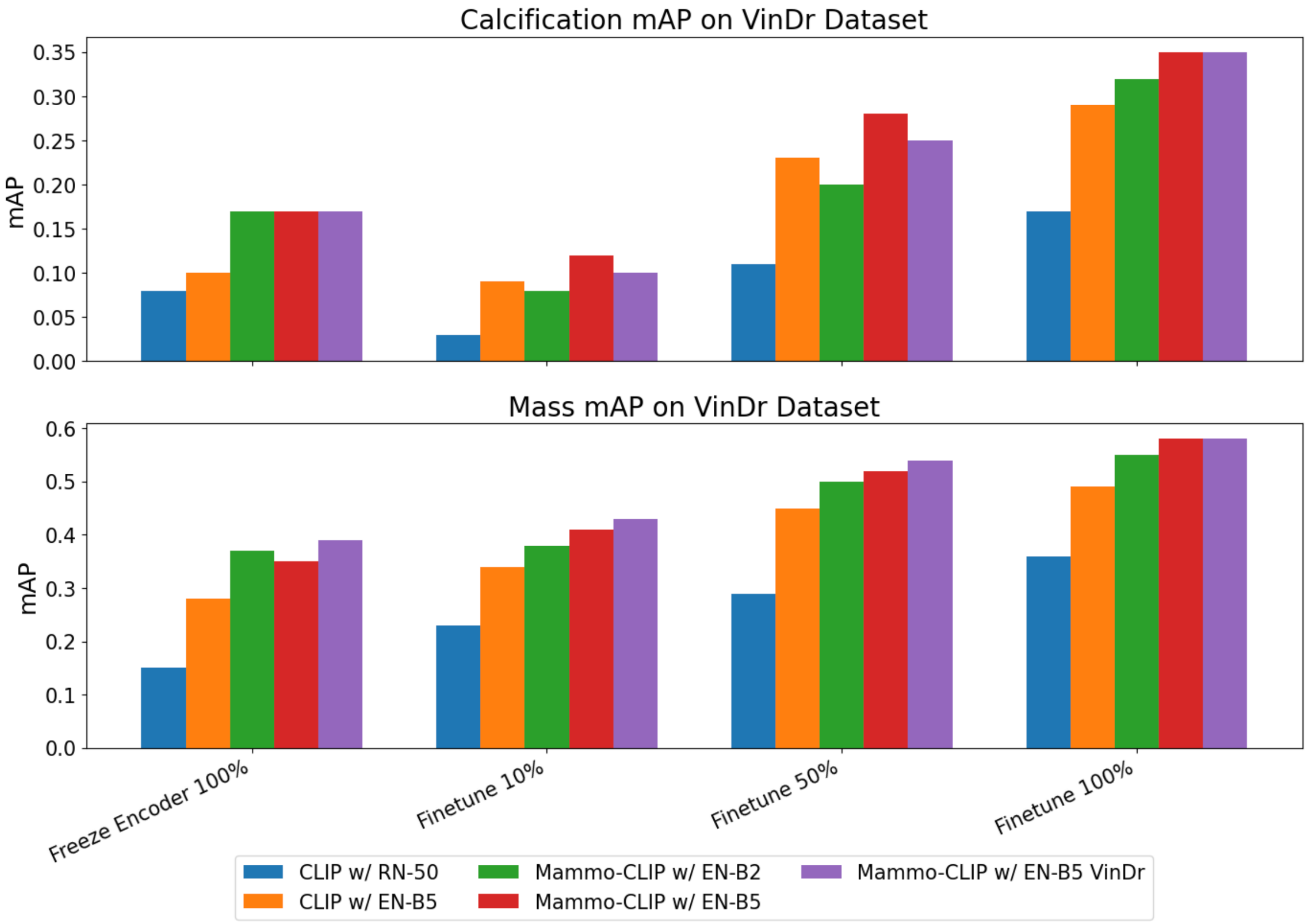
Cancer classification on RSNA



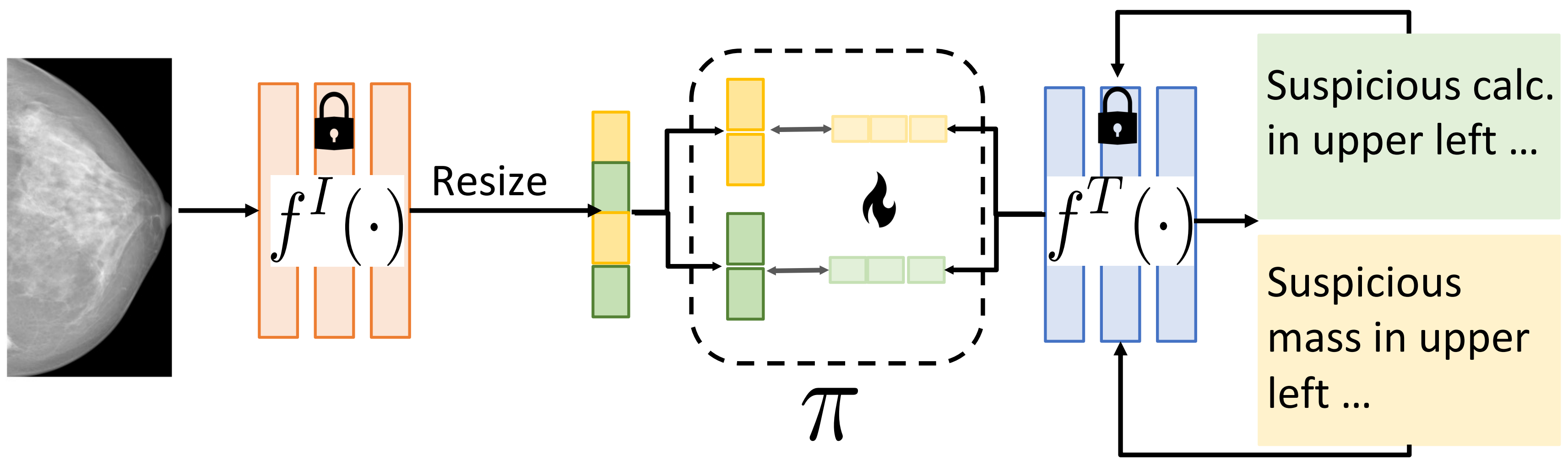
Findings classification on VinDr



Findings localization on VinDr



Mammo-FActOR



Suspicious calc. in upper left ...

Suspicious mass in upper left ...

Mammo-FActOR localization