

Supplementary Material

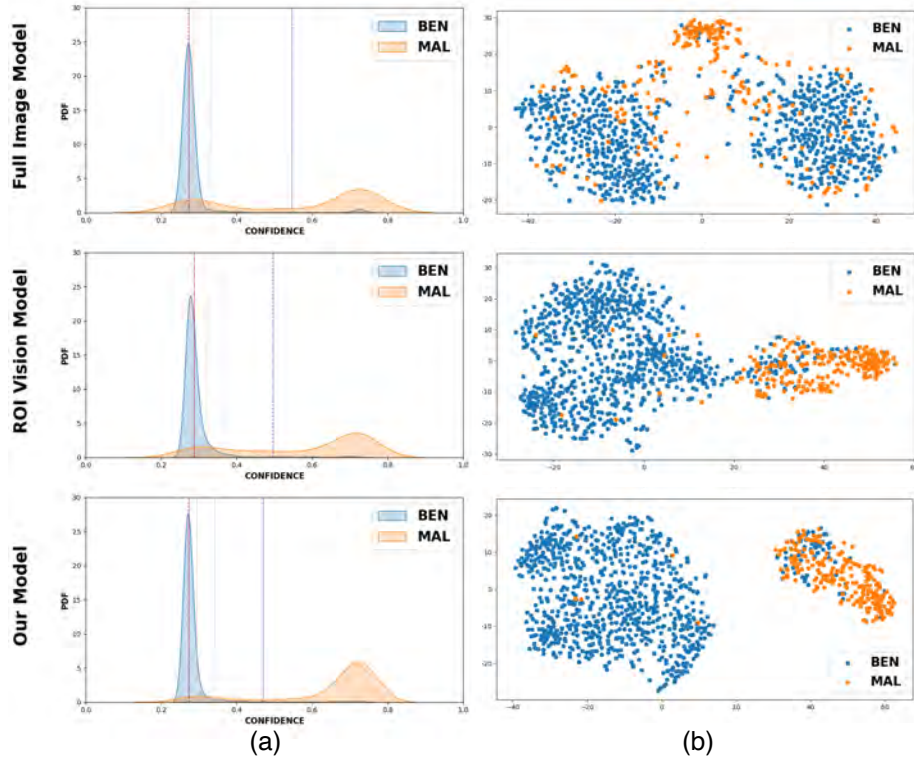


Fig. 1. The figure displays three different models in each row: (1) trained solely on whole mammograms, (2) one utilizing proposed ROIs, and (3) our proposed method. Column (a) presents confidence density plots, while the column (b) depicts t -SNE plots. Confidence plots feature dotted lines representing various FPI (False Positives per Image) levels (0.3, 0.1, 0.05, and 0.025) in decreasing order, demonstrating significant shifts in confidence distribution between full-image and ROI-based models. Moreover, the density plot of our proposed method displays heightened peaks, indicating enhanced confidence separation. In t -SNE plots, ROI-based modeling significantly improves inter-class distance in embeddings, further enhanced by the incorporation of clinical background information in our proposed method.

Table 1. Here we assess the impact of different strategies for ROI extraction on our model’s performance. We evaluate various object detection modules, including FocalNet, YOLO-v8, and Faster-RCNN. Results indicate that employing more sophisticated detection modules improves model performance, with extracted ROIs being more distinct and informative.

Model	Conference	AIIMS 1			AIIMS 2		
		Accuracy	F1-Score	AUC	Accuracy	F1-Score	AUC
Faster-RCNN [17] + OURS	TPAMI’17	0.95	0.76	0.961	0.96	0.66	0.934
YOLO-v8 [10] + OURS	Ultralytics’23	0.96	0.78	0.973	0.95	0.68	0.950
FocalNet-DINO [21] + OURS	NeurIPS’22	0.96	0.82	0.973	0.95	0.68	0.950

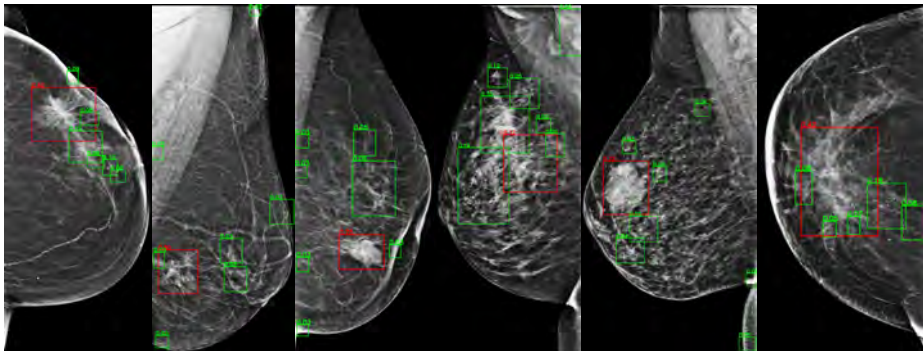


Fig. 2. The figure shows the ROI bounding boxes and the respective attention scores obtained from our proposed cross-attention layer for additional samples (some already included in the main paper). The ROI with the highest attention is drawn in red, and remaining in green.

Table 2. In the table we investigate the impact of varying the number of ROIs provided by the detection module on model’s performance. Increasing ROIs offers a broader context for learning but escalates computational demands, and can create more false positives. Conversely, reducing ROIs risks losing crucial regions. Our results in the table highlights the trade-off. Notably, utilizing 8 ROIs strikes the optimal balance between efficiency and performance.

#ROIs	AIIMS 1			AIIMS 2		
	Accuracy	F1-Score	AUC	Accuracy	F1-Score	AUC
2	0.96	0.77	0.962	0.96	0.71	0.939
8	0.96	0.82	0.973	0.95	0.68	0.950
16	0.96	0.80	0.973	0.96	0.68	0.939
25	0.96	0.81	0.973	0.95	0.66	0.943