

Supplementary Material

A Ablation Study

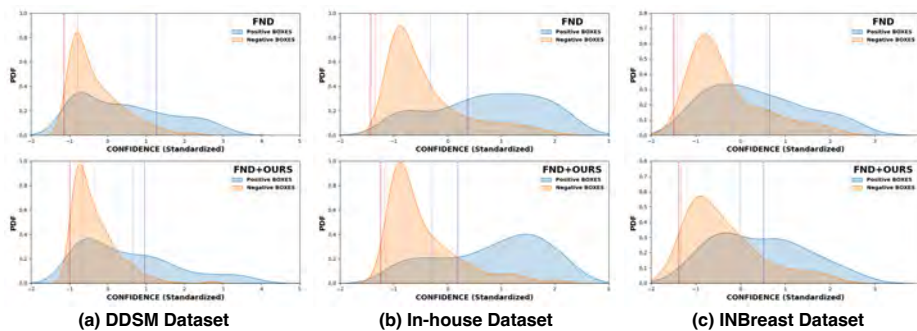


Fig. 1. Density plots compare our model’s performance across datasets, showing confidence distributions of bounding boxes from the baseline and our proposed method. Dotted lines denote various FPI levels (1, 0.5, 0.3, 0.1, 0.05, 0.025). Our proposed method exhibits superior separation between positive and negative bounding boxes compared to the baseline model. Notably, distribution drifts between classes are observed in both the DDSM and AIIMS datasets, with higher peaks in negative proposals, indicating increased model confidence in classifying malignancy. In the INBreast dataset, the distribution of negative samples is visibly shifted left, indicating improved model confidence.

Table 1. In this ablation study we explore incorporating a different cross-attention mechanism than the one proposed in the main paper. Here, we calculate the MCS through cross-attention between MLO and CC proposals. Confidence updates were determined based on attention weights between proposals. This module achieved comparable performance to our original architecture. This underscores the generalization and robustness of our approach.

FPI	0.3	0.5	0.1
OURS	0.93	0.95	0.96
OURS+Attention	0.92	0.94	0.96

Table 2. Ablation study results for the proposed CEN model utilizing different texture cues and excluding location information. Table contains detection and classification results on our AIIMS dataset. Firstly, we compare texture features [22] of proposals with learned embeddings. Results demonstrate that incorporating additional texture cues with CEN yields improved performance compared to single-view detection models, with further enhancement observed when integrating learned embeddings. Secondly, we evaluate the impact of removing location information and utilizing only proposal size. Findings suggest minimal contribution of location information to detection performance, aligning with clinical observations.

Model Name	Detection Results						Classification Results		
	R@0.025	R@0.05	R@0.1	R@0.3	R@0.5	R@1	Accuracy	F1-Score	AUC-Score
FND + texture[22]	0.69	0.79	0.86	0.90	0.93	0.95	0.958	0.752	0.924
FND + OURS(only size)	0.69	0.78	0.86	0.92	0.95	0.96	0.956	0.754	0.975
FND + OURS	0.70	0.80	0.86	0.93	0.95	0.96	0.958	0.747	0.976

Table 3. Comparison table with detection and classification metrics illustrating portability of our proposed method on AIIMS dataset. We replace the FND detection model with other prominent detection methods. These models include two DETR based models [9,10], as well as a CNN-based single-shot detection model YOLO-V8[6]. We compare their performance after adding our CEN model. These enhancements and comparative analyses collectively strengthen the robustness of our proposed approach and its adaptability to diverse detection models.

Model Name	Venue	Detection Results						Classification Results			
		R@0.025	R@0.05	R@0.1	R@0.3	R@0.5	R@1	Accuracy	F1-Score	AUC-Score	
DDSM	DAB Def.[10]	ICLR'22	0.21	0.28	0.4	0.54	0.58	0.63	0.728	0.638	0.772
	DAB Def + OURS		0.24	0.3	0.42	0.57	0.61	0.65	0.719	0.670	0.789
	DN DEF.[9]	CVPR'22	0.32	0.37	0.4	0.46	0.48	0.53	0.756	0.675	0.803
	DN DEF. + OURS		0.33	0.36	0.42	0.48	0.52	0.57	0.763	0.701	0.821
YOLO-V8	YOLO-V8[6]	Ultralytics'23	0.11	0.14	0.2	0.32	0.36	0.42	0.644	0.394	0.601
	YOLO-V8 + OURS		0.14	0.17	0.32	0.32	0.39	0.46	0.413	0.574	0.609
AIIMS	DAB DEF.[10]	ICLR'22	0.62	0.76	0.81	0.86	0.86	0.89	0.950	0.698	0.948
	DAB DEF. + OURS		0.64	0.78	0.86	0.9	0.9	0.92	0.947	0.729	0.961
	DN DEF.[9]	CVPR'22	0.64	0.72	0.74	0.78	0.81	0.82	0.952	0.680	0.949
	DN DEF. + OURS		0.67	0.77	0.8	0.84	0.84	0.85	0.956	0.732	0.948
YOLO-V8	YOLO-V8[6]	Ultralytics'23	0.19	0.27	0.38	0.51	0.61	0.71	0.916	0.212	0.779
	YOLO-V8 + OURS		0.3	0.4	0.48	0.61	0.67	0.74	0.904	0.489	0.812
INBreast	DAB DEF.[10]	ICLR'22	0.13	0.26	0.4	0.46	0.5	0.55	0.830	0.546	0.810
	DAB DEF. + OURS		0.13	0.28	0.36	0.47	0.53	0.55	0.832	0.577	0.822
	DN DEF.[9]	CVPR'22	0.17	0.26	0.39	0.44	0.47	0.51	0.825	0.526	0.772
	DN DEF. + OURS		0.19	0.28	0.38	0.48	0.51	0.54	0.798	0.566	0.794
YOLO-V8	YOLO-V8[6]	Ultralytics'23	0.02	0.03	0.06	0.08	0.1	0.12	0.258	0.346	0.541
	YOLO-V8 + OURS		0.04	0.06	0.07	0.07	0.09	0.18	0.470	0.356	0.579