Cross-view Correspondence Reasoning based on Bipartite Graph Convolutional Network for Mammogram Mass Detection (Supplementary Material)

Yuhang Liu[1] Fandong Zhang[2] Qianyi Zhang[1] Siwen Wang[1] Yizhou Wang[3] Yizhou Yu[1]
[1]Deepwise AI Lab [2] Center for Data Science, Peking University
[3] Center on Frontiers of Computing Studies, Dept. of Computer Science & Technology, Advanced Institute of Information Technology, Peking University
{liuyuhang, zhangqianyi, wangsiwen, yuyizhou}@deepwise.com
{fd.zhang, yizhou.wang}@pku.edu.cn

In this document, we provide additional materials to supplement our main submission. In the first section, we provide detailed mechanism of mammography screening. In the second section, we show additional detection results on DDSM dataset and in-house dataset, which demonstrates that our model outperforms our baseline methods.

1. Mammography Screening

Digital mammography is a specific type of low dose high-resolution X-ray imaging designed to detect breast cancer. Like all other X-ray imaging modalities, mammography displays the absorbance of ionizing radiation of the examined body part; in simplest terms, it measures the sum of atoms (or integral of tissue densities) along the projection ray. Due to the subjective nature of 3D to 2D X-ray projection, cancer tissue is easily overlooked and often obscured by normal breast tissue, especially in the case of dense breast type. To alleviate this, conventional mammography is taken from two standard angles: the cranial-caudal (CC) view is shot top-down and the mediolateral-oblique (MLO) view is shot from side at a certain angle. Yet, finding region-wise correspondence across the two views are still challenging for two reasons: tissue deformation under pressure and lack of unique visual cues. One can only roughly estimate the relationship based on the straight-line method, which assumes the perpendicular distance to nipple against the chest wall to be invariant under different projections. As illustrated in Figure 1, masses in Figure 1(a) and Figure 1(b) can satisfy the assumption perfectly, while Figure 1(c) and Figure 1(d) can not. The proposed method models cross-view geometric prior correspondences explicitly. Therefore, it provides stronger capacity than rule-based methods.

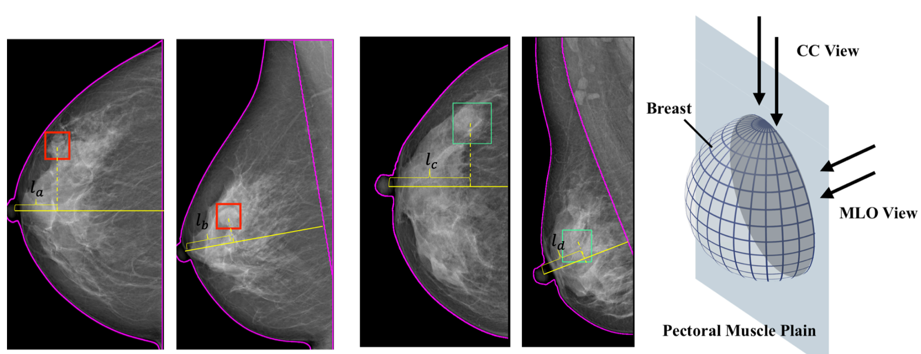


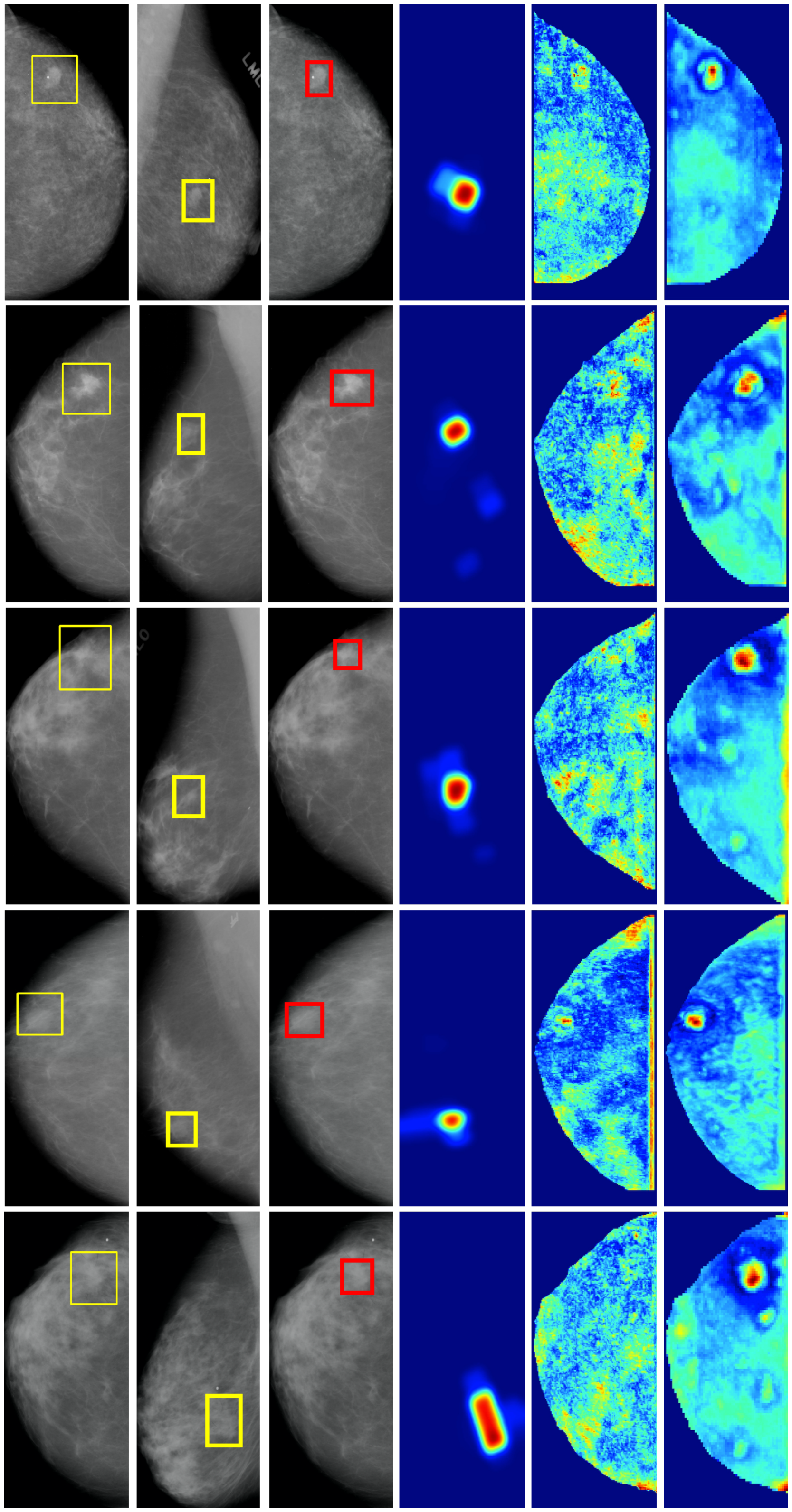
Figure 1. Relations between CC and MLO views. Figure (a)-(b) and (c)-(d) indicate two cases of CC and MLO views. The masses are annotated by the bounding boxes. Figure (e) is an ideal projection model. l_* indicates the perpendicular distance to nipple against the chest wall, where $* \in \{a, b, c, d\}$. Specially, the length l_a and l_b are roughly same, while l_c and l_d are not.

2. Additional Experimental Results

We next provide additional detection results and visualization analysis on both DDSM dataset and in-house dataset.

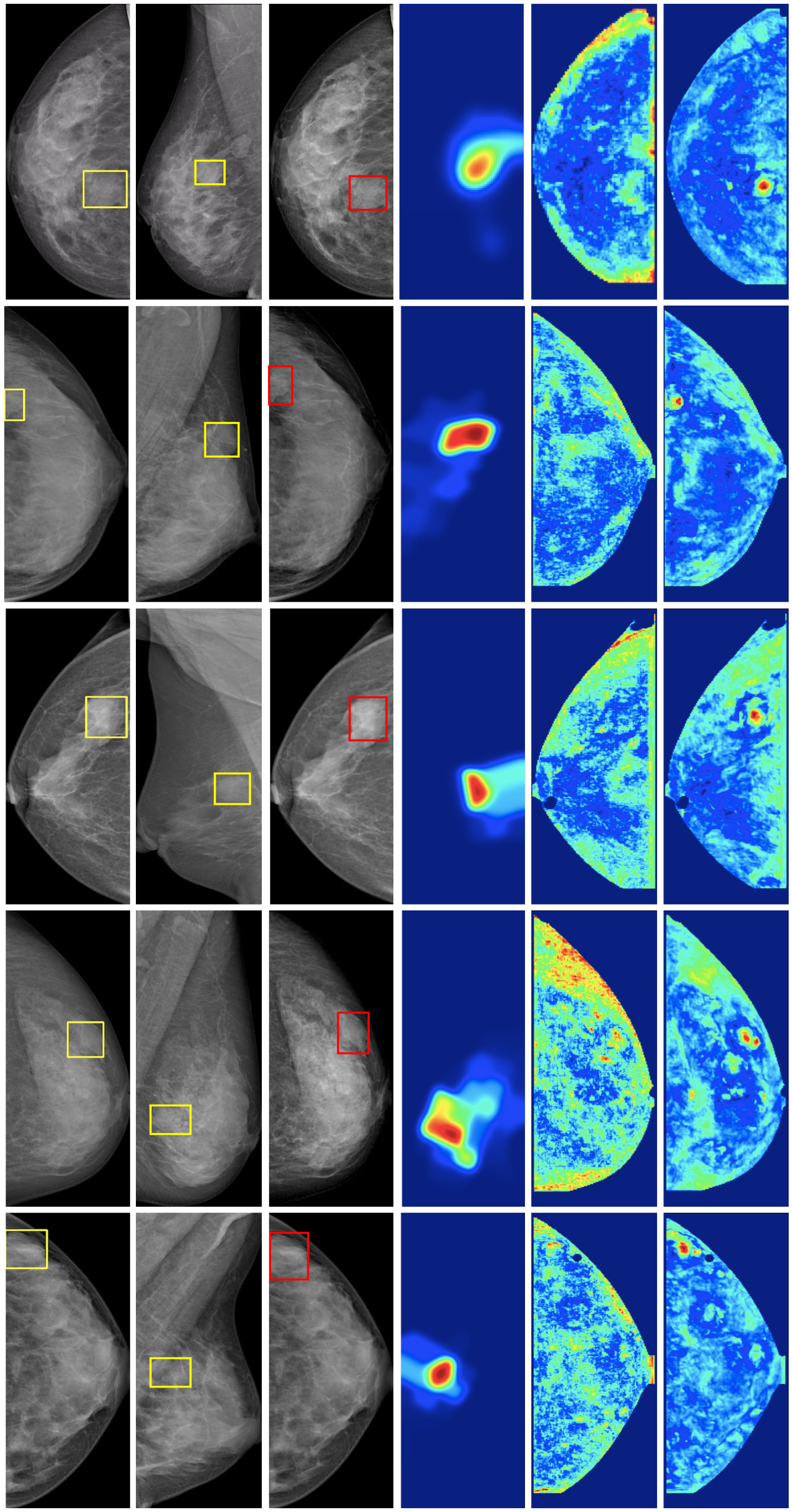
DDSM dataset: We first present additional results on DDSM in Figure 2. These results demonstrate that our model is ideally suited for cross-view mammogram mass detection.

In-house dataset: We show our results on in-house dataset in Figure 3 and Figure 4. Compared with DDSM dataset, the in-house dataset has higher image qualities.



(a) (b) (c) (d) (e) (f)

Figure 2. Additional results of the DDSM dataset. Each row shows a representative case. Column (a)-(b) refer to the examined image and its auxiliary view image with annotations. Column (c) indicates detection results by BG-RCNN. Column (e) visualizes the attention area on the auxiliary view. Column (f)-(g) visualize the response maps before and after correspondence visual reasoning.



(a) (b) (c) (d) (e) (f)

Figure 3. Additional results of the in-house dataset. Each row shows a representative case. Column (a)-(b) refer to the examined image and its auxiliary view image with annotations. Column (c) indicates detection results by BG-RCNN. Column (e) visualizes the attention area on the auxiliary view. Column (f)-(g) visualize the response maps before and after correspondence visual reasoning.

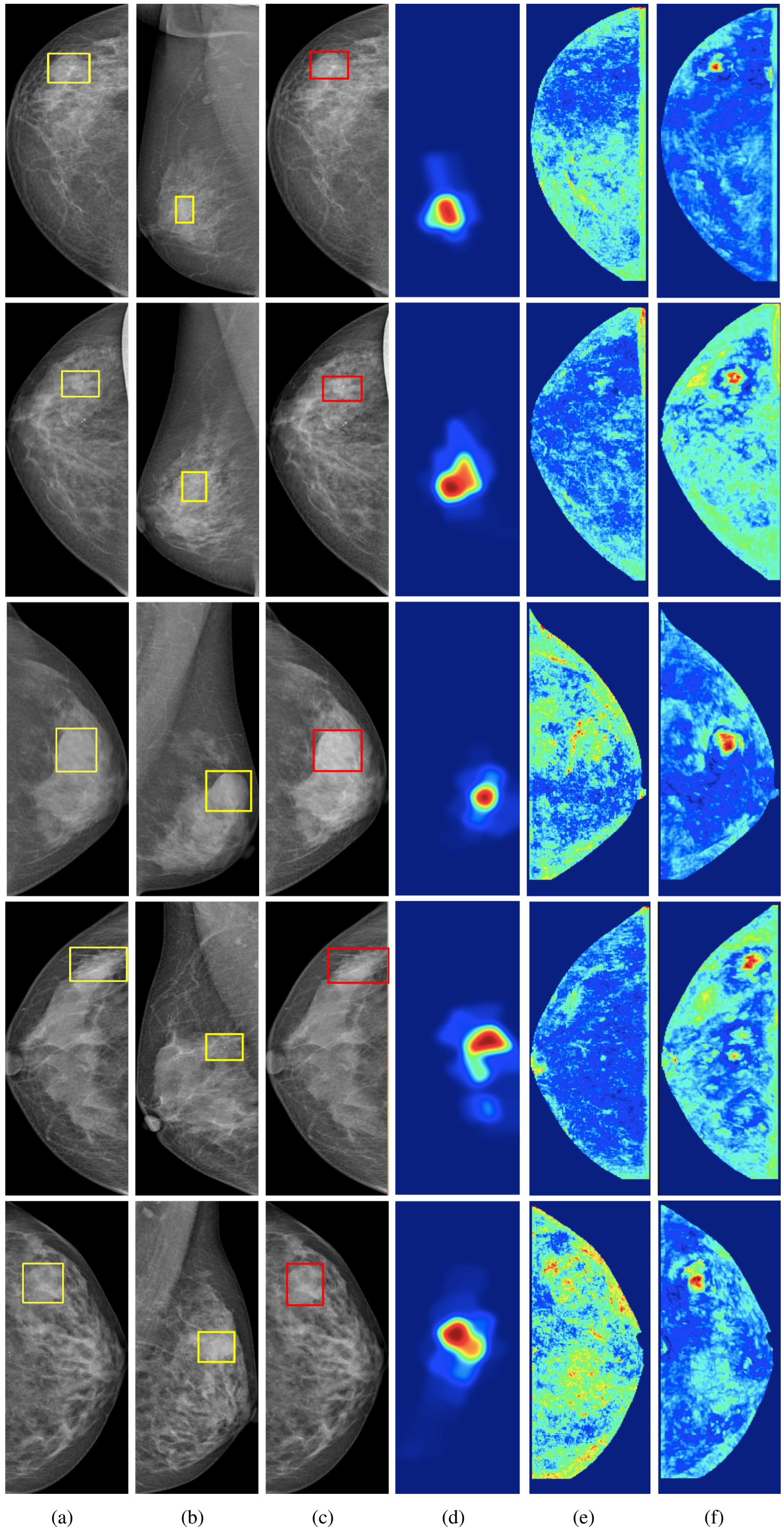


Figure 4. Additional results of the in-house dataset. Each row shows a representative case. Column (a)-(b) refer to the examined image and its auxiliary view image with annotations. Column (c) indicates detection results by BG-RCNN. Column (e) visualizes the attention area on the auxiliary view. Column (f)-(g) visualize the response maps before and after correspondence visual reasoning.