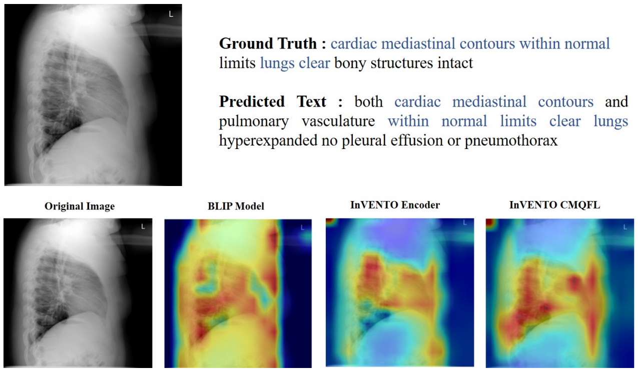
InVERGe: Intelligent Visual Encoder for Bridging Modalities in Report Generation

Supplementary Material

1. Extra Results and Visualization :

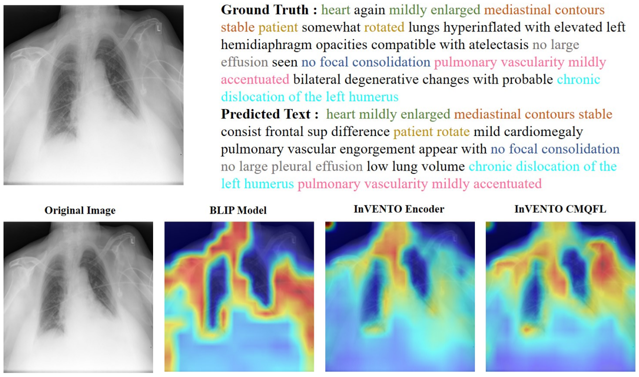
In this supplementary section, we present additional results and visualizations of the attention maps that provide further insight into our medical report generation system which shown in Figure 1. These results extend beyond the scope of the original paper and aim to offer a more comprehensive understanding of our method and its performance.

Below are some results in Figure 2, 3 from the breast dataset, which demonstrate the usefulness of our model in generating reports. We present attention maps generated by both BLIP and our InVERGe model's Encoder and CMQFL Layer in Figure 4. We also show the attention maps in individual words.



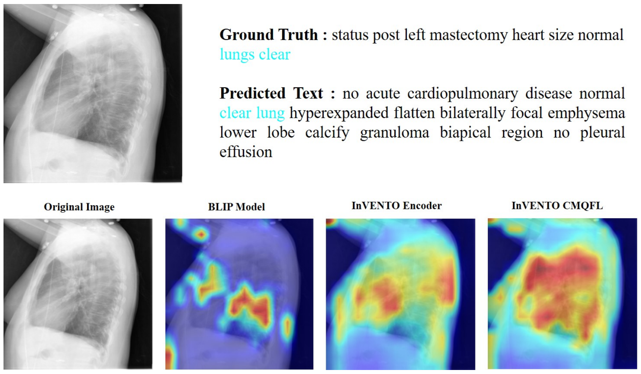
Ground Truth : cardiac mediastinal contours within normal limits lungs clear bony structures intact

Predicted Text : both cardiac mediastinal contours and pulmonary vasculature within normal limits clear lungs hyperexpanded no pleural effusion or pneumothorax



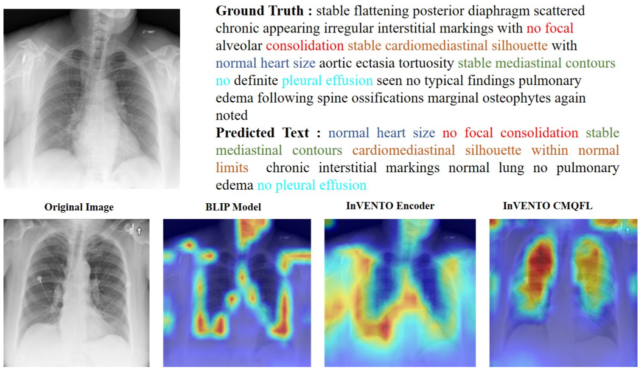
Ground Truth : heart again mildly enlarged mediastinal contours stable patient somewhat rotated lungs hyperinflated with elevated left hemidiaphragm opacities compatible with atelectasis no large effusion seen no focal consolidation pulmonary vascularity mildly accentuated bilateral degenerative changes with probable chronic dislocation of the left humerus

Predicted Text : heart mildly enlarged mediastinal contours stable consist frontal sup difference patient rotate mild cardiomegaly pulmonary vascular engorgement appear with no focal consolidation no large pleural effusion low lung volume chronic dislocation of the left humerus pulmonary vascularity mildly accentuated



Ground Truth : status post left mastectomy heart size normal lungs clear

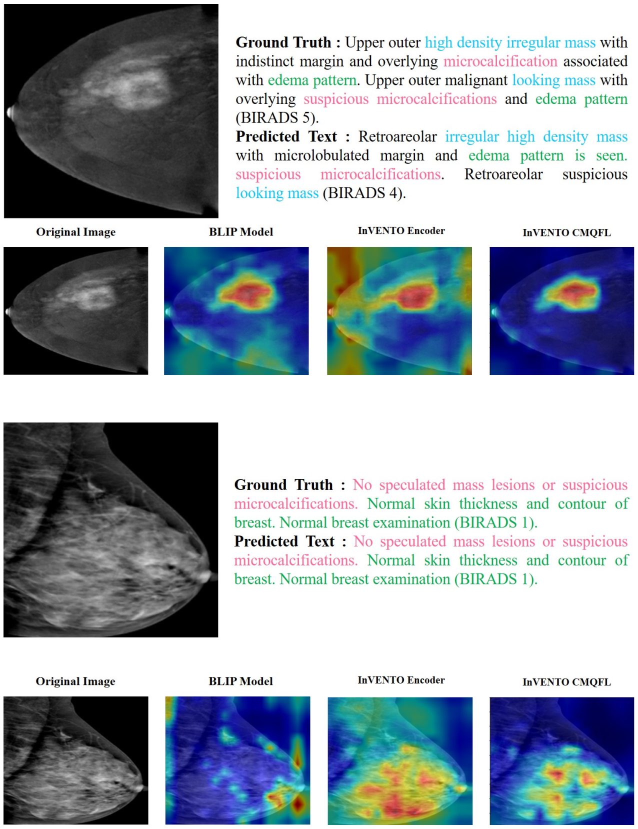
Predicted Text : no acute cardiopulmonary disease normal clear lung hyperexpanded flatten bilaterally focal emphysema lower lobe calcify granuloma biapical region no pleural effusion



Ground Truth : stable flattening posterior diaphragm scattered chronic appearing irregular interstitial markings with no focal alveolar consolidation stable cardiomediastinal silhouette with normal heart size aortic ectasia tortuosity stable mediastinal contours no definite pleural effusion seen no typical findings pulmonary edema following spine ossifications marginal osteophytes again noted

Predicted Text : normal heart size no focal consolidation stable mediastinal contours cardiomediastinal silhouette within normal limits chronic interstitial markings normal lung no pulmonary edema no pleural effusion

Figure 1. In this illustration, we present some sets of reports generated by the proposed InVERGe model, each accompanied by the corresponding ground truth reports for the X-ray images. Matched text is highlighted in the same colour to emphasise the alignment between ground truth and predicted reports. Partially, we also show the attention maps generated by the BLIP and our InVERGe model's Encoder and CMQFL Layer.



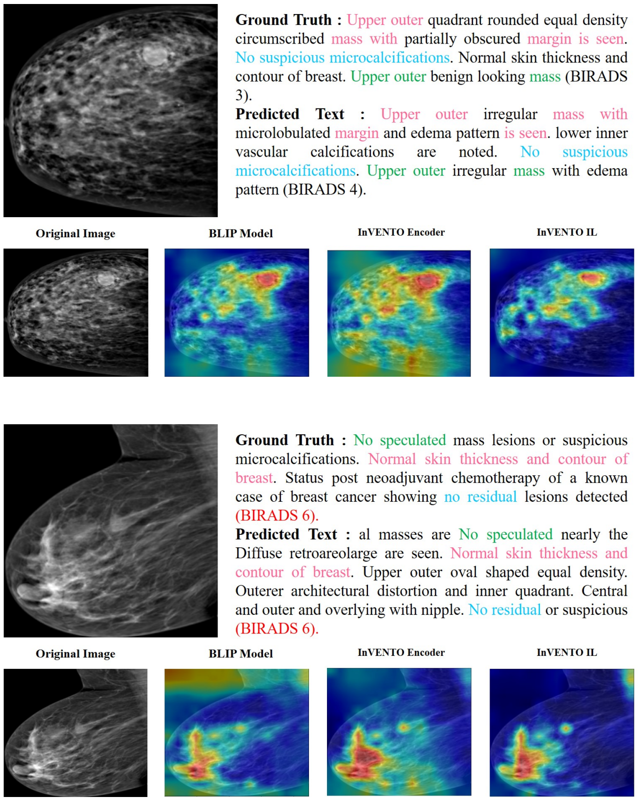
Ground Truth : Upper outer high density irregular mass with indistinct margin and overlying microcalcification associated with edema pattern. Upper outer malignant looking mass with overlying suspicious microcalcifications and edema pattern (BIRADS 5).

Predicted Text : Retroareolar irregular high density mass with microlobulated margin and edema pattern is seen. suspicious microcalcifications. Retroareolar suspicious looking mass (BIRADS 4).

Ground Truth : No speculated mass lesions or suspicious microcalcifications. Normal skin thickness and contour of breast. Normal breast examination (BIRADS 1).

Predicted Text : No speculated mass lesions or suspicious microcalcifications. Normal skin thickness and contour of breast. Normal breast examination (BIRADS 1).

Figure 2. These are the results on the low-energy images.



Ground Truth : Upper outer quadrant rounded equal density circumscribed mass with partially obscured margin is seen. No suspicious microcalcifications. Normal skin thickness and contour of breast. Upper outer benign looking mass (BIRADS 3).

Predicted Text : Upper outer irregular mass with microlobulated margin and edema pattern is seen. lower inner vascular calcifications are noted. No suspicious microcalcifications. Upper outer irregular mass with edema pattern (BIRADS 4).

Ground Truth : No speculated mass lesions or suspicious microcalcifications. Normal skin thickness and contour of breast. Status post neoadjuvant chemotherapy of a known case of breast cancer showing no residual lesions detected (BIRADS 6).

Predicted Text : al masses are No speculated nearly the Diffuse retroareolarge are seen. Normal skin thickness and contour of breast. Upper outer oval shaped equal density. Outerer architectural distortion and inner quadrant. Central and outer and overlying with nipple. No residual or suspicious (BIRADS 6).

Figure 3. These are the results on the subtraction images.

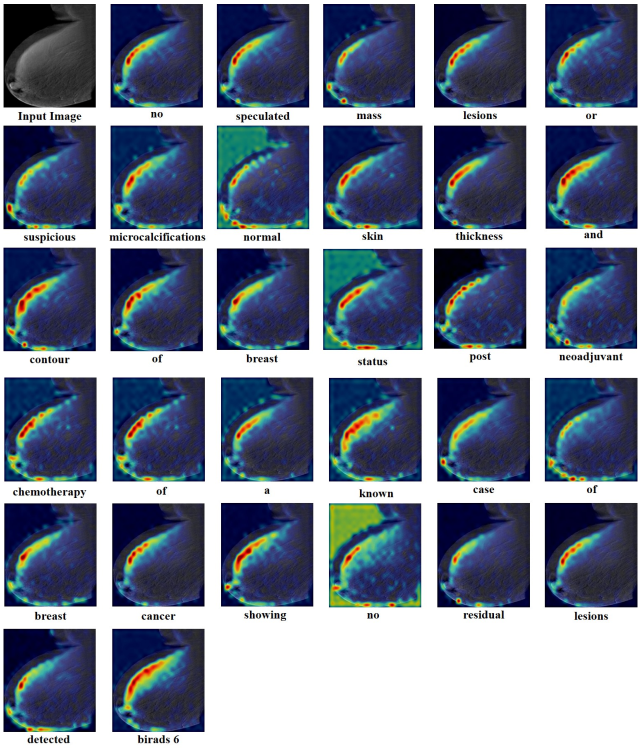


Figure 4. Individual words Visualization of each words of a report.