

# Supplementary Material for Learning to Read Chest X-Rays: Recurrent Neural Cascade Model for Automated Image Annotation

## 1. More Examples of the Chest X-ray Dataset

More examples of chest x-ray image, report, and annotations available on OpenI [1] is shown in Figure 1.

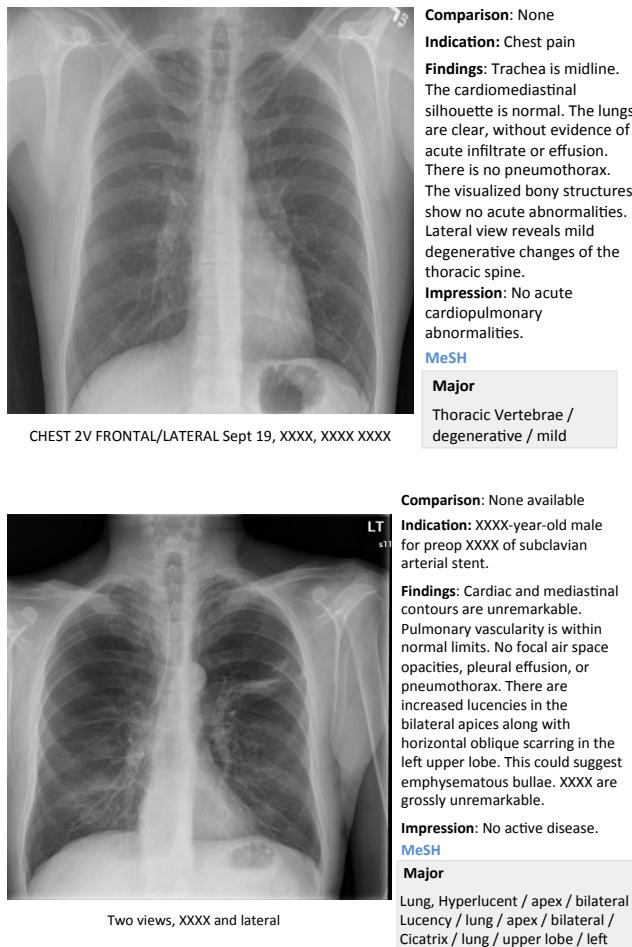


Figure 1. Examples of the OpenI [1] chest x-ray image, report, and annotations.

## 2. Visualizations of Joint Image/Text Context Vectors

Figure 2 shows the  $h_{im:text}$  vectors of about fifty annotations describing *calcified granuloma* are projected

onto two-dimensional planes via dimensionality reduction ( $\mathbb{R}^{1 \times 1024} \rightarrow \mathbb{R}^{1 \times 2}$ ), using the t-SNE [3] implementation of [2]. Figure 3 shows the visualization for the annotations describing *opacity*. We can see that different annotations describing a disease are thereby separated into different categories by the  $h_{im:text}$ .

## 3. More Annotation Generation Examples

More annotation generation examples are provided in Figures 4 and 5. Overall, the system generates promising results on predicting disease (labels) and its context (attributes) in the images. However, rare disease cases are more difficult to detect. For example, the cases *pulmonary atelectasis*, *spondylosis*, and *density* (Figure 4), as well as *foreign bodies*, *atherosclerosis*, *costophrenic angle deformity* (Figure 5) are much rarer in the data than *calcified granuloma*, *cardiomegaly*, and all the frequent cases listed in Table 1 of the main paper.

Furthermore, the (left or right) location of the disease cannot be identified in a lateral view (obtained by scanning the patient from the side), as shown in Figure 5. Since our dataset contains a limited number of disease cases, we treat each x-ray image and report as a sample, and do not account for different views.

For future work, we plan to improve prediction accuracy by both (a) accounting for the different views, and (b) collecting a larger dataset to better account for rare diseases.

## References

- [1] Open-i: An open access biomedical search engine. <https://openi.nlm.nih.gov>.
- [2] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011.
- [3] L. Van der Maaten and G. Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 2008.

```

calcified_granuloma lung lower_lobe left multiple
    calcified_granuloma lung upper_lobe left small
        calcified_granuloma lung base right
            calcified_granuloma lung bilateral multiple
                calcified_granuloma lung upper_lobe right multiple
                    calcified_granuloma lung base left
                        calcified_granuloma lung upper_lobe right

calcified_granuloma lung hilum right
    calcified_granuloma lung lower_lobe right large
        calcified_granuloma lung apex right small

calcified_granuloma lung hilum bilateral calcified_granuloma lung hilum right large
    calcified_granuloma lung multiple
        calcified_granuloma lung hilum right large
            calcified_granuloma lung upper_lobe right calcified_granuloma scattered spine degenerative
                calcified_granuloma lung upper_lobe right calcified_granuloma scattered spine degenerative
                    calcified_granuloma multiple thoracic_vertebrae_degenerative

calcified_granuloma lung hilum right calcified_granuloma lung middle_lobe right
    calcified_granuloma lung lower_lobe right large
        calcified_granuloma lung lymph nodes multiple
            calcified_granuloma multiple

calcified_granuloma lung lower_lobe left calcinosis lung hilum lymph nodes multiple
    calcified_granuloma lung middle_lobe right
        calcified_granuloma lung hilum right small

calcified_granuloma lung bilateral multiple

```

```

calcified_granuloma lung lower_lobe left calcinosis lung hilum lymph nodes multiple
    calcified_granuloma lung lower_lobe right calcified_granuloma lung lingula
        calcified_granuloma bilateral scattered multiple
            calcified_granuloma lung lower_lobe left thoracic_vertebrae_degenerative chronic
                calcified_granuloma lung lingula

calcified_granuloma lung middle_lobe right
    calcified_granuloma lung hilum lymph nodes multiple
        calcified_granuloma lung apex right small

calcified_granuloma lung bilateral multiple

calcified_granuloma lung right small
    calcified_granuloma lung hilum bilateral multiple large calcinosis lung hilum lymph_nodes bilateral multiple large
        calcified_granuloma lung upper_lobe right small

```

```

calcified_granuloma
calcified_granuloma
calcified_granuloma

```

Figure 2. Visualization of the  $h_{im:text}$  of about 50 samples of the annotations describing the disease *calcified granuloma* on 2D plane.

```

opacity lung upper_lobe hilum left pneumonia upper_lobe left

opacity lung lower_lobe left pulmonary_atelectasis lower_lobe left

opacity lung base right pulmonary_atelectasis base right
    opacity lung upper_lobe hilum left pneumonia upper_lobe left

opacity lung base left streaky opacity lung middle_lobe right
    opacity lung upper_lobe hilum left opacity lung middle_lobe right
        opacity lung lower_lobe right pulmonary_atelectasis upper_lobe right
            opacity lung base left mild pulmonary_atelectasis base left

opacity lung base bilateral patchy streaky
    opacity lung upper_lobe right focal pneumonia upper_lobe right
        opacity lung upper_lobe hilum left opacity lung upper_lobe right focal pneumonia upper_lobe right
            opacity lung upper_lobe right pneumonia upper_lobe right
                opacity lung lower_lobe left opacity lung middle_lobe right
                    opacity lung lower_lobe right opacity lung upper_lobe right focal
                        opacity lung base left focal
                            opacity lung lower_lobe left pulmonary_atelectasis lower_lobe left
                                opacity lung ribs left anterior round
                                    opacity lung bilateral interstitial diffuse reticular round severe
                                        opacity lung upper_lobe right round
                                            opacity lung lower_lobe bilateral interstitial
                                                opacity lung hilum left interstitial
                                                    opacity thorax opacity lung base left pulmonary_atelectasis base right

opacity lung lower_lobe left pulmonary_atelectasis base left focal
    opacity lung upper_lobe hilum left opacity lung base bilateral pulmonary_atelectasis base bilateral
        opacity lung base bilateral pulmonary_atelectasis base bilateral
            opacity lung lower_lobe right streaky opacity lung lower_lobe right
                opacity lung base bilateral pulmonary_atelectasis base bilateral
                    opacity lung base bilateral patchy streaky
                        opacity lung base bilateral
                            opacity lung lower_lobe bilateral interstitial
                                opacity lung lower_lobe bilateral interstitial
                                    opacity lung lower_lobe bilatera
                                        opacity lung lower_lobe right pulmonary_atelectasis lower_lobe right

```

Figure 3. Visualization of the  $h_{im:text}$  of about 50 samples of the annotations describing the disease *opacity* on 2D plane.

input image				
generated annotation	calcinos / aorta / thoracic_vertebrae_degenerative / mild thoracic_vertebrae_degenerative / mild	calcified_granuloma / lung / ribs / base / pleura / middle_lobe calcified_granuloma / lung / left	surgical_instruments / abdomen / lung / scattered / opacity / lung opacity / lung / base / left	calcified_granuloma / lung / hilum / lung / base / left calcified_granuloma / lung / bilateral / scattered
true annotation				
input image				
generated annotation	lung_bilateral / hyperdistention / lung / hilum / thorax / right calcinos / lung / hilum /right / granulomatous_disease	calcinos / lung / base / left / prominent calcified_granuloma / lung / left / calcinos / aorta	normal normal	normal normal
true annotation				
input image				
generated annotation	calcified_granuloma / scattered / lung / costophrenic_angle / right / moderate scoliosis / thoracic_vertebrae / right / moderate	granuloma / lung / ribs / posterior / calcified_granuloma / spine_degenerative calcified_granuloma / lung / base / right / small	calcified_granuloma / lung / hilum / thorax / right / mild cicatrix / lung / lingula	calcified_granuloma / lung / base / left calcinos / lung / hilum / right
true annotation				
input image				
generated annotation	thoracic_vertebrae_degenerative / mild pulmonary_atelectasis / base / left / mild	aorta_thoracic / tortuous / mild / thorax / left / thorax catheters_indwelling / left	airspace_disease / lung / middle_lobe / right / thoracic_vertebrae / mild spondylosis / thoracic_vertebrae	airspace_disease / lung / hilum / right / mild density / lung / middle_lobe / right
true annotation				

Figure 4. More examples of annotation generations (light green box) compared to true annotations (yellow box) for input images in the test set. Rare disease cases are more difficult to detect.

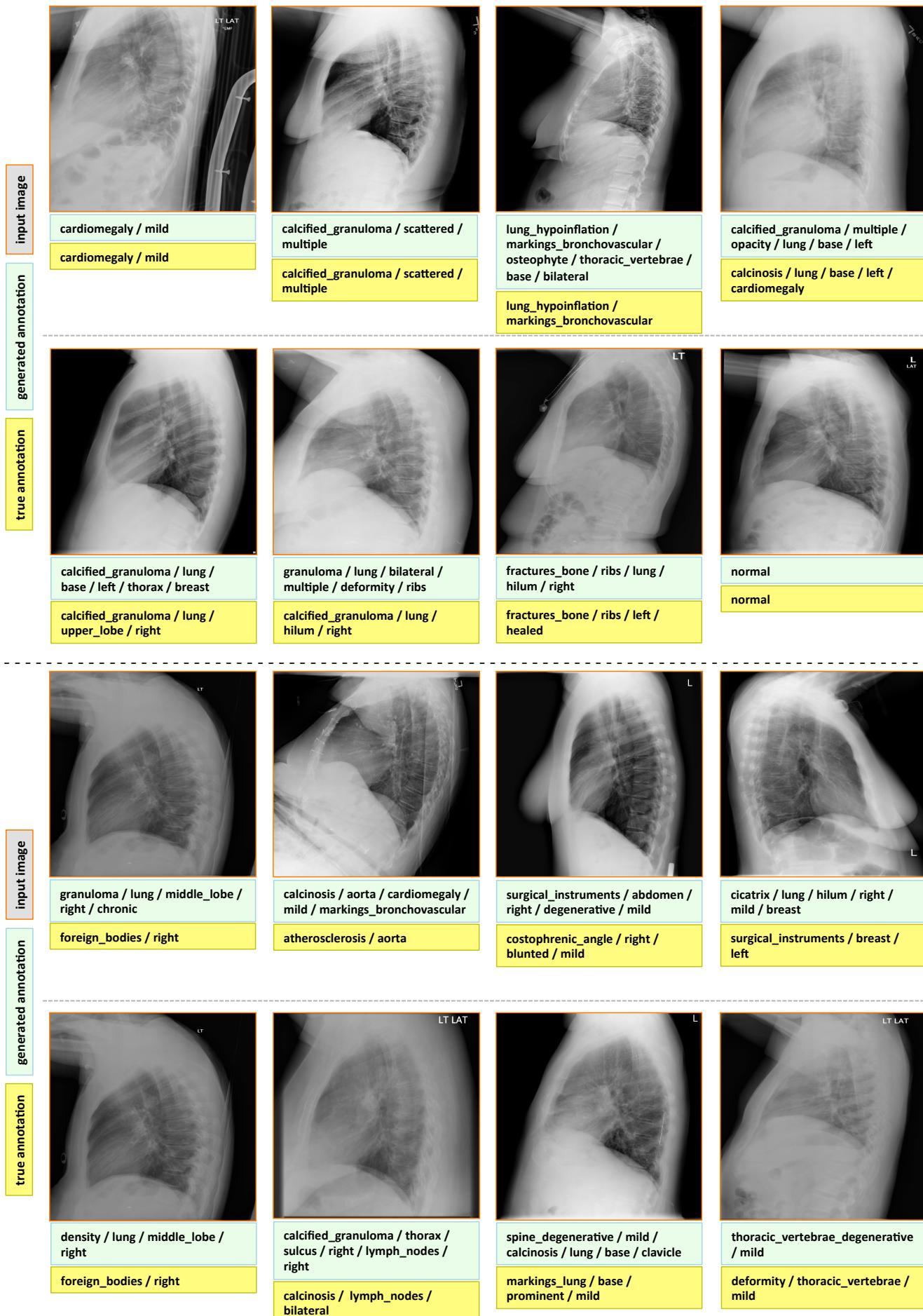


Figure 5. More examples of annotation generations (light green box) compared to true annotations (yellow box) for input images in the test set. In lateral (side) views, the (left or right) location of the disease can no longer be identified.