Representation Learning with a Transformer-Based Detection Model for Localized Chest X-Ray Disease and Progression Detection

Mehrdad Eshraghi Dehaghani *[1], Amirhossein Sabour *[1], Amarachi B. Madu[2], Ismini Lourentzou[3], and Mehdi Moradi **[1]

[1] McMaster University, Hamilton, ON, Canada
{eshragm, saboua4, moradm4}@mcmaster.ca
[2] Virginia Tech, Blacksburg, VA, USA
bmamarachi@vt.edu
[3] University of Illinois Urbana - Champaign, Champaign, IL, USA
lourent2@illinois.edu

Abstract. Medical image interpretation often encompasses diverse tasks, yet prevailing AI approaches predominantly favor end-to-end image-to-text models for automatic chest X-ray reading and analysis, often overlooking critical components of radiology reports. At the same time, employing separate models for related but distinct tasks leads to computational overhead and the inability to harness the benefits of shared data abstractions. In this work, we introduce a framework for chest X-ray interpretation, utilizing a Transformer-based object detection model trained on abundant data for learning localized representations. Our model achieves a mean average precision of ∼94% in identifying semantically meaningful anatomical regions, facilitating downstream tasks, namely localized disease detection and localized progression monitoring. Our approach also yields competitive results in localized disease detection, with an average ROC 89.1% over 9 diseases. In addition, to the best of our knowledge, our work is the first to tackle localized disease progression monitoring, with the proposed model being able to track changes in specific regions of interest (RoIs) with an average accuracy ∼67% and average F1 score of ∼71%. Code is available at https://github.com/McMasterAIHLab/CheXDetector.

Keywords: Disease Localization · Representation Learning · Longitudinal CXR Relationships · Transformer-based detection

1 Introduction

Chest X-ray is the most common medical imaging modality. In recent years, with the introduction of multiple large-scale annotated chest X-ray datasets [4,5], the field of automatic interpretation of chest X-ray images leveraging artificial intelligence has seen a great deal of activity. There are generally two

* Equal contribution
** Corresponding author

lines of work in this area. An early wave of works has focused on accurate detection/classification for a limited number of diseases or findings [13]. The obvious flaw is the limited scope, as radiology reports are not merely lists of findings. Instead, they consist of localized descriptions, often with comparative and localized references to the progress of disease from a previous scan. In response to these shortcomings, subsequent works have expanded the scope of these early works by including a larger number of findings in their classification models than those labeled in the publicly available datasets [17]. With the increased popularity of language models in medical imaging, the more recent wave of activity is focused on end-to-end training of image-to-sequence models that produce a complete radiology report given a chest X-ray image [7,11]. This line of work addresses the problem of limited scope and application of disease classifiers. However, these models are often evaluated for their readability and similarity to radiologist reports, and not for the accuracy of the findings they list or their comprehensiveness [18]. As is common in generative models, these models can often produce factually incorrect language. Additionally, there has been a growing interest in disease detection with semantically meaningful localization [1,10] as well as the monitoring of disease progression within image pairs, assessing whether a patient's condition has improved, deteriorated, or remained stable over time [6,9]. Despite these advancements, to the best of our knowledge, the challenge of *localized* disease progression monitoring, *i.e.*, predicting disease progression in specific anatomical regions, remains unexplored.

Given the weaknesses of direct image-to-text models and the variety of detection/classification tasks involved in chest X-ray interpretation, we propose to train and utilize a DEtection TRansformer (DETR) anatomical region detection model [2] to address multiple clinically relevant downstream tasks such as localized disease detection and localized disease progression monitoring. Specifically, previous works have provided large datasets of X-ray images with marked bounding boxes for anatomical regions (*e.g.*, 'lower lobe of right lung') [14,15]. We define the detection of these bounding boxes as the task for training a DETR model. When trained, this model provides a rich feature vector for each anatomical region that can be used for both localized disease detection and localized disease progression monitoring. For each task, we train relatively compact models, using the features from the upstream model. We show that the performance of our proposed framework is comparable to models specifically trained for these tasks. Our contributions can be summarized as follows:

(1) We introduce a novel approach for chest X-ray interpretation. By utilizing rich feature vectors generated by a DETR model trained for anatomical region detection, we address two clinically relevant downstream tasks simultaneously, *i.e.*, localized disease detection and localized progression monitoring.
(2) We introduce the task of disease progression monitoring at a localized level. Our experimental results show that a simple model that extracts anatomical region feature differences can achieve competitive accuracy on this new task. We additionally demonstrate that a simple MLP architecture can jointly achieve competitive performance in localized disease detection.

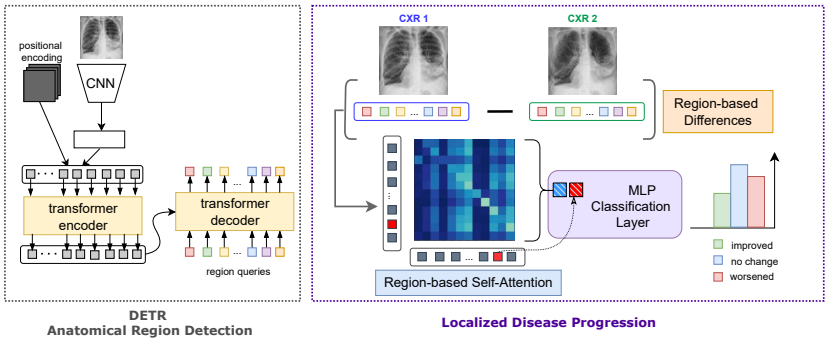

Fig. 1: We first train a DETR anatomical region detection model on a large collection of CXR images. Given a pair of CXR images, the pretrained DETR decoder extracts visual features for anatomical regions of interest (RoI) for each image. These features are then used to compute region-based visual differences between the two CXRs. The information encoded in the difference vector is summarized through the self-attention mechanism that captures the importance of each RoI vector in relation to other RoIs and helps the model focus on relevant RoI changes. The resulting summary vector is concatenated with the region of interest (RoI) vector and fed into a multi-layer perceptron (MLP) classification layer for predicting whether the condition localized on the specific RoI has improved, worsened, or remained unchanged.

(3) We further provide comprehensive ablation analysis with three model variations and qualitative examples to show the importance of anatomical regions in disease progression monitoring.

2 Methodology

2.1 Problem Definition

Let $\mathcal{C} = \{(\mathbf{X}, \mathbf{X}')_i\}_{i=1}^{N}$ be a set of CXR image pairs, where $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{H \times W \times C}$, and H, W and C are the height, width, and number of channels, respectively. Each image $\mathbf{X}$ is associated with a localized label set $\mathcal{Y}_i^1 = \{y_{i,k,m}\}_{k=1,m=1}^{K,M}$ where $y_{i,k,m}^1 \in \{0, 1\}$, indicating whether the label for the m-th finding appears in the k-th anatomical region of the image or not. In addition, each image pair $(\mathbf{X}, \mathbf{X}')_i$ is associated with a label set $\mathcal{Y}_i^2 = \{y_{i,k}\}_{k=1}^{K}$, where $y_{i,k}^2$ indicates whether the overall condition in the k-th anatomical region of the image pair $(\mathbf{X}, \mathbf{X}')_i$ has improved, worsened, or remained the same. The goal is to design a model that accurately predicts a set of labels indicative of the presence of pre-defined diseases at every anatomical Region of Interest (RoI) for an unseen image, and is also able to compare the two unseen images $(\mathbf{X}, \mathbf{X}')$ to predict localized progression labels as accurately as possible.

2.2 DETR region representation extraction backbone

To this end, we first utilize a DETR pre-trained region detection model for extracting feature representations of anatomical regions of interest (RoIs). The DETR region detection outputs a set of bounding boxes denoted as $\mathbf{B} = \{(b_1, c_1, s_1), (b_2, c_2, s_2), \ldots, (b_K, c_K, s_K)\}$, where $b_i = (x_{\min}, y_{\min}, x_{\max}, y_{\max})$ represents the coordinates of the i-th bounding box, c_i denotes the class label associated with the box (chosen from a predefined set of K anatomical regions), and s_i represents the confidence score associated with each region query, indicating the likelihood of corresponding to a valid RoI. For each region query, the output of the last hidden state of the decoder can be represented as $\mathbf{F} = \{f_1, f_2, \ldots, f_i, \ldots, f_K\}$, where f_i represents the last hidden state of the decoder for the i-th region query. These hidden states serve as learned anatomical region feature representations, capturing the contextual information extracted by the decoder regarding the corresponding anatomical region query.

2.3 Localized Disease Detection

By utilizing the extracted feature representations $\mathbf{F} = \{f_1, f_2, \ldots, f_i, \ldots, f_K\}$ for the K anatomical regions of interest (RoIs), we train a compact feed-forward network to predict the presence or absence of particular disease in the respective RoIs. Let $y_{i,k,m}$ represent the ground truth label indicating the actual presence (1) or absence (0) of the m-th disease in the k-th RoI for the i-th sample. Similarly, let $\hat{y}_{i,k,m}$ represent the predicted probability for the presence or absence of the m-th disease in the k-th RoI for the i-th sample. The binary cross-entropy loss $\mathcal{L}$ for a batch of N images can be defined as:

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{k=1}^{K} \sum_{m=1}^{M} \left(\hat{y}_{i,k,m} \cdot \log(y_{i,k,m}) + (1 - \hat{y}_{i,k,m}) \cdot \log(1 - y_{i,k,m}) \right), \quad (1)$$

where N denotes the batch size and M the number of diseases.

2.4 Localized Disease Progression Monitoring

For localized disease progression monitoring, the goal is to predict, for each image pair $(\mathbf{X}, \mathbf{X}')_i$, whether the condition of a particular k-th anatomical region has improved, worsened, or remained unchanged. Let $\mathbf{F}_1 = \{f_{1,1}, f_{1,2}, \ldots, f_{1,k}, \ldots, f_{1,K}\}$ represent the feature vectors extracted from the regions of interest (RoIs) in the first image $\mathbf{X}_i$, and $\mathbf{F}_2 = \{f_{2,1}, f_{2,2}, \ldots, f_{2,k}, \ldots, f_{2,K}\}$ represent the feature vectors extracted from the corresponding RoIs in the second image $\mathbf{X}'_i$, where K is the number of RoIs. The model computes the differences of region vectors between the two images $\mathbf{F}_{\text{diff}} = [\mathbf{F}_2 - \mathbf{F}_1] = [f_{2,1} - f_{1,1}, \ldots, f_{2,k} - f_{1,k}, \ldots, f_{2,K} - f_{1,K}]$. To summarize the RoI information and capture the relationships between different RoIs, we employ a self-attention mechanism. The self-attention operation on $\mathbf{F}_{\text{diff}}$ can be denoted as:

$$\boldsymbol{\alpha} = \text{Softmax}\left(\frac{\mathbf{F}_{\text{diff}} \mathbf{F}_{\text{diff}}^T}{\sqrt{d_k}} \right), \quad (2)$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_k]$ represents the attention weights for each RoI, d_k is the dimension of the key vectors, and the softmax function is applied along the rows of the matrix. Next, the weighted sum of the difference feature vectors corresponding to each RoI k is computed using the attention weights, *i.e.*, $\boldsymbol{\alpha}\mathbf{F}_{\text{diff}}$, providing a summarized representation of the global RoI information considering the interdependencies between different RoIs. To perform the final localized disease progression prediction, we report results in both "Global" and "Region-focused" attention variants. In the "Global" approach, we average the rows of the self-attention output to obtain a single global vector to be used for all the RoIs of the chest x-ray, which is concatenated with $\mathbf{F}_k$ and passed through a classification layer. In the "Region-focused" approach, the k-th row of the self-attention output, corresponding to the k-th RoI, is concatenated with $\mathbf{F}_k$ and used directly for prediction, *i.e.*, $\hat{y}_k = g\left([\mathbf{att}_k; \mathbf{F}_k]\right)$, where $\mathbf{att}_k$ is the self-attention output for the k-th RoI and $\mathbf{F}_k = [f_{2,k} - f_{1,k}]$ is the difference vector for the k-th RoI. The disease progression model loss for a batch of N images is defined as:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K}\left[y_{i,k}\cdot\log(\hat{y}_{i,k}) + (1-y_{i,k})\cdot\log(1-\hat{y}_{i,k})\right], \tag{3}$$

where $\mathcal{L}$ is the total loss over the entire batch, N is the batch size, K is the number of RoIs, and $y_{i,k}$, $\hat{y}_{i,k}$ are the ground truth label and model prediction for the k-th RoI of the i-th image.

3 Experiments

Implementation Details. We employ DETR with ResNet-50 backbone as the initial model for fine-tuning and extracting feature representations of anatomical regions of interest. The model was trained using PyTorch-Lightning [3], with AdamW optimizer [8], weight decay of 10^{-4}, backbone learning-rate of 10^{-5} and a learning-rate of 10^{-3}. To avoid the gradient exploding problem, we use gradient clipping of 0.1. The initial batch size for training is 13. For further increasing the batch size while considering the memory limitations, we use accumulated gradients of 5. The model was trained for 25 epochs. To filter out unwanted detections, we apply a threshold $\tau > 0.85$ to the scores associated with each region query. Only region queries with scores exceeding this threshold are considered valid detections. For the disease classification module, we use a feed-forward network (FFN) similar to DETR's FFN with an additional batch norm between layers. The input dimension for the FFN is 256, hidden dimension of 256, output dimension of 9 (equal to the number of studied finding classes). The model is trained for 100 epochs using Adam optimizer with a learning rate of 5×10^{-4} and a weight decay of 10^{-5}. For the localized disease progression detection module, a similar feed-forward network with an input dimension of 512, hidden dimension of 256 with two layers and output dimension of 3 is used. The model is trained for 100 epochs using Adam optimizer with a learning rate of 10^{-3} and a weight decay of 10^{-5}.

Table 1: Dataset characteristics for progression labels per anatomical regions at: right upper lung zone (RULZ), right mid lung zone (RMLZ), right lower lung zone (RLLZ), right costophrenic angle (RCA), right hilar structures (RHS), right apical zone (RAZ), left upper lung zone (LULZ), left mid lung zone (LMLZ), left lower lung zone (LLLZ), left costophrenic angle (LCA), left hilar structures (LHS), cardiac silhouette (CS)

Progression Label	RULZ	RMLZ	RLLZ	RCA	RHS	RAZ
Improved	957	2,338	5,681	5,249	7,799	744
Worsened	1,301	3,537	8,333	6,406	7,699	543
No Change	37,149	34,617	28,673	30,529	27,883	37,915
Total	39,407	40,492	42,687	42,184	43,381	39,202

Progression Label	LULZ	LMLZ	LLLZ	LCA	LHS	CS
Improved	678	2,382	6,305	5,225	7,749	1,722
Worsened	927	3,414	8,996	6,399	7,599	3,097
No Change	37,596	34,549	27,667	30,478	27,954	35,036
Total	39,201	40,345	42,960	42,102	43,302	39,855

Dataset. We make use of the Chest ImaGenome dataset [15]. This dataset consists of two different sub-datasets: 1) Locally labeled data using a combination of rule-based natural language processing (NLP) and CXR atlas-based bounding box detection techniques [14,16] to generate the annotations (silver dataset). This subset comprises $237,827$ frontal MIMIC-CXRs [5]. 2) Manually validated and corrected studies of 500 patients as ground truth. Chest ImaGenome is represented as an anatomy-centered scene graph with $1,256$ combinations of relation annotations between 29 CXR anatomical locations and their attributes. Each image is structured as one scene graph, resulting in approximately $670,000$ localized comparison relations between the anatomical locations across sequential exams. Rich representation features play a key role in performing the downstream task efficiently. To train the upstream model, we utilize the entire silver dataset of Chest ImaGenome with 70/10/20 split.

For localized disease progression, we consider the localized comparison relation data within Chest ImaGenome that pertains to cross-image relations for nine diseases of interest. Each comparison relation in the Chest ImaGenome dataset includes the DICOM identifiers of the two CXRs being compared, a set of comparison labels per some anatomical regions, and a set of disease names. In some cases, more than one progression label was assigned to one region. We excluded these samples from the dataset to focus on more accurate labels. The comparison is labeled as "no change", "improved" or "worsened", which indicates whether the patient's condition has remained stable, improved, or worsened, respectively. In contrast to [9] which reports global classification, we solve this problem at the local level, acquiring one progress label per anatomical location. We use $35,646$ CXR pairs in total that pertain to the nine diseases of interest. The distribution of the data is improved $(53,134)$, worsened $(58,251)$, and no change $(390,046)$. We use 70/10/20 train/validation/testing split across studies. Due to the large gap between the number of "no change" labels and two other

Table 2: Area under PR curves for the 12 anatomical locations with an mPA of 93.5%, calculated using [12]. Format for each cell: (Anatomical location: Area under PR curve, IoU threshold = 0.5)

RULZ	**RMLZ**	**RLLZ**	**RCA**	**RHS**	**RAZ**
0.984	0.961	0.962	0.818	0.963	0.968
LULZ	**LMLZ**	**LLLZ**	**LCA**	**LHS**	**CS**
0.983	0.964	0.955	0.743	0.960	0.959

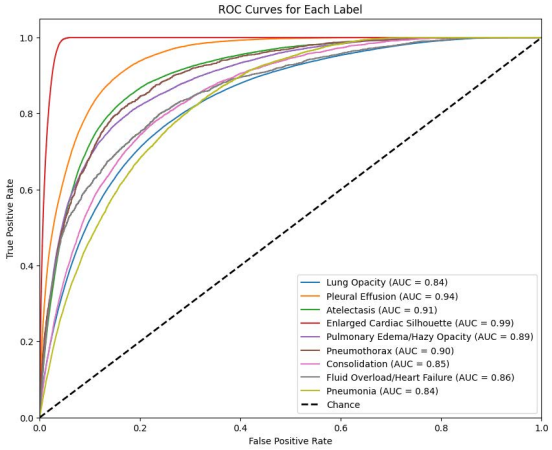

Fig. 2: Localized Multi-label Disease Detection Results

classes, only for training, we consider a random subset of "no change" labels equal to the maximum number of labels in the other two classes. Table 1 provides further details. For localized disease classification, we utilize the silver dataset (237, 827 CXRs) to train the model. We use 70/10/20 train/validation/testing split across studies. High-level statistics of the generated dataset based on findings and anatomical regions of interest are included in the supplementary material.

4 Results

Upstream detection network: Table 2 reports the area under precision-recall curves for the twelve target anatomical regions. The mean average precision (mAP) is 93.5%. For 10 of the 12 regions, the area under ROC curve is at or above 96%, with only right and left costophrenic angle being the exceptions.

Localized disease detection: Figure 2 shows the ROC curves obtained for the nine findings from the localized disease detection model. Despite a simple MLP architecture, the model provides an average AUC score of 89.1%. The closet benchmark to this work is [1] which reports localized disease detection on the same dataset with average AUC scores in the range of 89% to 93% for various models trained for the specific task of localized disease detection.

Table 3: Localized Disease Progression Results (Accuracy/Weighted F1)

Anatomical location	Attention (Global)	Attention (Region-focused)	MLP
Right Upper Lung Zone	66.66/77.07	78.93/84.94	96.17/94.30
Right Mid Lung Zone	62.25/69.12	63.40/70.12	79.76/80.07
Right Lower Lung Zone	56.38/57.47	58.33/59.69	17.48/7.15
Right Costophrenic Angle	61.72/64.12	56.62/60.69	46.67/51.78
Right Hilar Structures	61.17/59.93	59.24/60.29	16.41/6.89
Right Apical Zone	66.69/78.11	83.81/89.53	97.44/96.17
Left Upper Lung Zone	66.56/77.75	81.01/86.85	97.15/95.75
Left Mid Lung Zone	63.80/70.09	69.23/74.30	83.34/81.82
Left Lower Lung Zone	55.56/55.91	54.55/55.76	19.98/8.73
Left Costophrenic Angle	61.86/64.43	61.50/64.38	38.19/42.64
Left Hilar Structures	67.88/63.42	61.87/63.20	16.59/9.29
Cardiac Silhouette	65.38/73.40	67.59/75.29	90.28/88.00
Weighted Average	**67.36/70.60**	**66.65/70.86**	**60.35/57.54**

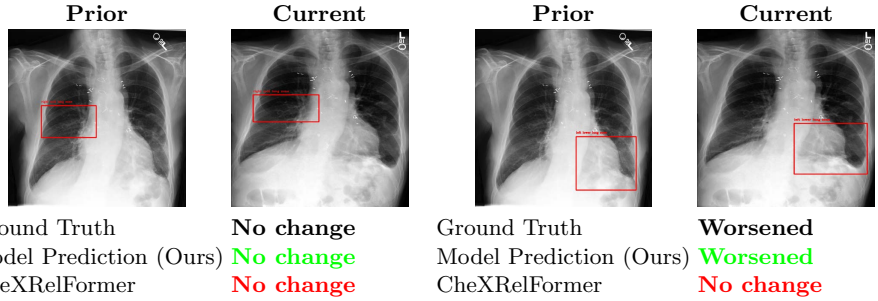

Fig. 3: Examples of model predictions obtained by our model compared against the ground-truth labels and CheXRelFormer model [9]. Lung Opacity pathology (left) and Pneumonia pathology (right).

Localized disease progression: The results for localized progression labeling are presented in Table 3 for the three model variations. It is clear that the introduction of the attention layer in this classifier has improved the results compared to the baseline of MLP. Both the global and region-focused attention architectures outperform the MLP model, with a slight edge for global attention which provides an average accuracy $\sim 67\%$ and an average F1 score of $\sim 71\%$. For this application, we do not have a current direct comparison from previous work. We trained a CNN Siamese network equivalent to the one in [15] with 3 classes, on 9 diseases and 12 anatomical regions, maintaining our original train/test splits. This simple model delivered a weighted average accuracy of only $\sim 34\%$ and F1 score of $\sim 32\%$. Authors in CheXRelFormer [9] report global disease progression with an average accuracy of 49% across the diseases. Figure 3 highlights the advantage of localized over global progress classification. The progression labels in different anatomical locations can be inconsistent. As the figure shows, while our

local model correctly classifies the progression label for both regions of interest, the global label by definition provides a single label that cannot be correct for both regions.

5 Conclusions

In this study, we presented a novel approach for interpreting chest X-rays, leveraging rich feature vectors derived from a DETR model trained specifically for anatomical region detection. By harnessing these feature vectors, we concurrently addressed two clinically significant downstream tasks: localized disease detection and localized progression monitoring. Furthermore, we introduce the novel task of disease progression monitoring at a localized level, demonstrating that extracting anatomical region feature differences can achieve competitive accuracy in this domain. Our experiments showcase the effectiveness of representation learning combined with simple architectures in achieving competitive performance for localized disease detection and progression.

Acknowledgments. This study was funded by Discovery Grant from Natural Sciences and Engineering Research Council of Canada (RGPIN-2022-04657).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Agu, N.N., Wu, J.T., Chao, H., Lourentzou, I., Sharma, A., Moradi, M., Yan, P., Hendler, J.: Anaxnet: Anatomy aware multi-label finding classification in chest x-ray. In: de Bruijne, M., Cattin, P.C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., Essert, C. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. pp. 804–813. Springer International Publishing, Cham (2021)
2. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S.: End-to-end object detection with transformers. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.M. (eds.) Computer Vision – European Conference on Computer Vision (ECCV) 2020. pp. 213–229. Springer International Publishing, Cham (2020)
3. Falcon, W., The PyTorch Lightning team: PyTorch Lightning (Mar 2019). https://doi.org/10.5281/zenodo.3828935, `https://github.com/Lightning-AI/lightning`
4. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R., Shpanskaya, K., Seekins, J., Mong, D.A., Halabi, S.S., Sandberg, J.K., Jones, R., Larson, D.B., Langlotz, C.P., Patel, B.N., Lungren, M.P., Ng, A.Y.: Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison (2019)
5. Johnson, A.E., Pollard, T.J., Berkowitz, S.J., et al.: Mimic-cxr, a de-identified publicly available database of chest radiographs with free-text reports. Scientific data pp. 1–8 (2019)
6. Karwande, G., Mbakwe, A.B., Wu, J.T., Celi, L.A., Moradi, M., Lourentzou, I.: Chexrelnet: An anatomy-aware model for tracking longitudinal relationships between chest x-rays. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part I. pp. 581–591. Springer (2022)

7. Li, M., Lin, B., Chen, Z., Lin, H., Liang, X., Chang, X.: Dynamic graph enhanced contrastive learning for chest x-ray report generation. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3334–3343 (2023)
8. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. In: International Conference on Learning Representations (ICLR) (2018)
9. Mbakwe, A.B., Wang, L., Moradi, M., Lourentzou, I.: Hierarchical vision transformers for disease progression detection in chest x-ray images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 685–695. Springer (2023)
10. Müller, P., Meissen, F., Brandt, J., Kaissis, G., Rueckert, D.: Anatomy-driven pathology detection on chest x-rays. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 57–66. Springer (2023)
11. Nguyen, H., Nie, D., Badamdorj, T., Liu, Y., Zhu, Y., Truong, J., Cheng, L.: Automated generation of accurate & fluent medical X-ray reports. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. pp. 3552–3569. Association for Computational Linguistics (Nov 2021)
12. Padilla, R., Passos, W.L., Dias, T.L.B., Netto, S.L., da Silva, E.A.B.: A comparative analysis of object detection metrics with a companion open-source toolkit. Electronics **10**(3) (2021). https://doi.org/10.3390/electronics10030279, `https://www.mdpi.com/2079-9292/10/3/279`
13. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M.P., Ng, A.Y.: Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning (2017)
14. Wu, J., Gur, Y., Karargyris, A., Syed, A.B., Boyko, O., Moradi, M., Syeda-Mahmood, T.: Automatic bounding box annotation of chest x-ray data for localization of abnormalities. In: Proceedings of the 17th International Symposium on Biomedical Imaging (ISBI). pp. 799–803. Institute of Electrical and Electronics Engineers (IEEE) (2020)
15. Wu, J.T., Agu, N.N., Lourentzou, I., Sharma, A., Paguio, J.A., Yao, J.S., Dee, E.C., Mitchell, W.G., Kashyap, S., Giovannini, A., Celi, L.A., Moradi, M.: Chest imagenome dataset for clinical reasoning. In: Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2) (2021)
16. Wu, J.T., Syed, A., Ahmad, H., et al.: Ai accelerated human-in-the-loop structuring of radiology reports. In: Proceedings of the Americal Medical Informatics Association (AMIA) Annual Symposium (2020)
17. Wu, J.T., Wong, K.C.L., Gur, Y., Ansari, N., Karargyris, A., Sharma, A., Morris, M., Saboury, B., Ahmad, H., Boyko, O., Syed, A., Jadhav, A., Wang, H., Pillai, A., Kashyap, S., Moradi, M., Syeda-Mahmood, T.: Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents. Journal of the American Medical Association (JAMA) Network Open **3**(10), e2022779–e2022779 (10 2020)
18. Yu, F., Endo, M., Krishnan, R., Pan, I., Tsai, A., Reis, E.P., Fonseca, E.K.U.N., Lee, H.M.H., Abad, Z.S.H., Ng, A.Y., et al.: Evaluating progress in automatic chest x-ray radiology report generation. Patterns **4**(9) (2023)