Supplementary Material - Detailed Annotations of Chest X-Rays via CT Projection for Report Understanding

Constantin Seibold[1]
constantin.seibold@kit.edu
Simon Reiß[1]
simon.reiss@kit.edu
Saquib Sarfraz[1,2]
muhammad.sarfraz@kit.edu
Matthias A. Fink[3]
matthias.fink@uni-heidelberg.de
Victoria Mayer[3]
victoria.mayer@med.uni-heidelberg.de
Jan Sellner[4]
j.sellner@dkfz-heidelberg.de
Moon Sung Kim[5]
moon-sung.kim@uk-essen.de
Klaus H. Maier-Hein[4]
k.maier-hein@dkfz-heidelberg.de
Jens Kleesiek[4]
jens.kleesiek@uk-essen.de
Rainer Stiefelhagen[1]
rainer.stiefelhagen@kit.edu

[1] Institute of Anthropomatics and Robotics
Karlsruhe Institute of Technology
Karlsruhe, Germany

[2] Autonomous Systems
Daimler TSS
Karlsruhe, Germany

[3] University Hospital Heidelberg
Heidelberg, Germany

[4] Medical Image Computing
German Cancer Research Center
Heidelberg, Germany

[5] Institute for Artificial Intelligence in Medicine
University Clinic Essen
Essen, Germany

1 Dataset: Additional Information

1.1 Visual Comparison between PAXRay and OpenI

We show a more exploratory comparison between our projected (right columns) X-rays and samples of OpenI dataset (left columns) for frontal (top row) and lateral (bottom rows) views in Fig. 1. Here, the position within frontal/lateral corresponds to the same position for the other view of the same patient. The differences in the frontal view stem from the overall difference in shoulder girdle positioning. While in the real images the arms are typically placed alongside the body, the arms are raised in the projected images due to the nature of the CT. In the lateral view, the real images display a larger variety in orientation and pose. Due to the projected images resulting from CTs usually taken while lying the pose between

the different images is similar. This further results in visual differences between images of female patients between the two domains, e.g. 3rd column and second row of real, and first column and second row of projected for both frontal and lateral.

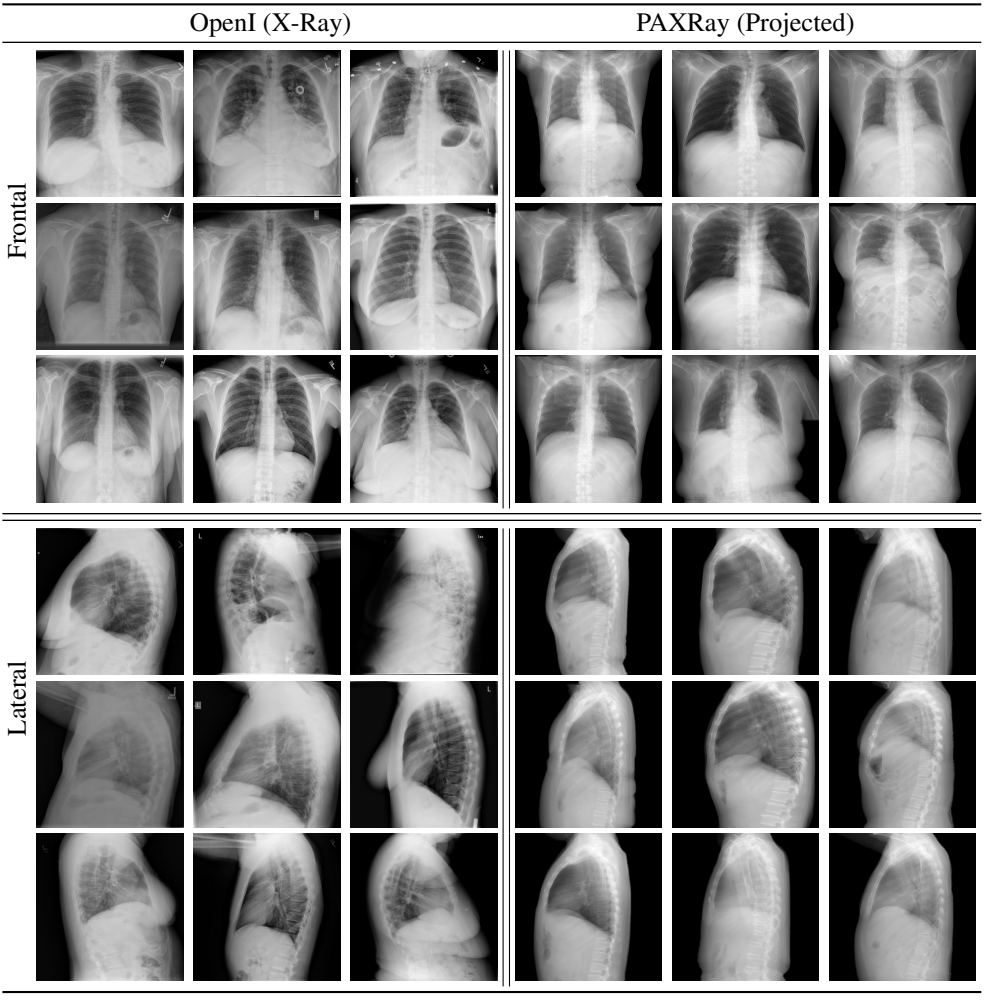

Figure 1: Comparison of different real and projected samples for the frontal and lateral view

1.2 Further Visual Examples of Generated Labels

We show additional examples for the annotations of our dataset in Fig. 2. We can see that lung halves can overlap in the second frontal row. While this might seem contradictory at first, when we consider the lung as a 3D-volume the classes can overlap along one dimension i.e. the lung halves in front and behind the heart as can be seen in Fig. 3. As we intend to capture the entire anatomy present in the upper body we consider this labeling assumption as a more holistic approach.

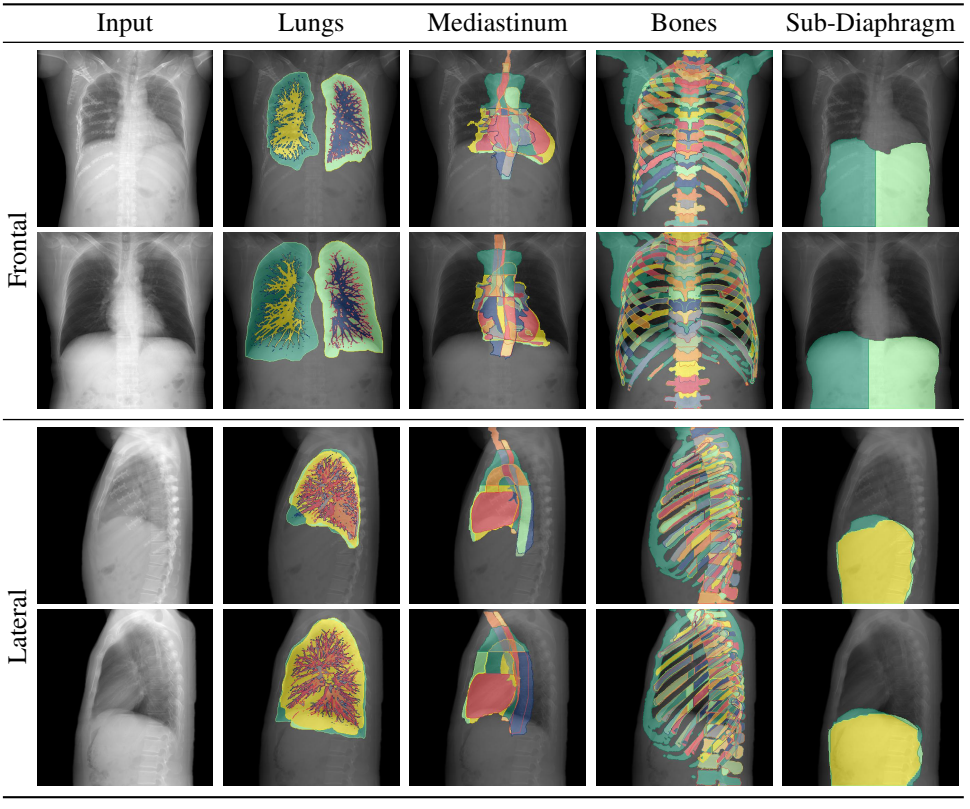

Figure 2: Different ground truth annotations of our PAX-ray dataset for frontal and lateral view

Figure 3: Example of a axial slice of a 3D Volume of the RibFrac dataset, where the two lung halves overlap along the sagittal plane

1.3 Complete Label Set

1. Lung
   (a) Right Lung
      i. Right Lobe Upper
      ii. Right Lobe Middle
      iii. Right Lobe Lower
      iv. Right Lung Vessel
   (b) Left Lung
      i. Left Lobe Upper
      ii. Left Lobe Lower
      iii. Left Lung Vessel

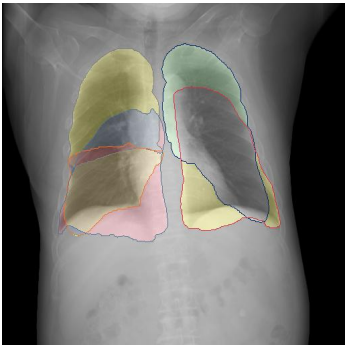

Figure 4: Lung Lobes

2. Mediastinum
   (a) Inferior Mediastinum
      i. Anterior Mediastinum
      ii. Middle Mediastinum
      iii. Posterior Mediastinum
   (b) Superior Mediastinum
   (c) Heart
   (d) Airways
   (e) Esophagus
   (f) Aorta
      i. Ascending Aorta
      ii. Aortic Arch
      iii. Descending Aorta

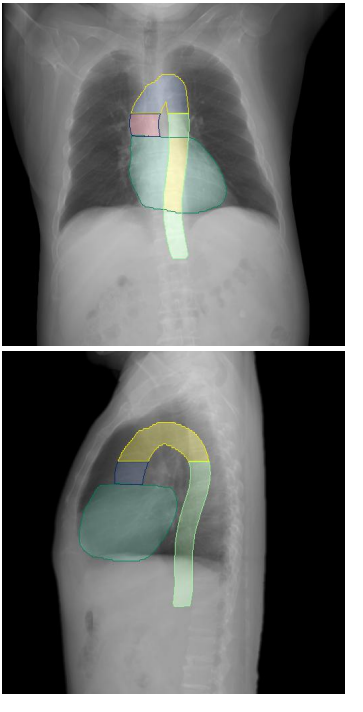

Figure 5: Heart and Aorta

Figure 6: Mediastinum

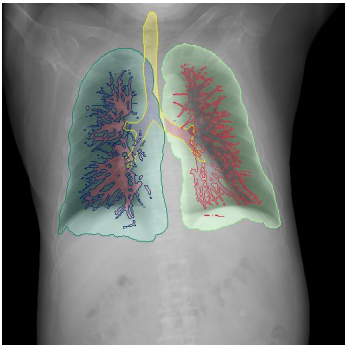

Figure 7: Airways and Vessels

3. Bones
    (a) Spine
        i. Cervical Spine
            A. C1
            B. C2
            C. C3
            D. C4
            E. C5
            F. C6
            G. C7

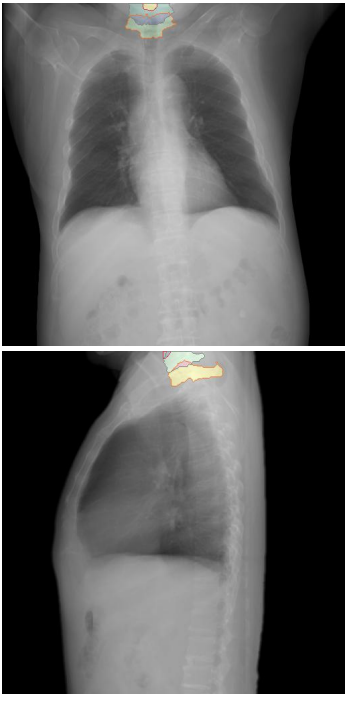

Figure 8: Cervical Spine

i. Thoracic Spine
   A. T1
   B. T2
   C. T3
   D. T4
   E. T5
   F. T6
   G. T7
   H. T8
   I. T9
   J. T10
   K. T11
   L. T12
   M. T13

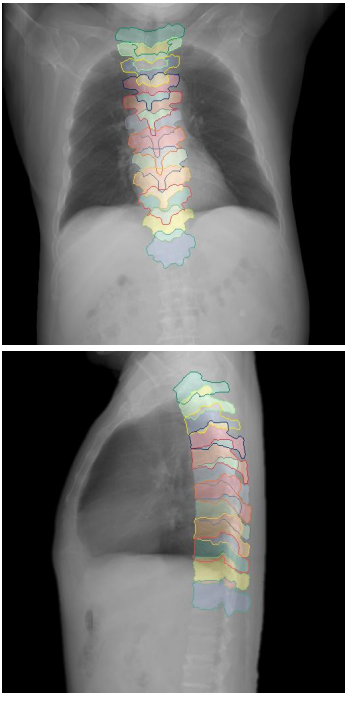

Figure 9: Thoracic Spine

i. Lumbar Spine
   A. L1
   B. L2
   C. L3
   D. L4
   E. L5
ii. Sacrum
iii. Cocygis

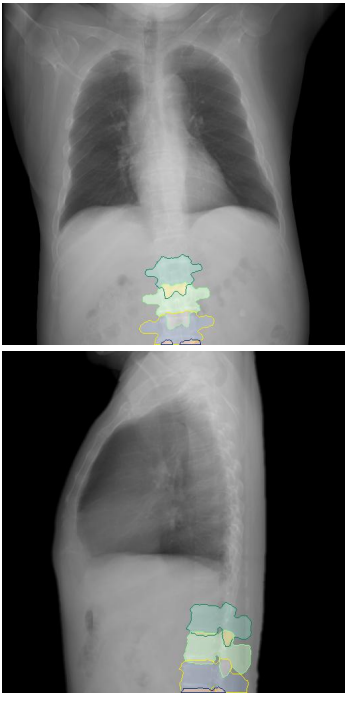

Figure 10: Lumbar Spine

(a) Ribs

i. 1st Rib

A. 1st Rib Posterior
B. 1st Rib Anterior
C. Left 1st Rib
D. Left 1st Rib Posterior
E. Left 1st Rib Anterior
F. Right 1st Rib
G. Right 1st Rib Posterior
H. Right 1st Rib Anterior

ii. 2nd Rib

A. 2nd Rib Posterior
B. 2nd Rib Anterior
C. Left 2nd Rib
D. Left 2nd Rib Posterior
E. Left 2nd Rib Anterior
F. Right 2nd Rib
G. Right 2nd Rib Posterior
H. Right 2nd Rib Anterior

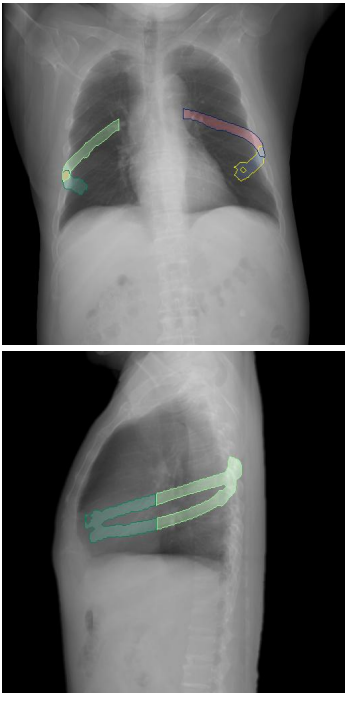

Figure 11: 6th Rib

i. 3rd Rib
   A. 3rd Rib Posterior
   B. 3rd Rib Anterior
   C. Left 3rd Rib
   D. Left 3rd Rib Posterior
   E. Left 3rd Rib Anterior
   F. Right 3rd Rib
   G. Right 3rd Rib Posterior
   H. Right 3rd Rib Anterior

ii. 4th Rib
   A. 4th Rib Posterior
   B. 4th Rib Anterior
   C. Left 4th Rib
   D. Left 4th Rib Posterior
   E. Left 4th Rib Anterior
   F. Right 4th Rib
   G. Right 4th Rib Posterior
   H. Right 4th Rib Anterior

iii. 5th Rib
   A. 5th Rib Posterior
   B. 5th Rib Anterior
   C. Left 5th Rib
   D. Left 5th Rib Posterior
   E. Left 5th Rib Anterior
   F. Right 5th Rib
   G. Right 5th Rib Posterior
   H. Right 5th Rib Anterior

iv. 6th Rib
   A. 6th Rib Posterior
   B. 6th Rib Anterior
   C. Left 6th Rib
   D. Left 6th Rib Posterior
   E. Left 6th Rib Anterior
   F. Right 6th Rib
   G. Right 6th Rib Posterior
   H. Right 6th Rib Anterior

v. 7th Rib
   A. 7th Rib Posterior
   B. 7th Rib Anterior
   C. Left 7th Rib

i. 8th Rib
   A. 8th Rib Posterior
   B. 8th Rib Anterior
   C. Left 8th Rib
   D. Left 8th Rib Posterior
   E. Left 8th Rib Anterior
   F. Right 8th Rib
   G. Right 8th Rib Posterior
   H. Right 8th Rib Anterior

ii. 9th Rib
   A. 9th Rib Posterior
   B. 9th Rib Anterior
   C. Left 9th Rib
   D. Left 9th Rib Posterior
   E. Left 9th Rib Anterior
   F. Right 9th Rib
   G. Right 9th Rib Posterior
   H. Right 9th Rib Anterior

iii. 10th Rib
   A. 10th Rib Posterior
   B. 10th Rib Anterior
   C. Left 10th Rib
   D. Left 10th Rib Posterior
   E. Left 10th Rib Anterior
   F. Right 10th Rib
   G. Right 10th Rib Posterior
   H. Right 10th Rib Anterior

iv. 11th Rib
   A. 1st Rib Posterior
   B. 1st Rib Anterior
   C. Left 11th Rib
   D. Left 11th Rib Posterior
   E. Left 11th Rib Anterior
   F. Right 11th Rib
   G. Right 11th Rib Posterior
   H. Right 11th Rib Anterior

v. 12th Rib
   A. 12th Rib Posterior
   B. 12th Rib Anterior
   C. Left 12th Rib

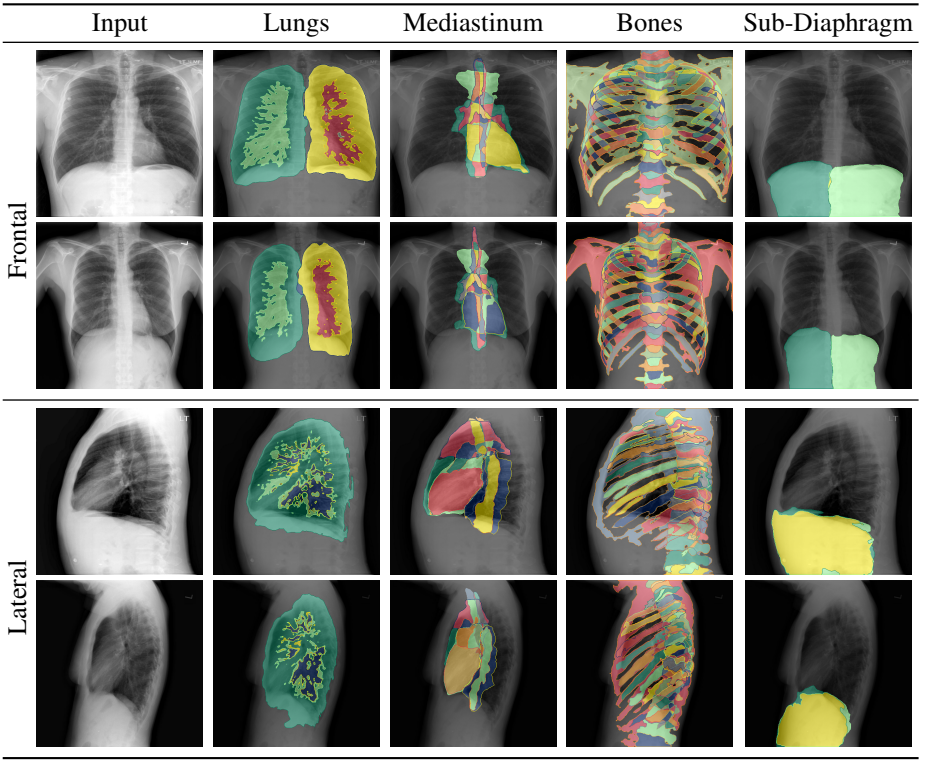

Figure 12: Qualitative results of a UNet with ResNet50 backbone trained on our PAX-ray dataset for multiple patients in OpenI.

4. Diaphragm
   (a) Hemidiaphragm Left
   (b) Hemidiaphragm Right

2 Additional Segmentation Results

2.1 3D Segmentation Results

In this section, we show the training performance of nnUNet [4] across the considered datasets. The training and validation losses as well as the validation performance for all splits of all datasets are displayed in Fig. 13. For SegThor, the final validation performance ranges between 0.91 and 0.93 across all splits. For Verse, the final validation performance ranges between 0.82 and 0.84 across all splits. Hofmanninger's *et al.* [3] report a dice score for their lung segmentation model of 0.99, 0.94, and 0.98 on the LTRC [1], LCTSC [7], and VESS12 [6] datasets respectively. Koitka *et al.* [5] show a performance for their BCA of 0.96 dice on their internal dataset.

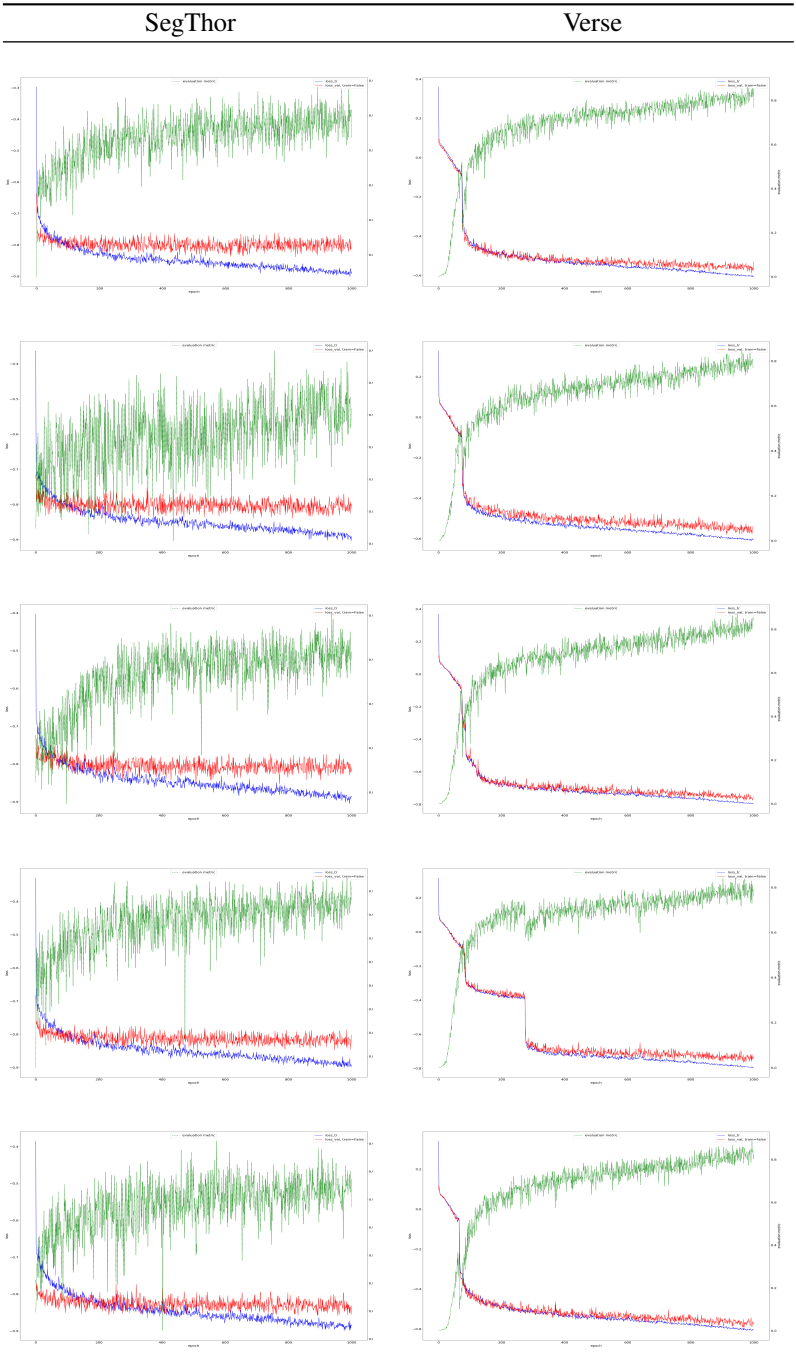

Figure 13: Visualization of training (Blue) and validation (Red) losses as well as validation performance (Green). For SegThor, the final validation performance ranges between 0.91 and 0.93 across all splits. For Verse, the final validation performance ranges between 0.82 and 0.84 across all splits.

2.2 2D Segmentation Results

We show the segmentation performance on the PAXRay dataset for all classes in Table 2-7.

2.3 Further Visual Examples of Predictions on Real X-Rays

We show additional examples for network predictions on the OpenI dataset for both frontal and lateral views in Fig. 12. The segmentation results for the mediastinal classes, the subdiaphragm, lung regions as well as most of the spine and ribs appear quite fitting for both views. One can notice that the prediction of individual ribs at times can be lacking despite the overall bone structure being segmented fine. Such errors can be seen in the overlapping areas of heart and ribs, possible attributable to the overall brighter region. Lungs vessel predictions in comparison to the annotations are quite coarse which is due to vessels being hardly visible in a chest X-ray.

3 Grounding Dataset

For our evaluation of medical phrase grounding, we use the OpenI dataset [7] which consists of medical reports paired with frontal and lateral chest X-rays. When creating the validation dataset for CXR phrase grounding we tried to incorporate a large variety of phrases and anomalies. We tasked two radiologist to highlight phrases within 100 medical reports for both the lateral and frontal (Anterior-Posterior/Posterior-Anterior) view in the OpenI dataset resulting in 178 frontal and 146 lateral bounding box annotations. We use the joined annotations as ground truth for the task. Following the NER-results of Stanza, these 100 reports resulted in annotated phrases with 90 different observations, 70 different observation modifiers, 88 different anatomies and 42 different anatomy modifiers.

References

[1] https://ltrcpublic.com/.

[2] Dina Demner-Fushman, Marc D Kohli, Marc B Rosenman, Sonya E Shooshan, Laritza Rodriguez, Sameer Antani, George R Thoma, and Clement J McDonald. Preparing a collection of radiology examinations for distribution and retrieval. *Journal of the American Medical Informatics Association*, 23(2):304–310, 2016.

[3] Johannes Hofmanninger, Forian Prayer, Jeanny Pan, Sebastian Röhrich, Helmut Prosch, and Georg Langs. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. *European Radiology Experimental*, 4 (1):1–13, 2020.

[4] Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2):203–211, 2021.

[5] Sven Koitka, Lennard Kroll, Eugen Malamutmann, Arzu Oezcelik, and Felix Nensa. Fully automated body composition analysis in routine ct imaging using 3d semantic segmentation convolutional neural networks. *European radiology*, 31(4):1795–1804, 2021.

Table 1: PAXRay Segmentation Performance for UNet and SegFPN with a ResNet50 backbone in IoU in % (1/6)

	Lungs	Right Lung	Left Lung
UNet	95.44	90.63	82.73
SegFPN	94.91	90.2	82.09
	Left Upper Lobe	Left Lower Lobe	Right Upper Lobe
	89.08	85.23	88.11
	88.05	85.17	87.33
	Right Middle Lobe	Right Lower Lobe	Right Lung Vessel
	81.68	87.46	54.65
	79.6	86.88	55.65
	Left Lung Vessel	Mediastinum	Inferior Mediastinum
	47.28	91.98	89.56
	48.37	90.76	87.77
	Superior Mediastinum	Anterior Mediastinum	Middle Mediastinum
	85.4	54.47	89.17
	81.89	50.22	87.49
	Posterior Mediastinum	Heart	Airways
	81.34	89.73	74.12
	79.22	88.76	70.09
	Esophagus	Aorta	Ascending Aorta
	68.2	81.2	62.49
	65.19	78.24	59.94
	Aortic Arch	Descending Aorta	Bones
	77.1	49.29	90.87
	71.53	45.99	87.77
	Spine	c1	c2
	92.6	0.0	0.0
	90.38	0.0	0.0
	c3	c4	c5
	0.0	1.71	11.05
	0.0	00.62	8.18

Table 2: PAXRay Segmentation Performance (2/6)

c6	c7	t1
37.83 30.45	67.71 55.54	78.87 67.88
t2	t3	t4
79.2 69.61	77.95 68.22	78.01 67.35
t5	t6	t7
77.96 65.37	78.05 64.86	77.92 64.73
t8	t9	t10
77.48 65.58	77.89 65.21	78.66 65.37
t11	t12	l1
79.05 66.18	80.16 66.58	82.34 66.32
l2	l3	l4
81.35 64.36	76.28 60.97	47.52 33.68
l5	l6	sacrum
10.94 8.4	0.0 0.0	0.0 0.0
cocygis	t13	ribs
0.0 0.0	0.0 0.0	84.98 78.72
Rib - 1	Rib - 2	Rib - 3
76.43 70.31	73.17 61.44	70.99 57.1
Rib - 4	Rib - 5	Rib - 6
72.6 56.97	70.12 54.81	70.61 55.04

Table 3: PAXRay Segmentation Performance (3/6)

Rib - 7	Rib - 8	Rib - 9
70.7 56.19	71.4 55.85	72.19 56.89
Rib - 10	Rib - 11	Rib - 12
71.77 57.83	59.75 47.93	54.64 43.11
Rib - Anterior -1	Rib - Posterior -1	Rib - Anterior -2
72.36 68.09	71.61 64.2	68.9 57.34
Rib - Posterior -2	Rib - Anterior -3	Rib - Posterior -3
73.35 62.73	66.23 52.7	73.35 60.97
Rib - Anterior -4	Rib - Posterior -4	Rib - Anterior -5
67.39 52.28	74.86 61.38	63.06 48.69
Rib - Posterior -5	Rib - Anterior -6	Rib - Posterior -6
73.6 60.37	63.96 48.63	71.48 58.12
Rib - Anterior -7	Rib - Posterior -7	Rib - Anterior -8
61.29 47.53	71.3 57.65	59.96 46.81
Rib - Posterior -8	Rib - Anterior -9	Rib - Posterior -9
72.83 57.68	60.56 46.57	72.67 57.71
Rib - Anterior -10	Rib - Posterior -10	Rib - Anterior -11
54.3 40.88	70.42 57.98	9.54 4.1
Rib - Posterior -11	Rib - Anterior -12	Rib - Posterior -12
59.65 48.01	4.79 1.27	53.23 41.58

Table 4: PAXRay Segmentation Performance (4/6)

Rib - Left - 1	Rib - Right - 1	Rib - Left - 2
67.22 63.02	68.72 63.59	62.72 53.12
Rib - Right - 2	Rib - Left - 3	Rib - Right - 3
65.49 55.17	61.35 49.12	61.3 49.99
Rib - Left - 4	Rib - Right - 4	Rib - Left - 5
61.12 48.29	62.71 50.89	57.52 44.62
Rib - Right - 5	Rib - Left - 6	Rib - Right - 6
61.17 49.57	59.32 45.67	62.55 50.27
Rib - Left - 7	Rib - Right - 7	Rib - Left - 8
59.8 46.84	63.92 51.72	60.29 46.39
Rib - Right - 8	Rib - Left - 9	Rib - Right - 9
63.31 51.95	61.27 48.18	62.39 51.52
Rib - Left - 10	Rib - Right - 10	Rib - Left - 11
59.04 46.69	61.27 52.6	46.06 35.51
Rib - Right - 11	Rib - Left - 12	Rib - Right - 12
52.32 43.08	42.28 35.04	46.77 34.1
Rib - Left - Anterior -1	Rib - Left - Posterior -1	Rib - Left - Anterior -2
65.15 61.68	60.8 56.58	65.12 60.91
Rib - Left - Posterior -2	Rib - Left - Anterior -3	Rib - Left - Posterior -3
65.14 59.28	59.25 48.4	61.25 54.35

Table 5: PAXRay Segmentation Performance (5/6)

Rib - Left - Anterior -4	Rib - Left - Posterior -4	Rib - Left - Anterior -5
62.13 50.95	65.1 55.9	58.13 44.78
Rib - Left - Posterior -5	Rib - Left - Anterior -6	Rib - Left - Posterior -6
60.12 50.25	56.42 45.63	64.76 54.84
Rib - Left - Anterior -7	Rib - Left - Posterior -7	Rib - Left - Anterior -8
57.04 44.34	62.28 52.02	58.09 46.96
Rib - Left - Posterior -8	Rib - Left - Anterior -9	Rib - Left - Posterior -9
65.33 55.15	50.97 39.15	61.23 50.43
Rib - Left - Anterior -10	Rib - Left - Posterior -10	Rib - Left - Anterior -11
55.18 45.3	63.21 53.55	53.43 39.72
Rib - Left - Posterior -11	Rib - Left - Anterior -12	Rib - Left - Posterior -12
60.69 47.98	57.17 45.03	62.79 52.99
Rib - Right - Anterior -1	Rib - Right - Posterior -1	Rib - Right - Anterior -2
51.21 38.86	60.34 48.28	54.15 42.56
Rib - Right - Posterior -2	Rib - Right - Anterior -3	Rib - Right - Posterior -3
65.62 53.87	48.55 37.94	62.19 48.39
Rib - Right - Anterior -4	Rib - Right - Posterior -4	Rib - Right - Anterior -5
52.04 40.85	65.25 53.91	48.98 35.43

Table 6: PAXRay Segmentation Performance (6/6)

Rib - Right - Posterior -5	Rib - Right - Anterior -6	Rib - Right - Posterior -6
62.69 49.37	50.69 39.66	63.8 53.08
Rib - Right - Anterior -7	Rib - Right - Posterior -7	Rib - Right - Anterior -8
36.76 26.17	58.41 46.67	43.77 34.46
Rib - Right - Posterior -8	Rib - Right - Anterior -9	Rib - Right - Posterior -9
61.3 52.86	5.13 02.42	45.54 35.21
Rib - Right - Anterior -10	Rib - Right - Posterior -10	Rib - Right - Anterior -11
6.41 03.1	50.9 41.91	0.27 0.0
Rib - Right - Posterior -11	Rib - Right - Anterior -12	Rib - Right - Posterior -12
37.64 30.32	0.11 0.0	43.84 31.95
Diaphragm	Hemidiaphragm - Right	Hemidiaphragm - Left
96.8 96.15	94.08 93.87	93.67 93.25
Mean		
60.69 51.87		

[6] Rina D Rudyanto, Sjoerd Kerkstra, Eva M Van Rikxoort, Catalin Fetita, Pierre-Yves Brillet, Christophe Lefevre, Wenzhe Xue, Xiangjun Zhu, Jianming Liang, Ilkay Öksüz, et al. Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the vessel12 study. *Medical image analysis*, 18(7):1217–1232, 2014.

[7] Jinzhong Yang, Greg Sharp, Harini Veeraraghavan, Wouter van Elmpt, Andre Dekker, Tim Lustberg, and Mark Gooding. Data from lung ct segmentation challenge. *The cancer imaging archive*, 20, 2017.