Pneumothorax Detection and Localization in X-Ray Images Given Richer Annotation Information

Qian Zhao PhD[1], Min Zhang PhD[1], Gopal Avinash PhD[1],
Vasantha Venugopal MD[2], Vidur Mahajan MBBS[2], Clinton Mielke PhD[3], Scott Hammond MD[3], and Rachel Calcutt MD[3]

[1] GE Healthcare, Milwaukee, WI, USA
[2] CARING, Mahajan Imaging, New Delhi, INDIA
[3] The University of California, San Francisco (UCSF)

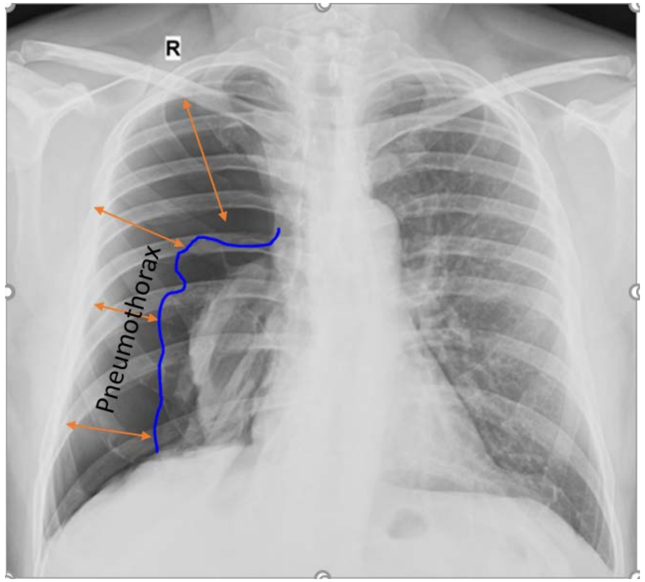
#CMIMI18

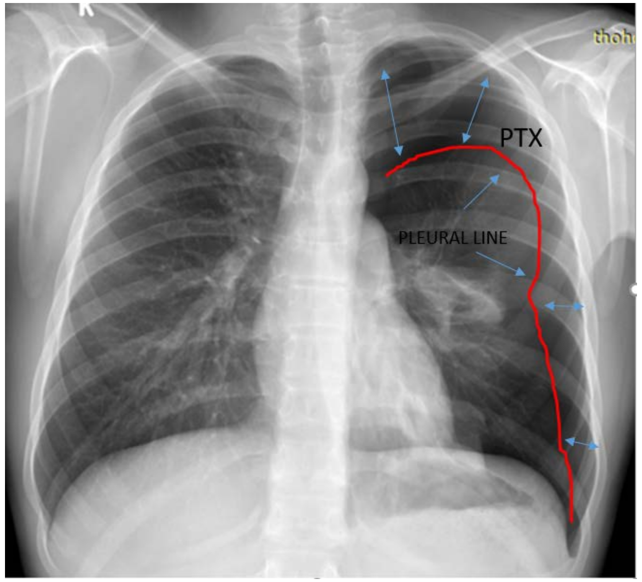
Outline

- Background
- Model Architecture and Design
- Data
- Experiments
- Results
- Conclusions

Pneumothorax

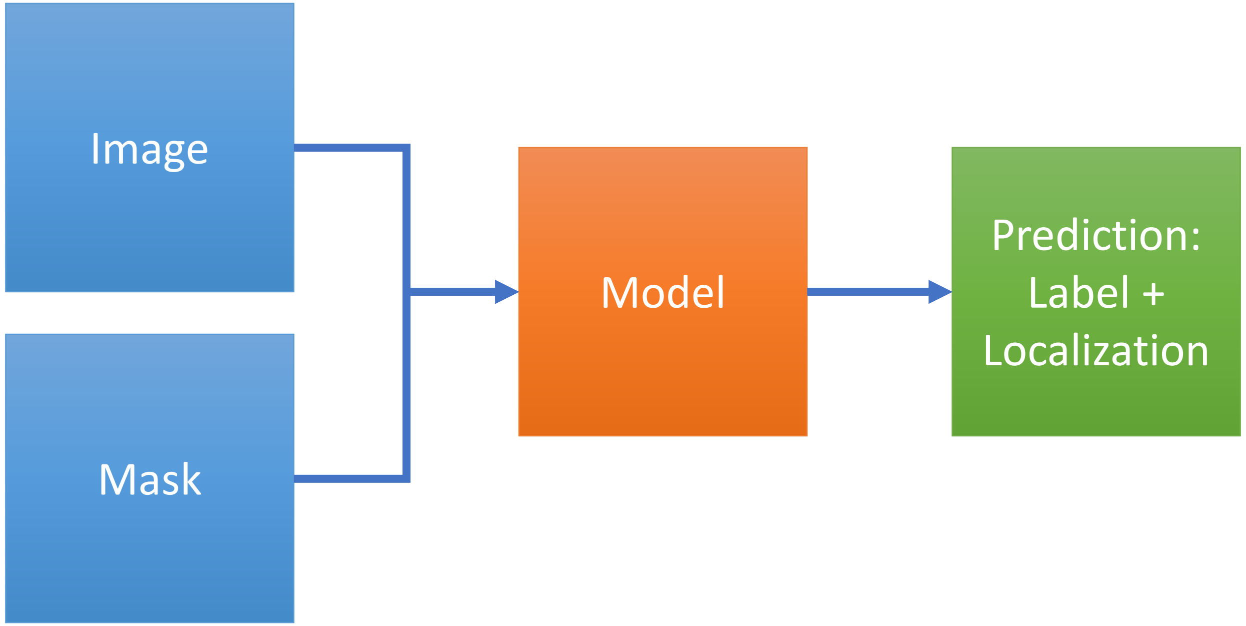
- Pneumothorax refers to collection of air in the pleural cavity.
- Large Pneumothorax especially those of the constantly enlarging type (tension pneumothorax) causing compression of mediastinal structures can be life-threatening
- Chest radiography is the first line of investigation- shows the outer margin of the visceral pleura (and lung) separated from the parietal pleura (and chest wall) by a lucent gas space devoid of pulmonary vessels.





Problem Statement

- How can we improve the performance of a classifier if masks pertaining to regions of interest are available in addition to an image level label?



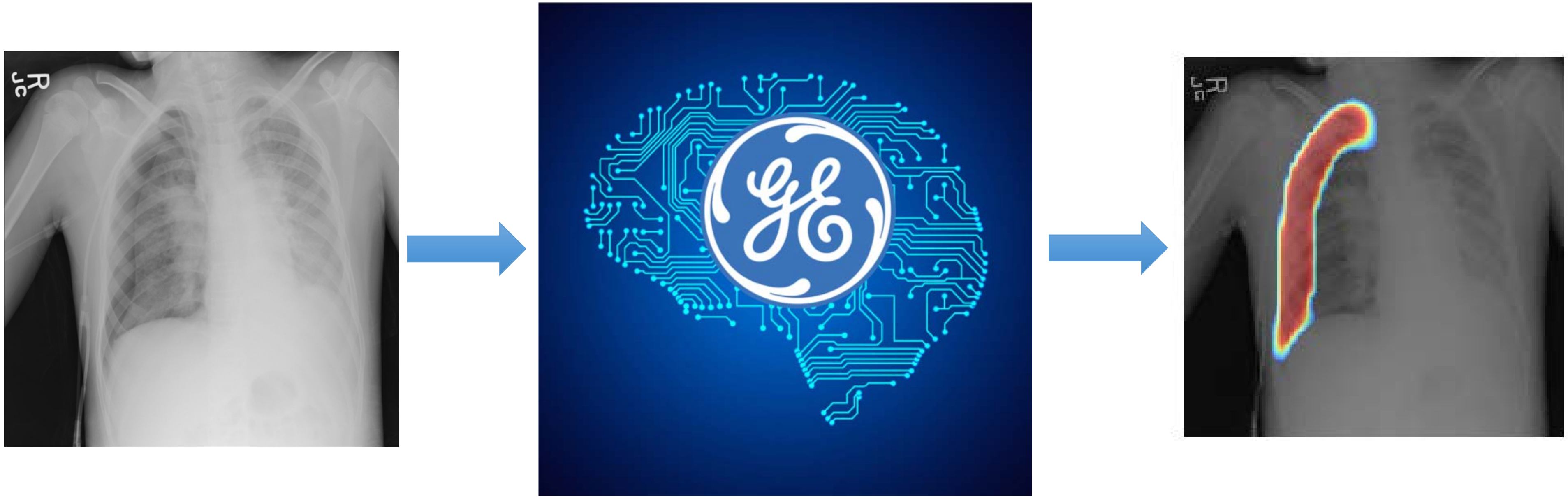
Related Work

- Most of the published literatures are improving the performance of a classifier by using more data.
- They focus to solve large data problem
- Li et. al (Google) proposed a semi-supervised method to recognize and localize thoracic diseases using bounding box

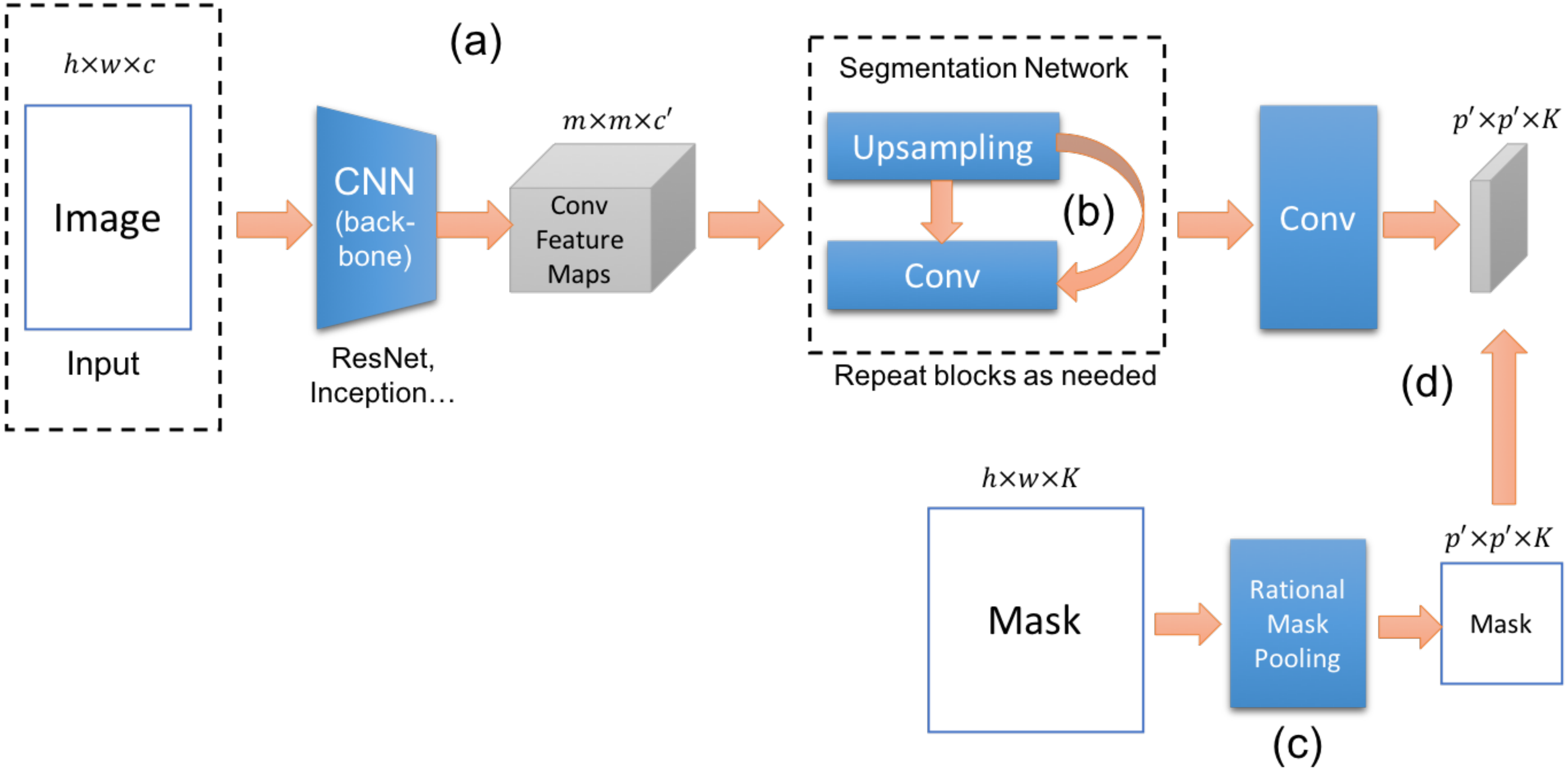
Li et. al, Thoracic Disease Identification and Localization with Limited Supervision, [arXiv: 1711.06373](https://arxiv.org/abs/1711.06373).

Contributions

- Use mask as input of classifier to improve the classification accuracy
- Generate high resolution localization maps



Model Architecture -- ClaSegNet



Model Design

- Backbone CNN outputs shared feature maps
  - - ResNet
  - - Inception
  - - …
- Loss function

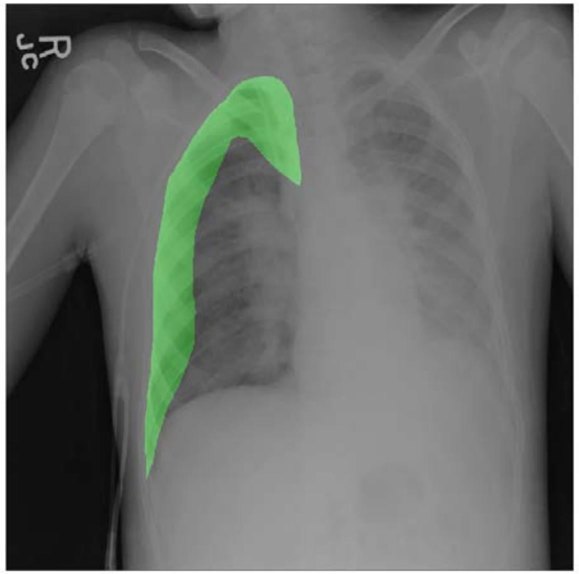
$$Lossmask = -\sum_{i} log\left(p(y_k|x_i, mask_i^k)\right)$$

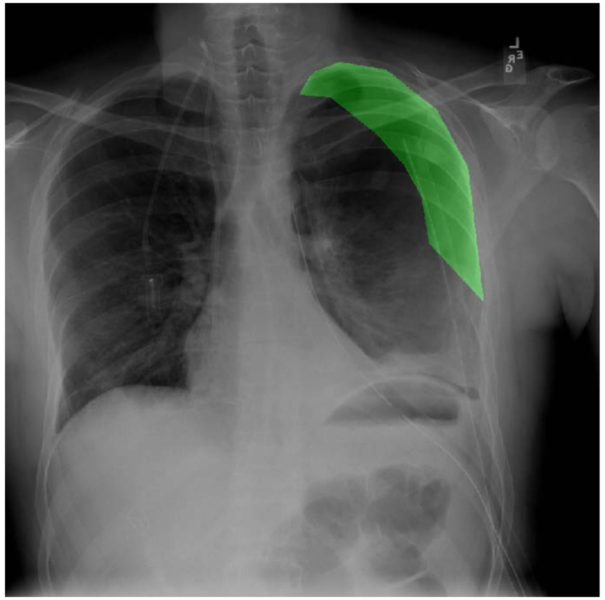
$$p(y|x_i, mask_i^k) = \prod_{j\in N} p_{ij}^k \cdot \prod_{j\in M\setminus \mathrm{N}} (1 - p_{ij}^k)$$

where $p(y|x_i, mask_i^k)$ is the probability of an image i being positive for class k, M the total area in image i and N the region covered by the mask.

Data and Annotations

NIH Data (# images)	Train (80%)	Val (10%)	Test (10%)
Pneumothorax	722	90	91
Non-pneumothorax	722	90	91
Total	1444	180	182





#CMIMI18

Experiments

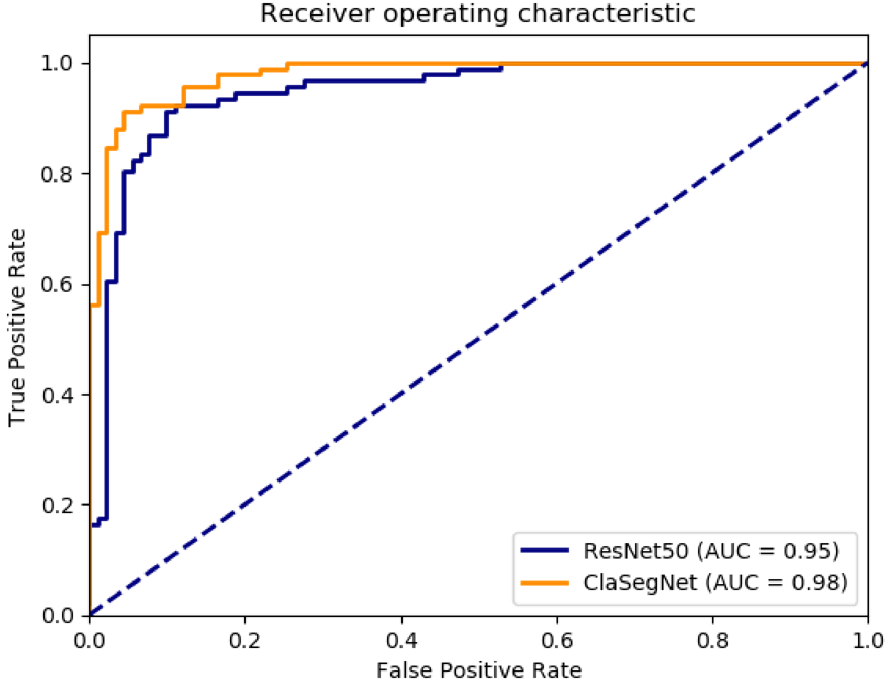
- Training
  1. Input image size is 512 x 512
  2. Normalized input image to [0, 1]
  3. Data augmentation (rotation, shift and horizontal flip)
  4. Initial learning rate is 0.0001
  5. Adam optimizer + learning rate decay
  6. Early-stopping
  7. Pre-trained backbone network
- Inferencing
  1. Input image size is 512 x 512
  2. Normalized input image to [0, 1]
  3. No data augmentation

Results

Model	val Accuracy	val Precision	val Recall	val AUC	val Dice
ResNet50	0.894	0.899	0.889	0.941	-
ClaSegNet	0.95	0.966	0.933	0.98	0.518

Model	test Accuracy	test Precision	test Recall	test AUC	test Dice
ResNet50	0.896	0.875	0.923	0.945	-
ClaSegNet	0.923	0.953	0.89	0.979	0.5

Test ROC

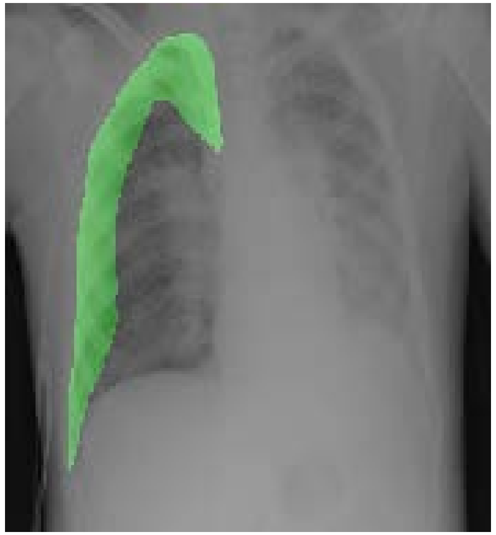


ResNet50		Predictions	
		nonPTX	PTX
Ground Truth	nonPTX	77	14
	PTX	8	83

ClaSegNet		Predictions	
		nonPTX	PTX
Ground Truth	nonPTX	87	4
	PTX	10	81

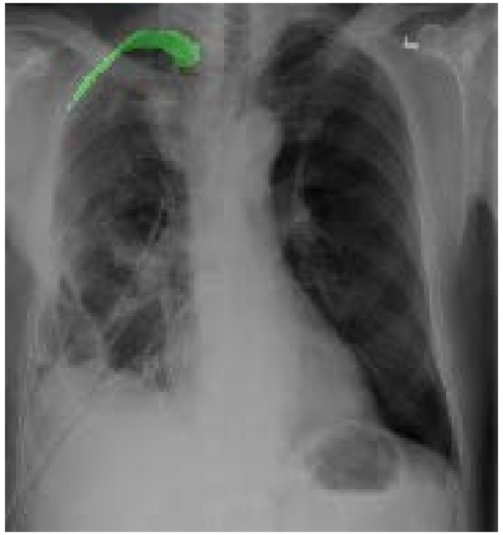
Localization maps -- TP

ground truth



prediction: 1.000

ground truth



prediction: 0.966

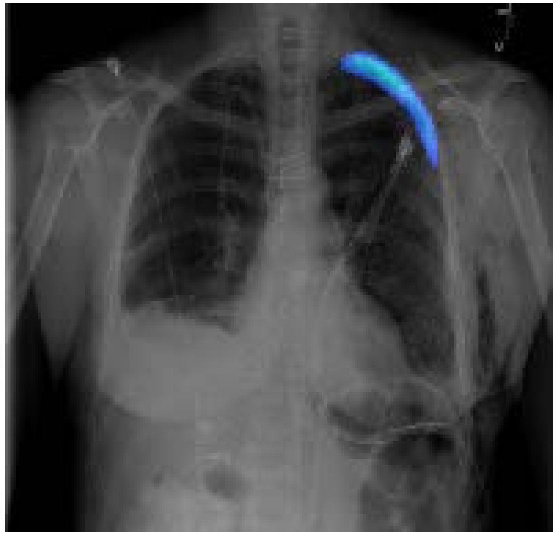
Localization maps -- FN

ground truth

prediction: 0.269

ground truth

prediction: 0.392

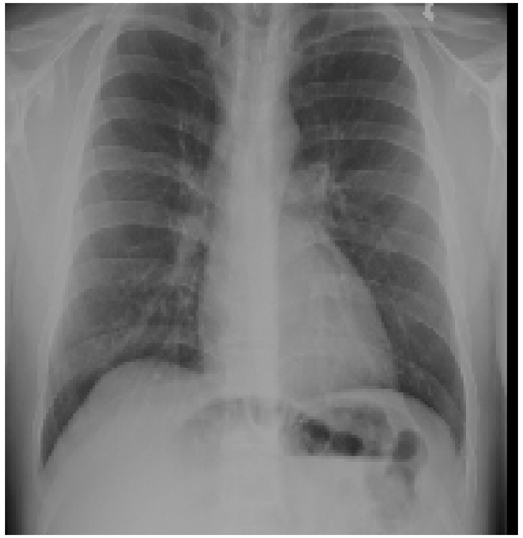


Localization maps -- FP

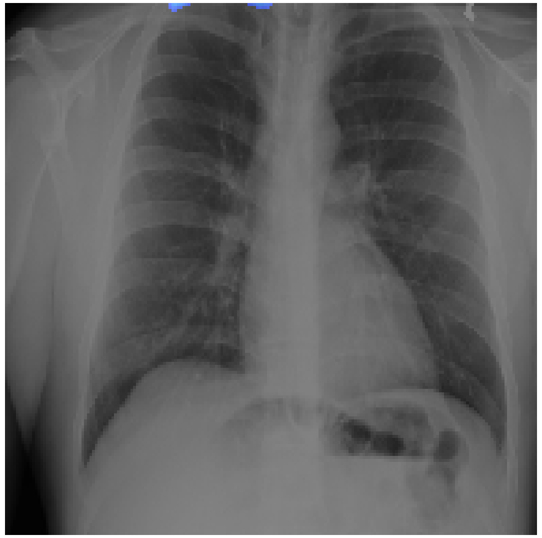
ground truth

prediction: 0.615

ground truth

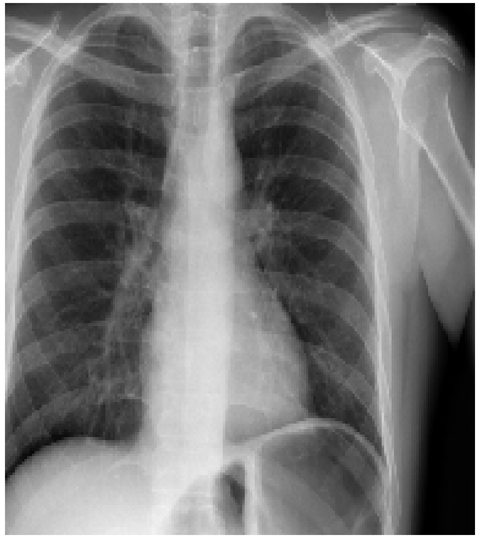


prediction: 0.526

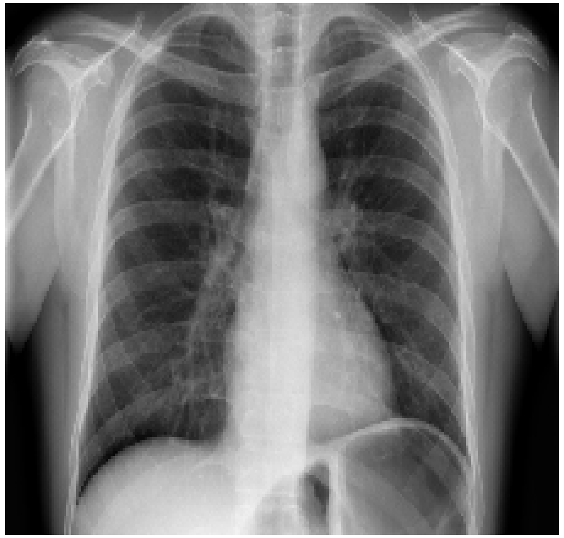


Localization maps -- TN

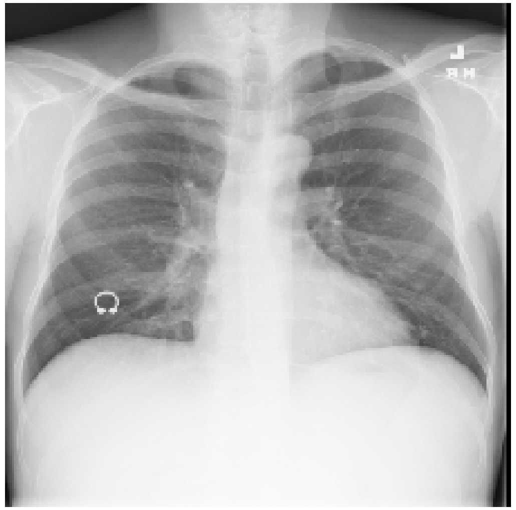
ground truth



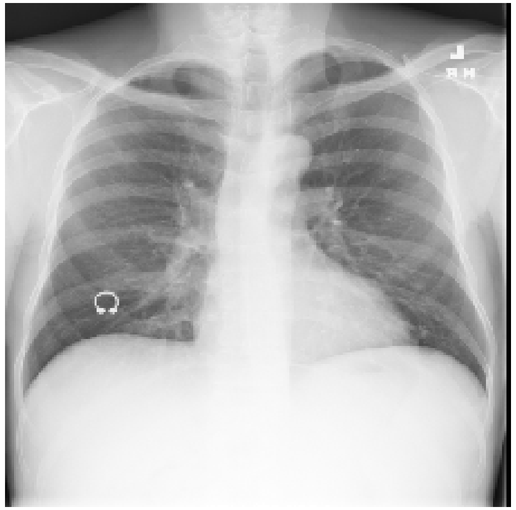
prediction: 0.021



ground truth



prediction: 0.008



Conclusions

- Utilizing richer annotation information (mask) improves the classification and localization accuracy
- Adding segmentation branch (decoder) improves the localization accuracy (resolution)

Acknowledgements

- We would like to acknowledge the contributions of:

- GEHC colleagues Zita Herczeg, Pal Tegzes, Gireesha Rao, Katelyn Nye, Monika Jain, Litao Yan for their support on developing, testing, training on the annotation platform, format conversions, and strategy; and
- Dr Rachel Callcut, Dr Scott Hammond, and the UCSF Center for Digital Health Innovation for ongoing AI collaborations with GEHC, and Drs John Mongan and Andrew Taylor (UCSF Radiology) for clinical insights and pneumothorax use-case refinement.
- Radiologist team from Mahajan Imaging for data annotations on the NIH data set.

Thank you!