Application of Artificial Intelligence to Help Detect Rib Fractures in Chest X-Rays

Gaoxiang Luo*, Andrew Walker*, Ju Sun
University of Minnesota - Twin Cities

Abstract

We intend to develop an accurate and reliable automated rib fracture detection method, based on chest x-rays and CT scans, that can be deployed in clinical environments to reduce the delays and errors that come along with the current practice of manually identifying fractures. This is important because nearly every two out of three patients who experience chest trauma have rib fractures, and they are often associated with significant morbidity and mortality. But oftentimes, clinicians wait for the critical imaging results from radiologists to inform the next decision. The motivation is to have AI help radiologists diagnose the non-complicated part of medical images to make the entire healthcare system more efficient. This research is challenging because rib fractures are elongated and sparse in space, and there hasn't been much related literature on this AI use case. We build an end-to-end deep learning classification pipeline for chest x-rays that achieves 82.7% test AUROC. We also generate heatmaps of critical regions, where AI thinks is important for decision making, to radiologists to further validate our model's reliability. Moving forward, we are incorporating electronic health records (EHR) as well as limited annotation of fractures to further improve our model.

Data Preparation

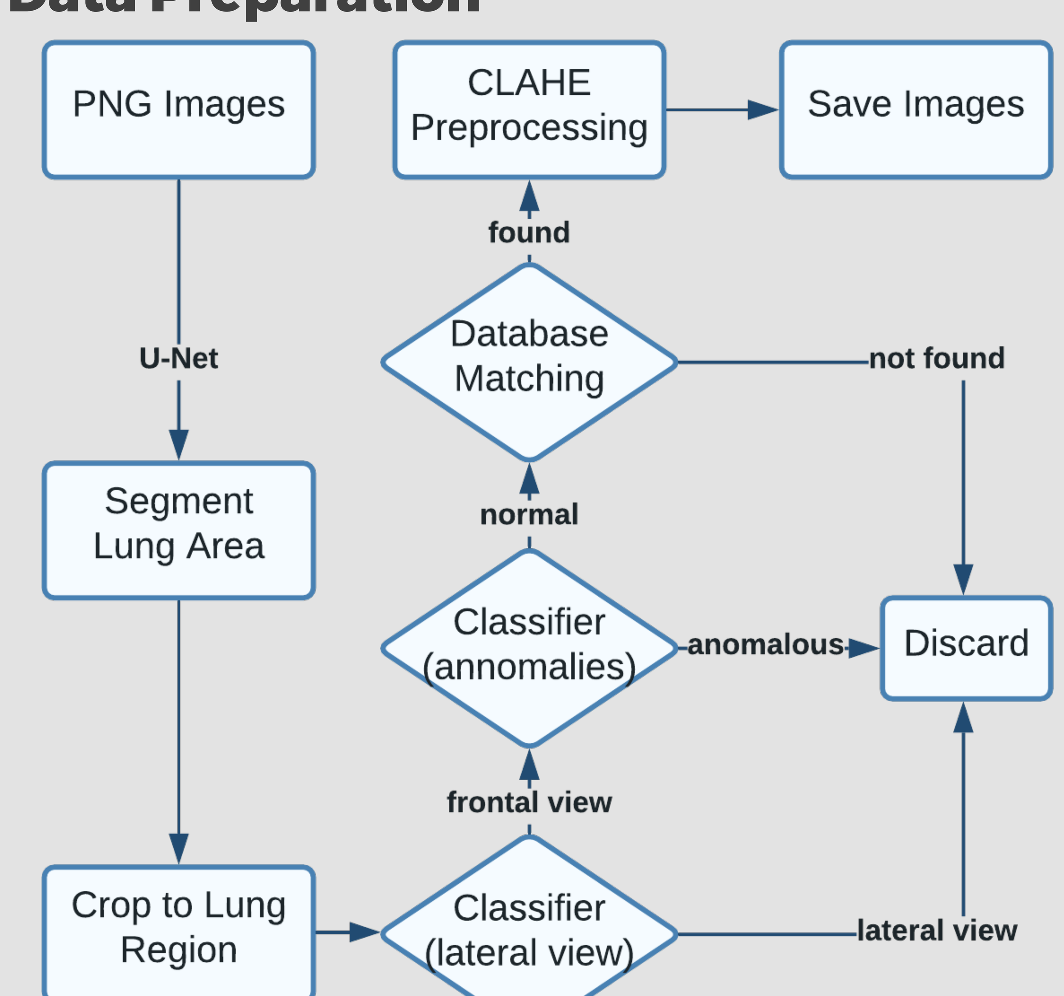


Figure 1: This module ensures the quality of the input data by removing data that is incorrectly formatted. It detects the lung region on the x-ray with a U-Net (Ronneberger, 2015) and crops to that area before feeding into the following submodules. From here, a Resnet-18 model is used to discard lateral view images, a conditional-GAN (Mirza, 2014) is used to discard other anomalous images, and the label is double-checked against the database.

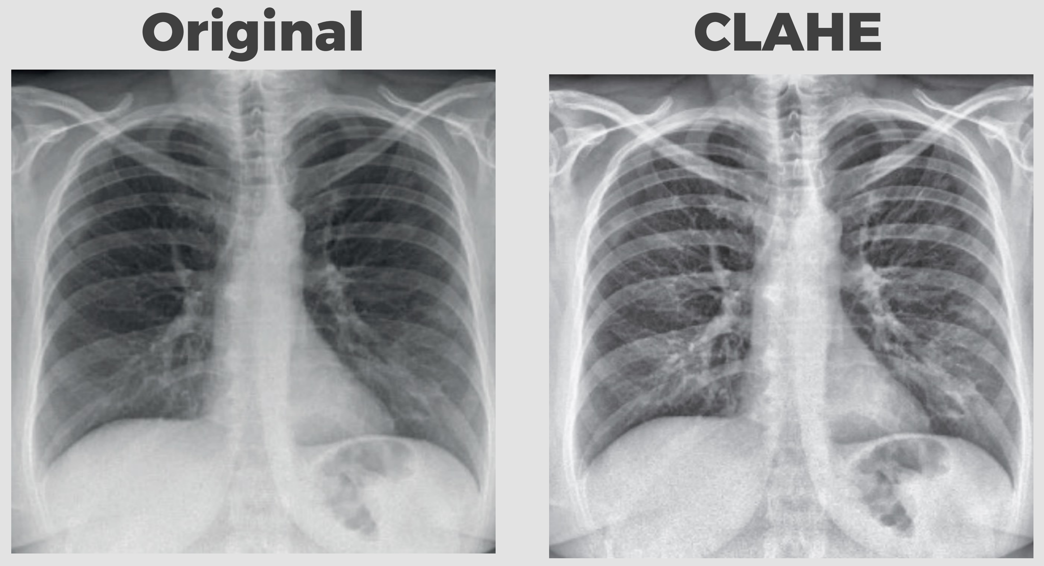
Introduction

Rib fractures is the most common injury after chest trauma, occurring in 40-80% of patients (Lin, 2016). Radiologists are often required to quickly read images (e.g., x-ray and CT) and identify life-threatening injuries. X-ray images alone can miss more than 85% of rib fractures, and CT scans (the current standard) can miss 20.7% (Cho, 2012). Proper detection of rib fractures can lead to different interventions for the patient; in one study, using CT instead of X-ray led to changes to more appropriate interventions for 20% of patients.

At M Health Fairview, patients with rib fractures have the poorest risk-adjusted outcomes in the regional Midwest trauma center benchmarking (Macheel, 2020). They desire an automated tool to help deliver the proper interventions to patients, as adherence to evidence-based practices can reduce mortality in rib fracture cases by 3x. An automated rib fracture detection tool can shorten diagnosis time, reduce workload, and improve accuracy (Zhou, 2020).

CLAHE

Figure 2: CLAHE (Contrast Limited Adaptive Histogram Equalization) is an image normalization tool used to adaptively boost contrast. It improves local contrast in an image and limits the over-amplification of noise in homogenous regions compared to adaptive histogram equalization. It is often used on digital x-ray images to improve edge definition. CLAHE improved the AUPRC about 4% in our preliminary experiment on CheXpert dataset for rib fracture classification. Image credit to Saiz (2020).



Dataset

	Train	Val	Test
Positive	1815	456	550
Negative	10240	2558	3200

Table 1: Among the positive train data points, 70 of them have segmentation annotations.

Model

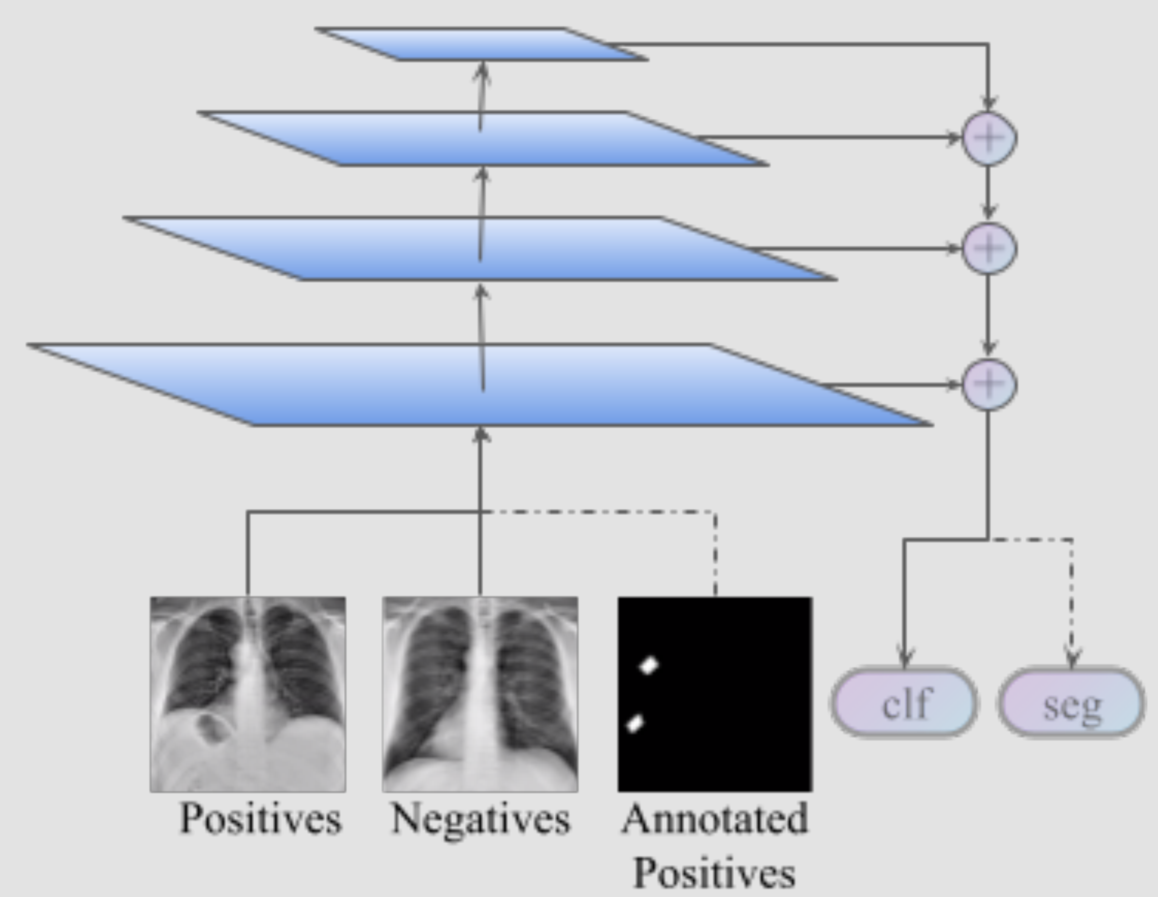


Figure 3: Our implementation of multi-task feature pyramid network (Lin, 2017).

Results

Table 2: The performance of DenseNet 121 if we perform center crop during preprocessing.

	DenseNet121 not use center crop	DenseNet121 use center crop
AUROC	**0.827**	0.750
AUPRC	**0.544**	0.352

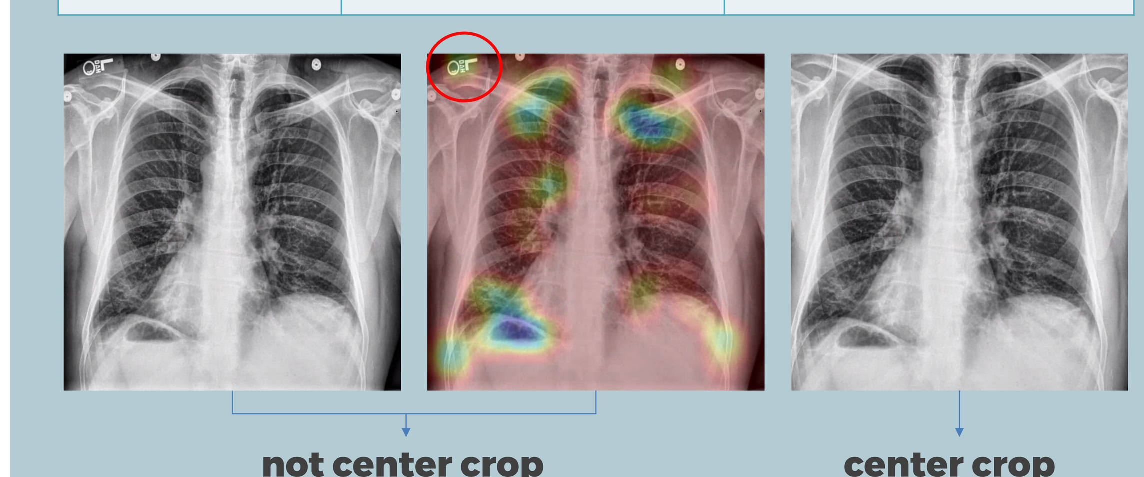


Table 3: The performance of feature pyramid network with ResNet-50 (He, 2015) backbone on classification head only, and on classification head as well as segmentation head.

	ResNet50-FPN clf only	ResNet50-FPN clf and seg
AUROC	0.663	**0.701**
AUPRC	0.232	**0.276**

Note: Although this is not our best performing model, it shows the benefit of extra supervision coming from the annotation of radiologists.

Heatmaps (human in the loop)

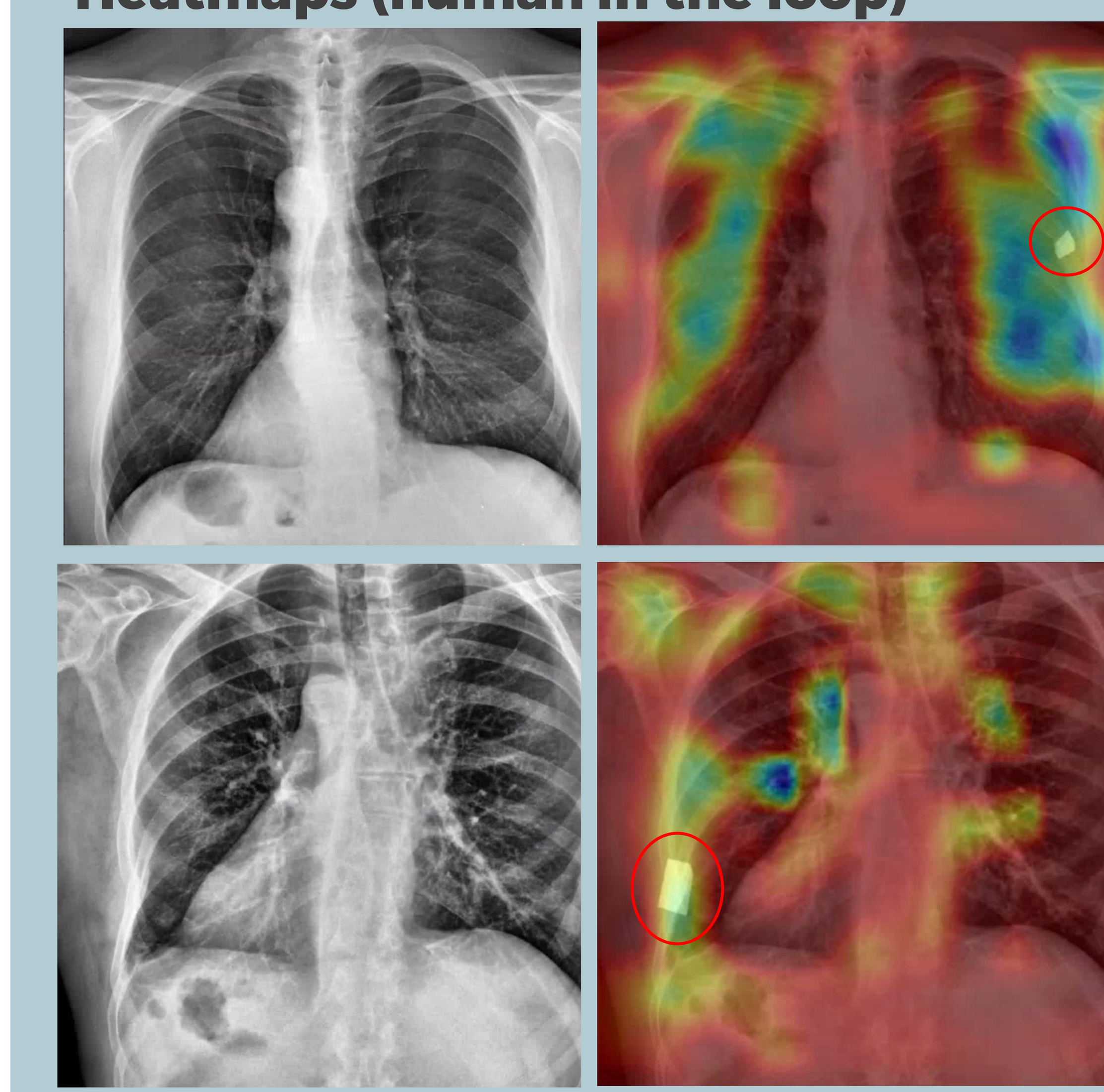


Figure 4: The heatmaps generated from the last activation layer of DenseNet121 using GradCAM++ (Chattopadhay, 2018). The blue areas are the regions of high responses (i.e., the model pays attention in decision making). The white area is where the rib fracture is located, annotated by a radiologist.

Reference

[1] Lin, F. C., Li, R. Y., Tung, Y. W., Jeng, K. C., & Tsai, S. C. (2016). Morbidity, mortality, associated injuries, and management of traumatic rib fractures. Journal of the Chinese Medical Association : JCMA, 79(6), 329–334.
[2] Cho, S. H., Sung, Y. M., & Kim, M. S. (2012). Missed rib fractures on evaluation of initial chest CT for trauma patients: pattern analysis and diagnostic value of coronal multiplanar reconstruction images with multidetector row CT. The British journal of radiology, 85(1018), e845–e850.
[3] Macheel, C., Reicks, P., Sybrant, C., Evans, C., Farhat, J., West, M. A., & Tignanelli, C. J. (2020). Clinical Decision Support Intervention for Rib Fracture Treatment. Journal of the American College of Surgeons, 231(2), 249–256.e2.
[4] Zhou, Q. Q., Wang, J., Tang, W., Hu, Z. C., Xia, Z. Y., Li, X. S., Zhang, R., Yin, X., Zhang, B., & Zhang, H. (2020). Automatic Detection and Classification of Rib Fractures on Thoracic CT Using Convolutional Neural Network: Accuracy and Feasibility. Korean journal of radiology, 21(7), 869–879.
[5] Pizer, S. M., Johnston , R. E., Ericksen, J. P., Yankaskas, B. C. and Muller, K. E., "Contrast-limited adaptive histogram equalization: speed and effectiveness," [1990] Proceedings of the First Conference on Visualization in Biomedical Computing, 1990, pp. 337-345, doi: 10.1109/VBC.1990.109340.
[6] Saiz, F., & Barandiaran, I. (04 2020). COVID-19 Detection in Chest X-ray Images using a Deep Learning Approach. International Journal of Interactive Multimedia and Artificial Intelligence, In Press.
[7] Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham.
[8] Mirza, M., & Osindero, S. (11 2014). Conditional Generative Adversarial Nets.
[9] Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature Pyramid Networks for Object Detection. arXiv
[10] Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2018). Densely Connected Convolutional Networks. arXiv
[11] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition.
[12] Chattopadhay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018). Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV).

Acknowledgements

We thank Cisco Research for funding this project, and we thank M Health Fairview for providing the dataset and the GPU server. We thank Le Peng, Hengyue Liang, and Prof. Ju Sun (advisor) from GLOVEX at the University of Minnesota for providing technical assistance and mentoring.

* Both Gaoxiang and Andrew are the co-leaders of the project, and they contributed equally.