Disease Detection from Chest X-ray

By Team_18
Danliang Wang, Youjung Kim and Brandon Leventhal

Overview

- Problem Introduction
- Background
- Data
- Related Work
- Setup and Approach
- Initial Outcome
- Optimizations
- Other Experiments
- Conclusion

Problem

X-ray data is abundant but the process of interpreting the outputs is both costly and time-consuming. How do we automate this process with deep-learning to reduce cost and improve results?

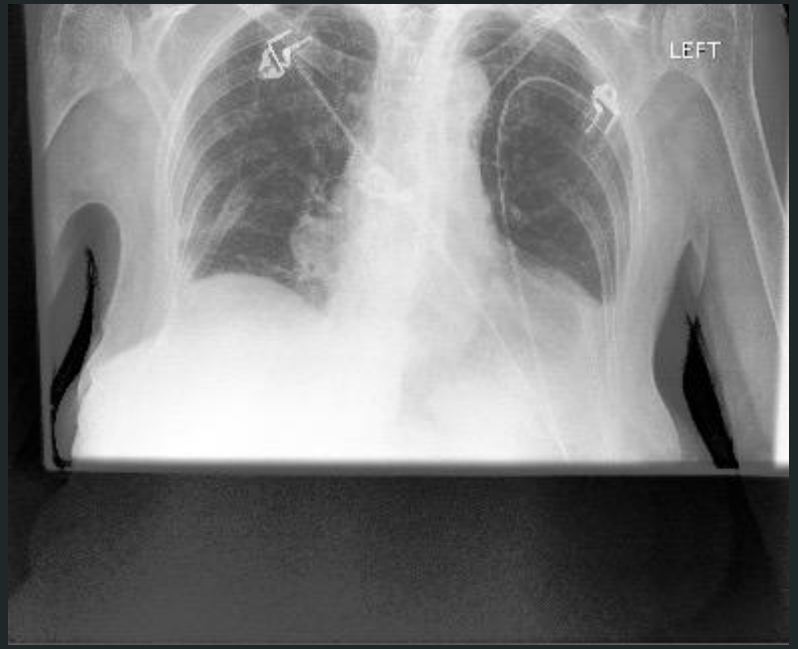
Background

- Advances in image-processing has made this viable
- Previous works has achieved strong results
- The use of DenseNet and CNN have been particularly successful

Data

- CheXpert by Stanford Machine Learning Group
- 224,316 chest radiographs of 65,240 patients labeled for the presence of 14 common chest radiographic observations

Example



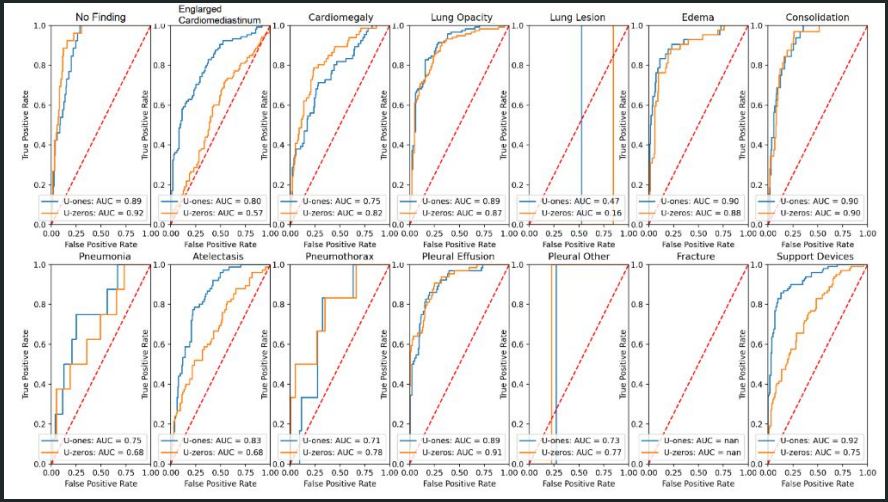
Related Work

- DenseNet121 by Gao Huang
- Residual Neural Network by Kaiming He
- Modeling improvements with DenseNet by Rajpurkar

Setup And Approach

- Preprocessing with pyspark and Python Image Library
- DenseNet121 and ResNet101 with Torch
- Cuda for parallelization

Initial Outcome



- 3 symptoms were poorly classified
- Excluding the 3 outliers we achieved an average AUC of .82
- Room for improvement

Optimizations

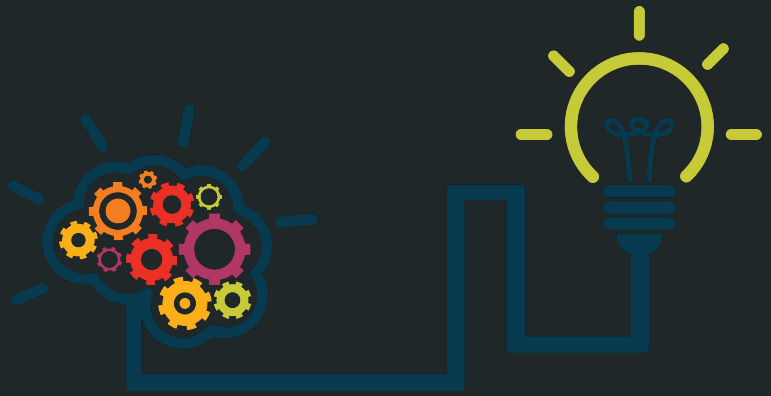
- DenseNet vs ResNet

	Average	No Finding	Enlarged Cardiomediastinum	Cardiomegaly	Lung Opacity	Edema
ResNet	0.76	0.9	0.6	0.68	0.83	0.81
DenseNet	0.84	0.89	0.8	0.75	0.89	0.9
	Consolidation	Pneumonia	Atelectasis	Pneumothorax	Pleural Effusion	Support Devices
ResNet	0.88	0.77	0.78	0.6	0.79	0.67
DenseNet	0.9	0.75	0.83	0.71	0.89	0.92

Other Experiments

- Localization using a heat map

Conclusion



- Preprocessing is important
- Experimentation of many models leads to best results
- High quality data is important for high quality results
- There are many improvements yet to be made