Chest x-rays and machine learning

AKA we have no idea what we're doing
please send help

Alistair Johnson

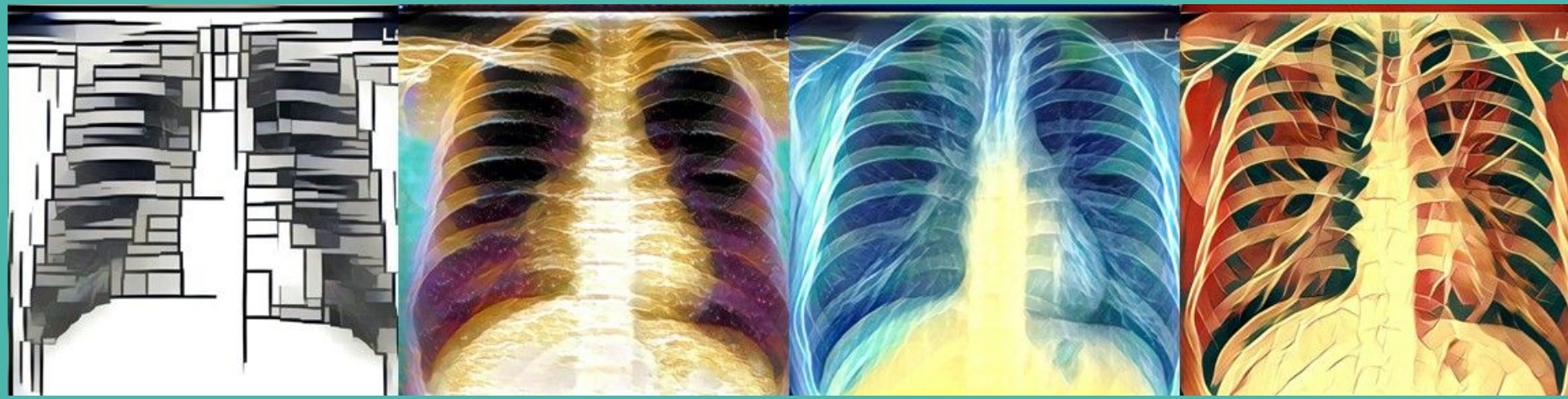
aewj@mit.edu

Overview

- The story of publicly available chest x-rays
- What we have available to us (as of a few weeks ago!)
- Interesting projects to try
- A demo with a published model

aewj@mit.edu

Wait.. what is a chest x-ray?



Chest radiographs

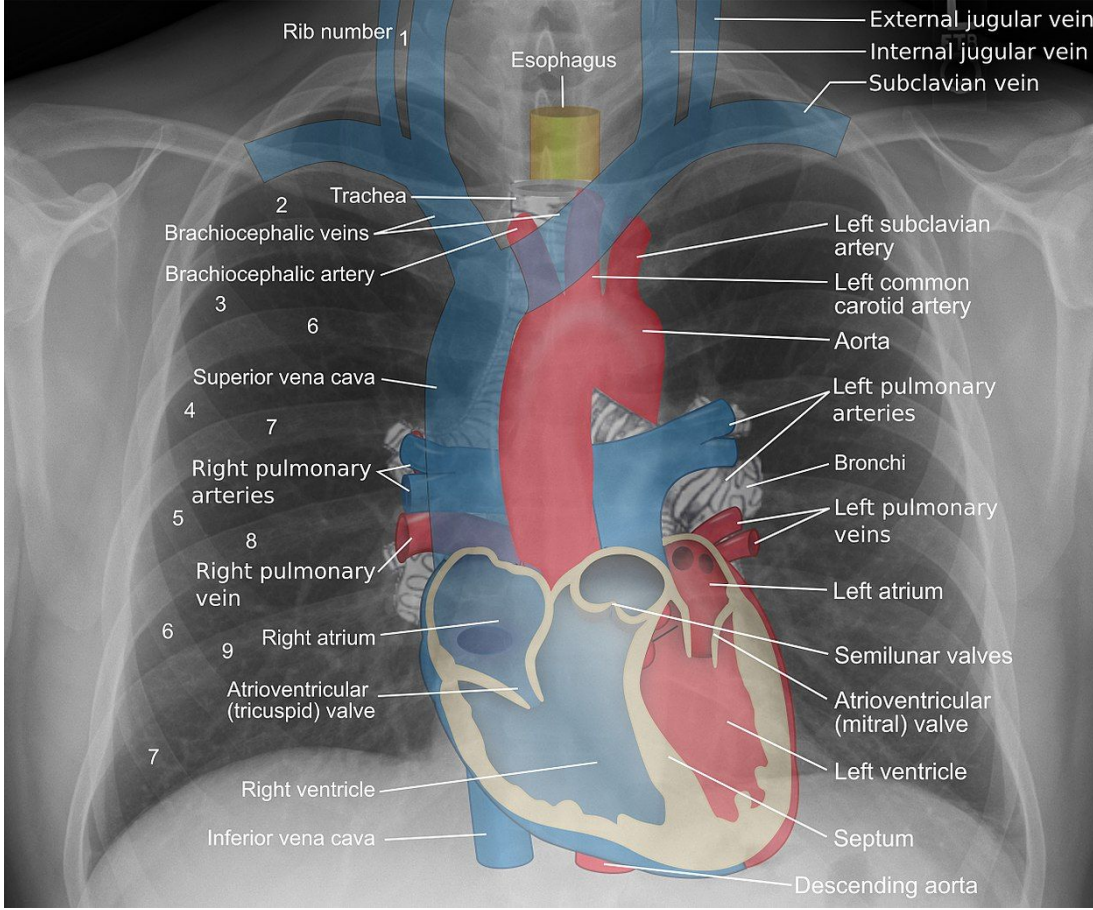
- Commonly called X-rays
  - (“X” stands for “I have no idea what this thing is but look I see through people’s skin”)
- Visualizes the lungs
- Visualizes the heart

aewj@mit.edu

Chest radiographs

- Overlay of the heart
- Many of these are not actually segmentable

What do radiologists look for in these x-rays?



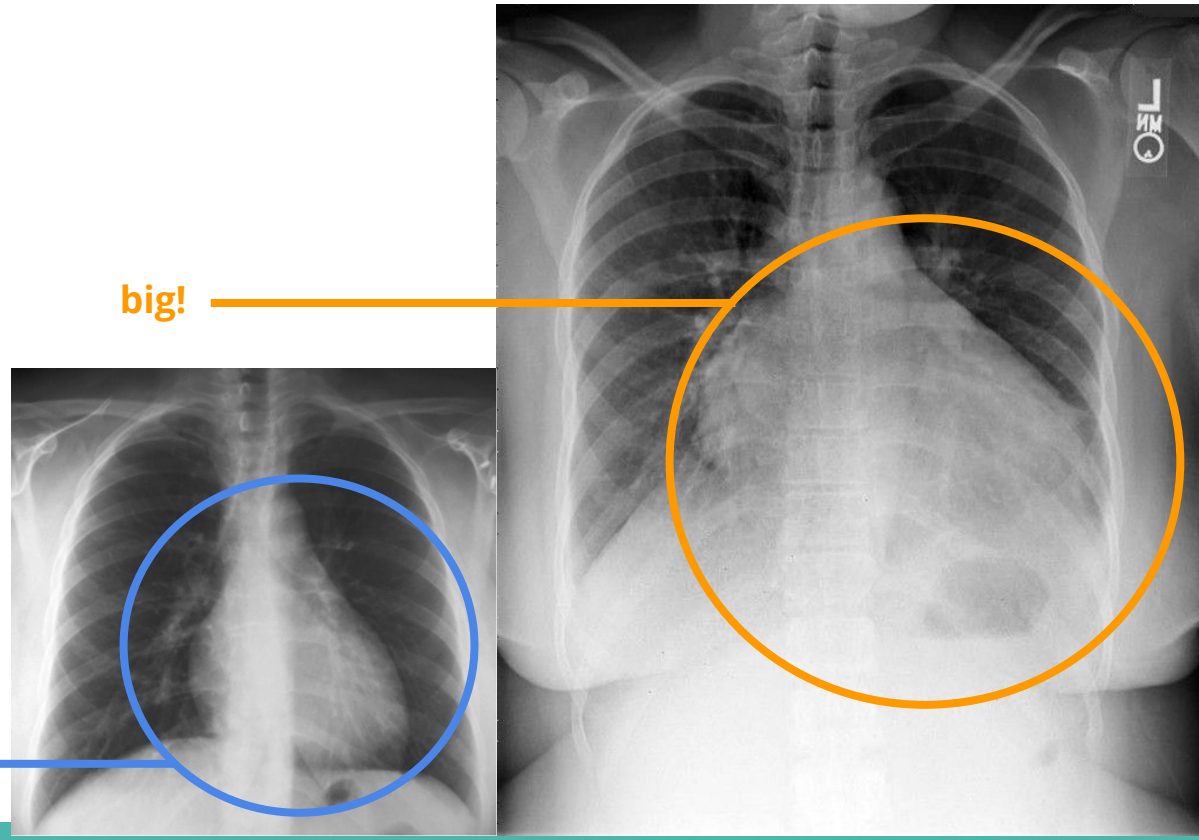
Mikael Häggström, M.D. "A pictorial essay: Radiology of lines and tubes in the intensive care unit". Indian Journal of Radiology and Imaging 21 (3): 182.

aewj@mit.edu

What do we look for in a chest x-ray?

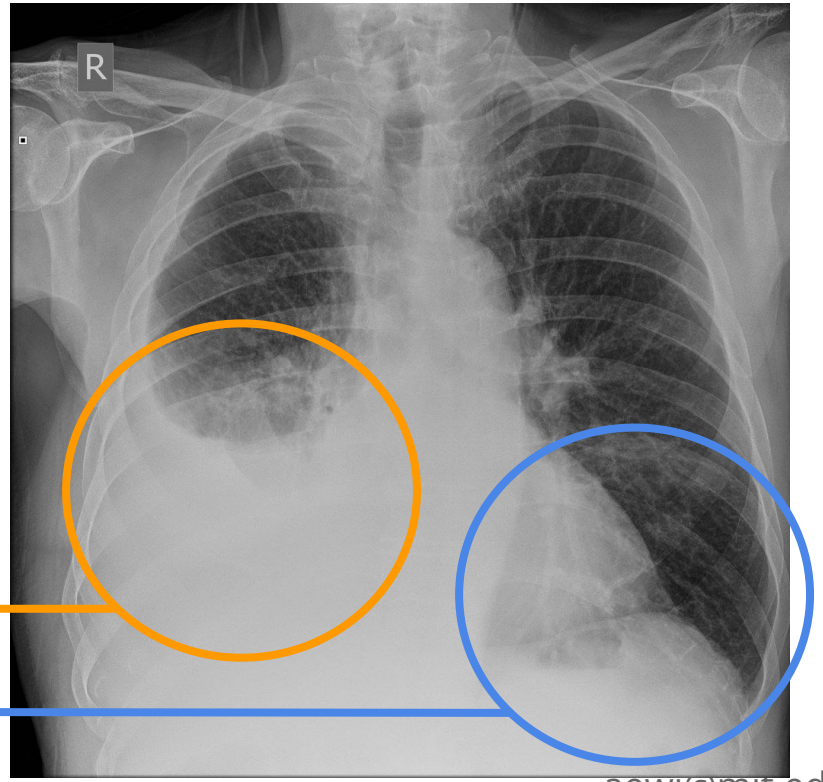
A big heart, “cardiomegaly”

(and not in a good way!)



What do we look for in a chest x-ray?

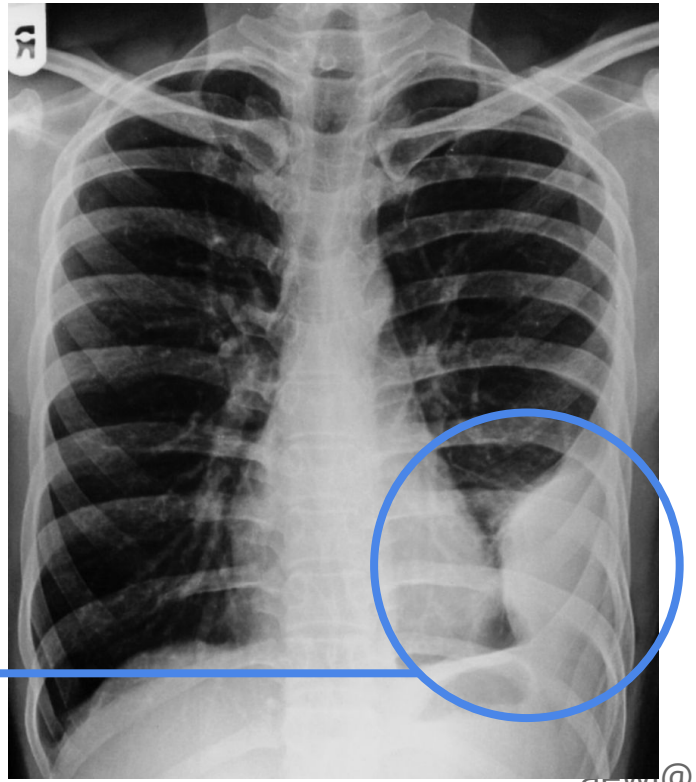
- Water where it is not supposed to be



aewj@mit.edu

What do we look for in a chest x-ray?

- ~~Water~~ Fluid where it is not supposed to be



Blood? Water? Pus?

aewj@mit.edu

Why should we automate x-ray scanning?

“My mom is a radiologist! Don’t take her job!”

- 12 million people in Rwanda
  - 11 radiologists
- 4 million people in Libya
  - 2 radiologists
- Most radiologists are in urban settings
- Delays in medical image interpretation are *bad*

Abi Rimmer. Radiologist shortage leaves patient care at risk, warns royal college. BMJ: British Medical Journal.

aewj@mit.edu

Public x-rays: trials and tribulations

Medical machine learning *is hard*

I could fill a book with why it is hard; but relevant to this talk:

- Clinical data is multimodal, and complex
  - Text, images, genomics, laboratory values, electrocardiograms, ...
- Researchers are far away from the data collection
  - This isn't 6.814! I want to build neural networks!
- **There isn't much data**
  - can't emphasize this one enough: there really isn't much.

In 2016:

- ImageNet - 14,000,000 images
- Chest x-rays - 4,000 images

aewj@mit.edu

NIH to the rescue!

- 30,000+ patients
- 100,000+ images
- Each associated with 14 labels
  - Derived automatically from the free-text report
- Freely, publicly available

Media Advisory Wednesday, September 27, 2017

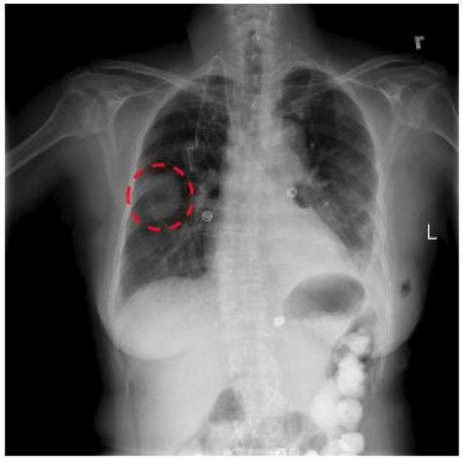
NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community

The dataset of scans is from more than 30,000 patients, including many with advanced lung disease.

What

The NIH Clinical Center recently released over 100,000 anonymized chest x-ray images and their corresponding data to the scientific community. The release will allow researchers across the country and around the world to freely access the datasets and increase their ability to teach computers how to detect and diagnose disease. Ultimately, this artificial intelligence mechanism can lead to clinicians making better diagnostic decisions for patients.

NIH compiled the dataset of scans from more than 30,000 patients, including many with advanced lung disease. Patients at the NIH Clinical Center, the nation's largest hospital devoted entirely to clinical research, are partners in research and voluntarily enroll to participate in clinical trials. With patient privacy being paramount,

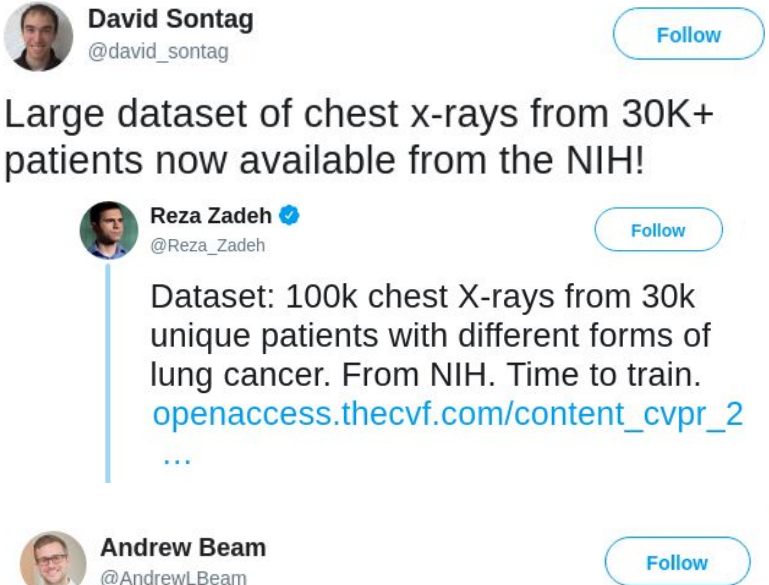


A chest x-ray identifies a lung mass.

https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available-chest-x-ray-datasets-scientific-community

aewj@mit.edu

NIH to the rescue!



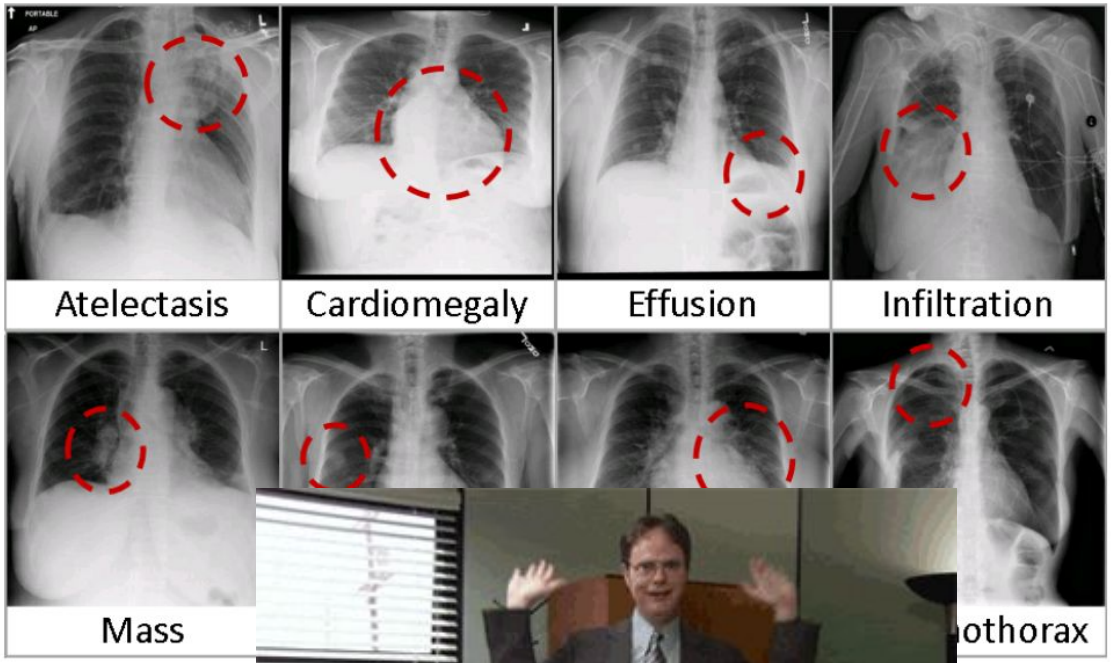
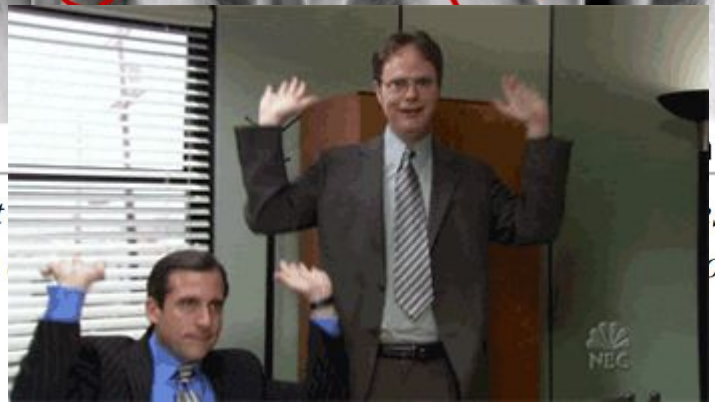


Figure 1. *Eight* *st X-rays* *that validate a* *osis.*



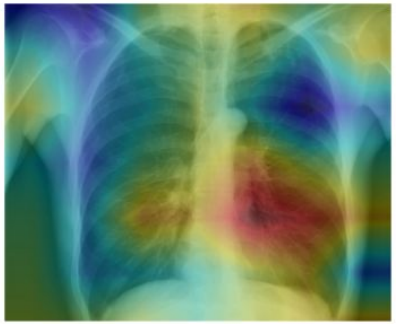
aewj@mit.edu

Deep learning success!

Input
Chest X-Ray Image

CheXNet
121-layer CNN

Output
Pneumonia Positive (85%)



… dramatic pause...

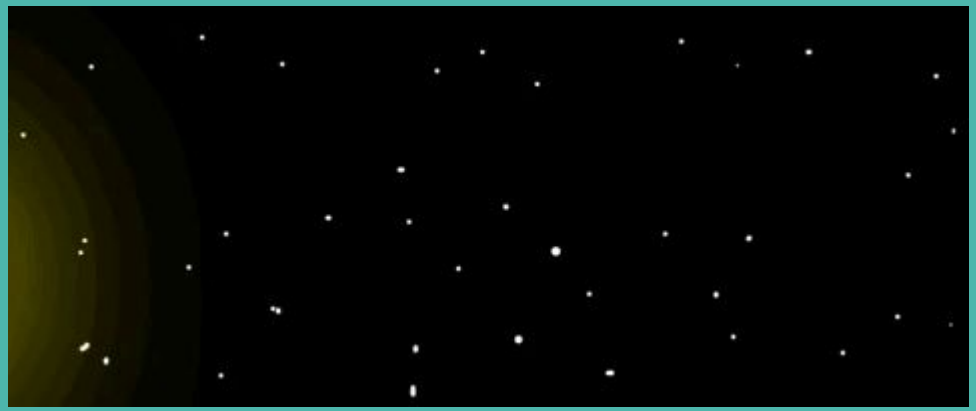
aewj@mit.edu

It's never that easy

- Radiologist manually reviewed the images
- Disagreed with the labels
- Radiologist is big in the twitterverse*, so this was a big deal

* Warning to future academics, twitter is a thing you have to use, and it's as terrible as you imagine.

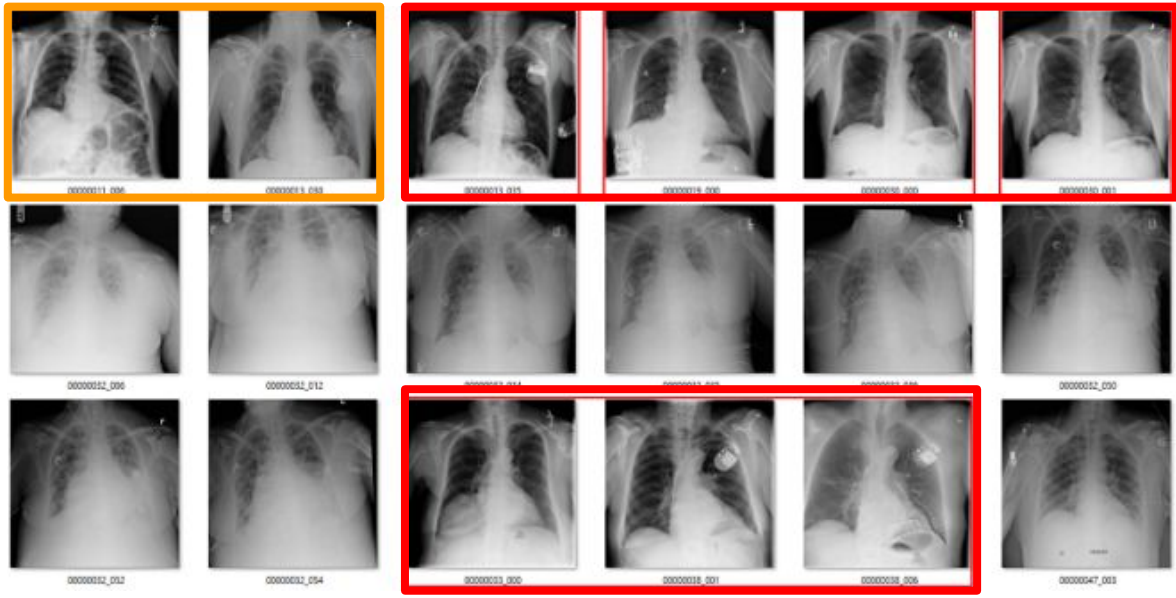
What went wrong?



Problems with the data...

- All of these were supposed to be *atelectasis*
- But many clearly aren't
- Consistent across labels

Label	PPV (visual)
Consolidation	35%
Cardiomegaly	80%
Pneumothorax	60%
Pneumonia	35%
Fibrosis	24%
No finding	60%



Red = I disagree

Orange = eh, I'm not sure

aewj@mit.edu

Problems with the data...

[**Hospital 7**] MEDICAL CONDITION:
78 year old woman with seizures, intubated
REASON FOR THIS EXAMINATION:
ett placement, pna

FINAL REPORT

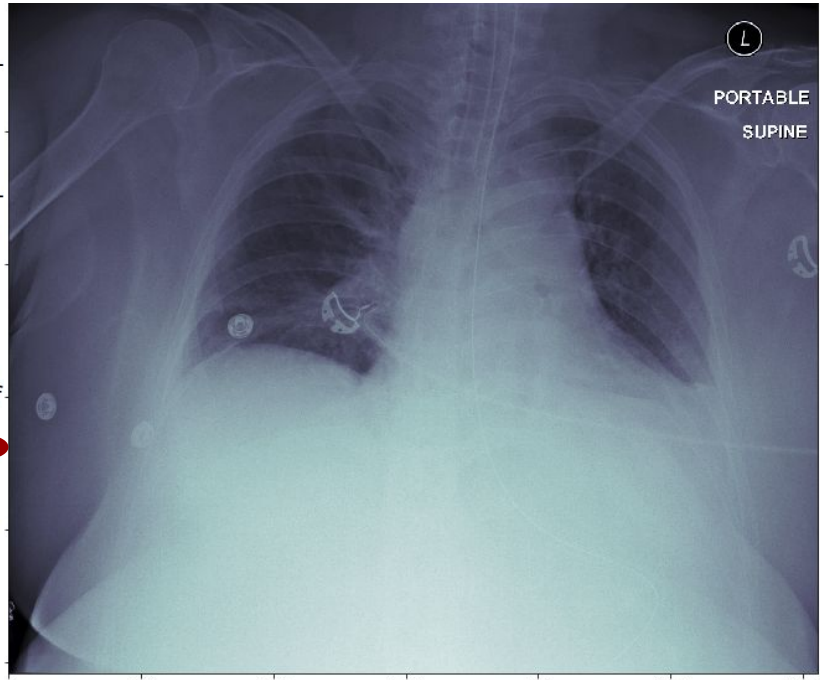
CLINICAL INFORMATION: 78-year-old female with seizures, intubated for airway protection.

COMPARISON: None.

FINDINGS: There is an endotracheal tube in place with the tip located 3.4 cm from the level of the carina. An orogastric tube is not seen with the tip off the inferior aspect of the film. There are low lung volumes, with foci of atelectasis and bibasilar opacity. There is no effusion or pneumothorax. The cardiac silhouette is not well evaluated due to low lung volumes.

IMPRESSION: Endotracheal tube appropriately positioned with low lung volumes with atelectasis and bibasilar opacities.

THE STUDY AND THE REPORT WERE REVIEWED BY THE STAFF RADIOLOGIST.



aewj@mit.edu

Problems with the data...

1. *unremarkable* cardiomediastinal silhouette

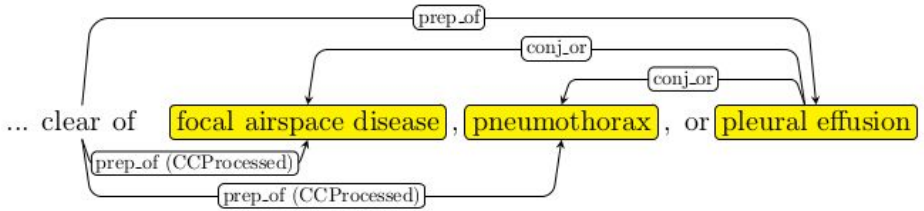
2. diffuse reticular pattern, which can be seen with an atypical infection **or** chronic fibrotic change. *no* focal consolidation.

3. *no* pleural effusion or pneumothorax

4. mild degenerative changes in the lumbar spine and old right rib fractures.

Observation	Labeler Output
No Finding	
Enlarged Cardiom.	0
Cardiomegaly	
Lung Opacity	1
Lung Lesion	
Edema	
Consolidation	0
Pneumonia	u
Atelectasis	
Pneumothorax	0
Pleural Effusion	0
Pleural Other	
Fracture	1
Support Devices	

This is way harder than it looks.

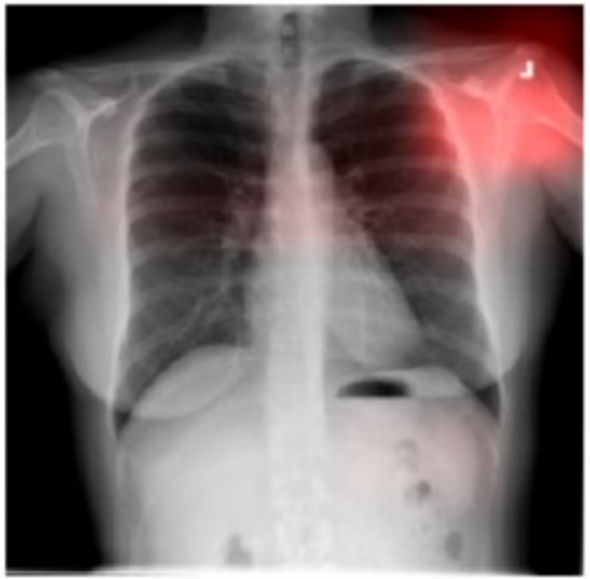


Problems with the model...

Confounding variables can degrade generalization performance of radiological deep learning models

John R. Zech, Marcus A. Badgeley, Manway Liu, Anthony B. Costa, Joseph J. Titano, Eric K. Oermann

Models can discriminate the *hospital which the patient was admitted to* with **> 95% certainty**



What do we do now?

The NIH have moved on

Now releasing CT scans

J Med Imaging (Bellingham). 2018 Jul;5(3):036501. doi: 10.1117/1.JMI.5.3.036501. Epub 2018 Jul 20.

DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning.

Yan K[1], Wang X[1], Lu L[2], Summers RM[1].

Media Advisory Friday, July 20, 2018

NIH Clinical Center releases dataset of 32,000 CT images

The National Institutes of Health's Clinical Center has made a large-scale dataset of CT images publicly available to help the scientific community improve detection accuracy of lesions. While most publicly available medical image datasets have less than a thousand lesions, this dataset, named DeepLesion, has over 32,000 annotated lesions identified on CT images.

… but we haven't!

CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison

Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silviana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, Jayne Seekins, David A. Mong, Safwan S. Halabi, Jesse K. Sandberg, Ricky Jones, David B. Larson, Curtis P. Langlotz, Bhavik N. Patel, Matthew P. Lungren, Andrew Y. Ng

(Submitted on 21 Jan 2019)

MIMIC-CXR: A large publicly available database of labeled chest radiographs

Alistair E. W. Johnson, Tom J. Pollard, Seth J. Berkowitz, Nathaniel R. Greenbaum, Matthew P. Lungren, Chih-ying Deng, Roger G. Mark, Steven Horng

(Submitted on 21 Jan 2019 (v1), last revised 23 Jan 2019 (this version, v2))

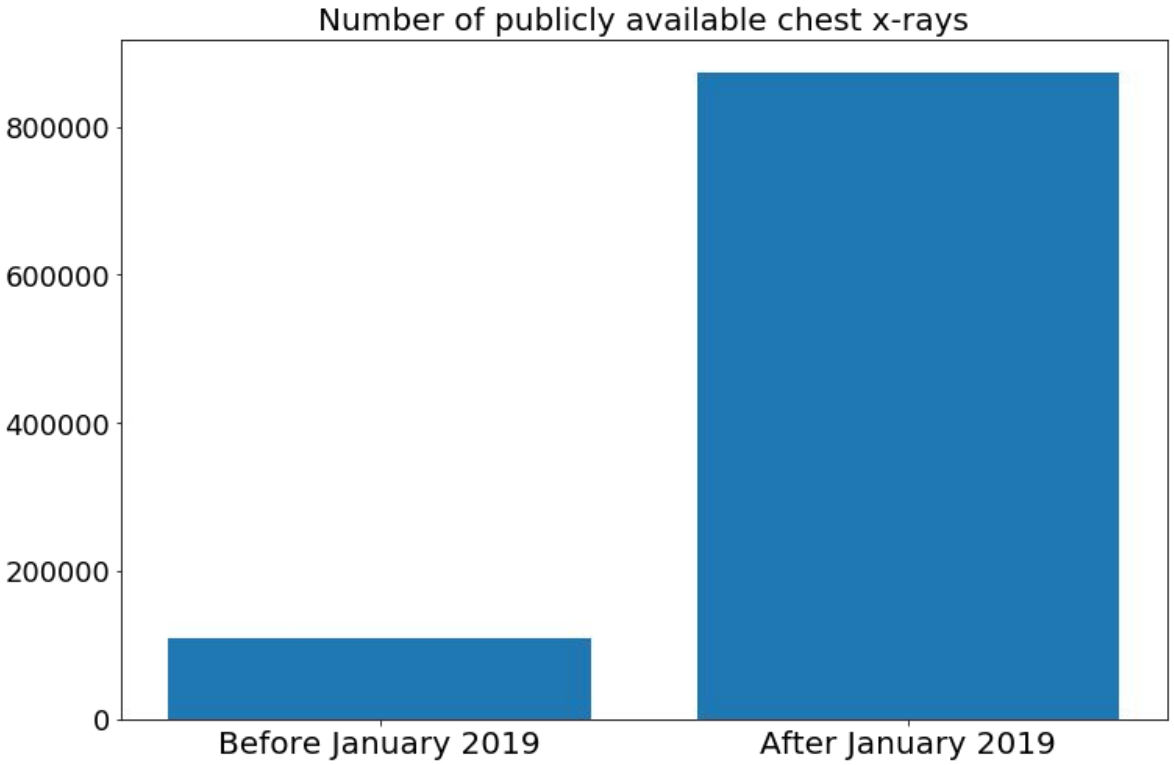
PadChest: A large chest x-ray image dataset with multi-label annotated reports

Aurelia Bustos, Antonio Pertusa, Jose-Maria Salinas, Maria de la Iglesia-Vayá

(Submitted on 22 Jan 2019 (v1), last revised 7 Feb 2019 (this version, v2))

aewj@mit.edu

Chest x-rays available



aewj@mit.edu

What can we do with this data?

Generate x-rays using a GAN

- Why? Because we can!

https://thispersondoesnotexist.com/ - GAN faces

https://thiscatdoesnotexist.com/ - GAN cats

https://thischestxraydoesnotexist.com/ - OPPORTUNITY HERE!!!

aewj@mit.edu

Replace the radiologist!*

- Each image is given with 14 labels
- Try to predict them!
- Participate in the CheXpert challenge

https://stanfordmlgroup.github.io/competitions/chexpert/

Leaderboard

Will your model perform as well as radiologists in detecting different pathologies in chest X-rays?

Rank	Date	Model	AUC
		Leaderboard Activating February 2019	

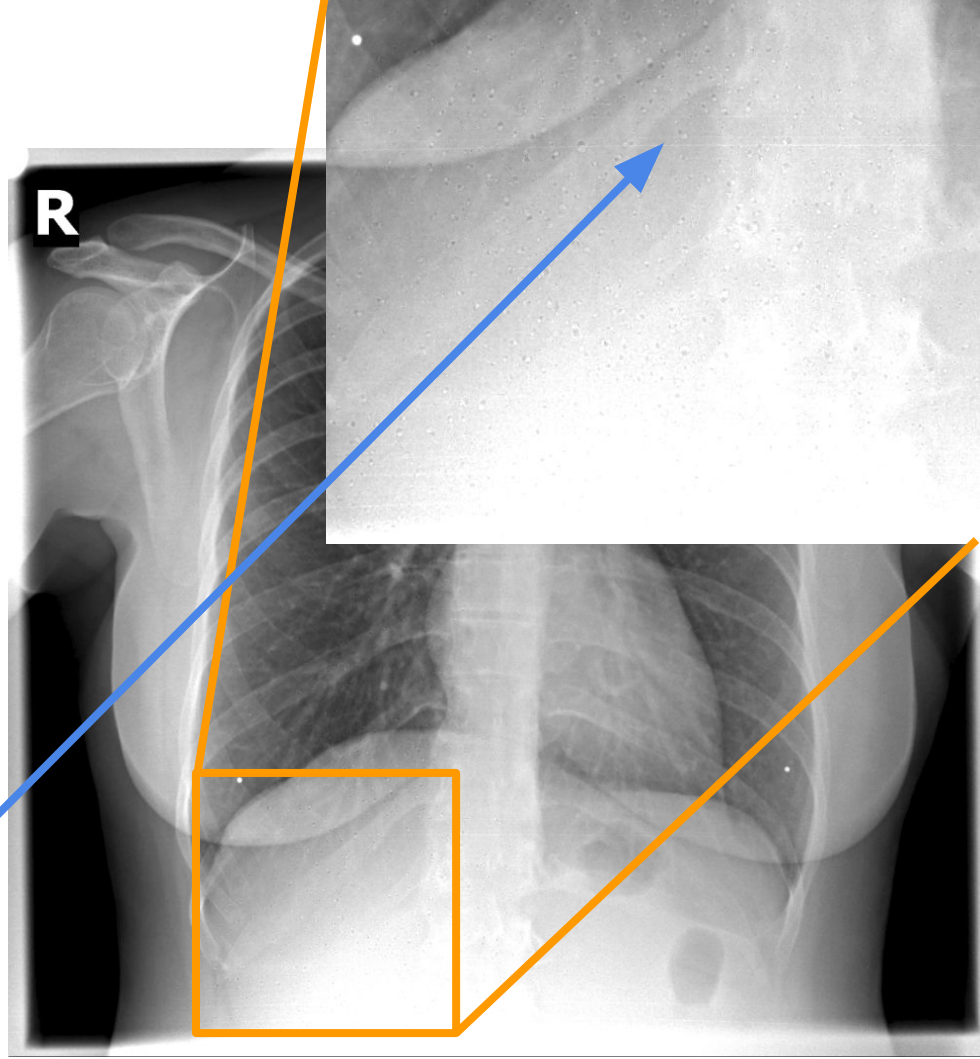
* Don't actually say you're better than a radiologist.

aewj@mit.edu

Model generalizability

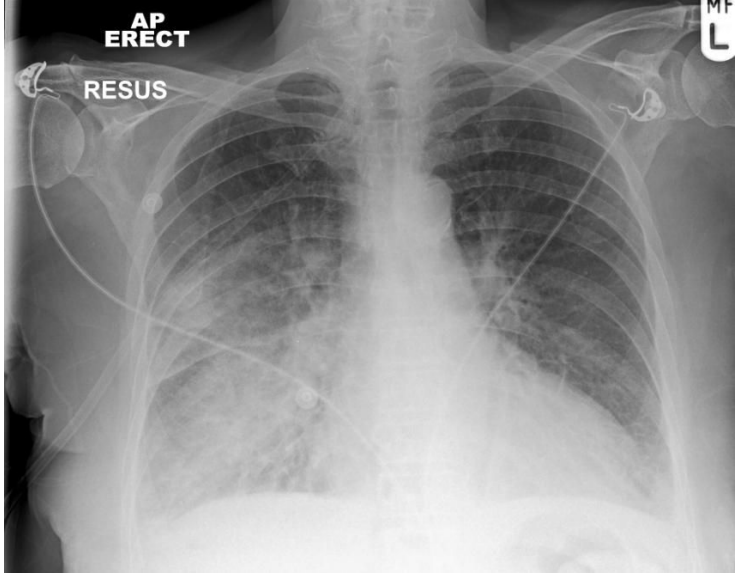
- Each dataset was collected at a distinct institution
- Different demographics
  - NIH clinical center gets “complicated” cases
  - Stanford/BIDMC/San Juan are tertiary academic medical centers
  - No instances of emphysema in MIMIC-CXR
- Different devices for collection

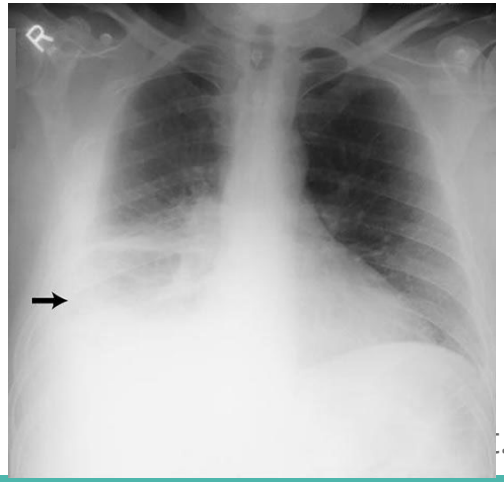
Speckled dots from underexposure not present in all databases



Multi-task learning

- Conditions vary in frequency
  - Airspace Opacity - 20% of images
  - Pleural Effusion - 20% of images
  - Lung Lesion - 3% of images
- Conditions are very similar
  - A pleural effusion looks a lot like pulmonary edema
  - A classifier may incorrectly classify effusions as edema unless it is aware of both classes





Multi-angle images

- Chest x-rays most often come in pairs
  - Frontal
  - Lateral

aewj@mit.edu

Multi-angle images

frontal

lateral

aewj@mit.edu

Multi-angle images

frontal

lateral

aewj@mit.edu

Multi-angle images

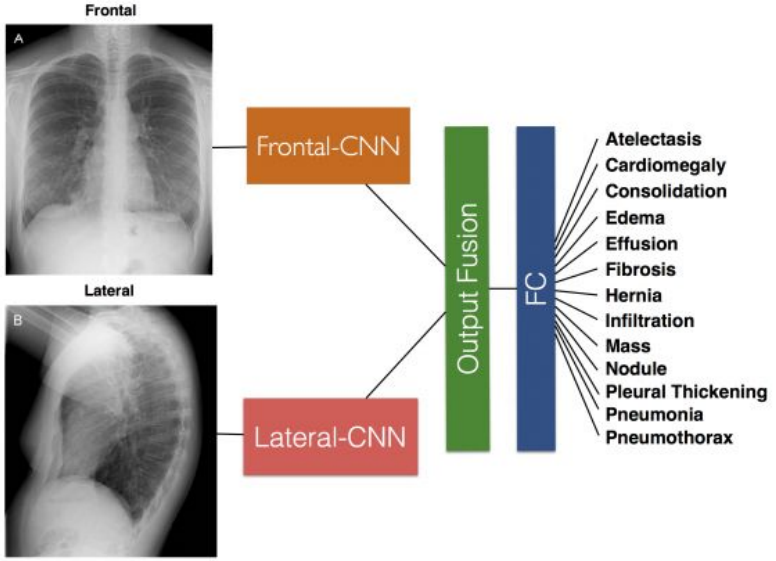
- Images are different resolutions
- Some pathologies better visualized on the lateral film
- How can we build a model that uses both views?

Large Scale Automated Reading of Frontal and Lateral Chest X-Rays using Dual Convolutional Neural Networks

Jonathan Rubin, Deepan Sanghavi, Claire Zhao, Kathy Lee, Ashequl Qadir, Minnan Xu-Wilson

(Submitted on 20 Apr 2018 (v1), last revised 24 Apr 2018 (this version, v2))

https://arxiv.org/abs/1804.07839

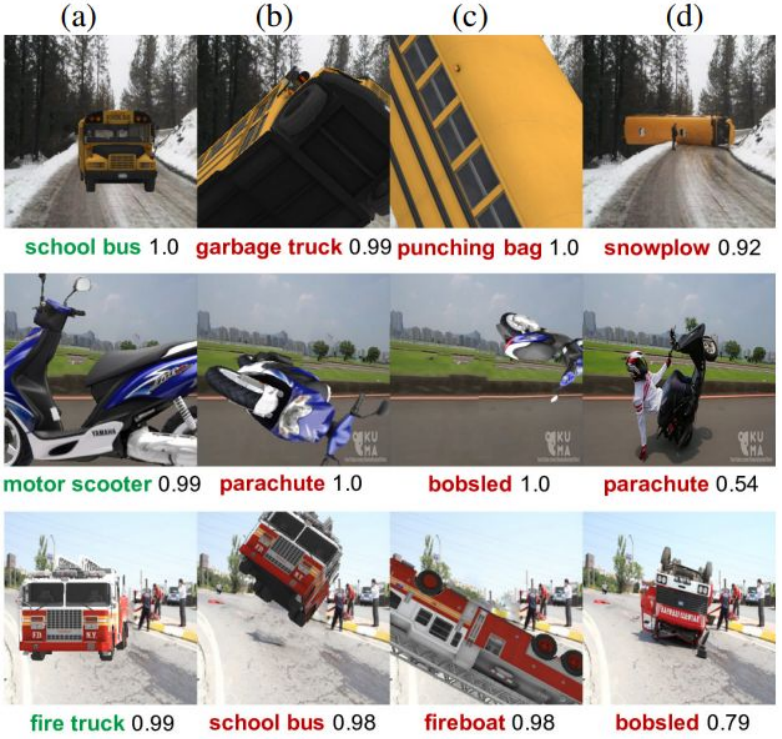


Consistent views

- Deep neural networks are fooled by odd poses
- In medical imaging, poses are (reasonably) consistent
- How do we incorporate this information into the model?

Strike (with) a Pose: Neural Networks Are Easily Fooled by Strange Poses of Familiar Objects

Michael A. Alcorn, Qi Li, Zhitao Gong, Chengfei Wang, Long Mai, Wei-Shinn Ku, Anh Nguyen

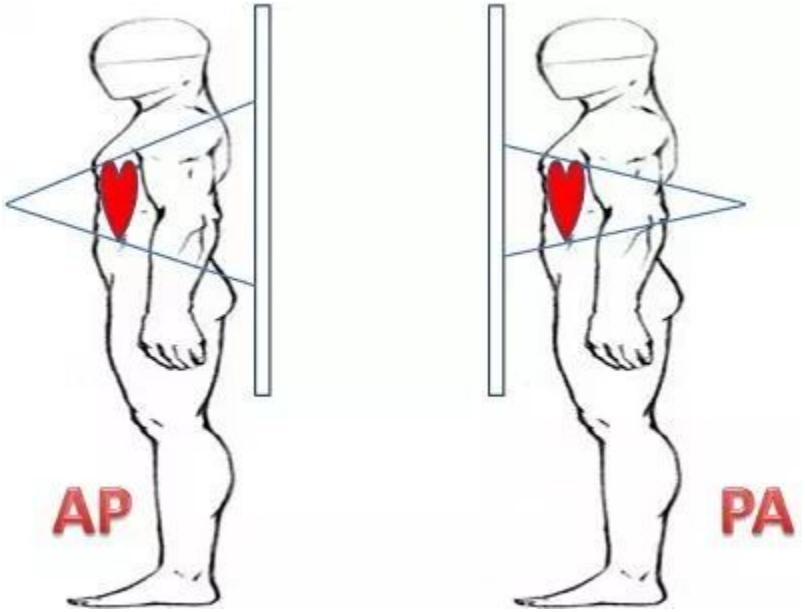


https://arxiv.org/abs/1811.11553

Consistent views

aewj@mit.edu

Consistent views



aewj@mit.edu

Consistent views

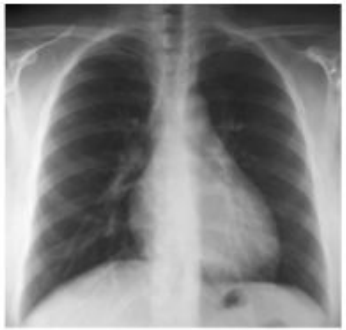
(a) P-A

(b) Lateral

(c) Lordotic

(d) A-P supine

(e) A-P



(f) P-A

(g) Lateral

(h) Lordotic

(i) A-P supine

(j) A-P

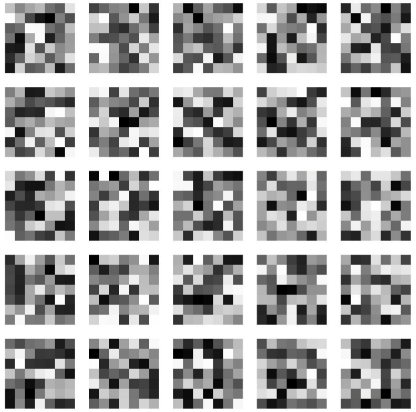
Are medical images the same as natural images?

- Train a neural network on *ImageNet*
- Look at the underlying filters at each epoch

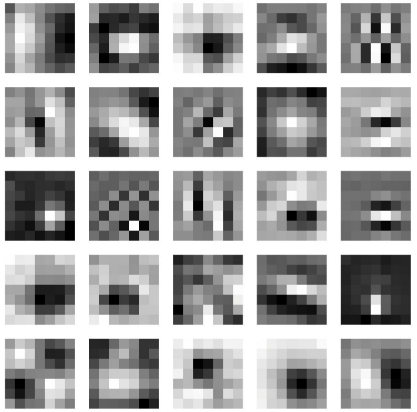
- We see Gabor filters

Transfusion: Understanding Transfer Learning with Applications to Medical Imaging

Maithra Raghu, Chiyuan Zhang, Jon Kleinberg, Samy Bengio



start



trained

https://arxiv.org/abs/1902.07208

aewj@mit.edu

Are medical images the same as natural images?

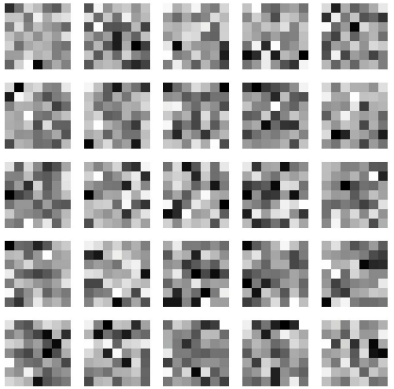
- Train a neural network on *medical images*
- Look at the underlying filters at each epoch

- We see … not Gabor??

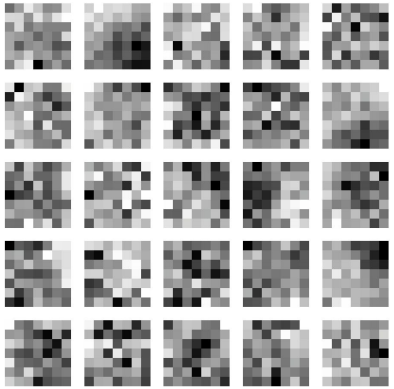
Transfusion: Understanding Transfer Learning with Applications to Medical Imaging

Maithra Raghu, Chiyuan Zhang, Jon Kleinberg, Samy Bengio

Rand Init Untrained



Rand Init Trained



https://arxiv.org/abs/1902.07208

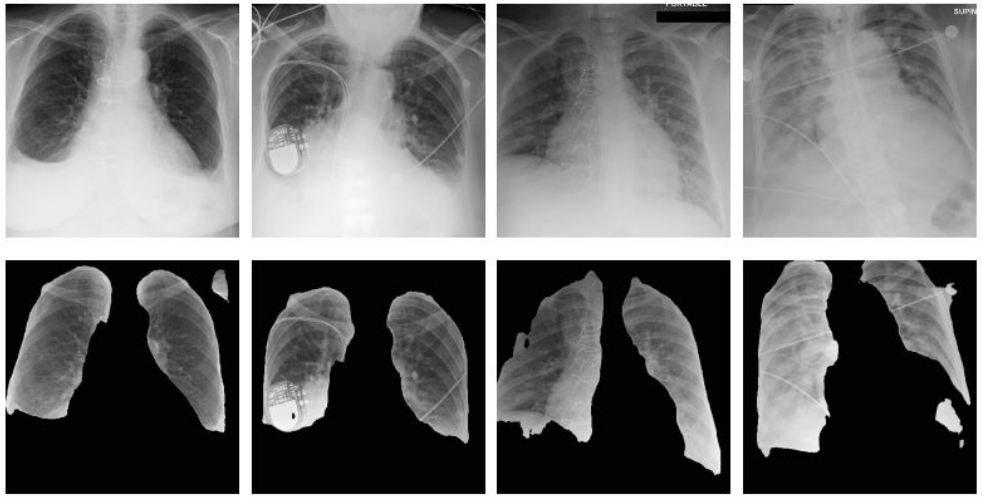
start

trained

aewj@mit.edu

Segmentation, quantification

- Radiologists qualify the image
  - “Lungs are clear”
  - “Low lung volumes”
  - “Mild pulmonary edema”
- We have no scale for mild
- Can we *quantify* lung opacity?
- Can we come up with a new measure of opacity which correlates with clinical outcome?



aewj@mit.edu

… and your own ideas!

Have a craic

GitHub repository for reproducing CheXNet

https://github.com/jrzech/reproduce-chexnet

Google Colab notebook of the above

https://drive.google.com/open?id=1pkWyWHDUyNR0P6xNho0miVK8fjF5u2X5

aewj@mit.edu