Detecting and Understanding Pneumonia with Deep Learning

Nikolaos Altiparmakis
Stanford University
https://youtu.be/6Dq7Zj8VIOw

MOTIVATION

- Pneumonia is one of the most dangerous and prevalent diseases worldwide.
- This results in increased workload in hospitals.
- Research has shown automation of pneumonia detection is possible [1] [2].
- In this research, we make use of small dataset and evaluate data augmentation and regularization methods to provide reliable results.
- We expect recall to be equal to 1, since doctors always edge on the worst-side outcome, when unsure of the patient's condition.

DATASET

- Our dataset includes 5,863 images of frontal-view chest X-ray images.
- Since the dataset is small, keep as many images on the training set as possible. Split using ratios 85%/7.5%/7.5%.
- Use data augmentation techniques:
  - Image rotation by up to 15 degrees.
  - Random brightness scaling by a factor a in [0.9, 1.2]
  - Horizontal flip.
- Explore regularization methods:
  - L2-regularization and dropout.
  - Neural style transfer.

ARCHITECTURE

- ResNet50 model with a final dense node with a sigmoid activation.
- Use mini-batches of size 64.
- Train the model using Adam with β1 = 0.9 and β2 = 0.999. the learning rate is equal to λ = 0.003 with a decay equal to α = 3e-6,.

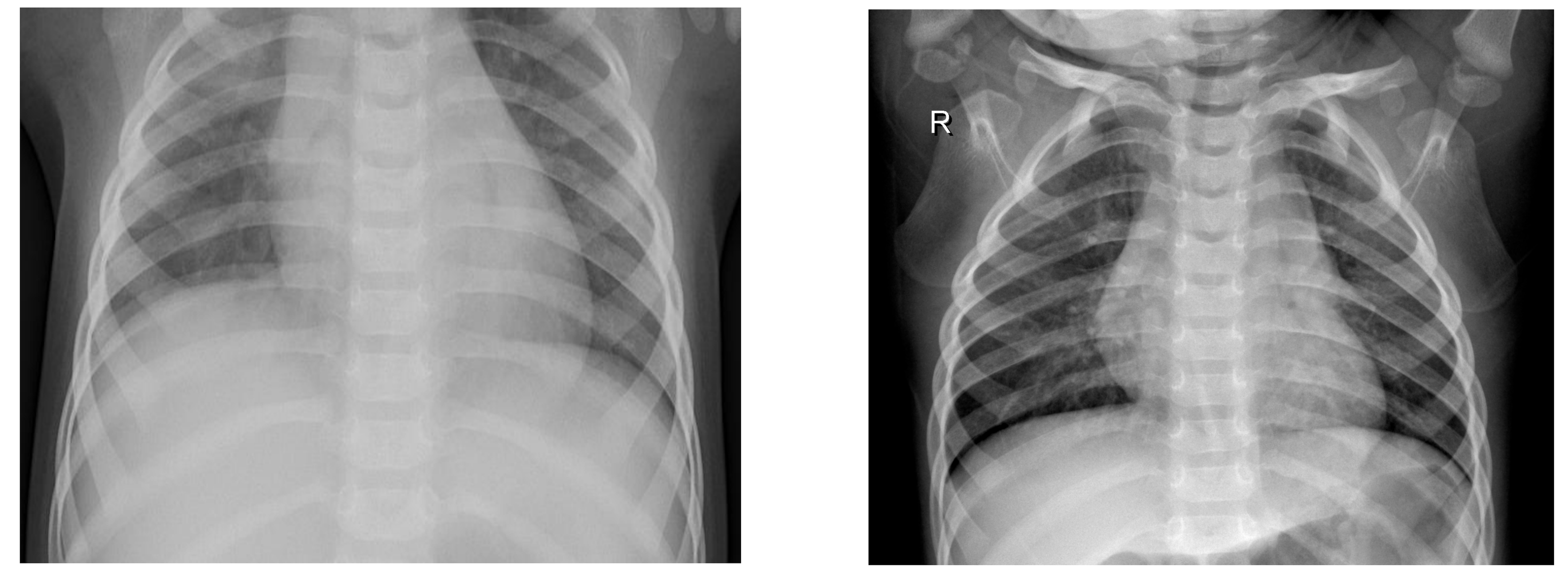

Figure 1: On the left: A chest X-ray from patient with pneumonia. On the right: A normal X-ray image.

RESULTS

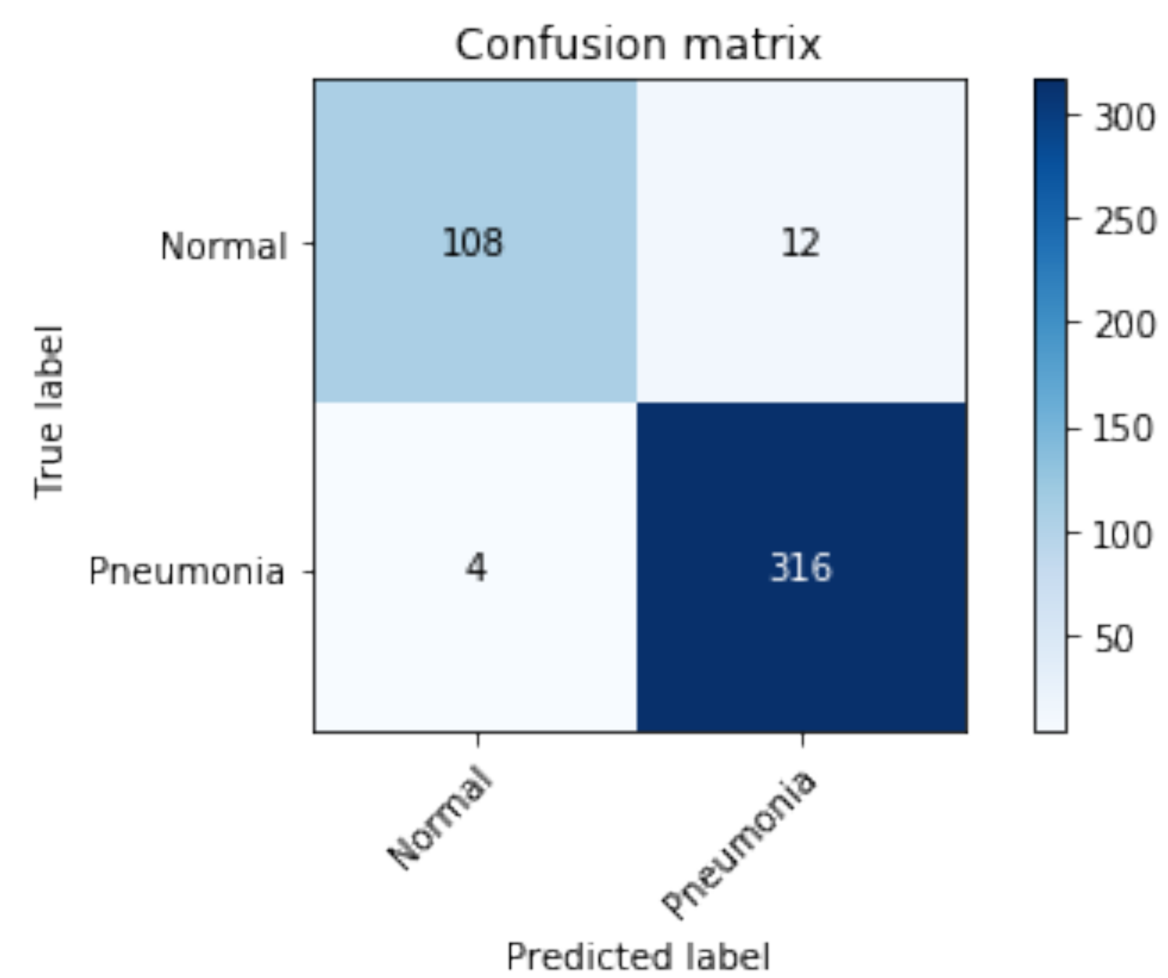

Figure 2: Confusion matrix for the test set.

- Generally good metrics: 98.36% accuracy on the training set, 96.36% on the validation and 96.61% on the test test.
- The use of dropout and L2-regularization had worse results, so were not used.
- Neural-style transfer did not work in this case due to the complexity of the problem.
- F1-score 94.17% on the test set, which generally suggests the model is not trivial (Fig.2).

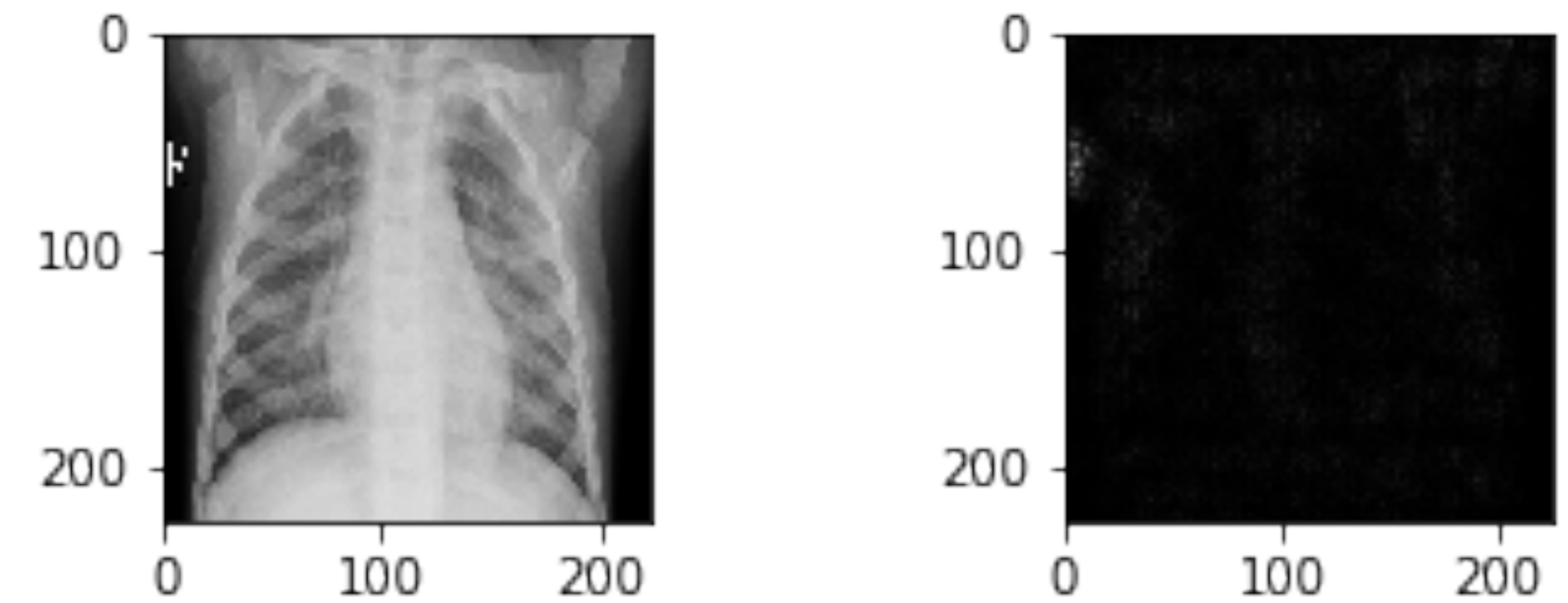

Figure 3: Estimation based on an irrelevant landmark in the picture

- Using the integrated gradients method (Fig. 3), we conclude that the model is not always basing its prediction on relevant data.

CONLUSION

- The dataset was too small to build a reliable tool.
- The general decisions were correct and would have a better result in a larger dataset.
- We should always take visualization into consideration during the training process.
- Use a more complicated dataset which includes not just the images, but the relevant areas of each image.
- We could always use visualization techniques for medical issues we do not know the causes of, since neural networks provide useful insights through their ability to visualize results.

REFERENCES

[1] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.

[2] Okeke Stephen, Mangal Sain, Uchenna Joseph Maduh, and Do-Un Jeong. An efficient deep learning approach to pneumonia classification in healthcare. *Journal of healthcare engineering*, 2019, 2019