Module 4 Project: Classification of X-Ray Images Using Convolutional Neural Networks

Seeta Rajpara

Self-Paced Data Science Bootcamp

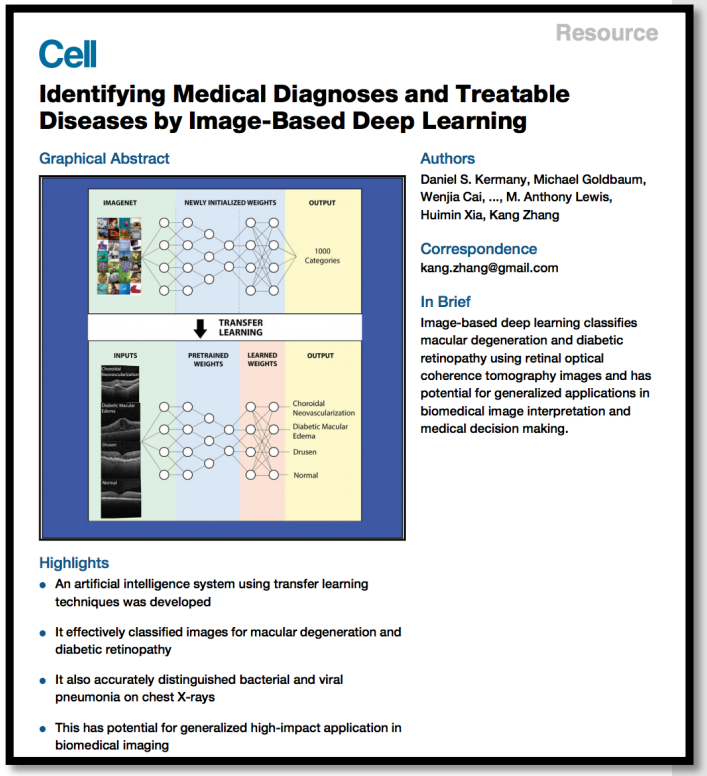
Problem Summary

Background:

- Study described ML method for diagnosing chest x-ray images of patients' lungs as being either healthy (normal) or having pneumonia.[3]
- Pneumonia causes air sacs in the lungs to fill with fluid or pus[1]
- Cause: bacterial, viral, or fungal infection
- Diagnosis:
  - Physical exam and chest X-ray imaging
  - Radiologists look for infiltrates (white spots) in the lungs for signs of infection
  - Lung inflammation is visualized by having denser, more opaque areas in imaging[2]

Objective:

- Build a machine learning model that will classify x-ray images of the lungs as either healthy (normal) or having pneumonia



Resource

Cell

Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning

Graphical Abstract

Authors

Daniel S. Kermany, Michael Goldbaum, Wenjia Cai, ..., M. Anthony Lewis, Huimin Xia, Kang Zhang

Correspondence

kang.zhang@gmail.com

In Brief

Image-based deep learning classifies macular degeneration and diabetic retinopathy using retinal optical coherence tomography images and has potential for generalized applications in biomedical image interpretation and medical decision making.

Highlights

- An artificial intelligence system using transfer learning techniques was developed
- It effectively classified images for macular degeneration and diabetic retinopathy
- It also accurately distinguished bacterial and viral pneumonia on chest X-rays
- This has potential for generalized high-impact application in biomedical imaging

Sources:

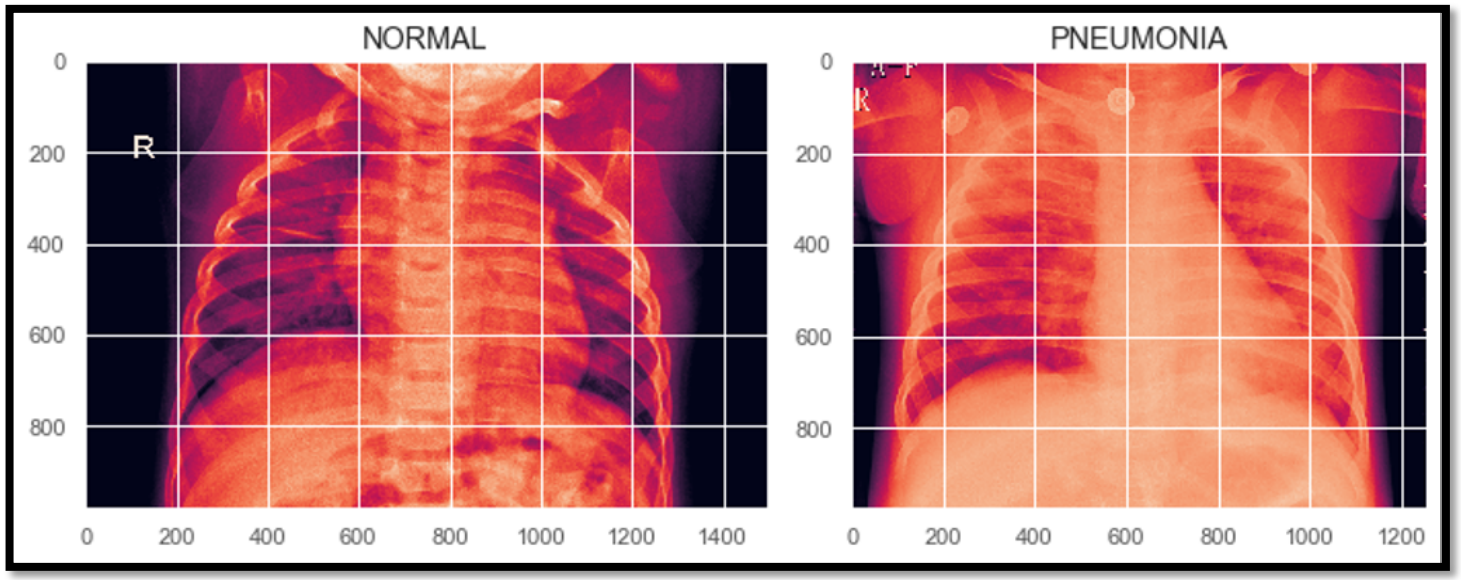
1) https://www.nhlbi.nih.gov/health/pneumonia
2) https://www.mayoclinic.org/diseases-conditions/pneumonia/multimedia/chest-x-ray-showing-pneumonia/img-20005827
3) http://www.cell.com/cell/fulltext/S0092-8674(18)30154-5

The Dataset

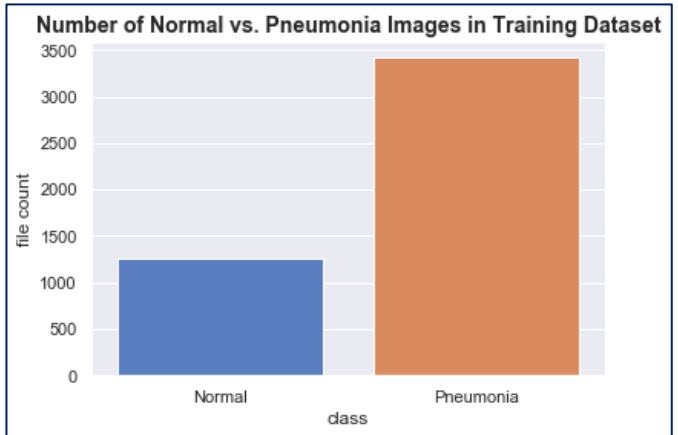
- Two folders (train/test) contain subfolders for each image category (Pneumonia/Normal)
- Total 5,863 X-Ray images (JPEG) and 2 categories (pneumonia/normal)
- Chest X-ray images
  - Pediatric patients (1-5 y/o) from Guangzhou Women and Children's Medical Center, Guangzhou.
  - All chest X-ray imaging was performed as part of patients' routine clinical care
- Radiographic preprocessing of image:
  - Removed all low quality or unreadable scans
  - Diagnoses for images were graded by two expert physicians before being cleared for training the AI system
  - In order to account for any grading errors, the evaluation set was also checked by a third expert

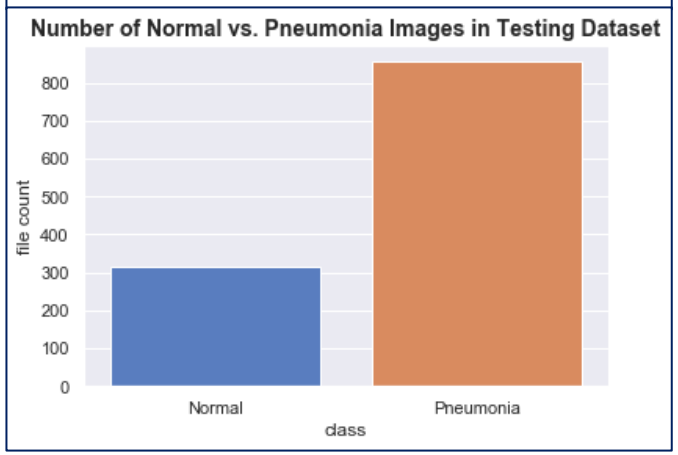
Sources:
1) https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

Exploratory Analysis



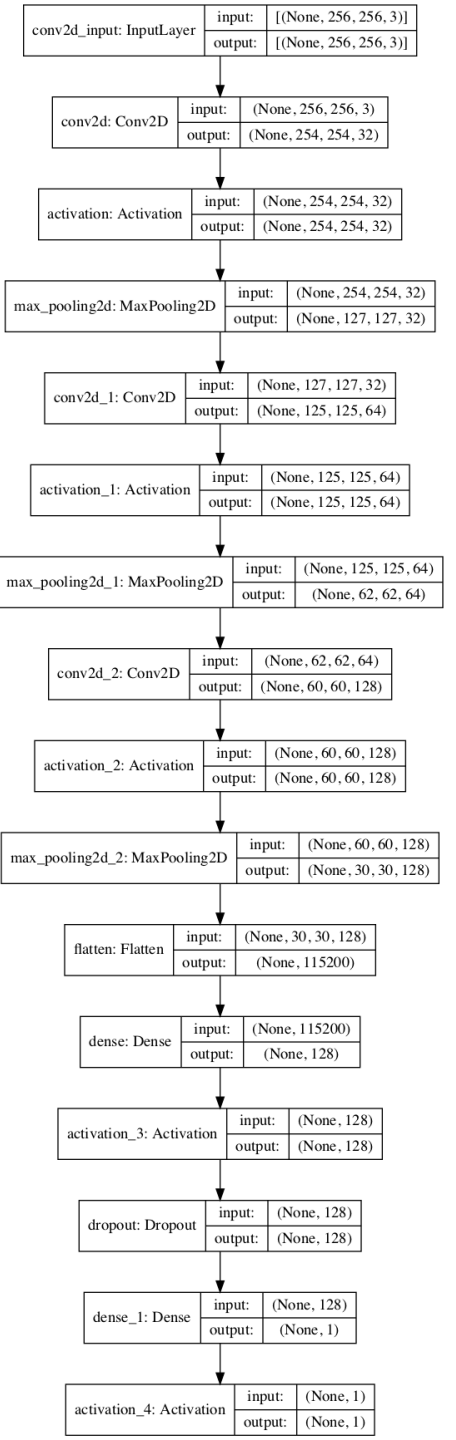
Examples of x-ray images from each class in the dataset are shown above



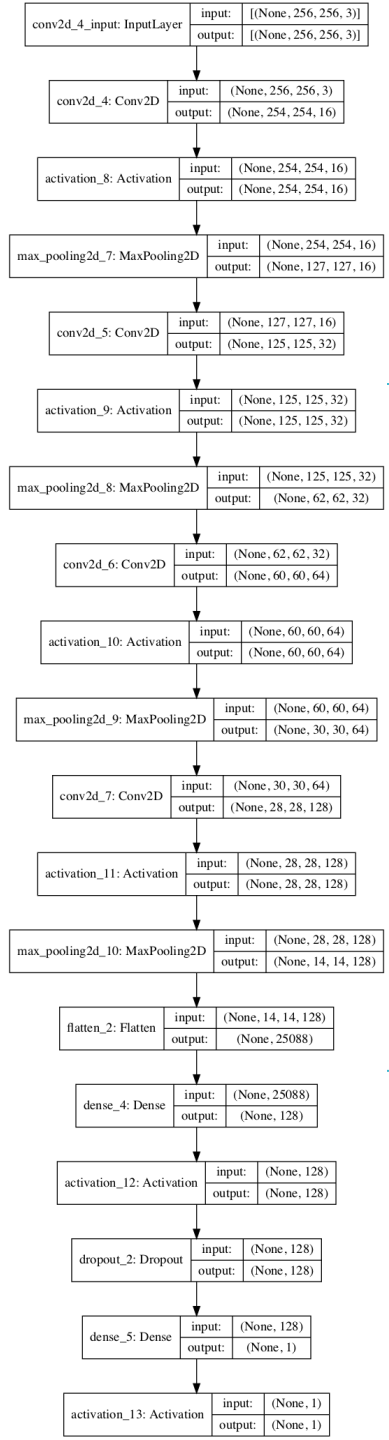


- In parsing through the datasets and directory structure, we have more images under the pneumonia category than normal (70:30 ratio)
- This could pose an issue with classification, but can be mitigated by data augmentation or fine tuning the layers of the network in order to accommodate for this imbalance

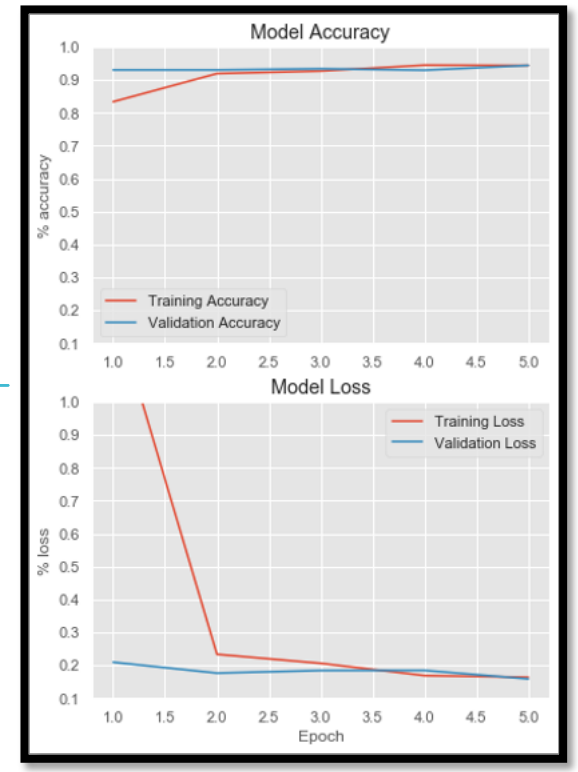
Model Results



`model_0`



`model_2`




```
results_model_2 = model_2.evaluate(test)
```

```
16s 403ms/step - loss: 0.1586 - accuracy: 0.9437
```

Conclusions and Future Steps

- Using a convolutional neural network with specific parameters, we were able to achieve **93% accuracy** in classifying images as normal or pneumonia.
- Although this was a relatively high accuracy rate, there was some "loss," around 20%, so some optimizations could be made on the model
- Further data augmentation methods, like rotations or other transformation in preprocessing, could mitigate the slight imbalance in data between the 2 classes
- Running more epochs could likely see an increase in accuracy and decrease in loss, so this would probably be the next best step in the immediate future
- Overall, we are facing a paradigm shift in medicine, using more AI and deep learning concepts in diagnostics.
- Further research on this shows that, while humans are still going to be important in diagnostics, allowing a sophisticated algorithm to make objective decisions will transform health care for the better[1]

Sources:
1) https://internal-journal.frontiersin.org/articles/10.3389/fmed.2019.00185/full

Thank you!