

Machine Learning Engineering Nanodegree

Capstone Proposal

ChestR

Disease predictor for chest
X-Rays

Shishir Horane

January 30, 2019

1 Domain Background

In medical industry chest x-ray is the most commonly performed diagnostic x-ray examination. It produces images of the heart, lungs, airways, blood vessels and the bones of the spine and chest. An x-ray(radiograph) is a non-invasive medical way for physicians to diagnose and treat medical conditions. Performing such an x-ray involves exposing the the chest to a small dose of ionizing radiation to produce the pictures of the body. X-rays are the most frequently used form of medical imaging.

The chest x-ray is commonly the first imaging test used to help diagnose symptoms such as chest pain, breathing difficulties, bad or consistent cough etc. Physicians will use these imaging to diagnose or monitor treatments for conditions such as pneumonia, heart problems, lung cancer, fluid or air collection in lungs etc. Although this isn't necessarily a conclusive diagnosis for all kind of diseases it is the first tool of diagnosing any possible illnesses for the above said symptoms.[1]

2 Problem Statement

Chest X-ray exams are one of the most frequent and cost-effective medical imaging examinations available. However, clinical diagnosis of a chest X-ray can be challenging and sometimes more difficult than diagnosis via chest CT.[2]

Reading and diagnosing chest x-ray images may be a relatively simple task for radiologists but, in fact, it is a complex reasoning problem which often requires careful observation and knowledge of anatomical principles, physiology and pathology. Such factors increase the difficulty of developing a consistent and automated technique for reading chest X-ray images while simultaneously considering all common thoracic diseases.[3]

There is a need to develop an automated technique that could analyze a given chest x-ray and which could predict the possible common thoracic diseases with some certainty. Also, it should aid radiologists or physicians by highlighting the suspected common areas of disease in order to pay close attention to certain areas of a chest x-ray, so as to not misdiagnose these diseases. This automated technique could benefit patients in developing countries who do not have access to the radiologists to read these x-rays.

3 Datasets and Inputs

National Institutes of Health (NIH) has released over 100,000 chest x-rays from more than 30,000 patients, including many with advanced lung disease. This data has been heavily compiled, cataloged and labeled for the diseases.[3] The data has been aggregated and available for download from the Kaggle website.[4]

The dataset contains 112,120 x-ray images of chest. Each of size 1024 X 1024 pixels in dimensions. These images have been categorized and labeled in *Data-entry-2017.csv* file for the diseases seen in the image, along with all the relevant patient information. The dataset contains labels for 14 diseases that has been tagged for over 50,000 chest x-ray images, along with remaining 60,000 x-rays with no findings label.

4 Solution statement

My proposal is to build a Convolutional Neural Network (CNN) trained on the above mentioned labeled chest x-rays dataset which would predict the disease with some certainty, named **ChestR**. This would be the perfect application for CNN, given that we want to find a pattern in the images and the given dataset perfectly encapsulates all the

required information to train a convolutional neural network. This well trained neural network is expected to predict better than any randomized predictions on for the thorax diseases. Also, the visualization of these CNNs would help us highlight the pattern/disease areas in the chest x-ray, which would be a good use case for aiding and assisting physicians or radiologists in looking for the suspect disease in the x-ray.

5 Benchmark Model

There are a few benchmark models for doing such disease predictions. ChexNet: Radiologist-level Pneumonia Detection on Chest X-rays with Deep Learning[5] is a 121 layer convolutional neural network that was trained on the same 112,000 xrays provided by NIH. This model was able to achieve state of the art results on predicting Pneumonia as well as the rest of the 14 diseases classified in the above mentioned dataset along with the F1 scores of the model for pneumonia prediction and the radiologists that participated in the study, along with the ROC metric for all the 14 disease predictions.

Another such benchmark model is one of the available kernel on kaggle that has simple x-ray trained CNN[6]. This kernel also expresses the results as ROC curve as a metric for measuring the quality of the model.

6 Evaluation Metrics

The ROC metric will be used to quantify the model, ChestR, results for the predictions of all the 14 diseases. These ROC values will be compared against the above mentioned benchmark models.

7 Project Design

Following will be the design approach for ChestR

- **Data analysis**

A thorough analysis of the data will be done to see the pattern of the disease label. If required the data normalization will be done here. Also, the data will be split into train, validation and test datasets if required. This dataset will be consistently used for comparing results when training models and measuring the accuracy of the model.

- **Base MLP Model**

Develop a base Multi Layered Perceptron model to see what kind of predictions we get here. This will be treated as the base model for comparing results with our fully trained CNN model.

- **Develop fully trained CNN Model**

Develop and train a deep convolutional neural network. Use the checkpoint method to train the model with varying parameters and layers to find the best result for the predictions on the testing dataset.

- **Transfer learning**

Develop and train a CNN by applying learnings from the previously trained ResNet50 or Inception or any of the well trained models on the ImageNet dataset. This transfer learning trained model will provide another comparison datapoint for our trained CNN model. The accuracy on the testing dataset will be used to compare this model.

- **Evaluation metrics**

Generate the ROC and accuracy metrics for the best trained CNN model. Compare these measurements against the benchmark models, if possible.

- **Visualize CNN**

The best accuracy model from above will be used to visualize the layers to find the pattern of diagnosing a disease. This will highlight the disease areas for predictions.

- **Visualize Predictions**

A subset the test dataset will be used to visualize the predictions, including false positives. The observations for success or failure in predictions will also be made based on this analysis.

References

- [1] <https://www.radiologyinfo.org/en/info.cfm?pg=chestradi>
- [2] <https://www.kaggle.com/nih-chest-xrays/data/home>
- [3] <https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available>
- [4] <https://www.kaggle.com/nih-chest-xrays/data/downloads/data.zip/3>
- [5] <https://arxiv.org/pdf/1711.05225.pdf>
- [6] <https://www.kaggle.com/kmader/train-simple-xray-cnn>