DETECTING AND CLASSIFYING LUNG DISEASES USING X-RAY IMAGES

Lu Lechuan Ong Kuan Yang Rhynade See Suyash Shekhar Tan Jin Wei Teo Ming Yi

ABSTRACT

In this project, we are using chest X-ray images to detect the presence of lung diseases and classifying the images with diseases into 14 different types of lung diseases. Our goal is to build upon and improve the current best performing model - CheXNet.

CheXNet

CheXNet is a 121-layer Dense Convolutional Network (DenseNet) developed by Rajpurkar et al. from the Stanford ML group.

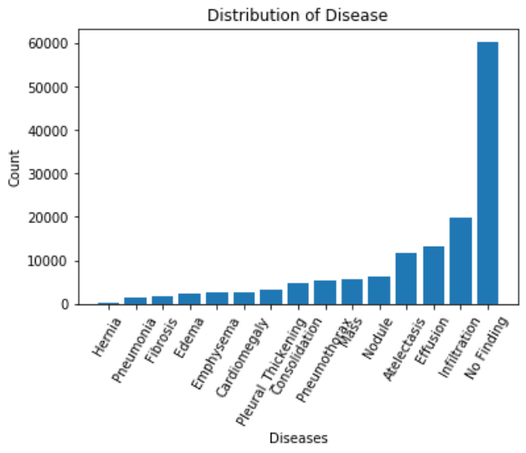
The last fully connected layer produces a 14-dimensional output, after which an elementwise sigmoid nonlinearity is applied to get the output, which is the predicted probability of the presence of each of the 14 disease classes.

DATASET

NIH Chest X-ray dataset 112,120 frontal X-ray images 30,805 unique patients 14 disease image labels

This dataset contains multi-label images - there are 20,796 images labelled with more than one disease.

This dataset is severely imbalanced:
53% of the dataset is “No Finding”
Minority class - "Hernia" only makes up 0.2% of the dataset.



PREPROCESSING

Different image preprocessing techniques were explored to extract relevant features:
- Histogram equalisation
- Contrast stretch
- Edge detection

Original Image | Image with histogram equalisation | Image with edge detection

MODEL

To tackle the imbalanced dataset problem, our team decided to use a 2 stage classifier.
- First Stage: Binary classifier, Finding vs No Finding.

- Second Stage: Further classify Finding to one or more of the 14 different diseases.

1st Stage Network Architecture:
Baseline Model
- Naive Bayes

Deep Learning Model
- Original CheXNet model, replacing last fully connected layer with 1 output and adding a sigmoid activation function.
- Optimizer: SGD with learning rate = 0.01, momentum = 0.9 and weight decay = 1e-4
While training, we reduce the learning rate by a factor of 2 when validation loss is not improved. Training loss function used is Binary Cross Entropy.

Simulated Annealing
- Avoid local minima by allowing acceptance of worse epochs.
- Probability of acceptance decreases along each iteration Ideally it should reach the global minimum in the final epoch.

2nd Stage Network Architecture:
Imbalanced dataset is still a problem within the disease-only data.

19,894 images of the Infiltration yet only 227 images of Hernia. Imbalanced ratio of approximately 100:1.

To reduce the skewness in the dataset, our team decided to implement a two-phase convolutional neural network(CNN).
- Minority classes are oversampled while majority classes are undersampled
- In the first phase, the model is trained on a balanced dataset.
- Allows to model to learn from an equal representation of each disease class and extract relevant features.
- In the second phase, the model is then trained on the actual dataset
Only weights in the final layers are modified

EVALUATION METRICS

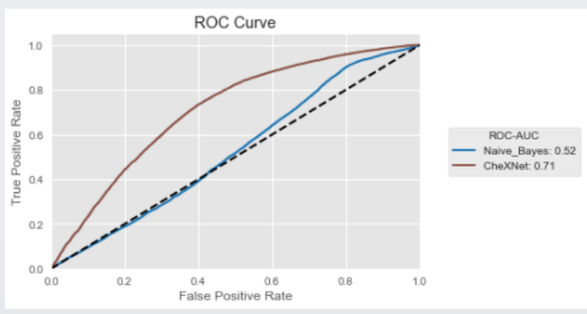
Area Under the Receiver Operating Characteristic curve (AUROC):
- ROC curve plotted with True Positive Rate (TPR) as the y-axis & False Positive Rate (FPR) as the x-axis
- Allows us to evaluate performance of our model over all possible thresholds and visualize the tradeoff between the TPR and FPR

In this problem, we are more concerned with False Negatives than False Positives. As such, True Positive Rate should be maximized as much as possible while maintaining a reasonable False Positive Rate.

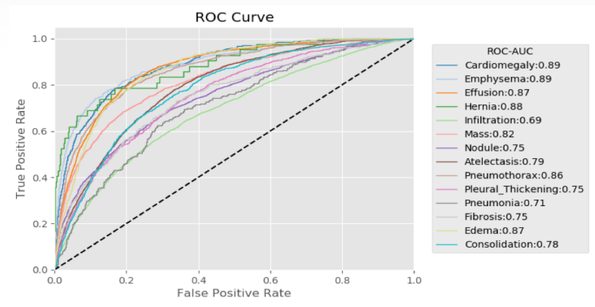
RESULTS

First Stage

Our baseline model obtained an AUC of 0.52 while the CheXNet architecture achieved an AUC of 0.71.

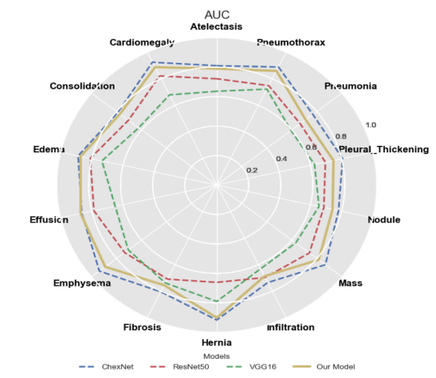


Second Stage



Comparison Across Different Models

Our model performed better than the baseline model [Wang et. al (2017)] but slightly worse than the current state-of-the-art model, CheXNet.



FINDINGS

Preprocessing

- No significant improvement observed in preprocessed images.
- Original images are already near “equalised”, hence image enhancement did not add much value to it.
- Certain preprocessing techniques might not be suitable. For e.g, corner detection might not have a significant impact as disease clusters are not as distinct as bones and ribs, and are therefore not detected.

First Stage

Naive Bayes (NB) model performed poorly compared to CheXNet. NB model has 2 main weaknesses:
1. Does not preserve spatial relationship
2. Assumption of feature independence

CheXNet addresses both problems by learning image features through applying filters to the image in a sequential order.

By applying different filters, CNNs also learn a lot of low level features such as edges. These features are further stacked together to make intricate patterns (e.g shape of lungs) and this is useful in detecting Xray images that consist of lesion regions.

Second Stage

Performance better than previous state of the art, Wang et al. 2017.
- Wang et al. used ResNet50 while our model uses a 121-layer DenseNet.
- In DenseNet, direct connections between layers allow information to flow backwards through the network.
- Combines low-level layers that recognize more granular features such as the presence of an edge and high-level layers that recognize more general features.

Performance slightly worse than current state of the art, ChexNet 2018.
- Skewness in data did not impair the performance as much as we hypothesized, two phase CNN did not improve results.
- Analysis of image shows ambiguity in data which limits performance of model.

No Finding

Mass I Pleural Thickening

[1] Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification: https://arxiv.org/pdf/1803.02315.pdf
[2] CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning https://medium.com/@jrzech/reproducing-chexnet-with-pytorch-695ff9c3bf66
[3] Weakly supervised localization of thorax diseases https://arxiv.org/pdf/1705.02315.pdf