# RSNA Pneumonia Detection

Osaetin Evbuoma

*Machine Learning Nanodegree – Udacity, 2018*

**Abstract**

In this project, I will attempt to build a machine learning model using deep learning that will to detect cases of pneumonia infections in chest radiographs. The model will leverage state-of-the-art object detection and instance segmentation algorithms that will help accurately detect the ailment.

## Domain Background

Pneumonia is an infection that inflames the air sacs of one or both lungs and fills it with fluid or pus (purulent material), causing cough with phlegm or pus, fever, chills and difficulty in breathing. It is the leading infectious cause of death internationally for children under 5, killing approximately 2,400 children a day. Pneumonia accounts for approximately 16% of the 5.6 million under-five deaths, killing around 880,000 in 2016 and around 900,000 children in 2015. Most of these victims are under 2 years old.

In the USA, pneumonia accounts for over 500,000 emergency ward visits with over 50,000 deaths in 2015, keeping the ailment among the top 10 deaths in the country. In Sub-Saharan Africa, over 500,000 infants died of the ailment in 2015 and the region accounts for roughly half of pneumonia deaths worldwide. It is imperative by these ridiculously high numbers that pneumonia is a problem that need to be tackled as early and quickly as possible.

## Problem Statement

Properly diagnosing pneumonia can be a tall order because it requires the review of chest radiograph (CXR) by highly trained specialists. The specialist confirms a case of pneumonia by also examining the patient's clinical history, vital signs and laboratory examination results. Pneumonia usually reveals itself in the lungs as an area(s) of increased opacity on CXR, however, diagnosis on CXR can be complicated by several other lung conditions. These include fluid overload (pulmonary edema), bleeding, volume loss (atelectasis), lung cancer or post-radiation/surgical changes. Outside of the lungs, fluid in the pleural space (pleural effusion) also appears as increased opacity on CXR.

CXRs are the most commonly performed diagnostic imaging study. Several factors such as positioning of the patient and depth of inspiration can alter the appearance of the CXR, complicating interpretation further. In addition, clinicians are faced with reading high volumes of images every shift.

The project will attempt to develop a model that will detect visual signals of pneumonia in medical images. It will automatically locate lung opacities on CXRs.

## Datasets and Inputs

The dataset for this solving this problem and building the model is provided by the Radiological Society of North America (RSNA) in collaboration with the US National Institute of Health, The Society of Thoracic Radiology and MD.ai. The dataset is made up of training and testing images which are DICOM files of CXRs and two CSV files. One CSV file contains labeled training data and the other is made up of details of the training data of the labeled CSV file.

The labeled training data CSV contains data provided as a set of patient IDs (which correspond to the filenames of the training images) and bounding boxes. The bounding boxes indicate the areas of the image that are infected with pneumonia and are defined as follows: *x*, *y*, *width*, *height*. The *x* and *y* are coordinates of the upper left corners of the bounding boxes while the *width* and *height* are the lengths of the edges of the bounding boxes. The file also contains a *Target* column that indicates if a patient ID has pneumonia. There may be multiple rows per patient ID. This means that there may be more than one CXR per patient.

| | patientId | x | y | width | height | Target |
|---|---|---|---|---|---|---|
| 0 | 0004cfab-14fd-4e49-80ba-63a80b6bddd6 | NaN | NaN | NaN | NaN | 0 |
| 1 | 00313ee0-9eaa-42f4-b0ab-c148ed3241cd | NaN | NaN | NaN | NaN | 0 |
| 2 | 00322d4d-1c29-4943-afc9-b6754be640eb | NaN | NaN | NaN | NaN | 0 |
| 3 | 003d8fa0-6bf1-40ed-b54c-ac657f8495c5 | NaN | NaN | NaN | NaN | 0 |
| 4 | 00436515-870c-4b36-a041-de91049b9ab4 | 264.0 | 152.0 | 213.0 | 379.0 | 1 |
| 5 | 00436515-870c-4b36-a041-de91049b9ab4 | 562.0 | 152.0 | 256.0 | 453.0 | 1 |
| 6 | 00569f44-917d-4c86-a842-81832af98c30 | NaN | NaN | NaN | NaN | 0 |
| 7 | 006cec2e-6ce2-4549-bffa-eadfcd1e9970 | NaN | NaN | NaN | NaN | 0 |
| 8 | 00704310-78a8-4b38-8475-49f4573b2dbb | 323.0 | 577.0 | 160.0 | 104.0 | 1 |
| 9 | 00704310-78a8-4b38-8475-49f4573b2dbb | 695.0 | 575.0 | 162.0 | 137.0 | 1 |

*Figure 1: Training Data Label*

The dataset was obtained via Kaggle and has a total size of 3GB.

## Solution Statement

This pneumonia detection problem will be tackled using a deep learning (computer vision) algorithm known as Mask Region-based Convolutional Neural Network (Mask R-CNN). Mask R-CNN is an instance segmentation algorithm that allows for the identification of pixel-wise locations of a desired class (pneumonia opacity). This means segmenting individual cases on pneumonia in a CXR. Mask R-CNN combines Faster R-CNN's detection of region proposals and recognition of objects in each region by predicting bounding boxes with Fully Convolutional Network's (FCN) semantic segmentation to detect instances of an object class in an image.

Mask R-CNN will be used on the training images (with validation done on a subset of the training images) to learn what pneumonia on CXRs looks like. Images with bounding boxes are expected to be learned and differentiated from images without bounding boxes.

## Benchmark Model

Mask Region-based Convolutional Neural Network outperforms all state-of-the-art models in object detection and instance segmentation on the Common Object and COntext (COCO) dataset and using varying Intersection over Union (IoU) metrics (this is explained further in the *Evaluation Metrics* section). State-of-the-art object detection model include R-FCN[1], SSD[2], YOLOv2[3] and NASNet[4]. The following table gives a breakdown of comparison of Mask R-CNN against the other models.

| **Model** | **COCO 2015/2016 (IoU = 0.5)** | **COCO 2015/2016 (IoU = 0.75)** | **COCO 2015/2016 (Official metrics)** |
|---|---|---|---|
| R-FCN | 53.2% | | 31.5% |
| SSD | 48.5% | 30.3% | 31.5% |
| YOLOv2 | 44.0% | 19.2% | 21.6% |
| NASNet | | 43.1% | |
| Mask R-CNN | 62.3% | 43.3% | 39.8% |

## Evaluation Metrics

The solution model will be evaluated on the mean Average Precision (mAP) at different Intersection over Union (IoU) thresholds. IoU is a measure of the overlap between two bounding boxes containing objects. It calculates the size of the overlap between the two objects, divided by the total area of the objects combined.

$$IoU(A, B) = \frac{A \cup B}{A \cap B}$$

*A: predicted bounding box (object); B: ground truth box (object)*

The metric sweeps over a range of IoU thresholds, at each point calculating an average precision value. The threshold values range from 0.4 to 0.75 with a step size of 0.05: *(0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75)*. In other words, at a threshold of 0.5, a predicted object is considered a "*hit*" if its IoU with a ground truth object is greater than 0.5.

At each threshold value *t*, a precision value is calculated based on the number of true positives (TP), false negatives (FN), and false positives (FP) resulting from comparing the predicted object to all ground truth objects:

$$\frac{TP(t)}{TP(t) + FN(t) + FP(t)}$$

[1] Region-based Fully Connected Network (F-RCN) [https://arxiv.org/pdf/1605.06409.pdf]
[2] Single-Shot Detector (SSD) [https://arxiv.org/pdf/1512.02325.pdf]
[3] You Only Look Once (YOLO) [https://arxiv.org/pdf/1506.02640.pdf]
[4] Neural Architecture Search Net (NASNet) [https://arxiv.org/pdf/1611.01578.pdf]

A true positive is counted when a single predicted object matches a ground truth object with an IoU above the threshold. A false positive indicates a predicted object had no associated ground truth object. A false negative indicates a ground truth object had no associated predicted object. If there are no ground truth objects at all for a given image, any number of predictions (FPs) will result in the image receiving a score of zero and being included in the mAP.

The average precision of a single image is calculated as the mean of the above precision values at each IoU threshold:

$$\frac{1}{|thresholds|}\sum_{t}\frac{TP(t)}{TP(t)+FN(t)+FP(t)}$$

Finally, the score returned by the evaluation metric is the mean taken over the individual average precisions of each image in the test dataset.

$$\frac{1}{n}\sum\left(\frac{1}{|thresholds|}\sum_{t}\frac{TP(t)}{TP(t)+FN(t)+FP(t)}\right)$$

*n: total images predicted (test images)*

# Project Design

The model to be developed for identifying cases of pneumonia is CXRs will be solved using deep learning and more specifically, using the theories of convolutional neural networks.

***Dataset Analysis***

The training dataset contains images of patient x-ray scans. Some images have cases of pneumonia while others do not. Roughly half of the dataset are entries with pneumonia. Additionally, the dataset contains some rows with the same patient ID but different coordinates that indicate pneumonic areas in the CXR. This also indicates that the images are the same (except for the coordinates of pneumonia instances). Therefore, the dataset will be parsed to combine the coordinates of pneumonia for all patients. This will mean that patients with multiple cases of pneumonia will have the pneumonia coordinates in one patient object. Considering that the coordinates of patients with pneumonia have been provided, there will be no need to manually annotate cases of pneumonia.

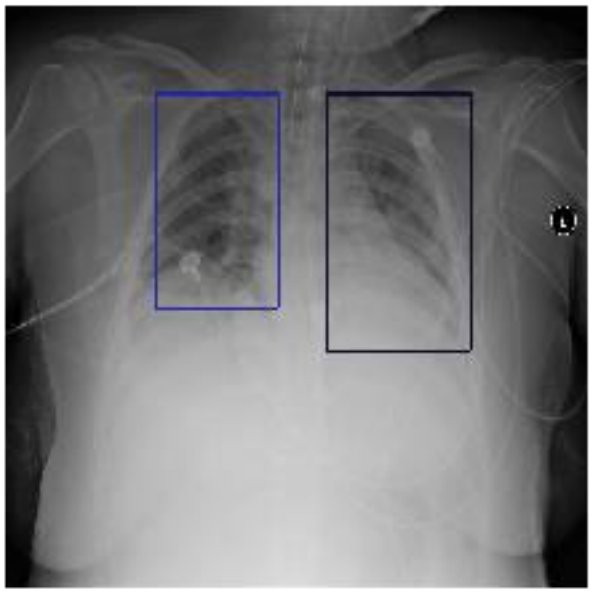

*Figure 2: Sample patient CXR using annotated data ({'dicom': 'stage_1_train_images/00436515-870c-4b36-a041-de91049b9ab4.dcm', 'label': 1, 'boxes': [[152.0, 264.0, 379.0, 213.0], [152.0, 562.0, 453.0, 256.0]]})*

***Training and prediction***

The parsed data will be trained with the Mask R-CNN algorithm as implemented by Matterport (https://github.com/matterport/Mask_RCNN). Only parsed data with annotated coordinates of pneumonia will be trained. Images without pneumonia will not be required since the model will not have any annotations to learn from. Matterport's implementation is chosen because of its integration with Tensorflow and its ease of implementation. The training will be done an Amazon Machine Image (AMI) EC2 instance running on GPU and will be run for at least 20 epochs. This might take a couple of hours to complete. The dataset will be split in a 9:1 ratio representing training and validation datasets respectively which will be used in training the model.

To kick off the training, already trained weights of Mask R-CNN on the COCO dataset will be used. This should help the model kick off training faster as against starting with randomly selected weights. At each training period, new weights will be calculated and compared to previous weights. If the new weights offer better learning results, they will be saved and used for further training and eventual inference.

The final trained model will be used on the test dataset to predict images with cases of pneumonia infections and the results will be evaluated based on the evaluation metrics to determine its performance and accuracy.

# References


1. Rui P, Kang K. National Ambulatory Medical Care Survey: 2015 Emergency Department Summary Tables. Table 27. Available from: www.cdc.gov/nchs/data/nhamcs/web_tables/2015_ed_web_tables.pdf

2. Deaths: Final Data for 2015. Supplemental Tables. Tables I-21, I-22. Available from: www.cdc.gov/nchs/data/nvsr/nvsr66/nvsr66_06_tables.pdf
3. Franquet T. Imaging of community-acquired pneumonia. J Thorac Imaging 2018 (epub ahead of print). PMID 30036297
4. Kelly B. The Chest Radiograph. Ulster Med J 2012;81(3):143-148
5. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 2017, http://openaccess.thecvf.com/content_cvpr_2017/papers/Wang_ChestX-ray8_Hospital-Scale_Chest_CVPR_2017_paper.pdf
6. UNICEF: https://www.unicef.org/media/media_89995.html
7. Arthur Ouaknine. Review of Deep Learning Algorithms for Object Detection. https://medium.com/comet-app/review-of-deep-learning-algorithms-for-object-detection-c1f3d437b852
8. Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick. Mask R-CNN. https://arxiv.org/pdf/1703.06870.pdf