

Chest X-Ray Abnormalities Detection with a focus on Infiltration

Linh Ta

Computer Science Department at Earlham College
Richmond, Indiana
lta18@earlham.edu

ABSTRACT

Navigating chest X-rays is an obligatory step to determine lung and heart diseases. Since many people now believe that Chest X-ray radiographs can detect COVID-19, the disease of decade, the problem of Chest X-ray Abnormalities Detection has gained an increasing attention from researchers. Numerous machine learning algorithms have been developed to address this problem to raise reading accuracy, improve efficiency, and save time for both doctors and patients. In this work, I propose a model to determine whether a Chest X-ray image has Infiltration and to detect the abnormalities in that image using YOLOv3. The model will be trained and tested with the VinBigData dataset. Overall, I will use the existing tool, YOLOv3, on a new problem of detecting Infiltration in Chest X-ray radiographs.

KEYWORDS

computer vision, medical image processing, chest radiograph, chest x-ray, image detection, abnormal chest X-ray identification

1 INTRODUCTION

Health has always been an essential social concern. According to the Peter G. Peterson Foundation, the cost of healthcare in the United States has dramatically increased over the past few decades, "from 5 percent of gross domestic product (GDP) in 1960 to 18 percent in 2018." [3] The continuous demand has motivated research and the creation of new medical devices to reduce the examination time and cost and to increase diagnostic accuracy. Along with physical medical equipment such as infusion pumps, medical lasers, and LASIK surgical machines, researchers now also pay attention to developing AI/ML-based software as a medical device. Multiple topics are under study with a view of AI/ML, including Chest X-ray (CXR) Abnormalities Detection. As evidence of the considerable concern in this topic, over 736,000 results are found in Google Scholar with this keyword. Besides the natural importance of chest observation for common heart, lungs, and bone diseases, the outbreak of Covid-19 in 2020 has further driven interest in the CXR Abnormalities Detection topic. It is widely believed that the CXR images might be useful in the process of diagnosing Covid-19 [1].

My research will apply different algorithms to detect Infiltration, the filling of air spaces with a substance denser than air (pus, blood, or protein) in a region of the lung, in CXR radiographs. The goal is to determine a model with high accuracy and speed on localizing the Infiltration in CXR images. My contribution in this research will be scrutinizing common approaches nowadays in the Object Detection field and their applications in CXR abnormalities detection to build a model with a fair performance in terms of speed and accuracy for

this specific problem. Specifically, I will preprocess the dataset in regards to the main focus of this project, Infiltration. After that, I will build a CXR abnormalities detection model using YOLOv3 [8] structure.

2 RELATED WORK

CXR Abnormalities Detection is a subproblem in the field of Object Detection. In simple terms, Object Detection means classifying if the image contains the target object and identifying the object's position within the image. Although there are numerous approaches towards CXR Abnormalities Detection, I will focus on the most common algorithms, YOLO (v1, v2, and v3) and R-CNN.

2.1 YOLOv1

YOLO [9] (You Only Look Once) was first introduced in 2016 by Redmon et al. as a new approach for object detection. According to Redmon et al., earlier methods (like R-CNN) are "slow and hard to optimize" because of their complexity. On the other hand, as its name suggests, YOLO requires only one look at the object to detect what and where the object is. It makes use of the single convolutional network inspired by GoogLeNet and sum-squared error for the loss function. Therefore, YOLO is an innovative way to detect objects with fast speed, small background errors, and the ability to learn generalizable representations of objects. According to Du [2], "On Pascal VOC2007, YOLOv1 processes images at 45 frames per second (FPS) which is twice to nine times faster than that of Faster R-CNN." However, some drawbacks of YOLO are incorrect localizing some objects, especially small ones in groups, such as flocks of birds, and the lack of generalizing objects capacity in unusual aspect ratios or configurations.

2.2 YOLOv2

Aware of those shortcomings, Redmon et al. established an updated version of YOLO, which is called YOLO9000 or YOLOv2 [7], in 2017. This version focuses on solving the issue of incorrect localization while maintaining classification accuracy. Redmon used three crucial ideas for YOLO's improvement: Batch Normalization, High Resolution Classifier, and Convolutional With Anchor Boxes [5]. YOLOv2 utilizes WordTree to combine data sources, ImageNet and COCO to train data simultaneously with their joint optimization technique, and Darknet-19 as the network architecture. Although the idea of jointed images from both detection and classification data increases performance for unknown objects, YOLOv2's performance on categories like "sunglasses" or "swimming trunks" significantly decreases due to COCO's lack of bounding box labels.

2.3 YOLOv3

In 2018, Redmon et al. published YOLOv3 [8] with multi-labels detection, meaning that an object can have more than one label. Therefore, YOLOv3 applies independent logistic classifiers rather than softmax in object classification and binary cross-entropy loss in class predictions. Furthermore, YOLOv3 also provides objectness score (a measure of how well the detector identifies the locations and classes of objects during navigation) for each bounding box using logistic regression and modifying the cost function. Unlike the previous versions, YOLOv3 makes use of Darknet-53 as the network architecture for feature extraction, Feature Pyramid Networks (3 predictions per location) for object predictions, and K-means clustering for determining the priors. As a result, YOLOv3 shows remarkable improvement in detecting small objects. However, the performance for medium and large objects decreases instead. As we can see in Figure 1, there are 106 layers fully convolutional underlying architecture, in which 53 layers are trained on Imagenet while the other 53 are for the task of detection. In the training process, YOLOv3 will downsample the dimensions of the input image by 32, 16 and 8 respectively and make prediction.

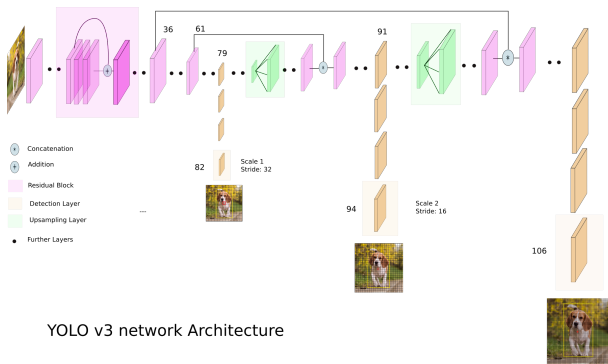


Figure 1: YOLOv3 Network Design [6]

2.4 Faster R-CNN CXR Abnormalities Detection

In their work, Rahmat et al. [10] apply Faster R-CNN, an updated version of R-CNN - one of common approaches for Object Detection, in detecting CXR abnormalities. The work uses NIH Clinical Center Dataset, where 80% of the dataset was used for training, and the remaining was used for testing. The model has an accuracy of 62% while that of a medical student is 56%, and the general practitioner is 50%. This is a reasonably good result, although there is still room for improvement, as indicated in the paper. Therefore, it can be said that Faster R-CNN is a practical and effective algorithm for detecting CXR abnormalities.

3 DATASET

I will use VinBigData dataset for this research. The dataset [4] is collected by the Vingroup JSC, the Vingroup Big Data Institute (VinBigData) since August 2018. The whole dataset was provided by two Vietnamese hospitals: Hospital 108 and the Hanoi Medical University Hospital. It contains 18,000 frontal-view CXR images

in DICOM format, which were labeled by a panel of experienced radiologists. The original dataset is divided into 14 critical types of chest abnormalities radiographs, labeled respectively from 0 to 13, and 14 for no finding observation. The labels provide information about disease classification and the coordinates of the object's bounding box. However, due to time constraints, I will focus on one abnormality type only - Infiltration. I choose this dataset because it labels both bounding boxes and classes for every image and has high accuracy.

4 METHODS

4.1 Framework

For this project, I use Python for building the Infiltration detector model and Slurm for automation. Figure 2 describes the overview of the software architecture. The model aims to identify the Infiltration and detect abnormal regions if exists. 80% of the dataset will be the train set and the other 20% will be the test set.

First, the images in both training and testing sets will go through data preprocessing, in which they are relabeled, resized, and reshaped using cv2 (Python). Since this step takes a huge amount of time, the images will be processed in parallel using Slurm to fully take advantage of the hardware resources. After that, I will train the model, YOLOv3, by feeding the processed images to the model while using Slurm. After training, the model should be able to classify if the Infiltration exists and generate bounding boxes for abnormalities in CXR radiographs. Subsequently, the model will be evaluated using Average Precision (AP) metric. This evaluation metric is chosen since it is a common measure of accuracy on prediction tasks. In order to improve the accuracy, I will try several techniques such as data augmentation, and learning rate scheduler. Specifically, data augmentation increases the images' variety by horizontal flipping, resizing, and color filtering, while the learning rate scheduler reduces the learning rate as a constant number of training epochs increases.

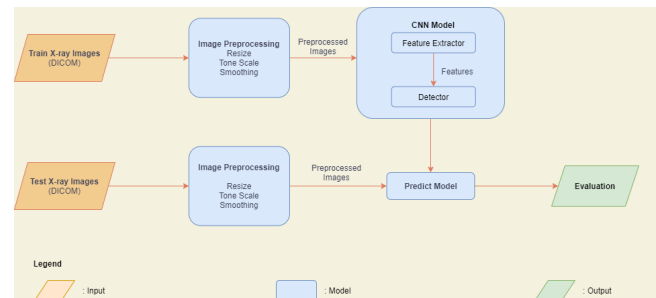


Figure 2: Software Architecture Diagram for CXR Abnormalities Detection

4.2 Preprocessing Methods

4.2.1 Format Conversion. The original format type of the VinBigData dataset is Digital Imaging and Communications in Medicine (DICOM), the standard for medical related imaging data. Since the model cannot read this specific format, format conversion from DICOM into readable types such as JPG and PNG is necessary.

4.2.2 Resizing. Since all the images in the dataset need to have the same shape for the training process, the images in both training and testing sets will be resized into (416, 416). However, cropping the images into the targeted ratio might result in the scenario where the lung part is cut. Therefore, we want to resize the images while preserving the width and height ratio. Another problem with resizing images is that values of existing annotation for bounding boxes are only correct with the current image size. As a result, the annotation file is also needed to be recalculated based on ratio between old and new width and height values.

4.2.3 Tone Scaling. Tone scaling transforms all images into gray-scale tone and inverts them if necessary. This step will boost the speed performance and accuracy.

4.2.4 Smoothing. The Gaussian data smoothing technique is applied to reduce noises from original CXR images. This step helps in creating a dataset with more accurate intensity surface.

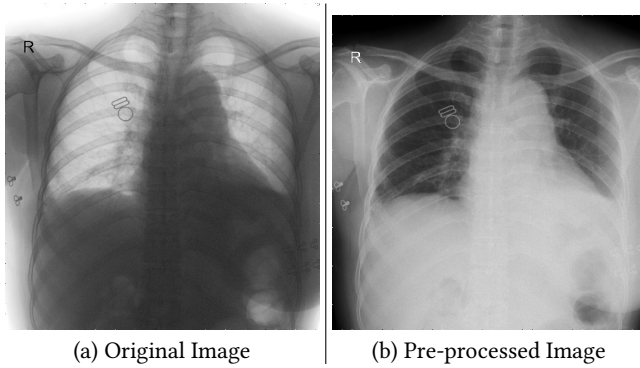


Figure 3: Example of before and after pre-processed images

4.3 YOLOv3 Model

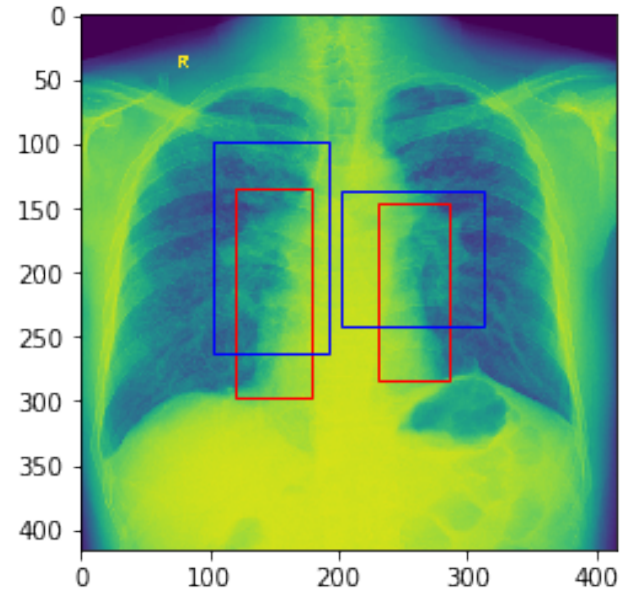
YOLOv3 model consists of two main parts: feature extraction and prediction. For feature extraction, the used network is the hybrid between the network from YOLOv2 and newfangled residual network. For prediction step, the bounding boxes are predicted using dimension clusters as anchor boxes with logistic regression. YOLOv3 predicts bounding boxes at 3 different scales, which are chosen arbitrarily. The classes for each bounding box are then decided by multilabel classification technique with independent logistic classifiers.

4.4 Evaluation Methods

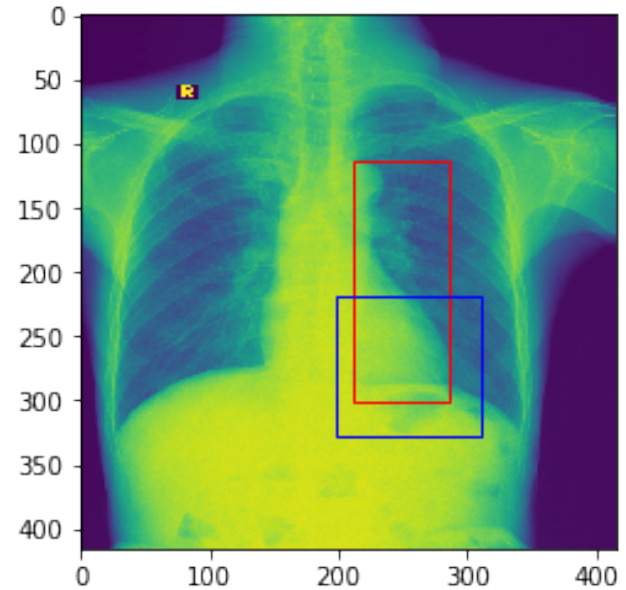
Average Precision (AP) is the evaluation metric for accuracy. It will be calculated using precision, recall, and intersection over union (IoU). While precision measures the percentage of correct positive predictions over all positives, recall measures the percentage of correct positives over all correct positives and negatives. Intersection over union measures overlap areas between ground truth and detection boundaries in order to classify the prediction is true or false positive.

The model will be tested and evaluated with Average Precision after training.

5 RESULTS



YOLOv3 running on sample 1



YOLOv3 running on sample 2

Figure 4: Samples of Chest X-ray images running through YOLOv3 are shown. Red bounding boxes are labeled by radiologists while blue bounding boxes are generated by YOLOv3.

The YOLOv3 model was trained with the VinBigData dataset multiple times. The batch size for each training session is 1000 samples.

After each training session, the weights used for the model are recalculated based on previous sessions and recorded for later usages. The overall average precision is about 83%.

6 FUTURE WORK

If I have more time and data, I would like to try out different models and algorithms, such as Faster R-CNN, YOLOv5, and Detectron, for this specific problem. I also want to expand this model to predict other types of lung diseases besides Infiltration.

7 ACKNOWLEDGEMENTS

I would like to express my special thanks of gratitude to Dr. Charlie Peck and Dr. David Barbella for their guidance and support in completing my project.

8 APPENDIX

Gitlab URL: <https://code.cs.earlham.edu/lta18/lta18-senior-capstone>
The repository contains description, software architecture diagram, video link, dataset's location, and instructions to run the code.

REFERENCES

- [1] BMJ. *The role of chest radiography in confirming covid-19 pneumonia*. July 2020. URL: <https://www.bmj.com/content/370/bmj.m2426>.
- [2] Juan Du. *Understanding of Object Detection Based on CNNFamily and YOLO*. 2018. URL: <https://arxiv.org/pdf/1804.02767>.
- [3] Peter G. Peterson Foundation. *WHY ARE AMERICANS PAYING MORE FOR HEALTHCARE?* Apr. 2020. URL: <https://www.pgpf.org/blog/2020/04/why-are-americans-paying-more-for-healthcare>.
- [4] Vingroup Big Data Institute. *VinBigData Chest X-ray Abnormalities Detection*. 2018. URL: <https://www.kaggle.com/c/vinbigdata-chest-xray-abnormalities-detection/data>.
- [5] Amro Kamal. *YOLO, YOLOv2 and YOLOv3: All You want to know*. 2019. URL: https://medium.com/@amrokamal_47691/yolo-yolov2-and-yolov3-all-you-want-to-know-7e3e92dc4899.
- [6] Ayoosh Kathuria. *What's new in YOLO v3?* 2018. URL: <https://towardsdatascience.com/yolo-v3-object-detection-53fb7d3bfe6b>.
- [7] Joseph Redmon and Ali Farhadi. *YOLO9000: Better, Faster, Stronger*. Dec. 2016. URL: <https://arxiv.org/pdf/1612.08242>.
- [8] Joseph Redmon and Ali Farhadi. *YOLOv3: An Incremental Improvement*. Apr. 2018. URL: <https://arxiv.org/pdf/1804.02767>.
- [9] Joseph Redmon et al. *You Only Look Once: Unified, Real-Time Object Detection*. May 2015. URL: <https://arxiv.org/pdf/1506.02640>.
- [10] Sharifah Aliman Taufik Rahmat Azlan Ismail. *CHEST X-RAY IMAGE CLASSIFICATION USING FASTER R-CNN*. May 2019. URL: https://www.researchgate.net/publication/335501390_Chest_X-ray_Image_Classification_using_Faster_R-CNN.