Automatic Detection of Pneumonia from Chest X-rays Using Deep Learning

Md Fazlay Rabbi Masum Billah
Department of Computer Science
University of Virginia
mb2vj@virginia.edu

Tamjid Al Rahat
Department of Computer Science
University of Virginia
tr9wr@virginia.edu

Faysal Hossain Shezan
Department of Computer Science
University of Virginia
fs5ve@virginia.edu

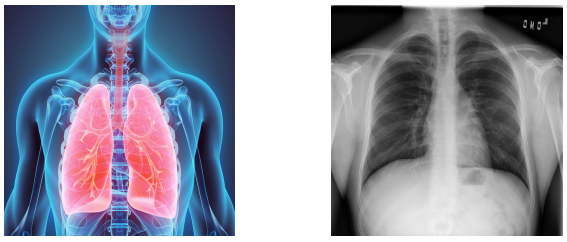
***Abstract*—Pneumonia, one of the deadliest disease and the main reason of child death in several countries of the world, is getting the concentration of the health specialist among the different countries of the world. Recently, the statistics of the number of death people, died from pneumonia in Virginia, creates a gruesome effect among the Virginia residents. One way to detect pneumonia is by analyzing the radiograph (X-ray) images of a pneumonia affected patient's lungs. There are certain characteristics present in a pneumonia affected patient's lungs. However, sometimes it becomes difficult to figure out those characteristics manually, even by some professional health specialist or doctor. In this paper, we have demonstrated an automatic machine learning classifier that will help to detect pneumonia automatically by investigating Chest Radiographs. With the help of latest deep learning techniques, we will be able to extract those important features (present in a pneumonia affected patient's) from the radiograph images, and thus successfully detect pneumonia affected patient.**

I. Introduction.

Pneumonia, an infection in one or both lungs, which can be caused by bacteria, viruses, or fungi. Of them all, bacterial pneumonia is mostly common these days. Pneumonia accounts for over 15% of all deaths of children under 5 years old internationally. In 2015, 920,000 children under the age of 5 died from the disease. In the United States, pneumonia accounts for over 500,000 visits to emergency departments [14] and over 50,000 deaths in 2015 [1], keeping the ailment on the list of top 10 causes of death in the country.

In 2016, **1,177 people died in the state of Virginia** [4] due to Pneumonia making it one of the top 10 leading causes of death. We firmly believe that automatic detection of pneumonia will help both doctors and patients to reduce this number of deaths at a significant level. Chest X-rays are considered to be the best method for detecting pneumonia [11]. However, this is still a challenging task and thus, radiologists sometimes fail to detect pneumonia.

In the era of technological advancement, modern hospital system is adorned with sophisticated medical equipment. In these days, we are not dependent only on the human doctor. We have online medicare and apps that serves the service of a human doctor. Major operations are considered to be operated successfully by the surgical robots. Moreover, it is predicted that robot doctor will perform better than human. As a matter of fact it becomes difficult to diagnosis a disease perfectly sometimes. And with the enough knowledge, we can build an automated tool that can be used to such diagnosis. Considering the impact of pneumonia, automated tool can be of great use in terms of detecting pneumonia.



(a) (b)

Fig. 1: (a) Human Lung (b) Radiograph of Human Lung

In this project, our primary goal will be the detection of bounding boxes corresponding to the diagnosis of pneumonia (e.g., lung infection) from chest radio-graphs. We have collected a pneumonia dataset RSNA [3]. Our main focus is to enable the tool to detect pneumonia by going through patient's x-ray report. We are not limiting ourselves only to that part. On top of that, we will investigate in which part of the patient's lungs exhibits the characteristics of pneumonia. As our intuition here is to make the robot able to detect as well as give proper treatment to pneumonia, finding the exact location of the pneumonia is necessary. In our dataset, we have x-ray images with the boundary pneumonia location in a lungs (if the corresponding x-ray image of the lungs contains pneumonia). So, here we use IoU(Intersection over union) metric for analyzing the performance of our model. IoU will measure the intersection are of the actual boundary box and the predicted boundary box. Our initial model is able to achieve that performance with 73% IoU (Intersection over union). After that, we have built more sophisticated model for this operation. We achieve 77.507% IoU which outperforms our initial model.

In this paper we are making the following contributions-

- Building an automatic CNN-based tool which can detect pneumonia whenever the patient's lungs x-ray images exhibits such characteristics.
- For giving a proper treatment, it is always necessary to

detect the exact location of the pneumonia. Our tool is able to detect the exact location of pneumonia with a convincing performance.

II. Related works.

Detecting pneumonia using Chest X-rays plays an important role in medical system and epidemiological studies [5], [7]. Recently, automated diagnosis from x-ray images is getting more and more attention from the researchers [8], [10]. Although deep learning based medical image analysis has been a focus for the researchers for last few years, there has been less works on pneumonia detection from chest x-rays. *Wang et al.* [17] introduced a deep learning technique to classify and localize Thorax diseases from chest X-ray images. However, their dataset is labeled by applying automatic extraction methods on associated radiology reports. For medical images, automatic labeling can be misleading and may cause significant errors. Rajpurkar et al. [12] has done a similar works on Pneumonia detection using the same dataset. Unlike the existing works, we aim to localize and classify pneumonia from a chest X-rays dataset that is labeled by human experts. Yao et al. [18] introduced statistical dependencies between labels which in turns make the prediction more accurate and outperforms Davis et al. [6]. Rajpurkar et al. built a model which consists of 121 convolution neural network to detect the existence of pneumonia in a patient's chest x-ray image [12]. However, we are not only detecting the presence of pneumonia. We are also localizing the position of the pneumonia.

III. Methods

The primary goal of this project is the detection of bounding boxes corresponding to the diagnosis of pneumonia (e.g. lung infection) on chest radiographs, a special 2D high resolution grayscale medical image. Under the hood of pneumonia detection, we actually do *Lung Opacity Detection* in chest radiographs. Lung opacity are not the same as pneumonia, rather these are vague, fuzzy clouds of white in the darkness of the lungs, which makes detecting them a real challenge.

A. Dataset.

We will use RSNA Pneumonia Detection Challenge dataset [3] in our project. This dataset contains $30,227$ bounding boxes with lung opacities for chest radio-graph images from $26,684$ patients for training purpose and additional $1,000$ patient cases for testing purpose. The dataset contains images of three classes- a) *Normal* b) *Lung Opacity* c) *No Lung Opacity / Not Normal*. *Normal* class refers to the images from patients with no Pneumonia and *Lung Opacity* class refers to the patients with Pneumonia. Figure 2 shows the difference between the images of these two classes in terms of opacities. Clearly, the radiograph with lung opacity has more fuzzy area than the normal one.

However, not all opacities in the chest radiographs imply the sign of pneumonia. There are different kind of opacities. Some are related to pneumonia and some are not. Lung diseases other than pneumonia such as bleeding, tumor, abnormal nodule and lung cancer can also cause the present of opacities in chest radiographs. This extra third class indicates that while pneumonia was determined not to be present, there was some type of abnormality on the image, and often times these abnormality may mimic the appearance of true pneumonia. Presence of this class of images make the pneumonia detection task more challenging. In our dataset, this class of images are labeled as *No Lung Opacity / Not Normal*. In Figure 3, we can see the differences between two classes of images. The opacities in the second image (right) are more likely caused by lung nodules, not pneumonia. The primary purpose of this class is to reduce the number of False Positives for the pneumonia detection task. As we focus only the Pneumonia detection task in this project, we have ignored *No Lung Opacity / Not Normal* class.

Class	Target	Patient Count
Lung Opacity	1	9555
No Lung Opacity / Not Normal	0	11821
Normal	0	8851

TABLE I: Classes and Number of patients in our dataset

Pt.Id	x	y	width	height	Target	Class
acb0...	-	-	-	-	**0**	Normal
1889...	105.0	486.0	295.0	339.0	**1**	Lung Opacity
6e9b...	-	-	-	-	**0**	No Lung Opacity

TABLE II: Bounding box annotations for the radiograph images in the dataset

Fields	Description
patientId	Each patientId corresponds to an unique image.
x	The upper-left x coordinate of the bounding box.
y	The upper-left y coordinate of the bounding box.
width	The width of the bounding box.
target	Indicates whether this sample has evidence of pneumonia.
class	Name of one of the three classes based on opacities

TABLE III: Description of different fields in the dataset

Lung Opacity In chest radiographs, lung opacity refers to an area that is more white than it should be. If we compare the chest radiographs of the two patients in the **Figure 2**, we can see that some areas of the lung of patient 2 (right) are obscured by opacity. There is clear difference between the black lung and tissue below it in the lung of patient 1, whereas in the second image there is just fuzziness. Usually the lungs are full of air. When someone has pneumonia, the air in the lungs is replaced by other material - fluids, bacteria, immune system cells, etc. That's why areas of opacities are areas that are grey but should be more black. The lung tissue in that area probably implies that the patient is not healthy. In our dataset, the hazziness of the pneumonia affected area

in the lungs are labeled with bounding boxes.

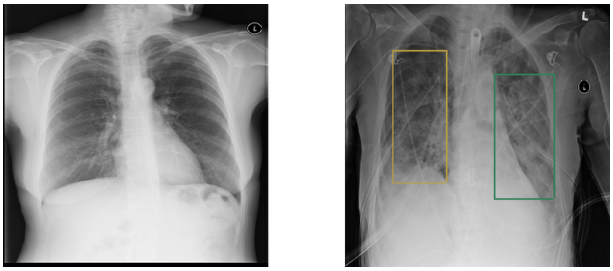


Fig. 2: Examples of **Normal** class (left) and **Lung Opacity** class (right) chest radiograph images.

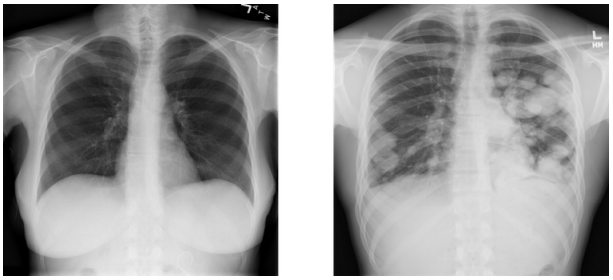
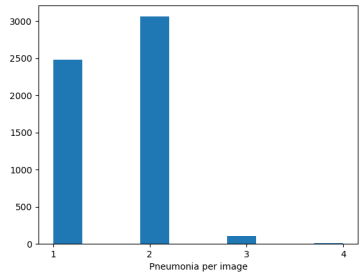
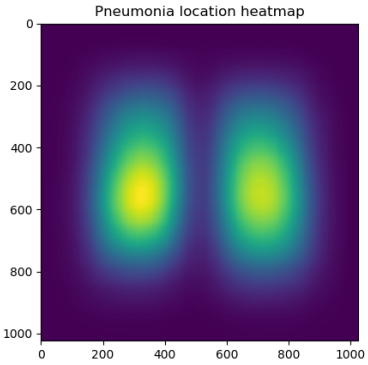


Fig. 3: Examples of **Normal** class (left) and **No Lung Opacity /Not Normal** class (right) chest radiograph images.



(a) Number of bounding box in each figure



(b) Heatmap of lung opacities location

Fig. 4: Visualization of the dataset

Pneumonia is one of many possible disease processes that can occur on a chest radiograph, and that any given single image may contain 0, 1 or many boxes corresponding to possible pneumonia locations. From Figure 4(a), we can see that most defected images contain one or two bounding boxes, which also refers to the defected region in the image. Also, Figure 4(b) shows the heatmap of the affected region in our dataset. The task of localizing the pneumonia affected region in the chest-radiograph images falls under the category of Object Detection task, which is different from Image Classification. In our dataset, each pneumonia defected image is annotated with the bounding box representing the defected region in the images. Usually, object detection model outputs mode boxes than actual objects (i.e., pneumonia defected region). Therefore, to assess the spatial precision, we need to remove the boxes with low confidence. So, we use **Intersection over Union(IoU)** area, a value that corresponds to the overlapping area between predicted box and ground-truth box. The higher IoU, the better the predicted location of a box for a given object. We keep all the boxes with an IoU greater than a threshold value 0.5.

Unlike the most common object detection problems, we have only one object to detect, which is the pneumonia defected region. So, we use mean IoU value as our performance metric.

Each image in our training dataset is a 1024x1024 dimensional pixel array. For output, we represent each bounding box in a mask array same as the size of input image. We represent the bounding boxes for each image using value 1 in the mask array, where other pixels not overlapped with the bounding box are 0.

IV. Experiments.

A. Data Pre-processing.

For finding the X-ray image that exhibits the characteristics of pneumonia, we have trained a CNN (Convolutional Neural Network) model with 25684 number of x-ray images. Among them, 5659 numbers are pneumonia image. We have extracted the position of the pneumonia from those images, and thus stored the location with the filename for further processing. However, if there is no pneumonia in a single image, we have assigned an empty pneumonia location list to them. In this way, we have generated the mapping between the pneumonia image and corresponding pneumonia location. Figure 4b depicts the heatmap of pneumonia in a x-ray image. Finally, we are able to distinguish between images with pneumonia against non-pneumonia images by going through that mapping.

B. Model Architecture.

Resnet. We have implemented a smaller version of the regular resnet model [2]. We have created four resnet blocks, where each block contains two convolutional layers. Each convolutional layer in from four resent blocks contains 64,128,256 and 512 channels respectively. We use

three kernels for each convolutional layer in our model architecture. For activation function, we use leaky Relu. We resize the input images in reflect mode into 256x256 to reduce the computation. Before feeding the inputs into each block, it was downsampled by using another convolutional layer followed by a maxpooling layer. The result is upsampled again in the output layer. We use *Intersection over Union (IoU)* as our loss function. Finally, we use AdamOptimizer for optimization during back-propagation.

YOLO V3. Our second model is built on top of the popular YOLO V3 object detection model [13]. We used their model structure and also model weight of different layers. This model predicts boxes at three different scales. Actually, YOLO model is built with several convolution layers. They added several structure of cnn network multiple times which benefits them at the end. Basically, YOLO V3 is a hybrid model built with network used in YOLO V2, Darknet-19, and newfangled residual networks.

V. Results.

We have tested our model in a GPU machine. It contains one Nvidia Tesla K40c GPU, 32GB of memory, 4TB hard drive, one 12 core Intel Core i7-5820K CPU - 3.30GHz. We explain the results of our models in the following sections:

Resnet. Initially, we have started with small learning rate and small epoch. In Figure 5, we have demonstrated the initial performance of our model. Here, we have plotted a graph between different epoch and the corresponding IoU value. We can observe that after 10 epochs, we are getting around 73% IoU. We have also observed the different loss in each epoch. From Figure 6, we can say that training and validation loss is decreasing with the increasing value of epoch.

To evaluate our model on the test images, we consider the prediction value threshold as 0.5. Figure 7 shows some of the results of our bounding box prediction. In Figure 8 we also show examples of false positives and false negatives found during the evaluation. We hope that the rate of false positive and false negative can be reduced by employing images from the third class of our dataset- No Lung Opacity/ Not Normal.

Transfer Learning using YOLO V3. After building our first model, we have used the same training data for building our YOLO model. In figure-10, we have showed the loss metric of our model on the validation data. It is calculated according to the base model's loss metric. From the figure-6 & 10, we can observe that our YOLO model's loss value is impressive compare to the Resnet model. It decrease to 1% after 1000 iteration and stays almost stable after that.

Figure-9 depicts the performance measurement in terms of IoU value. In object detection mechanism, achieving a good IoU value is always desirable. From figure 9, we can observe that after 1800 iterations our model's IoU value reach 73% which is equal to our Resnet model. However, our YOLO model outperforms our RESNET model in terms of IoU metric

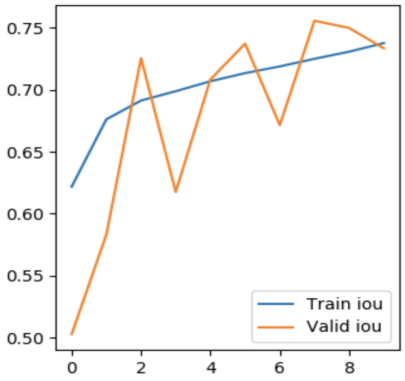


Fig. 5: Performance measurement of Resnet Model (epoch v/s IoU)

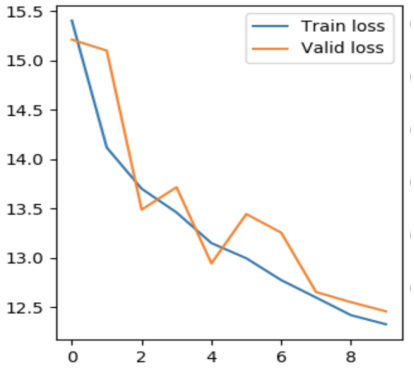


Fig. 6: Performance measurement Resnet Model (epoch v/s Loss)

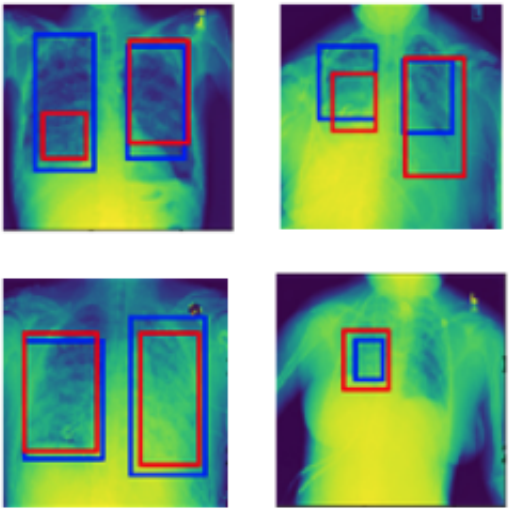


Fig. 7: Examples of predicted result by Resnet model. Predicted boxes are marked as **red** and ground truths are marked as **blue**.

also. Our YOLO model achieves 77.507% IoU at iteration 4932. This is really impressive and compare to our base

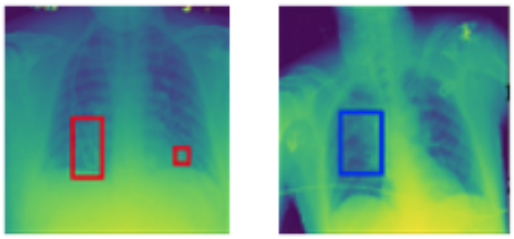


Fig. 8: Examples of *False Negative* (left) and *False Positive* (right) by Resnet model. Predicted boxes are marked as **red** and ground truths are marked as **blue**.

model (which is RESNET model), it outperforms that by 4% (approx.).

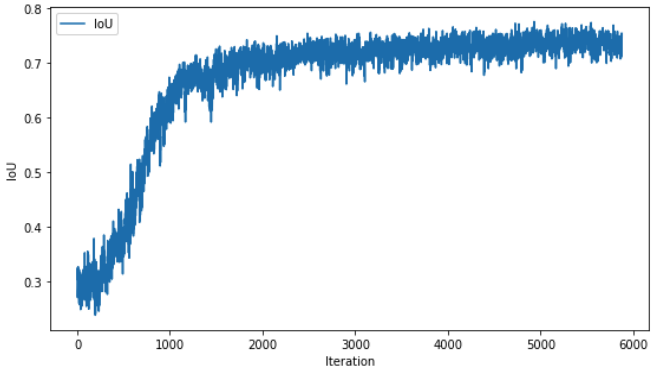


Fig. 9: Performance measurement of YOLOV3 Model (iteration v/s IoU)

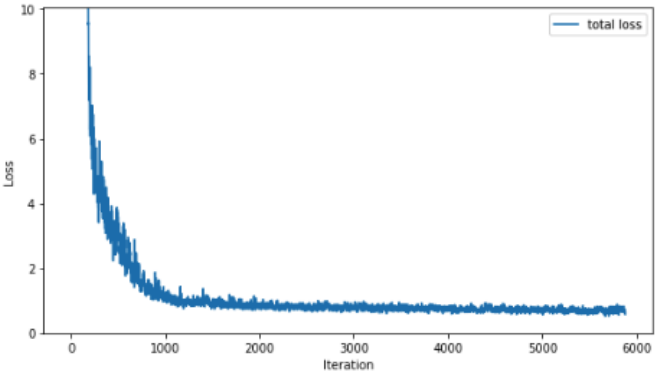


Fig. 10: Performance measurement YOLOV3 Model (iteration v/s Loss)

We have used our YOLO model on pneumonia affected patient's x-ray image. Figure 11 shows such a result where our models localize the location of pneumonia in a patient x-ray image.

VI. Future Work.

Although Resnet and YOLO model performed pretty well, there exists a lot of other model. In our view, performing bagging or boosting over these models could result in improved accuracy, precison and recall. Thus, in future we would like to try different models such as Faster R-CNN, SSD, R-FCN etc. on our dataset and evaluate performance of these models using different approaches.

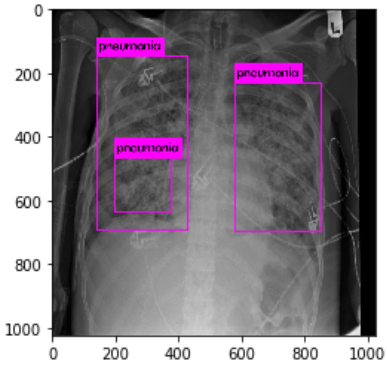


Fig. 11: Pneumonia Detection using YOLOV3 Model

A. Faster R-CNN

Region-based Convolutional Network or R-CNN begins with the region search and then performs the classification [15]. R-CNN is an alternative to exhaustive search in an image to capture object location. It initializes small regions in an image and merges them with a hierarchical grouping. Thus the final group is a box containing the entire image. The best R-CNNs models have achieved a 62.4% mAP score over the PASCAL VOC 2012 test dataset and a 31.4% mAP score over the 2013 ImageNet dataset.

B. Single-Shot Detector (SSD)

Single-Shot Detector (SSD) predicts all at once the bounding boxes and the class probabilities with a end-to-end CNN architecture [9]. The best SSD Over the test-dev dataset of the 2015 COCO challenge have had a score of 48.5% for an IoU = 0.5, 30.3% for an IoU = 0.75 and 31.5% for the official mAP metric.

C. Faster R-FCN

Region-based Fully Convolutional Network or R-FCN methodologies consist in detecting region proposals and recognize an object in each region [16]. R-FCN is a model with only convolutional layers allowing complete backpropagation for training and inference. The authors noticed that the R-FCN is 2.520 times faster than the Faster R-CNN counterpart.

VII. Contribution.

Three members of our team equally contributed on different parts of this project. Tamjid worked with data pre-processing, visualization and Resnet model. He also worked with implementing the IoU loss and performance metrics. Faysal worked with pre-processing the dataset for YOLO V3 and also trained the model using transfer learning. Fazlay Rabbi worked with the evaluation process with test data. He also studied the

existing works to find the suitable models for improving the performance of our task. Every member of the group also equally contributed in the writing and video making tasks.

VIII. Conclusion.

In this project, our main goal was to detect and localize pneumonia by investigating the chest x-ray image of a pneumonia affected patient. We achieved a convincing performance in terms of such operations. Our automatic CNN-based tool gets 73% IoU (built on RESNET) and 77.507% (built on YOLOV3) which is reliable for detecting as well as localize the position of pneumonia. We believe that our proposed technique will help doctors to detect pneumonia and thus, give proper treatment to the pneumonia affected patient.

References

[1] 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.

[2] Preliminary model architecture. available at: https://tinyurl.com/yc4eq9py.

[3] Rsna pneumonia detection dataset.

[4] Stats of the State of Virginia. 2016.

[5] Thomas Cherian, E Kim Mulholland, John B Carlin, Harald Ostensen, Ruhul Amin, Margaret de Campo, David Greenberg, Rosanna Lagos, Marilla Lucero, Shabir A Madhi, et al. Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. *Bulletin of the World Health Organization*, 83:353–359, 2005.

[6] H Dele Davies, Elaine E-l Wang, David Manson, Paul Babyn, and Bruce Shuckett. Reliability of the chest radiograph in the diagnosis of lower respiratory infections in young children. *The Pediatric infectious disease journal*, 15(7):600–604, 1996.

[7] T Franquet. Imaging of pneumonia: trends and algorithms. *European Respiratory Journal*, 18(1):196–208, 2001.

[8] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *CVPR*, volume 1, page 3, 2017.

[9] Kaiming He Jian Sun Jifeng Dai, Yi Li. R-fcn: Object detection via region-based fully convolutional networks. *arXiv:1605.06409*.

[10] Paras Lakhani and Baskaran Sundaram. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, 284(2):574–582, 2017.

[11] World Health Organization et al. Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children. 2001.

[12] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.

[13] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[14] Kang K Rui P. 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*.

[15] R. Girshick J. Sun S. Ren, K. He. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 2017. doi:10.1109/TPAMI.2016.2577031*.

[16] R. Girshick J. Sun S. Ren, K. He. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 2017. doi:10.1109/TPAMI.2016.2577031*.

[17] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, pages 3462–3471. IEEE, 2017.

[18] Li Yao, Eric Poblenz, Dmitry Dagunts, Ben Covington, Devon Bernard, and Kevin Lyman. Learning to diagnose from scratch by exploiting dependencies among labels. *arXiv preprint arXiv:1710.10501*, 2017.