Detecting Cancerous Nodules From Chest X-rays Using Deep Learning Techniques

CSE 499 report submitted in partial fulfillment of the requirements for the degree

of

Bachelor of Science in Computer Science and Engineering

Md. Tareq Mahmud	Nafis Iqbal	Shayam Imtiaz
1811174042	1811250642	1811710642

Done Under the Supervision of

Dr. Sifat Momen

Associate Professor

May 16, 2022

DECLARATION

Project Title	Detecting Cancerous Nodules From Chest X-rays Using Deep Learning Techniques
Author	*Md. Tareq Mahmud, Nafis Iqbal and Shayam Imtiaz*
Student IDs	1811174042, 1811250642 and 1811710642
Supervisor	Dr. Sifat Momen

This report is prepared as a requirement of the Capstone Design Project CSE499 A & B which is a two semester long senior design course. We declare that this **cse499** report entitled **Detecting Cancerous Nodules From Chest X-rays Using Deep Learning Techniques**, has not been accepted for any degree and is not concurrently submitted in candidature of any other degree. We would like to request you to accept this report as a partial fulfillment of Bachelor of Science degree under Electrical and Computer Engineering Department of North South University.

Md. Tareq Mahmud
1811174042

Department of Electrical Computer Engineering
North South University

Nafis Iqbal
1811250642

Department of Electrical Computer Engineering
North South University

Shayam Imtiaz
1811710642

Department of Electrical Computer Engineering
North South University

Date: May 16, 2022

Department of Electrical & Computer Engineering
North South University
Bashundhara, Dhaka - 1229, Bangladesh

APPROVAL

This is to certify that the cse499 report entitled **Detecting Cancerous Nodules From Chest X-rays Using Deep Learning Techniques**, submitted by **Md. Tareq Mahmud** (Student ID: 1811174042), **Nafis Iqbal** (Student ID: 1811520642) and **Shayam Imtiaz** (Student ID: 1811710642) are undergraduate students of the **Department of Electrical Computer Engineering**. This report partially fulfills the requirements for the degree of Bachelor of Science in Computer Science and Engineering on May 16, 2022, and has been accepted as satisfactory.

Dr. Sifat Momen
Associate Professor

Department of Electrical Computer Engineering
North South University

Dr. Mohammad Rezaul Bari
Associate Professor & Chairman

Department of Electrical Computer Engineering
North South University

Place: Dhaka, Bangladesh
Date: May 16, 2022

ACKNOWLEDGEMENTS

• We acknowledge the continuous guidance we received from our supervisor **Dr. Sifat Momen**.

Md. Tareq Mahmud, Nafis Iqbal, Shayam Imtiaz
North South University
May 16, 2022

Contents

List of Figures

Chapter 1

Introduction

Cancer is the leading cause of death worldwide, accounting for nearly 10 million deaths in 2020, or almost one in six deaths(37). The most common cancers are breast, lung, colon, rectum, and prostate cancers. Lung cancer is a type of cancer that develops in the tissues of the lungs, most commonly in the cells lining the airways(6). It is the leading cause of cancer death among men and the second leading cause of cancer death among women worldwide(32). More than 2.2 million new lung cancer cases were reported in 2020 alone(3). According to the National Cancer Institute of America (15), lung cancer is responsible for 51.6 percent of cancer-related deaths among males and 34.4 percent of cancer-related deaths among women in the United States. It demonstrates that lung cancer patients had the highest mortality rate.

1.1 Problem Statement

Pulmonary nodules are the first signs and symptoms of lung cancer. Pulmonary nodules are frequently discovered due to regular examination or CXR imaging for reasons unrelated to lung cancer. Nevertheless, long before clinical symptoms or any indicators appear(4), a chest radiograph (CXR) can reveal them, and chest radiography is the most common radiological exam globally. Although CXR can detect early-stage lung cancer, radiologists face a challenging task due to the various sizes of lung nodules(7). Furthermore, the density of pulmonary nodules varies significantly across the human body. As a result, CXR is crucial in correctly identifying nodules in the quest for early lung cancer diagnosis(4). Additionally, bones and organs frequently conceal lung nodules(4). Simultaneously, as the number of new cases grows, The danger of inaccuracy in detecting malignant nodules is a very complex undertaking. Thus CXR interpretation can quickly become complicated for radiologists(4). To detect the characteristics of lung nodules on CXRs, researchers devised a computer-aided detection (CADe) technique.

1.2 Motivation

Accurate early discovery of a malignant lung nodule can considerably improve the patient's chances of survival; yet, detection of early-stage lung cancer has been a serious concern for

the previous several decades. The ability to diagnose pulmonary nodules rapidly and reliably is critical in treating lung cancer(33). X-rays are often used to diagnose lung cancer(34). The motivation behind this work was to use X-rays to detect malignant nodules, which will help with better lung cancer diagnosis and treatment.

1.3 Goals and Objectives

Lung cancer is a high-risk disease that affects people worldwide, and lung nodules are the most common symptom of early lung cancer. Early lung nodule recognition decreased radiologists' effort and the risk of misdiagnosis and missed diagnoses. We propose accurately detecting cancerous lung nodules from chest X-rays using deep learning techniques discussed below.

1.4 Importance in countries like Bangladesh

Detecting lung nodules from the chest using medical imaging has been used for a long time. Primarily, lung nodules are detected on a chest using a medical imaging technique called computerized tomography (CT) scan(7). CT scan creates images of a cross-section of the patient's body using X-rays and computers(1). It takes photos of patients' bones, muscles, organs, and blood arteries in very thin slices, allowing healthcare providers to observe the patient's body in incredible detail(2). X-rays are one kind of radiation called electromagnetic waves(1). X-ray imaging creates pictures of the inside of one (1). As X-rays travel through the body, they are absorbed in different amounts by different tissues(1). Higher density tissue creates a whiter image than other tissues against the black background of the film. X-rays produce 2D images. X-rays are less detailed and less powerful than CT scans(5). Compared to CT scans, X-rays are more extensively utilized and less expensive. An X-ray is easily accessible and cheaper than a CT scan in countries such as Bangladesh, where medical services are not as good as in the United States or Canada.

1.5 Detection Approaches

Deep learning techniques have been widely used for object detection from images. Previously researchers have used deep learning to detect cancerous nodules from CT scans (14). Because nodule detection is done using CXR images in this study, it is far more difficult than nodule detection from CT scan slices. This work proposes using state of the art object detection models to detect cancerous nodules from CXR images. Initially, we tried to detect nodules using the Faster R-CNN model, which has shown to be one of the states of the art object detection model(24). Along with Faster-RCNN, Mask-RCNN (16), and You only look once (YOLO) (23) models are also used for detection. These object detection models generate bounding boxes around target objects, allowing them to be located in an image or video. These models typically use an encoder that takes an image and passes it through a series of blocks and

layers that learn to extract statistical information that can be used to find and classify objects. The encoder's outputs are then given to a decoder, which predicts each object's bounding boxes and labels.

1.6 Report Layout

The rest of the paper is arranged as follows: Part 2 discusses important papers, Part 3 discusses recommended techniques, and Section 4 discusses the experimental setup for all models. The evaluation metrics used to evaluate the works are discussed in Chapter 5. Chapter 6 contains the results, analysis, and debate. Model deployment is described in detail in Chapter 7. Finally, Chapter 8 concludes and discusses the future works.

Chapter 2

Literature Review

The study of detecting lung nodules from images has been ongoing for a long time to make lung nodule detection faster and more accurate. Their works (28) use faster R-CNN to detect lung nodules from CT images in their works. They show a comparative analysis between RCNN vs Fast R-CNN vs Faster R-CNN and how faster R-CNN performs better in detecting lung nodules. (38) It proposes a computer-aided automated pulmonary nodule detection method based on 2D CNN, which is composed of nodule candidate detection and false-positive reduction.(25) proposed a deep neural network-based automated lung cancer detection framework in their works where 3D lung CT image was analyzed for the presence of malignant nodules. (8) presents a deep learning framework to segment lung fields on clinical chest X-ray images for the first time. They propose an architecture which uses a simple seven-layer (six-layer for the fixed-scale) network that gives a state of the art result in segmenting lung fields from CXR images at that time. In their works, (10) proposes a CADe scheme that detects Pulmonary nodule on chest radiographs using a balanced convolutional neural network and classic candidate detection. This scheme includes segmentation of the lung parenchyma using the M-ASM algorithm, nodule enhancement techniques, watershed segmentation of lung nodules, synthetic training images, and transfer-learning methods for nodule classification. (19) develops three CNN Frameworks and an E-CNN model for detecting lung nodules in CXRs. They constructed three CNN models labelling them as CNN1, CNN2, and CNN3 and an ensemble model using the E-CNNs for the detection task. An adaptive loop filter was utilized to detect a tumour's candidate site in the method (36). They chose an ideal feature from 210 features to discriminate between normal and abnormal regions. Smooth loop filters were employed to smooth other tissues, and a wavelet-snake approach was used to extract and input geometry and data into the neural network for nodule classification by (17). (29) devising a multi-resolution massive training artificial neural network image processing technique that helps suppress rib and clavicle contrast in CXRs. Many studies have focused on decreasing false positives based on the classification premise.(39) suggested a way to reduce the FP's in the system by reducing the effects of rib features on lung nodule detection by employing the symmetry principle of the left and right lung regions. To increase the detection rate of nodules of various sizes (26) proposed a multiscale technique. The K-nearest neighbour classifier extracted features from the candidate regions using Gaussian filter banks of multiple sizes. In our works, we propose to train and build a deep learning-based model that will both detect and classify cancerous nodules from CXR using Faster R-CNN(24), YOLOv5(23) and EfficientDet (31) model. We will compare the models' performance and use the best.

2.1 Summary

Authors	Work	Results
Ying Su	uses faster R-CNN to detect lung nodule from CT images.	When compared to existing networks for detecting lung nodules, the optimized and upgraded approach suggested in their research enhances detection accuracy by more than 20%
Hongtao Xie	Computer-aided automated pulmonary nodule detection method based on 2D CNN	Extensive tests using the LUNA16 dataset revealed that the sensitivity of nodule candidate detection is 86.42 percent. At 1/8 and 1/4 FPs/ scan, the sensitivity for false positive reduction was 73.4 percent and 74.4 percent, respectively. This demonstrates that the proposed approach accurately detects lung nodules.
Jun Sang	A deep neural network based automated lung cancer detection	The proposed approach outperformed existing methods in terms of sensitivity (94 percent), specificity (90 percent), and area under the receiver operating curve (0.99) when tested on the LIDC-IDRI data set.
Mohammad R Arbabshiran	a deep learning model with an architecture that uses a simple seven-layer(six-layer for the fixed-scale) network	In comparison to manual segmentation, the suggested system achieved an IOU of 0.96 on the testing dataset. By a large margin, the proposed methodology outperforms state-of-the-art registration-based segmentation.
Sheng Chen	proposes a CADe scheme that detects Pulmonary nodule on chest radiographs using balanced convolutional neural network and classic candidate detection	The CADe technique demonstrated high performance of the published literatures (a sensitivity of 91.4% and 97.1%, respectively, with 2.0 false positives per image (FPs/image) and 5.0 FPs/image) for nodule cases using the Japanese Society of Radiological Technology database.
Chaofeng Li	ensembled 3 CNN models for the detecting lung nodules	In a five-fold cross-validation test, the suggested Ensembled-CNNs achieve a sensitivity of 94 percent and 84 percent, respectively, with an average of 5.0 false-positives (FPs) per picture and 2.0 FPs per image, which greatly beats the state of the art.
Jun Wei	Utilizied adaptive loop filter to detect a tumor's candidate site	The best feature set achieves a low number of false positives per image of 5.4 per image, with an 80 percent true positive detection rate.
Bilgin Keserci	used a wavelet-snake approach to extract and input geometry data into the neural network for nodule classification	The combined features resulted in a significant decrease in false positives, resulting in high performance in distinguishing between true and false positives.
Kenji Suzuki	Proposed a multi-resolution massive training artificial neural network image processing technique that helps suppress rib and clavicle contrast in CXRs	When this technique was applied to non-training chest radiographs, the visibility of nodules and lung vessels was significantly reduced but the visibility of ribs and clavicles was maintained.
Hiroyuki Yoshida	Proposed a way to reduce the FPs in the system by reducing the effects of rib features on lung nodule detection by employing the symmetry principle of the left and right lung regions	The method proved efficient in eliminating normal anatomic structures and minimizing the frequency of false detections in the CAD scheme for lung nodule detection when applied to clinical chest radiographs.
Arnold MR Schilham	Proposed a multi-scale technique to increase the detection rate of nodules of various size	Accepting an average of two false positives per image, the highest performance results in the identification of 51% of all nodules. This rises to 67 percent after four false positives. This is similar to the radiologists' previously reported detection rate of 70%.

Chapter 3

Proposed Methodology

This section discusses the methods and the algorithms used to accomplish the goal of the project discussed in this paper. The aim is to detect cancerous lung nodules from chest radiographs using an object detection model. For this, three states of the art object detection model are chosen. They are Faster RCNN with a Resnet50 backbone with a Resnet50 backbone (24), You only look once version 5(Yolov5) (23) and EfficientDet (31).

3.1 Faster RCNN with Resnet 50

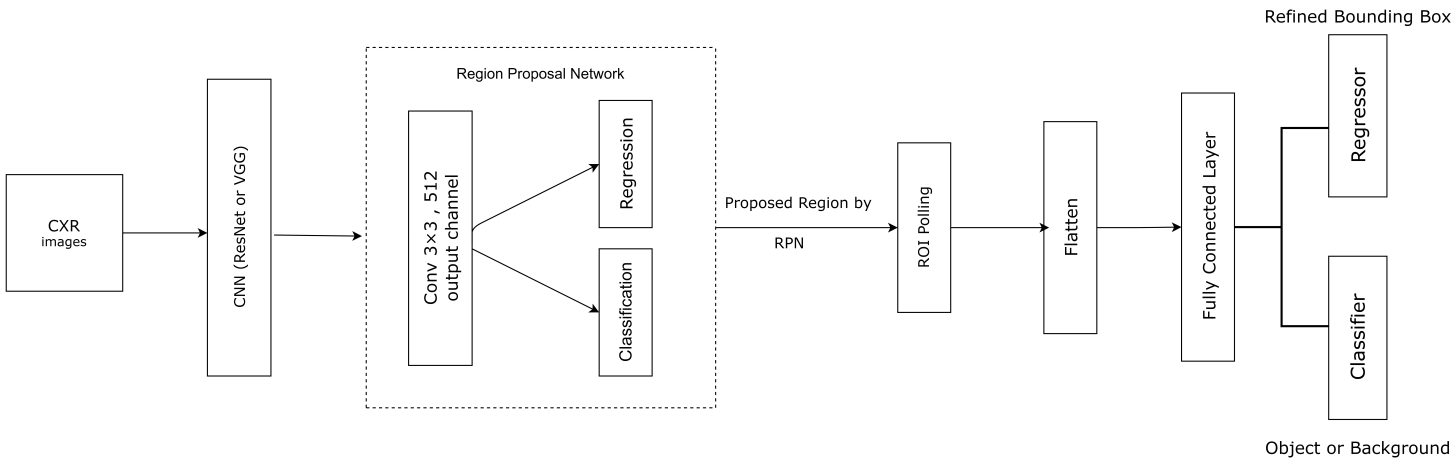


Figure 3.1: Faster RCNN with Resnet 50 model architecture

Fig 3.1 is a simple demonstration of the Faster-RCNN model used to detect cancerous lung nodules. The node21 dataset contains CXR images. Thus images should be converted to tensors at first to be loaded into GPU and then passed through a Convolutional Neural Network(CNN)(18).CNN such as VGG or ResNet can be chosen as the backbone feature extraction model. This algorithm implies a ResNet50 CNN model for feature extraction from CXR pictures, resulting in feature maps that are subsequently given to a Region proposal network (RPN). This RPN layer proposes areas where the object can be possibly found. As soon as the area where the cancerous lung nodule is located is detected, that area is labeled as a foreground class. The image area where the cancerous lung nodule is not present is labeled as a background class. The area labeled as foreground class will be passed to the next stage of the algorithm. This can be achieved by generating anchor boxes by the RPN layer. Anchor boxes are some set of predefined bounding boxes. Different sizes of anchor boxes are generated to detect objects of different sizes. In order to label an area as a foreground class, the RPN needs to calculate Intersection over Union(IOU). Here the images will have actual boxes to denote the cancerous nodule, and the RPN will generate some predicted boxes. The IOU is simply the overlapping area between the actual box and the predicted box. If the IOU value is greater than or equal to 0.5, the area will be labeled as a foreground class, otherwise a background class. The algorithm will learn from the particular area where the IOU is greater than 0.5. So for this purpose, the feature maps produced from the ResNet layer are simply passed to a CNN layer with a kernel size of 3x3 and output channel 512. The output of this convolutional layer is passed to two layers known as the classification layer and the regression layer. The classification class classifies the foreground class and background class, while the regression layer is used to generate the bounding boxes containing the area where the object is located. The areas or anchor boxes labeled as foreground class will move to the next layer known as the region of interest pooling layer(ROI). The output of the RPN layer will be in the form of feature maps of those anchor boxes labeled as foreground class. The RPN produces different sizes of feature maps which are reduced to the same size by the ROI layer. It takes the feature map for each region proposal to produce a fixed-size feature map of the region proposal using Max pooling. So the feature maps of all the region proposals will be of the same size. This is done because the region proposed containing the cancerous nodules needs to be flattened and passed to a fully connected layer. The ROI layer takes two inputs, the region of interest or the region proposal, and the feature maps produced by ResNet used to produce feature maps before passing to the RPN layer. The output of the ROI layer is flattened and then passed to fully connected layers. The output of these layers is then passed to a Classifier layer and a Regressor layer. The classifier gives the object class, and the Regressor will form and adjust the bounding boxes as the final output. Faster-RCNN uses a loss function called RPN loss which is given by:

$$\begin{aligned} L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) \\ +\lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}\ (t_i, t_i^*). \end{aligned} \tag{3.1}$$

In equation 3.1, The classification loss over two classes is the first term (There is an object or not). The second term is the regression loss of bounding boxes only when there is an object (i.e. p i* = 1).

.

3.2 YOLOv5

This section discusses the internal architecture of Yolov5. The algorithm of YOLO examines an image only once and finds all of the items there and their locations. Yolo outperformed many of the object detection algorithms such as RCNN, Fast-RCNN, Faster-RCNN, and sliding Window Object detection in the recent past. How Yolo works in order to detect cancerous lung nodules from chest radiographs is discussed below. In the given dataset, there are coordinates of the bounding boxes in the image and class probabilities such as 0 and 1, where 0 means no cancerous lung nodule in the image and 1 means cancerous lung nodule present in the image. In the given data set, the x and y coordinates determine the coordinate of the centre of the bounding box, and the width and height determine the width and height of the bounding box. In order to find the coordinate of the width and height, the values given in the dataset should be added to the value of x and y, which can be used to draw the bounding box in the image. Images containing no cancerous lung nodule have a value of 0 for the coordinates of x, y, height and width. These values of the coordinates are normalized before using it for training. A separate text file is created containing the normalized coordinates of the centre of the bounding boxes, the height and width, and the class probabilities. For each image, there might be multiple cancerous lung nodules, so for a specific image, all the coordinates of the cancerous lung nodule are given in one text file, and the text file is used as a label to train the model. The model used to achieve our goal is the large YOLOV5 model YOLOv5L. YOLOv5L contains a lot more parameters than other small YOLO models(23). YOLO uses a Convolutional neural network(CNN) to extract features from image data. It works by dividing photos into a grid, and each grid cell identifies a single object. For each grid cell, a vector is created that contains the grid cell coordinates and the class probability of whether a cancerous lung nodule is present in the grid. If a center coordinate is located in a particular grid cell, that grid cell is used to identify the cancerous lung nodule. In order to determine the accuracy of each prediction, the grid cells predict all of the bounding boxes and assign a confidence score to each of them. For each cancerous lung nodule, many bounding boxes can be created. Here the concept of IOU is used where IOU means intersection over the union. Here, the algorithm takes the original bounding box and the other bounding boxes and calculates the IOU to find the overlapping area. IOU is calculated by dividing the intersection area of two bound boxes by the union of areas of the two bound boxes, and so the value will be between 0 and 1. For all the bounding boxes generated for a cancerous lung nodule, we find the IOU value of all the bounding boxes with the original bounding box and keep only that bounding box with a maximum value of IOU. Getting these unique bounding boxes are known as Non-max suppression.

Following these procedures, the algorithm creates a single bounding box containing a single cancerous lung nodule.

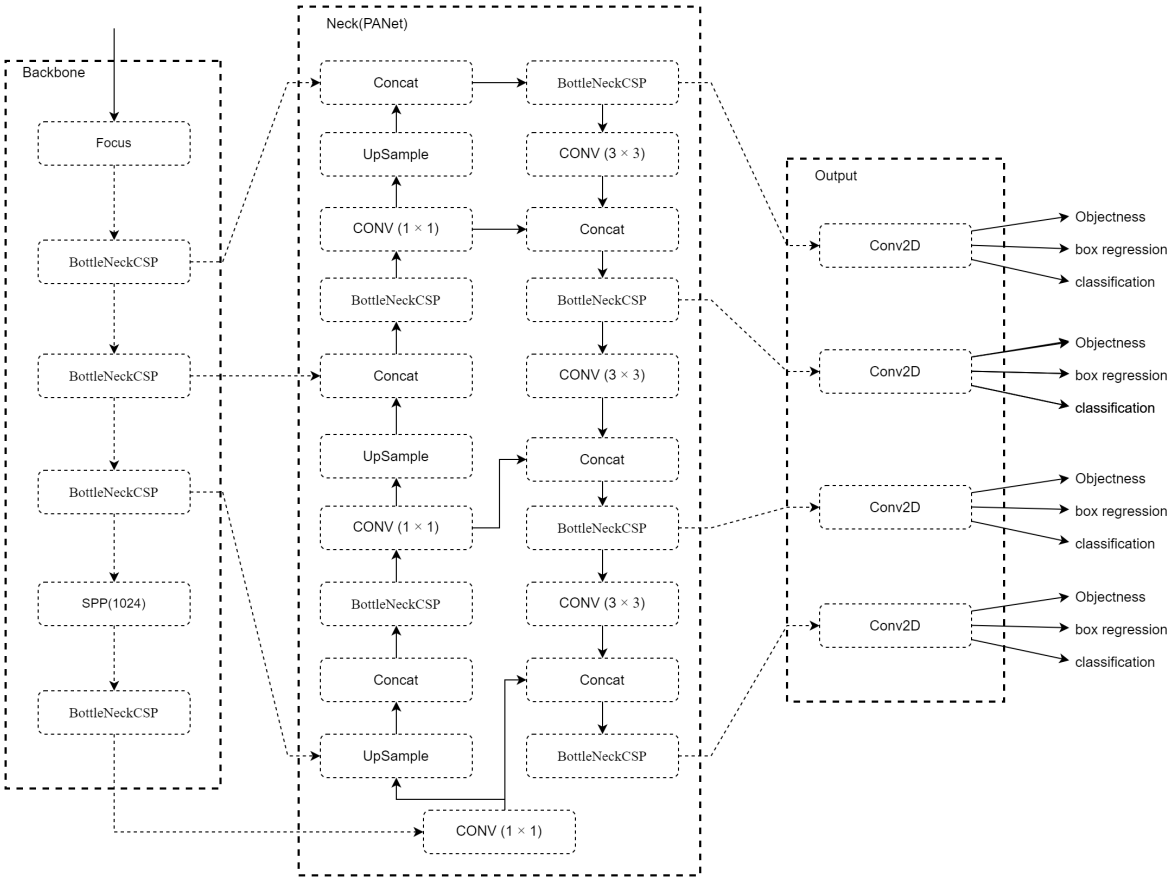


Figure 3.2: YOLOv5 model architecture

Figure 3.2 describes the network architecture of Yolov5. It is constructed in three parts. The first part shows the backbone network, which is CSPDarknet. The second part contains Neck: PANet and the final part contains the main Yolo Layer. Firstly input data goes through CSPDarknet for feature extraction. Then these data are fed to PANet for feature fusion. Lastly, the Yolo Layer gives detection results such as class, score, location, size etc. For loss computation of class probability and object score, YOLOv5 uses PyTorch's Binary Cross-Entropy with Logits Loss function. The Binary Cross entropy loss function is given by:

$$-(y\log(p)+(1-y)\log(1-p)) \tag{3.2}$$

In equation 3.2, log refers to the natural log, y is the binary indicator (0 or 1) if class label c is the correct classification for observation o and p is the predicted probability observation o is of class c.

3.3 EfficientDet

EfficientDet(31) is an object detection model developed by the Google research team in 2020. Despite the fact that various models, such as the AmoebaNet-based NASFPN(40) detector,

reaches a state of the art accuracy for object detection, it comes with a high cost of computation and latency. On the other hand, a large amount of work has been done to develop more efficient detector architectures, such as one-stage detectors like the Single Shot detector (22) and Yolo (23), which come with a sacrifice of accuracy. EfficientDet claims to solve this problem by using a systematic approach to finding an efficient multiscale feature fusion, model scaling and Multiscale Feature Representation. EfficientDet introduces an efficient bidirectional cross-scale connection and weighted feature fusion (BiFPN) to deal with the multiscale feature fusion problem.

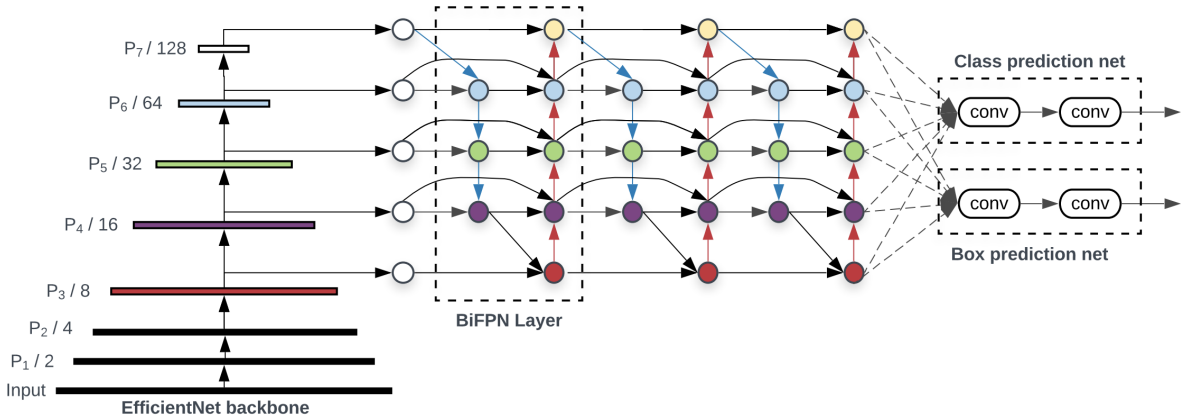


Figure 3.3: EfficientDet model Architecture

Figure 3.3 illustrates the EfficientDet architecture. Unlike FasterRCNN(24) and Retinanet(20) and many other models, which traditionally use the ResnetX (18,32,50). Architecture as the backbone network, EfficientDet uses EfficientNetB0-B6 (30) as the backbone network. The backbone network is followed by the introduced BiFPN network and the box/class prediction network. In our work, we use a PyTorch implementation of Efficient- Det. We convert our data to coco format and use the default image resolution to feed them into the network. We use the pre-trained D-2 version of weights that are trained on Imagenet for 80 classes. The model tries to learn and detect bounding boxes from the images and predict them when a new image is given as an input to the model. EfficientDet uses Focal Loss, which is given by:

$$FL = -\sum_{i=1}^{C=2} (1 - s_i)^{\gamma} t_i \log(s_i) \tag{3.3}$$

In equation 3.3, (1 s i) is a modulating factor, with the focusing parameter >= 0, to limit the influence of correctly identified samples in the loss. Focal Loss is the same as Binary Cross Entropy Loss when = 0.

Chapter 4

Experimental Setup

4.1 Dataset

2135 frontal chest x-ray images are used in training all 3 models. These CXR images are taken from the Node 21 dataset (28). Frontal chest radiographs with annotated bounding boxes around nodules make up the dataset(4). It comprises of 4882 frontal chest radiographs, with 1134 CXR images (1476 nodules) marked with bounding boxes around nodules, and the remaining 3748 images representing the negative class. The images in this set come from public databases that we can alter and share. They are gathered from JSRT (27), PadChest (9) ,Chestx-ray14 (35), Open-I(11).

4.2 Training environment and Machine configuration

On a Google Colab Pro environment with a p100 GPU and 16 GB of RAM, all three models are trained for 62 epochs. Pytorch support for Google Colab has been added, allowing GPU training. The google colab environment is run on the google chrome browser on a computer with a Core i5 architecture based CPU.

4.3 Training on Faster-RCNNResnet50

CXR images are trained using the Faster-RCNNResnet50 model, and annotations provided by the dataset are used. Annotations are given in a CSV file with the image name, class label, and the x,y and height and width coordinates of the nodules' positions. Images with no cancerous nodules contain the image names only, and everything else has a value of 0. 1922 images are used for training, and 213 images are used for validation. The input image shape is (1024x1024x3). We trained the model for 62 epochs with the following parameters. The initial learning rate is set to 0.005. Stochastic gradient descent(SGD) is used as the optimizer with a weight decay of 0.005 and momentum of 0.9. All the images files are in the native .mha format as collected from the node21 dataset.

4.4 Training on Yolov5l

To train CXR images on the YOLOv5 model the given annotations are converted into YOLO annotation format. 1922 images are used for training, and 213 images are used for validation. The input image shape is (1024x1024x3). Images have 1024 pixel height and width with a 3 channel input. For Yolov5l, we trained the model for 62 epochs with the following parameters. The initial learning rate is set to 0.01. Stochastic gradient descent(SGD) is used as the optimizer with a scaled weight decay of 0.0005 and momentum of 0.937. Image files for the model are converted from .mha to .png format as .png format is more compatible with existing model architecture and provided reliable detection. detection.

4.5 Training on EfficientDet

To train CXR images on the EfficientDet model the given annotations are first converted into COCO annotation format. 1922 images are used for training, and 213 images are used for validation. The input image shape is the same as used for the YOLO models. For EfficientDet, we trained the model for 62 epochs with the following parameters. The initial learning rate is set to 0.0001. Adam (Adam) is used as the optimizer with a scaled weight decay of 0.0005 and momentum of 0.9. All the image files are converted to .png format as it is more compatible with the existing model architecture.

Chapter 5

Evaluation Metrics

This section discusses the metrics used to evaluate the performance of the trained model. We used 4 standard evaluation metrics: Precision, recall, F1 score and mean Average Precision.

5.1 Precision

Precision is a popular evaluation metric used for classification task. Formula for Precision for a classification task is given as

$$Precision = \frac{TP}{TP+FP} \tag{5.1}$$

TP = True Positive FP = False Positive

True positive refers to an outcome where the model correctly predicts the positive class. False positive refers to an outcome where the model incorrectly predicts the positive class. Precision talks about how precise/accurate a model is out of the predicted positive, how many of them are actual positive.

5.2 Recall

Recall is another popular evaluation metric used in object detection. Formula for recall for a classification task is given as

$$Recall = \frac{TP}{TP+FN} \tag{5.2}$$

TP = True Positive FP = False Negative

True positive refers to an outcome where the model correctly predicts the positive class. False Negative refers to an outcome where the model incorrectly predicts the negative class. By designating a model as Positive, Recall calculates how many of the Actual Positives it captured as (True Positive).

5.3 F1 score

The F1 score is made up of two components: precision and recall. The F1 score aims to combine the precision and recall measurements into a single number. At the same time, the F1 score is created to operate well with data that is unbalanced.

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \tag{5.3}$$

5.4 mean Average Precision

Although Precision, Recall and F1 score are good metrics for evaluating a object detection model, recently another metric called mean Average Precision(mAP) is widely used too. Especially popular datasets like COCO(21) and PASCAL VOC(13) uses mAP to evaluate the performance of the model. To calculate mAP first we need to find the Average precision. Finding the area under the precision-recall curve is the general definition for the Average Precision (AP). The formula for average precision is given by

$$AP = \sum_{n} (R_n - R_{n-1}) P_n \tag{5.4}$$

Mean average precision is said to be the average precision for all classes. Formally,the mean average precision (mAP) of a set of queries is defined by

$$mAP = \frac{\sum_{k=1}^{Q} AveP(q)}{Q} \tag{5.5}$$

where Q is the number of queries in the set and AveP(q) is the average precision (AP) for a given query, q.

Chapter 6

Results and analysis

6.1 Results

All 3 models were trained for 62 epochs and evaluated using Precision, Recall,f1 and mean Average precision metrics. According to this (12), as the input for images containing no cancerous nodules was provided with no annotation, we consider all the scores for images with no cancerous nodules to be 1. The average score of both classes '0' and '1' is shown in the table below.

Model	Epochs	Precision	Recall	mAP	F1 score
FasterRCNN-Resnet50	62	0.75	0.18	0.17	0.29
EfficientDet	62	0.65	0.567	0.18	0.63
Yolov5l	62	0.83	0.82	0.81	0.82

Table 6.1: Validation Results of all the object detection models in detecting lung nodules from X-rays

6.2 Result Analysis and Comparison

Figure 6.1 illustrates the comparison of the results that were achieved after testing the 3 trained models. Between the 2 Yolo models, YOLOv5l was chosen for comparison as it gives more accurate detection than YOLOv5s on the training and validation set. Although all 3 models achieved decent Precision, but YOLOv5l seems to give a significant better recall and mean average precision score than the other two models. Both YOLO models predict the bounding boxes with high accuracy during validation. FasterRCNN-Resnet50 backbone predicts the bounding boxes with decent Precision but comes short in recall and mean Average Precision. EfficientDet gives a very poor mean average precision score but has an average recall score. The bounding boxes are likewise poorly predicted by EfficientDet.

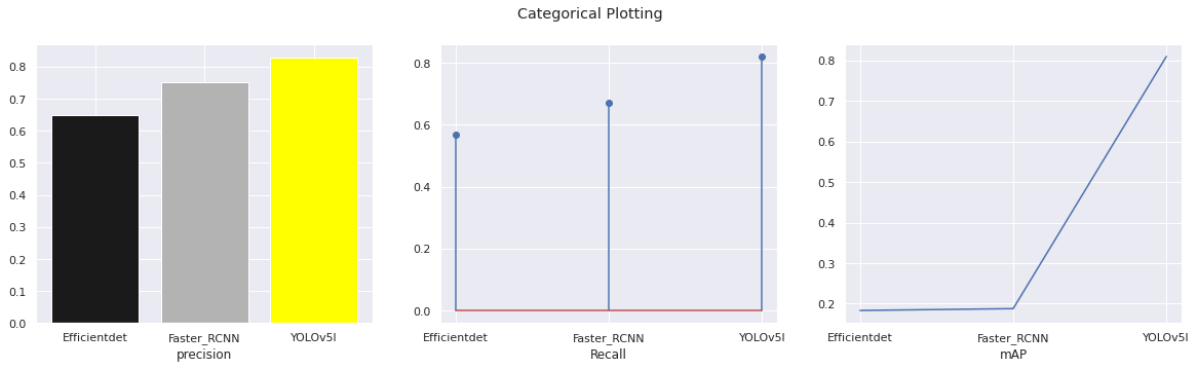


Figure 6.1: Result comparison of the 3 models

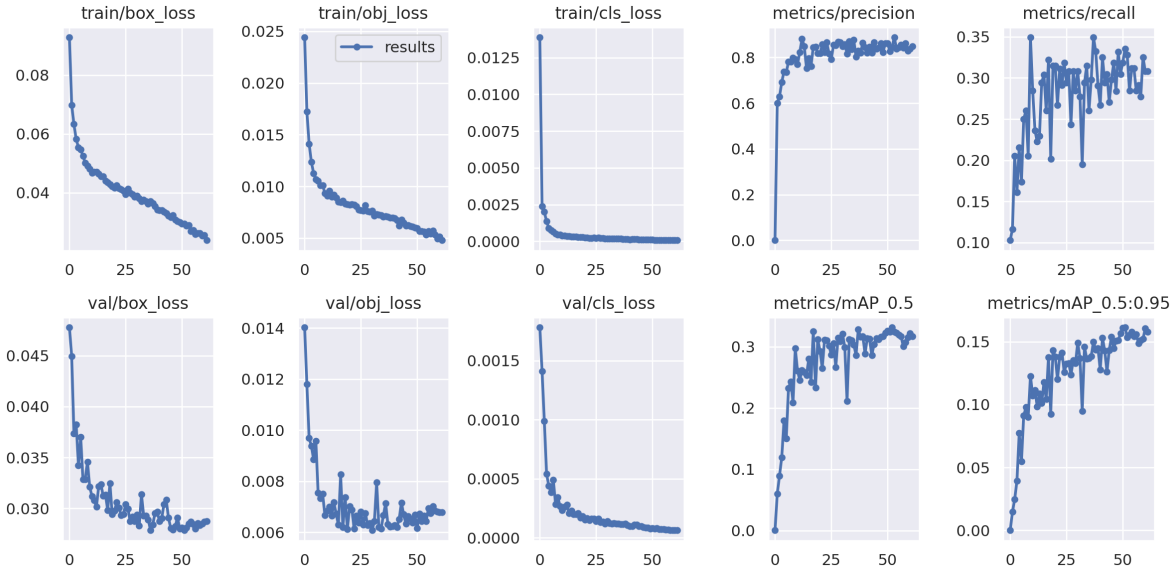
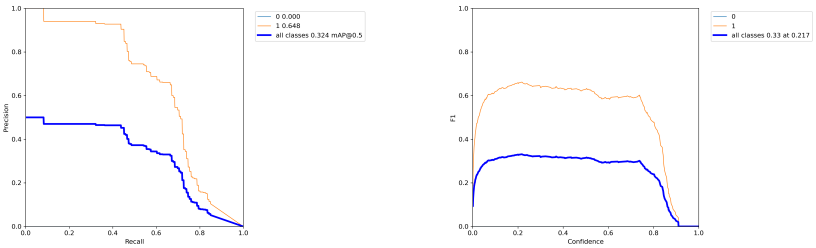


Figure 6.2: Results after training yolov5l for 62 epochs



(a) Precision vs Recall curve for Yolov5s

(b) F1 curve for Yolov5s

Figure 6.3: Precision Recall and F1 score for Yolov5l

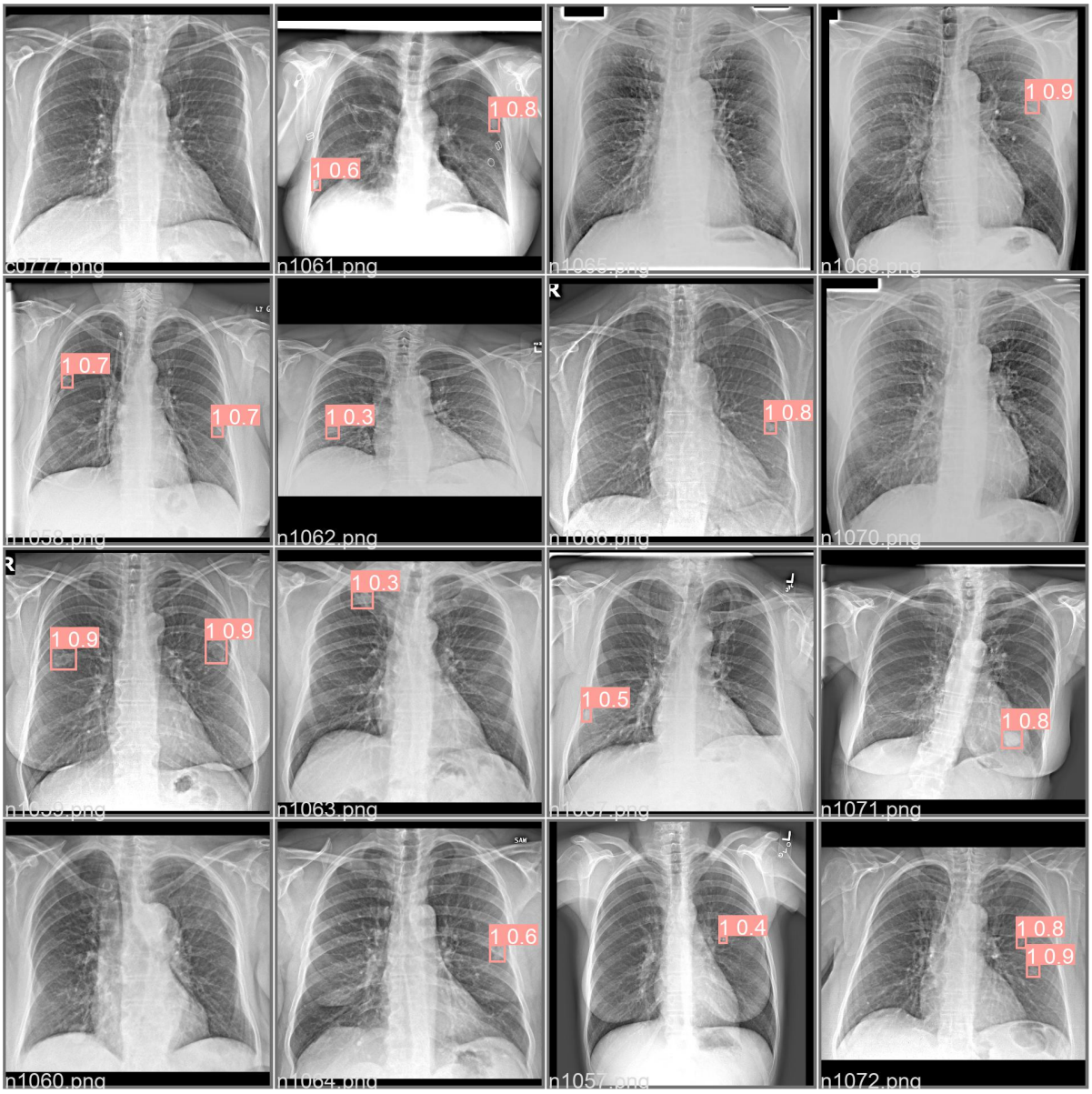


Figure 6.4: Yolov5l detection on validation set

6.3 Discussion

You only look once (YOLO) is known as one of the best object detection models and is proven to be the stand out model in this study too. YOLO can detect and predict the cancerous nodules for CXR images with high Precision. Also, the False-positive rate is very low, so YOLOV5 is reliable. YOLO is exceptionally good at object detection because of its complicated architecture and ongoing bug fixes with incremental enhancements. Although Faster-RCNN is one of the good early object detection models, it fails to detect small nodules from images in this experiment. The Resnet architecture Faster-RCNN uses cannot extract enough features from the training images as the nodules are oddly shaped and very small in size. Thus Faster-RCNN cannot detect the nodules from the images as well as YOLO. Despite performing very well for datasets like COCO(21) and Pascal VOC(13), EfficientDet can't seem to detect the nodules from CXR images in this study very well either. Training for higher epoch and choosing

a different EfficientNet backbone to extract features could result in better nodule detection by EfficientDet, but it is unlikely to surpass YOLO. One stage detector models like Single Shot Detector(SSD)(22) and Retinanet(20) could be two other models that can be used for nodule detection too. But they might produce similar results as EfficentDet as they already did in COCO(21) and Pascal VOC(13) datasets.

Chapter 7

Deploying the model

The best results for this experiment were achieved using the YOLOv5 models. Yolov5l is chosen for deployment because of its more reliable detection of lung nodules on the training and validation dataset. The best weights from the YOLOv5l model were chosen to deploy on a web server. Figure 7.1 illustrates the deploying process. Firstly, the dataset is trained using the yolov5l model. Then, the best parameters were chosen, which are automatically saved in a .pth file during training. For ease of use, a web app we created using Html, CSS and Javascript, which has a simple interface. In the backend Flask and Pytorch were used to load the .pth files and detect the cancerous nodules from images using the file parameters. Finally, the detection results were shown on a webpage.

Figure 7.1: Deploying the model on a web server

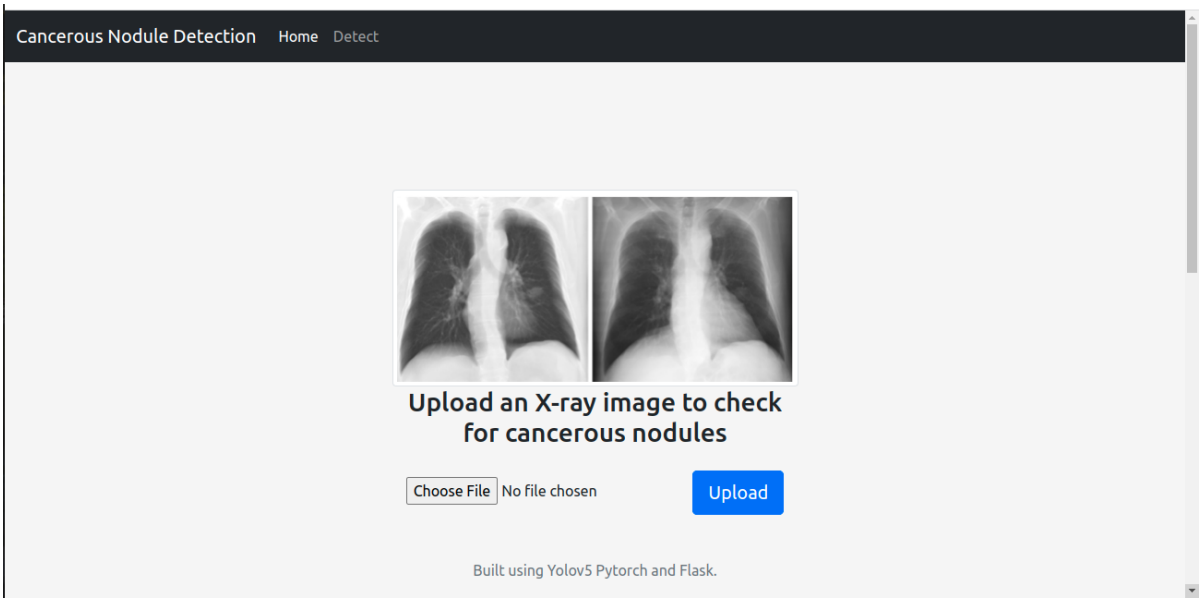


Figure 7.2: User interface of the webapp

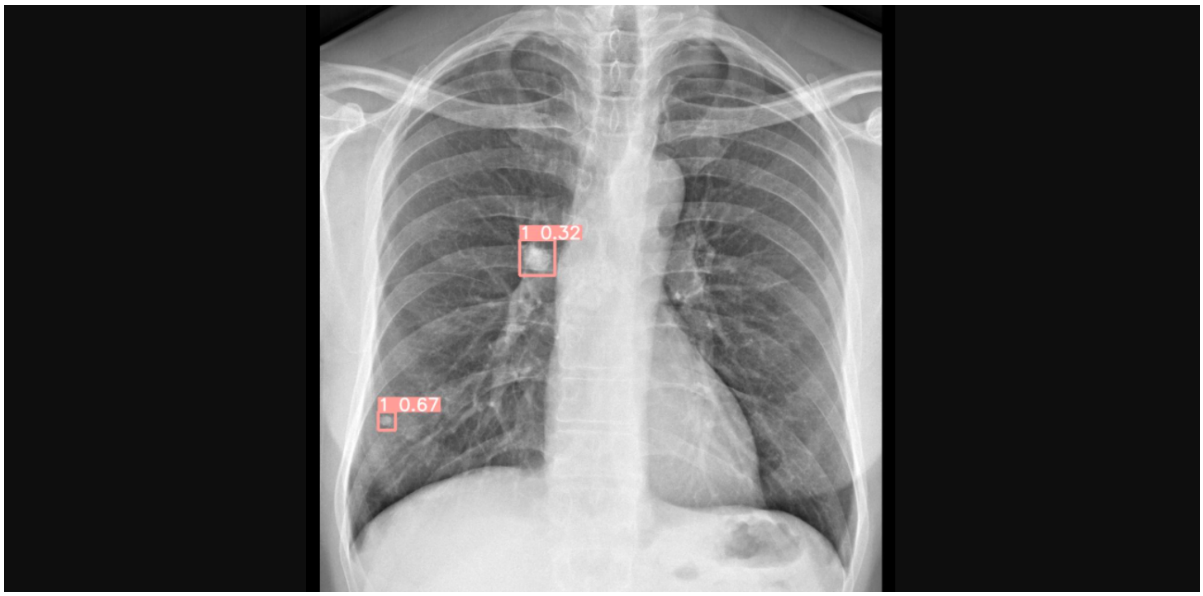


Figure 7.3: Showing detection results on webapp

Chapter 8

Conclusion and Future Work

Lung cancer is one of the world's most grievous cancers. It is the most common cancer-related death worldwide. However, if caught early enough, it can be curable. The cure rate for people with early-stage lung cancer can be as high as 80% to 90%. A lung nodule (or mass) is a tiny abnormal spot. The earliest stage of primary lung cancer can be cancerous nodules. It can be discovered during a chest CT scan or from chest radiographs. CT scans take longer and are more expensive than X-ray imaging tests. X-rays are cheaper than other forms of imaging testing due to the wide availability of technology and accessibility. So in this paper, we present the detection of cancerous lung nodules from Chest radiographs. For this purpose, three object detection techniques are used. We propose that these object detection models be improved and optimized. The highest Precision, recall and mean average Precision was achieved with two versions of YOLOv5. YOLO also predicts the bounding boxes with decent accuracy. Aside from YOLO, Faster-RCNN and EfficientDet were unable to detect the nodules in the chest x-rays accurately. In future, We want to use other object detection models like Retinanet(20) and Single Shot Detection (22) for detecting cancerous nodules and compare the results with the already used model in this study.

Bibliography

[1] How ct scan works? URL: https://www.nibib.nih.gov/science-education/science-topics/computed-tomography-ct. 1.4

[2] How x-rays works? URL: https://www.nibib.nih.gov/science-education/science-topics/x-rays. 1.4

[3] Lung cancer statistics | world cancer research fund international. URL: https://www.wcrf.org/cancer-trends/lung-cancer-statistics/. 1

[4] Node 21, grand challenge. URL: https://node21.grand-challenge.org/. 1.1, 4.1

[5] Ct scan versus mri versus x-ray, Oct 2021. URL: https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/ct-vs-mri-vs-xray. 1.4

[6] Lung cancer | lung cancer symptoms, Apr 2022. URL: https://medlineplus.gov/lungcancer.html. 1

[7] What tests can help detect lung cancer?, Feb 2022. URL: https://www.medicalnewstoday.com/articles/lung-cancer-detection. 1.1, 1.4

[8] Mohammad R Arbabshirani, Ahmed H Dallal, Chirag Agarwal, Aalpan Patel, and Gregory Moore. Accurate segmentation of lung fields on chest radiographs using deep convolutional networks. In *Medical Imaging 2017: Image Processing*, volume 10133, page 1013305. International Society for Optics and Photonics, 2017. 2

[9] Aurelia Bustos, Antonio Pertusa, Jose-Maria Salinas, and Maria de la Iglesia-Vayá. Padchest: A large chest x-ray image dataset with multi-label annotated reports. *Medical image analysis*, 66:101797, 2020. 4.1

[10] Sheng Chen, Yaqi Han, Jinqiu Lin, Xiangyu Zhao, and Ping Kong. Pulmonary nodule detection on chest radiographs using balanced convolutional neural network and classic candidate detection. *Artificial Intelligence in Medicine*, 107:101881, 2020. 2

[11] Dina Demner-Fushman, Sameer Antani, Matthew Simpson, and George R Thoma. Design and development of a multimodal biomedical information retrieval system. *Journal of Computing Science and Engineering*, 6(2):168–177, 2012. 4.1

[12] Dice-Group. Precision, recall and f1 measure · dice-group/gerbil wiki. URL: https://github.com/dice-group/gerbil/wiki/Precision,-Recall-and-F1-measure. 6.1

[13] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2):303–338, 2010. 5.4, 6.3

[14] Rasool Fakoor, Faisal Ladhak, Azade Nazi, and Manfred Huber. Using deep learning to enhance cancer diagnosis and classification. In *Proceedings of the international conference on machine learning*, volume 28, pages 3937–3949. ACM, New York, USA, 2013. 1.5

[15] Silke Gillessen, Gerhardt Attard, Tomasz M Beer, Himisha Beltran, Anders Bjartell, Alberto Bossi, Alberto Briganti, Rob G Bristow, Kim N Chi, Noel Clarke, et al. Management of patients with advanced prostate cancer: report of the advanced prostate cancer consensus conference 2019. *European urology*, 77:508–547, 2020. 1

[16] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017. 1.5

[17] Bilgin Keserci and Hiroyuki Yoshida. Computerized detection of pulmonary nodules in chest radiographs based on morphological features and wavelet snake model. *Medical Image Analysis*, 6(4):431–447, 2002. 2

[18] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012. 3.1

[19] Chaofeng Li, Guoce Zhu, Xiaojun Wu, and Yuanquan Wang. False-positive reduction on lung nodules detection in chest radiographs by ensemble of convolutional neural networks. *IEEE Access*, 6:16060–16067, 2018. 2

[20] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017. 3.3, 6.3, 8

[21] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014. 5.4, 6.3

[22] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer, 2016. 3.3, 6.3, 8

[23] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016. 1.5, 2, 3, 3.2, 3.3

[24] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28:91–99, 2015. 1.5, 2, 3, 3.3

[25] Jun Sang, Mohammad S Alam, Hong Xiang, et al. Automated detection and classification for early stage lung cancer on ct images using deep learning. In *Pattern Recognition and Tracking XXX*, volume 10995, page 109950S. International Society for Optics and Photonics, 2019. 2

[26] Arnold MR Schilham, Bram Van Ginneken, and Marco Loog. A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database. *Medical Image Analysis*, 10(2):247–258, 2006. 2

[27] Junji Shiraishi, Shigehiko Katsuragawa, Junpei Ikezoe, Tsuneo Matsumoto, Takeshi Kobayashi, Ken-ichi Komatsu, Mitate Matsui, Hiroshi Fujita, Yoshie Kodera, and Kunio Doi. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *American Journal of Roentgenology*, 174(1):71–74, 2000. 4.1

[28] Ying Su, Dan Li, and Xiaodong Chen. Lung nodule detection based on faster r-cnn framework. *Computer Methods and Programs in Biomedicine*, 200:105866, 2021. 2, 4.1

[29] Kenji Suzuki, Hiroyuki Abe, Heber MacMahon, and Kunio Doi. Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (mtann). *IEEE Transactions on medical imaging*, 25(4):406–416, 2006. 2

[30] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR, 2019. 3.3

[31] Mingxing Tan, Ruoming Pang, and Quoc V Le. Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10781–10790, 2020. 2, 3, 3.3

[32] Lindsey A. Torre, Rebecca L. Siegel, and Ahmedin Jemal. Lung cancer statistics. *Advances in Experimental Medicine and Biology*, 893:1–19, 2016. doi:10.1007/978-3-319-24223-1_1. 1

[33] Bram Van Ginneken, Samuel G Armato III, Bartjan de Hoop, Saskia van Amelsvoort-van de Vorst, Thomas Duindam, Meindert Niemeijer, Keelin Murphy, Arnold Schilham, Alessandra Retico, Maria Evelina Fantacci, et al. Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the anode09 study. *Medical image analysis*, 14(6):707–722, 2010. 1.2

[34] Shuhang Wang, Stefan Zimmermann, Kaushal Parikh, Aaron S. Mansfield, and Alex A. Adjei. Current diagnosis and management of small-cell lung cancer. *Mayo Clinic Proceedings*, 94(8):1599–1622, Aug 2019. URL: https://linkinghub.elsevier.com/retrieve/pii/S0025619619301260, doi:10.1016/j.mayocp.2019.01.034. 1.2

[35] 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 *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2097–2106, 2017. 4.1

[36] Jun Wei, Yoshihiro Hagihara, Akinobu Shimizu, and Hidefumi Kobatake. Optimal image feature set for detecting lung nodules on chest x-ray images. In *CARS 2002 computer assisted radiology and surgery*, pages 706–711. Springer, 2002. 2

[37] WHO. Cancer. URL: https://www.who.int/news-room/fact-sheets/detail/cancer. 1

[38] Hongtao Xie, Dongbao Yang, Nannan Sun, Zhineng Chen, and Yongdong Zhang. Automated pulmonary nodule detection in ct images using deep convolutional neural networks. *Pattern Recognition*, 85:109–119, 2019. 2

[39] Hiroyuki Yoshida. Local contralateral subtraction based on bilateral symmetry of lung for reduction of false positives in computerized detection of pulmonary nodules. *IEEE Transactions on Biomedical Engineering*, 51(5):778–789, 2004. 2

[40] Barret Zoph, Ekin D Cubuk, Golnaz Ghiasi, Tsung-Yi Lin, Jonathon Shlens, and Quoc V Le. Learning data augmentation strategies for object detection. In *European conference on computer vision*, pages 566–583. Springer, 2020. 3.3