



LUNG NODULE DETECTION
USING DEEP LEARNING MODEL

WINARTO WIJAYA

Supervised by Dr. V. Sivakumar

Capstone Project
Research Methodology in Computing and Engineering
MSC. IN ARTIFICIAL INTELLIGENCE

ASIA PACIFIC UNIVERSITY OF TECHNOLOGY & INNOVATION (APU)

SEPTEMBER 2021

DECLARATION OF SUPERVISOR(S)

“We hereby declare that We have read this thesis and in our
opinion this thesis is sufficient in terms of scope and quality for the
award of the degree of
Master of Science in Artificial Intelligence

Name of Supervisor: **DR. V. SIVAKUMAR**

Signature:

Date: 15 September 2021

Name of Supervisor (II) **TS. DR. SHANKAR DURAIKANNAN**

Signature:

Date: 15 September 2021

DECLARATION OF ORIGINALITY AND EXCLUSIVENESS

I declare that this thesis entitled
LUNG NODULE DETECTION USING DEEP LEARNING
is the result of my own research work except as cited in the references.

This thesis has not been accepted for any degree and it is not
concurrently submitted in candidature of any other degree.

Signature:



Name:

WINARTO WIJAYA

Date:

15 September 2021

To My Beloved Family

ACKNOWLEDGEMENT

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main capstone project supervisor, Dr. V. Sivakumar for encouragement, guidance, critics and friendship. I am also very thankful to Prof. Dr. Mandava Rajeswari for their guidance, advice, and motivation. Without their continued support and interest, this thesis would not have been the same as presented here.

My fellow postgraduate students should also be recognised for their support. My sincere appreciation also extends to all my colleagues and others who have aided at various occasions. Their views and tips are useful indeed. Unfortunately, it is not possible to list all of them in this limited space. I am grateful to all my family members.

ABSTRACT

Lung cancer is one of the leading causes of death in the world, early detection of lung cancer has proven to increase mortality rate. The radiologist uses Computer-Aided Diagnostic (CAD) system to recognize lung nodules in lung CT images. The interest in lung nodule detection can be in terms of the size, texture or location wise of nodule which can be a valuable information for doctors to initiate proper treatment. The current problem of using a CAD is the time-consuming process of evaluation with manual screening by a physician of the CT medical images. The evaluation process may deviate one's subjective consciousness which may lead to misdiagnosis and inefficiency. Therefore, the need for efficient lung nodule detection is in the demand. This paper intends to develop a prototype level of lung nodule detection using deep learning detection model. 3D deep learning with dual path network with Squeeze and Excitation was adapted to the Region Proposal Network (RPN) from past studies and influenced by the new trend of attention mechanism. The study is motivated by developing a more efficient deep learning model that employs 2 strategies: lung nodule localization and False-Positive Reduction components. The increasing of feature representational power using squeeze and excitation layer does perform on par and even better compared to the baseline model. However, due to restrained time-resource and computation power, both localization model and the classification results in underperform. Based on the approach, the result from classification model obtained a decent performance about 77% accurate which caused by restricted time-resource and computation power.

TABLE OF CONTENTS

DECLARATION OF SUPERVISORS	i
DECLARATION OF ORIGINALITY AND EXCLUSIVENESS	ii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
LIST OF FIGURES	vi
LIST OF TABLES	ix
LIST OF ABBREVIATIONS	xi
CHAPTER 1: INTRODUCTION TO THE STUDY	1
1.1 Introduction	1
1.2 Research Problem	2
1.3 Aim and Objectives	3
1.4 Justification for the Research	3
1.5 Feasibility, Scope & Limitation	4
1.6 Contribution of the study	5
CHAPTER 2: LITERATURE REVIEW	6
2.1 Introduction	6
2.1.1 Past Theories	6
2.1.2 Theoretical Framework	7
2.1.3 Techniques	7
2.2 Conclusion	11
CHAPTER 3: RESEARCH METHODOLOGY	12
3.1 Research Design	12
	vi

3.2	Investigation on Model Architecture, Deployment Platform and Dataset	12
3.2.1	Selection on Model Architecture	13
3.2.2	Selection on Deployment Platform	16
3.2.3	Selection Dataset	17
3.3	Model Architecture Derivation	18
3.3.1	3D Dual-Path Network	18
3.3.2	3D Res-Net18SE	22
3.4	Proposed Methodology	22
3.5	Summary	24
CHAPTER 4: FINAL DESIGN & SYSTEM IMPLEMENTATION		25
4.1	Introduction	25
4.2	System Implementation	25
4.2.1	Data Pre-processing	26
4.2.2	Modelling	29
4.2.3	Training and Testing	34
4.2.4	Results and Evaluation	35
4.3	Results	36
4.3.1	3D DPN and 3DPN_SE	36
4.3.2	3 D CNN Classifier	40
4.4	Project Timeline	42
4.5	Summary	43
CHAPTER 5: CONCLUSION AND RECOMMENDATIONS		45
5.1	Conclusion relating to the Research Objectives	45
5.2	Limitations	45
5.3	Recommendations and Suggestions for further research	46

5.3.1	Decoupled False Positive Reduction	46
5.3.2	Attention module from Transformer model	47
5.4	Summary	47
REFERENCES		48
APPENDICES		52
Appendix A:	Ethics form – Fast track	53
Appendix B:	Disclaimer form	60
Appendix C:	Log sheet	62
Appendix D:	Project Gantt Chart	66

LIST OF FIGURES

Fig 2.1: Lung nodule detection without localization framework (Zhang et al., 2020).	6
Fig 2.2: Lung nodule detection with localization – Faster RCNN (EL-Bana, Al-Kabbany and Sharkas, 2020).	6
Fig 2.3: Theoretical Framework Diagram (Ma et al., 2019).	7
Fig 2.4: Dual Path Convolutional Network Structure (Essaf, Li, Sakho, and Gadosey, 2020).	9
Fig 2.5: 3D Fusion Model Structure.	9
Fig 3.1: Lung Nodule Detection Pipeline Diagram.	12
Fig 3.2: The Faster RCNN on Region Proposal Network (Ren et al., 2015).	13
Fig 3.3: Region Proposal Network (RPN) (Ren et al., 2016).	14
Fig 3.4: YOLO architecture (Redmon et al., 2015)	14
Fig 3.5: YOLO Detection System. (Redmon et al., 2016)	15
Fig 3.6: Architecture Comparison of different networks and formation of DPN (Chen, Y.,et. Al, 2017).	18
Fig 3.7: Illustration of dual path connection (Zhu, W., et al., 2017).	19
Fig 3.8: Squeeze-and Excitation block (Hu, J., Shen, L. and Sun, G., 2018).	20
Figure 3.9: The architecture of 3D Res-Net 18 SE structure (Li, Y. and Fan, Y., 2020).	21
Fig 3.10: Overview of the framework.	23
Fig 3.11: Proposed Dual Path SE block connection.	24
Fig 4.1: System Implementation Routine.	25
Fig 4.2: 2D example of spacing and origin (Coordinate systems - Slicer Wiki, 2021).	26
Fig 4.3: Slice and Increment illustration (What is the difference between slice thickness and slice increment?, 2021).	27
Fig 4.4: Reconstruction of the image planes (Coordinate systems - Slicer Wiki, 2021).	27
Fig 4.5: Flowchart of pre-processing routine.	29
Fig 4.6: Proposed Model Network Architecture and Original DPN Model based on (Zhu, W., et al.,2017).	31
Fig 4.7: CNNT5 Classifier model for False Positive Reduction.	33
Fig 4.8 Training and Testing Routine.	34
	ix

Fig 4.9: Regression Loss comparison between 3 models.	36
Fig 4.10: Height Coefficient Prediction Loss.	37
Fig 4.11: Width Coefficient Prediction Loss.	38
Fig 4.12: Top Left Corner – X Coordinate Coefficient Prediction Loss.	38
Fig 4.13: Top Left Corner – Y Coordinate Coefficient Prediction Loss.	39
Fig 4.14: Confusion Matrix.	40
Fig 4.15: Confusion Matrix in percentage.	40
Fig 4.16: CNNT5 Model Training and Validation Loss Performance.	41
Fig 4.17 Gantt chart research plan 1 of 2.	42
Fig 4.18 Gantt chart research plan 2 of 2.	43

LIST OF TABLES

Table 3.1: Comparison table between Faster RCNN and YOLO.	15
Table 3.2: Comparison table between AWS Sagemaker and Google Colab.	16
Table 3.3: Comparison table of Public Lung Nodule Dataset.	17
Table 3.4: Annotation description for fig 3.7.	19
Table 3.5: annotations description for Squeeze and Excitation block.	21
Table 4.1: Hounsfield Density Range.	28
Table 4.2: Proposed Model Summary for lung nodule localization.	30
Table 4.3: CNNT5 Model Adaptation from (Dobko, M., 2019).	33
Table 4.4: Six-fold cross-validation results.	41

LIST OF ABBREVIATIONS

2D	2 Dimensions
3D	3 Dimensions
AI	Artificial Intelligence
AMST	Adaptive Morphology-based Segmentation Technique
ASE	Adaptive Structuring Element
AWS	Amazon Web Service
CADs	Computer Aided-Diagnostic system
CIP	Central Intensity Pooling
CNN	Convolutional Neural Network
CT	Computed Tomography
DPN	Dual Path Network
DSSD	Deconvolution Single Shot Detector
DT	Decision Tree
FCM	Fuzzy C-Mean
FLDA	Fisher Linear Discriminant Analysis
FPR	False-Positive Reduction
GAN	Generative Adversarial Network
GGO	Ground-Glass Opacity
GPU	Graphics Processing Unit
k-NN	k-Nearest Neighbor
LDP	Local Difference Pattern
LIDC-IDRI	Lung Image Database Consortium
LUNA 16	Lung Nodule Analysis 2016
NMS	Non-Maximum Suppression
OS	Operating System
PBB	Probability Bounding Box
PSO	Particle Swarm Optimization

RAM	Read Access Memory
R-CNN	Region based - Convolutional Neural Network
Res-Net	Residual Network
RF	Random Forest
RPN	Region Proposal Network
ROI	Region of Interest
SE	Squeeze Excitation
SGD	Singapore Dollar
SNN	Siamese Neural Network
SVM	Support Vector Machine
YOLO	You Only Look Once

CHAPTER 1

INTRODUCTION TO THE STUDY

1.1 Introduction

According to the world cancer research fund, lung cancer is reported to be the most commonly occurring cancer in men and third on women with 2 million cases in 2018 worldwide (Lung cancer statistics, 2018). It has been proven that the early detection of lung nodules cancer improves the survivability rate for at least five years given a proper prognosis treatment plan (Aberle et al,2011). Pulmonary nodules or coin lesions are detectable using Computed Tomography (CT) screening. The current medical practice on evaluating CT scans uses Computer-Aided Diagnosis (CAD) system, whereby the system will provide structure reports such as volume, localization and suggestions for further diagnosis and treatment. The CAD system use computer vision techniques to extract important features from each scan to provide doctors/experts insights.

CAD system for lung nodule detection and classification has been extensively studied in the past where earlier model approaches uses semi-automatic (handcraft image processing techniques) to fully automated (deep-learning techniques). To some extent semi-automated techniques would suffer from interpatient differences which require cumbersome calibration, essentially demonstrate the weak robustness and limitation. The current most advanced lung nodules detection in CT images is based on deep learning model whereby the system can perform diagnosis for lung cancer detection with state-of-the-art performance. Over the years, deep learning-based model had picked up the trend of research from 2D to 3D image approaches.

In general, deep learning-based methods can approach lung nodule detection in different ways and can be categorized into 2 distinct approaches: (1) Lung nodule detection with localization – multi-task problem, (2) Lung nodule detection without localization – classification task problem. The former problem consists of 2 components namely, localization and False Positive Reduction (FPR) typically involves detector model such as region-based lung detection model – Faster R-CNN, others may use YOLO with CNN classifier for FPR, while the latter problem consist of 1 component – Classification by itself, generally starts off with lung segmentation to obtain lung nodules, in most research U-Net model was used and passes the segmented nodule to CNN classifier for benign or malignant classification (Chen et al., 2021) (Cao et al., 2020) (EL-Bana,

Al-Kabbany and Sharkas, 2020) (Safarov and Whangbo, 2021). In this work, lung nodule detection with localization will be the focus of the research.

The lung nodule detection with localization and FPR components can be described as an end-to-end approach in which that the input starts with 2D/3D CT scans which produces output of bounding box detection and classification of the nodules. Deep learning detection model such as Faster R-CNN appears in most pulmonary nodule detection research whereby the Region Proposal Network was adopted to medical images to generate region proposal. Research by (Zhu, W., et al., 2017) built a faster R-CNN with a dual-path network harnessing Residual Network (Res-Net) and Dense-Net capability in exploring new features and feature refinement for lung nodule detection.

Recent advanced models have been moving towards the implementation of attention module which originates from a model called transformer (Vaswani, et al., 2017), initially proposed to tackle Natural Language Processing problem and was brought over and applied in computer vision domain. Evidently, a 3D RPN network was built with a ResNet-18 backbone with the additional spatial attention module Squeeze and Excitation module called Deep SEED 3D model by (Li, Y. and Fan, Y., 2020,). Challenges in the development of the deep learning lung nodule systems were usually found in the availability of datasets referring to data privacy concerns which leads to an unbalance dataset i.e., a large number of negative samples compared to positive samples (Banik and Bhattacharjee, 2021).

1.2 Research Problem

The research problem in this work essentially is to further asses on how to improve the lung nodule detection system to be more robust and efficient, in addition to view the gap from the previous study. The previous work on lung nodule detection provides sufficient resource to perform the research study in terms of enriching features and focus feature context – local and global context, etc.

The multi-modality approaches show a few pros in lung nodule detection; the performance yields considerably better than single modality approach, as it promotes the robustness of prediction model with the rich features extraction. The cons would be slow in adaptation of the system in terms of preparation and pre-processing operations, in addition, large computation prowess is required. For Generative Adversarial Network (GAN) based network approaches, benefits from the generation of new samples in prevention of overfitting and yet still suffer from

complicated implementation and high computation requirement. The Dual-Path Network (DPN) approach enable feature re-usage and new feature exploration using Res-Net and Dense-Net, this approach aligns with the desired goal which is to enrich feature map on feature extraction. This raises the questions of which algorithm is a better fit to use for lung nodule detection system.

In the current approach of lung nodule detection, the work from (Sun, J. et al., 2020) performs a lung nodule detection using Convolutional Neural Network (CNN) with shape attentive U-Net architecture, similarly the work by (Lin, Zheng and Hu, 2020) proposed a 3D CNN with shortcut connections for lung nodule classification. Another work from (EL-Bana, Al-Kabbany and Sharkas, 2020) uses a two-staged framework for automated lung nodule classification and detection, the work applies a lung nodule segmentation using deeplab-V3 and Faster RCNN with InceptionV2 with an accuracy of 97%. Based on the literature, self-attention mechanism which addresses the selective context on extracted feature were demonstrated. Hence another research gap of the inclusion of global context feature map can be study and explored.

1.3 Aim and Objectives

Aims: To develop an automated detection model for lung nodules on CT images

Objectives:

1. To evaluate different deep learning strategy for implementation on lung nodule detection.
2. To design and develop a deep learning framework for detection of lung nodule malignancy.
3. To appraise the performance by benchmarking with the adopted model on the same public dataset.

1.4 Justification for the Research

The significance of the research will help radiologist/physician to ease and streamline interpretation which promotes efficiency in evaluating lung nodule in CT image modality. The study intends to improve the existing system to the latest technology in the development of CAD system for lung nodule detection especially in localization route and false-positive reduction stages. The impact on advancing the existing system may in turn provide accessible framework for the future work and perhaps act as a supporting argument or claims on other report with the relevant challenges faces in this work.

Different approach of algorithm plays an important role in the system as it provides a different level of performance and robustness in handling interpatient differences computer vision application especially in medical imaging where accuracy is at stake. The common types of model architecture will be the 3D Faster RCNN and YOLO object detection model whereby implementation can be differed in the backbone construction e.g., Res-Net, Dense-Net , Inception etc. The existing lung nodule detection system are built with Dual Path Network (DPN), ResNet-18 SE whereby each backbone was implemented on 3D Faster RCNN architecture. The recent advancement in the transformer model for machine translation has started it`s wave by invading other domain with the self-attention mechanism adopted from transformer model and adapted to computer vision domain.

In a nutshell, the lung nodule detection model using deep learning can be further enhanced by applying the recent advancement on self-attention with the concept of increasing representational feature map in any feature-based operations such as convolution. The existence of feature map plays an important role in computer vision whereby most components in lung nodule detection essentially involves feature map. Therefore, lung nodule detection system can be further implemented with better feature representation.

1.5 Feasibility, Scope & Limitation

The development of the proposed system is limited to a medical image processing on lung CT-image modality whereby the designs of the systems will serve as simulation / prototype model which are not ready-use manner for public commercial usage. The proposed system will use a public dataset lung image due to the privacy concern and the availability of the data. Another reason of the use on lung public dataset will be the consistency of the annotated evaluation, in which the LUNA 16 dataset had been evaluated by at least 4 expert radiologists (Setio et al., 2017).

In addition to limitation factor, the proposed system will be trained under a low sets epochs with and half the size of the original size dataset. This limitation is more of infrastructure computation prowess where the model is trained using Amazon Web Service (AWS) Sagemaker paid notebook. The reason behind training under a reduced datasets is due to the restrained timely-resources and costs of training. Therefore, comparison of models will be sophisticated and subject to the current investigation. Moreover, the output of proposed lung nodule detection will be a classification of benign or malignant lung nodule, which will be the output from FPR model.

1.6 Contribution of the study

The contribution of the work directed towards a more enhanced feature representation on automated lung nodule detection development using deep learning for the adoption in clinician practice to improve health care for lung cancer. The consequence of improved health care systems may in turn provides more available and accessible options for the community.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In this section, an in-depth review of past theory relating to the topic discussed and theoretical framework was drawn to show the focus area of the literature review. An overview of lung nodule detection based on academic journals or conference papers will be discussed in following sections of literature review. The techniques section will discuss existing lung nodule detection model with the method of implementation, including the types of architecture for the nodule localization and detection.

2.1.1 Past Theories

In the past theories, deep learning approaches consists mostly of two strategies, 1) lung nodule detection with and without localization feature which different in the number of stages, to illustrate the past theories/ framework observe the figure 2.1 and 2.2.

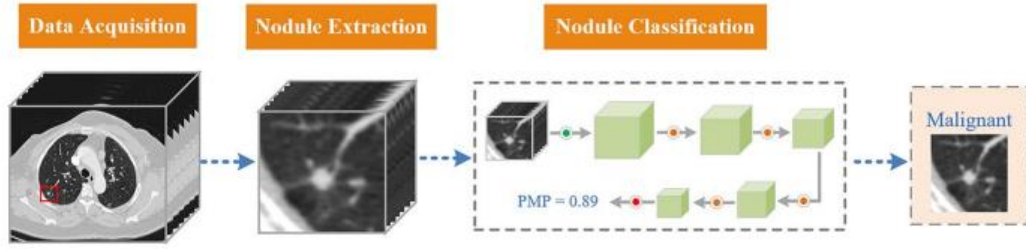


Fig 2.1: Lung nodule detection without localization framework (Zhang et al., 2020).

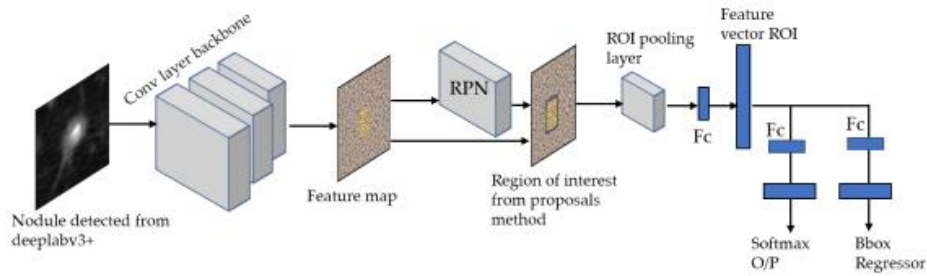


Fig 2.2: Lung nodule detection with localization – Faster RCNN (EL-Bana, Al-Kabbany and Sharkas, 2020).

Figure 2.1 and 2.2 shows the lung nodule detection approaches framework in most research papers for deep learning-based lung nodule detection. The approach differs in the ability to provide bounding box which also known as localization feature by implementing object detection framework namely, Faster RCNN, YOLO, and other detection model. On the other hand, the approach without localization feature performs solely using CNN classifier model and passes through FCN for classification output.

2.1.2 Theoretical Framework

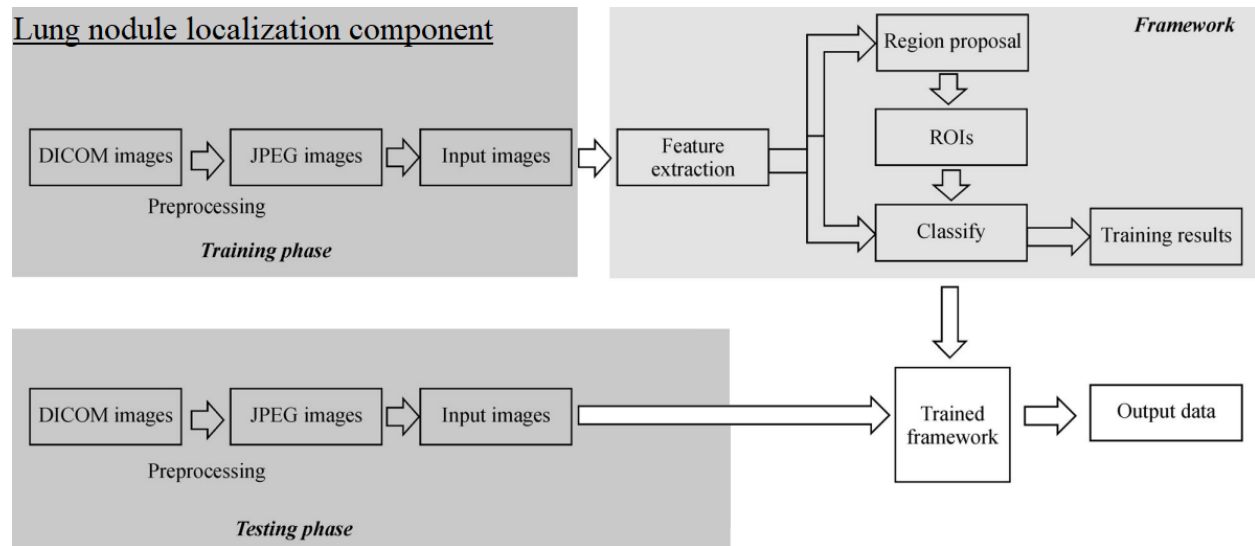


Fig 2.3: Theoretical Framework Diagram (Ma et al., 2019).

Figure 2.3 shows the theoretical framework diagram for the research from a survey paper, to direct the relevant research background in the literature review section. In the past research, exists sub-parts/ components that is shown in the theoretical framework which may be divided into individual component.

2.1.3 Techniques

Basic image processing technique approaches namely image filtering, binary image operations are considered as a non-deep learning approach in lung nodule detection and classification. Classification of lung nodules includes the use of machine learning techniques such as Support Vector Machine (SVM), Random Forest (RF), Fisher Linear Discriminant Analysis (FLDA), Decision Tree (DT), k-NN, and so on.

The work by (Kumar and Latte, 2017) proposed the improved OTSU's algorithm with the Particle Swarm Optimization (PSO) in the search for segmenting pulmonary parenchyma in CT lung images based on fractional calculus and natural selection. (Halder et al., 2020) developed an adaptive morphology-based segmentation technique (AMST) improves the segmentation of the lung nodule region by employing adaptive structuring element (ASE) and passes the output to the machine learning classifier SVM for nodule detection. In recent advancement of nodule detection, there has been growing popularity with regards to harnessing 3D space feature, in the work of (Gong et al., 2016) uses the 3D dot filtering and thresholding method for segmentation operation in which the pulmonary nodule candidates will be detected by the use of 3D dynamic self-adaptive template matching and finally passes to the FLDA classifier. The result of the work produces 90.24% sensitivity with 100 CT scans and 302 nodules.

Most of semi-automated lung nodule classification algorithm uses SVM classifier in detecting pulmonary nodules can be found in the work of (Halder et al., 2020) (El-Askary, Salem and Roushdy, 2019). The variation of these methods can be found in the image pre-processing method referring to the segmentation technique that specifies the target of ROI, in brief, the variation of lesions texture (solid and non-solid), localization (juxta-vascular, juxta-pleural, isolated, and pleural-tail). The traditional processing method aims at filtering/ extracts the foreground and background image in this case the lung (pulmonary nodules) and the soft tissue surrounding the lung (blood vessels). The limitation in the non-deep learning approaches can be pointed out at the inflexible of extracting pulmonary nodules in different sizes, shapes, and scale variability.

In a comparison of deep learning approaches, the performance has been proven to produce more superior performance than the non-deep learning method due to the high-scale variability in data-driven learning ability in which gradually became the trending approach. This is made possible due to available public datasets and the maturation in deep learning algorithms altogether. Deep learning-based algorithm revolves around the concept of Convolutional Neural Network (CNN) in which various architecture can be applied to produce different accuracy and sensitivity (EL-Bana, Al-Kabbany, and Sharkas, 2020).

The work of (Essaf, Li, Sakho, and Gadosey, 2020) developed an improved CNN model for nodule detection and classification which produces an accuracy result of 95.85% with 4.3/scan FPR based on 820 CT scans. The work performed dual path of convolutional neural network which

essentially a FPR component by taking input of a segmented lung in 2D image CT scan shown in figure 2.4. The proposed work was compared to the Lenet-5 network and other traditional methods.

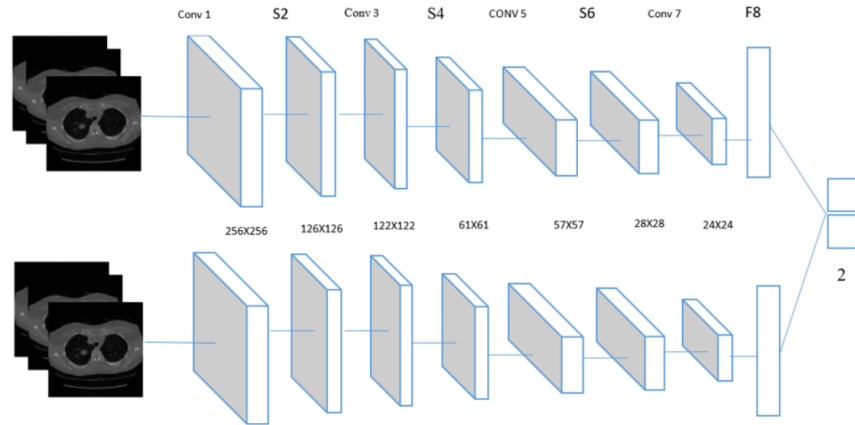


Fig 2.4: Dual Path Convolutional Network Structure (Essaf, Li, Sakho, and Gadosey, 2020).

The work by (Cao et al., 2018) developed a 3D CNN fusion model for lung nodule detection by harnessing the 3D features of volumetric space, the work performed a data augmentation to handle the imbalance classes and produce sensitivity of 95.1% at 8/scan FPR. The work uses LUNA 16 datasets and additional 335 patients of CT scans from the hospital to boost the performance of the approach based on the real situation. The model essentially employs two pre-trained model with one model trained on LUNA 16 and additional augmentation and the other additional was trained under a new dataset. The additional model intend to fuse the results of the two best performing models to enhance the result.

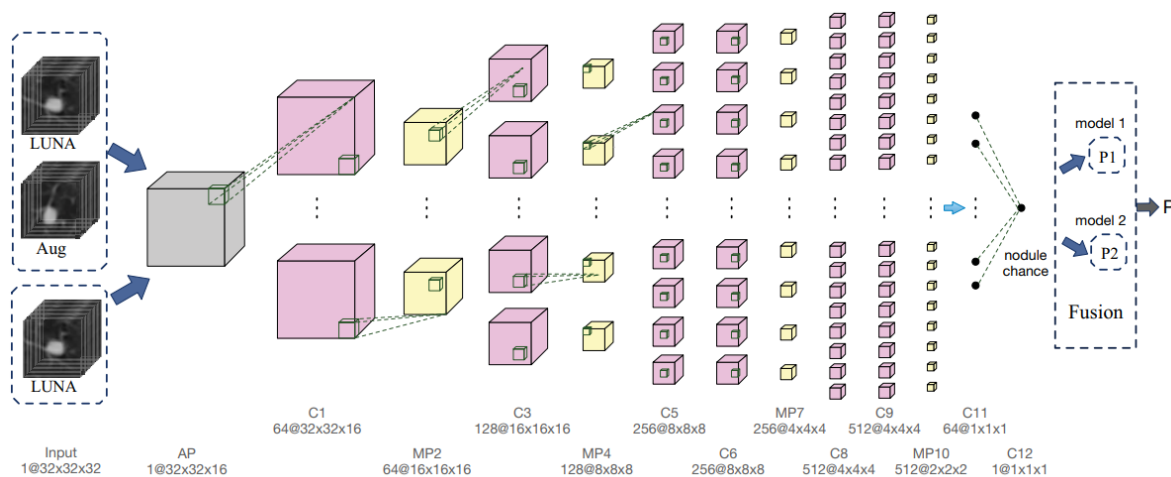


Fig 2.5: 3D Fusion Model Structure.

(Carvalho, Moreira, Figueiredo and Papanikolaou, 2019) proposed an automatic lung nodule detection CNN and segmentation using 2D residual deep CNN in which the work employs a sliding window method for nodule detection. The proposed work essentially employs 2 strategy method starting from the lung nodule detection component with 2D CNN and follows by the segmentation component which uses U-Net residual network. The work by (Rafael-Palou et al., 2021) proposed a nodules re-identification using Siamese Neural Network (SNN) which is an appealing characteristic of SNN whereby distance metric will be computed on features that had been extracted automatically by deep learning. The work had a different approach where the idea came from so call follow-up scenario where the patient will be diagnosed regularly and notice the difference in CT scan results. The work produces 89% accuracy in cross-validation and 92% in the test.

In the work by (Cao et al., 2020) proposed a lung nodule segmentation using Dual-Branch ResNet backbone with the capability of simultaneously capturing multi-view and feature-scale CT images. The proposed work intends to address the need of robust lung nodule segmentation framework with the proposed method of central intensity pooling (CIP) shown in figure 2.6, which was applied to multi-level feature. The work proposed to implement a Deconvolution Single-Shot Detector (DSSD) for its future work and was theorized that the model might be even more enhanced with the current proposed method.

The work by the author (Li, Y. and Fan, Y., 2020) proposed a lung nodule detection framework with a single strategy called Deep SEED model, essentially performs lung nodule localization framework with the 3D ResNet-18 with Squeeze and Excitation module based on the work (Hu, J., Shen, L. and Sun, G., 2018). The work adopts 3D RPN structure as backbone for its lung nodule detection framework. The work also adopts focal loss function for the model to combat imbalance dataset which especially common case for medical image dataset.

The work by (Guo et al., 2019) did a study on tumor segmentation masks with fusion strategy which the paper claims that by adopting different modality will produce a better result for segmentation. The work concluded with the verdict of multi-modality performs better than a single modality model and the work extended its experiments with low grade images, in which it appears to affect the quality of segmentation. The second finding was that the result performance of segmentation based on lower quality image is the same level or even higher than the performance of single modality network based on original image (Guo et al., 2019).

Another relevant research on lung nodule detection with localization by (Liu et al., 2020) did a study on automatic detection of pulmonary nodules on CT images using YOLOv3. The work proposed two distinct properties: one with automatic multi-scale feature extractor for nodule feature screening and two, the feature based bounding box generator. The study obtained sensitivity of 99.3% (FPs = 1) and was improved to 100%. In addition, the work reported that it achieved high computation efficiency.

2.2 Conclusion

Based on the literature review, the topic on attention module in lung nodule detection especially with the approach of 2 strategy: localization and FPR is lacking. The approach by (Li, Y. and Fan, Y., 2020) perform the squeeze and excitation which can be intriguing to develop with the current existing model. In medical image problem task, one of the challenges faced were imbalance dataset which is due to privacy concern and simply rare to collect, to solve this issue, many of the approach perform multi-scale feature extraction and other self-enriching image data such as Generative Adversarial Network (GAN) based model.

The identified research gap in this study is the use of attention module which essentially perform a more meaningful convolutional operation whereby feature map can be selectively calibrated to which may result in a better feature representation. The intended proposed work will be distinctive to the current literature review by applying squeeze and excitation layer on top of existing model.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Research Design

The main components for lung nodule detection with deep learning are: (1) lung nodule localization, (2) false positive reduction. The lung nodule localization requires object detection algorithm such as region-based detection model for multi-task (regression and classification) while for false positive reduction refers to CNN based model for classification task.

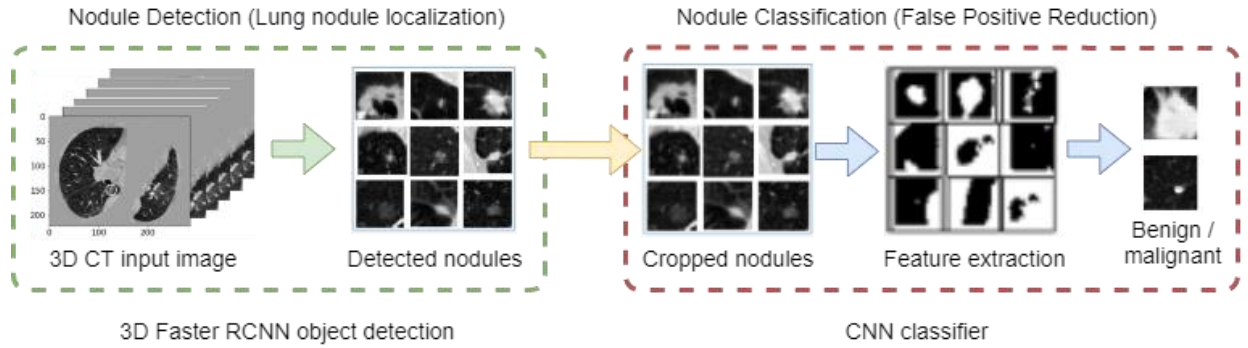


Fig 3.1: Lung Nodule Detection Pipeline Diagram.

Figure 3.1 shows the Lung nodule detection pipeline, where in lung nodule localization component, 3D lung pre-processed image in CT modality will be the input for 3D faster RCNN object detection model and outputs detected nodules. Then based on detected nodule's location, candidate will be cropped from the original CT image and passed as an input to the CNN classifier for further false positive reduction. Finally, the nodule classification will produce output of benign or malignant based on the proposed candidate. This block diagram determines the flow of the lung CT image system from input to the output.

3.2 Investigation on Model Architecture, Deployment Platform and Dataset

In this section, the investigation on algorithm, deployment platform and datasets are discussed in separate sections in this chapter. The investigations will help to narrow down the choices and requirement for the system implementation prototype. The first part of the is the model

Architecture and comparison table is done to highlight the difference in the characteristics. The following part will be the deployment platform where comparison is done between AWS Sagemaker Notebook and Google Collab.

3.2.1 Selection on model architecture

The selection of model architecture was decided based on a comparison between 2 computer vision detection models highlighted in table 3.1 to compare its characteristics. Faster RCNN and YOLO was developed around the year 2015 which was further developed still in the following year with the new backbone module inception net, mobile net and so on. The Region based CNN network is one of the most popular object detections which comprises of 2 parts in the network, Region Proposal Network and CNN classifier on top. The algorithm first started with an input 2D image where size was kept at least about 1000, 600 which then will be down sampled into a smaller dimension in order to keep the computation low for region proposal, shown in figure 3.1.

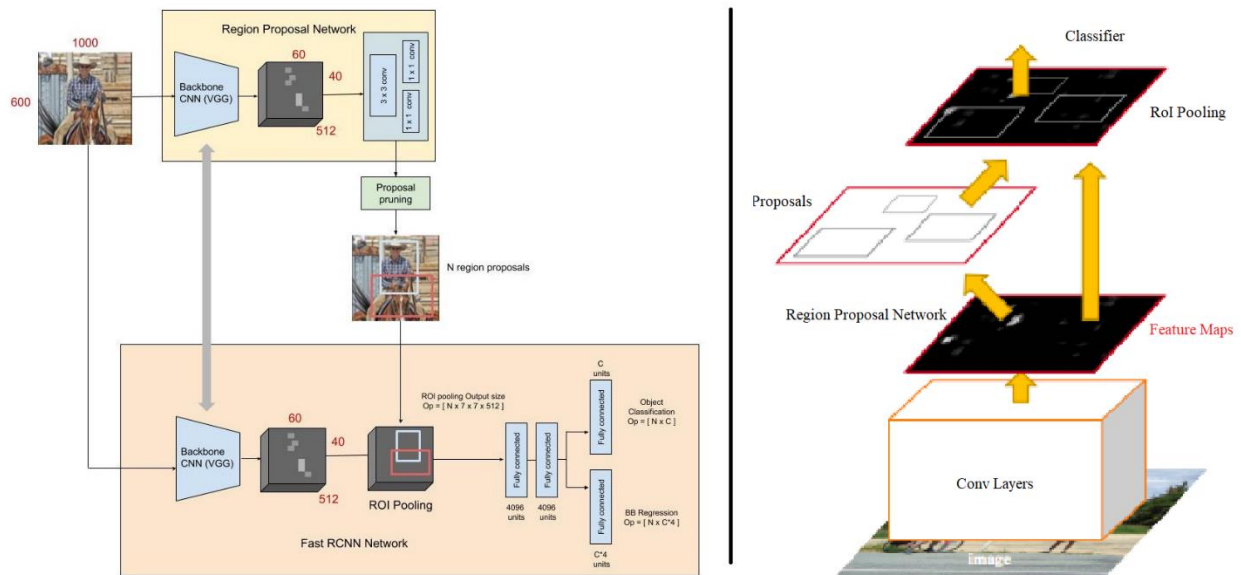


Fig 3.2: The Faster RCNN on Region Proposal Network (Ren et al., 2015).

To generate the so-called proposals for the region where the object may lie, a small network will be slide over the convoluted feature maps from the input image. Based on the figure 3.2, it shows how the faster RCNN work with the region proposal network and the architecture above can be further broken down into the mechanism of region proposal.

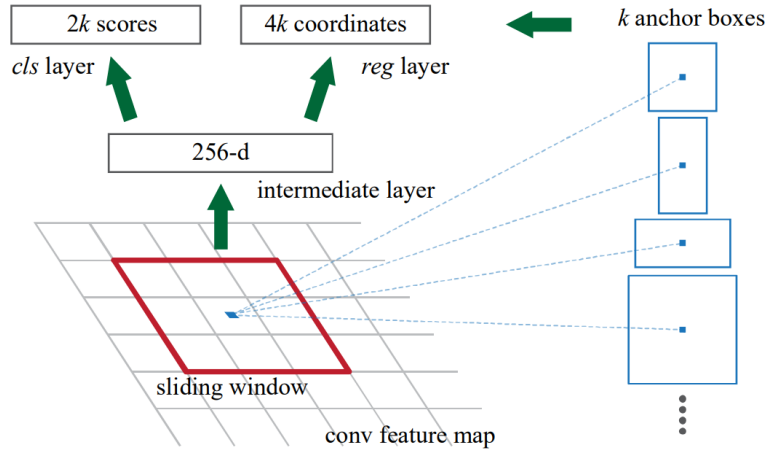


Fig 3.3: Region Proposal Network (RPN) (Ren et al., 2016).

As per mentioned earlier the region proposal network will slides over the convoluted feature map from the input image where the author introduces the concepts of anchors that is described as the central point of the sliding window, shown in figure 3.3. The sliding window will produce 9 proposals based on 3 scale and 3 aspect-ratio of the anchor boxes. The classifier is used to determine the probability of the object exist in the anchor boxes while regression layer is to determine the coordinates of the proposal. The mechanism of the anchor's probability score is dependent on the parameter Intersection-over-union whereby the anchors will be measured based on highest intersection-over-union overlap with ground truth box.

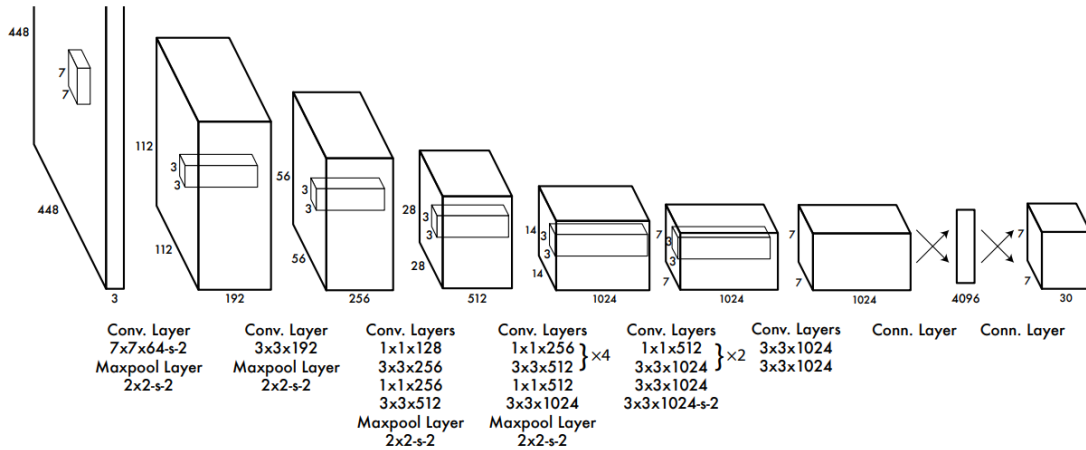


Fig 3.4: YOLO architecture (Redmon et al., 2015).

In comparison to YOLO architecture, the prediction of the object exist in the image occurs only once where in contradictory, region-based model (Faster RCNN) perform detection on

various region proposals which then results in performing prediction on multiple times. YOLO architecture works in a simpler way compared to region-based, in figure 3.4, the architecture performs resizing of input image and continues to run a convolutional network which finally performs Non-Maximum Suppression (NMS) based on a given threshold, shown in figure 3.5.

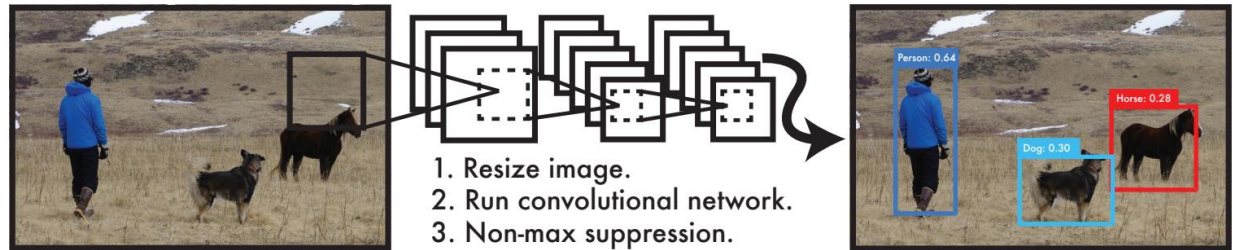


Fig 3.5: YOLO Detection System. (Redmon et al., 2016).

The comparison of models can be summarized in table 3.1, essentially the region proposal network in Faster RCNN allows object detection in almost real-time manner with approx. time taken of 0.12 seconds per image however despite the efficiency, the algorithm is not fast enough to be used on real-time applications (Sanchez, Romero and Morales, 2020). In contradictory, YOLO location of object is very efficient however it has lower accuracy detection in smaller objects compared to Faster RCNN model. To align the needs of the current project, lung nodule detection does need a real-time speed efficiency however, accuracy of detection carries more weight in the decision factor. Hence, Faster RCNN slower computation can be compensated with the higher accuracy compared to YOLO architecture. Ultimately, Faster RCNN model is chosen for the proposed model.

Table 3.1: Comparison table between Faster RCNN and YOLO.

Model Architecture	Faster RCNN	YOLO
Characteristic		
Computation Speed	Slower	Faster
Prediction Accuracy	Very Accurate	Accurate
Efficiency	Efficient	Very Efficient

3.2.2 Selection on Deployment Platform

The selection of deployment deep learning machine platform analyses the infrastructure capability to run the model which includes the hardware specifications for storage, accelerating computation component – Graphics Processing Unit (GPU) instance and memory, Random Access Memory (RAM), and costs. In table 3.2 shows the comparison table between AWS Sagemaker Notebook instance, Google Colab.

Generally, the choice of deployment setup will be Jupyter notebook with python programming language, local or virtual machine with considerable large storage unit of 300 GB for storing essential components such as dataset, scripts, other than that higher processing unit with higher RAM of 12 GB and compatibility with the Operating System (OS) usually in Linux. Deep Learning cloud services such as AWS provides a variety choice, especially with the AWS Sagemaker notebook instance service which capable of adjustment of instance type, and storage unit according to project needs. An instance type usually covers a certain fixed hardware specification and categorized into 4 groups namely, compute optimize, memory optimize, accelerated computing, and storage optimized. Each groups have different levels of RAM, GPU size, instance size and cost according to the given hardware specification. In comparison to Google Colab free tier, have a fixed hardware capability, in addition to the storage unit is still somewhat separated – referring to google drive.

Table 3.2: Comparison table between AWS Sagemaker and Google Colab.

DL Cloud Service Specification	AWS Sagemaker Notebook	Google Colab
Hardware Specifications	Compute Optimize – C5.xlarge: <ul style="list-style-type: none">• 4 vCPU• 4 GiB memory Accelerated Computing – p3.2xlarge: <ul style="list-style-type: none">• 8 vCPU• 61 GiB memory•	8 vCPU, 12 GiB memory
Pricing	C5.xlarge – SGD \$2/hour P3.2xlarge – SGD \$5/hour	Free
Bandwidth, storage unit, upload/download rate	Preferable	Not suggested

The two distinct services provide considerable options whereby one service is better than the other, however both options can be used to meet the needs. Ultimately, the proper implementation of the project will be using AWS Sagemaker Notebook instance while google Colab can be used for minor applications. In terms of reliability, google Colab is far from reliable as the current project involves a large dataset, rather than typical data science purpose/usage. Therefore, AWS Sagemaker Notebook instance is chosen.

3.2.3 Selection on Lung Nodule Public Dataset

The dataset that will be used for this project can be described as a CT lung nodule image which usually in the format of dicom / raw meta format file extension. There were a few available public lung nodule datasets which were made for past research namely, 1) Lung Image Database Consortium (LIDC-IDRI) which is a web-accessible international resource for development of computer-assisted diagnostic (CAD) system, 2) LUNA 16 stands for Lung Nodule Analysis which is a CT lung nodule image dataset made available for grand challenge development in 2016.

Table 3.3: Comparison table of Public Lung Nodule Dataset.

Dataset	Description	Feasibility
LIDC-IDRI	Lung image database consortium with 1318 studies, 1010 patients	Dataset downloading procedure requires complicated steps.
LUNA 16	A dataset derived based on LIDC-IDRI and divided into 10 subsets with 88 scans each fold. The subset has been hand picked from 1010 to 888 patients.	Dataset download is simpler than LIDC-IDRI and filtered.

Out of the two listed dataset, LUNA 16 was chosen due to two factors, firstly, the dataset is slightly cleaner whereby the patients scan was filtered from 1010 to 888 patients concerning the storage value, and secondly the LUNA 16 has more relevant approach than LIDC-IDRI projects. Ultimately, the LUNA 16 dataset is chosen.

3.3 Model Architecture Derivation

In this subsection, the Faster RCNN model on Dual Path Network (DPN) and Res-Net 18 SE will be derived based on the existing and proposed backbone. This is done to demonstrate the understanding of it and how it is derived.

3.3.1 3D Dual-Path Network

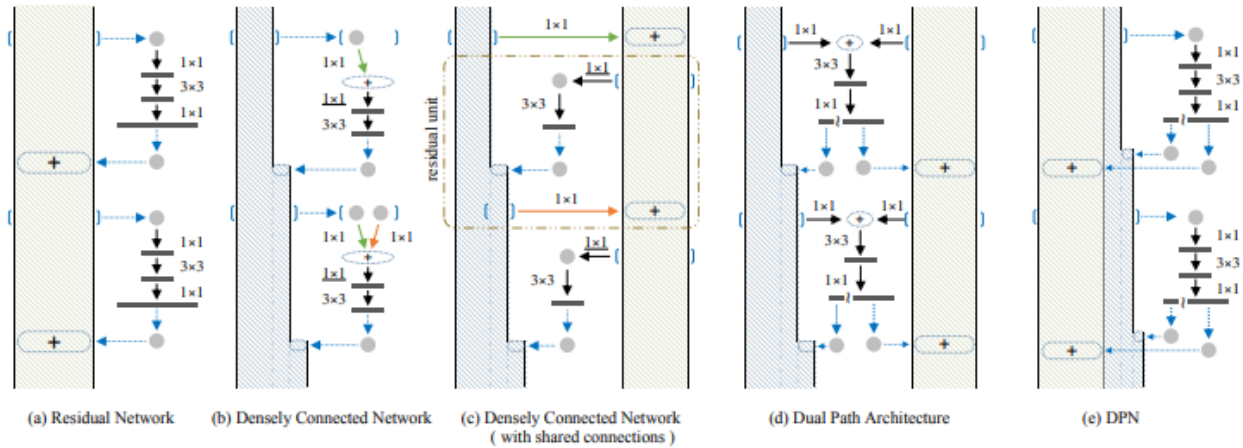


Fig 3.6: Architecture Comparison of different networks and formation of DPN (Chen, Y.,et. Al, 2017).

Figure 3.6 shows the architecture comparison of Res-Net, Dense-Net and the combination of weight sharing forming Dual-Path Network. The work of based on (Chen, Y.,et. Al, 2017) enabling the use of effective feature re-usage and re-exploitation from the advantage of residual and densely connected network which was studied to have a higher parameter efficiency, lower computational cost and lower memory consumption. As you may observed from the figure 3.6, the (a) residual network indicated was showing the re-usage of lower features to the later layer in the pipeline of residual network whereas (b) the densely connected layer increases the thickness of the pipeline by introducing input layer to the output layer which then showing the ability to re-exploit new features to the later layers. The study then performs the combination of both network and created a sharing connection and developed dual-path network at (c).

The work of dual path network inspired (Zhu, W., et al., 2017) to adopt the network for the lung nodule detection purpose. The author proposed a 3D faster RCNN with the backbone of dual

path network for detection and implement gradient boosting machine for lung nodule classification. The connections of dual path network for the lung nodule detection can be illustrated in figure 3.7.

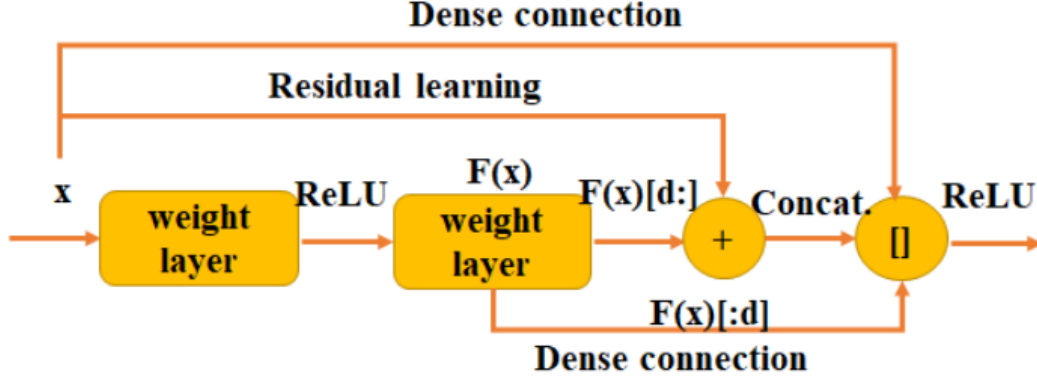


Fig 3.7: Illustration of dual path connection (Zhu, W., et al., 2017).

Table 3.4: Annotation description for fig 3.7.

Annotation	Descriptions
x	Input of the dual path block.
d	Number of features.
$F(x)$	Convolutional layer function.
$F(x)[d:]$	Residual learning.
$F(x)[:d]$	Dense connection.
$G(x)$	Rectified Linear Unit (ReLU).

Dual path feature map:

$$y = G([x[:d], F(x)[:d], F(x)[d:] + x[d:]]), \quad (1)$$

As per mentioned earlier, the dual path network employs residual and densely connected layer for feature re-using and re-exploitation, where the implementation of dual path network shares its feature map for residual learning and dense connection denoted with $F(x)[d:]$ and $F(x)[:d]$ respectively. The parameter d denotes the number of features to be exploited. The feature map for dual connection can be formulated as equation 1.

Where the output feature map y is the ReLU activation of the residual feature, dense connection feature and the input feature map. The dual path network will be used for the building of faster RCNN network with a U-net-like encoder decoder structure especially for region proposal network.

3.3.1 3D Res-Net 18 SE

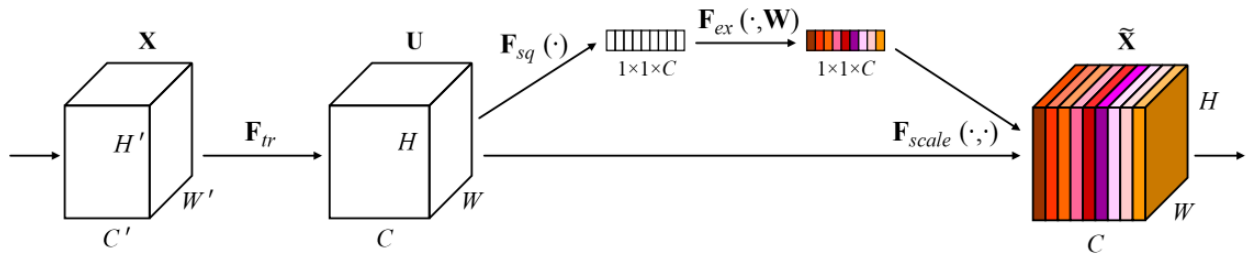


Fig 3.8: Squeeze-and Excitation block (Hu, J., Shen, L. and Sun, G., 2018).

Figure 3.8 shows the Squeeze-and-Excitation block by (Hu, J., Shen, L. and Sun, G., 2018) that enables networks to construct informative features by fusing both spatial and channel-wise information within local receptive field at each layer. The SE block can adaptively recalibrate channel-wise feature responses by two operations Squeeze and excitation. The two steps operation essentially performs a global average pooling and pass them through a simple gating mechanism with a sigmoid activation function to generate channel-wise statistics and return calibrated feature map to original channel dimension. The SE block intend to provide the access of global information and recalibrate the feature response for a better feature learning.

The author designed SE block with the intention to improve representational power of a network by performing a dynamic channel-wise calibration. The SE block had been used for enhancing the feature learning in which it had inspired (Li, Y. and Fan, Y., 2020) to performs a study on 3D Res-Net SE called deep SEED which is a model adoption for lung nodule detection using faster RCNN architecture with SE block on encoder-decoder structure for RPN network. Figure 3.9 shows the architecture of Res-Net 18 with Squeeze and Excitation (SE) structure in between the encoder-decoder structure by (Li, Y. and Fan, Y., 2020).

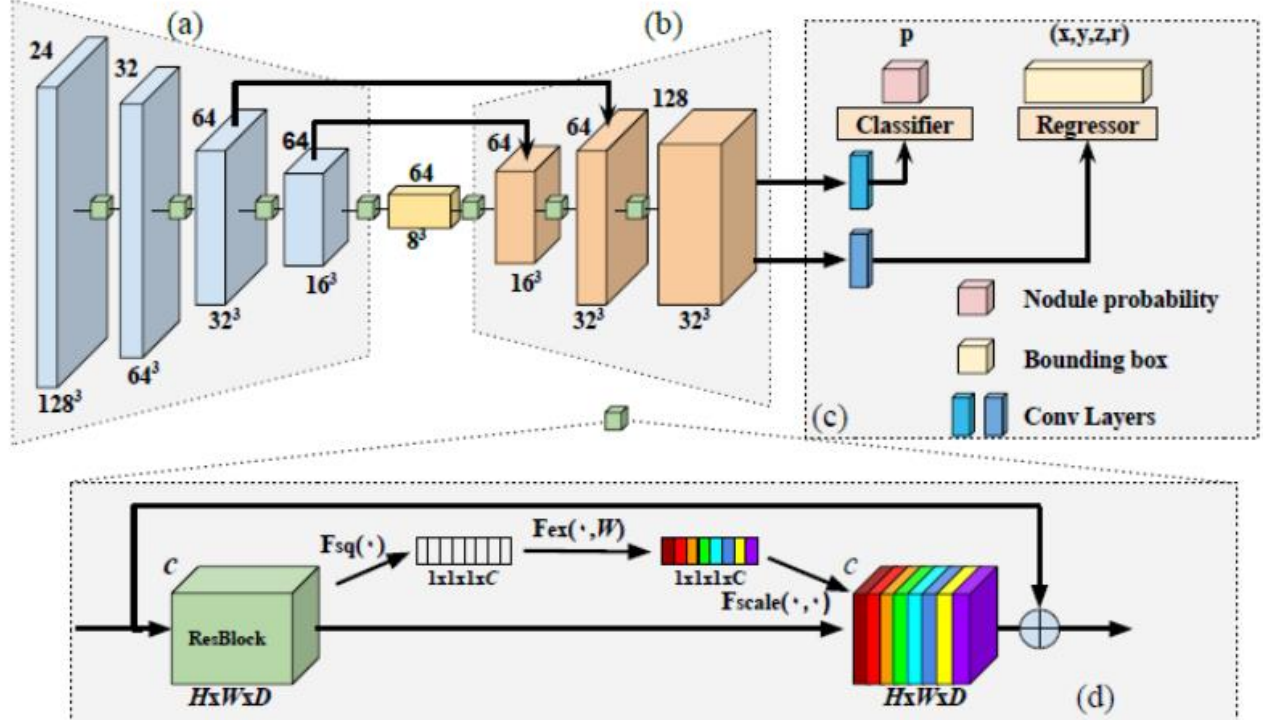


Fig 3.9: The architecture of 3D Res-Net 18 SE structure (Li, Y. and Fan, Y., 2020).

Due to the success in image classification task using SE block, the author intends to improve the lung nodule detection model by adaptively adjusting the weights of each feature map for lung nodule detection task.

Table 3.5: annotations description for Squeeze and Excitation block.

Annotation	Descriptions
x	Input feature.
U_i	Output feature after convolution layer.
Z_i	Convolutional layer function.
S_i	Residual learning.
F_{scale}	Dense connection.
$\sigma(x)$	Sigmoid activation function.
$\delta(x)$	Rectified Linear Unit (ReLU).

Convolution operation:

$$U = F(X) \quad (1)$$

Squeeze operations:

$$z_i = F_{sq}(u_i) \quad (2)$$

Excitation operations

$$\begin{aligned} s_i = F_{ex}(u_i) &= \sigma(g(Z, W)) \\ &= \sigma(W_2 \delta(W_1 Z)) \end{aligned} \quad (3)$$

As per mentioned earlier, the squeeze and excitation block were implemented after the feature map from the convolution layer which can be denote as the notation U where the $F(x)$ is the function of convolution. Next the output of squeeze operation is denoted by Z where Z is the global average operations on the feature map U by the function F_{sq} . The excitation operation is denoted by S where the output of a squeeze operations will be passed into a simple gating mechanism with sigmoid activation function and a nested ReLU activation function denoted by g will be performed on the output of squeeze operation with a parameter W where W_1/W_2 is a dimensionality reduction/increasing layer respectively. The output of SE block can be formulated as equation 3.

3.4 Proposed Methodology

The proposed system recommends the use of SE block and the efficient re-use and re-exploitation of feature maps which all directed to the region proposal network on 3D faster RCNN model. The proposed model intends to enhance the model performance by using SE block where generation of candidate proposal may be further improved with hope it reflects the result in classification output. The overview of the framework can be illustrated in figure 3.10 where the SE block will be placed at the last layer of convolution just before the residual learning in DPN network. The design essentially to test and confirm the enhancement of lung nodule detection component before it passes to the FPR component.

Proposed Model

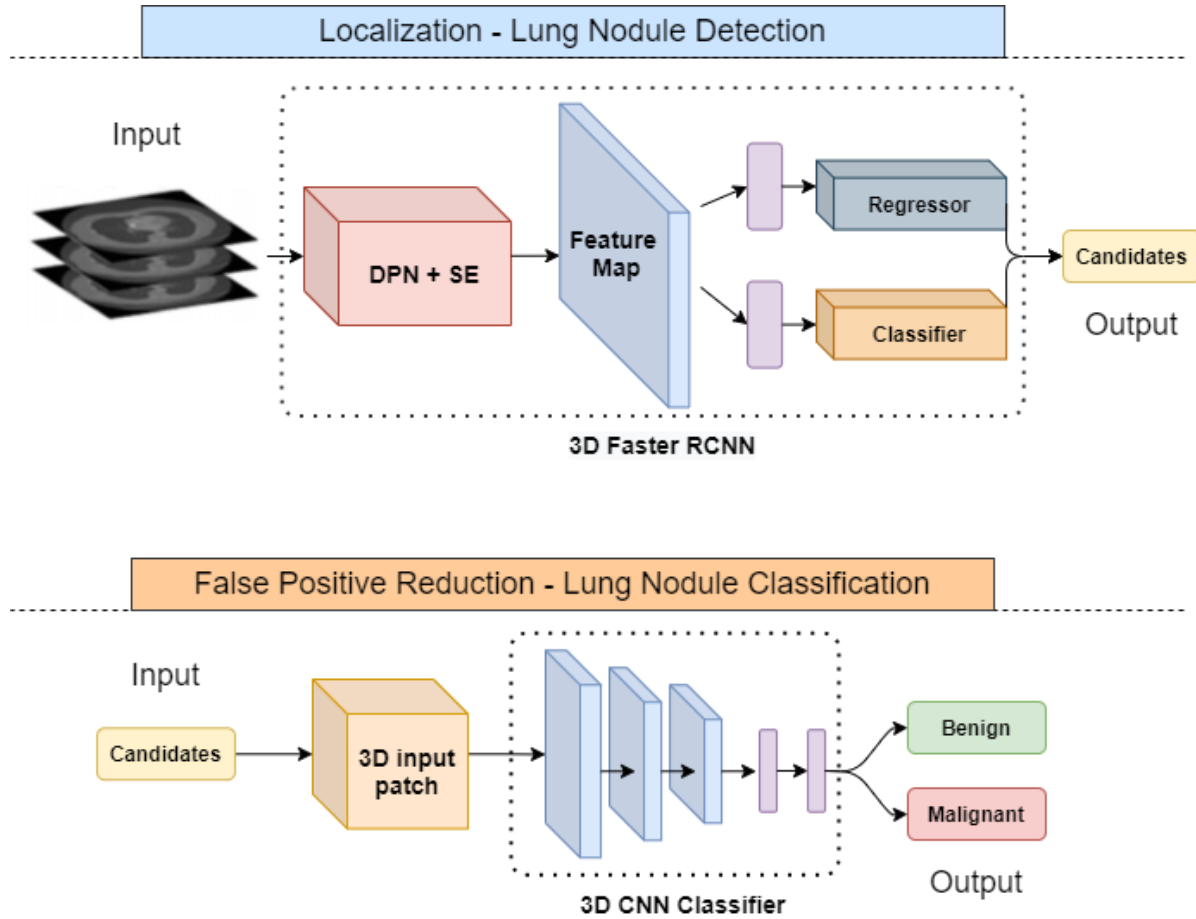


Fig 3.10: Overview of the framework.

Figure 3.11 shows the proposed placement of the SE block on the DPN network where the SE block will recalibrate the feature response based on the previous feature map and the block of the dual path network will be repeated according to the U-Net-like architecture which eventually improves the feature learning. Besides the implementation of SE block, the approach by the (Li, Y. and Fan, Y., 2020), the 3D Res-Net SE model adopts focal loss for the RPN network which is the improved version of cross entropy that intends to solve class imbalance problem in this case to correct misclassified samples nodule region. The focal loss function will be adopted in the current proposed model to handle imbalance dataset which is generally prominent across lung nodule detection dataset.

Proposed Dual Path SE Network Connection

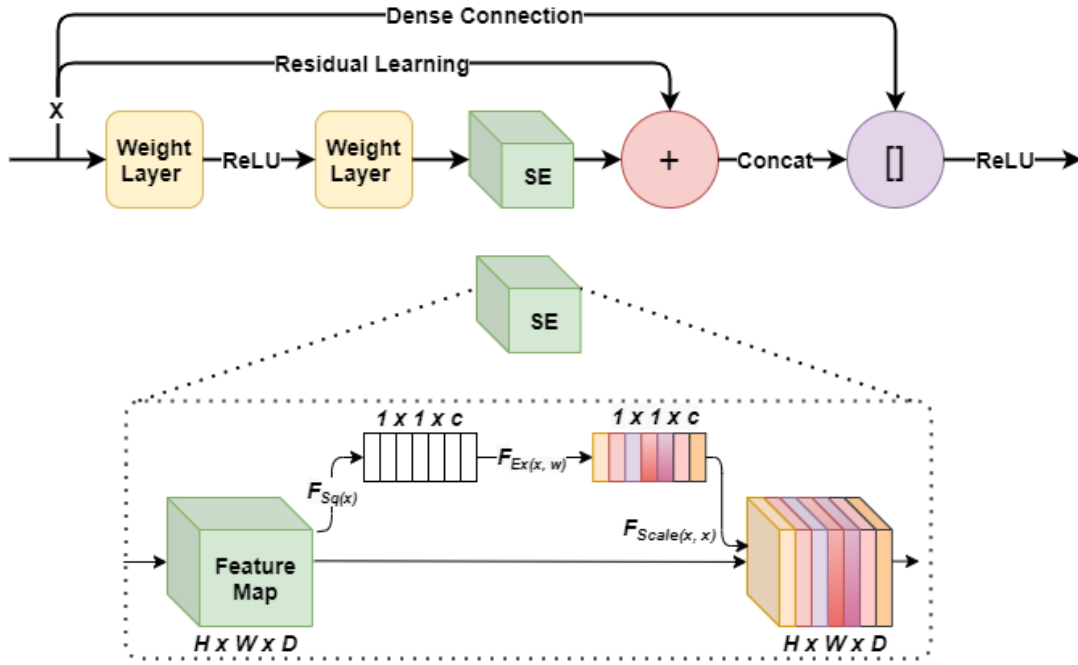


Fig 3.11: Proposed Dual Path SE block connection.

The loss function focal loss by (Ross, et al.,2017) for classification and regression loss function in RPN is derived by the equation below.

3.5 Summary

This section discussed the research design and methodology of the project where the research gap can be observed and to have a clear view and the mapping of the proposed model. The identification and evaluation of model characteristic was carried out where Faster RCNN was chosen to be the better choice of model in the context of medical usage with priority on accuracy rather than computation speed. The platform of deep learning cloud service was compared and concluded to deploy using AWS Sagemaker to cater and compute large dataset with pre-processing, training, and testing jobs in mind. The lung nodule localization model 3D dual path and 3D Res-Net SE was derived. Prototype proposal was pitched with the detail of the dual path SE block connection, and the adoption of the focal loss to combat the imbalance class.

CHAPTER 4

FINAL DESIGN & SYSTEM IMPLEMENTATION

4.1 Introduction

In this chapter, the final design and the implementation of the proposed concept will be shown along with the obtained results. This section documents system implementation and its justification of the approach with appropriate evaluation method. Based on the proposed method, the system was implemented and tested with several runs to determine the validity of the output and working mechanism of deep learning model. The justification of each approach is presented in the accordance to the subsection. The successful approach will then be used to proceed for further development.

4.2 System Implementation

In depth and thorough research was done in chapter 3 to help streamline the process implementation and to prevent any discrepancies especially in lung nodule tasks. The system implementation/construction can be grouped into four major parts: 1) data acquisition and pre-processing, 2) modelling, 3) training and testing, finally, 4) Evaluation. Based on the research, method applied in the implementation was crosschecked with past research studies especially on data pre-processing, which is still relevant back from 3 – 4 years past experiments. In terms of timely duration, the most timely used was in the pre-processing stage whereby acquiring and cleaning the requires large storage unit and memory allocation. Second will be the training and testing due to the large number of datasets which slows down the processing time.



Fig 4.1: System Implementation Routine.

4.2.1 Data Pre-processing

In this section, documents the method applied for data cleaning / pre-processing on raw CT scan images acquired from LUNA 16 dataset grand challenge website. The essential module required in data pre-processing includes data reading, data normalization, conversion from pixel unit to Hounsfield, re-sampling of 3D voxel/patches, augmentation, region saving and others.

To start off the data pre-processing, the understanding of the CT scans must be established by acknowledging the way of CT scans work and how the LUNA 16 dataset formed from LIDC-IDRI dataset. A brief information regarding LUNA 16 dataset, the dataset consists of 888 CT scans which were gather from LIDC-IDRI only with slice thickness less than 3mm. There are total of 36, 378 annotations of nodules that were marked by more than one radiologist with distribution of 2290, 1602, 1186, and 777 nodules annotated by at least 1, 2, 3, or 4 radiologists respectively. Based on the marking, nodules detected by 3 radiologists were deemed to be true nodules.

The nature of any relevant medical image in CT modality follows an image coordinate system where medical scanners create rectangular arrays of point which start at the upper left corner. The essential axis of anatomical coordinates denoted with notation (I J K) which represent I axis to the right, J axis to the bottom and K axis backwards. Based on an article of coordinate system (Coordinate systems - Slicer Wiki, 2021) explains the anatomical representation with origin position representing voxel (0 0 0) consists of a certain size e.g.(100mm, 50mm, -25mm) with a certain spacing specified e.g. (1.5mm, 0.5mm, 0.5mm) illustrated in figure 4.2. The origin and spacing information, image position / image coordinate voxel in anatomical coordinates can be calculated.

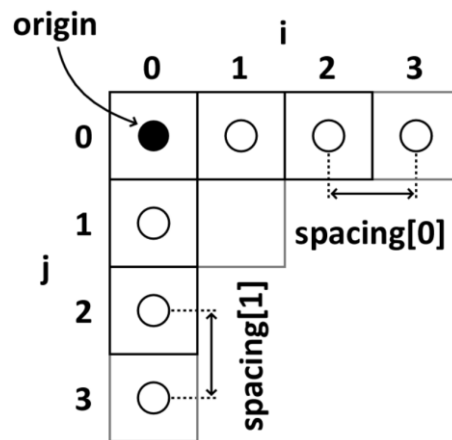


Fig 4.2: 2D example of spacing and origin (Coordinate systems - Slicer Wiki, 2021).

Slice thickness and slice increment in CT scans is an essential concept to understand the mechanism of the working principle. Slice thickness often refers to the axial and depicted as resolution of the scan, whereas slice increment refers to the movement of the scanner/table for scanning to the next slice, illustrated in figure 4.3. The image reconstruction planes can be referred in figure 4.4.

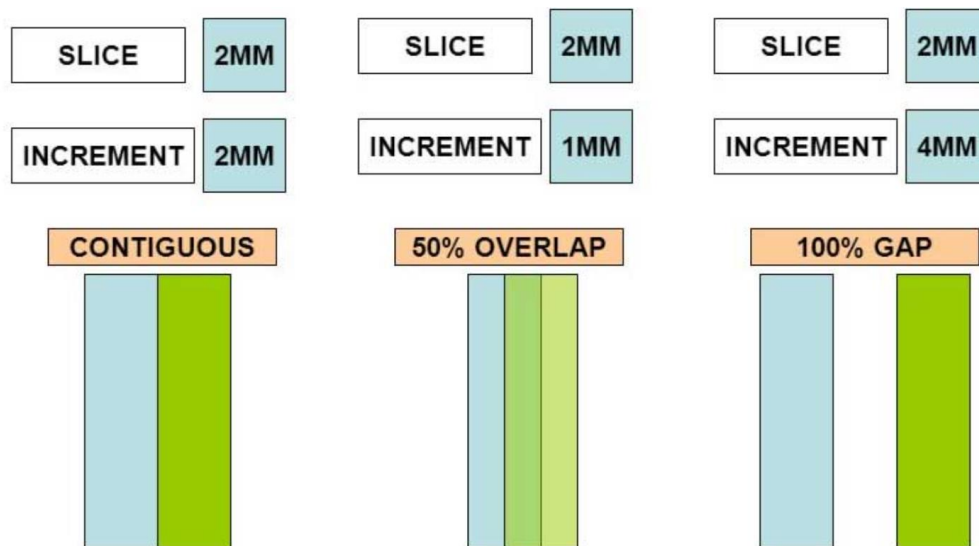


Fig 4.3: Slice and Increment illustration (What is the difference between slice thickness and slice increment?, 2021).

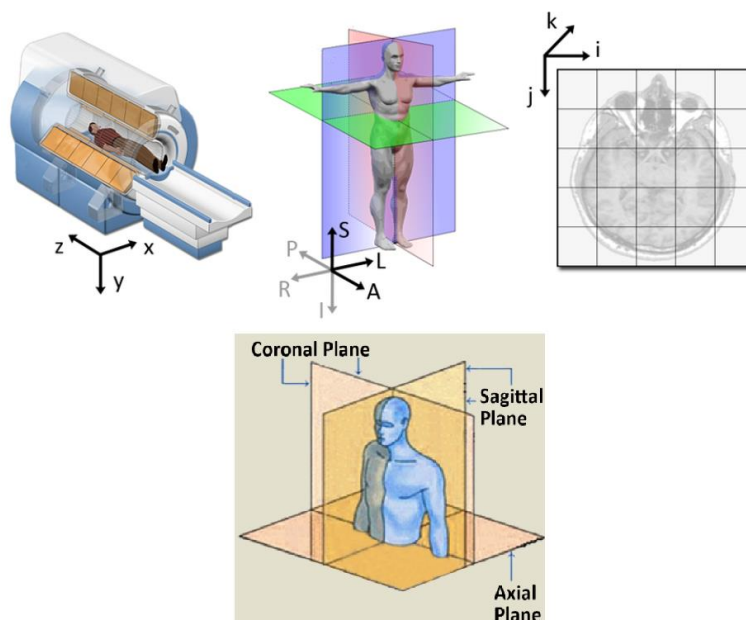


Fig 4.4: Reconstruction of the image planes (Coordinate systems - Slicer Wiki, 2021).

Based on the review of medical images CT scan, essential pre-processing must be done before feeding to the model, where steps of cleaning/pre-process can be illustrated by a flowchart in fig 4.5. The flow starts with loading image using a python library module called SimpleITK (SITK) stands for Simple Insight Tool Kit which allows reading image file in .mhd – raw meta format file extension. Image will then be resampled into isomorphic form of (1, 1, 1) due to inconsistency in different device configurations, doing so will streamline 3D images spacing.

Once the image has been resampled, the next step is to clean up the unnecessary components body tissue/organ, starting with binarizing the 3D slice images. Removing blobs that is connected to the border of the image and label the connected points followed by keeping the labels with 2 largest areas and proceed to segment the two lungs. Next, filling the holes inside the mask of lungs and fill the convex hull of each lung. Once the hull is filled, join the two separated right and left lungs, and perform a closure operation with disk of radius 10 to keep the nodules attached to the lung wall. The last few steps will be superimposing the binary mask to the input image, normalize the image by setting Hounsfield unit bounds and perform a scaling to 0.0 to 255.0. Normalization of the image involves the Hounsfield Density unit range which refers to the Hounsfield density range table, in table 4.1. Finally, zero centering the image with the mean coefficient of the whole LUNA 16 dataset. The following pre-processing techniques was adopted by (Mostafa, A, 2020) and cross-checked with pre-processing technique employed by (Liao, F., et al., 2017). The pre-process results can be illustrated in figure 4.6 with the label indicating the steps.

Table 4.1: Hounsfield Density Range.

Matter	Density (HU)
Air	-1000
Lung parenchyma (inspiration)	-850 to -910
Fat	-50 to -100
Water	0
White matter	20 to 30HU
Kidney	20 to 40HU
Spleen	35 to 55HU
Grey matter	37 to 45HU
Blood	45 to 65HU
Liver	45 to 65HU
Hematoma	40 to 90HU
Bone	700 to 3000

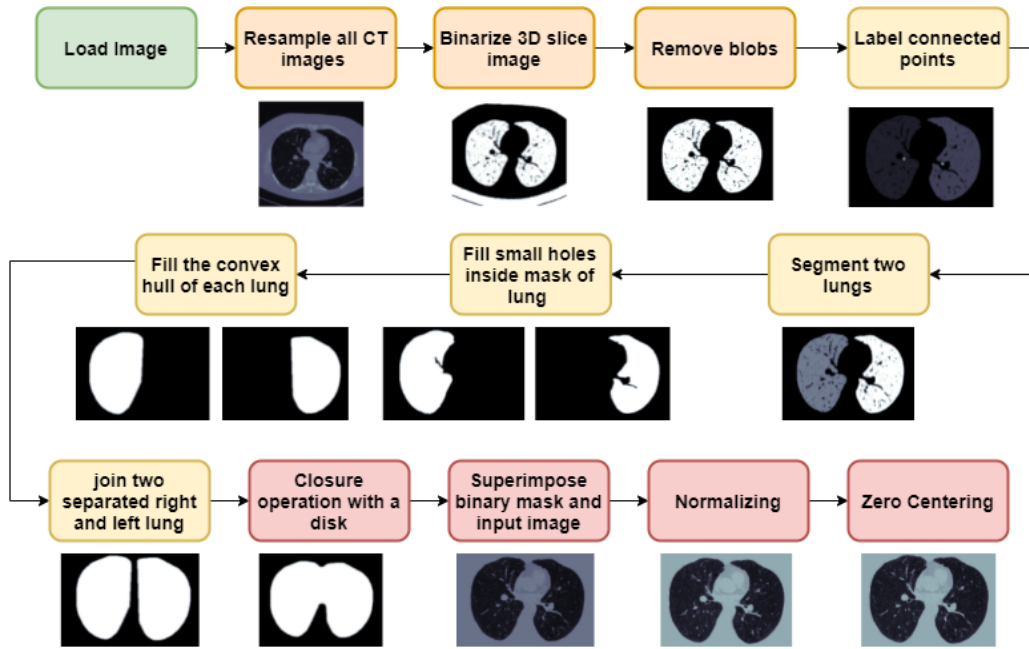


Fig 4.5: Flowchart of pre-processing routine.

4.2.2 Modelling

This section documents the modelling of the architecture namely, 3D faster RCNN for lung nodule localization and CNN classifier for lung nodule false positive reduction. The section is further divided into 2 parts: data loader, model architecture, and loss function. The model construction for lung nodule localization uses PyTorch API deep learning frameworks integrated with python 3.9, reason being was to access custom configurations with lower-level framework and to access the past study practices. PyTorch API offers a more complex options and requires more time in the building the framework however it also offers flexibility in optimization, loss calculation during training. On the contrary, the FPR model was less focused therefore simple TensorFlow is suffice.

The continuation after the data pre-processing, data must be pre-pared for the loading to the model in which PyTorch API uses so called data-loader method where a separate script will be called to access and provide data loading processes. The data loading script consists of a data loader class and comprises of 3 main functions: 1) `__init__`, 2) `__len__`, and 3) `__getitem__` (Pytorch.org., 2021). The implementation requires csv file for annotations and directory path for the `__init__` function to initialize the dataset image and it's annotation.

Moving on to the `__len__` function which works by returning the number of samples in our dataset while `__getitem__` function loads and returns the sample from the dataset at the given index

initialized earlier in the `__init__` function (Pytorch.org., 2021). Doing so will ease the loading of the dataset and provides readability aside from the model creation.

4.2.2.1 Proposed Lung Nodule Localization

The model architecture can be expanded in terms of block diagram which illustrated in the figure 4.6. The proposed model architecture was adapted upon dual path network model by the author (Zhu, W., et al.,2017) where the proposed model summary is described in the table 4.2.

Table 4.2: Proposed Model Summary for lung nodule localization.

Stage	Output	Original DPN Layer	Proposed Layer
Pre-dual path	$96 \times 96 \times 96, 24$	$3 \times 3 \times 3, 24$	$3 \times 3 \times 3, 24$
Dual path block 1	$48 \times 48 \times 48, 48$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 24 \\ 3 \times 3 \times 3, 24 \text{ stride } 2 \\ 1 \times 1 \times 1, 32 \end{cases}$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 24 \\ 3 \times 3 \times 3, 24 \text{ stride } 2 \\ 1 \times 1 \times 1, 32 \end{cases}$ Squeeze and Excitation layer
Dual path block 2	$24 \times 24 \times 24, 72$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 48 \\ 3 \times 3 \times 3, 48 \text{ stride } 2 \\ 1 \times 1 \times 1, 56 \end{cases}$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 48 \\ 3 \times 3 \times 3, 48 \text{ stride } 2 \\ 1 \times 1 \times 1, 56 \end{cases}$ Squeeze and Excitation layer
Dual path block 3	$12 \times 12 \times 12, 96$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 72 \\ 3 \times 3 \times 3, 72 \text{ stride } 2 \\ 1 \times 1 \times 1, 80 \end{cases}$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 72 \\ 3 \times 3 \times 3, 72 \text{ stride } 2 \\ 1 \times 1 \times 1, 80 \end{cases}$ Squeeze and Excitation layer
Dual path block 4	$6 \times 6 \times 6, 120$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 96 \\ 3 \times 3 \times 3, 96 \text{ stride } 2 \\ 1 \times 1 \times 1, 104 \end{cases}$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 96 \\ 3 \times 3 \times 3, 96 \text{ stride } 2 \\ 1 \times 1 \times 1, 104 \end{cases}$ Squeeze and Excitation layer
Deconvolution 1	$12 \times 12 \times 12, 216$	$2 \times 2 \times 2, 216$	$2 \times 2 \times 2, 216$
Dual path block 5	$12 \times 12 \times 12, 152$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 128 \\ 3 \times 3 \times 3, 128 \text{ stride } 2 \\ 1 \times 1 \times 1, 136 \end{cases}$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 128 \\ 3 \times 3 \times 3, 128 \text{ stride } 2 \\ 1 \times 1 \times 1, 136 \end{cases}$ Squeeze and Excitation layer
Deconvolution 2	$24 \times 24 \times 24, 224$	$2 \times 2 \times 2, 152$	$2 \times 2 \times 2, 152$
Dual path block 6	$24 \times 24 \times 24, 248$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 224 \\ 3 \times 3 \times 3, 224 \text{ stride } 2 \\ 1 \times 1 \times 1, 232 \end{cases}$	$2 \times \text{block of } \begin{cases} 1 \times 1 \times 1, 224 \\ 3 \times 3 \times 3, 224 \text{ stride } 2 \\ 1 \times 1 \times 1, 232 \end{cases}$ Squeeze and Excitation layer
Output	$24 \times 24 \times 24, 3 \times 5$	Dropout rate = 0.5 $1 \times 1 \times 1, 64$ $1 \times 1 \times 1, 15$	Dropout rate = 0.5 $1 \times 1 \times 1, 64$ $1 \times 1 \times 1, 15$

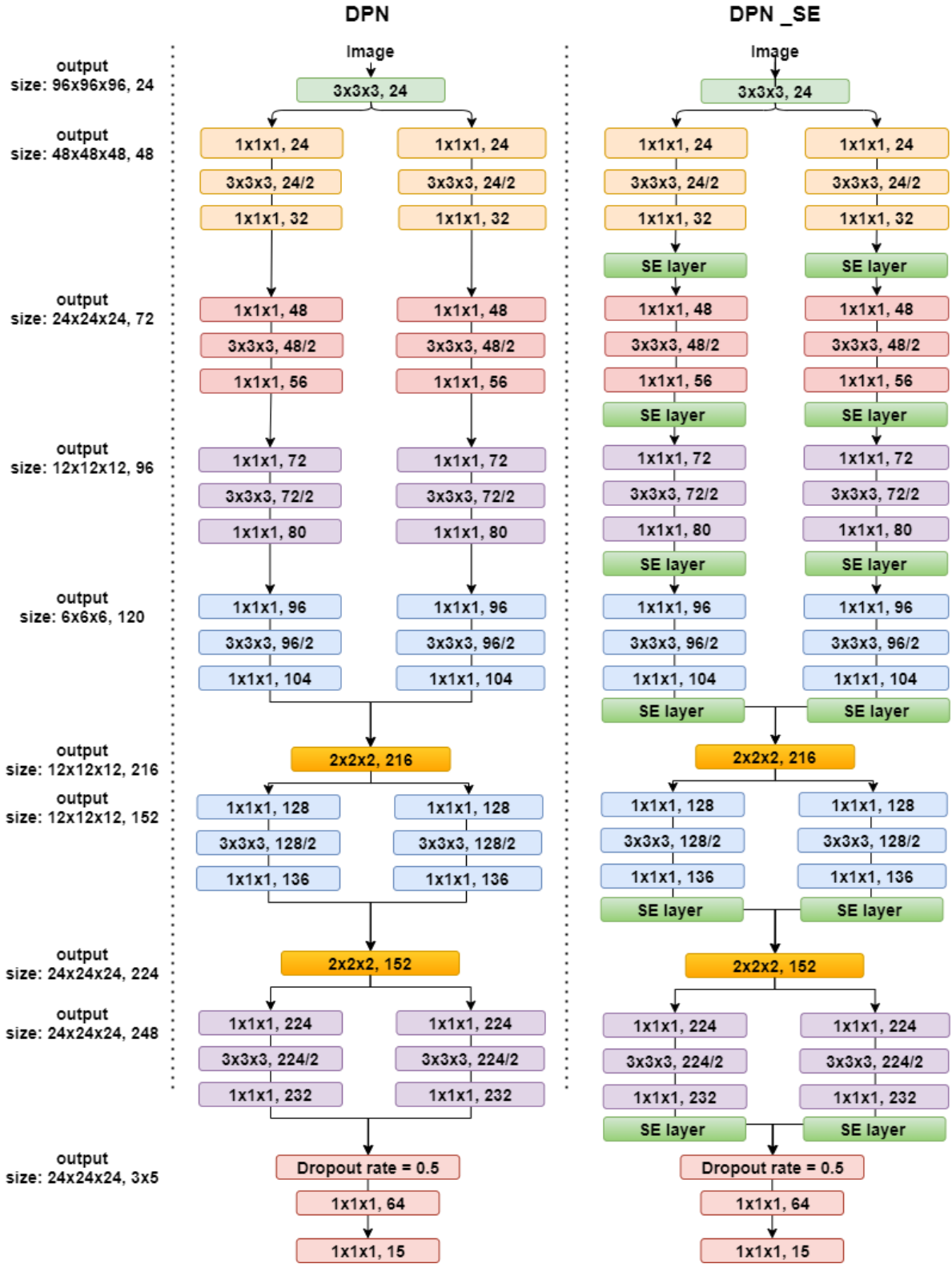


Fig 4.6: Proposed Model Network Architecture and Original DPN Model based on (Zhu, W., et al.,2017).

Figure 4.6 shows the proposed model with the network architecture detail, where the difference of model construction between the proposed and the baseline model is the SE layer attached in the end of every layer. The model construction consists of 2 basic blocks with residual and dense connections, the common stacked layers in between namely, 1) convolutional 3D, Batch normalization 3D will be stacked forming a bottleneck block. The convolution layer and batch normalization are set to a specified set of dimensions mimicking the U-Net-like architecture where encoder-decoder will have a symmetrical and a bottle-neck feature.

The concept of encoder-decoder essentially illustrates the contraction of the input image causing the increase of the feature map on encoder side and returns its dimension of the input image size back again in expansion/ decoder side. Notice the color-coded model in figure 4.6 which intend to indicate the symmetrical image size. The specified set of dimensions was adapted based on the previous study which is (24, 48, 72 and 96) where the depth of the dense net is set to 8, 8, 8, 8 and number of stacked blocks is 2.

4.2.2.2 False Positive Reduction CNNT5 Classifier

The CNN classifier model for False Positive Reduction component uses TensorFlow deep learning framework which was adopted from the work on lung nodule classification by (Dobko, M., 2019) where the author adapted the model into 3D input. The architecture for the current model was kept simple to fulfil the purpose of running test on the result from lung nodule localization. The model consists of convolutional layer with kernel size of $5 \times 5 \times 5$ with 16 number of filters followed by max pooling with pooling layer of 2 and strides of 2. The layer is then stacked twice and further flatten to feed in the fully connected layer in this case dense layers for classification. In the dense connection, the first layer filled up with 150, 100, 50 and 2 layers sequentially along the layers, dropout layers were implemented in between the dense layers to prevent overfitting with drop rate of 0.3 while the activation function used in the first three layers were ReLU and the last was softmax. The details of the training setup will be derived in the next section.

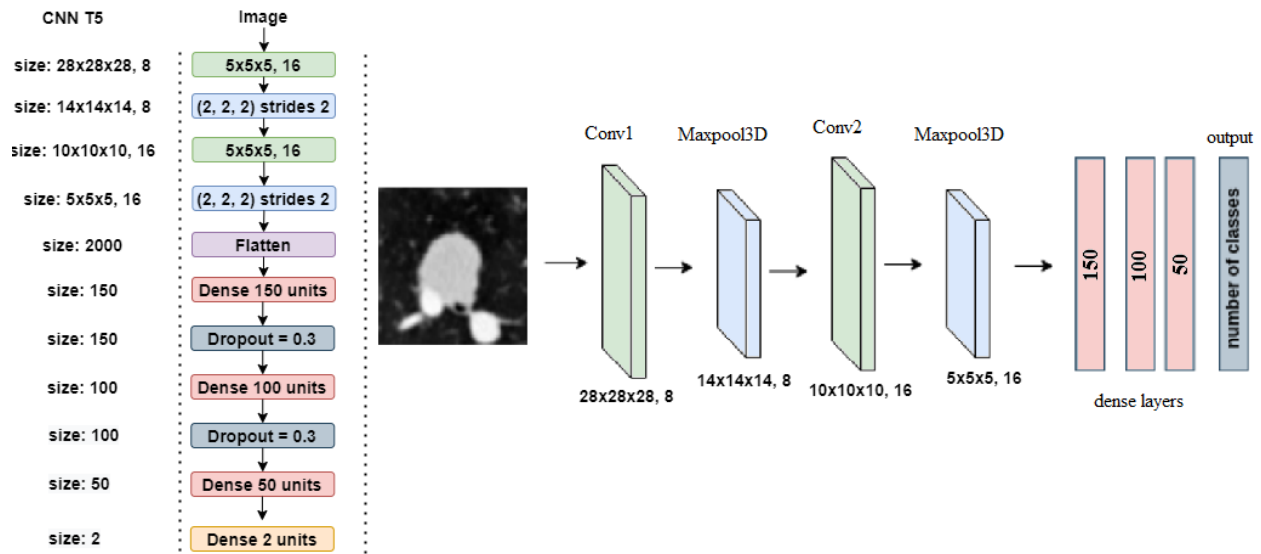


Fig 4.7: CNNT5 Classifier model for False Positive Reduction.

Table 4.3: CNNT5 Model Adaptation from (Dobko, M., 2019)

Stage	Output	No of Filter	Kernel size	Activation Function
Convolution 3D	$28 \times 28 \times 28, 8$	8	$5 \times 5 \times 5$	ReLU
Max Pooling 3D	$14 \times 14 \times 14, 8$	-	$2 \times 2 \times 2$ (<i>stride of 2</i>)	-
Convolution 3D	$10 \times 10 \times 10, 16$	16	$5 \times 5 \times 5$	ReLU
Max Pooling 3D	$5 \times 5 \times 5, 16$	-	$2 \times 2 \times 2$ (<i>stride of 2</i>)	-
Flatten	2000	-	Flatten	-
Dense layer 1	150	-	150	ReLU
Dropout 1	0.3	-	0.3	-
Dense layer 2	100	-	100	ReLU
Dropout 2	0.3	-	0.3	-
Dense layer 3	50	-	50	ReLU
Dense layer 4	2	-	2	Softmax

4.2.3 Training and Testing

In this section, training and testing operations will be explained on the setup phases in terms of the flow of the routine and parameters used. The training and testing operation were done independently by using both PyTorch and TensorFlow. The flow of the training and testing can be illustrated in figure 4.8.

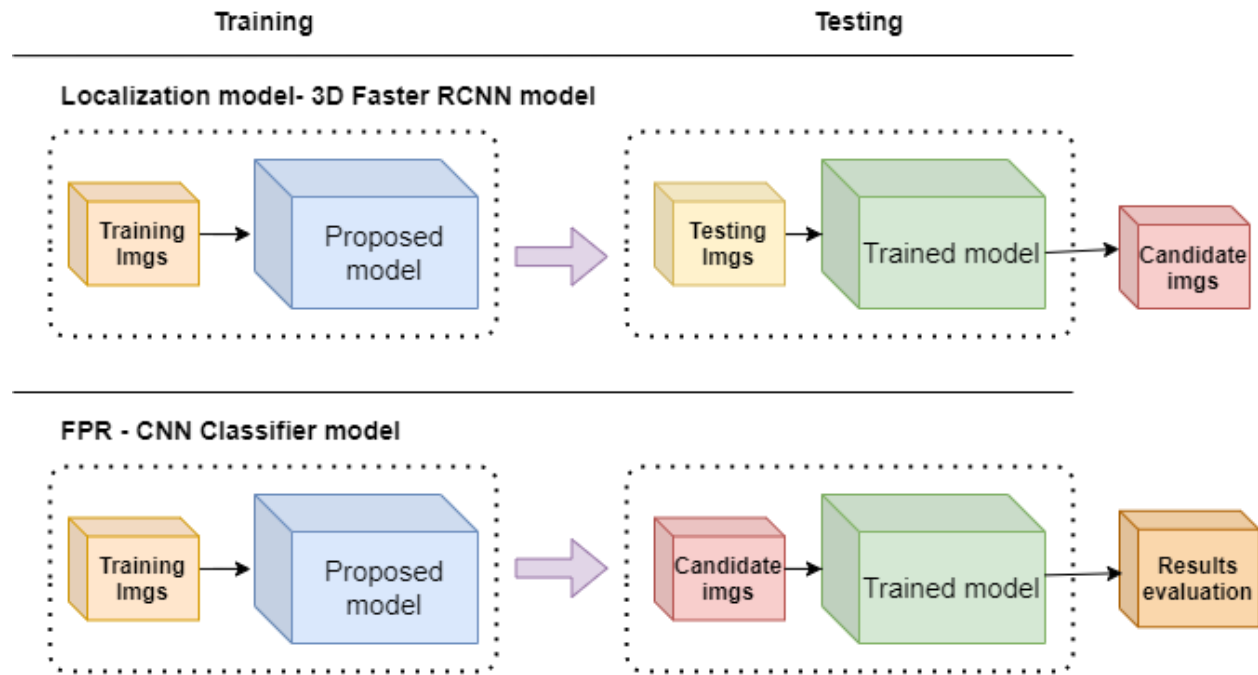


Fig 4.8 Training and Testing Routine.

The figure 4.8 shows the training and testing routine where the training of lung nodule detection and classification was performed independently at the same time. The testing operation however requires the output from the trained localization model in order to test the trained lung classifier model. Therefore, lung nodule localization carries a higher priority than the false positive reduction model, besides the classifier model has a simpler architecture than localization model which in turn takes a shorter duration for testing.

The dataset used in the training for lung nodule localization model was kept to only 4 out of 10 subsets due to the time-resource and computation constraints, with ratio of 80:20 which translates to 264 scans for training and 88 scans for validation. The training input images was preprocessed into patches of 96 x 96 x 96 and set batch size to 1 in accordance with the limited GPU memory, though generally lower batch size would result in a poor performance, in this case,

it was considered as a hardware limitation. In the training, for each model, the training uses 20 epochs in total with stochastic gradient descent optimization with momentum of 0.9 in the accordance with the work by (Zhu, W., et al., 2017). The weight decay was set to 1×10^{-4} and the learning rate scheduler was set with a default learning rate of 0.01, 0.001 after 13 epochs and 5×10^{-4} after 16 epochs.

The dataset used in training the lung nodule classifier for false positive reduction components uses a 3D patch of 32, 32, 32 with half distribution of positive and negative sample from the 10 subsets of preprocessed lung CT scan with masks generated. The model was trained with 25 epochs and uses 128 batch size and RMSprop with learning rate of 1×10^{-4} and momentum of 0.95. The training results was compared to the validation performance and the result appears to be decently acceptable whereby there was no sign of overfitting, yet accuracy was not overly accurate. The outcome of the performance was expected whereby simple model architecture was generating decently stable performance.

4.2.3 Results and Evaluation

The evaluation method employed for the obtained result will use the classification metrics with confusion matrix, accuracy, F1 score, and precision. The second evaluation method will be the plot of training logs where accuracy and loss during training will be visualized and by observation the model will be evaluate either it is overfit, underfit or fit.

The comparison loss plot will be visualized using TensorBoard visualization module on training and validation for lung nodule localization model. The comparison of loss c Reason behind using a so-called slightly inappropriate validation method was due to the limitation of timely-resource and computation prowess which limits the training setup especially the number of epochs and number of subsets in dataset. The poor training setup causes the performance of lung nodule localization to underperform compared to the expected result whereby large size of Probability Bounding Box (PBB) was omitted rather than a smaller size. This occurs due to the low performance of model training causes the weights to be improperly trained.

4.3 Results

In this section the performance results and evaluation of the testing model will be discussed for lung nodule detection localization model where comparison performance between 3D RPN and proposed model will be laid out. In addition, the second part will cover the lung nodule classification performance result and it's evaluation.

4.3.1 3D DPN and 3DPN_SE

As per mentioned in the earlier section the appropriate validation method for the lung nodule localization will be the visualization of the training performance between 3D DPN act as the baseline model and proposed model using TensorBoard on training loss for training and validation set. The metrics that will be compared will be the results of 2 losses in the model namely, bounding box regression loss and classification loss, in addition for validating the early losses, bounding box loss.

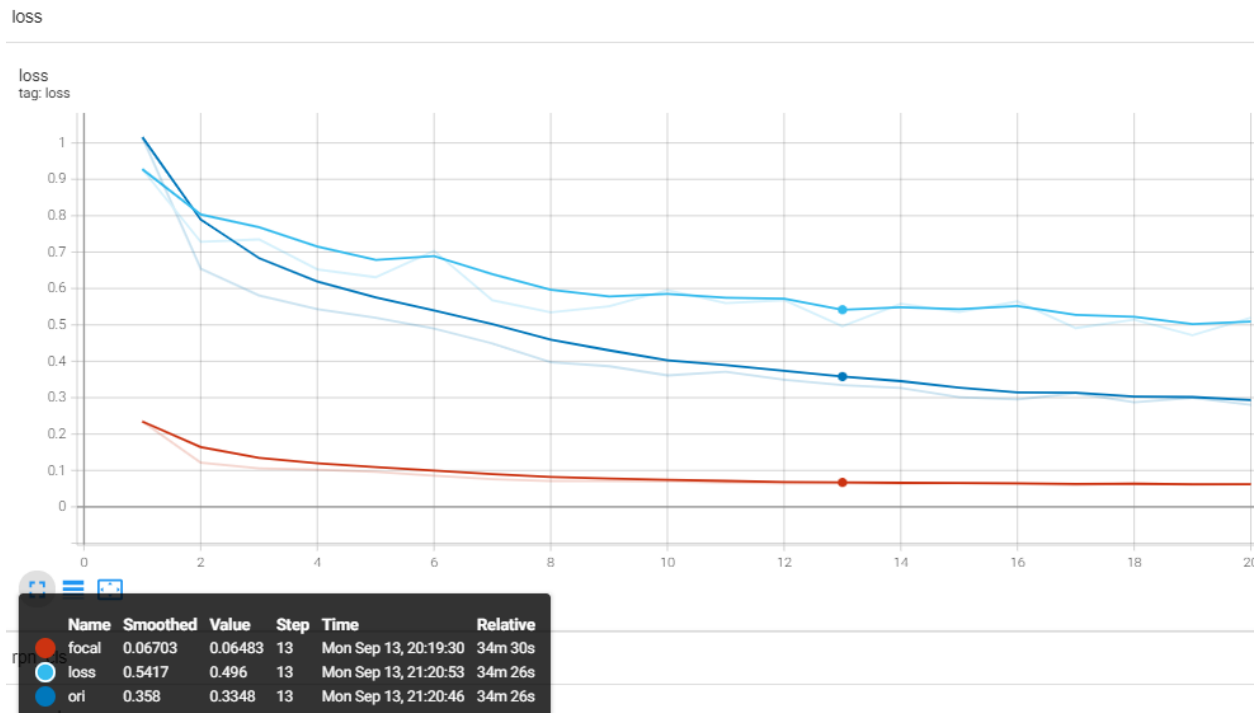


Fig 4.9: Regression Loss comparison between 3 models.

Figure 4.9 shows the results of the training session with indications of training and validation with brighter color and faded color respectively. There are 3 models trained during the

training session, 1) focal indicates the proposed model with focal loss with red color code, 2) loss indicate the proposed model with origin multi-task loss with lighter blue color code, and 3) the origin baseline model with multi-task loss with darker blue color code. Based on observation of each model training and validation results, shows no sign of overfitting nor underfitting, the result may look like an ideal case however the loss coefficient for proposed with multi-task loss and the baseline model remain prominent in the range of 0.48 – 0.59 unit. On the other hand, the proposed model with the focal loss does look unexpected whereby the convergence of loss was earlier, and the loss coefficient was lower than both other models. In comparison with the other two models, the difference is the focal loss adaptation with the “loss model” and both on loss function and SE layer to the baseline model. This raises the question of loss result validity whereby – does the focal loss truly made a huge impact ?, therefore, by reference to the prediction of bounding box location coefficient with four notations (x, y, w, h) represent coordinates of top left corner, width, and height.

The obtained results were visualized in figure 4.10 – 4.13 which shows the 4-coefficient loss graph plotted against the 3 models. Based on observations, the results on the 4-coefficient loss were aligned with the earlier regression whereby the result on the proposed model with multi-task loss evidently shows a poor performance, whereas the loss function on proposed model is on par with the baseline model.

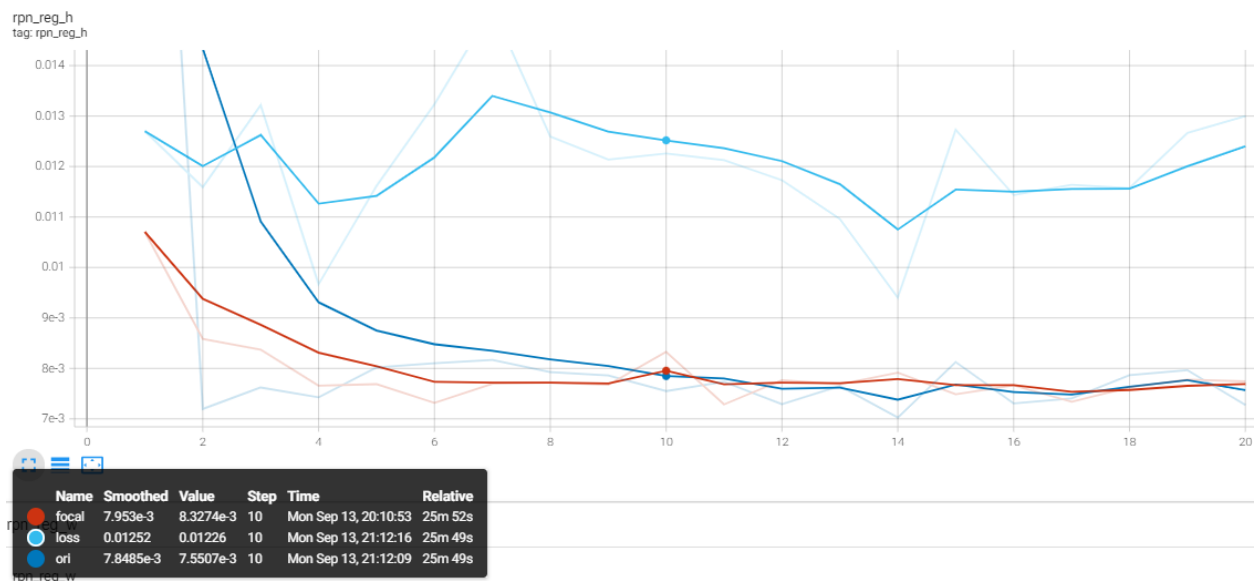


Fig 4.10: Height Coefficient Prediction Loss.

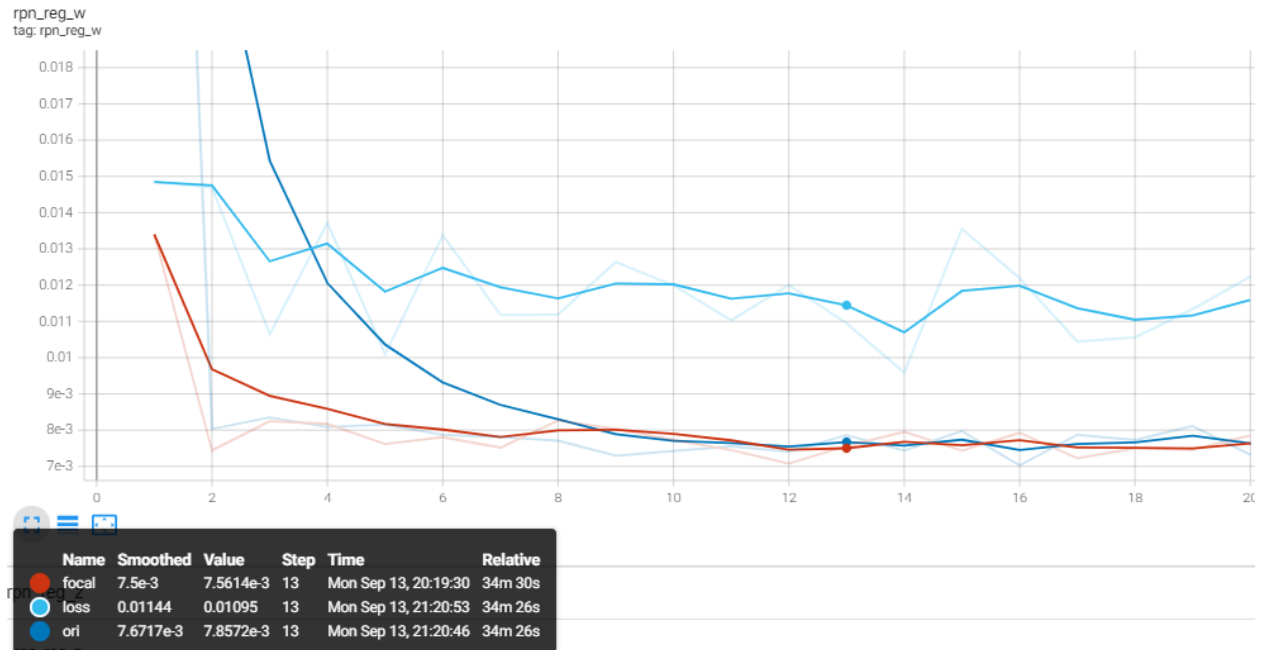


Fig 4.11: Width Coefficient Prediction Loss.

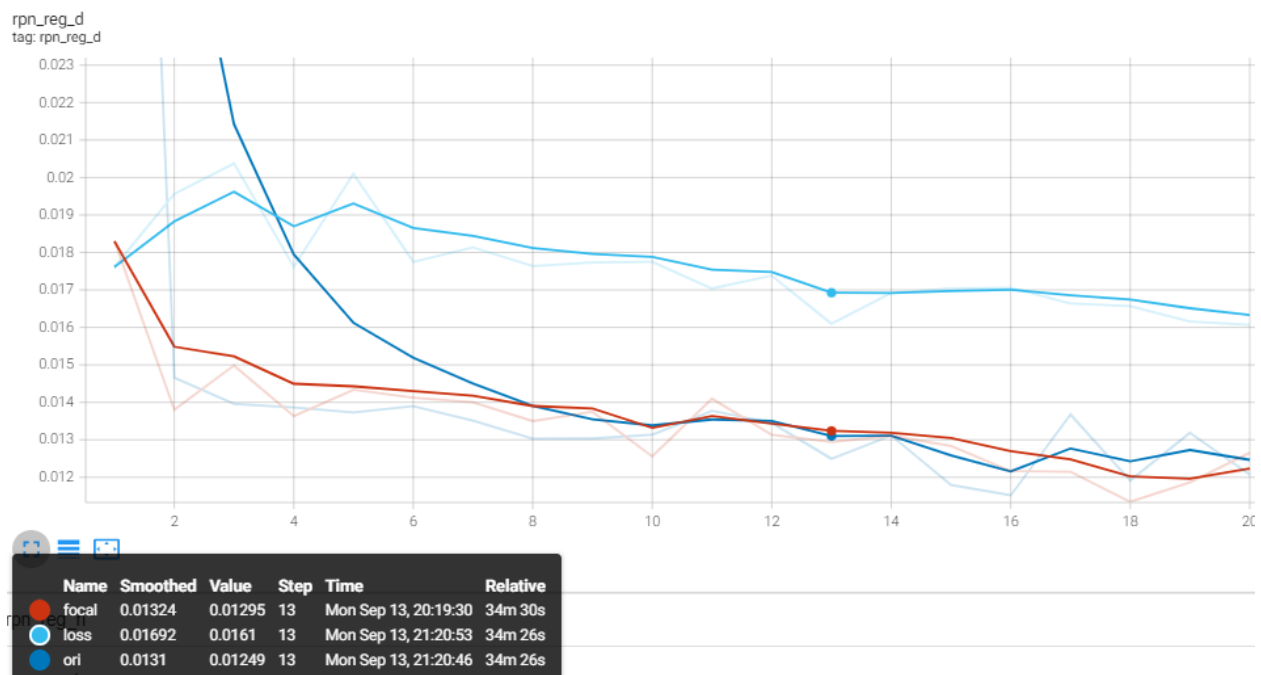


Fig 4.12: Top Left Corner – X Coordinate Coefficient Prediction Loss.

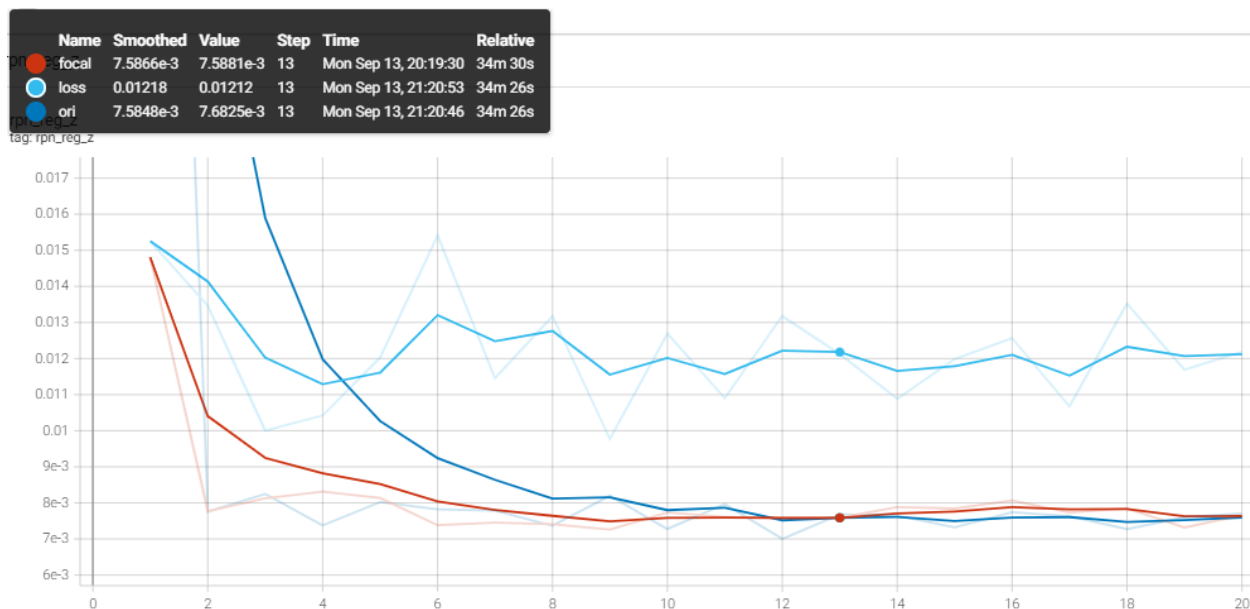


Fig 4.13: Top Left Corner – Y Coordinate Coefficient Prediction Loss.

Based on observation, the trend of the curves in baseline model and proposed with focal loss shares a similar contour except for the proposed model with multi-task loss, this insight fuel the question whether is it truly that the proposed model behaves with earlier loss convergence compared to the baseline model and perhaps it may give a promising development with the current setup. Unfortunately, the scope of this project has a limited time-resource and restricted computation prowess, leaving the investigation in an inconclusive manner due to inability in providing assertive proof. Despite the current circumstances, the finding on proposed lung nodule detection on localization component, the Squeeze and excitation layer with focal loss does indeed improve the model however fine-tuning processes is strictly required with appropriate training setup such as proper epoch training which dependent on dataset (Afaq, S. and Rao, S., 2020), and higher batch size would be recommended to give a better performance result.

As per mentioned earlier, the current approach was short-handed with the resource which causes the current deficiency on the obtained result. Clearly, that the proposed model pointing towards the better performance however further validation is required, especially the need to benchmarking with appropriate training setup. Other than local quantitative validation method, the findings can be cross-checked with relevant literature which will be discussed in the next section discrepancy on literature review.

4.3.1 3D CNN Classifier

In the second component of lung nodule detection, results on False Positive Reduction component will be evaluated on two evaluation method, 1) the accuracy metrics with confusion matrix along with the sensitivity and precision, 2) training and validation loss curve to evaluate on the case of training results – overfit, underfit, or fit.

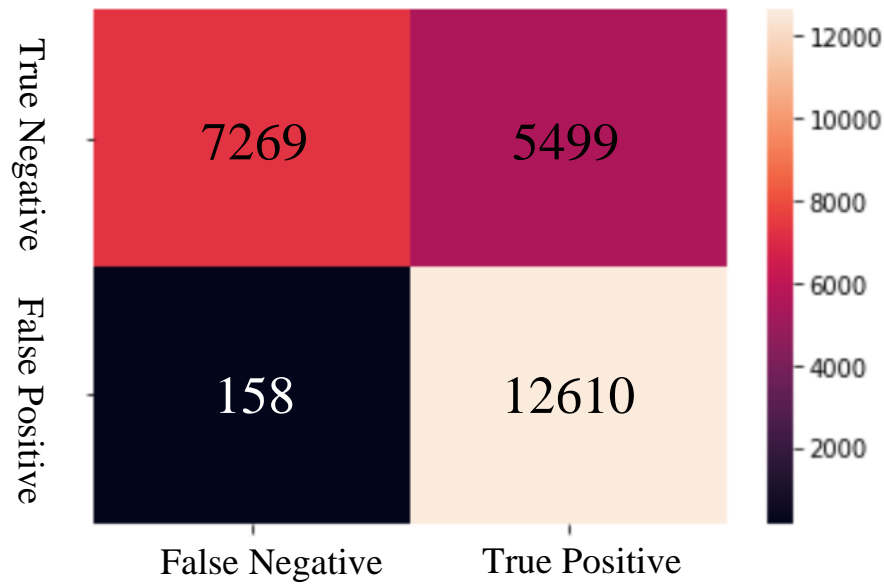


Fig 4.14: Confusion Matrix.

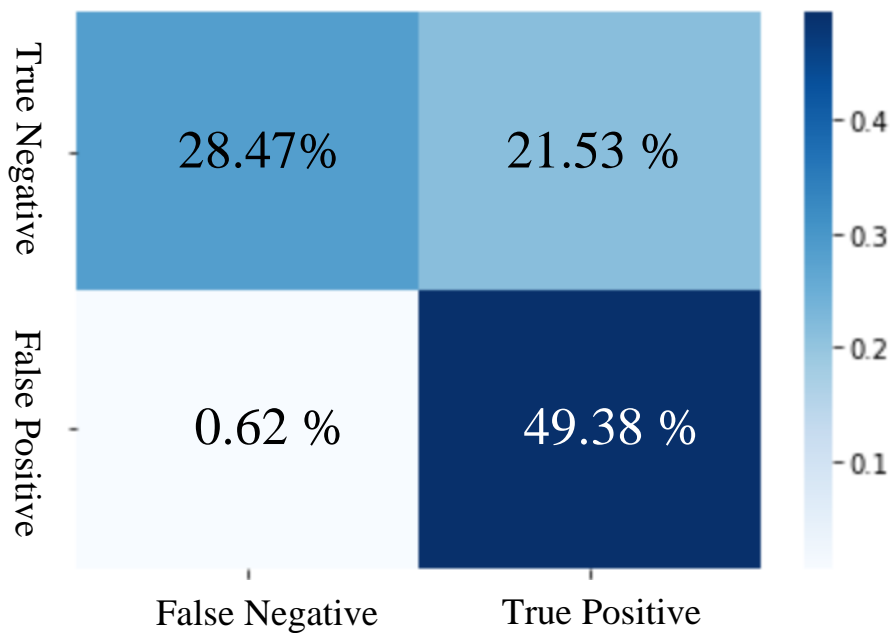


Fig 4.15: Confusion Matrix in percentage.

The figure 4.13 and 4.14 shows the confusion matrix that was done on the 1-fold validation, in the table 4.3 shows the six-folds cross-validation results where the model would be validated six times and averaged the performance result.

Table 4.4: Six-fold cross-validation results.

Fold	TP	FP	TN	FN	Accuracy	Sensitivity	Precision
1	12610	5499	7269	158	77.84%	98.76%	69.63%
2	12611	5498	7271	156	77.85%	98.77%	69.63%
3	12612	5497	7271	156	77.86%	98.77%	69.64%
4	12617	5494	7270	157	77.87%	98.77%	69.64%
5	12616	5495	7270	157	77.86%	98.77%	69.65%
6	12615	5496	7271	156	77.86%	98.77%	69.65%
Average	12614	5497	7270	157	77.85%	98.76%	69.64%

Table 4.3 shows the six-fold cross-validation results whereby the model was trained using the training was the half portion of the LUNA 16 subsets and validated on 20% portion of it. The model accuracy is rather low, and it was expected due to the simple model-built architecture explained in the earlier chapter.

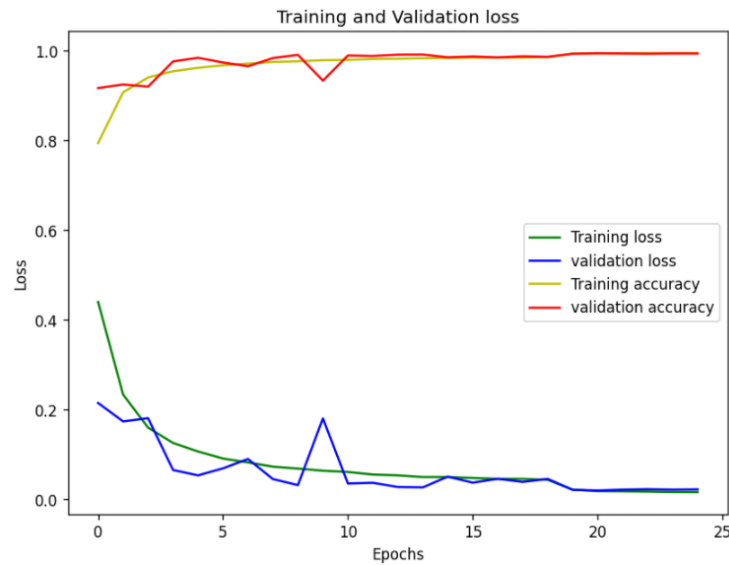


Fig 4.16: CNNT5 Model Training and Validation Loss Performance.

The figure 4.14 shows the model training and validation loss performance whereby based on observations, the result shows no sign of overfitting however converging close to each other starting at the epoch 15. The performance of the model is deemed acceptable and expected due to the simple mode of architecture with few blocks of convolutional layer and with the help of dropout layers resulting in the prevention of the model being overfit. Undeniably the classification mode could be further improved by using the most advanced existing model, however in this project, lung nodule classification was not the focused development.

4.4 Project Timeline

The plan of this research consists of several tasks within an objective, which laid out with Gantt chart in figure 4.17 covering activities from starting date of 17th May 2021 to submission on 17th September 2021. The research plan spans almost 4 months duration of about 16 weeks strictly speaking. The plan of this research covers 4 objectives namely: system hardware setup, building model 2, model 3, and final documentation & model adjustment. The first objective covers the initiation of cloud service, dataset cleaning, first model built and first test run, with a span of 31 days. Second objective covers the building of second model with the fine tuning of the first for the purpose of performance comparison, with a span of 21 days.

Milestone description	Category	Assigned to	Progress	Start	Days	Milestone description	Category	Assigned to	Progress	Start	Days
System hardware setup						2nd model					
Cloud service setup	Med Risk	Cloud service setup	1%	17-05-21	3	2nd model built	High Risk	coding	13%	17-06-21	14
Dataset setup	Med Risk	pre-processing	2%	20-05-21	3	2nd model training	Low Risk	training	16%	01-07-21	2
First model built	High Risk	coding	5%	23-05-21	21	fine tune first model	Low Risk	fine tuning	18%	03-07-21	2
first model training	Low Risk	training	8%	13-06-21	3	2nd test run	Milestone	secondday test	20%	05-07-21	1
test run	Milestone	preliminary test	10%	16-06-21	1	fine tune second model	On Track	fine tuning	22%	06-07-21	2

Fig 4.17 Gantt chart research plan 1 of 2.

the third objective covers similar aspects to second objective including the milestone with progress show case with supervisor roughly at the middle of the semester 25th July 2021 and report writing. The last objective would be the final touch up on documentation and final model adjustment, in

addition, final meetings may occur in the last few days before submission for the final checking. The third and last objective span about 23 and 35 days, respectively.

Milestone description	Category	Assigned to	Progress	Start	Days	Milestone description	Category	Assigned to	Progress	Start	Days
3rd model						documentation					
3rd model built	Med Risk	coding	26%	08-07-21	14	report writing chapter 3	Low Risk	documentation	46%	31-07-21	5
3rd model training	Low Risk	training	30%	22-07-21	2	fine tune on model 1	Low Risk	coding	50%	05-08-21	5
3rd test run	Milestone	tertiary test	34%	24-07-21	1	report writing chapter 4	Low Risk	documentation	54%	10-08-21	5
supervisor show case	Milestone	present	40%	25-07-21	1	report writing chapter 5	Low Risk	documentation	58%	15-08-21	5
report writing chapter 2	Low Risk	documentation	42%	26-07-21	5	Final model + final documentation	On Track	documentation	70%	20-08-21	14
						submission	Goal	documentation	100%	03-09-21	1

Fig 4.18 Gantt chart research plan 2 of 2.

Based on the allocated days on the Gantt chart, the longest span of time in an objective would be the building of the model which includes the pre and post process of adjustment, error troubleshooting, fine tuning opportunities, additional research, and result data collection. The idea of building additional model serves the purpose of comparison of model which may in the case of state-of-the-art model. Again, tailoring of pre and post adjustment may be made to align consistency of test condition. The full Gantt chart can be referred in the next 2 pages.

4.5 Summary

Based on the obtained results, the experiment concludes that the lung nodule localization with 3D DPN network architecture in addition on SE layers and the adaptation of focal loss does help in enhancing the performance result however, result obtained were from a handicapped training setups may result in discrepancy. Despite the short-handed circumstances, the comparison between models were kept cautious manner where training setups were strictly controlled to make an equivalent contrast. The training setups were 20 epochs with 1 batch size over 4 out of 10 subsets of LUNA16 dataset and trained on three different model approaches: 3D DPN baseline, 3D DPN_SE with multi-task loss, and 3 DPN baseline with focal loss. The insights concluded that the SE layer and focal loss made a drastic impact in which trend was seen to be align however further investigation is needed to provide a more assertive proof.

In the lung nodule classification – False Positive Reduction component performs with the expected result due to the simple building of architecture. The lung nodule achieved 77% accuracy trained with half portion of the LUNA 16 dataset and validated with 20% portion of the half. The second evaluation on the model was the training session where accuracy and loss were plot against training epoch. There was no sign of overfitting however the model validation loss was close enough with the training curve.

This conclude the results and evaluation section whereby model was built and evaluated with essential evaluation method. Although the result may not be supportive enough to claim that the project has been successful, the result shows that the model is still far from flawless.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion to the Research Objectives

In a nutshell, the listed objectives had been achieved with respect to the expected output results of the research. The first objective relates to the identification/ evaluation of different approach of lung nodule detection system using Faster RCNN model with DPN and DPN SE network. The in-depth literature review helps in understanding the framework and finally selects appropriate approach with the model architecture, Faster RCNN.

The second objective of the project relates to the designing and development of the lung nodule detection system based on previous objective. For this objective, the model of the lung nodule detection was dissected in chapter three and four of which the system comprises of two separate model: 1) localization and 2) False Positive Reduction. The model architecture, loss function, and optimizer were carefully considered for both lung nodule detection and classification model.

The final objective of the project relates to the performance evaluation on the developed system. The system was appraised by reviewing the number of epochs performed and the performance of the detection model. The compared detection models were to provide localization feature and performance was compared with the loss function convergence rate of training session between proposed and baseline.

5.2 Limitations

The limitations faced in this project was not able to perform a complete experiment on all sets of the dataset in comparison to other research works, which turns the work to only be comparable own approach. The complete experiment on all sets would allows comparison on other similar research work.

Another limitation of the work was the adoption of the model in terms of training time, since the model was developed separately referring to the lung nodule localization and FPR,

training must be done independently. During the training for lung nodule localization – faster RCNN model took at least 3 hours for 20 epochs, while based on the guidelines of previous work, number epochs run time should be at least 200 epochs with duration rough estimate about 3 days per 1 attempt model. Hence, in this work model was trained under restrained configuration and kept in a controlled parameter. Other limitation of the system adoption, pre-processing of images must be perform prior feeding to the models in order to train and predict. Therefore, data collection and data pre-processing took the most of testing and trials as it must fit and transform the whole batches of dataset. In the current work duration for data acquisition and preprocessing took at least 2 hours per subset out of 10 subsets from LUNA 16. In terms of data storage for the dataset used and pre-processed, took at least 125 GB for the pure dataset and 250 for pre-processed and augmented dataset. Besides storage, losses of data during download and upload may occur depending on internet connections especially on the first failed attempt on google drive and second success attempt using AWS Sagemaker notebook platform.

5.3 Recommendations and Suggestions for future work

Although the deep learning lung nodule detection system for this research achieved it's aim and objectives, there still room for improvement and enhancement. The following sections below are some of suggestions on how the system can be further enhanced for future work.

5.3.1 Decoupled False Positive Reduction

The current model approach was developed independently referring to the training and testing of the models which is due to the nature of the strategy in dealing with lung nodule detection. The current strategy approach is by developing lung nodule localization and false positive reduction (FPR) which in turns using Faster R-CNN object detection models and a separate CNN classifier model. The suggestion for future work involves related topic on decoupled false positive reduction, the work by (Tang, H., Zhang, C. and Xie, X., 2019) performs a similar work using Nodule Net and demonstrated a study about training a separate models will result in a suboptimal performance, therefore the author built both model on lung nodule localization and false positive reduction on training on the same time and shares some of the early feature map for classifier training. This approach could possibly improve the performance of the model, where the approach of decoupled

false positive reduction promotes efficient built and training time, ultimately improving lung nodule detection model.

5.3.2 Attention module from Transformer model

The recent new model creation rhymes with the use of attention module based on transformer model, formerly developed for the purpose of machine translation (Vaswani, et al., 2017). The current work can be further enhanced by focusing on the global-local context for either feature extraction or segmentation task. Besides, the current model only covers detection and classification of lung nodule, the current system can be further developed with additional segmentation on the lung nodule bounding box. Doing so will open a new venue for classification in terms of lung nodule type output rather than benign or malignant output. Coming back to the attention module, the module may be further explored on different parts of the lung nodule components which then may result in enhancing the performance.

5.4 Summary

In this study, the aim of this research was to develop and implement a deep learning lung nodule detection. To achieve the aim of the study, three objectives must be met whereby lung nodule detection algorithm were evaluated, develop, and design a simulation of a lung nodule detection and to assess the performance between existing models. In depth literature studies were reviewed while techniques and suggestions were applied to the current work. The aim and objectives were achieved and provided with appropriate justification. Results were attained with two different testing namely, FROC analysis and CAD nodule analysis evaluation based on LUNA 16 challenge. Based on the results obtained, the study concluded the inclusion of the Squeeze and Excitation module which promotes the attention/selective weight of feature maps does improve the feature extraction and consequently enhanced the region proposal network (RPN) in generating candidate proposal. In addition to the study, the implementation of focal loss does improve the training time to reach stable plateau without overfit while in contradictory to multi-task loss.

Reference

- Afaq, S. and Rao, S., 2020. Significance Of Epochs on Training A Neural Network. *International Journal of Scientific and Technology Research*, 19(6), pp.485-488.
- Armato 3rd, S.G., McNitt-Gray, M.F., Reeves, A.P., Meyer, C.R., McLennan, G., Aberle, D. R., et al., 2007. The Lung Image Database Consortium (LIDC): an evaluation of radiologist variability in the identification of lung nodules on CT scans. *Acad. Radiol.* 14, 1409–1421.
- Banik, D. and Bhattacharjee, D., 2021. Mitigating Data Imbalance Issues in Medical Image Analysis. *Data Preprocessing, Active Learning, and Cost Perceptive Approaches for Resolving Data Imbalance*, pp.66-89.
- Cao, H., Liu, H., Song, E., Hung, C., Ma, G., Xu, X., Jin, R. and Lu, J., 2020. Dual-branch residual network for lung nodule segmentation. *Applied Soft Computing*, 86, pp.1-11.
- Cao, G., Huang, T., Hou, K., Cao, W., Liu, P. and Zhang, J., 2018. 3D Convolutional Neural Networks Fusion Model for Lung Nodule Detection on Clinical CT scans. *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*,.
- Chen, Y., Li, J., Xiao, H., Jin, X., Yan, S. and Feng, J., 2017. Dual path networks. *arXiv preprint arXiv:1707.01629*.
- Chen, L., Gu, D., Chen, Y., Shao, Y., Cao, X., Liu, G., Gao, Y., Wang, Q. and Shen, D., 2021. An artificial-intelligence lung imaging analysis system (ALIAS) for population-based nodule computing in CT scans. *Computerized Medical Imaging and Graphics*, 89, p.101899.
- Dobko, M., 2019. *Lung Nodule Detection in Computed Tomography Scans Using Deep Learning* (Doctoral dissertation, Ukrainian Catholic University).

El-Askary, N., Salem, M. and Roushdy, M., 2019. Feature Extraction and Analysis for Lung Nodule Classification using Random Forest. *Proceedings of the 2019 8th International Conference on Software and Information Engineering*.

Essaf, F., Li, Y., Sakho, S. and Gadosey, P., 2020. Improved Convolutional Neural Network for Lung Cancer Detection. *Proceedings of the 2020 International Conference on Computing, Networks, and Internet of Things*.

EL-Bana, S., Al-Kabbany, A. and Sharkas, M., 2020. A Two-Stage Framework for Automated Malignant Pulmonary Nodule Detection in CT scans. *Diagnostics*, 10(3), p.131.

GitHub. 2021. *GitHub - s-mostafa-a/Luna16: Developing a well-documented repository for the Lung Nodule Detection task on the Luna16 dataset."grt123"*. [online] Available at: <<https://github.com/s-mostafa-a/Luna16>> [Accessed 16 September 2021].

Gong, J., Liu, J., Wang, L., Zheng, B. and Nie, S., 2016. Computer-aided detection of pulmonary nodules using dynamic self-adaptive template matching and a FLDA classifier. *Physica Medica*, 32(12), pp.1502-1509.

Halder, A., Chatterjee, S., Dey, D., Kole, S. and Munshi, S., 2020. An adaptive morphology-based segmentation technique for lung nodule detection in thoracic CT image. *Computer Methods and Programs in Biomedicine*, 197, p.105720.

Hu, J., Shen, L. and Sun, G., 2018. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).

Kumar, S. and Latte, M., 2017. Modified and Optimized Method for Segmenting Pulmonary Parenchyma in CT Lung Images, Based on Fractional Calculus and Natural Selection. *Journal of Intelligent Systems*, 28(5), pp.721-732.

Kalinovsky, A., Kovalev, V., 2016. Lung image segmentation using deep learning methods and convolutional neural networks.

Liao, F., Liang, M., Li, Z., Hu, X. and Song, S., 2019. Evaluate the malignancy of pulmonary nodules using the 3-d deep leaky noisy-or network. *IEEE transactions on neural networks and learning systems*, 30(11), pp.3484-3495.

Li, Y. and Fan, Y., 2020, April. DeepSEED: 3D squeeze-and-excitation encoder-decoder convolutional neural networks for pulmonary nodule detection. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (pp. 1866-1869). IEEE.

Liu, C., Hu, S., Wang, C., Lafata, K. and Yin, F., 2020. Automatic detection of pulmonary nodules on CT images with YOLOv3: development and evaluation using simulated and patient data. *Quantitative Imaging in Medicine and Surgery*, 10(10), pp.1917-1929.

Lung cancer statistics, World Cancer Research Fund, 2018. [Online]. Available: <https://www.wcrf.org/dietandcancer/cancer-trends/lung-cancer-statistics>. [Accessed: 16-Mar-2018].

Ma, J., Song, Y., Tian, X., Hua, Y., Zhang, R. and Wu, J., 2019. Survey on deep learning for pulmonary medical imaging. *Frontiers of Medicine*, 14(4), pp.450-469.

Materialise. 2021. *What is the difference between slice thickness and slice increment?*. [online] Available at: <<https://www.materialise.com/en/faq/what-difference-between-slice-thickness-and-slice-increment>> [Accessed 16 September 2021].

Pereira, S., Pinto, A., Alves, V., Silva, C.A., 2016: Brain tumor segmentation using convolutional neural networks in mri images. *IEEE transactions on medical imaging* 35(5), 1240–1251.

Pytorch.org. 2021. *Datasets & DataLoaders — PyTorch Tutorials 1.9.0+cu102 documentation*. [online] Available at: <https://pytorch.org/tutorials/beginner/basics/data_tutorial.html> [Accessed 16 September 2021].

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, pp.91-99.

Ross, T.Y. and Dollár, G.K.H.P., 2017, July. Focal loss for dense object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2980-2988).

Setio, A., Traverso, A., de Bel, T., Berens, M., Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M., Geurts, B., Gugten, R., Heng, P., Jansen, B., de Kaste, M., Kotov, V., Lin, J., Manders, J., Sónora-Mengana, A., García-Naranjo, J., Papavasileiou, E., Prokop, M., Saletta, M., Schaefer-Prokop, C., Scholten, E., Scholten, L., Snoeren, M., Torres, E., Vandemeulebroucke, J., Walasek, N., Zuidhof, G., Ginneken, B. and Jacobs, C., 2017. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Medical Image Analysis*, 42, pp.1-13.

Slicer.org. 2021. *Coordinate systems - Slicer Wiki*. [online] Available at: <https://www.slicer.org/wiki/Coordinate_systems#Image_coordinate_system> [Accessed 16 September 2021].

Safarov, S. and Whangbo, T., 2021. A-DenseUNet: Adaptive Densely Connected UNet for Polyp Segmentation in Colonoscopy Images with Atrous Convolution. *Sensors*, 21(4), p.1441.

Tang, H., Zhang, C. and Xie, X., 2019, October. Nodulenet: Decoupled false positive reduction for pulmonary nodule detection and segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 266-274). Springer, Cham.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008).

Zhang, G., Yang, Z., Gong, L., Jiang, S., Wang, L. and Zhang, H., 2020. Classification of lung nodules based on CT images using squeeze-and-excitation network and aggregated residual transformations. *La radiologia medica*, 125(4), pp.374-383.

Zhu, W., Liu, C., Fan, W. and Xie, X., 2018, March. Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 673-681). IEEE.

Appendices

Appendix A: Ethics Form

Office Record

Date Received:

Received by whom:

Receipt – Fast-Track Ethical Approval

Student name: Winarto Wijaya

Student number: TP041218

Received by:

Date:

APU / APIIT FAST-TRACK ETHICAL APPROVAL FORM (STUDENTS)

Tick one box (level of study):

- ☒ POSTGRADUATE (PhD / MPhil / Masters)
- ☐ UNDERGRADUATE (Bachelors degree)
- ☐ FOUNDATION / DIPLOMA / Other categories

Tick one box (purpose of approval):

- ☒ Thesis / Dissertation / FYP project
- ☐ Module assignment
- ☐ Other: _____

Title of Programme on which enrolled

Tick one box: ☒ Full-Time Study or ☐ Part-Time Study

Title of project / assignment: Lung Nodule Detection Using Deep Learning model

Name of student researcher: Winarto Wijaya... ..

Name of supervisor / lecturer: Dr. V. Sivakumar

Student Researchers- please note that certain professional organisations have ethical guidelines that you may need to consult when completing this form.

Supervisors/Module Lecturers - please seek guidance from the Chair of the APU Research Ethics Committee if you are uncertain about any ethical issue arising from this application.

		YES	NO	N/A
1	Will you describe the main procedures to participants in advance, so that they are informed about what to expect?		<input checked="" type="checkbox"/>	
2	Will you tell participants that their participation is voluntary?		<input checked="" type="checkbox"/>	
3	Will you obtain written consent for participation?		<input checked="" type="checkbox"/>	
4	If the research is observational, will you ask participants for their consent to being observed?		<input checked="" type="checkbox"/>	

5	Will you tell participants that they may withdraw from the research at any time and for any reason?		✓	
6	With questionnaires and interviews will you give participants the option of omitting questions they do not want to answer?		✓	
7	Will you tell participants that their data will be treated with full confidentiality and that, if published, it will not be identifiable as theirs?		✓	
8	Will you give participants the opportunity to be debriefed i.e. to find out more about the study and its results?		✓	

If you have ticked **No** to any of Q1-8 you should complete the full Ethics Approval Form.

		YES	NO	N/A
9	Will your project/assignment deliberately mislead participants in any way?		✓	
10	Is there any realistic risk of any participants experiencing either physical or psychological distress or discomfort?		✓	
11	Is the nature of the research such that contentious or sensitive issues might be involved?		✓	

If you have ticked **Yes** to 9, 10 or 11 you should complete the full Ethics Approval Form. In relation to question 10 this should include details of what you will tell participants to do if they should experience any problems (e.g. who they can contact for help). You may also need to consider risk assessment issues.

		YES	NO	N/A
12	Does your project/assignment involve work with animals?		✓	
13	Do participants fall into any of the following special groups? Note that you may also need to obtain satisfactory clearance from the relevant authorities		✓	



Signed..... Print Name..... Date.....

(Student Researcher)



Please note that any variation to that contained within this document that in any way affects ethical issues of the stated research requires the appending of new ethical details. New ethical consent may need to be sought.

The completed form (and any attachments) should be submitted for consideration by your Supervisor/Module Lecturer

SUPERVISOR/MODULE LECTURER

PLEASE CONFIRM THE FOLLOWING:

**Please Tick
Box**

I consider that this project/assignment has no significant ethical implications requiring a full ethics submission to the APU Research Ethics Committee	
i) I have checked and approved the key documents required for this proposal (e.g. consent form, information sheet, questionnaire, interview schedule)	
Or	
ii) I have checked and approved draft documents required for this proposal which provide a basis for the preliminary investigations which will inform the main research study. I have informed the student researcher that finalised and additional documents (e.g. consent form, information sheet, questionnaire, interview schedule) must be submitted for approval by me before they are used for primary data collection.	

SUPERVISOR AND SECOND ACADEMIC SIGNATORY

STATEMENT OF ETHICAL APPROVAL (please delete as appropriate)

1) THIS PROJECT/ASSIGNMENT HAS BEEN CONSIDERED USING AGREED APU/SU PROCEDURES AND IS NOW APPROVED

2) THIS PROJECT/ASSIGNMENT HAS BEEN APPROVED IN PRINCIPLE AS INVOLVING NO SIGNIFICANT ETHICAL IMPLICATIONS, BUT FINAL APPROVAL FOR DATA COLLECTION IS SUBJECT TO THE SUBMISSION OF KEY DOCUMENTS FOR APPROVAL BY SUPERVISOR (see Appendix A)

Signed... *V. Sivakumar* ...
(Supervisor/Lecturer)

Print Name... DR. V. SIVAKUMAR ... Date 16 September 2021

Signed...  ...
(Second Academic Signatory)

Print Name TS. DR. SHANKAR DURAIKANNAN Date 16 September 2021

Office Record	Receipt – Appendix A (Fast-Track Ethics Form)
Date Received:	Student name:
Received by whom:	Student number:
	Received by:
	Date:

APPENDIX A

AUTHORISATION FOR USE OF KEY DOCUMENTS

Completion of Appendix A is required when for good reasons key documents are not available when a fast track application is approved by the supervisor/module lecturer and second academic signatory.

I have now checked and approved all the key documents associated with this proposal e.g. consent form, information sheet, questionnaire, interview schedule

Title of project/assignment: Lung Nodule Detection using Deep learning model... ..
...

Name of student researcher : Winarto Wijaya... ..

Student ID: TP041218.....
...

Intake: APUMF2009AI

Signed... .. Print Name... .. Date... ..
(Supervisor/Lecturer)

Appendix B: Disclaimer Form

Office Record

Date Received:

Received by whom:

Receipt

Student name: Winarto Wijaya

Student number: TP041218

Received by:

Date:

ACADEMIC RESEARCH ETHICS

DISCLAIMER

**Declaration about ethical issues and implications of
research project/assignment proposals to be included on
project/assignment application forms**

Project/Assignment Title:**Lung Nodule Detection using Deep Learning Model.....**

.....

The following declaration should be made in cases where research project/assignment applicants for a particular project/assignment and the supervisor(s)/lecturer(s) for that project/assignment conclude that it is not necessary to apply for ethical approval for the research project/assignment.

We confirm that the University's guidelines for ethical approval have been consulted and that all ethical issues and implications in relation to the above project/assignment have been considered.

We confirm that ethical approval need not be sought.

Winarto Wijaya

_____
Name of Research Project/Assignment Applicant_____
Signature_____
Date_____
Dr. V. SIVAKUMAR_____
*V. Sivakumar*_____
Name of Research Project Supervisor/
Assignment Lecturer_____
Signature_____
Date

Appendix C: Log Sheets

Project Log Sheet – Supervisory Session

Notes on use of the project log sheet:

1. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at minimum SIX (6) during the course of the project (SIX mandatory supervisory sessions).
2. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any) since the last session, and noting these in the relevant sections of the form, effectively forming an agenda for the session.
3. A log sheet is to be brought by the STUDENT to each supervisory session.
4. The actions by the student (and, perhaps the supervisor), which should be carried out before the next session should be noted briefly in the relevant section of the form.
5. The student should leave a copy (after the session) of the Project Log Sheet with the supervisor and to the administrator at the academic counter. A copy is retained by the student to be filed in the project file.
6. It is recommended that students bring along log sheets of previous meetings together with the project file during each supervisory session.
7. The log sheet is an important deliverable for the project and an important record of a student's organisation and learning experience. The student **must** hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's name: Winarto wijaya..... **Date:** 5th March 2021..... **Meeting No:** 1...

Project title: Lung Nodule Detection using Deep learning

Supervisor's name: Dr. V. SIVAKUMAR **Supervisor's signature:** *V. Sivakumar*

Items for discussion (noted by student before mandatory supervisory meeting):

1. Topic of choice – to focus on with lung nodule detection
2. Expectation on the RMCE paper for literature review
3. RMCE related questions like research gap, problem statement

Record of discussion (noted by student during mandatory supervisory meeting):

1. Look for survey papers – topic is still open and not set to stone
2. try to find something that would be interesting
3. try to find resource such as dataset and past practices

Action List (to be attempted or completed by student by the next mandatory supervisory meeting):

1. massive research on lung nodule detection perhaps segmentation task
2. finish up the RMCE paper 1
3. possibly find out what is the option for the lung nodule project

Project Log Sheet – Supervisory Session

Notes on use of the project log sheet:

8. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at minimum SIX (6) during the course of the project (SIX mandatory supervisory sessions).
9. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any) since the last session, and noting these in the relevant sections of the form, effectively forming an agenda for the session.
10. A log sheet is to be brought by the STUDENT to each supervisory session.
11. The actions by the student (and, perhaps the supervisor), which should be carried out before the next session should be noted briefly in the relevant section of the form.
12. The student should leave a copy (after the session) of the Project Log Sheet with the supervisor and to the administrator at the academic counter. A copy is retained by the student to be filed in the project file.
13. It is recommended that students bring along log sheets of previous meetings together with the project file during each supervisory session.
14. The log sheet is an important deliverable for the project and an important record of a student's organisation and learning experience. The student **must** hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's name: Winarto Wijaya..... **Date:** 5th March 2021..... **Meeting No:** 2...

Project title: Lung Nodule Detection using Deep learning

Supervisor's name: Dr. V. SIVAKUMAR **Supervisor's signature:** *V. Sivakumar*

Items for discussion (noted by student before mandatory supervisory meeting):

1. Issue with the topic of choice
2. Late realization for implementation and side-tracked
3. updates with the progress implementation

Record of discussion (noted by student during mandatory supervisory meeting):

1. questions regarding presentation and tips for the final paper
2. strengthening on gap and problem statement
3. provide proper justification

Action List (to be attempted or completed by student by the next mandatory supervisory meeting):

1. back to implement and re-do some of the implementation
2. find proper justification for report
- 3.

Project Log Sheet – Supervisory Session

Notes on use of the project log sheet:

15. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at minimum SIX (6) during the course of the project (SIX mandatory supervisory sessions).
16. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any) since the last session, and noting these in the relevant sections of the form, effectively forming an agenda for the session.
17. A log sheet is to be brought by the STUDENT to each supervisory session.
18. The actions by the student (and, perhaps the supervisor), which should be carried out before the next session should be noted briefly in the relevant section of the form.
19. The student should leave a copy (after the session) of the Project Log Sheet with the supervisor and to the administrator at the academic counter. A copy is retained by the student to be filed in the project file.
20. It is recommended that students bring along log sheets of previous meetings together with the project file during each supervisory session.
21. The log sheet is an important deliverable for the project and an important record of a student's organisation and learning experience. The student **must** hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's name: Winarto Wijaya..... **Date:** 5th March 2021..... **Meeting No:** 3...

Project title: Lung Nodule Detection using Deep learning

Supervisor's name: Dr. V. SIVAKUMAR **Supervisor's signature:** *V. Sivakumar*

Items for discussion (noted by student before mandatory supervisory meeting):

1. updating progress implementation
2. asking for signing forms ethics and disclaimer
- 3.

Record of discussion (noted by student during mandatory supervisory meeting):

1. check on PL for confidentiality form
2. Strengthening justification
- 3.

Action List (to be attempted or completed by student by the next mandatory supervisory meeting):

1. Back on report and codes implementation
2. submission
- 3.

Appendix D: Gantt Chart

Supervisor Dr. V. Sivakumar
Project Start Date: 17-05-21

Milestone description	Category	Assigned to	Progress	Start	Days
3rd model					
3rd model built	Med Risk	coding	26%	08-07-21	14
3rd model training	Low Risk	training	30%	22-07-21	2
3rd test run	Milestone	tertiary test	34%	24-07-21	1
supervisor show case	Milestone	present	40%	25-07-21	1
report writing chapter 2	Low Risk	documentation	42%	25-07-21	5
documentation					
report writing chapter 3	Low Risk	documentation	46%	31-07-21	5
fine tune on model 1	Low Risk	coding	50%	05-08-21	5
report writing chapter 4	Low Risk	documentation	54%	10-08-21	5
report writing chapter 5	Low Risk	documentation	58%	15-08-21	5
Final model + final documentation	On Track	documentation	70%	20-08-21	14
submission	Goal	documentation	100%	03-09-21	1

