

**UNIVERSITI TEKNOLOGI MARA**

**AN APPLICATION OF FASTER  
R-CNN FOR CHEST X-RAY DIGITAL  
IMAGE CLASSIFICATION**

**MOHD TAUFIK BIN RAHMAT**

Dissertation submitted in partial fulfillment  
of the requirements for the degree of  
**Master of Data Science**

**Faculty of Computer and Mathematical Sciences**

**January 2019**

## ABSTRACT

Non-digitalized chest X-ray is an effective, low-cost screening tool, and it is important to indicate pathologies. However, there are some cases of misinterpretation in the diagnostic process. Reading and interpret chest X-ray may be a simple task for a radiologist, but not every doctor can do it the same. This paper aims to evaluate the performance of chest X-ray image classification using Faster R-CNN architecture. To develop a chest x-ray classifier model, Tensorflow package was used with python. The results show the propose model performance accuracy is 62%. The model then was compared to random selected one medical student and general practitioner. The model shows better in term of performance to classify chest x-ray images with 62% accuracy compared to selected medical students and general practitioners with their accuracy score of 56% and 50% respectively. In term of chest X-ray interpretation in this study, the result shows that the model performance is more reliable to use for chest x-ray images classification. Tough the model performance is better, but in medical field reality, it is still far from the standard to be applied. With 62% accuracy, the model is unsafe to use. The future works are to gain more knowledge from radiologist expert to improve chest -x-ray classifier performance.

## ACKNOWLEDGEMENT

Firstly, I wish to thank God for giving me the opportunity to embark on my project and for completing this long and challenging journey successfully. My gratitude and thanks go to my supervisor Dr. Azlan Ismail and Dr Sharifah Aliman.

My appreciation goes to the all lecturers who provided knowledge and assistance during the project. Special thanks to my colleagues and friends for helping me with this project.

Finally, this project is dedicated to the loving memory of my very dear late father and mother for the vision and determination to educate me. This piece of victory is dedicated to both of you. Alhamdulillah.

# TABLE OF CONTENT

	<b>Page</b>
<b>CONFIRMATION BY PANEL OF EXAMINERS</b>	<b>ii</b>
<b>AUTHOR'S DECLARATION</b>	<b>iii</b>
<b>ABSTRACT</b>	<b>iv</b>
<b>ACKNOWLEDGEMENT</b>	<b>v</b>
<b>TABLE OF CONTENT</b>	<b>vi</b>
<b>LIST OF TABLES</b>	<b>ix</b>
<b>LIST OF FIGURES</b>	<b>x</b>
<b>CHAPTER ONE INTRODUCTION</b>	<b>12</b>
1.1 Research Background	12
1.2 Problem Statement	13
1.3 Research Question	14
1.4 Research Objective	14
1.5 Significance of Study	14
<b>CHAPTER TWO LITERATURE REVIEW</b>	<b>16</b>
2.1 Introduction	16
2.2 Data Mining in Healthcare Systems	16
2.3 Data Mining for X-ray Diagnosis	19
2.4 Various Data Mining Techniques for Image Classification	23
2.5 Convolutional Neural Network (CNN)	25
2.5.1 2.5.1 Introduction	25
2.5.2 2.5.2 Architecture of Convolutional Neural Network	27
2.6 Faster R-CNN	30
2.6.1 2.6.1: Comparison between R-CNN, Fast R-CNN, and Faster R-CNN	30
2.6.2 2.6.2 The architecture of Faster R-CNN	31
2.6.3 2.6.3 RPN as a regional proposal	32
2.6.4 RPN loss function	32

2.7	Review of Previous Studies on Chest X-ray Image Analysis	32
<b>CHAPTER THREE RESEARCH METHODOLOGY</b>		<b>37</b>
3.1	Introduction	37
3.2	Data Preparation	39
3.2.1	Data Source and research understanding	39
3.2.2	Chest X-ray images data preparation	40
3.2.3	Creating a label map for chest X-ray images	42
3.3	Chest X-ray model classifier model development using Faster R-CNN	43
3.3.1	Installing TensorFlow-GPU	43
3.3.2	Setting up Object Detection directory and Anaconda Virtual Environment	44
3.3.3	Download repository of object detection TensorFlow from GitHub	45
3.3.4	Set up new Anaconda virtual environment	45
3.3.5	Configure PYTHONPATH environment variable	45
3.3.6	Compile Protobufs and run setup.py	45
3.3.7	Generate Training and Testing Data	46
3.3.8	Create a Label Map and Configure Training	46
3.3.9	Train the model	46
3.3.10	Setting up platform	47
3.3.11	Object detection	47
3.3.12	Dataset for validation	47
3.4	Validation and Model Deployment	48
3.4.1	Export Inference Graph	48
3.4.2	Validating with 100 chest X-ray images	48
3.5	Result Assessment	48
<b>CHAPTER FOUR RESULT AND DISCUSSION</b>		<b>49</b>
4.1	Introduction	49
4.2	Results and Discussion of findings	49
4.2.1	Visualize training and testing performance using Tensorboard	49
4.2.2	The result of Chest X-ray image classifier	51
4.2.3	Model validation and comparison	53