



**UNIVERSITI TEKNIKAL MALAYSIA MELAKA**

**DEVELOPMENT OF COMPUTED TOMOGRAPHY**

**LUNG CANCER ANALYSIS SYSTEM**

This report is submitted in accordance with the requirement of the Universiti Teknikal Malaysia Melaka (UTeM) for the Bachelor of Electronics Engineering Technology with Honours.



**NURLYANA BINTI KARIM**

**B071710899**

**960803-14-5528**

**FACULTY OF ELECTRICAL AND ELECTRONIC ENGINEERING**

**TECHNOLOGY**

2021

**BORANG PENGESAHAN STATUS LAPORAN PROJEK SARJANA MUDA**

Tajuk: DEVELOPMENT OF COMPUTED TOMOGRAPHY LUNG CANCER  
ANALYSIS SYSTEM

Sesi Pengajian: 2021

Saya **NURLYANA BINTI KARIM** mengaku membenarkan Laporan PSM ini disimpan di Perpustakaan Universiti Teknikal Malaysia Melaka (UTeM) dengan syarat-syarat kegunaan seperti berikut:

1. Laporan PSM adalah hak milik Universiti Teknikal Malaysia Melaka dan penulis.
2. Perpustakaan Universiti Teknikal Malaysia Melaka dibenarkan membuat salinan untuk tujuan pengajian sahaja dengan izin penulis.
3. Perpustakaan dibenarkan membuat salinan laporan PSM ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. \*\*Sila tandakan (X)

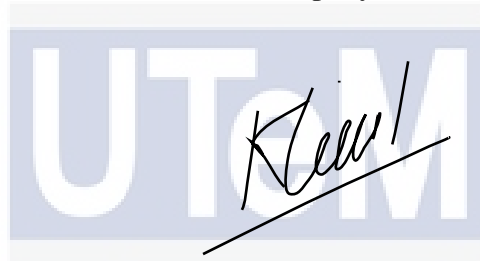

- SULIT\* Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia sebagaimana yang termaktub dalam AKTA RAHSIA RASMI 1972.
- TERHAD\* Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan.
- TIDAK TERHAD

Yang benar,

Disahkan oleh penyelia:



NURLYANA BINTI KARIM



Ts. KHAIRUL AZHA BIN A AZIZ

Ts. KHAIRUL AZHA BIN A AZIZ

Pensyarah Kanan

Alamat Tetap:

No.261, Jalan Asean 14,

Taman Asean,

75250 Melaka

Tarikh: 14/01/2021

Tarikh: 14/01/2021

\*Jika Laporan PSM ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan PSM ini

## DECLARATION

I hereby, declared this report entitled DEVELOPMENT OF COMPUTED TOMOGRAPHY LUNG CANCER ANALYSIS SYSTEM is the results of my own research except as cited in references.

Signature: *Lyana*

Author: NURLYANA BINTI KARIM

Date: 14/01/2021



اونيورسيتي تيكنيكل مليسيا ملاك

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

## APPROVAL

This report is submitted to the Faculty of Electrical and Electronic Engineering Technology of Universiti Teknikal Malaysia Melaka (UTeM) as a partial fulfilment of the requirements for the degree of Bachelor of Electronics Engineering Technology with Honours. The member of the supervisory is as follow:

Signature: .....  
Supervisor : Ts. KHAIRUL AZHA BIN A AZIZ

Ts. KHAIRUL AZHA BIN A AZIZ  
Pensyarah Kanan  
Jabatan Teknologi Kejuruteraan Elektronik Dan Komputer  
Fakulti Teknologi Kejuruteraan Elektrik Dan Elektronik  
Universiti Teknikal Malaysia Melaka

اونيورسيتي تيكنيكل مليسيا ملاك  
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

## ACKNOWLEDGEMENTS

In preparation of the project, I had to take the help and guidance of some respected persons, who deserve my deepest gratitude. As the completion of this project gave my pleasure, I would like to show my gratitude to Mr. Khairul Azha bin Aziz, supervisor, on Universiti Teknikal Malaysia Melaka (UTeM) for giving me a good guideline for this project throughout numerous consultations. I would also like to expand my gratitude to all those who have directly and indirectly guided us in writing this assignment.

Most importantly, none of this could have happened without my family. I should express my very profound gratitude to my parents, sisters, brothers and friends for providing me with unfailing support and continuous encouragement throughout years of study and through the process of researching and writing this project report. This accomplishment would not have been possible without them. Thank you.

I am using this opportunity to express our gratitude to everyone who supported us throughout the project. I am thankful for their aspiring guidance, invaluable constructive criticism and friendly advice during the project work. I am sincerely grateful to them for sharing their truthful and illuminating view on a number of issues related the project.

## ABSTRAK

Kanser paru-paru ialah salah satu bentuk kanser yang paling popular. Diagnosis awal kanser paru-paru dengan pendekatan penjagaan yang betul adalah penting untuk kehidupan individu. Dalam projek ini, kaedah ambang untuk segmentasi dan pengekstrakan lesi paru paru telah dicadangkan dalam Sistem Analisis Kanser Paru-Paru Tomografi Dikira. Laporan ini memberikan kaedah pengesanan kanser yang diperoleh daripada imbasan CT DICOM Lung untuk mengenal pasti lesi kanser. Imej pra-proses (ambang) dikumpulkan selepas mengimbas imej CT Paru. Seterusnya, fokus pada ciri-ciri tekstur pada imej paru-paru dalam perisian MATLAB dan aplikasi ALIZA. Tujuan projek ini adalah untuk menunjukkan keadaan perbandingan ukuran dan prestasi lesi menggunakan pemprosesan gambar barah paru-paru dan menilai ketepatan ambang iaitu Teknik Ambang dan Kaedah Fuzzy. Sistem ini menggunakan konvensyen visual untuk mewakili perkembangan system analisis barah paru-paru Tomografi Dikira. Gambar asal dan gambar segmen dari proses segmentasi adalah antara gambar di dalam Sistem GUI. Nombor dan ukuran lesi pesakit ditunjukkan juga di dalam Sistem GUI. Gambaran kaedah ambang menjadi jelas bahawa ukuran dan keluaran harus saling menghampiri, dibandingkan dengan dua kaedah iaitu Teknik Ambang dan Kaedah Fuzzy. Akhirnya, prestasi telah dikira dari pengukuran lesi di MATLAB dan dibandingkan dengan ukuran lesi dari aplikasi ALIZA.

## ABSTRACT

Lung cancer is one of the most popular forms of cancer. Early diagnosis of lung cancer with proper care is important for the life of the individual. In this project, threshold methods for segmentation and extraction of lung lesions have been proposed in the Computed Tomography Lung Cancer Analysis System. This report provides a cancer detection method texture features derived from the DICOM Lung CT scan for the recognition of cancerous lesions. The pre-process images by thresholding collected after scanning Lung CT images. Next, focus on features of lesion in the lung image that have cancer in MATLAB Software and ALIZA Application. The purpose of the project is to present the state of comparison of seize and performance lesion using image lung cancer processing and evaluating the accuracy of segmentation by thresholding that is Otsu Thresholding and Fuzzy C Means. This system use visual conventions to represent the development of Computed Tomography lung cancer analysis system on Graphical User Interface. The original image and segment image from segmentation process are among the images in the GUI system. Patient number and size of lesion displayed also in GUI System. The image of the thresholding method became clear that the size and output should be approaching, comparing from the two methods, OTSU Thresholding and Fuzzy C Means. Finally, the performance had calculated from the lesion size measurements in MATLAB and compared to the size of lesion from ALIZA Application.



## TABLE OF CONTENTS

	<b>PAGE</b>
<b>TABLE OF CONTENTS</b>	<b>ix-xiv</b>
<b>LIST OF TABLES</b>	<b>xv</b>
<b>LIST OF FIGURES</b>	<b>xvi-xvii</b>
<b>LIST OF ABBREVIATIONS</b>	<b>xviii-xix</b>
<b>LIST OF APPENDICES</b>	<b>xx</b>
<b>CHAPTER 1 INTRODUCTION</b>	<b>1</b>
1.1 Background	1
1.2 Problem Statement	1-2
1.3 Objectives	2-3
1.4 Project Scope	3-4
<b>CHAPTER 2 LITERATURE REVIEW</b>	<b>5</b>
2.0 Introduction	5
2.1 Cancer	5
2.1.1 Historical Cancer	5
2.2 Type of Common Cancer	6
2.2.1 Lung Cancer	6

2.2.2	Breast Cancer	6
2.2.3	Brain Cancer	6-7
2.2.4	Heart Cancer	7
2.2.5	Liver Cancer	7
2.3	Computed-Aided Detection	8
2.4	Lung Segmentation	8-9
2.4.1	Segmentation Method	9
2.4.1.1	Thresholding-based Method	9-11
2.4.1.2	Region-based Method	11-12
2.4.1.3	Shape-based Method	12-13
2.4.1.4	Neighboring Anatomy-guided Method	13-14
2.4.1.5	Machine Learning-based Method	14-15
2.4.2	Comparison Image Segmentation Methods	15-17
2.5	Other Cancer Segmentation	17
2.5.1	Brain Image Segmentation	18
2.5.1.1	Manual Segmentation Method	18
2.5.1.2	Unsupervised Method	18
2.5.1.3	Semi-supervised and Supervised Method	19
2.5.2	Breast Image Segmentation	19

2.5.2.1	Thresholding	19
2.5.2.2	Morphology-based	19-20
2.5.2.3	Active Contour	20
2.5.2.4	Texture-based	20
2.5.2.5	Pixel-wise Clustering	20
2.5.3	Heart Image Segmentation	20
2.5.3.1	Boundary-Driven Techniques	21
2.5.3.2	Region-Based Techniques	21
2.5.3.3	Graph-Cuts Techniques	21-22
2.5.3.4	Model-Fitting Techniques	22
2.5.4	Liver Image Segmentation	22
2.5.4.1	Manual Segmentation	22
2.5.4.2	Semi-Automated Segmentation	23
2.5.4.3	Fully-Automated Segmentation	23
2.6	Lung Nodule Detection	23-24
2.6.1	Clinical Applications	24-25
2.7	Lung Nodule Size	25
2.7.1	Table Pulmonary Nodule	26
2.8	Segment Image	26

2.8.1	Medical Resonance Imaging (MRI)	26-27
2.8.2	Computed Tomography Imaging(CT scan)	27-28
2.8.3	Positron Emission Tomography (PET)	28-30
2.9	Image Digital Imaging and Communication in Medicine	30
2.9.1	DICOM Standard	31-32
2.10	Hounsfield Unit	32-33
2.11	CT Image Quality	33-34
2.11.1	Spatial Resolution CT	34-35
2.11.2	Signal to Noise Ratio (SNR)	35
2.12	Images Segmentation Technique	35
2.12.1	Otsu Thresholding	35
2.12.2	Fuzzy C-Means	36
2.13	Summary	36
<b>CHAPTER 3      METHODOLOGY</b>		<b>37</b>
3.1	Introduction	37
3.2	Project Work Flow	37-38
3.3	Planning	38
3.3.1	Flow Chart Project Methodology	39
3.3.2	Flow Chart of Algorithm	40

3.3.3	Block Diagram Process	41
3.4	Research and Data Collection	42
3.5	Setting Software	42
3.5.1	Uploading the Data Set in Aliza Software	43-44
3.5.2	Segmentation of Medical Images	44-45
3.5.3	Uploading the Data Set in MATLAB	46-47
3.5.4	Cancer Segmentation at MATLAB	47
3.5.4.1	DICOM Image	47
3.5.4.2	Clipping Method	48
3.5.4.3	Edge Detection Using Kirsch Algorithms	48-49
3.5.4.4	Otsu Thresholding Technique	49-50
3.5.4.5	Fuzzy C Means	51-52
3.6	Size Check at MATLAB	53
3.7	Size Check at ALIZA Software	53-54
3.8	Comparison Size and Performance Lesion	54-56
3.9	Graphical User Interface (GUI)	56
3.9.1	Editor at MATLAB	57
3.9.2	Properties GUI	57-58
3.10	Image Grayscale Pixel Values and Hounsfield Values	58-62

<b>CHAPTER 4</b>	<b>RESULTS</b>	<b>63</b>
4.1	Introduction	63
4.2	Results	63-65
4.3	Development of Graphical User Interface (GUI) System	65-66
4.4	Size and Performance	66-78
4.5	Overall Result	79-80
4.6	Commercialization Relevance	80
<b>CHAPTER 5</b>	<b>CONCLUSION</b>	<b>81</b>
5.1	Conclusion	81-82
5.2	Recommendation	82
<b>REFERENCES</b>		<b>83-88</b>
<b>APPENDIX</b>		<b>89-92</b>

## LIST OF TABLES

TABLE	TITLE	PAGE
Table 2.1:	CT Lung Segmentation approaches five methods	15-17
Table 2.2:	Table Pulmonary Nodule	26
Table 2.3:	Variable as assign values from DICOM Info	31-32
Table 2.4:	Table of the matter and Hounsfield Unit	32-33
Table 3.1:	Comparison image between ALIZA and MATLAB software	55
Table 3.2:	Comparison between Grayscale Pixel Value and Hounsfield Unit	59-60
Table 3.3:	Comparison between Hounsfield Unit and Grayscale Pixel Value of Lesion	61-62
Table 4.1:	Size and Performance Results between ALIZA and MATLAB using Otsu Thresholding	68-72
Table 4.2:	Size and Performance Results between ALIZA and MATLAB using Fuzzy C means	73-78

## LIST OF FIGURES

FIGURE	TITLE	PAGE
Figure 2.1:	Flowchart of the threshold-based process of segmentation of the lungs	10
Figure 2.2:	Segmentation results with measured lung borders	10
Figure 2.3:	Flowchart of the region-based method of lung segmentation	12
Figure 2.4:	Examples of cases in which neighboring anatomy-guided segmentation method are produced	14
Figure 2.5:	Examples of successful machine learning-based segmentation	15
Figure 2.6:	Sample of original and segmentation image	18
Figure 2.7:	Seeded region growing is an intensity-based technique	23
Figure 3.1:	Project Work Flow	38
Figure 3.2:	Project Flow Chart	39
Figure 3.3:	Flow Chart Algorithm	40
Figure 3.4:	Block Diagram Process	41
Figure 3.5:	Basic function in Aliza Software	43
Figure 3.6:	The Application folder at Aliza Software	44
Figure 3.7:	Other folder inside the Segmentation Folder at Aliza Software	45
Figure 3.8:	The Mode to discard values at Aliza Software	45
Figure 3.9:	Connected Threshold at Aliza Software	45
Figure 3.10:	Basic Function in MATLAB	46



Figure 3.11: Workspace browser at MATLAB	46
Figure 3.12: Editor Window at MATLAB	46
Figure 3.13: Patient R040 images at MATLAB	47
Figure 3.14: Output for Clipping Code Process	48
Figure 3.15: Output image Kirsch Algorithm	49
Figure 3.16: Output image after Otsu Thresholding Segmentation	49
Figure 3.17: Otsu Thresholding Process Flowchart	50
Figure 3.18: Output image after Fuzzy C Means Segmentation	51
Figure 3.19: Clustering the image using 'k'	51
Figure 3.20: Fuzzy C Means Process Flowchart	52
Figure 3.21: Image Lung with Lesion at ALIZA Software	54
Figure 3.22: The Measurement of Lesion at ALIZA Software	54
Figure 3.23: The Programming At Editor MATLAB	57
Figure 3.24: The Image of Properties GUI	57
Figure 3.25: The Output at GUI	58
Figure 3.26: Gantt chart for FYP 2	62
Figure 4.1: Patient R040 images at MATLAB	64
Figure 4.2: Output for Clipping Code Process	64
Figure 4.3: Output image Kirsch Algorithm	64
Figure 4.4: Output image after Otsu Thresholding Segmentation	65
Figure 4.5: The interface for selection at GUI System	65
Figure 4.6: Overview results display on Graphical User Interface	67

## LIST OF ABBREVIATION

<b>2D</b>	-	2 Dimension
<b>3D</b>	-	3 Dimension
<b>4D</b>	-	4 Dimensions
<b>AC</b>	-	Alternating Current
<b>CM</b>	-	Centi Meter
<b>MM</b>	-	Milli Meter
<b>HU</b>	-	Hounsfield Unit
<b>CAD</b>	-	Computed Aided Detection
<b>CT</b>	-	Computed Tomography
<b>LIDC</b>	-	Lung Image Database Consortium
<b>FCM</b>	-	Fuzzy C Means
<b>FMM</b>	-	Finite Mixture Model
<b>SOM</b>	-	Self-Organizing Maps
<b>MLP</b>	-	Multi-Layer Perceptron
<b>SVM</b>	-	Support Vector Machine
<b>MRF</b>	-	Markov Random Fields
<b>DICOM</b>	-	Digital Imaging and Communications On Medicine
<b>SSMs</b>	-	Statistical Shape Models
<b>FPNs</b>	-	False Positive Nodules
<b>TPNs</b>	-	True Positive Nodules
<b>MRI</b>	-	Magnetic Resonance Imaging

- PET** - Positron Emission Tomography
- ROI** - Release of Information
- GUI** - Graphical User Interface
- RGB** - Red Green Blue
- SNR** - Signal to Noise Ratio
- FOV** - Field of View



## LIST OF APPENDICES

APPENDIX	TITLE	PAGES
Appendix A:	Code of the Read DICOM Images at MATLAB	89
Appendix B:	Code of the clipping process at MATLAB	89
Appendix C:	Code of the Kirsch Image at MATLAB	90
Appendix D:	Code of the OTSU Thresholding image at MATLAB	91-92
Appendix E:	Code of the call back for push button at GUI MATLAB	92



# CHAPTER 1

## INTRODUCTION

### 1.1 Background

Lung Cancer is the primary cause of the disease related death globally, with 30%-40% happening in developing countries. More than 1.8 million cases of lung cancer globally have resulted in 1.6 million deaths in 2012. Prevalence of lung cancer and consequent death from lung cancer is expected to increase over the next decade due to an increase in smoking rate. (Siang and John, 2016)

Proof that early detection of lung cancer has resulted more effective. Within this research, a Computer-Aided Detection (CAD) method is introduced is an effort to identify lung cancer areas utilizing Computed Tomography (CT) images. (El-Baz *et al.*, 2013)

CT scan is one of the imaging techniques used to examine disease or lesion. CT stands for computed tomographic modality, which depicts the scanned organ in a black and white setting. This modality often includes graphical readings at its edges in order to pinpoint the diseased region precisely. Big shifts, such as when the illness is at malignant level, display a noticeable improvement in the CT scan. In fact, the cancer checks are carried out and a CT scan of the lung is obtained for examination.

### 1.2 Problem Statement

Medical neglect is a major public health issue and a leading cause of death. According to specific physical conditions such as human visual system impairment,

exhaustion and agitation, physicians do not allow optimum use of CT images data. It is difficult to find a consistent and feasible solution that minimises the chances of a repeat case. Differentiating the infected of the cancer in the lungs and providing the proper solution to the problem are also the toughest in the medical job. Classification problems and the related solutions to get a better image of nodule. Image analysis may be solvent for this form of challenge, in particular to distinguish cancer affected areas in the lungs.

This research is suggested in order to increase the quality and precision of manual disease diagnosis. With this implementation manual work can be minimized and categorizing the distinction computer applications in any CT datasets can be performed automatically. It may serve to prevent a medical mistake and a misdiagnosis. It can be done by compiling a sample of correct CT scan readings and adding a sample classification algorithm. This research also carried out a review of the evidence on lung cancer and recommends the best method use to read the image of the nodule in lung.

### 1.3 Objective

The proposed CAD method has three key phases: segmentation the CT images, classification of identified areas that is lesions and compare the performance between methods used. The research is qualified, tested and verified using images of lung cancer.

Results obtain in the assessment of the lesion using MATLAB Software. In addition, the propose research can identify which methods is more accurate or closer to the original measurement.

The key objectives of the project are:

- To develop a Computed Tomography Lung Cancer Analysis System

- To proposed Thresholding method for the lung cancer segmentation and extraction.
- To present the state of comparison of lesion size and performance between Thresholding Methods.

#### **1.4 Project Scope**

The purpose out of this effort is the new approach that can be used to classify lung nodules. And it has been known to be reasonable accuracy. The algorithm is a series of simple procedure for image processing. In order to build a strong lung cancer screening method, the method utilizes at MATLAB, a high performance programming tool. Toolboxes allow the learning and implementation of specialized technology.

The suggested CAD method begins with the pre-processing of 3D CT scans utilizing segmentation, normalization, down sampling and zero-centering. The final classification stage deals with patterns that are depicted as point in a feature space. Seeking judgement boundaries in such vector spaces is a core concern in the theory of pattern recognition. The dataset for lung cancer images is from The Cancer Imaging Achieved (TCIA).

The purpose of the project is to present the state of comparison using image lung cancer processing and evaluating the accuracy of segmentation by thresholding methods that is Otsu Thresholding and Fuzzy C Means. Thresholding segmentation approaches have been shown to be successful in the measurement of lung lesion and this implementation by Computed Tomography (CT scan). However, the image data collection from the CT scan as a device image had very unique style as DICOM format.

It is because it does not only place the pixel of the image data but also data sets which are made up of attributes. This is because it not only stores image pixel info, but also data sets that are made up of attributes. To process the image requires a few images with only a cancer appearance in the lung, not a process for all pieces of slice image.

