

**UNIVERSITI TEKNOLOGI MARA**

**IMAGE FEATURE EXTRACTION  
FOR COLORECTAL CANCER  
CELLS CLASSIFICATION**

**MOHD. YAMIN BIN AHMAD**

Thesis submitted in fulfillment  
of the requirement for the degree of  
**Master of Science**

**Faculty of Computer and Mathematical Sciences**

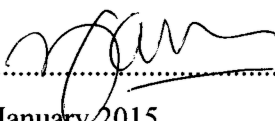
January 2015

## **AUTHOR'S DECLARATION**

I declare that the work in this thesis was carried out in accordance with the regulation of Universiti Teknologi MARA. It is original and is the result of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student : Mohd. Yamin Bin Ahmad  
Student I.D. No. : 2009798325  
Programme : Master of Science  
Faculty : Faculty of Computer and Mathematical Sciences  
Thesis Title : Image Feature Extraction For Colorectal Cancer  
Cells Classification

Signature of Student :   
Date : January 2015

## **ABSTRACT**

Manual screening of colorectal biopsy tissue under microscope to conform the presence of cancerous cell is difficult, arduous and time consuming. The criteria in diagnosing colorectal cancer cell are gland shape and nucleus size. In this study, we proposed a method of image pre-processing to extract the important feature of colorectal tissue images. Images captured under microscope may vary in colour brightness due to different H&E stain concentration and the size of biopsy tissue. To overcome this problem a method using HSV colour model to remove element outside the area of nucleus is used. A novel method named Pixel Mask Analyzer is proposed to clean the image and remove noises. Meanwhile, the gland boundary tracking and segmentation is proposed to extract the gland shape. By using the result of gland tracking, nucleus size that forms the glands are measured. By combining result of gland shapes and nucleus size, the image classification is performed. The result shows that classification achieves 96.9% accuracy by using the proposed methods. With the high accuracy results and findings of this study, it is hope that the study can contribute a very substantial amount of outcomes that would greatly benefit the research areas especially in image processing and classification of colorectal cancer.

## **ACKNOWLEDGEMENT**

In the name of Allah, Most Gracious, Most Merciful All praise and thanks are due to Allah, and peace and blessings be upon His Messenger.

I would like to express my most sincere gratitude and appreciation to my supervisor, Professor Dr. Azlinah Mohamed who has patiently guided and advised me through many discussions during the preparation of this study. My appreciation also goes to Professor Dr. Yasmin Anum Mohamad Yusof from Department of Biochemical Faculty of Medicine University Kebangsaan Malaysia and Professor. Dr. Siti Aisyah Md Ali from Department of Pathology Pusat Perubatan University Kebangsaan Malaysia for all the knowledge, guidance and opinions for this research.

Special thanks to my beloved family for their support, patience and endurance shown over the accomplishing of this study, and not forgetting all my friends that have provided me with useful tips, guidance and support in completing this study.

And last but not the least, may Allah, Beneficent Load of Mercy save us all from the evil inclinations of our souls and make us hate disbelief, transgressions and sins, and may He endear to our hearts faith and good works.

## **TABLE OF CONTENTS**

<b>AUTHOR'S DECLARATION</b>	ii
<b>ABSTRACT</b>	iii
<b>ACKNOWLEDGEMENT</b>	iv
<b>TABLE OF CONTENTS</b>	v
<b>LIST OF TABLES</b>	ix
<b>LIST OF FIGURES</b>	x
<b>CHAPTER ONE: INTRODUCTION</b>	1
1.1 Background	1
1.2 Statement of the Problem	3
1.3 Research Objectives	4
1.4 Scope of the Study	4
1.5 Significance of the Study	4
1.6 Overview of Research Methodology	5
1.7 Limitation of Study	6
1.8 Overview of the Thesis	6
<b>CHAPTER TWO: REVIEW OF LITERATURE</b>	8
2.1 Colon	8
2.2 What is Cancer?	9
2.3 Colorectal Cancer	10
2.3.1 Conventional Screening Process	11
2.3.2 Normal Colon Cells	13
2.3.3 Adenoma in Colon	14
2.3.4 Adenocarcinoma	15
2.4 Image Processing	16
2.4.1 Noise	16
2.4.2 Noise Removal	17
2.4.2.1 Salt and Pepper Noise Removal Using kFill Algorithm	17
2.4.2.2 Mean Filter	18