

REGION-GROWING BASED SEGMENTATION
AND BAG OF FEATURES CLASSIFICATION
FOR BREAST ULTRASOUND IMAGES

LEE LAY KHOON

MASTER OF COMPUTER SCIENCE

UNIVERSITI MALAYSIA PAHANG



SUPERVISOR'S DECLARATION

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of the Master of Computer Science (2015/2016)

(Supervisor Signature)

Name of Supervisor : DR. LIEW SIAU CHUIN

Position : SENIOR LECTURER

Date :



STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citation which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang or any other institutions.

(Author's Signature)

Full Name : LEE LAY KHOON

ID Number : MCC14001

Date :

REGION-GROWING BASED SEGMENTATION AND BAG OF FEATURES
CLASSIFICATION FOR BREAST ULTRASOUND IMAGES

LEE LAY KHOON

Thesis submitted in fulfilment of the requirements
for the award of the degree of
Master of Computer Science

Faculty of Computer System and Software Engineering
UNIVERSITI MALAYSIA PAHANG

JANUARY 2017

ACKNOWLEDGEMENTS

Faithfully thanks to my supervisor Dr. Eric Liew Siau Chuin for his guidance on my work, Region-growing based segmentation and classification for breast ultrasound images. Thanks to my supervisor for providing precious suggestions and opinions on helping to enhance my research work. Thanks to my parents for their supports and love throughout my life. Thanks to my friends on providing their precious opinions and suggestions, which helped much on my research.

TABLE OF CONTENTS

DECLARATION	
TITLE PAGE	
ACKNOWLEDGEMENTS	i
ABSTRACT	ii
ABSTRAK	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	xii
LIST OF ABBREVIATIONS	xiii
LIST OF SYMBOLS	xiv
CHAPTER 1 INTRODUCTION	1
1.1 Background	1
1.2 Breast Cancer Status	4
1.3 Ultrasound Imaging	5
1.3.1 Risks and Benefits of Ultrasound Imaging	6
1.4 Problem Statement	7
1.5 Research Objectives	8
1.6 Research Scopes	8
1.7 Organization of Thesis	9
CHAPTER 2 LITERATURE REVIEW	10
2.1 Introduction	10
2.2 Foundation	10

2.2.1	Roi (Region of Interest)	10
2.2.2	Roi (Region of Interest) in Segmentation	11
2.2.3	Image Segmentation	11
2.2.4	Medical Image	12
2.3	Modality	12
2.3.1	Magnetic Resonance Imaging (Mri)	12
2.3.2	Ultrasound	13
2.3.3	Mammogram	13
2.4	Breast Structure	17
2.5	Breast Tumor	17
2.5.1	Benign Tumors	18
2.5.2	Malignant Tumors	18
2.5.3	Comparison of Benign and Malignant	18
2.5.4	Tumor Grades	19
	2.5.4.1 Metastatic Cancer	20
2.6	Existing Segmentation Techniques	20
2.6.1	Gray-Level Thresholding	21
2.6.2	Ostu's Method	21
2.6.3	Gaussian mixture method	21
2.6.4	Region Growing	22
2.6.5	Region Splitting and Merging	22
2.6.6	Edge Detection	22
2.6.7	Prewitt Edge Detection	23
2.6.8	Laplacian of Gaussian	23
2.6.9	Watershed	23
2.6.10	Fuzzy C- Mean Clustering Method	24
2.6.11	K-Means Clustering	24
2.6.12	Hierarchical Clustering	24
2.6.13	Mean-Shift	24
2.6.14	Level-Set Method	25
2.6.15	Artificial Neural Network	25
2.7	Region Growing	33
2.7.1	Region Growing in Brain MRI Segmentation	33

2.7.2	Region Growing in Road Segmentation	33
2.7.3	Summary of Region Growing	34
2.8	Evaluating Performance of Segmentation Algorithm	35
2.8.1	Dice Coefficient	35
2.8.2	Jaccard Index	36
2.8.3	Cosine Efficient	37
2.8.4	Comparison Between Matrices	37
2.9	Preprocessing	38
2.9.1	Mean Filtering	39
2.9.2	Median Filtering	39
2.9.3	Gaussian Filtering	40
2.9.4	Comparison of Filtering	40
2.10	Noise Estimation	40
2.11	Evaluating Performance of Filtering Algorithms	41
2.11.1	Peak Signal to Noise Ratio (PSNR)	41
2.11.2	Signal Noise Ratio (SNR)	42
2.11.3	Mean Absolute Error (MAE)	42
2.11.4	Root Mean Square Error (RMSE)	43
2.12	Classification of Tumor	43
2.13	Classification Techniques	44
2.13.1	SVM	45
2.13.2	Genetic Algorithm	45
2.13.3	KNN-Classifier	45
2.13.4	ANN	46
2.13.5	Summary of Advantages and Disadvantages Techniques	46
2.13.6	Summary of Breast Tumor Classification.	47
2.14	Bag of Features	48
2.14.1	Surf Point's Extraction	49
2.14.2	K-Means Clustering	51
2.15	Confusion Matrix	51
2.16	Image Analysis and Preparation	52
2.17	Summary	53
	CHAPTER 3 METHODOLOGY	55

3.1	Introduction	55
3.2	Proposed Algorithm	56
3.3	Preprocessing Step	57
	3.3.1 Median Filtering	57
3.4	Flow Of Segmentation With Region Growing	57
	3.4.1 Selection of Initial Seeds	58
	3.4.2 Growing Formula Based On Stopping Criterion	61
3.5	Procedure of Bag of Features Classification	62
	3.5.1 Categorization By SVM	63
	3.5.2 Training Dataset	64
	3.5.3 Testing Dataset	65
	3.5.4 Classification	65
3.6	Summary	66
CHAPTER 4 RESULTS AND DISCUSSION		67
4.1	Introduction	67
4.2	Experimental Setup	67
	4.2.1 The Dataset and Experimental Procedure	67
4.3	Preprocessing Result	68
	4.3.1 Discussion on Preprocessing	69
4.4	Segmentation Experimental Result	70
	4.4.1 Experiment 1: Malignant cases	70
	4.4.2 Experiment 2: Benign cases	72
4.5	Discussion on Segmentation	73
	4.5.1 Relationship Between DC Value and Median Filtering	76
	4.5.2 Compared with Previous Works	77
4.6	Classification Experimental Result	78
4.7	Discussion on Classification	79
	4.7.1 Result of Bag of Features Classification	79
	4.7.2 Misclassifying	80
	4.7.3 Comparison of Work	82

4.8	Summary	82
CHAPTER 5 CONCLUSION AND RECOMMENDATIONS		84
5.1	Conclusion	84
5.2	Contributions	85
5.3	Future Work	85
REFERENCES		87
APPENDIX A FULL RESULT OF SEGMENTATION AND CLASSIFICATION (MALIGNANT CASES)		101
APPENDIX B FULL RESULT OF SEGMENTATION AND CLASSIFICATION (BENIGN CASES)		104
APPENDIX C RESULT TO ESTIMATION NOISE USING MEDIAN FILTERING, MEAN FILTERING, GAUSSIAN FILTERING AND WITHOUT FILTERING		107
PUBLICATION		109

LIST OF TABLES

Table 2.1	Comparison of Mammogram, Ultrasound, MRI	14
Table 2.2	Comparison of Tumor's Type	18
Table 2.3	Grade of Tumors	19
Table 2.4	Summary of Segmentation Methodologies	26
Table 2.5	Comparison of Similar Work, Indicating Authors, Years of Publication, Modalities, Techniques, Experiment's Findings And Result.	29
Table 2.6	Summary of Researchers Implemented Region Growing From The Year of 2011 To 2015	34
Table 2.7	Comparison of Matrices	38
Table 2.8	Comparison of Filtering	40
Table 2.9	Advantages And Disadvantages of SVM, GA, KNN And ANN	47
Table 2.10	Summary of Breast Tumor Classification	48
Table 2.11	Advantages And Disadvantages of BOF	49
Table 3.1	Pseudocode of Region Growing	60
Table 3.2	Dataset of Training And Testing	63
Table 4.1	Estimation of Noise By Using Median, Mean And Gaussian Filtering	69
Table 4.2	Average Value of SNR, PSNR, RMSE, MAE	70
Table 4.3	Result of Segmentation	74
Table 4.4	Filtering And Dice Coefficient	76

Table 4.5	Comparison of Our Segmentation Results With Related Publications	77
Table 4.6	Confusion Matrix of BOF	79
Table 4.7	Result of Classification.	80
Table 4.8	Comparison of Proposed Method With Previous Work By Using MIAS Dataset	82

LIST OF FIGURES

Figure 1.1	Type of Image Engineering	2
Figure 1.2	Breast Cancer Total Cases From 1993 To 2003.	5
Figure 1.3	Process of Image Ultrasound	6
Figure 2.1	Ultrasound Image Indicating ROI	11
Figure 2.2	Ultrasound Scanning Process	15
Figure 2.3	Breast Structure	17
Figure 2.4	Categories of Breast Tumour	17
Figure 2.5	Cell Tissue Versus Cancer	19
Figure 2.6	Grade And Cancer	20
Figure 2.7	Methodologies of Image Segmentation	21
Figure 2.8	Flow of Classification	43
Figure 2.9	Feature Extraction	50
Figure 2.10	Confusion Matrix Table	51
Figure 2.11	Type of Image In Compression	52
Figure 3.1	Flow Chat of Proposed Algorithm (Segmentation)	56
Figure 3.2	Flow Chart of Proposed Algorithm (Classification)	56
Figure 3.3	5 X 5 Median Filtering	57
Figure 3.4	Step of Region Growing In Segmentation	58
Figure 3.5	Region Growing And Seed Selection	62
Figure 3.6	Classification With Bag of Features	64
Figure 3.7	Traning And Testing Data In Bag of Features	65
Figure 4.1	Median Filtering With 5x5 Sub-Regions.	68
Figure 4.2	Median Filtering With 5 X 5 Sub-Regions	68

Figure 4.3	Malignant Breast Tumour Segmented I	70
Figure 4.4	Malignant Breast Tumour Segmented II	71
Figure 4.5	Breast Carcinoma Segmentation	71
Figure 4.6	Normal Breast Segmented	72
Figure 4.7	Fibroadenoma Breast Segmentation	72
Figure 4.8	Ducts Breast Ultrasound	73
Figure 4.9	Time Graph For Tumor Segmentation	75
Figure 4.10	Benign Breast Classification's Histogram Frequencies.	78
Figure 4.11	Malignant Breast Classification's Histogram Frequencies	79
Figure 4.12	Misclassification of Tumor (Malignant Misclassified As Benign)	81
Figure 4.13	Misclassification of Tumor (Benign Misclassified As Malignant)	81

LIST OF ABBREVIATIONS

BoF	Bag of Features
CT	Compute Tomography
Db	Decibel
DC	Dice Coefficient
DCIS	Ductal Carcinoma In Situ
DICOM	Digital Imaging And Communications In Medicine
IPT	Image Processing Tool
JPEG	Joint Picture Experts Group
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
PNG	Portable Network Graphics
RGB	Rgb Color Model (Red, Green, Blue)
ROI	Region of Interest
RONI	Region of Non-Interest
SURF	Speeded Up Robust Features
XML	Extensible Markup Language
US	Ultrasound

LIST OF SYMBOLS

δ	Size of the regular grid
$\cos \theta$	θ is the measure of the angle
K	Size of the dictionary
ϵ	Region overlapped
J	Jaccard
QS	Quotient of similarity and ranges between 0 and 1.
T	Thresholding
$\tau(I)$	Noise level function
S_{xy}	Processing Window (Size)
Z_{\min}	Minimum Grey Level Value in S_{xy}
Z_{\max}	Maximum Grey Level Value in S_{xy}
Z_{med}	Median of Grey Levels in S_{xy}
Z_{xy}	Grey Level at Coordinates (x, y)
S_{\max}	Maximum Allowed Size of S_{xy}