UNIVERSITI TEKNOLOGI MARA

TECHNICAL REPORT

CLASSIFICATION OF MICROCALCIFICATION IN MAMMOGRAM IMAGES USING ENHANCED SUPPORT VECTOR MACHINE (ESVM)

P03M19

MUHAMMAD AKMAL FIRDAUS (2017412702)
SITI MADINAH SHADAN (2017412712)
NUR SYAHIZAH MOHD FAUDZAI (2017412726)

Report submitted in partial fulfilment of the requirement for the degree of
Bachelor of Science (Hons.) Mathematics
Faculty of Computer and Mathematical Sciences

JULY 2019

ACKNOWLEDGEMENT

IN THE NAME OF ALLAH, THE MOST GRACIOUS, THE MOST MERCIFUL

First and foremost, we would like to express our sincerest gratitude to Allah S.W.T the Almighty for showering us with all the blessings continuously, we will forever be grateful.

Nonetheless, we would like to express our deepest and sincerest gratitude to our respected supervisor Puan Aminah Abdul Malek, for providing invaluable guidance and encouraging suggestion throughout the completion of this research. Her dynamism, vision, sincerity and motivation have deeply inspired us. It was a great privilege and honour to work and study under her guidance. We would also like to thank to our Mathematical Modelling (MAT530) and Final Year Project (MSP660) lecturers, Dr. Maznah Banu Habiboo Raman and Dr. Mat Salim Selamat for their motivation and consistent advice and guidance.

We express our deep sense of appreciation to each of our dedicated group members as well as our classmates for giving full cooperation, extremely great commitment, timely assistance and incredible support throughout this work directly or indirectly. It is a pleasant occasion to show our courtesy in remembering our father, mother and fellow family members who showered us with all the love, prayers, care as well as the whole-hearted support extended to us throughout our studies. Let us also take this opportunity to thank Universiti Teknologi Mara (UiTM) Seremban for providing us a conducive environment and sophisticated facilities which helped us in the progress of completing the work assigned.

TABLE OF CONTENTS

ACKNOWLEDGEMENT i

TABLE OF CONTENTS ii

LIST OF TABLES iv

LIST OF FIGURES v

ABSTRACT vi

CHAPTER 1: INTRODUCTION 1

1.1 Overview 1

1.2 Introduction 1

1.3 Problem Statement 3

1.4 Objectives 4

1.5 Significance and Project Benefit 4

1.6 Scope and Limitations 5

1.7 Definition of Terms and Abbreviations 6

CHAPTER 2: BACKGROUND THEORY AND LITERATURE REVIEW 7

2.1 Overview 7

2.2 Feature Extraction 7

2.2.1 Wavelet Transform (WT) 7

2.2.2 Discrete Wavelet Transform (DWT) 8

2.2.3 Choice of Wavelet Transform Methods 10

2.3 Feature Reduction using Principal Component Analysis (PCA) 10

2.3.1 Geometric Interpretation of Principal Component Analysis (PCA) 11

2.4 Image Classification 12

2.4.1 Support Vector Machine (SVM) 13

2.4.2 Evolution and Application of Support Vector Machines (SVM) 14

2.4.3 Kernel Functions in Support Vector Machine (SVM) 17

CHAPTER 3: METHODOLOGY AND IMPLEMENTATION 18

3.1 Overview 18

3.2 Phase 1: Data Acquisition 19

3.3 Phase 2: Image Segmentation 20

3.4 Phase 3: Feature Extraction using DWT 30

3.5 Phase 4: Feature Reduction using PCA 40

3.6 Phase 5: Classification using SVM 43

3.7 Phase 6: Accuracy of Classification Results using ROC Curve 47
CHAPTER 4: RESULTS AND DISCUSSION 50
4.1 Overview 50
4.2 Classification Result by Using ESVM 50
4.2.1 Parameter Analysis of ESVM 54
4.3 Performance Evaluation 58
CHAPTER 5: CONCLUSION AND RECOMMENDATIONS 62
5.1 Overview 62
5.2 Conclusion 62
5.3 Recommendations 63
REFERENCES 64
APPENDICES 70
APPENDIX A: Source Code of ESVM Implemented in MATLAB R2016a 70
APPENDIX B: Another 14 Samples of Parameters Used in OT Method 72

ABSTRACT

Support Vector Machine (SVM) is a supervised machine learning algorithm with the ability to build a classification model from a labeled dataset. SVM has been broadly used in image classification of medical imaging such as mammogram images for breast cancer detection due to its higher classification precision, higher prediction accuracy, better generalization capability and better overall performance. Breast cancer detection is critically dependent on early detection and accurate diagnosis. However, the existed histopathological classification of breast cancer has clinical utility that is limited due to insufficient prognostic and predictive power. The diagnostic decisions by experienced physicians can be increased by an effective medical decision support system. Therefore, this research concentrates on Enhanced Support Vector Machine (ESVM) that combines Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) for better data classification accuracy. DWT is used for the extraction of statistical and textures features of mammogram image. Meanwhile, the employment of PCA is to reduce the dimensionality of the datasets to avoid overfitting thus better accuracy can be achieved. The ESVM focuses on classification of microcalcification in 30 mammogram images obtained from the National Cancer Society Malaysia (NCSM). The performance of the ESVM method is measured in term of classification rate (CR), specificity (SP), sensitivity (SV) and the accuracy based on Receiving Operating Characteristics (ROC) curve obtained by comparing the images classification results of ESVM and expert findings. The result revealed that the classification rate is 90% while for the accuracy based on the ROC Curve is 0.9375. Based on the result obtained, it is proven that ESVM is excellent in classifying the microcalcification in mammogram images and act as an assistance for radiologists that provides "second judgment" on medical image readings.