# Mass Detection and False Positive Reduction in Mammographic Images

A Thesis submitted for the degree of

**Doctor of Philosophy**

By

**Man To Wong**

In

Faculty of Engineering and Information Technology

School of Computing and Communications

**UNIVERSITY OF TECHNOLOGY SYDNEY**

**AUSTRALIA**

Submitted DECEMBER 2015

**UNIVERSITY OF TECHNOLOFY SYDNEY**
**SCHOOL OF COMPUTING AND COMMUNICATIONS**

The undersigned hereby certifies that he has read this thesis entitled "**Mass Detection and False Positive Reduction in Mammographic Images"** by **Man To Wong** and that in his opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

Principal Supervisor

Prof. Xiangjian (Sean) He

# CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

---

Signature of Author

# Abstract

## Mass Detection and False Positive Reduction in Mammographic Images

Breast cancer is the most common type of cancer for women in America. Currently the most effective method for early detection of breast cancer is mammography. Mammography is the only widely accepted imaging method used for routine breast cancer screening. Masses are one of the important signs of breast cancer. However it is difficult to detect masses because masses have different size and shape and their features can be obscured or similar to the normal breast parenchyma. Reading mammograms is a demanding job for radiologists. A computer aided detection (CAD) system can provide a consistent second opinion to a radiologist and greatly improve the mass detection accuracy.

In this thesis, a computer aided detection system is developed which can segment the breast region from the background in the whole mammographic image, detect the suspicious regions from the breast region and then classify the suspicious regions to mass or normal breast tissue. The suspicious regions in the full mammographic image can be found by contrast limited adaptive histogram equalization and thresholding. These suspicious regions can be masses or normal breast tissue (false positives). To reduce the number of false positives in mass detection, a feature selection and classification approach using particle swarm optimization (PSO) and support vector machine (SVM) is proposed. Firstly, texture features are derived from the gray level co-occurrence matrix (GLCM) of each suspicious region. A PSO and SVM based feature selection is proposed to determine the significant features. The significant features found by PSO-SVM based feature selection are used by the SVM classifier to classify the suspicious region to mass or normal breast tissue.

One advantage of the proposed mass detection system is that it can detect different types of masses, including spiculated, circumscribed and ill-defined masses from the whole mammographic image. The number of false positives in mass detection can be reduced by the PSO and SVM based feature selection and mass classification method proposed

in this thesis. Experimental results show that the proposed PSO-SVM based feature selection technique can find the significant features that can improve the classification accuracy of SVM and perform better than other widely used feature selection methods. The proposed mass classification approach using PSO and SVM has better or comparable performance when compared to other state-of-the-art mass classification techniques, using sensitivity and specificity as the evaluation criteria.

In order to perform accurate image segmentation of the mass from the suspicious region, a mass segmentation method by PSO based image clustering is proposed. Two new fitness functions are proposed which can improve the performance of image clustering by generating more compact clusters and larger inter-cluster distance. The proposed PSO based image clustering, with the new fitness function, can improve the segmentation of the mass from mammographic image. It has been shown experimentally that PSO based image clustering can have better mass segmentation performance when compared to K-means, a widely used clustering technique.

# Acknowledgement

This research would not have been possible without the guidance and the help of many people. My gratitude to my supervisor, Prof. Xiangjian He, for his guidance, support and encouragement. His comments and suggestions during preparation of this thesis have been valuable.

I also appreciate the following people for providing various assistance for the completion of this research work: Zhiyuan Tan, Aruna Jamdagni, Chao Zeng, Ruo Du, Muhammad Abul Hasan, Sheng Wang, Prof. Hung Nguyen, Dr Wenjing Jia, Dr Qiang Wu, Dr Min Xu and Prof. W.C Yeh.

Last but not the least, I would like to express my love and gratitude to my daughter Katherine for her understanding and tolerance so that I can concentrate on this thesis.

# Author's Publication for the Ph.D

## Published papers

### Conference papers

1. M. T Wong, X. He, W. C Yeh, Z. Ibrahim, and Y. Y Chung. "Feature Selection and Mass Classification Using Particle Swarm Optimization and Support Vector Machine". In 21st International Conference on Neural Information Processing (ICONIP 2014), Part III, Lecture Notes in Computer Science, Vol. 8836, pages 439-446, Springer International Publishing, Kuching, Malaysia, November 3-6, 2014. (Tier A Conference)
2. M. T Wong, X. He, H. Nguyen, and W.C Yeh. "Mass Classification in Digitized Mammograms Using Texture Features and Artificial Neural Network". In 19th International Conference on Neural Information Processing (ICONIP 2012), Part V, Lecture Notes in Computer Science, Vol. 7667, pages 151-158, Springer-Verlag, Berlin, Heidelberg, Doha, Qatar, November 12-15, 2012. (Tier A Conference)
3. M. T Wong, X. He, H. Nguyen, and W. C Yeh. "Particle Swarm Optimization Based Feature Selection in Mammogram Mass Classification". In 2012 International Conference on Computerized Healthcare (ICCH), pages 152-157, Hong Kong, December, 2012.
4. M. T Wong, X. He, and W. C Yeh. "Image Clustering Using Particle Swarm Optimization". In 2011 IEEE Congress on Evolutionary Computation (CEC), pages 262-268, New Orleans, USA, June, 2011. (Tier A conference)

# Table of Contents

# List of Tables

# List of Figures