Breast Cancer Risk Assessment using Mammographic Image Texture Analysis

A thesis submitted for the degree of
Doctor of Philosophy

By

Xi-Zhao Li

School of Computer Science, Engineering and Mathematics
Faculty of Science and Engineering
Flinders University

November, 2014

Contents

List of Figures

List of Tables

List of Abbreviations

BSSA —– BreastScreen South Australia

CC —– craniocaudal (top) view of the breast

DDSM —– Digital Database of Screening Mammograms

ER —– Estrogen-Receptor

FNF —– false negative fraction

FPF —– false positive fraction

MLO —– mediolateral oblique (side) view of the breast

MR8 —– Maximum Response 8 filter bank

ROC —– the recever operating characteristic

ROIs —– regions of interest

SCC —– six-category classification

SD —– Standard deviation

TNF —– true negative fraction

TPF —– true positive fraction

Summary

Breast cancer is one of the most common cancers among women and early detection plays an important role in reducing the mortality and morbidity due to breast cancer. Importantly, early breast cancer detection is facilitated by accurate breast cancer risk assessment. This thesis aims to develop computer methods for analyzing tissue texture in screening mammograms in order to assess the risk of breast cancer.

According to the literature, the breast density is a strong indicator of breast cancer risk and is independent of non-mammographic risk factors (age, race, family history, etc.). In addition, texture from screening mammograms is also considered to play an important role in predicting breast cancer risk. However, the contribution of texture alone to breast cancer risk is unclear and the role of texture for assessing breast cancer risk over time is also unknown. The focus of this thesis is on studying the role of texture, independent of density, in breast cancer risk assessment.

In this thesis, the emphasis is on characterizing texture through the use of textons. Textons can be described as ubiquitous local texture patterns. The distribution of conventional textons (referred to as first-order textons in this thesis) has been shown to characterize texture in visual images and has been successful in tasks such as separating regions corresponding to grass from regions representing trees or animals. An important contribution of this thesis is the introduction of higher-order textons. The notion of higher-order textons is to extend the power of the first-order textons. Higher-order textons allow quantitative analysis of commonly occurring patterns of patterns, offering a mechanism for understanding more complex texture structure in images. In this thesis, textons and higher-order textons are used to distinguish mammograms from women having a high risk of breast cancer from women having a low risk of breast cancer.

A number of experiments were conducted to determine the best implementation of textons and higher-order textons for breast cancer risk assessment. Results indicate that texture analysis based on higher-order textons predicts risk at least as well as any method currently available for estimating breast cancer risk from mammograms. Risk of breast cancer can be measured using texture at least four years prior to the cancer becoming apparent mammographically.

In addition, a number of discoveries were made in the course of the study. Tex-

ture features from CC view mammograms (top view) perform better than texture features from MLO view mammograms (side view). Better risk assessment is obtained by measuring texture over the full breast than any particular local region of the breast. Texture features calculated from 3×3 local neighborhoods perform as good or better than texture features based on larger patches. Texture information relevant to breast cancer risk is more pronounced in the breast in which cancer eventually occurs than in the breast without known cancer of the same woman. These discoveries have potential impact on the fields of image analysis and computer-aided mammography and so form natural seeds for future work.

Publications arising from the Study

Referred Conference Paper

[1] Xi-Zhao Li, Simon Williams, and Murk J. Bottema. Intensity independent texture analysis in screening mammograms. In *11th International Workshop on Breast Imaging, IWDW2012, Philadelphia, PA, USA,* pages 474-481, July 2012.

[2] Xi-Zhao Li, Simon Williams, Gobert Lee, and Min Deng. Computer-aided mammography classification of malignant mass regions and normal regions based on novel texton features. In *12th International Conference on Control, Automation, Robotics and Vision, Guangzhou, China, ICARCV2012,* pages 1431-1436, December 2012.

[3] Xi-Zhao Li, Simon Williams, Peter Downey and Murk J. Bottema. Temporal breast cancer risk assessment based on higher-order textons. In *12th International Workshop on Breast Imaging, IWDW2014, Gifu, Japan,* pages 565-572, June-July 2014.

Referred Journal Paper

[4] Xi-Zhao Li, Simon Williams and Murk J. Bottema. Background intensity independent texture features for assessing breast cancer risk in screening mammograms. *Pattern Recognition Letters,* 34(9):1053-1062, Feb 2013.

[5] Xi-Zhao Li, Simon Williams and Murk J. Bottema. Texture and region dependent breast cancer risk assessment from screening mammograms. *Pattern Recognition Letters,* 36(15):117-124, Jan 2014.

[6] Xi-Zhao Li, Simon Williams and Murk J. Bottema. Constructing and applying higher order textons: Estimating breast cancer risk. *Pattern Recognition,* 47(3):1375-1382, Mar 2014.

Declaration

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Xi-Zhao Li, Candidate

Murk J. Bottema, Principal Supervisor
Simon Williams, Co-supervisor

Acknowledgements

I would like to thank my supervisor, A/Prof. Murk J. Bottema for his support and encouragement throughout my PhD study. Importantly, he helped me improve my research skills in medical image analysis, academic writing and mathematical theories for my project. Specifically, I really thank him for giving me the opportunity to study in Flinders University.

I would like to also thank my co-supervisor, Dr. Simon Williams for his guidance and supervision throughout my PhD study. Specifically, he provided me great help in using Lyx for academic writing and opportunity for practicing teaching.

I thank Dr. Gobert Lee for her support in academic fundamental knowledge, suggestions for my research and guidance in using university resources for doing research. Special thanks to Dr. Mariusz Bajger for his technical help; Dr. Ray Booth and A/Prof. Alan Branford for giving me the opportunity of doing some mathematics and statistic related part-time jobs in the school. A special thank to Dr. Shu-Chuan Chu for her support in my study as a friend.

Special thanks to Dr. Adham Atyabi for his help in using the super computer for running my time consuming programs. Particularly, I thank Dr. Adham Atyabi for all his precious suggestions of doing research.

I would also like to thank BreastScreen South Australia (BSSA) for providing mammogram images for my final stage research.

Gratefully, I thank China Scholarship Council (CSC) and Flinders University for all their financial support for my whole study.

Finally, I thank all my family for their great support and love all the way through.