UNIVERSITI PUTRA MALAYSIA

DYNAMIC RULE REFINEMENT STRATEGY OF ASSOCIATIVE CLASSIFIER FOR EFFECTIVE MAMMOGRAPHIC IMAGE CLASSIFICATION

NIRASE FATHIMA ABUBACKER

FSKTM 2016 31

DYNAMIC RULE REFINEMENT STRATEGY OF ASSOCIATIVE CLASSIFIER FOR EFFECTIVE MAMMOGRAPHIC IMAGE CLASSIFICATION

By

NIRASE FATHIMA ABUBACKER

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfillment of the Requirement for the Degree of Doctor of Philosophy

September 2016

DEDICATION

My Beloved Parents

Abstract of the thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirement for the degree of Doctor of Philosophy

DYNAMIC RULE REFINEMENT STRATEGY OF ASSOCIATIVE CLASSIFIER FOR EFFECTIVE MAMMOGRAPHIC IMAGE CLASSIFICATION

By

NIRASE FATHIMA ABUBACKER

September 2016

Chairperson : Azreen Azman, PhD
Faculty : Computer Science and Information Technology

Computer-aided diagnosis (CADx) has gained significant attention in helping radiologists in the interpretation of mammograms to assist in diagnostic decision-making. A more effective CADx increases the probability of cure. An effective mammogram classification technique benefit to the research of computer aided mammography for a better diagnostic assistance. However, the effectiveness of classifiers depends on the training data sets that are often small in data size and static, which does not adapt to changes. The main aim of this thesis is to propose an effective associative classifier using rule refinement technique that adapts changes in databases for building an effective CADx model in the classification of mammogram images.

The classifiers using Association Rule (AR) mining gain popularity compared to traditional classifiers due to their nature in reflecting close dependencies among single or multiple features for composing rules with its excellent interpretation. The existing associative classification techniques that are used in Computer Aided Diagnosis (CADx) have proved their efficiency in mammogram classification. The research aims to propose an improved associative classification model with its first step preprocessing that uses segmentation technique with filter that includes certain areas of the image for mammogram peripheral enhancement. The feature extraction is used to extract the most prominent features from mammogram images that represent various classes of the images to be used by classification techniques. A feature selection technique named Correlation Feature Selection (CFS) that involves a heuristic search is adopted for dimensionality reduction of feature space to improve efficiency and at the same time maintain the effectiveness of classification. The thesis discovers useful and interesting relations between features and class in the form of rules to build an efficient associative classifier from a large collection of mammogram images using association rule mining technique. An Associative Classifier that uses rules Highest Average Confidence (ACHAvC) is proposed for an effective classification of mammography. The classifier ACHAvC has achieved high

accuracy of 90% and specificity of 90%, however the sensitivity is 78.5% and not commendable in medical domain.

The effectiveness of an associative classifier depends largely on the generated rules based on training data. In previous works such as HiCARe, SACMINER, MINSAR, including ACHAvC the training data have been limited, which may produce the classification rules that are static and cannot adapt to a changing charecteristic of test images, as such it may not produce complete and accurate rules for classification. The classification performance can be further improved if the static rules are updated dynamically. The availability of radiologist ground truth for every case could be used to validate the classification result and refine the set of rules generated. A method Rule Refinement based on Incremental Modification (RRIM) is proposed that dynamically refines the rules every time when it is validated with the experts ground truth. As such these refined rules that adapt the changes in the data are then used for classification to further enhance the performance of the classifier ACHAvC with a reduced minimal error and with improved prediction accuracy.

The Performance of the proposed methods are evaluated for accuracy, sensitivity and specificity for the mammogram image data set, taken from the digital database for mammography from the University of South Florida, *Digital Database for Screening Mammography* (DDSM). The proposed method has achieved an overall classification accuracy of 96%, with sensitivity 92.56% and specificity 96.94% in testing stage which is comparatively better than the three benchmark approaches HiCARe, SACMINER, MINSAR that are chosen for the proposed research with accuracy of (91%, 85%, 79%), sensitivity (95%, 84%, 87%) and specificity (84%, 86%, 67%).

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

DYNAMIC RULE REFINEMENT STRATEGY OF ASSOCIATIVE CLASSIFIER FOR EFFECTIVE MAMMOGRAPHIC IMAGE CLASSIFICATION

Oleh

NIRASE FATHIMA ABUBACKER

September 2016

Pengerusi : Azreen Azman, PhD
Fakulti : Sains Komputer dan Teknologi maklumat

Bantuan diagnosis komputer (CADx) mendapat tumpuan penting bagi membantu ahli radiologi mentafsir mamogram untuk membuat keputusan diagnostik. CADx yang lebih efektif meningkatkan kebarangkalian untuk tahap kesembuhan. Teknik pengkelasan mamogram yang efektif memberi manfaat kepada penyelidikan mamografi berasaskan komputer untuk memberi bantuan diagnostik yang lebih baik. Walau bagaimanapun, keberkesanan pengkelasan ini bergantung kepada set data latihan yang sering kali bersaiz kecil dan statik, yang mana tidak peka kepada perubahan. Tujuan utama tesis ini adalah untuk mencadangkan satu pengkelasan menggunakan teknik peraturan perbaikan bersekutu yang peka kepada perubahan yang berlaku di pangkalan data bagi membina model CADx yang berkesan untuk klasifikasi imej mamogram.

Pengkelasan menggunakan Association Rule (AR) lebih terkenal berbanding pengkelasan secara tradisional di mana sifatnya yang mencerminkan kebergantungan antara ciri-ciri tunggal atau pelbagai untuk menghasilkan peraturan dengan tafsiran yang sangat baik. Langkah proses awal menggunakan teknik segmentasi dengan penapis yang merangkumi bahagian-bahagian tertentu imej untuk meningkatkan periferal mamogram. Ciri pengekstrakkan digunakan untuk ekstrak ciri-ciri yang paling tepat dari imej mamogram yang mewakili pelbagai kelas imej untuk digunakan oleh teknik pengkelasan. Teknik pemilihan ciri bernama Korelasi Ciri Pemilihan (CFS) yang melibatkan carian heuristik digunakan untuk pengurangan dimensi ruang ciri bagi meningkatkan kecekapan dan pada masa yang sama mengekalkan keberkesanan klasifikasi. Tesis mendapati hubungan yang berguna dan menarik antara ciri-ciri dan kelas untuk membina pengkelasan bersekutu yang efektif daripada koleksi imej mamogram yang banyak dengan menggunakan teknik penyatuan. Pengkelasan bersekutu yang menggunakan peraturan Highest Average Confidence (ACHAvC) dicadangkan untuk klasifikasi mamografi yang efektif. Pengkelasan ACHAvC mencapai 90% ketepatan dan 90% kekhususan, akan tetapi kepekaan yang agak kurang iaitu 78.5% dan tidak diterima di bidang perubatan.

Keberkesanan pengkelasan bersekutu bergantung kepada peraturan yang dihasilkan berdasarkan data latihan. Dalam kerja-kerja penyelidikansebeleum ini seperti HiCARe, SACMINER, MINSAR termasuk ACHAvC, data latihan adalah terhad, yang mana boleh menghasilkan peraturan klasifikasi yang statik dan tidak peka kepada ciri-ciri imej ujian yang berubah-ubah, oleh itu ia tidak boleh menghasilkan peraturan yang lengkap dan tepat untuk pengkelasan. Prestasi klasifikasi boleh diperbaiki jika peraturan statik dikemaskini secara dinamik. Kepakaran ahli radiologi boleh digunakan bagi setiap kes untuk megesahkan hasil klasifikasi dan menghalusi set peraturan yang dijana. Kaedah Rule Refinement berdasarkan Incremental Modification (RRIM) dicadangkan untuk memperbaiki peraturan secara dinamik setiap kali ia disahkan oleh pakar. Oleh itu peraturan-peraturan yang diperbaiki ini akan peka kepada perubahan yang berlaku pada data dan ianya digunakan untuk meningkatkan lagi prestasi pengkelasan ACHAvC dengan ralat yang minimum dan ketepatan ramalan yang lebih baik.

Prestasi kaedah yang dicadangkan dinilai dengan ketepatan, kepekaan dan kekhususan terhadap set data imej mamogram, yang diambil dari pangkalan data digital mamografi dari University of South Florida, Pangkalan Data Digital untuk Saringan Mamografi (DDSM). Kaedah yang dicadangkan telah mencapai ketepatan pengkelasan keseluruhan 96%, dengan 95.56% kepekaan dan kekhususan 96.94% dalam peringkat ujian di mana ia jauh lebih baik berbanding tiga kaedah lain yang diguna sebagai penanda aras yang mana ketepatannya (91%, 85%, 79%), kepekaan (95%, 84%, 87%) dan kekhususan (84%, 86%, 67%).

ACKNOWLEDGEMENTS

Praise be to Allah The Almighty, Lord of the universe for His blessings into mankind and guidance. I bow before Almighty Allah and express my humblest and sincerest words of gratitude to Him, Who bestowed upon the feeble author the potential and ability to make a material contribution to the already existing ocean of knowledge. My unreserved gratitude and appreciation to my father late E. Abubacker, may his soul rest in peace and my beloved mother Saffura Bi for their moral upbringing, financial and intellectual courage to successfully guide and support my academic careers.Also I express my hearty thanks to my brothers and my sister, who motivated throughout my study.

No words could effectively express my sincere thanks and appreciation to my supervisor, Dr. Azreen Azman for his kind assistance, readily available support, patience, encouragement, critical advice, suggestions, intellectual guidance and understanding throughout my study in Malaysia.

I am also greatly indebted to my Co-supervisors, Associate Professor. Dr. Masrah Azrifah Azmi Murad and Associate Professor Dr. Shyamala Doraisamy for their valuable guidance and encouragement throughout my Ph.D.

I extend my special thanks to my friends Dr Sankaraiah, Mohammed El Thahir Manna and Dr.Rekha SaravanaPerumal for their invaluable help during my experiments and thesis writing whenever I needed them and without those help, this work was not possible to be completed. I express my sincere thanks to my friends Leila Mansourian, Prem Kumar and my academic leader Mr. Nicholas Geremy for their valuable support.

I wish to express my deepest appreciation to numerous people who walked with me along the journey of this study and thesis preparations especially my brothers and sister. I enjoyed my time spent in Malaysia thoroughly and I would cherish these memories for the rest of my life.

I find no words to thank the patience and unconditional love and support of my husband Gulam Hussain and my children Mohammed Raseen and Mufliha Sabreen during my entire PhD studies. My heartfelt thanks goes to my husband whose sacrifices, support, love, caring inspired me to overcome all the difficulties throughout my academic life. This dissertation process would not be successful without having his patience, love, and dedication.

At last but not the least, I wish to thank The Ministry of Higher education (MOHE), Malaysia for granting me a scholarship for direct contribution in my Ph.D. study at Universiti Putra Malaysia (UPM), Malaysia.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Doctor of Philosophy. The members of the supervisory committee were as follows:

Azreen Azman, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Masrah Azrifah Azmi Murad, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Shyamala Doraisamy, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

____
BUJANG KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- This thesis is my original work;
- Quotations, illustrations and citations have been duly referenced;
- This thesis has not been submitted previously or concurrently for any other agree at any other institution
- Intellectual property of the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- Written permission must be obtained from the supervisor and the office of Deputy Vice Chancellor (Research and Innovation) before the thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other material as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- There is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ______________________________ Date: __________________

Name and Matric No.: Nirase Fathima Abubacker, GS32867

Declaration by Members of Supervisory Committee

This is to confirm that:

- The research conducted and the writing of this thesis was under our supervision;
- Supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ____
Name of
Chairman of
Supervisory
Committee: Dr. Azreen Azman

Signature: ____
Name of
Member of
Supervisory
Committee: Associate Prof. Dr. Masrah Azrifah Azmi Murad

Signature: ____
Name of
Member of
Supervisory
Committee: Associate Prof. Dr. Shyamala Doraisamy

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACR	American College Of Radiology
ACS	American Cancer Society
AMA	American Medical Association
BI	Blurred Image
CAD	Computer Aided Diagnosis
CADe	Computer Aided Detection
CADx	Computer Aided Diagnosis
CC	Cranio-Caudal
NN	Neural Network
DDSM	The Digital Database For Screening Mammography
FN	False Negative
FP	False Positive
GLCM	Gray Level Co-Occurrence Matrix
JPEG	Joint Photographic Expert Group
MIAS	The Mammographic Image Analysis Society
MLO	Mediolateral-Oblique
MRI	Magnetic Resonance Imaging
NCI	National Cancer Institute
NTP	Normalized Thickness Profile
ROC	Receiver Operating Characteristic
ROI	Region Of Interest
SI	Segmentation Image
SVM	Support Vector Machine
TN	True Negative
TP	True Positive

CHAPTER 1

INTRODUCTION

1.1 Introduction

Breast cancer is the most common leading cancer among women worldwide. In the western and developing countries, including Malaysia, breast cancer has become a major health problem. The most recent estimate by Forouzanfar et al. (2010) indicates that more than 1.6 million new cases of breast cancer occurred among women worldwide in 2010. Early detection of breast cancer is possible only when regular screening examinations are carried out (Anderson & Jakesz, 2008). All types of breast cancer diagnosis depend on a biopsy using mammography findings. A mammogram is a special type of x-ray photograph that uses high-resolution film, high contrast and low dose x-ray for imaging the breasts (Vainio & Bianchini, 2002). This helps to detect and diagnose breast cancer effectively in its early stage. Digital Mammography is the most excellent gold standard screening method for breast cancer detection in its early stage (Vainio & Bianchini, 2002). Several studies have found that women with a family history may benefit from regular breast cancer screening, reporting higher cancer detection rates (Halapy et al., 2004; Kerlikowske et al, 2000). A mortality benefit from screening mammography in women at average risk of developing breast cancer has been established (Nelson et al., 2009; Canadian Task Force on Preventive Health Care, 2011); Studies have demonstrated that digital mammography has higher cancer detection rates in women who are more likely to have their cancer missed by screen-film mammography (Burrell et al., 1996; D'Orsi & Newell, 2007; Ikeda et al., 1992; Rosenberg et al., 1998). However, not all breast cancers are detectable accurately using mammograms even though they are very sensitive due to its difficulty in reading that requires abundant experience (Joseph Y.L, 1999). Several researches have shown that 20% to 40% of breast cancer detection failure rate is due to image complexity structure and radiologist fatigue (Harvey J et.al, 1993; Beam.C et.al, 1996; Elmore.J et.al, 1994). Due to this, about 65% of cases that are referred to surgical biopsy are just the normal (Kopanas, 1992; Knutzen, 1993). Additionally, mammogram interpretation is a high demanding job. About 10% of normal mammograms are misidentified by physicians as abnormal leading to more stressful tests and unnecessary diagnostic procedures for a normal patient moreover the misinterpretation of abnormal mammograms results in high rate of death (Jackson, 1993).

Computer-aided diagnosis (CADx) has gained significant attention in helping radiologists in the interpretation of mammograms to assist in diagnostic decision-making (Yusof, N.M et.al, 2007). A more effective CADx invariabily increases the probability of cure. In particular, the CAD system for automated detection/classification could provide a second opinion that improves the chances of detecting tumors and also reduces the human workload associated with the diagnosis. Such a system is capable to automatically classify and suggest the pathological terms for a new mammography image. These suggested terms are presented to doctors as additional information to assist them in the diagnosis of breast cancer. As such, it is vital to select a tool for revealing unknown information using the features extracted from images. An effective mammogram classification technique benefit to the research of computer aided mammography for a better diagnostic assistance. The classifiers using the Association Rule (AR) mining gain popularity compared to traditional classifiers due to their nature in reflecting close dependencies among single or multiple features for composing rules with its excellent interpretation. However the classification rules are stagnant on the training set and not adaptable to a changing distribution of test images. The classification performance can be improved if the static rules are updated dynamically. The main aim of this thesis is to propose an effective classifier using rule refinement technique that updates rules incrementally to adapt changes in databases. As such the database becomes more knowledgeable and helps in optimizing the performance of an associative classifier, thus can be used in a hierarchically CADx system that classify the mammogram images firstly according to its features, after indicates if there is a lesion and finally what kind of lesion is.

This chapter consists of five sections. The Section 1.2 outlines the motivation to carry out this work and the background for the research. Subsequently, In Section 1.3, the research problem statement is presented. The Aim and objectives of the research are explained in the sections 1.4 and 1.5 respectively. In addition, section 1.6 presents the significance of the research work. Following this, the sections 1.7 and 1.8 presents the research contributions and organization of this work.

1.2 Motivation of the work

The advancement in digital images has created a new breakthrough in every field. Medical imaging has developed into one of the most important fields due to its continuous growth in computerised medical image visualisation and advances in analysis methods. However, exploring the ever increasing quantities of medical images for manual diagnosis is cumbersome and time consuming. Also, factors like image quality and eye fatigue may affect the diagnostic results. Furthermore, most often radiologists have to deal with urgent cases where they need to analyze and evaluate the images comprehensively in a shorter period of time, thus abnormality may be unobserved (Sampat et.al, 2005).

Assisting the radiologists to classify conspicuous cases may help to improve accuracy in the interpretation of medical images and deliver a better diagnosis. Two different radiologists interpretation of the same mammogram could enhance the prediction accuracy of mammographic screening more than 30% (Skaane, P., et.al, 2013). Nevertheless, this double reading could increase the radiologists cost, work load and makes it unfeasible. Also there are possibilities for a slight overlap on different judgements given by different radiologists for the same diagnosis. Hence, by making use of computer based systems along with the knowledge of radiologist's feedback could be useful to detect the abnormalities and thus making better diagnostic decisions; this process is called Computer Aided Diagnosis (CADx). Therefore, it is important to develop a computer-aided diagnostic system to aid radiologists in finding the abnormalities effectively and to reduce the number of unnecessary biopsies.

1.3 Problem Statement

Many CADx based systems were used by the radiologist for the diagnosis of breast masses. In order to automate the classification system for CADx, there are some limitations in the process of breast mass diagnosis (Jalalian. A et.al, 2013). Accurate segmentation of a breast mass could be one of the important step for the diagnosis of breast cancer in mammography. The mammograms may contain some labels namely the background and that must be removed before subsequent tasks such as feature extraction and classification step. In full-field digital mammograms background region is totally black and uniform. Therefore, the background is composed of all pixels with an intensity equal to zero. However, mammogram contains many artifacts (Mustra, M., & Grgic, M. , 2013**)** and due to this nature the dark regions of the border might be set to background which are likely to be ignored during the segmentation process. As such these segmentation techniques may result in some missed parts of the peripheral region when applied to mammogram images. Hence a filter that scales the gray level of the border regions to brighten the mammogram image is required to enhance the details of images and helps to obtain a better segmented output.

A mammogram uses a machine that takes a lower dose x-ray by flattening the breast between 2 plates to look only at the breast tissue that spreads apart. However, this flattening of the breast is subject to the deformation due to heavy force applied during compression that leads to a difference in the thickness of the breast (Kallenberg, Michiel GJ, et al., 2012). Hence a smoothly varying correction function is required to expand the perceptibility of the peripheral area. Peripheral enhancement is a technique t+hat greatly reduces the dynamic range of the mammogram (He, W. et.al, 2014).

In the context of mammogram classification, there are some traditional classification methods such as Decision Trees (Devi R.D.H et.al, 2015); Sequential Minimal Optimization (SMO) (Sharma, V., & Singh, S., 2014), Radial Basis Function (Pratiwi, M., Harefa, J., & Nanda, S., 2015) proved with good prediction accuracies. However, a few other traditional classification methods produced using neural network and probabilistic approaches, are difficult to understand (Setiawan, A. S., Wesley, J., & Purnama, Y. 2015). Methods using Naive Bayes generate strong feature independence assumptions (Karabatak, M., 2015). Therefore, the prediction accuracies using traditional methods may not be commendable. Hence there is a need to build an efficient classifier that generates strong associations between features and reveals hidden relationship that can be missed by other classification algorithms. The existing associative classification techniques HiCARe (Riberio et.al, 2008), SACMINER (Watanabe et.al, 2011), MINSAR (Traina AJ et.al, 2012), that are used in Computer Aided Diagnosis (CADx) are static and do not adapt to changes in the database over the time. The classification association rules generated using training samples form the base knowledge to come up with a target output for new images. However, these existing rules are stagnant and cannot adapt to a changing distribution of test images. Hence the process of knowledge (rule) refinement is required to have an effective knowledge base that dynamically adapts to changes for a more accurate classification. The information about the lesions for each patient that are provided in the ground truth by the experts are used for rule refinement.

1.4 Research Aim

The main aim of the proposed research work is to propose an efficient classifier in order to build an effective CADx model for the classification of mammogram images based on refined classification association rules with the validated apperances of instances over time.

1.5 Research Objectives

The primary focus of this research work is to propose methods and algorithms for improving the performance of mammogram classification. The following objectives are set to accomplish this.

1. To improve an effective segmentation technique with a filter that includes specific areas of the image for mammogram peripheral enhancement.
2. To propose a classification model using generated association rules for an effective mammogram image classification of new instances.
3. To propose a dynamic rule refinement strategy of the association rules based on experts ground truth for an effective classification of mammography images

1.6 Significance of Study

This study will be a significant endeavour in promoting assistance to radiologists in their interpretation of mammogram images with the following benefits.

- Factors like image quality, eye fatigue do not affect the diagnostic results.
- Reducing the number of missing detection cases, combined with an enhanced technique have resulted in the effective diagnosis and potential survival.
- Evaluates the mammogram images comprehensively in a short time.
- Easy analysis of disease specific information among patients using different modalities is possible.
- Adapt to changes by learning from prior known cases to arrive at correct conclusions

- Automated classification technique of abnormalities assists radiologists in their interpretation of mammogram images and overcomes the inconsistencies in manual grading.

1.7 Research Contributions of the Thesis

The technical and social contributions of the proposed research are as follows:

- Enhancing the preprocessing technique using peripheral enhancement that uses the modified segmentation technique
- Proposed a new Associative Classifier that uses a predefined weight as highest average confidence based on a category for classification decision.
- Enhancing the Association Rule based classifier that dynamically refines the classification rules

1.8 Organization of the Thesis

The Organization of the thesis comprises of eight chapters and shown in Figure 1.1. The detailed descriptions of each chapter are provided below:

Chapter 1 presents the introduction, motivation, aim, objectives, significance, contribution of the research and the organization of the thesis.

Chapter 2 discusses appropriate related works for Segmentation, Enhancement, Feature Extraction methods, Feature Selection methods, Data Mining Algorithm techniques, and Classification. Soft computing theories such as rule refinement techniques increment learning are also presented in this chapter.

Chapter 3 presents the overall implementation methodology with a detailed framework for the interpretation of the proposed automated classification of mammogram.

Chapter 4 presents the preprocessing, feature extraction and feature selection processes to identify more discriminating features irrespective of its excessive dimensionality. The selected features are effective in mammogram diagnosis to facilitate the improvement in the performance of a medical diagnostic process in minimizing the failure rate of the diagnostic process. Also the proposed system results are compared with the existing methods.

Chapter 5 focuses on generating Association Rules (AR) for an effective construction of different classifiers using association rules. Generated rules are used to build a classification model and is evaluated to study its predictive capability. The proposed Associative Classifier using Highest Average Confidence (ACHAvC) is compared with other traditional classifiers.

Chapter 6 focuses on association rule refinement for the performance optimization of classification models. The refined rules are used to optimize a classification model, which is evaluated with different prediction models to study its predictive capability.

Chapter 7 provides the conclusions of the research work and future work is described briefly.

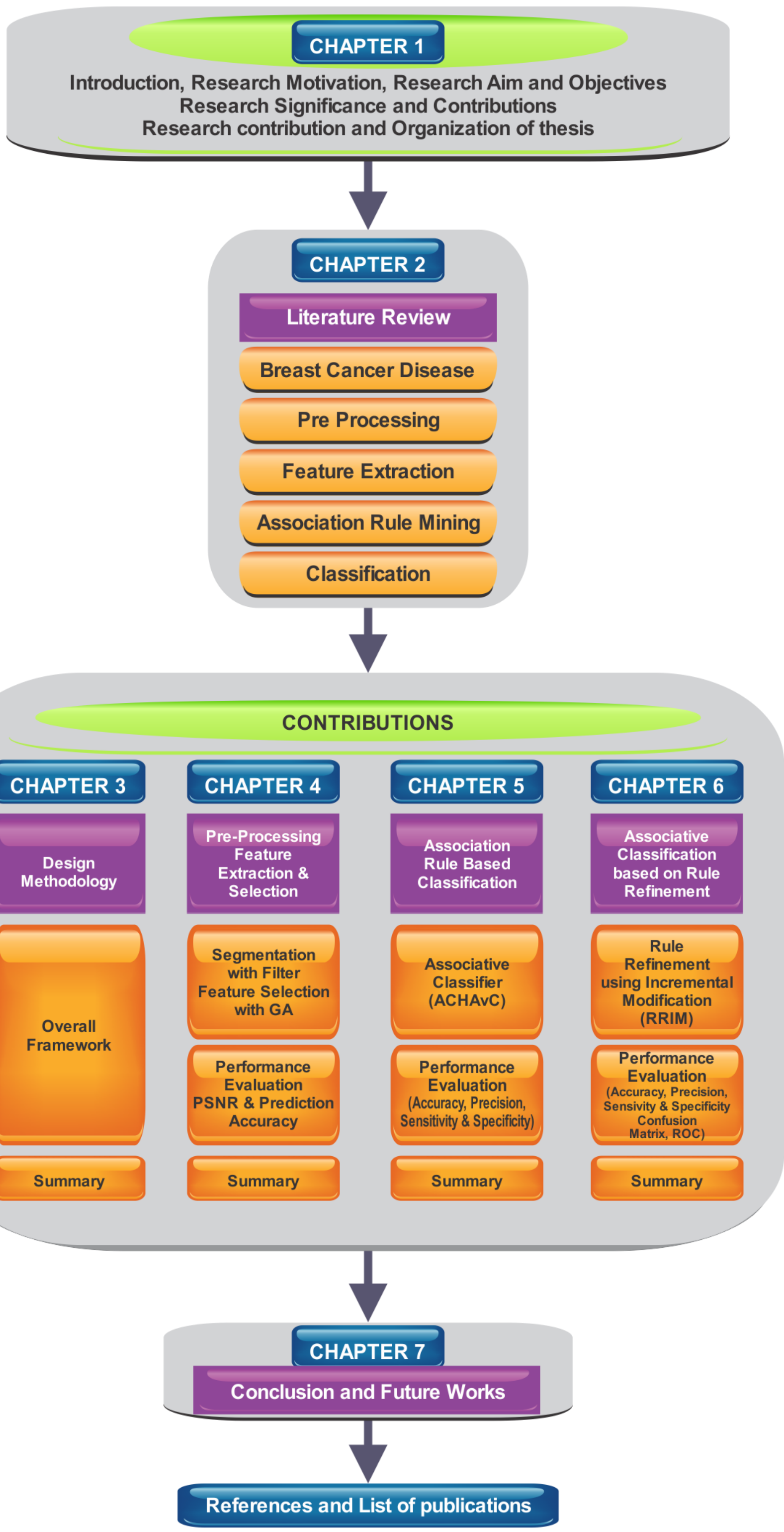

Figure 1.1: Organization of the thesis

1.9 Summary

This work is an integration of Computer Science and Medical Science, and highly contributes to the specified problem domains. This chapter gives a gist about the thesis, challenges faced, motivation for this work and major contributions. In the coming chapters more detailed technical explanation and experimental analysis are given.

REFERENCES

Abdelhamid, N., Ayesh, A., Thabtah, F., Ahmadi, S., & Hadi, W. (2012). MAC: A multiclass associative classification algorithm. *Journal of Information & Knowledge Management*, *11*(02), 1250011.

Abu-Amara, F., & Abdel-Qader, I. (2009). Hybrid mammogram classification using rough set and fuzzy classifier. *Journal of Biomedical Imaging*, *2009*, 17.

Abubacker, N. F., Azman, A., Doraisamy, S., Murad, M. A. A., Elmanna, M. E. M., & Saravanan, R. (2014). Correlation-Based Feature Selection for Association Rule Mining in Semantic Annotation of Mammographic Medical Images. In *Information Retrieval Technology* (pp. 482-493). Springer International Publishing.

Adnan, M., Alhajj, R., & Barker, K. (2005, September). Performance analysis of incremental update of association rules mining approaches. In Intelligent Engineering Systems, 2005. INES'05. Proceedings. 2005 IEEE International Conference on (pp. 129-134). IEEE.

Agrawal, R., & Srikant, R. (1994, September). Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB* (Vol. 1215, pp. 487-499).

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *ACM SIGMOD Record*, *22*(2), 207-216.

Akobeng, A. K. (2007), Understanding diagnostic tests 1: sensitivity, specificity and predictive values. Acta Paediatrica, 96: 338–341. doi: 10.1111/j.1651-2227.2006.00180.x

American Cancer Society. Cancer Facts & Figures 2013. Atlanta: American Cancer Society;

American College of Radiology. BI-RADS Committee. (1998). Breast imaging reporting and data system. American College of Radiology (Ed.). American College of Radiology.

Amornchewin, R., & Kreesuradej, W. (2007, December). Incremental association rule mining using promising frequent itemset algorithm. In*Information, Communications & Signal Processing, 2007 6th International Conference on* (pp. 1-5). IEEE.

Amornchewin, R., & Kreesuradej, W. (2008, October). Probability-based incremental association rule discovery algorithm. In *Computer Science and its Applications, 2008. CSA'08. International Symposium on* (pp. 212-215). IEEE.

Anderson, B., & Jakesz, R. (2008). Breast cancer issues in developing countries: An overview of the breast health global initiative. World J Surg., 32(12), 2578-2585.

Antonie, M. L., Zaiane, O. R., & Coman, A. (2003). Associative classifiers for medical images. In *Mining Multimedia and Complex Data* (pp. 68-83). Springer Berlin Heidelberg.

Aumann, Y., Feldman, R., Lipshtat, O., & Manilla, H. (1999). Borders: An efficient algorithm for association generation in dynamic databases. *Journal of Intelligent Information Systems*, *12*(1), 61-73.

Ayan, N. F., Tansel, A. U., & Arkun, E. (1999, August). An efficient algorithm to update large itemsets with early pruning. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 287-291). ACM.

B. Verma, P. McLeod, and A. Klevansky,(2009) "A novel soft cluster neural network for the classification of suspicious areas in digital mammograms," *Pattern Recognition*, vol. 42, pp. 1845-1852, 2009.

Ball, J. E., & Bruce, L. M. (2007, August). Digital mammographic computer aided diagnosis (CAD) using adaptive level set segmentation. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE* (pp. 4973-4978). IEEE.

Bankman, I. N., Nizialek, T., Simon, I., Gatewood, O. B., Weinberg, I. N., & Brody, W. R. (1997). Segmentation algorithms for detecting microcalcifications in mammograms. *Information Technology in Biomedicine, IEEE Transactions on*, *1*(2), 141-149.

Bassant, M., El-Horbaty, E. S. M., & Salem, A. B. M. (2013). A New Intelligent Telemedicine System based on Service-Oriented for Cancer Image Classification.

Beam, C. A., Layde, P. M., & Sullivan, D. C. (1996). Variability in the interpretation of screening mammograms by US radiologists: findings from a national sample. *Archives of internal medicine*, *156*(2), 209-213.

Beutel, J., Kundel, H. L., & Van Metter, R. L. (2000). Handbook of Medical Imaging, volume 1: Physics and Psychophysics.

Bhatnagar, S. (2012, December). Algorithm for finding association rules in distributed databases. In *Parallel Distributed and Grid Computing (PDGC), 2012 2nd IEEE International Conference on* (pp. 915-920). IEEE.

Bindu, C. H., & Prasad, K. S. (2012). An efficient medical image segmentation using conventional OTSU method. *Int. J. Adv. Sci. Technol*,*38*, 67-74.

Bing L. W. H. Y. (1998, August). Integrating classification and association rule mining. In *Proceedings of the fourth international conference on knowledge discovery and data mining*.

Bird, R. E., Wallace, T. W., & Yankaskas, B. C. (1992). Analysis of cancers missed at screening mammography. *Radiology*, *184*(3), 613-617.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford university press.

Biswas, S. K., & Mukherjee, D. P. (2011). Recognizing architectural distortion in mammogram: a multiscale texture modeling approach with GMM. *Biomedical Engineering, IEEE Transactions on*, *58*(7), 2023-2030.

Bovik, A. C., Markey, M. K., & Sampat, M. (2012). *U.S. Patent No. 8,164,039*. Washington, DC: U.S. Patent and Trademark Office.

Bovis, K., & Singh, S. (2000). Detection of masses in mammograms using texture features. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on* (Vol. 2, pp. 267-270). IEEE.

Bovis, K., & Singh, S. (2002, July). Classification of mammographic breast density using a combined classifier paradigm. In *Medical Image Understanding and Analysis (MIUA) Conference, Portsmouth*.

Bovis, K., Singh, S., Fieldsend, J., & Pinder, C. (2000). Identification of masses in digital mammograms with MLP and RBF nets. In *Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on* (Vol. 1, pp. 342-347). IEEE.

Braz Junior, A. Cardoso de Paiva, A. Corrêa Silva, and A. Cesar Muniz de Oliveira, (2009) "Classification of breast tissues using Moran's index and Geary's coefficient as texture signatures and SVM," *Computers in Biology and Medicine*, vol. 39, pp. 1063-1072.

Brian D. Ripley. (1996). *Pattern recognition and neural networks*. Cambridge university press.

Brzakovic, D., Luo, X. M., & Brzakovic, P. (1990). An approach to automated detection of tumors in mammograms. *Medical Imaging, IEEE Transactions on*,*9*(3), 233-241.

Buciu, I., & Gacsadi, A. (2011). Directional features for automatic tumor classification of mammogram images. *Biomedical Signal Processing and Control*, *6*(4), 370-378.

Burrell, H. C., Sibbering, D. M., Wilson, A. R., Pinder, S. E., Evans, A. J., Yeoman, L. J., ... & Robertson, J. F. (1996). Screening interval breast cancers: mammographic features and prognosis factors. Radiology, 199(3), 811-817.

Buttar, H., & Kaur, R. (2013). Association Technique in Data Mining and Its Applications. International Journal of Computer Trends and Technology (IJCTT)-volume4Issue4–April 2013.

Byng, J. W., Critten, J. P., & Yaffe, M. J. (1997). Thickness-equalization processing for mammographic images. *Radiology*, *203*(2), 564-568.

Canadian Task Force on Preventive Health Care. (2011). Recommendations on screening for breast cancer in average-risk women aged 40–74 years.*Canadian Medical Association Journal*, *183*(17), 1991-2001.

Cascio, D. O. N. A. T. O., Fauci, F., Magro, R., Raso, G., Bellotti, R., De Carlo, F., ... & Lauria, A. (2006). Mammogram segmentation by contour searching and mass lesions classification with neural network. *Nuclear Science, IEEE Transactions on*, *53*(5), 2827-2833.

Ceci, M., Appice, A., & Malerba, D. (2004). Spatial associative classification at different levels of granularity: a probabilistic approach. In *Knowledge Discovery in Databases: PKDD 2004* (pp. 99-111). Springer Berlin Heidelberg..

Chakaravarthy, V. T., Pandit, V., & Sabharwal, Y. (2009, March). Analysis of sampling techniques for association rule mining. In *Proceedings of the 12th international conference on database theory* (pp. 276-283). ACM.

Chan, H. P., Wei, D., Helvie, M. A., Sahiner, B., Adler, D. D., Goodsitt, M. M., & Petrick, N. (1995). Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space. *Physics in medicine and biology*, *40*(5), 857.

Chang, C. H., & Yang, S. H. (2003). Enhancing SWF for incremental association mining by itemset maintenance. In *Advances in Knowledge Discovery and Data Mining* (pp. 301-312). Springer Berlin Heidelberg.

Chen, J., Yin, J., & Huang, J. (2005). Mining correlated rules for associative classification. In *Advanced Data Mining and Applications* (pp. 130-140). Springer Berlin Heidelberg.

Chen, J., Yin, J., Zhang, J., & Huang, J. (2005). Associative classification in text categorization. In *Advances in Intelligent Computing* (pp. 1035-1044). Springer Berlin Heidelberg.

Chen, M. S., Han, J., & Yu, P. S. (1996). Data mining: an overview from a database perspective. *Knowledge and data Engineering, IEEE Transactions on*, *8*(6), 866-883.

Cheng, H. D., Shi, X. J., Min, R., Hu, L. M., Cai, X. P., & Du, H. N. (2006). Approaches for automated detection and classification of masses in mammograms. *Pattern recognition*, *39*(4), 646-668.

Cheung, D. W. L., Lee, S. D., & Kao, B. (1997, April). A General Incremental Technique for Maintaining Discovered Association Rules. In DASFAA (Vol. 6, pp. 185-194).

Chien, Y. W. C., & Chen, Y. L. (2010). Mining associative classification rules with stock trading data–A GA-based method. *Knowledge-Based Systems*,*23*(6), 605-614.

Christoyianni, I., Dermatas, E., & Kokkinakis, G. (2000). Fast detection of masses in computer-aided mammography. *Signal Processing Magazine, IEEE*, *17*(1), 54-64.

Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space. *Physics In Medicine And Biology*, *40*(5), 857-876.

D. Wang, L. Shi, and P. Ann Heng, (2009)"Automatic detection of breast cancers in mammograms using structured support vector machines," *Neurocomputing*, vol. 72, pp. 3296-3302, 2009

D'Orsi, C. J., & Newell, M. S. (2007, October). Digital mammography: clinical implementation and clinical trials. In *Seminars in roentgenology* (Vol. 42, No. 4, pp. 236-242). WB Saunders.

Dae-Won, K. I. M. (2012). Classification based on predictive association rules of incomplete data. *IEICE TRANSACTIONS on Information and Systems*, *95*(5), 1531-1535.

DaPonte, J. S., & Sherman, P. (1991). Classification of ultrasonic image texture by statistical discriminant analysis and neural networks.*Computerized medical imaging and graphics*, *15*(1), 3-9.

Das, A., Ng, W. K., & Woon, Y. K. (2001, October). Rapid association rule mining. In *Proceedings of the tenth international conference on Information and knowledge management* (pp. 474-481). ACM.

David W Cheung, Jiwei Han, Vincent T Ng and C. Y. Wong, Maintenance of Discovered Association Rules in Large Databases: An Incremental Updating Technique. *In proceedings of international conference on data engineering, pages 106-114.*

Davies, D. H., & Dance, D. R. (1990). Automatic computer detection of clustered calcifications in digital mammograms. *Physics in medicine and Biology*, *35*(8), 1111.

Davis, L. (1990, June). Hybrid genetic algorithms for machine learning. In *Machine Learning, IEE Colloquium on* (pp. 9-1). IET.

De Oliveira Martins, L., Junior, G. B., Silva, A. C., de Paiva, A. C., & Gattass, M. (2009). Detection of masses in digital mammograms using K-means and support vector machine. *ELCVIA: electronic letters on computer vision and image analysis*, *8*(2), 39-50.

Deb, K., & Pratap, A. (2002). A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II.

Devi, R. D. H., & Devi, M. I. (2015). Outlier detection algorithm combined with decision tree classifier for early diagnosis of breast cancer. *Int J Adv Engg Tech/Vol. VII/Issue II/April-June*, *93*, 98.

Dhanabhakyam, M., & Punithavalli, M. (2011). A survey on data mining algorithm for market basket analysis. *Global Journal of Computer Science and Technology*, *11*(11).

Dhawan, A. P. (1996). Analysis of mammographic microcalcifications using gray-level image structure features. *Medical Imaging, IEEE Transactions on*,*15*(3), 246-259.

Dheeba, J., & Tamil Selvi, S. (2010, October). Screening mammogram images for abnormalities using radial basis function neural network. In*Communication Control and Computing Technologies (ICCCCT), 2010 IEEE International Conference on* (pp. 554-559). IEEE.

Dias, A. H., & De Vasconcelos, J. (2002). Multiobjective genetic algorithms applied to solve optimization problems. *Magnetics, IEEE Transactions on*, *38*(2), 1133-1136.

Docusse, T. A., Pereira, A. S., & Marranghello, N. (2009). Microcalcification border characterization. *IEEE Engineering in Medicine and Biology Magazine*,*5*(28), 41-43.

Dogra, A., & BHALLA, P. (2015). Image sharpening by gaussian and butterworth high pass filter. *International Journal of Advanced Biological and Biomedical Research*, *3*(1), 93-96.

Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*, *31*(4), 198-211.

Domingos, P., & Pazzani, M. (1996). Beyond independence: conditions for the optimality of the simple Bayesian classifier. Proceedings of ICML ’96

Dominguez, A. R., & Nandi, A. K. (2007, April). Enhanced multi-level thresholding segmentation and rank based region selection for detection of masses in mammograms. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on* (Vol. 1, pp. I-449). IEEE.

Dom ńguez, A. R., & Nandi, A. K. (2009). Toward breast cancer diagnosis based on automated segmentation of masses in mammograms. *Pattern Recognition*, *42*(6), 1138-1148.

Dougherty, J., Kohavi, R., & Sahami, M. (1995, July) Supervised and unsupervised discretization of continuous features. In Machine learning: proceedings of the twelfth international conference (Vol. 12, pp. 194-202).

Drukker, K., Giger, M. L., Vyborny, C. J., & Mendelson, E. B. (2004). Computerized detection and classification of cancer on breast ultrasound 1. Academic radiology, 11(5), 526-535.

Eddaoudi, F., Regragui, F., Mahmoudi, A., & Lamouri, N. (2011). Masses detection using SVM classifier based on textures analysis. *Applied Mathematical Sciences*, *5*(8), 367-379.

El-Faramawy, N. M., Rangayyan, R. M., Desautels, J. E. L., & Alim, O. A. (1996, May). Shape factors for analysis of breast tumors in mammograms. In *Electrical and Computer Engineering, 1996. Canadian Conference on* (Vol. 1, pp. 355-358). IEEE.

Elfarra, B. K., & Abuhaiba, I. S. (2013). New feature extraction method for mammogram computer aided diagnosis. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, *6*(1), 13-36.

Elmore, J. G., Wells, C. K., Lee, C. H., Howard, D. H., & Feinstein, A. R. (1994). Variability in radiologists' interpretations of mammograms. *New England Journal of Medicine*, *331*(22), 1493-1499.

Evans, J. A. (1994). Screening mammography breast cancer diagnosis in asymptomatic women Scane, Potchen, Sierra, Azavedo. 1993 Mosby. £ 110. pp 546. ISBN 0 8016 6488 8.

F. Moayedi, Z. Azimifar, R. Boostani, and S. Katebi, (2015) "Contourlet-based mammography mass classification using the SVM family," *Computers in Biology and Medicine*, vol. 40, pp. 373-383, 2010.

Fang, L., & Qizhi, Q. (2012, May). The Study on the Application of Data Mining Based on Association Rules. In *Communication Systems and Network Technologies (CSNT), 2012 International Conference on* (pp. 477-480). IEEE.

Fekjær W.Harald (2007). Modelling breast cancer incidence, progression and screening test sensitivity using screening data. PhD thesis. Department of biostatistics. The University of Oslo, Series of Dissertations, No.570. (p-9).

Ferreira, A., Nascimento Jr, F., Tsang, I. R., Cavalcanti, G. D., Ludermir, T. B., & De Aquino, R. R. (2007, August).Analysis of mammogram using self-organizing neural networks based on spatial isomorphism. In *Neural Networks, 2007. IJCNN 2007. International Joint Conference on* (pp. 1796-1801). IEEE.

Forouzanfar, M., Foreman, K., Delossantos, A., Lozano, R., Lopez, A., Murray, C., & Naghavi, M. (2011). Breast and cervical cancer in 187 countries between

1980 and 2010: a systematic analysis. Lancet, 378(9801), 1461–1484. doi:10.1016/S0140-6736(11)61351-2.

Gao, X., Wang, Y., Li, X., & Tao, D. (2010). On combining morphological component analysis and concentric morphology model for mammographic mass detection. *Information Technology in Biomedicine, IEEE Transactions on*, *14*(2), 266-273.

Geetha, K., & Thanushkodi, K. (2008, January). New particle swarm optimization for feature selection and classification of microcalcifications in mammograms. In *Signal Processing, Communications and Networking, 2008. ICSCN'08. International Conference on* (pp. 458-463). IEEE.

Geronimo, J. S., Hardin, D. P., & Massopust, P. R. (1994). Fractal functions and wavelet expansions based on several scaling functions. *Journal of approximation theory*, *78*(3), 373-401.

Giger, M. L. (1993). Computer-aided diagnosis. *RSNA categorical course in physics*, 283-298.

Giger, M. L., Karssemeijer, N., & Armato, S. G. (2001). Guest editorial computer-aided diagnosis in medical imaging. *Medical Imaging, IEEE Transactions on*, *20*(12), 1205-1208.

Goldberg, D. E. Genetic algorithms in search, optimization, and machine learning. Reading, MA: Addison-Wesley; 1989. *Objective Function Value [Eq.(18)]*

Gopal, G., & Kanaga, E. G. M. (2013). A Study on Enhancement Techniques for Mammogram Images. International Journal of Advanced Research in Electronics and Communication Engineering, 2(1), pp-036.

Grim, J., Somol, P., Haindl, M., & Daneš, J. (2009). Computer-aided evaluation of screening mammograms based on local texture models. *Image Processing, IEEE Transactions on*, *18*(4), 765-773.

Halapy, E. E., Chiarelli, A. M., Klar, N., & Knight, J. (2004). Breast screening outcomes in women with and without a family history of breast and/or ovarian cancer. *Journal of medical screening*, *11*(1), 32.

Hall, M. A. (1999). Correlation-based feature selection for machine learning(Doctoral dissertation, The University of Waikato).

Han, J., & Kamber, M. (2001). Data mining: concepts and techniques.

Han, J., Cheng, H., Xin, D., & Yan, X. (2007). Frequent pattern mining: current status and future directions. *Data Mining and Knowledge Discovery*, *15*(1), 55-86.

Han, J., Pei, J., & Yin, Y. (2000, May). Mining frequent patterns without candidate generation. In *ACM SIGMOD Record* (Vol. 29, No. 2, pp. 1-12). ACM.

Han, W., Dong, J., Guo, Y., Zhang, M., & Wang, J. (2011, November). Identification of masses in digital mammogram using an optimal set of features. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2011 IEEE 10th International Conference on* (pp. 1763-1768). IEEE.

Hannu Toivinen, Sampling Large Databases for Association Rules. *In proceedings of the 22nd VLDB conference, pages 134-145.*

Haralick RM, Shanmugam K, Dinstein I (1973) Texture features for image classification. IEEE Trans Syst Man Cybern 8(6):610–621

Harris JR, Lippman ME, Morrow M & Osborne C.K (2004). Disease of the Breast. 3rd edition, USA, Lippincott Williams & Wilkins (capital 1).

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. Applied statistics, 100-108.

Harvey, J. A., Fajardo, L. L., & Innis, C. A. (1993). Previous mammograms in patients with impalpable breast carcinoma: retrospective vs blinded interpretation. 1993 ARRS President's Award. *AJR. American journal of roentgenology*, *161*(6), 1167-1172.

Hassanien, A. E., & Ali, J. M. Feature extraction and rule classification algorithm of digital mammography based on rough set theory. In *Available at www. wseas. us/ e-library/ conferences/ digest2003/ papers* (pp. 463-104).

He, W., Kibiro, M., Juette, A., Hogg, P., Denton, E. R., & Zwiggelaar, R. (2014). A Novel Breast Image Preprocessing For Full Field Digital Mammographic Segmentation and Risk Classification. In *MIUA* (pp. 79-84).

Heath, M., Bowyer, K., Kopans, D., Moore, R., & Kegelmeyer, W. P. (2000, June). The digital database for screening mammography. In Proceedings of the 5th international workshop on digital mammography (pp. 212-218).

Hidber, C. (1999). *Online association rule mining* (Vol. 28, No. 2, pp. 145-156). ACM.

Holmes G, Donkin A, Witten IH (1994) WEKA: a machine learning workbench. In: Proceedings second Australia and New Zealand conference on intelligent information systems, Brisbane, pp 357–361

Holte, R. C. (1993). Very simple classification rules perform well on most commonly used datasets. *Machine learning*, *11*(1), 63-90.

Hong TP,Wang CY, Tao YH (2001) A new incremental data mining algorithm using pre-large itemsets. Int Data Anal 5(2):111–129

Hu, K., Lu, Y., Zhou, L., & Shi, C. (1999). Integrating classification and association rule mining: A concept lattice framework. In *New Directions in Rough Sets,* *Data Mining, and Granular-Soft Computing* (pp. 443-447). Springer Berlin Heidelberg.

Hu, M. K. (1962). Visual pattern recognition by moment invariants.information Theory, IRE Transactions on, 8(2), 179-187.

Huo, Z., Giger, M. L., & Vyborny, C. J. (2001). Computerized analysis of multiple-mammographic views: Potential usefulness of special view mammograms in computer-aided diagnosis. *Medical Imaging, IEEE Transactions on*, *20*(12), 1285-1292.

Huo, Z., Giger, M. L., Vyborny, C. J., Wolverton, D. E., Schmidt, R. A., & Doi, K. (1998). Automated computerized classification of malignant and benign masses on digitized mammograms. *Academic Radiology*, *5*(3), 155-168.

J. Dheeba and S. T. Selvi, (2012) "A Swarm Optimized Neural Network System for Classification of Microcalcification in Mammograms," *Journal of medical systems*, vol. 36, pp. 3051-3061, 2012.

J. Ren, (2012) "ANN vs. SVM: Which one performs better in classification of MCCs in mammogram imaging," *Knowledge-Based Systems*, vol. 26, pp. 144-153, 2012.

Jackson, V. P., Hendrick, R. E., Feig, S. A., & Kopans, D. B. (1993). Imaging of the radiographically dense breast. *Radiology*, *188*(2), 297-301.

Jacquelyn. P., Palmer, C. H., Muller, C. C., Little, S. T., Smith, D. C., Fatouros, P. P., & de Paredes, E. S. (1999). Quality assurance in mammography: artifact analysis. *Radiographics*, *19*(2), 503-522.

Jain, S., & Kabra, S. (2012). Mining & Optimization of Association Rules Using Effective Algorithm. *International Journal of Emerging Technology and Advanced Engineering*, *2*(4).

Jalalian, A., Mashohor, S. B., Mahmud, H. R., Saripan, M. I. B., Ramli, A. R. B., & Karasfi, B. (2013). Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. *Clinical imaging*, *37*(3), 420-426.

Jaleel, J. A., Salim, S., & Archana, S. (2014, July). Textural features based computer aided diagnostic system for mammogram mass classification. In*Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014 International Conference on* (pp. 806-811). IEEE.

Jensen, D., and Cohen, P. (2000) Multiple comparisons in induction algorithms. Machine Learning 38(3), (pp. 309 – 338), 2000

Jian Chen, Jian Yin, Jun Zhang, and Jin Huang. Associative classi_cation in text categorization. In ICIC (1), pages 1035{1044, 2005.

Jiang, Y., Nishikawa, R. M., Wolverton, D. E., Metz, C. E., Giger, M. L., Schmidt, R. A., ... & Doi, K. (1996). Malignant and benign clustered microcalcifications: automated feature analysis and classification.*Radiology*, *198*(3), 671-678.

John Henry Holland. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press.

Jones, T. L., & Stevens, K. K. (2014). Dense Breast Notification: Anatomy, Imaging, and Patient Awareness. *Radiologic technology*, *86*(1), 17-22.

Junior, G. B., de Paiva, A. C., Silva, A. C., & de Oliveira, A. C. M. (2009). Classification of breast tissues using Moran's index and Geary's coefficient as texture signatures and SVM. *Computers in Biology and Medicine*, *39*(12), 1063-1072.

Kadambe, S., & Boudreaux-Bartels, G. F. (1992). Application of the wavelet transform for pitch detection of speech signals. IEEE Transactions on Information Theory, 38(2), 917-924.

Kallenberg, M. G., van Gils, C. H., Lokate, M., den Heeten, G. J., & Karssemeijer, N. (2012). Effect of compression paddle tilt correction on volumetric breast density estimation. *Physics in medicine and biology*,*57*(16), 5155.

Kallenberg, M., & Karssemeijer, N. (2010). Comparison of tilt correction methods in full field digital mammograms. In Digital Mammography (pp. 191-196). Springer Berlin Heidelberg.

Kallergi, M., Woods, K., Clarke, L. P., Qian, W., & Clark, R. A. (1992). Image segmentation in digital mammography: comparison of local thresholding and region growing algorithms. *Computerized medical imaging and graphics*, *16*(5), 323-331.

Karabatak, M. (2015). A new classifier for breast cancer detection based on Naïve Bayesian. *Measurement*, *72*, 32-36.

Karssemeijer, N., & Snoeren, P. R. (2010). Image Processing. In *Digital ammography* (pp. 69-83). Springer Berlin Heidelberg.

Kerlikowske, K., Carney, P. A., Geller, B., Mandelson, M. T., Taplin, S. H., Malvin, K., ... & Ballard-Barbash, R. (2000). Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Annals of Internal Medicine*, *133*(11), 855-863.

Kim, J. K., & Park, H. W. (1999). Statistical textural features for detection of microcalcifications in digitized mammograms. *Medical Imaging, IEEE Transactions on*, *18*(3), 231-238.

Kwon, S.K., Tamhankar, A., & Rao, K.R. (2006). Overview of H.264/MPEG-4 part 10. *Journal of visual communication and image representation*, 17, 186-216.

Le TP, Hong TP, Vo B, Le B (2011) Incremental mining frequent itemsets based on the trie structure and the prelarge itemsets.In: Proceedings of the 2011 IEEE international conference on granular computing, Kaohsiung, pp 369-373

Le TP, Hong TP, Vo B, Le B (2012) An efficient incremental mining approach based on IT-tree. In: Proceedings of the 2012 IEEE *Lecture Notes in Computer Science*, 4782, 577-585.

Lee, C. H., Lin, C. R., & Chen, M. S. (2001, October). Sliding-window filtering: an efficient algorithm for incremental mining. In Proceedings of the tenth international conference on Information and knowledge management (pp. 263-270). ACM.

Lehman, C. D., Blume, J. D., DeMartini, W. B., Hylton, N. M., Herman, B., & Schnall, M. D. (2013). Accuracy and interpretation time of computer-aided detection among novice and experienced breast MRI readers. *AJR. American journal of roentgenology*, *200*(6), W683.

Li, H. D., Kallergi, M., Clarke, L. P., Jain, V. K., & Clark, R. A. (1995). Markov random field for tumor detection in digital mammography. *Medical Imaging, IEEE Transactions on*, *14*(3), 565-576.

Li, H., Wang, Y., Liu, K. J., Lo, S. C. B., & Freedman, M. T. (2001). Computerized radiographic mass detection. I. Lesion site selection by morphological enhancement and contextual segmentation. *Medical Imaging, IEEE Transactions on*, *20*(4), 289-301.

Li, L., Qian, W., Clarke, L. P., Clark, R. A., & Thomas, J. A. (1999, May). Improving mass detection by adaptive and multiscale processing in digitized mammograms. In *Medical Imaging'99* (pp. 490-498). International Society for Optics and Photonics.

Li, W., Han, J., & Pei, J. (2001). CMAR: Accurate and efficient classification based on multiple class-association rules. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on* (pp. 369-376). IEEE.

Li, W., Han, J., & Pei, J. (2001). CMAR: Accurate and efficient classification based on multiple class-association rules. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on* (pp. 369-376). IEEE.

Li, X., Qin, D., & Yu, C. (2008, October). ACCF: Associative classification based on closed frequent itemsets. In *Fuzzy Systems and Knowledge Discovery, 2008. FSKD'08. Fifth International Conference on* (Vol. 2, pp. 380-384). IEEE.

Lin CW, Hong TP (2009) The Pre-FUFP algorithm for incremental mining. Expert Syst Appl 36(5):9498–9505.

Lin, W., Alvarez, S. A., & Ruiz, C. (2002). Efficient adaptive-support association rule mining for recommender systems. *Data mining and knowledge discovery*, *6*(1), 83-105.

Liu, H., & Motoda, H. (2012). *Feature selection for knowledge discovery and data mining* (Vol. 454). Springer Science & Business Media.

Lo, J. Y., & Floyd Jr, C. E. (1999). Application of artificial neural networks for diagnosis of breast cancer. In *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on* (Vol. 3). IEEE.

Lo, J. Y., Gavrielides, M. A., Markey, M. K., & Jesneck, J. L. (2003, May). Computer-aided classification of breast microcalcification clusters: Merging of features from image processing and radiologists. In *Medical Imaging 2003*(pp. 882-889). International Society for Optics and Photonics.

M. Fraschini, "Mammographic masses classification: novel and simple signal analysis method," *Electronics letters*, vol. 47, pp. 14-15, 2011.

M. Meselhy Eltoukhy, I. Faye, and B. Belhaouari Samir, "A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram,(2010) " *Computers in Biology and Medicine*, vol. 40, pp. 384-391, 2010.

Ma, B. L. W. H. Y. (1998, August). Integrating classification and association rule mining. In *Proceedings of the fourth international conference on knowledge discovery and data mining*.

Machado, F. P. (2003). CPAR: Classification based on predictive association rules.

Magro, r., raso, g., & cascio, d. (2004). Mammogram segmentation by contour searching and massive lesions classification with neural network.

Maitra, I. K., Nag, S., & Bandyopadhyay, S. K. (2011). Identification of abnormal masses in digital mammography images. *International Journal of Computer Graphics*, *2*(1), 17-30.

Marpe, D., Wiegand, T., & Sullivan, G. J. (2006). The H. 264/MPEG4 advanced video coding standard and its applications. *Communications Magazine, IEEE*, *44*(8), 134-143.

Mathew, l. S. (2013). Integrated associative classification and neural network model enhanced by using a statistical approach. International journal of data mining & knowledge management process, 3(4), 107.

Matsubara, T., Fujita, H., Kasai, S., Goto, M., Tani, Y., Hara, T., & Endo, T. (1997, December). Development of new schemes for detection and analysis of mammographic masses. In *Intelligent Information Systems, 1997. IIS'97. Proceedings* (pp. 63-66). IEEE.

Mavroforakis, M. E., Georgiou, H. V., Dimitropoulos, N., Cavouras, D., & Theodoridis, S. (2006). Mammographic masses characterization based on localized texture and dataset fractal analysis using linear, neural and support vector machine classifiers. *Artificial Intelligence in Medicine*, *37*(2), 145-162.

McInerney, M., & Dhawan, A. P. (1993). Use of genetic algorithms with backpropagation in training of feedforward neural networks. In *Neural Networks, 1993., IEEE International Conference on* (pp. 203-208). IEEE.

Meenalosini, S., Janet, J., & Kannan, E. (2012). A novel approach in malignancy detection of computer aided diagnosis. *American Journal of Applied*

Metz, C. E. (2008). ROC analysis in medical imaging: a tutorial review of the literature. *Radiological physics and technology*, *1*(1), 2-12.

Moayedi, F., Azimifar, Z., Boostani, R., & Katebi, S. (2010). Contourlet-based mammography mass classification using the SVM family. *Computers in biology and medicine*, *40*(4), 373-383.

Moftah, H. M., Azar, A. T., Al-Shammari, E. T., Ghali, N. I., Hassanien, A. E., & Shoman, M. (2014). Adaptive k-means clustering algorithm for MR breast image segmentation. *Neural Computing and Applications*, *24*(7-8), 1917-1928.

Mohan, S., & Ravishankar, M. (2013). Optimized histogram based contrast limited enhancement for mammogram images. Short Pap. ACEEE Int. J. Inf. Technol, 3(1), 66-71.

Mohanty, A. K., Senapati, M. R., & Lenka, S. K. (2013). A novel image mining technique for classification of mammograms using hybrid feature selection. *Neural Computing and Applications*, *22*(6), 1151-1161.

Mohanty, A. K., Senapati, M. R., Beberta, S., & Lenka, S. K. (2013). Texture-based features for classification of mammograms using decision tree. *Neural Computing and Applications*, *23*(3-4), 1011-1017.

Mohanty, A. K., Swain, S. K., Champati, P. K., & Lenka, S. K. (2011). Image mining for mammogram classification by association rule using statistical and GLCM features.

Morisaki, S., Monden, A., Matsumura, T., Tamada, H., & Matsumoto, K. I. (2007, May). Defect data analysis based on extended association rule mining. In *Proceedings of the Fourth International Workshop on Mining Software Repositories* (p. 3). IEEE Computer Society.

Mustra, M., & Grgic, M. (2013). Robust automatic breast and pectoral muscle segmentation from scanned mammograms. *Signal processing*,*93*(10), 2817-2827.

N. Arikidis, A. Karahaliou, S. Skiadopoulos, P. Korfiatis, E. Likaki, G. Panayiotakis, *et al.*, (2008) "Size-adapted segmentation of individual mammographic microcalcifications," in *BioInformatics and BioEngineering, 2008. BIBE 2008. 8th IEEE International Conference on*, 2008, pp. 1-5.

Nanavati, A. A., Chitrapura, K. P., Joshi, S., & Krishnapuram, R. (2001, October). Mining generalised disjunctive association rules. In *Proceedings of the tenth international conference on Information and knowledge management* (pp. 482-489). ACM.

National Cancer Institute http://www.cancer.gov/cancertopics/factsheet/detection/mammograms

Nelson, H. D., Tyne, K., Naik, A., Bougatsos, C., Chan, B. K., & Humphrey, L. (2009). Screening for breast cancer: an update for the US Preventive Services Task Force. *Annals of internal medicine*, *151*(10), 727-737.

Ni, X. 2008. Research of data mining based on neural networks. World Academy Of Science, Engineering And Technology.

Nishikawa, R. M., Jiang, Y., Giger, M. L., Vyborny, C. J., & Schmidt, R. (1992, October). Computer-aided detection of clustered microcalcifications. In *Systems, Man and Cybernetics, 1992., IEEE International Conference on*(pp. 1375-1378). IEEE.

Nithya, R., & Santhi, B. (2012). Mammogram analysis based on pixel intensity mean features. *Journal of Computer Science*, *8*(3), 329.

Olmezogullari, E., & Ari, I. (2013, June). Online association rule mining over fast data. In *Big Data (BigData Congress), 2013 IEEE International Congress on* (pp. 110-117). IEEE. Optics and Photonics.

Pal, N. R., Bhowmick, B., Patel, S. K., Pal, S., & Das, J. (2008). A multi-stage neural network aided system for detection of microcalcifications in digitized mammograms. *Neurocomputing*, *71*(13), 2625-2634.

Pao, Y. (1989). Adaptive pattern recognition and neural networks.

Pappas, T. N. (1992). An adaptive clustering algorithm for image segmentation.*Signal Processing, IEEE Transactions on*, *40*(4), 901-914.

Parekh, R. (2010). Using texture analysis for medical diagnosis. *IEEE MultiMedia*, (2), 28-37.

Park, J., & Sandberg, I. W. (1991). Universal approximation using radial-basis-function networks. Neural computation, 3(2), 246-257.

Park, S., Kim, B., Lee, J., Goo, J. M., & Shin, Y. G. (2011). GGO nodule volume-preserving nonrigid lung registration using GLCM texture analysis.*Biomedical Engineering, IEEE Transactions on*, *58*(10), 2885-2894.

Patil, S., Udupi, V. R., & Bhogle, P. M. (2013). Implementation of Breast Tumor Demarcation in Digital Mammograms using Local Thresholding and Mathematical Morphology. *International Journal of Computer Applications*, *66*(24).

Petrick, N., Chan, H. P., Sahiner, B., & Helvie, M. A. (1999). Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms. *Medical physics*, *26*(8), 1642-1654.

Petrick, N., Chan, H. P., Sahiner, B., & Wei, D. (1996). An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection. *Medical Imaging, IEEE Transactions on*, *15*(1), 59-67.

Petrosian, A., Chan, H. P., Helvie, M. A., Goodsitt, M. M., & Adler, D. D. (1994). Computer-aided diagnosis in mammography: classification of mass and normal tissue by texture analysis. *Physics in Medicine and Biology*, *39*(12), 2273.

Philpotts, L.E.:(2009) Can Computer-aided Detection be Detrimental to mammographic Interpretation. Radiology, 17–22.

Platt, J. (1998). Sequential minimal optimization: A fast algorithm for training support vector machines.

Pol, U. (2013). Design and Development of Apriori Algorithm for Sequential to concurrent mining using MPI. *International Journal of Computers & Technology*, *10*(7), 1785-1790.

Pradeep, N., Girisha, H., Sreepathi, B., & Karibasappa, K. (2012). Feature extraction of mammograms. *International Journal of Bioinformatics Research, ISSN*, 0975-3087.

Pratiwi, M., Harefa, J., & Nanda, S. (2015). Mammograms Classification Using Gray-level Co-occurrence Matrix and Radial Basis Function Neural Network. *Procedia Computer Science*, *59*, 83-91.

Qodmanan, H. R., Nasiri, M., & Minaei-Bidgoli, B. (2011). Multi objective association rule mining with genetic algorithm without specifying minimum support and minimum confidence. *Expert Systems with applications*, *38*(1), 288-298.

Quinlan, J. (1993) C4.5: Programs for machine learning. San Mateo, CA: Morgan Kaufmann.

Radhakrishnan, M., Kuttiannan, T., & Tiruchengode, N. (2012). Comparative analysis of feature extraction methods for the classification of prostate cancer from TRUS medical images. *IJCSI Int J Comput Sci*, (9), 1.

Rai, A., & Kumar, D. (2014). A Comparative Study on Wrapper and Filter Methods for Feature Selection in Data Mining. *International Journal Of Scientific Research And Education*, *2*(08).

Raja Keerthana, K. T., & Thangavel, K. (2011, February). Feature selection in mammogram image using rough set approach. In *Innovations in Emerging Technology (NCOIET), 2011 National Conference on* (pp. 147-152). IEEE.

Rangayyan, R. M. (2004). *Biomedical image analysis*. CRC press.

Ribeiro, M. X., Traina, A. J., Balan, A. G., Traina, C., & Marques, P. (2007, June). SuGAR: A framework to support mammogram diagnosis. In *Computer-Based Medical Systems, 2007. CBMS'07. Twentieth IEEE International Symposium on* (pp. 47-52). IEEE.

Ribeiro, M., Traina, A., Traina, Jr., C., & Azevedo-Marques, P. (2008). An Association Rule-Based Method to Support Medical Image Diagnosis With Efficiency. *IEEE Trans. Multimedia*, *10*(2), 277-285.

Richardson, I.E.(2010). *The H.264 advance video compression standard (*2nd Ed.), Wiley.

Rico, D., Yang, J., Augustine, B., Mawdsley, G. E., & Yaffe, M. J. (2003). Peripheral thickness correction for volumetric breast density estimation. In *Digital Mammography* (pp. 187-191). Springer Berlin Heidelberg.

Rish, I. (2001, August). An empirical study of the naive Bayes classifier. In*IJCAI 2001 workshop on empirical methods in artificial intelligence* (Vol. 3, No. 22, pp. 41-46). IBM New York.

Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Systems with Applications*, *42*(3), 990-1002.

Sahiner, B., Chan, H. P., Petrick, N., Helvie, M. A., & Goodsitt, M. M. (1998). Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis. *Medical Physics*,*25*(4), 516-526.

Sahiner, B., Chan, H. P., Petrick, N., Helvie, M. A., & Hadjiiski, L. M. (2001). Improvement of mammographic mass characterization using spiculation measures and morphological features. *Medical Physics*, *28*(7), 1455-1465.

Sahiner, B., Chan, H. P., Petrick, N., Wei, D., Helvie, M., Adler, D. D., & Goodsitt, M. M. (1996). Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images.*Medical Imaging, IEEE Transactions on*, *15*(5), 598-610.

Sahiner, B., Chan, H. P., Wei, D., Petrick, N., Helvie, M. A., Adler, D. D., & Goodsitt, M. M. (1998). Image feature selection by a genetic algorithm: Application to classification of mass and normal breast tissue. *Medical Physics*, *23*(10), 1671-1684.

Sahiner, B., Hadjiiski, L. M., Chan, H. P., Paramagul, C., Nees, A., Helvie, M., & Shi, J. (2008, March). Concordance of computer-extracted image features with BI-RADS descriptors for mammographic mass margin. In*Medical Imaging* (pp. 69151N-69151N). International Society for Optics and Photonics.

Salzberg, S. L. (1994). C4. 5: Programs for machine learning by jross quinlan. morgan kaufmann publishers, inc., 1993. Machine Learning, 16(3), 235-240.

Sameti, M., Ward, R. K., Morgan-Parkes, J., & Palcic, B. (2009). Image feature extraction in the last screening mammograms prior to detection of breast cancer. *Selected Topics in Signal Processing, IEEE Journal of*, *3*(1), 46-52.

Sampat, M. P., Markey, M. K., & Bovik, A. C. (2005). Computer-aided detection and diagnosis in mammography. *Handbook of image and video processing*, *2*(1), 1195-1217.

Sankur, B., & Sezgin, M. (2001). Image thresholding techniques: A survey over categories. Pattern Recognition, 34(2), 1573-1607.

Sarawagi, S., Thomas, S., & Agrawal, R. (1998). *Integrating association rule mining with relational database systems: Alternatives and implications* (Vol. 27, No. 2, pp. 343-354). ACM.

Sasikala, D., & Premalatha, K. (2013, January). Mining association rules from XML document using modified index table. In *Computer Communication and Informatics (ICCCI), 2013 International Conference on* (pp. 1-6). IEEE.

Schafer, R., & Sikora, T. (1995). Digital Video Coding Standards and Their Role in Video Communications. *Proceedings of the IEEE*, 83(6), 907-927.

Serban, G. A. B. R. I. E. L. A., Czibula, I. G., & Campan, A. L. I. N. A. (2006). A Programming interface for medical diagnosis prediction. *Studia Universitatis" Babes-Bolyai", Informatica, LI (1), pag*, 21-30.

Sethi, P., & Alagiriswamy, S. (2010). Association rule based similarity measures for the clustering of gene expression data. *The open medical informatics journal*, *4*, 63.

Setiawan, A. S., Wesley, J., & Purnama, Y. (2015). Mammogram Classification using Law's Texture Energy Measure and Neural Networks.*Procedia Computer Science*, *59*, 92-97.

Setiono, R. (1996). Extracting rules from pruned neural networks for breast cancer diagnosis. *Artificial Intelligence in Medicine*, *8*(1), 37-51.

Sharma, V., & Singh, S. (2014). CFS–SMO based classification of breast density using multiple texture models. *Med Biol Eng Comput*, *52*(6), 521-529.

Sharma, V., & Singh, S. (2014). CFS–SMO based classification of breast density using multiple texture models. *Medical & biological engineering & computing*, *52*(6), 521-529.

Shekhawat P.B., Dhande S.S (2011) A classification technique using associative classification. *International journal of computer application* (0975-8887) vol. 20-No.5,20-28.

Shen, L., Rangayyan, R. M., & Desautels, J. L. (1993). Detection and classification of mammographic calcifications. *International Journal of Pattern Recognition and Artificial Intelligence*, *7*(06), 1403-1416.

Shilby Thomas, Sreenath Bodagala and Sanjay Ranka, An Efficient Algorithm for Imcremental Updation of Association Rules in Large Databases. *In proceedings of international conference on Knowledge Discovery and Data Mining, pages 263-266.*

Shinde, S. R., Sabale, S., Kulkarni, S., & Bhatia, D. (2015, January). Experiments on content based image classification using Color feature extraction. In *Communication, Information & Computing Technology (ICCICT), 2015 International Conference on* (pp. 1-6). IEEE.

Signs of disease, Mammographic Image Analysis Homepage, 2009, http://www.mammoimage.org/signs-of-disease/ [Accessed April 2013].

Singh, B. K. (2011). Mammographic image enhancement, classification and retrieval using color, statistical and spectral Analysis. *International Journal of Computer Applications*, *10*, 18-23.

Singh, V. P., Srivastava, A., Kulshreshtha, D., Chaudhary, A., & Srivastava, R. (2016). Mammogram Classification Using Selected GLCM Features and Random Forest Classifier. *International Journal of Computer Science and Information Security*, *14*(6), 8

Skaane, P., Bandos, A. I., Gullien, R., Eben, E. B., Ekseth, U., Haakenaasen, U., ... & Hofvind, S. (2013). Prospective trial comparing full-field digital mammography (FFDM) versus combined FFDM and tomosynthesis in a population-based screening programme using independent double reading with arbitration. *European radiology*, *23*(8), 2061-2071.

Snoeren, P. R., & Karssemeijer, N. (2004). Thickness correction of mammographic images by means of a global parameter model of the compressed breast. Medical Imaging, IEEE Transactions on, 23(7), 799-806.

Solbjør, M (2008) Women's experiences of mammography screening, Decision making, participation and recall. Norges teknisk-naturvitenskapelige universitetet.

Song, S. E., Seo, B. K., Cho, K. R., Woo, O. H., Son, G. S., Kim, C., ... & Kwon, S. S. (2015). Computer-aided detection (CAD) system for breast MRI in assessment of local tumor extent, nodal status, and multifocality of invasive breast cancers: preliminary study. *Cancer Imaging*, *15*(1), 1.

Srikant, R. and Agrawal, R. 1996. Mining quantitative association rules in large relational tables. *SIGMOD-96*.

Stefanoyiannis, A. P., Costaridou, L., Sakellaropoulos, P., & Panayiotakis, G. (2000). A digital density equalization technique to improve visualization of breast periphery in mammography. *The British journal of radiology*, *73*(868), 410-420.

Stefanoyiannis, A. P., Costaridou, L., Skiadopoulos, S., & Panayiotakis, G. (2003). A digital equalisation technique improving visualisation of dense mammary gland and breast periphery in mammography. *European journal of radiology*, *45*(2), 139-149.

Suganthi, M., & Madheswaran, M. (2012). An improved medical decision support system to identify the breast cancer using mammogram. *Journal of medical systems*, *36*(1), 79-91.

Sun, S., Wang, D., & Chen, S. (2007). A highly efficient parallel algorithm for H.264 encoder

Suri, J. S., & Rangayyan, R. M. (2006). Recent advances in breast imaging, mammography, and computer-aided diagnosis of breast cancer.

Székely, N., Tóth, N., & Pataki, B. (2006). A hybrid system for detecting masses in mammographic images. *Instrumentation and Measurement, IEEE Transactions on*, *55*(3), 944-952.

T. Subashini, V. Ramalingam, and S. Palanivel, (2010) "Automated assessment of breast tissue density in digital mammograms," *Computer Vision and Image Understanding*, vol. 114, pp. 33-43, 2010.

Tai, S. C., Chen, Z. S., & Tsai, W. T. (2014). An automatic mass detection system in mammograms based on complex texture features. *Biomedical and Health Informatics, IEEE Journal of*, *18*(2), 618-627.

Tang, Z., & Liao, Q. (2007). A New Class Based Associative Classification Algorithm. *IMECS, 2007*, 685-689.

Tarassenko, L., Hayton, P., Cerneaz, N., & Brady, M. (1995). Novelty detection for the identification of masses in mammograms.

Tassa, T. (2014). Secure mining of association rules in horizontally distributed databases. *Knowledge and Data Engineering, IEEE Transactions on*, *26*(4), 970-983.

Teng, W. G., & Chen, M. S. (2005). Incremental Mining on Association Rules. In Foundations and Advances in Data Mining (pp. 125-162). Springer Berlin Heidelberg.

Thabtah, F., Cowling, P., & Peng, Y. (2005). MCAR: multi-class classification based on association rule. In *The 3rd ACS/IEEE International Conference onComputer Systems and Applications, 2005.* (p. 33). IEEE.

Thomas, S., Bodagala, S., Alsabti, K., & Ranka, S. (1997, August). An Efficient Algorithm for the Incremental Updation of Association Rules in Large Databases. In *KDD* (pp. 263-266).

Toivonen, H. (1996, September). Sampling large databases for association rules. In *VLDB* (Vol. 96, pp. 134-145).

Tortajada, M., Oliver, A., Martí, R., Ganau, S., Tortajada, L., Sentís, M. & Zwiggelaar, R. (2014). Breast peripheral area correction in digital mammograms. *Computers in biology and medicine*, *50*, 32-40.

Traina, A. J., Watanabe, C. Y., Ribeiro, M. X., & Traina, C. (2012, December). A statistical associative classifier with automatic estimation of parameters on computer aided diagnosis. In *Machine Learning and Applications (ICMLA), 2012 11th International Conference on* (Vol. 1, pp. 564-567). IEEE.

Tsai, C. F., Tsai, C. W., Chen, C. P., & Lin, F. C. (2002). A Multiple-Searching Approach to Genetic Algorithms for Solving Traveling Salesman Problem. In *JCIS* (pp. 362-366).

Tsay, Y. J., & Chiang, J. Y. (2005). CBAR: an efficient method for mining association rules. *Knowledge-Based Systems*, *18*(2), 99-105.

Tseng, V. S., Wang, M. H., & Su, J. H. (2005, April). A new method for image classification by using multilevel association rules. In *Data Engineering Workshops, 2005. 21st International Conference on* (pp. 1180-1180). IEEE.

Vafaie, H., & De Jong, K. (1992, November). Genetic algorithms as a tool for feature selection in machine learning. In *Tools with Artificial Intelligence, 1992. TAI'92, Proceedings., Fourth International Conference on* (pp. 200-203). IEEE.

Vainio, H., & Bianchini, F. (2002). Breast cancer screening, IARC handbooks of cancer prevention. *IARC, Lyon*, 157-170.

Varela, C., Tahoces, P. G., Méndez, A. J., Souto, M., & Vidal, J. J. (2007). Computerized detection of breast masses in digitized mammograms.*Computers in Biology and Medicine*, *37*(2), 214-226.

Vasantha, M., Bharathi, V. S., & Dhamodharan, R. (2010). Medical image feature, extraction, selection and classification. *International Journal of Engineering Science and Technology*, *2*(6), 2071-2076.

Veldkamp, W. J., Karssemeijer, N., Otten, J. D., & Hendriks, J. H. (2000). Automated classification of clustered microcalcifications into malignant and benign types. *Medical Physics*, *27*(11), 2600-2608.

Veloso, A., Meira Jr, W., de Carvalho, M., Pôssas, B., Parthasarathy, S., & Zaki, M. J. (2002, April). Mining Frequent Itemsets in Evolving Databases. In SDM (pp. 494-510).

Verma, B., & Zhang, P. (2007). A novel neural-genetic algorithm to find the most significant combination of features in digital mammograms. *Applied soft computing*, *7*(2), 612-625.

Viton, J. L., Rasigni, M., Rasigni, G., & Llebaria, A. (1996). Method for characterizing masses in digital mammograms. *Optical Engineering*, *35*(12), 3453-3459.

Vyborny, C. J., Giger, M. L., & Nishikawa, R. M. (2000). Computer-aided detection and diagnosis of breast cancer. *Radiologic Clinics of North America*, *38*(4), 725-740.

W. P. (1993). Comparative evaluation of pattern recognition techniques for detection of microcalcifications in mammography.*International Journal of Pattern Recognition and Artificial Intelligence*, *7*(06), 1417-1436.

Wadhwani, S., Wadhwani, A. K., & Saraswat, M. (2013). Classification of Breast Cancer Detection Using Artificial Neural Networks. *Current Research in Engineering, Science and Technology Journals*, *1*(3), 85-91.

Wan, J. W., & Dobbie, G. (2004, January). Mining association rules from XML data using XQuery. In *Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation-Volume 32* (pp. 169-174). Australian Computer Society

Wang, K., Tay, W. and Liu, B. 1998. An interestingnessbased interval merger for numeric association rules. *KDD-98*.

Wang, X., Georganas, N. D., & Petriu, E. M. (2011). Fabric texture analysis using computer vision techniques. *Instrumentation and Measurement, IEEE Transactions on*, *60*(1), 44-56.

Watanabe, C. Y., Ribeiro, M. X., Traina Jr, C., & Traina, A. J. (2011). SACMiner: A new classification method based on statistical association rules to mine medical images. In *Enterprise Information Systems* (pp. 249-263). Springer Berlin Heidelberg.

Watanabe, C. Y., Ribeiro, M. X., Traina, A. J., & Traina, C. (2012, December). A statistical associative classifier with automatic estimation of parameters on computer aided diagnosis. In *Machine Learning and Applications (ICMLA), 2012 11th International Conference on* (Vol. 1, pp. 564-567). IEEE.

Weszka, J. S., Dyer, C. R., & Rosenfeld, A. (1976). A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *4*(SMC-6), 269-285.

Wikipedia (2009). Breast cancer. (Available at. http://en.wikipedia.org/wiki/Breast_cancer (Accessed 08.June 2013).

Wirth, M. A., & Stapinski, A. (2003, June). Segmentation of the breast region in mammograms using active contours. In *Visual Communications and Image Processing 2003* (pp. 1995-2006). International Society for Optics and Photonics.

Woods, K. S., DOSS, C. C., BOWYER, K. W., SOLKA, J. L., PRIEBE, C. E., & KEGELMEYER JR, W. P. (1993). Comparative evaluation of pattern recognition techniques for detection of microcalcifications in mammography.*International Journal of Pattern Recognition and Artificial Intelligence*, *7*(06), 1417-1436.

Wu, C. H., Wang, J. Y., & Chen, C. J. (2012, July). Mining condensed rules for associative classification. In *Machine Learning and Cybernetics (ICMLC), 2012 International Conference on* (Vol. 4, pp. 1565-1570). IEEE.

Wu, T., Moore, R. H., & Kopans, D. B. (2010). U.S. Patent No. 7,764,820. Washington, DC: U.S. Patent and Trademark Office.

Y. Zhang, N. Tomuro, J. Furst, and D. S. Raicu, (2012) "Building an ensemble system for diagnosing masses in mammograms," *International journal of computer assisted radiology and surgery*, vol. 7, pp. 323-329, 2012.

Yang, Q., Li, T., & Wang, K. (2004). Building association-rule based sequential classifiers for web-document prediction. *Data mining and knowledge discovery*, *8*(3), 253-273

Yoda, K., Fukuda, T. Morimoto, Y. Morishita, S. and Tokuyama, T. (1997). Computing optimized rectilinear regions for association rules. *KDD-97*.

Yu, L., & Liu, H. (2003, August). Feature selection for high-dimensional data: A fast correlation-based filter solution. In *ICML* (Vol. 3, pp. 856-863).

Yu, S., & Guan, L. (2000). A CAD system for the automatic detection of clustered microcalcifications in digitized mammogram films. *Medical Imaging, IEEE Transactions on*, *19*(2), 115-126.

Yuan, Y., Giger, M. L., Li, H., & Sennett, C. (2008, March). Correlative feature analysis of FFDM images. In *Medical Imaging* (pp. 69151L-69151L). International Society for Optics and Photonics.

Yun, J., Zhanhuai, L., Yong, W., & Longbo, Z. (2005, November). Joining associative classifier for medical images. In *Hybrid Intelligent Systems, 2005. HIS'05. Fifth International Conference on* (pp. 6-pp). IEEE.

Yusof, N. M., Isa, N. A., & Sakim, H. A. (2007). Computer-aided detection and diagnosis for microcalcifications in Mammogram: a review. *International Journal of Computer Science and Network Security*, *7*, 202-208.

Zaıane, O. R., Antonie, M. L., & Coman, A. (2002). Mammography classification by an association rule-based classifier. *MDM/KDD*, 62-69.

Zhao, Q., & Bhowmick, S. S. (2003). Association rule mining: A survey. Nanyang Technological University, Singapore.

Zheng, B., Chang, Y. H., Wang, X. H., & Good, W. F. (1999, July). Comparison of artificial neural network and Bayesian belief network in a computer-assisted diagnosis scheme for mammography. In *Neural Networks, 1999. IJCNN'99. International Joint Conference on* (Vol. 6, pp. 4181-4185). IEEE.

Zheng, B., Mello-Thoms, C., Wang, X. H., & Gur, D. (2007, April). Improvement of visual similarity of similar breast masses selected by computer-aided diagnosis schemes. In *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on* (pp. 516-519). IEEE.

Zheng, B., Qian, W., & Clarke, L. P. (1996). Digital mammography: mixed feature neural network with spectral entropy decision for detection of microcalcifications. *Medical Imaging, IEEE Transactions on*, *15*(5), 589-597.

Zheng, B., Qian, W., & Clarke, L. P. (1996). Digital mammography: mixed feature neural network with spectral entropy decision for detection of microcalcifications. *Medical Imaging, IEEE Transactions on*, *15*(5), 589-597.

Zhou, Q. Q., Kasabov, N., Benwell, G. L., Purvis, M., & Zhang, F. (1998). Neuro-fuzzy methods for environmental modelling.

Zou, F., Zheng, Y., Zhou, Z., & Agyepong, K. (2008, June). Gradient vector flow field and mass region extraction in digital mammograms. In *Computer-Based Medical Systems, 2008. CBMS'08. 21st IEEE International Symposium on* (pp. 41-43). IEEE.

BIODATA OF STUDENT

The student was born on 5th March, 1974 in Dindugul District, in the state of Tamilnadu, India. She obtained her early education at her native town and finished higher secondary education from Our Lady of Lourdes Girls Higher Secondary School, Dindugul, Tamilnadu, India. Afterwards, she proceeded to M.V.M. College for Women for her Bachelor Degree in Mathematics (B.Sc) in 1994. After completion of Bachelor Degree, she joined Masters in Computer Application (MCA) in the year 1994 and completed in the year 1997.

During the year from 1999 to 2014 she has been working as a Lecturer in few colleges in Malaysia and Saudi Arabia. In the meantime she has completed her Masters in Philosophy (M.Phil) Computer Science in Bharathiyar University, Coimbotore, Tamil Nadu, India and later in the year 2012 she joined in Universiti Putra Malaysia for her Doctor of Philosophy (PhD) programme in the field of Intelligent System . At present she is working as a Lecturer in Asia Pacific University, Malaysia. The Author is married in 1999 and is blessed with two kids.

LIST OF PUBLICATIONS

A. International Journals (Published)

Nirase Fathima Abubacker A. Azman, M.A.A. Murad and S. Doraisamy (2015), "Effective Rule Based Classifier using Multivariate Filter and Genetic Miner for Mammographic ImageClassification", Research Journal of Applied Sciences, Engineering and Technology, Maxwell Scientific Organization, *10*(5), pp.591-59

B. International Journals (Published)

Nirase Fathima Abubacker A. Azman, M.A.A. Murad and S. Doraisamy (2016), *"An Integrated Method of Associative Classification and Neuro Fuzzy Approach for Effective Mammographic Classification",* Journal of Neural Computing & Applications

C. Book Chapter (Published)

Nirase Fathima Abubacker A. Azman, M.A.A. Murad and S. Doraisamy (2014), "*Correlation-Based Feature Selection for Association Rule Mining in Semantic Annotation of Mammographic Medical Images*" 14th International Conference on Asia Information Retrieval Society' 2014, Springer LNCS 8870, pp 482-493, Springer International Publishing Switzerland 2014

D. International Conferences (Published)

Nirase Fathima Abubacker A. Azman, M.A.A. Murad and S. Doraisamy (2013), "An Approach For The Skull Fracture Detection In Dicom Images Using Neighboring Pixels" *Intelligent Systems Design and Applications (ISDA), 2013 13th International Conference on* , vol., no., pp.177,181, 8-10 Dec. 2013, doi: 10.1109/ISDA.2013.6920731

E. International Conferences (Accepted)

Nirase Fathima Abubacker A. Azman, M.A.A. Murad and S. Doraisamy (2016), *"Adaptive Associative Classifier for Mammogram Classification"* SAI Intelligent Systems Conference 2016, September 21-22, 2016 | London, UK

F. International Journals (Submitted)

Nirase Fathima Abubacker, A. Azman, M.A.A. Murad and S. Doraisamy (2016), *"An Improved Peripheral Enhancement of Mammogram Image by using Filtered Region Growing Segmentation"* Japanese Journal The Institute of Electronics, Information and Communication Engineers (IEICE)

Nirase Fathima Abubacker, A. Azman, M.A.A. Murad and S. Doraisamy (2016), *"An Effective Rule Refinement Technique for Adaptive Classifier in Mammogram Image Classification"* International Journal of Advances in Electrical and Computer Engineering (AECE)

UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT

ACADEMIC SESSION : ________________

TITLE OF THESIS / PROJECT REPORT :

DYNAMIC RULE REFINEMENT STRATEGY OF ASSOCIATIVE CLASSIFIER FOR EFFECTIVE MAMMOGRAPHIC IMAGE CLASSIFICATION

NAME OF STUDENT: NIRASE FATHIMA ABUBACKER

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (√)

☐	CONFIDENTIAL	(Contain confidential information under Official Secret Act 1972).
☐	RESTRICTED	(Contains restricted information as specified by the organization/institution where research was done).
☐	OPEN ACCESS	I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for :

☐	PATENT	Embargo from ________ until ________ (date) (date)

Approved by:

____________________	____________________
(Signature of Student) New IC No/ Passport No.:	(Signature of Chairman of Supervisory Committee) Name:
Date :	Date :

[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted.]