Cairo University

COMPUTER AIDED DIAGNOSTIC SYSTEM FOR DIGITAL MAMMOGRAPHY

By

Hani Mahmoud Mohammed Bu-Omer

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Biomedical Engineering and Systems

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

COMPUTER AIDED DIAGNOSTIC SYSTEM FOR DIGITAL MAMMOGRAPHY

By
Hani Mahmoud Mohammed Bu-Omer

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Biomedical Engineering and Systems

Under the Supervision of

Prof. Dr. Abou-Bakr M. Youssef

………………………………..

Professor of Biomedical Engineering
Biomedical Engineering and Systems
Faculty of Engineering, Cairo University

Prof. Dr. Yasser M. Kadah

………………………………..

Professor of Biomedical Engineering
Biomedical Engineering and Systems
Faculty of Engineering, Some University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

COMPUTER AIDED DIAGNOSTIC SYSTEM FOR DIGITAL MAMMOGRAPHY

By
Hani Mahmoud Mohammed Bu-Omer

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
Biomedical Engineering and Systems

Approved by the
Examining Committee

Prof. Dr. Abou-Bakr M. Youssef **(Thesis Main Advisor)**

Prof. Dr. Yasser M. Kadah **(Thesis Advisor)**

Prof. Dr. Ayman M. Eldeib **(Internal Examiner)**

Prof. Dr. Mohamed I. Eladawi **(External Examiner)**
Professor of Electronics Engineering, Faculty of Engineering, Helwan University

---

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2017

Engineer's Name: Hani Mahmoud Mohammed Bu-Omer
Date of Birth: 21 / 04 / 1986
Nationality: Yemeni
E-mail: hani1393313@eng1.cu.edu.eg, hani2011ye@yahoo.com
Phone: +201154430998 / +9675305007
Address: 4 Haret Alorobah, Altokhy St., Faisal, Giza.

Registration Date: 01 / 03 / 2013
Awarding Date: …./…./ 2017
Degree: Master of Science
Department: Biomedical Engineering and Systems

Supervisors:

Prof. Dr. Abou-Bakr M. Youssef
Prof. Dr. Yasser M. Kadah

Examiners:

Prof. Dr. Abou-Bakr M. Youssef	(Thesis main advisor)
Prof. Dr. Yasser M. Kadah	(Thesis advisor)
Prof. Dr. Ayman M. Eldeib	(Internal examiner)
Prof. Dr. Mohamed I. Eladawi	(External examiner)
- Faculty of Engineering, Helwan University	

Title of Thesis:

Computer Aided Diagnostic System for Digital Mammography

Key Words:
Digital Mammography; Breast Cancer; Computer-Aided Diagnosis; Machine Learning; Medical Image Processing.

Summary:
Digital mammography has emerged as the most popular screening technique for early detection of breast cancer and other abnormalities in human breast tissue and it continues to be the standard screening tool for breast cancer detection. In this study, we developed Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) systems and we tested and experimentally verified these systems using the publicly available MIAS and DDSM mammographic databases, respectively. We first applied preprocessing technique to enhance the breast peripheral regions, then we extracted a set of 543 different textural features from the regions of interest, and after that the most relevant features were selected using four selection methods. The selected features are then used as input to different classifiers to differentiate between normal and abnormal breast masses in CADe system, and to distinguish between normal, benign and malignant breast tissues in CADx system. All cases were correctly detected in the proposed CADe system, while CADx system achieved 98.67% overall accuracy.

Acknowledgments

After an intensive period of a year, today is the day: writing this note of thanks is the finishing touch on my thesis. It has been a period of intense learning for me, not only in the scientific arena, but also on a personal level. Writing this thesis has had a big impact on me. I would like to reflect on the people who have supported and helped me so much throughout this period.

Foremost, all praises and thanks be to **ALLAH**, the Lord of the worlds, for the good health, wellbeing and guidance that were necessary to complete this thesis.

I am using this opportunity to express my profound gratitude to my thesis advisors, **Prof. Abou-Bakr M. Youssef** and **Prof. Yasser M. Kadah**, for the continuous support in my M.Sc. study and research, for their patience, motivation, enthusiasm, immense knowledge and for challenging my potential to innovate throughout my academic career. Their guidance helped me in all the time of research and writing of this thesis.

Besides my advisors, I would like to thank the rest of my thesis defense committee: **Prof. Ayman M. Eldeib** and **Prof. Mohamed I. Eladawy**, for their encouragement, insightful comments, and hard questions.

I would also like to express my sincere gratitude to the Head of Biomedical Engineering and Systems Department, **Prof. Ahmed M. Badawi** for providing me with all the facilities during the period of my M.Sc. study.

My sincere thanks also goes to **Eng. Sheikh/ Abdullah Ahmed Bugshan**, the Chairman of the Board of Trustees of Hadhramout Establishment for Human Development, for all his care and support of outstanding students. I appreciate his partial support during my M.Sc. study and putting me through the best education possible. He gave me the chance to attend the 11th IEEE EMBS International Summer on Biomedical Imagine (France, 2014).

I would also like to sincerely thank **Mugahed A. Al-antari** and **Asaad M. Anaam** who provided technical support and encouragement that greatly assisted my study.

I don’t forget to thank all my close friends **Riyadh Baashn**, **Atef Alhuraibi**, **Faez Alamodi**, **Meena Makary, Ahmed Abdullah, Faez Alhamed**, **Muhammad Bakhrebah**, **Abdulrahman Baalawi**, **Jamal Alakbari**, **Ahmed Albity**, **Murad Baamodi**, **Abdullah Alkalaly**, **Amer Binsadoon, Fahed Baemeran, Abdullah Bawazir** and all my friends from Hadhramout, Yemen and Egypt for their love, support and encouragement.

Finally, I must express my very profound gratitude to my **Parents** and to my **Sister**, **Brothers** and all my **Family** for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

Dedication

To the loving memories of my **Father**, **Uncle** and **Grandmothers**.

To my **Mother**; your sacrifices, support, encouragement, and continuous love have sustained me throughout my life and made everything I have done possible.

To my **Sister**, **Brothers** and all my **Family** members; nothing can describe your love and care.

To my advisor, **Prof. Yasser M. Kadah**; your support and encouragement have made this achievement possible.

Table of Contents

List of Tables

List of Figures

Nomenclature

ADEWNN / Ada-DEWNN	Adaptive Differential Evolution Wavelet Neural Network
ACR	American College of Radiology
ANN	Artificial Neural Networks
ARCH	Architectural distortion
ASYM	Asymmetry
AUC	Area Under Curve
BI-RADS	Breast Imaging-Reporting and Data System
BI	Blurred Image
BBS	Branch and Bound Selection
CAD	Computer Aided Diagnosis
CADe	Computer Aided Detection
CADx	Computer Aided Diagnosis
CBIR	Content-Based Image Retrieval
CC	Cranio-Caudal
CIRC	Well-defined/circumscribed masses
DBN	Deep Belief Network
DCT	Discreet Cosine Transform
DDSM	The Digital Database For Screening Mammography
DF-BrCanD	diverse features based breast cancer detection
DWT	Discrete Wavelet Transform
ELM	Extreme Learning Machine
FCM	Fuzzy C-Mean
FFDM	Full Field Digital Mammography
FDA	Food and Drug Administration
F-GA	Frequency-Genetic Algorithm
FN	False Negative
FNR	False Negative Rate
FP	False Positive
FPR	False Positive Rate
FSLR	Forward stepwise linear regression
GLCM	Gray Level Co-Occurrence Matrix
GLPF	Gaussian Low Pass Filter

IRMA	Image Retrieval in Medical Applications
K-NN	K-voting Nearest Neighbor
LCP	Local Configuration Pattern
LDA	Linear Discriminant Analysis
MIAS	The Mammographic Image Analysis Society
MLO	Mediolateral-Oblique
MLP	Multi-Layer Perceptron
MISC	Ill-defined Masses
NLM	Non-Local Means
NPV	Negative Predictive Value
NTP	Normalized Thickness Profile
OWBPE	Opposite Weight Back Propagation per Epoch
OWBPI	Opposite Weight Back Propagation per Pattern in Initialization
OWBPP	Opposite Weight Back Propagation per Pattern
pAUC	Partial Area Under the Curve
PCA	Principal Component Analysis
PCET	Polar Complex Exponential Transform
PNN	Probabilistic Neural Network
PPV	Positive Predictive Value
QDA	Quadratic Discriminant Analysis
ROC	Receiver Operating Characteristic
RBF	Radial Basis Function
ROI	Region of Interest
S3VM	Semi Supervised Support Vector Machine
SBS	Sequential Backward Selection
SFFS	Sequential Floating Forward Selection
SFS	Sequential Forward Selection
SI	Segmented Image
SPIC	Spiculated masses
SSL	Semi-Supervised Learning
SVM	Support Vector Machines
TN	True Negative
TP	True Positive
WBCD	Wisconsin Breast Cancer Dataset
WPD	Wavelet Packet Decomposition

Abstract

Breast cancer is the most common cause of death in women and the second leading cause of cancer deaths worldwide. Digital mammography has emerged as the most popular screening technique for early detection of breast cancer and other abnormalities in human breast tissue and it continues to be the gold standard screening tool for breast cancer detection.

In this work, Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) systems are developed and tested using the public and freely available mammographic databases named MIAS and DDSM databases, respectively. CADe system is used to differentiate between normal and abnormal tissues and it assists radiologists avoid missing a breast abnormality, while CADx is developed to distinguish between normal, benign and malignant breast tissues and it helps radiologists to decide whether a biopsy is needed when reading a diagnostic mammogram or not.

Any CAD system is constituted of common stages including: preprocessing and segmentation of mammogram images, extraction of regions of interest (ROI), features extraction, features selection and classification. In both proposed CAD systems, ROIs are selected using window size of 32×32 pixels then a total of 543 features from four different feature categories are extracted from each ROI and then normalized. After that the selection of the most relevant features is performed using four different selection methods from MATLAB Pattern Recognition Toolbox v.5 (PRtool5) named Sequential Backward Selection (SBS), Sequential Forward Selection (SFS), Sequential Floating Forward Selection (SFFS) and Branch and Bound Selection (BBS) methods. We also utilized Principal Component Analysis (PCA) as the fifth method to reduce the dimensions of the features set. After that we used different classifiers such as Support Vector Machines (SVM), K-voting Nearest Neighbor (K-NN), Quadratic Discriminant Analysis (QDA) and Artificial Neural Networks (ANN) for the classification. Both CAD systems have the same implementation stages but different output. CADe systems is designed to detect breast abnormalities while CADx system indicates the likelihood of malignancy of lesions. Finally, we independently compared between performance of all classifiers with each selection method in both systems.

The evaluation of the proposed CAD systems is done using performance indices such as sensitivity, specificity, the area under curve (AUC) of the Receiver Operating Characteristic (ROC) curves, the overall accuracy and Cohen-k factor.

Both CAD systems provided encouraging results. These results were different corresponding to selection method and classifier.

Chapter 1 : Introduction

1.1. Overview and Problem Definition

Before the twentieth century breast cancer was considered as shameful disease and women felt shy to complain from health problems related to their breast regions and tended to silently suffer rather than talking to their doctors. After surgery has been advanced, and long-term rates of survival are improved, the awareness of this disease started to raise among women around the world and the successful treatment possibility has been improved [1].

Breast cancer is the abnormal growth of breast cells, which usually starts in the lobules of the inner lining of the milk ducts. Different types of breast cancer exist with different stages, truculence and genetic makeup. 10-years disease-free survival rates vary from 98% to 10% with best treatment plan. Breast cancer treatment includes various strategies such as surgery, chemotherapy, hormone therapy and radiation. Breast cancer is still being a significant public healthcare problem among women and it is the most common cancer around the world. The cause of this disease still remain undetermined and this make primary prevention to be impossible. It is believed that early detection of breast cancer is the most promising way to lower the number of women suffering from it and this improves the chances to provide proper treatment options so that treatment will work with better results [2].

Among women in U.S., breast cancer is the most commonly diagnosed cancer besides skin cancer and its death rates are the highest among other cancers besides lung cancer. In 2016, it is expected to diagnose an approximately 246,660 new cases of invasive breast cancer in women, and the expected diagnosed cases of non-invasive breast cancer are 61,000 women in the U.S. Also About 2,600 new patients of invasive breast cancer in men are expected to be discovered [3].

Mammography is a special type of radiography, using low radiation levels to acquire images for breast to diagnose a consequent exist of abnormal structures that implies a disease like cancer. The early detection of strange mammary pathologies like non-palpable breast masses and calcifications is extremely important for successful cure of breast cancer patients. Mammography is the standard screening tool that is used to perform the task of breast cancer detection and it results to reduce at least 30% in breast cancer deaths in the world [4], [5].

The benefit of mammography screening has undergone some recent contention since definite evidence relating mammography with mortality may not be proven. In contrast, an Institute of Medicine Report on Mammography suggests that using mammography screening for earlier detection of breast cancer may be an important factor in decreasing mortality from breast cancer in recent years [6].

The computer-based systems may provide a second supportive alternative in detecting breast abnormalities by completing the expert knowledge of radiologists, and