DEVELOPMENT OF A DISCRETE WAVELET TRANSFORM AND ARTIFICIAL NEURAL NETWORK BASED CLASSIFICATION SYSTEM FOR MAMMOGRAM IMAGES

by

LUQMAN MAHMOOD MINA

Thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

September 2016

سورة المجادلة (11)

In the Name of Allah, the most Beneficent, the most Merciful

{Allah will raise those who have believed among you and those who were given knowledge, by degree. And Allah is acquainted with what you do}

Surah Al-Mujaadila (11)

DEDICATION

To my supervisor Professor Dr. Nor Ashidi Mat Isa

To my parents, who have made me, the man I am today

To my dearest friends Dr. Khamees Khalaf and Ph.D. student Mr. Salam Mohammed

To my dear wife Trifa for her unlimited love, support, patience and encouragement;

my three children: Helen, Aren and Asin for their understanding, patience, helping and bearing my absence at home;

my brothers and sisters who embrace me with their love, kindness and unconditional support as their eldest brother.

Luqman

ACKNOWLEDGEMENTS

First and foremost, I would like to thank and praise Allah, the almighty Merciful, who bestowed upon me by facilitating the successful accomplishment of this PhD research work. Respectfully, I acknowledge that my success is not only related to me but its possession to all the people whom guide me, teach me, pray for me and encourage me. From them, I learned the knowledge, patient, wisdom, humility and also how to gain access to my goals.

I would like to seize this opportunity to specially acknowledge to my supervisor, Prof. Dr. Nor Ashidi Mat Isa for his invaluable suggestions, dedication support, constructive effort and beneficial comments that have remarkably influenced to bring this work into light. His advices and words of encouragement helped me to overcome most of the difficulties I have faced. I am proud to conduct this research under his supervision.

The acknowledgement would be incomplete if I do not express my sincere gratitude to my co-supervisors, Prof. Dr. Kamal Zuhairi Zamli, for his support and encouragement to discuss and share his views about any issues related to this research. He always finds the time for listening to the little problems and roadblocks that unavoidably crop up in the course of performing research.

Special thanks to Universiti Sains Malaysia for their cooperation and providing the facilities required for this research. Also, special unreserved appreciation goes to the friendly staff of the school of Electrical and Electronic Engineering, for their cooperation and friendly attitude. My sincere appreciation is extended to Dr. Khamees Kh. Hasan for his invaluable support and encouragement during the study. I would like to extend my deepest appreciation to my best friends, the Phd student Mr. Salam Kareem for their support, prayers and endless friendship and encouragement.

I would like to express the deepest appreciation to the Ministry of High Education and Scientific Research in Kurdistan\Iraq for its support. I wish acknowledge to Ministry of Electricity for their cooperation and facilities on approval for giving me the opportunity to seek for this Ph.D. degree.

Last but not the least, I would like to dedicate my gratitude to the dearest and nearest people to my heart, my family for their unremitting support, encouragement and boundless patience with me throughout the years of the research, especially my beloved wife and my wonderful kids, who patiently underwent the alienation and distance from the family without complaints and are supportive of me. This thesis is for you. Finally, to all my family members, I am forever indebted for your understanding, support, endless patience and encouragement when it matters the most. You are all like bright candles in the dark days and difficult times, so thank again.

TABLE OF CONTENTS

CHAPTER THREE: METHODOLOGY

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Description
1-D	One Dimension
1D-DWT	One Dimension Discreet Wavelet Transform
2-D	Two Dimensions
2D-DWT	Two Dimension Discreet Wavelet Transform
3-D	Three Dimension
A	Approximation Subband
AMBE	Absolute Mean Brightness Error
AAMBE	Average Absolute Mean Brightness Error
ABC	Artificial Bee Colony
AMLT	Adaptive Multilevel Threshold
ANN	Artificial Neural Network
ANCE	Adaptive Neighborhood Contrast Enhancement
AMBE	Absolute Mean Brightness Error
AUC	Area Under the Curve
BP	Backpropagation
BPANN	Backpropagation Artificial Neural Network
BPN	Backpropagation Network
CAD	Computer Aided Diagnosis
CBT	Clustering-based Thresholding
CLAHE	Contrast Limited Adaptive Histogram Equalization
CT	Computed Tomography
D	Diagonal subband

dB	Decibel
DCIS	Ductal Carcinoma In Situ
DDSM	Digital Database for Screening Mammography
DWT	Discreet Wavelet Transform
ECGs	Electrocardiography
EEGs	Electroencephalogram
ELMANN	Extreme Learning Machin Artificial Neural Network
EM	Expectation Maximization
FFDM	Full Field Digital Mammogram
FN	False Negative
FNF	False Negative Fraction
FANC	Fine Needle Aspiration Cytology
FP	False Positive
FROC	Free Response Operating Characteristic
GA	Genetic Algorithm
GLCM	Gray Level Co-occurrence matrix
GLRLM	Gray Level Run-Length Method
GMRF	Gaussian Markov Random Field
GN	Genetic Network
H	Horizontal subband
HE	Histogram Equalization
HH	High-High
HL	High-Low
HMLP	Hybrid Multilayer Perceptron
HT	Histogram Shaped-based Thresholding

HWT	Haar Wavelet Transform
IARC	International Agency for Research Center
IT	Information-based Thresholding
KNN	K-Nearest Neighborhood
LDA	Linear Discriminate Analysis
LH	Low-High
LL	Low-Low
LM	Levenberg Marquard
LMLO	Left Medio-Lateral Oblique
LMS	Least Mean Square
LT	Locally Adaptive Thresholding
MC	Microcalcification
MIAS	Mammographic Images Analysis Society
MLP	Multilayer Perceptron
Mod-max	Modules-maximum
MPM	Maximizer of the Posterior
MPV	Mean Pixel Value
MRA	Multi-Resolution Analysis
MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
MWA	Multiresolution Wavelet Analysis
OAT	Object Attribute Thresholding
PCA	Principal Component Analysis
PET	Positron Emission Tomography

PNN	Probabilistic Neural Network
POSWNN	Particle Swarm Optimization Wavelet Neural Network
PSNR	Peak Signal to Noise Ratio
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
RMLO	Right Medio-Lateral Oblique
ROC	Receiver Operating Curve
ROI	Region of Interest
SFM	Screen Film Mammogram
SGLDM	Spatial Gray Level Dependency Matrix
SONN	Swarm Optimization Neural Network
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
TPF	True Positive Fraction
TWSVM	Twin Support Vector Machine
UK	United Kingdom
US	Ultrasonography
V	Vertical subband
WHO	World Health Organization

LIST OF SYMBOLS

Symbol	**Description**
d	Difference between mean and median
f_r	Reduction Factor
R_{inner}	Inner Binary Mask
R_{outer}	Outer Binary Mask
$R_{skin\text{-}air}$	Skin-air Interface Region
$R'_{skin\text{-}air}$	Binary Breast Profile LMLO
$R''_{skin\text{-}air}$	Binary Breast Profile RMLO
T_0	Initial Threshold
T_n	Current Threshold
T_{inner}	Inner Threshold
T_{outer}	Outer Threshold
μ	Mean Value
δ	Median Value
$\Phi(x)$	Haar Scaling Function
$\Psi(y)$	Haar Wavelet Function
$\Phi(x,y)$	Approximation Subband
$\Psi(x,y)$	Detail Subband

PEMBANGUNAN SISTEM PENGKELASAN BERASASKAN JELMAAN GELOMBANG KECIL DISKRET DAN RANGKAIAN NEURAL BUATAN UNTUK IMEJ MAMOGRAM

ABSTRAK

Pada masa ini, terdapat pelbagai sistem diagnosis bantuan komputer (CAD) yang dibangunkan sejak beberapa tahun lalu untuk membantu ahli radiologi dalam pengecaman lesi mamografi yang boleh menunjukkan kehadiran kanser payudara. Walau bagaimanapun, prestasi CAD terhad oleh dua isu utama iaitu (i) kawasan yang tidak diingini (seperti label segi empat tepat berintensiti tinggi, pita, artifak, antara muka kulit dan air, dan lain-lain) yang boleh mengganggu pengecaman kanser payudara dan mengurangkan kadar ketepatan CAD, (ii) ketidakteraturan tekstur mamogram yang meliputi ciri-ciri seperti entropi, tenaga, kepencongan, kurtosis, min dan sisihan piawai yang berhubung kait dalam domain ruang dan tidak penting untuk pengelasan. Oleh itu, bagi menangani masalah yang dinyatakan di atas, sistem CAD yang lebih baik untuk imej mamogram dicadangkan. CAD yang dicadangkan ini terdiri daripada tiga peringkat utama, iaitu prapemprosesan, pengekstrakan ciri dan pengelasan imej mamogram. Pada peringkat prapemprosesan, *Adaptive Multilevel Threshold* (AMLT), yang berjaya menyingkirkan kawasan yang tidak diingini seperti yang dinyatakan sebelum ini, dicadangkan. Hal ini memberikan kelebihan kepada sistem dengan membolehkan pencarian terhadap keabnormalan terkekang pada lingkungan tisu payudara tanpa menjejaskan kawasan yang tidak diingini dalam latar belakang imej. Pada peringkat pengekstrakan ciri, dua ciri baharu iaitu median maksimum dan minimum subjalur berfrekuensi tinggi dicadangkan untuk pengkelasan imej mamogram kepada kategori normal, benigna dan malignan. Analisis plot kotak membuktikan bahawa kedua-dua ciri baharu tiada hubung kait dan penting untuk pengelasan imej mamogram berbanding dengan ciri-ciri konvensional. Pada peringkat pengelasan, rangkaian perseptron berbilang lapis (MLP) digunakan untuk mengelaskan mamogram normal dan tidak normal pada fasa pertama dan mamogram benigna dan malignan pada fasa kedua. Keputusan purata yang terhasil daripada 322 imej mamogram pada fasa pertama merumuskan bahawa pendekatan yang dicadangkan berjaya mencapai keputusan yang boleh harap dengan ketepatan sebanyak 96,27%, kepekaan sebanyak 94,78% dan kekhususan sebanyak 96.60%. Di samping itu, keputusan purata yang terhasil daripada 115 imej yang tidak normal mempunyai ketepatan, kepekaan dan kekhususan, masing-masing sebanyak 95.65%, 96.18% dan 95.38%. Keputusan eksperimen akhir menunjukkan bahawa sistem pengelasan mamogram yang dibangunkan mampu mencapai pengelasan tertinggi berbanding dengan sistem terkini yang lain. Prestasi pengelasan yang menggalakkan ini menunjukkan bahawa sistem yang dicadangkan tersebut boleh digunakan untuk membantu ahli patologi dalam menjalankan proses diagnosis.