

بسم الله الرحمن الرحيم

SUDAN UNIVERSITY OF SCIENCES & TECHNOLOGY

FACULTY OF GRADUATE STUDIES

DEPARTMENT OF BIOMEDICAL ENGINEERING

**Breast Tumors Classification Using Adaptive Neuro-Fuzzy
Inference System (ANFIS)**

**تصنيف أورام الثدي باستخدام نظام الاستدلال العصبي-الضبابي المنطقي
المتكيف**

**A Dissertation submitted in partial fulfillment of the
requirements for the Degree of Master of Science in
Biomedical Engineering**

Prepared by:

Sahar Haj Ali Abu Likailik

Supervisor:

Dr. Zeinab Adam Mustafa

March 2016

Dedication

To...

Our most loved ones....

Parents, families and closest friends,

With their loving and promising support,

Their unbounded motivation that has given us the perseverance we need to accomplish our goals.

Special thanks go to my husband, my guardian angel, my reason of success, for his support in all my life.

And mostly,

To all those women who have struggled with breast cancer

Acknowledgement

In the name of the Almighty Allah, to whom we are grateful for, for all the blessings, guidance and constant perseverance that we have been gifted with throughout this experience.

I confer our utmost gratitude to **Dr. Zeinab Adam Mustafa** for all her support, guidance and the benefit of her experience that she has bestowed upon us. Most importantly, for the confident faith she has entrusted me with every step of the way.

A token of my appreciation and thanks is handed out to **Dr. Ahmed Abdallah**. His time to provide me with his guidance.

Last but not least, the unconditional love and support from dear families and friends and specially my dear husband cannot be forgotten, and without which, accomplishing this task would not have been possible.

مستخلص البحث

سرطان الثدي هو حاليا السبب الأول في وفيات النساء في جميع أنحاء العالم. في السودان نسبة لزيادة النسبة، والكشف في المراحل المتأخرة والبدائية المبكرة للمرض يجعل الكشف المبكر وتشخيص سرطان الثدي مهمة طاعية.

حاليا تعتبر الاشعة السينية لتصوير الثدي هي الأكثر فعالية والأقل تكلفه والأكثر حساسية لاكتشاف الاورام الصغيرة، ولكن تعتمد حساسية الجهاز على كثافة انسجة الثدي مما يؤثر على دقة التشخيص وبالتالي، هناك ضرورة كبيرة لتطوير طرق الكشف الألي وتصنيف المناطق المشبوهة في تصوير الثدي بالأشعة السينية، كوسيلة لمساعدة أطباء الأشعة لتحسين فعالية برامج الفحص و تجنب الخزعات غير الضرورية.

الهدف من هذه الدراسة هو إنشاء نظام واجهة كمبيوتر لتحديد المكان ، وكشف وتصنيف كتلة الثدي باستخدام نظام الاستدلال العصبي-الضبابي المنطقي المتكيف.

نظام الاستدلال العصبي-الضبابي المنطقي المتكيف المستخدم في الدراسة للكشف عن سرطان الثدي يعتمد على خمسة مدخلات تميز السرطان. هذا النظام يجمع بين خاصية قدره علي التكيف عند الشبكات العصبية و النهج النوعي للمنطق الضبابي.

ويتم تنفيذ هذه المراحل باستخدام برنامج MATLAB. تم تصميم واجهة المستخدم الرسومية (GUI) لكل مرحلة من مراحل الخوارزمية. أظهرت نتائج المنهجيات المقترحة بان لها قدرة عالية في تحسين صور الثدي، توطين، كشف وتصنيف الورم الثدي. وكان النظام قادرا على تحقيق دقة حساسية 94.4٪، و 100٪ خصوصية، 97.1٪ القيمة التنبؤية الإيجابية، و 100٪ القيمة التنبؤية السلبية، قيمة Az تساوي 0.972 ودقة التصنيف الكلية 98٪.

اثبت النظام فعاليته في تحسين صور الثدي والتميز بين الأورام الخبيثة والحميدة مع مستوى فعال من الدقة.

Abstract

Breast cancer is currently the foremost cause in women's mortality worldwide. In Sudan, the increasing incidence, detection at late stages and the early onset of the disease makes early detection and diagnosis of breast cancer an overbearing task.

Currently, X-ray mammography is the single most effective, low-cost, and highly sensitive technique for detecting small lesions. However, the sensitivity of mammography is highly challenged by the presence of dense breast parenchyma, which deteriorates both detection and characterization tasks. Thus, there is a significant necessity for developing methods for automatic detection and classification of suspicious areas in mammograms, as a means of aiding radiologists to improve the efficacy of screening programs and avoid unnecessary biopsies.

The objective of this study is to create a computer interfacing system for the localization, detection and classification of breast masses using ANFIS.

The ANFIS classifier was used to detect the breast cancer when five features defining breast cancer indications were used as inputs. The proposed ANFIS model combined the neural network adaptive capabilities and the fuzzy logic qualitative approach.

Results demonstrated the proposed methodologies have high potential in enhancing breast images, localizing, detecting and classifying the breast tumor. The system was able to achieve an accuracy of 94.4% sensitivity, 100% specificity, 97.1% positive predictive value, 100% negative predictive value, an Az value of 0.972 and an overall classification accuracy of 98%.

The created systems have therefore, proved to effectively enhance the quality of the breast images and discriminate between malignant and benign tumors with an effective level of precision.

CONTENTS

Dedication	ii
Acknowledgement	iii
Abstract	v
List of Tables	viii
List of Figures	viii
Glossary	x

1. INTRODUCTION	1
1.1 General View	2
1.2 Hypothesis	3
1.3 Objectives	3
1.4 Dissertation Layout	4
2 THEORETICAL BACKGROUND	5
2.1 Female Breast	6
2.1.1 Breast Cancer	7
2.1.2 Breast Cancer Lesions	8
2.2 Mammography	11
2.3 Digital Mammography	11
2.4 Mammographic Interpretation by Radiologists	11
2.5 Digital Images	12
2.5.1 Digital Image Processing	13
2.5.2 Applications of Digital Image Processing	13
2.5.3 MATLAB and the Image Processing Toolbox	16
2.6 Image Enhancement Techniques	16
2.7 Features	17
2.7.1 Texture Features	18
2.7.1.1 Gray Level Co-occurrence Matrices (GLCM)	18
2.8 ANFIS Classification	20
2.8.1 ANFIS Architecture	20
2.8.2 ANFIS Learning Algorithm	22
3 LITERATURE REVIEW	24
4 ALGORITHM METHODOLOGY AND IMPLEMENTATION	28
4.1 System Overview	29
4.2 Proposed Flowchart Algorithm	29
4.3 Data Source	30
4.4 Image Enhancement	30
4.5 Pattern Recognition and Classification	31
4.5.1 Image Segmentation	31
4.5.1.1 The Watershed Transform	31
4.5.2 Feature Selection	32
4.5.2.1 Extracted Features	32

4.5.3	Classification Stage.....	33
5	RESULTS AND DISCUSSIONS.....	36
5.1	Results.....	37
5.1.1	Results for the Image Processing and Enhancement System.....	37
5.1.2	Results for the Pattern Recognition and Classification System....	38
5.1.2.1	Extracting the ROIs.....	38
5.1.2.2	Classification Stage.....	38
5.1.2.2.1	Performance Measures.....	38
5.1.2.2.1.1	Confusion Matrix.....	39
5.1.2.2.1.2	ROC and the Area Under Curve (AUC).....	40
5.1.2.2.1.3	Obtained Results.....	40
5.2	Graphical User Interface (GUI) of the Systems.....	41
6	CONCLUSIONS AND RECOMMENDATIONS.....	42
6.1	Conclusions.....	43
6.2	Recommendations.....	43
7	LIST OF REFERENCES.....	44

LIST OF TABLES

Table Number	Table Name	Page number
Table 5.1	Summary of classification performance	40

LIST OF FIGURES

Figure Number	Figure Name	Page number
2.1	Anatomy of the Female Breast	7
2.2	Microcalcifications	8
2.3	Side and overhead view of an architectural distortion	9
2.4	Morphologic spectrum of mammographic masses	9
2.5	Mass examples with different shapes and borders	10
2.6	Mass examples with different densities	10
2.7	Comparing an analog with a digital image of the same breast	11
2.8	Digitization of a continuous image	13
2.9	CAT slice image	14
2.10	Example of Gamma-ray imaging for a bone scans	15
2.11	Mammogram image representing side view of left breast with corresponding histogram	17
2.12	Example of how to generate a Co-occurrence matrix with eight grey levels	19
2.13	Adjacency of pixel in four directions (horizontal, vertical, left and right diagonals)	19
2.14	Basic structure of ANFIS	20
2.15	ANFIS Architecture with nine rules	22
4.1	illustrates the proposed algorithm flow chart	29
4.2	shows the ANFIS structure that used in the study	34
5.1	Image before and after processing	37
5.2	watershed segmentation result	38
5.3	Confusion matrix	39
5.4	Pattern recognition and classification system GUI with the illustration of a malignant mass in	41

	a breast image being detected and classified	
5.5	Pattern recognition and classification system GUI with the illustration of a benign mass in a breast image being detected and classified	41

GLOSSARY

ANFIS	: Adaptive Neuro-fuzzy Inference System
MF	: Membership Function
AUC	: Area Under Curve
BI-RADS	: Breast Imaging-Reporting and Data System
BLAS	: Basic Linear Algebra Subprograms
BSE	: Breast Self Examination
CAD	: Computer Aided Diagnosis
CAT	: Computerized Axial Tomography
CDF	: Cumulative Distribution Function
CLAHE	: Contrast Limited Adaptive Histogram Equalization
DWT	: Discrete Wavelet Transform
DWT	: Dyadic Wavelet Transforms
EISPACK	: Eigen System Package
FCLAHE	: Fuzzy Contrast-Limited Adaptive Histogram Equalization
FCM	: Fuzzy C-Means
GA-LFS	: Linear Forward Selection and Genetic Algorithm
GLCM	: Gray-Level Co-occurrence Matrices
GUI	: Graphical User Interface
K-nn	: K-nearest neighbor
LAPACK	: Linear Algebra Package
LINPACK	: Linear System Package
LRM	: Local Range Modification
LDA	: Linear Discriminant Analysis
LSMD	: Least-Squares Minimum Distancex
MATLAB	: Matrix Laboratory
MC	: Microcalcification
MCPCNN	: Multiple Circular Path Convolution Neural Network
MIAS	: Mammographic Image Analysis Society
MLP	: Multi-Layer Perceptron
MSE	: Mean Square Error
PSNR	: Peak Signal to Noise Ratio
RBF	: Radial Basis Function
RDWT	: Redundant Dyadic Wavelet Transform
RICK	: Radiology and Isotope Center of Khartoum
ROC	: Receiver Operating Characteristic
ROI	: Region Of Interest
SGLD	: Spatial Gray Level Dependence Matrices
SVM	:Support Vector Machine
PGM	: Portable Gray Map