بسم الله الرحمن الرحيم

MONA MAGHRABY

شبكة المعلومات الجامعية

التوثيق الالكتروني والميكروفيلم

MONA MAGHRABY

شبكة المعلومات الجامعية

التوثيق الإلكتروني والميكروفيلم

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات

يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار

MONA MAGHRABY

DIGITAL INFRARED IMAGING FOR BREAST CANCER DETECTION USING SEQUENTIAL MINIMAL OPTIMIZATION, KERNEL LOGISTIC REGRESSION AND MULTILAYER PERCEPTRON

By

Shaimaa Adel Abd El Halim

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
2020

DIGITAL INFRARED IMAGING FOR BREAST CANCER DETECTION USING SEQUENTIAL MINIMAL OPTIMIZATION, KERNEL LOGISTIC REGRESSION AND MULTILAYER PERCEPTRON

By

Shaimaa Adel Abd El Halim

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

Assoc. Prof. Dr. Amr A.Sharawi

Faculty of Engineering, Cairo University

Assist. Prof. Dr. Nader A.Rahman

Faculty of Engineering, Misr University for Science and Technology

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2020

Engineer's Name: shaimaa Adel Abd El Halim
Date of Birth: 30/8/1986
Nationality: Egyptian
E-mail: Engshimaa_adel@yahoo.com
Phone: 01094667029
Address: …….

Registration Date: 1/10/2012
Awarding Date: …./…./……..
Degree: Master of Science
Department: …….

Supervisors: Assoc. Prof. Dr. Amr A.Sharawi

Assist. Prof. Dr. Nader A.Rahman

Examiners:

Prof. Dr. Mai Saeed Mabrouk (External examiner)
Faculty of Engineering, Misr University for Science and Technology
Prof. Dr. Manal Abd El-Wahed (Internal examiner)
Faculty of Engineering, Cairo University
Assoc. Prof. Dr. Amr A.Sharawi (Thesis main advisor)
Faculty of Engineering, Cairo University

Title of Thesis:
DIGITAL INFRARED IMAGING FOR BREAST CANCER DETECTION USING SEQUENTIAL MINIMAL OPTIMIZATION, KERNEL LOGISTIC REGRESSION AND MULTILAYER PERCEPTRON

Key Words:
KERNEL LOGISTIC REGRESSION ,BREAST CANCER , MULTILAYER PERCEPTRON,SEQUENTIAL MINIMAL OPTIMIZATION

Summary
Breast cancer is the main reason for women's death. In Egypt, breast cancer is considered to be the most common cancer among females. Death percentage by breast cancer is up to 37.7% out of 12.621 new cases in 2008. Breast cancer starts when cells in the breast begin to grow out of control. However, there are many techniques help to discover the cancer in a more safe way. One of them is Digital Infrared Imaging (thermography); it is based on the metabolic activation and vascular circulation in both pre-cancerous tissue and the cancerous one. Radiographic images obtained from thermography equipment are one of the most f techniques that used for helping in early detection of breast cancer. This thesis proposes a method for breast cancer detection that uses image processing techniques. These techniques are applied to 142 breast digital thermal images; 77 of them are normal images and 65 are abnormal ones. Matlab, is used for detecting region of interest (ROI) and feature extraction. In addition, Weka is used for classification and feature selection. In the (ROI) extraction phase, active contour techniques are used, and then 72 statistical and textural features are extracted and used to feed classifiers. It is worth mentioning, there are three different classifiers. These classifiers are Sequential minimal optimization (SMO), kernel logistic regression, and Multilayer perception and their accuracy had reached to 99.29%, 98.59%, 96.4%. Afterwards it comes to the feature selection phase, at which the best dominant features are selected to minimize the processing time and effort. Accordingly, there are two types of wrapper feature selection methods are called best first and Greedy Step wise. These methods are used to minimize the number of features. Finally, only 6 features were selected, which gave us an accuracy of 99.29% using Sequential Minimal Optimization (SMO), 99.29% using Kernel Logistic Regression, and 97.18 using Multilayer Perceptron.

Disclaimer

I hereby declare that this thesis is my own original work and that no part of it has been submitted for a degree qualification at any other university or institute. I further declare that I have appropriately acknowledged all sources used and have cited them in the references section.

Name: Date:

Signature:

Dedication

For my father’s soul, for my wonderful mother and finally for my dear daughter **Mariam**

Acknowledgments

First of all,thanks to Allah the most merciful for guiding me through and giving me the strength to complete this work the way it is.

I would like to express my deepest gratitude to my supervisors Assoc. Prof. Dr. Amr A.Sharawi and Assist. Prof. Dr. Nader A.Rahman not only for providing the opportunity for this research but also for their directive supervision, continuous support, advice and encouragement.

I would to like to thank my sister and my brothers their support and advice during my research. Special thanks to my brother Dr.Hosam Adel

Lastly thanks go to “Center Advanced Software & Biomedical Engineering Consultations” staff for their support and their great effort especially Prof.Dr. Ahmed Hisham

Table of Contents

DISCLAIMER ..... **I**

DEDICATION ..... **II**

ACKNOWLEDGMENTS ..... **III**

LIST OF FIGURES ..... **VI**

ABSTRACT ..... **VII**

CHAPTER 1: INTRODUCTION **1**

1.1 PROBLEM DEFINITION 1
1.2 OBJECTIVE OF THE THESIS 1
1.3 SCREENING FOR BREAST CANCER 1
1.4 TECHNIQUES USED FOR BREAST CANCER DETECTION. 2
1.4.1 Breast Exam. 2
The doctor will examine both breasts of the patient. And lymph nodes which is in armpit to check for any lumps or any abnormalities. 2
1.4.2 Breast MRI (Magnetic Resonance Imaging). 2
1.4.3 CT (CAT) Scans (Computerized Tomography). 3
1.4.4 Ultrasound 4
1.4.5 Mammograms 4
1.4.6 Thermography 5
1.4.7 Removing a sample of breast cells for testing (biopsy). 6
1.5 ORGANIZATION OF THESIS 6

CHAPTER 2: GENERAL BACKGROUND **7**

2.1 DEFINITIONS OF THERMOGRAPHY 7
2.2 MEDICAL HISTORY OF THERMOGRAPHY 8
2.3 PHYSICS BEHIND THERMAL IMAGING 8
2.4 IMAGE PROCESSING TECHNIQUES AND MACHINE LEARNING 10
2.4.1 Segmentation: 10
2.5 ACTIVE CONTOURS 11
2.5.1 History 12
2.5.2 Parametric models (snakes) 12
2.6 COLOR MODEL 13
2.6.1 HSV color model 13
2.7 FEATURE EXTRACTION 14
2.7.1 Statistical features 15
2.7.2 Texture features 15
2.8 FEATURE SELECTION 16
2.8.1 Feature Ranking 17

2.8.2 Feature subset selection....17
2.9 CLASSIFIER....18
2.9.1 Sequential Minimal Optimization (SMO)....18
2.9.2 Multilayer perceptron....18
2.10 KERNEL LOGISTIC REGRESSION....19
2.11 RELATED WORK....19
CHAPTER 3: MATERIALS AND METHODS....22
3.1 DATABASE....22
3.2 SEGMENTATION....24
3.3 FEATURE EXTRACTION....25
3.4 FEATURE SELECTION....26
3.4.1 Greedy Hill Climbing Search....27
3.4.2 Best First Search....27
3.5 CLASSIFIERS....27
3.5.1 Artificial neural network....27
3.5.2 Sequential Minimal Optimization (SMO)....28
3.5.3 Kernel logistic regression (KLR)....30
3.6 K-FOLD CROSS-VALIDATION....30
CHAPTER 4: RESULTS AND DISCUSSIONS....31
4.1 PERFORMANCE EVALUATION.....31
4.1.1 Classifiers Evaluation for HSV, RGB and gray Image.....32
4.1.2 Classifiers Evaluation before Feature Selection....32
4.1.2 Feature selection results....33
4.2.3 Evaluation of classifier after feature selection....36
4.2.4 Evaluation of the classifier using common features....38
DISCUSSION....39
CHAPTER 5: CONCLUSION AND FUTURE WORK....41
5.1 CONCLUSION....41
5.2 FUTURE WORK....41
ملخصالرسالة....أ

List of Figures

FIGURE 3. 1 BREAST IMAGE IN FRONT POSITION 23
FIGURE 3. 2 BREAST IMAGE IN THE RIGHT SIDE AT 45° 23
FIGURE 3. 3 BREAST IMAGE IN THE RIGHT SIDE AT 90° 23
FIGURE 3. 4 BREAST IMAGE IN THE LEFT SIDE AT 45° 23
FIGURE 3. 5 BREAST IMAGE IN THE LEFT SIDE AT 90° 23
FIGURE 3. 6: FLOW CHART SHOWING THE COMPONENTS OF THE PROPOSED PATTERN RECOGNITION SYSTEM TO CLASSIFY BREAST IMAGES. 24
FIGURE 3. 7 RIGHT RGB IMAGE 25
FIGURE 3. 8: LEFT BREAST MASK 26
FIGURE 3. 9:RIGHT BREAST MASK 26
FIGURE 3. 10: SEGMENTED LEFT BREAST 26
FIGURE 3. 11: SEGMENTED RIGHT BREAST 26
FIGURE 3. 12: HSV OF THE RIGHT BREAST 26
FIGURE 3. 13: HSV OF THE LEFT BREAST 26
FIGURE 3. 14: BASIC NEURAL NETWORK ARCHITECTURE 28
FIGURE 3. 15: LINEAR SVM 29

FIGURE 4. 1:SHOW THE ACCURACY AT DIFFERENT VALUE OF C 33
FIGURE 4. 2: NUMBERS OF COMMON FEATURES 38

Abstract

Breast cancer is the main reason for women's death. It is the second most frequent cause of cancer death for women after lung cancer. Worldwide. In Egypt, breast cancer is considered the most-popular cancer among females. In 2008 the proportion of deaths due to breast cancer was 37.7% among 12,621 new cases. When breast cells begin to develop out of control it can cause breast cancer. However, there are many techniques that help you discover cancer more safely. One of these is digital infrared imaging (thermography); it is relying on metabolic activation and vascular circulation in both pre-cancerous and cancerous tissue. Radiographic images obtained from thermography equipment are one of the most f techniques that used for helping in early detection of breast cancer. a method for detecting breast cancer using image processing techniques was proposed by this thesis. These techniques are applied to 142 digital thermal images of the breast; 77 of these are normal images and 65 are abnormal. Matlab is used to detect the region of interest (ROI) and feature extraction. Besides that, Weka is used for the classification and selection of functions. The active contour techniques are used in the extraction phase (ROI) and therefore 72 statistical and textural features are extracted and used to supply classifiers. It is worth mentioning, there are three different classifiers. These classifiers are Sequential minimal optimization (SMO), kernel logistic regression, and Multilayer perception and their accuracy had reached to 99.29%, 98.59%, 96.4%. Afterwards it comes to the feature selection phase, at which the best dominant features are selected to minimize the processing time and effort. Accordingly, there are two types of wrapper feature selection methods are called best first and Greedy Step wise. These methods are used to minimize the number of features. Finally, only 6 features were selected, which provided us with 99.29% accuracy using Minimum Sequential Optimization (SMO), 99.29% by using kernel logistic regression and 97.18 with Multilayer Perceptron.

Chapter 1: Introduction

1.1 Problem definition

Breast cancer is a major health problem in Western countries as well as our country. In the USA a study developed by the American Cancer Society estimated that among 8 to 12 women one will develop breast cancer during her life [1].Breast cancer is the 2nd mostly common diagnosed cancer world-wide [2].There are two main types of breast cancer namely cranio-caudal (CC) and media-lateral oblique (MLO).Breast cancer usually appears in the ducts, tubes that convey milk to the nipple, and lobules, organs that create milk. It happens in all men and women, although male breast cancer is an uncommon disease. A few works have reported that the development rate of a tumor is corresponding to its temperature [3].Screening test searches for cancer disease before an individual exhibits any symptom. It can help to discover cancer tumor at an early stage. Many masses are benign, implying that abnormal growth is mainly limited to a circumscribed, single and growing mass of cells [4].Some cancers are dangerous, which means that abnormal growth attacks the surrounding tissues that may metastasize or spread to remote areas of the body [5].The masses have the potential to lead to complications, while malignant tumors are an extremely severe cancer. Most breast cancers will have metastasized before reaching a tangible size. In this way, different quantities of imaging techniques are found for breast tumor recognition at the tissue level. These are arranged as Infrared Imaging and Mammography. In an early stage, Infrared imaging or thermography is one of the best breast cancer detection routines and is also useful for the recording of advanced stages of breast cancer.

1.2 Objective of the Thesis

This thesis suggests a technique for detecting breast cancer in thermal Images. The technique has four main parts. In the first part, the image processing techniques are used to prepare the image for feature and pattern Extraction processes using active contour technique. The second part is extracting some feature. In third part we select the best features that have the more effect on the system by using wrapper method. Finally in fourth part we used the selected features to feed classifiers to get better accuracy and to reduce the number of false positive cases in the diagnosis of breast cancer.
The approach of this research is based on detecting breast cancer from breast thermal images using image processing techniques.

1.3 Screening for Breast Cancer

In an uninfected population, the purpose of any disease screening is to reduce mortality from this disease [6]. It is aimed at reducing the risk of developing or dying from disease