A LOCAL DIRECTIONAL TERNARY PATTERN TEXTURE DESCRIPTOR FOR MAMMOGRAPHIC BREAST CANCER CLASSIFICATION

MARY WALOWE MWADULO

A Thesis Submitted in Partial Fulfillment of the Requirements of the Degree of Doctor of Philosophy in Information Technology of Masinde Muliro University of Science and Technology

October 2020

DECLARATION

This thesis is my original work prepared with no other than the indicated sources and support and has not been presented elsewhere for a degree or any other award.

………   ……….   …………12 OCT 2020………………….

Mary Walowe Mwadulo   Date

SIT/LH/0005/2015

CERTIFICATION

The undersigned certify that they have read and hereby recommend for acceptance of Masinde Muliro University of Science and Technology a thesis entitled "**A Local Directional Ternary Pattern Texture Descriptor for Mammographic Breast Cancer Classification.**"

…… …………………………………   …………12 OCT 2020 ………………….

Dr. Stephen Makau Mutua   Date
Department of Computer Science
School of Computing and Informatics
Meru University of Science and Technology

………………………………………   ………12 OCT 2020 ……………………….

Dr. Raphael Angulu
Department of Computer Science
School of Computing and Informatics
Masinde Muliro University of Science and Technology

ACKNOWLEDGMENT

To God Almighty for giving me a sanity mind as I worked through the project.

To my supervisors Dr. Raphael Angulu and Dr. Stephen Makau Mutua for their unlimited guidance and consistent counsel during the entire research project.

To my husband Eric Gitonga, children Prudence, Christian, and Patience who supported and encouraged me during the entire research project.

I feel indebted to all who supported me directly and indirectly to complete this project. May God bless you abundantly.

DEDICATION

This thesis is dedicated to my loving husband Eric, charming daughters Prudence and Patience, and ever jolly son Christian.

ABSTRACT

Breast cancer is a top killer illness for women globally, but early and effective screening can increase their survival rate. Mammography is the tool used by a radiologist to screen for breast cancer, however, a radiologist is susceptible to human observer variability, and therefore, reading and interpretation of mammography test results depend on the expertise of the radiologist administering the test. To improve the reading and interpretation accuracy of the test, researchers' developed computer-aided extraction descriptors that extract discriminant features. These descriptors include the Local Binary Patterns (LBP), the Local Ternary Patterns (LTP), and the Local Directional Patterns (LDP), however, they have not yet yielded satisfactory results in differentiating breast cancer tumor types. The LBP descriptor is inadequately dependable in capturing breast cancer discriminant features because it is easily affected by noise. The LTP descriptor uses a fixed threshold value for all images in a dataset, making it not invariant to pixel value transformation. It is also not practically easy to select an optimum threshold value in real application domains. The LDP descriptor relies on top k significant directional responses and ignores the remaining $8\text{-}k$ directional responses. Disregarding the remaining directional responses reduces the computation efficiency since each pixel in an image carries subtle information. Given the limitations identified among the mentioned local texture descriptors, developing an effective texture descriptor becomes a viable and challenging research problem. Therefore, this study seeks to develop an improved local texture descriptor that considers all directional responses and applies an adaptive threshold in encoding image gradient. The new Local Directional Ternary Pattern (LDTP) texture descriptor calculates the absolute difference between the value of the center pixel and the values of its local neighboring pixels for a 3x3 image region. To get edge responses in eight directions, the absolute differences are convolved with a kirsch mask, then the pixels are transformed into zeros and ones using mini-max normalization. We then passed the normalized values through a soft-max function to get the probability of an edge in a certain direction. Then, two threshold values are calculated and used to split the probability space into three parts for -1, 0, +1 bits to generate a ternary pattern. The resultant Local Directional Ternary Pattern (LDTP) code is then split into a positive and negative LDTP code. Histograms of negative and positive LDTP encoded images are fused to get texture features. We validated the LDTP texture descriptor on the Mammographic Image Analysis Society (MIAS) breast cancer dataset using Support Vector Machine (SVM) and Artificial Neural Network (ANN) classifiers for normal/abnormal and benign/malignant classes. When the LDTP texture descriptor was compared against LDP, LTP, and other existing texture descriptors, it showed robustness and reliability in encoding an image gradient. The highest classification accuracy was attained by the SVM classifier, with 97.32% and 93.93% for normal/abnormal and benign/malignant classes, respectively.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AD: Architectural Distortion

ANN: Artificial Neural Network

AUC: Area Under Curve

BA: Bilateral Asymmetry

BCDR: Breast Cancer Digital Repository

BI-RADS: Breast Imaging Reporting and Data System

BRCA: BReast CAncer

BSE: Breast Self-Examination

CAD: Computer-Aided Detection

CBE: Clinical Breast Examination

CLAHE: Contrast Limited Adaptive Histogram Equalization

DCIS: Ductal Carcinoma In Situ

DDSM: Digital Database for Screening Mammography

GLCM: Gray-Level Co-occurrence Matrix

IARC: International Agency for Research on Cancer

IRMA: Image Retrieval in Medical Applications

KNN: K-Nearest Neighbour

LBP: Local Binary Pattern

LDA: Linear Discriminant Analysis

LDTP: Local Directional Ternary Pattern

LOOCV: Leave One Out Cross-Validation

LPOCV: Leave P Out Cross-Validation

LTP: Local Ternary Pattern

MCCV: Monte Carlos Cross-Validation

MIAS: Mammographic Image Analysis Society

MLC: Machine Learning Classifier

MLO: Media-Lateral Oblique

MLP: Multi-Layer Perceptron

MR1: Magnetic Resonance Imaging

ROC: Receiver Operating Characteristics

ROI: Region Of Interest

SVM: Support Vector Machine

WHO: World Health Organization

DEFINITION OF OPERATIONAL TERMS

Classification: A data mining function that assigns labels intending to accurately predict the target class for each case in the data

Data Set: A collection of data whose content is a single database table

Feature Descriptor: An algorithm that takes an image as input and gives out a feature vector that encodes interesting information into a series of numbers and acts as a numerical fingerprint that can be used to differentiate one feature from another.

Feature Extraction: A stage in the CAD system used to extract discriminant features

Mass: A breast abnormality indicator which occupies some space, found in at least two different projections defined by a wide range of features that can indicate benign changes but can also be a part of malignant change

Microcalcification: Tiny flecks of calcium typically in the range of 0.lmm- 1.0mm found in the soft tissue of the breast that can serve as an early indicator of breast cancer.

CHAPTER ONE

INTRODUCTION

1.1 Background

Cancer is a large group of related illnesses caused by the unrestrained division of body cells which spread, crowd out normal cells, and develop into a tumor that is benign or malignant [1][2]. Benign tumors develop slowly and do not attack neighboring tissues or extend to other parts of the body; therefore, they are cancerous. However, malignant tumors are cancerous and since with time they spread to neighboring tissues such as the lymph nodes which can cause the failure of major organs. Cancer can begin anywhere in the human body; therefore, the name of cancer depends on the affected area. Among cancer types, breast cancer is a top killer for women [3]. Although breast cancer can also develop in men, however, the highest risk and incidence is in women above 50 years of age [4][5] [6][7][8].

Breast cancer occurs because of the abnormal development of breast cells where the cells divide faster than healthy cells, accumulate, and form a mass. This cancer can either begin within the lobules glands which produce milk, or the milk ducts, which are used to transport milk to the nipple. Though unlikely, breast cancer can begin within the stromal tissues, that embody the fatty and fibrous connective tissues of the breast[9].

Breast cancer is the second most common killer for women, [10]. Bray and Soerjomataram [11] reported about 2,088,849 (11.6%) new breast cancer cases and 626,679 (6.6%) breast

cancer deaths in [11]. The survival rates of breast cancer patients vary in the world because it depends on factors such as age, geographical factors, and race, however, a relative survival estimate is 91% at 5 years diagnosis, 86% after 10 years, and 80% after 15 years [12]. A report by International Agency for Research on Cancer (IARC) showed that there are more deaths in less developed regions than the developed regions because a shift in lifestyle is causing an increase in incidence, and also because clinical advances to combat the disease are expensive and sometimes unavailable. Because of several contributing factors, there has been a general increase in breast cancer cases in recent years. A study conducted in 2017 on cancer incidences from 2005 to 2015 showed an increase in breast cancer cases by 33%. Out of which 12.6% of the incidences were because of population growth, 16.4% because of an aging population, and 4.1% was because of increasing age-specific incident rates [13]. According to the Globocan report of 2018, breast cancer caused 74072 deaths and 168690 incidences in Africa. The age-standardized incidence rate stood at 37.9 per 100,000 people, varying from 6.9 per 100,000 people in Gambia to 69.6 per 100,000 people in Mauritius. The age-standardized mortality rate stood at 17.2 per 100,000 people, varying from 4 per 100,000 persons in Gambia to 29.1 per 100,000 persons in Somalia.

Globocan 2018 statistics estimated breast cancer incidence rate at 40.3 per 100,000 persons with a mortality rate of 17.8 per 100000 persons. According to Bray and Soerjomataram [11], the annual incidence of breast cancer in Kenya is about 5985 (12.5% of all new cancer cases) and the annual mortality is about 2553 (7.7% of all cancer

deaths)[11][11][11]. Gakunga et al. [14] predicted that by 2025 the annual incidence of breast cancer in Kenya will increase to 8052 persons and there will be annual mortality of 3448 persons which is an increase of 35% for both incidence and mortality rates[14][14][14]. Korir et al. [15] researched cancer incidences among Nairobi women between 2004 to 2008 and showed that breast cancer age-standardized incidence rate stood at 51.7 per 100,000. The incidence rate is among the highest recorded in the African registry. Also, while in developed countries such as the USA, the mean age at breast cancer diagnosis is 64.1 years, in Kenya the mean age at diagnosis is at 51.9 years [16].

The alarming trend in the prevalence of breast cancer in Kenya underscores the need to develop an evidence-based intervention that can handle this volatile epidemic. The growing demand for breast cancer treatments with the nation's very limited supply capacity for diagnosis and treatment poses serious health-care policy challenges to the Kenyan government [17]. The problem is that on the demand side, the number of patients has been growing and is expected to continue, however, on the supply side, the facilities, equipment, and experts available appear not to be growing at a commensurate rate[18].

Physical and imaging tests are used to screen for breast cancer tumors. The physical examinations comprise; Breast Self-Examination (BSE) [19] and Clinical Breast Examination (CBE) [20] while the imaging tests comprise; Ultrasound, Magnetic Resonance Imaging (MRI), and Mammography. Unlike physical examination tests, imaging tests are more reliable since they allow a radiologist to have an internal view of

3

the breast. Among the imaging tests, Mammography is an effective screening test since it can recognize the breast cancer cells in their early stages before physical signs develop [21]. However, reading and interpretation of a mammography test are performed by a radiologist who is susceptible to human observer variability. Therefore, the results depend on the expertise of the radiologist administering and reading the mammography test. As a measure towards achieving a more accurate reading and interpretation of mammography test results, some radiologists resulted to double reading of mammogram test results. Nevertheless, this is not a viable solution since it is economically costly and time intensive[22][23]. Moreover, dust particles on a mammogram image and breast surgery scars on a patient can obstruct a radiologist which can result to false interpretation even if double reading was conducted [24].

Therefore, the adoption of a breast cancer Computer-Aided Detection (CAD) system that reads a mammogram image, extracts significant features and predicts a breast cancer tumor type independent of a radiologist is a viable solution. The ability of a CAD system to preprocess a mammogram image can aid in improving image quality by eliminating noise and artifacts found in mammogram images.

Researchers have developed several local texture descriptors for extracting discriminant features. The Local Binary Patterns (LBP) descriptor thresholds an image pixel based on the value of the central pixel and encodes the image gradient to a binary bit. It is computationally simple, enables image analysis in real-time, and can withstand monotonic

gray-scale changes. However, it is inadequately reliable in capturing breast cancer discriminant features because it is sensitive to noise. The Local Ternary Patterns (LTP) descriptor is better at handling noise than LBP because it uses three bits to encode an image gradient. However, it uses a fixed threshold defined by the user for all datasets or all images in a dataset, therefore it does not make it dynamically suitable for all images in a dataset. The threshold should be dynamically selected. Disregarding the remaining directional responses makes it miss discriminant texture features which reduce the computation efficiency.

1.2 Motivation for Study

The current advance in modern medical technologies and the evolvement of diseases has led to an increased amount of imaging data requiring analysis and a need to improve disease treatment. One such disease is breast cancer, which is rated as the highest killer among women. In 2018, Bray and Soerjomataram [11] reported an estimated 2,088,849 new breast cancer incidences and a mortality rate of 6.6%. A research conducted among Nairobi women between 2004 to 2008 showed that breast cancer accounted for 0.23 of all incidences, which is the highest rate so far recorded in the Africa cancer registry [15]. Survival rates for women with breast cancer vary in the world because of differences in the age bracket at risk, geographical factors, and race. There is a relative survival estimate of 91% when diagnosed within 5 years of the occurrence, 86% when identified after 10 years of the occurrence, and 80% when recognized after 15 years of occurrence [12]. Early breast cancer screening tools and strategies of identifying risk factors conditioning the

5

occurrence of breast cancer have been put in place, even though there are no efficient methods to prevent it because its causes are unidentified.

Existing local texture descriptors are effective in pattern recognition tasks because they can withstand position or light variations than global descriptors. However, their performance in breast cancer classification has not yet yielded satisfactory results. Sensitivity to noise is the principal limitation of the existing local texture descriptors, therefore, making them unreliable in capturing breast cancer discriminant features and they also disregard some useful information leading to loss of discriminative texture features which reduces the computation effectiveness. The need to uncover an effective way of extracting significant features in identifying breast cancer tumor and contribute positively towards providing better health care to humanity forms the foundation for the work presented.

1.3 Statement of the Problem

 A mammography test can recognize a breast cancer tumor in its early stages before physical indicators develop, and thus it is the recommended imaging test performed by a radiologist for breast cancer screening. The reading and interpretation of a mammography test results are therefore dependent on the skills and experience of a radiologist. However, a radiologist is susceptible to human observer variability and therefore can make a false reading and interpretation. To improve the reading and interpretation accuracy of a mammography test, some radiologists resorted to double reading of a mammography test result which is time-consuming, economically costly and increases recall rate [25], [26].

The current trend is to use a breast cancer CAD system that extracts significant breast cancer features, which are then used to classify the breast cancer tumor. Techniques used to extract the local descriptors suggested in literature include; the Local Binary Patterns (LBP), the Local Ternary Patterns (LTP), and the Local Directional Patterns (LDP). These local descriptors have shown their effectiveness in pattern recognition tasks for face recognition, however, they have not yet yielded satisfactory results in breast cancer tumor classification. The Local Binary Patterns (LBP) produces binary codes sensitive to noise and in some circumstance, they miss the local structure, since they do not consider the effect of the central pixel [27][28], consequently, they are insufficiently reliable in capturing breast cancer discriminant features. The Local Ternary Patterns (LTP) uses a fixed threshold defined by the user for all datasets or all images in a dataset [29] which does not make it dynamically appropriate for all images in a dataset. Further, the fixed threshold makes it not invariant to grayscale transformation. The Local Directional Patterns (LDP) rely on top k significant directional responses and ignores the remaining 8-k directional responses[30]. Disregarding the remaining directional responses reduces the computation efficiency, since each pixel in an image carry subtle information.

Given the limitations identified among local texture descriptors, developing an effective local texture descriptor for breast cancer classification becomes a viable and challenging research problem. This study, therefore, seeks to develop a new local texture descriptor

that considers all directional responses and an adaptive threshold in encoding image gradient for breast cancer classification.

1.4 Research Objectives

1.4.1 General Objective

This study sought to develop a Local Directional Ternary Pattern texture descriptor that considers all directional responses and an adaptive threshold in encoding image gradient for breast cancer classification.

1.4.2 Specific Objectives

  i.   To analyze existing techniques on breast cancer detection

 ii.   To develop a Local Directional Ternary Pattern texture descriptor that considers all directional responses and an adaptive threshold in encoding image gradient.

iii.   To validate the Local Directional Ternary Pattern texture descriptor in breast cancer detection.

1.5 Research Questions

  i.   How do the existing techniques detect breast cancer?

 ii.   How can a Local Directional Ternary Pattern texture descriptor that considers all directional responses and an adaptive threshold when encoding image gradient be developed?

iii.   How valid is the developed Local Directional Ternary Pattern texture descriptor in breast cancer detection?

1.6 Significance of the Study

Breast cancer is amongst the most severe health problems in recent times. A World Health Organization (WHO) report of 2018 estimated a breast cancer mortality rate of 15% of all cancer types among women. According to a report by Globocan in 2018, Kenya accounted for 7.7% of the breast cancer mortality rate, making it the third leading cause of death.

The findings of this study benefit the research community through contribution to knowledge by uncovering a new local texture descriptor that considers all directional responses and an adaptive threshold in encoding image gradient. The application of the developed local texture descriptor highlights an improved way of identifying a breast cancer tumor. Also, it contributes to practice by developing an improved approach that aids a radiologist to make a more accurate interpretation of mammogram results by eliminating observer oversight, which reduces unnecessary biopsies, increases the survival rate of women, and provides better health care to humanity.

1.7 Scope of the Study

Although image tests like Ultrasound, MRI, and Mammography exist, this research focused on extracting features captured through mammography because mammography is the recommended imaging test for breast cancer screening, since it can identify a breast

cancer tumor in its early stages before physical symptoms appear. Additionally, even though global and local features exist in an image, this research focused on local texture features which have been proven to be more effective than global features in pattern recognition. Several breast cancer datasets exist, however, this research used the Mammographic Image Analysis Society (MIAS) dataset. The choice was anchored on a thorough literature review that revealed that the MIAS dataset is a reliable source of mammographic images. Similarly, because of the number of different permutations of classifiers possible, this study used two classifiers; SVM and ANN because they have shown high levels of classification accuracy.

1.8 Assumptions in the Study

This research presumed that significant and reliable features could be extracted by considering local texture features because they consider the internal structural properties of an image instead of the region and boundary features used by a radiologist. Also, the tools and material used in this research elicited reliable and valid responses.

1.9 Limitations of Study

The choice of machine learning classifiers was based on information got from the literature review and therefore may not have been an effective representative sample in terms of the contribution that each classifier would have made in the research. However, to ensure the sample was representative, the choice of the machine learning classifiers was from different categories of classifiers. Further, validation of the new local texture descriptor

was based on binary classification and the performance was evaluated based on the presence or absence of breast cancer tumors. Therefore, this research cannot be extended to multi-class classification, however, to provide a universal generalization, the results got in the study were compared against results got by other researchers. Finally, because of the complexities involved and time constraints in obtaining direct patient data, the researcher used a publicly available dataset. However, to ensure the sourced dataset was valid, complete, and comprehensive, appropriate cross-validation was done.

1.10 Thesis Contributions

This research made the following contributions:

- Provided a comprehensive literature review on breast cancer detection as outlined in Chapter two of this thesis.

- Developed a new local texture descriptor that considers all directional responses and an adaptive threshold in encoding image gradient as outlined in section 4.2 of this thesis

- Performed experimental validation of the new local texture descriptor using breast cancer data as outlined in section 4.3 of this thesis

- Showed robustness and effectiveness of the new local texture descriptor by evaluating its performance against state of art local descriptors and comparing it against existing local descriptor as outline in section 5.2 of this thesis

1.11 Thesis Outline

The structure of this thesis is summarized below:

Chapter 1 presented a background detail of breast cancer formation, recent statistics on breast cancer mortality and survival rate, breast cancer screening tests existing local texture descriptors with their limitations, the motivation of study, the statement of the problem, the main and specific objectives of the research, Significance, and Scope of the research, Assumptions, and limitations considered in the research, the contribution of thesis and thesis outline.

Chapter 2 analyzed the literature associated with the first objective. The chapter presented a discussion on risk factors conditioning occurrence of breast cancer cells, screening tests for identifying breast cancer, mammographic indicators of breast cancer cells, breast cancer modeling approaches, breast cancer CAD system, feature extraction descriptors, feature selection methods, classification techniques, breast cancer datasets, and breast cancer evaluation protocols. Also, a conceptual and theoretical framework of the research was provided.

Chapter 3 addressed the research methodology adopted. It began by demystifying the research philosophy adopted and justified the choices taken during the entire research process using Saunders's research onion. It also explained in detail the experimental set-up followed when developing and validating the LDTP texture descriptor. Further, this

chapter presented the research questions, research tools, and material, and a brief discussion on ethical considerations mirrored in the research.

Chapter 4 explained the development of the LDTP texture descriptor in encoding an image gradient. The chapter explained the process of encoding an image gradient in two steps; Computing the edge responses for all the eight directions as explained in section 4.2.1 and calculating an adaptive threshold as explained in section 4.2.2 of this thesis.

Chapter 5 elucidated and discussed the experimental validation using the breast cancer dataset. The experimental validation was presented based on the experimental setup defined in section 3.4. In detail, the chapter outlined the reading of mammogram images into the MATLAB environment, mammogram image preprocessing procedure, the process of data augmentation, extraction using the developed LDTP texture descriptor, classification process using SVM and ANN classifiers, and validation of each of the classifiers into Normal/abnormal class and benign/malignant class using the breast cancer dataset. The chapter also explained the experimental results achieved by the developed LDTP descriptor based on breast cancer data. An analysis of results on the performance of LDTP descriptor for normal/abnormal and benign/malignant classes, accuracy levels comparisons of LDP, LTP, and LDTP descriptors using ANN and SVM classifiers, specificity and sensitivity comparison of SVM and ANN classifier for LDTP descriptor, the statistical significance of accuracy level achieved by LDTP descriptor and Accuracy level comparison of LDTP descriptor against existing local descriptors was presented.

13

Chapter 6 concluded by highlighting the research findings, the achievement of each research objective, the contributions made by this research concerning knowledge and practice, and insights for future improvement.

CHAPTER TWO

LITERATURE REVIEW

2.1 Overview

This chapter analyzes literature on principal areas associated with objective one in the study of breast cancer detection. The key areas discussed include risk factors conditioning occurrence of breast cancer cells, screening tests for identifying breast cancer, mammographic indicators for depicting breast cancer, image representation for modeling breast cancer tumors, various techniques for preprocessing and segmenting mammographic breast images, several techniques for extracting, selecting and classifying breast cancer features, breast cancer evaluation protocols, and breast cancer datasets. It also presented a review on related studies for breast cancer detection, conceptual and theoretical framework.

2.2 Risk Factors Conditioning the Occurrence of Breast Cancer Cells

A breast cancer risk factor helps to identify women who are susceptible to the disease and they need to undergo further analysis. However, the presence of a risk factor in an individual does not necessarily ascertain that the cancer is inevitable. Breast carcinoma can also develop in females who have no detectible risk factors. Studies conducted to determine factors conditioning the occurrence of breast cancer cells reported that the disease is not attributed to one factor, but a combination of several factors.

Risk factors that condition the occurrence of breast cancer cells are broadly grouped into extrinsic and intrinsic factors. Extrinsic factors are caused by influences outside the human body such as environmental conditions and an individual's lifestyle, whereas intrinsic factors are caused by influences inside the human body that occur naturally due to genetic makeup and internal structures.

2.2.1 Intrinsic Risk Factors

A major intrinsic risk factor is being female. Although breast cancer cells can also develop among the male gender, females have a higher risk and incidence than males. The likelihood of a female human species getting breast cancer is estimated to be 100 times more than in males [31] [32] [33]. This is because the female human species have more breast tissues than males. Furthermore, the production of the female hormone estrogen in the life of a female promotes the development of breast cancer [33].

Aging is another intrinsic risk factor that is inevitable and contributes to 90% of breast cancer occurring among older women. As a woman grows older, the risk of developing breast cancer cells intensifies and consequently, more breast cancer incidences occur among middle-aged and elderly women than in young women. Studies by [4] [5] [6][7][8] showed that women between 50-64 years of age are at the highest risk. In 2016 Susan Komen [34] analyzed the chances of developing breast cancer by considering age as a major factor. She reported that at age 20 the likelihoods are 1 in 1674 women, at age 30 the likelihoods are 1 in 225 women, at age 40 the probabilities are 1 in 69 women, at age 50

16

the probabilities are 1 in 44 women, at age 60 the likelihoods are 1 in 29 women and at age 70 the probabilities are 1 in 26 women.

Family history is an intrinsic risk factor than contributes to 5-10% chances of breast cancer occurring. The risk is more among women who have close female blood relatives like a mother, a daughter, or a sister who has the illness [33]. Furthermore, the risk increased if the female relative developed cancer before attaining 50 years or developed cancer in the two breasts [33].

Even though 15 in every 100 women with breast cancer acquired it through family history, other times it is a faulty gene linked to breast cancer the was inherited[35]. The human body contains Breast Cancer 1 (BRCA 1) and Breast Cancer 2 (BRCA 2) genes that protect it against cancer. They correct any damaged Deoxyribonucleic Acid (DNA) during cell division. However, when the genes mutate and a faulty version is inherited, then the possibility of developing the breast cancer cell is triggered because the altered gene is incapable of repairing the damaged cells which later build up to form a tumor. Malone et al [36] reported that women with a faulty BRCA1 or BRCA 2 gene have a 60% to 80 % chance of developing breast cancer cells in their life.

A woman's reproductive history is also a contributing risk factor for the occurrence of breast cancer. A woman's long exposure to hormones, because of early menstrual [37] or late menopauses caused by estrogen produced in the body, raises the probability of having

17

breast cancer. In a year, those women who begin menstrual periods earlier than average have a 5% increased probability of acquiring breast cancer, whereas women who begin menopause late have a 2.8% increased probability of acquiring the disease [38]. Having children reduces a woman's exposure to her estrogen which lowers the chances of developing the breast cancer cell by 7% [39]. Breastfeeding decreases the chances of developing breast cancer cells by 4.3% for every 12 months a woman breastfeeds [39] [40].

Another intrinsic factor is high breast density. A woman with a lot of dense breast tissues is at a higher risk of developing breast cancer cells than one without [41] [42] [43]. Dense breasts have more connective tissues than normal breasts which appear white. Tumors and other abnormalities also appear white, making it challenging for a radiologist to spot a tumor on a mammogram consequently leading to false interpretation. Furthermore, as the density of the breast tissues increases, the sensitivity of a mammogram reduces [44]. In a study meant to assess the mammographic density and risks of developing breast cancer among women with a family history of breast cancer cases, Duffy et al. [45] reported a 3% increased risk per 10cm^3 of dense tissue.

The race is also another intrinsic risk factor conditioning the occurrence of the breast cancer cell. A study by ban et al. [46] showed a notable difference in occurrences of breast cancer cells and mortality rates amongst women in distinct races. The study showed that the black population had the highest mortality rate of 18.2% against 12.3% and 14.8%

among the white and Hispanic populations, respectively. Even though the slight difference in mortality could be because of differential access to health care, genetics, and socioeconomic factors for both early diagnosis and treatment, the increase in mortality among the black population is because of the high progression level of the disease and higher risk of triple-negative phenotype linked to poor prognosis of the disease[47]. In another study by Iqbal et al. [48] they attributed the high mortality rate among the black population to delayed treatment, misuse of treatment, and underuse of treatment.

2.2.2 Extrinsic Risk Factors

An extrinsic risk factor for the occurrence of the breast cancer cell is dietary habits. Consuming foods rich in fat or enhanced with food preservatives and flavors promote neoplastic transformation in the mammary glands, which triggers excessive weight gain that could lead to obesity [47]. Excessive weight gain or obesity, especially among postmenopausal women, which is linked to an increased chance of breast cancer cells occurring. Saxe et al. [49] confirmed that consuming a low-fat diet especially among women who have reached menopause significantly decreases the chances of neoplastic transformation of the mammary glands and therefore reduces the chances of breast cancer cells developing.

Not being engaged in physical exercise is another extrinsic risk factor conditioning the occurrence of the breast cancer cell. Researchers have associated regular physical exercise either moderately or intensively to low breast cancer risk. Physical exercises control blood

sugars and reduce levels of estrogen, which affects how breast cells grow. A case-controlled study by Kobayashi et al. [50] showed that women who engaged in moderate to vigorous physical activities reduce their chances of developing the breast cancer cell by 30%. Another study by Razvi et al. [51] showed that active females have a breast cancer reduction rate of between 15-20%. However, the effect differs across the life course and menopausal status [50]. Another study showed a 20-40% breast cancer risk reduction when physical activities are done 3-5 times every week [52].

Another extrinsic risk factor conditioning the development of breast cancer cells is using Hormone Replacement Therapy. Women who consume HRP drugs for over five years during menopause, not only raise their likelihood of acquiring breast cancer but also risk the cancer being discovered in late stages [53]. Also, consuming birth control pills is associated with a higher probability of developing breast cancer cells. Studies have also shown that a woman's chances for developing breast cancer cells increase with high consumption of alcohol. Other factors like excessive smoking, exposure to cancer-causing chemicals, and hormonal changes because of working in night shifts also may raise the chances of developing the breast cancer cell.

2.3 Breast Cancer Screening Tests

Globally, breast cancer is the second most common killer for women [10]. In developing countries, the death rate is higher than in developed countries because of a limited capacity for diagnosis and treatment. However, the long-term survival rate for women can be improved through effective screening during the early stages of breast cancer development

[19]. Breast cancer screening is performed through physical examination tests or imaging tests. Physical examination tests include Breast Self-Examination (BSE) and Clinical Breast Examination (CBE) while imaging tests include; Breast Ultrasound, Magnetic Resonance Imaging, and Mammography.

Breast Self-Examination (BSE) is a cost-effective test since it only requires a woman to physically inspect her breasts regularly for any abnormal lumps using her fingers. This test is recommended for women who are above 20 years of age. However, its effectiveness in reducing the number of breast cancer death is contentious because being healthy does not compel a woman to perform the test, and evidence from clinical trials is limited [19] [54]. some reasons that most women do not practice Breast Self-Examination include; attitude, lack of awareness, lack of time to perform the test, lack of self-assurance in performing the procedure properly, distress experienced in discovering a tumor, and awkwardness related with breast inspection [19] [55][56].

Clinical Breast Examination (CBE) test is recommended for the early identification of palpable breast tumors [20]. The test is suitable for women 40 years and younger [57]. It involves a thorough physical inspection of the breasts by a clinician. During the procedure, the clinician makes a visual check-up for palpation of the breast and inspects auxiliary lymph nodes [20], [58]. The CBE test aims to detect irregularities in the breasts that allow for further evaluation. Even though the Clinical Breast Examination (CBE) test is better at identifying breast abnormalities than Self Breast Examination (SBE), on its own, it cannot

accurately differentiate between a benign tumor and a malignant tumor [57]. Also, understanding the visual and palpable observations of CBE is difficult [57] because of the different combinations of a patient's physical features like age, parity, menopausal status, breast tissue density, and modularity which can affect the interpretation of the observation made [59]. Unlike the physical examination tests, imaging tests allow a radiologist to have an internal view of the breast, which makes the results got more certain and reliable. Breast imaging tests include; Ultrasound, MRI, and Mammography.

An ultrasound test uses sound waves to generate a picture of the internal structure of the breast [60]. It helps determine if a breast tumor is a mass or cyst. It is primarily used to complement other imaging tests such as mammography and also guides on biopsy of suspicious lesions. Even though the use of Ultrasound is helpful when used in women who have dense breasts and are at a high risk of developing the disease, it has a low contrast for soft tissues than mammography and contrast-enhanced MRI. Therefore, it becomes challenging when differentiating between a benign and a malignant tumor because of overlapping characteristics [61]. Also, it is highly operator dependent, which means it requires real-time adjustments of focal zones, pressure, and patient positioning [62].

Breast Magnetic Resonance Imaging (MRI) creates a detailed picture of the internal structure of the breast by using a magnetic field and radiofrequency pulses. It is useful for breast lumps that cannot be seen by mammography or ultrasound, especially for women with high breast density [60] or in a patient who has had a biopsy and the doctor wants to see the extent of the disease. Unlike mammography and Ultrasound, MRI can cover a large

22

portion of the breast and is unlikely to cause an allergic reaction experienced by ionized based substances in X-rays and CT scans [61]. However, MRI is relatively expensive, time-consuming, the procedure is uncomfortable for the patient, it produces blurred images and cannot be performed on patients who have metal implants because it uses magnets [62]. Besides, it has more false-positive results compared to mammography because of the inability to distinguish between a malignant and a benign tumor [61].

A mammography is a screening tool that uses a low dose of x-ray to assist a radiologist in examining the internal structure of the breast for any suspicious tumor in women without physical signs of breast cancer [63]. This screening involves exposing the breasts to a low dose of ionizing radiation to get a picture of the internal parts of the breasts. There are two types of mammography, Digital and Film mammography. Digital mammography is better than Film mammography at finding breast cancer cells among women who are premenopausal, below the age of 50 years, and have densely structured breasts [61]. However, there are no significant differences in the results obtained, when either Digital or film mammography is used for women not in any of the above-mentioned groups since its effectiveness is the same. Even though a mammography test is economically affordable and it is a procedure that takes less than five minutes, the procedure is, however, uncomfortable for the patient because it causes compression of the breast tissue, the effectiveness of a mammography test is greatly reduced when performed on women with densely populated tissues and it exposes a patient to x-ray ionizing radiation which may induce cancerous cells especially for women younger than 20 years of age [64]. However,

23

there is emerging evidence that when MRI and mammography tests are performed together in women who have a high probability of developing the disease than for women who have an average probability [61] the detection of breast cancer cells is increased significantly.

Several factors, including the age of the patient, elapsed time since they performed the last examination, the density of the breasts, and the skills of the radiologist interpreting the results can influence the sensitivity and specificity of screening tools [65]. An age-dependent analysis by Linda et al. [54] revealed that sensitivity for BSE ranges from 12% to 41% whereas CBE has a sensitivity of 40% to 69% and a specificity of 88% to 99% [20]. A major shortcoming of BSE and CBE as physical examination tests, of the breast, is that the sensitivity to detect breast cancer is low, which in most cases results in women undergoing biopsies.

Imaging tests give higher sensitivity values than physical examination tests, which minimizes the need for biopsies. The clinical guidelines released in 2015 showed that performing an MRI on high-risk women without cancer resulted in an age-dependent sensitivity and specificity ranging from 71% to 100% and 81% to 97% respectively [66]. Linda et al. [54] showed the age-dependent sensitivity of a mammography test is between 77% to 95% whereas specificity ranges from 94% to 97% while Pisano et al. [65] showed the sensitivity of mammography for dense breast tissues to be as low as 30% to 48%. Despite Ultrasound and MRI tests having higher sensitivity values than mammography in non-fatty breasts, both Ultrasound and MRI have a higher probability of overestimating the

extent of the tumor [67]. Therefore, a lower MRI specificity compared to mammography might be associated with high biopsies rate and high chances of over-diagnosis, especially when used in women with a low probability of developing the breast cancer cell. Additionally, the accuracy of mammography alone is higher than the accuracy of Ultrasound, MRI alone, or a combination of mammography with clinical-based examination tests [67].

A review of ten selected articles between 2001 and 2018 is shown in Table 2.1 to determine among mammography, Ultrasound, and MRI the imaging test with the highest sensitivity and/ or specificity for breast cancer diagnosis. We based the inclusion criteria of the articles on studies that compared mammography, Ultrasound, and MRI image testing modalities using performance measures of sensitivity and/ or specificity and showed the number of patients used in the study.

Table 2. 1: Sensitivity and specificity analysis of imaging tests

Author	No. of Patients	Mammography (%)		Ultrasound (%)		MRI (%)	
		Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Warner et al,2001 [31]	196	33	99.50	60	93	100	91
Warner et al, 2004 [32]	236	36	99.8	33	96	77	95
Sim et al, 2004 [33]	245	53.9	85.7	83.3	65.5	93.3	63.6
Kuhl et al,2005 [34]	529	32.6	96.8	39.5	90.5	90.7	97.2
Crowshaw et al, 2011[35]	61	81	48	90	33	86	79
Valente et al, 2012[36]	244	21	99.5	43.5	96.2	37.1	96.7
Shao et al 2013[37]	90	72.7	62.9	80.0	60.0	90.9	82.7
Huzarski et al, 2017[38]	2995	57	-	59	-	86	-
Hossam and Hurb 2018 [39]	50	67.8	88.2	98.4	94.0	99.0	97.0
Huang et al 2018 [40]	107	88.6	71.4	90.9	79.4	95.5	81.0

From the ten reviewed articles, performance comparison analysis of the imaging tests showed that in four out of the nine articles that considered specificity, mammography gave a higher specificity than Ultrasound and MRI. Also from the ten reviewed articles, it can be noted that with regards to sensitivity, MRI had the highest sensitivity in nine out of the ten articles. Table 2.1 shows that none of the imaging tests had a perfect sensitivity and specificity value in diagnosing breast cancer cells.

Besides considering the efficiency of the screening tool based on sensitivity and specificity scores, another fundamental consideration is how to ensure there are fewer errors when the radiologist is reading and interpreting the mammographic images. This can be achieved by adopting a Computer-Aided Detection (CAD) system for breast cancer diagnosis.

2.4 Mammographic Signs of Breast Cancer

Signs of breast cancer on a mammogram image are seen as clusters of microcalcification, presences of masses, Architectural Distortions (AD), and Bilateral Asymmetry (BA) of the breast. Mass and microcalcification are the most common signs used to indicate breast cancer because they are detected during the early stages of cancer development. While masses are commonly seen among patients diagnosed with invasive breast cancer, microcalcification is reported in a higher percentage among patients diagnosed with Ductal Carcinoma In situ (DCIS) breast cancer. A study conducted by Gadjos et al. [78] indicated that 95% of breast cancer present as masses was invasive cancer, while 68% of breast cancer present as microcalcification was associated with DCIS. A study conducted by

Venkatesan et al. [79] showed that Bilateral Asymmetry and Architectural distortion are rarely linked to the presence of breast cancer cells. In this study, out of 1552 breast cancer cases, 56% indicated presences of masses, 29% indicated presences of microcalcification, 12% indicated Bilateral Asymmetry and 4% indicated Architectural Distortion.

2.4.1 Microcalcification

Calcification is an accumulation of calcium in the breast which can either be macro calcification or microcalcification. Macro calcification is large calcium deposits considered noncancerous therefore, they are not linked to breast cancer and consequently, no special consideration is dedicated to them [80]. On the other hand, microcalcifications are clusters of calcium deposits that are small bright white dots of varying sizes and shapes in the breast tissue [81]. Benign microcalcifications have a regular shape and are found in isolation while malignant microcalcification has an irregular shape and is clustered [82]. Microcalcification is identified by its size, shape, number, and distribution [83] [84]. The more, bigger, and closely clustered they are, the higher the chance is for them to be identified. Microcalcification is a challenge to discover because they are not separated from the surrounding normal tissue. When malignant cells start to invade the tissues the microcalcification viewed on a mammogram will be seen as a light patch on normal tissue.
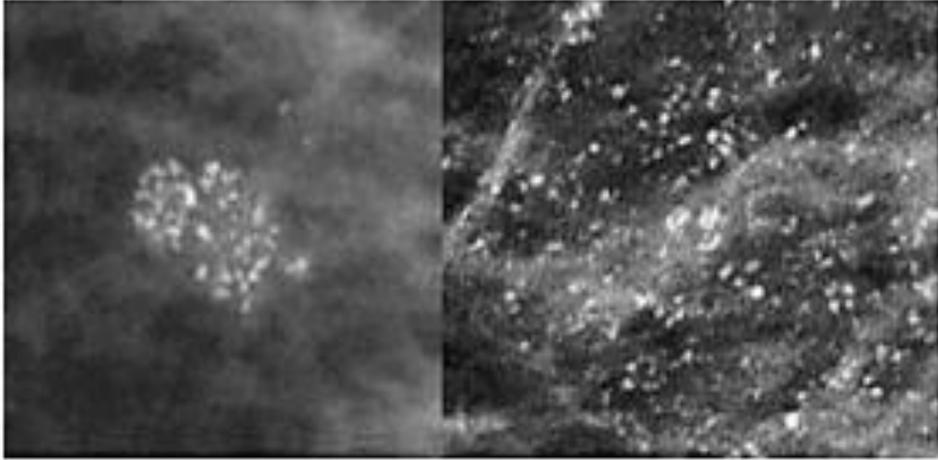
Figure 2. 1: Microcalcification in the breast (Source: [41])

Developing a model to detect microcalcification is easier because their presences are depicted by their numbers and how they are distributed, however their small size, presences of overlapping breast tissue, low contrast, and breast density especially in young women increase the probability that they can be missed or misinterpreted [83],[85]. Studies in [86] [87] [88][89] [90] developed wavelet transform-based methods, for detecting microcalcification. The main reason being that wavelet transform can specifically locate the image region with high spatial frequencies than transforms like Fourier than give the content of frequencies but cannot locate in the image the specific spatial frequencies [83]. Also, because microcalcification is seen as bright white dots on the mammogram, the wavelet transform methods can easily identify them as discontinued points. Studies in [91] [92] opted to detect microcalcification through shape features, though they achieved good results, they cannot be used alone and it is a challenge to detect microcalcification based on shape because of their small size. Other Studies in [93] [94] opted to use texture features, but even with good classifiers like SVM and ANN the results were not good enough.

2.4.2 Masses

A mass is a lump that varies in shape and size and can be seen in two projection; shape and margin properties [95]. A mass that is round, smooth, and has circumscribed margins has an increased possibility to be benign whereas a mass that is spiculated, rough, and has blurry margins has an increased likelihood to be malignant [95]. A majority of algorithms used for mass detection have two stages: (1) detecting suspicious regions (2) classifying suspicious regions.

2.4.2.1 Detecting suspicious regions

When a suspicious region is correctly identified, it is expected that the sensitivity to correctly classify the region during the classification stage is increased. However, sometimes it results in a high number of false positives. Algorithms applied for stage one detection are of two forms; Pixel-based detection methods and region-based detection methods [96].

Pixel-based detection is based on the extracted features of its local neighborhood, which is achieved by either defining a threshold value or using a sophisticated classification technique [96]. The suspicious pixels which have been detected are then grouped into regions by connected pixels. Regions labeled as suspicious by the algorithm do not automatically indicate malignancies. Categorization of the suspicious region into malignancy or benign is done during the classification stage.

30

Researchers have mainly focused on pixel-based methods. Liu et al. [97] designed a multiresolution system to detect spiculated lesions in mammograms based on a binary tree classifier. Experimental evaluation using the MIAS dataset showed the scheme achieved a low positive rate. Sampat and Bovik [98] presented a two-stage algorithm for detecting spiculated lesions in mammograms. The lesions were first enhanced then their location was detected. The algorithm was tested on the DDSM dataset and results achieved indicate that the algorithm could correctly locate the mass region. Zwiggelaar et al. [99] developed a model for detecting spiculated lesions based on local scale oriented signatures built using recursive median filtering and Principle Component Analysis. They achieved a sensitivity value of 80%. Even though a majority of researchers have focused on spiculated masses because of their high probability of showing malignancy, Other researchers such as [100],[101],[102],[103], and [104], focused on masses without considering a specific type of margin.

An advantage of pixel-based methods is accessibility to many pixels per image for training a classifier. However, the ability to use multiple pixels could pose a challenge. When multiple pixels are present on the margin and center of a mass, it becomes a challenge for the classifier to differentiate between these two regions, since two regions may be grouped to the same class, yet they may not always be homogenous. These pose a challenge of discriminating against a mass from normal surrounding tissue.

Region-based detection methods, use filtering or segmentation techniques to get the Region Of Interest (ROI). Shape and texture features are then extracted from each ROI and then they are classified as either suspicious or not. An advantage of region-based detection methods is taking into consideration spatial information which improves its discriminate power of a mass from its surrounding tissues. Also, the features extracted directly correlate with the shape and margins of the extracted regions. These ensure regions with similar characteristics can be categorized together. Further, in comparison to pixel-based methods, region-based are less computationally expensive. However, region-based detection methods have few samples for training if a classifier is to be used.

2.4.2.2 Classifying suspicious region into normal or abnormal tissue.

This stage classifies the suspicious region as normal/abnormal tissue and decreases the number of false positives generated in stage one. Olivera et al. [105] used the Support Vector Machine classifier to differentiate between normal and abnormal breast regions and attained an average accuracy of 98.88% with the DDSM dataset. To develop a system that emulates image features used by a radiologist Brake et al. [106] relied on intensity, density, location, and contrast features to get approximately 75% accuracy of all detected cancer. Wang et al. [107] relied on size, shape, contrast, homogeneity, and speculation features and got AUC of 0.786±0.026.

Even though studies in [108], [109], [110], [111], [112] established various approaches for identification of masses, detecting a mass is more challenging than detecting

microcalcification, because masses have poor image contrast [113]. Therefore, to increase the visibility of a mass, proper segmentation is important to separate the tumor from its background. Also, because benign and malignant tumors develop form one spot and spread outward, the shape of the tumor is relatively not specific. Therefore, differentiating the shape of a benign tumor from a malignant tumor is a challenging task.
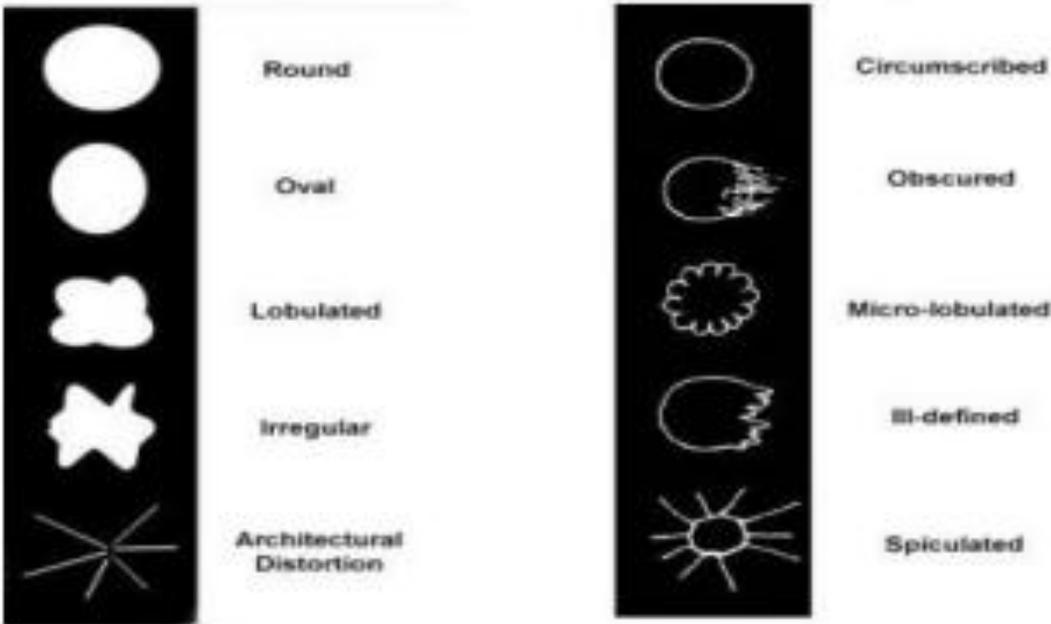
Figure 2. 2: Mass shapes and margins [42]

Most 80 -85% of breast cancer cells are identified as masses, clusters of microcalcification, or a combination of both [114]. However, 15-20% of breast cancer cells may not exist as a malignant mass or a malignant microcalcification. In such circumstances, focal nondescript abnormalities that include bilateral asymmetry and architectural distortion may be the only clue that a malignant tumor is present.

33

2.4.3 Architectural Distortion (AD)

Architectural Distortion (AD) occurs when the normal breast architecture deforms without the presence of a defined mass. The distortion often appears star-shaped. Even though (AD) are less predominate in indicating breast cancer than masses and microcalcification, they can predict breast cancer with high accuracy at the screening and diagnostic stage [79]. Approximately, 12% - 45% of missed breast cancer cases in mammogram screenings are because of the breast cancer cells manifesting themselves inform of breast Architectural Distortion [115]. Therefore, as a measure to reduce the number of missed breast cancer incidences, it is significantly important to accurately identify ADs. See Figure 2.3.
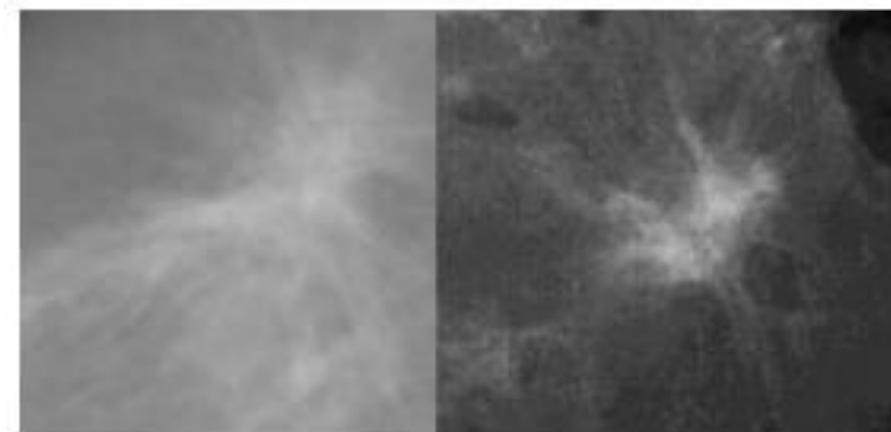
Figure 2. 3: Architectural distortion of the breast (Source: [41])

Even though a high percentage of DCIS manifest themselves as malignant appearing microcalcification, Architectural distortion accounts for 10.8% [116]. However, AD's is predominantly associated with invasive ductal and lobular breast cancer.

34

2.4.4 Bilateral Asymmetry (BA)

Bilateral Asymmetry is an indicator used by radiologists to identify the existence of breast cancer cells viewed on a mammogram. To detect Bilateral Asymmetry (BA), compare the right and left breasts for any inequality. Examining the breast asymmetry can provide insight into symptoms like parenchymal distortion and asymmetric points which cannot be evaluated by other methods [117]. Before the left and the right breasts of a patient are compared, flip one breast so that both breasts are on the same orientation. On a computer monitor, different colors are used to distinguish contrast regions on the left and regions on the right breast. The regions on the right breast that are different from the corresponding regions on the left are highlighted using green color, while regions on the left breast that are different from corresponding regions on the right are highlighted using a blue color [80].

Microcalcification, masses, AD's, and BA are all breast cancer signs viewed on mammography. However, their performance in terms of sensitivity, specificity, and accuracy varies. A research carried out by [118] revealed that mass detection has the highest sensitivity of 94.7% followed by microcalcification with 93.7%. The sensitivity was good in both microcalcification and masses but poor for Architectural distortion and Bilateral Asymmetry. Based on specificity values, Architectural Distortion got the highest score of 79.1% and Bilateral Asymmetry being the least with 52.4%.In terms of performance accuracy, masses had the highest with 84.8% followed closely by

35

microcalcification with 82.1%. Bilateral Asymmetry had the least accuracy of 67.4%. The outstanding performance of masses and microcalcification makes them the most popularly used mammographic signs of breast cancer detection. Even though many publications have focused on detecting and analyzing microcalcification and masses, very few researchers focused on detecting Architectural Distortion in mammograms. Broeders et al. [119] and Rangayyan et al. [120] suggested that a more effective breast cancer prognosis could be realized if more attention is geared towards the detection of AD's.

2.5 Breast Cancer Modelling

Locating and correctly interpreting a breast tumor is a viable though challenging task in computer vision because of the inherent variability of the tools used and breast structure. While tools used to capture the breast image are susceptible to noise and illumination, breast structure among individuals varies. While some individuals may have fatty breasts others have dense breasts. This section presents various approaches discussed in the literature for image representation when detecting breast cancer cells.

2.5.1 Shape Models

The shape is a significant visual cue used by radiologists to distinguish the two distinct classes of breast tumors. Benign tumors have circular and symmetric shapes while malignant tumors have random and asymmetric shapes [95]. To identify the shape of the tumor, the radiologist uses the shape interior region or contours defined by the tumor boundary. An effective shape model should successfully discover similar shapes in a

pattern matching problem, irrespective of whether the shape being recognized is rotated or flipped. Also, it should be robust such that it can effectively compare two images even though one may be noise affected or distorted.

When the interior region of shape is used, it takes all the pixels within the shape into consideration for shape representation. Further, the interior region can be considered as global since it considers the entire region or structural since it partitions the interior region into sub-parts for shape image representation [121]. Shape contours are defined by tumor boundaries. The boundary defines the edges, which define a transition between tumor and surrounding tissues. Understanding tumor boundary aid in knowing to what extent the cancer cells have spread, which brings out the size of the tumor. The shape contours can be represented globally in which the complete margin information is used to develop the feature vector or structural in which the shape boundary information is broken into sub-parts [121].

Shape features such as roundness, eccentricity, and compactness are used to characterize a tumor [122]. Caulkin et al. [123] defined shape vector $xshape = (x_i \dots x_n, y_i \dots y_n)$ by placing k equally spaced points on margins of each mass then defined the origin as the point of intersection between the border and the line connecting the nipple. This approach was based on the understanding that masses are formed in the breast ducts that originate from the nipple. However, this introduced rotational variances in the model, because the model did not consider interdependencies between size and feature vector. To improve the

37

model Berks et al. [124] defined a compact model of shape variation built by applying PCA to the aligned vectors Such that $xshape \approx x'shape + Pshape\ bshape$ where $Pshape$ is a matrix of principal modes and $bshape = p'shape(xshape(-x'shape))$is the vector of model parameters stored for each shape. This model had negligible variance discarded. An important characteristic of using shape to locate breast tumors is low computation complexity and robustness of shape features [121]. However, using shape features requires enough landmark points to reveal the complete shape of the tumor.

2.5.2 Texture Models

Texture models, mainly model a tumor by looking at structural properties which are not visible to the human eye. They have been used extensively in modeling tumors for breast cancer detection. They take into consideration the texture of a tumor which makes the model appropriate for breast cancer detection at all stages. Texture is a feature of homogeneity of images using the pixel for tonal variation, which has a certain scale, regularity, and directionality [125]. Texture highlights important details about the structural arrangement and environmental relationship of the object in an image and reveals important discriminatory characteristics related to variability patterns [126]. While shape models are inspired by the properties for which a radiologist looks for, appearance models are structured in a way that they can capture important properties not noticeable to the human eye [83]. Therefore, appearance models can capture properties that are valuable but not extracted visually. Also, they do not occur over a point, but over a region. Even though appearance models are used to analyze many images in natural and medical

science, extracting the features is challenged by changes that could result from illumination, position view, which cause a large variation in the image appearance of texture. Also, two different textures under different imaging conditions can appear to be similar [127]. Furthermore, to effectively use appearance models, there is a need to first process each region to remove background breast tissue [124].

2.5.3 Hybrid Models

Hybrid modeling seeks to improve on the limitations of shape and texture models by combining them to get a more robust model for breast cancer detection. Hybrid models aim to improve the detection of masses and microcalcification by increasing the rate of true positive and reducing the rate of false-positive which improves the CAD system [122]. The performance of hybrid models is expected to be higher than when using either shape or appearance models. However, care should be taken to ensure we combine no two or more similar models, and also the combination does not cancel the effects of each other. Models could be combined in parallel, in which individual outputs of the models are concatenated and used for final breast cancer detection, or hierarchical in which the output of the first model is fed as input for the second model. This approach takes advantage of the strength in each of the modeling techniques.

2.6 Breast Cancer CAD System

Breast Cancer Computer-Aided Detection is a clinically proven technology aimed at aiding a radiologist in identifying abnormal cells viewed on a mammographic image by decreasing their observational oversight [128],[83]. It serves as a second reader by improving a radiologist's ability to make an objective and more accurate interpretation of

breast mammographic images which reduces rates of false-negative [128]. The objective of a CAD system is not diagnosing, but to bring the attention of the radiologist to a specific area whose analysis might determine the need for further analysis [21].

Studies by [129], [130], and [131] showed that when a CAD system is used as an aid by the radiologist, it increases breast cancer detection accuracy level. Further, Dheeba et al. [63] noted that using a CAD system improves the rate of breast cancer detection by 14.8% more than when radiologist does the detection. The approach is cost-effective and the accuracy levels got are comparable to the double reading of mammograms. Because CAD systems have high sensitivity rates, they are suitable especially when observing abnormalities in dense breast tissues, which are common among young women. A breast cancer CAD system can detect anomalies and suspicious areas of the breast during image preprocessing, feature extraction, and classification. Figure 2.4 shows an overview of the CAD system for breast cancer detection. Below is a detailed explanation of each of the breast cancer CAD stages.

```
┌─────────────┐      ┌──────────────┐      ┌──────────────┐
│ Input Image │ ===> │ Image        │ ===> │ Segmentation │
│             │      │ Preprocessing│      │              │
└─────────────┘      └──────────────┘      └──────────────┘
                                                   ‖
                                                   v
                                           ┌──────────────┐
                                           │ Feature      │
                                           │ Extraction   and │
                                           └──────────────┘
                                                   ‖
                                                   v
                                           ┌──────────────┐
                                           │ Classification│
                                           └──────────────┘
                                                   ‖
                                                   v
                                           ┌──────────────┐
                                           │ Performance  │
                                           │ Evaluation   │
                                           └──────────────┘
```
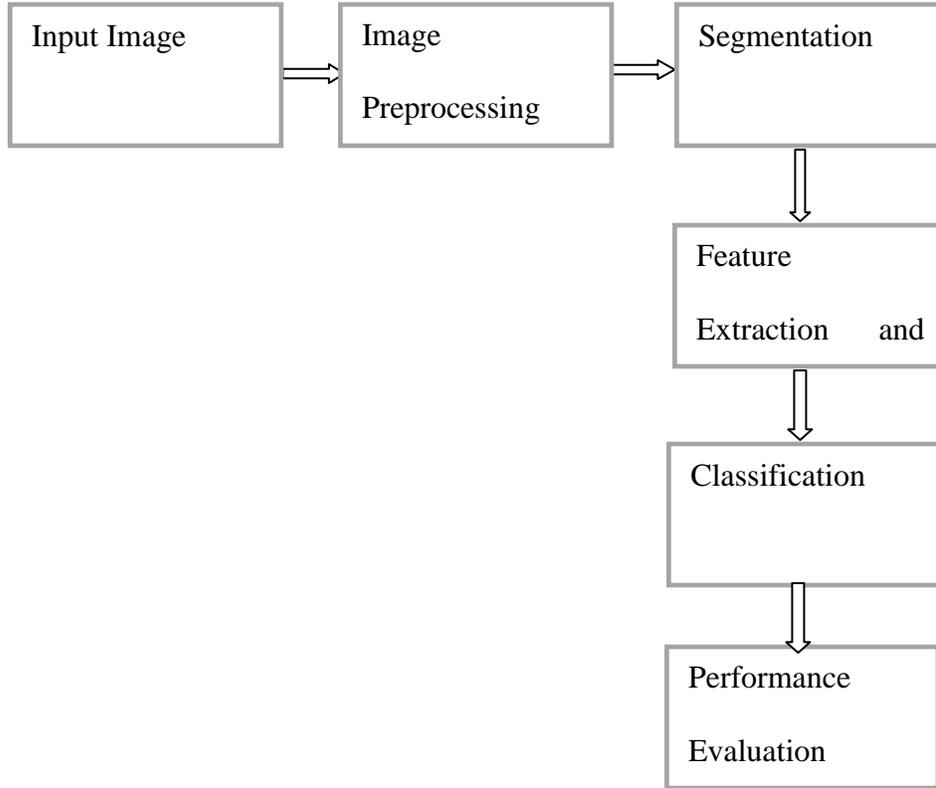
Figure 2. 4: Flow Diagram of Breast cancer CAD system [43]

2.7 Breast Cancer Preprocessing Techniques

Preprocessing is a foremost task in medical imaging. Its goal is to prepare an image for further analysis by enhancing image quality, eliminating noise and artifacts, without distorting image features [130]. Mammogram preprocessing is done in two different ways: De-noising and Intensity enhancement [132] [133]. The section below explains the most popular De-noising and Intensity enhancement techniques for mammograms.

41

2.7.1 Mammogram De-noising Techniques

Noise is unwanted signals that are seen as grain particles in an image and cause a change in image intensity [134]. Noise is caused by changes in light or thermal energy of heat inside an image. The presence of noise alters the image pixels, which changes intensity values resulting in an image that appears to have dots.

Salt and pepper, Speckle, Poisson, and Gaussian noises affect mammogram images. The Salt and pepper noise appears as white and black dots on a mammogram image and is caused by an abrupt change in image signal and accumulation of specks of dust in the image when the image is being captured. When a mammogram image has this noise, it substitutes the original image pixels values with noisy pixel values. Speckle noise is multiplicative, and it is created when dust particles settle on the image. Poisson noise is caused by the statistical behavior of electro-magnetic waves of an X-ray machine. Gaussian noise is caused by electronic circuits and sensor noise. To remove these types of noises, researchers have devices techniques for de-noising. Figure 2.5 illustrates the categorization of image de-noising techniques.
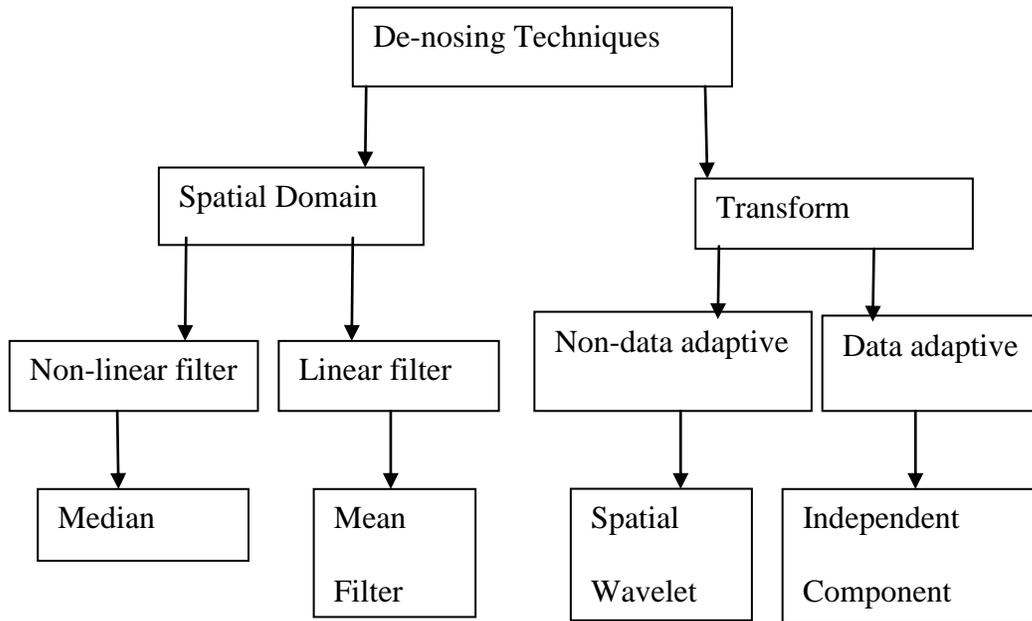
Figure 2. 5: Classification of image de-noising techniques (source [44])

De-noising techniques are classified as spatial or transform domain. The spatial domain aims at removing or isolating frequencies in the images. This technique considers the original noisy image and speed of the processing tool while the transform domain turns out a certain transformation of noisy image data and then applies the noise reduction process in the transformed image [136]. Table 2.2 shows the merits and demerits of different spatial and transform techniques.

Table 2. 2: Merits and Demerits of de-noising techniques

De-noising technique		Description	Examples	Merits	Demerits
Spatial Domain	Linear Filter	They are linearly constrained and generate input signals that vary based on time	Wiener filter Mean filter	Sharpen the edges of an image Correct unequal illumination Reduce salt and pepper noise	Blur shape images Destroy lines and image details Perform poorly in the presence of signal independent noise
	Non-linear Filter	They do not follow the linearity relationship	Median Filter Adaptive Filter	Reduce the noise without any attempt to explicitly identify it	They are complex to use design and implement than linear filters
Transform domain	Data adaptive	Characteristics of the image inside the filter	Independent Component Analysis (ICA)	it uses a non-Gaussian signal	Computationally costly
	Non-data adaptive	Coefficients are fixed and there is no difference in function between direct and transposed form.	Spatial Wavelet	can be applied to the signal without computing some statistics, Can be applied to feed-forward filters	Dependent on noise and signal characteristics which are unknown

2.7.2 Mammogram Intensity Enhancement Techniques

Intensity enhancement aims to improve image quality which makes it easily comprehensible by a human eye [60].To achieve mammogram image enhancement, the contrast of a mammogram is increased by creating a large intensity difference between the image of interest and background. When a mammogram image is enhanced the background noise is removed which makes it easy to determine the Region of Interest and also the complexity for reading and interpreting a mammogram by a radiologist is reduced [137].

An effective mammogram enhancement technique aims to improve an image by enhancing features in masses [137]. Researchers have used several techniques. However, the histogram modification approach is the most popularly applied technique. It is based on creating a new distribution of intensities that are highly uniform by reassigning pixel intensity values [31]. A narrow histogram implies a low contrast image while a uniform histogram depicts a high-contrast image [137]. I discuss some commonly used histogram enhancement techniques below.

The histogram equalization (HE) technique works by uniformly distributing intensity values and maximizing information visibility in mammograms. Each pixel value is substituted by the integral of the histogram image in that pixel. Because HE efficiently spreads out high-intensity values to allow low-intensity values to get better contrast [138]. This technique is suitable for use with images that have a dark or bright foreground and background, especially when the image is viewed on an x-ray [139]. In a mammogram, HE

is employed to adjust the contrast of the image which allows better visibility of image anomalies. A desirable outcome of the HE technique is in its effectiveness to equally distribute the image pixels in the entire image. However, this technique only works effectively when the image has only one object or if there is no dissimilarity between the object and its surroundings [31].

Adaptive histogram equalization is different from HE in that it calculates multiple histograms from different sections of an image instead of the entire image. Kayode et al. [137] ascertained that AHE is suitable for enhancing edges in regions and improving region contrast. However, Sivaramakrishna et al. [140] argued that AHE tends to generate a significant level of noise. Further, Kayode et al. [137] noted that AHE suffers from overestimation since noise is heightened because almost uniform regions generate highly peaked histograms that result in large values.

The generation of noise experienced by Adaptive Histogram Equalization (HE) can be overcome by using the Contrast Limited Adaptive Histogram equalization technique. With CLAHE, the contrast in a homogeneous area is limited to avoid amplifying noise. Contrast Limited Adaptive Histogram equalization technique operates on small areas called tiles. Every tile is enhanced, such that the histogram of the output region is approximately similar to the histogram defined by the distribution parameter [141]. It requires getting a local histogram defined by neighbors of each pixel and using a defined limit to trim the histogram, then redistribute the pixels and finally integrate the histograms [137]. Because

mammogram images comprise more than one object CLAHE would be a suitable technique because it does not over enhance.

2.8 Breast Cancer Segmentation Techniques

Image segmentation techniques partition a digital image into many sets of components to highlight significant regions of interest for further investigation [130]. They define suspected abnormalities and isolate them from the surrounding regions using tumor edges or pixel regions. The output of the segmentation process is a set of segments jointly covering the entire image. They segment images automatically or manually based on similar properties of color, shape, or intensity. However, in either case, segmentation is challenging because of overlapping tissues. Providing a segmentation technique that can detect and localize breast anomalies is highly desirable. Researchers have applied segmentation techniques in breast cancer detection for confining suspicious areas, tracking the progression of breast tumors, and analyzing anatomic structures. Segmentation techniques like thresholding, using edges, region growing, and clustering is used with mammograms.

2.8.1 Thresholding Techniques

The threshold segmentation technique separates the foreground and background by converting a grayscaled image into a binary image based on an optimum threshold [142]. It uses the threshold value for generating a binary image by allocating zeros to all pixels

values below the threshold and assigning ones to all pixel values above the threshold. The threshold-based segmentation technique can be performed globally or locally. In the global threshold, there is a search for a threshold value T_g by relying on the intensity of nearby tissues, while in local thresholding it defines several local threshold values Tl for every pixel based on the intensity values of its neighboring pixel [143]. Global thresholding produces poor results when background illumination is uneven. Even though threshold-based segmentation techniques do not perform well in images with the close color spectrum and do not guarantee object coherence [144], they are simple to implement, fast especially if repetition is on similar images and are used for real-time application [145]. A commonly used threshold-based segmentation technique is Otsu.

The Otsu threshold algorithm creates a one intensity threshold that splits the pixels into two distinct classes. It selects an optimal threshold value by minimizing the weighted intensity variance within the class intensity by applying equation (2.1) given by:

$$\partial_w^2(t) = w_0(t)\partial_0^2(t) + w_1(t)\partial_1^2(t) \qquad (2.1)$$

Where weights w_0 and w_1 are probabilities of two classes whose value ranges from 0 to 255 at threshold t and ∂_0^2 and ∂_1^2 are class variances for the two classes, while ∂_w^2 (t) is the minimized value. Equation (2.2) defines the class probability for the two classes:

$$w_0(t) = \sum_{i=1}^{t} (h(i)) \quad\quad (2.2)$$

$$w_1(t) = \sum_{i=t+1}^{I} (h(i))$$

The mean for each of the class $\mu_1(t)$ and $\mu_2(t)$ is calculated by applying equation (2.3).

$$\mu_1(t) = \sum_{i=1}^{t} \frac{ih(i)}{w_0(t)} \quad\quad (2.3)$$

$$\mu_2(t) = \sum_{i=1+t}^{1} \frac{ih(i)}{w_1(t)}$$

Then the variance is obtained by applying equation (2.4)

$$\partial_0^2(t) = \sum_{t=1}^{t} [i - \mu_1(t)]^2 \frac{h(i)}{w_0(t)} \quad\quad (2.4)$$

$$\partial_1^2(t) = \sum_{i=1+t}^{1} [i - \mu_2(t)]^2 \frac{h(i)}{w_1(t)}$$

The Otsu method achieves outstanding performance when the histogram has a bimodal distribution. However, if the object area is small compared to its background area, the histogram will no longer be bimodal. When the intensity variance between the object and background area is larger compared to the mean difference, it degrades the ability to separate between the gray level histogram.

49

2.8.2 Edge-Based Techniques

Edges are points in an image where the gray values change from one pixel to another. Edge-based segmentation methods use the abrupt pixel change of an image to extract useful information. Because edges occur at image location which represents image boundaries, edges can segment the tumor region in mammographic images to aid separation of tumor from surrounding tissues. Further, it can show breast abnormalities.

Gradient and Laplacian are the two main edge-based segmentation techniques. To find an edge, a gradient is used to detect edges by calculating the first derivative of image intensity. Examples of gradient techniques include; Roberts, Prewitt, and Sobel. Laplacian techniques locate an edge where the second derivative has zero-crossing [146]. Even though edges are easy to perceive from the human eye and work well with good contrast they are highly sensitive to noise and perform poorly in images with low contrast [144].

2.8.3 Region-Based Techniques

The region-based segmentation techniques divide a region into homogenous regions of connected pixels by using pre-defined criteria like intensity, color, or shape. They include region growing, split and merge, and watershed transform method.

The region growing approach works by selecting an initial seed point from which the region grows depending on similarity with the neighboring pixels until it achieves a

homogeneous and connected region. It can correctly segment a region that has the same properties, provides clear edges, and the user can accurately define the initial seed point and criteria for growing the seed [147]. However, it is computationally expensive, sensitive to noise, and lacks a global view of the problem since it is a local technique [147].

The split and merge technique divides an image into four by four squares by using Region Adjacency Graph (RAG). It then merges the squares depending on similarity with neighboring pixels [148]. It repeats the procedure until it gets a homogeneous connected region. The advantage of this technique is it allows the user to set the criteria for a split which can differ from the merge criteria. Also, the user can progressively determine the number of split levels. However, this technique suffers from blocky segments caused by the number of split levels used [149]. Even though making high-level splits reduce blocky segments, the process leads to high computation time.

Watershed transform uses mathematical morphology to compute the gradient of the original image. Watershed is simple to implement, produces unique solutions for a particular image, adapts to different digit grid, and extends to a k-dimensional image [150]. Since a mass tumor has a higher intensity than its surrounding, watershed segmentation is suitable for mass segmentation. However, when the watershed transform is applied directly to a mammogram image, it can cause over-segmentation because of the presence of noise in the image. Also, because it does not produce smooth boundaries, it is inefficient and thus may not be a final segmentation technique [151]. Even though region-based methods apply multiple criteria at the same time and perform well in the presences of noise [144],

their performance depends on the selection criteria, in addition, region-based techniques are computationally expensive in terms of time and storage and can lead to over or under segmentation.

2.8.4 Cluster-Based Technique

Cluster-Based techniques group objects based on attribute similarity. They generate clusters using K mean since it is easy and fast to implement efficient, and scalable to large datasets [145]. However, it is a challenge to define the number of clusters because there are no explicit selection criteria [152]. We find a discussion on some segmentation techniques for mammographic images in [153] [154] [155].

Table 2. 3: Mammogram segmentation techniques

Segmentation techniques	Description	Advantages	Disadvantages
Edge-based methods	It aims at identifying sharp discontinuities in the image	Easy to be perceived by humans Suitable for images with good contrast	Unsuitable for low contrast images Highly sensitive to noise
Thresholding based methods	Selects an appropriate threshold value according to image properties then assigns pixel images to specific regions	Can be used for real-time applications Simple to implement Fast especially if repeating on similar images	Does not work well in images with a close color spectrum No guarantee of object coherence
Region growing methods	Divide a region into homogeneous regions of connected pixels based on predefined criteria such as intensity, color, or texture.	More than one criteria can be applied concurrently It is simple to implement Performs well in the presence of noise	It is dependent on the selected criteria It is computationally expensive in terms of time and memory It suffers the problem of under or over-segmentation
Cluster-based methods	Groups objects based on their similarity	Easy to implement Allow 3-dimensional segmentation	Challenged on how to define the number of clusters

2.9 Mammogram Feature Extraction Techniques

Feature extraction is a dimensionality reduction technique that aims at transforming a high dimensional input feature set to a low dimensional feature set containing features with good discriminant ability between one class and another [157]. The reduction of features

53

implies not only imposing a random or predetermined ceiling on the number of features for the newly generated feature set, but the choice of features also depends on the importance of the feature based on discriminate power. Since the extracted or selected are the only inputs that guide the classifier, it is paramount to have a feature set that contains the most significant and discriminate features that could positively influence the chances of achieving high classification results. Features extraction techniques for breast cancer detection using mammograms are categorized based on descriptors for shape, texture, and Hybrid [158]. Below is a detailed discussion of the descriptors.

2.9.1 Shape-Based Descriptors

The shape is a significant visual cue applied in medical image processing for the identification of images. Extracting shape features from an image is a low computation process that involves few properties of an image resulting in a robust feature set [121]. Shape features play a significant role in applications used for retrieving, identifying, and classifying shapes of an image [130]. For breast cancer detection, the shape of a breast tumor is extracted then classified as a benign or malignant tumor. The shape of an image is defined by its margin properties. The margin defines the edges, which define a transition between mass and surrounding tissues. Understanding the margins aid in knowing to what extent the cancer cells have spread, which brings out the size of the mass. Coupling this knowledge with the mass property of having defined shapes is used to define mass features. Shape-based descriptors are extracted based on the entire region or contours of the image.

Region-based features make use of the interior content of an image. It considers all the pixels within the shape region for representing shape. Region-based features can be global or structural, depending on whether they divide the shape into subparts. The global region-based approach considers the entire shape region for shape representation, while the structural approach partitions the shape region into sub-parts used for representing shape [121].

Contour based shape features use boundary features by using either a global or structural approach. The global approach does not split the shape into sub-components rather, the whole boundary information is used to get the feature vector for the matching process. Structural methods split the shape boundary information into small segments, called primitives. The final representation of the structural method is a string or a graph used for matching during the process of image retrieval [121].

Shape features extraction techniques are designed for different purposes based on how accurately the features promote effective recognition. However, a good shape feature extraction technique can effectively discover shapes similar in a pattern matching problem, irrespective of whether the shape being recognized is rotated or flipped. Also, it should be robust such that it can effectively compare two images even though one may be noisy or distorted. Kallergi and Übersichtsarbeit [159] reported morphology to be a significant clinical factor used by clinicians' detection of microcalcification. However, a major challenge with extracting shape features is that it becomes more complicated in the presence of noise, occlusion, and arbitrary distortion making recognition difficult [121].

Also, the accuracy of the diagnosis algorithm relies on the effectiveness of the segmentation process [83].

2.9.2 Texture-Based Descriptors

Texture is a feature of homogeneity of images using the pixel as the fundamental for tonal variation, which has a certain scale, regularity, and directionality [125]. Texture highlights the structural arrangement and environmental relationship of the object in an image and reveals important discriminatory characteristics related to variability patterns [126]. The section below explains the local texture feature extraction techniques applied for breast cancer detection.

2.9.2.1 Gabor Filter

The Gabor filter is a texture analysis technique that uses different frequencies and orientations modulated by a Gaussian function [160]. It computes a 2D Gabor filter as shown by equation (2.5)

$$g(a,b) = \left(\frac{1}{2\pi\sigma^a\sigma^b}\right)exp\left[-\frac{1}{2}\left(\frac{a^2}{\sigma_a^2} + \frac{b^2}{\sigma_b^2}\right) + 2\pi jSx\right] \qquad (2.5)$$

In which σ_a and σ_b are standard deviations of the distribution while S is the radial frequency.

The general equation for a Gabor filter bank is given by equation (2.6):

$$g_f(a,b) = z^{-m}g(\bar{a},\bar{b}) \qquad (2.6)$$

where,

$$\bar{a} = acos\theta + bsin\theta \, and \, \bar{b} = -asin\theta + bcos\theta$$

$$\theta k = \pi\frac{k-1}{n}, k = 1,2,3 \dots n$$

where n is the number of orientations used and $z^{-m} = 0,1,2 \dots S$ for S scales.

The 2D band-pass filter has shown optimal localization properties in both spatial and frequency domains. Hence, making it suitable for extracting image features positioned in a particular orientation and within a certain frequency because shape defines breast cancer masses and margin for an orientation. Researchers have extensively used the Gabor filter for mass segmentation and edge detection in breast cancer mass detection. Salabat et al. [161] Initialized different filters in Gabor bank to different scales and orientations to extract patterns in ROI to that differentiated between normal and abnormal breast tissues.

2.9.2.2 Linear Discriminant Analysis

The Linear Discriminant Analysis (LDA) is a dimensionality reduction strategy that aims at transforming a high dimensional feature vector to a low dimensional space by increasing the ratio between intraclass scatter S_w and inter-class scatter S_b.

S_w and S_b measures are given by equation (2.7):

$$S_w = \sum_{b=1}^{c}\sum_{a=1}^{N_b}(x_a^b - \mu)(x_a^b - \mu)^T \tag{2.7}$$

Where x_a^b is the ath instance and μb is the mean of class b, while b,c is the number of classes. N_b is the number of instances in class b calculated using equation (2.8)

$$S_b = \sum_{b=1}^{c} (\mu_b - \mu)(\mu_b - \mu)^T \qquad (2.8)$$

2.9.2.3 Gray-Level Co-occurrence Matrix (GLCM)

The Gray-level co-occurrence matrix is a texture measure that uses image distance and orientation to examine the texture of a grayscale image by comparing every pixel with its neighboring pixel. The image distance between the reference pixel value and its neighboring pixel values forms a square shape [162] quantized in 0^0, 45^0,90^0,180^0 orientations [109]. Also, it is balanced about the diagonal, such that if there is a difference of 2 cells from the diagonal, then it is a two-level gray difference [163].

The GLCM value at *GLCM (x,*y) is defined by probability measure of reference pixel *x* and gray value *y* at neighboring pixel by distance *d*, orientation θ given by equation (2.9)

$$P_{d,\theta} = P(x,y) \qquad (2.9)$$
where $P(x,y)$ is defined by

$$P(x,y) = \frac{GLCM\ (x,y)}{\sum_{x=0}^{N}\ \sum_{y=0}^{N} GLCM\ (x,y)}$$

58

Even though GLCM is easy to implement and it has good performance in terms of the processing time [164], and it gives good results in a large field of application, however, the large dimensionality forces it to reduce the number of gray levels and may not be effective in images with a lot of noise. Also, the image quantization process leads to loss of information, making the extracted features not reliable [83]. Also, there is no established way of choosing the displacement vector (d) and calculating Co-occurrences matrices for different values [83].

2.9.2.4 Local Binary Pattern

The Local Binary Pattern (LBP) is a local texture descriptor proposed by Ojala et al. [165]. Given an N × N image, LBP operator thresholds p neighboring pixels with the central pixel results to an 8-Bit binary code. The LBP operator considers p neighboring pixels along a circular path either clockwise or counter-clockwise and R distance, which is the radius of comparison as shown in formula (2.10):

$$s(x) = (x_c, y_c) = \sum_{i=0}^{7} s\,(g_i - g_c)2^i \qquad (2.10)$$

$$s(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & if\ x < 0 \end{cases}$$

Where g_i the value of its neighbors is, g_c is the gray value of the referenced pixel, x_c is the total number of neighbors, and y_c is a neighborhood radius. First, it identifies the center pixel value, then compares that central pixel with the neighboring values using a defined radius. The thresholding and encoding steps define the LBP extraction algorithm.

During thresholding, the LBP compares the value of the central pixel with all the neighboring pixel values P based on an identified distance R. If the value of the neighboring pixel is higher than the value of the central pixel, then it allocates value 1 to that position otherwise value 0. The values are then read clockwise or counter-clockwise into a binary value. Thresholding aims to get the local binary differences [166] that result in an eight-bit binary number.

The LBP operator is computationally simple and can withstand monotonic gray-scale changes [167]. Figure 2.6 shows a sample 3*3 input image and its corresponding LBP code. The generated LBP code is then used to get a global histogram.

7	6	5
7	7	4
7	9	8

1	0	0
1		0
1	1	1

1	2	4
128		8
64	32	16

Binary Pattern = 11110001_2

Decimal= $1+16+32+64+128=241_{10}$

    (a)   (b)   (c)   (d)
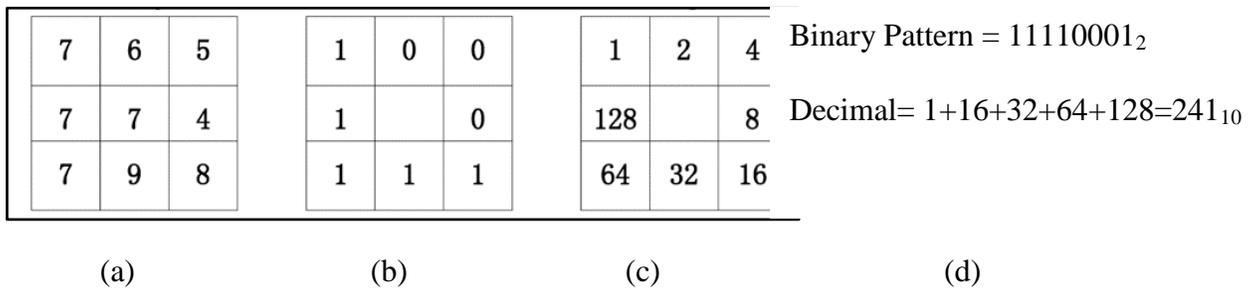
Figure 2. 6: Sample LBP code (a) Sample image region (b) Threshold image (c) Weights assigned to image (b) (d) Binary pattern and its equivalent decimal value

The ability to discriminate and computational simplicity of LBP has made it a common technique for breast segmentation and classification, therefore researchers have adopted LBP for identifying breast abnormalities in mammograms [168] which has facilitated LBP feature extraction during segmentation and classification of breast cancer tumors.

2.9.2.5 Local Ternary Pattern

Relying on the central pixel as a threshold in LBP makes it sensitive to noise. A minor change of central pixel significantly changes the LBP code. To overcome this challenge, LTP extends LBP by thresholding the pixels into (0, 1,-1) instead of (0, 1). Using three value pixels makes LTP robust to noise than LBP. Consider threshold constant c, center pixel r, and neighbor pixel n. Equation (2.11) gives the LTP formula.

$$S(N) = \begin{cases} 1, if\ n > r + c \\ 0, if\ n > r - c\ and\ r + c \\ -1\ if\ n < r - c \end{cases} \tag{2.11}$$

Where $S(N)$ s is the nth neighbor containing the LTP code value.

After thresholding, to get rid of the negative values, the ternary pattern generated is divided into positive and negative patterns.
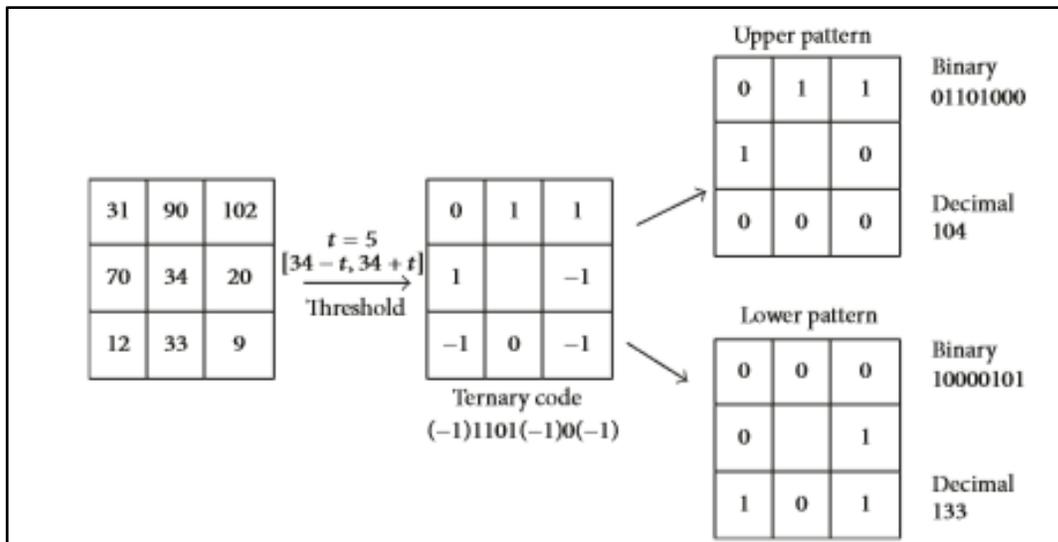
Figure 2. 7: LTP Operator

While LTP can resist noise, its major disadvantage is that modification done is invariant under grayscale transform of intensity values. Also, the discriminate property is still inadequate since it can classify two or more distinct patterns in the same class.

2.9.2.6 Local Directional Pattern

To resolve the challenge of relying on neighboring pixel intensity which makes LBP unstable, Jabid et al. [30] proposed a Local Directional Pattern descriptor that encodes image texture by computing edge response values of a pixel in a different direction. Popular edge detectors like Frei-Chen, Kirsch, Sobel, and Prewitt edge detectors are used in this case [169]. Among the edge detectors, Kirsch is the most popular because it identifies different directional edge responses more accurately by considering all the eight neighbors of a pixel [169]. Figure 2.8 shows the Kirsch mask.

$$
\begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix}
\begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}
\begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}
\begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}
$$

$M_0(0°) \qquad M_1(45°) \qquad M_2(90°) \qquad M_3(135°)$

$$
\begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix}
\begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}
\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix}
\begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}
$$

$M_4(180°) \qquad M_5(225°) \qquad M_6(270°) \qquad M_7(315°)$
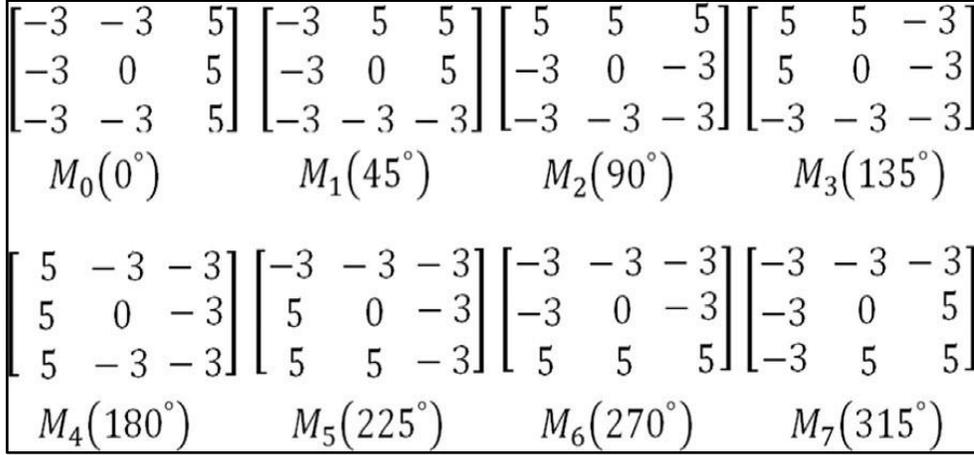
Figure 2. 8: Kirsch Mask

The original LDP does not consider all response values because they are not all equally important. Therefore the k most top prominent directional values b_i are selected while the remaining $(8-k)$ are set to 0 as shown by the formula (2.12).

$$
LDP_k \sum_{i=0}^{7} bi \ (mi - mk)x \ 2^i \tag{2.12}
$$

Where bi is

$$
bi(a) = \begin{cases} 1 \ a \geq 0 \\ 0 \ a < 0 \end{cases}
$$

85	32	26		313	97	503		0	0	1
53	50	10	→	537	X	399	→	1	X	1
60	38	45		161	97	303		0	0	0

LDP binary code: 00010011   LDP decimal code:19

Figure 2. 9: A sample LDP code using $k = 3$

LDP operator is more robust to changes caused by noise than LBP because it applies orientation responses that are more stable than intensity values used by LBP [30]. Even though LDP is more stable than LBP, it depends on the top three responses which result in information loss around a local neighborhood. Also, according to Shabat and Tapamo [170], LDP is computationally expensive when compared to other techniques.

2.9.2.7 Local Quinary Pattern

While LBP and LTP encode the pixels into (0,1) and (0,1,-1) respectively, Local Quinary Pattern (LQP) encodes the pixels into 5 values by using two threshold values τ_1 and τ_2 set by the user [171] using equation (2.13).

64

$$\begin{cases} 2 & p \geq gc + \tau_2 \\ 1 & gc + \tau_1 \leq p \leq gc + \tau_2 \\ 0 & gc - \tau_1 \leq p \leq gc + \tau_1 \\ -1 & gc - \tau_2 \leq p \leq gc - \tau_1 \\ -2 & Otherwise \end{cases} \qquad (2.13)$$

The Local Quinary pattern is split into four binary patterns. Figure 2.10 illustrates how a Quinary code is split into four binary patterns.
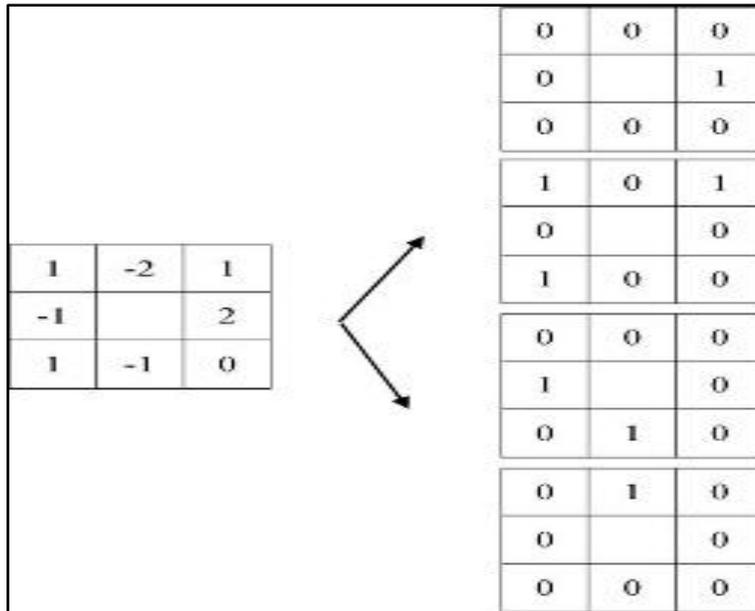
Figure 2. 10: Quinary Code split into four binary code

2.9.3 Hybrid-Based Descriptors

Hybrid feature extraction techniques combine two or more feature extraction techniques to help improve the detection of masses and microcalcification in breast cancer Computer-Aided detection systems [122]. When hybrid feature extraction techniques are used, the performance is expected to be higher than when using individual feature extraction techniques. However, care should be taken to ensure no two or more similar techniques are

65

combined, as not much improvement will be achieved. Also during, combination care should be taken such that no two combined techniques cancel the effects of each other.

2.10 Breast Cancer Feature Selection Techniques

Feature selection is a process of searching for the best feature subset among competing candidate subsets, according to some defined evaluation measure. The selection process involves generating a new feature set by disregarding features with insignificant or no predictive information while ensuring detection accuracy does not decrease [172]. Feature selection infers not only to a random or pre-defined threshold on the number of features to be considered, rather it further defines the type of features to be selected based on their characteristics, their presence, or lack of interaction with other features and selection algorithm. This implies that features are picked or discarded depending on their importance to the target class.

The process of feature selection involves performing a search to generate a subset of candidates, then testing the goodness of the generated subset by checking against a stopping criterion and then determining the result through a validation process. It then uses the selected features for classification. Filters, wrappers, embedded and hybrid is a way of categorizing feature selection techniques [173].

2.10.1 Filter Technique

The filter technique evaluates the importance of a feature by considering its intrinsic characteristics independent of the classification algorithm [174], [175]. It selects a feature

based on a ranking criterion used to score all the features and a defined threshold value used to eliminate all features below the threshold value [174]. Subsequently, the features obtained are fed to a classification algorithm. Table 2.4 shows a generalized filter algorithm.

Given dataset D, the algorithm begins with an initial subset S_0, which could be an empty/full set or a randomly generated subset. Subsequently, the initial subset S_0, is evaluated based on an independent measure M and the result is stored in Ω_{best}. After that, a new subset S is created and evaluated (by the same M independent measure) for further comparison with the previous best one Ω_{best}. If it is found to be better, it is considered being the best current subset. The process is repeated until a preset stopping criterion ε is attained. The algorithm then outputs the last best current subset S_{best} as the final set of best features.

Table 2. 4: Pseudocode of a generalized filter algorithm (Source: [45])

Filter Algorithm

Input:

       D (f_0,f_1,..,F_{N-1})   // initial dataset with N features

       S_0 // an empty subset

       ε // stopping criteria

Output:

       S_{best} // best subset

Begin

Ω_{best} =eval (S_0,M);   // Evaluate S_0 by independent measure M

 do begin

S = generate (D);   // Generate a subset of features for evaluation

Ω=eval (S, D, M);   // Evaluate the current subset S by an independent measure M

If ($\Omega>\Omega_{best}$)

$\Omega_{best} = \Omega$;

$S_{best} = S;$

end until (ε is reached);

return $S_{best;}$

end

Filter techniques are categorized into univariate/ attribute evaluation and Multivariate/ subset evaluation techniques. Univariate filters consider each feature separately during the selection process [177]. Each feature is given a score based on an evaluation measure and amount of information within a feature that can help describe the target class. The features are then ranked based on a score. Features with the lowest score are ranked last. A threshold can be defined to help eliminate the features with the least score. Univariate filter techniques are scalable, independent on any classifier, thus exhibit generality and are fast [177]. However, they ignore feature dependencies since every feature is evaluated separately depending on its relevance with the target class. Consequently, a feature can be discarded as being useless on its own, yet may be useful when considered with others

68

[178]. Also, they do not interact with a classification algorithm which makes the technique produce results not tailored to a specific classifier.

Apart from evaluating class relevance like univariate filters, multivariate filters calculate feature dependencies between each feature pair to find relationships among features [174]. Even though multivariate filters are less computational complex than wrappers techniques, they are slow and less scalable when compared with univariate filters [174].

2.10.2 Wrapper Technique

Wrapper techniques use a predetermined selection algorithm to evaluate and determine the features to select [178]. They define a search procedure in the space of likely feature subset then generate and evaluate various subsets of features. To get an evaluated subset of features, they define a specific selection model which makes this technique custom-made to a particular algorithm [174]. Wrappers are better at obtaining optimal feature by allowing for the specific biases and heuristics of the learning algorithm and the training set [179]. Table 2.5 shows a generalized wrapper algorithm.

Table 2. 5: Pseudocode of a generalized wrapper algorithm (source: [45])

Wrapper Algorithm

Input:

D (f$_0$,f$_1$,..,F$_{N-1}$)   // initial dataset with N features

S_0 // an empty subset to start the search

ε // A stopping criterion

MLC   // A specified Machine Learning Classifier

Output:

S_{best} // an optimal subset of features

Begin

θ_{best} =eval (S_0,D,MLC);   // Evaluate S_0 by a *MLC*

do begin

S = generate (D);   // Generate a subset of features for evaluation

θ=eval (S, D, MLC);   // Evaluate the current subset S by classifier MLC

If ($\theta>\theta_{best}$)

$\theta_{best} = \theta$;

$S_{best} = S$;

end until (ε is reached);

return S_{best};

end

2.10.3 Embedded Techniques

The embedded techniques select features using a particular learning algorithm during the training process. Their interaction with the learning algorithm is at a lower computational cost than with the wrapper techniques [178], [179]. The embedded techniques capture feature dependencies considering not only relations between one input feature and the output feature but also look for local features that allow higher local discrimination. They

70

select the optimum feature subset for a known number of elements in a set using an independent criterion. They use the selection algorithm to choose the ultimate best feature subset among the optimal feature subsets across different cardinalities.

2.10.4 Hybrid Techniques

Hybrid techniques combine two or more similar or different selection techniques. See a generalized hybrid algorithm in Table 2.6. They first apply a filter technique to decrease the number of features in the original feature set. Then they use the resultant feature subset as input to the second stage, where they apply a wrapper technique to choose an appropriate number of features. Another way of implementing a hybrid technique is by applying the filter technique twice to a dataset.

Table 2. 6: Pseudocode of a generalized hybrid algorithm (Source: [45])

Hybrid Algorithm

Input:

D (f$_0$,f$_1$,..,F$_{N-1}$)   // initial dataset with N features

S_0 // A subset to start the search

Output:

S_{best} // an optimal subset of features

Begin

C_0 = card (S_0);   // Calculate the cardinality of S_0

Ω_{best} = eval (S_0, D, M);   // Evaluate S_0 by an independent measure M

Ω_{best} = eval (S_0,D, MLC);   // Evaluate S_0 by a classifier MLC

71

for $c = C_0 + 1$ *to N* **begin**

for i= 0 to N-c **begin**

$S = S_{best} \cup \{f_i\};$ //Generate a subset of feature with cardinality c for evaluation

Ω=eval $(S,D,M);$ // Evaluate the current subset S by an independent measure M

If $(\Omega > \Omega_{best})$

$\Omega_{best} = \Omega$;

$S_{best} = S;$

end

θ=eval $(S'_{best}, D, MLC);$ // Evaluate S'$_{best}$ by a classifier *MLC*

If $(\theta > \theta_{best})$

$S_{best} = S'_{best};$

$\theta_{best} = \theta;$

else

break and return S_{best}

end

return $S_{best};$

end

2.10.5 Comparative Analysis of Feature Selection Techniques

Feature selection aims to generate a set of features with significantly important features while ensuring it selects not too many or too few features than required. The problem of feature selection has been addressed through univariate filter techniques such as discussed

by Heshmati et al. [180] and Alharbi et al. [181]. Researchers are relying heavily on univariate filters that perform ranking to order features because they are simple, scalable, and have good empirical success [182].

The ranking is done by looking at the relevance of a feature in relation to the class. Therefore, they discard a feature that is not significant to the target class. This brings out two issues. First, some features when considered alone may not be useful but become useful when considered together with other features. Also, two features that are not useful individually can become useful when combined with other features [182]. Secondly, ranking is based on relevance alone. This means they do not consider redundancy therefore, choosing two highly ranked correlated features results in a feature subset of closely correlated features. According to Guyon and Elisseff [182], correlated features are redundant because there is no additional information gained by including two or more of these features. Also, the best feature set does not necessarily comprise the best individually selected features. Therefore, the selection of features should be done by looking at features as a set rather than on contribution made by an individual feature. This makes univariate filter techniques inappropriate for selecting the best feature set. However, in most real-world situations, the best set of features and the number of features in such a set is unknown.

The lack of feature dependencies experienced in univariate filter techniques provoked the introduction of multivariate filter techniques. Apart from considering the relevance of the features as in the univariate filter techniques, multivariate filter techniques also consider

redundancy. Redundancy among features is identified through correlation criteria as in [183], [184] or mutual information criteria in [185], [181]. Some correlation criteria can only detect linear dependencies between features and the target class. This implies that some degree of redundancy will still be present. Also, a dataset can contain uncorrelated features that complement each other.

Correlated criteria cannot discriminate complimentary features therefore two complementary features are presumed to be different. Guyon and Elisseff [182] proved that absences of correlation do not imply the absence of feature complementarities. The demerit of mutual information criteria is that it considers the quantity of dependency rather than the type of dependency which can cause the inaccurate ordering of the features because a feature with more information does not necessarily mean it is more useful than one with less information.

Wrapper techniques addressed the lack of interaction with the classifier experienced by filter techniques. Wrapper techniques consider feature dependencies and selection algorithms, which makes them tailored to the specific learning algorithm. Researchers in [94], [85], [95] used wrapper techniques. However, they used forward or backward selection techniques, which have a high likely hood of falling in the local minimum [181]. Vasantha [95] noted that Sequential Forward and Sequential Backward selection present sub-optimal solutions and suffer from the nesting effect. Such that initially selected /removed features cannot be discarded or reintroduced later. Wrapper techniques are

generally computationally expensive, especially when dealing with large datasets. Wrapper techniques also suffer from overfitting, however, this can be avoided by using cross-validation [179].

To take advantage of the simplicity of filter techniques and the high accuracy of wrapper techniques, researchers developed hybrid techniques. Perez [9] and Huang et al. [186] developed hybrid techniques that exploit the efficiency of filters and the accuracy of wrappers. To circumvent computation expenses experienced by wrappers, hybrid techniques perform feature selection by first using filter techniques to rank the features and also reduce the input feature set for the application of the wrapper in the second step. Even though hybrid techniques have reported outstanding performance, sometimes the performance improvement is not significant enough to warrant the effort, which depends on the way they form the hybrid. While some hybrid techniques conduct one filtration after another, others attempted to find the intersection of two or more filters and use the output for the multivariate filter technique as in [186].

A major challenge facing researchers is how to choose an appropriate algorithm for the selection process. Some researchers have chosen selection algorithms without considering their stability. Considering the stability of a feature selection algorithm makes the technique robust to minor variations, especially when new learning samples are added or removed from the learning set [187], [188], [189]. Naseriparsa et al. [190] attempted to handle the instability of the selection algorithm by considering sample size and carefully

choosing the selection algorithm. However, this has not been exhaustive. The stability of a selection algorithm should not be examined alone, rather in combination with the predictive performance of the features, because factors such as inherent characteristics of an algorithm, underlying data distribution, and sample size affect stability. For instance, multivariate filter-based feature selection techniques are less stable than Univariate filters. [187]. Haury et al. [188] and Kalousis et al. [189] showed that ensemble feature selection is a better approach for achieving stable feature subsets. Table 2.7 shows the merit and demerit of feature selection techniques discussed.

Table 2. 7: Merits and Demerits of feature selection techniques

Feature selection method	Examples	Merit	Demerit
Filter	Maximum Difference Feature Selection (MDFS) [46]; Correlation-based Feature Selection [47]; ReliefF [48]	Computationally efficient, Independent of any algorithm Fast and scalable	lack interaction with the classifier Ignore feature dependencies
Wrapper	Sequential Forward Selection (SFS)[49]; Modified Artificial Bee Colony Feature Selection (MBCFS) [50]; Enhance Cuckoo Search Feature Selection [51]	Includes interaction between feature subset search and model choice Takes into account feature dependencies Less computationally intensive	Depend on a specific classifier They are likely to overfit than filter methods They are computationally intensive
Embedded	PSO-KNN [52]; PSO-KNN [53]	No need to split the training data Allow interaction with classification model Less prone to overfitting	Classifier dependent Computationally expensive than filter-based methods
Hybrid	New Particle Swarm Optimization and Genetic Algorithm [54]; GA [55]; SFS, SBS, and F-score [56]	-Have better accuracy than filter and wrapper methods	-Classifier dependent selection -A lot of effort can result in insignificant improvement

2.11 Breast Cancer Classification Techniques

Classification is a vital stage in a CAD system for differentiating and labeling abnormalities by mapping data to a pre-defined target class. A classification algorithm aims to build a classifier that takes some input features during training and learns the pattern, then uses the knowledge learned to predict new features during the testing phase

77

[196]. Breast cancer is purely a classification problem in which the classifier assigns a label based on the attributes tested. It can either be a binary classification problem where two distinctive features are used to distinguish absences or presences of cancer cells or a multi-class problem that uses more than two labels. Since radiologists are often prone to making mistakes when analyzing mammograms, using classification algorithms is a suitable approach for automating the analysis of breast tumors because classifiers can learn complex relationships and patterns [197].

Classification techniques used in breast cancer detection are supervised or unsupervised. In supervised classification techniques, the classifier learns based on a training set, and they use the gained knowledge for classification. Geometric and statistical distance measures are used to define how close a point is to each of the training samples. The approach has two main phases: A training phase in which data is analyzed by a classifier and a testing phase in which a classifier estimates accuracy using test data [198]. Supervised techniques rely on a training set that distinguishes spectral distinctiveness of the classes by avoiding two or more classes being similar so that the rate of misclassification is low. To achieve this, they require a well-developed training set to ensure a good representation of data which aids outstanding performance for the classifier. Some examples of commonly used supervised classification techniques for breast cancer classification include Support Vector Machine, Artificial Neural Networks, K-Nearest Neighbor, Random Forest, Bayesian and Naïve Bayes.

In unsupervised classification, there are no class samples provided, rather they generate clusters they then assign the clusters classes (classified). It is faster than a supervised technique because it does not require prior learning. Also, they create classes based on spectral information only, and therefore, they are not subjective. However, sometimes the spectral classes may not correspond with the information classes, consequently, more time is used to interpret and label the classes. The section below explains classifiers used for breast cancer classification.

2.11.1 Support Vector Machine

Support Vector Machine (SVM) is a supervised classifier that uses statistical learning theory to find a hyperplane that separates two classes. They aim at maximizing the margin between the separating hyperplane and data points on either side of the hyperplane. If the training data is separable linearly, then there exists a pair (v,b) calculated using equation (2.14).

$$V^T x_i + b \geq 1 \; for \; all \; x_i \in P \qquad (2.14)$$
$$V^T x_i + b \leq -1 \; for \; all \; x_i \in N$$

In which the decision rule is computed by;

$$f_{v,b} = sgn \; (V^T x + b) \qquad (2.15)$$

where v is the weight vector and b is the bias

The SVM classifier creates the hyperplane, using a selected kernel. Choosing a suitable kernel function is essential because the kernel defines the feature space in which the learning set instances are classified. The two most popular kernels are polynomial and Gaussian kernel. The equation (2.16) represents the polynomial kernel where d is the polynomial degree while they calculate the Gaussian kernel using equation (2.17).

$$k_{a,b} = (a.b + 1)^d \qquad (2.16)$$

$$k_{a,b} = exp\left(\frac{\|a - b\|}{2\partial^2}\right)^2 \qquad (2.17)$$

The advantages of SVM are (1) they are highly accurate, especially in high-dimensional space because they can model complex nonlinear decision boundaries. (2) They are less prone to over fit than other supervised classifiers. (3) They provide a compact description of the learned model [199]. (4) They have good generalization properties compared to conventional classifiers because during training they aim at minimizing misclassification risk [85], [186].

2.11.2 K- Nearest Neighbor

Nearest-neighbor is a nonparametric technique in which places a new instance in a class closest to it depending on a pre-defined distance measure. Given a learning set $L = \{(y_i, x_i), i = 1,2,3 .....n\}$ where y_i denotes class membership and x_i represents predictive

80

value. To determine the nearest neighbor $(y1, x1)$ for the learning set L, on distance function d for a new instance (y, x) is given by formula (2.18).

$$d(x, x_i) = \min\big(d(x, x_i)\big) \qquad (2.18)$$

Popular distance metrics used to define the closeness of an instance to the nearest neighbor are Euclidean distance [199], Manhattan [200], and Murkowski. Even though KNN requires large storage, it is highly sensitive on the choice of similarity function and defines no mechanism for selecting k, it however effectively handles noisy and large training data.

2.11.3 Artificial Neural Network

Artificial Neural Network (ANN) is a classifier that has neurons arranged into input and output layers, as sometimes hidden layers which convert an input vector into some output. Each neuron takes an input, applies a nonlinear function to it, and then passes the output on to the next layer The ANN is appropriate when the problem to solve is too complex to use the conventional techniques or challenging to get algorithmic solutions [9]. However, ANN has a generic layered structure which proves to be time-consuming and results in very poor performance. Furthermore, this technique uses "black-box" technology. Therefore, attempting to understand and explain exactly how it performs the classification process is almost impossible [201].

2.11.4 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is a classification approach developed by R. A. Fisher whose goal is to transform a feature space with n-dimensional samples to a smaller feature subspace k where $k \leq n-1$ while maintaining the class-discriminatory information [9]. Even though LDA is mathematically robust and can model complex problems with good accuracy, it, however, produces C-1 feature projections such that if the classification error estimates establish that more features are needed, some other method must be employed to provide those additional features. Also, because it assumes a unimodal Gaussian likelihood when the distributions are significantly non-Gaussian, consequently it is not able to preserve any complex structure of the data, which may be needed for classification.

2.11.5 Ensemble of Classifiers

Some of the individual classifiers such as discussed in the above section are limited because of high time and space complexity, especially when the sample size is large, others are limited in terms of high error rates leading to high rates of misclassification. Based on this, studies have proposed the hybridization of individual classifiers as a way to improve their classification performance. One such approach is to ensemble classifiers either from the same family or different families. The ensemble of classifiers not only achieves better classification performance than individual classifiers but also, increases the reliability and confidence of results achieved [57]. Because of the diversity of individual classifiers, they are prone to make different errors. Therefore allowing an individual classifier to participate

in an ensemble, enables the classification error of one classifier to be compensated by another classifier. However, caution must be taken on what classifier to ensemble to ensure that a classifier does not cancel the positive effects of another classifier.

An ensemble of classifiers can be constructed by either influencing the input feature space or classifier output targets. An ensemble of the classifier using input feature space considers the input features given to the classifier. Classifiers in the ensemble can either work with the same set of training features or different subsets of the training set that are apportioned to the classifiers. In terms of classifier output targets, the final decision could be based on a majority voting rule which gives equal weights to the decisions of each classifier and carries out the prediction with the highest number of votes[58]. Alternatively, weights can be assigned to individual classifiers and the final decision is based on the classifier with the highest votes.

2.12 Performance Evaluation

Performance of a breast cancer model is measured by True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) using a Receiver Operating Characteristic (ROC) [204] and/or a confusion matrix [205]. Receiver Operating Characteristic is a graphical plot applied for binary classifiers. Using different threshold values, it creates a graphical plot of True Positive Rate against False Positive Rate. An advantage of ROC is that it presents the performance of a classifier in a visual format that is interpreted graphically or numerically [130]. A Confusion matrix visualizes the

83

performance of a classifier by displaying total correct and incorrect predictions made by the model in a table form [206]. It defines the final decision by using True Positive, True Negative, False Positive, and False Negative measures. The final decision is the values along the diagonal of the confusion matrix.

Alongside evaluation tools, there are specific evaluation metrics like accuracy, error rate, sensitivity, specificity, Precision, and F-score used to measure the true performance of a classifier.

Accuracy assesses the effectiveness of a classifier by looking at the percentage of instances predicted correctly. Even though it is a simple and intuitive evaluation measure for a classifier, it is not appropriate for imbalanced data [207]. Equation (2.19) defines it.

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{FN} + \text{FP} + \text{TN} + \text{TP}} \qquad (2.19)$$

The compliment of Accuracy is the Error rate, which seeks to find the percentage of instances predicted incorrectly. Accuracy and Error rate is a general metric that adapts to the multi-class classification problem. Equation (2.20) defines the error rate.

$$\text{Error rate} = \frac{\text{FN} + \text{FP}}{\text{FN} + \text{FP} + \text{TN} + \text{TP}} \qquad (2.20)$$

The sensitivity metric evaluates the effectiveness of a classifier. By looking at the proportion of samples belonging to the positive class, which was correctly predicted as positive against all true positive and false negative samples. Equation 2.21 shows the calculation for the sensitivity metric.

$$\text{Sensitivity} = \frac{TP}{FN + TP} \qquad (2.21)$$

The specificity metric evaluates the effectiveness of a classifier. By looking at the proportion of samples belonging to the negative class correctly predicted as negative against all true negative and false-positive samples. Equation (2.22) shows the specificity metric.

$$\text{Specificty} = \frac{TN}{FP + TN} \qquad (2.22)$$

The precision metric correctly predicts a positive prediction. Unlike the accuracy measure Precision and sensitivity can work well with imbalanced data, however, they present a severe shortcoming in that they cannot classify the negative examples, and they are more useful when combined with accuracy or when applied to both positive and negative classes [207]. Equation (2.23) defines it:

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (2.23)$$

85

F-score is a common metric for imbalance problems [208]. F-score values range from 0 to 1, where 1 is perfect classification and 0 is a total failure. It combines sensitivity and precision which are good metrics for retrieving information where imbalance problem occurs. F score depends on the ß factor, which is a parameter that takes values from 0 to infinity. ß factor is used to control the influence of sensitivity and precision separately. When ß = 0 the f score reduces to precision and conversely when ß approaches infinity, the f score reduces to sensitivity [207]. To integrate the two measures, ß = 1 is used, making it represent a harmonic mean between sensitivity and precision. Equation (2.24) defines it.

$$\text{F-score} = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision Sensitivity}} \qquad (2.24)$$

2.13 Breast Cancer Datasets

Breast cancer datasets are used to facilitate research in breast cancer analysis and aid in developing algorithms used in teaching and training. There are several breast cancer datasets available for research. While some are publically available others are privately owned therefore accessible by specific institutions. Even though the research community has attempted to define an ideal breast cancer dataset, each dataset is unique in terms of the kind and number of cases handled, the conditions the images were taken, and the information provided about each case. Because of differences in strengths and weaknesses, it is a challenge to compare performance differences of models, approaches, and techniques based on these datasets. However, a study conducted by Nahid and Kong [1]

86

showed that DDSM and MIAS datasets have the highest number of research papers published.

2.13.1 Digital Database for Screening Mammography (DDSM)

Digital Database for Screening Mammography is a publically available dataset [209] for researchers analyzing mammogram images. The dataset was initially accessible through http://marathon.csee.usf.edu/Mammography/Database.html [210] which is no longer available, however, the cases are organized into several volumes and are available online by anonymous ftp of ftp://figment.csee.usf.edu/pub/DDSM/cases.

The cases in the dataset are from Massachusetts General Hospital, Wake Forest University School of Medicine, and Washington University. Originally, the dataset had 596 cases with 373 non-cancerous and 223 cancerous cases [210]. However, with time they improved it by adding more cases, therefore it now has 2620 cases. Out of the 2620 cases, 695 are normal, 95 Benign, and 101 Malignant. It also has an enhanced interface to enable previewing of each case. The database uses the Joint Pictures Expert Group (JPEG) file image type, and it is 231GB in size distributed in 43 volumes. Every case comprises two images of each breast, patient information from different ethnic and racial backgrounds, and the age of a patient. It also describes the breast density using (ACR) and Breast Imaging Reporting and Data System (BI-RADS) annotation [211] [212].

2.13.2 Mammographic Image Analysis Society (MIAS)

Mammographic Image Analysis Society is a 2.3 GB dataset generated by UK research groups that wanted to understand mammograms. It is publically accessible and contains 322 mammographic images of both breasts from 161 patients. The 322 digitized mammograms are in the mediolateral view out of which 51 are in malignant class, 64 are in benign class, and 207 in the normal class [213]. Further, the MIAS dataset gives information on the location and radius of the tumor, type of abnormality, and breast tissue affected [85]. The image file format is Portable Gray Map (PGM) and the mammograms are available through a Pilot European Image Processing (PEIPA) archives at the University of Essex.

Even though the dataset contains breast density information, Rungayyan [214] observed that the classification of breast density information was not according to any standard. Also, although it defines the MIAS annotation based on the region of interest centered on a circular radius, this was not sufficient for studies such as that conducted by Oliver et al. [215] which required manual segmentation for all circumscribed and spiculated masses. Also, the resolution used to digitize the images makes it unsuitable for experiments for detecting microcalcification, yet the dataset had the highest percentage of microcalcification [212].

2.13.3 Image Retrieval in Medical Applications (IRMA)

Image Retrieval in Medical Applications dataset contains images collected from the Department of Diagnostic Radiology, Department of Medical Informatics, Division of Medical Image Processing, and the Chair of Computer Science VI at the Aachen University of Technology [216]. The project aimed to develop and implement methods for Content-Based Image Retrieval. The dataset contains 10,509 reference images split into normal cases (12 volumes), cancer cases (15 volumes), and benign cases (14 volumes): each case may have one or more associated Pathological Lesions (PLs) segmentations, usually in Medio-lateral-Oblique (MLO) and Cranio-caudal (CC) images of the same breast [9].

2.13.4 Breast Cancer Digital Repository (BCDR)

Breast Cancer Digital Repository (BCDR) is a Portuguese breast cancer image dataset with real female patient information from medical records supplied by the Faculty of Medicine at the University of Porto, Portugal [217] [209]. They released it for the public domain in 2012, even though it is still under development. It has 1734 patient cases against 5776 Mediolateral-Oblique (MLO) and Cranio-Caudal (CC) image views [217].

They split the BCDR dataset into Breast Cancer Digital Repository–Film Mammography (BCDR-FM) and Breast Cancer Digital Repository-Digital Mammography (BCDR–DM). Out of the 1734 cases, the BCDR-FM makes up 1010 cases with an age bracket

89

between 20 -90 years [217]. The BCDR–DM comprises 724 cases of which 723 are female and 1 male with ages between 27 and 92 years old.

2.13.5 INBreast

The mammogram images in the INBreast dataset were gotten from the breast center in CHSJ, Porto, allowed by the Hospital Ethics Committee and National Committee of Data protection. It is a publicly available dataset with 115 cases and 410 MLO and CC images from which 90 cases are from women with both breasts affected (four images per case) and 25 cases from mastectomy patients (two images per case) [212]. They saved the images in DICOM format with all confidential information removed. The images are FFDM taken from screening, diagnostics, and follow up cases. The dataset is available at http://medicalresearch.inseporto.pt/breastresearch/GetINbreastDatabase.html.

2.13.6 Nijmegen Dataset

The Nijmegen dataset contains 40 digitized mammograms of both Cranio-Caudal (CC) and Medio-Lateral Oblique (MLO) views recorded from 21 patients by the National Expert and Training Centre for Breast Cancer Screening and the Department of Radiology at the University of Nijmegen Netherlands [218]. They digitized the images with a screen-film resolution of 2048 2048 pixels [218]. Each mammogram image shows one or more microcalcification clusters verified by histology. The total number of clusters in the dataset is 105. The truth circle marked by expert radiologists shows the location and size of each

cluster. The dataset also contains look-up tables for rescaling of the image data and text files, storing the center and diameter of the truth circles in pixel coordinates [218].

2.13.7 Banco web LAPIMO dataset

It is the most recent dataset that requires users to register, gain access and contribute using http://lapimo.sel.eesc.usp.br.bancoweb/. It has 320 cases, 1473 images in MLO and CC view. It avails patient background information and breast density annotation with BI-RADS.

2.13.8 Mammography Image reading for Radiologists and Computers Learning (MIRAcle) dataset

It is a web-accessible mammographic dataset with a dynamic repository for machines and radiologists training and evaluation module [212]. It has 204 mammograms collected from 196 patients and it is available for classification or education evaluation by radiologists [219].

Other mammographic datasets such as Lawrence Livermore National Laboratories (LLNL) [216] [220], Washington University Digital Mammography Database [220], Trueta [221] [222], Malaga [222], and Rheinisch-Westfälische Technische Hochschule (RWTH) [216] exist, however, most of them are not available in the public domain, others are privately owned while others are still at the experimental stage therefore much information about them is unknown. We can find more explanation about them in [51, 59-

61]. Table 2.8 presents a brief description of the characteristics of breast cancer mammographic datasets.

Table 2. 8: Characteristics of breast cancer mammographic datasets

Dataset	No. of images	Mode of image acquisition	Image view	Lesion type	No. of patients	Access type	Ground Truth
DDSM	9916	Screen Film	MLO and CC	All kinds	2620	Public	Boundary chain code of findings
MIAS	322	Screen Film	MLO	All kinds	161	Public	Centre & radius of the circle around the area of interest
IRMA	10509	Screen Film	MLO and CC	All kinds	Unknown	Public	Boundary chain code of findings
BCDR	5776	FFDM	MLO, CC, and Ultrasound	All kinds	1734	Public	Lesion Contour
INBreast	410	FFDM	MLO and CC	All kinds	115	Public	Lesion Contour
Nijimegen	40	Screen Film	MLO and CC	MCs	21	Private	Centre & radius of the circle around the area of interest
Banco Web	1400	Screen Film	MLO, CC, and	All kinds	320	Public	ROI available in some images

LAPIMO			others				
MIRAcle	204	Unknown	Unknown	Unknown	196		ROI of findings
LLNL	198	Unknown	MLO and CC	MCs	50	Public	Binary images of MCs clusters &area of some MCs
Trueta	320	FFDM	Unknown	Unknown	89	Private	Centre & radius of the circle around the area of interest
Malaga	Unknown	Unknown	MLO and CC	Masses	35	Unknown	Not available

2.14 Breast Cancer Evaluation Protocols

Evaluation protocols are statistical methods that estimate the generalization ability of a model based on some new data [223]. While there are several evaluation protocols, they have common characteristics. For instance, they all keep a subset of data for testing the model and train the model using the remaining set of data. However, they differ in calculating the generalization error, the number of samples kept for testing, and whether the process repeats many times or performed only once [223]. Because breast cancer models are likely to over fit, it is, therefore, important to test the model on unseen data which enhances generalization.

Cross-validation is a popular evaluation strategy that works by averaging together multiple runs of model tests. It partitions data into a training set and a testing set. Depending on how data is partitioned, cross-validation can be exhaustive (Leave-One-Out Cross-Validation and Leave P-Out Cross Validation) in which it tests all possible ways to divide the data or non-exhaustive (k-fold, hold out and monte Carlos) which approximate data split.

Hold out is an exhaustive cross-validation which involves simple and efficient computations [224]. Given a dataset of size *N*, it randomly partitions the dataset into a training set (d_0) and testing set (d_1). Even though the size of each set is arbitrarily, typically the testing set should be smaller than the training set. A common split is to have *2/3 (N)* for training and *1/3(N)* for validation. It builds a model using the training set and evaluates

95

performance on the testing set. It performs the entire process of training and testing only once, therefore making it computationally efficient. However, it has two inherent biases; bias from individual observation and bias owing to a small training set leading to overestimation. Also, the lack of averaging multiple runs may lead to highly misleading results because multiple iterations help to smoothen out the error leading to a more stable model [225].

Another commonly used strategy is the *K*-fold Cross-Validation. It lays a foundation for other types of cross-validations. Other types of validations are special cases of *k*-fold or comprise repetitive rounds of *k*-fold Cross-Validation. In *k*-fold cross-validation, it generates *k* equal subsets from the original sample and *k* iterations performed on training and testing set. For every repetition, it set aside a different fold for testing, and the remaining *k-1* set aside for training. It gets an estimated error by averaging all testing errors and also gets the confidence range of the estimated error by calculating the measure of variance such as standard deviation [226].

An advantage of *k* fold lies in its ability to eventually use all samples for both training and testing the model. Also, even though value k is an unfixed parameter, the most commonly used values for *k* are 5 and 10, which give an estimate that is statistically likely to be accurate for 5 to 10 times longer computation time. [226]. For instance, 10-fold cross-validation is a popular alternative to LOOCV because it needs less computation time than LOOCV when the numbers of samples are many. However, *K*-fold Cross Validation is

slightly more positively biased than LOOCV, since it develops the classifier on 90% of the full dataset and it provides error estimates with less variance [227].

Leave One Out Cross Validation (LOOCV) is a special *K*-fold validation where *K* (the number of folds) is equal to *N* the number of samples [228]. It is also called the jackknife method [227]. It works by taking all samples for training except those used for testing. Therefore, for *n* samples, there are *n* different training sets and *n* different test sets. Given a dataset with *C* classes, it performs *C-1* validation experiments. For each experiment, it uses data from *C-1* classes for training, and data from the class left out for testing. Therefore, given a dataset of *n* breast images, LOOCV performs *n-1* validation experiments. In each experiment *I,* it uses n_i breast images for testing and the remaining *n-1* for training. This strategy uses each image for both training and testing. This way, it validates in the same way as its application scenario.

The advantages of this strategy are its implementation simplicity, and it does not waste a lot of time, since it removes only one sample for testing and the remaining *n-1* for training. In terms of accuracy, LOOCV has high variance as an estimator for the test error which decreases by finding the average of all *n* training–testing partitions [227]. Also, it requires less computation time than Leave P Out cross-validation (LPOCV) since LOOCV requires $C\binom{n}{1}$ passes while LPOCV requires $C\binom{n}{p}$ passes. Moreover, it eliminates the positive bias of the partition because each training set comprises one less sample than the full dataset, [227]. However, *n* passes may still require a large computation time, in which case another

strategy such as *K* fold cross-validation may be more appropriate [225]. Also, because of the high variance, it is likely to generate unreliable estimates [229].

Leave P Out cross-validation (LPOCV) is the same as LOOCV when the value of *p=1*. Given a set of *N* images, LPOCV uses *p* observations for validation and *N-p* for training. It repetitively leaves out every possible subset of *p* images where $p \le (N-1)$. Then it averages the error for all trials, to give overall effectiveness. Even though this strategy is exhaustive, it can become computationally infeasible to train and validate models for all possible combinations, especially when p is moderately large [230].

Another type of cross-validation is Monte Carlo Cross-Validation (MCCV) commonly known as Repeated Random Sub-sampling. It generalizes the split sample method. Given a dataset of *N* samples, MCCV generates several random splits of the dataset into a training set (d_0) and testing set (d_1). For every split, it trains the model on the training data and tests the performance on testing data. Then it averages the outcomes over the splits. For every training-testing partition, every instance appears either in the training or testing set, but not in both. And because the training and testing split is random, an instance may appear in a test set several times. An instance may appear in *m* training set and not in the test set, whereas another instance may appear in all the test sets, but not in the training set. Compared to k fold cross-validation, MCCV training /testing split does not depend on the number of folds. However, because of overlap, MCCV can oversample some observations while not sampling others, unlike in k fold which has disconnected test sets

[227]. Therefore, with a repetitive analysis using different random splits, the overlap results in varying outcomes.

Molinaro et al. (2005) noted that even though MCCV does not decrease the bias of the split sample method, it significantly decreases the variance of the split sample error estimate. They also reported that the random training and testing splits are sufficiently significant to decrease the variance. As the number of random splits tends to infinity, the results of MCCV validation tend towards LPOCV validation.

Bootstrap is a re-sampling strategy introduced by Efron and Tibshirani commonly used when dealing with small datasets [224]. [231]. A bootstrap set of instances forms by uniformly picking instances from the original data n to form a training set with the same number of samples as the original dataset. The samples that are not selected for training form a testing set. Since it samples data with replacement, it gives the probability of any data instance not being chosen by $\left(1 - \frac{1}{n}\right)^{n} \approx e^{-1}$ which is approximately 0.368, while the probability of a data instance being picked for the training set is (1-0.368) = 0.632n. Therefore, the estimated accuracy of bootstrap sampling increases as the number of repeats increase.

2.15 Related Work

This section presents previous studies on texture descriptors for breast cancer tumor identification. It highlights the feature descriptor used, the strengths and weakness of the descriptors, and the results achieved. Oliver et al. [232] used an LBP descriptor for the false-positive reduction in mammographic mass detection by first getting a global descriptor for each ROI, then analyzing its spatial texture information. They then divided the ROI into small regions and local texture description computed using Local Binary Pattern. The combined Local descriptor in a spatially enhanced histogram defined the feature descriptor they used. Using the SVM classifier with 1792 ROI's extracted from the DDSM dataset and testing the experiment using Leave-One-Out cross-validation, the used descriptor reduced the false-positive rate. Even though they showed that LBP features are effective for false-positive reduction using different ROI image sizes, however, the LBP descriptor could assign the same pattern to a pixel in a tumor region and to another pixel in a normal region which leads to noticeable reductions in the number of false positives. Also, by extending the basic LBP histogram into a spatially enhanced histogram leads to long histograms which affect the efficiency of the descriptor.

To reduce the effect of noise inherent in the basic LBP descriptor Chen et al. [233] developed a Robust Local Binary Pattern (RLBP) descriptor which maps a non-uniform pattern to a uniform pattern by partitioning each 8 bit LBP binary code into a sequence of three consecutive bits. If they find 010 or 101 bits in the binary code, they replace them with 000 or 111 bits, respectively, consequently reducing the impact of noise. The RLBP

descriptor is simple to comprehend, and the experimental results from two datasets showed that it is robust in the presence of noise. However, mapping the non-uniform pattern to a uniform pattern consequently distorts the overall description and any wrong mapping affects the quality of the final histogram.

Sansare and Kinge [167] used Gabor filter and LBP to classify benign, malignant, and normal breast cancer cells using the SVM classifier on 158 MIAS images. The images underwent preprocessing of noise removal image enhancement and pectoral muscle removal. They achieved an accuracy of 96.72%, 84%, and 81.90% for the benign, malignant, and normal class respectively. A major weakness of LBP is its dependence on the intensity difference of the pixels, which makes it very sensitive to noise and illumination changes, which makes it unsuitable in differentiating between a benign and a malignant tumor.

Rabidas et al. [234] compared the performance of LBP, LBP Variance, and Complete LBP descriptors for benign/malignant breast mass classification. By using a stepwise logistic regression method for feature selection and the Fisher Linear Discriminant Analysis classifier on 200 mammograms from the DDSM database, they achieved a classification accuracy of 92.95% 87.7%, and 90.6% with LBP, LBPV, and CLBP respectively using tenfold cross-validation. Even though LBP achieved the highest accuracy, it was not rotation invariant and therefore not suitable for benign/ malignant classification, which requires edge orientation at the margin.

101

To achieve a local pattern that is rotation invariant, Rabidas et al. [235] developed Discriminative Robust LBP (DRLBP) and Discriminative Robust LTP (DRLTP) texture descriptors for classification of mammographic masses into benign and malignant classes. They based the methodology on the hypothesis that both texture and edge information of a malignant mass differs from a benign mass. Using the Fisher Linear Discriminant Analysis classifier, stepwise logistic regression for feature selection, and tenfold cross-validation on 58 masses from the Mini MIAS dataset, they achieved the best results with DRLBP having an AUC of 0.98. Even though the results of DRLBP and DRLTP were very close to each other, the authors did not test the statistical significance of that difference. Also, the authors used very few images to warrant a reliable conclusion.

Gardezi and Faye [236] fused Completed Local Binary Patterns (CLBP) with curvelet features for Normal/Abnormal classification using images from MIAS and IRMA datasets. First, they computed CLBP features using rotationally invariant mapping; then, they computed curvelet features from the curvelet sub-band coefficient. They extracted CLBP and curvelet features, they then fused them and passed them to a classifier. They got a classification accuracy of 96%. Even though the author tested the statistical significance of the developed descriptor against other descriptors, we attribute the high accuracy to the fusion of different texture features and using images from two different datasets.

Ponraj et al. [237] developed a Local Binary Textual Patterns that considers the central pixel. They compare the central pixel with the neighboring pixel to generate a binary bit.

They then convert the binary bit to a decimal number by counting the number of zeros and ones. Then textual features are used to classify a mammogram into either a normal or an abnormal class. By using an SVM classifier and 70 images from the MIAS database, they achieved an accuracy of 97.2% and 96.4% for Binary pattern one and Binary pattern two, respectively. The major criticism of the methodology used is that it leads to long histograms that increase the computational speed. Further, there was no comparative analysis of the developed descriptor with Original LBP or existing LBP variants as evidence of significant contribution. Also, the authors only stated that they performed preprocessing but did not explain how they did it.

Abdel-Nasser et al. [238] developed a Uniform Local Directional Pattern (ULDP) descriptor for classifying breast tissues into a mass or normal and breast density into fatty, glandular, or dense tissue. They used two publicly available mammographic datasets; Mini MIAS and INbreast. Using NLSVM, LSVM, LDA, and MLP classifiers, they got AUC of 0.92, 0.93, 0.91, and 0.92 respectively. Although the developed descriptor achieved excellent results, however, it is important to note that the CAD system was not fully automatic since the selection of the Region Of Interest (ROI's) was manual. Also, the selection of ROI's was done from the middle of the breast, yet a tumor can occur anywhere within the breast. Therefore, it is uncertain that the developed descriptor achieved good mass/ normal classification. Further, even though the authors concluded that the developed descriptor can properly discriminate between different tissues regardless of their shape,

there is no evidence that they investigated the influence of the shape of the masses on the performance of the descriptor.

Rampun et al. [239] used the Uniform LTP descriptor that modeled the appearance of the fibro-granular disk region. They used SVM on the MIAS dataset with stratified ten-fold cross-validation. They achieved an accuracy of 82.33%. Even though the authors used Uniform LTP which is more stable, less prone to noise, and has fewer labels than the original LTP. However, extracting features from both lower and upper patterns in eight different orientations leads to the generation of very long histograms which slowed downed the classification phase. Also, this method highly depends on the parameter settings of the LTP operator.

Muramatsu et al. [240] developed the Radial Local Ternary Patterns (RLTP) that considered not only pattern orientation about the center of a mass but also robustness to image rotation. Using 376 ROI's from Nagoya medical dataset and ANN, SVM, and RF classifiers, for benign /malignant classification, they achieved the highest Area Under Curve (AUC) of 0.90 with the ANN classifier. They performed the test using Leave-One-Case-Out cross-validation, the training and test dataset were not completely independent. Therefore, to test the effectiveness of the developed descriptor, there is a need to validate it with an independent dataset.

Paramkusham et al. [241] used the Local Quinary Pattern (LQP) to classify mass into normal/abnormal classes and geometric features for classifying abnormalities into either a

benign or a malignant tumor. The images got from the IRMA dataset underwent noise removal through a median filter and automatic segmentation using the k-mean algorithm. They used the SVM classifier, which attained classification accuracy of 99.27% and 79.13% for the normal, abnormal, and benign, malignant classes respectively. Even though the developed Local Quinary Pattern (LQP) showed stability in accuracy levels for all the five folds used, however, the authors did not test the statistical significance for the Benign/Malignant classification, yet the difference was small.

Apart from the Local texture descriptors, the research community has also developed other descriptors. Faya et al. [242] used wavelet-based feature extraction, where they define a wavelet coefficient for each image set. They then selected the threshold value and used the Euclidean distance to differentiate between two classes. They tested the approach on the MIAS dataset for normal, abnormal, and then benign, malignant classes. With a training/testing ratio of approximately 50:50, they achieved a classification accuracy of 98.55% and 98% for normal, abnormal, and benign, malignant classes respectively.

Eltoukhy et al. [243] proposed a texture feature extraction technique based on the curvelet transform. The mammogram image is first cropped, then passed through the curvelet transform. They then pass the resultant feature vector to a feature selector, then they performed normal/abnormal and benign/malignant classifications using the Nearest Neighbor classifier on images taken from the MIAS dataset. They got a classification accuracy of 97.03% and 91.68% for the normal/abnormal and benign/malignant

classification, respectively. Herwanto and Arymurthy [110] developed a system for diagnosing breast masses using GLCM features. The images first underwent preprocessing of image cropping, artifact removal, and enhancement. Using 73 images taken from the MIAS dataset, they achieved a classification accuracy of 88% for normal/abnormal classification. Gardezi and Faye [236] fused CLBP and curvelet features for classifying tumors into either normal or abnormal class by using images from MIAS and IRMA datasets. First, they computed CLBP features using rotationally invariant mapping. They then computed curvelet features from the curvelet sub-band coefficient, then pass the fused features to One Nearest Neighbour classifier. They attained an accuracy of 96%. Pratiwi et al. [244] extracted GLCM based features using RBFNN and BPFNN classifiers on the MIAS dataset. Experimental results showed that RBFNN achieved a higher accuracy of 93.98% and 94.27% for normal/abnormal and benign/malignant classification respectively than BPFNN. Biswas et al. [245] extracted GLCM features using 20 images from the MIAS dataset and KNN, SVM, and ANN classifiers. The mammograms first underwent preprocessing of artifact and noise removal, ROI extraction, and image enhancement. The experimental results showed that the 3NN classifier outperformed SVM and ANN with 95% accuracy for normal/ abnormal classification. Htay and Maung [246] extracted GLCM features from a mammogram image that had undergone preprocessing of noise removal, image cropping, and segmentation by Otsu thresholding. The KNN classifier was used to classify a mammogram image into either a normal or an abnormal class. Using 120 images from the MIAS dataset, they achieved an accuracy of 92%.

From the previous work on the application of texture descriptors for breast cancer classification using mammogram images, literature has revealed that the extraction of breast cancer features sparingly uses local texture features, because of low classification accuracy levels achieved especially for benign/malignant classification. Studies that produced high accuracies used either privately owned datasets, which are not accessible for verification, or they used a few images for testing. Therefore, there is a need to develop a more effective local texture descriptor for extracting breast cancer features, which consequently improves classification accuracy.

2.16 Theoretical Framework

I anchor this study on the feature analysis theory of face recognition studied by Shepherd et al. [59] in 1979. It is a bottom-up theory of pattern recognition because it looks at details first in terms of features (nose, mouth, and hair) of the face, then the entire picture when trying to recognize or describing the face. . Feature analysis theorizes that humans and animals have neurons and neural networks that function as detectors, which aid in recognizing objects by observing and assembling their features to determine the object. According to this theory, the visual system breaks down the incoming stimuli into its features and processes the information. However, some features may be more important than others. Therefore, there is a need to discern the most discriminant features and use them for more accurate recognition.

For instance an image with a two-dimensional (N by N) array of pointwise or pixel-wise intensity values and p possible pixels then the number of possible images is a set N of size pN^2. To distinguish all possible images having N by N pixels, there is a need for a space of N^2 dimensions which is too large in practice to search for a particular image. Therefore, the core idea behind feature analysis is the ability to transform the original feature set into a new set with fewer but discriminative features. The transformation can be by way of feature selection or feature extraction. While feature selection maintains the original set, feature extraction transforms the original set into a new feature set. Shepherd et al [59] aimed to see how features are used when recalling unfamiliar faces. Participants were briefly shown faces of people they had never seen before and then they had to describe the faces. The results showed that the participants describe the individuals based on features such as the hair and the eyes, indicating that faces of the unfamiliar individual tend to be recalled using the faces salient features.

In this study, the aim is to define a set of features that aid in differentiating between a normal, benign, and malignant tumor in breast cancer classification. I pick the set of features to be defined from mammogram images that contain very many features. I used the feature analysis theory to help in correctly differentiating tumor types by first reducing the dimension of the features through feature extraction. Even though basic logic portrays that the more features an image contains, the higher the classification power, however, this is not always the case. Presences of many features, some of which may be irrelevant and/or redundant degrade classification performance, especially when the training examples (a

108

common characteristic of medical data) and computation power is limited, which may lead to model overfitting. Since features are inputs to a classifier if the features are not informative enough the classifier performs poorly [248]. Therefore, the feature analysis theory aids in actualizing the feature extraction process.

Given N-dimensional features represented by $\{x_1, x_2 \ldots \ldots x_n\}$ feature extraction aims to project a higher dimensional space to a lower-dimensional space M whose feature vector is $\{z_1, z_2 \ldots \ldots z_m\}$, where $m < n$ and each of the m features is a summation of the input feature set $\sum\{x_1, x_2 \ldots x_n\}$. The aim is to project so that the lower dimensional feature set is smaller than the original which will facilitate better classification. In linear feature extraction, the data is assumed to lie on the lower dimensional linear space. The projection is done using matrix factorization [60]. Given a dataset $M: \times D,$ there exists a projection matrix $U: D \times K$ and a projection $N: Z \times K$ where $M = N.U$. Using $UU^T = 1$(Orthogonal property of eigenvectors), we get $M = N.UT$. A graphical representation is presented in figure 2.11. The resultant feature set M is expected to have: (1) fewer features than feature set N, (2) uncorrelated features in their reduced form (3) large variances between features.
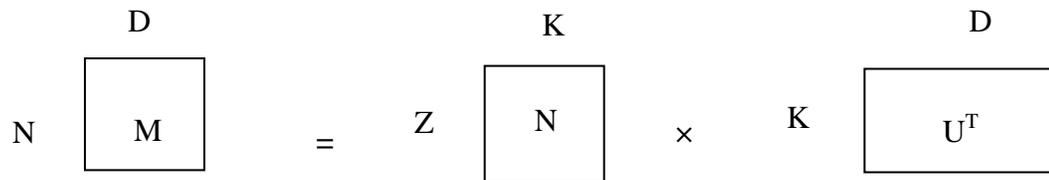
Figure 2. 11: Theoretical Framework

When feature set *M* has fewer features than feature set *N*, it not only implies setting a threshold on the number of features used to build feature set *M* but also selects features from feature set *N* based on their significant discriminant ability to the target class. Also, since correlation promotes feature redundancy, there is a need to have the features uncorrelated. Correlated features do not contribute additional information rather they promote redundancy, which does not add value to the decision taken by the classification algorithm [249]. Finally, a large variance between features promotes discrimination. If a feature takes similar values for all instances, then it cannot be discriminant enough. Since we need the features to distinguish between the different instances, therefore we need a large variance between the features, otherwise, the features will not be informative.

2.17 Conceptual Framework

The focus of this study was to develop a local texture descriptor for breast cancer classification. The study defined the local texture features as the independent variables and the level of classification accuracy of breast tumors as the independent variable. The moderating variables were the MIAS dataset, the MATLAB coding environment, and SVM and ANN classifier. I conducted cross-validation to ensure that the sourced dataset was valid, complete, and comprehensive Also, to ensure the classifiers were an effective representative sample, the selection of the classifiers was from different categories. A conceptual framework showing the relationship among variables in the study is shown in Figure 2.12.

Independent Variable

Dependent variable

| Local texture features | → | Breast tumor classification |

Moderating Variables

| MIAS Dataset | | Coding Environment MATLAB | | Classifiers: SVM and ANN |
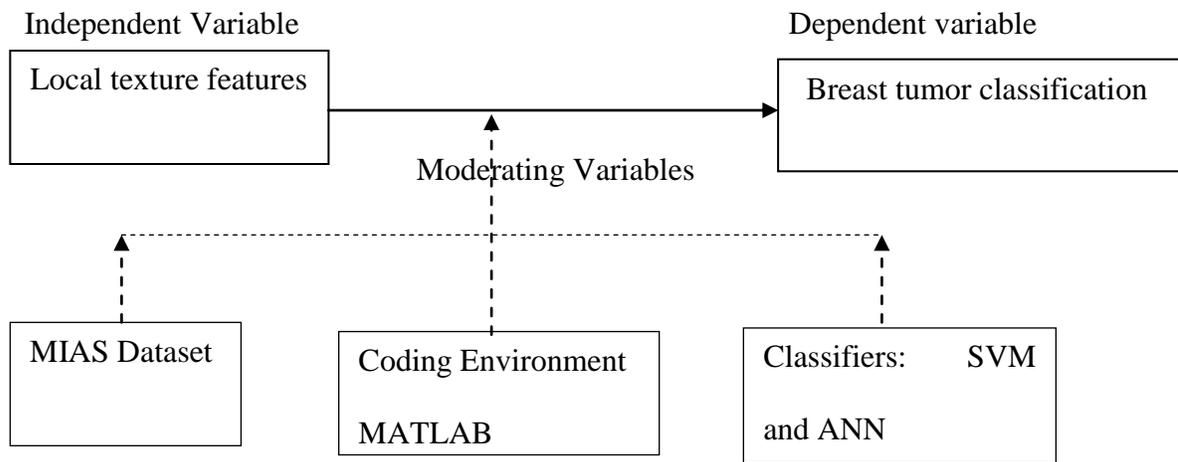
Figure 2. 12: Conceptual Framework

2.18 Summary

A comprehensive literature review on breast cancer detection revealed that enormous effort directed towards modeling breast cancer detection by developing feature extraction, feature selection, and classification techniques. In feature extraction, shape, texture, and hybrid descriptors are used to model the breast cancer detection system. Even though shape descriptors are a visual cue used by a radiologist in identifying breast tumors, they however make the recognition process difficult, especially when the image has noise, occlusion, and arbitrary distortion. Texture descriptors provide a better recognition because they take into consideration the structural arrangement and environmental relationship of the image. The hybrid based descriptors ought to yield better results since they combine two or more different descriptors, however, care must be taken to ensure the presences of one type of descriptor does not compromise or cancel the effect of the other descriptor.

Breast cancer detection is purely a classification problem. The classification is binary or multiclass. In binary classification, two attributes such as normal or abnormal, benign or malignant are used to distinguish a breast cell. Multiclass classifications use more than two labels, such as classifying a tumor into a normal, malignant, or benign class. The classification process is also viewed as either supervised or unsupervised. Both supervised and unsupervised breast cancer classification is used.

There are several datasets for breast cancer detection in mammograms as discussed in section 2.13 of this thesis. The datasets have a commonality in indicating the number of

112

images contained, the mode of acquiring the images, the lesion focused on, the access type, and grounded truth. However, they differ in terms of the image format and how much patient information is available. I observed that MIAS and DDSM datasets are the most popularly used and publically accessible datasets.

There are several evaluation protocols for breast cancer detection as discussed in section 2.14 of this thesis. They all have a common feature in which they partition the data into a testing and training set. However, these evaluation protocols differ in terms of; (1) the method for calculating the generalization error (2) criteria for partitioning the training and testing sets (3) the method adopted for carrying out repeats. Among the evaluation protocols, literature revealed the k- fold cross-validation is the most frequently used and reliable evaluation protocol.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Overview

This chapter addressed the research methodology adopted for the study. It began by demystifying the research philosophy adopted and provided an explanation and justification for the choices taken during the entire research process using Saunders's research onion. It also explained the experimental set-up for the methodology followed, a review of research objectives with a focus on the source of data and data analysis technique used. It also discussed the research tools, material, and model validation techniques used. The chapter concluded with a brief discussion on ethical considerations.

3.2 Research Philosophy

Research philosophy is a belief and assumptions on the development of knowledge in a specific field [250].The assumption created justify direction taken for conducting research, gathering, and interpreting knowledge, the researcher views the world in terms of values upheld in research and reality [251]. The selection of a research philosophy depends on the nature of knowledge examined. Thus, comprehending the research philosophy used aids in explaining the assumptions built within the research.

A research philosophy takes a Positivism, Realism, or Interpretive approach. Positivism adopts the philosophical stand of the natural science that emphasizes on universal truth and

quantification [250]. Positivist work with facts based on an observable reality which leads to the creation of reliable data, hypothesizing existing theories, and generalizing the results gotten rather than impression [250]. The researcher sees himself as a neutral recorder, and similar research conducted using the same instrument should reach the same conclusion.

This research adopted a positivist philosophical stance, which relies on facts and existing theory. It is a fact breast cancer exists and a mammogram can capture its features. Mammographic images are the observable reality and they anchor their credibility on the grounded truth provided for each image. This research is anchored on the theory of feature analysis which theorizes that humans and animals have neurons and neural networks that aid in recognizing objects by observing and assembling their features to determine the object. According to this theory, the visual system breaks down the incoming stimuli into its features and processes the information. However, some features may be more important than others. Therefore, the feature extraction process is used to discriminate features. The success of this research was evaluated based on studies conducted by other researchers in a similar area by measuring how closely the findings match, irrespective of the imperfection of the data collection instruments.

3.3 Research Design

This study used simulation and experimental design by simulating the training and testing the model on an existing dataset and performing validation experiments using the local texture descriptor developed. To demystify the entire research process, this study applied the research onion established by Saunders et al. [252]. The research onion was progressively used to illustrate the analysis of the different components of research so that I achieve a good research design. Its usefulness lies in its applicability and flexibility to any type of research methodology in a variety of contexts. The research onion covers research philosophy, research approaches, research strategy, research choices, time horizon, data collection and analysis, research design, and sampling techniques. The outer layer being the research philosophy.
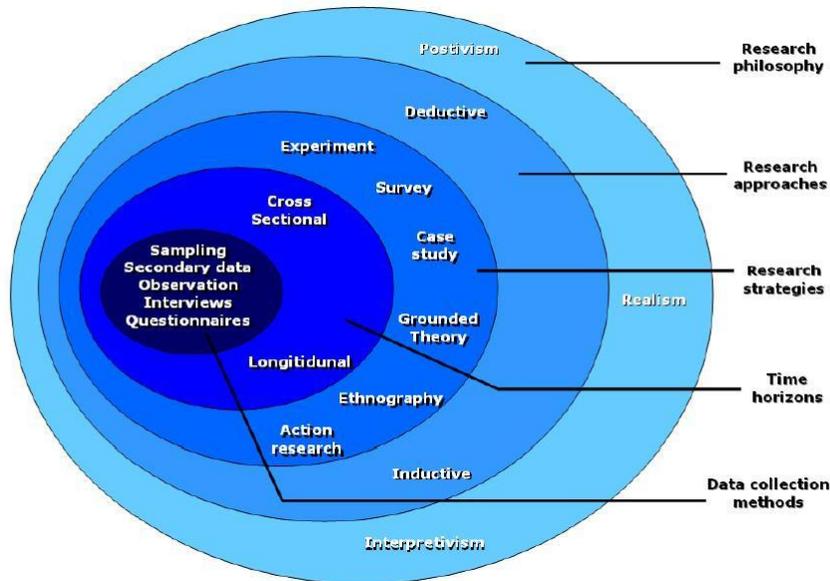
Figure 3. 1: Research Onion by Saunders 2012 [61]

116

3.3.1. Research Approach

The research approach is the second layer of the research onion. The research approach is meant to help the researcher make an informed decision on strategies to be used in research and to adapt to a research design that caters to constraints [250]. A research approach can either be inductive or deductive. In the inductive approach, the researcher starts from a specific aspect and moves towards general aspects of research while in deductive the researcher starts from general to specific by first making an assumption based on some pre-existing idea from which a research approach is formulated and tested.

This research adopted a deductive approach by basing it on the pre-existing theory of feature analysis, making it is possible to define a set of new features that are significant discriminate through the process of feature extraction. Further, this research formulated a research relationship between breast tissue features in a mammogram and the occurrence of breast cancer based on previous research. Finally, it is particularly suitable for a positivist approach, which is the philosophical stance adopted in this research.

3.3.2 Research Strategy

The third layer of the research onion is the research strategy. It is a step by step worked out a plan that directs the thoughts of a researcher [250], it can be experimental, a case study, action research, or based on grounded theory and ethnography. The choice depends on the research objectives stipulated.

Experimental research analyzes the results achieved against the expected results [250]. This research achieved objective two through simulation-based experiments. The research aimed to prove a pre-existing theory of feature analysis by extracting discriminant and significant breast tissue features from an image. The features were then used to predict the occurrence of breast cancer. Further, the development of the breast cancer classification models was by using discriminant features. Using an experiment, the researcher established a causal link between mammographic breast tissue features and the occurrence of breast cancer.

3.3.3 Time Horizon

The time horizon is the fourth layer in the research onion. It defines the data collection period within which the project is planned for completion [250]. The time horizon is defined as longitudinal or cross-sectional. In a longitudinal time horizon, the researcher repeatedly gathers data over an extended period because the phenomenon under observation changes over time [250]. In cross-sectional time horizon data is collected at a specified time and therefore the researcher is required to gather the data only once. Therefore, this research applied a cross-section time horizon because the mammographic images were taken at a certain time in the life of a patient. The datasets used showed the age of the patient when the mammogram images were taken.

3.3.4 Data Source

Data collection is the innermost layer of the research onion. Data collected can either be primary or secondary. This research utilized images collected from the MIAS breast cancer dataset. Section 3.3.4.1 of this thesis gives a detailed explanation of the dataset. Because the dataset contained few images, the number of images was increased through the process of data augmentation as discussed in section 3.4.3 of this thesis.

The mammographic images used to validate the proposed model were obtained from the Mammographic Image Analysis Society (MIAS) breast cancer dataset. The dataset was generated by a UK research group that wanted to understand mammogram images. It is publically accessible for scientific research and has 322 digitized mammographic images of both breasts from 161 patients. It provides background information on the class of abnormalities present in a mammogram image. The class of abnormalities comprises normal/ abnormal class and based on abnormality severity, the abnormal class is further split into malignant and benign class. Out of 322 mammogram images, 207 images are normal and 115 images are abnormal. Among the 115 abnormal images, 64 are benign and 51 are malignant [213]. The abnormalities are grouped into microcalcification, circumscribed mass, ill-defined mass. Spiculated mass, architectural, and asymmetry as shown in Table 3.1. In this research, all the 322 mammogram images were used for training and testing the model.

Table 3. 1: Image distribution in the MIAS dataset[62]

Class	Benign	Malignant	Total
Microcalcification	12	13	25
Circumscribed mass	19	4	23
Ill-defined mass	7	7	14
Spiculated mass	11	8	19
Architectural	9	10	19
Asymmetry	6	9	15
Normal	-	-	207
Total	**64**	**51**	**322**

The choice of the MIAS dataset was based on its public accessibility and online availability making it a reliable source of mammogram images. It also provides metadata corresponding to the background tissue, class, severity of the abnormality as well as coordinates to the center of the abnormality, and the approximate radius of a circle enclosing it. Such coordinates and radius allow the extraction of ROI's from images. It is also an old dataset published in 1994 with very little updates [253] therefore, the images are in their original form. However, a major challenge in using the MIAS dataset was that images are in 8 bits which compromises image quality. To ensure the quality of the images was improved, preprocessing by removing noise, artifacts and image enhancement was

performed. A detailed explanation of the preprocessing procedure is presented in section 3.4.2 of this thesis. Also, the MIAS dataset has very few images to warrant effective pattern learning. As a measure to facilitate better training, the number of images in the dataset was increased through the process of data augmentation discussed in section 3.4.3 of this thesis.

3.4 Experimental Setup

This section outlines the experimental simulation-based setup used in this research. Areas covered include: a description of the dataset used, preprocessing techniques applied, feature extraction technique employed, Classifiers, and performance evaluation measures applied. Figure 3.2 shows the steps in the experimental setup.

```
┌─────────────────┐      ┌─────────────────┐      ┌─────────────────┐
│ Read  mammogram │ ───> │ Mammogram image │ ───> │ Data Augmentation│
│ image           │      │ Pre processing  │      │                 │
└─────────────────┘      └─────────────────┘      └─────────────────┘
                                                           │
                                                           v
                                                  ┌─────────────────┐
                                                  │ Feature extraction│
                                                  └─────────────────┘
                                                           │
                                                           v
                                          ┌─────────────────────────────┐
                                          │ Classification,  Validation and│
                                          │ Performance Evaluation       │
                                          └─────────────────────────────┘
                                               │                    │
                                               v                    v
                                      ┌──────────────┐      ┌──────────────┐
                                      │  Abnormal    │      │   Normal     │
                                      └──────────────┘      └──────────────┘
                                         │        │
                                         v        v
                                  ┌──────────┐ ┌──────────┐
                                  │  Benign  │ │ Malignant│
                                  └──────────┘ └──────────┘
```
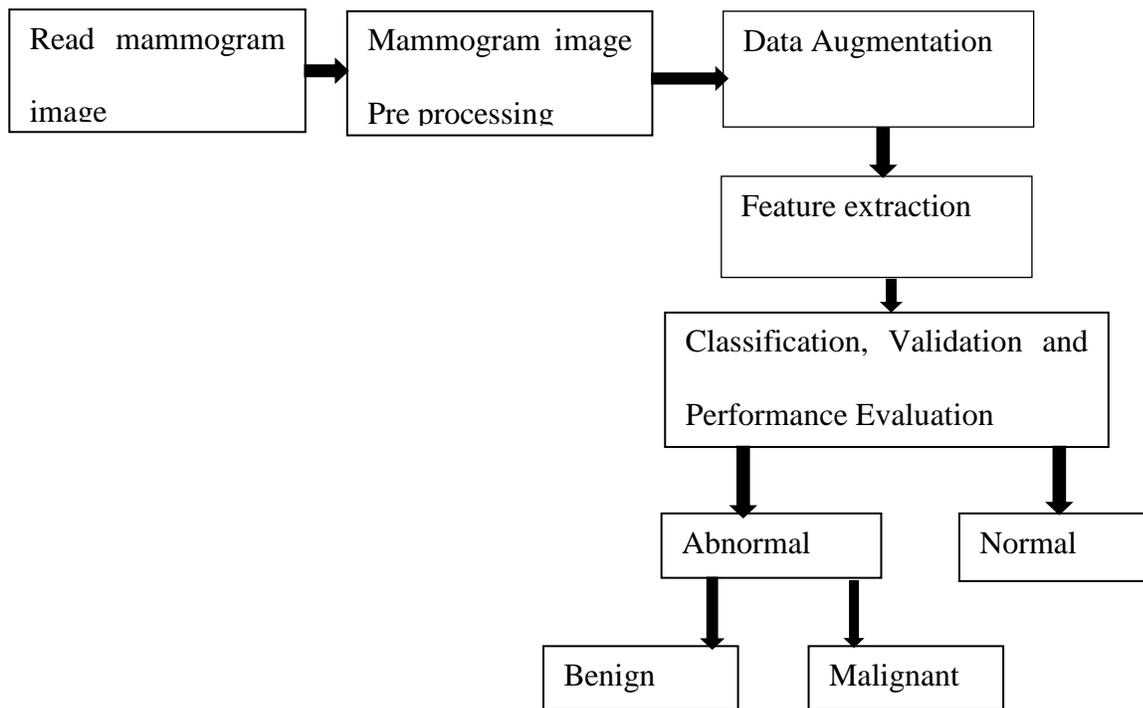
Figure 3. 2: Experimental setup of the research methodology

3.4.1 Reading a Mammogram Image

I took the input images used in this study from the MIAS dataset. First, I downloaded the dataset from Pilot European Image Processing (PEIPA) archives. The images are in Portable Gray Map (PGM) and categorized into Normal, Benign, and malignant classes. To allow classification into normal/abnormal and benign/malignant classes. The dataset has the images split into normal/abnormal classes which contained 322 images and benign/malignant classes which contained 115 images. I then read the images into the MATLAB environment using the *imread* function. Below is a brief explanation of the mammogram image reading process in MATLAB.

122

On the *command window* section, type:

img=imread ('mdb001.pgm') ;

imshow (img);

To read one image at a time and display it.

3.4.2 Mammogram Image Preprocessing

Mammogram preprocessing aims to improve image quality which avoids undue influences on the abnormality from the background of a mammogram [254]. This research preprocessed the image for artifacts and noise removal, mammogram cropping, and image enhancement. An Adaptive filter was used to remove artifacts and noise. Since the characteristics of mammographic images vary from one area to another and an adaptive filter can change with the varying characteristics of the mammogram, noise, and artifact. Once the noise and artifacts are removed, the mammogram images were enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE) technique.

The CLAHE technique was initially developed to minimize the shadow of edges and noise emitted in a homogeneous area of a medical image [256]. The technique has been used for enhancing digital mammogram [256] [254], [257], [258], and has shown good improvement to mammogram visual quality. An *N x N* input image *I* is split into small blocks. CLAHE is then applied to every block to enhance contrast. Finally, using bilinear interpolation, the individually enhanced small neighboring blocks are joint together to form the entire image. The choice of CLAHE technique was based on the advantage that it

123

can perform two in one operation of noise reduction and contrast enhancement, it prevents edge shadowing effect on mammogram making the edges clearer and prevents over-amplification of noise that may be present in the image especially in a homogeneous area [141] [259]. Finally, the enhanced image is cropped. The intension of cropping the images was to remove undesirable parts of the mammogram and extract the region of interest [255]. Cropping was done using the *imcrop* MATLAB built-in function. The algorithm below presents the entire procedure for preprocessing a mammogram image:

Input: Mammogram image

Output: Preprocessed mammogram image

1. Step 1: Using MATLAB, open the image batch processor App by selecting the App icon from the apps gallery

2. Step 2: Load mammogram image into the Image batch processor App by specifying the name of the folder containing the images

3. Step 3: Using the *new function* specify the Wiener-2 adaptive filter function to remove noise and artifacts

4. Step 4: Run the noise and artifact removal function on all the images

5. Step 5: Obtain the results of the Batch processing operation and save the results into a file

6. Step 6: Load the file containing the noise and artifact removed mammogram images into the Image batch processor

7. Step 7: Enhance the image by applying the CLAHE technique using MATLAB *adapthisteq* inbuilt function

8. Step 8: Run the *adapthisteq* inbuilt function on all the images

9. Step 9: Obtain the results of the Batch processing operation and save the results into a file

10. Step 10: Open the file containing the enhanced mammogram images

11. Step 11: To perform manual cropping and obtain the Region of Interest, use the crop image tool

12. Step 12: Specify the optional return value *rect* in which *imcrop* function returns the four elements position vector of the rectangle drawn

13. Step 13: Move the cursor over the image and draw a rectangle over the portion of the image you want to crop

14. Step 14: Perform the crop operation by double-clicking in the crop rectangle or select *Crop image* on the context menu

15. Step 15: The crop image tool returns the cropped area

16. Step 16: Save the cropped image

17. Step 17: Repeat step 10 to 16 for all mammogram images

3.4.3 Data Augmentation

Data augmentation is a strategy used by practitioners' to artificially increase the size of training samples [260]. The importance of data augmentation is to increase samples for

training a model without collecting new data. Mirroring and cropping are the most commonly used data augmentation techniques because they do not alter the image. In image mirroring, they represent the image data in different orientations resulting in copies of the same image but from different perspectives or visual angles [261]. It can mirror an image along the x, y, -x, -y-axis. Cropping entails getting copies of the same image by focusing on different areas of the same image, which results in a reduced image [261].

The dataset used in this research contains 322 images which are not sufficient for learning and generalization of a model. Therefore, to increase the size of the dataset, a data augmentation process was used. The number of images was increased by mirroring the images along the x and y-axis. Each image is transformed into three different images: the original image, the image mirrored along the x-axis, and the image mirrored along the y-axis. These ensure model generalization since the images are increased from 322 to 966 images. Table 3.2 shows the number of images after mirroring the images along the x and y-axis. The algorithm below presented the data augmentation process.

Input: Cropped image

Output: augmented image

1. Step 1: Load the folder containing the cropped images into the Image batch processor App by specifying the name of the folder

2. Step 2: Create the *flip-down function* on the image batch processor App using the *new* function

3. Step 3: Run the *flip-down function* to mirror the cropped images along the x-axis

4. Step 4: Store the flipped down images

5. Step 5: Load the folder containing the cropped images into the Image batch processor App by specifying the name of the folder

6. Step 6: Create the *flip right function* on the image batch processor App using the *new* function

7. Step 7: Run the *flip right function* to mirror the cropped image along the y-axis

8. Step 8: Store the flipped right images

Table 3. 2: Number of images after data augmentation

Class	Benign	Malignant	Total
Microcalcification	36	39	75
Circumscribed mass	57	12	69
Ill-defined mass	21	21	42
Spiculated mass	33	24	57
Architectural	27	30	57
Asymmetry	18	27	45
Normal	-	-	621
Total	**192**	**153**	**966**

3.4.4 Mammogram Feature Extraction

Feature extraction aims to transform a feature set from high-dimensional space to lower-dimensional space, which is a more convenient representation of the actual feature set [191]. The goal of feature extraction is to represent an input image to its reduced form, which is a more convenient representation of the image to facilitate effective classification. As discussed in the literature mammogram feature extraction techniques include; Texture, Shape, and Hybrid. This research work used Local Direction Patterns (LDP) and Local Ternary Pattern (LTP) texture features extraction techniques to come up with a new Local texture descriptor termed as Local Directional Ternary Pattern (LDTP). The algorithm below presents a brief explanation of the Local Directional Ternary Pattern (LDTP) descriptor.

Input: 3 by 3 image region

Process: Run LDTP descriptor

Output: LDTP encoded image region

1. Step 1: Read a 3×3 mammogram image region

2. Step 2: Calculate the absolute differential value between the neighboring pixel and the central pixel using equation 4.1 in this thesis

3. Step 3: Convolve the resultant absolute differential matrix with Kirsch mask

4. Step 4: Perform mini max normalization using equation 4.2 in this thesis

5. Step 5: Perform softmax normalization using equation 4.3 in this thesis

6. Step 6: Calculate the upper and lower threshold values using equation 4.4 and equation 4.5 respectively

128

7. Step 7: Using the threshold values obtain a Local Direction Ternary Pattern code using equation 4.6 in this thesis

8. Step 8: Split the ternary pattern into a positive and negative pattern using equation 4.7 in this thesis

9. Step 9: Concatenate the eight bits into a binary number for the positive pattern

10. Step 10: Convert the binary number to its decimal equivalent and allocate it to the central pixel as its LDTP.

11. Step 11: Repeat step 1 to 10 for the entire region

12. Step 12: Save the positive LDTP pattern to a .csv file

3.4.5 Classification

To validate the effectiveness of the new local texture descriptor, I used two classifiers SVM and ANN. In this research work, I modeled the breast cancer classification problem as a binary classification problem performed on two levels. By using 966 mammogram images obtained after data augmentation, in the first level of classification, I used a classifier to distinguish between normal and abnormal classes. In the second level classification, 345 abnormal tumors were used for training the classifiers to categorize the tumor into a malignant or benign class. The algorithm below shows the process of classification.

1. Step 1: Using the *imageSet function* in MATLAB, recursively load the file containing the feature extracted images in the three categories; Normal, benign, and malignant

2. Step 2: From each category split the images into a training set and a test set using the 70: 30 ratio

3. Step 3: Load the .csv file for *Normal/ Abnormal class* of the extracted features into the classification learners App in MATLAB.

4. Step 4: Select the classifiers used for training and click on the *train button*

5. Step 5: To inspect the accuracy of the predictions in each class, click on the confusion matrix

6. Step 6: Export the model to make predictions with new data

7. Step 7: Load the .csv file for *Benign/ Malignant class* of the extracted features into the classification learners App in MATLAB.

8. Step 8: Select the classifiers used for training and click on the train button

9. Step 9: To inspect the accuracy of the predictions in each class, click on the confusion matrix

10. Step 10: Export the model to make predictions with new data

3.4.6. Validation and Evaluation Protocol

The model was validated using 10-fold cross-validation when selecting SVM and ANN parameters. In 10-fold cross-validation, it generates 10 equal subsets from the original sample and 10 iterations performed on training and testing set. For every repetition, it sets aside one different fold for testing and the remaining 9 folds for training.

The performance of the developed descriptor was measured using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) using a two-class confusion

matrix. As discussed in the literature metrics used to test performance include Sensitivity, Specificity, Precision, Accuracy, and F-score. In this study, the performance of classifiers was measured using Accuracy, Sensitivity, and Specificity. Accuracy is the simplest and most intuitive evaluation measure for classifiers. In this study, accuracy was used to aid comparison between the performance of the developed local descriptor and other existing descriptors, because most of the researchers tested the classifiers based on it, therefore it made performance comparison less challenging.

3.5 Review of Research Questions

Table 3. 3: Review of Research questions

Research Questions	Research Design
i. How do the existing techniques detect breast cancer??	Exploratory
ii. How can a Local Directional Ternary Pattern texture descriptor that considers all directional responses and an adaptive threshold when encoding image gradient be developed?	Simulated Experiment
iii. How valid is the developed Local Directional Ternary Pattern texture descriptor in breast cancer detection?	Simulated Experiment

3.6 Research Tools and Material

This study used the MATLAB software package for image processing. It has powerful and easy-to-use features, especially when dealing with images. I implement all algorithms

developed using MATLAB (R2018a) on a PC equipped with an Intel Core i5 processor, 8GB RAM, and Windows 10 Operating System. I based the choice of MATLAB on its capability to carry out complex computations; it provided a visualization environment and easy to use mathematical notations [262]. Further, it provided inbuilt toolboxes for image processing, computer vision, and batch processing which are appropriate for this research. A detailed explanation of the MATLAB software installation and activation procedure in section 3.6.1 of this thesis.

Getting actual images to carry out this research was highly difficult because of privacy and legal issues and technical hurdles. Therefore, the MIAS dataset was used to study the effectiveness of the developed local descriptor as it was a benchmark dataset available online for research. To use the MIAS dataset, I downloaded raw images. Section 3.6.2 of this thesis presents an explanation on the process for downloading the dataset.

3.6.1 MATLAB Installation and Activation procedure

To install MATLAB R2018a, insert the MATLAB flash drive then open the *bin\win* folder and double click on the setup then follow the steps given below.

Step 1: Select the installation method by logging into your account if you have a MathWorks account or use a file installation key. In this case, I used a file installation key.

Step 2: In the MathWorks installer window, select the button *Next* to use a File installation key. After the selection of " *Use a file installation key*", click the *Next* button.

Step 3: In the Licence agreement window, select the *Yes* button and then accept the terms of the license agreement by clicking on the *Accept* radio button. Then select the *Next* button.

Step 4: In the File installation Key window, click the radio button on "*I have the activation key for my license*" then enter the file installation key found in the *readme.txt* file. Then click the *Next* button.

Step 5: In the folder selection window, select the place where you want to install the MATLAB software then click the *Next* button.

Step 6: In the product selection window, select all the products to be installed by clicking on all the checkboxes then at the end click the *Next* button.

Step 7: In the license file browser, select the license file and click *Next*.

Step 8: In the installation window, choose the *desktop* option to install a shortcut then click the *Next* button.

Step 9: In the confirmation window, click the *Install* button, when the installation is complete, click on the *finish* button

Once MATLAB R2018a was installed successfully, it required activation. There are two options for activation. Either to login to MathWorks Account or to enter the full path to license file. Since I had the license file, I used the option of using a license file. Therefore to activate the MATLAB R2018a the following steps were followed;

Step 1: On the provide License file window, click the radio button beside the option for "*Activate automatically using the Internet*" and then click the *Next* button.

Step 2: Click on browse and locate the license file.

Step 3: Once you provide a valid license file, you will be prompted with a success message of "*Activation Complete*".

3.6.2 MATLAB Image processing App

After installing and activating the MATLAB R2018a software two major applications; the image batch processor App and the classification learners' App tools were used in this research software.

The image batch processor App was used to read in images that needed to be preprocessed. The preprocessing functions carried out in this study included image noise removal and image enhancement. To access the image batch processor App follow the steps outlined below.

Step 1: On the *Apps* tab, in the image processing and computer vision toolbox, click *Image Batch Processor*.

Step 2: Load the images into the image batch processor app by clicking on *Load images*.

Step 3: In the *load images* section, specify the folder containing the images you want to load. By default, the app includes images in subfolders. The Image batch processor app creates thumbnails of the images loaded

Step 4: Specify the name of the function to use by typing the name in the *function name box* in the batch function section.

Step 5: Test the new function by running the batch processor on one of the images

Step 6: Once the results of the test are successful, then execute the function on all the other images

Step 7: To save the results, click on the *Export* tab. You can export the results to a file or workspace.

3.6.3 MATLAB Classification Learners App

Another App that was used in this study was the classification learners App. This App was used to train the SVM classifier. To access and use the classification learners App in MATLAB follow the steps given below.

Step 1: In the Apps tab, click on the *classification Learner* section or select it from the machine learning toolbox

Step 2: Click *New Session* and import your data from the workspace or from a file

Step 3: If you import your data from a file select the file, wait for it to loaded then click on *import selection* tab

Step 4: Specify a response variable and predictor variables

Step 5: Choose the type of validation under the *validation* section, then click on *start session* tab

Step 6: You can visualize the distribution of your data by looking at the scatter plot

Step 7: On the model type section select the classifiers

Step 8: Click the *train* tab to train your data based on the classifier select in step 7

Step 9: View the results of the training in the *history list* window.

Step 10: You can further analyze the results of the model by looking at the *confusion matrix* and the *ROC curve.*

3.6.4 Procedure for downloading the mammogram images from the MIAS dataset

The MIAS dataset contains 322 images out of which 270 are non-cancerous and 52 are cancerous was used in this study. The dataset is available at https://www.repository.cam.ac.uk/handle/1810/250394 The images have 50 μm per pixel and clipping/padding to a fixed size of 1024×1024 pixels in Portable Gray Map (PGM) format and associated truth data [63]. The procedure for downloading the images was described as follows:

Step 1: Type the URL https://www.repository.cam.ac.uk/handle/1810/250394 on browser

Step 2: Click on *view/open* files icon on the Mammographic Image Analysis Society (MIAS) database v1.2.1, then wait for the download to complete

Step 3: Extract the images to a specified folder

3.7 Ethical Consideration

The researchers understand that mammographic images contain sensitive information and that they are properties of the research community. The researcher exercised utmost confidentiality by using the mammographic images purely for academic research. Since the

mammographic images were anonymously coded with no patient demographic information, they could not be traced back to individual patients. Also, an effort was made to source for research permit approval from the National Commission of Science Technology and Innovation (NACOSTI) before proceeding with data collection.

3.8 Summary

This section began by demystifying the research philosophy adopted and the research process followed using Saunders's research onion. It also explained the experimental setup for the development of the new local descriptor citing the process of mammogram image reading, mammogram image preprocessing, data augmentation, feature extraction, classification, and performance evaluation of the classifiers. A review of research questions, research tools and material used, model validation, and ethical consideration during the research process was also presented in this section.

CHAPTER FOUR

LDTP TEXTURE DESCRIPTOR

4.1 Overview

This chapter explains the process of developing the LDTP texture descriptor based on the algorithm outlined in section 3.4.4 of this thesis. The chapter explains the image gradient encoding process in two major steps; (1) computing the edge responses for all the eight directions as explained in section 4.2.1 and (2) calculating an adaptive threshold as explained in section 4.2.2 of this thesis.

4.2 Local Directional Ternary Pattern (LDTP) Texture Descriptor

The state of the art LTP descriptor uses a static threshold τ defined by a user for all datasets or images in a dataset. The chosen threshold is invariant to grey-value variations, static and lacks a defined way of selecting an optimum value for the threshold. The Local Directional Patterns (LDP) only considers top k directional responses when encoding local texture in an image. Although the difference between a reference pixel and its neighbors derives an image gradient, it does not consider the central pixel, thereby deteriorating the discriminative power of the features extracted. Further, LDP disregards $8-k$ responses, which leads to a loss of subtle texture features. Herein, a Local Directional Ternary Patterns (LDTP) texture descriptor is developed. Unlike LDP and LTP, the LDTP descriptor considers the central pixel, takes into account all directional responses, and uses two adaptive threshold values for a 3×3 image regions when deriving an image gradient.

138

The LDTP applies the Kirsch mask to compute responses for all the eight directions for the central reference pixel in a 3×3 image region. It uses the sign of the directional responses to increase the discriminative nature of the encoded image gradient.

4.2.1 Computing directional responses

Given a 3×3 image region, LDTP first finds the absolute difference in grey-values between neighboring pixels and central reference pixel as defined by equation (4.1)

$$P_{i,j} = \left|P_{i,j} - P_c\right| \tag{4.1}$$

Where $P_{i,j}$, is the gray value at row i, column j, and P_c is the gray value of the central pixel. The absolute differences are then convolved with the Kirsch masks to get directional responses. Figure 4.1 illustrates a 3×3 image region, the absolute difference in grey values, and corresponding directional responses.

85	32	26
53	50	10
60	38	45

(a)

35	18	24
3	0	40
10	12	5

(b)

7	175	215
-57	0	111
-241	-225	15

(c)

Figure 4. 1: LDTP process (a) Image region (b) Absolute differential values (c) Directional responses using kirsch masks

While the original LDP descriptor picks the top three directional responses, Here we consider all the eight directional responses because each response carries important information and therefore should not be ignored.

The directional responses are then normalized using the min-max normalization technique as shown in equation (4.2)

$$xi^{norm} = \frac{x_i - min}{Max - Min} \tag{4.2}$$

Where $x_{i,j}$ is the absolute response at index i, j and *Max* and *Min* are the maximum and minimum responses respectively, and $x_{i,j}^{norm}$ is the normalized value for $x_{i,j}$ responses.

The normalized values are in the range of $0\ldots1$, but they add up to a value of more than 1. However, to ensure the values add to a maximum value of 1, the normalized values are passed through a softmax function given by Equation (4.3) and illustrated in Figure 4.2(b) in comparison to the min-max technique in figure 4.2 (a).

$$P_{i,j} = \frac{e^{x_{i,j}^{norm}}}{\sum e^{x_{i,j}^{norm}}} \tag{4.3}$$

Where $P_{i,j}$ is the likelihood for the presence of an edge towards a given direction, $e^{x_{i,j}^{norm}}$ is the exponential value of the normalized absolute responses at index i,j and $\sum e^{x_{i,j}^{norm}}$ is the summation of all the exponential values.

0.5318	0.9091	1.0000
0.3818	0	0.7636
0.0364	0.0000	0.5455

0.1192	0.1739	0.1904
0.1026	0	0.1503
0.0726	0.0700	0.1209

(a) (b)

Figure 4. 2: Normalization process with (a) Min-Max Normalization and (b) Soft-max

technique

4.2.2 Calculating an adaptive threshold

Once the normalization is done, the probability space is split into three parts for -1,0,+1 bits for the generated ternary pattern and calculates two thresholds, T_p and T_n as shown in equation (4.4) and (4.5);

$$T_P = \frac{1}{4} \sum_{i=0}^{3} L_i^T \qquad (4.4)$$

$$T_n = \frac{1}{4} \sum_{i=0}^{3} L_i^B \qquad (4.5)$$

Where L_i^T is the ith top likelihood, and L_i^B is the ith bottom likelihood. For the likelihoods shown in Figure 4.2 (b), the thresholds are $T_p=0.1589$ and $T_n =0.0911$, respectively.

Where T_p is calculated as the average of the four highest values given by 0.1739, 0.1904, 0.1503, and 0.1209, while T_n is the average of the four lowest values given by 0.1192, 0.1026, 0.0726 and 0.0700

LDTP code is calculated as shown in equation (4.6);

LDTP code is calculated as;

$$\text{LDTP} = \sum_{i=0}^{7} p(L_i - (T_p|T_n)) \times 2^i) \qquad (4.6)$$

$$p(\dots) = \begin{cases} 1, & \text{if } L_i - T_p \geq 0 \\ -1, & \text{if } L_i - T_n \leq 0 \\ 0, & \text{Otherwise} \end{cases}$$

Where L_i is a likelihood i as shown in Figure 4.2 (b)

For the likelihoods in Figure 4.2 (b), Figure 8 shows the resultant ternary pattern and the corresponding positive and negative LTP codes. The LDTP code is split into a positive and negative LTP code, as shown in Figure 4.3(b) and Figure 4.3(c), respectively.

0	+1	+1
0		0
-1	-1	0

(a)

0	1	1
0		0
0	0	0

(b)

0	0	0
0		0
1	1	0

(c)

Figure 4. 3: Resultant Ternary Pattern code with (a) Resultant LDTP code (b) Positive

LDTP code (c) Negative LDTP code

The positive and negative gradients of an $M \times N$ image shown in Figure 4.3 (b) and Figure

4.3 (c) are represented as H_p and H_n and defined as expressed in eq. (4.7);

$$H_p(i) = \sum_{m=0}^{M} \sum_{n=0}^{N} f\big(LDTP_p(m,n), i\big) \qquad (4.7)$$

$$H_n(i) = \sum_{m=0}^{M} \sum_{n=o}^{N} f(LDTP_n \ (m,n), i)$$

$$f(x, i) = \begin{cases} 1, & \text{if } x = i \\ 0 & \text{Otherwise} \end{cases}$$

Where f (x, i) is a logical function that compares if the LDTP code at location x (m, n) of

the LDTP encoded image is equal to current LDTP pattern i for all i in the image$0 \le i \le$

C_k^8.The resultant histogram has dimensions$1 \times C_k^8$, which represent the image. The two

histograms are then fused and used for pattern recognition application.

4.3 Summary

This chapter explained the process of developing the LDTP texture descriptor based on the LDTP algorithm defined in section 3.4.4 of this thesis. The development of the LDTP texture descriptor was explained in two steps of calculating the edge responses for all the eight directions as explained in section 4.2.1 and calculating an adaptive threshold as explained in section 4.2.2 of this thesis.

CHAPTER FIVE

EXPERIMENTAL VALIDATION, RESULTS AND DISCUSSION

5.1 Overview

This chapter elucidates and discusses the experimental validation in terms of the experimental setup defined in section 3.4. It provides details on how the mammogram images were read into the MATLAB environment, how the images were preprocessed to improve their quality, how data augmentation was performed to increase the number of images, how extraction was performed using the developed LDTP texture descriptor, how the classification was performed using SVM and ANN classifiers and validation of each of the classifiers into Normal/abnormal class and benign/malignant class using the breast cancer dataset. Also the results achieved by the developed LDTP texture descriptor in terms of performance evaluation of LDTP descriptor using ANN and SVM classifiers, Accuracy comparison of LDP, LTP, LDTP descriptors using SVM and ANN classifiers, sensitivity and specificity comparison between SVM and ANN classifiers, the statistical significance of the achieved accuracy using Wilcoxon signed-rank test. In addition a detailed discussion on classification comparison between LDTP with related studies based on classifying the breast cells into Normal/Abnormal and Benign/Malignant classes is presented.

5.2 Experimental Validation of Local Directional Ternary Pattern (LDTP)

Steps of experimental validation of the LDTP descriptor are as shown in Figure 3.2. First, a mammogram image is read from the MIAS dataset. Then, the read image undergoes preprocessing which involves noise and artifacts removal, image enhancement, and cropping. Then, data augmentation is performed to increase the number of images that are then passed to the LDTP descriptor for feature extraction. Then classification is performed using SVM and ANN classifiers using the MIAS dataset. The performance of LDTP is then evaluated using accuracy, sensitivity, and specificity metrics. A detailed explanation of all the steps followed during the experimental validation is given below.

5.2.1 Mammogram Image Acquisition and Reading

The study used mammogram images taken from the MIAS dataset. The images are split into malignant, benign, and normal classes each containing 51, 64, 207 images respectively. Figure 5.1 shows three sample images from each of the categories in the MIAS dataset.


| (a) | (b) | (c) |
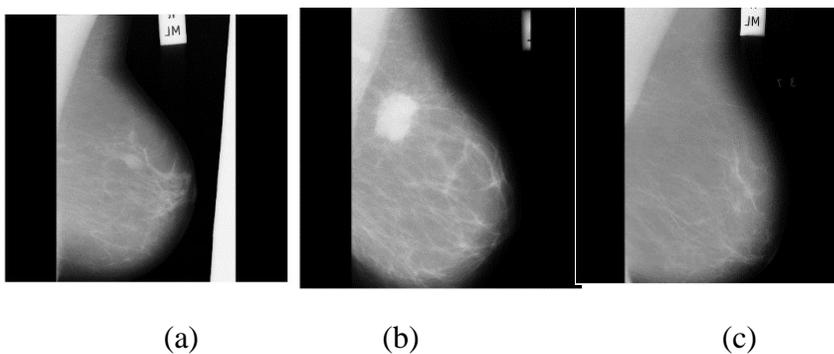
Figure 5. 1: Sample MIAS images (a) mdb012 benign image (b) mdb184 malignant image

(c) mdb006 Normal image.

146

To load the images into the MATLAB environment, store the images into three separate folders designated for benign, malignant, and normal cases. Use the image batch processor App under the image processing and computer vision toolbox to load the images into MATLAB. Figure 5.2 shows how images are loaded through the image batch processor app.
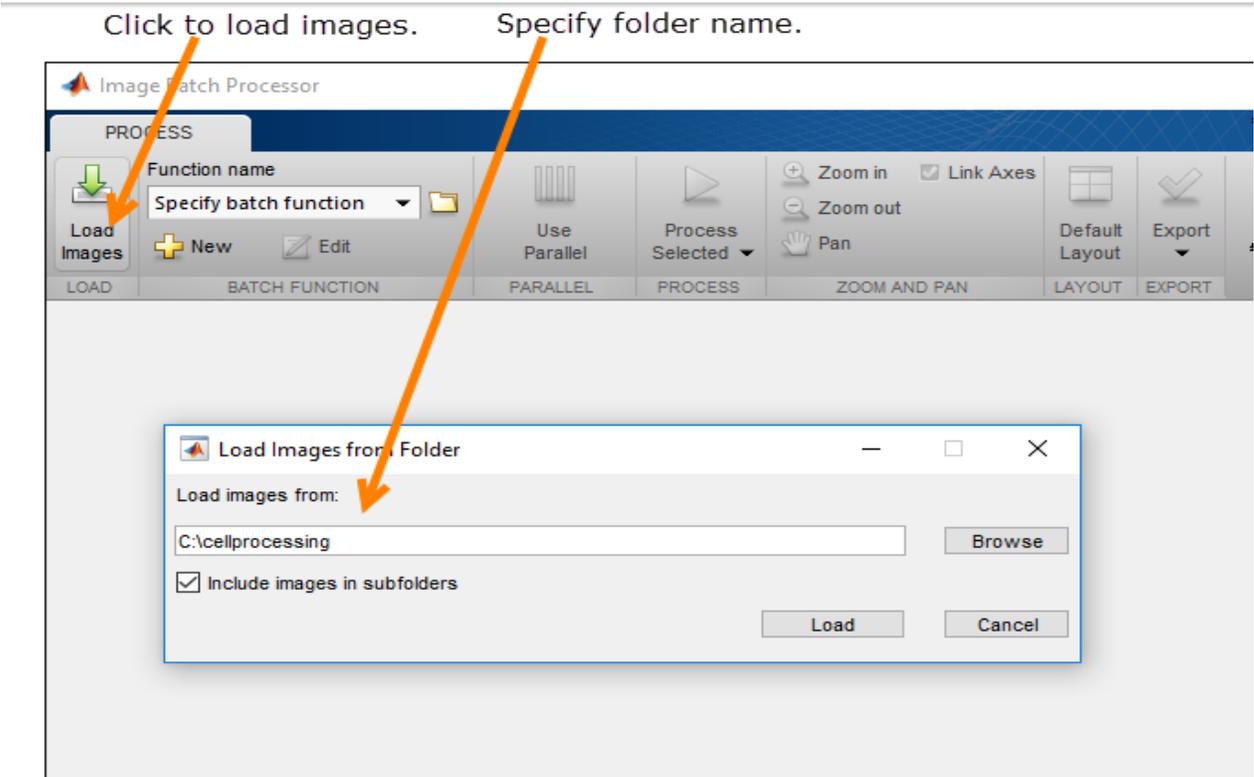

Figure 5. 2: The process of loading images into the image batch processor app

147

5.2.2 Mammogram Preprocessing

The purpose of preprocessing a mammogram is to improve its quality by removing artifacts, labels, and noise that may be present in the mammogram image. Also, through enhancement techniques image contrast is enhanced. In this research, preprocessing was done in two stages. Stage one consisted of filtering noise, removing artifacts, and enhancing the mammogram images. The wiener-2 adaptive filter function is applied to filter noise and eliminate artifacts and labels. To enhance the mammogram images CLAHE technique is applied. The merit of this technique is that it does not amplify any noise present in a mammogram image.

The second stage of preprocessing entails image cropping. Cropping is performed to localize the affected breast region without being influenced by the unwanted region. Because of the different orientation and position of a breast tumor, the images were manually cropped.

To prepare the mammogram images for feature extraction, each of the images underwent noise removal, enhancement, and cropping. The output of one stage became the input for the next stage of preprocessing. .Figure 5.3: shows sample images of the input images and the preprocessed images from each of the three categories.
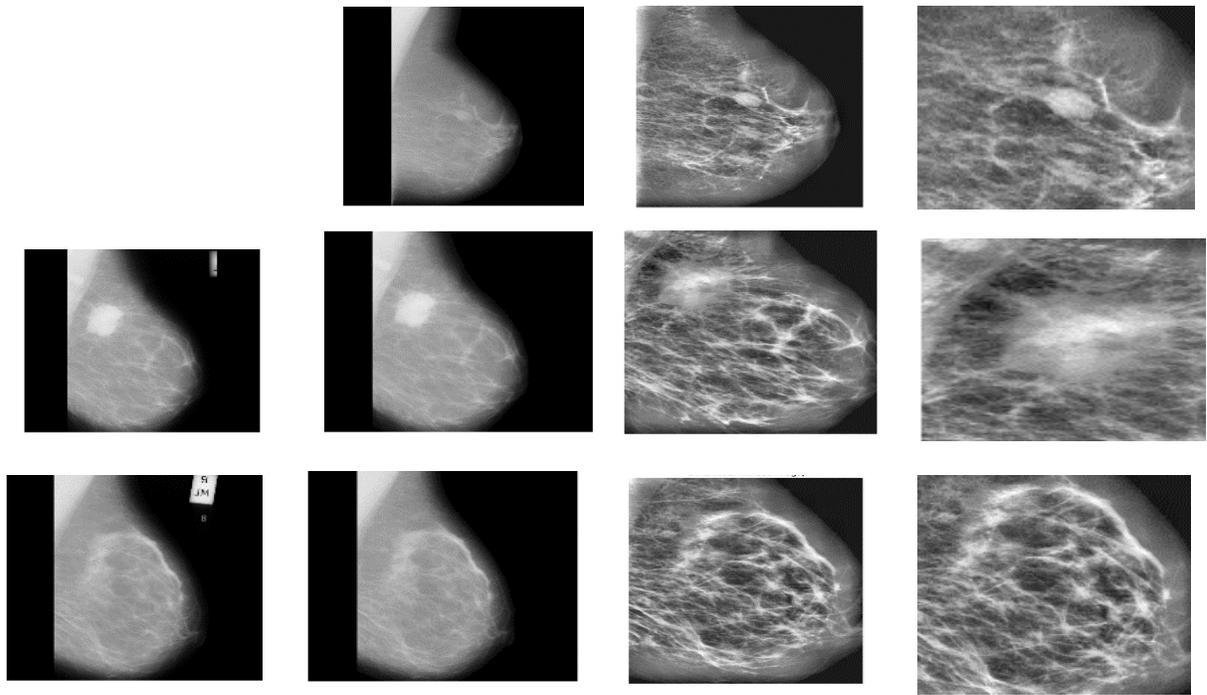
(a)   (b)   (c)   (d)

Figure 5. 3: preprocessed Benign, Malignant and Normal images (a) Input image (b)

artifact and noise removed image   (c) enhanced image (d) Cropped image

5.2.3 Data Augmentation

The MIAS dataset used in the study has 322 images which are very few to warrant effective pattern learning, therefore as a measure to facilitate better learning the number of images was increased through the process of data augmentation. Each mammogram image was mirrored along the x-axis and the y-axis. Figure 5.4 shows a mammogram image mirrored along the x and y-axis respectively.

149

|		
(a)	(b)	(c)
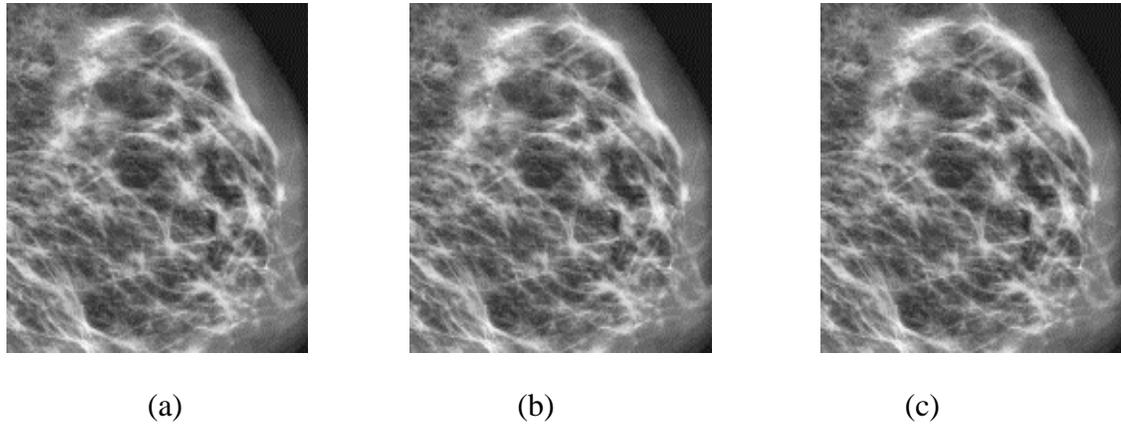
Figure 5. 4: mdb184 malignant image mirrored on x and y-axis (a) preprocessed mdb184 malignant image (b) mdb184 malignant image mirrored along the x-axis (c) mdb184 malignant image mirrored along the y-axis

5.2.4 Mammogram Feature Extraction using LDTP descriptor

In this research, a new Local Directional Ternary Pattern (LDTP) texture descriptor that considers all directional responses and an adaptive threshold when encoding image gradient was developed for breast cancer classification. The mathematical explanation of how the texture descriptor was developed was presented in section 4.2 of this thesis. Figure 5.5 shows a sample preprocessed benign image with its corresponding LDTP extracted image and histogram.

(a)   (b)   (c)
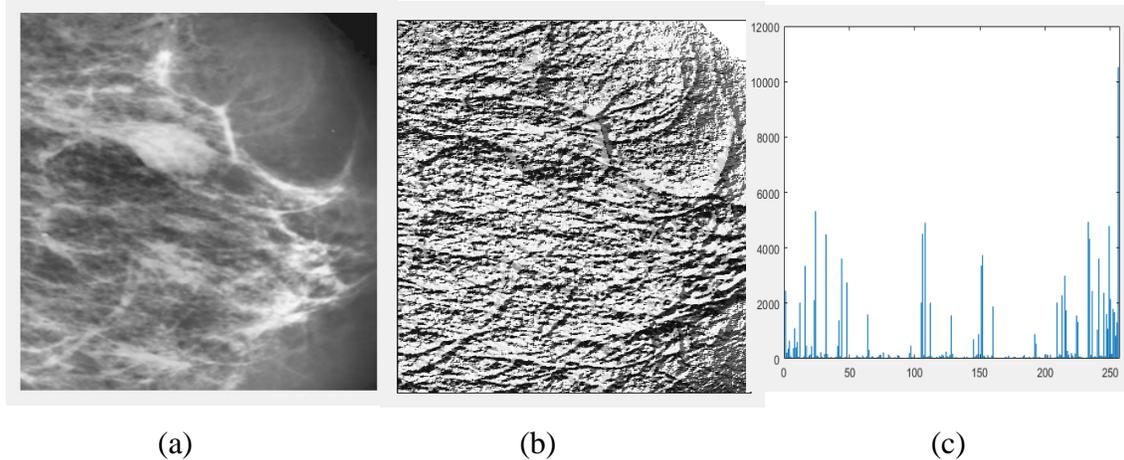
Figure 5. 5: LDTP images for Mdb012 benign image (a) Input mammogram image (b) LDTP mammogram image (c) Histogram for the LDTP image

5.2.5 Classification

The classification stage was conducted in two levels using a total of 966 instances. Out of 966 instances, 621 were normal and 345 were abnormal. The first level of classification aimed to differentiate a normal instance from an abnormal instance. Here, all the 966 instances were used with 30% of the instances used for testing. Therefore a total of 290 instances were used for testing from which 186 instances belonged to the normal class while 104 instances belonged to the abnormal class. The classification process was carried out using SVM and ANN classifiers, from which confusion matrices were generated. Table 5.1 and Table 5.2 shows the confusion matrix for normal /abnormal classification using ANN and SVM classifiers respectively.

151

Table 5. 1: Confusion matrix for normal/abnormal classification using ANN classifier

Class	Normal	Abnormal
Normal	182	4
Abnormal	4	100

Table 5. 2: Confusion matrix for normal/abnormal classification using SVM classifier

Class	Normal	Abnormal
Normal	185	1
Abnormal	5	99

In the second level of classification, the aim was to distinguish between a benign and a malignant tumor, a total of 345 instances were used, out of which 192 were benign and 153 were malignant. From each class, 30% of the instances were used for testing. Therefore, 58 instances were selected from the benign class, and 46 instances were selected from the malignant class. Table 5.3 and Table 5.4 shows the confusion matrix for benign/malignant classification using ANN and SVM classifiers respectively.

Table 5. 3: Confusion matrix for benign/malignant classification using ANN classifier

Class	Benign	Malignant
Benign	55	3
Malignant	6	40

Table 5. 4: Confusion matrix for benign/malignant classification using SVM classifier

Class	Benign	Malignant
Benign	56	2
Malignant	4	42

5.2.6 Performance Validation and Evaluation

To assess the performance of the developed LDTP descriptor, its performance was evaluated using SVM and ANN classifiers. To show the robustness of the developed LDTP descriptor, an accuracy metric was used to evaluate its classification performance against LDP and LTP descriptors. To understand the effectiveness of SVM and ANN classifiers to correctly classifying a breast tumor, sensitivity, and specificity measures were calculated. Further, to show the statistical significance of the achieved accuracy the Wilcoxon signed-rank test was conducted.

5.2.6.1 Performance evaluation of LDTP descriptor using ANN and SVM classifiers

Two levels of classifications were performed to determine the accuracy of ANN and SVM classifiers for the LDTP descriptor. The first level of classification determined if the tumor was normal or abnormal, while the second level of classification determined if the tumor was benign or malignant. Figure 5.6 shows the results using accuracy measure when LDTP was passed through SVM and ANN classifiers. The figure depicts that SVM achieved higher accuracy than ANN with 97.32% and 93.93% for the normal/ abnormal and benign/malignant classes respectively.
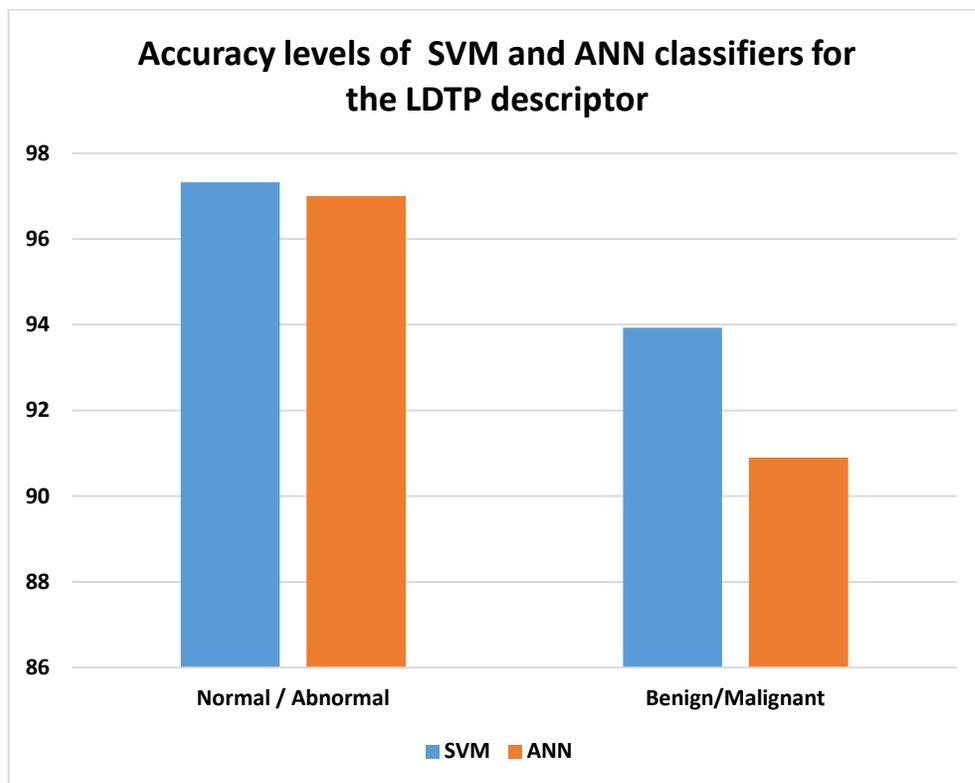

Figure 5. 6: Accuracy levels of SVM and ANN classifiers for LDTP descriptor

5.2.6.2 Accuracy levels comparison of SVM and ANN classifiers for LDP, LTP, and LDTP descriptor

The accuracy levels of SVM and ANN classifier for the LDTP descriptor was compared against SVM and ANN classifiers' accuracy for LDP and LTP descriptors for the normal/abnormal and Benign/malignant classes. In both cases, LDTP performed better than LDP and LTP with accuracy levels of 97% and 90.90% for the normal/abnormal and benign/malignant class. Figure 5.7 and Figure 5.8 shows the comparisons using ANN and SVM classifiers.
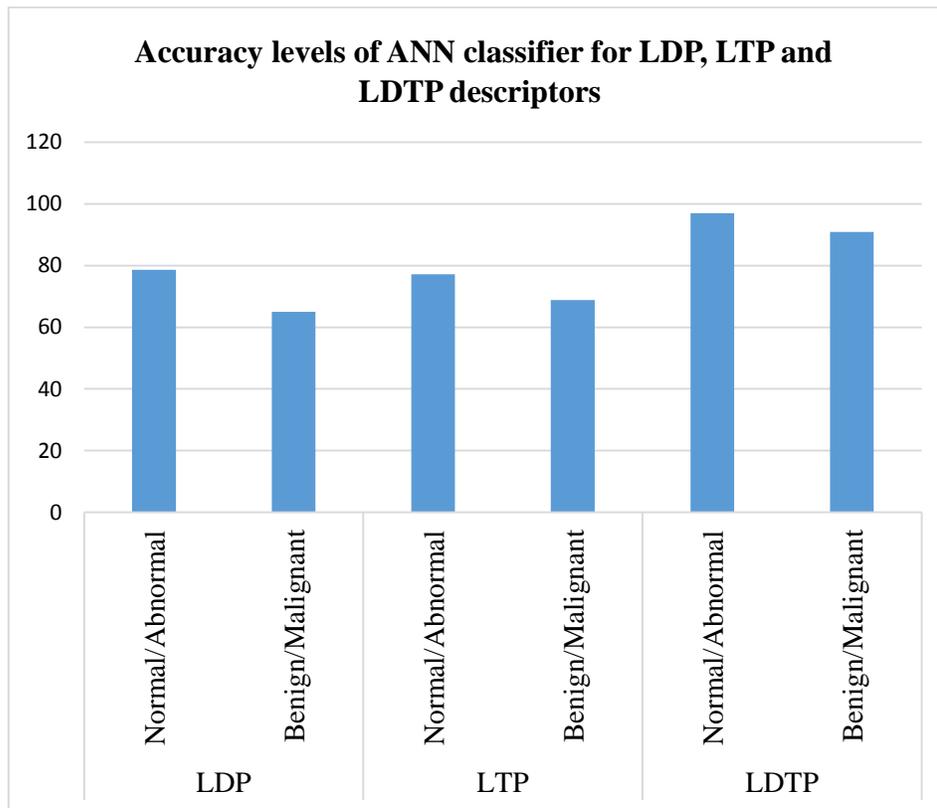

Figure 5. 7: Accuracy comparison of ANN classifier for LDP, LTP, and LDTP descriptors

Accuracy levels of SVM classifier for LDP, LTP and LDTP descriptors
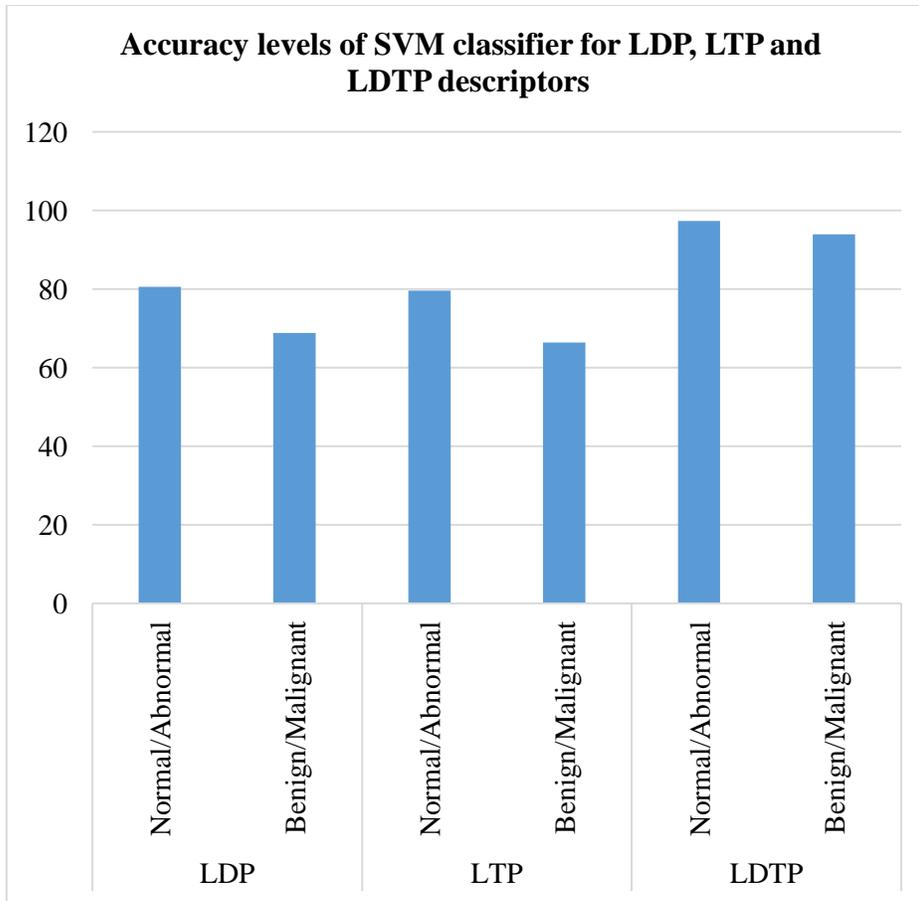
Figure 5. 8: Accuracy comparison of SVM classifier for LDP, LTP, and LDTP descriptors

5.2.6.3 Sensitivity and Specificity comparison of ANN and SVM classifiers for LDTP descriptor

To be able to determine the effectiveness of the SVM and ANN classifiers in identifying a normal, malignant or benign tumor, sensitivity and specificity measures were calculated. The results of the sensitivity test determined the percentage of normal tumors that were correctly classified as normal. Figure 5.9 shows the sensitivity comparison of SVM and ANN classifiers for the LDTP descriptor.

156
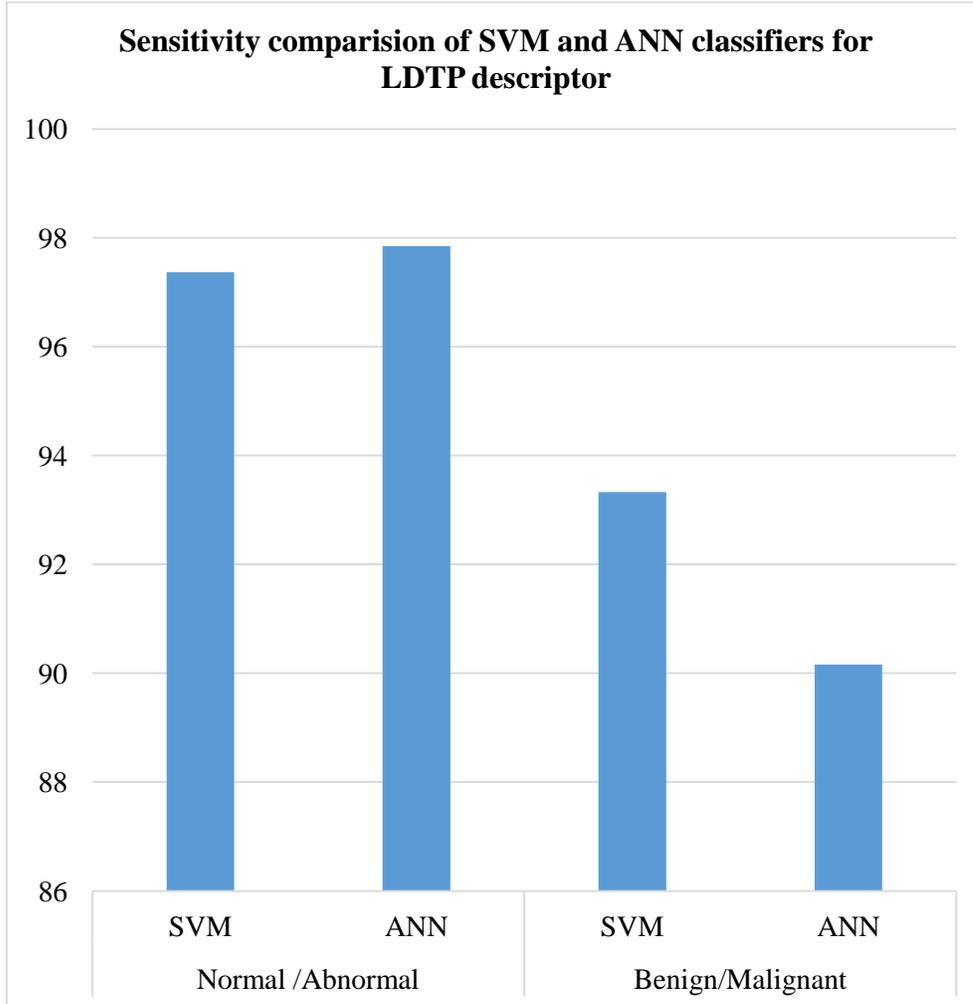
Figure 5. 9: Sensitivity measure of SVM and ANN classifiers for the LDTP descriptor

The specificity measure determined the percentage of abnormal tumors correctly classified as abnormal. Figure 5.10 shows a specificity comparison of SVM and ANN classifiers for the LDTP descriptor.
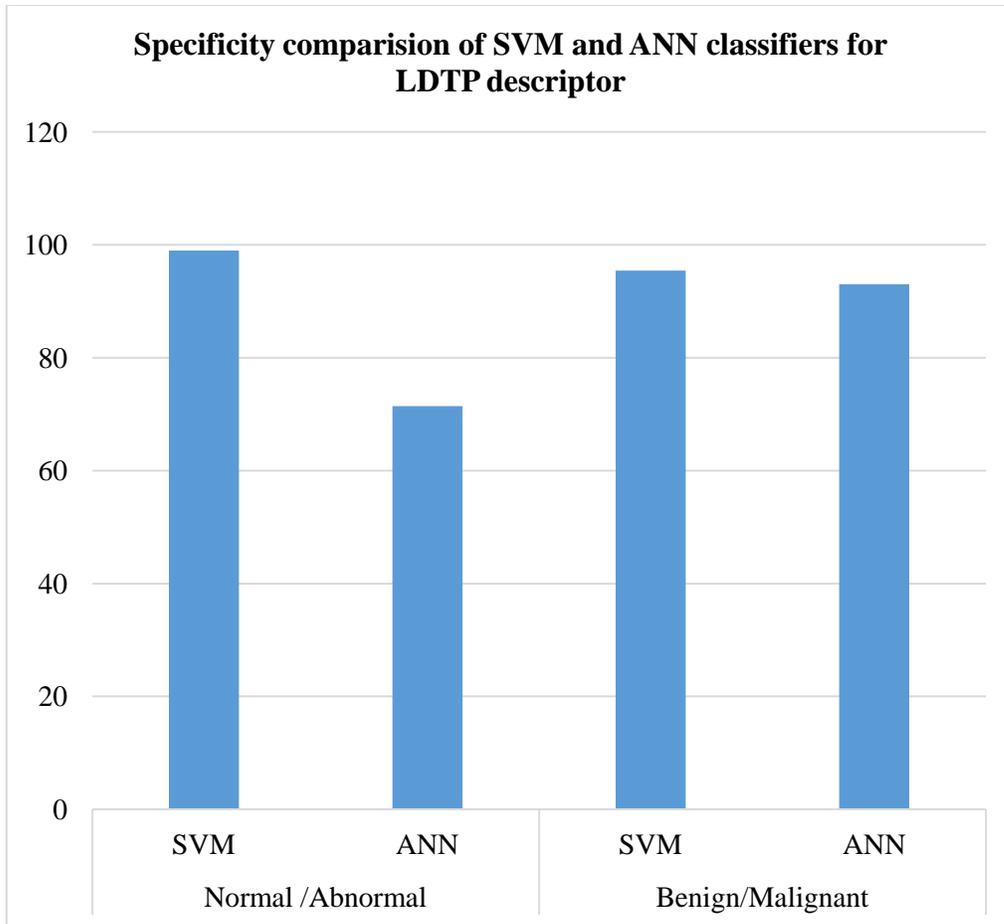
Figure 5. 10: Specificity measure of SVM and ANN classifiers for the LDTP descriptor

5.2.6.4 Statistical Significance of the achieved accuracy improvement

To test for statistical significance of the accuracy levels obtained by SVM classifiers for the LDTP descriptor against LDP and LTP descriptors, a Wilcoxon signed-rank test was used. The Wilcoxon signed ran test provided an estimate for acceptance and rejection of the significance level among the different descriptors used. In this study, the Wilcoxon signed-rank test is performed at a significance level of $\propto=0.05$ to analyze the relative

comparisons of pairs of descriptors. Let n_1 and n_2 be respectively the accuracy vectors of

two descriptors that are under comparisons. Each of the n_1 and n_2 vectors has 10 elements.

The difference scores $d= n_1 - n_2$ is calculated and if the test detects significance difference

we consider that one of the two descriptors outperforms (underperforms) the other.

The null hypothesis states that "there is no improvement in the accuracy levels between the

LDTP descriptor against the LDP descriptor and the LDTP descriptor against the LTP

descriptor". The rejection of the assumption would mean that there was a statistically

significant improvement of the accuracy achieved by the LDTP descriptor. The procedure

is performed for each pair of descriptors using MS-Excel software. Table 5.5 shows the

results of the Wilcoxon test for the LDTP descriptor against LDP and LTP descriptors at a

significance level of $\propto=0.05$.

Table 5. 5: The results of the Wilcoxon test for LDTP descriptor against LDP and LTP descriptors

Descriptor	Class	P-value	Null Hypothesis
LDTP vs LDP	Normal/Abnormal	0.000279953	Reject
	Benign/Malignant	0.000279953	Reject
LDTP vs LTP	Normal/Abnormal	0.000279953	Reject
	Benign/Malignant	0.000279953	Reject

5.3 Analysis of Results

The results of the experimental validation were evaluated based the confusion matrix values gotten after executing the LDTP descriptor using ANN and SVM classifiers. Also the results obtained were presented inform of bar graphs to give insight into the performance of the LDTP descriptor when compared to LDP and LTP descriptors based on Accuracy sensitivity and specificity measures. A detailed analysis of the results achieved is presented below.

5.3.1 Performance Analysis of the LDTP descriptor for the Normal/ Abnormal class

Table 5.1   shows the confusion matrix for the normal/abnormal classes using the ANN classifier. For the normal class, it could correctly classify 182 out of 186 instances, while for the abnormal class, it correctly classified 100 out of 104 instances resulting in a classification accuracy of 97.85% and 96.16% respectively. Table 5.2 depicted confusion matrix for normal/ abnormal classification using SVM classifier. 185 out of 186 instances were correctly classified as normal while out of 104 instances, it correctly classified 99 as abnormal. These correspond to the classification accuracy of 99.46% and 95.19% for the normal and abnormal classes, respectively. Therefore, from the confusion matrix, it can be established that the average accuracy for the SVM classifier was 97.32% while the average accuracy for the ANN classifier was 97%, implying that the SVM classifier achieved a higher classification accuracy than the ANN classifier.

5.3.2 Performance Analysis of LDTP descriptor for the Benign/Malignant class

Table 5.3 depicted the classification of Benign/malignant class using ANN classifier. The classifier correctly classified 55 out of 58 benign instances and 40 out of 46 instances as malignant. This resulted in a percentage accuracy of 94.84 and 86.96 respectively.

Table 5.4 shows the confusion matrix for classifying an abnormal tumor belonging to either a benign or malignant class using an SVM classifier. The SVM classifier correctly classified 56 out of 58 benign instances and 42 out of 46 malignant instances, which resulted in a percentage accuracy of 96.55 and 91.30 respectively. The accuracy of the benign class was slightly higher than the malignant class by 5.25%. This implies that the SVM classifier could distinguish the benign class with higher accuracy than the malignant class.

As shown in Figure 5.6 in terms of accuracy levels of classifiers, the SVM classifier achieved a slightly higher accuracy level than the ANN classifier for both normal/abnormal and benign/malignant classes.

5.3.3 Accuracy level Comparison of ANN classifier for LDP, LTP, and LDTP descriptors

The accuracy level of the ANN classifier for the LDTP descriptor was compared against the accuracy level of the ANN classifier for the LDP and LTP descriptors. In both levels of classifying a tumor into either a normal/abnormal class or a benign/malignant class, the

161

LDTP descriptor outperformed the LDP and LTP descriptors. Figure 5.7 shows that all three descriptors could distinguish between normal and abnormal tumors with higher accuracy than when differentiating a benign tumor from a malignant one.

5.3.4 Accuracy level Comparison of SVM classifier for LDP, LTP, and LDTP descriptors

When the accuracy level of the SVM classifier was compared among the three descriptors, it was established that the LDTP descriptor outperformed the LDP and LTP descriptor for both normal/ abnormal and benign/malignant classification. Figure 5.8 depicted low accuracy levels in benign/malignant classification for both LDP and LTP descriptors. Further, for both normal/abnormal and benign/malignant classes, the LTP descriptor achieved the lowest classification accuracy.

5.3.5 Sensitivity and Specificity comparison of SVM and ANN Classifiers for LDTP descriptor

Figure 5.9 shows that the ANN classifier achieved a higher sensitivity value than the SVM classifier for the normal/abnormal class, while the SVM classifier had a slightly higher sensitivity value than the ANN classifier for benign/malignant class. The implication of the sensitivity results achieved implies that the ANN classifier was better at distinguishing between a normal and abnormal tumor, while SVM was better at differentiating between a benign and malignant tumor. In terms of specificity, Figure 5.10 shows that the SVM classifier had a higher specificity value than the ANN classifier for both normal/abnormal

162

and benign/malignant classes. This suggests that the SVM classifier was better than the ANN classifier, at designating a patient who does not have cancer cells at a higher accuracy which consequently increases the true negative rate.

5.3.6 Statistical Significance of the accuracy level achieved by LDTP descriptor

The goal of calculating the statistical significance of the achieved accuracy between the LDTP descriptor against LDP and LTP descriptors was to show the statistical significance of accuracy improvement achieved by the developed LDTP descriptor. Table 5.5 shows that the LDTP descriptor against LDP and LTP descriptors achieved a P-value of 0.000279953 and since the P-value was below 0.05, this is clear that there was a significant improvement in terms of accuracy levels for the LDTP descriptor.

5.3.7 Accuracy level comparison of LDTP descriptor with Existing Local descriptors

The accuracy level of the LDTP descriptor was compared against the accuracy levels of existing local descriptors in classifying breast cells into normal/abnormal and benign/malignant classes. Table 5.6 shows the comparison of the LDTP descriptor with existing local descriptors. The comparison analyzed the local descriptors applied, classifier used, the dataset implemented, the number of images used, the average classification accuracy achieved for normal/abnormal and/or benign/malignant classes, the author, and the year the work was published.

Table 5. 6: Accuracy level comparison of LDTP descriptor with existing local descriptors

Descriptor	Classifier	Dataset	No. of images	Classification Accuracy/AUC		Author/Year of Publication
				Normal/Abnormal	Benign/Malignant	
LBP	SVM	MIAS	70	97.2% 96.4%	-	Ponraj *et al* [237] 2017
ULDP	NLSVM, LSVM LDA, MLP	MIAS INBreast	312 417	0.92, 0.93 0.91, 0.92	-	Abdel-Nasser *et al* [238] 2015
LTP RLTP	SVM, ANN, RF	Nagoya Medical Centre	376	-	LTP-0.765, 0.773, 0.712 RLTP-0.895, 0.900,0.810	Muramatsu *et al* [240] 2016
LBP, LBPV, CLBP	Fisher Linear Discriminant Analysis	DDSM	200	-	92.95%, 87.7% 90.6%	Rabidas *et al* [234] 2016
DRLBP DRLTP	Fisher Linear Discrimin	MIAS	58	-	0.98, 0.96	Rabidas *et al* [235]2016

	ant Analysis					
CLBP with Curvelet	INN	MIAS IRMA	-	96%	-	Gardezi and Faye [236] 2015
LQP and Gabor filter	SVM	IRMA	137	99.27%	79.41%	Paramkusham *et al* [241] 2016
LDTP	**SVM ANN**	**MIAS**	**290-N/A 104-B/M**	**SVM-97.32% ANN-97.00%**	**SVM-93.93% ANN-90.90%**	**Developed LDTP**

5.4 Discussion

This research developed a local texture descriptor that encodes an image gradient by considering all directional responses and an adaptive threshold for breast cancer classification. The developed LDTP descriptor focused on classifying breast cells into normal, benign, or malignant tumors using SVM and ANN classifiers. The experiment was conducted using the MIAS dataset. Since the MIAS dataset has too few images to warrant effective generalization, the number of images was increased through data augmentation. Each image was mirrored along the x and y-axis to generate two additional images, therefore the images were increased from 322 to 966. The training/testing ratio of the

images for each category was 70:30. For the testing sample, 290 images were used for normal/abnormal class and 104 for benign/malignant class. The results based on accuracy level comparison showed that the developed LDTP descriptor produced higher accuracy for normal, abnormal, and benign, malignant classes when compared to LDP and LTP descriptors. The highest accuracy was 97.32% attained by the SVM classifier.

When the developed LDTP descriptor was compared against results achieved by other studies, it was evident from Table 5.6 that LDTP performed better than existing descriptors. Due to variations in datasets, the classifiers, the number of images and different texture descriptors used by researchers, a one-to-one comparison of the results with existing descriptors was a challenge. However, to minimize bias, the comparison was based majorly on studies that used the MIAS dataset, local descriptors, and performed classification based on normal/abnormal class and/or benign/malignant class.

When the performance of the LDTP descriptor for the normal/abnormal and benign/malignant classification is compared with the results achieved by researchers shown in table 5.6, it can be concluded that LDTP descriptor outperformed all the selected descriptors except for the work published by Paramkusham et al [241] on normal/abnormal class and work published by Rabidas et al. [235] on benign/malignant class. The higher results could be attributed to using too few images which could have caused overfitting. Sometimes, the descriptors achieved high classification for one classification group and very low results for the other group as witnessed by work done by

166

Paramkusham et al. [241]. As for the developed LDTP descriptor, the accuracy levels for both normal/abnormal and benign/malignant classification were relatively high, with an average of over 95% accuracy level.

5.5 Summary

This chapter explained the experimental validation based on the experimental setup defined in section 3.4. The experimental validation process included mammogram images reading, mammogram image preprocessing procedure, data augmentation process, feature extraction using LDTP descriptor, classification using SVM and ANN classifiers, and validation of each of the classifiers into Normal/abnormal class and benign/malignant class using the breast cancer dataset. Also the chapter explained the results obtained by the Local Directional Ternary Pattern (LDTP) descriptor. Performance analysis of the LDTP descriptor for normal/abnormal and benign/malignant classes, Accuracy level comparisons of ANN and SVM classifiers for LDP, LTP, LDTP descriptor, sensitivity and specificity comparison of SVM and ANN classifiers for LDTP descriptor, the statistical significance of accuracy level achieved by LDTP descriptor and a detailed discussion on classification comparison between LDTP descriptor with related studies based on classifying the breast cells into either Normal/Abnormal and/or Benign/Malignant classes.

CHAPTER SIX

CONCLUSION AND FUTURE WORK

6.1 Overview

This section concludes by highlighting significant findings of the research, explained the achievement of the research objectives, provided a brief description of contribution to knowledge and practice, and highlighted future work.

6.2 Summary of Findings

Effective screening can extend the survival rate for women diagnosed with breast cancer cells. Mammography is the recommended imaging test for breast cancer identification because it can recognize breast cancer cells many years in advance before physical indicators appear. However, many suspicious findings on a mammogram are benign tumors that eventually require a patient to undergo unnecessary biopsies, consequently causing anxiety to patients and increase the cost of diagnosis.

This study developed a local texture descriptor used on breast cancer data to classify a tumor as normal, benign, or malignant. This study modeled the problem of breast cancer as a two-class classification problem that classified the breast tumor into a normal or abnormal tumor and the abnormal tumor into a benign or malignant tumor. The developed LDTP descriptor showed impressive results and can effectively predict breast cancer tumor type with good precisions than LDP and LTP descriptors. Further, the LDTP descriptor

performed better for normal/abnormal class than benign/malignant class. In terms of classifier performance for the LDTP descriptor, the SVM classifier outperformed the ANN classifier using accuracy, sensitivity, and specificity measure for the benign/malignant class. The implication being that the SVM classifier was good at distinguishing a benign tumor from a malignant tumor, yet other researchers have shown that differentiating a malignant tumor from a benign tumor is a challenging task.

6.3 Achievement of Research Objectives

The objective of this research was to develop a Local Directional Ternary Pattern texture descriptor that considers all directional responses and an adaptive threshold in encoding image gradient for breast cancer classification. The objective was broken down into three specific objectives as outlined in Section 1.4.2 of this thesis. Below is a brief explanation of how the objectives were met.

To achieve the first objective of breast cancer detection techniques, a thorough literature review was conducted.

A comprehensive literature review presented in chapter two of this thesis revealed that researchers have directed significant effort towards developing a more effective breast cancer classification system. Even though local descriptors have effectively been used for pattern recognition especially in face recognition, section 2.9 in this thesis revealed that researchers sparingly used the local descriptors for breast cancer classification, because of low classification accuracy levels achieved especially for benign/malignant classification.

Also, studies that used the local descriptor and achieved high accuracies either used privately owned datasets that cannot be accessed for verification or used a few images for testing. Also, even though shape descriptors have visual cues used by the radiologist in identifying breast tumor, they, however, make the recognition process difficult especially when the image has noise, occlusion, and arbitrary distortion therefore, texture descriptors provide a better recognition because they take into consideration the structural arrangement and environmental relationship of the image.

To achieve objective two, a Local Directional Ternary Pattern texture descriptor that considers all directional responses and an adaptive threshold when encoding image gradient was developed.

The experimental setup defined in section 3.4 of this thesis, addressed the methodology followed in developing the LDTP texture descriptor. This section explained the procedure for reading the mammogram image into MATLAB, the procedure for preprocessing the mammogram image, the process of data augmentation, algorithmic steps of the new local descriptor which was explained in further details in section 4.2 of this thesis, and the process of classification, validation, and evaluation.

To achieve objective three, simulation based experimental validation of the LDTP texture descriptor on breast cancer data was conducted.

The experimental validation of the developed Local Directional Ternary Pattern was presented in section 4.3 of this thesis. This section addressed the procedure for mammogram image reading, mammogram preprocessing, data augmentation, feature extraction using the developed LDTP descriptor, classification, performance validation, and evaluation. Further, the performance analysis of the LDTP descriptor was presented in section 5.2 of this thesis.

6.4 Contribution to Knowledge

This study contributed to knowledge by developing a new Local Directional Ternary Pattern (LDTP) descriptor that considers all directional responses and an adaptive threshold in encoding image gradient for breast cancer classification. Experimental validation and evaluation of the LDTP descriptor showed better performance compared to LDP and LTP descriptors from which the new descriptor was based on. Even though local texture descriptors have not been widely applied in breast cancer classification as compared to the face recognition system, development of the LDTP descriptor and its good performance in comparison to existing methods shows a future direction that subsequent researchers can exploit.

The work was further availed to the research community by publishing it in a peer-reviewed journal. Further, the developed LDTP descriptor affirmed the theory of feature analysis as used in face recognition. This showed that theory of feature analysis can

effectively be applied in breast cancer tumor differentiation. Consequently building a better breast cancer classification system that improves the classification accuracy.

6.5 Contribution to Practice

The current imaging test recommended for breast cancer screening is mammography. The mammographic test is performed by a radiologist who reads and interprets the test results, however, the radiologist is susceptible to human observer variability and the reading and interpretation of the test results depend on the proficiency of the radiologist administering the test. This research contributed to practice by presenting a more effective way of reading and interpreting the mammographic test results through a Computer-Aided Detection System (CAD). The CAD system could aid the radiologist in making a more accurate reading and interpretation of mammogram results by eliminating observer oversight, which consequently reduces unnecessary biopsies, increases the survival rate of women, and provides better health care to humanity.

Currently, a radiologist relies on visual cue and the shape interior region or contours defined by the tumor boundary to identify the shape of the tumor and classify it. Shape features are not sufficient to describe a tumor, because they do not provide structural arrangements and environmental relationships of the image. This research makes a second contribution to practice by presenting a technique of extracting discriminant features independent of a radiologist. This technique extracts features by looking at the structural properties which are not visible to the human eye unlike a radiologist who relies on the

shape features. Therefore using the features extracted by the developed descriptor provides insight for the radiologist since it eliminates observer variability.

Dust particles on a mammogram and breast surgery scars on a patient can obstruct the radiologist and lead to false interpretation. This study makes a third contribution to practice by providing a way of improving a mammogram image by preprocessing the image before use. In this study, three ways of improving and enhancing image quality are discussed and implemented. The images were cleansed of noise, labels and artifacts then enhanced and cropped.

6.6 Future Work

The prospects of this research are to further show the robustness and feasibility of the developed LDTP descriptor by fusing different texture features and ensembling different classifiers to arrive at a final decision. Fusing features and classifier decisions is an area that has not been exhaustively exploited for breast cancer classification. Fusing features in different categories of features could result in an enriched feature set that can differentiate between breast cancer tumors especially for the benign/malignant class. Also, implementing decision fusion from multiple classifiers is expected to generate a classifier that is more robust from individual classifiers. To further test the feasibility of the LDTP descriptor, there would be a need to implement and compare its performance on other mammographic breast cancer datasets with many images. Also, in the future, the issue of balancing between effectiveness and efficiency should be taken into consideration.

Currently, a lot of emphases has been focused on how to effectively improve breast cancer detection at the expense of efficiency. Finally in future it would be viable to develop a CAD feature extraction tool based on the LDTP descriptor.

REFERENCES

[1]   A. Nahid and Y. Kong, "Involvement of Machine Learning for Breast Cancer Image Classification : A Survey," *Hindawi Comput. Math. Methods Med.*, vol. 2017, no. i, p. 29, 2018.

[2]   A. C. Society, "Cancer Facts & Figures 2019," 2019.

[3]   R. L. Siegel and K. D. Miller, "Cancer Statistics , 2019," vol. 69, no. 1, pp. 7–34, 2019.

[4]   P. Skaane, "Radiology Screen-Film Mammography versus Full-Field Digital Mammography with Soft-Copy Reading : Randomized Trial in a Population-based Screening Program — The Oslo II Study 1," pp. 197–204, 2004.

[5]   N. M. Hambly, N. Phelan, and F. L. Flanagan, "Women's Imaging • Or iginal Research Comparison of Digital Mammography and Screen-Film Mammography in Breast Cancer Screening: A Review in the Irish Breast Screening Program," no. October, pp. 1010–1018, 2009.

[6]   N. M. Hambly, M. M. McNicholas, N. Phelan, G. C. Hargaden, A. O'Doherty, and F. L. Flanagan, "Comparison of digital mammography and screen-film mammography in breast cancer screening: A review in the Irish Breast Screening Program," *Am. J. Roentgenol.*, vol. 193, no. 4, pp. 1010–1018, 2009.

[7]   M. R. Del Turco *et al.*, "Full-field digital versus screen-film mammography: Comparative accuracy in concurrent screening cohorts," *Am. J. Roentgenol.*, vol. 189, no. 4, pp. 860–866, 2007.

[8]   J. M. Lewin *et al.*, "Clinical comparison of full-field digital mammography and

screen-film mammography for detection of breast cancer," *AJR Am J Roentgenol*, vol. 179, no. 3, pp. 671–677, 2002.

[9] N. P. Pérez, "Improving Variable Selection and Mammography-based Machine Learning Classifiers for Breast Cancer CADx," 2015.

[10] J. Ferlay *et al.*, "Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer*, vol. 136, no. 5, pp. E359–E386, 2015.

[11] F. Bray, J. Ferlay, and I. Soerjomataram, "Global Cancer Statistics 2018 : GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA Cancer*, pp. 394–424, 2018.

[12] "Breast Cancer," 2018.

[13] "Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability- Adjusted Life-years for 32 Cancer Groups, 1990 to 2015 A Systematic Analysis for the Global Burden of Disease Study," 2017.

[14] R. Gakunga *et al.*, "Social determinants and individual health-seeking behaviour among women in Kenya : protocol for a breast cancer cohort feasibility study," *BMJ Open*, 2019.

[15] A. Korir, N. Okerosi, V. Ronoh, G. Mutuma, and M. Parkin, "Incidence of cancer in Nairobi, Kenya (2004-2008)," *Int. J. Cancer*, vol. 137, no. 9, pp. 2053–2059, 2015.

[16] R. T. Sawe et al., "Aggressive breast cancer in western Kenya has early onset , high proliferation , and immune cell infiltration," *BMC Cancer*, pp. 1–15, 2016.

[17] F. W. Wambalaba, B. Son, A. E. Wambalaba, D. Nyong, and A. Nyong, "Prevalence and Capacity of Cancer Diagnostics and Treatment : A Demand and Supply Survey of Health-Care Facilities in Kenya," vol. 26, pp. 1–12, 2019.

[18] C. N. Tenge, R. T. Kuremu, N. G. Buziba, K. Patel, and P. A. Were, "Burden and Pattern of Cancer in Western Kenya," *East Afr. Med. J.*, vol. 86, no. 1, 2009.

[19] M. Tafa Segni and D. M. Tadesse, "Breast Self-examination: Knowledge, Attitude, and Practice among Female Health Science Students at Adama Science and Technology University, Ethiopia," *Gynecol. Obstet.*, vol. 06, no. 04, 2015.

[20] S. McDonald, D. Saslow, and M. H. Alciati, "Performance and Reporting of Clinical Breast Examination: A Review of the Literature," *CA. Cancer J. Clin.*, vol. 54, no. 6, pp. 345–361, 2004.

[21] A. Jalalian et al., "Review article : Foundation and Methodologies in Computer Aided diagnosis systems for breast cancer detection," pp. 113–137, 2017.

[22] M. Posso and T. Puig, "Cost-Effectiveness of Double Reading versus Single Reading of Mammograms in a Breast Cancer Screening Programme," 2016.

[23] M. C. Posso, T. Puig, M. J. Quintana, J. Solà-roca, and X. Bonfill, "Double versus single reading of mammograms in a breast cancer screening programme : a cost-consequence analysis," 2016.

[24] B. C. Maria Rizzi, Matteo D'Aloia, "Review: Health care CAD system for breast microcalcification cluster detection," *J. Med. Biol. Eng.*, vol. 32, no. 3, pp. 147–156, 2011.

[25] M. Posso, M. Posso, and T. Puig, "Cost-Effectiveness of Double Reading versus

Single Reading of Mammograms in a Breast Cancer Screening Programme," no. July, 2016.

[26]   A. Group, B. S. Service, and S. Clinic, "The Efficacy of Double Reading Mammograms in Breast Screening," vol. Clinical R, pp. 248–251, 1994.

[27]   T. H. Rassem and B. E. Khoo, "Completed Local Ternary Pattern for Rotation Invariant Texture Classification Completed Local Ternary Pattern for Rotation Invariant Texture Classification," no. April, 2014.

[28]   L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Local binary features for texture classi fi cation : Taxonomy and experimental study," *Pattern Recognit.*, vol. 62, pp. 135–160, 2017.

[29]   M. Raja, "Optimized Local Ternary Patterns: A new texture model with set of optimal patterns for texture analysis," no. August, 2017.

[30]   T. Jabid, H. Kabir, and O. Chae, "Local Directional Pattern (LDP) – A Robust Image Descriptor for Object Recognition," *Seventh IEEE Int. Conf. Adv. Video Signal Based Surveill.*, pp. 482–487, 2010.

[31]   H. M. H. Alharbi, P. Kwan, A. Jayawardena, and A. S. M. Sajeev, "Fuzzy Image Segmentation for Mass Detection in Digital Mammography," *Multidiscip. Comput. Intell. Tech.*, no. January, pp. 378–402, 2012.

[32]   W. F. Anderson, I. Jatoi, J. Tse, and P. S. Rosenberg, "Male Breast Cancer : A Population-Based Comparison With Female Breast Cancer," 2010.

[33]   Y. Ponirovskaya, "Risk Factors for Breast Cancer," 2000.

[34]   W. He, P. Hogg, A. Juette, E. R. E. Denton, and R. Zwiggelaar, "Breast image pre-

processing for mammographic tissue segmentation," *Comput. Biol. Med.*, vol. 67, pp. 61–73, 2015.

[35] M. King, M. King, J. H. Marks, and J. B. Mandell, "Breast and Ovarian Cancer Risks Due to Inherited Mutations in BRCA1 and BRCA2," vol. 643, no. 2003, 2012.

[36] K. E. Malone, J. R. Daling, J. D. Thompson, C. A. O'Brien, L. V. Francisco, and E. A. Ostrander, "BRCA1 mutations and breast cancer in the general population: Analyses in women before age 35 years and in women before age 45 years with first-degree family history," *J. Am. Med. Assoc.*, vol. 279, no. 12, pp. 922–929, 1998.

[37] S. Rinaldi et al., "Menstrual and reproductive factors and risk of breast cancer : A case-control study in the Fez region , Morocco," pp. 1–12, 2018.

[38] J. Chang-claude et al., "Age at Menarche and Menopause and Breast Cancer Risk in the International BRCA1 / 2 Carrier Cohort Study," vol. 16, no. April, pp. 740–747, 2007.

[39] V. Siskind, K. Tajima, J. M. Liff, and A. Morabia, "Breast cancer and breastfeeding : collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries , including 50 302 women with breast cancer and 96 973 women without the disease," no. March, 2015.

[40] E. H. Anstey, M. L. Shoemaker, C. M. Barrera, E. O. Neil, A. B. Verma, and D. M. Holman, "HHS Public Access," vol. 53, pp. 1–11, 2018.

[41] S. E. Schetter, "Breast Density as an Independent Risk Factor for Cancer," vol. 3,

no. 1, pp. 10–19, 2014.

[42] N. F. Boyd, L. J. Martin, M. J. Yaff, and S. Minkin, "Mammographic density and breast cancer risk : current understanding and future prospects," pp. 1–12, 2011.

[43] L. Sun, M. Sc, J. Stone, M. Sc, and E. Fishell, "Mammographic Density and the Risk and Detection of Breast Cancer," pp. 227–236, 2007.

[44] P. E. Freer, "Mammographic Breast Density : Impact on Breast Cancer Risk and Implications for Screening 1," 2015.

[45] S. W. Duffy *et al.*, "Mammographic density and breast cancer risk in breast screening assessment cases and women with a family history of breast cancer," *Eur. J. Cancer*, vol. 88, pp. 48–56, 2018.

[46] K. A. Ban, "Epidemiology of Breast cancer R," *Surg. Oncol. Clin. NA*, vol. 23, no. 3, pp. 409–422, 2014.

[47] M. Kamińska, T. Ciszewski, K. Łopacka-szatan, P. Miotła, and E. Starosławska, "Breast cancer risk factors," vol. 14, no. 3, pp. 196–202, 2015.

[48] J. Iqbal, O. Ginsburg, P. A. Rochon, P. Sun, and S. A. Narod, "Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States," *JAMA - J. Am. Med. Assoc.*, vol. 313, no. 2, pp. 165–173, 2015.

[49] G. A. Saxe, C. L. Rock, M. S. Wicha, and D. Schottenfeld, "Diet and risk for breast cancer recurrence and survival," *Breast Cancer Res. Treat.*, vol. 53, no. 3, pp. 241–253, 1999.

[50] L. C. Kobayashi and I. Janssen, "Moderate-to-vigorous intensity physical activity across the life course and risk of pre- and post-menopausal breast cancer," pp. 851–

861, 2013.

[51] N. Razvi, A. Maqbool, E. Commission, and N. Jahan, "A review of breast cancer risk factors," no. March 2018, 2017.

[52] E. Monninkhof, S. Elias, and F. Vlems, "Physical Activity and Breast Cancer : A Systematic Review," *epidemiology*, vol. 18, no. 1, pp. 137–157, 2007.

[53] E. E. Calle *et al.*, "Breast cancer and hormone replacement therapy: Collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer," *Lancet*, vol. 350, no. 9084, pp. 1047–1059, 1997.

[54] M. Linda L. Humphrey, MD, MPH; Mark Helfand, MD, MS; Benjamin K.S. Chan, MS; and Steven H. Woolf, MD, P. Services, and T. Force, "Breast Cancer Screening: A summary of the Evidence for the U.S. Preventive Services Task Force," 2002.

[55] K. Birhane et al., "Practices of Breast Self-Examination and Associated Factors among Female Debre Berhan University Students," vol. 2017, 2017.

[56] O. Article, "Knowledge , Attitude and Practice of Breast Self‐ examination among Females in Medical and Non‐ medical Colleges in Qassim University," pp. 219–224, 2017.

[57] J. Osuch, C. Coleman, and M. Dolan, "Clinical Breast Examination : Practical Recommendations for Optimizing Clinical Breast Examination : Practical Recommendations for Optimizing Performance and Reporting," no. November, 2004.

[58] D. Saslow et al., "Clinical Breast Examination: Practical Recommendations for Optimizing Performance and Reporting," *CA. Cancer J. Clin.*, vol. 54, no. 6, pp. 327–344, 2004.

[59] A. Mavi et al., "The effect of age, menopausal state, and breast density on 18 F-FDG uptake in normal glandular breast tissue," *J. Nucl. Med.*, vol. 51, no. 3, pp. 347–352, 2010.

[60] P. Hiral and P. Poorvi, "A Review on Different Techniques Used for Feature Extractions from Mammogram of Breast Cancer," *J. Image Process. Pattern Recongnition Prog.*, vol. 1, no. 3, pp. 0–7, 2015.

[61] P. Mehnati and M. J. Tirtash, "Comparative Efficacy of Four Imaging Instruments for Breast Cancer Screening," *Asian Pacific J. Cancer Prev.*, vol. 16, pp. 6177–6186, 2015.

[62] S. Islam, N. Kaabouch, and W. C. Hu, "A Survey of Medical Imaging Techniques Used for Breast Cancer Detection," *Electro/Information Technol. (EIT), 2013 IEEE Int. Conf.*, pp. 10–14, 2013.

[63] J. Dheeba, N. Albert Singh, and S. Tamil Selvi, "Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach," *J. Biomed. Inform.*, vol. 49, pp. 45–52, 2014.

[64] C. M. Ronckers, C. A. Erdmann, and C. E. Land, "Radiation and breast cancer : a review of current evidence," pp. 21–32, 2005.

[65] E. D. Pisano *et al.*, "Diagnostic Accuracy of Digital versus Film Mammography: Exploratory Analysis of Selected Population Subgroups in DMIST," *Radiology*, vol.

246, no. 2, pp. 376–383, 2008.

[66]   CDE, "Clinical Guidelines," 2015.

[67]   W. A. Berg et al., "Diagnostic Accuracy of Mammography , Clinical Examination , US , and MR Imaging in Preoperative Assessment of Breast Cancer 1," 2004.

[68]   B. E. Warner et al., "Comparison of Breast Magnetic Resonance Imaging , Mammography , and Ultrasound for Surveillance of Women at High Risk for Hereditary Breast Cancer," *J. Clin. Oncol.*, vol. 19, no. 15, pp. 3524–3531, 2001.

[69]   E. Warner *et al.*, "Surveillance of BRCA1 and BRCA2 Mutation," vol. 292, no. 11, pp. 1317–1325, 2004.

[70]   L. S. J. Sim, J. Hendriks, and S. M. C. Fook-chong, "Breast Ultrasound in Women With Familial Risk of Breast Cancer," pp. 600–606.

[71]   C. K. Kuhl et al., "Mammography , Breast Ultrasound , and Magnetic Resonance Imaging for Surveillance of Women at High Familial Risk for Breast Cancer," *J. Clin. Oncol.*, vol. 23, no. 33, pp. 8469–8476, 2005.

[72]   R. Croshaw, H. Shapiro-wright, E. Svensson, K. Erb, and T. Julian, "Accuracy of Clinical Examination , Digital Mammogram , Ultrasound , and MRI in Determining Postneoadjuvant Pathologic Tumor Response in Operable Breast Cancer Patients," *Ann. Surg. Oncol.*, pp. 3160–3163, 2011.

[73]   S. A. Valente et al., "Accuracy of Predicting Axillary Lymph Node Positivity by Physical Examination , Mammography , Ultrasonography , and Magnetic Resonance Imaging," *Ann. Surg. Oncol.*, pp. 1825–1830, 2012.

[74]   H. Shao, B. Li, X. Zhang, Z. Xiong, Y. Liu, and G. Tang, "Comparison of the

diagnostic efficiency for breast cancer in Chinese women using mammography , ultrasound , MRI , and different combinations of these imaging modalities," *J. Xray. Sci. Technol.*, vol. 21, pp. 283–292, 2013.

[75] T. Huzarski, B. Górecka-szyld, and J. Huzarska, "Screening with magnetic resonance imaging , mammography and ultrasound in women at average and intermediate risk of breast cancer," *Hered. Cancer Clin. Pract.*, pp. 1–8, 2017.

[76] A. Hossam and H. M. Harb, "Performance Analysis of Breast Cancer Imaging Techniques," *Int. J. Comput. Sci. Inf. Secur.*, no. May 2017, 2018.

[77] H. Sun, H. Li, S. Si, and S. Qi, "Performance evaluation of breast cancer diagnosis with mammography, ultrasonography and magnetic resonance imaging," *J. Xray. Sci. Technol.*, pp. 1–9, 2018.

[78] C. Gajdos et al., "Mammographic Appearance of Nonpalpable Breast Cancer Reflects Pathologic Characteristics," *Ann. Surg.*, vol. 235, no. 2, pp. 246–251, 2002.

[79] A. Venkatesan, P. Chu, K. Kerlikowske, E. A. Sickles, and R. Smith-bindman, "Positive Predictive Value of Specific Mammographic Findings according to Reader and Patient Variable," vol. 250, no. 3, 2009.

[80] J. Bozek, M. Mustra, K. Delac, and M. Grgic, "A Survey of Image Processing Algorithms in Digital Mammography," pp. 631–657, 2009.

[81] H. Boulehmi, H. Mahersia, and K. Hamrouni, "A New CAD System for Breast Microcalcifications Diagnosis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 4, pp. 133–143, 2016.

[82] I. Zyout, I. Abdel-qader, and C. Jacobs, "Bayesian Classifier with Simplified

Learning Phase for Detecting Microcalcifications in Digital Mammograms," *Int. J. Biomed. Imaging*, vol. 2009, 2009.

[83] M. P. Sampat, M. K. Markey, and A. C. Bovik, *Computer-Aided Detection and Diagnosis in Mammography*, Second Edi. Elsevier Inc.

[84] S. O. Grady and M. P. Morgan, "BBA - Reviews on Cancer Microcalci fi cations in breast cancer : From pathophysiology to diagnosis and prognosis," vol. 1869, no. March, pp. 310–320, 2018.

[85] M. a. Alolfe, W. a. Mohamed, A.-B. M. Youssef, Y. M. Kadah, and A. S. Mohamed, "Feature selection in computer aided diagnostic system for microcalcification detection in digital mammograms," in *26th National Radio Science Conference (NRSC2009)*, 2009.

[86] T. Balakumaran, "Detection of Microcalcification in Mammograms Using Wavelet Transform and Fuzzy Shell Clustering," vol. 7, no. 1, pp. 121–125, 2010.

[87] M. G. Mini, V. P. Devassia, and T. Thomas, "Multiplexed Wavelet Transform Technique for Detection of Microcalcification in Digitized Mammograms," *J. Digit. Imaging*, vol. 17, no. 4, pp. 285–291, 2004.

[88] D. Gunawan, "Microcalcification Detection Using Wavelet Transform," pp. 694–697.

[89] S. Bouyahia, J. Mbainaibeye, and N. Ellouze, "Wavelet Based Microcalcifications Detection in Digitized Mammograms," no. January, pp. 23–31, 2009.

[90] N. B. Karayiannis, "Detection of Microcalcifications in Digital Mammograms Using Wavelets," no. September, 2014.

185

[91] I. Zyout, "Computer-Aided Diagnosis of Microcalcification Clusters Using Morphology Based Features and PSO-SVM Parameter Selection," vol. 2, no. 2, pp. 126–144, 2016.

[92] I. Zyout, I. Abdel-qader, and C. Jacobs, "Embedded Feature Selection using PSO-kNN: Shape-Based Diagnosis of Microcalcification Clusters in Mammography.," *Juspn*, vol. 3, no. 1, pp. 7–11, 2011.

[93] P.-P. T. 1 Yi-Jhe Huang 1, Ding-Yuan Chan 2, Da-Chuan Cheng 3,*, Yung-Jen Ho 3,*, "Automated Feature Set Selection and Its Application to MCC Identification in Digital Mammograms for Breast Cancer Detection," pp. 4855–4875, 2013.

[94] K. Geethal and K. T. A. Kishore, "New Particle Swarm Optimization for Feature Selection and Classification of Microcalcifications in Mammograms Method Anlyl," *IEEE-International Conf. Signal Process. Netw. Madras Inst. Technol. Anna Univ. Chennai India,* vol. 4, no. 6, pp. 458–463, 2008.

[95] M. Vasantha, "Classifications of Mammogram Images Using Hybrid Features," 2015.

[96] M. P. Sampat, M. K. Markey, and A. C. Bovik, "Computer-Aided Detection and Diagnosis in Mammography," *Handb. Image Video Process.*, pp. 1195–1217, 2005.

[97] S. Liu, C. F. Babbs, and E. J. Delp, "Multiresolution detection of spiculated lesions in digital mammograms," *IEEE Trans. Image Process.*, vol. 10, no. 6, pp. 874–884, 2001.

[98] M. P. Sampat and A. C. Bovik, "Detection of Spiculated Lesions in Mammograms," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, vol. 1, no. July, pp. 810–813, 2003.

[99] R. Zwiggelaar et al., "Model-based detection of spiculated lesions in mammograms," *Med. Image Anal.*, vol. 3, no. 1, pp. 39–62, 1999.

[100] K. U. Sheba and S. G. Raj, "An approach for automatic lesion detection in mammograms," *Cogent Eng.*, vol. 2016, pp. 1–16, 2018.

[101] J. Wei *et al.*, "Computer-aided detection of breast masses on mammograms: Dual system approach with two-view analysis," *Med. Phys.*, vol. 36, no. 10, pp. 4451–4460, 2009.

[102] A. Elmoufid *et al.*, "Automatic Diagnosing of Suspicious Lesions in Digital Mammograms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 5, pp. 510–518, 2016.

[103] S. Punitha, A. Amuthan, and K. S. Joseph, "Benign and malignant breast cancer segmentation using optimized region growing technique," *Futur. Comput. Informatics J.*, vol. 3, no. 2, pp. 348–358, 2018.

[104] K. Hu, X. Gao, and F. Li, "Detection of suspicious lesions by adaptive thresholding based on multiresolution analysis in mammograms," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 2, pp. 462–472, 2011.

[105] F. Soares Sérvulo de Oliveira, A. Oseas de Carvalho Filho, A. Corrêa Silva, A. Cardoso de Paiva, and M. Gattass, "Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM," *Comput. Biol. Med.*, vol. 57, pp. 42–53, 2014.

[106] G. M. Te Brake, N. Karssemeijer, and J. H. C. L. Hendriks, "An automatic method to discriminate malignant masses from normal tissue in digital mammograms," *Phys. Med. Biol.*, vol. 45, no. 10, pp. 2843–2857, 2000.

[107] Y. Wang, F. Aghaei, A. Zarafshani, Y. Qiu, W. Qian, and B. Zheng, "Computer-aided classification of mammographic masses using visually sensitive image features," *J. Xray. Sci. Technol.*, vol. 25, no. 1, pp. 171–186, 2017.

[108] N. R. Mudigonda, R. M. Rangayyan, and J. E. L. Desautels, "Detection of Breast Masses in Mammograms by Density Slicing and Texture Flow-Field Analysis," vol. 20, no. 12, pp. 1215–1227, 2001.

[109] R. Nithya and B. Santhi, "Mammogram classification using maximum difference feature selection method," *J. Theor. Appl. Inf. Technol.*, vol. 33, no. 2, 2011.

[110] Herwanto and A. Arymurthy, "A System for Computer Aided Diagnosis of Breast Cancer Based on Mass Analysis," *Int. Conf. Robot. Biomimetics, Intell. Comput. Syst. Yogyakarta,* pp. 247–253, 2013.

[111] X. Liu, X. Xu, and J. Liu, "A new automatic method for mass detection in mammography with false positives reduction by supported vector machine," pp. 33–37, 2011.

[112] H. A. Khan, A. Al Helal, K. I. Ahmed, and R. Mostafa, "Abnormal Mass Classification in Breast Mammography using Rotation Invariant LBP," no. September, 2016.

[113] R. Nithya and B. Santhi, "Computer Aided Diagnosis System for Mammogram Analysis : A Survey," vol. 5, no. 4, 2015.

[114] R. J. McKenna, "The abnormal mammogram radiographic findings, diagnostic options, pathology, and stage of cancer diagnosis," *Cancer*, vol. 74, no. 1 S, pp. 244–255, 1994.

[115] S. Banik, R. M. Rangayyan, and J. E. Leo Desautels, "Computer-aided detection of architectural distortion in prior mammograms of interval cancer," *J. Digit. Imaging*, vol. 47, no. 5, pp. 1–193, 2013.

[116] M. K. Shetty, "Mammographic Signs of Breast Cancer," pp. 93–117, 2015.

[117] V. Lattanzio and G. Simonetti, *Mammography: Guide to Interpreting, Reporting and Auditing Mammographic Images*, vol. 240, no. 2. 2006.

[118] R. Bhanumathi and G. . Suresh, "Performance Analysis in Computer Aided Detection of Breast Cancer by Mammography," *Int. J. IT Eng.*, vol. 01, no. 02, 2013.

[119] M. J. M. Broeders, N. C. Onland-Moret, H. J. T. M. Rijken, J. H. C. L. Hendriks, A. L. M. Verbeek, and R. Holland, "Use of previous screening mammograms to identify features indicating cases that would have a possible gain in prognosis following earlier detection," *Eur. J. Cancer*, vol. 39, no. 12, pp. 1770–1775, 2003.

[120] R. M. Rangayyan, F. J. Ayres, and J. E. Leo Desautels, "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs," *J. Franklin Inst.*, vol. 344, no. 3–4, pp. 312–348, 2007.

[121] M. N. Patel and P. Tandel, "A Survey on Feature Extraction Techniques for Shape based Object Recognition," *Int. J. Comput. Appl.*, vol. 137, no. 6, pp. 16–20, 2016.

[122] E. Omidiora, S. O. Olabiyisi, A. Temitope, and A. Temilola, "Feature Extraction Techniques for Mass Detection in Digital Mammogram (Review)," *J. Sci. Res. Reports*, no. February 2019, 2017.

[123] S. Caulkin and S. Astley, "Generating Realistic Mass Lesions In Digital

Mammograms Using Statistical models," pp. 285–294, 1999.

[124] M. Berks, S. Caulkin, R. Rahim, C. Boggis, and S. Astley, "Statistical Appearance Models of Mammographic Masses," pp. 401–408.

[125] X. Zhang, J. Cui, W. Wang, and C. Lin, "A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm," *Sensors (Switzerland)*, vol. 17, no. 7, 2017.

[126] P. Delogu, M. Evelina, P. Kasae, and A. Retico, "Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier," vol. 37, pp. 1479–1491, 2007.

[127] M. Varma and A. Zisserman, "A Statistical Approach to Texture Classification from Single Images," 2004.

[128] R. A. Castellino, "Computer aided detection (CAD): An overview," *Cancer Imaging*, vol. 5, no. 1, pp. 17–19, 2005.

[129] M. Giger, "Computer Aided Diagnosis of breast lesions in medical images," *Comput. Med.*, pp. 39–45, 2000.

[130] A. Jalalian *et al.*, "Review article : Foundation and Methodologies in Computer-Aided Doagnosis systems for breast cancer detection," pp. 113–137, 2017.

[131] H. S. Sheshadri and A. Kandaswamy, "Computer aided decision system for early detection of breast cancer," no. August, pp. 149–154, 2006.

[132] I. A. Lbachir, R. Es-salhi, I. Daoudi, and S. Tallal, "A new mammogram preprocessing method for Computer-Aided Diagnosis systems," 2017.

[133] Kshema, G. Jayesh, and S. Dhas, "Preprocessing Filters for Mammogram images: A

review," in *IEEE Conference on Emerging Devices and smart systems (ICEDSS 2017)*, 2017, no. March, pp. 3–4.

[134] H. A. Alghaib, M. Scott, and R. R. Adhami, "An overview of mammogram analysis," *IEEE Potentials*, vol. 35, no. 6, pp. 21–28, 2016.

[135] B. Kaur and M. Shukla, "Image De-noising and its Methods : A Survey," vol. 3, no. 6, pp. 234–238, 2014.

[136] C. A. H. J and N. M. N. S, "A Survey on Image Denoising methods," vol. 3, no. 1, pp. 153–156, 2013.

[137] A. A. Kayode, B. S. Afolabi, and B. O. Ibitoye, "An Explorative Survey of Image Enhancement Techniques Used in Mammography," vol. 12, no. 1, pp. 72–79, 2015.

[138] D. N. Ponraj, M. E. Jenifer, and J. S. Manoharan, "A Survey on the Preprocessing Techniques of Mammogram for the Detection of Breast Cancer," *J. Emerg. Trends Comput. Infromation Sci.*, vol. 2, no. 12, pp. 656–664, 2011.

[139] M. Vasantha, S. Bharathi, and R. Dhamodharan, "Medical Image Feature , Extraction , Selection And Classification," *Int. J. Eng. Sci. Technol.*, vol. 2, no. 6, pp. 2071–2076, 2010.

[140] R. Sivaramakrishna, N. A. Obuchowski, W. A. Chilcote, G. Cardenosa, and K. A. Powell, "Mammographic Enhancement Algorithms : A Preference Study," no. July, pp. 45–51, 2000.

[141] C. L. Y. Sivakumari, "Comparison of Diverse Enhancement Techniques for Breast Mammograms," vol. 1, no. 7, pp. 400–407, 2013.

[142] P. Sivakumar and S. Meenakshi, "A review on image segmentation techniques," *Int.*

J. Adv. Res. Comput. Eng. Technol., vol. 5, no. 3, 2016.

[143] Y. Ali and S. Hamed, "Early Breast Cancer Detection using Mammogram Images : A Review of Image Processing Techniques," vol. 12, no. March, pp. 225–234, 2015.

[144] N. Tokas, S. Karkra, and M. K. Pandey, "Comparison of Digital Image Segmentation Techniques- A Research Review," *Int. J. Comput. Sci. Mob. Comput.*, vol. 5, no. 5, pp. 215–220, 2016.

[145] S. Yuheng and Y. Hao, "Image Segmentation Algorithms Overview," vol. 1.

[146] P. A. Mlsna and J. J. Rodríguez, *Chapter 19 - Gradient and Laplacian Edge Detection*, 1st ed. Elsevier.

[147] S. Kamdi and R. K. Krishna, "Image Segmentation and Region Growing Algorithm," *Int. J. Comput. Technol. Electron. Eng.*, vol. 2, no. 1, pp. 103–107.

[148] I. Ait, R. Es-salhi, I. Daoudi, S. Tallal, and H. Medromi, "A Survey on Segmentation Techniques of Mammogram Images," *Adv. ubiqutious Netw.*, 2016.

[149] M. Hosseinzadeh and P. Khoshvaght, "A Comparative Study of Image Segmentation Algorithms," *Int. J. Sci. Eng. Technol.*, vol. 9, no. 8, pp. 1966–1971, 2015.

[150] R. Dubey, "Review of various techniques of mammogram image segmentation," no. May, 2015.

[151] A. A. Hefnawy, "An Improved Approach for Breast Cancer Detection in Mammogram based on Watershed Segmentation," *Int. J. Comput. Appl.*, vol. 75, no. 15, pp. 26–30, 2013.

[152] K. Chuang, H. Tzeng, S. Chen, J. Wu, and T. Chen, "Fuzzy c-means clustering with

spatial information for image segmentation," *Comput. Med. imaging Graph.*, vol. 30, pp. 9–15, 2006.

[153] R. Besar, "Identification of masses in digital mammogram using gr a y level co-occurrence matrices," *Biomed. Imaging Interv. J.*, 2009.

[154] D. Xie, M. L. Id, Y. Xie, D. Liu, and X. Li, "A fast threshold segmentation method for froth image base on the pixel distribution characteristic," pp. 1–18, 2019.

[155] J. Nagi and S. A. Kareem, "Automated breast profile segmentation for ROI detection using digital mammograms," no. January, 2011.

[156] H.P.Narkhede, "Review of Image Segmentation Techniques," *Ijisme, Issn2319-6386*, vol. 1, no. 8, pp. 2015–2018, 2013.

[157] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," *Proc. 2014 Sci. Inf. Conf. SAI 2014*, pp. 372–378, 2014.

[158] A. P. Esmita Gupta ME Student, "Process Mining A Comparative Study," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 3, no. 11, pp. 17–23, 2014.

[159] M. Kallergi and R. A. Übersichtsarbeit, "Computer-aided diagnosis of mammographic microcalcification clusters," *Int. J. Med. Phys. Res. Pract.*, vol. 314, no. 2004, 2012.

[160] Y. Zheng, "Breast Cancer Detection with Gabor Features from Digital Mammograms," *Open Access*, pp. 44–62, 2010.

[161] S. Khan and M. Hussain, "A comparison of different Gabor feature extraction approaches for mass classification in mammography," 2015.

[162] M. A. Berbar, "Hybrid methods for feature extraction for breast masses classification," *Egypt. Informatics J.*, vol. 19, no. 1, pp. 63–73, 2018.

[163] E. K. Sharma, E. Priyanka, E. A. Kalsh, and E. K. Saini, "GLCM and its Features," *Int. J. Adv. Res. Electron. Commun. Eng. Vol.*, vol. 4, no. 8, pp. 2180–2182, 2015.

[164] P. Mohanaiah, P. Sathyanarayana, and L. Gurukumar, "Image Texture Feature Extraction Using GLCM Approach," vol. 3, no. 5, pp. 1–5, 2013.

[165] T. Ojala, M. Pietikainen, and D. Harwood, "A Comparative Study on Texture Measures with Classification based on Feature Distributions," *Pattern Recognit.*, vol. 29, no. l, 1996.

[166] T. H. Rassem and B. E. Khoo, "Completed Local Ternary Pattern for Rotation Invariant Texture Classification," *Sci. World J.*, vol. 2014, 2014.

[167] A. V Sansare and S. R. Kinge, "Classification of Breast cancer Using Local Binary Pattern and Gabor Filter," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 5, no. X, pp. 1389–1395, 2017.

[168] S. Naresh and M. Tech, "Breast Cancer Detection using Local Binary Patterns," vol. 123, no. 16, pp. 6–9, 2015.

[169] S. Lee, "Multilayer Cluster Neural Network for Totally Unconstrained Handwritten Numeral Recognition," *Neural Networks*, vol. 8, no. 5, pp. 783–792, 1995.

[170] A. M. Shabat and J. Tapamo, "A comparative study of the use of local directional pattern for texture-based informal settlement classification," *J. Appl. Res. Technol.*, vol. 15, no. 3, pp. 250–258, 2017.

[171] L. Nanni, A. Lumini, and S. Brahnam, "Artificial Intelligence in Medicine Local

binary patterns variants as texture descriptors for medical image analysis," *Artif. Intell. Med.*, vol. 49, no. 2, pp. 117–125, 2010.

[172] K. C. Khor, C. Y. Ting, and S. P. Amnuaisuk, "From feature selection to building of Bayesian classifiers: A network intrusion detection perspective," *Am. J. Appl. Sci.*, vol. 6, no. 11, pp. 1949–1960, 2009.

[173] M. Data, Z. M. Hira, and D. F. Gillies, "A Review of Feature Selection and Feature Extraction Methods Applied on A Review of Feature Selection and Feature Extraction Methods Applied on Microarray Data," no. July, 2015.

[174] Y. Saeys, I. Inza, and P. Larra??aga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[175] S. Srivastava, N. Sharma, and S. K. Singh, "Fusion of SFS-SVM feature selection methods using robust rank aggregation for optimal feature subset selection for mammogram classification," vol. 3, no. 1, pp. 171–176, 2014.

[176] N. P. Pérez, "Improving Variable Selection and Mammography-based Machine Learning Classifiers for Breast Cancer CADx," 2015.

[177] J. S. S.Uma, "Human Interaction Pattern Mining Using Enhanced Artificial Bee Colony Algorithm S.," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 3, no. 10, pp. 10131–10138, 2015.

[178] V. Kumar, "Feature Selection: A literature Review," *Smart Comput. Rev.*, vol. 4, no. 3, 2014.

[179] S. Vanaja, "Analysis of Feature Selection Algorithms on Classification : A Survey," *Int. J. Comput. Appl. `(0975 – 8887)*, vol. 96, no. 17, pp. 28–35, 2014.

[180] A. Heshmati, R. Amjadifard, and J. Shanbehzadeh, "ReliefF-based feature selection for automatic tumor classification of mammogram images," in *2011 7th Iranian Conference on Machine Vision and Image Processing, MVIP*, 2011.

[181] H. Alharbi, G. Falzon, and P. Kwan, "A novel feature reduction framework for digital mammogram image classification," *2015 3rd IAPR Asian Conf. Pattern Recognit.*, pp. 221–225, 2015.

[182] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," *J. Mach. Learn. Res.*, vol. 3, no. 3, pp. 1157–1182, 2003.

[183] W. Han, J. Dong, Y. Guo, M. Zhang, and J. Wang, "Identification of masses in digital mammogram using an optimal set of features," *Proc. 10th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. Trust. 2011, 8th IEEE Int. Conf. Embed. Softw. Syst. ICESS 2011, 6th Int. Conf. FCST 2011*, pp. 1763–1768, 2011.

[184] L. Choridah, "Identification of Malignant Masses on Digital Mammogram Imagesbased on Texture Feature and Correlation based Feature Selection," in *6th Intwernational Conference on Information Technology and Electrical Engineering (ICTEE)*, 2014.

[185] X. Liu and J. Tang, "Mass Classification in Mammograms Using Selected Geometry and Texture Features, and a New SVM-Based Feature Selection Method," *Syst. Journal, IEEE*, vol. 8, no. 3, pp. 910–920, 2014.

[186] Y.-J. Huang *et al.*, "Automated Feature Set Selection and Its Application to MCC Identification in Digital Mammograms for Breast Cancer Detection," *Sensors*, vol. 13, no. 4, pp. 4855–4875, 2013.

[187] T. M. Khoshgoftaar, A. Fazelpour, H. Wang, and R. Wald, "A survey of stability analysis of feature subset selection techniques," *2013 IEEE 14th Int. Conf. Inf. Reuse Integr.*, pp. 424–431, 2013.

[188] A. C. Haury, P. Gestraud, and J. P. Vert, "The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures," *PLoS One*, vol. 6, no. 12, 2011.

[189] A. Kalousis, J. Prados, and M. Hilario, "Stability of Feature Selection Algorithms," *Fifth IEEE Int. Conf. Data Min.*, pp. 218–225.

[190] M. Naseriparsa, A.-M. Bidgoli, and T. Varaee, "A Hybrid Feature Selection Method to Improve Performance of a Group of Classification Algorithms," *Int. J. Comput. Appl.*, vol. 69, no. 17, pp. 975–8887, 2013.

[191] A. I. Technology, "Mammogram Classification Using Maximum Difference Feature Selection Method," *J. Theor. Appl. Inf. Technol.*, vol. 33, no. 2, 2011.

[192] S. Shanthi and V. M. Bhaskaran, "Modified Artificial Bee Colony Based Feature Selection : A New Method in the Application of Mammogram Image Classification," vol. 3, no. 6, pp. 1664–1667, 2014.

[193] M. N. Sudha and S. Selvarajan, "Feature Selection Based on Enhanced Cuckoo Search for Breast Cancer Classification in Mammogram Image," *Circuits Syst.*, vol. 7, no. 04, p. 327, 2016.

[194] M. Aswini, S. K. Lenka, and S. Manas, "A novel image mining technique for classification of mammograms using hybrid feature selection," *Int. J. Adv. Res. inComputer Sci.*, vol. 2, no. 3, pp. 1151–1161, 2013.

[195] S. K. L. Aswini Kumar Mohanty, Sukanta Kumar Swain, Pratap Kumar Champati, "Image Mining for Mammogram Classification by Association Rule Using Statistical and GLCM features," *J. Comput. Sci.*, vol. 8, no. 5, pp. 309–318, 2011.

[196] H. Karim and K. Zand, "A Comparative Survey on data mining techniques for breast cancer diagnosis and prediction," *Int. J. Fundam. Appl. Life Sci.*, vol. 5, no. 2005, pp. 4330–4339, 2015.

[197] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, "Supervised Machine Learning: A Review of Classification Techniques," *Informatica*, vol. 31, pp. 501–520, 2007.

[198] J. Han and M. Kamber, *Data mining : Concepts and Techniques*. 2006.

[199] K. M. Han Jiawei, *Data mining concepts and techniques*. 2006.

[200] M. D. Malkauthekar, "Analysis of Euclidean distance and Manhattan distance measure in face recognition," *IET Conf. Publ.*, vol. 2013, no. CP646, pp. 503–507, 2013.

[201] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, and D. I. Fotiadis, "Machine learning applications in cancer prognosis and prediction," *Comput. Struct. Biotechnol. J.*, vol. 13, pp. 8–17, 2015.

[202] S. Sarhan, "An Ensemble Classifier Based on SVM,KNN and NB for diagnosing Erythemato-Squamous Disease," *Res. Gate*, no. January, 2018.

[203] G. Texier, R. S. Alldoji, L. Diop, J. B. Meynard, L. Pellegrin, and H. Chaudet, "Using decision fusion methods to improve outbreak detection in disease surveillance," *BMC Med. Inform. Decis. Mak.*, vol. 19, no. 1, pp. 1–11, 2019.

[204] S. H. Park, "Receiver Operating Characteristic (ROC) Curve : Practical Review,"

vol. 5, no. March, 2004.

[205] T. Shultz and S. Fahlma, "Confusion Matrix. Encyclopedia of machine learning," 2011.

[206] C. Oprea and P. Ş Ti, "Perofrmance Evaluation of the Data Mining Classification Methods," pp. 249–253, 2014.

[207] Q. Gu, L. Zhu, Z. Cai, and C. Science, "Evaluation Measures of the Classification Performance of Imbalanced Data Sets," *Comput. Intell. Intell. Syst.*, pp. 461–471, 2009.

[208] C. J. V. R. B. Sc and D. Ph, "Information Retrieval," 1979.

[209] D. C. Moura and M. A. Guevara, "An evaluation of image descriptors combined with clinical data for breast cancer diagnosis," pp. 561–574, 2013.

[210] K. Bowyer, D. Kopans, P. K. Jr, R. Moore, K. Chang, and S. Munishkumaran, "Current Status of the Digital Database for Screening Mammography," pp. 457–460.

[211] P. K. M. Heath, K. Bowyer, D. Kopans, R. Moore, "The Digital Database for Screening Mammography." pp. 1–10.

[212] C. Moreira, A. Cardoso, and S. Jaime, "INbreast:Toward a Full- Field Digital Mammographic Database," 2012.

[213] J. Sunkling, "The mammographic image analysis society digital mammogram database." 2014.

[214] R. M. Rangayyan and N. R. I. Mudigonda, "Boundary modelling and shape analysis methods for classification of mammographic masses," *Med. Biol. Eng. Comput.*,

vol. 38, 2000.

[215] A. Oliver, J. Freixenet, J. Martí, E. Pérez, J. Pont, and E. R. E. Denton, "A review of automatic mass detection and segmentation in mammographic images," *Med. Image Anal. Elsevier*, vol. 14, no. 2, pp. 87–110, 2010.

[216] J. E. E. Oliveira, M. O. Gueld, A. D. A. Araújo, B. Ott, and T. M. Deserno, "Towards a Standard Reference Database for Computer-aided Mammography," in *Proceedings of SPIE*, 2008, vol. 6915, pp. 1–9.

[217] M. Angel, G. Lopez, and D. C. Moura, "BCDR : A Breast Cancer Digital Repository," *Res. Gate*, no. September 2017, 2012.

[218] T. Netsch and H. Peitgen, "Scale-Space Signatures for the Detection of Clustered Microcalcifications in Digital Mammograms," vol. 18, no. 9, pp. 774–786, 1999.

[219] Z. C. Antoniou *et al.*, "A web- accesible mammographic image database dedicated to combine training and evaluation of radiologists and machines," 2009.

[220] J. asjit S. Suri, S. K. Setarehdan, and S. Singh, *Advanced Algorithms Approaches to Medical Image Segmentation*. 2002.

[221] A. Oliver, J. Pont, E. R. E. Denton, and J. Freixenet, "A Statistical Approach for Breast Density Segmentation," vol. 23, no. 5, pp. 527–537, 2010.

[222] A. O. Malagelada, *Automatic mass segmentation in mammographic images*, no. May. 2016.

[223] M. Budka, B. Gabrys, and S. Member, "Density-Preserving Sampling : Robust and Efficient Alternative to Cross-Validation for Error Estimation," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 24, no. 1, pp. 22–34, 2013.

[224] J. R. Nolan, "Computer Systems that learn: An empirical study on the effect of noise on the performance of three classification methods."

[225] A. M. Molinaro, R. Simon, and R. M. Pfeiffer, "Prediction Error Estimation : A Comparison of Resampling Methods," 2005.

[226] R. Struart and N. Peter, *Artificial Intelligence A Modern Approach*. 2010.

[227] W. Dubitzky, M. Granzow, and D. Berrar, *Fundamentals of data mining in genomics and proteomics*. 2006.

[228] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data mining, Iference and Prediction*. 2001.

[229] R. Angulu, J. R. Tapamo, and A. O. Adewumi, "Age estimation via face images : a survey," *EURASIP J. Image Video Process.*, 2018.

[230] A. Celisse, "Optimal cross-validation in density estimation with the," *Ann. Stat.*, vol. 42, no. 5, pp. 1879–1910, 2014.

[231] B. Efron and R. Tibshirani, "Bootstrap Methods for Standard Errors , Confidence Intervals , and Other Measures of Statistical Accuracy," *Stat. Sci.*, vol. 1, no. 1, pp. 54–75, 1986.

[232] A. Oliver, X. Llad, J. Freixenet, and J. Mart, "False Positive Reduction in Mammographic Mass Detection Using Local Binary Patterns," *Med. Image Comput. Comput. Interv.*, pp. 286–293, 2007.

[233] J. Chen, G. Zhao, and M. Pietikainen, "RLBP : Robust Local Binary Pattern," in *British machine vision conference 2013*, 2013, no. November 2018.

[234] R. Rabidas, A. Midya, J. Chakraborty, and W. Arif, "A Study of Different Texture

Features Based on Local Operator for Benign-malignant Mass Classification," in *6th International Conference on Advances in Computing & Communications, ICACC*, 2016, pp. 389–395.

[235] R. Rabidas, A. Midya, A. Sadhu, and J. Chakraborty, "Benign-Malignant Mass Classification in Mammogram using Edge Weighted Local Texture Features," in *SPIE*, 2016, vol. 9785, pp. 1–6.

[236] S. Jamal, S. Gardezi, and I. Faye, "Fusion of Completed Local Binary Pattern Features with Curvelet Features for Mammogram Classification," *Appl. Math. Inf. Sci. An Int. J.*, vol. 12, no. 6, pp. 1–12, 2015.

[237] N. Ponraj, J. Winston, Poongodi, and M. Mercy, "Novel Local Binary Textural Pattern for Analysis and classification of mammogram using Support Vector Machine," in *International Conference on Signal Processing and Communication (ICSPC'17)*, 2017, pp. 380–383.

[238] M. Abdel-nasser, H. Rashwan, D. Puig, and A. Moreno, "Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern," *Expert Syst. Appl.*, no. August, 2015.

[239] A. Rampun, P. Morrow, B. Scotney, and J. Winder, "Breast Density Classification Using Local Ternary Patterns in Mammograms," in *Research Gate*, 2017, no. June.

[240] C. Muramatsu, T. Hara, T. Endo, and H. Fujita, "Breast Mass Classification on mammograms using Radial Local Ternart Patterns," *Comput. Biol. Med.*, 2016.

[241] S. Paramkusham, K. . Rao, and P. Rao, "Novel technique for the detection of abnormalities in Mammograms using texture and geometric features," in

International Conference on Microwave, Optical and Communication Engineering, ICMOCE, 2016, pp. 150–153.

[242] I. Faye, B. B. Samir, and M. M. M. Eltoukhy, "Digital mammograms classification using a wavelet based feature extraction method," *2009 Int. Conf. Comput. Electr. Eng. ICCEE 2009*, vol. 2, no. 1, pp. 318–322, 2009.

[243] M. M. Eltoukhy, I. Faye, and B. B. Samir, "Curvelet Based Feature Extraction Method for Breast Cancer Diagnosis in Digital Mammogram."

[244] M. Pratiwi, Alexander, J. Harefa, and S. Nanda, "Mammograms Classification Using Gray-level Co-occurrence Matrix and Radial Basis Function Neural Network," *Procedia Comput. Sci.*, vol. 59, no. Iccsci, pp. 83–91, 2015.

[245] R. Biswas, A. Nath, and S. Roy, "Mammogram classification using gray-level co-occurrence matrix for diagnosis of breast cancer," *Proc. - 2016 Int. Conf. Micro-Electronics Telecommun. Eng. ICMETE 2016*, pp. 161–166, 2017.

[246] T. T. Htay and S. S. Maung, "Early Stage Breast Cancer Detection System using GLCM feature extraction and K-Nearest Neighbor (k-NN) on Mammography image," in *18th International Symposium on Communications and Information Technologies (ISCIT)*, 2018, no. Iscit, pp. 171–175.

[247] H. D. Ellis, J. W. Shepherd, and G. M. Davies, "Identification of familiar and unfamiliar faces from internal and external features : some implications for theories of face recognition," vol. 8, no. 1977, pp. 431–439, 1979.

[248] K. Gaurav and K. B. Pradeep, "A Detailed Review of Feature Extraction in Image Processing Systems," in *Fourth International Conference on Advanced Computing*

& Communication Technologies, 2016, no. February.

[249] I. Guyon, "An Introduction to Variable and Feature Selection 1 Introduction," vol. 3, pp. 1157–1182, 2003.

[250] M. Saunders, P. Lewis, and A. Thornhill, *Research methods for business students fi fth edition*. 2009.

[251] Flick, "Introducing Research Methodology," *SAGE*, no. 800, p. 9774, 2015.

[252] M. Saunders, P. Lewis, and A. Thornhill, *Research Methods for Business Students*. 2016.

[253] B. R. N. Matheus and H. Schiabel, "Online mammographic images database for development and comparison of CAD schemes," *J. Digit. Imaging*, vol. 24, no. 3, pp. 500–506, 2011.

[254] I. K. Maitra, S. Nag, and S. K. Bandyopadhyay, "Technique for preprocessing of digital mammogram," *Comput. Methods Programs Biomed.*, vol. 107, no. 2, pp. 175–188, 2011.

[255] M. Dong, Z. Wang, C. Dong, X. Mu, and Y. Ma, "Classification of Region of Interest in Mammograms Using Dual Contourlet Transform and Improved KNN," *J. sensors*, vol. 2017, 2017.

[256] X. Wang, B. S. Wong, and T. C. Guan, "<title>Image enhancement for radiography inspection</title>," *PROC.SPIE*, vol. 5852, pp. 462–468, 2005.

[257] A. Daskalakis, D. Cavouras, and P. Bougioukos, "an Efficient Clahe-Based, Spot-Adaptive, Image Segmentation Technique for Improving Microarray Genes quantification," *Exp.*, pp. 4–7, 2007.

[258] J. E. Ball and L. M. Bruce, "Digital mammogram spiculated mass detection and spicule segmentation using level sets," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, pp. 4979–4984, 2007.

[259] A. A. Kayode, B. S. Afolabi, and B. O. Ibitoye, "An Explorative Survey of Image Enhancement Techniques Used in Mammography," *Int. J. Comput. Sci.*, vol. 12, no. 1, pp. 72–79, 2015.

[260] I. Pozzi, R. Nusselder, and D. Zambrano, "Further Advantages of Data Augmentation on Convolution Neural Networks," in *27th international Conference on Artificial Neural Networks*, 2018, no. June 2019, pp. 284–293.

[261] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, 2019.

[262] S. Chapman, *MATLAB Programming for Engineers*, 4th editio. 2008.

APPENDICES

Appendix 1: Proposal approval letter from the board of postgraduate studies of MMUST

MASINDE MULIRO UNIVERSITY OF SCIENCE AND TECHNOLOGY (MMUST)

Tel: 056-30870
Fax: 056-30153
E-mail: directordps@mmust.ac.ke
Website: www.mmust.ac.ke

P.O Box 190
Kakamega – 50100
Kenya

Directorate of Postgraduate Studies

Ref: MMU/COR: 509099

Date: 22ᴺᴰ November, 2018

Mary Walowe Mwadulo,
SIT/LH/005/2015
P.O. Box 190-50100,
KAKAMEGA.

Dear Ms. Walowe,

RE: APPROVAL OF PROPOSAL

I am pleased to inform you that the Directorate of Postgraduate Studies has considered and approved your Ph.D proposal entitled: "*Breast Cancer Detection Using Enhanced Mammographic Features.*" and appointed the following as supervisors:

1. Dr. Stephen Makau Mutua   - School of Computing and Informatics, MMUST
2. Dr. Raphael Angulu   - School of Computing and Informatics, MMUST

You are required to submit through your supervisor(s) progress reports every three months to the Director Postgraduate Studies. Such reports should be copied to the following: Chairman, School of Computing and Informatics Graduate Studies Committee and Chairman, Computer Science Department. Kindly adhere to research ethics consideration in conducting research.

It is the policy and regulations of the University that you observe a deadline of three years from the date of registration to complete your Ph.D thesis. Do not hesitate to consult this office in case of any problem encountered in the course of your work.

We wish you the best in your research and hope the study will make original contribution to knowledge.

Yours Sincerely,

Prof. John Obiri
DIRECTOR, DIRECTORATE OF POSTGRADUATE STUDIES

206

Appendix II: Research Permit from NACOSTI

NATIONAL COMMISSION FOR
SCIENCE, TECHNOLOGY & INNOVATION

Ref No: 362668

Date of Issue: 09/August/2019

RESEARCH LICENSE

This is to Certify that Ms. MARY MWADULO of Masinde Muliro University of Science and Technology, has been licensed to conduct research in Meru on the topic: COMPUTER AIDED TEXTURE FEATURE EXTRACTION APPROACH FOR BREAST CANCER DETECTION USING MAMMOGRAPHY IMAGES for the period ending : 09/August/2020.

License No: NACOSTI/P/19/179

362668

Applicant Identification Number

Director General
NATIONAL COMMISSION FOR
SCIENCE, TECHNOLOGY &
INNOVATION

Verification QR Code

NOTE: This is a computer generated License. To verify the authenticity of this document, Scan the QR Code using QR scanner application.

207