A survey of the methods used to classify breast density in mammograms and ultrasound images

Ângela Oliveira[1,2], *José Alberto Fonseca-Moutinho*[3], *Manuela Pereira*[1] *and Mário M. Freire*[1]
[1]Instituto de Telecomunicações, Department of Computer Science, University of Beira Interior, Covilhã, Portugal
[2] Department of Informatics, School of Technology of the Polytechnic Institute of Castelo Branco, Portugal
[3] Faculty of Health Sciences, University of Beira Interior, Covilhã, Portugal

Corresponding Author: Ângela Oliveira

Address: Instituto de Telecomunicações,
Department of Computer Science, University of Beira Interior,
Rua Marquês de Ávila e Bolama
6200-001 Covilhã,
Portugal
Phone number: 00351 962743531

Abstract

Introduction: Breast cancer is a serious disease that affects millions of people, most of whom are women. Breast density has been shown to be an important risk factor and must be considered in breast cancer screening and prevention. This review article highlights the importance of breast density as a risk factor associated with breast cancer, both as a single factor and associated with other known risk factors.

Objective: The objective of this article is to analyze the methods used to evaluate breast density in the most common complementary diagnostic procedures used by radiologists: mammography and breast ultrasound.

Conclusion: Many methods are used to calculate breast density using mammography, but there are fewer methods for evaluating breast density using ultrasound. The set of computational methods used to evaluate breast density in ultrasounds is difficult to apply in practice. Given the importance of ultrasound in the diagnosis of breast cancer, the specification of breast density calculation methods for this type of supplementary means of diagnosis is relevant. Some of the most commonly used methods in mammography do not provide satisfactory results when they are applied in breast ultrasound. Nevertheless, this analysis provides a starting point to further research in breast density assessment in ultrasound.

Keywords: breast cancer; breast density; breast density evaluation; breast density classification.

1. Introduction

2. Cancer is a scourge that affects many people [1]. Breast cancer is associated with an image of severe gravity because it affects a body part full of symbolism in motherhood and femininity. Breast cancer is a malignant tumor that develops in the cells of the breast tissue. It is much more common in women, although it can also affect men [2].
3. In Europe, the incidence of breast cancer in 2008 within the 27 member states of the European Union (EU) [3] Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovenia, Slovakia, Finland, Sweden and the United Kingdom and three European Free Trade Association (EFTA) countries, i.e., Iceland, Norway and Switzerland, totaled 332,771 cases, which corresponds to approximately 7% of the population, and a mortality rate of 89,797, which

corresponds to approximately 3% in 27 countries and a total of approximately 500 million people. The distribution of breast cancer in these 27 countries is illustrated in the chart of Figure 1.

4.The most recent data on the incidence of cancer in Portugal, which is included in the National Cancer Registry 2001 [4], shows that breast cancer ranks second and represents 14% of registered cancers. According to the Portuguese Institute of Oncology (IPO) in Porto [5] data, from the period of 1989 to 2009 in Northern Portugal, breast cancer has significant values of malignant tumors, which first appear in 30-40% of the cases.

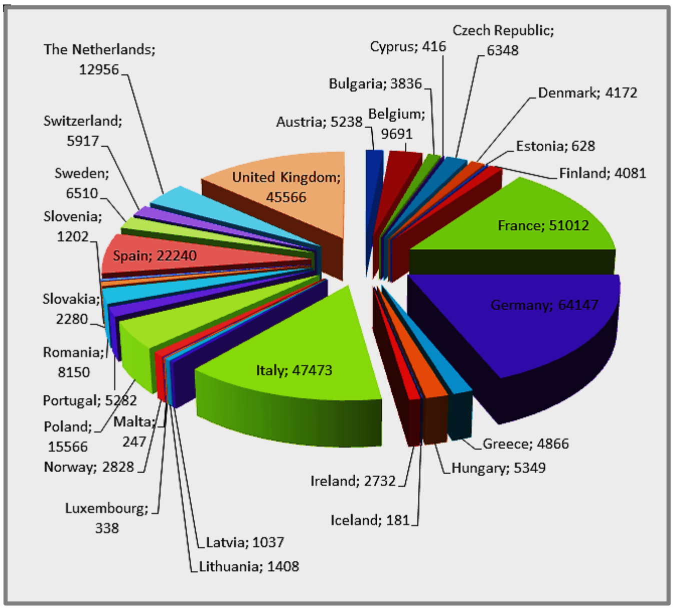


Figure 1- Statistical Data from EEC and EFTA – Number of cases of breast cancer in EEC and EFTA [3].

In Figure 2, the percentage of breast cancer compared with other types of cancer is presented; breast cancer is in 2nd place in the list of the highest number of cases in 1990 and in 1st place in the remaining years.

The number of patients who received diagnoses during 1989-2009 increased. Society has shown a greater concern with the possibility of having any type of cancer. This interest is important because early detection and subsequent treatment indicates a lower mortality rate. Figures 2 and 3 illustrate not only the importance of breast cancer compared with other types of cancer but also a significant increase in the cases treated over the years and in the number of cases where a malignant tumor exists.

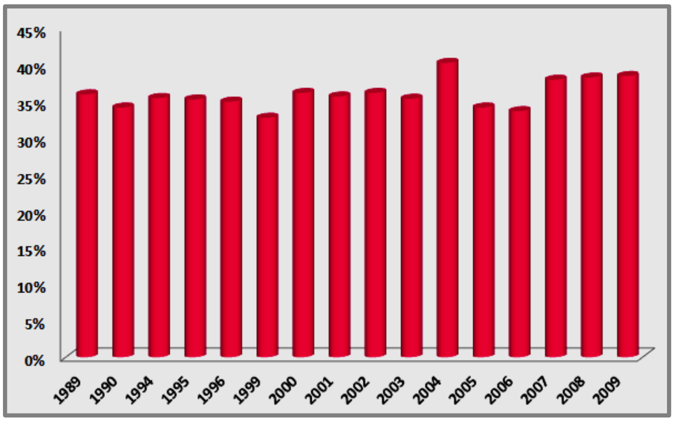
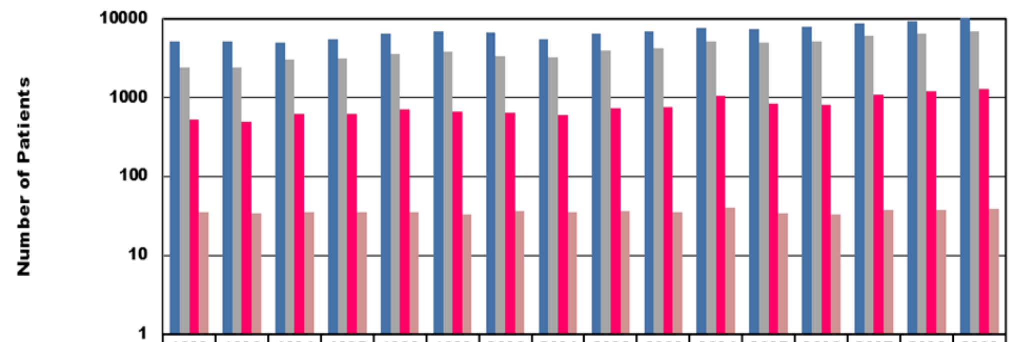


Figure 2- Percentage of breast cancer in different years based on the Statistical Data from the Portuguese Institute of Oncology of Porto [5].



	1989	1990	1994	1995	1996	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Patients Received	5195	5075	5050	5453	6502	6931	6789	5584	6501	6878	7573	7500	7808	8823	9452	10650
Malignant Tumors	2410	2430	3036	3176	3614	3829	3408	3256	3975	4277	5125	5052	5171	6085	6460	6880
Breast Cancer	532	493	614	622	710	669	655	610	732	772	1048	842	820	1086	1223	1281
Breast Cancer (%)	35,9	34,1	35,4	35,2	34,9	32,7	36,1	35,6	36,1	35,3	40,2	34,1	33,6	37,9	38,2	38,4

Figure 3- - The number of patients receiving a diagnosis with malignant tumors, namely breast cancer, in women - Statistical Data from the Portuguese Institute of Oncology in Porto [5].

According to IPO, in 2000-2001, women who had been observed during the previous five-year period had a relative survival of 83% in 99% of the breast cancer cases detected in those years [5].

Considering the relationship between the demographics in European countries [6] and the number of breast cancer cases in each country in 2008, breast cancer generally occurred more often in more densely populated countries, where there is an increased number of breast cancer cases. Figure 4 illustrates the relationship between the number of breast cancer cases and the population of each of the 27 countries of the European Union plus the 3 countries of the European Free Trade Association.

The American Cancer Society - Breast Cancer Facts and Figures 2011 [7] shows that 230,480 new cases of breast cancer were estimated to occur in 2011 in the United States. This cancer ranks second in cause of death. However, the same study stated that although the incidence of mortality of breast cancer continues to be relevant, the decreasing rate is directly related to early diagnosis. In a similar study in 2007 by the European Cancer Organization [8], similar conclusions were obtained: it is essential to early detect and screen to prevent and cure disease. Ferlay [3] also found that breast cancer ranked second, with a percentage of 13%.

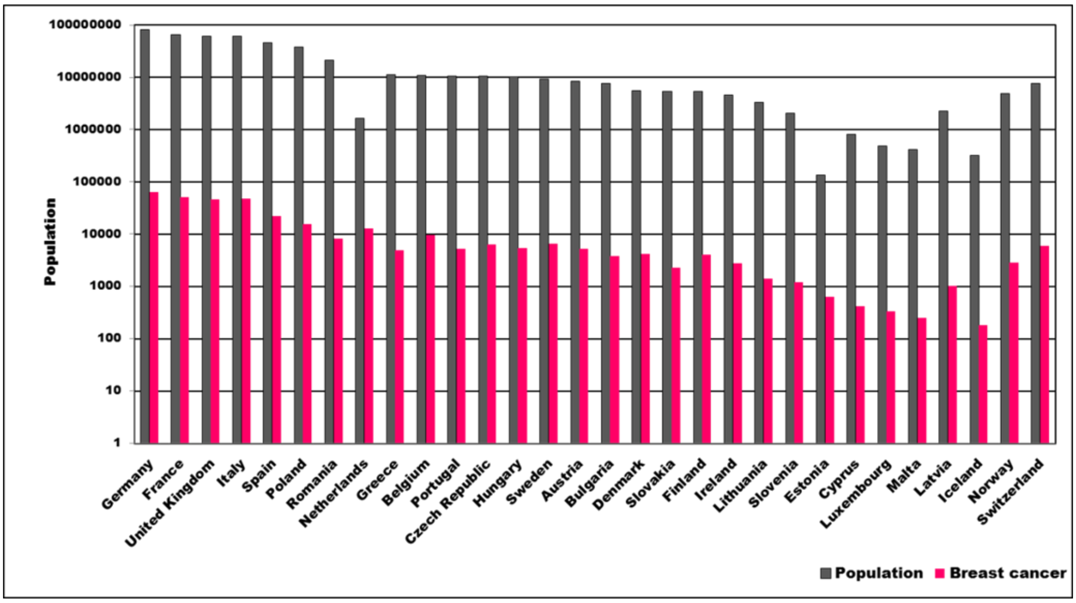


Figure 4- - Relation between population from EU and the number of breast cancer cases [3].

The National Breast and Ovarian Cancer Center, in a 2009 publication [9], considered the following risk factors: moderate to strongly increased, slightly increased or decreased. The moderate to strongly increased group accounts for factors such as **sex** because women are naturally prone to breast cancer: they are 100 times more likely to have breast cancer than are men. A**ge** is another factor that influences the propensity to having breast cancer; older women are at a higher risk. Studies indicate that 75% of breast cancer occurs in women over the age of fifty. **Affluent** countries that are highly populated show a higher number of breast cancer cases, as shown in Figure 1. Genetics are another risk factor, as indicated by **family history.** Women who have had first-degree relatives, such as a mother or sister, with breast cancer show a higher propensity towards the disease. This risk increases proportionally with the number of first-degree relatives who have breast cancer. Breast condition is also a risk factor and includes **breast density**. Women with high breast density present a four to six times increased risk compared with women who have low breast density. Women presenting higher levels of estrogen have a two-fold increased risk.

For the slightly increased or decreased group, **hormonal factors,** such as reproductive history, menstrual history, menopausal status and exogenous hormone, and **personal lifestyle** are important factors to consider. However, they depend on women's habits, namely overweight and obesity, alcohol consumption, and physical activity.

Other factors that have been considered risk factors have no evidence of support in this study, including factors such as pregnancy termination or abortion, smoking or environmental pollutants. Because the causes of breast cancer are still undetermined, early detection using medical examinations is important. The earlier that signs of the disease are detected, the greater the likelihood of successful healing is. In addition to self-examination, there is a set of medical tests that can be used to identify the presence or absence of the disease: a clinical breast examination and imaging tests, which may include mammography or ultrasound and can be performed with some frequency, especially after the age of 40. These are primarmily used for early detection, whereas tests such as MRI (Magnetic Resonance Imaging) and biopsy are performed when there are signs of cancer.

William Black and Gilbert Welch [10] reported the importance of tests such as mammography and ultrasound. The article discusses the advantages of tests using technology of many types, focusing on the tests used to detect cancer. A reduction of mortality from breast cancer in women was found for those who had mammograms, which is one of the most commonly performed tests. The risk factor that has become increasingly important is breast density, which is therefore the focus of this paper.

2. Breast density as a risk factor

Presentation of the breast varies from woman to woman and depends on breast composition, including both glandular and fat tissue. Therefore, in exams such as mammography or breast ultrasound, breast tissue presents itself differently: darker regions indicate fat, and clearer regions indicate glandular tissue. Breast density is a way to describe the types of tissue that make up the breast. The breast is made up of glandular or ductal tissue, fibrous connective tissue and fatty tissue. The amount of each of these tissues varies in women. Women who have more fibrous connective and glandular tissue than fatty tissue have greater breast density. Breast density depends on factors such as the number of children, weight and age. Breast density is measured according to the presence of a higher or lower amount of fat in the breast tissue. Because the most common exam is mammography, almost every study was developed considering mammograms.

As a relevant risk factor, studies published have considered breast density to be important since 1976, when Wolfe [11] established a relationship between the mammary gland density and the risk of breast cancer. Several studies have been developed in this area, thus giving relevance to breast density based on different approaches:

- “Mammographic Densities and Breast Cancer Risk” [12] analyzes the literature published during 1976-1997 under terms such as mammography and breast cancer risk. Qualitative and quantitative methods for classifying parenchymal patterns are presented and compared. In this study, mammographic densities and other risk factors for breast cancer are analyzed, and mammographic densities are identified as an independent risk factor.
- The literature review “Applications and literature review of the BI-RADS classification” [13] concerns the usefulness and limitations of the BI-RADS lexicon.
- “Breast Density and Parenchymal Patterns as Markers of Breast Cancer Risk: A Meta-analysis” [14] states that mammographic features are associated with the risk of breast cancer. This association varies considerably between studies, and it is uncertain whether this relevance is modified when associated with other risk factors. Reviews of other studies have been performed, and the findings revealed that breast density is one of the strongest risk factors. Therefore, more consideration should be given to the routine measurement of mammographic density because this marker has potential to be used for the research and prevention of breast cancer.
- In the study “A Comparison of Breast Tissue Classification Techniques” [15], different strategies for extracting features from tissue and their classification systems are reviewed, and the feasibility of estimating breast density by using automatic computer vision techniques and the benefits of segmentation of the breast based on internal tissue information are demonstrated.
- The study “Breast Image Registration Techniques: a survey” [16] gives an overview of the current state-of-the-art in the breast image registration techniques: Image registration; and reviews literature on intra-modality breast image registration on the design of co-registered multimodality breast imaging acquisition systems and validation of breast registration methods.

- The study “Comparison between Wolfe, Boyd, BI-RADS and Tabár Based Mammographic Risk Assessment” [17] provides a comparative study of the Wolfe, Boyd, BI-RADS and Tabár-based assessment approaches for mammographic image classification methods.
- In the study “Comparing Measurements of Breast Density” [18], the authors undertake a theoretical analysis of physical breast density definitions and area versus volumetric estimation techniques and analyze both the images and the results of applying the various techniques.
- In the study “Mammographic Density. Measurement of mammographic density” [19], Martin Yaffe reviews the techniques for measuring density and gives some consideration for strengths and limitations.
- The paper “Automated breast cancer detection and classification using ultrasound images: A survey” [20] reviews Computer-aided design (CAD) systems for breast cancer detection and classification using ultrasound images and summarizes the techniques developed. The advantages and disadvantages are discussed, different performance evaluation metrics are studied and future developments and trends are also investigated.
- The paper “Automatic Breast Density Segmentation: an integration of different approaches” [21] states that in most studies, breast density is assessed by using a user-assisted threshold method that is both time-consuming and subjective. In this study, the authors develop a breast density segmentation method that is fully automatic and is based on pixel classification, considering different approaches known in literature, such as breast density segmentation.
- The paper “A review of automatic mass detection and segmentation in mammographic images” [22] reviews the existing approaches for automatic detection and segmentation of masses in mammographic images. The advantages and disadvantages of the various approaches are demonstrated.

The above studies show the interest in breast density and its relevance for risk of breast cancer. The studies mention classification systems that involve breast density and image analysis techniques, thus revealing a great interest in image processing and breast density evaluation. Since Wolfe’s publication in 1976, several studies have considered the relationship between breast density and the risk of breast cancer.

Both Boyd [23] and Ursin [24] reported that women with a high mammary density have a greater probability of developing breast cancer. A new model based on the Gail Model [25] was built, and breast density was added as a risk factor, as presented by Chen [26]. The conclusions determined that with the newly developed model, women with a higher breast percentage density have a higher propensity for breast cancer.

In 2006, Titus-Ernstoff presented [27] a study that evaluates the risk factors of breast cancer that are associated with breast density. Although the density of the breast is a risk factor for breast cancer, the study suggested that with more efficient mammary density measurements, a higher amount of consistent studies can be produced.

In 2007, Boyd presented [28] a study using mammography that related breast density to the risk of cancer. In the conclusion of that study, breast density was reported to be strongly associated with an increased risk of breast cancer: the higher the density, the higher the risk of cancer; further, this risk persists for a considerable period of time. The calculation of assigned risk shows that breast density explains a significant proportion of cases of breast cancer in young women and demonstrates that a large percentage of women have more than 50% mammary density. In 2007, Vachon presented [29] a study where several models were evaluated and to which another risk factor is added, namely breast density. The conclusion was that breast density is important not only in women who are considered to be at risk and are receiving mammograms but also in women at younger ages. In 2008, Jeffrey presented [30] a breast density analysis in a group of women. The conclusions were that women with low breast density had a lower risk, unless they had a family history of breast cancer.

The impact of breast density associated with several risk factors has been widely analyzed to demonstrate that breast density is a risk factor of breast cancer. In 2011, Boyd stated [31] that in future studies, breast density should be improved by calculating the percentage of breast density, which should be included in the definition of individual risk.

In America, there are 19 states with breast density notification laws [32], which require physicians to notify women who present mammographic breast density. This information is sent to women with dense breasts. Working groups of experts in breast imaging and breast cancer risk conducted several studies related to breast density to provide women and radiologists with accurate information [33], [34].

Table 1 shows several studies that aim to address the combination of breast density and other risk factors of breast cancer and thus consider breast density to be a risk factor.

3. Qualitative classification of breast density

Breast density is generally quantified by a technician or doctor who performs the exam, and there are several approaches to this classification [11]. However, this classification in most cases depends on the skills of the person who made the observation and therefore does not allow a uniform assessment.

D’Orsi [46] found that the evaluation of body part thickness shows a thicker density in larger areas. When calculating the size and thickness of the breast, the same standards are followed. According to the same author, the breast can be classified into three broader categories, depending on the relative amounts of glandular tissue versus adipose tissue [46]:

- Glandular breast: In general, a young breast is denser because it contains a relatively small amount of fat tissue. This usually occurs in women under 30 years; however, an older woman who has never carried a pregnancy to completion may also be included in this category. Nevertheless, pregnant and lactating women can still be included in this category.
- Fatty and glandular breast: With increasing age, the tendency is for the fat in the breast to increase. There is an approximation in the amount of fatty tissue and glandular tissue. Women between 30 and 50 years old are usually included in this category.
- Fatty breast: Women aged 50 and/or upon the occurrence of menopause are included in this category. With the end of reproductive life, the breast loses fibrous mass and turns into fat.

Although it may be based on the above categories, the process of quantifying breast density is not exact, and several approaches have been defined over time. In most situations, the quantification of breast density is performed by the technician who is performing the test. There are several ways of classifying it [28], the most common of which involves two radiologists analyzing breast density and distinguishing it as fitting into one of the following six categories: 0%, <10%, from 10% and <25%, from 25% and <50%, from 50% and <75%, and >75%.

.

Table 1- Summary of studies considering breast density associated with other risk factors.

Risk Factors \ Studies	Maskarinec G et al [35]	Maskarinec G et al [36]	Vachon C et al [37]	Barlow W et al [38]	Palomares M et al [39]	Mitchell G et al [40]	Boyd N et al [41]	. Boyd N et al [42]	Kerlikowske K et al [43]	Yaghjyan L et al [44]
Ethnicity	✓	✓		✓						
Age		✓	✓	✓	✓		✓	✓	✓	
Residence		✓	✓							
Exam Date			✓							
Menopause Status			✓	✓			✓	✓		
Age at Menopause				✓			✓			
Type of Menopause				✓						
Number of Mammograms			✓							
Postmenopausal										✓
Interval Between Mammograms			✓	✓						
Age at Menarche				✓	✓		✓	✓		
Age at the Birth of First Child				✓	✓		✓	✓		
Number of Live Births							✓			
Use of Hormone Therapy				✓			✓			
Not use of Hormone Therapy									✓	
Personal History of Breast Cancer				✓	✓					✓
Family History of Breast Cancer				✓	✓			✓		
First-degree Relatives with Breast Cancer					✓		✓			
Number of Breast Biopsies					✓					
Atypical Hyperplasia					✓					
BRCA1 Mutation Carriers or Not						✓				
BRCA2 Mutation Carriers or Not						✓				
Height (cm)							✓	✓		
Weight (kg)							✓	✓		
BMI (Body Mass Index)							✓			

In conclusion, breast density is an important risk factor, but a method for assessing breast density is also important. As shown in Tables 2 and 3, several algorithms have been developed to evaluate breast density by ultrasound .

Among the most widely used classifications, the Wolfe classification [11] is based on mammograms and is a visual classification method that can be defined as follows:

- N1 corresponds to fatty normal breast;
- P1 corresponds to prominent ducts occupying less than 25% of the breast;
- P2 corresponds to prominent ducts occupying between 25% and 75% of the breast; and
- Dy corresponds to breast dysplasia and is extremely dense.

The classification of Breast Imaging Reporting and Data System - BI-RADS [17, 46], which is based on standard reports for viewing mammograms developed by the American College of Radiology (ACR) [47], is divided into the following categories:

- Category 1: Breast is mostly made up of fat <25% breast density.
- Category 2: Approximately 25% to 50% breast density.
- Category 3: Approximately 51% and 75% breast density.
- Category 4: Extremely dense > 75% breast density.

In Table 2 of [29], the classification of BI-RADS and Wolfe, are presented as qualitative classifications. Three methods of quantitative classification are also presented:

- First, acetate is placed in superposition on the mammography image, and a technician searches for areas of breast density. The total amount of breast density is measured using a delimitation tool. The percentage of breast density is assigned on a scale of 0% to 100% and is then fitted into five levels of 0%, 1% to 24%, 25% to 49%, 50% to 74% and larger or equal than 75%.
- Second, a computer-assisted method is used, in which mammograms are digitized, and two initial points are selected. The first separates the background image of the breast, and the second identifies the boundary of the dense tissue. In the calculated pixels, some represent the total breast area and others represent the dense area, thus providing a formula to calculate the breast density percentage.
- Third, breast density is classified by experts in radiology [28].

4. Breast density evaluation

Digital Image Processing, dating from the 1960s, was developed by research projects at NASA in the United States of America. Shortly thereafter, studies that require the knowledge provided by this method, such as medicine, microscopy, and meteorology, appeared.

Since the discovery of X-rays by Wilhelm Konrad Roentgen in 1895, medical images have become an important resource and are widely used in the practice of medicine. There are many methods, approaches and objectives for medical image processing. The methods allow doctors to noninvasively inspect the human body for abnormalities and allow for fast diagnostic decisions [48].

4.1 Computer-based Image Processing

Computer vision addresses theories and algorithms for automating the process of visual perception and involves tasks such as noise removal, smoothing, edge sharpening, image segmentation to isolate object regions, and interpretation. Therefore, image processing may be defined as applying a series of processes of acquisition, correction, improvement, or image compression and processing to improve image quality and information. To perform breast density evaluation, the following steps, as illustrated in Figure 5, were defined [49].

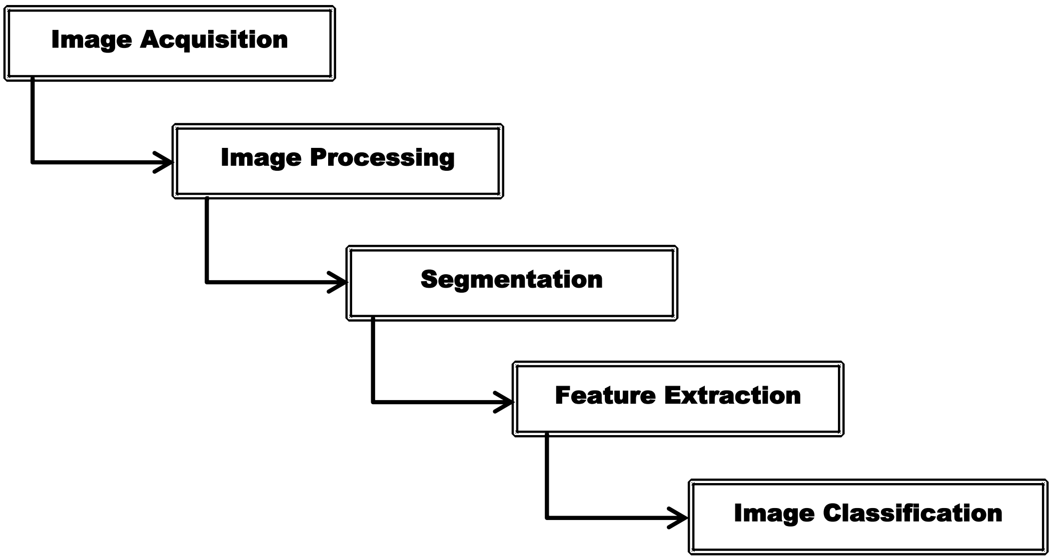


Figure 5- Steps for processing medical images [49].

Image acquisition – Generally, scanning an image means making it computationally manageable. When transforming an image to a digital form, it is necessary to convert it into a signal. The definition of this signal, which represents the image, is a process where each pixel is represented by an integer value proportional to the brightness and color at the corresponding point in the image.

Image Preprocessing - Processing techniques are used to improve some aspects of the image such as mitigating noise and enhancing contours; edge detection; image registration; and improving the characteristics of intensity, color and texture.

Segmentation – Dividing an image into distinct regions where the pixels of each region have similar characteristics. The success of image analysis depends on an effective image segmentation process. There are different approaches and different ways of performing the segmentation process [50].

Some researchers have proposed the divisions of segmentation as follows:

- Texture segmentation
- Region segmentation

In texture segmentation, segmentation as a graph-cut problem is formulated. Other researchers have considered a partition of a color image based on different modes within the estimated empirical distribution by extracting regions of interest in the image [51-54].

In region segmentation, important information about the structure of the objects in the image is given. Several methods have been proposed to combine color and texture with the contours of the image [55-57].

Another approach considers the divisions to be as follows [58]:

- Non-contextual segmentation techniques
- Contextual segmentation techniques

In a non-contextual technique, the relationships among features of an image are not considered, and image segmentation is performed by considering the global attributes. In contextual segmentation, the features are relevant for the segmentation process. The simplest process for a non-contextual process is thresholding. The input to a thresholding operation is typically a grayscale or color image, and the output is a binary image that represents the segmentation. The binary map contains two values: if the pixel's intensity is higher than the threshold, then it is labelled with a value of one and the pixel is set to white. Conversely, if the pixel's intensity is lower than the threshold, then it is labelled with a value of zero and is set to black. The segmentation depends on both the image property being thresholded and the chosen threshold.

Adaptive thresholding or color thresholding can also be used. In adaptive thresholding, the thresholds change dynamically over the image [59]. In color thresholding, there is more information regarding the pixel levels; thus, it involves partitioning the color space [60]. Contextual segmentation includes a spatial analysis, i.e., each pixel is analyzed, as are its neighboring pixels. In general, context segmentation includes methods such as region growing and merging or splitting techniques [61].

Region growing is a region-based segmentation in which pixels that have similar properties are grouped into a large region. The pixels are grouped together and are marked by principals of similarity and spatial proximity. Region splitting and region merging are opposite methods. The splitting process starts with the whole image, which is recursively divided into sub-regions until a homogeneity condition is satisfied. The merging process starts with a small region and merges regions with similar characteristics.

Another approach separates traditional image segmentation methods into three categories [61]:

- Pixel-based segmentation
- Edge-based segmentation
- Region-based segmentation (described previously)

Pixel-based segmentation corresponds to the thresholding segmentation that were previously presented in non-context segmentation. Edge-based segmentation consists of detecting edges between regions. Some authors consider a fourth method: clustering-based segmentation, which clusters tokens with high similarity (small distance in the feature space).

Feature Extraction - This is the process by which parameters are obtained for use in the classification process, which are, in most situations, derived from segmentation. Image classification is the biggest task after extracting the image characteristics because it classifies the extracted object into a category.

Image Classification - This process depends on the feature that it aims to classify. Different ways of dealing with the variability lead to different ways of classifying images, but two basic image classification strategies are presented [62]:

- Supervised classification: the algorithms for supervised classification are conventional pixel-labelling algorithms. Examples include multidimensional thresholding; Minimum-distance classification; maximum likelihood classification; and support vector machine.

Unsupervised classification – the algorithms for unsupervised classification examine a large number of unknown pixels and divide them into a number of classes based on natural groupings present in the image values. Examples include K-means, fuzzy K-means, hierarchical, and histogram-based clustering.

The process is illustrated in Figure 6.

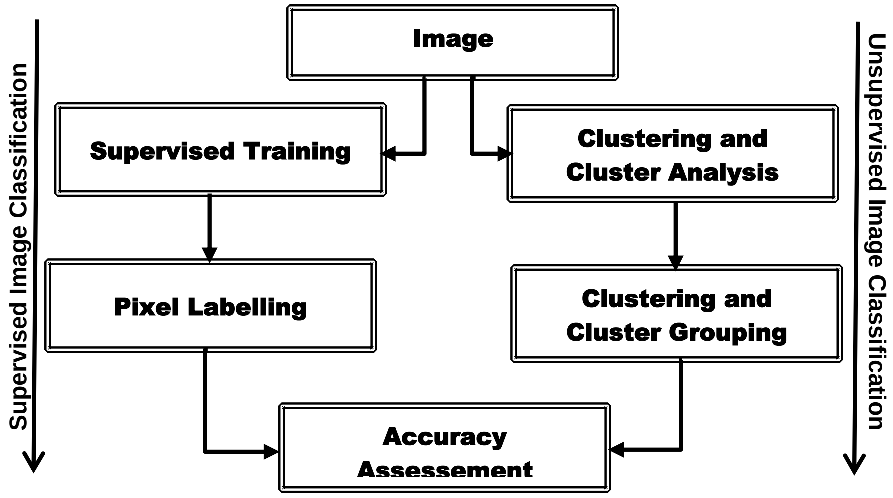


Figure 6- Supervised and unsupervised image classification process.

4.2 Algorithms for breast density evaluation

Breast density was first evaluated according to a qualitative classification scheme, with its origin in the work developed by Wolfe [11]. Quantitative approaches using visual estimation, plan metrics, and computer-assisted methods were later developed. Several studies and different methods for classifying breast density based on mammograms have been proposed. Over time, several approaches to achieve an improvement in the final evaluation have been presented. These studies also use a comparison of the density value calculated by semi-automatic or automatic methods with a value assigned by an expert in accordance with a system of classification to fit the results into a category. In breast density evaluation, image classification often involves a classification scheme and a classification metric. The classification scheme fits one of the classifications of Wolfe [11] or BI-RADS [17], [46]. Metric classification uses statistical classification or a classifier such as KNN (K - Nearest Neighbors).

Table 2 shows different proposals for breast density classification in mammographic images, and Table 3 shows different proposals for breast density classification in ultrasound. As shown by these tables, only a few algorithms have been proposed to evaluate breast density in ultrasound images.

Table 2- Different approaches for breast density classification in mammograms.

Author/Year	Segmentation Feature and/or Feature Extraction	Classification Metric	Classification Scheme
Taylor P et al 1994 [63]	Threshold	Statistical and texture measures	Wolfe categories
Byng J et al. 1994 [64]	Threshold	Threshold	Six categories
Suckling J et al 1995 [65]	Feature vector	Neural networks	Comparison between algorithm and radiologist
Byng J et al 1996 [66]	Fractal analysis	Threshold	Six categories
Byng J et al 1997 [67]	Threshold	Histogram and Fractal geometry Proportional hazards regression model	Six density categories
Karssemeijer N et al 1998 [68]	Threshold	K-Nearest Neighbors classifier	Four categories
Byng J et al 1998 [69]	Threshold	Percent density	Six categories
Zhou C et al 2001[70]	Histogram	. Rule-based classification	Four categories
Sivaramakrishna R. et al 2001 [71]	Threshold	Percent density	Comparison between algorithm and radiologist
Saha P et al 2001 [72]	Fuzzy methods	Sum of intensities of pixels	Comparison between algorithm and radiologist
Bovis K et al 2002 [73]	Threshold	Fourier transform; Laws' texture masks; Discrete Wavelet Transform	BI-RADS
Muhimmah I et al 2005 [74]	Histograms	Feature vectors and k-nearest-neighbor approach: an Euclidean distance, Bayesian Probability, major voting	Six categories and radiologist.
Torres-Mejía G et al 2005 [75]	Histograms	Percent density	Percentage of dense pixels compared with Wolfe categories
Oliver A et al 2005 [76]	Fuzzy methods	k-Nearest Neighbours algorithm and a Decision Tree classifier	Three categories
Martin k et al 2006 [77]	Threshold	Percent density	Histogram Classification
Oliver A et al 2006 [78]	Histograms	. Bayesian classifier :with k-Nearest Neighbours algorithm and the C4.5 decision tree	BI-RADS
Muhimmah I et al 2006 [79]	Histogram e	Feature vectors in combination with a multiclass Directed Acyclic Graph – Support Vector Machine	Three categories
Lu L et al 2007 [80]	Histogram	Percent density	BI-RADS
Heine J et al 2008 [81]	Threshold	Percent density	BI-RADS
Oliver A et al 2008 [82]	Fuzzy methods	Bayesian combination of a number of classifiers	BI-RADS
Oliver A et al 2010 [22]	Fuzzy methods	Karhunen–Loeve transform	Two categories
Subashini T et al 2010 [83]	Threshold	Vector machine	Three density categories
Liu L et al 2010 [84]	Histogram	Feature vectors	Three categories
Mustra M et al 2010 [85]	Covariance matrix	IB1 Classifier	Three categories
Bueno G et al 2011 [86]	Covariance matrix	k-NN, SVM and LBN	BI-RADS

Table 3- Different approaches for breast density classification in ultrasound.

Author/Year	Segmentation Feature and/or Feature Extraction	Classification Metric	Classification Scheme
Chang R et al 2006 [87]	Adaptive Speckle Noise	Threshold	BIRADS
Chen J et al 2009 [88]	Adaptive Speckle Noise	Threshold	BIRADS
Chang R. et al 2010 [89]	Volumetric breast density	Fuzzy methods, Percent density	No Scheme

The methods proposed in [87] and [88] and those mentioned in Table 3 follow the same approach: data acquisition, preprocessing data for speckle reduction and density classification.

In both cases, image acquisition was performed using an SSD-5500 ultrasound machine with a linear 6 cm ASU-1004 transducer. In this ultrasound system, the probe was immersed in a water bath coupling, and three passes were performed to cover the entire breast. The acquired images were stored in a DICOM (Digital Imaging and Communications in Medicine) file, and a DICOM reader decomposed them into serial 2D images. The second step consists of preprocessing, which includes speckle noise reduction and an adaptive threshold, which detects the region of interest through an algorithm that distinguishes the pixels of the different regions, and roughly divides the regions into fibroglandular tissue and fat tissue. Finally, in the third step, two methods, threshold- and proportion-based, were applied to provide a measure of breast density and the corresponding classification according to BI-RADS.

Thus, by analyzing the procedure used in the methods described in [87] and [88], the specific form of image acquisition through breast submersion, the pre-processing of the obtained images in DICOM format and the respective qualitative classification are similar.

Chang et al. [89] used a three-dimensional ultrasound technique called automated whole breast ultrasound (ABUS) that is used to automatically scan a large area of breast with two to five passes such that the whole breast is scanned completely. After segmenting the breast region, the fuzzy c-mean classifier was used to differentiate the fibroglandular and fatty tissues in the ABUS images. The percent density and fibroglandular tissue volume were compared and correlated in both ABUS and MRI imaging modalities with the linear regression analysis.

5. Methods for feature extraction in ultrasound images

6.Based on the analyses discussed above, three methods have been used for feature extraction in mammography: basic histogram thresholding, fuzzy c-means and gray-level co-occurrence matrix. We performed a preliminary investigation on the applicability of these features for classifying breast density in ultrasound images.

.

5.1 Thresholding in ultrasound images

An ultrasound image is selected and converted into grayscale. To apply the thresholding to the image, the histogram is generated from a selected area that represents the glandular area of the breast ultrasound image, as illustrated in Figure 7.

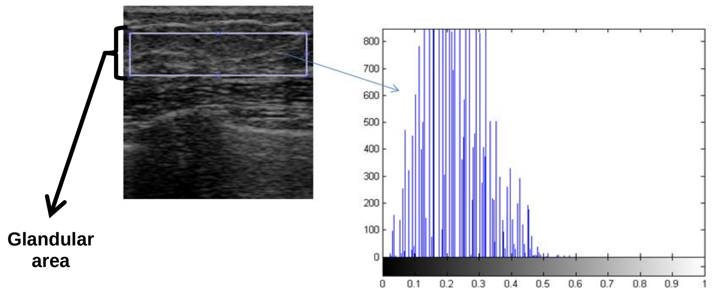


Figure 7- Selected area in the breast glandular area and the respective image histogram.

As shown in Figure 7, the histogram from the selected breast area has a maximum value close to 0.6. All values are concentrated in the dark area of the histogram. A thresholding that divides the range of grayscale [1:256] into two ranges [1:128] and [129:256] is considered. The values of each range are counted; dark pixels are identified in [1:128] and white pixels are identified in [129:256]. The formula to evaluate breast density is the sum of white pixels divided by the sum of white pixels and dark pixels. The obtained value is close to zero. Nevertheless, there is space for further research on this issue.

5.2 Fuzzy C-Means in ultrasound images

Fuzzy c-means segmentation of an image was used to convert an input image into two segments to represent the dark area in one cluster and the white area in another. For a selected area that represents the glandular area of the breast ultrasound image, the fuzzy c-means algorithm is applied.

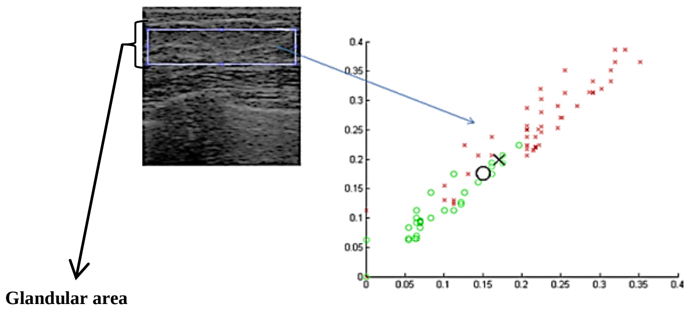


Figure 8 - Fuzzy C- Means applied to the selected breast area.

As observed in Figure 8, the two clusters cannot be identified clearly, and the center of the clusters represented by “O” and “X” are close, which indicates that it is not possible to extract the two features for further classification.

5.3 Gray-level co-occurrence matrix in ultrasound images

To calculate the gray-level co-occurrence matrix for a grayscale image, the MATLAB® *graycomatrix* [90] package is used to evaluate the following values:

- Contrast – Provides the measure of the intensity between each pixel and its neighbor. If the value of contrast is zero, it means there is no variance in grayscale intensity.
- Homogeneity - Returns a value that measures the closeness of the distribution of elements, each element in relation to an element in the diagonal. Large values of homogeneity mean that the image contains similar levels of gray.
- Correlation **-** Returns a measure of how each pixel is correlated with the neighboring pixels.

In this case, four gray-level co-occurrence matrixes are calculated, i.e., for angles 0º, 45º, 90º and 135º, and the final matrix is the mean of the four matrixes. Two gray levels are considered.

The calculations of the contrast, homogeneity and correlation are performed for a set of 85 breast ultrasound images, and the obtained results are similar, which suggests that the images have a lower contrast range between 0.000 and 0.035. There is a weak variance in the grayscale intensity. The correlation is close to one and ranges from 0.875 to 1.000, which means that the pixels are strongly correlated, and a high value for the homogeneity, which ranges from 0.300 to 0.749, indicates that the image contains similar levels of gray, which makes these features unsuitable for further classification.

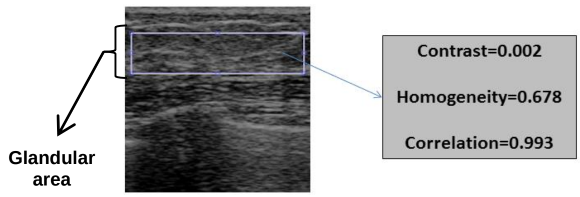


Figure 9- Results from gray-level co-occurrence matrix in the selected area**.**

Figure 9 provides an example illustrating this situation. Based on this example, it is difficult to analyze the ultrasound images with these methods.

Conclusions

In this paper, the relevance of breast density as a risk factor and the gravity associated with breast cancer are discussed. Several studies have analyzed breast density with other risk factors and concluded that breast density may be considered a risk that is as relevant as the other known risks. For this reason, the assessment of breast density value is important. Quantitative and qualitative approaches for evaluating breast density are also discussed.

For mammography, several algorithms are used to obtain, in most cases, a qualitative assessment of breast density. However, few algorithms exist for ultrasound images. Three of the methods that have previously been applied to mammography were applied to ultrasound images. Although the obtained results are not satisfactory, they may be a starting point for further research on the assessment of breast density in ultrasound images.

Authors' contributions

All authors conceived and designed the project. JFM acquired the set of breast ultrasound images and provided the corresponding clinical interpretation. AO, MP and MF were responsible for developing the algorithms. AO wrote the code and ran the experiments under the supervision of JFM, MP and MF. AO wrote the first draft of the manuscript and JFM, MP and MF critically revised it for important intellectual content. All authors approved the final version of the manuscript.

Acknowledgements

We are grateful to Dr. Carlos Gomes from Cova da Beira's Hospital in Covilhã, Portugal, for providing the second manual evaluation of the breast density for the set of 82 images considered in this study.

Statement on conflicts of interest

The authors state that they have no competing interests to declare.

References

[1] Jemal, A; Siegel R; Jiaquan X et al.; Cancer Statistics 2010, CA Cancer J Clin 2010;60:277-300.

[2] National Cancer Institute, What You Need to Know About Breast Cancer, NIH Publication 2009; 09:1556.

[3] Ferlay, J; Estimates of cancer incidence and mortality in Europe in 2008, Eur J Cancer 2010; 46(4):765-781.

[4].Portuguese Institute of Oncology Francisco Gentil, National Cancer Registry 2001, Mediana SA, 2008.

[5] Portuguese Institute of Oncology of Porto, Cancer Registry 1989 - 2009

[6] Eurostat European Commission, Population and social conditions – First Demographic estimates for 2008, Eurostat European Commission, 2008.

[7] American Cancer Society, Cancer Facts & Figures 2011, Atlanta: American Cancer Society, 2011.

[8].European Cancer Organization (ECO), available: http://www.ecco-org.eu.

[9] National Breast and Ovarian Cancer Centre. Breast cancer risk factors: a review of the evidence. National Breast and Ovarian Cancer Centre, Surry Hills, NSW, 2009.

[10] Black, W; Welch, H; Screening for disease, Am J Roentgenol 1997; 168:3–11.

[11] Wolfe, J; Risk for breast cancer development determined by mammographic parenchymal pattern, PubMed 1976; 37(5):2486-92.

[12] Boyd, N; Lockwood, G et al.; Mammographic Densities and Breast Cancer Risk, Cancer Epidem Biomar 1998; 7:1133-1144.

[13] Obenauer, S; Hermann, K; Grabbe, E; Applications and literature review of the BI-RADS classification, EUR RADIOL 2005; 15:1027-1036.

[14] McCormack, V, Silva, I; Breast Density and Parenchymal Patterns as Markers of Breast Cancer Risk: A Meta-analysis, Cancer Epidem Biomar 2006; 15:1159-1169.

[15] Oliver, A; Freixenet, J; Martí, R; Zwiggelaar, R.; A Comparison of Breast Tissue Classification Techniques, Med Image Comput Assist Interv 2006; 9(2):872-9.

[16] Guo, Y; Sivaramakrishna, R; et al.; Breast image registration techniques: a survey, Med Biol Eng Comput 2006; 44:15-26.

[17] Muhimmah, I; Oliver, A; et al.; Comparison Between Wolfe, Boyd, BI-RADS and Tabár Based Mammographic Risk Assessment, Lect Notes Comput SC 2006; 4046: 407-415.

[18] Highnam, R; Jeffreys, M; Comparing measurements of breast density, Phys Med Biol 2007; 52:5881–5895.

[19] Yaffe, M; Mammographic density. Measurement of mammographic density, Breast Cancer Res 2008; 10:209.

[20] Cheng, H; Shan, J et al.; Automated breast cancer detection and classification using ultrasound images: A survey, Pattern Recognition 2010; 43: 299–317.

[21] Kallenberg, M; Lokate, M; Automatic breast density segmentation: an integration of different approaches, Phys Med Biol 2011; 56(9): 2715.

[22] Oliver, A; Freixenet, J et al.; A review of automatic mass detection and segmentation in mammographic images, Med Image Anal 2010; 14(2):87–110.

[23] Boyd, N; Dite, G et al.; Heritability of Mammographic Density, a Risk Factor for Breast Cancer, New Engl J Med 2002; 347(12):886-894.

[24] Ursin, G; Hovanessian-Larsen, L et al.; Greatly increased occurrence of breast cancers in areas of mammographically dense tissue, Breast Cancer Res 2005; 7(5):605-8.

[25] Gail, M; Brinton, L et al.; Projecting individualized probabilities of developing breast cancer for white females who are being examined annually, J Natl Cancer Inst 1989; 81:1879–86.

[26] Chen, J; Pee, D et al.; Projecting Absolute Invasive Breast Cancer Risk in White Woman With a Model That Includes Mammographic Density, J Natl Cancer I 2006; 98(17):1215-26.

[27] Titus-Ernstoff, L; Tosteson, A et al; Breast cancer risk factors in relation to breast density (United States), Cancer Causes Control 2006; 17(10):1281-1290.

[28] Boyd, N; Guo, H et al.; Mammographic Density and the Risk and Detection of Breast Cancer, New Engl J Med 2007; 356:227-236.

[29] Vachon, C; Gils, C et al.; Mammographic density, breast cancer risk and risk prediction, Breast Cancer Res 2007; 9:217.

[30] Tice, J; Cummings, S et al.; Using Clinical Factors and Mammographic Breast Density to Estimate Breast Cancer Risk: Development and Validation of a New Predictive Model, Ann Intern Med 2008; 148:337-347.

[31] Boyd, N; Martin, L et al.; Mammographic density and breast cancer risk: current understanding and future prospects, Breast Cancer Res 2011; 13:223.

[32] Dehkordy, S; Carlos, R; Dense Breast Legislation in the United States: State of the States, Journal of the American College of Radiology 2013; 10(12):899–902.

[33] Smith, ML; The Density Conundrum: Does Legislation Help or Hurt?; Journal of the American College of Radiology 2013;10(12):909–912.

[34] Price, E; Hargreaves, J et al.; The California Breast Density Information Group: A Collaborative Response to the Issues of Breast Density, Breast Cancer Risk, and Breast Density Notification Legislation, Radiology 2013; 69(3):887-892.

[35] Maskarinec, G; Pagano, I; A Longitudinal Investigation of Mammographic Density: The Multiethnic Cohort, Cancer Epidem Biomar 2006; 15:732-739.

[36] Maskarinec, G; Pagano, I et al.; Mammographic Density and Breast Cancer Risk The Multiethnic Cohort Study, Am J Epidemiol 2005; 162(8):743-52.

[37] Vachon, C; Pankratz, V et al.; Longitudinal Trends in Mammographic Percent Density and Breast Cancer Risk, Cancer Epidem Biomar 2007; 16(5):921-928.

[38] Barlow, WE; White, E et al.; Prospective Breast Cancer Risk Prediction Model for Women Undergoing Screening Mammography, J Natl Cancer I 2006; 98(17):1204-1214.

[39] Palomares, M; Machia, J et al., Mammographic Density Correlation with Gail Model Breast Cancer Risk Estimates and Component Risk Factors, Cancer Epidem Biomar 2006;15:1324-1330.

[40] Mitchell, G; Antoniou, A et al.; Mammographic Density and Breast Cancer Risk in BRCA1 and BRCA2 Mutation Carriers, Cancer Res 2006; 66:1866-1872.

[41] Boyd, N; Martin, J et al.; Body Size, Mammographic Density, and Breast Cancer Risk, 2006, Cancer Epidem Biomar 2006; 15:2086-2092.

[42] Boyd, N; Lockwood, G et al.; Mammographic Densities and Risk of Breast Cancer Among Subjects With a Family History of This Disease, J Natl Cancer Inst 1999; 91(16):1404–1408.

[43] Kerlikowske, K; Ichikawa, L et al.; Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk, J Natl Cancer Inst 2007; 99(5):386-395.

[44] Yaghjyan, L; Colditz, G; Mammographic Breast Density and Subsequent Risk of Breast Cancer in Postmenopausal Women According to Tumor Characteristics, J. Natl Cancer Inst. 2011; 103 (15):1179-1189.

[45] Bontrager, K; Radiographic Positioning and Related Anatomy Workbook and Laboratory Manual 1993, Elsevier Mosby.

[46] D'Orsi, C; Bassett, L et al.; Illustrated Breast Imaging Reporting and Data System: illustrated BI-RADS. 3rd ed. Reston, American College of Radiology 1998.

[47] American College of Radiology (ACR), available: http://www.acr.org.

[48] Zhou, X; Gordon, R; Detection of early breast cancer: an overview and future prospects, Crit Rev Biomed Eng 1989; 17(3):203–255.

[49] Van Bemmel, J; Musen, M; Handbook of Medical Informatics, New York: Springer (Springer-Verlag edition) 1997.

[50] Mobahi, H; Rao, R et al.; Segmentation of Natural Images by Texture and Boundary Compression, Computer Vision and Pattern Recognition 2010.

[51] Comanicu, D; Meer, P; Mean shift: a robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002; 24:603–619.

[52] Felzenszwalb, P; Huttenlocher, D; Efficient graph-based image segmentation, Int J Comput Vision 2004; 59(2):167–181.

[53] Cour, T; Benezit, F; Shi, J; Spectral segmentation with multiscale graph decomposition, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2005.

[54] Wang, J; Jia, Y et al.; Normalized tree partitioning for image segmentation, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2008.

[55] Zhu, Q; Song, G; Shi, J; Untangling cycles for contour grouping. In Proceedings of the International Conference on Computer Vision 2007.

[56] Kim, J; Fisher, A et al.; A nonparametric statistical method for image segmentation using information theory and curve evolution, IEEE Transactions on Pattern Analysis and Machine Intelligence 2007; 14(10):1486–1502.

[57] Ren, X; Fowlkes, C; Malik, J; Learning probabilistic models for contour completion in natural images. International, Journal of Computer Vision 2008; 77:47–63.

[58] Efford, N; Digital Image Processing: A Practical Introduction Using JavaTM. Pearson Education 2000.

[59] Gonzalez, R; Woods, R; Digital Image Processing, Addison-Wesley Publishing Company 1992;7.

[60] Raof, R; Salleh, Z et al.; Color thresholding method for image segmentation algorithm of Ziehl-Neelsen sputum slide images, Electrical Engineering, Computing Science and Automatic Control, in Proceedings 2008 5th International Conference 212,217:12-14

[61] Blaschke, T; Burnett, C; Pekkarinen, A; Image Segmentation Methods for object-based Analysis and Classification, Sensing Image Analysis: Including the Spatial Domain, Kluwer Academic Publisher 2004; 211–236.

[62] Kamavisdar, P; Saluja, S; Agrawal, S; A Survey on Image Classification Approaches and Techniques, IJARCCE 2013; 2(1).

[63] Taylor, P; Hajnal, S et al.; Measuring image texture to separate "difficult" from "easy" mammograms, Brit J Radiol 1994; 67(797):456-63.

[64] Byng, J; Boyd, N et al.; The quantitative analysis of mammographic densities, Phys Med Biol 1994; 39:1629–38.

[65] Suckling, J; Dance, D et al.; Segmentation of mammograms using multiple linked self-organizing neural networks, Med Phys 1995; 22:145.

[66] Byng, J; Boyd, N et al.; Automated analysis of mammographic densities, Phys Med Biol 1996; 41(5):909-23.

[67] Byng, J; Martin, J et al.; Automated Analysis of Mammographic Densities and Breast Carcinoma Risk, CANCER 1997; 80(1).

[68] Karssemeijer, N; Automated classification of parenchymal patterns in mammograms, Phys Med Biol 1998; 43:365–378.

[69] Byng, J; Yaffe, M; Analysis of Mammographic Density and Breast Cancer Risk from Digitized Mammograms, Radiographics 1998; 18(6):1587-1598.

[70] Zhou, C; Chan, H et al; Computerized image analysis: Estimation of breast density on mammograms, Medical Physics 2001; 28(6):1056–1069.

[71] Sivaramakrishna, R; Obuchowski, N et al.; Automatic segmentation of mammographic density, Acad Radiol. 2001; 8(3):250-6.

[72] Saha, P; Udupa, J et al.; Breast Tissue Density Quantification Via Digitized Mammograms, IEEE Trans. Med. Imaging 2001; 20(8):792-803.

[73] Bovis, K; Singh, S; Classification of mammographic breast density using a combined classifier paradigm, in Proceedings of the 4th Int Workshop on Digital Mammography 2002; 177–180.

[74] Muhimmah, I; Denton, E; Zwiggelaar, R; Histograms in Breast Density Classification: Are 4 Better than 1, in Seminar Nasional Aplikasi Teknologi Informasi 2005.

[75] Torres-Mejía, G; Stavola, B et al.; Mammographic Features and Subsequent Risk of Breast Cancer: A Comparison of Qualitative and Quantitative Evaluations in the Guernsey Prospective Studies, Cancer Epidem Biomar 2005; 14:1052-1059.

[76] Oliver, A; Freixenet, J; Zwiggelaar, R; Automatic Classification of Breast Density, in Image Processing 2005. ICIP 2005. IEEE International Conference.

[77] Martin, K; Helvie, M et al.; Mammographic Density Measured with Quantitative Computer-aided Method: Quantitative Computer-aided Method: Comparison with Radiologists 'Estimates and BI-RADS Categories, Radiology 2006; 240(3):656-665.

[78] Oliver, A; Freixenet, J et al., A Comparison of Breast Tissue Classification Techniques, LNCS 2006; 4191:872–879.

[79] Muhimmah, I; Zwiggelaar, R; Mammographic Density Classification using Multiresolution Histogram Information, in Proceedings of the International Special Topic Conference on Information Technology in Biomedicine (ITAB '06), Ioannina, Greece, October 2006.

[80] Lu, L; Nishino, T et al.; Computing mammographic density from a multiple regression model constructed with image-acquisition parameters from a full-field digital mammographic unit, Phys. Med. Biol. 2007; 52: 4905.

[81] Heine, J; Carston, M; Scott, C et al.; An Automated Approach for Estimation of Breast Density, Cancer Epidemiol Biomarkers Prev 2008;17:3090-3097.

[82] Oliver, A; Freixenet, J et al.; A Novel Breast Tissue Density Classification Methodology, IEEE Trans Inf Technol Biomed 2008; 12(1): 55-65.

[83] Subashini, T; Ramalingam, V; Palanivel, S; Automated assessment of breast tissue density in digital mammograms, Comp Vis Image Und 2010; 114:33-44.

[84] Liu, L; Wang, J; He, K; Breast Density Classification Using Histogram Moments of Multiple Resolution Mammograms, in Proceedings 3rd International Conference on Biomedical Engineering and Informatics 2010.

[85] Mustra, M; Grgic M, Delac K, Feature Selection for Automatic Breast Density Classification in Proceedings 52nd International Symposium ELMAR-2010, Croatia.

[86] Bueno, G; Vallez, N et al.; Automatic breast parenchymal density classification integrated into a CADe system, IJCARS 2011; 6(3): 309-318.

[87] Chang, R; Chang-Chien, K et al. Breast density analysis in 3-D whole breast ultrasound images. Conf Proc IEEE Eng Med Biol Soc. 2006; 1:2795-8.

[88] Chen, J; Huang, C et al. Breast density analysis for whole breast ultrasound images. Med Phys. 2009; 36(11):4933-43.

[89] Chang, R; Shen, Y et al., Volumetric Density Analysis Using Automated Whole Breast Ultrasound, Med Phys 2010; 37:3418.

[90] MATLAB documentation, available: http://www.mathworks.com/help/images/ref/graycomatrix.html 1994-2014 The MathWorks, Inc.