Aberystwyth University

A review on automatic mammographic density and parenchymal segmentation
He, Wenda; Juette, Arne; Denton, Erika R. E.; Oliver, Arnau; Marti, Robert; Zwiggelaar, Reyer

Review Article

A Review on Automatic Mammographic Density and Parenchymal Segmentation

Wenda He,[1] Arne Juette,[2] Erika R. E. Denton,[2] Arnau Oliver,[3] Robert Martí,[3] and Reyer Zwiggelaar[1]

[1]*Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK*
[2]*Department of Radiology, Norfolk & Norwich University Hospital, Norwich NR4 7UY, UK*
[3]*Department of Architecture and Computer Technology, University of Girona, 17071 Girona, Spain*

Correspondence should be addressed to Wenda He; hewenda@gmail.com and Reyer Zwiggelaar; rrz@aber.ac.uk

Breast cancer is the most frequently diagnosed cancer in women. However, the exact cause(s) of breast cancer still remains unknown. Early detection, precise identification of women at risk, and application of appropriate disease prevention measures are by far the most effective way to tackle breast cancer. There are more than 70 common genetic susceptibility factors included in the current non-image-based risk prediction models (e.g., the Gail and the Tyrer-Cuzick models). Image-based risk factors, such as mammographic densities and parenchymal patterns, have been established as biomarkers but have not been fully incorporated in the risk prediction models used for risk stratification in screening and/or measuring responsiveness to preventive approaches. Within computer aided mammography, automatic mammographic tissue segmentation methods have been developed for estimation of breast tissue composition to facilitate mammographic risk assessment. This paper presents a comprehensive review of automatic mammographic tissue segmentation methodologies developed over the past two decades and the evidence for risk assessment/density classification using segmentation. The aim of this review is to analyse how engineering advances have progressed and the impact automatic mammographic tissue segmentation has in a clinical environment, as well as to understand the current research gaps with respect to the incorporation of image-based risk factors in non-image-based risk prediction models.

1. Introduction

Breast cancer is the most common cause of death in female cancer sufferers in both developed and developing countries [1]. The exact cause(s) of the majority of breast cancers still remains unknown [2]. By far the most effective way to tackle the disease is through early detection, precise identification of women at risk, and applying preventative measures. Within screening mammography, both mammographic density and parenchymal pattern have been established as image-based risk factors [2]. Current studies suggest that mammographic density is a major risk factor, even though it adds only little to the Gail model [3], and it is still not clear how to incorporate the factors as biomarkers into the risk prediction models [4].

Substantial advances have been made with computer aided mammography in breast cancer research and treatment. Mammographic risk can be assessed in a clinical environment based on subjective appraisal of mammograms using protocols such as BI-RADS (American College of Radiology's Breast Imaging Reporting and Data System) [5], which can lead to inter- and intraobserver variability [6]. Within computer aided mammography, the idea of developing a fully automatic and repeatable breast tissue segmentation using computer vision and machine learning techniques is to facilitate cancer risk classification. The major challenge is to segment a given mammogram according to density and/or parenchymal patterns so that the distribution and characteristic mixture of the breast tissue can be determined, leading to an accurate and objective breast cancer risk estimation. Many studies have been conducted in an attempt to achieve this; however there are still significant gaps in translating the newly acquired knowledge into clinical improvements [2]. In the recent publication, Harvey et al. [7] evaluated mammograms for approximately 3400 women.

Breast density was calculated using automated software and used as an additional image-based risk factor for cancer risk prediction. Results indicated that reading of individual risk can be more accurate by incorporating image-based risk factors into established non-image-based risk model (e.g., the Gail [8] and the Tyrer-Cuzick [9] models). The initial results, based on a limited dataset, are promising; however, more investigations and evidence are required in order to fully establish and incorporate density into risk models in a clinical environment and to establish their place in changing clinical practice. A recent case-control study [10] compared six established digital mammographic density assessment methods and their ability to predict breast cancer risk. The selected approaches included three area based approaches and three fully automated volumetric methods. Area based approaches are referred to as mammographic image analysis over 2D breast tissue projection, whilst volumetric methods focus on assessment of the true physical volume of breast composition. With respect to patient care and economic sustainability, it is critical to have an effective and cost-effective analysis to evaluate the potential for personalised screening and prevention programmes and reduce anxiety and stress to the patient resulting from overscreening/treatment [11].

This paper presents a comprehensive review on automatic mammographic breast density and parenchymal segmentation methodologies developed over the past two decades, from its infancy in the early 90s to date. Note that mammographic breast density and parenchymal segmentation is referred to as mammographic tissue segmentation in the rest of the paper. The aim of this review is to analyse how engineering advances have progressed over these two decades and the impact automatic mammographic tissue segmentation has had in the clinical environment. Automatic mammographic tissue segmentation and measuring breast tissue composition for risk stratification (to adapt screening interval to risk) may be of prognostic value in distinguishing women with certain mammographic appearance likely to develop breast cancer, leading to a successful prevention and/or treatment. It should be noted that the focus of this review is to discuss how automated (as opposite to manual and semiautomated) segmentation of breast density and parenchyma is able to assist in the prediction of risk of developing cancer. These techniques do not aim to assess the risk of increased density masking breast cancer [12] when present although the two issues are clearly related. Note that the review does not cover aspects of abnormality detection or abnormality segmentation.

The remainder of the paper is organised as follows: Section 2 describes the established breast parenchymal patterns as seen in breast images and their interconnected relationships between different schemes, Section 3 briefly discusses mammograms and variations, Section 4 critically reviews the existing automatic mammographic tissue segmentation strategies in the literature, Section 5 provides discussion with respect to the development of automatic mammographic tissue segmentation, its progress, current state, and future beyond translational research in clinical practices, and Section 6 concludes the review.

2. Breast Parenchymal Patterns and Density Categories

Within screening mammography, a number of (mammographic risk assessment) schemes have been developed to estimate the likelihood of women developing breast cancer: Wolfe (1976) [13, 14], Tabár (1982) [15, 16], Boyd (1995) [17, 18], and BI-RADS (1993) [5, 19].

2.1. Wolfe. Wolfe [13] empirically categorised mammograms into four parenchymal patterns:

(1) N1 (primarily fatty): mammogram is composed of fat and a few fibrous tissue strands.

(2) P1 ($\leq$25% prominent ducts): mammogram shows a prominent duct pattern and a beaded appearance can be found either in the subareolar or the upper axillary quadrant.

(3) P2 (>25% prominent ducts): mammogram indicates severe involvement of a prominent duct pattern, which may occupy from one-half up to all of the volume of the parenchyma.

(4) DY (dense fibrovascular tissue): mammographic features show an increase in density of the parenchyma, which may be homogeneous with a minor component of prominent ducts.

Wolfe's classification was found to be associated with breast cancer risk, and data analysis has revealed a progressive increase of as much as 37 times higher future cancer risk from N1 (the lowest risk) to DY (the highest risk) [14, 20]. Wolfe's pioneering study generated considerable controversy due to the (first) randomised controlled mammographic screening trial [21] and limited mammographic capability (e.g., breast positioning and compression) of the early 70s [14]. It should be noted that Wolfe's method cannot be repeated for its subjectivity of assessment. Wolfe's 1976 conclusions [13] have been criticised as erroneous and unrealistic, but the realisation of the association between parenchymal patterns and mammographic risk is indisputable, which contributed tremendously in research of image-based mammographic risk classification and leads the way for future clinical advances. The reader is referred to [14] for the key aspects of this study which caused debate of the controversies [13].

2.2. Tabár. Strongly influenced by Wolfe's work [21], Tabár and Dean [15] proposed a model based on a mixture of four mammographic building blocks composing the normal breast anatomy: nodular (N) densities mainly correspond to terminal ductal lobular units; linear (L) structures correspond to either ducts or fibrous or blood vessels; homogeneous (H) structureless densities correspond to fibrous tissues; radiolucent (R) areas are related to adipose fatty tissues. Mammograms were subdivided into five risk categories based on the distributions of the four building blocks (e.g., [N%, L%, H%, R%]) [16]:

(1) T_I is composed of [25%, 15%, 35%, 25%], the lowest risk. Mammogram shows normal fibroglandular

tissue with partial fatty replacement, where pathological changes can be easily perceived despite the fact that the breast may be "dense" radiologically.

(2) T_{II} is composed of [2%, 14%, 2%, 82%]. Mammogram is characterised by the overrepresentation of radiolucent fatty tissue, which provides excellent background for radiologists to detect abnormalities.

(3) T_{III} is similar in composition to T_{II}, except that the retroareolar prominent ducts are often associated with periductal fibrosis. Neither of these patterns (i.e., T_{II} and T_{III}) has nodular densities or diffuse fibrosis, and the overrepresentation of radiolucent fatty tissue makes pathological lesions relatively easier to detect through mammography.

(4) T_{IV} is composed of [49%, 19%, 15%, 17%]. Mammogram is dominated by prominent nodular, linear densities and appears to be resistant to the process of involution, which makes perception of pathological lesions difficult on mammograms.

(5) T_V is composed of [2%, 2%, 89%, 7%], the highest risk. Mammogram is dominated by extensive homogeneous structureless fibrous tissue, which limits the capabilities of mammography to demonstrate the normal anatomy and reveal small pathological lesions.

Tabár's approach is based on mammographic anatomic (pathologic) correlations rather than pattern reading alone (e.g., Wolfe classification), and the primary difference between Wolfe and Tabár's classification is Tabár's T_I [22, 23]. Tabár's definitions of mammographic risk patterns are more discriminatory than Wolfe's, helping to increase reproductive accuracy of the classification. The reader is referred to [23, 24] for a detailed comparison with respect to these two risk schemes.

2.3. Boyd

In a subsequent study to [13], Wolfe et al. [85] described a stronger association between mammographic density and breast cancer risk. This led Boyd et al. [17] to develop a method to measure mammographic percentage density using a computer aided technique, which marked movement away from describing patterns to objective assessment using tissue percentages. In particular, the risk categories are defined using a thresholding method [86], also known as the Cumulus interactive threshold software (University of Toronto), which is considered to be the "gold standard" tool for density measurement [81, 87]. With the Cumulus, the reader can identify the boundaries of the breast tissue, define the threshold for dense tissue on the mammogram, and measure the total area of the breast. The percentage of mammographic density is examined to estimate the proportion of fibroglandular tissue (as opposed to fat) of the breast and is divided into six class categories (SCC):

(1) SCC_1, density = 0%,

(2) SCC_2, density < 10%,

(3) SCC_3, 10% ≤ density < 25%,

(4) SCC_4, 25% ≤ density < 50%,

(5) SCC_5, 50% ≤ density < 75%,

(6) SCC_6, density ≥ 75%.

Subsequent follow-up studies [17, 88] have established a significant independent association between overall breast density and future breast cancer risk, which indicated a much lower magnitude (4-5 times increase) than the 37 times increase initially proposed by Wolfe [13].

2.4. BI-RADS

A series of studies (e.g., [85, 88, 89]) indicated that the percentage of dense breast tissue is highly associated with breast cancer risk. BI-RADS [5] was designed as a quality assurance tool to standardise mammography reporting, in order to reduce confusion in breast imaging interpretations and to facilitate outcome monitoring. Four breast compositions were identified as follows:

(1) B_1, the breast being almost entirely fat (<25% glandular);

(2) B_2 (25%–50% scattered fibroglandular densities);

(3) B_3 (50%–75% heterogeneously dense breast tissue);

(4) B_4, the breast being extremely dense (>75% glandular).

BI-RADS, widely used throughout North America and much of Europe [90], covers the significant relationship between increased breast density and decreased mammographic sensitivity in detecting cancer [14, 91].

Area based assessments can be subjective and imprecise. Volumetric density is a relatively new concept. The developed volumetric measuring techniques work out what kind of tissues must have been present by assessing breast composition at each pixel by calculating the X-ray attenuation between that pixel and the X-ray source [83]. Volumetric breast density and visual assessment are strongly correlated [10]; however, visual density categories as described in BI-RADS based on 2D projection cannot be directly used for volumetric breast density assessments. For example, visual density areas as shown in side on views of breasts under compression can be the same/similar due to superimposition; the measurements can be very different if measured based on volumetric breast density. The fifth edition of BI-RADS Atlas [92] included changes in breast density reporting categories with emphasis on volumetric assessments. The updated BI-RADS helps clarify what breast density assessment should be and distinguishes density from other BI-RADS assessments. The latest edition indicates that, in a clinical environment, focus of density as a risk factor has now moved towards density as a masking factor.

2.5. Correlations between Different Schemes

Muhimmah et al. [24] investigated the (Spearman's) correlations (r_s) between the four different schemes (i.e., Wolfe, Boyd, BI-RADS, and Tabár) using the MIAS (Mammographic Image Analysis Society) database [93], which is a publicly available, digitised database. Results indicate risk classification mappings between the four schemes as follows.

(i) Wolfe $\rightarrow$ Boyd: N1 $\rightarrow$ SCC_1, P1-2 $\rightarrow$ SCC_{2-5}, and DY $\rightarrow$ SCC_6; $r_s = 0.93$; this is in line with reported results ($r_s = 0.81$) by Brisson et al. [91];

(ii) Wolfe $\rightarrow$ BI-RADS: N1 $\rightarrow$ B_1, P1 $\rightarrow$ B_2, P2 $\rightarrow$ B_3, and DY $\rightarrow$ B_4; $r_s = 0.93$.

(iii) Wolfe $\rightarrow$ Tabár: N1 $\rightarrow$ T_{II}, P1 $\rightarrow$ T_{III}, P2 $\rightarrow$ T_{IV}, and DY $\rightarrow$ T_V; $r_s = 0.93$ excluding T_I due to the weak correlation with any of the Wolfe patterns [24]; this is in line with results in [22, 23];

(iv) Boyd $\rightarrow$ BI-RADS: SCC_1 $\rightarrow$ B_1, SCC_{2-3} $\rightarrow$ B_2, SCC_{4-5} $\rightarrow$ B_3, and SCC_6 $\rightarrow$ B_4; $r_s = 0.91$;

(v) Boyd $\rightarrow$ Tabár: SCC_1 $\rightarrow$ T_{II}, SCC_{2-3} $\rightarrow$ T_{III}, SCC_{4-5} $\rightarrow$ T_{IV}, and SCC_6 $\rightarrow$ T_V; $r_s = 0.93$;

(vi) BI-RADS $\rightarrow$ Tabár: B_1 $\rightarrow$ T_{II}, B_2 $\rightarrow$ T_{III}, B_3 $\rightarrow$ T_{IV}, and B_4 $\rightarrow$ T_V; $r_s = 0.96$ excluding T_I [24]; a recent study [94] indicated a strong direct correlation between these two schemes with T_I taken into account, where B_1 $\rightarrow$ $T_{II/III}$, B_2 $\rightarrow$ T_I, B_3 $\rightarrow$ T_{IV}, and B_4 $\rightarrow$ T_V; $r_s = 0.92$; note that Tabár's T_{II} and T_{III} have the same tissue composition; see Section 2.2.

Strong correlations were established between Wolfe, Boyd, and BI-RADS categories [24] but the correlations with Tabár categories are less straight forward due to Tabár's T_I. According to the literature, these correlations have not been investigated using other mammographic modalities (e.g., digital mammography and tomosynthesis), with the exception of one study [94] which investigated the correlation between BI-RADS and Tabár categories using a private digital mammography database. Mammographic parenchymal pattern and percentage density (PD) have shown being strongly associated with breast cancer risk in the literature [95–97]. However, Brisson et al. [91] suggested that PD provides more information on breast cancer risk than Wolfe's parenchymal patterns. Once PD is taken into account (e.g., using BI-RADS scheme), parenchymal pattern information is redundant. Tabár's scheme seems to capture something more than just density assessment, which could be useful in temporal analysis of breast parenchymal changes. However, it is still unclear whether this additional information from Tabár's scheme is related to breast cancer risk [23]; therefore further investigations are warranted. Example mammographic images assessed based on Tabár and BI-RADS are shown in Figures 1 and 2, respectively.

3. Screen Mammography

Early detection through breast screening programmes is by far the most effective way to improve survival rate [98]. There are multiple breast imaging modalities; mammography, ultrasound, thermography, PET (positron emission tomography), MRI (magnetic resonance imaging), CT (computed tomography), scintimammography, optical imaging, and electrical impedance tomography (EIT) based imaging have all been used for different purposes. The clinical role for nonmammographic modalities is often to provide additional information when the results of mammograms are indeterminate or of limited utility [99]. This review focuses on evaluation of density using (X-ray related) mammographic modalities, and other modalities are outside the current scope. The reader is referred to [98] for a detailed overview and [100] for a comparison of breast imaging modalities, and the associated biomarkers used for early detection can be found in [101].

Mammography is the gold standard method in detection of early stage breast cancer before lesions become clinically palpable. The most criticised aspect of screening mammography is that, for women with dense breasts, the sensitivity of mammography is significantly reduced leading to concerns over unnecessary biopsies and treatment [102, 103]. This may lead to another harm to the patient (e.g., physical, emotional, financial, and/or psychological) [104]. Despite these limitations, mammography as an initial examination can have very good sensitivity and specificity when compared to other modalities (e.g., PET, CT, and MRI). However, the reading and assessment of mammogram densities can be highly subjective with only moderate agreement among radiologists [24, 105, 106]. Note that modalities such as ultrasound and MRI can have better sensitivity in some cases but with a corresponding loss of specificity. Clinically these modalities are most often used as an adjunct to mammography, whether 2D or 3D, and as follow-up examination to investigate specific mammographic findings. Mammography is a low cost simple procedure, which is affordable for all population screening [98]. Therefore, there is a huge incentive to develop computer aided mammographic analysis approaches, in order to deliver objective and accurate results in an effective and cost-effective way and to facilitate early detection and precise identification in a clinical environment.

3.1. Mammographic Image Quality Aspect.
Mammographic image quality can be considered an indication of clarity with which radiologically significant details can be perceived in an image. Superb image quality is imperative for reliable detection and accurate characterisation of breast parenchyma and is expected to be beneficial for breast imaging analysis. Broadly speaking, there are two types of mammography including 2D mammographic projection (e.g., screen-film mammography (SFM) and full field digital mammography (FFDM)) and 3D digital breast tomosynthesis (DBT).

3.1.1. SFM and FFDM.
Clinical evaluation has indicated that SFM and FFDM are similar in their ability to detect cancer [107]; however, FFDM is more effective at finding cancer in certain groups of the population, such as women who are premenopausal or perimenopausal, under the age of 50, and have dense breasts. This indicates that in this subgroup some anatomical regions are better visualised by FFDM than SFM. In particular, FFDM demonstrated improved image quality with significantly better depiction of the nipple, skin, pectoral muscle, and especially contrast in parenchymal and fatty tissue [108]. Note that digital mammography imaging generates two types of images for analysis, raw ("for processing") and vendor postprocessed ("for presentation"), of which postprocessed images are commonly used in clinical practice.

3.1.2. DBT.
Routine screening mammography relies upon a select number of views, for example, Craniocaudal (CC)

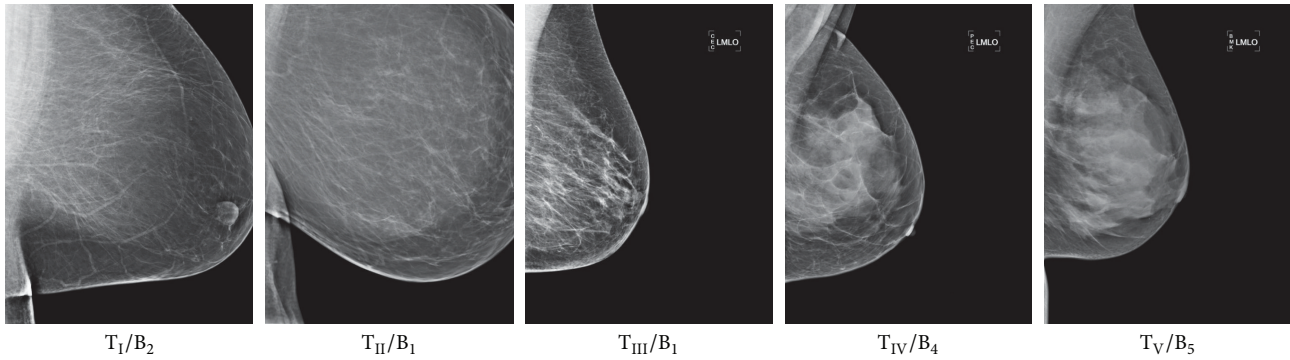T_I/B_2 T_{II}/B_1 T_{III}/B_1 T_{IV}/B_4 T_V/B_5

FIGURE 1: Example mammographic images with Tabár (T) risk classifications from low to high and their equivalent according to BI-RADS (B) scheme.

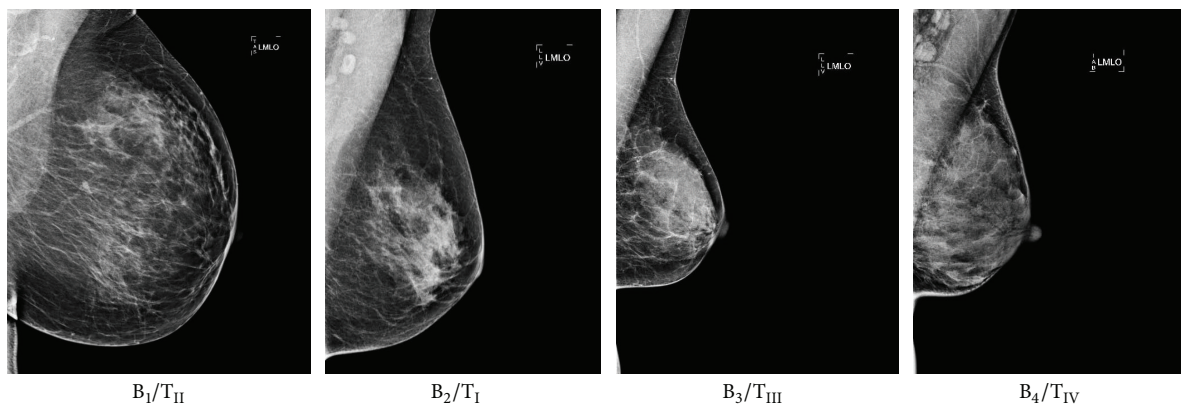B_1/T_{II} B_2/T_I B_3/T_{III} B_4/T_{IV}

FIGURE 2: Example mammographic images with BI-RADS (B) density categories from low to high and their equivalent according to Tabár (T) risk classification.

and Mediolateral Oblique (MLO) view, to assess breast tissue structures [109]. With 2D projections, it can be difficult to separate normal glandular tissue from tumours. One of the biggest challenges to screening radiologists is to interpret superimposed fibroglandular tissue (anatomical noise) in the image, in which pathological structures can be obscured and remain undetected. In some cases this can mimic lesions leading to false positive results, unnecessary recalls for additional screening, and/or biopsy. DBT is a recent advanced image acquisition technology, in which the conventional mammography technique has been modified to acquire a 3D view of the breast; a series of thin-resolution images are aggregated to generate a 3D image of the breast. The reconstructed format eliminates image superimposition and allows images of submillimetre cross sections to be analysed, increasing the conspicuity of features that are often obscured by overlapping structure in a single-projection view [110]. A newly developed 2D imaging "modality" called C-View (synthetic 2D view) can be generated from the 3D DBT data during the mammography exam [111], eliminating the need for additional 2D exposures. Initial clinical studies have shown that screening with C-View imaging may result in clinical performance superior to that of a conventional 2D mammogram [112]. However, further evaluation and validation are needed to verify the usage of C-View in a clinical environment. Note

that density assessment between area based 2D projection and 3D DBT can be very different. Clinical investigation is currently under way to establish the relationships between density assessment using these two very different modalities.

4. Automatic Breast Density and Parenchymal Segmentation

In principle, there are three approaches to mammographic tissue segmentation, manual, semiautomatic (interactive), and fully automatic. Although more effective than manual segmentation with respect to speed and efficiency, semiautomatic methods have several limitations; the interactive segmentation and labelling of mammograms can still be subjective and time consuming and require operator training. Such approaches make large studies and clinical usage costly. Defining parameters in a subjective manner could introduce observer bias. However, manual and semiautomatic approaches can sometimes be a good idea for obtaining "ground truth" to facilitate other studies. The focus of this review is on fully automatic approaches to mammographic tissue segmentation, with manual and semiautomatic approaches excluded.

A large quantity of literature was reviewed in order to illustrate the progress and advances in mammographic tissue segmentation; this section is separated into four categories
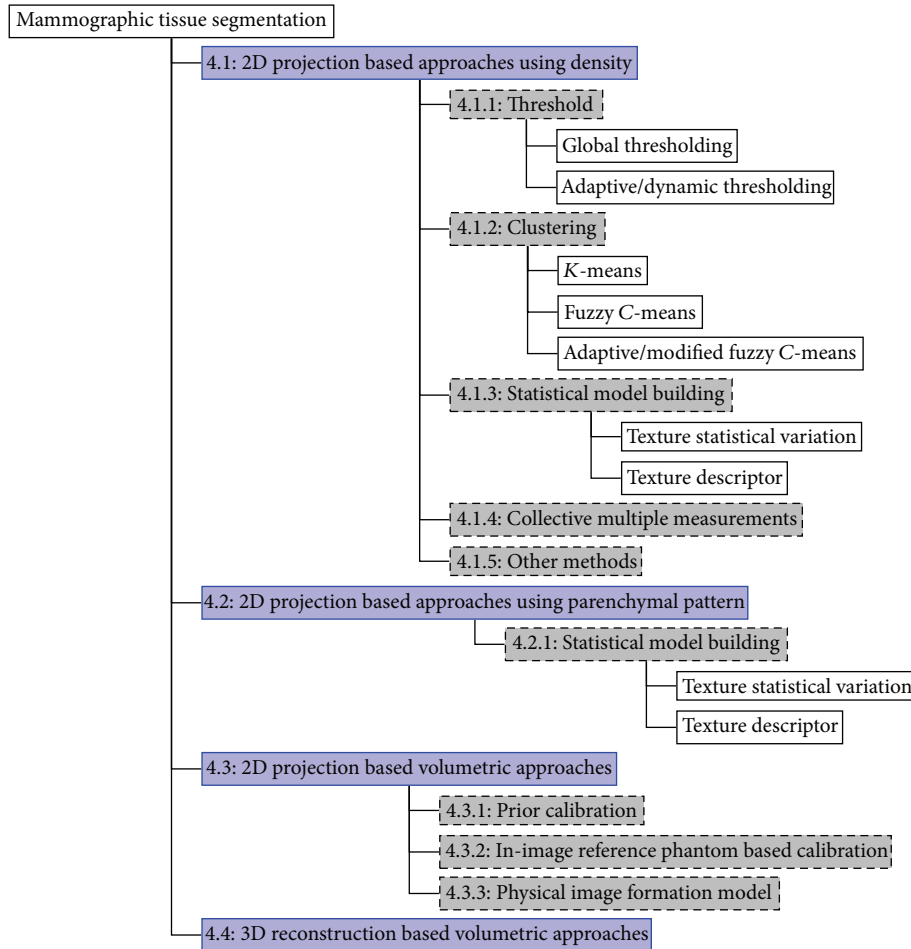
FIGURE 3: A tree representation for mammographic tissue segmentation categorisation based on combination (e.g., shaded in light purple) of technological advances and image modalities. Subcategorisation (e.g., shaded in light grey) is based on core segmentation principles.

according to a combination of two technological advances, imaging modalities (i.e., 2D projection and 3D reconstruction) and segmentation principles (i.e., based on 2D projection or volumetric data). It is ideal to use categorisation based on a combination of the technological advances over the past two decades, so that the separation within the literature is according to the timeline of the developments. The reader is referred to Figure 3 which shows a tree representation for mammographic tissue segmentation categorisation used in this review.

Mammographic tissue segmentation is often used as an intermediate stage prior to deriving certain features from the segmentation for risk classification. Some studies are closely related and in some cases the same methodologies evolve over time with improved results. In order to fully focus on breast density and parenchymal segmentation, other aspects such as comparisons of classification techniques used for the follow-up risk classification are briefly noted, as these are considered to be separate modules/steps in the process pipeline. However, risk classification results are sometimes used to indicate the correctness of the segmentation when the evaluation of the segmentation quality is not available. The

reviewed mammographic tissue segmentation is explicitly developed for breast tissue separation within the breast region and does not cover the detection/segmentation of abnormalities (e.g., mass and microcalcification), not does it cover the separation of the breast area in mammograms. Note that, in the literature, breast tissue density and dense tissue are often referred to as parenchymal patterns, fibroglandular disk, and parenchymal density.

4.1. 2D Projection Based Approaches Using Density. The majority of mammographic tissue segmentation uses digitised SFM or FFDM. Based on the core segmentation principles, they can be categorised into five groups: thresholding, clustering, statistical model building, collective multiple measurements, and other methods. Figure 4 shows the distribution of the studies conducted with respect to the five groups.

4.1.1. Thresholding. To overcome observer bias in choosing threshold values subjectively, various studies were conducted to explore techniques to determine a threshold value automatically in order to achieve consistent segmentation

Distribution of mammographic density segmentation
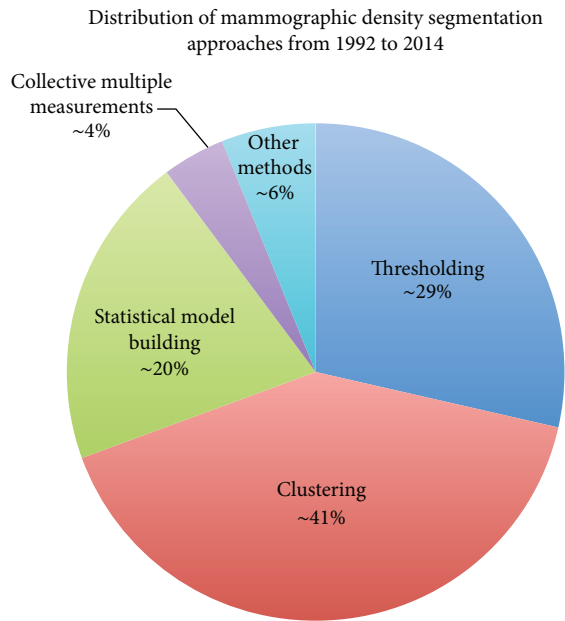approaches from 1992 to 2014

FIGURE 4: Distribution of mammographic density segmentation using 2D projection based approaches.

objectively in a high-throughput manner. High intensity dense and low intensity fatty breast tissue can be separated intuitively based on the mammographic pixels values. This led to the development of automatic threshold determination for mammographic tissue segmentation. Single threshold value approaches are difficult due to mammographic density inhomogeneity leading to inaccurate mammographic tissue segmentation and inaccurate breast density estimation. To improve this, adaptive/dynamic threshold based methods have emerged. A summary of representative studies using thresholding techniques can be found in Table 1. Note that automated thresholding is simple but it has not worked reliably to date across databases. Various studies indicated good results with the developed approaches versus risk; however, they may have been heavily trained on specific databases.

Global Thresholding. Early studies investigated various discriminant functions for breast density based segmentation, and the following studies are some examples. Matsubara et al. [25] applied a two-stage segmentation technique. Fatty and dense tissue were segmented using variance histogram analysis. The segmented dense tissue was then further grouped into either dense or semidense tissue by a binary technique and the threshold values were determined by discriminant analysis. Alternatively, Saha et al. [26] developed an automatic method to segment fatty and dense breast tissue using scale-based fuzzy connectivity methods [113]. Dense breast tissue was segmented as a set of fuzzy connected objects using automatically determined threshold values. The optimal threshold was derived by minimising a second-order statistic threshold energy function, computed by considering spatial arrangements of pixel intensities. Sivaramakrishna et al. [27] used a modified version of Kittler's optimal threshold procedure [114] for dense and fatty breast tissue segmentation,

whilst Olsén and Mukhdoomi [28] used minimum cross-entropy [115] to obtain an optimum threshold for segmenting glandular tissue automatically. Early investigations assumed that regions of glandular tissue have significant differences in brightness compared to the surrounding fatty tissue. Many histogram-based techniques focused on detecting intensity peaks which may indicate dense tissue or abnormalities, regardless of glandular tissue intensity and texture variation. Such an approach is limited because fibrous tissue is often embedded in fatty tissue, and the appearance of glandular tissue can be ranging from bright, fluffy blobs, to sparse lines. It is an inherent disadvantage for 2D mammographic projections to have superimposed tissue and overlapping structures which cause variation in glandular/fatty intensity.

Ferrari et al. [31] presented a technique to segment the fibroglandular disc in mammograms based on the statistical modelling of breast density using a weighted Gaussian mixture model. The parameters of the model and the number of breast tissue classes were determined using the expectation-maximization (EM) algorithm. Ferrari et al. [32] improved the technique by segmenting the fibroglandular disc in mammograms using four weighted Gaussians. The different density functions of the model were represented by the mixture of Gaussians corresponding to specific density classes in the breast. The parameters of the model and the number of breast tissue classes were also determined using the EM algorithm, where the number of Gaussian kernels of the model was determined by an information theoretic approach (i.e., minimum description length method [116]). The shortcoming in identifying the fibroglandular disc in the related studies [31, 32] was the parameter initialisation in the EM algorithm which can lead to convergence to local maxima, causing poor segmentation results. Similar to Ferrari et al. [32], El-Zaart [33] used a statistical approach for detecting the fibroglandular disc. The key difference between the two approaches is that Ferrari used Gaussian mixture modelling while El-Zaart used Gamma mixture modelling. The EM technique with a Gamma distribution was developed to estimate the statistical histogram parameters. Evaluation indicated that Gamma based method detected the fibroglandular disc regions more precisely, while the Gaussian based method falsely detected more regions that are not part of the glandular discs. The mixture of Gaussian distributions was used to model various breast parenchyma and is ideal for symmetric data; however, experimental results showed that the image histograms are not always symmetric. For this reason, Gamma distributions are more appropriate for modelling symmetric and nonsymmetric histograms of mammographic images, from which the thresholding values were selected at the valleys.

Lu et al. [30] adapted the semiautomatic breast density segmentation proposed by Boyd et al. [17], and modified the algorithm for automatic measuring of breast density from raw FFDM. A multiple regression model analysis was developed utilising image acquisition parameters and pixel intensity statistics to derive threshold values for mammographic tissue segmentation. Such an approach can be affected by various factors that influence the signal intensity recorded by the digital detector, including the estimation error of the compressed

TABLE 1: Summary of representative studies using thresholding based methods for mammographic tissue segmentation. r/R denotes correlation coefficient. Note that (1) largely identical studies are excluded in the list and (2) in case of multiple results, only the best reported results are listed.

Study	Year	Number of density categories	Modalities	Number of views	Number of images	Segmentation evaluation	Risk/density estimation accuracy
Global thresholding							
Matsubara et al. [25]	2000	Fatty, mammary gland diffuseness, nonuniform high density, and high density	Digitised SFM	MLO	148	Visually assessed	90% (ratios of the four densities)
Saha et al. [26]	2001	Fatty and dense	Digitised SFM	MLO and CC	174	Visually assessed (acceptable)	N/A
Sivaramakrishna et al. [27]	2001	Fatty and dense	Digitised SFM	CC	32	Visually assessed	Spearman's $r = [0.92, 0.95]$ (automatic-manual)
Olsén and Mukhdoomi [28]	2007	Fatty and glandular	Digitised SFM	MLO and CC	160 (MIAS + DDSM)	Visually assessed	N/A
Tzikopoulos et al. [29]	2011	Fatty and fibroglandular	Digitised SFM	MLO and CC	322 (mini-MIAS)	Visually assessed	86% (fatty, fatty-glandular, and dense-glandular) intraclass $r = 0.94$ (BI-RADS)
Lu et al. [30]	2007	Fatty and dense	FFDM (raw)	CC	172	Visually assessed	N/A
Ferrari et al. [31]	2000	Fatty and fibroglandular	Digitised SFM	MLO	66 (mini-MIAS)	Visually assessed (84% successful)	N/A
Ferrari et al. [32]	2004	Uncompressed-fatty, fatty, nonuniform-dense, and high-dense	Digitised SFM	MLO	84 (mini-MIAS)	Visually assessed (81% excellent or good)	N/A
El- Zaart [33]	2010	Fatty and fibroglandular	Digitised SFM	MLO	N/A	Visually assessed	N/A
Adaptive/dynamic thresholding							
Zhou et al. [34]	2001	Fatty and dense	Digitised SFM	MLO and CC	260	Visually assessed (94% correct); r (CC, MLO) = 0.94, 0.91 (automatic-manual)	42% (BI-RADS)
Neyhart et al. [35]	2002	Radiolucent and radiodense	Digitised SFM	MLO and CC	N/A	Visually assessed	N/A
Kim et al. [36]	2010	Fatty and dense	FFDM	MLO and CC	80	Visually assessed; $r = 0.99$ (automated-manual)	N/A
Nickson et al. [37]	2013	Dense and fatty	Digitised SFM	CC	5919 women	Visually assessed; 41% "perfect" agreement (Cumulus-AutoDensity)	Pairwise $r = 0.63$ (Cumulus-AutoDensity)

breast height, the heel effect, quantum mottle (noise), beam hardening, and detector nonuniformity. A FFDM unit may exhibit temporal changes due to ageing of the X-ray tubes and variation between FFDM unit manufacturers; moreover differences between radiology facilities can also be an issue. The approach can be problematic as the parameters and parameter estimates of regression models may change from installation to installation, site to site, and analyst to analyst.

Adaptive/Dynamic Thresholding. Zhou et al. [34] investigated an approach in which the image was first classified by hand according to the characteristic features of the grey-level histogram into one of the BI-RADS density categories using a rule based classification. With the known BI-RADS density category and the shape of the histogram (e.g., unimodal, bimodal, or multimodal), the density segmentation was achieved by using combined discriminant analysis [117] and maximum entropy principle [118] based threshold selection methods. Successful segmentation strongly depended on a correctly classified mammogram, which would result in selecting the optimal threshold. The developed method was used in a clinical study [119], and results suggested that misclassified histograms occurred more often on extremely dense and fatty mammograms, but the overall breast density estimation was more accurate than the radiologists' visual estimation. However, because the configured parameters for threshold determination were dataset specific and the segmentation was performed after density classification, it was concluded in [119] that such an automatic approach cannot yet be used as a stand-alone density measurement tool.

Kim et al. [36] developed a scheme for breast density estimation using statistical (e.g., standard deviation) and boundary information (e.g., edge magnitude) to compute an optimal intensity threshold between dense and fatty tissue in order to define dense and nondense areas. One key hypothesis in the study is that boundaries between the dense and fatty portions of the breast usually have high values of gradient magnitude. Therefore an iterative search was used and the optimal threshold was determined by combining these features to best divide the fat and dense regions. Nickson et al. [37] adopted this method for mammographic density segmentation and noticed that often a false optimal threshold is produced, which separates a narrow band of low intensities along the breast skin line from the rest of the breast. Nickson et al. modified the method, named AutoDensity, to converge to an intensity level that separates dense from fatty tissues. This is achieved by performing an iterative search for the optimal threshold while decreasing a region of interest within the breast until a stopping condition is met. The modified method was used on a large population screening programme, and the automated measurement of breast density from digitised SFM using AutoDensity performs similarly (modest correlation) to the Cumulus (a semiautomated user-assisted PD estimation method) [17].

4.1.2. Clustering. Clustering techniques have been widely used for mammographic density segmentation. The principle is to segment the breast tissue based on pixels with similar tissue appearance. Three groups can be identified, general (hard clustering), fuzzy (soft) clustering, and EM hierarchical clustering. A summary of representative studies using clustering techniques can be found in Table 2.

General Clustering. Oliver et al. [38] employed the K-means algorithm [120] for mammographic density segmentation, and breast density was categorised into fatty, glandular, and dense tissue. A set of morphological and cooccurrence matrix [121] based texture features were extracted from the segmented areas; these features were used for mammographic risk classification based on a leave-one- (image) out methodology. Results indicated that the risk classification accuracies decreased dramatically when including the glandular class; as dense breast area increases, more variations appear between the tissue clusters. This suggested that the K-means clustering based segmentation may be limited in dealing with inter- (density) class variation. Other texture and/or intensity based features can also be used in such a clustering based mammographic tissue segmentation; however, computational efficiency may decrease as the number of (combined) features increases. Dimensionality reduction can be used to remove redundant features. Strange et al. [39] used a manifold learning technique to preserve certain manifold properties using a large scale Kernel Principal Components Analysis (PCA) [122], as a means of reducing the data dimensionality prior to a clustering based mammographic tissue segmentation. K-means clustering was used to assign either fatty or dense tissue class to each image pixel based on the nearest cluster in low-dimensional manifold space. Results suggested that such a manifold learning was good for BI-RADS category 4 but less satisfactory for the other BI-RADS density categories.

Fuzzy Clustering. Oliver et al. [41, 42] used fuzzy C-means (FCM) [123] (an extension of K-means) to allow a cluster to be associated with one or several classes (i.e., fatty, glandular, and dense tissue); this is referred to as fuzzy membership function. In this improved approach, the classification was performed using the DDSM (Digital Database of Screening Mammographies) data [124], a publicly available digitised mammogram database. When compared with [38] the risk classification accuracies were improved. This may indicate that FCM produced more accurate segmentation, which resulted in more discriminative breast tissue features. Oliver et al. [43] quantitatively compared FCM with normalised cuts [125] and mean shift [126] for clustering based mammographic tissue segmentation. With respect to the subsequent risk classification, results indicated that FCM outperformed the other approaches. Some aspects regarding the actual density segmentation were lacking in related studies (e.g., [38, 41]) but were addressed in [45, 46]. In particular, the FCM approach to mammographic tissue segmentation was compared with other techniques. First, a fractal technique [127] was used to recursively split an image into quadrants, where the stopping criteria were determined based on local histogram measures on the consistency of uniform tissue. The quality of the fractal based approach resulted in pixelated segmentation due to its quad tree structure based splitting and analysis. Second, a statistical region based approach

TABLE 2: Summary of representative studies using clustering based methods for mammographic tissue segmentation. r/R denotes correlation coefficient. Note that (1) largely identical studies are excluded in the list and (2) in case of multiple results, only the best reported results are listed.

Study	Year	Number of density categories	Modalities	Number of views	Number of images	Segmentation evaluation	Risk/density estimation accuracy
General clustering							
Oliver et al. [38]	2005	Fatty and dense	Digitised SFM	MLO	180 (MIAS)	N/A	73% (fatty, glandular, and dense)
Strange et al. [39]	2013	Dense and fatty	FFDM	CC	12	Visually assessed	N/A
Marias et al. [40]	2005	Dense, semidense, and fatty	Digitised SFM	MLO	146	Visually assessed	65% (BI-RADS I and II) 86% (BI-RADS III and IV)
Oliver et al. [41]	2005	Dense and fatty	Digitised SFM	MLO	300 (DDSM)	N/A	47% (BI-RADS)
Oliver et al. [42]	2005	Dense and fatty	Digitised SFM	MLO	320 (MIAS) and 300 (DDSM)	N/A	48% MIAS, 47% DDSM (BI-RADS)
Oliver et al. [43]	2007	Dense and fatty	Digitised SFM	MLO	322 (MIAS)	N/A	78% (BI-RADS)
Oliver et al. [44]	2008	Dense and fatty	Digitised SFM	MLO and CC	322 (MIAS) and 831 (DDSM)	Visually assessed	77% MIAS, 86% DDSM (BI-RADS)
Oliver et al. [45]	2006	Dense and fatty	Digitised SFM	MLO	322 (MIAS)	Visually assessed	82% (BI-RADS)
Torrent et al. [46]	2008	Dense and fatty	FFDM	MLO and CC	300	Visually assessed	82% MLO, 75% CC (BI-RADS)
Tortajada et al. [47]	2012	Dense and fatty	Digitised SFM and FFDM	MLO and CC	322 (MIAS), 831 (DDSM), and 236 (digital DB)	Visually assessed	86% MIAS, 77% DDSM, and 92% digital DB (BI-RADS)
Adaptive/modified fuzzy C-means							
Chen and Zwiggelaar [48]	2010	4 densities	Digitised SFM	MLO	N/A	Visually assessed	N/A
Keller et al. [49]	2011	2–9 densities	FFDM (processed)	MLO	160	Visually assessed; $r = 0.75$ (automatic-manual)	N/A
Keller et al. [50]	2012	2–13 densities	FFDM (processed and raw)	MLO	160	Visually assessed; r (raw, processed) = 0.82, 0.85 (automatic-manual)	N/A
Expectation-maximisation							
Aylward et al. [51]	1998	Dense, fatty, and uncompressed-fatty	Digitised SFM	MLO and CC	70	Visually assessed	N/A
Zwiggelaar et al. [52]	2002	4–6 densities	Digitised SFM	MLO	263 (MIAS)	Visually assessed	67% (SCC)
Zwiggelaar et al. [53]	2003	4–6 densities	Digitised SFM	MLO	263 (MIAS)	Visually assessed	86% (SCC)
Zwiggelaar et al. [54]	2003	4–6 densities	Digitised SFM	MLO	263 (MIAS)	N/A	86% (BI-RADS)
Zwiggelaar and Denton [55]	2004	4 densities	Digitised SFM	MLO	36 (MIAS)	Visually assessed	75% (4 densities: 0%–10%, 11%–25%, 26%–50%, and 51%–75%)
Selvan et al. [56]	2006	5/8 densities	Digitised SFM	MLO	112 (mini-MIAS)	Visually assessed (92% good/excellent)	N/A

supported by the Fisherfaces algorithm [128] was used, in which a set of mammographic patches containing either dense or fatty tissue was used to model breast tissue. Subsequently, a model driven mammographic tissue segmentation was performed. Such an approach resembles rather large and coarse segmentation due to the block tissue analysis that could be less effective when tissue variations are small. Third, a multiple thresholding based on information theory (i.e., excess entropy [129]) was investigated. The excess entropy based algorithm loses accuracy in areas where the dense and fatty tissue boundary is less clear. The FCM outperformed the other approaches and produced more anatomically realistic segmentation, which led to satisfactory risk classifications. Tortajada et al. [47] extended and consolidated the segmentation framework in [46] for mammographic density segmentation using FFDM. A novel peripheral enhancement technique was developed to enhance texture appearance of uncompressed-fatty tissue near the breast skin line. Mammographic segmentation was significantly improved near the breast peripheral areas, leading to more correct features derived from the segmented dense tissue, and higher risk classification accuracies were achieved. Example mammographic segmentation can be found in Figure 5.

Adaptive/Modified Fuzzy C-Means. Chen and Zwiggelaar [48] proposed a modified FCM algorithm for mammographic density segmentation using the EPIC (European Prospective Investigation on Cancer) database [130]. The modified FCM incorporates local spatial and intensity information based on an adaptive local window filter; the weighting coefficients within the local window are used to differentiate the neighbouring pixels. The clustering performance on the intensity histogram of the filtered image is faster than conventional FCM. The conventional FCM algorithm uses grey-level information at a single pixel as the feature space and this contains no spatial contextual information, which makes it very sensitive to noise and intensity inhomogeneities. Segmentation results indicated the robustness of the modified FCM in dealing with intensity inhomogeneities with different density categories. Visual assessment indicated that local window filtering was able to eliminate intensity inhomogeneities and avoid excessive blur.

Keller et al. [49] presented a novel multiclass FCM algorithm for automated identification and quantification of breast density, which is optimised for the imaging characteristics of digital mammography. The proposed algorithm involves an adaptive histogram-based method to estimate the number of clusters, which uses the tissue properties of the specific mammogram followed by segmentation through clustering using linear-discriminant analysis (LDA). The classifier combines imaging and patient characteristics to achieve optimal segmentation through cluster merging. A set of FFDM was used in the evaluation, and a strong correlation was observed between the estimated PD and radiological "ground truth" using BI-RADS density categories. Results also indicated relatively poor performance when compared to identifying breast dense tissue based on two-class FCM paradigm using digitised SFM. This may be due to the fact that the majority of grey-level intensity profiles of breast

tissue as extracted from FFDM tend to be multimodal. This finding is not in line with the results reported in [46], in which a good visual agreement was observed between FCM and expert annotations. This may be due to the fact that the clustered features are different; cooccurrence matrices and grey-level intensities features were used in [46] and [49], respectively. Keller et al. [50] extended the adaptive multicluster FCM approach to estimate breast PD in both raw and postprocessed FFDM images. This is expected to be beneficial in terms of direct clinical application and retrospective analysis.

Expectation-Maximisation. Aylward et al. [51] investigated a mixture modelling technique to differentiate fatty and dense breast tissue. The statistical modelling of the breast components was based on pixel intensity distribution sampling, using Gaussian mixture models, but no texture aspects were taken into account. The parameters of the mixture were iteratively determined using the EM algorithm, maximising the log-likelihood of the data representing the distribution [131]. To quantify accuracy, automatic dense tissue estimates were compared with the "ground truth" provided by experts. A set of images from three different mammography unit manufacturers was used in the evaluation. Whilst breast density estimates were satisfactory, in some cases, the separation of fat and dense tissue was not achieved. Nevertheless, results from this early study showed possible intensity distribution discrepancy (e.g., variation in breast tissue contrast) when using images acquired from different manufacturers.

Zwiggelaar et al. [52] investigated a combination of statistical modelling and EM algorithm for a texture based approach to mammographic tissue segmentation. The investigation consists of basic grey-level information and spatial correlation information, both combined to achieve texture modelling. The statistical modelling was used for data generalisation and noise removal purposes. Segmentation was derived using an information theoretic approach [132]. The segmented breast parenchyma is in line with breast anatomical structures, and the accuracy was influenced by the size and shape of the local neighbourhood. In a follow-up study, Zwiggelaar et al. [53] incorporated a set permutation-occurrence matrices to encode texture features. The study emphasised that, for texture based mammographic tissue segmentation, it is important to incorporate both grey-level value and spatial correlation information. The use of cooccurrence matrices in texture based segmentation has a major disadvantage, because in principle the dimensionality for the derived texture features is infinite, leading to redundant features and time consuming analysis. Zwiggelaar and Denton [55] addressed this issue by qualitatively selecting a subset of a large set of cooccurrence matrices. A transportation measure [133] was used to determine the difference between cooccurrence matrix based texture features. Segmentation results showed improvements over tissue specific areas when compared to previous studies [52–54]. This indicated that the segmentation performance is directly related to the transportation measure of an ordered set of cooccurrence matrices.

When using the conventional EM based segmentation, the efficiency of the stochastic model depends on the accuracy of estimation of the model's parameter set. Selvan et
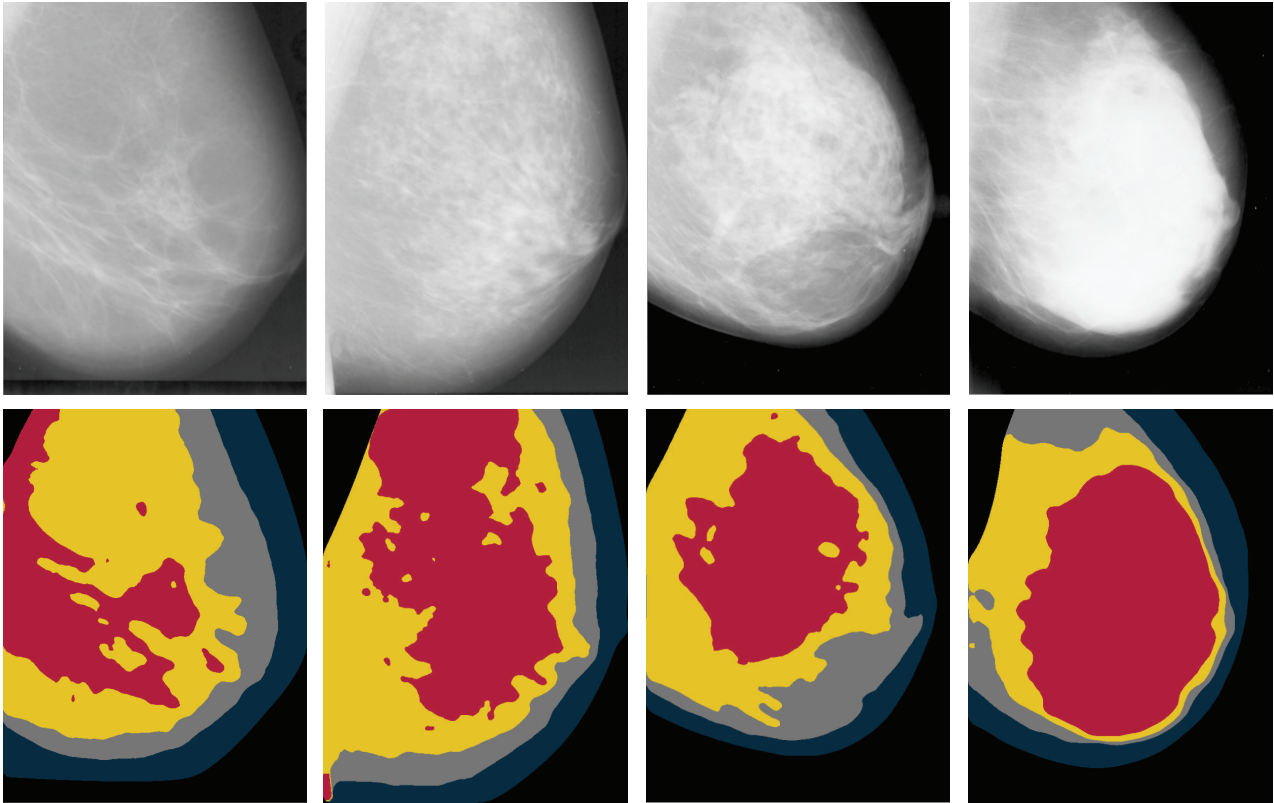
FIGURE 5: Example mammographic density segmentation. Images (digitised mammograms) from left to right are rated with BI-RADS 1–4, respectively. Fatty, semifatty (e.g., scattered fibroglandular tissue), semidense (e.g., heterogeneous dense tissue), and dense tissue are colour coded as navy, grey, yellow, and red, respectively.

al. [56] proposed an advanced EM approach to estimate the model parameter set more accurately for mammographic tissue segmentation. Two heuristic search techniques were investigated, particle swarm optimization [134] and evolutionary programming [135]. Experimental results indicated better approximation of a given histogram by a mixture of density functions and reduced computational time when compared to the popular EM algorithm. Visual assessment showed that when compared to the reconstructed image using standard EM parameter estimates, irrelevant image details are largely reduced in the resegmented image where parameters are estimated by the heuristic approach. The usefulness of the optimisation approach depends on accurate estimation of the model parameter set and incorporation of context information.

4.1.3. Statistical Model Building. There are two main types of statistical model building techniques to model breast tissue, those based on texture statistical variation or those which use texture descriptors. These techniques are closely related and can be combined. A summary of representative studies using statistical model building techniques can be found in Table 3.

Texture Statistical Variation. Miller and Astley [57] were among the first to perform mammographic tissue segmentation using texture features. Two approaches were investigated,

morphological operation based granulometric techniques [136] and Laws texture energy method [137]. In the granulometric based multiresolution method, morphological structuring elements of various sizes were used to cover the range of anatomical structures. Their shape determines the texture characteristic to be measured; in particular, a set of circular structuring elements were used to produce images in which intensity represents texture coarseness, whilst for Laws texture energy based approach, a Bayesian classifier [138] was trained on a leave-one- (image) out basis to produce glandular probabilities for an unseen mammogram. The investigation indicated that breasts with dense-glandular patterns and relatively smooth disc appearances were falsely classified as fatty tissue; the misclassification was linked to poor textural detail as seen on digitised SMF. This may have been due to the quality of digitisation, spatial and grey-level resolution.

Suckling et al. [58] proposed a method to segment mammograms into four classes (i.e., background, pectoral muscle, fibroglandular tissue, and fatty tissue). The first- and second-order grey-level histogram features were calculated and used in multiple linked self-organising neural networks for pixel classification [139]. The developed approach did not use separate data explicitly to train breast tissue models; training was performed over the novel images themselves, at tissue specific areas based on prior knowledge. It was assumed that there is a certain consistency in the patient positioning,

Table 3: Summary of representative studies using statistical model building based methods for mammographic tissue segmentation. r/R denotes correlation coefficient. Note that (1) largely identical studies are excluded in the list and (2) in case of multiple results, only the best reported results are listed.

Study	Year	Number of density categories	Modalities	Number of views	Number of images	Segmentation evaluation	Risk/density estimation accuracy
Texture statistical variation							
Miller and Astley [57]	1992	Fatty and dense	Digitised SFM	MLO	40	Visually assessed; 80% overlapping areas matched	N/A
Suckling et al. [58]	1995	Fatty and fibroglandular	Digitised SFM	MLO	30 (15 pairs)	Visually assessed; 69% ± 12% overlapping areas matched	N/A
Heine and Velthuizen [59]	2000	Fatty and dense	Digitised SFM	MLO and CC	50	Visually assessed	N/A
Heine et al. [60]	2008	Dense and fatty	Digitised SFM	MLO and CC	369 cases and 712 controls	Visually assessed	Pearson R (CC) = 0.70 (automatic–Cumulus) Spearman's r (CC) = 0.49 (automatic–BI-RADS)
Petroudi and Brady [61]	2006	Dense, fatty, and breast edge	Digitised SFM	MLO and CC	32	Visually assessed (88% very satisfactory); ~94.4% overlapping areas matched Visually assessed; ~87.9%	N/A
Gong et al. [62]	2006	Dense, dense with structures, fatty, and fatty breast edge	Digitised SFM	MLO and CC	43	overlapping areas matched Visually assessed; accuracy 0.916 ± 0.038, area overlap based on 15 images	~87.9% (Wolfe)
Oliver et al. [63]	2010	Dense and fatty	Digitised SFM and FFDM	MLO and CC	322 (MIAS) and 250 (the Trueta DB)	0.900 ± 0.122, and Dice coefficient 0.943 ± 0.077 (60 MLO from the Trueta DB)	N/A
Texture descriptors							
Zwiggelaar and Denton [64]	2006	4 densities	Digitised SFM	MLO	60	Visually assessed	72% (Wolfe)
Adel et al. [65]	2007	Fibroglandular, fatty, and uncompressed and fatty	Digitised SFM	MLO and CC	50 (mini-MIAS)	Visually assessed (68% good); >60% agreement with manual segmentation	N/A
Zwiggelaar [66]	2010	4 densities	Digitised SFM	MLO	322 (MIAS)	Visually assessed	64% (BI-RADS)

and the positioning of anatomical structures is in the same regions. Segmentation results indicated that breast peripheral boundaries were difficult to identify due to X-ray photon penetration which is almost complete over these areas. Visual assessment indicated reasonably good segmentation with a tendency of oversegmentation; however, anatomic structures with variable appearance of the fibroglandular tissue were not identified correctly. The position of the parenchyma alters dramatically across the series of images, but the algorithm does seem to produce consistent parenchymal shape over several images. It was concluded that high density tissue was found to be more prone to misclassification, and the number of neurons used had a direct effect on the performance of the network. This early study indicated that potential image acquisition related information (e.g., the thickness of the compressed beast, the attenuation coefficients of the breast tissue, and breast anatomical models) may be useful for the task but was difficult or unattainable when using digitised SFM.

Petroudi and Brady [61] proposed a statistical modelling approach to mammographic tissue segmentation based on a framework previously developed for mammographic risk classification [140]. The approach focused on texture analysis over a more localised area, and a pixel based classification can be used as a means of segmenting breast parenchymal into different densities. To incorporate both contextual and spatial neighbourhood (structural) information, multivector Gaussian HMRF (Hidden Markov Random Field) [141] and texton (texture primitives) [140] technique were used. The multivector image representation [142] is achieved using a filter bank [143], whilst all the parameters were estimated using the EM algorithm. The key aspect in the study is the use of texton based statistical modelling for rotation invariant mammographic texture. The hypothesis is that pixels from similar breast tissue have similar texture properties, even when there are relatively large anatomical changes due to involution and use of hormone replacement therapy. However, issues still remain with the automatic determination of the appropriate filter bank, the number of textons, and the size of the neighbourhood. The texton selection aspect was investigated in a related study [62], which again used textons and HMRF for mammogram tissue segmentation. Similar segmentation results were achieved when a greedy algorithm was incorporated in the textons learning process to remove similar textons.

Oliver et al. [63] presented a mammographic density segmentation which utilises modelling of a set of patches of either fatty or dense parenchyma and statistical analysis. Two modelling strategies were investigated; one is a Karhunen-Loeve-based model with PCA [144] and the other is a linear-discriminant-based model using LDA [128]. Once the tissue models are learnt, pixel based two-class (i.e., fatty and dense) segmentation was performed using a nearest neighbour classifier. The evaluation was performed using both digitised SFM and FFDM. There were noticeable variations observed in tissue appearance for the two databases; therefore, the breast tissue modelling was independently performed for both databases. The study indicated that the number of patch samples used is vitally important in the training process; it

should be large enough to provide sufficient data variation per tissue class but small enough to avoid overfitting of the classifier. It may be difficult to subsample the square or rectangular patches with the precision needed within a mixture of tissue, leading to oversegmentation. A further detailed annotation over patches could be employed to address this issue; however, this procedure can be labour intensive and time consuming. It was also noticed that the algorithm may not be able to correctly identify fatty tissue, some ducts, and linear structures that are much brighter than the fatty breast tissue. This is due to texture modelling variation when using digitised and digital mammograms. Many of the issues related to patch based breast tissue modelling are in line with a number of studies reviewed in Section 4.2.

Texture Descriptors. Zwiggelaar and Denton [64] developed a texture classification approach using contrast information, based on the concept that texture can be discriminated from the contrast between key structural elements and their repeating patterns. The developed texture analysis method is related to local binary patterns (LBP) [145] and is similar to SUSAN [146]. Mammographic texture is modelled by estimating local aspects using a set of binary images which can be generated by thresholding using the corresponding grey-level bands. Unlike the LBP and SUSAN approaches, this method does not utilise histogram information to extract texture information, only one model per grey-level band exists, and these models can be compared directly. A hybrid metric based on probability density distributions and transportation estimation was used to classify unseen pixels as a means of segmentation. Results showed a strong correlation between the various texture regions in the mammographic images and the segmented areas. The region boundary effects seem to play a significant role, as most segmented areas were mainly occupied by two classes; this is partly due to unbalanced distribution over the density classes.

Zwiggelaar [66] investigated a breast tissue segmentation methodology based on local grey-level appearance histograms. The mammographic texture modelling incorporated both grey-level and spatial aspects. Variation in local grey-level appearance is represented in histogram format for which the distribution varies with BI-RADS breast density categories. Visual assessment indicated realistic mammographic tissue segmentation but is predominately two-class segmentation associated with BI-RADS II and III, whilst the breast boundary regions are often associated with high risk BI-RADS III and IV. It should be noted that the underlying information within the local window is the same as that being used in texton based approaches [61], where the texture models are formed by cluster centre related histograms. A distinct difference compared to LBP based texture analysis is that the developed methodology uses the full grey-level range instead of reduction to binary patterns, which means the resulting histograms might contain more sparsely populated texture information.

4.1.4. Collective Multiple Measurements. According to the literature, different approaches to breast density estimation can be performed equally well. However, manual/semiautomatic

segmentation can remain superior despite some of the inherited limitations (e.g., losing accuracy due to fatigue and time consuming process). It may be that a human observer is able to combine context, morphology, and textural information, whereas many automated methods only focus on one of these characteristics [67]. Therefore, it is conceivable to utilise various principles for mammographic tissue segmentation. A summary of representative studies using collective multiple measurements can be found in Table 4.

Kallenberg et al. [67] developed a breast density segmentation method based on pixel classification, in which different approaches known in the literature to segment breast density are integrated and extended. The features used include, for example, location of clustered high density tissue in relation to the skin line and nipple, dense tissue intensity information, Gaussian derivatives, cooccurrence matrix based Haralick texture features, and global context features calculated from the whole image. A neural network classifier was trained based on segmentations obtained using the Cumulus. The sequential floating forward selection (SFFS) algorithm [147] was applied to select the optimal subset from a large pool of features, which removed irrelevant and redundant features from the data, and a threshold was defined for the classifier output to obtain a dense tissue segmentation. The PD results show a high correlation ($r = 0.9$) between the automatic measurements and the Cumulus results. It was concluded that a combination of segmentation strategies outperforms the application of single segmentation techniques.

Li et al. [68] investigated image-processing software based on ImageJ [148] for automated analysis of mammographic density and penalised regression to construct a measure that mimics the Cumulus. A set of automated thresholding methods were applied to separate the dense breast tissue, and these vary according to the type of pixel intensity information (e.g., histogram shape, clustering, and entropy). A watershed algorithm was employed to subdivide dense tissue areas into smaller objects from which a variety of measurements were obtained for the breast as a whole, as well as for the "objects" of dense tissue. A total of 1008 measurements were obtained as output from ImageJ, but only 772 variables were informative. The feature dimensionality was further reduced to 123 using PCA. To compare breast density as measured by the Cumulus, a large dataset containing cases (having condition/disease) and controls (not having condition/disease) was used for the evaluation. Results indicated that the mammographic density measurement has a correlation equal to $r = 0.875$ which was similar to that reported by Kallenberg et al. [67] ($r = 0.895$) and substantially higher than the work done by Heine et al. [60] ($r = 0.70$). Note that generalisation of the developed method is currently limited to MLO images. So far, this study provided the strongest evidence that mammographic images contain additional information to percentage density which improves the ability to discriminate between breast cancer disease statuses.

4.1.5. Other Methods. There are some methods in the literature which have not been widely adopted, most conducted as proof of concept or feasibility studies. Although no quantitative and qualitative evaluations were performed with respect

to mammographic tissue segmentation, some of the ideas are worth noting which may inspire further improvements and investigations. A summary of representative studies can be found in Table 4.

Lao and Huo [149] developed a hierarchical approach to segment breast dense tissue from mammographic images based on unsupervised learning and multiple levels of detail. The method initially segments the breast based on entropy maximum thresholding. Next, the resultant dense and fatty tissue segmentation was used to facilitate FCM membership function initialisation, followed by clustering based segmentation. Finally, the grouped dense tissue was used to generate dense tissue intensity and homogeneity features; a pixel based thresholding technique was used to produce the final dense breast segmentation. Chen et al. [70] developed a mammographic tissue segmentation method using topographic maps of breast regions at multiple intensity levels that represent both topological and geometrical structures of different dense breast tissue. A topographic map is a morphological and multiscale decomposition of an image relying on the connected components of level sets. The topological and geometric structures are represented by a shape tree, from which dense tissue regions are detected by analysing the saliency and independency of the shapes. This is an unsupervised method as it does not require a learning stage or prior knowledge. The analysis is based on components of the topographic map instead of the image pixels, which could significantly reduce the dimensionality of the data to be analysed. Postprocessing is required to compute geometric moments of the level sets in order to remove incorrect dense regions. Visual assessment indicated that the derived segmentation resembled the results obtained using threshold based methods; however, segmentation seems to omit anatomical structures which do not appear as dense as they should be.

4.2. 2D Projection Based Approaches Using Parenchymal Pattern. Another group of mammographic tissue segmentation approaches focuses on breast tissue separation beyond different densities (e.g., dense and fatty) and using Tabár's tissue modelling instead to group breast tissue into four classes (i.e., linear, nodular, homogeneous, and radiolucent); see Section 2.2 for details of this scheme.

4.2.1. Statistical Model Building. These methods are all related to statistical model building using texture statistical variation or descriptors. A summary of representative studies with respect to Tabár's tissue modelling can be found in Table 5.

Texture Statistical Variation. Muhimmah et al. [71] applied a texton related technique for mammographic tissue segmentation based on Tabár's tissue modelling. A texton selection strategy was incorporated using a combination of visual assessment and minimum spanning tree topological information. The textons for each tissue type (i.e., nodular, linear, homogeneous, and radiolucent) were generated from mammographic patches containing tissue specific samples. In the texton selection process, the (Euclidean) minimum spanning tree was used to indicate a topologically probable

TABLE 4: Summary of representative studies using other less popular methods (e.g., collective multiple measurements) for mammographic tissue segmentation. r/R denotes correlation coefficient; AUC denotes area under ROC curve. Note that (1) largely identical studies are excluded in the list and (2) in case of multiple results, only the best reported results are listed.

Study	Year	Number of density categories	Modalities	Number of views	Number of images	Segmentation evaluation	Risk/density estimation accuracy
Collective multiple measurements							
Kallenberg et al. [67]	2011	Dense and fatty	Digitised SFM	MLO	1300	Visually assessed; Pearson R (percent density, dense area) = 0.911, 0.895 (automatic-Cumulus) Visually assessed; $r = 0.884$ (automatic-Cumulus);	N/A
Li et al. [68]	2012	Dense and fatty	Digitised SFM	MLO	765 cases and 747 controls	AUC = 0.589 (four densities: <5%, 25%–50%, 50%–75%, and >75%)	N/A
Other methods							
Chen et al. [69]	2012	Dense and fatty	Digitised SFM	MLO	321 (MIAS)	Visually assessed	70% (BI-RADS)
Chen et al. [70]	2013	Dense and fatty	Digitised SFM	MLO and CC	321 (MIAS) and 831 (DDSM)	Visually assessed	76% MIAS, 81% DDSM (BI-RADS)

TABLE 5: Summary of representative studies using statistical model building methods for mammographic tissue segmentation. Note that (1) this list consists of studies mainly related to Tabár tissue modelling and (2) in case of multiple results, only the best reported results are listed.

Study	Year	Number of mammographic building blocks	Modalities	Number of views	Number of images	Segmentation evaluation	Risk/density estimation accuracy
			Texture statistical variation				
Muhimmah et al. [71]	2007	Linear, nodular, homogeneous, and radiolucent	Digitised SFM	MLO	320 (MIAS)	Visually assessed	N/A
He et al. [72]	2008	Linear, nodular, homogeneous, and radiolucent	Digitised SFM	MLO	320 (MIAS)	Visually assessed	N/A
He et al. [73]	2009	Linear, nodular, homogeneous, and radiolucent	Digitised SFM	MLO	320 (MIAS)	Visually assessed (65% good/very good) Tabár (38% good/very good) BI-RADS	53% (Tabár), 70% (BI-RADS)
He et al. [74]	2011	Linear, nodular, homogeneous, and radiolucent	Digitised SFM	MLO	320 (MIAS)	Visually assessed	53% (Tabár)
He et al. [75]	2012	Linear, nodular, homogeneous, and radiolucent	Digitised SFM	MLO	320 (MIAS)	Visually assessed	85% (Tabár), 78% (BI-RADS)
He et al. [76]	2014	Nodular, homogeneous, and radiolucent	FFDM	MLO and CC	360	Visually assessed	79% (Tabár), 80% (BI-RADS)
			Texture descriptors				
He et al. [77]	2010	Linear, nodular, homogeneous, and radiolucent	Digitised SFM	MLO	320 (MIAS)	Visually assessed	78% (Tabár), 75% (BI-RADS)
He and Zwiggelaar [78]	2013	Nodular, homogeneous, and radiolucent	FFDM	MLO and CC	360	Visually assessed	N/A

correct connectivity, in high dimensional space. Distinct textons (higher discriminative power) tend to be situated towards the outer edges of the tree, whilst common texture, noise, and intensity aspects tend to be modelled by textons in the central part of the tree. Subsequently, a model driven based mammographic tissue segmentation was performed using the selected textons. An alternative texton selection process using texton ranking, outlier detection, and visual assessment can be found in [150]. Both studies partially addressed the texton selection issue discussed in [61] and showed realistic segmentation results.

He et al. [73, 74] investigated a mammographic tissue segmentation based on spatial moments and prior information (e.g., shape related texture features) of mammographic building blocks. A set of geometric moments (spatial moments) of different orders were computed over local textures. They were used as global shape descriptors to encode different spatial characteristics of mammographic texture intensity distribution. Visual assessment indicated that the spatial moments based texture features are able to capture orientation sensitive texture properties, as well as more complicated textural properties, resulting in realistic mammographic tissue segmentation. However, the segmentation has a tendency of oversegmenting linear structures and undersegmenting small areas from nodular and homogeneous regions. Note that a low intensity background (e.g., low or zero grey-level values) can lead to "meaningless moments" [151] which can cause misclassification between high density homogeneous tissue and low density radiolucent tissue. The developed methodology is capable of modelling complex mammographic images and can deal with intraclass variation and noise aspects. It is particularly suitable for modelling Tabár's mammographic building blocks as different breast tissue types have texture patterns related not only to their periodicity but also to shape features.

Clinical observations indicated that breast peripheral areas may not be fully compressed which may cause unexpected intensity and texture variation within these areas. Such breast parenchymal appearance discrepancies may not be desirable for tissue modelling within computer aided mammography. He et al. [76] developed a mammographic image preprocessing (peripheral enhancement) method to improve the image quality before analysis (similar issue was addressed in [152]). Visual assessment indicated significant improvement on segmented anatomical structures and tissue specific areas when using the processed images.

Texture Descriptors. He et al. [77] developed a texture signature [153] based methodology for mammographic tissue segmentation in an attempt to incorporate spatial and geometric texture features. The framework is similar to the moments based approach [74], but for each pixel a texture signature is generated which consists of three distinct subsignatures. It is effectively a stack of three 2D histograms, encoding directional texture features (e.g., intensity, orientation, and elongation variation). Visual assessment indicated good and consistent segmentation results, and it was robust in dealing with the "meaningless moments" issues encountered in [74]. Two aspects were investigated in subsequent studies [75, 78] to

further improve the framework. First, a binary model matching pattern based Bayes classifier was investigated in [75]. An idea similar to "fuzzy membership" was used to allow a pixel to be associated with multiple classes at the same time. The process generates a binary pattern of the "fuzzy membership" through model matching and is subsequently used in training the Bayes classifier. Second, a feature and classifier selection technique was proposed in [78]. Results indicated that at the tissue modelling stage, over- and/or undertraining can cause tissue composition fluctuation between nodular and homogeneous tissue, whilst the percentages of radiolucent tissue were less sensitive to the algorithm's parameter configurations. This resulted in inadequate compositions for nodular and homogeneous tissue according to Tabár's parenchymal patterns. Despite the shortcomings in some cases, these improvements helped to produce more realistic segmentation and improved risk classification accuracies. Example mammographic segmentation can be found in Figure 6.

4.3. 2D Projection Based Volumetric Approaches. Due to technological limitations, 2D projection area measures are known to be flawed [154, 155] as they do not provide measures of the fibroglandular tissue volume (FVG). With X-ray mammograms, the brightness at any given point in the image is dependent on the thickness of glandular and other dense tissues projected onto that point. The arrangement of glandular tissue within the breast depends on the way in which the breast is compressed. Therefore, measures of the projected dense tissue area vary depending on compression. Substantial efforts have been directed towards volumetric breast density estimation techniques which may allow for improved assessment of breast fibroglandular tissue; this may in turn lead to improved breast cancer risk assessment [156]. Three groups of 2D projection based volumetric approaches can be identified, prior calibration, in-image phantom based calibration, and physical image formation models. A summary of representative studies using 2D projection based volumetric approaches can be found in Table 6.

4.3.1. Prior Calibration. Highnam and Brady [157] developed the Standard Mammogram Form (SMF) to measure glandular tissue composition objectively, with the details in Section 4.3.3. However, only "relative" tissue measurements were obtained unless a careful calibration was performed. Following their study, Kaufhold et al. [158] proposed a calibration approach to glandular tissue composition estimation in digital mammography. The approach was evaluated using breast phantoms of varying thickness and FFDM. From these images, mean signal, noise levels, and computed calibration curves were extracted for quantitatively estimating tissue composition on a pixel-wise basis. In terms of different error sources on the estimates of tissue composition, the initial study concluded that these error sources include compressed breast height estimation error, residual scattered radiation, quantum noise, and beam hardening. Errors in the estimated compressed breast height contributed the most error in tissue composition. Heine et al. [79] proposed a calibration method as an extension to [158] which coupled with first-order histogram statistics as an automated surrogate for breast density

TABLE 6: Summary of representative studies using 2D projection based volumetric methods for mammographic tissue segmentation. r/R denotes correlation coefficient. Note that in case of multiple results, only the best reported results are listed.

| Study	Year	Number of density categories	Modalities	Number of views	Number of images	Segmentation evaluation	Risk/density estimation accuracy
Prior calibration							
Heine et al. [79]	2011	Dense and fatty	FFDM (raw)	CC	106 cases and 106 controls	Visually assessed; linear R = 0.78 (automatic-Cumulus)	Risk estimates associated with the lowest to highest quartiles, odds ratios: 1.0, 3.4, 3.6, and 5.6
In-image reference phantom based calibration							
Alonzo-Proulx et al. [80]	2012	Dense and fatty	FFDM	CC	55087	Pearson r (left, right volumetric density-breast volume) = 0.92, 0.91	N/A
Ourselin et al. [81]	2014	Dense and fatty	FFDM	CC	480	Percent breast fibroglandular volume, r^2 = 0.8 (automatic-predicted)	N/A
Physical image formation model							
Hartman et al. [82]	2008	Dense and fatty	FFDM and MRI	MLO and CC	550 (275 pairs) and 88 MRI	Breast density volumes, Pearson (left-right breast, Quantra-MRI) r = 0.923, 0.937	N/A
Highnam et al. [83]	2010	Dense and fatty	FFDM and MRI	MLO and CC	2217 and MRI from 26 younger women	Breast density volumes, Pearson r (left-right breast, CC-MLO view, Volpara-MRI) = 0.923, 0.915, 0.94 Pearson r (volumetric breast density	N/A
Gubern-Mérida et al. [84]	2014	Dense and fatty	FFDM and MRI	MLO and CC	680 and 168 MRI	Volpara-MRI, fibroglandular tissue volume Volpara-MRI) r = 0.91, 0.84	Density grade, r = 0.40 (Volpara-BI-RADS)
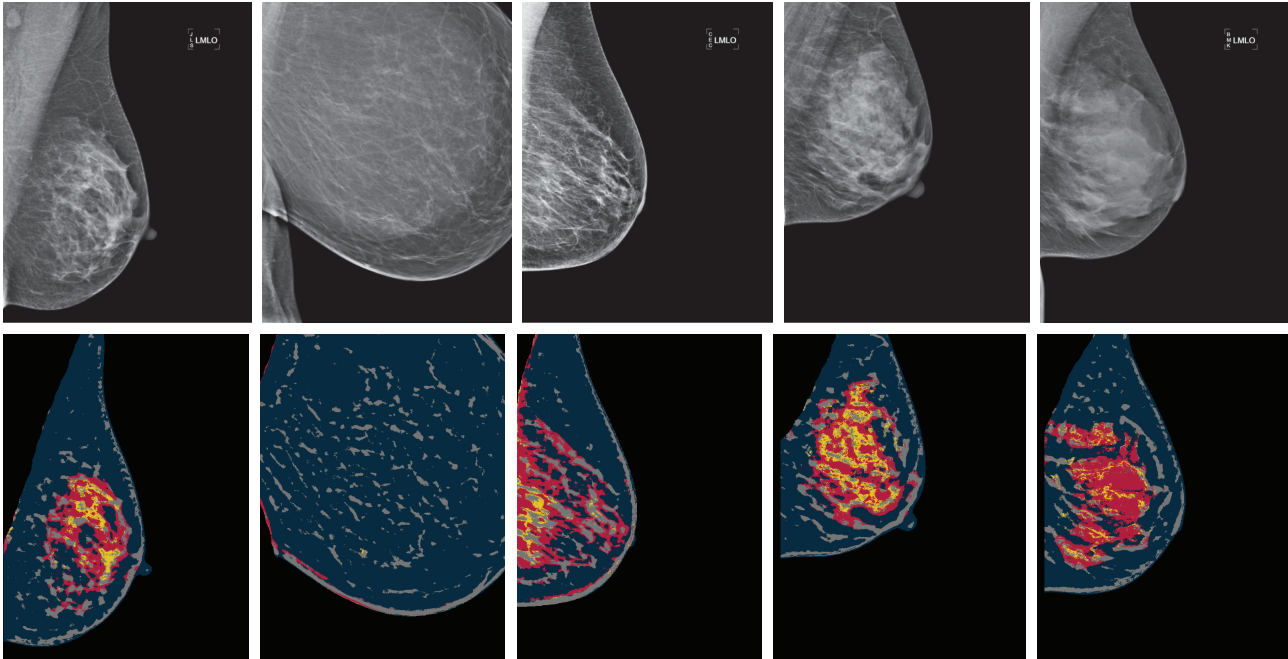
FIGURE 6: Example mammographic parenchymal segmentation. Images (digital mammograms) from left to right are rated with Tabár I–V, respectively. Nodular, linear structure, homogeneous, and radiolucent areas are colour coded as yellow, grey, red, and navy, respectively.

assessment from FFDM for the purposes of estimating cancer risk directly. It should be noted that this method requires the use of raw FFDM explicitly. The calibration process produces a normalised effective X-ray attenuation coefficient scale at the pixel level referred to as the percent glandular representation. Mammograms are standardised automatically in this way; the calibration resulted in normalised images with pixel values ranging from 0 to 100, where increasing value represents increased X-ray attenuation. The PD measure is an estimate of the number of pixel values above a fixed idealised X-ray attenuation fraction.

4.3.2. In-Image Reference Phantom Based Calibration. Pawluczyk et al. [159] described a volumetric analysis of the mammographic density from digitised SFM. An initial calibration of the imaging system is required with a tissue-equivalent plastic device. The subsequent correction for exposure factor variations and film processing characteristics is performed through images of an aluminium step-wedge placed adjacent to the breast during imaging. In order to calculate volumetric breast density (VBD), information about the compressed breast thickness and system parameters used for taking the mammogram and information from the calibration device are all taken into account. Results obtained for the known density phantoms showed that VBD can be estimated to be within 5% accuracy of the actual value. However, thickness correction is needed to correct inaccuracy in the compression thickness indicator of the mammography units. Alonzo-Proulx et al. [80] adapted the work of Pawluczyk et al. [159] and proposed a volumetric method which incorporated a more robust algorithm to estimate the thickness of the breast [160, 161] using a prior

calibration of the digital image signal versus tissue thickness and composition. The thickness of the compressed breast is estimated using an empirical model that corrects the thickness readout of the mammography system as a function of compression force. The developed VBD measurements are in good agreement with the VBD measured on a dedicated breast CT system [162].

Hufton et al. [163] developed a quantitative method for determining the volume of dense breast tissue from digitised SFM, similar in principle to [159]. This automated technique used a calibrated step-wedge placed alongside the breast during mammography; markers on the compression plate enable accurate measurement of compressed breast thickness and the thickness of dense tissue to be determined at each pixel. Data from the step-wedge and acquired images were then used to measure total breast volume, glandular volume, and the derived percentage of volumetric dense breast tissue. Such a method is insensitive to variations in exposure parameters and film processing. Note that where the breast loses contact with the compression plate (e.g., peripheral area), a rank order filter based on noise estimation is used to find the outer edge and a semicircular profile model is used to determine breast thickness in this region. The original design of the step-wedge (35 mm) was too big to fit alongside larger breasts and the lead lining made it a relatively heavy, unwieldy device that could not be easily attached to the bucky [164]. Identification of step positions in the digital image is required for analysis, but one end of the wedge was overexposed and the other underexposed; therefore finding the ends accurately was challenging.

Shepherd et al. [165] developed a method to measure the percentage fibroglandular tissue (PFV) and FGV using single

X-ray absorptiometry (SXA) in order to obtain tissue volumetric measures during image acquisition using a reference phantom. This can generate an accurate thickness profile of the breast at each image position and calculate an image specific calibration of the pixel grey scale values to breast composition. It is a promising method which may be utilised in commercial breast image equipment. Closely related work can be found in [166, 167]. A modified calibration approach for the SXA was proposed in [168]. The method takes into account both geometric and image related factors that impact the calibration of grey-level into absolute tissue composition. Breast density is calculated using the thickness correction factors and recalibration procedures, where all necessary parameters are extracted automatically. Note that it is essential to conduct weekly phantom scanning as part of quality control monitoring, with a specially designed calibration phantom to control thickness and grey-level conversion stability. This weekly maintenance could be a limitation to the SXA approach. To address this, Ourselin et al. [81] developed a method to automatically quantify volumetric breast density without the use of phantoms. It applies a shape and appearance model to digital mammograms and combines volumetric density measures derived by the SXA method to build a statistical model based on image parameters extracted from the image.

4.3.3. Physical Image Formation Models. Highnam and Brady [157] proposed a normalised representation of the breast which was referred to as the Standard Mammogram Form (SMF). Such an approach transforms digitised SFM to an "interesting" tissue representation. The key assumption in the normalisation method is that the X-ray attenuation coefficients of fibroglandular and abnormal tissue are almost equal and are quite different from that of fatty tissue. The technique utilises mammographic information which relates to the physical volume of the breast; the height of nonfatty tissue under compression that corresponds to each pixel is calculated in order to estimate the volume of "interesting" tissue. The volumetric values of "interesting" tissue have a good anatomical correspondence with the expected tissue density of each mammographic risk. A mammogram correction is required for scattered radiation and the dependency of image formation parameters (e.g., tube voltage (kVp), spectrum, and exposure time (mAs)). Nevertheless, such a volumetric breast density estimation is fully automated and is able to compensate for calibration data errors (e.g., mAs, kVp, and spectra) by using image-based calibration [169]. Due to relatively weak control over the image acquisition process, it is difficult to eliminate variability in image characteristics (e.g., contrast and brightness), leading to variation in the mammographic intensity distribution, and increased difficulty in identification of grey-level texture based mammographic patterns [170]. However, closely related studies [171, 172] noted that possible errors can occur in the calculation of tissue height, which influences the overall accuracy of the measures. The main disadvantage of SMF is that a complete and substantial set of calibration data would be needed to generate realistic breast composition measures and yet there are many trials that have retrospectively collected images

with no calibration data. The issue of measuring breast composition using retrospective digitised SMF with little or no calibration data available was later addressed in [169]. Such a volumetric technique was criticised as less reliable than 2D threshold based methods [173] and did not provide a stronger predictor of breast cancer risk [174].

van Engeland et al. [175] presented a method for estimation of dense breast tissue volume using FFDM. Such an approach maps the thickness of dense tissue to a pixel, determined using a physical model of image acquisition. Linear attenuation coefficients of breast tissues were derived from empirical data as a function of tube voltage, anode material, filtration, and compressed breast thickness. The developed volumetric breast density estimation can be severely hampered or rendered impossible for contrast enhancement images, because such images violated the assumption made that pixel values are proportional to exposure. Results were significantly worse when applied to images after logarithm transformation ("for presentation/processed" images) than those obtained from the raw images. Over- or underestimation of breast density volume can be seen when the method was applied to thickness corrected images, whilst very dense breasts can cause difficulty in obtaining a reliable calibration location. In contrast to the SMF approach [157], many aspects of the image formation process and parameters related to the acquisition procedure were required, whilst in [175] many calibration parameters can be obtained from the DICOM header. Note that all volumetric techniques from 2D projection data assume two components (i.e., fatty and dense tissue) in the breast; however, more recent volumetric developments (e.g., van Engeland's work) are far less reliant on the image physics, although there is still some sensitivity to compressed breast thickness [176].

4.3.4. Commercial Software. Hartman et al. [82] presented a volumetric breast density technique as an extension to SMF [157]. The software is named Quantra, which significantly improves Highnam and Brady's SMF approach [157]. The performance of SFM based volumetric breast density estimation and breast cancer risk has been reported to be equal to visual assessment methods (e.g., Wolfe and SCC) but poorer than the Cumulus. This may be due to error compensation in the calibration data by finding the breasts fatty peripheral areas which is difficult for very dense breasts, leading to underestimation in volumetric density and undermining breast cancer risk predictions [169].

Highnam et al. [83] built upon the work in [175] and developed software named Volpara that uses relative physics modelling together with additional information derived from the image to substantially reduce dependence on imaging physics data. The volume of dense tissue is found by a ratio of integrated (dense tissue) values over the breast areas, and the volume is calculated by multiplying the area of the breast using the recorded breast thickness. One of the difficulties in volumetric based breast density estimation, as pointed out in [173], is in finding a reference area of the breast which is entirely fat, especially when the breast is very dense.

Both commercial software programmes have been under scrutiny, and reports from different studies are not always

in line with each other. For example, a validation study [84] objectively evaluated Volpara breast density assessment on FFDM using measurements obtained from breast MRI. It was concluded that accurate volumetric breast density assessment is feasible in FFDM and has potential to be used in objective breast cancer risk models and screening stratification. Similarly, Wang et al. [177] assessed agreement of three techniques (i.e., SXA, Quantra, and Volpara) with MRI for PFG, absolute FGV, and total breast volume. It was concluded that automated volumetric fibroglandular tissue measures from FFDM were in substantial agreement with MRI and if associated with breast cancer could be used in clinical practice to enhance risk assessment and prevention. However, Jeffreys et al. [178] compared the new volumetric breast density method using Volpara with the Cumulus. Results indicated that there is a very strong correlation between the Cumulus PD and VBD, but there is less relationship between the Cumulus PD and the absolute volume of dense tissue. A recent study [179] has shown disagreement in measured volumetric density when using Quantra and Volpara based technologies. To date there are no reports of breast cancer associations for Volpara or Quantra measures of volumetric breast density techniques. Therefore further validations of these new methods against breast cancer outcomes are needed. It should be noted that volumetric breast density can be measured from 2D projections, but the precise location of the dense tissue cannot be identified. Therefore, to date there are no volumetric segmentation results available. However, Volpara can produce a "density map" which allows dense mammography patterns to be visualised. Promising results [10, 180] emerged recently showing that volumetric density measures have the potential to surpass area based measures in breast cancer risk assessment.

4.4. 3D Reconstruction Based Volumetric Approaches. A limited number of studies exist regarding mammographic tissue segmentation using 3D DBT; so far there are few investigations in the literature which lead to developing "true" 3D dense tissue segmentation algorithms for estimating volumetric breast density.

Kontos et al. [181] conducted a feasibility study to distinguish dense and fatty tissue in DBT using texture features (e.g., skewness, coarseness, contrast, energy, homogeneity, and fractal dimension). The selected texture features have been used in previous mammographic tissue segmentation studies (e.g., [44]) and indicated that these features tend to correlate to breast density when computed from DBT images [182]. When computing grey-level texture statistics, the conventional 2D texture descriptors were extended to 3D by considering 3D voxel neighbourhoods rather than 2D pixel neighbourhoods. The study concluded that the fractal dimension was superior in DBT, while contrast was best in 2D projections. It was suggested that for 2D tomographic separation of the breast tissue layers in DBT, the dominant contribution to the grey-level values in the images is the X-ray attenuation at the specific voxel in the breast volume. This indicated that volumetric parenchymal properties such as self-similarity reflected by fractal dimension could be more accurately estimated by the corresponding DBT texture

features. Recent improvements on Kontos's method have led to fully automated 3D fibroglandular tissue segmentation and VBD estimation from DBT images, and the study showed strong agreement with existing volumetric techniques based on FFDM and MRI images [183]. The updated algorithm exploits the geometry of the acquisition of DBT sequences as well as the relationship between image intensity and tissue density; 3D segmentation of the fibroglandular tissue is achieved by analysing both 2D projection images and reconstructed DBT slices. Results indicated that VBD estimations were highly correlated for DBT and FFDM ($r = 0.88$), DBT and MRI ($r = 0.76$), and FFDM and MRI ($r = 0.73$). In terms of clinical relevance, such a fully automated quantitative VBD estimation from DBT could result into more accurate measures of the fibroglandular tissue in the breast which may lead to more accurate measure of breast cancer risk.

Shafer et al. [184, 185] conducted a multimodality study which utilised a 3D hidden-Markov model (HMM) based breast tissue modelling technique for DBT segmentation of adipose and glandular tissue. Much work has been done previously using 2D HMRF to estimate density and other features from mammograms [61, 186]. Glandular segmented MRI were used in training a 3D HMM; the model was validated and used to segment DBT breast volumes. The breast density for DBT images is calculated as the ratio of glandular voxels to all voxels in the breast volume. All MR and DBT images were processed to optimise the available range of values for the breast tissue segmentation task. It was also assumed that the possible tissue type of each DBT voxel is either adipose or glandular. The evaluation showed mixed results; some are encouraging and others are dissatisfactory. This indicated that the cross-modality training and testing scheme using HMM needs to be further investigated to refine the process and increase accuracy and reliability for breast DBT segmentation.

5. Discussion

This review covers mammographic tissue segmentation existing in the literature covering the period from 1992 to 2014, focusing on automatic approaches, and excluding manual and semiautomatic approaches. Only a fully automatic approach can deliver consistent, standardised, fully reproducible, and comparable measurements for high-throughput breast screening programmes across sites. Figure 7 shows the distribution of various mammographic tissue segmentation approaches, and Figure 8 shows the trend lines for area and volumetric based mammographic density/parenchymal segmentation approaches. This indicates a decline in the number of studies using 2D projection and an increase in the number of studies in 3D volumetric segmentation. Without automation and accurate estimation of breast density/parenchymal patterns, mammographic segmentation developed in a research environment is unlikely to become useful in clinical practices for assessing mammographic risk or in clinical trials evaluating the influence of different variables (drugs) on breast tissue [187]. Eng et al. [10] concluded that fully automated methods are valid alternatives to the labour intensive "gold standard" Cumulus for quantifying
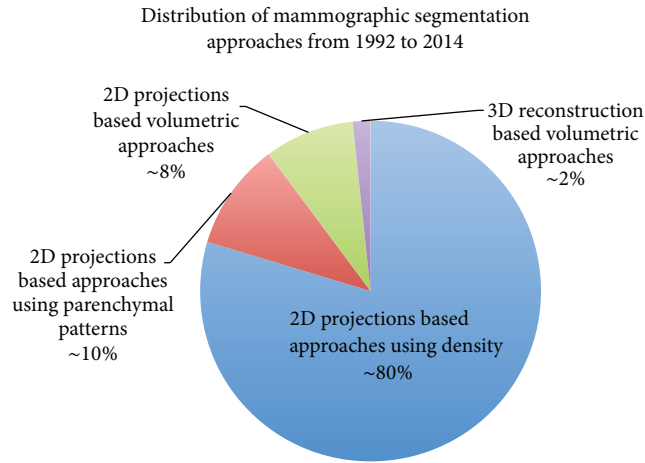
Distribution of mammographic segmentation
approaches from 1992 to 2014

FIGURE 7: Distribution of various mammographic tissue segmentation approaches from 1992 to 2014.
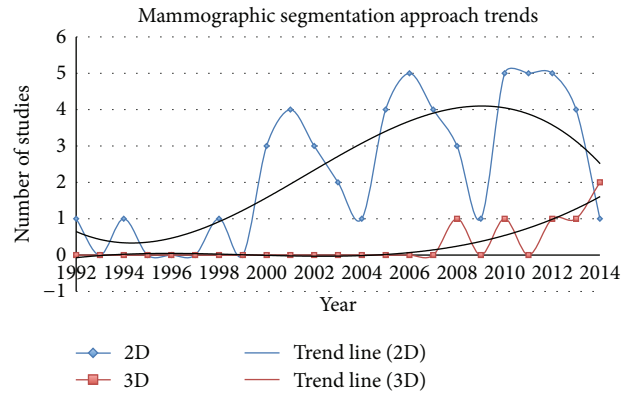
Mammographic segmentation approach trends

FIGURE 8: Polynomial based trend lines for both area and volumetric based mammographic density/parenchymal segmentation approaches from 1992 to 2014. Note that the (blue) 2D trend line shows harmonic (e.g., interval peaks) behaviour which may be because a surge of manuscripts was submitted for publication at the same years (e.g., biannual) as the major breast imaging related conferences.

density. However, the choice of a particular method will depend on the aims and settings, and it is essential that the same density assessments approach is required for a research design or survey in which the same subjects are observed repeatedly over a period of time. Despite various methodology issues related to the developed techniques as reviewed in Section 4, there are several other obstacles for most of the developed approaches to demonstrate that image-based risk prediction can be incorporated in the current non-image-based risk prediction models (e.g., the Gail and the Tyrer-Cuzick models).

5.1. Database. There are only two popular publicly available databases (i.e., MIAS and DDSM) containing digitised SFM that have been used in some studies covered in the literature. Only one database (i.e., INBreast [188]) containing FFDM recently became publicly available. Currently there is no tomosynthesis database publicly available, making it difficult to directly compare the robustness of newly developed mammographic tissue segmentation with the existing methods in the literature. Mammographic image appearance varies when using different mammography units and imaging acquisition processes; differences in imaging parameters can affect image brightness, contrast, and textural similarities between very fatty and very dense breasts. The digitisation processes employed for the MIAS, the DDSM, and other private databases are distinctively different, which can also influence mammographic image appearance. A developed mammographic tissue segmentation may work well for one database but not necessarily for other databases; a tuning process may be required to adapt the approach for a different database [189]. As a result, it is difficult to determine the robustness and reliability of a segmentation approach without quantitative and qualitative analysis using mammographic images obtained from different manufactures and perhaps digitised differently. Note that digitised SFM lacks the linearity of the characteristic curve of film, which in combination with the lack of uniformity of thickness of breast tissue during compression can influence the accuracy of the relationship,

between mammographic findings of apparently dense breast tissue and the X-ray attenuation of breast tissue [190]. Based on this review and within the scope of automatic mammographic tissue segmentation, ~73% of the studies used digitised SFM, ~25% used FFDM, and <2% used DBT. For the evaluation or validation, about ~35% of the studies used publicly available databases with the rest using private datasets, size varying from a handful of images to in excess of a thousand images.

5.2. Imaging Modality. There are doubts about the relationship between mammographically determined breast density/parenchymal patterns and breast cancer risk [155]. One of the inherent limitations for area based measurements is that they do not take thickness of dense tissue into account. It is biologically more plausible that breast cancer risk is related to the volume of dense tissue in the breast rather than to its projection. Conventional 2D mammography is affected by tissue superimposition, whilst volumetric analysis of breast tissue through emerging tomographic breast imaging modalities has been suggested as necessary to advance breast cancer risk modelling [155]. It should be noted that all the mammographic risk/density schemes as discussed in Section 2 were established based on digitised SFM. Although some image-based models like the BI-RADS scheme are reviewed in a clinical environment, there are no updates in the literature to indicate that when using FFDM or DBT, the relationship between breast density/parenchymal patterns and breast cancer risk is the same as that established for SFM. To date, all the studies presume the relationship between breast density/parenchymal patterns of different modalities and breast cancer risk is the same, with no clinical evidence or proof. Previously unknown relationships between breast density/parenchymal patterns and breast cancer risk may be discovered by using more advanced breast imaging techniques and modalities, especially when FFDM is likely to

replace standard SFM and DBT is fundamentally different to SFM and FFDM.

5.3. *Raw or Processed FFDM*. Physical image formation model based volumetric approaches use known X-ray attenuation of breast tissue to calculate tissue composition at a given pixel. Successful application of these methods to digitised SFM was initially met with difficulty; however, with the introduction of FFDM, the development of robust methods and commercial products (i.e., Quantra and Volpara) became possible. How the raw images are processed is largely unknown to research institutions as they are manufactured at the proprietors discretion. Experimental results indicated that such a processing is likely to distort the relationship between the image signal and X-ray transmission, and it could interfere with the ability to derive density information from the images [86]. Visually speaking, radiologists have observed that the processed mammograms appear to be less dense than when imaged with film mammography (e.g., see Figures 5 and 6). Therefore, using thresholding or physical image formation model based approaches to measure breast density on a processed FFDM is likely to be problematic, especially if such measurements are to be compared to those obtained from SFM [86].

5.4. *Study Population Matters*. During evaluations, many studies only used a small amount of data from the general populations of western developed countries, and few studies subgrouped data based on different criteria. Although it is well established that density measurements are associated with breast cancer risk, still several aspects are questioned [87]. For example, hormone therapy is known to increase breast density; certain regimes are known to relate to breast cancer risk [191], but the degree to which density and risk are related during hormone therapy is not known [192, 193]. Several studies [96, 97, 194, 195] on the assessment of breast cancer risk in postmenopausal women showed no clear correlation between the disease and tissue density, and the magnitude of the association varied. Therefore studies involving all ranges and subgroups of the populations based on age, body mass, ethnic background, and so forth could be beneficial in providing additional evidence and linking breast density/parenchymal patterns with other breast cancer risk factors established in non-image-based risk prediction models.

5.5. *Segmentation Schemes*. The vast majority of mammographic tissue segmentation approaches are focused on the well-established density as biomarker. However, some of the studies use BI-RADS density categories and others may use the SCC or Wolfe scheme. The relation between breast cancer risk and density is at best hypothesised [87], and the optimal approach to derive information from mammograms relating to risk of breast cancer is still argued [196, 197]. Various classification/categorisation mappings between different schemes have been studied [24, 90, 94], which is beneficial for experimental comparison when using different image-based breast cancer risk prediction models. Recent studies have indicated that not just density but also heterogeneity in mammograms can be associated with risk [196, 198, 199]. This

suggested that mammographic tissue segmentation based on Tabár's parenchymal patterns (i.e., mammographic building blocks) could be more useful than density based approaches.

5.6. *Evaluation*. Clinically obtained "ground truth" is often used in algorithm evaluation and validation. Many studies have shown issues of large interobserver variations in breast tissue annotated by radiologists with different mammogram reading experiences. This makes area overlapping based evaluation unreliable. Such variations also exist for mammographic risk assessment by radiologists, as manual mammographic risk assessment is intrinsically subjective, and there is substantial inter- and intraobserver variability [119, 200]. This subjectiveness includes visual assessment. Therefore, inconsistent results could be obtained if a different group of radiologists is used to obtain "ground truth" for evaluation or validation, even when the methodology remains the same. There is no easy solution to overcome this, but using consensus data is probably by far the best way to deal with this issue.

6. Conclusions

A comprehensive review on automatic breast density and parenchymal segmentation is presented, suggesting that there are many issues related to the developed techniques, evaluations, and practicalities that hinder the progress of bridging translational research and clinical utilisation. Various fully automatic mammographic tissue segmentation approaches have been developed or are currently under investigation to overcome subjective estimation of tissue composition, reduce inter- and intraobserver variability, and eliminate ambiguous outcomes. The likely transition to use 3D technology with tomosynthesis will provide a platform for work on breast tissue segmentation with 3D rendering to facilitate new areas of work such as 3D image guided biopsy or image guided surgical excision. Mammographic tissue segmentation has a strong impact on breast cancer risk estimation; however, it has limited influence in clinical decision making and has not yet been included in any established risk prediction model. It may be ideal to incorporate image-based risk factors in non-image-based risk prediction models, so that automatically measured breast tissue composition can be used for risk stratification to adapt screening interval to risk. Further investigations are needed in order to develop a robust, reliable, and fully automatic mammographic segmentation which can help to provide prognostic value in distinguishing which women with certain mammographic appearance are likely to develop breast cancer. Only precise identification of women at risk can lead to a successful prevention and/or treatment and a reduction in the incidence of and mortality from breast cancer.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] J. Ferlay, H.-R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin, "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008," *International Journal of Cancer*, vol. 127, no. 12, pp. 2893–2917, 2010.

[2] S. A. Eccles, E. Aboagye, S. Ali et al., "Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer," *Breast Cancer Research*, vol. 15, article R92, 2013.

[3] H. Darabi, K. Czene, W. Zhao, J. Liu, P. Hall, and K. Humphreys, "Breast cancer risk prediction and individualised screening based on common genetic variation and breast density measurement," *Breast Cancer Research*, vol. 14, no. 1, article R25, 2012.

[4] N. F. Boyd, L. J. Martin, M. J. Yaffe, and S. Minkin, "Mammographic density and breast cancer risk: current understanding and future prospects," *Breast Cancer Research*, vol. 13, no. 6, article 223, 2011.

[5] American College of Radiology, *Breast Imaging Reporting and Data System (BI-RADS)*, American College of Radiology, Reston, Va, USA, 2003.

[6] A. Karellas, "Mammographic pattern analysis: an emerging risk assessment tool," *Academic Radiology*, vol. 14, no. 5, pp. 511–512, 2007.

[7] J. A. Harvey, G. J. Stukenborg, W. F. Cohn et al., "Volumetric breast density improves breast cancer risk prediction," *Oncology Times*, vol. 37, no. 6, p. 33, 2015, Proceedings of the San Antonio Breast Cancer Symposium 2014.

[8] M. H. Gail, L. A. Brinton, D. P. Byar et al., "Projecting individualized probabilities of developing breast cancer for white females who are being examined annually," *Journal of the National Cancer Institute*, vol. 81, no. 24, pp. 1879–1886, 1989.

[9] J. Tyrer, S. W. Duffy, and J. Cuzick, "A breast cancer prediction model incorporating familial and personal risk factors," *Statistics in Medicine*, vol. 23, no. 7, pp. 1111–1130, 2004.

[10] A. Eng, Z. Gallant, J. Shepherd et al., "Volumetric breast density improves breast cancer risk prediction," *Breast Cancer Research*, vol. 16, no. 5, article 439, 2014.

[11] D. G. R. Evans, J. Warwick, S. M. Astley et al., "Assessing individual breast cancer risk within the U.K. National Health Service Breast Screening Program: a new paradigm for cancer prevention," *Cancer Prevention Research*, vol. 5, no. 7, pp. 943–951, 2012.

[12] C. H. van Gils, J. D. M. Otten, A. L. M. Verbeek, and J. H. C. L. Hendriks, "Mammographic breast density and risk of breast cancer: masking bias or causality?" *European Journal of Epidemiology*, vol. 14, no. 4, pp. 315–320, 1998.

[13] J. N. Wolfe, "Breast patterns as an index of risk for developing breast cancer," *American Journal of Roentgenology*, vol. 126, no. 6, pp. 1130–1139, 1976.

[14] E. A. Sickles, "Wolfe mammographic parenchymal patterns and breast cancer risk," *American Journal of Roentgenology*, vol. 188, no. 2, pp. 301–303, 2007.

[15] L. Tabár and P. B. Dean, "Mammographic parenchymal patterns: risk indicator for breast cancer?" *The Journal of the American Medical Association*, vol. 247, no. 2, pp. 185–189, 1982.

[16] L. Tabár, T. Tot, and P. B. Dean, *The Art and Science of Early Detection with Mamography: Perception, Interpretation, Histopatholigic Correlation*, Georg Thieme, 1st edition, 2004.

[17] N. F. Boyd, J. W. Byng, R. A. Jong et al., "Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study," *Journal of the National Cancer Institute*, vol. 87, no. 9, pp. 670–675, 1995.

[18] N. F. Boyd, L. J. Martin, L. Sun et al., "Body size, mammographic density, and breast cancer risk," *Cancer Epidemiology Biomarkers & Prevention*, vol. 15, no. 11, pp. 2086–2092, 2006.

[19] E. S. Burnside, E. A. Sickles, L. W. Bassett et al., "The ACR BI-RADS c experience: learning from history," *Journal of the American College of Radiology*, vol. 6, no. 12, pp. 851–860, 2009.

[20] E. Sala, R. Warren, J. McCann, S. Duffy, N. Day, and R. Luben, "Mammographic parenchymal patterns and mode of detection: implications for the breast screening programme," *Journal of Medical Screening*, vol. 5, no. 4, pp. 207–212, 1998.

[21] J. C. Bailar III, "Mammography: a contrary view," *Annals of Internal Medicine*, vol. 84, no. 1, pp. 77–84, 1976.

[22] I. T. Gram, E. Funkhouser, and L. Tabár, "The Tabár classification of mammographic parenchymal patterns," *European Journal of Radiology*, vol. 24, no. 2, pp. 131–136, 1997.

[23] I. T. Gram, Y. Bremnes, G. Ursin, G. Maskarinec, N. Bjurstam, and E. Lund, "Percentage density, Wolfe's and Tabár's mammographic patterns: agreement and association with risk factors for breast cancer," *Breast Cancer Research*, vol. 7, no. 5, pp. R854–R861, 2005.

[24] I. Muhimmah, A. Oliver, E. R. E. Denton, J. Pont, E. Pérez, and R. Zwiggelaar, "Comparison between wolfe, boyd, BI-RADS and Tabár based mammographic risk assessment," in *Digital Mammography*, vol. 4046 of *Lecture Notes in Computer Science*, pp. 407–415, Springer, 2006.

[25] T. Matsubara, D. Yamazaki, H. Fujita, T. Hara, T. Iwase, and T. Endo, "An automated classification method for mammograms based on evaluation of fibroglandular breast tissue density," in *Proceedings of the International Workshop on Digital Mammography*, pp. 737–741, 2000.

[26] P. K. Saha, J. K. Udupa, E. F. Conant, D. P. Chakraborty, and D. Sullivan, "Breast tissue density quantification via digitized mammograms," *IEEE Transactions on Medical Imaging*, vol. 20, no. 8, pp. 792–803, 2001.

[27] R. Sivaramakrishna, N. A. Obuchowski, W. A. Chilcote, and K. A. Powell, "Automatic segmentation of mammographic density," *Academic Radiology*, vol. 8, no. 3, pp. 250–256, 2001.

[28] C. Olsén and A. Mukhdoomi, "Automatic segmentation of fibroglandular tissue," in *Image Analysis*, vol. 4522 of *Lecture Notes in Computer Science*, pp. 679–688, Springer, Berlin, Germany, 2007.

[29] S. D. Tzikopoulos, M. E. Mavroforakis, H. V. Georgiou, N. Dimitropoulos, and S. Theodoridis, "A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry," *Computer Methods and Programs in Biomedicine*, vol. 102, no. 1, pp. 47–63, 2011.

[30] L.-J. W. Lu, T. K. Nishino, T. Khamapirad, J. J. Grady, M. H. Leonard Jr., and D. G. Brunder, "Computing mammographic density from a multiple regression model constructed with image-acquisition parameters from a full-field digital mammographic unit," *Physics in Medicine and Biology*, vol. 52, no. 16, pp. 4905–4921, 2007.

[31] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, and A. F. Frère, "Segmentation of mammograms: identification of the

skin-air boundary, pectoral muscle, and fibroglandular disc," in *Proceedings of the 5th International Workshop on Digital Mammography (IWDM '00)*, pp. 573–579, Toronto, Canada, June 2000.

[32] R. J. Ferrari, R. M. Rangayyan, R. A. Borges, and A. F. Frère, "Segmentation of the fibro-glandular disc in mammograms using Gaussian mixture modelling," *Medical and Biological Engineering and Computing*, vol. 42, no. 3, pp. 378–387, 2004.

[33] A. El-Zaart, "Expectation-maximization technique for fibro-glandular discs detection in mammography images," *Computers in Biology and Medicine*, vol. 40, no. 4, pp. 392–401, 2010.

[34] C. Zhou, H. P. Chan, N. Petrick et al., "Computerized image analysis: estimation of breast density on mammograms," *Medical Physics*, vol. 28, no. 6, pp. 1056–1069, 2001.

[35] J. T. Neyhart, M. Kirlakovsky, L. M. Coleman, R. Polikar, M. Tseng, and S. A. Mandayam, "Automated segmentation and quantitative characterization of radiodense tissue in digitized mammograms," *AIP Conference Proceedings*, vol. 615, no. 1, pp. 1866–1873, 2002.

[36] Y. Kim, C. Kim, and J. H. Kim, "Automated estimation of breast density on mammogram using combined information of histogram statistics and boundary gradients," in *Medical Imaging 2010: Computer-Aided Diagnosis*, vol. 7624 of *Proceedings of SPIE*, p. 76242F, San Diego, Calif, USA, March 2010.

[37] C. Nickson, Y. Arzhaeva, Z. Aitken et al., "AutoDensity: an automated method to measure mammographic breast density that predicts breast cancer risk and screening outcomes," *Breast Cancer Research*, vol. 15, no. 5, article R80, 2013.

[38] A. Oliver, J. Freixenet, A. Bosch, D. Raba, and R. Zwiggelaar, "Automatic classification of breast tissue," in *Proceedings of the 2nd Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA '05)*, vol. 3523 of *Lecture Notes in Computer Science*, pp. 431–438, June 2005.

[39] H. Strange, E. R. E. Denton, M. Kibiro, and R. Zwiggelaar, "Manifold learning for density segmentation in high risk mammograms," in *Pattern Recognition and Image Analysis*, vol. 7887 of *Lecture Notes in Computer Science*, pp. 245–252, Springer, Berlin, Germany, 2013.

[40] K. Marias, M. G. Linguraru, M. A. G. Ballester, S. Petroudi, M. Tsiknakis, and S. M. Brady, "Automatic labelling and BI-RADS characterisation of mammogram densities," in *Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS '05)*, pp. 6394–6398, IEEE, September 2005.

[41] A. Oliver, J. Freixenet, and R. Zwiggelaar, "Automatic classification of breast density," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '05)*, vol. 2, pp. 1258–1261, September 2005.

[42] A. Oliver, J. Martí, J. Freixenet, J. Pont, and R. Zwiggelaar, "Automatic classification of breast density according BIRADS categories using a clustering approach," in *Proceedings of the Computed Aided Radiology and Surgery (CARS '05)*, Berlin, Germany, June 2005.

[43] A. Oliver, X. Lladó, R. Martı, J. Freixenet, and R. Zwiggelaar, "Classifying mammograms using texture information," in *Proceedings of the Medical Image Understanding and Analysis Conference*, pp. 223–227, July 2007.

[44] A. Oliver, J. Freixenet, R. Martí et al., "A novel breast tissue density classification methodology," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 55–65, 2008.

[45] A. Oliver, J. Freixenet, R. Martí, and R. Zwiggelaar, "A comparison of breast tissue classification techniques," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006*, vol. 4191 of *Lecture Notes in Computer Science*, pp. 872–879, Springer, Berlin, Germany, 2006.

[46] A. Torrent, A. Bardera, A. Oliver et al., "Breast density segmentation: a comparison of clustering and region based techniques," in *Digital Mammography*, vol. 5116 of *Lecture Notes in Computer Science*, pp. 9–16, Springer, Berlin, Germany, 2008.

[47] M. Tortajada, A. Oliver, R. Martí et al., "Adapting breast density classification from digitized to full-field digital mammograms," in *Breast Imaging*, vol. 7361 of *Lecture Notes in Computer Science*, pp. 561–568, Springer, Berlin, Germany, 2012.

[48] Z. Chen and R. Zwiggelaar, "A modified fuzzy c-means algorithm for breast tissue density segmentation in mammograms," in *Proceedings of the 10th International Conference on Information Technology and Applications in Biomedicine (ITAB '10)*, pp. 1–4, IEEE, Corfu, Greece, November 2010.

[49] B. Keller, D. Nathan, Y. Wang et al., "Adaptive multi-cluster fuzzy c-means segmentation of breast parenchymal tissue in digital mammography," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*, vol. 6893 of *Lecture Notes in Computer Science*, pp. 562–569, Springer, 2011.

[50] B. M. Keller, D. L. Nathan, Y. Wang et al., "Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation," *Medical Physics*, vol. 39, no. 8, pp. 4903–4917, 2012.

[51] S. R. Aylward, B. M. Hemminger, and E. D. Pisano, "Mixture modeling for digital mammogram display and analysis," in *Proceedings of the International Workshop on Digital Mammography*, pp. 305–312, 1998.

[52] R. Zwiggelaar, P. Planiol, J. Martí et al., "EM texture segmentation of mammographic images," in *Proceedings of the International Workshop on Digital Mammography*, pp. 223–227, Bremen, Germany, June 2002.

[53] R. Zwiggelaar, L. Blot, D. Raba, and E. R. E. Denton, "Set-permutation-occurrence matrix based texture segmentation," in *Pattern Recognition and Image Analysis*, vol. 2652 of *Lecture Notes in Computer Science*, pp. 1099–1107, Springer, Berlin, Germany, 2003.

[54] R. Zwiggelaar, L. Blot, D. Raba, and D. R. E. Erika, "Texture segmentation in mammograms," in *Medical Image Understanding and Analysis*, 2003.

[55] R. Zwiggelaar and E. R. E. Denton, "Optimal segmentation of mammographic images," in *Proceedings of the International Workshop on Digital Mammography*, pp. 751–757, 2004.

[56] S. E. Selvan, C. C. Xavier, N. Karssemeijer, J. Sequeira, R. A. Cherian, and B. Y. Dhala, "Parameter estimation in stochastic mammogram model by heuristic optimization techniques," *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 4, pp. 685–695, 2006.

[57] P. Miller and S. Astley, "Classification of breast tissue by texture analysis," *Image and Vision Computing*, vol. 10, no. 5, pp. 277–282, 1992.

[58] J. Suckling, D. R. Dance, E. Moskovic, D. J. Lewis, and S. G. Blacker, "Segmentation of mammograms using multiple linked self-organizing neural networks," *Medical Physics*, vol. 22, no. 2, pp. 145–152, 1995.

[59] J. J. Heine and R. P. Velthuizen, "A statistical methodology for mammographic density detection," *Medical Physics*, vol. 27, no. 12, pp. 2644–2651, 2000.

[60] J. J. Heine, M. J. Carston, C. G. Scott et al., "An automated approach for estimation of breast density," *Cancer Epidemiology Biomarkers & Prevention*, vol. 17, no. 11, pp. 3090–3097, 2008.

[61] S. Petroudi and M. Brady, "Breast density segmentation using texture," in *Digital Mammography*, vol. 4046 of *Lecture Notes in Computer Science*, pp. 609–615, Springer, Berlin, Germany, 2006.

[62] Y. C. Gong, M. Brady, and S. Petroudi, "Texture based mammogram classification and segmentation," in *Digital Mammography*, vol. 4046 of *Lecture Notes in Computer Science*, pp. 616–625, Springer, 2006.

[63] A. Oliver, X. Lladó, E. Pérez et al., "A statistical approach for breast density segmentation," *Journal of Digital Imaging*, vol. 23, no. 5, pp. 527–537, 2010.

[64] R. Zwiggelaar and E. R. E. Denton, "Texture based segmentation," in *Digital Mammography*, vol. 4046 of *Lecture Notes in Computer Science*, pp. 433–440, Springer, Berlin, Germany, 2006.

[65] M. Adel, M. Rasigni, S. Bourennane, and V. Juhan, "Statistical segmentation of regions of interest on a mammographic image," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 49482, 8 pages, 2007.

[66] R. Zwiggelaar, "Local greylevel appearance histogram based texture segmentation," in *Digital Mammography*, vol. 6136 of *Lecture Notes in Computer Science*, pp. 175–182, Springer, Berlin, Germany, 2010.

[67] M. G. J. Kallenberg, M. Lokate, C. H. van Gils, and N. Karssemeijer, "Automatic breast density segmentation: an integration of different approaches," *Physics in Medicine and Biology*, vol. 56, no. 9, pp. 2715–2729, 2011.

[68] J. Li, L. Szekely, L. Eriksson et al., "High-throughput mammographic-density measurement: a tool for risk prediction of breast cancer," *Breast Cancer Research*, vol. 14, no. 4, article R114, 2012.

[69] Z. Chen, E. R. E. Denton, and R. Zwiggelaar, "Topographic representation based breast density segmentation for mammographic risk assessment," in *Proceedings of the 19th IEEE International Conference on Image Processing (ICIP '12)*, pp. 1993–1996, October 2012.

[70] Z. Chen, A. Oliver, E. R. E. Denton, and R. Zwiggelaar, "Automated mammographic risk classification based on breast density estimation," in *Pattern Recognition and Image Analysis*, vol. 7887 of *Lecture Notes in Computer Science*, pp. 237–244, Springer, Berlin, Germany, 2013.

[71] I. Muhimmah, W. He, E. R. E. Denton, and R. Zwiggelaar, "Segmentation based on textons and mammographic building blocks," in *Medical Image Understanding and Analysis*, pp. 228–232, 2007.

[72] W. He, I. Muhimmah, E. R. E. Denton, and R. Zwiggelaar, "Mammographic risk assessment based on Tabár mammographic building blocks," in *Pattern Recognition and Image Analysis: New Information Technologies*, pp. 219–222, 2008.

[73] W. He, E. R. E. Denton, and R. Zwiggelaar, "Mammographic segmentation based on mammographic parenchymal patterns and spatial moments," in *Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine*, pp. 1–4, November 2009.

[74] W. He, E. R. E. Denton, K. Stafford, and R. Zwiggelaar, "Mammographic image segmentation and risk classification based on mammographic parenchymal patterns and geometric moments," *Biomedical Signal Processing and Control*, vol. 6, no. 3, pp. 321–329, 2011.

[75] W. He, E. R. E. Denton, and R. Zwiggelaar, "Mammographic segmentation and risk classification using a novel binary model based bayes classifier," in *Breast Imaging*, vol. 7361 of *Lecture Notes in Computer Science*, pp. 40–47, Springer, Berlin, Germany, 2012.

[76] W. He, E. R. E. Denton, and R. Zwiggelaar, "A novel breast image preprocessing for full field digital mammographic segmentation and risk classification," in *Medical Image Understanding and Analysis*, pp. 40–47, 2014.

[77] W. He, E. R. E. Denton, and R. Zwiggelaar, "Mammographic image segmentation and risk classification using a novel texture signature based methodology," in *Digital Mammography*, vol. 6136 of *Lecture Notes in Computer Science*, pp. 526–533, Springer, Berlin, Germany, 2010.

[78] W. He and R. Zwiggelaar, "Breast parenchymal pattern analysis in digital mammography: associations between tabár and birads tissue compositions," in *Computer Analysis of Images and Patterns*, vol. 8048 of *Lecture Notes in Computer Science*, pp. 386–393, Springer, 2013.

[79] J. J. Heine, K. Cao, D. E. Rollison, G. Tiffenberg, and J. A. Thomas, "A quantitative description of the percentage of breast density measurement using full-field digital mammography," *Academic Radiology*, vol. 18, no. 5, pp. 556–564, 2011.

[80] O. Alonzo-Proulx, R. A. Jong, and M. J. Yaffe, "Volumetric breast density characteristics as determined from digital mammograms," *Physics in Medicine and Biology*, vol. 57, no. 22, pp. 7443–7457, 2012.

[81] S. Ourselin, M. A. Styner, S. Malkov, K. Kerlikowske, and J. Shepherd, "Automated volumetric breast density derived by shape and appearance modeling," in *Medical Imaging*, vol. 9034 of *Proceedings of SPIE*, pp. 90342T–90342T–7, San Diego, California, USA, March 2014.

[82] K. Hartman, R. Highnam, R. Warren, and V. Jackson, "Volumetric assessment of breast tissue composition from FFDM images," in *Digital Mammography*, vol. 5116 of *Lecture Notes in Computer Science*, pp. 33–39, Springer, Berlin, Germany, 2008.

[83] R. Highnam, M. Brady, M. J. Yaffe, N. Karssemeijer, and J. Harvey, "Robust breast composition measurement—Volpara," in *Digital Mammography*, vol. 6136 of *Lecture Notes in Computer Science*, pp. 342–349, Springer, Berlin, Germany, 2010.

[84] A. Gubern-Mérida, M. Kallenberg, B. Platel et al., "Volumetric breast density estimation from full-field digital mammograms: a validation study," *PLoS ONE*, vol. 9, no. 1, Article ID e85952, 2014.

[85] J. N. Wolfe, A. F. Saftlas, and M. Salane, "Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study," *The American Journal of Roentgenology*, vol. 148, no. 6, pp. 1087–1092, 1987.

[86] M. J. Yaffe, "Measurement of mammographic density," *Breast Cancer Research*, vol. 10, no. 3, article 209, 2008.

[87] M. Nielsen, G. Karemore, M. Loog et al., "A novel and automatic mammographic texture resemblance marker is an independent risk factor for breast cancer," *Cancer Epidemiology*, vol. 35, no. 4, pp. 381–387, 2011.

[88] C. Byrne, C. Schairer, J. N. Wolfe et al., "Mammographic features and breast cancer risk: effects with time, age, and menopause status," *Journal of the National Cancer Institute*, vol. 87, no. 21, pp. 1622–1629, 1995.

[89] A. F. Saftlas, R. N. Hoover, L. A. Brinton et al., "Mammographic densities and risk of breast cancer," *Cancer*, vol. 67, no. 11, pp. 2833–2838, 1991.

[90] K. Taylor, P. Britton, S. O'Keeffe, and M. G. Wallis, "Quantification of the UK 5-point breast imaging classification and mapping to BI-RADS to facilitate comparison with international literature," *The British Journal of Radiology*, vol. 84, no. 1007, pp. 1005–1010, 2011.

[91] J. Brisson, C. Diorio, and B. Mâsse, "Wolfe's parenchymal pattern and percentage of the breast with mammographic densities: redundant or complementary classifications?" *Cancer Epidemiology Biomarkers and Prevention*, vol. 12, no. 8, pp. 728–732, 2003.

[92] American College of Radiology, *Breast Imaging Reporting and Data System (BI-RADS)*, American College of Radiology, Reston, Va, USA, 2014.

[93] J. Suckling, J. Parker, D. Dance et al., "The mammographic images analysis society digital mammogram database," in *Proceedings of the 2nd International Workshop on Digital Mammography*, pp. 375–378, 1994.

[94] W. He, M. Kibiro, A. Juette, E. R. E. Denton, and R. Zwiggelaar, "A revisit on correlation between Tabár and Birads based risk assessment schemes with full field digital mammography," in *Breast Imaging—12th International Workshop (IWDM '14), Gifu City, Japan, June-July 2014. Proceedings*, vol. 8539 of *Lecture Notes in Computer Science*, pp. 327–333, Springer, 2014.

[95] A. M. Oza and N. F. Boyd, "Mammographic parenchymal patterns: a marker of breast cancer risk," *Epidemiologic Reviews*, vol. 15, no. 1, pp. 196–208, 1993.

[96] N. F. Boyd, G. A. Lockwood, J. W. Byng, D. L. Tritchler, and M. J. Yaffe, "Mammographic densities and breast cancer risk," *Cancer Epidemiology Biomarkers and Prevention*, vol. 7, no. 12, pp. 1133–1144, 1998.

[97] P. M. Vacek and B. M. Geller, "A prospective study of breast cancer risk using routine mammographic breast density measurements," *Cancer Epidemiology Biomarkers & Prevention*, vol. 13, no. 5, pp. 715–722, 2004.

[98] S. V. Sree, E. Y. Ng, R. U. Acharya, and O. Faust, "Breast imaging: a survey," *World Journal of Clinical Oncology*, vol. 2, no. 4, pp. 171–178, 2011.

[99] E. L. Rosen, W. B. Eubank, and D. A. Mankoff, "FDG PET, PET/CT, and breast cancer imaging," *Radiographics*, vol. 27, supplement 1, pp. S215–S229, 2007.

[100] S. Malur, S. Wurdinger, A. Moritz, W. Michels, and A. Schneider, "Comparison of written reports of mammography, sonography and magnetic resonance mammography for preoperative evaluation of breast lesions, with special emphasis on magnetic resonance mammography," *Breast Cancer Research*, vol. 3, no. 1, pp. 55–60, 2001.

[101] V. V. Levenson, "Biomarkers for early detection of breast cancer: what, when, and where?" *Biochimica et Biophysica Acta—General Subjects*, vol. 1770, no. 6, pp. 847–856, 2007.

[102] W. A. Berg, L. Gutierrez, M. S. NessAiver et al., "Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer," *Radiology*, vol. 233, no. 3, pp. 830–849, 2004.

[103] T. M. Kolb, J. Lichy, and J. H. Newhouse, "Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations," *Radiology*, vol. 225, no. 1, pp. 165–175, 2002.

[104] H. Thornton, A. Edwards, and M. Baum, "Women need better information about routine mammography," *British Medical Journal*, vol. 327, no. 7406, pp. 101–103, 2003.

[105] K. Kerlikowske, D. Grady, J. Barclay et al., "Variability and accuracy in mammographic interpretation using the American college of radiology breast imaging reporting and data system," *Journal of the National Cancer Institute*, vol. 90, no. 23, pp. 1801–1809, 1998.

[106] M. B. I. Lobbes, J. P. M. Cleutjens, V. Lima Passos et al., "Density is in the eye of the beholder: visual versus semi-automated assessment of breast density on standard mammograms," *Insights into Imaging*, vol. 3, no. 1, pp. 91–99, 2012.

[107] E. D. Pisano, R. E. Hendrick, M. J. Yaffe 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.

[108] A. Fischmann, K. C. Siegmann, A. Wersebe, C. D. Claussen, and M. Müller-Schimpfle, "Comparison of full-field digital mammography and film-screen mammography: image quality and lesion detection," *British Journal of Radiology*, vol. 78, no. 928, pp. 312–315, 2005.

[109] F. J. Gillbert, K. C. Young, S. M. Astley, P. Whelehan, and M. G. C. Gillan, *Digital Breast Tomosynthesis*, NHSBSP Publication, 2010.

[110] T. Wu, A. Stewart, M. Stanton et al., "Tomographic mammography using a limited number of lowdose cone-beam projection images," *Medical Physics*, vol. 30, pp. 365–380, 2003.

[111] E. H. L. Mungutroy, J. M. Oduko, J. C. Cooke, and W. J. Formstone, *Practical Evaluation of Hologic Selenia Dimensions Digital Breast Tomosynthesis System*, NHSBSP Publication, 2014.

[112] J. Marshall, A. Kshirsagar, S. Narin, and N. Gkanatsios, "Will new technologies replace mammography CAD as we know it?" in *Breast Imaging*, vol. 8539 of *Lecture Notes in Computer Science*, pp. 30–37, Springer, Cham, Switzerland, 2014.

[113] P. K. Saha, J. K. Udupa, and D. Odhner, "Scale-based fuzzy connected image segmentation: theory, algorithms, and validation," *Computer Vision and Image Understanding*, vol. 77, no. 2, pp. 145–174, 2000.

[114] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognition*, vol. 19, no. 1, pp. 41–47, 1986.

[115] A. D. Brink and N. E. Pendock, "Minimum cross-entropy threshold selection," *Pattern Recognition*, vol. 29, no. 1, pp. 179–188, 1996.

[116] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[117] A. K. C. Wong and P. K. Sahoo, "Gray-level threshold selection method based on maximum entropy principle," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 19, no. 4, pp. 866–871, 1989.

[118] J. N. Kapur, "Twenty-five years of maximum-entropy principle," *Journal of Mathematical and Physical Sciences*, vol. 17, no. 2, pp. 103–156, 1983.

[119] K. E. Martin, M. A. Helvie, C. Zhou et al., "Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and BI-RADS categories," *Radiology*, vol. 240, no. 3, pp. 656–665, 2006.

[120] J. MacQueen, "Some methods of classification and analysis of multivariate observations," in *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, 1967.

[121] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.

[122] K. Zhang and J. T. Kwok, "Clustered Nystrm method for large scale manifold learning and dimension reduction," *IEEE Transactions on Neural Networks*, vol. 21, no. 10, pp. 1576–1587, 2010.

[123] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY, USA, 2013.

[124] M. Heath, K. Bowyer, D. Kopans et al., "Current status of the digital database for screening mammography," in *Proceedings of the 4th International Workshop on Digital Mammography*, pp. 457–460, Nijmegen, The Netherlands, June 1998.

[125] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[126] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[127] D. Raba, J. Martí, R. Martí, and M. Peracaula, "Breast mammography asymmetry estimation based on fractal and texture analysis," in *Computed Aided Radiology and Surgery (CARS '05)*, 2005.

[128] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[129] D. P. Feldman and J. P. Crutchfield, "Structural information in two-dimensional patterns: entropy convergence and excess entropy," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 67, no. 5, Article ID 051104, 2003.

[130] E. Riboli, K. J. Hunt, N. Slimani et al., "European prospective investigation into cancer and nutrition (EPIC): study populations and data collection," *Public Health Nutrition*, vol. 5, no. 6, pp. 1113–1124, 2002.

[131] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B: Methodological*, vol. 39, no. 1, pp. 1–38, 1977.

[132] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[133] Y. Rubner, C. Tomasi, and L. J. Guibas, "Metric for distributions with applications to image databases," in *Proceedings of the IEEE 6th International Conference on Computer Vision*, pp. 59–66, January 1998.

[134] R. C. Eberhart and Y. Shi, "Comparison between genetic algorithms and particle swarm optimization," in *Evolutionary Programming VII*, vol. 1447 of *Lecture Notes in Computer Science*, pp. 611–616, Springer, Berlin, Germany, 1998.

[135] X. Yao, Y. Liu, and G. Lin, "Evolutionary programming made faster," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 2, pp. 82–102, 1999.

[136] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, New York, NY, USA, 1984.

[137] K. I. Laws, "Textured image segmentation," Tech. Rep. 940, Image Processing Institute, University of Southern California, Los Angeles, Calif, USA, 1980.

[138] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, The Morgan Kaufmann Series in Representation and Reasoning, Morgan Kaufmann, San Mateo, Calif, USA, 1988.

[139] D. J. Lewis, D. G. Corr, C. R. Gent, and C. P. Sheppard, "The use of multiple linked self-organizing artificial neural networks for classification of remotely sensed images," in *Proceedings of the Satellite Symposia 1 & 2: Navigation and Mobile Communication and Image Processing ESA-IY2*, pp. 333–345, 1992.

[140] S. Petroudi, T. Kadir, and M. Brady, "Automatic classification of mammographic parenchymal patterns: a statistical approach," *Engineering in Medicine and Biology Society*, vol. 1, pp. 798–801, 2003.

[141] P. Brault and A. Mohammad-Djafari, "Bayesian segmentation and motion estimation in video sequences using a Markov-Potts model," in *Proceedings of the 5th WSEAS International Conference on Applied Mathematics (Math '04)*, 2004.

[142] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, 2001.

[143] M. Varma and A. Zisserman, "Classifying images of materials: achieving viewpoint and illumination independence," in *Computer Vision—ECCV 2002*, vol. 2352 of *Lecture Notes in Computer Science*, pp. 255–271, Springer, Berlin, Germany, 2002.

[144] M. A. Turk and A. P. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[145] M. Pietikainen, "Image analysis with local binary patterns," in *Image Analysis*, vol. 3540 of *Lecture Notes in Computer Science*, pp. 115–118, Springer, Berlin, Germany, 2005.

[146] S. M. Smith and J. M. Brady, "SUSAN—a new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no. 1, pp. 45–78, 1997.

[147] R. Hupse and N. Karssemeijer, "The effect of feature selection methods on computer-aided detection of masses in mammograms," *Physics in Medicine and Biology*, vol. 55, no. 10, pp. 2893–2904, 2010.

[148] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, no. 7, pp. 671–675, 2012.

[149] Z. Lao and Z. Huo, "Quantitative assessment of breast dense tissue on mammograms," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '09)*, pp. 2605–2608, November 2009.

[150] W. He, I. Muhimmah, E. R. E. Denton, and R. Zwiggelaar, "Mammographic segmentation based on texture modelling of Tabár mammographic building blocks," in *Digital Mammography*, vol. 5116 of *Lecture Notes in Computer Science*, pp. 17–24, Springer, Berlin, Germany, 2008.

[151] M. Tuceryan, "Moment-based texture segmentation," *Pattern Recognition Letters*, vol. 15, no. 7, pp. 659–668, 1994.

[152] M. Tortajada, A. Oliver, R. Martí et al., "Breast peripheral area correction in digital mammograms," *Computers in Biology and Medicine*, vol. 50, pp. 32–40, 2014.

[153] W. He and R. Zwiggelaar, "Image classification: a novel texture signature approach," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 2725–2728, IEEE, Hong Kong, China, September 2010.

[154] R. P. Highnam, M. Jeffreys, V. McCormack, R. Warren, G. D. Smith, and M. Brady, "Comparing measurements of breast density," *Physics in Medicine and Biology*, vol. 52, no. 19, pp. 5881–5895, 2007.

[155] D. B. Kopans, "Basic physics and doubts about relationship between mammographically determined tissue density and breast cancer risk," *Radiology*, vol. 246, no. 2, pp. 348–353, 2008.

[156] C. M. Vachon, C. H. van Gils, T. A. Sellers et al., "Mammographic density, breast cancer risk and risk prediction," *Breast Cancer Research*, vol. 9, no. 6, article 217, 2007.

[157] R. Highnam and M. Brady, *Mammographic Image Analysis*, Kluwer Academic, New York, NY, USA, 1999.

[158] J. Kaufhold, J. A. Thomas, J. W. Eberhard, C. E. Galbo, and D. E. González Trotter, "A calibration approach to glandular tissue composition estimation in digital mammography," *Medical Physics*, vol. 29, no. 8, pp. 1867–1880, 2002.

[159] O. Pawluczyk, B. J. Augustine, M. J. Yaffe et al., "A volumetric method for estimation of breast density on digitized screen-film mammograms," *Medical Physics*, vol. 30, no. 3, pp. 352–364, 2003.

[160] A. H. Tyson, G. E. Mawdsley, and M. J. Yaffe, "Measurement of compressed breast thickness by optical stereoscopic photogrammetry," *Medical Physics*, vol. 36, no. 2, pp. 569–576, 2009.

[161] G. E. Mawdsley, A. H. Tyson, C. L. Peressotti, R. A. Jong, and M. J. Yaffe, "Accurate estimation of compressed breast thickness in mammography," *Medical Physics*, vol. 36, no. 2, pp. 577–586, 2009.

[162] O. Alonzo-Proulx, N. Packard, J. M. Boone et al., "Validation of a method for measuring the volumetric breast density from digital mammograms," *Physics in Medicine and Biology*, vol. 55, no. 11, pp. 3027–3044, 2010.

[163] A. Hufton, S. M. Astley, T. Marchant, and H. Patel, "A method for the quantification of dense breast tissue from digitised mammograms," in *Proceedings of the International Workshop on Digital Mammography*, pp. 430–435, 2004.

[164] M. Berks, J. Diffey, A. Hufton, and S. Astley, "Feasibility and acceptability of stepwedge-based density measurement," in *Proceedings of the International Workshop on Digital Mammography*, pp. 355–361, 2006.

[165] J. A. Shepherd, L. Herve, J. Landau, B. Fan, K. Kerlikowske, and S. R. Cummings, "Novel use of single X-Ray absorptiometry for measuring breast density," *Technology in Cancer Research & Treatment*, vol. 4, no. 2, pp. 173–182, 2005.

[166] S. Malkov, J. Wang, K. Kerlikowske, S. R. Cummings, and J. A. Shepherd, "Single X-ray absorptiometry method for the quantitative mammographic measure of fibroglandular tissue volume," *Medical Physics*, vol. 36, no. 12, pp. 5525–5536, 2009.

[167] J. A. Shepherd, K. Kerlikowske, L. Ma et al., "Volume of mammographic density and risk of breast cancer," *Cancer Epidemiology Biomarkers & Prevention*, vol. 20, no. 7, pp. 1473–1482, 2011.

[168] S. Malkov, J. Wang, F. Duewer, and J. A. Shepherd, "A calibration approach for single-energy x-ray absorptiometry method to provide absolute breast tissue composition accuracy for the long term," in *Breast Imaging*, Lecture Notes in Computer Science, pp. 769–774, Springer, 2012.

[169] R. Highnam, X. Pan, R. Warren, M. Jeffreys, G. Davey Smith, and M. Brady, "Breast composition measurements using retrospective standard mammogram form (SMF)," *Physics in Medicine and Biology*, vol. 51, no. 11, pp. 2695–2713, 2006.

[170] S. Petroudi, K. Marias, R. English, R. Adams, and M. Brady, "Classification of mammographic patterns using area measurements and the standard mammogram form (SMF)," in *Medical Image Understanding and Analysis*, pp. 197–200, 2002.

[171] K. Marias, S. Petroudi, R. English, R. Adams, and M. Brady, "Subjective and computer-based characterisation of mammographic patterns," in *Proceedings of the International Conference on Digital Mammography*, pp. 552–557, 2002.

[172] K. Marias, C. Behrenbruch, R. Highnam, S. Parbhoo, A. Seifalian, and M. Brady, "A mammographic image analysis method to detect and measure changes in breast density," *European Journal of Radiology*, vol. 52, no. 3, pp. 276–282, 2004.

[173] V. A. McCormack, R. Highnam, N. Perry, and I. D. S. Silva, "Comparison of a new and existing method of mammographic density measurement: intramethod reliability and associations with known risk factors," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 16, no. 6, pp. 1148–1154, 2007.

[174] J. Ding, R. Warren, I. Warsi et al., "Evaluating the effectiveness of using standard mammogram form to predict breast cancer risk: case-control study," *Cancer Epidemiology Biomarkers & Prevention*, vol. 17, no. 5, pp. 1074–1081, 2008.

[175] S. van Engeland, P. R. Snoeren, H. Huisman, C. Boetes, and N. Karssemeijer, "Volumetric breast density estimation from full-field digital mammograms," *IEEE Transactions on Medical Imaging*, vol. 25, no. 3, pp. 273–282, 2006.

[176] S. Lau, K. H. Ng, and Y. F. A. Aziz, "Are volumetric breast density measurements robust enough for routine clinical use?" in *Proceedings of the European Congress of Radiology*.

[177] J. Wang, A. Azziz, B. Fan et al., "Agreement of mammographic measures of volumetric breast density to MRI," *PLoS ONE*, vol. 8, no. 12, Article ID e81653, 2013.

[178] M. Jeffreys, J. Harvey, and R. Highnam, "Comparing a new volumetric breast density method (Volpara) to Cumulus," in *Digital Mammography*, vol. 6136 of *Lecture Notes in Computer Science*, pp. 408–413, Springer, Berlin, Germany, 2010.

[179] L. Beattie, E. Harkness, M. Bydder et al., "Factors affecting agreement between breast density assessment using volumetric methods and visual analogue scales," in *Breast Imaging*, vol. 8539 of *Lecture Notes in Computer Science*, pp. 80–87, Springer, Cham, Switzerland, 2014.

[180] J. S. Brand, K. Czene, J. A. Shepherd et al., "Automated measurement of volumetric mammographic density: a tool for widespread breast cancer risk assessment," *Cancer Epidemiology Biomarkers Prevention*, vol. 23, no. 9, pp. 1764–1772, 2014.

[181] D. Kontos, R. Berger, P. R. Bakic, and A. D. A. Maidment, "Breast tissue classification in digital breast tomosynthesis images using texture features: a feasibility study," in *Medical Imaging 2009: Computer-Aided Diagnosis*, Proceedings of SPIE, Lake Buena Vista, Fla, USA, February 2009.

[182] D. Kontos, P. R. Bakic, A. B. Troxel, E. F. Conant, and A. D. A. Maidment, "Digital breast tomosynthesis parenchymal texture analysis for breast cancer risk estimation: a preliminary study," in *Digital Mammography*, vol. 5116 of *Lecture Notes in Computer Science*, pp. 681–688, Springer, Berlin, Germany, 2008.

[183] E. M. S. Pertuz, S. P. Weinstein, E. F. Conant, and D. Kontos, "Fully automated volumetric breast density estimation from digital breast tomosynthesis images: multi-modality comparison with digital mammography and breast MRI," in *Proceedings of the Radiological Society of North America—Scientific Assembly and Annual Meeting*, 2014.

[184] C. M. Shafer, V. L. Seewaldt, and J. Y. Lo, "Validation of a 3D hidden-Markov model for breast tissue segmentation and density estimation from MR and tomosynthesis images," in *Biomedical Sciences and Engineering Conference (BSEC '11)*, pp. 1–4, March 2011.

[185] C. M. Shafer, V. L. Seewaldt, and J. Y. Lo, "Segmentation of adipose and glandular tissue for breast tomosynthesis imaging using a 3D hidden-Markov model trained on breast MRIs," in *Medical Imaging 2011: Physics of Medical Imaging*, vol. 7961 of *Proceedings of SPIE*, February 2011.

[186] H. D. Li, M. Kallergi, L. P. Clarke, V. K. Jain, and R. A. Clark, "Markov random field for tumor detection in digital mammography," *IEEE Transactions on Medical Imaging*, vol. 14, no. 3, pp. 565–576, 1995.

[187] J. A. Harvey and V. E. Bovbjerg, "Quantitative assessment of mammographic breast density: relationship with breast cancer risk," *Radiology*, vol. 230, no. 1, pp. 29–41, 2004.

[188] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "INbreast: toward a full-field digital mammographic database," *Academic Radiology*, vol. 19, no. 2, pp. 236–248, 2012.

[189] W. He, E. R. E. Denton, and R. Zwiggelaar, "A study on mammographic image modelling and classification using multiple databases," in *Breast Imaging*, vol. 8539 of *Lecture Notes in Computer Science*, pp. 696–701, Springer, 2014.

[190] H. F. Boehm, T. Schneider, S. M. Buhmann-Kirchhoff et al., "Automated classification of breast parenchymal density: topologic analysis of X-ray attenuation patterns depicted with digital mammography," *American Journal of Roentgenology*, vol. 191, no. 6, pp. W275–W282, 2008.

[191] R. T. Chlebowski, L. H. Kuller, R. L. Prentice et al., "Breast cancer after use of estrogen plus progestin in postmenopausal women," *The New England Journal of Medicine*, vol. 360, no. 6, pp. 573–587, 2009.

[192] C. M. Vachon, V. S. Pankratz, C. G. Scott et al., "Longitudinal trends in mammographic percent density and breast cancer risk," *Cancer Epidemiology Biomarkers and Prevention*, vol. 16, no. 5, pp. 921–928, 2007.

[193] N. F. Boyd, L. J. Martin, Q. Li et al., "Mammographic density as a surrogate marker for the effects of hormone therapy on risk of breast cancer," *Cancer Epidemiology Biomarkers & Prevention*, vol. 15, no. 5, pp. 961–966, 2006.

[194] B. Threatt, J. M. Norbeck, N. S. Ullman, R. Kummer, and P. Roselle, "Association between mammographic parenchymal pattern classification and incidence of breast cancer," *Cancer*, vol. 45, no. 10, pp. 2550–2556, 1980.

[195] D. B. Thomas, R. A. Carter, W. H. Bush Jr. et al., "Risk of subsequent breast cancer in relation to characteristics of screening mammograms from women less than 50 years of age," *Cancer Epidemiology Biomarkers & Prevention*, vol. 11, no. 6, pp. 565–571, 2002.

[196] A. Manduca, M. J. Carston, J. J. Heine et al., "Texture features from mammographic images and risk of breast cancer," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 18, no. 3, pp. 837–845, 2009.

[197] J. Couzin, "Dissecting a hidden breast cancer risk," *Science*, vol. 309, no. 5741, pp. 1664–1666, 2005.

[198] G. Torres-Mejía, B. de Stavola, D. S. Allen et al., "Mammographic features and subsequent risk of breast cancer: a comparison of qualitative and quantitative evaluations in the guernsey prospective studies," *Cancer Epidemiology Biomarkers & Prevention*, vol. 14, no. 5, pp. 1052–1059, 2005.

[199] H. Li, M. L. Giger, O. I. Olopade, and L. Lan, "Fractal analysis of mammographic parenchymal patterns in breast cancer risk assessment," *Academic Radiology*, vol. 14, no. 5, pp. 513–521, 2007.

[200] B. T. Nicholson, A. P. LoRusso, M. Smolkin, V. E. Bovbjerg, G. R. Petroni, and J. A. Harvey, "Accuracy of Assigned BI-RADS Breast Density Category Definitions," *Academic Radiology*, vol. 13, no. 9, pp. 1143–1149, 2006.

The Scientific World Journal

Gastroenterology Research and Practice

MEDIATORS of INFLAMMATION

Journal of Diabetes Research

Disease Markers

Journal of Immunology Research

International Journal of Endocrinology

PPAR Research

BioMed Research International

Hindawi

Submit your manuscripts at
http://www.hindawi.com

Journal of Ophthalmology

Stem Cells International

CAM
Evidence-Based Complementary and Alternative Medicine

Journal of Obesity

Journal of Oncology

Computational and Mathematical Methods in Medicine

Behavioural Neurology

Parkinson's Disease

AIDS Research and Treatment

Oxidative Medicine and Cellular Longevity