Microcalcification detection in full-field digital mammograms: A fully automated computer-aided system

T.M.A. Basile, A. Fanizzi, L. Losurdo, R. Bellotti, U. Bottigli, R. Dentamaro,
V. Didonna, A. Fausto, R. Massafra, M. Moschetta, P. Tamborra, S. Tangaro, D. La Forgia

PHYSICA MEDICA-EUROPEAN JOURNAL OF MEDICAL PHYSICS 64, pp.1-9, Aug 2019

Microcalcification Detection in Full-Field Digital Mammograms: a Fully Automated Computer-Aided System

Basile, T.M.A.[a,b,1,*], Fanizzi, A.[c,1], Losurdo, L.[c,1], Bellotti, R.[a,b], Bottigli, U.[d], Dentamaro R.[c], Didonna V.[c], Fausto A.[e], Massafra R.[c], Moschetta M.[f], Tamborra P.[c], Tangaro, S.[b], La Forgia, D.[c]

[a] *Department of Physics, University of Bari "Aldo Moro", Bari, Italy*
[b] *INFN National Institute for Nuclear Physics, Bari Division, Bari, Italy*
[c] *I.R.C.C.S. "Giovanni Paolo II" National Cancer Centre, Bari, Italy*
[d] *Department of Physical Sciences, Earth and Environment, University of Siena, Siena, Italy*
[e] *Department of Diagnostic Imaging, University Hospital of Siena, Siena, Italy*
[f] *Interdisciplinary Department of Medicine, University of Bari "Aldo Moro", Bari, Italy*

Abstract

Background: Microcalcification clusters in mammograms can be considered as early signs of breast cancer. However, their detection is a very challenging task because of different factors: large variety of breast composition, highly textured breast anatomy, impalpable size of microcalcifications in some cases, as well as inherent low contrast of mammograms. Thus, the need to support the clinicians' work with an automatic tool.

Methods: In this work a three-phases approach for clustered microcalcification detection is presented. Specifically, it is made up of a pre-processing step, aimed at highlighting potentially interesting breast structures, followed by a single microcalcification detection step, based on Hough transform, that is able to grasp the innate characteristic shape of the structures of interest. Finally, a cluster identification step to group microcalcifications is carried out by means of a clustering algorithm able to codify expert domain rules.

Results: The detection performance of the proposed method has been eval-

[*] *Corresponding author at:* Department of Physics, University of Bari - Via Amendola, 173, 70126 Bari, Italy.
Email address: teresamaria.basile@uniba.it (Basile, T.M.A.)
[1] These authors contributed equally.

uated on 364 mammograms of 182 patients obtaining a true positive ratio of 91.78% with 2.87 false positives per image.

Conclusions: Experimental results demonstrated that the proposed method is able to detect microcalcification clusters in digital mammograms showing performance comparable to different methodologies exploited in the state-of-art approaches, with the advantage that it does not require any training phase and a large set of data. The performance of the proposed approach remains high even for more difficult clinical cases of mammograms of young women having high-density breast tissue thus resulting in a reduced contrast between microcalcifications and surrounding dense tissues.

Keywords: Breast Cancer, Clustered Microcalcification Detection, Image Processing, Computer-Aided System, Hough Transform

---

1. Introduction

Worldwide breast cancer is the leading type of cancer in women [1]. However, 99 percent of women whose breast cancer was detected early (stage 1 or 0) survive beyond five years after diagnosis [2]. Indeed, when detected in good time, its treatment allows women to have a good prognosis reducing the rate of mortality and the incidence of surgery, radiation therapy and oncologic treatments [3]. Mammography is the most widely used technique in screening programs to detect breast cancer at a very early stage [4, 5], but it is not a perfect procedure. It is estimated that 11 percent of women undergoing annual mammogram testing to screen for breast cancer are unnecessarily called back for further testing; this high rate of false-positive exams may lead to added anxiety, unnecessary biopsies for the patient, and consequent excessive treatments such as surgical excision. Moreover, there is also an estimated rate of false negatives of about 13 percent, i.e. cases of breast cancer that radiologists fail to detect [6].

Among the main early indirect signs of breast cancer visible on mammograms are the microcalcifications, tiny spots of calcium deposits, whose diameter range from 0.1 to 1 mm. They can be localized or broadly diffused along the breast

ducts, scattered or clustered throughout the breast tissue [7]. Specifically, a microcalcification cluster reveals the presence of a breast cancer, so its detection is of higher clinical importance than the detection of isolated microcalcifications. Nevertheless, the identification of microcalcifications is a challenging task, not only because of the large variety of breast composition and highly textured breast anatomy, but also because of the inherent low contrast of mammograms and, in some cases, the impalpable size of microcalcifications. Also different types of breast structures (such as curvilinear structures) may appear along with the microcalcifications making their identification more difficult. In addition, the diagnostic capabilities of mammography decrease in breasts characterized by a high percentage of radiographically dense fibroglandular tissue, leading to increased follow-up studies, including biopsies [8, 9].

In addition to errors attributable to the perceptual difficulty of the task, a significant proportion of missed detection can be attributed to errors in perception (search failures), or alternatively to errors in interpretation. The low prevalence in mammography becomes an even more common source of error for inexperienced radiologists when they must face enormous numbers of mammograms generated in widespread screening. Indeed, in many cases, a disease detected in a current exam can also be seen in retrospect on the previous exam. These retrospectively visible or actionable cancers could have been detected in that previous exam with an expert reading [10, 11].

The early signs of breast cancer are often ambiguous or difficult to see, especially with regard to particular lesions such as microcalcifications.

For these reasons, it could be helpful to have an automatic tool able to support the clinicians in making more accurate detection, and consequently, diagnosis of breast pathologies. This is a field of research that has been active for several decades and various studies confirmed the benefits that the use of Computer Aided Detection (CAD) systems can bring to clinician work [12]. Indeed, in 2017, IBM Research co-organized a coalition of oncology and technology partners to pose this challenge to the science community and to find out if machine learning technologies could be used to increase the accuracy of mammography

screening [13].

Due to their particular nature and structure, the detection of microcalcifications requires a dedicated automatic tool and not one of generalized detection on the various possible lesions of the breast. Different techniques were used to solve this task by means of image analysis algorithms [14] stochastic modelling methods [15], methods for multiscale decomposition [16, 17] or more advanced techniques of machine learning [18, 19, 20, 21].

Nevertheless, most of the works presented in the literature lack an automatic localization of the suspicious regions, indeed they make use of manually selected regions. Other systems, whose scheme includes an automatic extraction of regions of interest, require a sizeable set of training data to learn a model able to efficiently detect clusters of microcalcifications. Here, we proposed a fully automated CAD system able to detect clusters of microcalcifications in full-field digital mammograms. Indeed, with respect to most of the state-of-art systems, the novelty of the present work is that it can detect regions containing clusters of microcalcifications, without any human intervention and without the need to carry out a laborious training phase that could require many labeled data.

The proposed method includes three main phases: an image pre-processing step followed by microcalcifications detection and cluster identification. As a first step, a set of image analysis algorithms are used in order to make potentially interesting breast structures more evident. Successively, a combination of a threshold-based method and Hough transform is used to detect the single microcalcifications. Finally, the single microcalcifications are automatically grouped into clusters in order to obtain the regions containing the lesions of interest. The proposed approach was tested on 364 full-field digital mammograms selected from the public database BCDR [22, 23] (Breast Cancer Digital Repository) obtaining results comparable with the state-of-art systems.

As development of a previously published work [24], an extensive experimental cross-validation evaluation was performed with the aim of assessing the optimal combination of the thresholds employed in the step of single microcalcification detection. Moreover, a more sophisticated automated procedure for

the identification of microcalcification clusters was developed: it is based on classification algorithms aimed at reducing the false positive rate.

2. Materials and Methods

2.1. Materials

In this study, images extracted by the Breast Cancer Digital Repository were used [22, 23]. The BCDR, a public database, is made up of two different repositories: (1) a Film Mammography-based Repository (BCDR-FM) and (2) a Full-Field Digital Mammography-based Repository (BCDR-DM). In the database all available medio-lateral oblique (MLO) and cranial caudal (CC) views of the left and right breast are included.

Only BCDR-DM images were used in this study due to their immediate advantage with respect to the analogical films and digitized mammograms: the lack of information loss and the reduced background noise. Finally, the use of full field digital images is mandatory for the exploitation of the implemented CAD on images obtained with the digital equipment owned by oncology institutes worldwide, in order to maintain functional compatibility for future clinical research activities.

The MLO and CC images are in TIF format with a resolution of 3328 (width) by 4084 (height) or 2560 by 3328 pixels, depending on the compression plate used in the acquisition (according to the patients breast size). All the images were of 0.35mm/pixel in spatial resolution. Furthermore, the BCDR-DM dataset provides, for each mammogram, clinical and demographic meta-data including patient age, breast density (there are four breast density categories reported, i.e. $< 25\%$, $[25-50]\%$, $[50-75]\%$, $> 75\%$) and `BIRADS` category (Table 1 summarizes the categories that are included in the international Breast Imaging Reporting And Data System (`BIRADS`) scoring system for the diagnosis of breast lesions [25]).

Category 0	incomplete (additional work-up is needed)
Category 1	negative (no findings found)
Category 2	benign finding(s)
Category 3	probably benign finding(s)
Category 4	suspicious abnormality
Category 5	highly suggestive of malignancy
Category 6	known biopsy with proven malignancy

Table 1: Breast Imaging Reporting And Data System (`BIRADS`) scoring.

2.1.1. Experimental Settings

The dataset used to asses the validity of the approach was made up of 364 mammograms randomly selected among those containing microcalcifications. Since the `BIRADS` label provided by the public database refers to the most important lesion found in the breast, it happens that none or only a single cluster were tagged even if more than one cluster of microcalcifications was present in the mammogram. Thus, as first step, we asked our expert radiologists to analyse the mammograms and tag each microcalcifications cluster, as we are interested in detecting all of this kind of lesion. In particular, a double blind reading was manually performed by two radiologists, dedicated to senologic diagnostics, from our Institute. They were asked to identify regions containing microcalcification clusters using a circular bounding box (as shown in Figure 2). Then, as a result of the comparison between the readings of the two radiologists, the clusters for which both radiologists agreed were considered.

In the dataset of 364 mammograms (236 containing clusters of microcalcifications), 353 clustered microcalcifications were identified with an average of 1.5 clusters for mammogram. To be precise, 172 clusters were classified as `BIRADS category 2`, 122 as `BIRADS 3` and 59 as `BIRADS 4-5`.

As regards the age distribution of the patients, the mammograms analyzed were from patients aged between 32 and 89, with an overall average age of 56.5. The age distribution of the patients in the dataset used in the analysis for each `BIRADS` category identified by radiologists is summarized in Table 2.

A further characterization of the patients considered for the study includes considerations about breast typology, and in particular breast density, with the

Age Range	BIRADS 2	BIRADS 3	BIRADS 4-5	Overall dataset
30-39	1.16% (2)	16.39% (20)	23.73% (14)	10.20% (36)
40-49	21.51% (37)	23.77% (29)	3.39% (2)	19.26% (68)
50-59	38.95% (67)	22.95% (28)	44.07% (26)	34.28% (121)
60-69	19.19% (33)	19.67% (24)	16.95% (10)	18.98% (67)
$\geq$ 70	19.19% (33)	17.21% (21)	11.86% (7)	17.28% (61)
Total	100% (172)	100% (122)	100% (59)	100% (353)

Table 2: Age distribution wrt BIRADS categories (% and # number of microcalcification clusters).

aim of assessing the validity of the proposed method even in the case of high-density breast tissues which is a more difficult case to analyze for both human experts and automatic tools.

A small percentage of dataset used, that is 6.04%, concerned a particular type of breast for which it was not possible to define its density because of its unusual characteristics (Fig. 1).

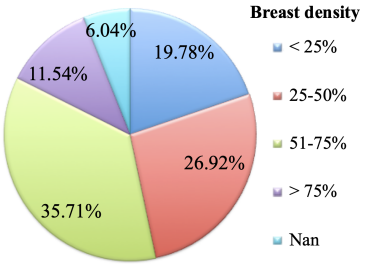


Figure 1: Breast density category distribution in the dataset.

2.2. Methods

As indicated in Section 1, the challenges in microcalcification detection are related to their dimension and morphology, but also to the hosting breast tissue both because of its density, that can vary among subjects, and because of the different surrounding/underlying structures that may appear along with the microcalcifications. Some significant examples are provided in Figure 2.

It is therefore essential to operate a segmentation of the image capable of isolating the microcalcifications from the rest of the breast tissue. With this

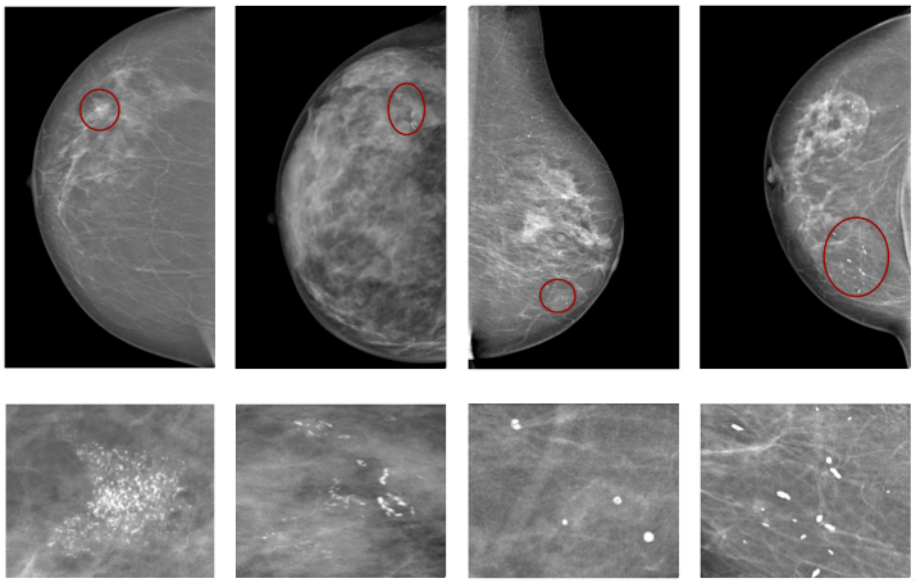


Figure 2: Some examples of diffused clusters of microcalcifications.

aim, the image is first processed to highlight the breast regions that potentially contain the interesting structures. Then, starting from the resulting elaboration, a two-stage segmentation step is performed in order to isolate the single microcalcifications. Once the candidate microcalcifications are obtained, a detection step devoted to group them in significant clusters is performed. Figure 3 outlines the general scheme of the proposed approach.

Regarding the first step, it is based on a set of well-defined image processing techniques appropriately selected for the task of obscuring irrelevant regions from the image thus resulting in a potentially interesting region finding process. The following phase is a two-stage step that interleaves saturation steps on the image, with the aim of removing noise from the images elaborated during the first step, with structures finding steps, the aim of which is to exactly delineate the single microcalcification contours. This is performed by means of the Hough transform. Finally, the microcalcifications are grouped in significant clusters by exploiting a set of codified domain expert rules automatically applied in the final step of the procedure. Each phase will be described in detail as follows.

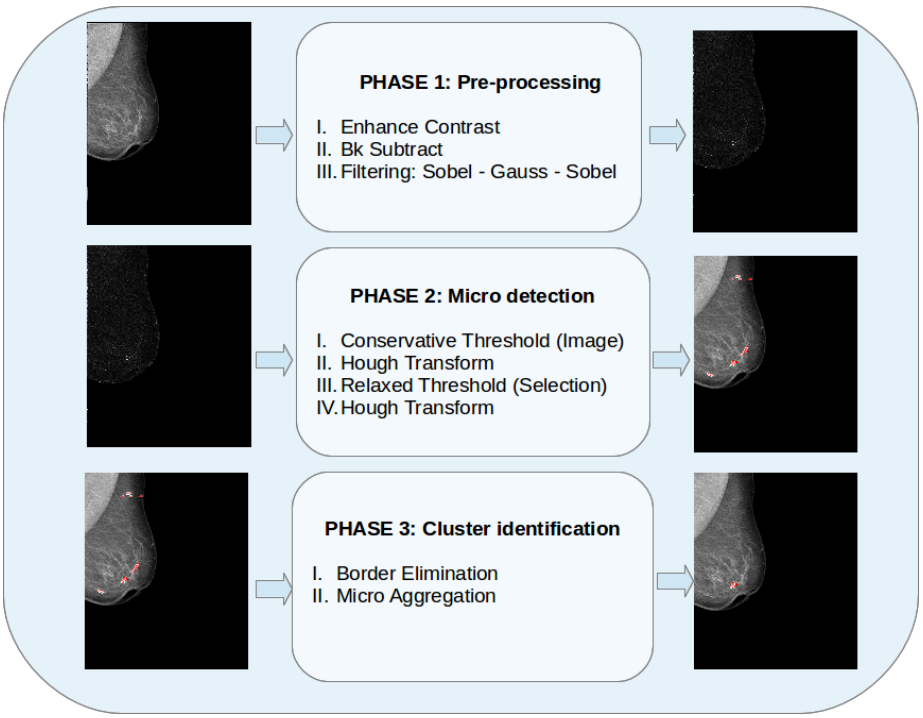


Figure 3: General scheme of the proposed approach.

2.2.1. Image Pre-processing

The goal of the pre-processing phase is to highlight the breast regions that potentially contain the interesting structures, i.e. the microcalcifications. This region highlighting process is performed using a set of purposely selected image analysis algorithms as reported below and sketched in Algorithm 1.

First, a contrast enhancement method is used to enhance image details. Due to the tight distribution of the grayscale in mammography images, the details are blurred. Transformation to the histogram of the original image turns them into uniform distribution thus allowing the enlargement of the grayscale interval. Consequently, the contrast is increased to distinguish the details. Basically, the image enhancement procedure first makes a transformation of the original image by dividing the interval of grayscales into two parts, and then reduces the range of intervals to enhance the contrast interval of the original image.

Algorithm 1 Pre-processing

Require: original image;

Ensure: pre-processed image (grayscale);

1: Enhance Contrast and Invert LUT (look up table);

2: Apply Background Subtraction Filter and Invert LUT (look up table);

3: Apply Sobel gradient;

4: Smooth with Gaussian filtering ($\sigma = 2$);

5: Apply Sobel gradient;

Successively, a background subtraction filter is run to make the breast tissue structures more evident. This procedure is based on the rolling ball algorithm [26] that is able to remove smooth continuous backgrounds from images. The underlying idea of such a method is that, imaging a 3D surface with the pixel values of the image being the height, a ball rolling over the back side of the surface creates the background. A local background value is determined for every pixel by averaging over a very large ball around the pixel. This value is then subtracted from the original image, removing large spatial variations of the background intensities.

The gradient image is then calculated to detect changes in contrast. Starting from the gray-level image, a Sobel gradient [27] is applied to detect edges. A Gaussian filtering is now necessary to smooth the image in order to obtain an image able to preserve the structure boundaries. At this point, the Sobel gradient of the image is again computed resulting in an image containing more evident boundaries of the structures of interest.

Figure 4 shows the application of the pre-processing phase on a sample image. As one can note, in the final step of the procedure (Fig. 4(f)) the region containing the microcalcifications is more evident with respect to the rest of breast structures.

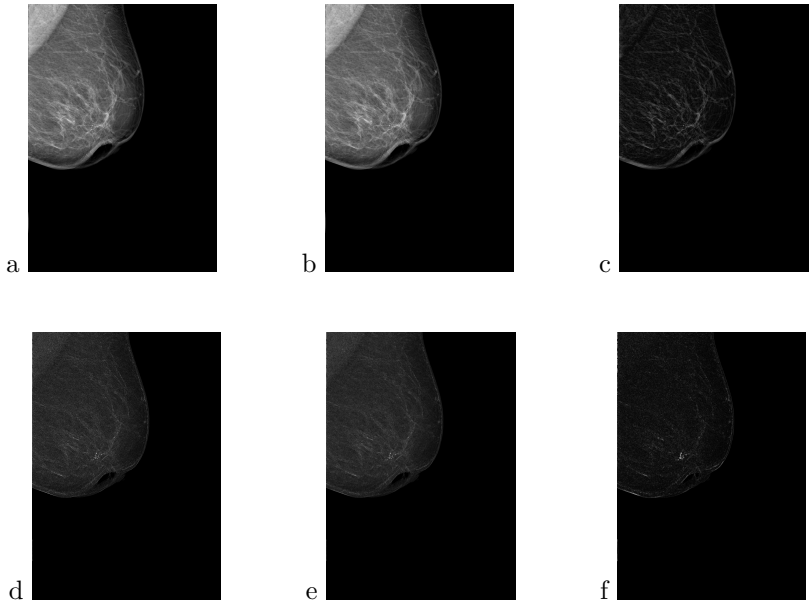


Figure 4: Pre-processing: (a) Original image. (b) Enhance Contrast (c) Subtract Background (d) Sobel edge detection (e) Gaussian Filter (f) Sobel edge detection.

2.2.2. Image Segmentation: microcalcification detection

The goal of this step is to provide a segmentation of the single microcalcifications contained in the image. It is a two-stage step that interleaves saturation phases on the image with structures (microcalcifications) discovering steps.

Specifically, we segment the image by considering that the shape of the microcalcifications can be approximated by a curved structure. Hence, we try to find the morphological structure that best approximates the microcalcification boundary by using the Hough transform [28, 29, 30]. This technique is used to transform a set of feature points in the image space into a set of accumulated votes in a parameter space. For each feature point, votes are accumulated in an accumulator array for all parameter combinations. The array elements that contain the highest number of votes indicate the presence of the shape.

The parameter space of the Hough transform is three-dimensional: it consid-

ers the coordinates of the center and the dimension of the structures. This could make a direct implementation of the Hough transform very complex. However, medical literature provides information about the dimension of microcalcifications. Hence, we can exploit this a priori knowledge to reduce the computational effort required by the application of the Hough transform. The Hough transform was therefore applied with the constraints of finding circles with a radius varying from $r_1 = 2$ pixels to $r_2 = 20$ pixels[2], as defined on the basis of domain expert knowledge and according to medical literature [31]. The circles that best match the microcalcification edges in the binary image are selected.

However, before applying the Hough transform to the pre-processed image, a noise removal step is necessary. Although microcalcifications in mammograms appear as relatively bright regions in comparison with the surrounding breast tissue or masses, when a lot of glandular tissue is present the mammograms are very bright even after the first general pre-processing phase, thus making small microcalcifications poorly visible [32]. For this reason, it is necessary to increase the contrast of the pre-processed image with respect to its distribution of gray levels.

Therefore, in order to adjust mammography intensity values, the grayscale pixel values of the pre-processed image are mapped into a new image such that values lower than a certain threshold become saturated. In particular, two levels of threshold on different selections of the image are defined. In the first one, a first threshold of saturation (α_1) is fixed that works on the whole image (see Alg. 2: line 1). In the second level, carried out after a preliminary application of the Hough transform (Alg. 2: line 2), microcalcifications found in the first step are brought to the average level and another saturation threshold level (α_2) is applied on 200x200 pixels windows selected in the neighborhood of each microcalcification detected by the Hough transform in its preliminary

[2]In the specific case of the dataset exploited for the study, the images have a spatial resolution of 72ppi (0.35mm/pixel) and hence the range 2-20 pixels corresponds to a value ranging from 0.7 to 7 mm.

application (i.e. Alg. 2: line 2). As last step, the Hough transform is applied again on each of the selected sub-images (Alg. 2: lines 5-13).

This cascaded segmentation process is aimed at reducing the number of false positives in the image.

Algorithm 2 Microcalcification Detection

Require: I = pre-processed image;
$Orig$ = Original mammogram;

Ensure: F = superimposition of detected microcalcifications on starting image;
$\{[R, C]\}$ = detected microcalcifications: radii and center coordinates;

1: I_1 = First Saturation on I (α_1);
2: $\{[R_1, C_1]\}$ = Apply Hough Transform on I_1 - radius in $\{r_1 \dots r_2\}$;
/* $\{[R_1, C_1]\}$ = {*[radii of detected microcalcifications, centers]*} */
3: $Radii = \{R_1\}$;
4: $Centers = \{C_1\}$;
5: **for** $[r, c] \in \{[R_1, C_1]\}$ **do**
6: W = Select a 200x200 pixel window on the neighborhood of c;
7: m_W = average gray-level on W;
8: W = Saturate to m_w the pixel gray-level value in the area with center c and radius r;
9: W_1 = Second Saturation on W (α_2);
10: $\{[R_c, C_c]\}$ = Apply Hough Transform on W_1 - radius in $\{r_1 \dots r_2\}$;
11: $Radii = \{Radii \cup R_c\}$;
12: $Centers = \{Centers \cup C_c\}$;
13: **end for**
14: $\{[R, C]\} = \{[Radii, Centers]\}$ (radii and centers of detected structures);
15: F = Detected microcalcifications depicted on original mammogram $Orig$;

2.2.3. Image Segmentation: Cluster Detection

Once all the single microcalcifications have been detected, a cluster identification step is performed. This step is based on a set of expert domain codified

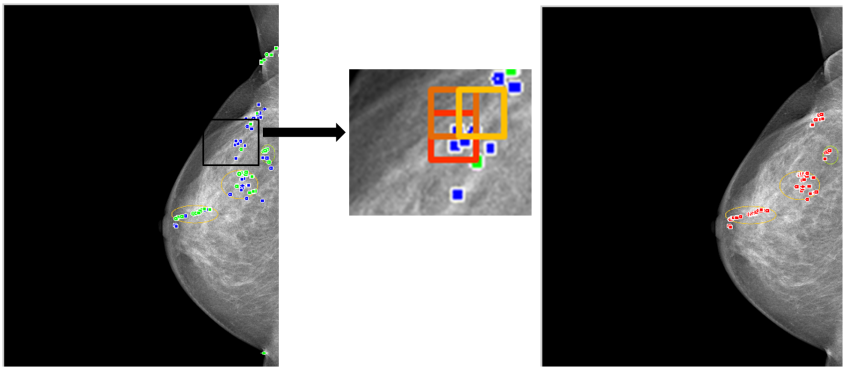


Figure 5: Microcalcification cluster detection. (a) A sliding window of 100X100 pixels is moved on the image and horizontally and vertically shifted by 50X50 pixels at each run. (The green circles correspond to the microcalcifications detected with the first saturation threshold (α_1) on the whole image, while the blue circles correspond to microcalcifications found with the second saturation threshold (α_2) on 200x200 pixel windows selected in the neighborhood of each microcalcification detected in the first detection step (see Alg.2)). (b) Only significant microcalcification clusters: only clusters containing more than three microcalcifications are significant.

rules: the first one regards the position of the microcalcification that can never appear on the breast borders, the second rule considers the cluster characterization, i.e. a cluster becomes significant, and hence deserving attention, if it contains a set of microcalcifications in a restricted area [33, 34].

This consists of a preliminary cleaning phase on the set of the structures found belonging to the breast border as they do not represent a breast lesion. Successively, the procedure scans the image looking for significant clusters, namely it aggregates microcalcifications by using a sliding window of 100x100 pixels on the image, that is shifted by 50x50 pixels at each run, so as to overlap its neighbors both horizontally and vertically (Fig. 5(a)). Since microcalcifications are suspect when they are in a cluster, in such a process, only clusters made up of more than three microcalcifications in a well-defined area are considered for the output result (Fig. 5(b)).

Finally, starting from the microcalcifications found, clustering techniques are

applied to automatically identify the microcalcification clusters in the image. In particular, in order to identify the number k of clusters, we first apply a Ward's hierarchical agglomerative clustering algorithm to the Euclidean distance between the coordinates of the microcalcifications within the image; successively, to optimize the classification result, we apply a non-hierarchical algorithm, i.e. k-means clustering [35].

Before reporting the results, an important comment on the measure of accuracy adopted is in order. A cluster is considered as a True Positive (TP) cluster, and hence considered in the evaluation, if the number of its detected foci is larger than three, as explained above, and it is located within the true marking (ground truth) of the radiologists. Compared to the manual segmentation by radiologists, the performance of the proposed system was evaluated in terms of sensitivity and false positives per image rate. Sensitivity (Sens) is the proportion of True Positive (TP) with respect to the total number of TP and False Negatives (FN). TP are cases correctly identified by the system, in accordance with expert annotations (the system and the experts are in accordance), while with FN the experts tagged a cluster that the system was not able to detect, i.e.

$$Sens = TP/(TP + FN)$$

The False Positives per image (FPi) Rate is the average of False Positive (FP) cases (the system identified a cluster that the experts did not tag) out of the total of the images analyzed (n), i.e.:

$$FPi = FP/n$$

3. Results

3.1. Saturation Thresholds Validation

As reported in Algorithm 2 and explained in Section 2.2.2, the detection of the individual microcalcifications is based on two levels of saturation, respectively on the whole image and on the area around the microcalcification

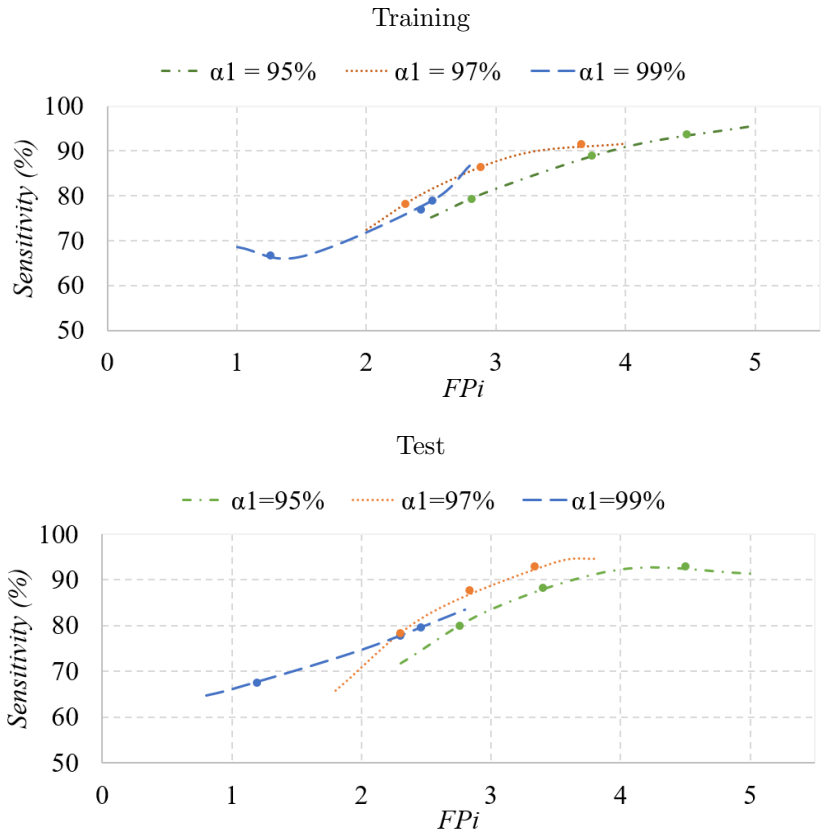
detected in the first step, that exploit as many saturation thresholds (α_1 and α_2 in Section 2.2.2).

Here we report the analysis performed in cross-validation on the whole dataset with the aim of evaluating the combination of the optimal saturation thresholds. In particular, we evaluated the classification performance on 100 ten-fold cross-validation rounds, varying the saturation thresholds of the first and second detection levels. We empirically assessed the effect of the combination of three values for the saturation thresholds (95%, 97%, 99%) applied to the whole image (α_1) and to the area of the microcalcifications (α_2) identified in the previous step (see Sec. 2.2.2). On the basis of the empirical evaluations carried out, we estimated the Free-response Receiver Operating Characteristic (FROC) curve for each of the three saturation threshold values on the first and second detection level (Fig. 6) obtained in training and in testing cross-validation. The combination of the saturation threshold that showed the best trade off between sensitivity and FPi was $\alpha_1 = 97\%$ and $\alpha_2 = 95\%$; this combination was able to detect the microcalcification clusters with a median sensitivity of 91.46% and 92.89%, and a median FPi value of 3.67 and 3.33 in training and in test cross-validation, respectively.

The first threshold of saturation $\alpha_1 = 97\%$ on the whole image was more conservative with respect to the second one, $\alpha_2 = 95\%$, which is applied only to the selected windows in the neighborhood of each microcalcification detected at the first step, thus allowing a reduction in the number of false positives in the image. Figure 7 shows the result of the double level of detection and its efficacy in recovering microcalcification structures potentially involved in the definition of a cluster.

3.2. Performance Evaluation

In the following, the results for $\alpha_1 = 97\%$ and $\alpha_2 = 95\%$ saturation thresholds are reported as this combination allowed us to obtain a good compromise between sensitivity and FPi. The proposed computer-aided detection system was able to detect the microcalcification clusters with a sensitivity of 91.78%



Training

	α_1 (Sens % – FPi # – Standard Error)					
	95%		97%		99%	
α_2	Sens	FPi	Sens	FPi	Sens	FPi
95%	93.73 ± 0.02	4.47 ± 0.01	**91.46 ± 0.02**	**3.67 ± 0.00**	78.97 ± 0.03	2.50 ± 0.00
97%	88.92 ± 0.02	3.74 ± 0.01	86.38 ± 0.03	2.88 ± 0.00	76.97 ± 0.04	2.41 ± 0.00
99%	79.26 ± 0.03	2.81 ± 0.00	78.18 ± 0.03	2.30 ± 0.00	66.77 ± 0.04	1.25 ± 0.00

Test

	α_1 (Sens % – FPi # – Standard Error)					
	95%		97%		99%	
α_2	Sens	FPi	Sens	FPi	Sens	FPi
95%	93.01 ± 0.21	4.50 ± 0.07	**92.89 ± 0.22**	**3.33 ± 0.02**	79.54 ± 0.31	2.46 ± 0.02
97%	88.19 ± 0.23	3.41 ± 0.07	87.67 ± 0.25	2.83 ± 0.04	77.78 ± 0.34	2.30 ± 0.02
99%	80.20 ± 0.31	2.75 ± 0.02	78.26 ± 0.31	2.28 ± 0.03	67.61 ± 0.37	1.19 ± 0.01

Figure 6: Free-response Receiver Operating Characteristic (FROC) curve (Graphics) and median values of the detection performance (Sens and FPi along with the Standard Error) (Tables) estimated in cross-validation for each of combination of the three saturation threshold values of the first (α_1) and second (α_2) detection steps.

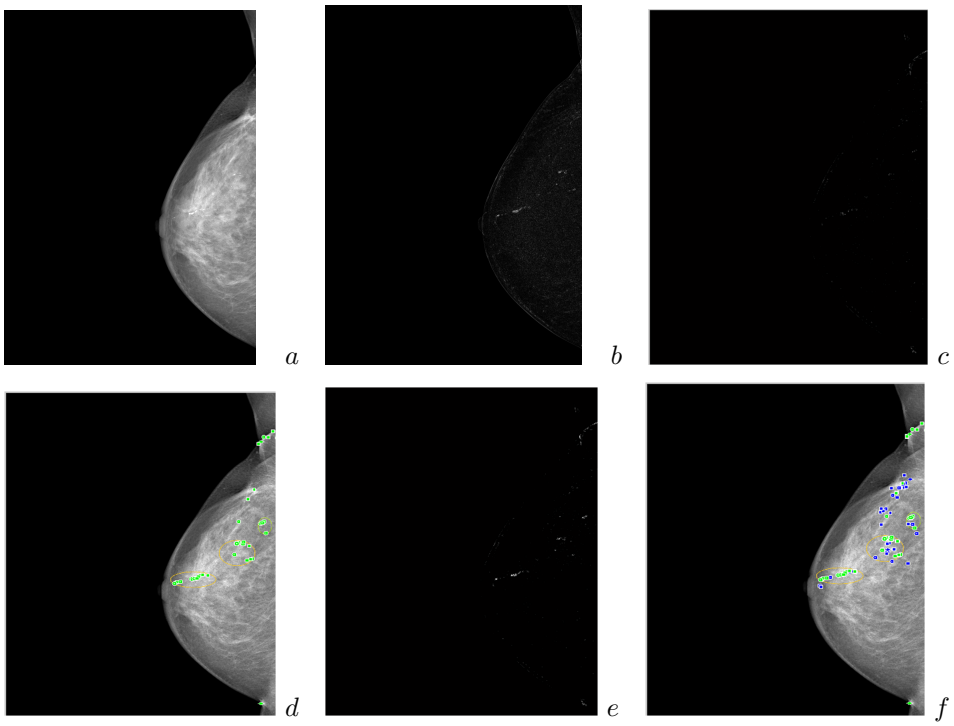


Figure 7: Single microcalcification detection. (a) Original Image (b) Pre-processed image. (c) Step1: conservative Saturation Threshold on Image with $\alpha_1 = 97\%$ (d). Step2: Hough Transform on image c (and superimposition of detected microcalcifications - green circles - on starting image). (e) Step3: relaxed Saturations Threshold on Selected Windows of 200x200 pixels with $\alpha_2 = 95\%$. (f) Step4: Hough Transform on thresholded 200x200 pixels selected windows (and superimposition of detected microcalcifications - blue circles - on starting image). Final detected microcalcifications.

with 2.87 FPi. In particular, the lowest sensitivity equal to 81.15% was obtained for the clusters tagged as `BIRADS category 3`, the dubious nature of lesions, while the system identified 98.26% and 94.92% of the clusters of `BIRADS 2` and `BIRADS 4-BIRADS 5`, respectively.

The performance of the system remained high for younger patients characterized by a dense breasts (Tab. 3). Specifically, the model detected the microcalcification clusters in women with breast density over 50% with a sensitivity higher than 90% and FPi less than 3.

Density classes	Mean Age	N. images	N. micro	Sensitivity	FPi
$< 25\%$	60.86 ± 11.06	72	46	88.89%	2.58
$25 - 50\%$	60.86 ± 9.62	98	80	95.00%	2.90
$51 - 75\%$	54.43 ± 10.99	130	182	90.66%	2.49
> 75	47.10 ± 5.07	42	40	92.50%	2.93

Table 3: Performance on different breast density classes (Images of breast with a density not available (Nan) were not considered).

4. Discussion

In this study, a fully automated CAD system for detecting clusters of microcalcifications in full-field digital mammograms has been proposed. It consists of three steps: a phase devoted to the enhancement of the structures of interest by means of image analysis techniques; the detection of individual microcalcifications performed by using a combination of a threshold-based method and Hough transform; finally, a cluster identification process that codifies some domain expert rules is carried out in order to group single microcalcifications into clusters. Experimental results showed that the proposed method is able to detect microcalcifications with a sensitivity of 91.78% and 2.87 FPi. In particular, the performance of our approach still remains high (over 89%) in the case of more difficult clinical cases of mammograms characterized by high-density breast tissue because of the predominance of fibroglandular tissues, thus resulting in a reduced contrast between microcalcifications and surrounding dense tissues. Indeed, it is well known that in particularly dense breast mammography sensitivity for early malignancy detection is reduced as a result of the effort required in locating cancer within an opaque and uniform background.

For the purpose of comparing the performance of the proposed approach with respect to the literature, the methods that use only full-field digital mammograms were selected, even if in some cases they use private databases. It is worth noting that since each approach uses a different set of images coming from different databases, the comparison is only made in a qualitative way.

As shown in Table 4, the result obtained by the proposed model is comparable to the state-of-art approaches, also with respect to the methods that

Methods	**Data (N. mammograms)** (with microcalcifications)	**Sensitivity**	**FPi**
DoG and SVM [14]	188 (188) (Private DB)	85%	0.28
Wavelet Decomposition[17]	138 (14) (Private DB)	92.9%	0.08
FCM-WF [18]	39 (39) (Private DB)	93%	0.04
Boosting classifier [19]	280 (90) (Private DB)	80%	$[1.28 - 3.02]$
Boosting classifier (CI) [19]	280 (90) (Private DB)	90%	$[3.54 - 4.09]$
ICA [20]	200 (100) (DDSM [36])	81.8%	2.55
ICA + Age Feature [20]	200 (100) DDSM [36]	91.8%	4.45
Deep Learning (SAE) [37]	1204 (1204) (Private DB)	93%	Not Available
Deep Learning (CNN) [21]	188 (188) (Private DB)	90%	1.17
Proposed Method	364 (236) (BCDR [22])	91.78%	2.87

Table 4: Performance comparison of the proposed approach wrt the literature.

use well known machine learning approaches in order to learn classifiers able to characterize individual microcalcifications [14, 19, 20]. These approaches perform an initial training step on regions of interest previously determined and labeled in order to automatically learn and select the most salient features, which are subsequently used in a classifier to perform the detection of individual microcalcifications. Our approach outperforms the best performance in terms of sensitivity reported in [14] that was obtained with the combination of Difference-of-Gaussians (DoG) detector and SVM. The same applies with respect to the method reported in [19], where an increment of sensitivity (90% wrt 80%) is reached with a consequent increase in the false positive rate ($[3.54-4.09]$ wrt $[1.28 - 3.02]$ confidence interval), and the one presented in [20] where an increment of 10% in sensitivity (91.8% wrt 81.8%), with an increase in the false positive rate from 2.55 to 4.45, is reached when the patient's age is taken into account in the learning phase. All these works, however, unlike our approach, for training the model require a large set of data that has to contain a significant number of cases representative of different situations that can occur in real domains.

As to the methods that use image analysis techniques to suppress the noise in the image and improve the contrast between the regions of interest and the background [17, 18], they show a sensitivity still comparable to that obtained by

our proposed method but declare less than one FPi. However, in [17] the performance was evaluated on a unbalanced dataset made up of 138 mammograms of which only 14 images contain microcalcifications. On the other hand, in [18] the results were obtained on dataset decidedly reduced compared to those in the literature; moreover, they used a private dataset purposely designed for the task. Indeed, in this last work, a simulated database was created by injecting microcalcifications in healthy mammograms, the number of false positives was computed as the number of healthy segmented microcalcifications that were erroneously associated with the cluster and the method was evaluated by counting the number of single microcalcifications correctly associated to the clusters.

The proposed method still performs when compared to deep learning approaches that have recently been used to characterize clusters of microcalcifications [37, 21]. Nevertheless, the drawback of such methods consists in requiring a large dataset for their discrimination [38].

5. Conclusion

Mammography is the most widespread screening method for early detection of breast cancer. However, it is a heavy work for radiologists to provide an accurate and uniform evaluation due to several factors.

Early detection of breast cancer is often difficult, especially with regard to particular lesions such as microcalcifications and in women with dense breast tissue. Therefore, improving the detection of early signs of cancer represent an important task which could be effectively supported by automated system tools.

This work presents a method able to detect microcalcifications clusters in full-field digital mammograms showing performance comparable to different methodologies (machine learning, image analysis, multi-scale decomposition, deep learning) used in the state-of-art approaches. The performance of our approach remained high also when the breast was characterized by high-density tissue, that is when the human reader sensitivity for early detection is reduced due to the effort required in locating cancer within an opaque and uniform back-

ground. The computational effort required by almost all the methods reported in literature,moreover, is higher with respect to our proposal.

Because of the particular nature and dimensions of microcalcifications, this type of lesion requires a dedicated CAD system. However, as a supporting system a functional CAD must be able to detect different kind of lesions such as masses or distortions. The model developed therefore could be integrated in a more complex diagnostic tool.

As a concluding remark, it is worth noting that although mammography is still the standard of care for breast cancer prevention, according to American College of Radiology (ACR) appropriateness criteria, also Digital Breast Tomosynthesis (DBT) is currently considered a usually appropriate screening modality. For non-calcified cancer, DBT has been shown to have higher accuracy compared to digital mammography [39, 40, 41], whereas for evaluation of microcalcification clusters there are discordant observations [42, 43, 44]. In any case, a CAD system for detecting microcalcification clusters may be useful because of the limitations in their visibility due to several factors related to specific characteristics of the instrumentation. DBT generates a pseudo 3-dimensional volume which is different from the typical 2D image of a digital mammography. However, in future work we will test our approach, opportunely revised, on the 3D images generated from DBT. In particular, we will analyze the possibility of applying the proposed approach on each of the 2D slices generated by the DBT exam and then combining the resulting analysis.

Acknowledgments

This work was supported by the Italian Oncology Institute "Giovanni Paolo II" I.R.C.C.S. Ricerca Corrente 2016 project: *"Analysis and development of a computer-aided detection system for breast cancer diagnosis based on multimodal imaging"*.

References

[1] National Cancer Institute, Cancer Stat Facts - Female Breast Cancer, https://seer.cancer.gov/statfacts/html/breast.html.

[2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, A. Jemal, Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA: a cancer journal for clinicians 68 (6) (2018) 394–424.

[3] K. D. Miller, R. L. Siegel, C. C. Lin, A. B. Mariotto, J. L. Kramer, J. H. Rowland, K. D. Stein, R. Alteri, A. Jemal, Cancer treatment and survivorship statistics, 2016, CA: a cancer journal for clinicians 66 (4) (2016) 271–289.

[4] A. Howell, The emerging breast cancer epidemic: early diagnosis and treatment, Breast Cancer Research 12 (4) (2010) S10.

[5] K. A. Cronin, A. J. Lake, S. Scott, R. L. Sherman, A.-M. Noone, N. Howlader, S. J. Henley, R. N. Anderson, A. U. Firth, J. Ma, et al., Annual report to the nation on the status of cancer, part i: National cancer statistics, Cancer 124 (13) (2018) 2785–2800.

[6] Breast cancer surveillance consortium. sensitivity, specificity, and false negative rate for 1,682,504 screening mammography examinations from 2007 - 2013. based on bcsc data through 2013, https://www.bcsc-research.org/statistics/screening-performance-benchmarks/screening-sens-spec-false-negative.html (2013).

[7] W. Dähnert, Radiology review manual, Lippincott Williams & Wilkins, 2011.

[8] C. M. Checka, J. E. Chun, F. R. Schnabel, J. Lee, H. Toth, The relationship of mammographic density and age: implications for breast cancer screening, American Journal of Roentgenology 198 (3) (2012) W292–W295.

[9] N. F. Boyd, L. J. Martin, M. Bronskill, M. J. Yaffe, N. Duric, S. Minkin, Breast tissue composition and susceptibility to breast cancer, Journal of the National Cancer Institute 102 (16) (2010) 1224–1237.

[10] K. K. Evans, R. L. Birdwell, J. M. Wolfe, If you dont find it often, you often dont find it: Why some cancers are missed in breast cancer screening, PloS one 8 (5) (2013) e64366.

[11] J. Mordang, A. Gubern-Mérida, A. Bria, F. Tortorella, R. Mann, M. Broeders, G. den Heeten, N. Karssemeijer, The importance of early detection of calcifications associated with breast cancer in screening, Breast cancer research and treatment 167 (2) (2018) 451–458.

[12] E. L. Henriksen, J. F. Carlsen, I. M. Vejborg, M. B. Nielsen, C. A. Lauridsen, The efficacy of using computer-aided detection (cad) for detection of breast cancer in mammography screening: a systematic review, Acta Radiologica 60 (1) (2019) 13–18.

[13] Ibm. dream challenge results: Can machine learning help improve accuracy in breast cancer screening?, https://www.ibm.com/blogs/research/2017/06/dream-challenge-results/ (2017).

[14] J. Wang, R. M. Nishikawa, Y. Yang, Improving the accuracy in detection of clustered microcalcifications with a context-sensitive classification model, Medical physics 43 (1) (2016) 159–170.

[15] S. Y. Shin, S. Lee, I. D. Yun, H. Y. Jung, Y. S. Heo, S. M. Kim, K. M. Lee, A novel cascade classifier for automatic microcalcification detection, PloS one 10 (12) (2015) e0143725.

[16] N. Arikidis, K. Vassiou, A. Kazantzi, S. Skiadopoulos, A. Karahaliou, L. Costaridou, A two-stage method for microcalcification cluster segmentation in mammography by deformable models, Medical Physics 42 (10) (2015) 5848–5861.

[17] X. Zhang, N. Homma, S. Goto, Y. Kawasumi, T. Ishibashi, M. Abe, N. Sugita, M. Yoshizawa, A hybrid image filtering method for computer-aided detection of microcalcification clusters in mammograms, Journal of Medical Engineering 2013.

[18] L. Vivona, D. Cascio, F. Fauci, G. Raso, Fuzzy technique for microcalcifications clustering in digital mammograms, BMC medical imaging 14 (1) (2014) 23.

[19] A. Oliver, A. Torrent, X. Lladó, M. Tortajada, L. Tortajada, M. Sentís, J. Freixenet, R. Zwiggelaar, Automatic microcalcification and cluster detection for digital and digitised mammograms, Knowledge-Based Systems 28 (2012) 68–75.

[20] R. Gallardo-Caballero, C. García-Orellana, A. García-Manso, H. González-Velasco, M. Macías-Macías, Independent component analysis to detect clustered microcalcification breast cancers, The Scientific World Journal 2012.

[21] J. Wang, R. Nishikawa, Y. Yang, Global detection approach for clustered microcalcifications in mammograms using a deep learning network., Journal of Medical Imaging 4 (2).

[22] R. Ramos-Pollán, M. A. Guevara-López, C. Suárez-Ortega, G. Díaz-Herrero, J. M. Franco-Valiente, M. Rubio-Del-Solar, N. González-De-Posada, M. A. P. Vaz, J. Loureiro, I. Ramos, Discovering mammography-based machine learning classifiers for breast cancer diagnosis, Journal of medical systems 36 (4) (2012) 2259–2269.

[23] D. C. Moura, M. A. G. López, P. Cunha, N. G. de Posada, R. R. Pollan, I. Ramos, J. P. Loureiro, I. C. Moreira, B. M. F. de Araújo, T. C. Fernandes, Benchmarking Datasets for Breast Cancer Computer-Aided Diagnosis (CADx), in: J. Ruiz-Shulcloper, G. Sanniti di Baja (Eds.), Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 326–333.

[24] A. Fanizzi, T. Basile, L. Losurdo, N. Amoroso, R. Bellotti, U. Bottigli, R. Dentamaro, V. Didonna, A. Fausto, R. Massafra, et al., Hough transform for clustered microcalcifications detection in full-field digital mammograms, in: Applications of Digital Image Processing XL, Vol. 10396, International Society for Optics and Photonics, 2017, p. 1039616.

[25] C. D'Orsi, E. Sickles, E. Mendelson, E. Morris, 2013 ACR BI-RADS Atlas: Breast Imaging Reporting and Data System, American College of Radiology, 2014.

[26] S. R. Sternberg, Biomedical Image Processing, Computer 16 (1) (1983) 22–34.

[27] R. C. Gonzalez, R. E. Woods, Digital Image Processing (3rd Edition), Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.

[28] P. V. Hough, Machine analysis of bubble chamber pictures, in: Conf. Proc., Vol. 590914, 1959, pp. 554–558.

[29] J. Sklansky, On the Hough technique for curve detection, IEEE Transaction on Computers 27 (10) (1978) 923–926.

[30] S. J. K. Pedersen, Circular Hough Transform, Aalborg University, Vision, Graphics, and Interactive Systems 2007.

[31] P. Henrot, A. Leroux, C. Barlier, P. Génin, Breast microcalcifications: the lesions in anatomical pathology, Diagnostic and interventional imaging 95 (2) (2014) 141–152.

[32] M. Ciecholewski, Microcalcification segmentation from mammograms: A morphological approach, Journal of Digital Imaging 30 (2) (2017) 172–184. doi:10.1007/s10278-016-9923-8.

[33] E. A. Sickles, Breast calcifications: mammographic evaluation., Radiology 160 (2) (1986) 289–293.

[34] Y.-Z. Shao, L.-Z. Liu, M.-J. Bie, C.-c. Li, Y.-p. Wu, X.-m. Xie, L. Li, Characterizing the clustered microcalcifications on mammograms to predict the pathological classification and grading: A mathematical modeling approach, Journal of digital imaging 24 (5) (2011) 764.

[35] A. Gordon, Classification, 2nd Edition, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, CRC Press, 1999.

[36] M. Heath, K. Bowyer, D. Kopans, R. Moore, W. P. Kegelmeyer, The Digital Database for Screening Mammography, in: M. Yaffe (Ed.), Proceedings of the Fifth International Workshop on Digital Mammography, Medical Physics Publishing, 2001, pp. 212–218.

[37] J. Wang, X. Yang, H. Cai, W. Tan, C. Jin, L. Li, Discrimination of breast cancer with microcalcifications on mammography by deep learning, Scientific reports 6 - 2016.

[38] T. Kooi, G. Litjens, B. van Ginneken, A. Gubern-Mérida, C. I. Sánchez, R. Mann, A. den Heeten, N. Karssemeijer, Large scale deep learning for computer aided detection of mammographic lesions, Medical image analysis 35 (2017) 303–312.

[39] S. Ciatto, N. Houssami, D. Bernardi, F. Caumo, M. Pellegrini, S. Brunelli, P. Tuttobene, P. Bricolo, C. Fantò, M. Valentini, et al., Integration of 3d digital mammography with tomosynthesis for population breast-cancer screening (storm): a prospective comparison study, The lancet oncology 14 (7) (2013) 583–589.

[40] B. M. Haas, V. Kalra, J. Geisel, M. Raghu, M. Durand, L. E. Philpotts, Comparison of tomosynthesis plus digital mammography and digital mammography alone for breast cancer screening, Radiology 269 (3) (2013) 694–700.

[41] P. Skaane, A. I. Bandos, R. Gullien, E. B. Eben, U. Ekseth, U. Haakenaasen, M. Izadi, I. N. Jebsen, G. Jahr, M. Krager, et al., Prospective trial

comparing full-field digital mammography (ffdm) versus combined ffdm and tomosynthesis in a population-based screening programme using independent double reading with arbitration, European radiology 23 (8) (2013) 2061–2071.

[42] M. L. Spangler, M. L. Zuley, J. H. Sumkin, G. Abrams, M. A. Ganott, C. Hakim, R. Perrin, D. M. Chough, R. Shah, D. Gur, Detection and classification of calcifications on digital breast tomosynthesis and 2d digital mammography: a comparison, American Journal of Roentgenology 196 (2) (2011) 320–324.

[43] D. Kopans, S. Gavenonis, E. Halpern, R. Moore, Calcifications in the breast and digital breast tomosynthesis, The breast journal 17 (6) (2011) 638–644.

[44] A. Tagliafico, G. Mariscotti, M. Durando, C. Stevanin, G. Tagliafico, L. Martino, B. Bignotti, M. Calabrese, N. Houssami, Characterisation of microcalcification clusters on 2d digital mammography (ffdm) and digital breast tomosynthesis (dbt): does dbt underestimate microcalcification clusters? results of a multicentre study, European radiology 25 (1) (2015) 9–14.