

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Incorporating Breast Asymmetry Studies into CADx Systems

---

José María Celaya Padilla,  
Cesar Humberto Guzmán Valdivia,  
Jorge Issac Galván Tejada, Carlos Eric Galván Tejada,  
Hamurabi Gamboa Rosales,  
Juan Rubén Delgado Contreras,  
Antonio Martínez-Torteya, Roberto Olivera Reyna,  
Jorge Roberto Manjarrez Sánchez,  
Francisco Javier Martínez Ruiz, Idalia Garza-Veloz,  
Margarita L. Martínez-Fierro, Victor Treviño and  
Jose Gerardo Tamez-Peña

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.69526>

---

## Abstract

Breast cancer is one of the global leading causes of death among women, and an early detection is of uttermost importance to reduce mortality rates. Screening mammograms, in which radiologists rely only on their eyesight, are one of the most used early detection methods. However, characteristics, such as the asymmetry between breasts, a feature that could be very difficult to visually quantize, is key to breast cancer detection. Due to the highly heterogeneous and deformable structure of the breast itself, incorporating asymmetry measurements into an automated detection system is still a challenge. In this study, we proposed the use of a bilateral registration algorithm as an effective way to automatically measure mirror asymmetry. Furthermore, this information was fed to a machine learning algorithm to improve the accuracy of the model. In this study, 449 subjects (197 with calcifications, 207 with masses, and 45 healthy subjects) from a public database were used to train and evaluate the proposed methodology. Using this procedure, we were able to independently identify subjects with calcifications (accuracy = 0.825, AUC = 0.882) and masses (accuracy = 0.698, AUC = 0.807) from healthy subjects.

**Keywords:** breast cancer, asymmetry, bilateral registration, CAD

---

## 1. Introduction

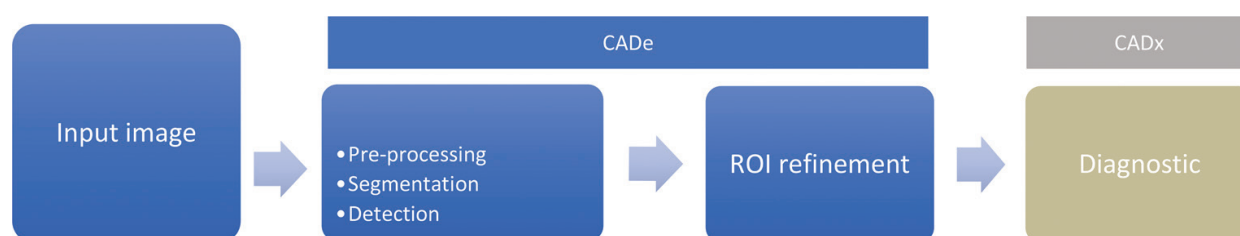
Cancer is one of the leading causes of death worldwide. In 2008, nearly 13% (7.6 million) of all deaths were cancer related. Among all types of cancer, lung, liver, colon, breast, and cervical are the most frequent ones. Recent studies predict 13.1 million cancer deaths for 2030 [1]. Among women, breast cancer is the deadliest type of cancer. Nearly 1.8% of all worldwide deaths are breast cancer related [2].

Till today, there is no cure for breast cancer, and since the trigger to develop any type of cancer is still a mystery, there is not an effective way to prevent the occurrence. Early detection of breast cancer plays a key role in a positive prognosis. There are several imaging technologies that might be used by specialists for the early detection of breast cancer, such as magnetic resonance imaging, ultrasound, and X-ray mammogram. The last technique is the primary tool used to diagnose and detect breast cancer worldwide, and it has been proved to be the best cost-effective tool to diagnose the disease [3].

In clinical practice, mammography allows for the detection of early signs of tumors before they become apparent [3]. Common signs of early cancer inside the breast tissue are micro-calcifications, architectural distortions, and masses [4]. During the screening procedure, radiologists use those signs to generate a standardized evaluation of the risk of cancer in a given patient, called Breast Imaging-Reporting and Data System (BI-RADS). This report helps oncologists to decide a course of action among women at risk of developing breast cancer [5].

The broad use of mammogram has driven the development of computer-aided detection (CADE) and computer-aided diagnosis (CADx) systems. While both approaches aim to assist radiologists to detect and diagnose breast cancer as early as possible, CADx systems are used as a second opinion [6] and CADE ones aim to improve visualization of the lesions (with up to 35% improvement in detection rate [7]). However, although it has been shown that CADE systems have helped radiologists to better interpret findings [8], it has also been demonstrated that in some cases they may make interpreting the images more difficult, reducing the accuracy of early cancer detection [7]. Furthermore, these systems also may increase the workload of the radiologists [8].

A typical CADE system, whose workflow is shown in **Figure 1**, consists of two algorithms applied sequentially, one to detect suspicious regions or regions of interest (ROI), and one to refine such regions. The former includes the preprocessing of the images, segmentation of



**Figure 1.** Typical workflow of a CADE/CADx system, adapted from Chen g et al. [13].

the breast tissue, and the detection of the ROI itself. The latter process is performed to reduce the number of false positives [8], and usually relies on machine learning techniques [9–11]. Lastly, the results are presented to the radiologist, highlighting in the original mammography the regions that the analysis deemed highly suspicious. As seen in the same image, CADx systems follow the same workflow as CADe ones. However, besides highlighting areas of higher risk to the radiologist, additional algorithms are used to analyze each ROI and generate a computer-based diagnosis. It is important to mention that, currently, few CADe and CADx commercial systems have been approved by the Food and Drug Administration of the United States of America [12].

Many methodologies used by CADx systems analyze only one breast, or even just a subregion of the breast, at a time. That is, they evaluate the left and right breasts as independent objects, unlike radiologists, who analyze images of both breasts simultaneously to evaluate their asymmetry. Radiologists do so because asymmetry is related to early signs of breast cancer (i.e. parenchymal distortion, bright spots, masses, etc.) [14, 15] and it may be used to reduce the rate of false positive detection of masses [16, 17]. Asymmetry can refer to either a longitudinal study, where current and prior mammograms are compared, or a bilateral study, where differences between the left and right breast are analyzed.

A few CADx systems have already tried to incorporate asymmetry studies to enhance diagnosis [14, 18–20]. Some researchers have studied the use of a feature-based asymmetry analysis, where the mammograms are processed individually and the differences between the individual analyses are used as a mean to quantify asymmetry [21]. This approach has also been used to characterize risk factors, such as breast density, and predict near-term breast cancer [14].

Another method that evaluates asymmetry, this one trying to mimic the approach used by radiologists, is the mammogram subtraction. In this approach, differences between mammograms are enhanced by performing a rigid registration (alignment) of the images. However, this methodology was originally employed only in longitudinal studies [22], comparing the same breast at two different times, since the highly heterogeneous and deformable tissue of the breast has hindered the inclusion of subtraction approaches in bilateral asymmetry studies [19].

Miller et al. [23] proposed a technique for the detection of bilateral symmetry using a semi-automated texture-based procedure that segments the glandular tissue, measuring the shape between views, and thus detecting the occurrence of asymmetries. The algorithm obtained an accuracy of 0.867 on a validation dataset of 30 screening mammogram pairs. Later, Miller et al. [24] presented a method for the detection of bilateral asymmetry based on measures of shape, topology, and distribution of brightness. This method was tested on 104 mammogram pairs, yielding a classification accuracy of 0.74.

Lau et al. [25] proposed a method for the detection of breast tumors that extracted measures of brightness, roughness, and directionality, and was based on localized asymmetry. This method was evaluated using 10 pairs of mammograms where asymmetry was a significant factor in the radiologist's diagnosis. A sensitivity of 0.92 was obtained, with 4.9 false positives per mammogram. However, the alignment was tuned manually using control points.

Ferrari et al. [26] characterized asymmetry as variations in oriented textural patterns, obtained using directional filtering with Gabor wavelets at different orientations and scales. Using a database with 80 images resulted in a classification accuracy of up to 0.744.

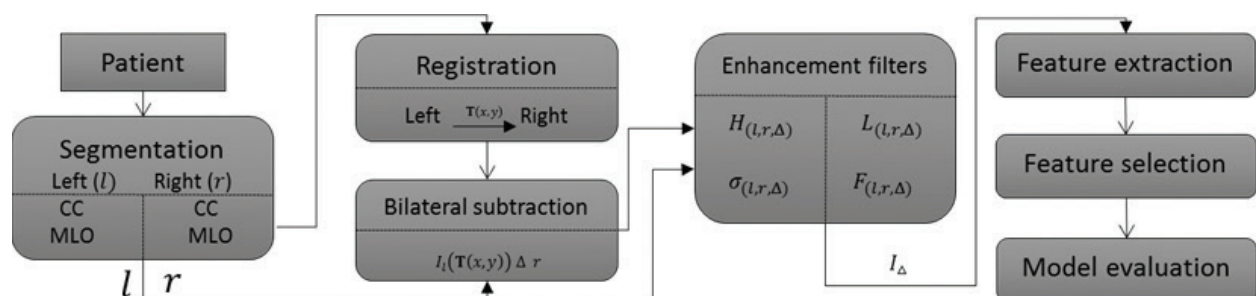
Rodriguez-Rojas et al. [21] presented a CADx system targeted to detect high-risk cancer patients. To do so, automated breast tissue segmentations were performed on 200 Mexican subjects labeled as either low- or high-risk according to their BI-RADS score. Then, 50 features were extracted, and bilateral differences between mammograms were defined by subtracting corresponding features in both mammograms. Finally, a genetic algorithm selected a predictive combination of features. Using this methodology, they were able to classify low-risk and high-risk cases with an area under the receiver operating characteristic (ROC) curve (AUC) of 0.88 on a 150-fold cross-validation set. The features included in the model were associated with the differences in signal distribution and tissue shape.

In summary, and as presented, most asymmetry detection methods are either feature-based, rely on simple bilateral subtraction techniques [14, 27], or depend on an ROI provided by a radiologist [24, 25]. Thus, in order to efficiently measure asymmetry, a better and automatic registration must be performed [28]. To do so, alignment has been improved by using the nipple as a reference point [29] and by co-registering both breasts using a robust point matching approach [22]. Nevertheless, none of those works include a fully automated bilateral registration. In this chapter, a methodology that incorporates an automatic asymmetry analysis with both a feature-based and a pixel-wise bilateral subtraction into a CADx system is presented.

## 2. Methodology

The proposed methodology follows the CADx workflow presented in the previous section. However, asymmetry measurements are used to aid in the diagnosis. To obtain such measurements, two additional stages are incorporated into the workflow: registration and pixel-wise subtraction. Additionally, a series of image transformations are incorporated to enhance different characteristics of the breast in the mammograms. This work is based on and follows previous efforts [30–32].

**Figure 2** shows how the bilateral asymmetry information was incorporated into the CADx system. Briefly, soft tissue is first segmented, the image of the left breast is then registered to its right counterpart and a bilateral subtraction of the co-registered images is performed;



**Figure 2.** Workflow of the proposed methodology.



images are then filtered and features are extracted; a multivariate model is selected using a train set; and finally, the model is evaluated on a validation set. A detailed explanation of each stage is presented in the following sections.

## 2.1. Materials

A total of 1796 digitalized film mammograms from 449 different subjects were used. From those, 45 were classified as healthy subjects (HS) (mean age of 59.3 and standard deviation (SD) of 9.8 years), 197 as subjects with malignant calcifications (CS) (mean age of 58 and SD of 10.9 years), and 207 as subjects with malignant masses (MS) (mean age of 64.1 and SD of 10.1 years). Each subject had the four standard mammograms taken, namely, left and right craniocaudal (CC), and left and right mediolateral oblique (MLO) projections.

In order to avoid problems associated with intra-scanner variability [17, 22, 33], all mammograms in this study were obtained from the Howtek dataset of the Digital Database for Screening Mammography public database [34], in which all mammograms were digitalized using a Howtek 960 scanner using a sampling rate of 43.5 micrometers per pixel and a 12-bit depth.

## 2.2. Segmentation

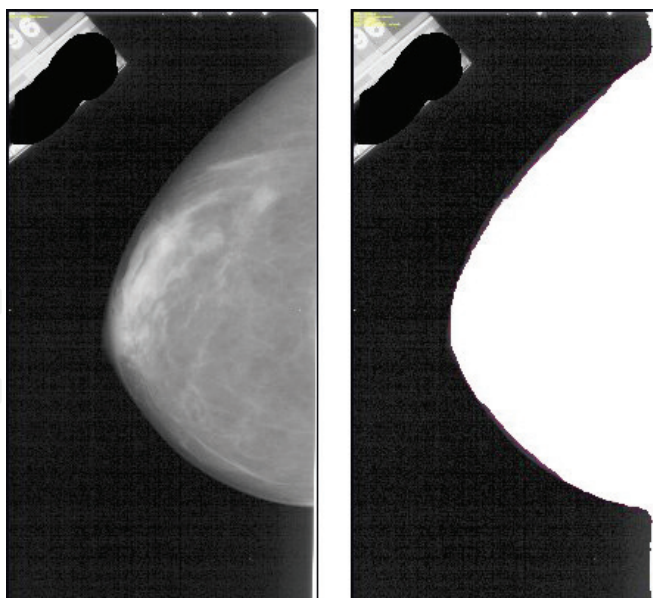
Segmentation, also called categorization by computer vision definitions, allows delimiting one or several parts of a given image assigning one class label (e.g. bone, muscle, fat, skin, calcification, and mass). This process is defined by the division or segmentation of the image into several homogeneous regions disjointed from their surroundings. A commonly used automatic segmentation of the breast tissue is based on the estimation of the background noise. For this study, an initial segmentation mask was created by estimating the background noise in the image and discarding all pixels below five standard deviations of the noise level. Then, holes were removed by applying closing morphological operations with a  $3 \times 3$  supporting region, as described by Eq. (1):

$$S(A) = (A(x, y) \oplus B(x, y)) \ominus B(x, y) \quad (1)$$

where  $\oplus$  and  $\ominus$  are the grayscale dilation and erosion morphological operations, respectively.  $B(x, y)$  is a  $3 \times 3$  structural element.  $A(x, y)$  is the image being segmented and  $S(A)$  is the resulting segmentation of the  $A(x, y)$  image. The largest connected region is used as the segmentation mask while all other high-intensity regions are removed from the images. **Figure 3** shows an example of the results of the segmentation procedure.

## 2.3. Registration

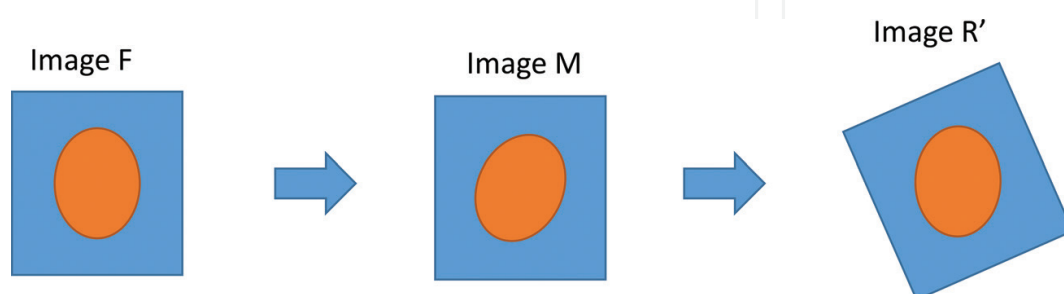
Image registration can be defined as the intensity and spatial mapping between two images [35]. Given two input images  $F$  and  $M$ , image registration can be expressed as  $R' = g[T(F)]$ , where  $T$  is a spatial transformation function,  $g$  an intensity transformation function, and  $R'$  the registered image. The transformation function is not always necessary; a lookup table can be used to pinpoint intensities. A visual example of image registration is presented in **Figure 4**, where an image  $M$  is being registered to match image  $F$ .



**Figure 3.** Segmentation of breast tissue. The image on the left is the original CC mammogram and the image on the right shows the superimposed segmentation mask in white (image from Ref. [32]).

Image registration has been widely used in medical applications [28, 36, 37]. However, the soft nature of the breast tissue makes them highly deformable, and rigid registration procedures, in which only rotation, translation, and scaling functions are used, are not sufficient. Therefore, nonrigid registration methods are necessary [38, 39]. There are many approaches to deal with medical imaging registration, the most recent comparison of algorithms based on a retrospective evaluation was published by West et al. [40], but it was constrained to do intra-patient rigid registration. Also recently, Diez et al. [28] and Celaya-Padilla et al. [30] compared registration algorithms with breast images as a source, and both concluded that the B-Splines approach was the most consistent.

Breast image registration based on a B-Splines transformation is defined as follows: given two input images ( $F$  = target image,  $M$  = image being registered),  $M$  is deformed by modifying a mesh of control points following a maximization of a similarity measure based on steepest descent gradient [6, 15]. The deformed image is compared to  $F$  using a similarity metric. If the images are similar enough, the process stops. Otherwise, the process reiterates.



**Figure 4.** Basic example of an image registration procedure.  $F$  is the target image,  $M$  is the image to be registered, and  $R'$  is the registered image.

**Figure 5** shows a multi-resolution pyramid approach [41] for the B-Spline implementation. There, the images are first registered using low-resolution images, the B-Spline transformation parameters are moved into the next higher resolution and parameter optimization is run again, and so on. This often avoids issues with local minima in the parameter search space and reduces computational time [15].

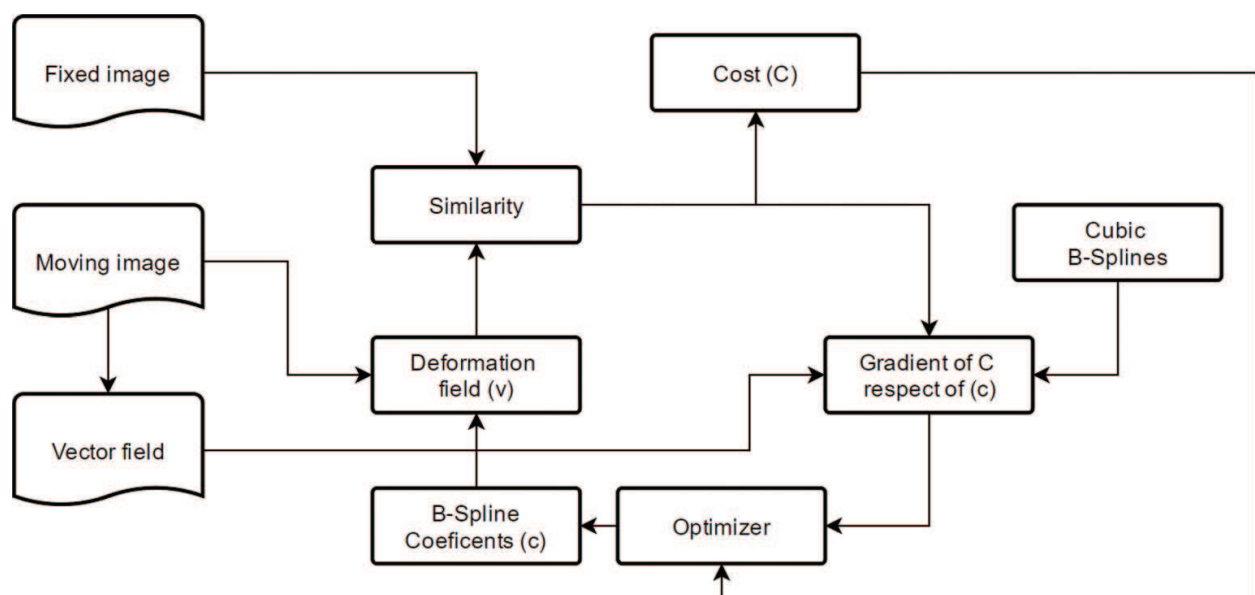
For this study, the image to be registered was first horizontally flipped. Then, both the moving image and the target image were resampled into a lower resolution image. Next, the pyramids for the multi-resolution were generated. Afterwards, the registration process detailed in **Figure 5** was carried out. And finally, the original moving image was deformed using the final parameters of the registration. For this implementation, mutual information [39] was used as the similarity metric. In **Figure 6**, the checkerboard of an example result from the B-Spline registration procedure is presented. There, it can be seen that the registered image was successfully aligned with its counterpart.

## 2.4. Image subtraction

Once the images were co-registered, a pixel-wise absolute difference was computed between the left and right images, as defined by Eq. (2) as follows:

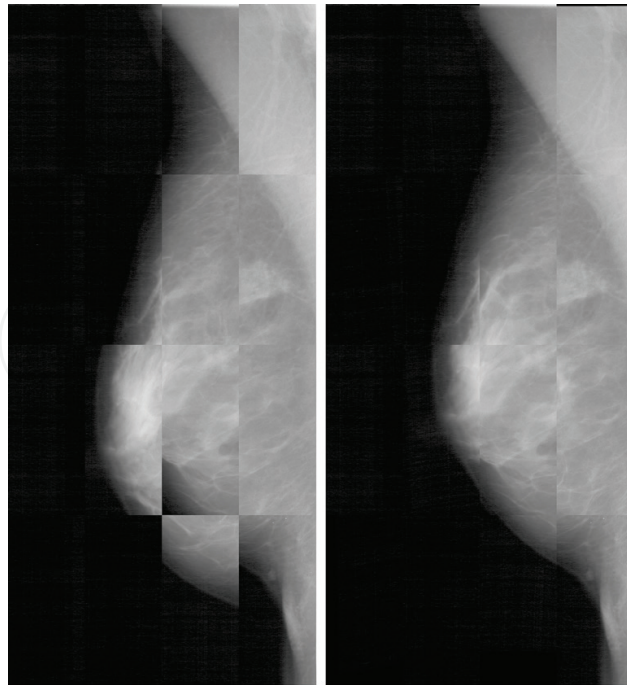
$$I_{\Delta}(x, y) = |I_r(x, y) - I_l(T(x, y))| \quad (2)$$

where  $I_r(x, y)$  represents the right image,  $I_l(T(x, y))$  represents the left image registered to the right image space, and  $I_{\Delta}(x, y)$  represents the map of absolute differences. **Figure 7** shows an example of the differential image for two given input images.

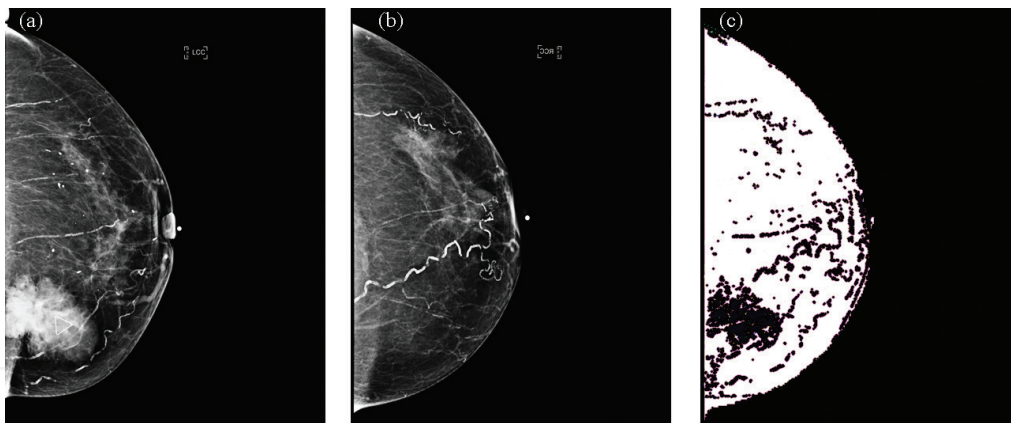


**Figure 5.** B-Spline registration typical framework.





**Figure 6.** Checkerboard comparison of images *pre* and *post* B-Spline registration. The image in the left shows a comparison between a left and horizontally flipped right breast before registration, and the right image shows the results of the registering process.



**Figure 7.** Image subtraction example. Left: unaltered CC view of left breast, middle: horizontally flipped CC view of right breast, and right: color map of the subtraction image  $I_{\Delta}$ . White and black pixels inside the breast tissue represent small and large intensity differences, respectively.

## 2.5. Image enhancement

To study the appearance of the architectural distortions, two enhancing filters were applied to the images: a morphological high-frequency enhancement filter ( $H$ ) designed to enhance fiber-like tissues, and a Laplacian of Gaussian filter ( $L$ ) that enhances high-frequency patterns inside the breast tissue. Additionally, since the texture between normal and abnormal tissues is different [42], two texture maps were created. The first map computed the local standard deviation ( $\sigma$ ) of the mammograms, and the second map computed the local fractal

dimension ( $F$ ). All image processing was implemented in C++ using Insight Segmentation and Registration Toolkit (ITK) libraries for image manipulation following previous efforts [32, 43].

## 2.6. Feature extraction

There are several features that may be quantified when aiming to detect early signs of cancer. For this analysis, 43 features were extracted from each image. These features can be grouped in three main categories: shape (i.e. area, perimeter, compactness, elongation, region centroid, region scatter), signal (i.e. mean, median, energy, variance, standard deviation, dynamic range, z mean, entropy, skewness, kurtosis, z range, fraction greater than z deviations, fraction lower than z deviations, value at fraction, 5% trimmed mean, 5% trimmed standard deviation, 5% trimmed z Mean), and morphology (i.e. total signal, signal centroid, signal scatter, and signal surface). Details of the full feature extraction procedure can be found in Ref. [32].

The enhancement filters and texture maps presented in Section 2.5 were applied to the four screening mammograms (i.e. left and right CC, and left and right MLO) and to the two bilateral subtraction images (CC and MLO), yielding a set of 15 images for both the CC and the MLO views:  $I_r, I_l, I_{\Delta}, H_r, H_l, H_{\Delta}, L_r, L_l, L_{\Delta}, \sigma_r, \sigma_l, \sigma_{\Delta}, F_r, F_l, F_{\Delta}$ , where  $I$  is the raw image,  $H, L, \sigma$ , and  $F$  are the enhanced images described in Section 2.5, and  $r, l$ , and  $\Delta$ , stand for the right, left, and bilateral subtraction images, respectively. Features were then extracted from this set of images. Additionally, to study the feature-based asymmetry analysis, the average and absolute difference of each left-right pair of measurements was also analyzed, resulting in 860 additional features, resulting in a total of 2150 features per subject.

## 2.7. Feature selection

The first step of the feature selection process consisted discarding highly correlated to avoid redundancy. For any pair of features with a Spearman correlation coefficient larger than 0.96, one feature was randomly selected to be kept, and the other removed from the selection. The dataset was normalized using the empirical distribution of the healthy subjects and a z-normalization was performed using the rank-based inverse normal transformation [44].

In order to select the most accurate and compact set of features from each dataset, the least absolute shrinkage and selection operator (LASSO) method was used [45]. The shrinkage and selection method minimizes the sum of squared errors and penalizes the regression coefficients, as described by Eq. (3) as follows:

$$\hat{\beta}^{\text{lasso}} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^N \left( y_i - \beta_0 - \sum_{j=1}^p x_{ij} \beta_j \right)^2 \text{ subject to: } \sum_{j=1}^p |\beta_j| \leq t \quad (3)$$

Given a set of input measurements  $x_1 \dots x_n$  and an outcome  $y$ , the lasso method fits a linear model where  $x_i$  is the covariate vector for the  $i^{\text{th}}$  case and  $y_i$  is the outcome,  $t$  is a tuning parameter that determines the amount of regularization, and  $N$  is the number of cases.

The multivariate search was performed using a class balanced data sample of 100 subjects for training and the remaining subjects as a blind test set. The models were calibrated using a leave-one-out cross-validation strategy, training the models at every split using  $N - 1$  subjects

and evaluating the model using the remaining subjects [46]. The final reported performance was obtained by applying the final model gathered on the training stage and evaluating it in the blind test set.

3. Results

A total of 1796 mammograms were successfully segmented. The image sets of nine subjects had to be removed from the experiment due to problems with the registration process, six were from MS, two from CS, and one from HS. All the remaining subjects were included in the subsequent stages of the analysis. The 2150 extracted features were filtered by the correlation process, removing 826 features.

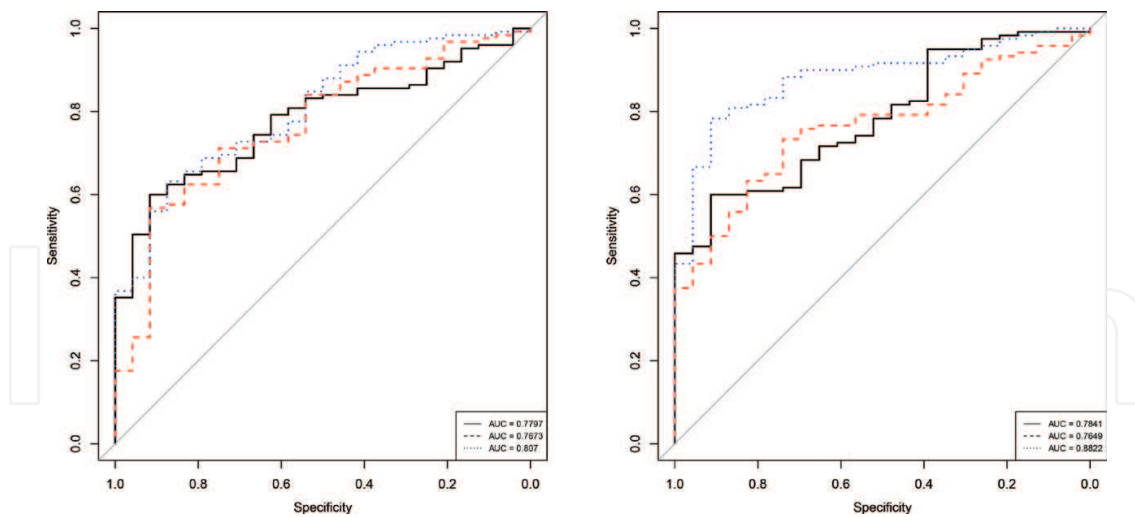
**Table 1** shows the features that were selected for each model: the CS versus HS ( $n = 12$ ), and the MS versus HS ( $n = 16$ ). The former achieved an accuracy of 0.825 with an AUC of 0.882 and the latter an accuracy of 0.698 with an AUC of 0.807. **Figure 8** shows the ROC curves for both the models.

CS versus HS				MS versus HS		
#	View	Image	Feature	View	Image	Feature
1	CC	$H_{\Delta}$	27	CC	$L_{\Delta}$	40
2	CC	$F_{\Delta}$	13	CC	$I_r$	29
3	CC	$I_r$	29	CC	$L_l$	40
4	CC	$H_l$	29	CC	$F_r$	6
5	CC	$H_l$	6	MLO	$H_l$	11
6	MLO	$I_l$	28	CC	$L_{\Delta\text{avg}}$	29
7	MLO	$H_l$	11	CC	$I_{\Delta\text{s}}$	28
8	MLO	$H_l$	21	CC	$\sigma_{\Delta\text{s}}$	38
9	CC	$I_{\Delta\text{s}}$	28	CC	$F_{\Delta\text{avg}}$	12
10	CC	$\sigma_{\Delta\text{s}}$	38	MLO	$I_{\Delta\text{s}}$	40
11	MLO	$L_{\Delta\text{avg}}$	27	MLO	$I_{\Delta\text{s}}$	28
12	CC	$H_{\Delta}$	27	MLO	$I_{\Delta\text{s}}$	29
13				MLO	$H_{\Delta\text{s}}$	31
14				MLO	$H_{\Delta\text{s}}$	7
15				MLO	$L_{\Delta\text{avg}}$	39
16				MLO	$L_{\Delta\text{avg}}$	27

Note: Features are grouped by dataset, symmetric features are denoted with:

$$I_{\Delta\text{avg}} = \frac{I_r + I_l}{2}, H_{\Delta\text{avg}} = \frac{H_r + H_l}{2}, L_{\Delta\text{avg}} = \frac{L_r + L_l}{2}, \sigma_{\Delta\text{avg}} = \frac{\sigma_r + \sigma_l}{2}, F_{\Delta\text{avg}} = \frac{F_r + F_l}{2},$$
$$I_{\Delta\text{s}} = |I_r - I_l|, H_{\Delta\text{s}} = |H_r - H_l|, L_{\Delta\text{s}} = |L_r - L_l|, \sigma_{\Delta\text{s}} = |\sigma_r - \sigma_l|, F_{\Delta\text{s}} = |F_r - F_l|.$$

**Table 1.** Features of the proposed models.



**Figure 8.** ROC curves for the classification models. Left: MS versus HS, right: CS versus HS. The dashed line represents the model with the features from only the difference images, the solid line the one with features from only the raw images, and the dotted line the one with the features from all images (from Ref. [32]).

## 4. Discussion

The proposed methodology is fully automated and does not require manual intervention as previous proposals [16, 17]. Although the approach is similar to others [22], we did not attempt to remove the pectoral muscle from the segmentation mask, since the presence of abnormal axillary lymph in this area is an indicator of occult breast carcinoma [47]. However, from the computational point of view, the feature extraction process may be affected if the region processed is not well focused [48].

The proposed registration process achieved a good performance having only 2.0% of the subjects that had to be discarded due to registration issues. This performance is remarkable when considering the amount of deformation undergoing in a mammography procedure. The B-spline deformation is an improvement over rigid or affine co-registration methods [33]. The advantage of the deformable registration has been recognized as a key element in breast analysis and has been successfully used in longitudinal studies [22]. Regarding digital subtraction, the differences in the X-ray projection, and image acquisition and digitizing artifacts may affect the detection of asymmetric patterns. Our results indicate that even in the presence of registration artifacts, the digital subtraction added information that was successfully incorporated during the feature selection process.

The B-Spline transformation algorithm, proposed for the bilateral mammogram registration presented, shows a clear improvement after the registration. Due to lack of temporal mammograms, temporal registration was not tested. Nevertheless, the methodology could be implemented in such task. However, the temporal registration should be re-optimized using a new set of parameters.

The enhanced images and texture maps enriched the feature set providing a four-fold increase in extracting features per patient, which were also incorporated in the final classification models. Regarding symmetry, the strategy of exploring bilateral symmetry has been explored by other researchers where a series of features (signal, texture, breast density, etc.) were

computed from each mammogram and the absolute difference between both breasts was obtained to measure breast tissue asymmetry, and used it to predict the likelihood of developing cancer [19]. We extended this idea by registering the left and right images using a deformable transformation, which increased the number of features per patient by 25%.

This study shows that healthy subjects, subjects with calcifications, and subjects with masses can accurately be classified through models generated via mammography registration and a feature selection methodology. The analysis of the feature selection strategy demonstrated that even when using a different approach for the feature selection strategy, the proposed methodology achieved similar results as the previously presented ones. Therefore, we can say that the methodology is robust to the feature selection strategy.

The methodology demonstrated that the image subtraction of registered images generates information that aids in the identification of subjects with lesions, such as malignant masses and calcifications. The methodology also incorporated the use of feature-based asymmetry into the CADx system. The combination performance achieved has the potential to be used to queue cases with a high chance of malignant findings, or may have the practical use of triaging mammograms in developing countries where there is a deficiency of expert readers.

## Author details

José María Celaya Padilla<sup>1\*</sup>, Cesar Humberto Guzmán Valdivia<sup>1</sup>, Jorge Issac Galván Tejada<sup>2</sup>, Carlos Eric Galván Tejada<sup>2</sup>, Hamurabi Gamboa Rosales<sup>2</sup>, Juan Rubén Delgado Contreras<sup>3</sup>, Antonio Martínez-Torteya<sup>4</sup>, Roberto Olivera Reyna<sup>2</sup>, Jorge Roberto Manjarrez Sánchez<sup>5</sup>, Francisco Javier Martínez Ruiz<sup>2</sup>, Idalia Garza-Veloz<sup>6</sup>, Margarita L. Martínez-Fierro<sup>6</sup>, Victor Treviño<sup>7</sup> and Jose Gerardo Tamez-Peña<sup>7</sup>

\*Address all correspondence to: jose.celaya@uaz.edu.mx

1 Autonomous University of Zacatecas/ CONACyT – Universidad Autónoma de Zacatecas (CONACyT – UAZ), Jardín Juárez, Centro, Zacatecas, Zacatecas, Mexico

2 Electric Engineering Department, Autonomous University of Zacatecas (UAZ), Jardín Juárez, Centro, Zacatecas, Zacatecas, Mexico

3 Superior Technical Institute of Zacatecas South (ITSZaS), Las lomas, Tlaltenango, Zacatecas, Mexico

4 Engineering Department, Monterrey University (UdeM), Morones Prieto Pte, Jesús M. Garza, San Pedro Garza García, Nuevo Leon, Mexico

5 Computer Engineering Systems Department, Superior Technical Institute of Jerez (ITSJ) Libramiento Fresnillo-Tepetongo, Fracc. Los Cardos, Jerez de García Salinas, Zacatecas, Mexico

6 Health Sciences Department, Human Medical School, Autonomous University of Zacatecas (UAZ), Jardín Juárez, Centro, Zacatecas, Zacatecas, Mexico

7 Bioinformatic group, Medical School, Monterrey Institute of Technology (ITESM), Eugenio Garza Sada, Monterrey, Nuevo Leon, Mexico



## References

- [1] World-Health. Cancer Fact sheet No. 297 [Internet]. 2015. Available from: <http://www.who.int/mediacentre/factsheets/fs297/en/index.html> [Accessed: 09/02/2015]
- [2] DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. CA: A Cancer Journal for Clinicians. 2014;**64**(1):52-62
- [3] Ng K, Muttarak M. Advances in mammography have improved early detection of breast cancer. Honk Kong College of Radiologist. 2003;**6**(3):126-131
- [4] Chan H-P, et al. Computer-aided classification of mammographic masses and normal tissue: Linear discriminant analysis in texture feature space. Physics in Medicine and Biology. 1995;**40**(5):857-876
- [5] D'orsi C, Bassett L, Berg W, Feig S, Jackson V, Kopans D. Breast Imaging Reporting and Data System: ACR BI-RADS-Mammography. Reston: American College of Radiology (ACR); 2003
- [6] Doi K. Current status and future potential of computer-aided diagnosis in medical imaging. The British Journal of Radiology. 2005;**78**(1):S3-S19
- [7] Sampat MP, Markey MK, Bovik AC, et al. Computer-aided detection and diagnosis in mammography. Handbook of Image and Video Processing. 2005;**2**(1):1195-1217
- [8] Dromain C, Boyer B, Ferre R, Canale S, Delaloge S, Balleyguier C. Computed-aided diagnosis (CAD) in the detection of breast cancer. European Journal of Radiology. 2013;**82**(3):417-423
- [9] Ramos-Pollán R, et al. Discovering mammography-based machine learning classifiers for breast cancer diagnosis. Journal of Medical Systems. 2012;**36**(4):2259-2269
- [10] Li M, Zhou Z-H. Improve computer-aided diagnosis with machine learning techniques using undiagnosed samples. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans. 2007;**37**(6):1088-1098
- [11] Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Computerized Medical Imaging and Graphics. 2007;**31**(4-5):198-211
- [12] Eadie LH, Taylor P, Gibson AP. A systematic review of computer-assisted diagnosis in diagnostic cancer imaging. European Journal of Radiology. 2012;**81**(1):e70-e76
- [13] Cheng H-D, Cai X, Chen X, Hu L, Lou X. Computer-aided detection and classification of microcalcifications in mammograms: A survey. Pattern Recognition. 2003;**36**(12):2967-2991
- [14] Zheng B, Sumkin JH, Zuley ML, Wang X, Klym AH, Gur D. Bilateral mammographic density asymmetry and breast cancer risk: A preliminary assessment. European Journal of Radiology. 2012;**81**(11):3222-3228
- [15] Scutt D, Lancaster GA, Manning JT. Breast asymmetry and predisposition to breast cancer. Breast Cancer Research. 2006;**8**(2):R14

- [16] Giger ML, Yin FF, Vyborny CJ. Comparison of bilateral-subtraction and single-image processing techniques in the computerized detection of mammographic masses. *Investigative Radiology*. 1993;**28**(6):473-481
- [17] Yin FF, Giger ML, Doi K, Metz CE, Vyborny CJ, Schmidt RA. Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images. *Medical Physics*. 1991;**18**(5):955-963
- [18] Tan M, Zheng B, Ramalingam P, Gur D. Prediction of near-term breast cancer risk based on bilateral mammographic feature asymmetry. *Academic Radiology*. 2013;**20**(12):1542-1550
- [19] Wang X, Lederman D, Tan J, Wang XH, Zheng B. Computerized prediction of risk for developing breast cancer based on bilateral mammographic breast tissue asymmetry. *Medical Engineering & Physics*. 2011;**33**(8):934-942
- [20] Wang X, Lederman D, Tan J, Wang XH, Zheng B. Computerized detection of breast tissue asymmetry depicted on bilateral mammograms: A preliminary study of breast risk stratification. *Academic Radiology*. 2010;**17**(10):1234-1241
- [21] Rodriguez-Rojas J, Garza-Montemayor M, Trevino-Alvarado V, Tamez-Pena JG. Predictive features of breast cancer on Mexican screening mammography patients. In: *Spie Medical Imaging. International Society for Optics and Photonics; Florida, USA*. 2013, pp. 867023-867023-9
- [22] Martí R, Díez Y, Oliver A, Tortajada M, Zwiggelaar R, Lladó X. Detecting abnormal mammographic cases in temporal studies using image registration features. In: *Breast Imaging. Springer; Gifu, Japan*. 2014. pp. 612-619
- [23] Miller P, Astley SM. Detection of breast asymmetry using anatomical features. In: *IS&T/ SPIE's Symposium on Electronic Imaging: Science and Technology. International Society for Optics and Photonics; Los Angeles California, USA*. 1993. pp. 433-442
- [24] Miller P, Astley S. Automated detection of breast asymmetry using anatomical features. *State of the Art in Digital Mammographic Image Analysis, Series in Machine Perception and Artificial Intelligence*. 1994;**9**:247-261
- [25] Lau TK, Bischof WF. Automated detection of breast tumors using the asymmetry approach. *Computers and Biomedical Research*. 1991;**24**(3):273-295
- [26] Ferrari RJ, Rangayyan RM, Desautels JL, Frère AF. Analysis of asymmetry in mammograms via directional filtering with Gabor wavelets. *IEEE Transactions on Medical Imaging*. 2001;**20**(9):953-964
- [27] Suri JS, Rangayyan RM. *Recent Advances in Breast Imaging, Mammography, and Computer-Aided Diagnosis of Breast Cancer*. SPIE press; Washington, USA. 2006. pp. 488-525
- [28] Díez Y, et al. Revisiting intensity-based image registration applied to mammography. *IEEE Transactions on Information Technology in Biomedicine*. 2011;**15**(5):716-725

- [29] Mendez AJ, Tahoces PG, Lado MJ, Souto M, Correa JL, Vidal JJ. Computer-aided diagnosis: Automatic detection of malignant masses in digitized mammograms. *Medical Physics*. 1998;**25**:957
- [30] Celaya-Padilla JM, Rodriguez-Rojas J, Trevino V, Tamez-Pena JG. Local image registration a comparison for bilateral registration mammography. In: Presented at the International Seminar on Medical Image Processing and Analysis; Mexico, DF; 2013
- [31] Celaya-Padilla JM, Rodriguez-Rojas J, Galván-Tejada JI, Martínez-Torteya A, Treviño V, Tamez-Peña JG. Bilateral image subtraction features for multivariate automated classification of breast cancer risk. In: SPIE Medical Imaging. International Society for Optics and Photonics; San Diego, CA, USA. 2014. pp. 90351T-90351T-7
- [32] Celaya-Padilla J, Martinez-Torteya A, Rodriguez-Rojas J, Galvan-Tejada J, Treviño V, Tamez-Peña J. Bilateral image subtraction and multivariate models for the automated triaging of screening mammograms. *BioMed Research International*. 2015;**2015**:1-12
- [33] Yin FF, Giger ML, Doi K, Vyborny CJ, Schmidt RA. Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral subtraction technique. *Medical Physics*. 1994;**21**(3):445-452
- [34] Heath M, Bowyer K, Kopans D, Moore R, Kegelmeyer P. The digital database for screening mammography. In Proceedings of the 5th International Workshop on Digital Mammography. San Diego, CA, USA. 2000. pp. 212-218
- [35] Guo Y, Suri J, Sivaramakrishna R. Image registration for breast imaging: A review. In: 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. IEEE; Toronto, Canada. 2006. pp. 3379-3382
- [36] Marias K, Behrenbruch C, Parbhoo S, Seifalian A, Brady M. A registration framework for the comparison of mammogram sequences. *IEEE Transactions on Medical Imaging*. 2005;**24**(6):782-790
- [37] Kok-Wiles SL, Brady M, Highnam R. Comparing mammogram pairs for the detection of lesions. In *Digital Mammography*. Springer; Nijmegen, Netherlands. 1998. pp. 103-110
- [38] Hill DL, Hawkes DJ. Across-modality registration using intensity-based cost functions. In *Handbook of Medical Imaging*. Academic Press, Inc; Orlando, FL, USA. 2000. pp. 537-553
- [39] Mattes D, Haynor DR, Vesselle H, Lewellyn TK, Eubank W. Nonrigid multimodality image registration. In: *Medical Imaging 2001*. International Society for Optics and Photonics; San Diego, CA, USA. 2001. pp. 1609-1620
- [40] West J, et al. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*. 1997;**21**(4):554-568
- [41] Rosenfeld A. *Multiresolution Image Processing and Analysis*. Springer Science & Business Media; Heidelberg, Germany. 2013
- [42] Rangayyan RM, Banik S, Desautels JL. Computer-aided detection of architectural distortion in prior mammograms of interval cancer. *Journal of Digital Imaging*. 2010;**23**(5):611-631

- [43] Ibanez L, Schroeder W, Ng L, Cates J. The ITK Software Guide. Kitware; New York, USA. 2003
- [44] Beasley TM, Erickson S, Allison DB. Rank-based inverse normal transformations are increasingly used, but are they merited? *Behavior Genetics*. 2009;**39**(5):580-595
- [45] Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1996;**58**:267-288
- [46] Friedman J, Hastie T, Tibshirani R, Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, Innsbruck, Austria, 2010;**33**(1):1-22
- [47] Ganesan K, Acharya UR, Chua KC, Min LC, Abraham KT. Pectoral muscle segmentation: A review. *Computer Methods and Programs in Biomedicine*. 2013;**110**(1):48-57
- [48] Raba D, Oliver A, Martí J, Peracaula M, Espunya J. Breast segmentation with pectoral muscle suppression on digital mammograms. In: *Pattern Recognition and Image Analysis*. Springer; Berlin, Heidelberg, Germany. 2005. pp. 471-478