Doctoral Thesis

Case-level detection of mammographic masses

Meritxell Tortajada Giménez

2017

Doctoral Thesis

Case-level detection of mammographic masses

Meritxell Tortajada Giménez

2017

Doctoral Programme in Technology

Supervised by:

Jordi Freixenet and Robert Martí

Work submitted to the University of Girona in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Publications

A list of publications of the author for the PhD candidacy is given below, ordered according to their topic.

Publications related to the mass detection and false positive reduction proposals:

- [**IWDM 2014**] R. Martí, Y. Díez, A. Oliver, M. Tortajada, R. Zwiggelaar, and X. Lladó. Detecting Abnormal Mammographic Cases in Temporal Studies Using Image Registration Features. International Workshop on Digital Mammography, LNCS 8539, pp 612-619. Gifu, Japan. June 2014.

- [**IBPRIA 2013**] Y. Díez, M. Tortajada, S. Ganau, L. Tortajada, M. Sentís, and R. Martí. Demons Methods for Digital Mammography Registration. Iberian Conference on Pattern Recognition and Image Analysis, LNCS 7887, pp 253-260. Madeira, Portugal. June 2013.

- [**ECR 2013**] M. Tortajada, A. Oliver, Y. Díez, J. C. Vilanova, J. Martí, J. Freixenet, S. Ganau, L. Tortajada, R. Aguilar, M. Vives, J. Fernández-Bayó, M. Sentís, and R. Martí. CAD system using case-based information. European Congress of Radiology. Vienna, Austria. March 2013.

- [**RSNA 2010**] M. Tortajada, A. Oliver, Y. Díez, R. Martí, J. C. Vilanova, and J. Freixenet. Integrating bilateral information in the eigendetection CAD

approach. Scientific Assembly and Annual Meeting of the Radiological Society of North America. Chicago, Illinois. December 2010.

- [**MICCAT 2010**] M. Tortajada, A. Oliver, Y. Díez, R. Martí, and J. Freixenet. Bilateral comparison for improving breast masses detection. Medical Image Computing in Catalunya. Girona, Spain. October 2010.

- [**EMBC 2010**] M. Tortajada, A. Oliver, Y. Díez, R. Martí, J. C. Vilanova, and J. Freixenet. Improving a CAD system using bilateral information. IEEE Conference of the Engineering in Medicine and Biology Society, pp 5054-5057. Buenos Aires, Argentina. September 2010.

Publications related to the breast tissue characterization:

- [**JDI 2015**] A. Oliver, M. Tortajada, X. Lladó, J. Freixenet, S. Ganau, L. Tortajada, M. Vilagran, M. Sentís, and R. Martí. Breast Density Analysis Using an Automatic Density Segmentation Algorithm. Journal Digital Imaging, vol 28, pp 604612. February 2015. IF 1.406, Q3(91/124) RNMMI.

- [**IWDM 2012**] M. Tortajada, A. Oliver, R. Martí, M. Vilagran, S. Ganau, L. Tortajada, M. Sentís, and J. Freixenet. Adapting breast density classification from digitized to full-field digital mammograms. International Workshop on Digital Mammography, LNCS 7361, pp 561-568. Philadelphia, Pennsylvania. July 2012.

- [**MICCAT 2011**] M. Tortajada, A. Oliver, R. Martí, J. Freixenet, S. Ganau, L. Tortajada, and M. Sentís. Breast tissue density classification in digital mammography. Medical Image Computing in Catalunya. Barcelona, Spain. October 2011.

Other publications related to mammographic works:

- [**CBM 2014**] M. Tortajada, A. Oliver, R. Martí, S. Ganau, L. Tortajada, M. Sentís, J. Freixenet, and R. Zwiggelaar. Breast peripheral area correction in digital mammograms. Computers in Biology and Medicine, vol 50, pp 32-40. March 2014. IF 1.240, Q3(64/102) CSIA.

- [**KBS 2012**] A. Oliver, A. Torrent, X. Lladó, M. Tortajada, L. Tortajada, M. Sentís, J. Freixenet, and R. Zwiggelaar. Automatic microcalcification and cluster detection in digital and digitised mammograms. Knowledge-Based Systems, vol 28, pp 68-75. 2012. IF 2.422, Q1(15/111) CSAI.

- [**IWDM 2010**] A. Oliver, A. Torrent, M. Tortajada, X. Lladó, M. Peracaula, L. Tortajada, M. Sentís, and J. Freixenet. A boosting based approach for automatic micro-calcification detection. International Workshop on Digital Mammography, LNCS 6136, pp 251-258. Girona, Spain. June 2010.

- [**IWDM 2008**] M. Tortajada, R. Martí, J. Freixenet, J. Fernández, M. Sentís. Image Correction and Reconstruction for Breast Biopsy. International Workshop on Digital Mammography, LNCS 5116, pp 545-552. Tucson, Arizona. July 2008.

- [**CARS 2007**] J. Fernández-Bayó, M. Sentís, M. Tortajada, M. Chmeissani, M. Lozano, S. Ganau, L. Tortajada, G. Blanchot, J. García, F. Kainberger, M. Maiorino, R. Martínez, J. P. Montage, G. Pellegrini, C. Puigdengoles, M. Ullán. X-ray machine for general radiology and mammography based on room temperature Solid State detector coupled to photon-counting electronics. Evaluation of results. Computer Assisted Radiology and Surgery 21st

International Congress and Exhibition, vol 2, 1 Suppl., pp S9-S11. Berlin, Germany. June 2007.

- [**NIMPRA 2006**] G. Blanchot, M. Chmeissani, A. Díaz, F. Díaz, J. Fernández, E. García, J. García, F. Kainberger, M. Lozano, M. Maiorino, R. Martínez, J.P. Montagne, I. Moreno, G. Pellegrini, C. Puigdengoles, M. Sentís, L. Terés, M. Tortajada, M. Ullán. Dear-Mama: A photon counting X-ray imaging project for medical applications. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol 569, pp 136-139. 2006. IF 1.185 Q1(5/32) NST.

- [**NSS-MIC 2006**] R. Martínez, G. Blanchot, M. Chmeissani, A. Díaz, F. Díaz, E. García, J. García, F. Kainberger, I. Moreno, G. Pellegrini, C. Puigdengoles, M. Sentís, L. Terés, M. Tortajada, M. Ullán. Detection of Early Markers in Mammography Project. IEEE Nuclear Science Symposium and Medical Imaging Conference. San Diego, California. October 2006.

- [**VERTEX 2005**] G. Blanchot, M. Chmeissani, A. Díaz, F. Díaz, J. Fernández, E. García, J. García, F. Kainberger, M. Lozano, M. Maiorino, R. Martínez, J. P. Montagne, I. Moreno, G. Pellegrini, C. Puigdengoles, M. Sentís, M. Tortajada, M. Ullán. Dear-Mama: A Photon Counting X-ray Imaging Project for Medical Applications. International Workshop on Vertex Detectors. Nikko, Japan. November 2005.

Other works related to medical imaging:

- [**SMIT 2008**] M. Gruber, P. Homolka, M. Tortajada, R. Martínez, M. Pretterklieber, F. Kainberger. Musculoskeletal imaging by a prototype of a digital

imaging x-ray machine based on direct capture of X-ray photons with pixel detectors coupled to photon counting readout electronics. International Conference of the Society for Medical Innovation and Technology. Vienna, Austria. August 2008.

- **[ECR 2008]** M. Gruber, P. Homolka, M. Tortajada, R. Martínez, M. Pretterklieber, F. Kainberger. Musculoskeletal imaging by a prototype of a digital imaging x-ray machine based on direct capture of X-ray photons with pixel detectors coupled to photon counting readout electronics. European Congress of Radiology. Vienna, Austria. March 2008.

List of Abbreviations

ACR American College of Radiology

BI-RADS Breast Imaging Reporting and Data System

CAD Computer-Aided Detection/Diagnosis

CC Cranio-Caudal

CT Computed Tomography

FFDM Full-Field Digital Mammography

FP False Positive

KNN K-Nearest Neighbors

MI Mutual Information

MLO Medio-Lateral Oblique

MR MultiResolution

MRI Magnetic Resonance Imaging

RF Random Forest

ROI Region of Interest

SSD Sum of Squared Differences

SVM Support Vector Machine

US Ultrasound

List of Figures

List of Tables

For Lidia, my North Star.

Acknowledgements

Being true to myself, I feel comfortable saying thanks to everyone who crossed with me during my life. What I am and what I get is the weighted sum of the contributions of everyone who share with me time, effort, words, knowledge, laughs and tears. So, I should name everybody from P4, from CIRS, from UDIAT, from my degree, from my high school, from CAB, from MUEC, from La Verneda, from Veguillas and also more than half the people from Barcelona, Girona, Granollers and Sabadell. And of course, all members of my blood family and of my family-in-law. It's a extremely long list, difficult to detail and still, it's possible to forget someone. So, I prefer to do not get lost in names, I want to say thank you to those who make my life easier and those who make it more difficult,

'Cause it makes me that much stronger
Makes me work a little bit harder
It makes me that much wiser
Made me learn a little bit faster
Made my skin a little bit thicker
Makes me that much smarter
So thanks for making me a fighter.

Contents

Resum

L'objectiu principal d'aquesta tesi és la detecció automàtica de masses en imatges mamogràfiques digitals fent servir tota la possible informació mamogràfica del pacient, això inclou informació bilateral, temporal i ipsilateral. Com a primer pas, les imatges digitals són preprocessades per tal de millorar la seva qualitat abans de començar a treballar amb elles dintre del marc de la detecció pròpiament dit. El nou mètode de millora de la qualitat de mamografies digitals s'aplica per a compensar la reducció del gruix mamari a la part perifèrica de la mama. A continuació, per a obtenir la informació bilateral i temporal del pacient, es fa servir el registre afí seguit del registre B-Splines, i la informació obtinguda s'incorporarà a la part de detecció del procés global. El registre aplicat és el que ha obtingut els millors resultats a nivell d'utilitat i d'usabilitat d'entre els diferents mètodes analitzats. Finalment, la correspondència entre les vistes CC i MLO basada en l'ús de línies epipolars és el mètode que s'utilitza durant la part de reducció de falsos positius. La informació ipsilateral permet distingir entre masses reals (quan existeix correspondència entre lesions trobades a les vistes CC i MLO) i masses falses (en cas contrari). A més, per tal d'afegir informació sobre la densitat de la mama analitzada, diferents mètodes d'avaluació de densitat mamària, tant qualitatius com quantitatius, han estat proposats i avaluats. Els resultats inicials mostren que el CAD multi-imatge incrementa la seva sensibilitat i redueix el nombre de falsos positius respecte els resultats obtinguts pel CAD uni-imatge.

Resumen

El objetivo principal de esta tesis es la detección automática de masas en imágenes mamográficas digitales, usando para ello toda la información mamográfica disponible del paciente, esto incluye información bilateral, temporal e ipsilateral. Como primer paso, las imágenes digitales son preprocessadas para mejorar su calidad antes de empezar a trabajar con ellas dentro del marco de la detección propiamente dicho. El nuevo método de mejora de la calidad de mamografías digitales es aplicado para compensar la reducción del grosor mamario en la parte periférica de la mama. A continuación, para obtener la información bilateral y temporal del paciente se usa el registro afín seguido del registro B-Splines y la información obtenida se incorporará en la parte de detección del proceso global. El registro aplicado es el que ha obtenido los mejores resultados a nivel de utilidad y de usabilidad de entre los diferentes métodos analizados. Finalmente, la correspondencia entre las vistas CC y MLO basada en el uso de líneas epipolares es el método usado durante la parte de reducción de falsos positivos. La información ipsilateral permite distinguir entre masas reales (cuando existe correspondencia entre lesiones halladas en las vistas CC y MLO) y masas falsas (en caso contrario). Además, con el fin de añadir información sobre la densidad de la mama analizada, diferentes métodos de evaluación de densidad mamaria, tanto cuantitativos como cualitativos, han sido propuestos y evaluados. Los resultados iniciales muestran que el CAD multi-imagen incrementa su sensibilidad y reduce el número de falsos positivos respecto a los resultados obtenidos por el CAD uni-imagen.

Abstract

This thesis is focused on the automatic detection of masses in FFDM images by using case-level information which includes bilateral, temporal and/or ipsilateral information. As a first step, FFDM images are preprocessed to improve image quality before working on the proper detection framework. A novel enhancement method is applied to compensate the thickness reduction in peripheral edges of the breast in FFDM. Following, B-Splines image registration with Affine initialisation is used to obtain bilateral and temporal information that is incorporated in the detection stage of the whole process. This registration approach is considered the optimal one that provides useful and usable case-level information among several investigated registration methods. Finally, CC/MLO correspondence approach based on using curved epipolar lines is used in the FP stage. Ipsilateral information allows to distinguish between real (when CC/MLO lesion correspondence exists) and non-real (when there is no CC/MLO lesion correspondence) masses. Furthermore, in order to add breast density information to the detection process, different methods for breast density assessment are analysed. Both, qualitative and quantitative methods are proposed and evaluated. Initial results show a better performance of the multi-image CAD approach relative to the single-image CAD approach. Sensitivity increases and the number of FPs is reduced.

Chapter 1

Introduction

The aim of this research work is the development of an automatic tool to detect masses in digital mammographic images by including case level information. This initial chapter is an introduction to breast cancer and the different imaging techniques for detecting it. As well as an introductory explanation about single and multiple view computer aided detection systems. In addition, a general overview of this thesis is provided.

1.1 Breast Cancer

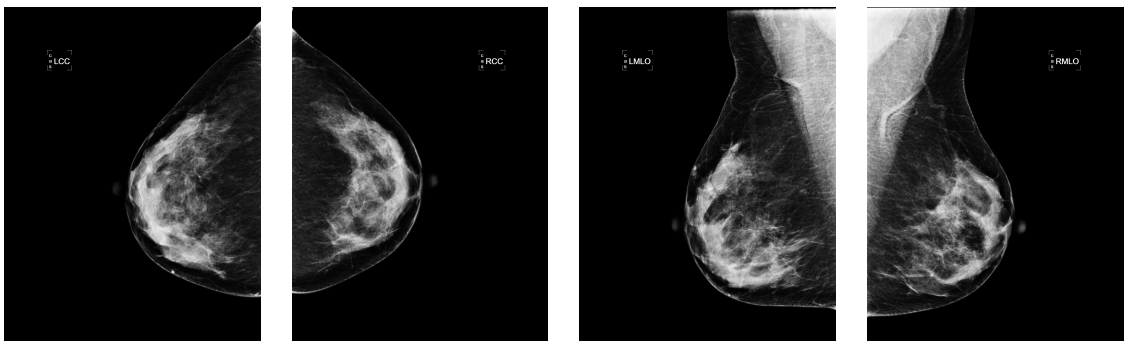
Breast cancer is a disease characterised by the uncontrolled growing of cells in the breast that attack the neighbouring healthy cells. Breast cancer is the second most frequent malignant tumour around the world, exceeded only by lung cancer. In the female population, breast cancer is by far, the most common cancer with an estimated 1.67 million worldwide new cancer cases diagnosed in 2012 (25% of all cancers). It is the most common cancer in women both in more and less developed regions with slightly more cases in less developed (883.000 cases) than in more developed (794.000 cases) regions [225].

Breast cancer ranks as the fifth cause of death from cancer overall (522.000 deaths) and while it is the most frequent cause of cancer death in women in less developed regions (324.000 deaths, 14.3% of total), it is now the second cause of cancer death in more developed regions (198.000 deaths, 15.4%) after lung cancer.

The range in mortality rates between world regions is less than that for incidence because of the more favorable survival of breast cancer in (high-incidence) developed regions [225]. Nowadays, breast cancer survival rate is growing thanks to factors such as mammographic screening programs (for an early detection) or more efficient treatments [85, 310].

1.2 Breast Cancer Detection Imaging Techniques

Mammography (x-ray image of the breast) is the most common method used for breast cancer detection. Mammography is used in breast cancer screening programs due to capacity of depicting cancers at a smaller size and earlier stage [275]. Current commercial machines can detect lesions up to 70μm using low-dose radiation. During an usual screening mammographic exam 4 images are acquired, CC and MLO views from left and right breasts. CC view images the breast from above and MLO view is taken from an oblique or angled view (usually 45°) (see Fig. 1.1). Although recently, trying to improve the performance of mammography, digital breast tomosynthesis (DBT) has been developed. DBT is a 3D imaging technology that acquires images of the compressed breast at multiple angles during a short scan. Images are then reconstructed into a series of 1mm thick high-resolution slices. In addition, CC and MLO images are synthetically reconstructed from the 3D dataset. These 2D reconstructed images are known as C-Views [281, 280, 361, 243].

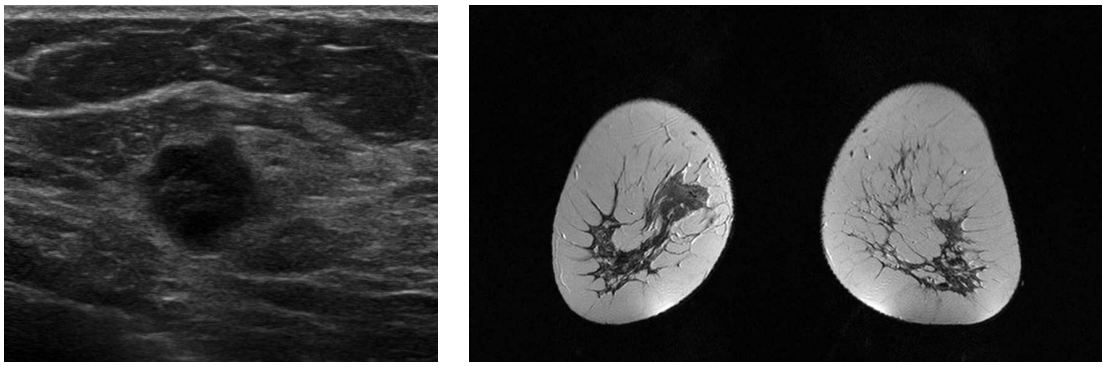


(a) (b) (c) (d)

Figure 1.1: Complete mammographic screening case: (a) CC view of the right breast, (b) CC view of the left breast, (c) MLO view of the right breast, and (d) MLO view of the left breast.

However, not all breast cancer can be detected with x-rays and of course, mammography has limitations, so other imaging techniques have to be used. **Breast US** (see Fig. 1.2(a)) is becoming a necessary technique in the screening environment, especially in dense breast cases. Recently, the automated breast ultrasound (ABUS) has been presented as an emerging approach that allows improved consistency and reproducibility of images, minimizes operator dependence, and aids with inclusion of the whole breast [46].

Breast MRI (see Fig. 1.2(b)) is another technique that can be used for breast cancer detection, although in most cases is a complement for mammography. Nevertheless, in high risk women with BRCA1, BRCA2 (genetic mutation) is used as a detection method [170] . In addition to the aforesaid breast imaging techniques, optical imaging, breast thermography, positron emission mammography, electric impedance imaging or breast CT are also techniques that can be used for breast cancer detection and/or evaluation [288, 101].



(a) (b)

Figure 1.2: Example of breast imaging techniques: (a) breast US, and (b) breast MRI.

1.2.1 Mammographic Abnormalities

According to BI-RADS® [275] there are several important findings in mammograms. Calcifications, asymmetries, architectural distortions, masses, intramammary lymph nodes, skin lesions or solitary dilated ducts are included in the aforesaid findings group.

Calcifications

Breast calcifications are calcium deposits within the breast that appear in mammograms as white spots similar to grains of salt. There are two types of breast calcifications according to their size: macrocalcifications and microcalcifications. Macrocalcifications look like large white dots or dashes and are almost always noncancerous. On the contrary, microcalcifications are very fine white specks and their irregular clustering is usually a sign of cancer [187].

Asymmetries

Altough both breasts are externally different, at an internal level they are structurally similar, so the fact of detecting an asymmetry of the breast parenchyma between left and right breasts may be indicative of the presence of a lesion or the development of a cancer [187]. Depending on the type of asymmetry, this is visible on only one mammographic projection or on more than one projection, although all types of asymmetries have concave-outward borders and usually are seen to be interspersed with fat tissue [275].

Architectural distortions

Architectural distortions are characterised by the distortion of the parenchyma with no definite mass visible. For mammography, this includes thin straight lines or spiculations radiating from a point, and focal retraction, distortion, or straightening at the anterior or posterior edge of the parenchyma. Architectural distortions may also be associated with asymmetry or calcifications. In the absence of an appropriate history of trauma or surgery, architectural distortions are suspicious for malignancy [275].

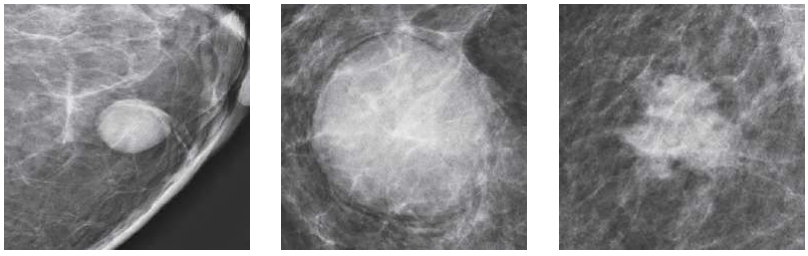
Masses

A lesion is classified as mass when is 3-dimensional and occupies space. Masses are seen on two different mammographic projections and have completely or partially convex-outward borders and radiologically appears denser in the center than at the periphery. If a potential mass is seen only on a single projection, it should be called an asymmetry until its 3-dimensionality is confirmed [275].

Masses can be classified by shape (see Fig. 1.3), by margin (see Fig. 1.4) or by density (see Fig. 1.5). Shape describes the roundness of the mass, while

margin characterises the edge or border of the lesion. Even though, both descriptors (margin and shape) are important predictors of whether a mass is benign or malignant. Finally, density is used to define the x-ray attenuation of the mass relative to the expected attenuation of an equal volume of normal fibroglandular breast tissue. Most breast cancers that present as a mass are of equal or higher density than an equal volume of normal fibroglandular tissue. However, breast density is a subjective evaluation that is least reliable among the mammographic features of masses (i.e., shape and margin) [275].

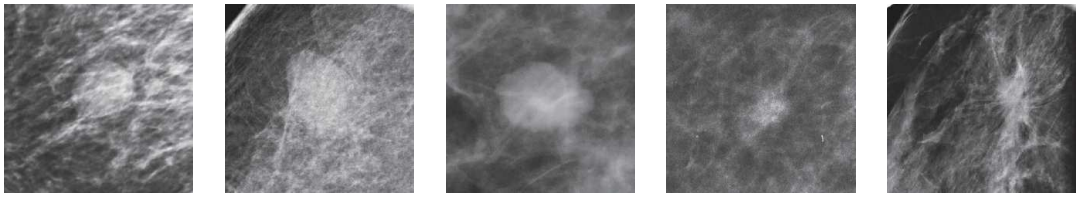
Masses by shape are classified in oval, round and irregular (see Fig. 1.3). An oval mass is elliptical or egg-shaped. A round mass is spherical, ball-shaped, circular, or globular in shape. Finally, the shape of an irregular mass is neither round nor oval [275].



(a) (b) (c)

Figure 1.3: Example of masses by shape: (a) oval, (b) round, and (c) irregular. Images extracted from $ACR\ \ BI-RADS^{®}Atlas$ [275].

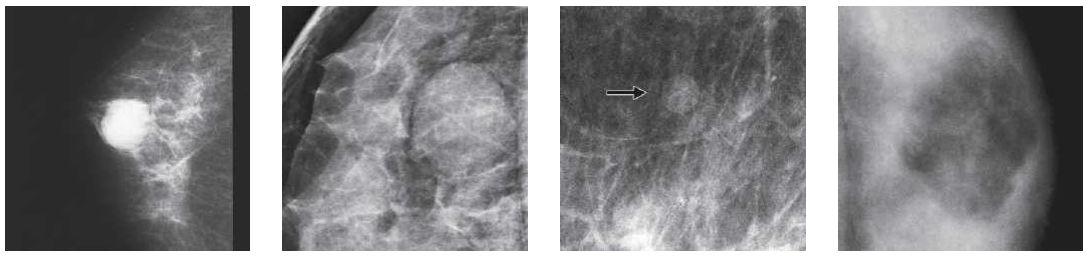
Masses by margin are classified in circumscribed, obscured, microlobulated, indistinct or spiculated (see Fig. 1.4). The margin of circumscribed masses is sharply demarcated with an abrupt transition between the lesion and the surrounding tissue. An obscured margin is one that is hidden by superimposed or adjacent fibroglandular tissue. The margin of microlobulated masses is characterised by short cycle undulations. In indistinct masses, there is no clear demarcation of the entire margin, or of any portion of the margin, from the surrounding tissue. Finally, spiculated mass margin is characterised by lines radiating from the mass [275].



(a) (b) (c) (d) (e)

Figure 1.4: Example of masses by margin: (a) circumscribed, (b) obscured, (c) microlobulated, (d) indistinct, and (e) spiculated. Images extracted from *ACR BI−RADS®Atlas* [275].

Masses by density are classified in high-density, equal-density, low-density and fat-containing (see Fig. 1.5). When the x-ray attenuation of the mass is greater than the expected attenuation of an equal volume of fibroglandular breast tissue is considered as high-dense. While, when x-ray attenuation of the mass is the same or lower as the expected attenuation of an equal volume of fibroglandular breast tissue are considered as equal-density and low-density respectively. Finally, fat-containing masses category includes all masses containing fat, such as oil cyst, lipoma, or galactocele, as well as mixed-density masses such as hamartoma [275].



(a) (b) (c) (d)

Figure 1.5: Example of masses by margin: (a) high density, (b) equal density, (c) low density, and (d) fat-containing. Images extracted from *ACR BI − RADS®Atlas* [275].

1.3 CAD Systems

Due to the difficulty of finding lesions in mammograms, computerized tools were and still are being developed to help radiologists in the evaluation of mammograms. This

is the case of computer aided detection (CADe) mammographic systems and computer aided diagnosis (CADx) mammographic systems. CADe systems indicate the presence of possible abnormalities whereas CADx systems classify potential lesions into malignant or benign often in terms of malignancy likelihood [99].

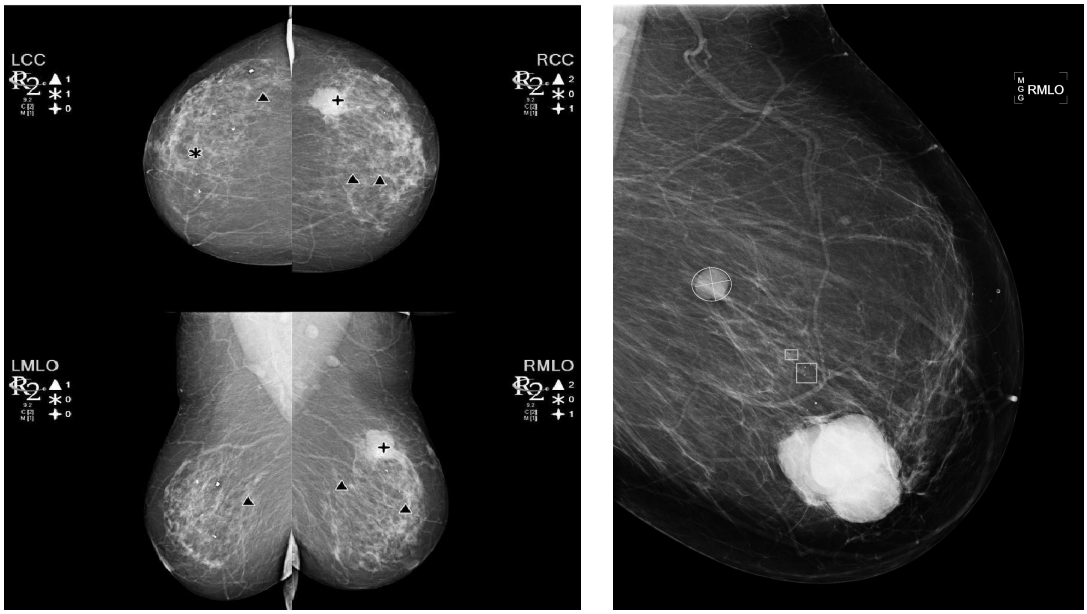
CADe systems have been used by radiologists in their daily work for more than 15 years, although the way of using this information can vary. One mode is to read mammographic images without displaying CADe marks and then compare the results with the output of CADe systems before taking a final decision. Radiologists decide to discard lesions detected in the initial read or take into account lesions overlooked in the initial read. But this mode probably increases the reading time. To reduce the reading time, another way of working is to display CADe marks first and then take the final decision, but it is uncertain if lesions may be missed when CADe does not prompt to the radiologists. Then, a reasonable solution for solving this possible effect is the increase of the sensitivity of the general CADe systems [99].

CADe mammographic systems search for breast abnormalities such as microcalcifications, masses, architectural distortions or asymmetries, although most of the works are focused on detecting clustered microcalcifications and masses [255]. Due to the difficulty of distinguishing masses from the surrounding breast parenchyma, the sensitivity of CADe systems for detecting microcalcifications is higher than the ones for detecting masses, although the number of FP is higher than desired in both cases. Current research works are focused on both, increasing the sensitivity and reducing the number FP of CADe systems.

CADx systems have been proposed to help radiologists in the discrimination of benign and malignant mammographic lesions and so, to improve the positive predictive value (biopsies positives for cancer / biopsies tested) of mammogram interpretation. The task of distinguish between benign and malignant lesions is usually considered as a two-class classification problem. Most CADx approaches start with the delimitation of a ROI that contains the lesion to classify. Then, systems have a four-stage process: lesion segmentation, feature extraction, feature selection, and lesion classification. CADx systems are slowly being introduced in the clinical workflow as CADe systems begin to give more information beyond just localization [99, 80].

1.3.1 Commercial CAD systems

There are various commercial CAD systems with Food and Drug Administration (FDA) approval [226]: the AccuDetect Mammography CAD, the Kodak Mammography CAD Engine, the M-Vu CAD System, the R2 ImageChecker CAD or the iCAD Second Look. **R2 ImageChecker CAD** [137] was the first CAD mammography system approved by FDA. ImageChecker analyses 2D digital and digitised mammographic images and marks suspicious masses (asterisk), clusters of calcifications (triangle) and masses and calcifications at the same location (four points of a compass) (see Fig. 1.6(a)). **iCAD Second Look** [142] is another remarkable CAD system. Second Look detects potential microcalcifications (rectangle) and masses (ellipse) (see Fig. 1.6(b))



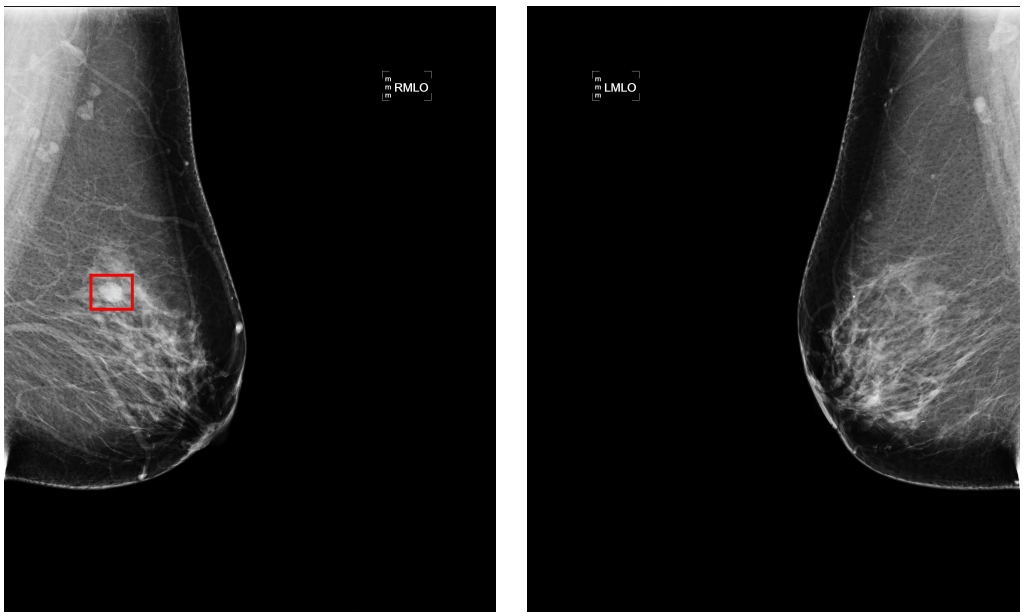
(a) (b)

Figure 1.6: Example of commercial CAD mammographic systems: (a) R2 ImageChecker CAD, and (b) iCAD Second Look.

1.3.2 Multi-Image CAD techniques

An open research line for improving the results of CAD systems is mimicking radiologists' way of reading mammograms and analysing multiple images of the same patient to have a global idea of the detected abnormalities. Due to the fact that most CAD schemes use the information of a single mammogram for detecting abnormalities, adding case-level information to CAD systems is presented as a promising approach [354, 223]. This case level-information can be used to detect real lesions or to discard false lesions (FPs).

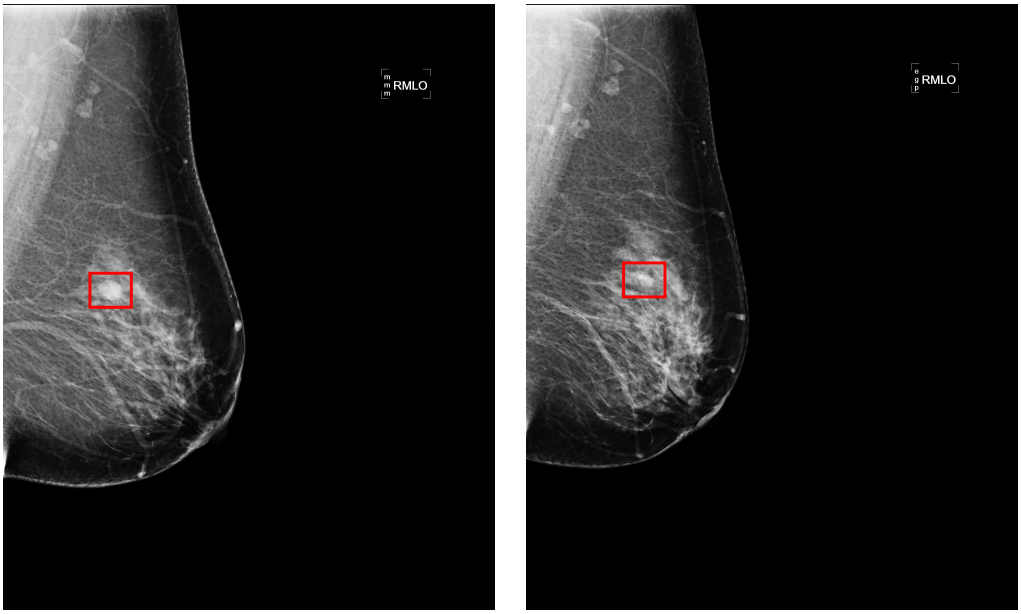
Based on the idea that both breasts are structurally similar [168], comparison of the same view of left and right breasts (bilateral comparison) can allow finding differences between breasts that are indicative of lesions (see Fig. 1.7) and can also allow finding similarities between corresponding regions that are indicative of normality.



(a) (b)

Figure 1.7: (a) Right MLO mammogram with lesion squared in red compared with (b) left MLO mammogram.

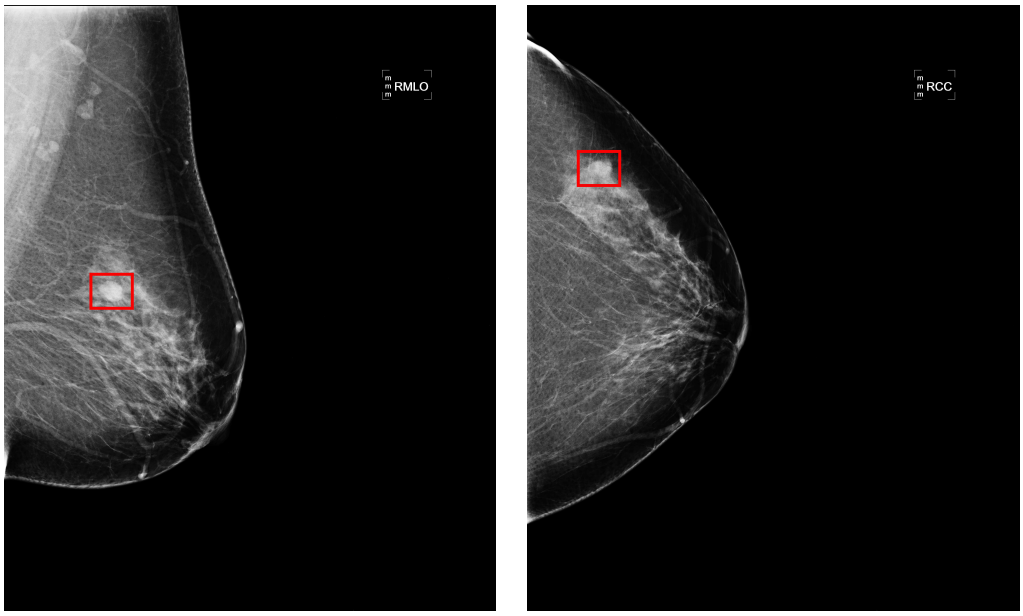
Current and prior homonym images examination (temporal comparison) can be used for analysing breast changes that could be considered as cancerous sign [342]. Abnormalities changes in size or shape (see Fig. 1.8) or the current presence of an abnormality are indicative of lesion.



(a) (b)

Figure 1.8: (a) Current right MLO mammogram with lesion squared in red compared with (b) prior right MLO mammogram with the same lesion squared in red.

The comparison of CC and MLO projections of the same breast (ipsilateral comparison) can allow matching suspicious findings [21]. When the same abnormality is present in CC and MLO corresponding regions is regarded as true lesion (see Fig. 1.9), however when an abnormality is located in only one view is considered as false lesion.



(a) (b)

Figure 1.9: (a) Right MLO mammogram with lesion squared in red compared with (b) right CC mammogram with the same lesion squared in red.

1.4 Objectives of this Thesis

The main objective of this thesis is the proposal of a framework for detecting masses independently of its shape, margin or density in FFDM images by using not only information from each mammogram but also case-level information.

In order to address the main goal of this thesis we divide the main objective into a set of sub-objectives. Sub-objectives are either directly or indirectly related with CAD systems.

FFDM images usually have an overexposed region in the outer part of the breast due to the thickness reduction in the periphery of the breast [150]. To compensate this effect, a determining sub-goal of this thesis is the development of an algorithm for **breast peripheral image enhancement**. In addition to breast peripheral overexposition issue, breast density is also related with image quality. FFDM accuracy is really dependent of the amount of breast dense tissue [48], in this sense, another significant sub-goal is the development of an algorithm for **breast density assessment**.

To obtain the case-level information several approaches can be used. The majority of **bilateral and temporal** strategies are based on estimating a geometric transformation between FFDM images to obtain a spatial correspondence between them [354]. So, an important sub-goal is the study of the different **image registration** algorithms to determine the optimal alignment between either bilateral or temporal images. Other approaches are based on defining a coordinate system to determine the relative coordinates of a ROI in its corresponding (bilateral/ipsilateral) view [229]. In this direction, another sub-goal is to develop an algorithm for **ipsilateral correspondence**.

Starting from a single-image CAD system that detects masses independently of its shape, margin or density, the key question is where to add the generated case-level information to transform the single-image CAD system into a multi-image one. This information can be added in the detection stage or in the FPs reduction stage. So, two required sub-goals are the analysis of using **case-level information** in both steps, **detection** and **FPs reduction**.

1.5 Overview of this Thesis

Even though the objective of this thesis is the development a mammographic CAD system that uses case-level information to detect masses, there are several issues that need to be considered previously: the necessity of a preprocessing stage before image analysis, the assessment of types and quantities of breast tissue or how to obtain this case-level information.

This thesis is organised as follows. Chapter 2 describes our novel approach for FFDM peripheral enhancement. In this thesis two types of images are used, digitised

and digital. Digitised images do not need a previous enhancement process however, digital images need a preprocessing step to deal with the overexposed area in the breast periphery. After overexposed area delimitation, the correction algorithm is applied and enhancement results are evaluated, showing the beneficial effects in the image quality.

Related also with the quality of the data provided in mammographic images, it is important to identify the amount of dense tissue in each mammogram, since the sensitivity of mammography is reduced in dense breasts. Chapter 3 presents an study about breast tissue characterisation. Breast is composed by fatty and dense tissue. Different strategies for breast segmentation according tissue types are evaluated. Finally both, qualitative and quantitative analysis are used to classify mammograms according the amount of dense tissue.

Another key point in the multi-image CAD development is the discussion about obtaining the case-level information. In Chapter 4 we analyse different known registration algorithms, as well as some registration methods combinations and multiresolution approaches. Registration results are evaluated in order to find the best approaches that could provide useful and suitable not only bilateral but also temporal information to our multi-image CAD system.

Chapter 5 presents a CAD scheme that includes bilateral, temporal and ipsilateral information. Based on our previously developed single-image digitised CAD mammographic system, we develop a multi-image FFDM CAD system. Single CAD parameters are tuned to adapt the system to digital inputs. Moreover, bilateral and temporal information is added in the detection framework and ipsilateral information is used during the FP reduction step. Results when using or not case-level information are compared.

Finally, the thesis concludes with Chapter 6. In this final chapter, conclusions are extracted and the future research work is described.

Chapter 2

Breast Peripheral Enhancement

During mammographic acquisitions, the breast is compressed with a tilting compression paddle, hence the breast thickness is non uniform across the mammogram, being thinner in the periphery and thus overexposing this area. In this chapter we present a new approach for breast peripheral enhancement in digital mammograms.

2.1 Introduction

As a consequence of the current digital revolution, traditional film-based hospitals are being converted to digital hospitals, where patient medical records, chart information, and test results are easily available electronically for physicians from anywhere in the hospital and beyond. By definition, FFDM is part of this digital scenario. In digital mammography, digital acquisition, digital storage, and digital display processes may be separated and individually optimised [343, 18, 276].

Digital detectors offer higher quantum efficiency and higher resolution than traditional screen-film receptors [282]. These translate into both lower dose and improved image quality mammograms. Besides, Berns et al. [23] showed that digital mammography acquisition is, a highly significant 35%, shorter acquisition in time. Once the images are acquired, the Digital Imaging and Communications in Medicine (DICOM) [14] standard handles the storage and communication protocol, enabling also the integration of the different imaging devices of the hospital. Thus displays, scanners, servers, workstations, printers, and network hardware can be integrated in

a fully digital system, usually referred as the Picture Archiving and Communication System (PACS). From here, the images are sent to the screening workspace, where usually, one or more experts will analyse and diagnose the case.

In contrast with the typical film-screen image, experts view the images on an electronic display (also called soft-copy display). In contrast to (static) film reading, soft-copy offers new opportunities. For instance, there is experimental evidence that alternating the current and prior mammograms on the same display allows better evaluation of temporal changes than conventional display of images next to each other [324]. However, a faulty inadequately calibrated or improperly set up display can compromise the overall quality of a diagnostic procedure [241].

In order to help radiologists during breast imaging evaluation, different image processing algorithms are developed to improve the visualisation of digital mammograms, either enhancing some image features to allow the detection of different types of lesions [278, 213] or improving the quality of the mammograms to compensate for possible acquisition limitations [292]. This chapter presents a proposal for image quality improvement, specifically for the correction of a known issue, the presence of an overexposed boundary area in the majority of mammograms, as it is shown in Figure 2.1(a).

In general, mammography x-ray beam consists of a central ray (perpendicular to the image receptor) and a diverging beam. The intensity of the x-ray beam is not uniform being lower in the anode side than in the cathode side (the called heel effect). This is because the rays which are parallel or near parallel to the anode get absorbed by the anode itself. For an appropriate exploitation of the heel effect, in mammography the x-ray tube is oriented with the cathode side placed over the chest wall area and the anode side placed over the nipple [56, 332]. However tube orientation is not enough to save the overexposition effect. This effect is due to the change of the breast thickness during mammographic acquisition and to the air gap effect where the skin is separated from the compression paddle and detector platform [284]. Unfortunately, this overexposition effect cannot be solved by modifying the typical window width/center parameters that viewers use (parameters that specify a linear conversion from stored pixel values to values to be displayed [13]).

Several methods have been proposed for overexposed area correction in mammography, which can be classified in non-parametric [26, 52, 154, 24], paramet-

ric [291, 284, 8, 150, 314] or other [104, 159] approaches. Non-parametric approaches try to adjust the intensity of the overexposing areas by means of traditional image processing techniques, like segmentation and equalisation. In contrast, parametric approaches adjust the intensity of the images according to a different model, which may be as specific as the type of digital detector [291] or as general as a 3D representation of the breast [284, 8, 150, 314]. The others group includes methods that do not belong to any of the above categories. This is the case of the approach proposed by Goodsitt et al. [104, 159] who designed physical filters to adjust the x-ray beam distribution for compensating the tissue thickness.

(a) (b)

Figure 2.1: Example of the peripheral enhancement: (a) original and (b) enhanced images. The images are obtained after the best manual window width and window center configuration.

In non-parametric approaches, the correction of the overexposed area follows an additive approach, where the intensities of this area (which are too dark) are added with another factor in order to obtain an intensity distribution more homogeneous throughout the breast. The different approaches may vary in the determination of the correction area and in the factor which is added. Regarding the area, algorithms could be applied to the whole area [26] or just to a region determined by a segmentation algorithm [292]. In the first case, the correction in the inner part of the breast

is inappreciable, since the additive factor is assumed to be 0 in that part. In the second one, authors segment the breast using intensity based features in order to locate those areas which were overexposed, and the correction algorithm is limited just to this part. Regarding the correction factor, Bick et al. [26] generated a curve based on the mean grey level of all points in the same distance to the skin-line. A second curve was generated by subtracting this curve from the mean intensity value of the image. The intensity value of this second curve at each distance was then added to the corresponding pixels. Karssemeijer and te Brake [154, 155] firstly computed a smoothed version of the mammogram. Subsequently, all pixels below a threshold were corrected by subtracting from the original intensity the smoothed intensity and adding the mean value of the inner part of the breast. A similar approach was also developed by Byng et al. [52]. The underlying assumption of these approaches is that thickness variations are smoother than density variations. Bessa et al. [24] use a gamma intensity correction based approach. Once the distance of each pixel to the skin line is computed, each gamma correction value is calculated as a sigmoid function of its distance to the skin line.

Our work focuses on non-parametric approaches, since it is a more general approach able to improve the quality of a single image without needing extra information, for instance the effective attenuation coefficients for breast tissues [284, 150] or the x-ray spectrum [8]. To compensate the thickness variations in the periphery of the breast, we propose a peripheral enhancement method that applies a multiplicative correction factor for each pixel of the overexposed area according to grey-level continuity constraints. Figure 2.1(b) shows an example of the application of this algorithm. Note that the correction of the overexposed area also affects the behaviour of the window width and window center adjustments. The proposed peripheral enhancement not only represents an improvement in the appearance of the images but also for breast cancer detection [61, 331]. The enhancement method is described in detail in Section 2.2. Results are shown in Section 2.3. Section 2.4 shows the benefits of our approach in two different applications. Chapter ends with discussion and conclusions.

2.2 Peripheral Area Correction

Figure 2.2 depicts our proposal for breast peripheral enhancement. The method can be divided into two main steps: determining the overexposed and non-overexposed areas, and equalising the mean intensities of both areas, in order to enhance the intensities of the overexposed area to be similar to the ones of the non-overexposed area. Before determining the overexposed area, an initial step is necessary in order to separate the breast region from other areas of the mammogram. Besides, for the intensity correction, the distance transform of the image is also computed. A final optional step is the integration of the pectoral muscle into the final mammogram.

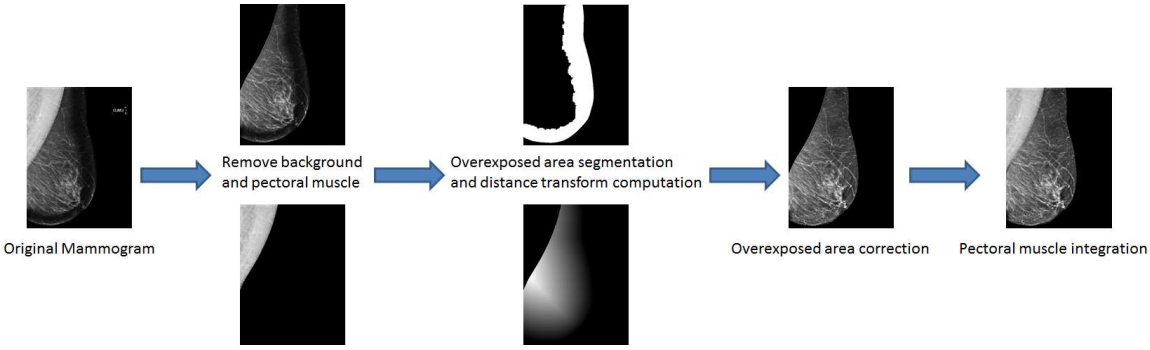


Figure 2.2: Graphical description for the overexposed area correction.

2.2.1 Breast Area Segmentation

In addition to the actual breast, mammograms contain other regions that must be removed before applying the enhancement technique, namely the background (which may contain some labels) and the pectoral muscle.

There are a lot of methods to segment breast area from background on mammographic images [289, 290, 320, 15, 349, 251, 139, 239], although majority of works are histogram [177, 350, 27] or gradient [206, 155, 267, 324, 198, 308, 307, 309] based techniques, either to determine one unique threshold that separates the x-ray exposed area from the rest [177, 27, 155, 265, 139, 338] or to find two limits, one that differentiates the x-ray exposed region from the rest and other that delimits the x-ray exposed area that contains or not the breast [351, 206, 207, 205].

In conventional film-screen mammograms, the background is noisy and inhomogeneous [254, 58], however in FFDM this region is totally black and uniform, and hence it does not require any pre-processing, a clear advantage over conventional screen-film. Therefore, the background is composed by all pixels with intensity equal to zero. Additionally, labels can be removed by keeping just the larger region of the image.

Different approaches have been proposed for detecting the pectoral muscle in MLO images [173, 87, 254, 218]. In this work we implemented the proposal of Kwok et al. [173] and according to their assumptions CC view is not considered because the pectoral muscle is only seen in about $30\% - 40\%$ of CC images. Roughly, initially the pectoral edge is estimated by a straight line which is automatically obtained by comparing the brighter and darker regions of the mammogram's upper corner. Subsequently, this straight line is transformed to a curve. This process consists in moving each pixel in the line through a few pixels along its perpendicular line and looking for the best cliff candidate, which is paced as a new edge. When all the pixels in the line are moved to each candidate position, the new edge is smoothed by least square fitting. This process is iteratively performed until convergence. The algorithm was initially developed for digitised mammograms, however the same approach has been successfully tested with digital mammograms [235].

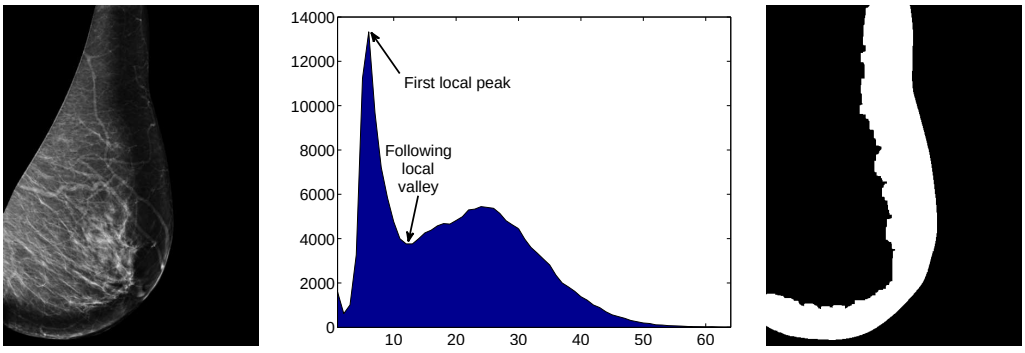
Therefore, mammograms are divided into breast area, background, and pectoral muscle, and only the breast area is kept for subsequent processing. Figure 2.3(a) shows the result of this step, where black areas are not considered in the rest of the processing steps.

2.2.2 Overexposed Area Determination

To determine the overexposed area, the histogram of the breast area is computed. Figure 2.3(b) shows the histogram of the breast shown in Figure 2.3(a). Two different areas can be seen in the histogram. The right part of the histogram corresponds to the brighter intensities, which are located in the inner part of the breast. In this region, no single intensity dominates over the rest and the wide peak reflects the inhomogeneous tissues of that region. Conversely, the left part of the histogram shows a narrow high peak, which corresponds to the outer part of the breast that

shows a darker intensity, due to the overexposing effect. This is the part that needs to be corrected.

In initial experiments, the overexposed area was determined by the automatic Otsu thresholding algorithm. However, in dense breasts, this algorithm tended to group the overexposed area with fatty tissue. So finally, to locate the overexposed part we analysed the histogram of the mammogram, automatically locating the first local peak and subsequently looking for its adjacent local valley. The intensities contained between the histogram origin and this valley are the ones that will be corrected. These intensities correspond to the pixels of the overexposed area, as it is shown in the binary map of Figure 2.3(c). This map has been refined using morphological transformations to obtain a smoother contour.



(a) (b) (c)

Figure 2.3: Example of the thresholding process: (a) breast area segmentation, (b) histogram of the breast area clustered in 64 bins, (c) overexposed area segmentation.

2.2.3 Distance Transform

Before the calculation of the correction factor, the Euclidean distance map of the mammogram is obtained by computing the minimum distance from each pixel to the breast skin-line. This image helps to speed up the whole process, since the use of the distance map makes it possible to deal with all pixels at the same distance at the same time.

The distance transform depends on the metric used, being the Euclidean or Manhattan metrics the most common ones. The latter provides a faster resolution of the problem, especially in multidimensional images, at the expense of obtaining just

an approximated solution. Rosenfeld et al. [259] provided the roots for fast computation of the Euclidean distance transform, and recently Fabbri et al. [83] reviewed different approaches for computing it, distinguishing among three classes according to their own implementation: ordered propagation algorithms, raster scanning and independent scanning. In ordered propagation algorithms [84], the algorithm starts from the seeds (0 distance) and progressively transmits the information to other pixels in order to increase the distance. In raster scanning algorithms [274], 2D masks are used to guide the processing of pixels line by line, top to bottom, then bottom to top. Finally, independent scanning schemes process each row of the image, independently of the other, and then process each column of the result (similarly to the implementation of the Fourier transform of an image by a sequence of 1D transforms in orthogonal directions). Independent scanning schemes can be divided into parabola intersection algorithms [70], mathematical morphology based algorithms [138], and algorithms based on Voronoi tessellation [202, 335], these latter providing a more general scheme and allowing parallel computing implementations.

The implementation used in this paper follows the approach of Maurer et al. [202], which allows fast computation of the Euclidean distance transform to multidimensional spaces. At each dimension level, the distance transform is computed by constructing the intersection of the Voronoi diagram (whose sites are the feature voxels) with each row of the image. This construction is performed efficiently by using the distance transform in the next lower dimension. Authors demonstrated that the algorithm has linear time complexity (O(N)), while the parallel version of the algorithm being used here runs in O(N/p) time, being p the number of processors.

2.2.4 Correction Factor

Once the overexposed area is located, the goal is to correct its intensities by considering the values of the non-overexposed area of the mammogram by following an iterative process. Firstly only the pixels belonging to the inner overexposed area are corrected looking at the intensities of its neighbouring pixels which are not overexposed. Once these pixels have been corrected, the overexposed area mask is reduced and the new inner overexposed pixels are corrected. This process is applied until all pixels are corrected.

Formally, given a pixel x of breast B with intensity $I(x)$, we define its neighbourhood as:

$$N(x) = \{t \in B : d(t, \mathrm{S}) = d(x, \mathrm{S}) \ \wedge \ d(t, x) \leq k\} \tag{2.1}$$

where t is a pixel of the breast, S is the skin-line contour, $d()$ refers to the distance (notice that the first distance is point-to-line while the second one is the common Euclidean distance between two points) and k represents the size of the neighbourhood (discussed in Section 2.3). Notice that with this definition, we are considering a neighbourhood $N(x)$ of pixel x the closest neighbours of x with same distance to the breast skin-line.

We also define $N_{in}(x)$ as the neighbourhood of x located one pixel further inside the breast:

$$N_{in}(x) = \{t \in B : d(t, S) = d(x, S) + 1 \ \wedge \ d(t, x) \leq k\} \tag{2.2}$$

With these definitions and from the furthest overexposed pixel from the skin-line boundary to the closest, each intensity of each pixel is iteratively corrected as:

$$I'(x) = I(x)\frac{\overline{I}_{N_{in}(x)}}{\overline{I}_{N(x)}}$$

where $\overline{I}$ refers to the mean intensity of each neighbourhood. The rationale behind this equation is twofold. By dividing by the mean value of the neighbours at the current distance we are normalising the overexposed values without losing information, while when multiplying by the mean of previous values we are giving more weight to the intensity continuity constraints. Notice that the use of the distance map allows us to speed up the process dealing at the same time with all the pixels which are located at the same distance of the skin-line.

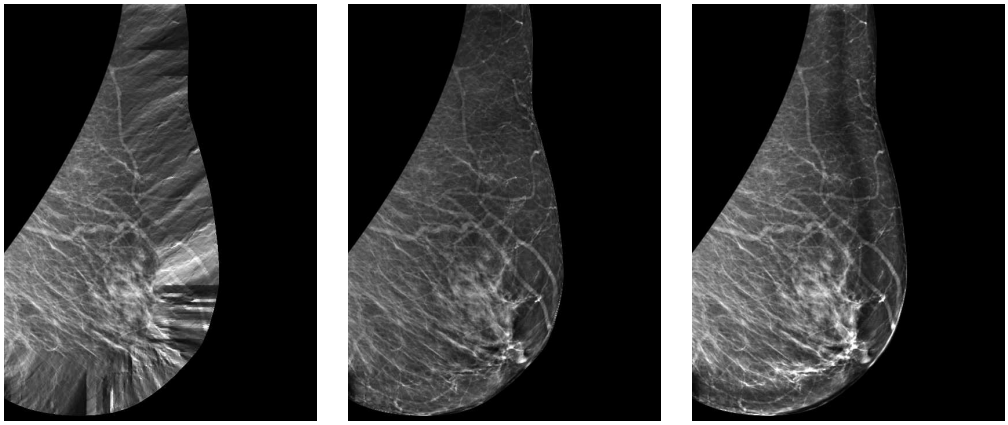
2.2.5 Pectoral Muscle Integration

The final step of the algorithm is the integration of the pectoral muscle into processed images. This step may be unnecessary for posterior computer-based analysis

such as the detection of abnormalities [95, 235] or the analysis of the breast density [230, 234]. However, we have experimentally observed that radiologists feel more comfortable when the pectoral muscle appears in the mammogram. This is done just adding the result of the pectoral muscle segmentation and the enhanced mammogram, since they are disjoint regions.

2.3 Results

A total of 334 FFDM acquired using a Hologic Selenia mammographic system, with resolution equal to 70 μm pixel, size 4096×3328 or 3328×2560, and 12-bit depth were used in this work. These images included CC and MLO views of 117 women, 50 full exams (CC and MLO of left and right breasts) while the rest were CC and MLO of either right or left breasts. After double-reading by the experts, 98 images contained masses (10 of them were located in the peripheral region).



(a) (b) (c)

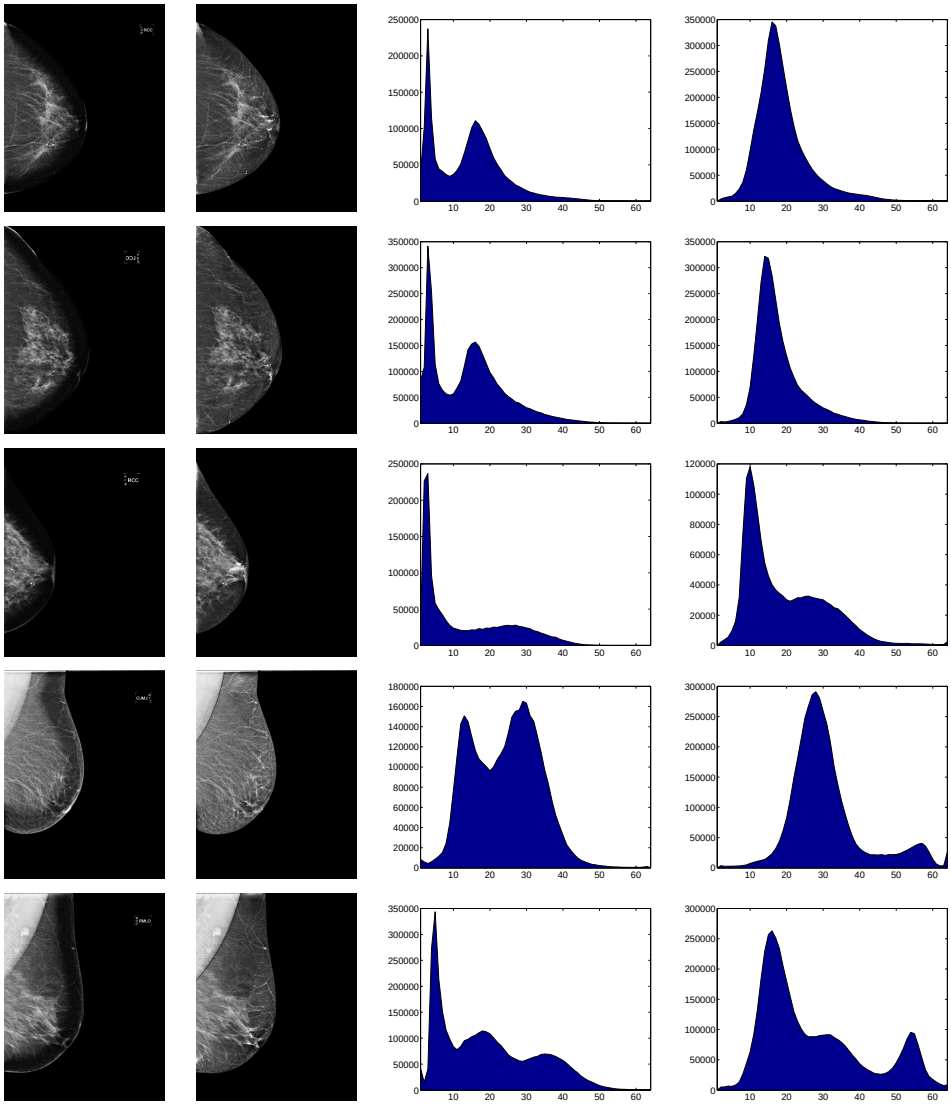
Figure 2.4: Example of the peripheral enhancement when the value of the size of the neighbourhood is varied: (a) $k = 20$, (b) $k = 100$, and (c) $k = 180$. The best results were obtained using values of k around 100 pixels.

The only parameter in our algorithm that needs to be tuned is the size of the neighbourhood used for correcting intensities of the overexposed pixels (parameter k). We fix this size to be 100 pixel (i.e. 7 mm). Figure 2.4 shows the resulting image when smaller, similar, and larger values of k are used. If k is too small the

neighbourhood is also small, and the new value assigned depends on just few pixels. Hence, the propagated value is always close to the reference pixels and corrected pixels are assigned with similar intensity. On the other hand, when k is large, the appearance of the mammograms is better, although a dark ribbon appears due to the fact that new value depends from pixels far from it. The best results were obtained with k close to 100 pixels. Actually, the tuning of this parameter allows for a certain flexibility, since we obtained similar results using values of k between 70 and 120 pixel (4.9 and 8.4 mm).

Figure 2.5 shows the result of the enhancement on five different mammograms, three CC and two MLO views (displayed without adjusting the contrast of the images). The first column corresponds to the original mammograms while the second one shows the corrected images. In MLO views the boundary automatically found between the pectoral muscle and the breast is depicted in black. Comparing the mammograms, the dark overexposed areas have been corrected, obtaining a more homogeneous intensity distribution along all the breast area, and obtaining a mammogram with much more visual detail in the peripheral area. Additionally, notice that the structures lying between the overexposed and non-overexposed areas are continuous and have not been altered or disrupted in any place. This means that the tissue shown in the improved area corresponds to the real tissue of that region, and that our approach is not introducing artefacts that could affect the subsequent analysis of the image.

To quantitatively analyse the improvement in terms of image quality, Figure 2.5(c) and (d) shows the histograms of the breast area (without including the pectoral muscle in MLO images) before and after the enhancement, respectively. As expected, the big peak in the overexposed area is not present in the corrected histogram, and in contrast, the smooth wide peak reflects that the intensities are distributed among all the breast area.

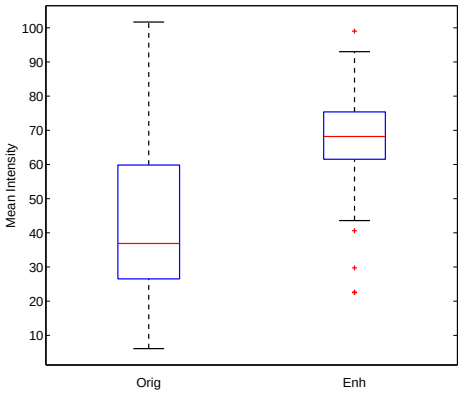


(a) (b) (c) (d)

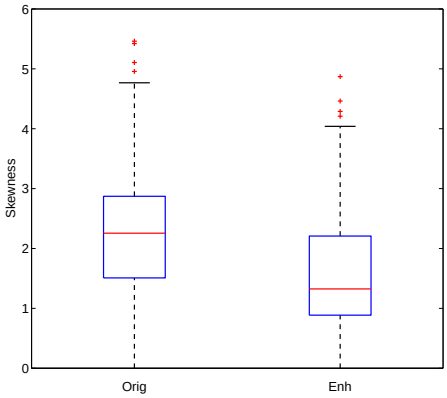
Figure 2.5: Example of the peripheral enhancement: (a) original mammogram and (b) corresponding corrected mammogram, where in MLO views the border between pectoral muscle and the breast is shown in black, (c) breast area histogram of the original mammogram, and (d) breast area histogram of the corrected mammogram. Notice that the narrow peak of darker intensities has been removed and the histogram is now more homogeneous.

To determine the quality of the processed images, visual assessment was performed by one radiologist with more than 10 years of experience in mammographic images. All original and corrected images were displayed side by side and the observer just labelled the images as correctly processed or not, where correctly processed means that the quality of the images substantially improved after the enhancement. A total of 90.42% of the images were evaluated as correctly processed. In the rest of cases, the overexposed area was not as dark as in the rest of the images, due to either less radiation or breast physiology and the correction was actually not necessary. When applying our approach to these images, the area selected for correction was the full mammogram, and the final output was the original image but globally enhanced, i.e. without obtaining any local effect. For all cases, the quality of the images did not decrease after the processing.

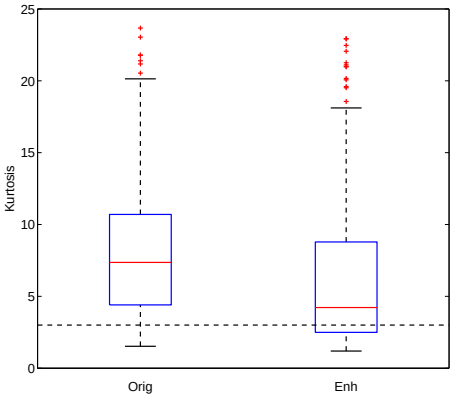
A quantitative analysis was also performed by comparing the mean intensity and the histograms of the original and enhanced mammograms. Regarding mean intensity, since overexposed areas have darker intensities, a "visually better" mammogram should obtain a larger mean intensity value. On the other hand, we also have to define what a "better" histogram means. Histograms from overexposed mammograms present a high peak in dark intensities. This peak is the one that should be reduced when enhancing the mammograms, obtaining a histogram with a well-defined central peak after enhancement. To measure this, we compute the skewness and the kurtosis of the histogram. The skewness measures how asymmetric the histogram is, 0 being a totally symmetric distribution. The kurtosis measures how wide the peak is, being 3 the kurtosis defined by a normal distribution. Figure 2.6 compares the mean intensity and both the skewness and kurtosis measures for the 334 analysed mammograms summarised in terms of boxplots. As expected, the mean intensity value has increased, almost doubling its value. On the other hand, skewness has been reduced, hence obtaining a more symmetric distribution, while kurtosis is also reduced, obtaining a wider peak of the breast area. Analysing the results using paired t-Student test, the improvements in each measure were significant with p-value< 0.0001.



(a)



(b)



(c)

Figure 2.6: Quantitative analysis of the original and enhanced histograms in terms of boxplots: (a) mean intensity of the breast area, (b) skewness, and (c) kurtosis.

2.4 Applications

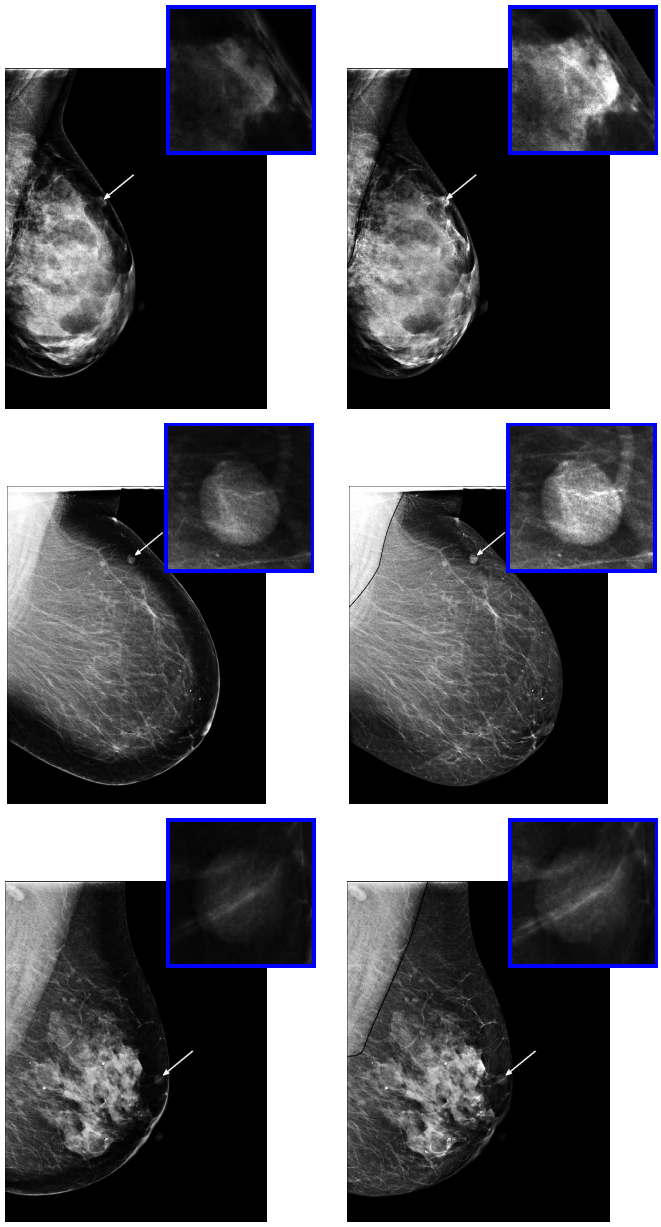
In this section two important applications are showed. The use of the peripheral enhancement allows improving the procedures: first, in expert detection of masses routinely performed in clinical practice, and secondly, in a computer aided application focusing on computing the breast density of the mammogram.

2.4.1 Mass Detection After Enhancement

A total of 334 FFDM were used for mass detection assessment. 98 of the 334 images contained masses and 10 of the masses were located in the peripheral breast region (see Section 2.3 for more detail).

To perform an estimation of the benefits of the approach for manual analysis of the breast we asked to the same experts that diagnosed the mammograms to analyse all the cases again but after the peripheral enhancement (the radiologists were different from the one that performed the qualitative evaluation). Notice that there was more than one year of difference between the first and the second diagnosis, and we can consider that radiologists did not remember the cases.

As expected, radiologists were able to diagnose again all the cases containing masses. After asking their opinion about the enhancement, they reported two main favourably reasons to the enhancement. First of all, masses were easier to detect, allowing a faster analysis of the images. For instance, Figure 2.7 shows three examples of mammograms with masses located at the overexposed area. Notice that the masses are highlighted after peripheral correction. And secondly, the border of the masses was also clearer compared to without the enhancement. This is an important result, since the border and the shape of the masses are two important features when diagnosing the mass as cancer or benign.



(a) (b)

Figure 2.7: Example of the peripheral enhancement in mammograms containing masses: (a) original mammogram and (b) corresponding enhanced mammogram. The white arrow indicates the location of the mass in each image, which is zoomed in the small box for a better visualisation. Notice that using the proposed enhancement the contrast of the lesion is now clearer.

2.4.2 Automatic Breast Density Estimation After Enhancement

Breast density is an important risk factor for breast cancer. It is well known that dense breasts are more likely to develop breast cancer than fatty breast, and besides, the analysis of dense breasts is more difficult, since the own tissue may mask the abnormalities. Breast density estimation allows for the classification of mammograms according to their internal tissue. Consequently, it is a first step to a more personalised analysis.

Oliver et al. [230] developed an automatic method for breast density analysis based on first segmenting the mammograms into two classes: fatty and dense. Subsequently, features from both sets were extracted and used to classify the mammograms. Authors obtained good results using two different digitised databases (kappa=0.81 and kappa=0.67, respectively). However, when using this approach in digital mammograms [313] (specifically in 236 FFDM of our digital database), the method fails to obtain good results due to overexposed area (kappa less than 0.4). In order to overcome this issue and obtain better results, the peripheral enhancement was applied. The results were improved to kappa=0.88 (see Chapter 3 for more details).

2.5 Discussion and Conclusions

Overexposed areas in a mammogram are observed as darker areas, and the structures lying in this region are hardly visible. In this chapter we presented a novel algorithm that automatically enhances the intensity of that region to obtain similar intensities to the rest of the mammogram. Our approach was inspired by the additive models used to enhance the mammograms [26, 52, 154]. However, we claim that our multiplicative model not only provides intensity continuity on the overexposed area (the same as the additive model) but also preserves the relationship between the pixels of the overexposed area.

The presence of the overexposed area cannot be solved by common visualisation contrast tools, due to the fact that most of these tools are based on histogram information. For instance, manually applying one of the most common tools such as the

adjustment of window width and window center [224, 175] the whole mammogram cannot be properly visualised. Actually, experts should tune the parameters for both overexposed and non-overexposed areas independently, which is not possible. When applying our enhancement approach, we are locally correcting the overexposed area without considering global histogram information. Instead, each pixel is corrected using just neighbourhood information. After applying the peripheral enhancement, and using the window width and center adjustment tool, the full mammogram can be properly visualised, as shown in Figure 2.1.

The proposed approach presents many benefits, either for manual or automatic analysis of the mammogram. In manual analysis, the proposed enhancement helps uncover different kind of abnormalities in the peripheral zone. Although expert radiologists may actually find them after thorough analysis of the mammogram, abnormalities are easier to detect after the enhancement. Moreover, the physical limits of the mass, which were not visible on the original mammograms, could properly be assessed after enhancement. On the other hand, automatic analysis, such as computer aided detection and breast density measurement systems can also benefit from this enhancement [151, 313].

We have provided a peripheral enhancement algorithm for digital mammograms. This step is necessary to balance the overexposure of the breast periphery. The algorithm has been tested on a large database of digital mammograms, considering visually better for diagnostic in 90.42% of the images. In these images, the correction of the overexposed area helps to improve the quality of the image and to obtain a better visualisation and assessment.

Chapter 3

Breast Density Classification

Mammographic breast density is strongly associated with breast cancer, being demonstrated its potential as a risk indicator for the development of this type of disease. Furthermore, breast density reduces the accuracy of mammography and the sensitivity of automatic breast lesion detection systems is significantly dependent on breast tissue characteristics. Here, we describe the development of an automatic breast tissue classification methodology for digital mammograms.

3.1 Introduction

Mammographic density represents the amount of fibroglandular tissue, which is radiographically dense, in contrast to fat tissue, that appears lucent in a mammogram [185]. The majority of studies about the relationship between breast density and breast cancer reported that women with high dense breast have greater risk of suffering from breast cancer than those with low dense breast [36, 203, 318], so breast density is considered one of the strongest risk factors for breast cancer [348, 38].

Besides the risk of developing breast cancer, density is also related to the difficulty of radiological reading of mammograms [304, 32, 10, 48] and the difficulty of automatic breast cancer detection [153, 229, 232, 48]. Specifically, studies show that even though breast density does not affect the sensitivity of microcalcification detection CAD systems, it significantly affects mass detection and the sensitivity of CAD systems for mass detection decreases in dense mammograms [258] whereas the

specificity of the system remains relatively constant [135] or is decreased [45]. Therefore, breast density assessment is regarded as an important tool to help radiologists and CAD systems to detect breast cancer.

Dense tissue segmentation is commonly used as the first step in breast density evaluation methods. This is a difficult task, very time consuming and prone to inter/intraraters subjectivity. Semiautomated [50, 51, 35] and automated techniques [153, 151, 121] are being developed, although the lack of a reliable ground truth is a hard problem to evaluate those approaches.

A huge variety of strategies can be found in the literature. For instance, some authors try to find a threshold to obtain a suitable segmentation [53, 262, 221, 87, 204, 236, 217]. Byng et al. [54] and Lu et al. [186] obtain an interactive threshold by using local skewness and fractal dimension measurements, although Byng et al. use digitised mammograms and Lu et al. adapt the method to analyze digital images. Sivaramakrishna et al. [279] use a modified version of Kittler's method. Zhou et al. [358] and Martin et al. [201] apply a rule based classification to classify the breast images into four classes and depending on the class use a discriminant analysis based method or a maximum entropy principle to calculate the optimal threshold. Torrent et al. [311] use a threshold technique based on the excess entropy. El-Zaart [78] considers that the histogram of an image is a mixture of gamma distributions and selects the thresholds at the valleys of the multimodal histogram. Kim et al. [161] and Nickson et al. [222] determine the best threshold that divide dense and nondense regions by combining statistical and boundary information, although Kim et al. work with FFDM and Nickson et al. adapted the method for digitised images.

However some authors consider that the use of thresholds may not be sufficient to distinguish all of the dense parts of the breast and use other segmentation approaches, for instance clustering techniques [19, 198, 311]. Strange et al. [293] and Elmoufidi et al. [79] use k-means clustering, although Elmoufidi et al. also use a recurrent seed based region growing technique. Oliver et al. [231, 233, 230], Chen et al. [64] and Keller et al. [157, 158] use a fuzzy C-means based algorithm. Oliver et al. divide the breast area in two regions (dense and fatty) whereas Chen et al. divide the breast in four regions (fatty, fatty glandular, dense glandular and dense) and Keller et al. [157, 158] calculate the appropriate number of clusters for each image. Zwiggelaar et al. [367, 364, 365] use a set of cooccurrence matrices to gener-

ate a texture feature vector at a pixel level and divide this data in six classes using the expectation maximisation clustering. Selvan et al. [269] assume that mammograms can be represented by a finite generalized gaussian mixture model (as many gaussians as regions) and compare the results of the estimation of the model parameters by using the expectation maximization algorithm and an heuristic parameter estimation approach. He et al. [124] employ an unsupervised clustering technique focused on incorporating an optimal centroids initialisation step and on reducing the number of missegmentation by using an adaptive cluster merging method.

Regardless of the dense tissue segmentation applied, texture analysis [192, 210, 295, 245, 366, 123, 363, 120] is one of the most common procedure used during the density assessment process. Authors work with moments [51, 54, 153, 183, 294], with fractal dimension [55, 51, 54, 312, 316], with local binary patterns [233, 319], with Laws' masks [210, 32, 169], with cooccurrence matrices [29, 367, 364, 365, 233, 219, 151, 319] or with textons [247, 246, 31, 102, 214] among other texture features. Nevertheless, there isn't a classifier that stands out among others, SVM [216, 183, 294, 48, 316, 158, 319, 1], neural networks [32, 48, 151, 319], nearest neighbours [153, 230, 219, 169, 253], bayesian [230, 48, 319] or probabilistic latent semantic analysis [31, 60] are examples of the classifiers used. Although, in breast tissue classification, it's important to review not only the classification technique applied but also the type of classification measurement (qualitative or quantitative) used.

For qualitative evaluation of mammographic density [155, 29, 32, 198, 131, 219, 294], the Wolfe categories [55, 304, 303, 312, 215, 366], the Tabár grade [215, 126, 123, 120, 122, 121, 127, 125], the Boyd classes [35, 215] or the ACR BI-RADS score [201, 215, 186, 60, 145, 62, 63, 319] can be used. Wolfe [348] takes into account the density of the parenchyma and the duct pattern. Tabár [300] uses the amount of nodular densities, linear structures, homogeneous fibrous tissues and radiolucent adipose fatty tissues. Boyd [35] works with the percentage density (PD) (dense area divided by total area) similar to BI-RADS [227]. However the BI-RADS is becoming a standard in the medical community on the assessment of mammographic images. According to the BI-RADS protocol, mammograms can be classified into four categories:

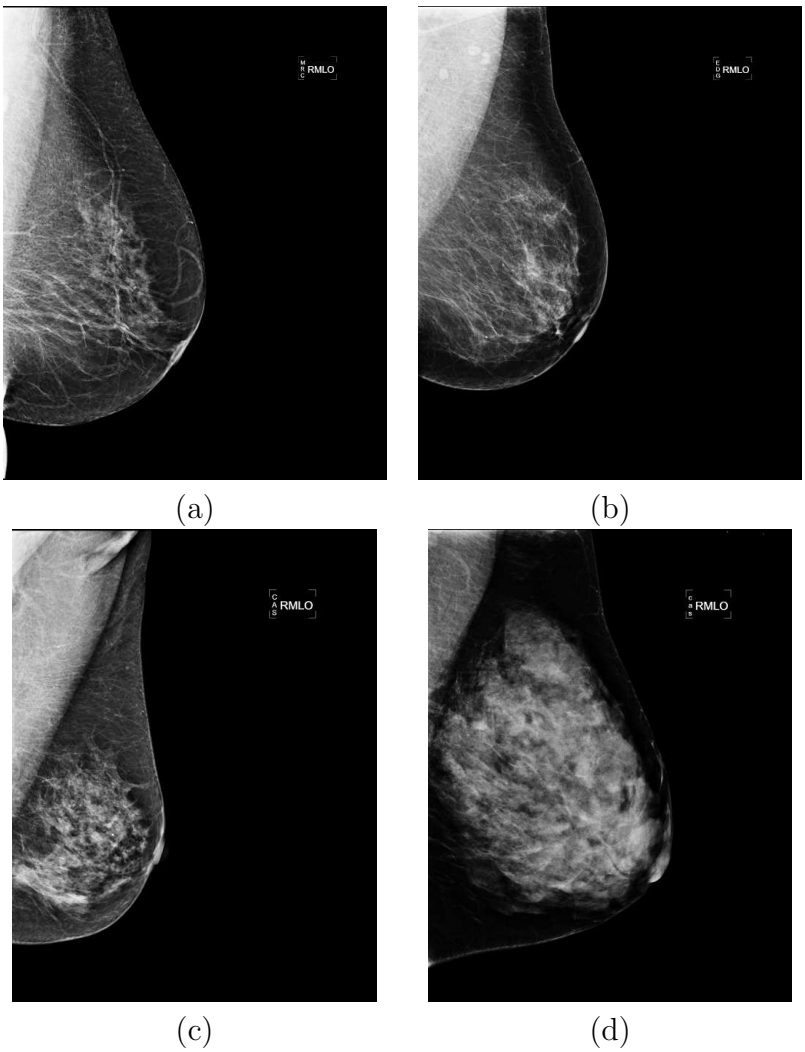
I. The breast is almost entirely fat ($< 25\%$ glandular).

II. There are scattered fibroglandular densities ($25 - 50\%$ glandular).

III. The breast tissue is heterogeneously dense ($51 - 75\%$ glandular).

IV. The breast tissue is extremely dense ($> 75\%$ glandular).

Figure 3.1 shows four mammograms, one of each density class. The amount of fibroglandular tissue increases with each category, which could obscure detection of small masses or lower the sensitivity of mammography.



(a) (b)

(c) (d)

Figure 3.1: Four FFDM of increasing internal density. (a) BI-RADS I, (b) BI-RADS II, (c) BI-RADS III, and (d) BI-RADS IV.

For quantitative assessment of mammographic density [133, 86, 30, 344, 2, 117, 263, 293], area based and volumetric approaches are developed. Area based methods estimate the PD, the absolute density and/or the absolute nondensity [358, 221, 131, 130, 151, 157, 132, 158]. This is the case of the semiautomated Cumulus software [50, 35], the fully automated ImageJ based methods [181, 287] or the fully automated AutoDensity software [222]. All these tools use digitised screen film (analogue) mammograms although Cumulus and ImageJ based methods also work in converting the raw digital images into "analogue-like" images. Volumetric approaches calculate the volumetric PD and the dense tissue volume [273, 242, 34, 184, 43, 106]. The Volpara software [323, 134, 270], the Quantra software [204, 4, 65], the CumulusV algorithms [10, 8, 9] or the Single X-ray Absorptiometry (SXA) method [271, 194, 272] are examples of fully automated volumetric tools to measure breast tissue composition from preprocessed (raw) digital mammograms.

Both, qualitative and quantitative assessment approaches are equaly valid and used for breast density evaluation, the choice of one or another depends on the intended use [81]. In this chapter, we present different approaches for mammographic density classification and analysis of both types. In Section 3.2, density assessment methods are described in detail. Section 3.3 contains the results of each approach and the proper result comparisons. Finally, discussion and conclusions are presented in Section 3.4.

3.2 Proposals

We present three different methods for breast density analysis. Two of them automatically classify mammograms according BI-RADS categories. Methods follow the same line of work, the breast area is segmented in regions, then morphological and texture features are extracted from fatty and dense regions and finally different classifiers are applied to classify the mammograms in BI-RADS I, II, III or IV. The most significant differences between them are the image preprocessing applied and the segmentation technique used. Obviously, these methods belong to the family of qualitative breast density assessment approaches. On the contrary, the third method focuses on quantifying the amount of breast dense tissue. This approach classifies each pixel in dense or fatty by using intensity, texture, and morphological

information.

Note, that all of the approaches have a common preprocessing step to extract the breast area and remove the background, labels, and pectoral muscle areas (see Fig. 3.2(a)-(b), Fig. 3.3(a)-(b)). The method applied for breast area segmentation is described in detail in Section 2.2.1. However, in few images this method failed and the pectoral muscle was manually removed.

Note also, that during mammographic acquisitions, the breast is compressed with a tilting compression paddle, so the breast thickness through the mammogram is nonuniform, being lower in the periphery. This implies that pixels near the skin line are overexposed, and in a mammogram those pixels appear darker than the rest of breast pixels (see Fig. 3.1, Fig. 3.2(a)-(b), Fig. 3.3(a)-(b), and Fig. 3.4(a)). Methods deal with this issue differently. Specifically, for qualitative methods, the region segmentation methodology used depends on the presence or not of this overexposed area.

3.2.1 Qualitative Methods

3.2.1.1 Region Based Classification Using Eigen Analysis

The proposed method is based on an initial three class segmentation of the breast region. After segmentation, a set of morphological and texture features of the tissues of interest are extracted, and finally, mammograms are classified according to BI-RADS categories. In the following sections we will describe each of these steps.

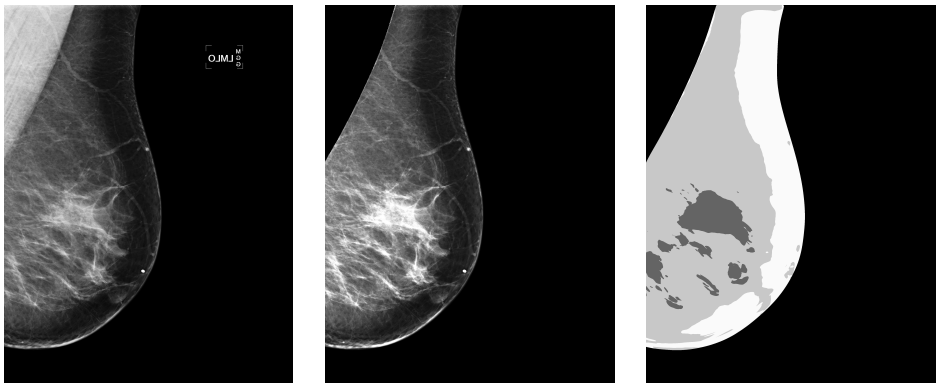
Segmentation

After breast area extraction, our segmentation method takes into account the possibility of presenting the already known overexposed region and distinguishes between 3 classes: fatty, dense and overexposed.

The segmentation approach consists of two main steps. Firstly (the training stage) the system learns what each tissue represents and secondly (the testing stage) the system is able to detect if a new area is fatty, dense or overexposed.

During the training stage, a database of ROIs that contain the three tissue types was created. This set of ROIs was used as input of an algorithm based on

the eigenfaces approach [315], and a set of templates that represents the three classes was obtained. Finally, during the testing stage, a Bayesian pattern matching approach [143] was used to decide when a pixel is part of a fatty or a dense or a overexposed region. Figure 3.2(c) shows an example of the final segmentation.



(a) (b) (c)

Figure 3.2: Example of the firsts steps of the tissue classification methodology: (a) original image, (b) preprocessed image (only breast region is preserved) and (c) three class segmentation.

Feature Extraction

The result of segmentation algorithm is the division of the breast into three clusters, although we focused on fatty and dense tissue that are really our tissues of interest. A set of features for both classes are extracted from the original images. In this work, we used morphological and texture features, similar to Oliver et al. [230]. The relative area, the first four histogram moments (mean intensity, the standard deviation, the skewness, and the kurtosis), a set of features derived from Laws' texture measures [178] and grey level cooccurrence matrices (GLCM) [116] were calculated. We generated 25 Laws' 5x5 operators that were convoluted to each digital image to generate 25 features. From GLCM, we extract 176 features. GLCM can be specified as a matrices of relative frequencies P_{ij}, in which two pixels separated by a distance d and angle θ have grey levels i and j. Here, we used four different directions: 0°, 45°, 90° and 135°, and four distances (1, 4, 7 and 10 pixels). For each cooccurrence matrix, contrast, energy, entropy, homogeneity, correlation, sum

average, sum entropy, sum variance, difference average, difference entropy and difference variance were calculated.

Each of these features were extracted from fatty and dense class, then we dealed with 412 features in total (2 relative areas, 8 moments, 50 from Laws' texture measures and 352 from cooccurrence matrices). Note that, all features were normalized to unit variance and zero mean.

Classification

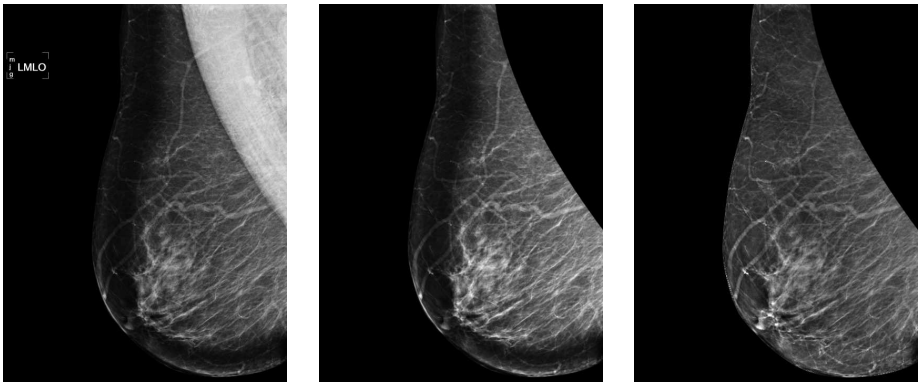
The classification of mammograms according to BI-RADS categories was performed using the following classifiers: Naives Bayes [146], KNN [3], RF [44], AdaBoost [96] and SVM [68]. Due to the large number of features, we included a feature selection step that automatically selected the most effective subset of features. Various feature selection techniques were proposed and evaluated such as Principal Component Analysis [148], Gain Ratio attribute evaluation [347] or SVM [111]. The data was analysed by using WEKA [115] data mining software.

3.2.1.2 Region Based Classification Using FCM Analysis

This classification method is based on a previously developed algorithm for breast tissue density classification [230] of digitised mammograms. It is an adaptation to digital images. The adapted method consisted in: (1) preprocessing, (2) segmentation in fatty and dense tissue, (3) feature extraction from both classes, and (4) classification according to BI-RADS categories.

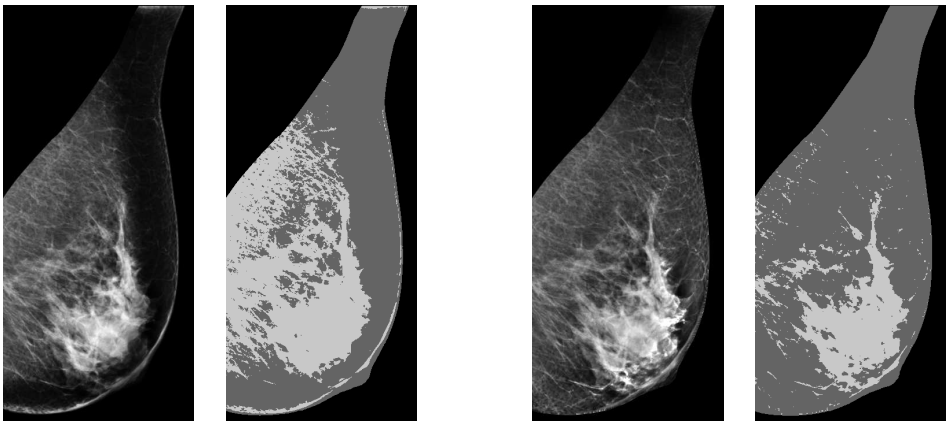
Preprocessing

Once breast area was extracted, the fuzzy C-Means (FCM) clustering approach was applied. The first results of the FCM segmentation were not accurate enough (see Fig. 3.4) due to the presence of the known overexposed area (see Fig. 3.3(a)-(b) and Fig. 3.4(a)). So, we decided to compensate the thickness variations in the periphery of the breast by using a peripheral enhancement method that is described in detail in Chapter 2. An example of the overall preprocessing process can be seen in Figure 3.3.



(a) (b) (c)

Figure 3.3: Example of the preprocessing process: (a) original image, (b) breast area segmentation and (c) peripheral enhancement.



(a) (b) (c) (d)

Figure 3.4: Example of the segmentation process: (a) original breast area, (b) FCM without previous peripheral enhancement, (c) breast area after peripheral enhancement and (d) FCM with previous peripheral enhancement.

Segmentation

Grey level information in combination with the FCM approach [25] is used to group the pixels of the breast area into fatty and dense tissue classes. We use the same criterion function and the same initial seed placement methodology as in the work of Oliver et al. [230].

Feature Extraction

Once the breast area is divided into two classes, a set of morphological and texture features for fatty and dense tissue are extracted, similar to the eigen region based approach. The relative area, the first four moments of the histogram, 25 Laws' texture features and 176 texture features derived from cooccurrence matrices are calculated for fatty and dense tissue, 412 features in total (see Feature Extraction Section in Sec. 3.2.1.1 for more details).

Classification

Due to the large number of features, a feature selection step is included selecting the most effective subset of features. Various feature selection techniques are evaluated (using WEKA [115] data mining software) such as Principal Component Analysis [148], Gain Ratio attribute evaluation [347] or SVM [111]. The classification of mammograms according to BI-RADS categories is also performed with WEKA. We used different popular classifiers like RF [44] or SVM [68] and some combinations of classifiers as AdaBoost [96] or a binary tree of SVM. The binary tree consists in firstly, classification of digital mammograms in low or high breast density category and then low dense cases are classified in BI-RADS I or II and independently, high dense cases in BI-RADS III or IV. The reason was to convert our multiclass classification problem into multiple binary classification problems because SVM is originally a binary classifier [7].

3.2.2 Quantitative Methods

3.2.2.1 Pixel Based Classification

Figure 3.5 graphically shows our proposal for breast density segmentation. Since it is a supervised approach based on a classifier, two main parts can be distinguished, the training and the testing part. In the training part, the classifier learns to distinguish between fatty and dense pixels, and in the testing part, the classifier assigns the fatty or dense label to each pixel of the input image. In order to use not only intensity but other information, a previous step is performed in order to compute features related to morphology or texture of each pixel. Moreover, before processing the images,

an initial preprocessing step is necessary in order to remove the pixels outside the breast area and also for enhancing the peripheral breast area. Each step is described in more detail below.

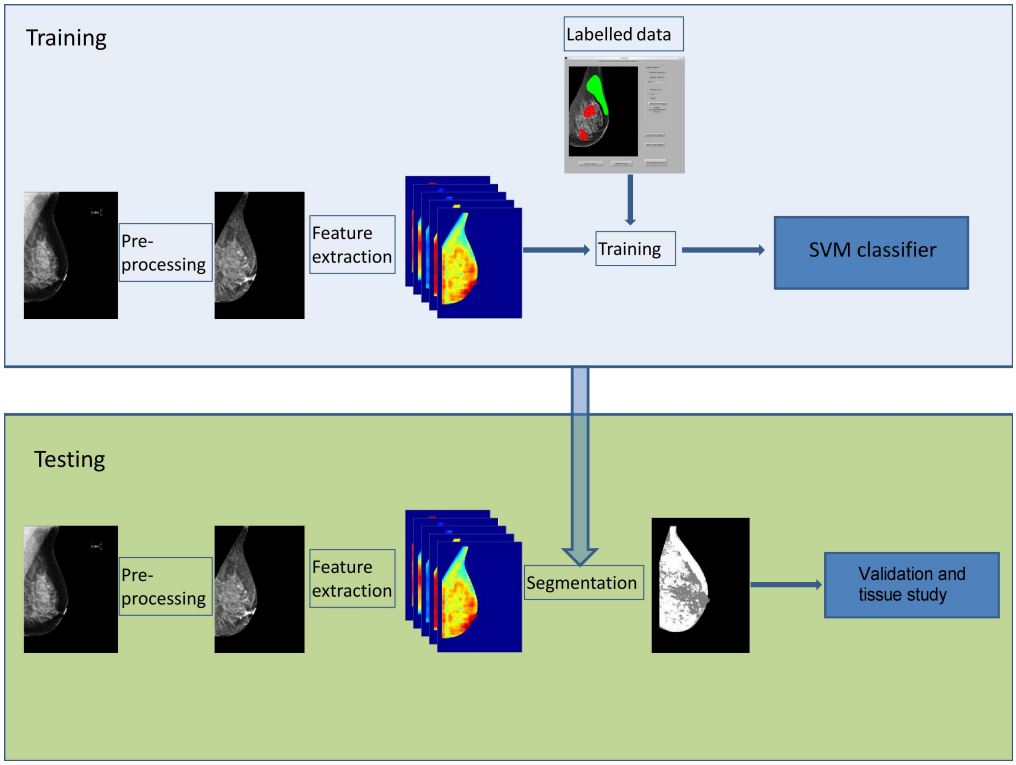


Figure 3.5: Our approach follows a pixel based classification scheme. Firstly, the algorithm learns to distinguish between fatty and dense tissue (blue box), and subsequently, it is used to segment new images (green box).

Preprocessing

The preprocessing step is the same as the one used in the FCM region based approach (Section 3.2.1.2). Firstly, the breast area is extracted and then the overexposed area is improved by a peripheral enhancement method (see Chapter 2).

Feature Extraction

A common feature used for breast density segmentation is the own image intensity [311]. However, this information alone might not be enough for

performing a correct classification. Hence, texture information allows to introduce information of the neighbourhood of the pixel into the segmentation algorithm [233]. Additionally, the use of morphological features, incorporates information on the likely locations of dense tissue into segmentation algorithms [151]. Our proposed tool also combines intensity, texture, and morphological features.

We computed all the features shown in Table 3.1 using two mammograms of each BI-RADS class (hence, 8 mammograms in total) and a large variety of scales. To compute the texture features, a search window was centered on a pixel to define a ROI. Textural features were computed inside this region and assigned to the corresponding central pixel. This process was repeated for all pixels on the image, except for pixels located at the border of the image where the window could not be placed; these pixels were excluded from further processing. The size of the window allowed experimentation with different texture scales. Small windows allowed detection of small tissue patterns whereas large windows allowed detection of bigger patterns. The range of sizes used enabled the detection of most textural patterns present in the breast density.

Windows from 1.96mm to 28.28mm were used to calculate the three first histogram moments, the histogram entropy, the three first Laplacian moments, the Laplacian homogeneity and entropy, the GLCM (at 0°, 45°, 90° and 135° directions) contrast, energy, entropy and homogeneity, the fractal dimension average, the magnitude of quaternion energy and standard deviation, the phase angle (associated with the direction k) standard deviation and weighted (magnitude of quaternion norm) standard deviation and the LBP histogram (10 bins). In total, 245 features including the original and corrected (i.e. pixel value after peripheral enhancement) intensities, the position, the distance to skin, the distance to nipple and the angle to nipple.

We created one image representation for each one of the 245 features. Each image representation was generated using the pixel computed feature as pixel value. In total,1960 image representations were created (245 image representations for each mammogram that is 8 image representations for each feature). The BI-RADS classification of each mammogram allowed us to, from the com-

Type	Feature
Intensity	Original Corrected
Morphologic	Position x, y Distance to skin Distance to nipple Angle to nipple
Texture	Histogram moments 1, 2, 3 Histogram entropy Laplacian moments 1, 2, 3 Laplacian homogeneity and entropy Cooccurrence matrices Fractal dimension Quaternion wavelets Local binary patterns

Table 3.1: Features initially analysed for breast density segmentation.

plete feature set, visually select the features that best distinguished between dense and fatty tissue (notice that the lack of a proper ground truth prevents the use of automatic feature selection algorithms). This led us to a smaller set of features, which are summarised in Table 3.2. In total, 51 features were used: original and corrected intensities, position x and y, distance to skin, the three first histogram moments and histogram entropy (computed at four different scales), the three first Laplacian moments and Laplacian homogeneity and entropy (also at four different scales), and Local Binary Patterns [228] (we used 10 bins), which were computed using an elliptical neighbourhood. As Nanni et al.[220] observed, using an anisotropic neighbourhood rather than an isotropic one is usually more useful in medical imaging, as anisotropic distributions are more common in these types of images.

Segmentation

The segmentation step was performed for classification at the pixel level by means of a SVM classifier [327]. SVM separates pixels in two possible categories: fatty pixel or dense pixel.

To train the SVM, dense and fatty regions were manually selected from a set of 8 mammograms (two for each BI-RADS category) that were removed from the

Feature	Window size (in mm)
Original intensity	N/A
Corrected intensity	N/A
Position x, y	N/A
Distance to skin	N/A
Histogram moments	1.96, 4.20, 9.80, and 19.88
Histogram entropy	1.96, 4.20, 9.80, and 19.88
Laplacian moments	1.96, 4.20, 9.80 and 19.88
Lapl. homog. & entr.	1.96, 4.20, 9.80 and 19.88
Local binary patterns	11.48 (10 bins)

Table 3.2: Features used for breast density segmentation along with the scale where they have been computed.

dataset and not used for further testing and subsequent analysis. Therefore, each pixel was characterised using its intensity, neighbourhood texture and morphological information and was used as an input for training the classifier. To extract the training data, we combined two different strategies:

- Manual selection of ROIs being clearly dense or fatty. However, using only these regions, the classifier had the lack of data from those regions where the tissue is not clearly dense or fatty.

- Manual selection of ROIs from regions containing fatty and dense tissue. The pixels in these ROIs were divided using an automatic threshold [237]. Pixels with intensity higher than this threshold were considered as dense, whereas pixels with lower intensity than the threshold were considered as fatty.

One of the main drawbacks of SVM is its large computational time for training. Notice that the training step is performed prior to the testing step, hence allowing a fast segmentation of the mammograms.

3.3 Results

3.3.1 Qualitative Methods

3.3.1.1 Region Based Classification Using Eigen Analysis

The method was applied to 233 mammograms from our local digital database. This subset is composed of left or right MLO mammograms from 233 healthy women. An expert mammographic reader classified all the images in year 2005 using BI-RADS.

We used a leave-one-case-out methodology, i.e., each mammogram is analyzed by a classifier trained using the mammograms of all other women in the database. The confusion matrices for the best three classifiers (RF, AdaBoost and SVM) are shown in Table 3.3, where each row corresponds to results based on the manual classification by the expert. Cohen's kappa (κ) coefficient [66] and the correct classification percentage (CCP) are also shown, defining κ as:

$$\kappa = \frac{P(D) - P(E)}{1 - P(E)} \tag{3.1}$$

where $P(D)$ is the percentage of classified mammograms that are correct and $P(E)$ the expected proportion by chance. A κ equal to 1 means a statistically perfect model whereas a value equal to 0 means every model value was different from the actual value.

And being CCP the sum of the diagonal elements of the confusion matrix divided by the total number of used mammograms.

Without applying feature selection seems that the best classifier for high dense breasts is RF, while for low dense breast SVM achieves the best results. But when feature selection is considered, only AdaBoost and SVM improve remarkably, mainly for BI-RADS III and IV.

3.3.1.2 Region Based Classification Using FCM Analysis

The method was applied to the whole set of 236 mammograms acquired with a Selenia FFDM system that form our local digital database. This database is composed of left and/or right MLO mammograms from 233 healthy women.

No Feature Selection

	RF $\kappa = 0.58$, $CCP = 70\%$				AdaBoost $\kappa = 0.55$, $CCP = 67\%$				SVM $\kappa = 0.57$, $CCP = 69\%$			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
I	30	19	0	1	34	14	0	2	34	15	0	1
II	10	65	8	3	10	66	8	2	7	70	8	1
III	0	13	32	7	0	14	27	11	0	15	27	10
IV	3	1	6	35	4	1	10	30	3	2	10	30

SVM Feature Selection

	RF $\kappa = 0.60$, $CCP = 71\%$				AdaBoost $\kappa = 0.63$, $CCP = 73\%$				SVM $\kappa = 0.67$, $CCP = 76\%$			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
I	34	13	1	2	35	12	1	2	35	14	0	1
II	13	66	6	1	8	69	9	0	6	72	8	0
III	0	13	31	8	0	15	30	7	0	11	34	7
IV	3	1	6	35	3	0	6	36	2	0	7	36

(A) (B) (C)

Table 3.3: Confusion matrices. (A) RF, (B) AdaBoost, and (C) SVM.

Segmentation

To determine the quality of the segmentation results, visual assessment was performed by one observer with more than 10 years of experience in mammographic images. To assess the segmentation improvements, the observer evaluated the differences in the segmentation results when images were enhanced or not. Around 92% of the segmentations obtained after image enhancement were considered similar or better than the ones obtained before enhancement. Specifically, the 45% were regarded as strictly better, therefore results show that there is a clear improvement in the segmentation results when images are previously peripheral enhanced (see Fig. 3.4).

Experts Classification

Four expert mammographic readers classified all the images using BI-RADS in year 2011 (readers are different from the ones that participated in the digitised study [230]). The ground truth was determined by majority vote. In case of

tie, the median value was considered as the consensus opinion (like in [230]).

Expert A ($\kappa = 0.94$, $CCP = 96\%$) — (A)

Consensus \ Expert A	I	II	III	IV
I	84	1	0	0
II	1	86	1	0
III	0	0	39	7
IV	0	0	0	17

Expert B ($\kappa = 0.78$, $CCP = 86\%$) — (B)

Consensus \ Expert B	I	II	III	IV
I	85	0	0	0
II	17	67	4	0
III	0	8	36	2
IV	0	0	4	13

Expert C ($\kappa = 0.70$, $CCP = 79\%$) — (C)

Consensus \ Expert C	I	II	III	IV
I	54	31	0	0
II	0	77	11	0
III	0	2	42	2
IV	0	0	3	14

Expert D ($\kappa = 0.61$, $CCP = 73\%$) — (D)

Consensus \ Expert D	I	II	III	IV
I	51	34	0	0
II	0	85	2	1
III	0	1	18	27
IV	0	0	0	17

Table 3.4: (A)-(D) Confusion matrices for four expert radiologists and their consensus opinion in 2011.

Table 3.4 shows the confusion matrices for the classification of FFDM for the four readers and the consensus opinion in year 2011. Like in the previous work carried out in our research group for digitised mammograms classification [230], the results show an evident interobserver variability, illustrating the difficulty of the breast tissue density classification task. In low dense breasts categories {BI-RADS I & II}, expert B tends to classify in BI-RADS I (17 mammograms were classified as BI-RADS I being BI-RADS II) whereas experts C and D tend to classify in BI-RADS II (31 and 34 mammograms respectively were classified as BI-RADS II being BI-RADS I). Note also that expert B repeats this underestimation assignment when classifying in BI-RADS II (8 mammograms were classified as BI-RADS II being BI-RADS III) and expert C repeats the overestimation when classifying in BI-RADS III (11 mammograms were classified as BI-RADS III being BI-RADS II). In high dense categories {BI-RADS III & IV}, expert D differs from the rest considering a few BI-RADS III mammograms (18/46) and a lot of BI-RADS IV (27 mammograms were classified as BI-RADS IV being BI-RADS III). When considering the individual BI-RADS classes, the CCP values for expert A are really high (99%, 98%, 85%, 100%, respectively). The results of the other experts are less homogeneous and lower, except for expert C in BI-RADS III with CCP = 91%. Using the κ values, the agreement of expert A with the consensus opinion belongs to the *almost perfect* category ($\kappa = 0.94$) whereas the agreement of experts B, C and D with the consensus opinion belong to the *substantial* category ($\kappa = 0.78, 0.70, 0.61$

respectively).

		Expert D (Year 2011) $\kappa = 0.61$ $CCP = 73\%$				Expert D (Year 2005) $\kappa = 0.41$ $CCP = 57\%$			
		I	II	III	IV	I	II	III	IV
Consensus	I	51	34	0	0	49	36	0	0
	II	0	85	2	1	0	52	35	1
	III	0	1	18	27	0	0	17	29
	IV	0	0	0	17	1	0	0	16
		(A)				(B)			

Table 3.5: (A) Confusion matrix for one expert radiologist and the consensus opinion in 2011 and (B) confusion matrix for the same expert radiologist and the consensus opinion in 2005.

Furthermore, a few years after the first experiment, one of the experts classified the same database according to BI-RADS. Table 3.5 shows the confusion matrices for the classification of FFDM for one reader and the consensus opinion, for two different periods in time. Results reveal intraobserver variability in BI-RADS II and III classification. In the past the reader classified 88 mammograms as BI-RADS II whereas now the number increases to 120. On the other hand, 52 mammograms were considered BI-RADS III compared to the current 20. Examining each class, there are no significant variations in CCP values for BI-RADS I (before: 58%, after: 60%), BI-RADS III (before: 37%, after: 39%), and BI-RADS IV (before: 94%, after: 100%), in contrast with the CCP values for BI-RADS II (before: 59%, after:97%).

Automatic Classification

To evaluate our algorithm, we used a leave-one-out methodology, i.e., each digital mammogram is analyzed by a classifier trained using the mammograms of all other women in the database. Table 3.6(C) shows the best confusion matrix after analyzing different feature selection and classification methods. Specifically the confusion matrix is obtained using SVM feature selection (SVS) followed by binary tree of SVM classification (BTSVC) and this combination

achieved a κ of 0.88 and a CCP of 92% (216/236). These values are higher than the values of experts B, C and D although they are lower than the ones of the expert A. When considering each BI-RADS classes, the CCP for BI-RADS I is 93% (79/85), for BI-RADS II is 89% (78/88), for BI-RADS III is 93% (43/46) and for BI-RADS IV is 94% (16/17). Note that BI-RADS III reaches the highest CCP value in comparison with the ones reached by the experts (A: 85%, B: 78%, C: 91%, D: 39%).

Bayesian MIAS ($\kappa = 0.81$, $CCP = 86\%$)

Consensus \	I	II	III	IV
I	79	1	3	4
II	3	86	6	8
III	0	2	85	8
IV	0	6	4	27

(A)

Bayesian DDSM ($\kappa = 0.67$, $CCP = 77\%$)

Consensus \	I	II	III	IV
I	58	25	23	0
II	15	295	26	0
III	12	46	196	1
IV	5	18	18	93

(B)

SVS + BTSVC ($\kappa = 0.88$, $CCP = 92\%$)

Consensus \	I	II	III	IV
I	79	6	0	0
II	5	78	5	0
III	0	2	43	1
IV	0	0	1	16

(C)

Table 3.6: Confusion matrices for MIAS, DDSM and digital databases classification and their respectively consensus opinion: (A) Bayesian combination of KNN and C4.5 classifiers in MIAS, (B) Bayesian combination of KNN and C4.5 classifiers in DDSM and (C) SVS + BTSVC in digital database.

Table 3.6(A)-(B) also shows the best confusion matrices of the previous work carried out in our research group in [230]. In this case the used classifier was a Bayesian combination of KNN and C4.5 classifiers and the classification method was tested using two public databases: the Mammographic Image Analysis (MIAS) database [296] and the Digital Database for Screening Mammography (DDSM) [128] which were obtained from scanned or digitised film images. Although a direct comparison with our previous results is difficult because the datasets used are different, in principle, the confusion matrix of the digital database (Table 3.6(C)) seems to be better than the others because there are less nonzeros off-diagonal elements. When comparing κ and CCP values in digital and digitised databases ($\kappa = 0.81$, CCP = 86% (277/322) for MIAS and $\kappa = 0.67$, CCP = 77% (642/831) for DDSM), they are slightly better in the digital case. Examining the individual BI-RADS classes, the

CCP for MIAS data set were 91%, 83%, 89% and 73% (respectively) and for DDSM were 55%, 88%, 77% and 69% (respectively). All these values are also somewhat better in the digital case (93%, 89%, 93% and 94%), although the highest difference is in BI-RADS IV. Using the two class classification (low vs high density), the CCP for low case is 91% for digital, 87% for MIAS and 80% for DDSM, whereas for high case is 94% for digital, 85% for MIAS and 74% for DDSM, so in both cases, the percentage is higher for digital database. These results make explicit the improvement reached adapting the method from digitised to digital.

Similar comparison can be made between the best confusion matrix obtained with the eigen region based approach (Table 3.3(C) with SVM feature selection) and the confusion matrix of the FCM region based approach (Table 3.6(C)). Although consensus opinion used in FCM region based work was obtained after the eigen region based work, therefore eigen and FCM approaches were evaluated using different consensus opinion.

Using the same criterions, Table 3.6(C) can be also considered better than Table 3.3(C) with SVM feature selection, because has less nonzeros off-diagonal elements and κ and CCP values are quite higher in FCM region based case (0.67 vs 0.88 and 76% vs 92%). Comparing the CCP for each BI-RADS class: 70% vs 93%, 84% vs 89%, 65% vs 93% and 80% vs 94% (respectively), all values are larger in FCM region based case, mainly in BI-RADS III and I. Obviously, when comparing low vs high density CCP (79% vs 91% and 72% vs 94%) results show the same behaviour. So, the preprocessing step and the segmentation algorithm used in the FCM region based method seem to be a significant upgrade from the initial method (i.e., the eigen region based approach).

Note that, for the sake of concreteness, results of the half-way points between the initial quantitative approach (eigen based segmentation with no enhancement) and the last quantitative approach (FCM based segmentation with enhancement) are not presented. Although, it should be pointed that FCM with no enhancement obtained worse results than eigen with no enhancement (see segmentation results in Fig. 3.6(a)-(c)). And eigen with enhancement obtained

a worse segmentation than FCM with enhancement (see segmentation results in Fig. 3.6(d)-(f)).

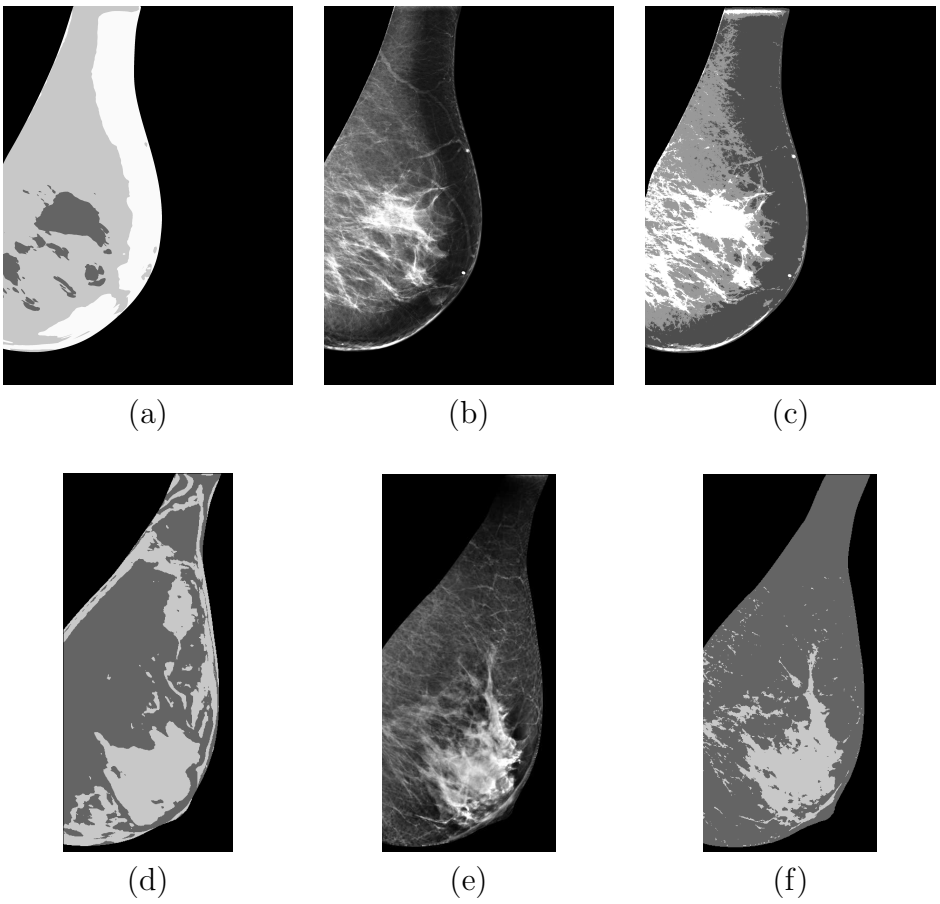


Figure 3.6: Example of the different segmentation processes: (a) eigen without previous peripheral enhancement, (b) original image, (c) FCM without previous peripheral enhancement, (d) eigen with previous peripheral enhancement, (e) original image and (f) FCM with previous peripheral enhancement.

3.3.2 Quantitative Methods

To evaluate the performance of the tool, we used three different strategies. Firstly, using a boxplot, we compared the percentage of dense tissue clustered according to its BI-RADS class. Ideally, the denser the class, the higher the mean percentage of dense tissue should be. This analysis allows us to manually correlate the ground truth labelled by the experts with the results of the tool, and it provides a strong

evaluation of the tool due to the absence of manually segmented images.

To provide a complementary evaluation, we estimated the dense tissue percentage in bilateral breasts (left and right) and in ipsilateral views (MLO and CC views). In the first case, it is well known that the internal tissue distribution is similar, despite being two different breasts, and therefore, the percentage of dense tissue should be highly correlated. In the second case, we are comparing the tissue distribution of the same breast but using different points of view, which should also be closely correlated. These analyses allow us to test the repeatability of the tool.

As a potential use of this tool, we analysed the evolution of breast density during three screening exams. The result of our study was compared with the three different models that Boyd et al. [38] proposed to describe density change over time. Specifically, these models use the 25th and 75th percentile (i.e. first and third quartiles) of the distribution of density percentage to describe the different density behaviours. In all models, the percentage of breast density decreases when age increases, but the change of the interquartile range (IQR) is different depending on the model. In model A the IQR increase with age, in model B does not change, and in model C IQR decreases.

Transversal Analysis

Figure 3.7 presents a box-and-whisker plot representation of the mammographic density percent according to the BI-RADS classification given by the experts. In this type of plot, a box is drawn enclosing the first and third quartiles of the data, and the median value divides the box into two parts. Moreover, the whiskers show the variability outside the upper and lower quartiles. Therefore, each box represents the variability of the automatic estimation of the dense percentage clustered according to the experts' manual annotation. It can be seen in the figure that the dense percentage increases as the BI-RADS class increases, showing a high agreement between manual annotation and automatic estimation.

Moreover, the distinction between BI-RADS I, II, and III is very clear (the boxes are well separated). In contrast, the boxes of BI-RADS III and IV are partly overlap, although the median for BI-RADS IV is outside the enclosed area in BI-RADS III. Additionally, this figure seems to indicate that it is

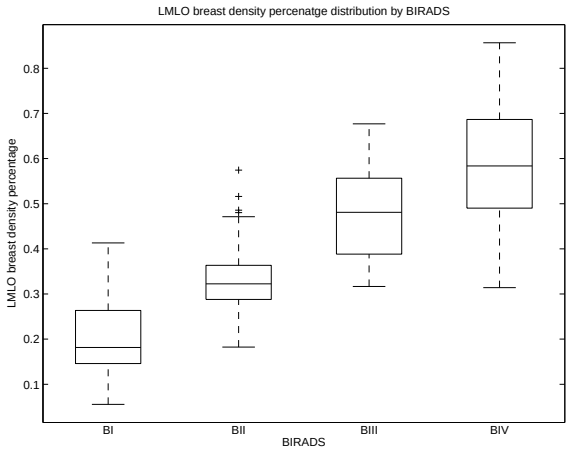
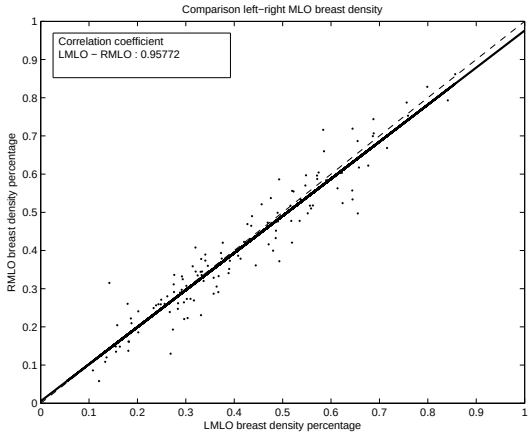


Figure 3.7: Boxplot between BI-RADS and density percent of the segmentation result.

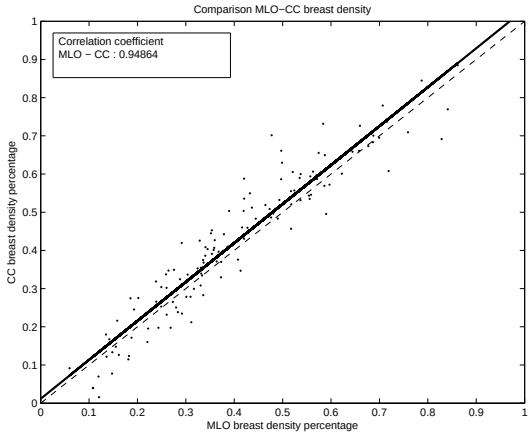
easier to segment dense tissues in low dense mammograms than in highly dense mammograms, where dense regions can be distributed along the mammogram. These claims are reinforced numerically based on analysis of the mean density of each BI-RADS class. Specifically, we obtained the following results: BI-RADS I: 0.20 ± 0.08, BI-RADS II: 0.32 ± 0.09, BI-RADS III: 0.47 ± 0.10 and BI-RADS IV: 0.58 ± 0.13. Upon analysis of the independence of each class using pairwise t tests, we found that the distributions were significantly different (p value < 0.01).

Figure 3.8 compares the percentages of dense tissue in the transversal study. Figure 3.8(a) shows the results of the comparison between dense percentage of all left and right MLO mammograms for the basal exploration (bilateral comparison), whereas Figure 3.8(b) shows the relationship between both views of the same breast (ipsilateral comparison). Note that in each graph, a point corresponds to a particular case. As expected, the segmentation results for bilateral comparison are highly correlated, as indicated by the plotted regression line $\rho = 0.958$ (p value < 0.05). These results can be extrapolated to the fact that mammograms with similar tissue density are segmented with a similar percentage of dense area. Regarding the ipsilateral comparison, the correlation between the dense area segmentation in both views is also very strong (p value < 0.05). However, in contradiction to the bilateral comparison, the slope of

the regression line is greater than 1, indicating that the density percentage in MLO mammograms is slightly lower than the CC view. This fact agrees with the results reported in other studies, where a high correlation between both views and a lower dense percentage for MLO view were observed [49, 318].



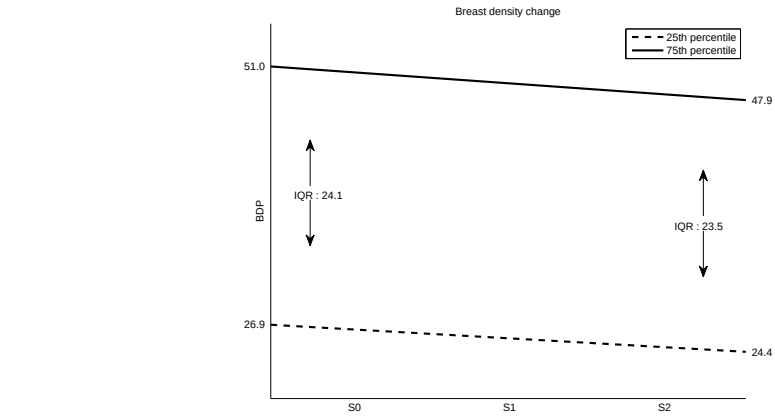
(a)



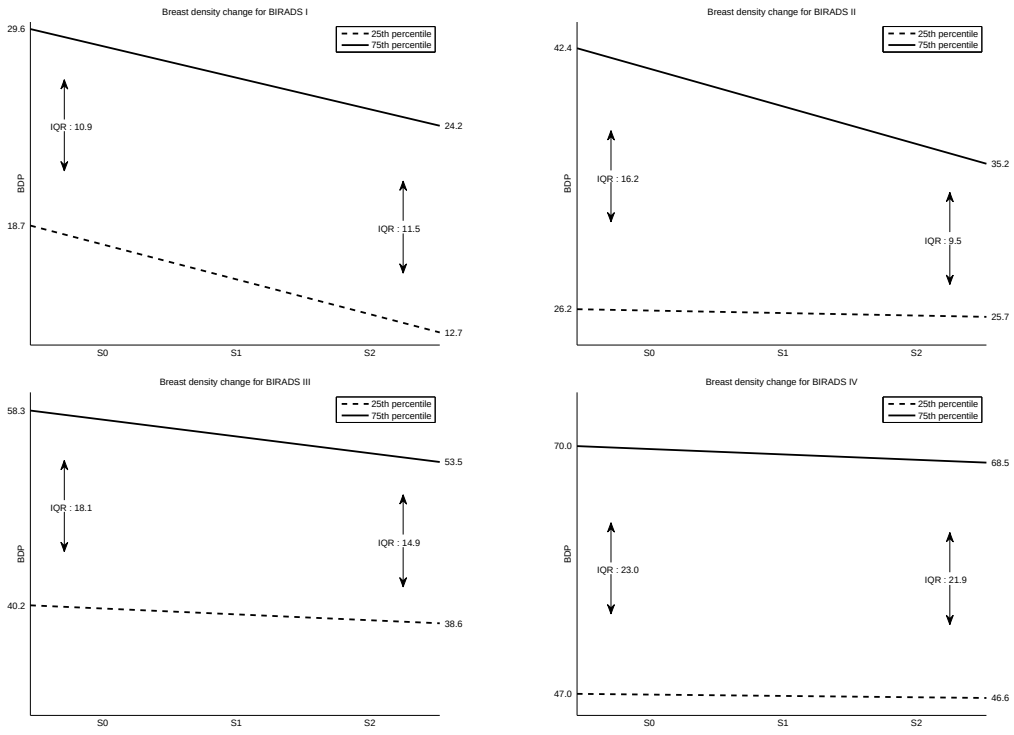
(b)

Figure 3.8: Comparison of the transversal density estimation results. (a) Bilateral comparison and (b) ipsilateral comparison.

Longitudinal Analysis



(a)



(b)

Figure 3.9: First (dotted line) and third (continuous line) quartile evolution of breast density percentage when using (a) all the database together or (b) the database splitted according to the BI-RADS categories (from left top to bottom right BI-RADS I, BI-RADS II, BI-RADS III, and BI-RADS IV). S0 contains the first follow up exams, S1 the second follow up, and S2 the third follow up. Notice that the behaviour is different depending on the density category.

Regarding the analysis of the density evolution, we computed the density percentage of all patients in the first, second and third acquired studies. Afterwards, and similarly to the boxplot analysis, we derived the first and third quartiles at each time in order to obtain the evolution of the breast density, as was performed by Boyd et al. [38]. Figure 3.9(a) shows the regression line for the first and third quartiles when including the whole dataset in the study. As expected, breasts tend to decrease their density with time, and it seems that the decrease in the first quartile is slightly slower than the decrease in the third quartile.

We repeated the same procedure for all of the patients but clustered the mammograms according to the BI-RADS density categories assigned by the experts in the first study. These results are shown in Figure 3.9(b). Notice that although both quartiles decrease their density in all categories, the behaviour is different in each of these categories. Specifically, for BI-RADS I, the decrease in the ratio of both quartiles is significant and almost parallel, although the decrease of the first quartile is slightly greater than that of the third one. For BI-RADS II and III, the decrease in the third quartile is significant, whereas the first quartile remains almost unaffected. Finally, for BI-RADS IV, there is only a slight decrease in both the first and third quartiles.

Furthermore, for each case within the database, we computed a line of best fit for density percentage variation over time, and we analysed the different slopes. As expected, the mean slope for all women in the database was negative (-0.02), indicating a density decrease. Upon analysing the data in terms of BI-RADS category, the mean slope was also negative. Specifically, for BI-RADS IV, we obtained the lowest mean slope (-0.003).

3.4 Discussion and Conclusions

3.4.1 Qualitative Methods

We have provided two qualitative approaches for breast density classification. Results of the initial approach demonstrate the improvement of using a feature selection

step to reduce the huge number of features (412), since all the classifiers obtained better results after feature selection.

To improve our initial results, we propose an alternative to deal with the over-exposed breast area and we explore other segmentation techniques. After using our own peripheral enhancement method (Chapter 2) and a FCM clustering algorithm, we obtained a κ of 0.88 and a CCP of 92% (that represents a better agreement in 3 out of 4 radiologists that classify the images). These results are better than the initial ones ($\kappa = 0.67$ and $CCP = 76\%$), which indicates that the included changes improve the overall method.

After testing different classifiers, SVM obtained the best results in both qualitative methods. Since SVM is originally designed for binary classification, we convert our multiclass problem in several binary class problems to optimize SVM results, and as expected, our binary tree of SVM obtains better results than the multiclass SVM.

The comparison with other automatic approaches, it's a difficult task because the datasets used are different, although the confusion matrices can be compared, and also the κ and the CCP values. Comparing these parameters, results of the FCM region based approach are also better than the previous work carried out in our research group [230] using MIAS ($\kappa = 0.81$ and $CCP = 86\%$) and DDSM ($\kappa = 0.67$ and $CCP = 77\%$).

3.4.2 Quantitative Methods

A novel automatic tool developed for breast density tissue segmentation has been validated in this work. The bilateral comparison of the results has shown a very strong correlation that agrees with previous studies that analysed manually segmented mammograms. Additionally, the ipsilateral comparison also showed a very strong correlation between both views and with almost the same slope reported in previous studies performed with manually segmented images.

Comparisons with other automatic approaches that segment the breast according to its density is not an easy task because it is not feasible to obtain a ground truth detailing all the dense regions in a large set of mammograms. In addition to this difficult computational task, inter and intraexpert variability is very high when

looking at breast densities [230]. However, by considering the location of dense areas that plays a key role in its classification, could be seen as an improvement over the threshold based approaches or the clustering approaches that just use image intensities. We also tested the use of different classifiers typically used in mammography, such as a neural network [151] and a boosting algorithm [235]. The performance of both classifiers was worse than that obtained with SVM, with slightly better results obtained using the neural network compared to the boosting classifier. The main drawback of SVM is its computational cost, mostly due to the amount of time needed to compute the features. Note that once the features are computed for each pixel, the classification can be quickly performed by parallelising techniques.

The longitudinal analysis confirms the well known fact that breast density decreases with age. Although there are other factors involved in breast density variations, such as menopause status, pregnancy and childbearing [37], in this work, we focused on the evolution with age, comparing our results with the theoretical models proposed by Boyd et al. [38] based on the IQR. According to the results obtained with our database and graphically shown in Figure 3.9(a), we observed that breast densities decreased over time, whereas the IQR slightly decreased. Therefore, our data seems to be congruent with model C, although the decrease is so slight that it could also be congruent with model B. However, depending on the BI-RADS score, no single model always described the density behaviour of all the women. Specifically, for BI-RADS I, densities decreased with age, whereas IQR increased; therefore, model A or B should have been assigned in this case. On the other hand, BI-RADS II and III were clearly well described by model C. Finally, the BI-RADS IV category label should be assigned to model B or C. Our database was composed by only three screening studies per woman, which limits the fitting of this model. More studies for each patient are required to provide a better adjustment. Additionally, the lower average slope obtained for high density breasts indicates that the group with an increased risk of breast cancer [51] has a slower decrease in the density of the breast.

There are some limitations in our study. The lack of manual annotations of dense regions in a large dataset of mammograms prevents not only the comparison between manual and automatic segmentations using quantitative overlapping measures but also the use of automatic feature selection algorithms. On the other hand, regarding the longitudinal analysis, 6 years is a short time to detect changes

in breast density, and this does not allow significant clinical conclusions to be made. However, automatic tools allow the detection of density changes in this short period of time.

Automatic tools based on computational approaches allow accurate estimation of breast density and quantification of changes with time. Longitudinal changes in breast density are dependent on the internal density of each breast, in addition to other factors. In our experiments, density changes in low density breasts presented a more heterogeneous behaviour than those in high density breasts, where the percentage of dense tissue seemed to be more stable over time.

Chapter 4

Image Registration

Image registration is a process that studies how to map one image into another. The complexity of the image registration method depends on the geometric differences between images. In this chapter, we evaluate different techniques to align mammographic images, as well as the possibility of use this information in other fields such as automatic mass detection. Left and right mammograms are registered to analyse bilateral dissimilarities. Current and prior mammographic images registration is used to evaluate temporal changes in the breast.

4.1 Introduction

Image registration is the process of finding an optimal transformation function that overlays one image into another. A registration framework is composed by a transformation, an interpolator, a metric and an optimizer (see Fig. 4.1). Two images are used as input data, the fixed or target image and the moving or template image which will be transformed to be as similar as possible to the fixed one. During the transformation step, the points of the moving image are mapped into the points of the fixed one and the interpolator calculates the intensities of the moving image at non-natural positions. Next, a metric measures the similarity of the fixed image and the transformed one. Similarity measurement is used as input by the optimizer to recalculate the transformation parameters in order to find a better mapping or otherwise, to finish the process.

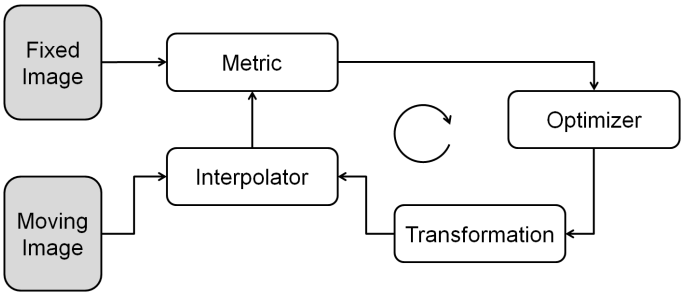


Figure 4.1: Registration framework.

- **Transformation**

  The geometric differences between fixed and moving images can be linear or non-linear, so transformation functions can be divided according the type of the geometric deformation in linear or global and non-linear or local [283].

  – *Linear Mapping*

    In linear mapping models all pixels suffer the same transformation. Translation, rigid transformation (translation and rotation), similarity transformation (translation, rotation and scaling) or affine transformation (translation, rotation, scaling and shearing) are examples of the used models [283], but rigid and affine transformations are the most commonly used [140].

    For medical images, **rigid transformation** is considered a good approximation since satisfies the rigid-body constraints [193].

    Given the original point (x,y), the rigid transformed point (X,Y) is defined as:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

    where θ is the rotation angle and (t_x, t_y) is the translation vector.

    Affine transformation is more general than the rigid one and allows additional information about scaling and shearing factors that can be useful when scaling parameters are unknown/incorrect or when images are skewed during the acquisition [193, 93].

    Affine transformation is defined as:

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

where (x, y) are the original points, a, b, c, and d are the affine parameters, (t_x, t_y) is the translation vector and (X, Y) are the transformed points.

– *Non-linear Mapping*

Although global mapping may be sufficient in many applications, the geometric difference between images can be complex and sometimes a more flexible transformation model is needed, allowing for local deformation [359]. In non-linear mapping models, pixels are transformed locally having a different transformation depending on their position. There are a huge variety of non-linear geometric transformations, for instance transformations derived from physical models, transformations inspired by interpolation theory or transformations based on prior knowledge. Physical models category includes elastic body models or diffusions models, among others. Within the group of interpolation theory, radial basis functions, free form deformations or locally affine models are found [286].

Diffeomorphic Demons belongs to the diffusion models group within the physical models category and has been used successfully in variety of registration scenarios including brain MRI [164, 277, 71], cardiac CT [248], breast CT [268] or breast MRI registration [238]. The Diffeomorphic Demons [330] algorithm derives from Thyrion's Demons [306]. The mapping function used in Diffeomorphic Demons registration is continuous, smooth and invertible and its inverse is also smooth (by definition of diffeomorphism). Furthermore, Diffeomorphic Demons registration formalises Demons optimization over the space of displacement fields and has different variants corresponding to the operation allowed in the space of deformation field: exponential, addition and composition.

Thin-Plate Spline belongs to the radial basis functions group within the interpolation theory category. Thin-Plate Spline is one of the most commonly used methods for non-linear medical image registration [147] and is applied, among other medical fields, in registration of brain images [147], breast images [334, 199], prostate images [329, 212] or liver

images [176]. Thin-Plate Spline provides smooth deformations with easily controlled behaviour [346, 199].

Given n control points x_i, the Thin-Plate Spline mapping function is defined as:

$$X = a_o + a_1 x + a_2 y + \sum_{i=1}^{n} c_i g(x, x_i)$$

Thin-Plate Spline is composed by an optional low-degree polynomial ($a_o + a_1 x + a_2 y$) plus a sum of weighted Thin-Plate Spline radial basis functions with the form:

$$g(x, x_i) = \|x - x_i\|^2 \ln(\|x - x_i\|).$$

B-Splines free form deformation belongs to the free-from deformations group within the interpolation theory category. Free form deformation modelled by B-Splines is popular in medical applications (brain CT [174], brain MRI [260], breast MRI [261], cardiac US [179], lung CT [326] or x-ray mammography [72, 59] image registration). Free form deformation using B-Splines is based on deforming an image by modifying a mesh of control points following a maximization of a similarity measure. These control points define a mesh of smooth and continuous B-Splines functions with limited support, i.e, control points can be locally controlled and modifying a control point only affects neighbouring points. The degree of deformation of the mesh can be modelled with the resolution of the mesh (coarse meshes are more suited for large scale transforms and finer meshes for local deformations) [72].

- **Interpolator**

The nearest neighbour interpolation [190, 180, 305, 98, 6, 40], the linear interpolation [180, 305, 98, 6, 5, 333, 40], the cubic interpolation [180, 305, 98, 40], the cubic splines interpolation [180, 6, 5] or the sinc interpolation [180, 305, 6, 5] belong to the most commonly used interpolators for calculating the new grey values of the moving image after transformation [359].

Linear interpolation is one of the simplest, fastest and popular interpolation methods [305, 12, 92]. Linear interpolation estimates the grey value of the

transformed image (y) by computing the straight line between the nearest neighbours.

Given x in the interval (x_0, x_1):

$$y = y_0 + (y_1 - y_0)\frac{x - x_0}{x_1 - x_0}$$

where y_0 and y_1 are the grey values of x_0 and x_1 respectively.

- **Metric**

  Different similarity measures can be minimised/maximised in order to optimize the transformation parameters [193]. For instance, sum of absolute differences [136, 77], cross-correlation [160, 17], SSD [171, 112, 113, 336, 73, 41] or MI [261, 267, 346, 322, 199, 337, 309, 73] among others.

  SSD is commonly used for registering images of the same modality based on the idea that corresponding pixels have equal grey values [286], and also due to its simplicity and fast computation [67].

$$SSD = \sum_{(i,j)\in F} (I_F(i,j) - I_M(i,j))^2$$

  where F is the fixed image, $I_F(i,j)$ is the intensity value of F in the (i,j) position and $I_M(i,j)$ is the intensity value of the moving image (M) in the (i,j) position.

  MI is also a robust similarity measure for image registration problems [249]. Although MI is less strict than SSD and assumes only a statistical relation between the intensities of the images [165].

$$MI = H(F) + H(M) - H(F,M)$$

  where $H(F)$, $H(M)$ denote the marginal entropies of the fixed image F and the moving image M respectively and $H(F,M)$ denotes their joint entropy, which is calculated from the joint histogram of F and M.

- **Optimizer**

  To obtain the optimal transformation parameters, several methods can be applied, the most commonly used [107, 333, 286] include: gradient descent method [171, 191, 261, 257, 256, 113, 336, 166], conjugate gradient method [119, 191, 336, 166], Levenberg-Marquardt algorithm [149, 12], Newton-Raphson method [362, 209], quasi-Newton method [191, 166], Neyer-Mead downhill simplex method [191, 208], and Powells method [191, 75, 333].

  Gradient descent method belongs to the first order derivative methods, since requires the first order derivative of the objective function at each iteration. Being $C(p)$ the cost function at p (transformation parameters), gradient descent finds p that minimises C by following the direction that decreases the energy. At each iteration, the current position is updated according to:

$$p_{n+1} = p_n - \alpha_n \frac{\partial C}{\partial p}(p_n)$$

  where α_n is the step size.

Although image registration is used in a wide variety of fields, in this chapter, we focus on medical imaging, specifically in mammographic images. Most of the mammographic published approaches focus on registering image features such as breast boundary [265, 355], nipple [290, 33] or internal regions, for example, the pectoral muscle [98], salient regions [167, 198] or internal linear structures [200]. But also, there are approaches that only use intensity information [47, 322] or combine feature-based and intensity-based techniques [346, 257, 20].

Registration methods are often used independently, however it is commonly accepted that results can be improved in terms of accuracy and robustness by combining different approaches [193, 346, 199, 118, 182]. Algorithm combination exploits the benefits of the different methods, for instance using a global and a local method. In this case, the global method recovers for main pose and scale differences and the local method accounts for localized non-linear deformations.

Results can also be improved by using a MR approach [317, 92] since by reducing the size of the images reduce the geometric difference between them and simplify the correspondence process [105]. Registration is first achieved between images at

the lowest resolution, propagating parameter estimation into a higher resolution and registering again [257, 256, 182]. This often avoids local minima in the parameter search space and reduces computational time.

Van Engeland et al [322] report that linear transformations obtain more robust results compared to non-linear methods in terms of producing less non realistic distortions in mammographic images. However Díez et al. [72] evaluated a larger number of linear and non-linear intensity-based algorithms and showed that although these unrealistic deformations can occur, well parametrised non-linear algorithms obtain a better overall performance. Next, in Section 4.2, different linear and non-linear intensity-based approaches are proposed, as well as some combination of linear and non-linear methods and some MR approaches. The results of the registration methods are evaluated according to different strategies in Section 4.3 and finally conclusions are presented.

4.2 Proposals

Image registration is regarded as an important tool for the analysis of bilateral and temporal mammograms [265, 196, 257, 97, 322, 91, 113]. Although the clinical use of the image registration results themselves is unusual due to the computational cost and the complexity of validating the results, especially in non-linear approaches [69]. However, image registration results are useful for automatic breast abnormalities detection systems [177, 360, 155, 33, 205, 22, 189, 338].

Note that in bilateral and temporal registration of digital mammograms, a preprocessing stage is applied. During this stage, the pectoral muscle was manually segmented. Background and pectoral muscle areas were also eliminated to preserve the breast area. To reduce computation time, images were resized (quarter size image) using bilinear interpolation. In addition, mammograms were flipped when necessary to match the orientation of the breasts in each mammogram. After that, a peripheral enhancement was applied to compensate thickness variations in breast periphery (see Chapter 2).

4.2.1 Bilateral Registration

Bilateral comparison of mammographic images is justified from a clinical point of view by Kopans [168], who makes two important observations when comparing bilateral mammograms of the same woman: though one breast may be larger than the other, internal structures are quite symmetric over broad areas, and overlapping tissue structures that form summation shadows and normal tissue variations on the mammogram highlight asymmetries. Also, the radiologists Tabár and Dean [299] developed a method for looking for lesions based on comparing corresponding regions of the left and right mammographic images to detect differences that could be lesions.

Based on the popularity of different algorithms for mammography registration, bilateral mammographic images were registered by using Rigid, Affine and B-Splines transformation methods. Furthermore, all these methods were used to study the role of some commonly used registration aspects such as Affine initialisation and MR approach. Specifically, Rigid (R), Affine (A), Affine with MR (mA), B-Spline (B), B with MR (mB), A in combination with B (AB) and mA in combination with mB (mAmB) were evaluated. In latter digital works, additive Diffeomorphic Demons was also evaluated for bilateral registration.

Transformation parameters were recovered by maximizing two different similarity measures: SSD and MI. Finally, linear interpolation and gradient descent method were used as interpolator and optimiser respectively.

4.2.2 Temporal Registration

Temporal comparison of mammograms is also an important part of the diagnostic procedure. When older mammograms are available, radiologists compare them with more recent images, since interval changes analysis allows detection of new lesion over time and assessment of the lesion growing [195, 308, 285].

The main goal of temporal work was to evaluate the suitability of Diffeomorphic Demons for temporal mammographic registration, in particular the additive variant. For comparison purposes, Affine and B-Splines were also considered, as well as Affine initialisation and MR approach. Specifically, A, Diffeomorphic Demons with MR

(mD), A in combination with Diffeomorphic Demons (AD), mA in combination with mD (mAmD), mB, AB and mAmB were evaluated.

Like in bilateral work, SSD and MI were used as metric, linear interpolation as interpolator and gradient descent as optimiser, except for Demons based methods that use an optimization function defined over the space of displacement fields.

4.3 Results

All registration methods were implemented using the Insight Toolkit (itk) libraries [141]. We used 128 histogram bins and 10000 samples for metric computations. A minimum step length stopping criteria was also used. For practical reasons we also fixed a maximum number of iterations for all methods to a maximum of 1000 iterations. In combined or MR methods these iterations were evenly distributed between the methods or MR levels.

Metric values are the most widely used tool to measure image similarity in computer vision. Concerning their application to intensity-based registration of medical images, their use is twofold: first they provide the quantification of similarity between images. Second, the registration process in usually formalized as an optimization problem and metric values determine both stopping conditions and solution update. This is true for Affine and B-Spline registration but not for Demons, as these methods use an optimization function defined over the space of displacement fields.

Consequently, metric measurements provide an objective value on how successful some methods were in terms of optimization and how similar two images are in their own particular terms. Unfortunately, a metric that is able to express exactly what medical experts perceive as a "better" registration does not yet exist, so other criteria are also necessary. In the results presented, SSD and MI are used as similarity measurement.

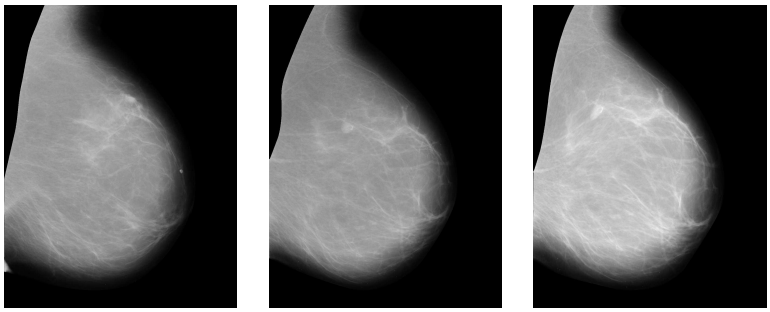
We used two different subsets of mammograms, one was composed by digitised images and the other by digital images.

4.3.1 Digitised Mammograms

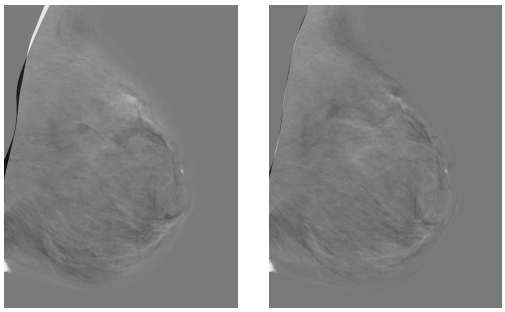
In initial experiments, we used 208 digitised mammograms extracted from MIAS database [296], including 104 pairs of left and right MLO mammograms. In total, there were 52 mammograms containing at least one mass.

4.3.1.1 Bilateral Registration

Figure 4.2 shows an example of the registration results of our initial experiments. Fig. 4.2(b) is the mammogram which is registered with Fig. 4.2(a). The resulting mammogram is shown in Fig. 4.2(c), while Fig. 4.2(d) and Fig. 4.2(e) shows the difference of the images after and before of the registration. Note that Fig. 4.2(e) is more homogeneous than Fig. 4.2(d), indicating that the mammogram after registration is more similar to the target mammogram without registration.



(a) (b) (c)



(d) (e)

Figure 4.2: mAmB registration example of mammographic images: (a) fixed image, (b) moving image, (c) registered image, (d) difference image before registration "(a)-(b)", and (e) difference image after registration "(a)-(c)".

For evaluating the results of registration methods we computed similarity metrics before (BEF) and after registration to obtain an indication of how similar the bilateral images were. A higher similarity after image registration was expected and the method with the highest similarity was considered the best.

Table 4.1 presents the numeric results for the complete digitised subset for both metrics. As can be seen, results are useful to compare registration methods using the same metric but is not possible to compare results between metrics due to the fact that the results are related to the metric optimized by the algorithm. With both metrics the standard deviation (SD) is large with respect to mean values. This can be explained by the heterogeneity of the images and, nevertheless, SD decreases after registration. The methods that obtained the best results were mB for the SSD metric and mAmB for the MI metric.

	SSD		MI	
	Mean	**SD**	**Mean**	**SD**
BEF	2369.33	1656.87	0.73	0.24
R	795.84	502.77	0.83	0.27
A	503.03	310.15	1.05	0.21
B	277.51	131.69	1.34	0.23
mA	788.25	476.13	1.05	0.21
mB	218.56	100.93	1.34	0.23
AB	276.69	143.14	1.37	0.21
mAmB	221.30	111.29	1.38	0.22

Table 4.1: Evaluation of the analysed registration methods for bilateral digitised mammograms.

Once the best methods for each metric were chosen according the numeric results, images after registration were reviewed by different observers to qualitative evaluate the registration results. MI obtained better results than SSD, because in some SSD registered images appear artefacts and unrealistic structures (see Fig. 4.3).

4.3.2 Digital Mammograms

Images used in latter experiments were obtained from our local digital database. We used 584 FFD mammograms from 86 patients with breast cancer. 30 patients have one study, 52 patients have two studies and 4 patients have three studies.

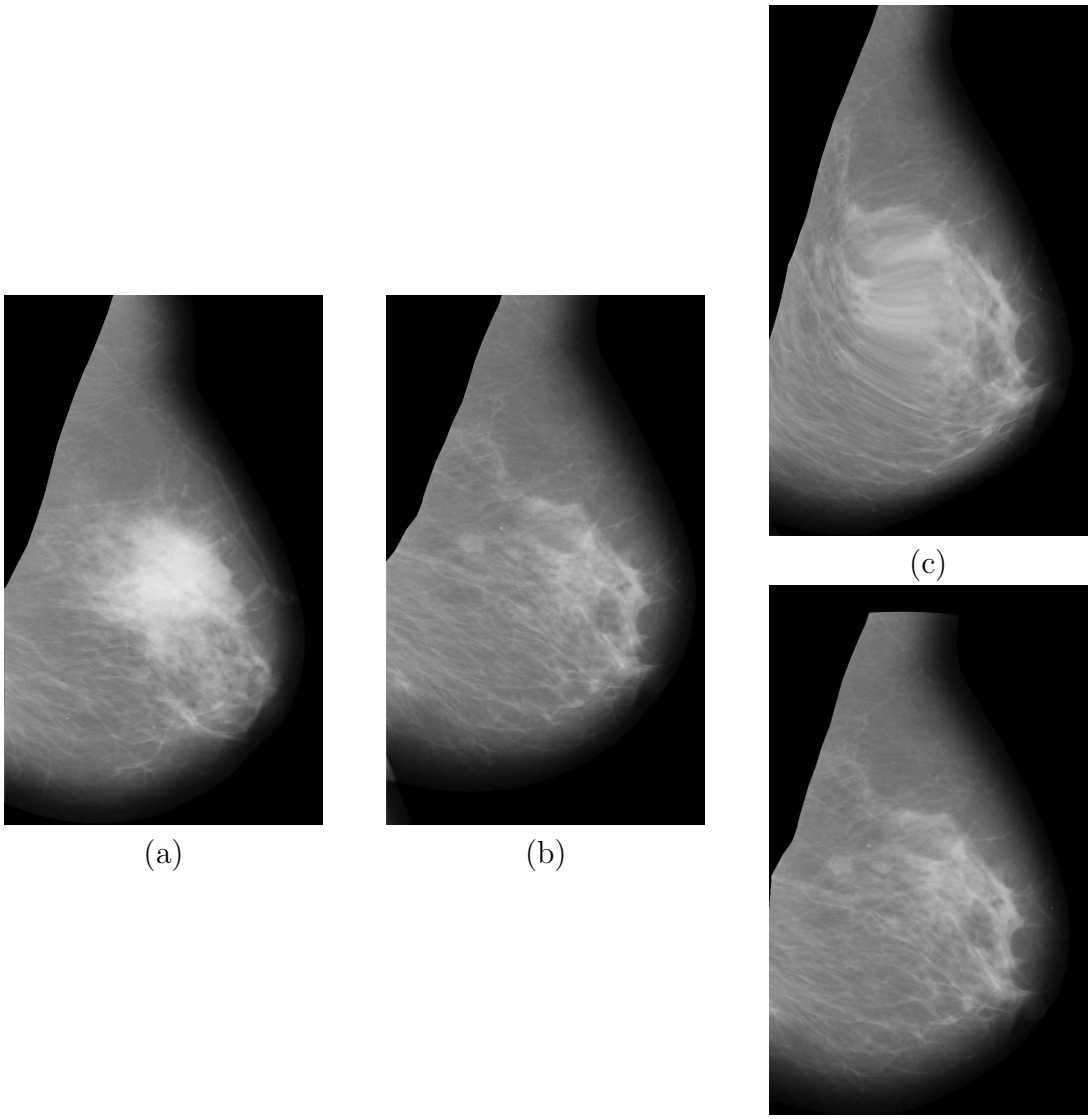


Figure 4.3: Example of SSD distortions: (a) fixed image, (b) moving image, (c) registered image with SSD as similarity metric, and (d) registered image with MI as similarity metric.

Each study contains four mammographic images, two views (CC and MLO) of left and right breasts. Mammograms were acquired using a Selenia FFD mammography system, with resolution 70 micron per pixel, size 4096x3328 or 2560x3328, and 12 bit depth. Presence of masses was annotated by expert radiologists and this allowed us to distinguish between those registration instances containing masses and those

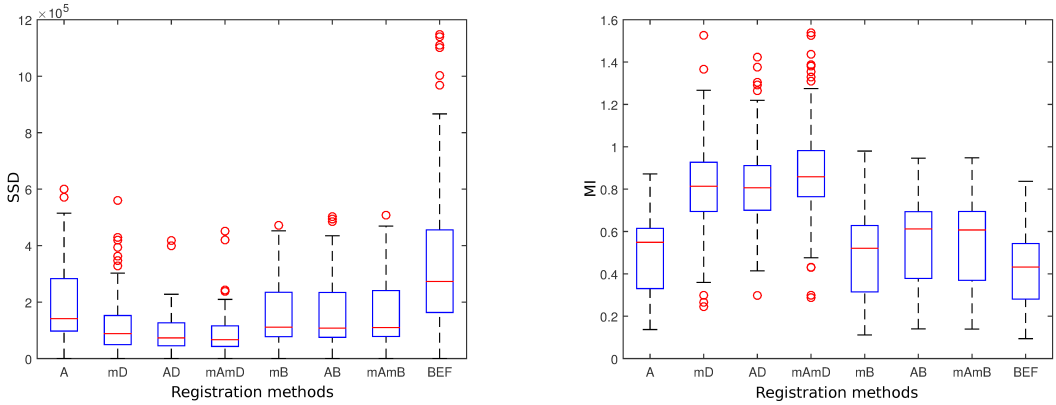
not containing them.

4.3.2.1 Temporal Registration

For temporal comparison of mammograms, each mammographic image was registered with its homonymous mammogram in posterior studies using different registration methodologies doing in total more than 3450 temporal registrations. We use boxplots as a compact way to describe thousands of data. We present several multiple boxplots where each box will group the data resulting from a registration method.

Figure 4.4 shows compared performances of the different registration approaches evaluating the similarity between resulting images with SSD (see Fig. 4.4(a)) and with MI (see Fig. 4.4(b)). We see how Demons based methods obtain better results than other approaches using both, SSD and MI measurements and we also observe how Demons improve using Affine initialisation and applying MR with both similarities.

For B-Splines based methods, Affine initialisation and MR approaches do not seem to be an improvement when SSD measure is calculated, however it obtains a better result for MI measurement. The three B-Splines based methods (mB, AB, mAmB) obtain similar SSD values. When comparing MI values, AB and mAmB



(a) (b)

Figure 4.4: Temporal registrations: (a) SSD (lower positive values stand for better results), and (b) MI (higher positive values stand for better results).

outperform mB and the difference among them is not really clear. Similar trends were observed in bilateral digitised registration results.

Taking into account that MI was considered better than SSD in bilateral digitised registration work, we have focused on MI data analysis. MI data is provided separately for registrations without and with masses in Figure 4.5. Demons based methods obtain better results than other approaches independently of the presence of masses. Despite separation, Demons and B-Splines do better when using Affine initialisation and applying MR. If only one of the two improvements is used, Demons based methods seem to benefit more from MR while B-Splines benefits from Affine initialisation.

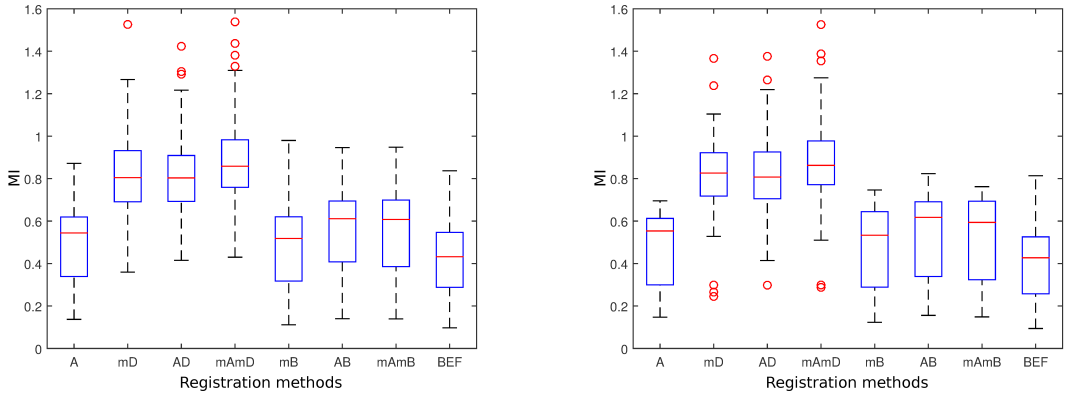


Figure 4.5: Temporal registrations without masses (a) and with masses (b). Values correspond to MI, therefore higher positive values stand for better results.

Concerning the presence of masses, both median and mean MI values slightly increase for Demons based methods when masses are present (see Table 4.2). The fact that images with masses obtain higher MI values than images with no masses could be considered as an indicator of mass presence. This suggests the possibility of using Demons based methods in registration applications such as mass detection. This possibility will be addressed further in Deformation Fields Analisys subsection in Section 5.3.2 and in Section 5.4.2.

Figure 4.6 shows an example of Demons and B-Splines registration methods. Although Demons based methods obtain better numeric results than B-Splines based methods, as expected, Demons results are not visually understandable or clinically

	Median		Mean	
	Without Masses	**With Masses**	**Without Masses**	**With Masses**
mD	0.80	0.83	0.80	0.82
AD	0.80	0.81	0.80	0.82
mAmD	0.85	0.86	0.87	0.88

Table 4.2: Evaluation of Demons registration methods for temporal digital mammograms.

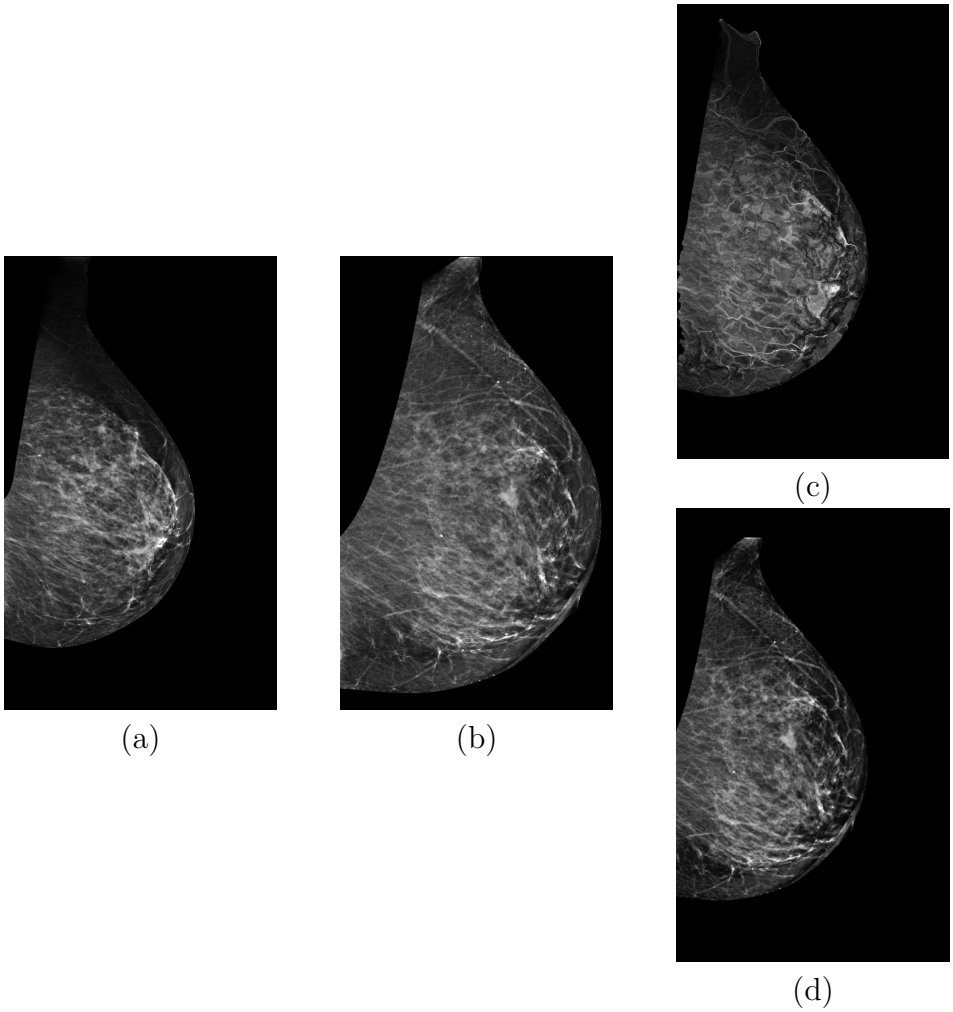


Figure 4.6: Example of Demons and B-Splines registrations: (a) fixed image, (b) moving image, (c) mAmD registered image, and (d) mAmB registered image.

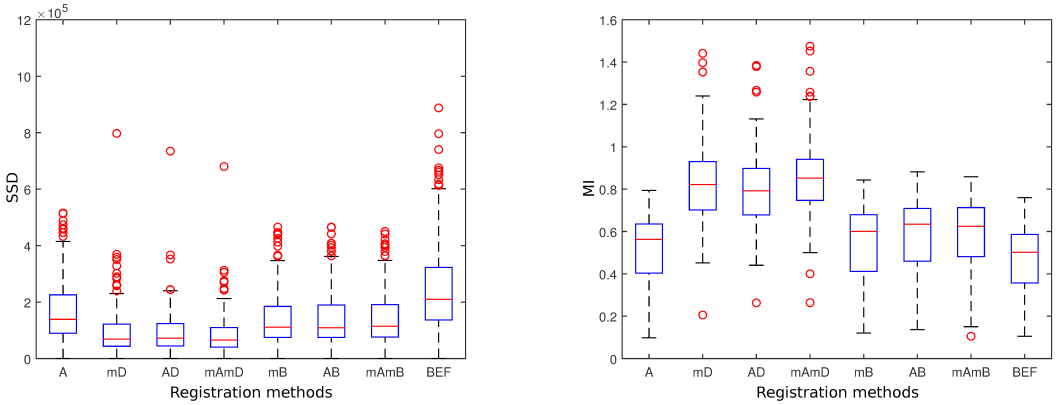
usable (see Fig. 4.6(c)), exactly the opposite of B-Splines (see Fig. 4.6(d)).

4.3.2.2 Bilateral Registration

For bilateral comparison of mammograms, each right mammographic image was registered with its left homonymous mammogram. More than 4000 bilateral registration were analised. In general, bilateral data values are quite similar to the temporal ones, there is no a method that stands out depending on the nature of the registered images (bilateral or temporal). The same as the temporal case, Demons based methods obtain better results than other approaches using both measurements, SSD and MI (see Fig. 4.7). Regarding the improvement of using an Affine initialisation and MR, we also see how Demons based methods improve their results when similarity assessment is calculated with MI, even though this is more subtle for SSD measure.

Although in bilateral and temporal digital cases, mAmD obtains the best results, the behaviour of AD and mD is different. For bilateral digital registrations mD does better or similar than AD (with MI and SSD assessment respectively), however for temporal digital registrations AD works better or similar than mD (when using SSD and MI as metric respectively).

Bilateral B-Splines based methods have similar performance to temporal ones.



(a) (b)

Figure 4.7: Bilateral registrations: (a) SSD (lower positive values stand for better results), and (b) MI (higher positive values stand for better results).

For SSD metric, the three approaches (mB, AB, mAmB) have closest results and for MI metric, AB and mAmB work similarly and better than mB. Therefore, B-Splines based methods seem to be more stable than Demons based method, independently of the nature of the images.

When we focus on MI data, as in the temporal digital case, Demons based methods obtain the best results in both cases, without and with masses (see Fig. 4.8). In addition, Demons and B-Splines work better after Affine initialisation and MR application. Similar to the temporal case, when techniques are analysed independently, MR approach benefits more Demons methods than Affine initialisation. Conversely, for B-Splines methods, the Affine initialisation improvement is higher than the MR one.

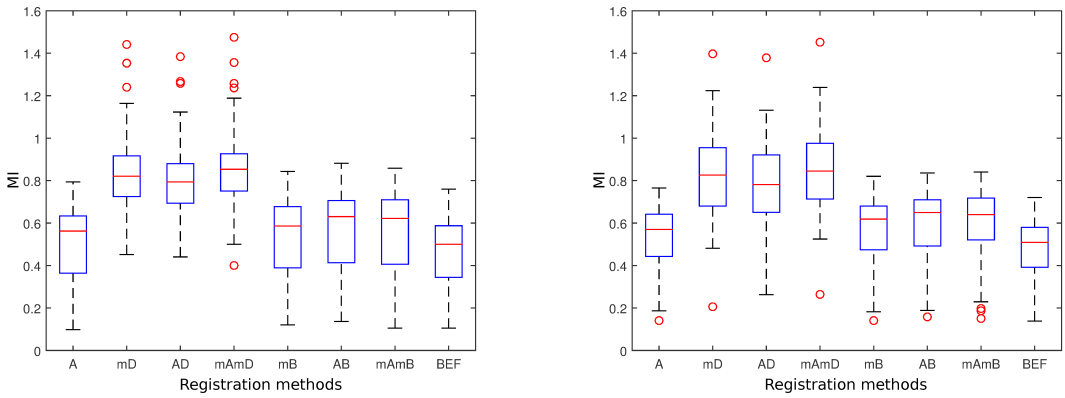


Figure 4.8: Bilateral registrations without masses (a) and with masses (b). Values correspond to MI, therefore higher positive values stand for better results.

	Median		Mean	
	Without Masses	**With Masses**	**Without Masses**	**With Masses**
mD	0.82	0.83	0.83	0.82
AD	0.79	0.78	0.78	0.79
mAmD	0.85	0.84	0.84	0.85

Table 4.3: Evaluation of Demons registration methods for bilateral digital mammograms.

In contrast with the temporal digital results, Demons median and mean values

do not increase when a mass is present (see Table 4.3). If Demons median values increase then mean values decrease and viceversa. So, in principle bilateral digital deformations do not allow the possibility of using them for mass detection.

For comparing digital and digitised bilateral registration results (see Table 4.1), we present a table with mean and SD values of SSD and MI metrics (see Table 4.4). Values are calculated by using only bilateral registrations that contain at least one mass, like in digitised bilateral registrations analysis. SSD values are larger than the digitised ones, because digital images are 12 bits instead of 8 bits. B-Splines based methods obtain the best results for both, SSD and MI. For SSD measurements, mAmB obtains the lowest mean value, although mB could be considered the best method, as in the digitised case, due to the negligible difference between the mean values and also since mB has a lower SD value. For MI metric, the highest values are obtained with AB and mAmB, the same trend as in the digitised experiment, although the difference between these approaches is not really significant.

	SSD		MI	
	Mean	**SD**	**Mean**	**SD**
BEF	240217.14	141286.79	0.47	0.15
A	169057.34	106847.13	0.53	0.16
mB	137611.93	90849.03	0.56	0.17
AB	139618.50	94333.26	0.59	0.18
mAmB	137427.87	92561.67	0.59	0.18

Table 4.4: Evaluation of the analysed registration methods for bilateral digital mammograms.

4.4 Discussion and Conclusions

We have presented results using over 7500 bilateral and temporal registrations corresponding to breast cancer patients. The analysis of bilateral registrations over digitised images shows how B-Spline based methods outperform Rigid and Affine registrations. We have also studied the use of MR and Affine initialisation. There is no a significant difference when both techniques are used together, however when they are applied individually, for SSD, MR approach improves registration results,

while for MI, Affine initialisation makes a difference. Since it is not possible to compare results between metrics due to the different nature of the measurements (direct pixels difference vs entropies analysis), qualitative assessment criteria was used to evaluate results according the metric used and MI is considered better than SSD.

According to our digital study, Demons based methods provide the best results in terms of measures commonly used (SSD and MI) to evaluate the quality of registration. However, B-Splines methods show more stability and similar trends, independently of the type of the registered images (temporal or bilateral digital mammograms).

In general, local methods (Demons and B-Splines) show the best results when combining Affine initialisation and MR approach. This is clearer for MI metric than for the SSD. However, both improvements do not seem to be significantly necessary, since only the MR approach seems to provide enough improvement for Demons methods and only Affine initialisation seems to offer suitable enhancement for B-Spline.

The analysis of digital registration data shows how Demons based methods are better than Affine and B-Splines based approaches. Although the unrealistic deformations obtained from Demons based methods, indicates that the direct image difference between registered images cannot be used by automatic breast abnormalities detection systems. However, temporal Demons based methods show application possibilities in specific fields such as mass detection. Therefore, other ways to use Demons temporal registration information has to be studied, such as deformation fields analysis (see Deformation Fields Analysis in Sec. 5.3.2 and in Sec. 5.4.2).

Having in mind the idea of using the information obtained from registrations for helping automatic mass detection systems, MI B-Splines based methods obtain less artefacts so direct registered image difference information could be used as an indicator of the presence of a lesion. Besides, for B-Splines based methods it is easy to keep the transformation computed and to apply it to the original digital image (without resizing) for calculating the image difference in the original size. According to our digital study (including both temporal and bilateral digital registrations), the bests MI B-Splines based methods are AB and mAmB, being not really remarkable the difference between them. So in order to improve the computational time, MR approach can be dismissed and AB can be considered as the best method for

calculating the direct registered image difference.

Chapter 5

Multi-Image CAD System

CAD systems are tools developed to help radiologists during the medical image reading and cancer diagnosing processes. These systems commonly analyse each image independently of the rest of the available data of the patient in order to find abnormalities. In contrast, we propose a new framework for automatic mass detection in FFDM images that takes into account bilateral, temporal and also ipsilateral information of the patient.

5.1 Introduction

Breast cancer is considered one of the most frequent tumour in women. However, an early detection considerably increases the life expectancy. So, a key point in the fight against this type of disease is to achieve its detection in early stages. CAD mammographic systems are automatic or semi-automatic software tools developed to assist radiologists in the detection and/or classification of mammographic abnormalities [94]. Although the idea of developing computer systems to assist radiologists in the detection and classification of breast cancer is not recent, the development of FFDM imaging systems has been a catalyst in the increase of such computer systems [172].

CAD mammographic systems are starting to be integrated into clinical workflow in hospitals and these systems are becoming part of the cancer diagnosis process. The success is, in part, due to the past and ongoing efforts on research and devel-

opment of CAD systems [74]. However, the idea of adding additional information from different views based on the workflow of the radiologists, that compare different views of the same patient when reading mammograms, is not used in these commercial CADs. Even though, in the literature, different works that use bilateral, temporal or ipsilateral comparison of mammograms for improving CAD systems can be found and evidence the benefits of incorporating all the available information from the patient.

Note that independently of the comparison information added to CAD systems, since the breast area is really the region to study and to provide a reference for the pairing process, greater number of works apply a preprocessing step for breast area segmentation [290, 155, 309, 251, 139, 239, 338] (see Section 2.2.1 for more detail).

After breast border detection, the background and the burned information on the image (usually labels) are removed and breast area is isolated. Some authors, also take into account the presence of the pectoral muscle in the segmented breast area, since the influence of pectoral muscle presence in the correspondence process is a well known issue [197]. The variations on the placement and tension of pectoral muscle in mammograms and also the intensity value differences between breast region pixels and pectoral muscle pixels may cause problems not only for matching mammograms but also for detecting abnormalities. In order to increase the reliability of CAD systems, pectoral area is removed [211, 265, 198, 239] or is processed to avoid the intensity level differences [155, 308, 307, 325, 309, 321, 139].

5.1.1 Bilateral Information

Bilateral breast comparison is based on exploiting architectural symmetry between left and right breasts [168]. Symmetry statement gives rise to two complementary hypothesis: differences between corresponding areas of both breasts indicate the presence of lesions and similarities between corresponding areas of both breasts indicate normality. Some CAD systems use these hypothesis for detecting abnormalities [28, 33, 15, 205, 16, 39] or for reducing the number of FP [22, 255, 349, 139, 57], although the detection works [211, 290, 155, 265] are far more common than the FP works. In the detection case, the system looks for breast areas misalignments or non-correspondence between ROIs and identifies these findings as abnormalities. In

the FP reduction case, the system analyses the probability of being an abnormality by taking into account both, the features of the candidate area and the features of its corresponding area in the bilateral breast.

Even though left and right breast images are usually taken under similar conditions of breast compression thickness, exposure parameters or breast position, bilateral mammograms are not truly a mirror of the other. Then, obtaining not only the differences but also the similarities between bilateral images it is not a trivial subtraction process, a general description about the whole process is described below.

Correspondence between bilateral images

In works related with bilateral breast comparison, two main different techniques can be found to establish a correspondence between images. One is based on finding an optimal geometric transformation between bilateral breast images to fit them (bilateral registration) and the other one is based on defining a coordinate system to determine the relative coordinates of a ROI in its bilateral image.

Image registration algorithms estimate a transformation between two images in order to align them. One image, the moving image, is deformed to match the other one, the fixed image. Some authors consider that transformation between bilateral images can be regarded as global by assuming that bilateral breasts are architectural symmetric and bilateral mammographic images are taken under similar conditions [177, 100, 351, 289, 207, 33, 205, 22]. As opposed to this idea, some authors consider that global registration is not enough for matching bilateral breasts, due to the fact that the breast is a deformable object. Although there are works that apply only non-linear registration [155, 196, 257, 59], it is common to firstly apply a global method for recovering the main translation, rotation, scaling and shearing differences and finally, apply a non-linear method for local refinement by recovering soft tissue deformations [167, 265, 345, 346, 199, 82].

Besides registration, to define a procedure that establishes the same coordinate system in left and right breast images can also be used for finding correspondences between points or ROIs from bilateral breasts. Once the coordinate

systems are defined and the coordinates of the ROIs are given, the relative coordinates of the above mentioned ROIs in the bilateral breast are calculated as proportion to the given ones. In the literature, either Cartesian [139] or Polar [349, 156] coordinate systems are used.

Adding bilateral information

After establishing relation between bilateral breasts areas, the information about comparing corresponding areas can be used by the systems for detecting potential lesions [211, 290, 28, 265, 15, 16, 39], either masses [351, 360, 155, 167, 33, 205, 82] or other asymmetries [177, 156, 316, 341]. A common practice for detecting potential lesions is to use grey-level differences information obtained from the subtraction of left and right images. For subtracting them, different methods are applied as simple linear subtraction, multiple linear subtraction [290] or non-linear subtraction [350]. However, bilateral linear subtraction is the most common approach to identify regions that differ from one image to another [355, 155, 207, 33, 156, 59]. There are also other methods not based on pixel by pixel comparison that can be used for detecting masses. This is the case of methods that compare features of left and right breasts or features of specific regions of bilateral breasts to measure the dissimilarity between these corresponding regions [177, 211, 167, 338].

Bilateral comparison information can be used not only for detection but also (although less common) for FP reduction [88, 113, 255, 57]. The main idea of FP reduction methods is to take into account features of the suspicious regions and features of the corresponding bilateral regions for classifying the possible masses in real or non-real [22, 301, 302, 349, 139].

5.1.2 Temporal Information

Temporal comparison of mammograms is used for assessing breast changes through time [342]. When a lesion is detected in the current mammogram but not in the prior one, this may be considered as a cancerous sign. In the same way, if in current and prior mammograms a lesion is detected and its size and/or shape have changed, this may be classified as sign of potential cancer. Otherwise, no changes may be

considered as non-cancer lesion. These assumptions can be used by CAD systems to either detect new lesions [250, 264, 360, 167, 339, 340] or to classify them as benign or malignant [114, 309, 188, 189], although the detection purpose seems to be more common than the classification one.

Although temporal comparison works with mammogramms of the same breast, acquisition parameters variation and breast tissue changes over the time cause important differences at radiographic appearance level. So, unfortunately, like in the bilateral case, comparing temporal mammograms is not a trivial subtraction process. An overview of temporal processes is presented below:

Correspondence between temporal images

The comparison of temporal mammograms usually is more complex than the bilateral one, because exposure parameters, patient position and breast compression in current and prior acquisitions normally are considerably different. And also, in spite of being the same breast and view, breast is on a continuous changing process [197]. Due to the substantial radiographic appearance differences and being a deformable body, majority of authors use non-linear geometric transformations for registering previous and current mammograms, instead of applying only global approaches. Thin-Plate Spline based techniques are abundant in the literature [345, 196, 346, 324, 198, 334, 199], although other approaches are also used as B-Splines based ones [118].

Adding temporal information

Although linear subtraction of prior and current post-registered mammogramms [264, 196, 346, 257] is often used for obtaining the breast variations over the time that could indicate the presence of a new/changeable lesion, features comparison of temporal registered images or specific ROIs [114, 324, 91, 309, 340] seems to be most commonly used either for detection of lesions or for benignity/malignancy classification of abnormalities. Note that classification in benign and malign could be considered as a FP reduction step, because lesions classified as benign can be suppressed [188, 189].

5.1.3 Ipsilateral Information

The majority of lesions are visible in both mammographic views (CC and MLO). Having in mind this idea, when a suspicious lesion is only present in one of the two views, it may be considered as a false lesion. Otherwise, when the suspicious lesion is detected in MLO and its presence is confirmed in the CC view (and viceversa), it is considered as a real lesion and a biopsy of the abnormality is recommended [21]. CAD systems follow this correspondence rule and when a candidate lesion is not detected in both views is definitely considered as a non-lesion, i.e, as a FP.

Correspondence between ipsilateral images

To locate the corresponding position of a candidate mass in its ipsilateral view, authors commonly use the distance from the nipple to the mass location [297, 11, 339, 252, 321, 266, 357]. Although more sophisticated techniques can be used, for instance, curved epipolar lines [162, 163]. CC and the MLO images are related to the geometry of stereo vision. In stereo vision, given a point in one image, a straight line which indicates where the point lies (epipolar line) can be defined in the other image. However this cannot be applied directly for CC and MLO images, since the breast is a non-rigid object and the deformation variations of the breast during the compression affect on the stereo problem. The solution is to calculate curved epipolar lines by developing a simulation of object deformation into the stereo geometry.

Adding ipsilateral information

To evaluate the similarity between matched regions, different features can be extracted [103, 109, 325, 251, 266]. The most commonly used are texture [240, 297, 11, 110, 321, 353, 239, 341], spatial [252, 90, 139, 89, 328] or shape [356, 108, 252, 90, 89] features. Usually, these features are the input of artificial neural networks based classifiers [298, 297, 325, 356, 321, 353, 139]. These classifiers are used either for distinguishing between benign and malignant lesions [109, 110, 108, 239] or for FP reduction [298, 240, 321, 266, 139, 244, 328].

5.2 Single-Image CAD System Proposal

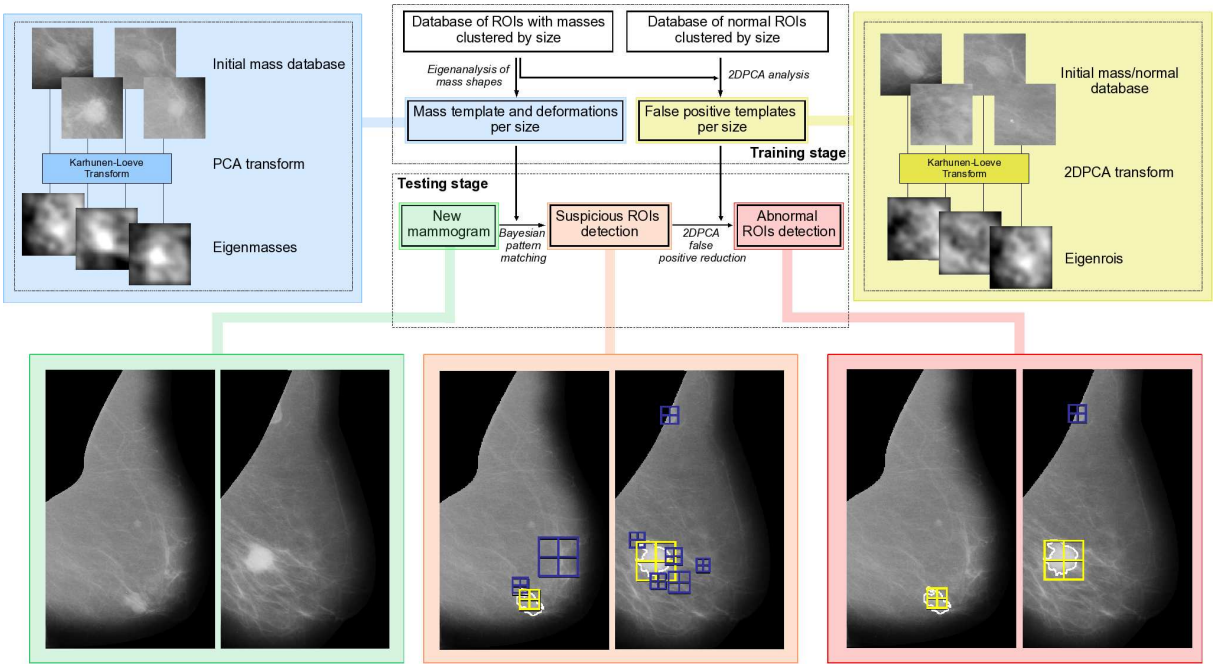


Figure 5.1: Layout of the single-image CAD system (top center flow diagram), depicting graphically the main steps with details in the specific boxes. The bottom row shows the results of the CAD system: the original images, the template matching and the FP reduction step. Note, however, that in the second mammogram a FP remains.

Fig. 5.1 depicts the single-image CAD previously developed in our research group [95]. The CAD follows a model-based approach where an initial training stage to learn the morphology of the masses on the current database is needed. After this step, the system is able to detect masses in new mammograms.

The training stage is divided in two steps (Fig. 5.1). First, using a database of Regions of Interest (ROI), which each contain a mass in the center, the system automatically learns their size and shape. The approach is based on using the Karhunen-Loeve transform to take the variation of the mass shapes into account. The output of this initial step is a template per each size. Templates are defined by their mean shape and possible deformations. The second step of the training stage is directed to FP reduction, which in this case means to distinguish between normal tissue and real masses. For this purpose the training images should now include

instances of real masses and also instances of ROIs being normal tissue. Hence, a second database is constructed by adding normal ROIs to the mass database. Regarding the number of mass ROIs, three times as many normal ROIs were included for each size-cluster. This ratio results in a good compromise between the performance of the FP reduction algorithm and the huge variability of FP ROIs. Therefore, the output of this second learning step is a classifier (2DPCA) trained to predict if a ROI contains a mass or normal tissue.

Once the system has been trained, it is ready to detect masses in unseen mammograms. This testing stage is, again, divided in two steps (Fig. 5.1). The first stage is focused on the detection of all suspicious regions, while the second stage (FP reduction) aims to classify the detected regions as normal tissue or masses. The detection step consists in matching the templates to regions in the mammograms. According to the used probabilistic framework, the objective function to minimize is (see the original work for details [95]):

$$\lambda = \sum_{k=1}^{N} (\xi_k - 1)^2 + \Omega(\psi^{s,\xi,d}, Y) \tag{5.1}$$

where ξ_i are the set of parameters that models the template deformations and $\Omega(\psi^{s,\xi,d}, Y)$ is the potential function that measures the agreement between the template deformed with parameters ξ_i and the image itself (Y). Therefore, this function λ consists of two terms: a first one that measures the deviation of the deformed template from the prototype (hence penalizing larger deviations from the template), and a second one which describes the fitness of the deformed template to the boundaries of the image (the closer the template to the image, the smaller the potential value). The output of this step is, hence, a set of detected suspicious regions. In order to ensure that these regions are real masses, the FP reduction step is applied. Each suspicious region is cropped from the image and used as the input to the 2DPCA classifier. Therefore, only the regions classified as being real masses are marked in the mammogram.

5.3 Multi-Image CAD System Proposal

We have developed a model-based multi-image CAD system for mass detection for FFDM, extending the single-image CAD system to work with digital images and to use case-level information. The multi-image CAD system includes information from the automatic registration of bilateral and temporal mammographic images and from the automatic correspondence between ipsilateral mammographic images. This information can be integrated either as *a priori* or *a posteriori* information:

- In the *a priori* case, the registration is performed before the detection starts. Mammogram registration comparison is inserted in the probabilistic detection framework by modifying the potential function (see Fig. 5.1).
- In contrast, in the *a posteriori* case, the detection step is executed independently of the case-level information. At the end of the algorithm, however, the non-correspondence between detected regions is used as a way to reduce the number of FPs.

According to the trends in multi-image works (see Sections 5.1.1, 5.1.2 and 5.1.3), in our multi-image CAD scheme, the comparison between bilateral and temporal mammograms is integrated into the probabilistic CAD detection framework as a priori information. And the correspondence between ipsilateral mammograms is integrated into the CAD FP reduction framework as a posteriori information.

Note that all the used FFDM images are preprocessed. During the preprocessing step, pectoral muscle is manually segmented and eliminated, and labels are also removed. Once the breast region is extracted, a peripheral enhancement is applied (see Chapter 2 for more detail).

5.3.1 Bilateral Information

Image registration is used in order to detect differences between bilateral images that could be due to the development of lesions. Specifically, before the detection starts, a registration algorithm combining two intensity-based methods is applied. We use an Affine transformation to recover the main pose and scaling differences and a B-Splines registration method for localized non-linear deformations according the

results obtained in Chapter 4. Once the images are aligned, the absolute difference between the registered and the fixed images is calculated (see in Fig. 5.2 an illustration of the bilateral image difference process). This subtracted image is normalized and multiplied (pixel by pixel) by the original image (only subtracted values bigger than 0 are used). In some sense we are dealing with this bilateral information as an enhancement procedure. Hence, as the result of the multiplication, regions with different structures are highlighted. To avoid adding noise to our CAD system, only the most significant differences (10%) are used to enhance the image that contains the lesion.

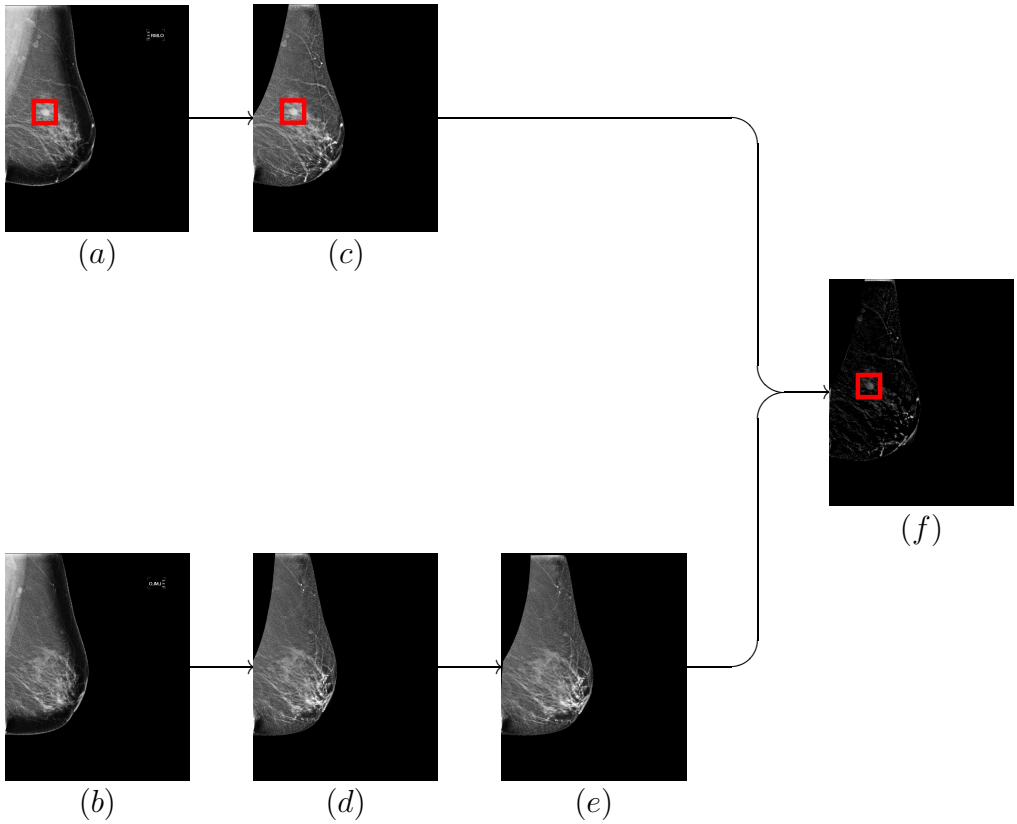


Figure 5.2: Example of the bilateral image difference calculation. (a) Original right MLO mammogram with lesion squared in red, (b) original left MLO mammogram, (c) preprocessed right MLO mammogram with lesion squared in red, (d) preprocessed left MLO mammogram, (e) left MLO mammogram after registration, and (f) difference image with possible lesion squared in red ($|(c) - (e)|$).

5.3.2 Temporal Information

Image Subtraction

In the same way as in bilateral comparison, the absolute image difference is considered as an option to add temporal information. Temporal subtraction process is equal to the bilateral one. After B-Splines registration with Affine initialisation, images are subtracted (see in Fig. 5.3 an illustration of the temporal image difference process). Then, subtracted image is multiplied by the original one to enhance the differences.

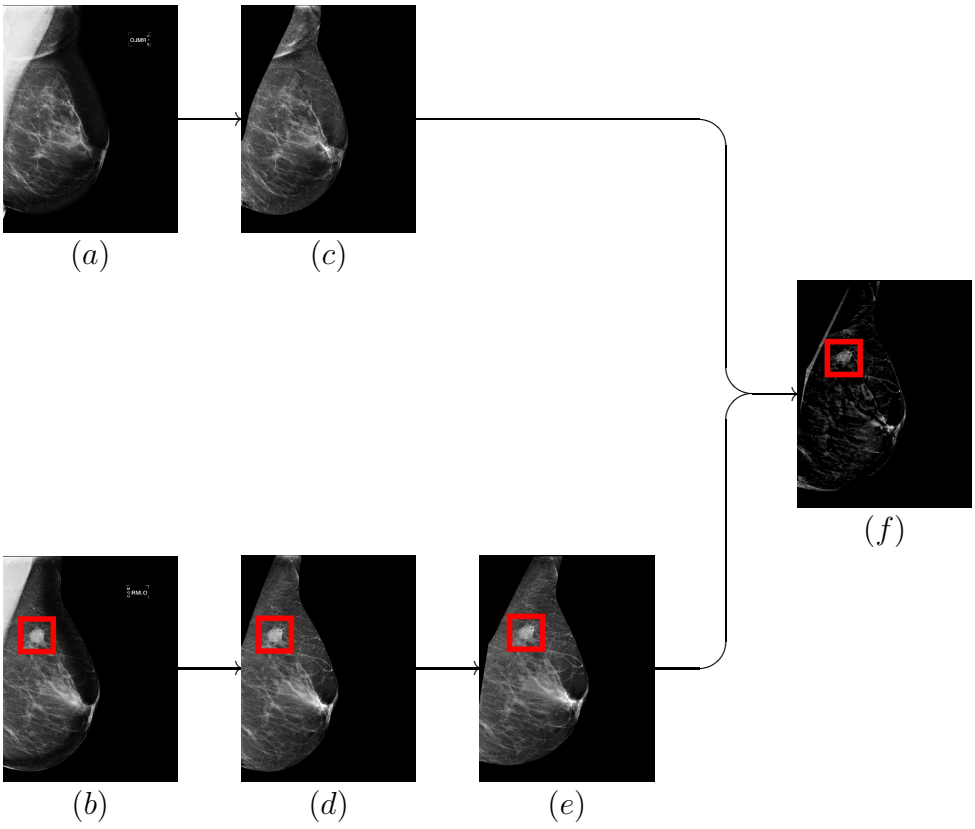


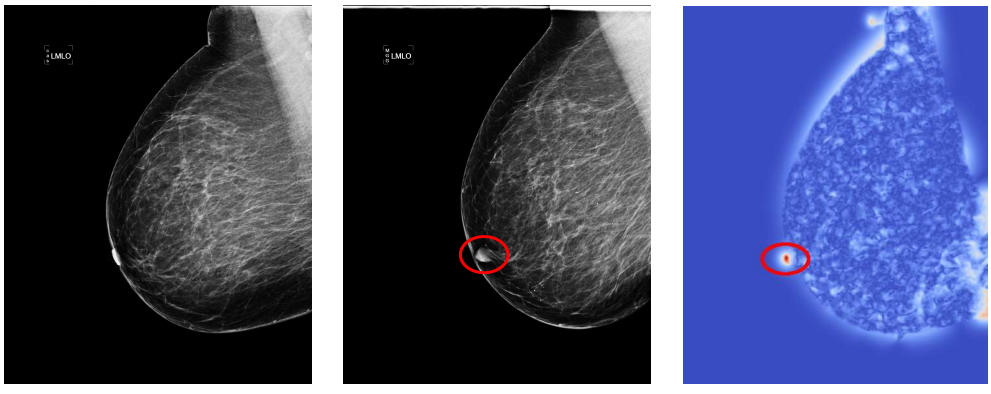
Figure 5.3: Example of the temporal image difference calculation. (a) Previous original right MLO mammogram, (b) current original right MLO mammogram with lesion squared in red, (c) preprocessed previous right MLO mammogram, (d) preprocessed current right MLO mammogram with lesion squared in red, (e) current right MLO mammogram after registration with lesion squared in red, and (f) difference image with possible lesion squared in red ($|(c) - (e)|$).

Deformation Fields

Following the opened research line in Chapter 4, we studied alternatives to image subtraction in order to incorporate mass detection information from Demons methods to our multi-image CAD system and we focused on the study of Demons deformation fields.

In Demons based registration approaches, every pixel is transformed in a way that does not necessarily be related to its neighbouring pixels. Each of these movements can be expressed by a displacement vector (the displacement suffered by each pixel normalised by the image size). The union of all these vectors stands for a displacement (or deformation) field that characterises the non-linear transformation. Diffeomorphic Demons with MR (mD), Affine in combination with Diffeomorphic Demons (AD) and Affine MR in combination with mD (mAmD) were evaluated.

Fig. 5.4 shows a visual example of a deformation field. For ease of visualization, only vector norms are presented in the image in the form of a color code. Pixels closer to red correspond to higher norm values and pixels closer to white to lower values. The figure shows the potential of deformation fields for applications such as mass detection. Pixels corresponding to the mass appearing in the second temporal study present the highest norm values in the whole of the image.



(a) (b) (c)

Figure 5.4: Deformation field example: (a) source image , (b) target image with visible mass and (c) corresponding AD deformation field.

Feature Based Classification

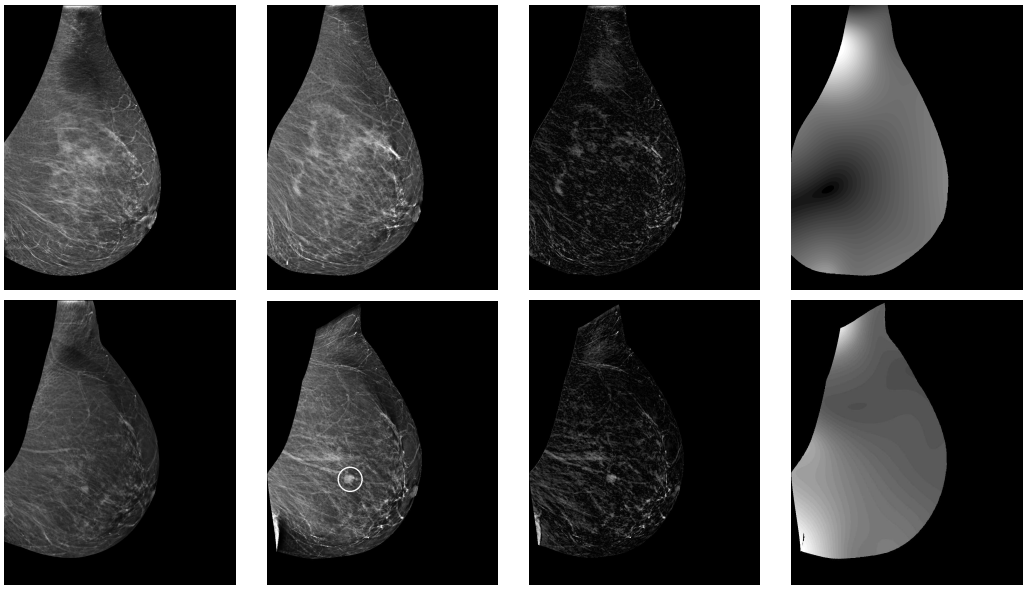
Besides, B-Splines and Demons, TPS is an interesting registration transformation as described in Chapter 4 (Sec. 4.1) and in Section 5.1. So, we also studied the utility of the TPS based registration results and also how to add them to our multi-CAD system.

After an initial affine registration, the proposed algorithm extracts interest points found in the boundary and applies a robust point matching (RPM) approach obtaining a non-linear transformation based on a TPS [199]. RPM is compared to using Affine registration and no registration at all (see results in Section 5.4.2).

From the registration results three sets of features are extracted which are then used to classify a patient into normal or abnormal. The first feature set is computed from the difference image while the second set is extracted from the deformation field. In these two sets (difference image and deformation field) the features computed are the first five statistical moments of the intensity or deformation distribution. Finally, the third set of features is composed of various similarity measures commonly used in image registration computed between the fixed and moving images: root mean squared error, cross-correlation, entropy of the difference image and mutual information [359], having a total of 14 features.

Figure 5.5 shows an example of image registration of a normal and abnormal case, with the transformed moving image, the difference image and the deformation field magnitude. While differences in the deformation field are difficult to appreciate, structural dissimilarities in the difference image are highlighted, including the lesion in the abnormal case.

The above described features are computed for each single temporal registration. As we are registering left and right temporal mammograms of the same patient independently, we also study the effect of combining the features hence obtaining a unique feature vector for each woman. The hypothesis is that this combination can help towards the classification as in normal cases those features are likely to be more stable compared to abnormal cases due to the development of breast cancer. Various simple combinations have been



(a) (b) (c) (d)

Figure 5.5: Image Registration: (a) Fixed and (b) transformed moving mammograms, (c) image difference and (d) deformation field magnitude. Top row shows a normal mammogram and bottom a mammogram with a lesion (white circle).

tested: mean, signed and absolute differences, and minimum and maximum.

Features have been used in a RF classifier in order to differentiate between normal and abnormal cases containing a mass. The parameters are empirically set to 500 decision trees and a feature subset size of 3 features for each tree. Although other classifiers (such as SVM, Adaboost and KNN) and feature selection methods have been tested, RF obtained the best results overall. PRTools software has been used for the implementation [76]. All features have been normalised to a zero mean and unit standard deviation. A leave-one-woman-out validation approach has been used for testing.

5.3.3 Ipsilateral Information

With the aim of reducing the number of FPs, we propose to take advantage of the correspondence between CC and MLO views. The suspicious lesions with no ipsilateral correspondence are considered less probable to be real lesions than the

ones with positive correspondence.

Detection is applied to CC and MLO images independently. After detection, a CC/MLO correspondence approach based on using curved epipolar lines [163] is applied over each suspicious region in CC. Given a point in the 2D CC image, it is reconstructed as a line in the 3D CC-compressed breast. This straight line is transformed in a curved line when the breast is uncompressed according the compression model proposed by Kita et al. [163]. This curved line is rotated to pass to MLO space and it is transformed according the aforementioned compression model and MLO-projected (see Fig. 5.6). If a MLO suspicious lesion is situated along the curved epipolar line generated by the CC suspicious lesion, this is considered as a positive correspondence between ipsilateral images and the CC abnormality is regarded as a true lesion. Otherwise is considered as a FP.

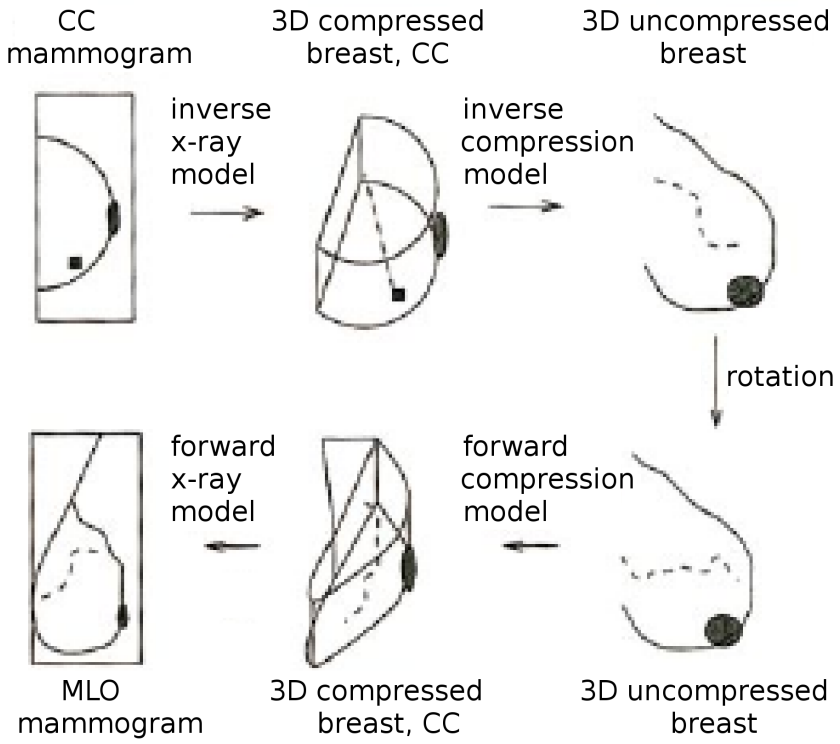


Figure 5.6: Overview of the CC/MLO correspondence process. Extracted from Kita et al. [163].

5.4 Results

5.4.1 Bilateral Information

Digitised

We used two different databases: the Mammographic Image Analysis Society (MIAS) database [296] and the Digital Database for Screening Mammography (DDSM) [129]. For the MIAS database we have obtained detailed annotations of the masses, while for the DDSM database this is not available. This determines that we can only use the MIAS database for detailed evaluation, but DDSM for training (see [95] for further details). The experimental results presented in this section were performed using a subset of 208 mammograms extracted from MIAS, including 104 pairs of left and right MLO mammograms. In total there were 52 mammograms containing at least one mass.

The registration methods that obtained the best results were B-Splines MR for SSD metric and B-Splines MR with Affine MR initialisation for MI metric. Once the best methods for each metric were chosen, another criteria, visual assessment, was performed to determine the metric to be used. Difference images after registration were reviewed by different observers to evaluate dissimilarities from a global point of view where MI obtained better results than SSD (see Chapter 4 for details, specifically Sec. 4.3). Therefore, B-Splines MR with Affine MR initialisation method with MI metric was the one used to test the proposed dual-image CAD system.

To perform the quantitative evaluation we used Receiver Operating Characteristic (ROC) and Free-Response ROC (FROC) analysis. Although before presenting results, for the sake of clarity, some concepts will be revised.

True Positive (TP): System detects an abnormality in an abnormal case.

True Negative (TN): System no detects abnormalities in a normal case.

False Positive (FP): System detects an abnormality in a normal case.

False Negative (FN): System no detects abnormalities in an abnormal case.

Lesion Localisation (LL): System marks a region as abnormal and is close to a lesion (within an agreed upon accuracy).

Non-Lesion Localisation (NL): System marks a region as abnormal but is no close to any lesion.

Derived form these definitions, the True Positive Fraction (TPF) and the False Positive Fraction (FPF) are defined as:

$$TPF = \frac{TP}{TP + FN} \qquad FPF = \frac{FP}{FP + TN}$$

And the Lesion Localisation Fraction (LLF) and the Non-Lesion Localisation Fraction (NLF) are defined as:

$$LLF = \frac{LL}{Total\,Number\,of\,Lesions} \qquad NLF = \frac{NL}{Total\,Number\,of\,Images}$$

In ROC analysis, a graphical curve represents the TP rate as a function of the FPs rate, i.e., the ROC curve is the plot of TPF along the y-axis vs FPF as the confidence level is varied. The Area Under the ROC curve (A_z) is the probability that an abnormal case is rated higher than a normal case, therefore, A_z is an indication for the overall performance of the observer, and is typically used to analyse the performance of the algorithms [42]. On the other hand, in FROC analysis the LLF is plotted against the NLF [352]. Note that in this analysis the definition of a detected region is needed. In this work we assume that a region is detected if the overlap between that region and the suspicious region is at least 50% [152].

Without taking the bilateral information into account and using the above mentioned dataset, our CAD system obtained $A_z = 0.716$. In contrast, when adding this information this value increased to $A_z = 0.852$. On the other hand, Fig. 5.7 shows the results of the FROC analysis without and with using the bilateral information. Note that at lower sensitivities this information was not useful. However, at higher sensitivities the dual-image CAD improved the single-image one. For instance, at a LLF of 80% the single-image CAD obtained 1.68 FPs per image, while the dual-image CAD reduced to 0.90, and at a LLF of 88% the FPs per image were 1.85 and 0.99, respectively. The obtained results show the benefits of including the bilateral information, increasing the sensitivity of the CAD at less FPs per image.

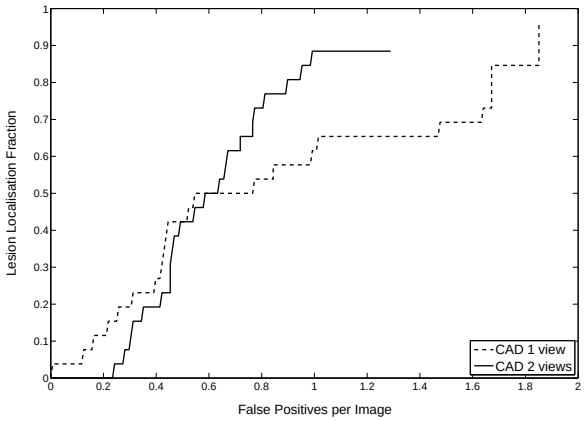


Figure 5.7: FROC analysis for the CAD system without (CAD single-image) and with (CAD multi-image) considering bilateral information.

Digital

To evaluate the performance of the multi-image CAD system, a set of 74 cases containing masses from our FFDM database is used. Each case contains left and right mammograms of a patient. The evaluation of our experiments is done by comparing the detection results of our CAD system when using or not bilateral information. Initial results indicate benefits of including bilateral information in contrast with our previous developed single-image CAD system (Fig 5.1). Detection sensitivity is improved by a factor of 10% (from 0.76 to 0.83). Using bilateral information, the main detection errors are found in subtle masses in extremely dense breasts where difference image is not able to highlight the lesion.

5.4.2 Temporal Information

Image Subtraction

To evaluate the performance of the multi-image CAD system, a set of 47 cases containing masses from our FFDM database is used. Each case contains current and previous mammograms of a patient. The evaluation of our experiments is done by comparing the detection results of our CAD system when using or not temporal information. Mass detection does not significantly im-

prove (sensitivity around 85% with both systems). Simple image subtraction is able to emphasize the lesion when the mass is not present in previous studies, similar to bilateral cases. However, when current and previous images contain a mass, lesion changes are subtle and can be obfuscated by the changes in the overall structure breast (i.e. dense tissue decrease) or acquisition parameters.

Deformation Fields

For the automatic study of deformation fields, several parameters can be considered [144]. We focus on two of them: average and maximum norms of the vectors in the deformation field. Our aim is to provide statistical support to the claim that Diffeomorphic Demons, can be used for mass detection in digital mammography images.

Concerning the average norm of the deformation field of the methods studied, we computed the mean values of this average norm for every method before and after registration. A noticeable increase was observed for Diffeomorphic Demons methods when FFDM contain masses. This behaviour was observed even more clearly for the maximum norm indicator (see Table 5.1). Note that only Demons based methods are shown in this table because the other methods did not provide significant information.

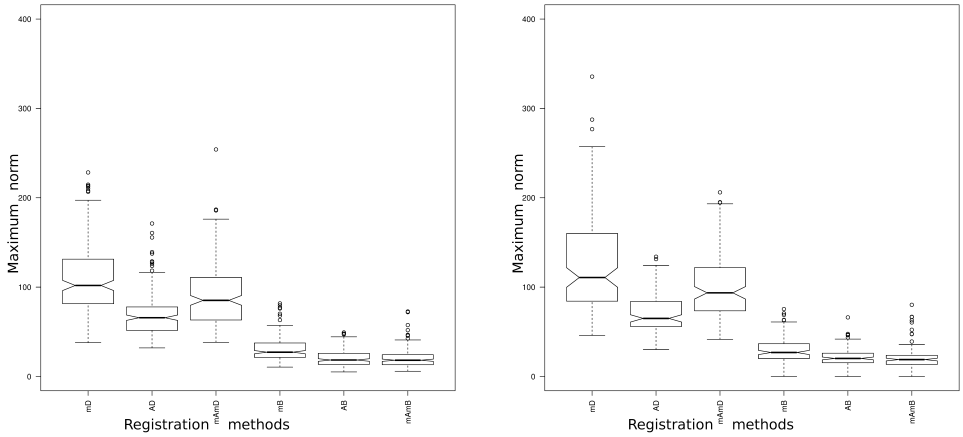
	Average Norm				**Max Norm**			
	Without Masses		**With Masses**		**Without Masses**		**With Masses**	
	μ	σ	μ	σ	μ	σ	μ	σ
mD	13.99	4.04	15.45	7.45	105.65	34.54	127.21	67.36
AD	4.04	1.43	4.19	1.32	69.24	24.44	70.30	21.35
mAmD	7.79	3.23	8.07	2.96	91.13	32.91	98.17	35.42

Table 5.1: Statistical summary of deformation fields for Demons methods. Registrations with and without masses are presented separately.

In order to provide inferential backing to these observations of descriptive statistics, a set of pairwise t-tests was performed to check if the maximum norm of a given method was significantly different for registrations with and without masses. The alternative hypothesis (H_1) was stated as *"the maximum norm is higher for registrations with masses than for those without masses"*. Results showed statistically significant difference for mD ($p - value = 0.00127$) and

for mAmD ($p - value = 0.05547$). This shows the potential of Diffeomorphic Demons methods for mass detection in digital mammography.

Results of the maximum norm of the vectors in the deformation field are also visualised in Fig. 5.8. Maximum norms for all Demons based methods studied are presented, data is provided separately for registrations with and without masses. We also see how, as expected, using Affine initialisation decreases the maximum norm. This happens as the initial Affine step helps to cover part of the transformation of each pixel in order to reach its corresponding pixel. Furthermore, the increase in maximum norm for registrations with masses is also visible.

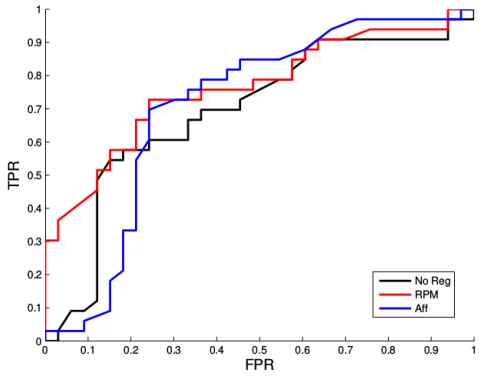


(a) (b)

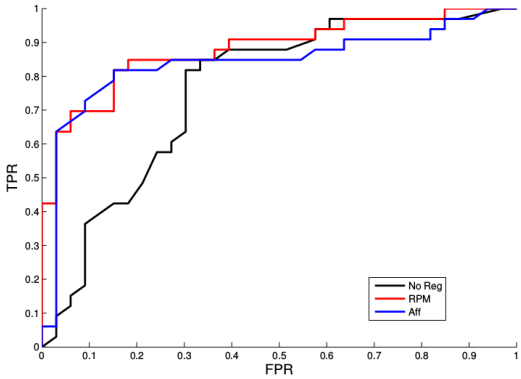
Figure 5.8: Deformation field maximum norm. Temporal registrations (a) without masses and (b) with masses.

Feature Based Classification

Figure 5.9 and Table 5.2 show classification results in terms of area under the curve (AUC) when using the proposed algorithm (RPM) compared to no registration (No Reg), and Affine transformation using MI (Aff). Features are computed for two cases: for a single registration (*Single*) or combining left and right temporal features using the maximum (experimental results have shown that combining using the maximum value obtained the best results) of both features (*Combined*). For the single case, only one mammogram is used for feature extraction: the one with the mass for abnormal cases and left or right randomly selected for normal ones.



(a)



(b)

Figure 5.9: Abnormal classification ROC curves using features from RPM and Aff also compared to no registration. Single features (a) are compared to their combination using the maximum operation (b).

	No Reg	RPM	Aff
Single	0.69	0.76	0.70
Combined	0.76	0.88	0.84

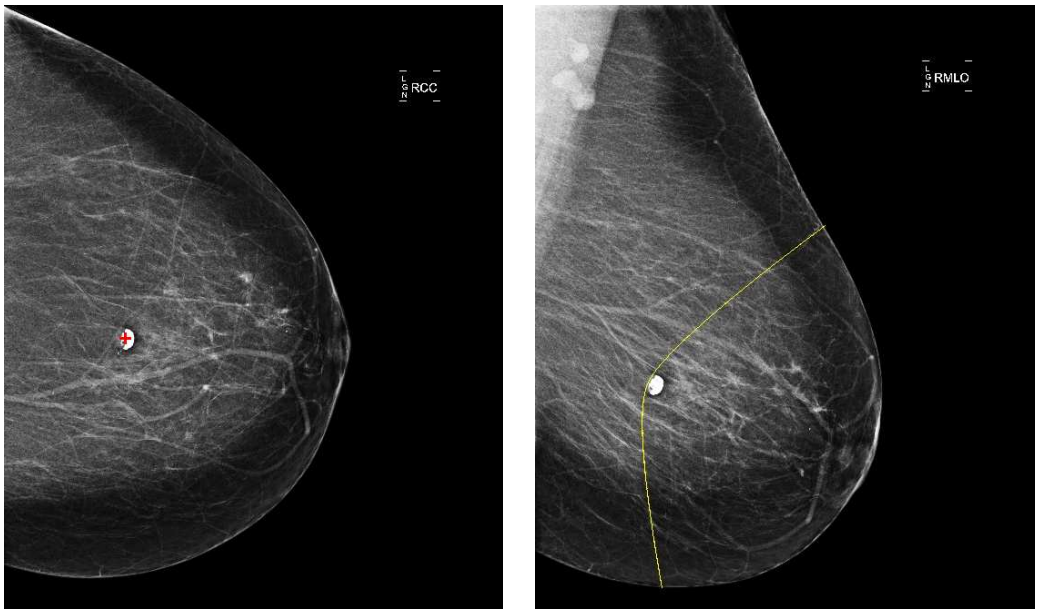
Table 5.2: AUC for classification of abnormal cases. Features used in the classifier are obtained after no registration (No Reg), Aff or RPM. Single features are compared to their combination using the maximum operation.

Regarding the ROC curves with single features, the use of image registration shows a clear improvement compared to no registration, specially for the RPM case. This is also reflected in the AUC values (0.69 compared to 0.70 of the Aff or 0.76 for the RPM). For each case, single ROCs are obtained hence statistical significance could no be computed. Regarding feature combination, it is also clear that results improve in all cases, including the no registration case. Differences are relevant with respect to the use of registration algorithms compared to no registration, although between Aff and RPM (0.84 vs 0.88) this difference is not that evident. This would need further investigation which will include the comparison with other non-rigid algorithms. Regarding feature analysis it has been observed that features based on the intensity similarity (moments of the difference image and mutual information) show better discriminant properties than the rest of the features. However, with the inclusion of other registration algorithms this could change in favour of other features such as the deformation field.

5.4.3 Ipsilateral Information

To evaluate the performance of the multi-image CAD system, a set of 74 cases containing masses from our FFDM database is used. Each case contains CC and MLO mammograms of a patient. Once detection is applied to CC and MLO images, curved epipolar lines are generated (one by each CC suspicious lesion) and an accurate ipsilateral correspondence between ROIs is obtained (see an example in Figure 5.10). A CC abnormality is considered as a true lesion if any MLO suspicious lesion is on its curved epipolar line, if along the curved epipolar line there is only one lesion, MLO abnormality is also considered as a true lesion. When no correspondence is found, the suspicious lesion is considered as a FP and is rejected.

Results indicate that FPs are reduced by a factor of 15%.



(a) (b)

Figure 5.10: Curved epipolar line. (a) Right CC mammogram, and (b) homonym right MLO mammogram. Point marked with a red cross in (a) is situated across the curved epipolar line (yellow line) in (b).

5.5 Discussion and Conclusions

Initial digitised work presents a comparison for a mass detection CAD system when using or not registration information of bilateral mammograms. According to the obtained results, the introduction of the registration information as *a priori* information is considered as an improvement of our previously developed single-image CAD system.

Later work integrates bilateral, temporal and ipsilateral information into our single-image CAD system. According to the obtained results, the introduction of the bilateral information as a priori information via image subtraction can be seen as an improvement of our single-image CAD system, however the introduction of the temporal information by the same way does not seem to be an improvement.

Regarding ipsilateral correspondence, results show a promising option for FP reduction.

In order to add useful temporal information to our CAD system, we investigate the use of other image similarity measures like the average norm or the maximum norm of the deformation fields. Access to annotations from an expert radiologist allowed us to discriminate those registrations containing masses from those not containing them. The analysis of differences between both scenarios using deformation fields showed statistically significant differences in behaviour for Demons methods.

A framework to classify mammograms into normal and abnormal cases is also presented. Framework is based on classifying image based features from temporal image registration results. Feature combination between left and right breast has been shown to obtain better results in terms of ROC analysis compared to using single features alone. This indicates that combining features with other views such as CC and MLO could further improve the results.

Chapter 6

Conclusions

In this last chapter we present the summary of the thesis as well as the obtained conclusions. Furthermore, we describe the future directions of our work.

6.1 Summary of the Thesis

CAD mammographic systems were and are being developed to help radiologists in the detection of lesions in mammograms. During the evaluation of the images, radiologists use information about all available images of each patient, that is, left and right images of the same view, CC and MLO views of the same breast or previous examinations of the patient. However, the majority of CAD systems use only each image independently to find abnormalities. The research presented in this thesis includes the development of a multi-image CAD system for detecting masses to outperform the single-image one.

Mass detection is a complex task because this type of lesions are extremely difficult to see and some of the times only appear very subtle signs of them in FFDM images. Fortunately, it is proved that CAD mammographic systems are useful tools that help radiologists to detect masses. In addition, mammogram preprocessing and the use of breast density information have been demonstrated that improve CAD detection rates, as well as adding information from other mammograms of the same patient. The first part of this thesis is addressed to enhance FFDM images. Subsequently, the second part is focused on breast density classification. The third

part analyses different mammogram registration approaches which are going to be applied to several multi-image CAD schemes.

During the FFDM acquisition the breast is compressed. Despite the compression is non-uniform is still needed to compensate the thickness reduction in outer breast edges. A novel peripheral enhancement algorithm was developed that keeps the non-overexposed area unchanged and balances the overexposed area using information from the non-overexposed region. Corrected images are considered visually better for diagnostic than non-corrected ones. Furthermore, peripheral breast enhancement improves expert detection of masses and automatic breast density classification.

Breast density is related with an increased risk of breast cancer. Women with high-dense breast have more probability of developing a cancer than women with low-dense breast, furthermore, breast density decreases FFDM accuracy and CAD systems sensitivity. Two algorithms for breast density classification were proposed and developed. The first one classified breasts density according BI-RADS. This qualitative approach was tested with FFDM images from our local database and outperformed other automatic approaches. The second one classified each pixel as dense or fatty. This quantitative method was also tested with digital images. Despite of the lack of a reliable ground truth, longitudinal analysis confirmed that breast density decreases with age and the transversal analysis confirmed the high density correlation between left and right breasts, and also between CC and MLO views.

Registration is an important tool to compare two images. In order to find differences and similarities between two mammograms a correct alignment is needed. We investigated which method obtained best results for bilateral and also for temporal mammogram pairs. Non-linear and linear based approaches were tested. Non-linear methods outperformed linear-methods, in addition, non-linear methods with linear initialisation as well as MR approaches improved registration results, however the more flexible the transformation model the lower the qualitative evaluation results.

CAD systems help radiologists during the reading process. Although current systems do not take into account all the available information of each patient. We developed a multi-image CAD system that uses case-level information to detect masses on FFDM images. Bilateral and temporal information are used during the detection step, while ipsilateral information is integrated on the FP reduction stage.

The proposal is tested with FFDM images from our local database. Multi-image CAD system performs better when case-level information is added, except when temporal comparison is used. As an alternative, deformation fields analysis or image based features analysis of temporal registration pairs seems to provide useful information regarding mass detection.

6.1.1 Contributions

Thus, we consider that the main contributions of this thesis are:

- A new algorithm to compensate the thickness reduction in breast periphery which balances the overexposure in this peripheral region.
- An algorithm for qualitative breast density classification based on FCM segmentation and SVM classification.
- A new algorithm for quantitative breast density classification based on pixel classification.
- An exhaustive analysis about inter- and intra-expert variability in breast density classification.
- An accurate comparison about several linear and non-linear registration methods, as well as some combinations of linear and non-linear methods and MR approaches.
- A proposal for automatic mass detection which takes into account case-level information in terms of bilateral, temporal and ipsilateral images.

6.2 Further Work

Both, multi-image CAD system and breast tissue classification issues need additional work to increase the reliability of the proposals. Furthermore, both issues would be interesting starting points for future research lines. Different directions are possible: to add the whole case-level information (bilateral, temporal and ipsilateral

information) at the same time in our multi-image CAD systems, to add temporal information by using either Demons deformation fields analysis or robust point matching based features analysis, to add breast density information, to reduce the computational time of our CAD system, to adapt our system to other x-ray imaging techniques or to go in depth in our qualitative breast density classification approach.

As we have shown, when bilateral comparison is added to CAD scheme, results are better than the ones obtained when this information is not used, the same as in the ipsilateral case. But, experiments using both data, bilateral image difference and ipsilateral correspondence, are not yet tested. Therefore, a logical direction in the automatic case-level detection framework would be the integration of firstly bilateral information during the detection stage and finally ipsilateral information for FP reduction.

Regarding the temporal comparison, even though adding this information does not penalize our multi-image CAD system but it does not improve the results either. Therefore, we should not to add the temporal image difference during the detection stage. In fact, according the results, the image difference is useful for enhancing obvious differences between images (when a mass is present in only one image), however changes in size or shape are too subtle to be detected by image subtraction. Nevertheless, helpful temporal information could be obtained in another way, for instance from Demons deformation fields analysis or from robust point matching based features analysis. Both analysis are useful to distinguish between normal and abnormal mammographic images and this conclusion would be important to reduce the number of FP. Therefore, in the final stage of our multi-image CAD system as well as to use ipsilateral correspondence, the system would use the temporal abnormal/normal scheme for FP reduction.

As well as adding case-level information to our multi-image CAD system, adding qualitative density information would be an improvement during detection process itself. The fact that density reduces FFDM accuracy makes necessary a different detection process depending on the amount of breast density. So, depending on the breast density classification (i.e, depending on the BI-RADS category), the algorithm parameters could be tuned different and the used training databases of ROIS could not be the same.

Although adding case-level and qualitative density information could improve our

CAD results, the large quantity of information to analyse and generate transforms our CAD in a slow system which is not helpful for radiologists. A key point for the near future would basically be the optimization of the registration steps since BI-RADS density classification is more efficient.

Taking into account that new x-ray mammographic imaging techniques are being developed and integrated in hospitals, the idea of a CAD that detects masses in FFDM images would be an obsolete challenge in near future, so a future work would be to adapt our multi-image CAD system to tomosynthesis and/or C-View images.

Breast density qualitative process seems to be an already robust framework however, the lack of a breast density segmentation ground truth makes incomplete the testing part of our breast density quantification approach. Therefore the first step in this future line, is to obtain a reliable breast density segmentation ground truth and test the algorithm segmentation results. Furthermore, breast density quantification approach should be optimised, because the classification of each pixel in fatty and dense is high time consuming and, the same for CAD system, the algorithm could be also adapted for working with C-Views.

Bibliography

[1] M. Abdel-Nasser, H. A. Rashwan, D. Puig, and A. Moreno. Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern. *Expert Systems with Applications*, 42(24):9499–9511, 2015.

[2] M. Adel, M. Rasigni, S. Bourennane, and V. Juhan. Statistical segmentation of regions of interest on a mammographic image. *EURASIP Journal on Advances in Signal Processing*, 2007(2):3, 2007.

[3] D. Aha and D. Kibler. Instance-based learning algorithms. *Machine Learning*, 6:37–66, 1991.

[4] Z. Aitken, V. A McCormack, R. P. Highnam, L. Martin, A. Gunasekara, O. Melnichouk, G. Mawdsley, C. Peressotti, M. Yaffe, N. F. Boyd, and I. dos Santos Silva. Screen-film mammographic density and breast cancer risk: a comparison of the volumetric standard mammogram form and the interactive threshold measurement methods. *Cancer Epidemiology, Biomarkers and Prevention*, 19(2):418–428, 2010.

[5] P. Aljabar, J. V. Hajnal, R. G. Boyes, and D. Rueckert. Interpolation artefacts in non-rigid registration. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, volume 3750, pages 247–254, 2005.

[6] P. Aljabar, D. Rueckert, and J. V. Hajnal. Interpolation artefacts in non-rigid image registration. In *Proc. of the International Society for Magnetic Resonance in Medicine*, page 2195, 2004.

[7] E. L. Allwein, R. E. Schapire, and Singer Y. Reducing multiclass to binary: A unifying approach for margin classifiers. *Journal of Machine Learning Research*, 1:113–141, 2000.

[8] O. Alonzo-Proulx, R. A. Jong, and M. J. Yaffe. Volumetric breast density characteristics as determined from digital mammograms. *Physics in Medicine and Biology*, 57(22):7443, 2012.

[9] O. Alonzo-Proulx, G. E. Mawdsley, J. T. Patrie, M. J. Yaffe, and J. A. Harvey. Reliability of automated breast density measurements. *Radiology*, 275(2):366–376, 2015.

[10] O. Alonzo-Proulx, N. Packard, J. M. Boone, A. Al-Mayah, K. K. Brock, S. Z. Shen, and M. J. Yaffe. Validation of a method for measuring the volumetric breast density from digital mammograms. *Physics in Medicine and Biology*, 55:3027–3044, 2010.

[11] M. Altrichter, Z. Ludányi, and G. Horváth. Joint analysis of multiple mammographic views in CAD systems for breast cancer detection. In *Proc. Scandinavian Conference on Image Analysis*, pages 760–769, 2005.

[12] J. Ashburner. A fast diffeomorphic image registration algorithm. *NeuroImage*, 38(1):95–113, 2007.

[13] National Electrical Manufacturers Association. *Digital Imaging and Communications in Medicine (DICOM)*. National Electrical Manufacturers Association, 3.1 edition, 2006.

[14] National Electrical Manufacturers Association. http://medical.nema.org. *Digital Imaging and COmmunications in Medicine*, accessed June 15, 2016.

[15] Y. Attikiouzel and R. Chandrasekhar. DSP in mammography. In *Proc. International Conference on Digital Signal Processing*, pages 29–34, 2002.

[16] K. August, B. Biswal, and S. Reisman. Process for reducing false negative mammography bilateral registration. In *Proc. IEEE Annual Northeast Bioengineering Conference*, pages 69–70, 2004.

[17] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26–41, 2008.

[18] D. Avrin, R. Morin, D. Piraino, A. Rowberg, N. Detorie, M. Zuley, J. A. Seibert, and E. D. Pisano. Storage, transmission, and retrieval of digital mammography, including recommendations on image compression. *Journal of the American College of Radiology*, 3:609–614, 2006.

[19] S. R. Aylward, B. H. Hemminger, and E. D. Pisano. Mixture modelling for digital mammogram display and analysis. *Proc. International Workshop on Digital Mammography*, pages 305–312, 1998.

[20] P. Bakic, F. Richard, and A. D. A. Maidment. Effect of breast compression on registration of successive mammograms. In *Proc. International Workshop on Digital Mammography*, 2004.

[21] L.W. Bassett, D. H. Bunnell, R. H. Jahashahi, R. H. Gold, R. D. Arndt, and J. Lisman. Breast cancer detection: one versus two views. *Radiology*, 165:95–97, 1987.

[22] N. D. Bedard, M. P. Sampat, P. A. Stokes, and M. K. Markey. Reducing false-positive detections by combining two stage-1 computer-aided mass detection algorithmscomputerized detection of breast tissue asymmetry depicted on bilateral mammograms: A preliminary study of breast risk stratification. In *Proceedings of the SPIE*, volume 6144, pages 1869–1876, 2006.

[23] E. A. Berns, R. E. Hendrick, M. Solari, L. Barke, D. Reddy, J. Wolfman, L. Segal, P. De Leon, S. Benjamin, and L. Willis. Digital and screen-film mammography: Comparison of image acquisition and interpretation times. *American Journal of Roentgenology*, 187(1):38–41, 2006.

[24] S. Bessa, I. Domingues, j. S. Cardoso, P. Passarinho, P. Cardoso, V. Rodrigues, and F. Lage. Normal breast identification in screening mammography: a study on 18 000 images. In *Proc. IEEE International Conference on Bioinformatics and Biomedicine*, pages 325–330, 2014.

[25] J. C. Bezdek. *Pattern Recognition With Fuzzy Objective Function Algorithms*. Plenum Press, New York, 1981.

[26] U. Bick, M. L. Giger, R. A. Schmidt, R. M. Nishikawa, and K. Doi. Density correction of peripheral breast tissue on digital mammograms. *Radiographics*, 16(6):1403–1411, 1996.

[27] U. Bick, M. L. Giger, R. A. Schmidt, R. M. Nishikawa, D. E. Wolverton, and K. Doi. Automated segmentation of digitized mammograms. *Academic Radiology*, 2(1):1–9, 1995.

[28] U. Bick, M. L. Giger, R. A. Schmidt, R. M. Nishikawa, D. E. Wolverton, and K. Doi. Computer-aided breast cancer detection in screening mammography. In *Proc. International Workshop on Digital Mammography*, pages 97–103, 1998.

[29] L. Blot and R. Zwiggelaar. Background texture extraction for the classification of mammographic parenchymal patterns. In *Proc. Medical Image Understanding and Analysis*, pages 145–148, 2001.

[30] L. Blot and R. Zwiggelaar. A volumetric approach to risk assessment in mammography: a feasibility study. *Physics in Medicine and Biology*, 50(3):695–708, 2005.

[31] A. Bosch, X. Muñoz, A. Oliver, and J. Martí. Modeling and classifying breast tissue density in mammograms. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1552–1558, 2006.

[32] K. Bovis and S. Singh. Classification of mammographic breast density using a combined classifier paradigm. In *Proc. Medical Image Understanding and Analysis*, pages 177–180, 2002.

[33] K. J. Bovis and S. Singh. Detection of masses in mammograms using texture measures. In *Proc. International Conference on Pattern Recognition*, volume 2, pages 267–270, 2000.

[34] N. Boyd, L. Martin, A. Gunasekara, O. Melnichouk, G. Maudsley, C. Peressotti, M. Yaffe, and S. Minkin. Mammographic density and breast cancer

risk: evaluation of a novel method of measuring breast tissue volumes. *Cancer Epidemiology, Biomarkers and Prevention*, 18(6):1754–1762, 2009.

[35] N. F. Boyd, J. W. Byng, R. A. Jong, E. K. Fishell, L. E. Little, A. B. Miller, G. A. Lockwood, D. L. Tritchler, and M. J. Yaffe. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian national breast screening study. *J. Natl Cancer Inst.*, 87(9):670–675, 1995.

[36] N. F. Boyd, G. A. Lockwood, J. W. Byng, D. L. Tritchler, and M. J. Yaffe. Mammographic densities and breast cancer risk. *Cancer Epidemiology, Biomarkers and Prevention*, 7:1133–1144, 1998.

[37] N. F. Boyd, L. J. Martin, J. Stone, L. Little, S. Minkin, and M. J. Yaffe. A longitudinal study of the effects of menopause on mammographic features. *Cancer Epidemiology, Biomarkers and Prevention*, 11(16):1048–1053, 2002.

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

[39] J. Bozek, E. Dumic, and M. Grgic. Bilateral asymmetry detection in digital mammography using b-spline interpolation. In *Proc. International Conference on Systems, Signals and Image Processing*, pages 1–4, 2009.

[40] J. Bozek, M. Grgic, and K. Delac. Comparative analysis of interpolation methods for bilateral asymmetry. In *Proc. IEEE International Symposium of the Croatian Society Electronics in Marine*, pages 1–7, 2010.

[41] J. Bozek, M. Grgic, and J. A. Schnabel. Validation of rigid registration of mammographic images. In *Proc. IEEE International Symposium of the Croatian Society Electronics in Marine*, pages 11–16, 2011.

[42] A. P. Bradley. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159, 1997.

[43] J. S. Brand, K. Czene, J. A. Shepherd, K. Leifland, B. Heddson, A. Sundbom, M. Eriksson, J. Li, K. Humphreys, and P. Hall. Automated measurement of

volumetric mammographic density: a tool for widespread breast cancer risk assessment. *Cancer Epidemiology, Biomarkers and Prevention*, 23(9):1764–1772, 2014.

[44] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[45] R. F. Brem, J. W. Hoffmeister, J. A. Rapelyea, G. Zisman, K. Mohtashemi, G. Jindal, M. P. DiSimio, and S. K. Rogers. Impact of breast density on computer-aided detection for breast cancer. *American Journal of Roentgenology*, 184(2):439–444, 2005.

[46] R. F. Brem, M. J. Lenihan, J. Lieberman, and J. Torrente. Screening breast ultrasound: past, present, and future. *American Journal of Roentgenology*, 204(2):234–240, 2015.

[47] D. Brzakovic, N. Vujovic, M. Neskovic, P. Brzakovic, and K. Fogarty. Mammogram analysis by comparison with previuos screening. In *Proc. International Workshop on Digital Mammography*, pages 131–140, 1994.

[48] G. Bueno, N. Vállez, O. Déniz, P. Esteve, M.A. Rienda, M. Arias, and C. Pastor. Automatic breast parenchymal density classification integrated into a cade system. *International Journal of Computer Assisted Radiology and Surgery*, 6:309–318, 2011.

[49] J. W. Byng. *Mammographic densities and risk of breast cancer*. PhD thesis, Graduate Department of Medical Biophysics, University of Toronto, 1997.

[50] J. W. Byng, N. F. Boyd, E. Fishell, R. A. Jong, and M. J. Yaffe. The quantitative analysis of mammographic densities. *Physics in Medicine and Biology*, 39(10):1629–1638, 1994.

[51] J. W. Byng, N. F. Boyd, E. Fishell, R. A. Jong, and M. J. Yaffe. Automated analysis of mammographic densities. *Physics in Medicine and Biology*, 41(5):909–923, 1996.

[52] J. W. Byng, J. P. Critten, and M. J. Yaffe. Thickness equalization processing for mammographic images. *Radiology*, 203(2):564–568, 1997.

[53] J.W. Byng, M.J. Yaffe, R.A. Jong, R.S. Shumak, G.A. Lockwood, D.L. Tritchler, and N.F. Boyd. Analysis of mammographic density and breast cancer risk from digitized mammograms. *Radiographics*, 18(6):1587–1598, 1998.

[54] J.W. Byng, M.J. Yaffe, G.A. Lockwood, L.E. Little, D.L. Tritchler, and N.F. Boyd. Automated analysis of mammographic densities and breast carcinoma risk. *Cancer*, 80(1):66–74, 1997.

[55] C. B. Caldwell, S. J. Stapleton, D. W. Holdsworth, R. A. Jong, W. J. Weiser, G. Cooke, and M. J. Yaffe. Characterization of mammographic parenchymal pattern by fractal dimension. *Physics in Medicine and Biology*, 35:235–247, 1990.

[56] R. R. Carlton and A. M. Adler. *Principles of radiographic imaging: an art and a science*. Cengage Learning, 2012.

[57] D. Cascio, V. Chietri, F. Fauci, M. Iacomi, R. Ienzi, R. Magro, G. Raso, S. Sorce, and M. Vasile Simone. Fast fourier transform recursive filtering for bilateral mammography comparison. In *European Congress of Radiology*, 2010.

[58] P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangeri, M. L. Pepe, and R. M. Rangayyan. Estimation of the breast skin-line in mammograms using multidirectional gabor filters. *Computers in Biology and Medicine*, 43(11):1870–1881, 2013.

[59] J. Celaya-Padilla, A. Martínez-Torteya, J. Rodríguez-Rojas, J. Galván-Tejada, V. Trevio, and J. Támez-Pea. Bilateral image subtraction and multivariate models for the automated triaging of screening mammograms. *BioMed Research International*, 2015:1–12, 2015.

[60] S. Chatzistergos, J. Stoitsis, K. S. Nikita, and A. Papaevangelou. Development of an integrated breast tissue density classification software system. In *Proc. IEEE International Workshop on Imaging Systems and Techniques*, pages 243–245, 2008.

[61] B. Chen, W. Wang, J. Huang, M. Zhao, G. Cui, J. Xu, W. Guo, P. Du, P. Li, and J. Yu. Comparison of tissue equalization, and premium view post-

processing methods in full field digital mammography. *European Journal of Radiology*, 76(1):73–80, 2010.

[62] Z. Chen, E. R. Denton, and R. Zwiggelaar. Topographic representation based breast density segmentation for mammographic risk assessment. In *Proc. IEEE International Conference on Image Processing*, pages 1993–1996, 2012.

[63] Z. Chen, A. Oliver, E. Denton, and R. Zwiggelaar. Automated mammographic risk classification based on breast density estimation. In *Proc. Iberian Conference on Pattern Recognition and Image Analysis*, pages 237–244, 20213.

[64] Z. Chen and R. Zwiggelaar. A modified fuzzy c-means algorithm for breast tissue density segmentation in mammograms. In *Proc. IEEE Conference on Information Technology Applications in Biomedicine*, pages 1–4, 2010.

[65] S. Ciatto, D. Bernardi, M. Calabrese, M. Durando, M. A. Gentilini, G. Mariscotti, F. Monetti, E. Moriconi, B. Pesce, A. Roselli, Stevanin C., Tapparelli M., and Houssami N. A first evaluation of breast radiological density assessment by quantra software as compared to visual classification. *Breast*, 21(4):503–506, 2012.

[66] J. Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20:27–46, 1960.

[67] A. A. Cole-Rhodes and R. D. Eastman. Gradient descent approaches to image registration. In J. Le Moigne, N. S. Netanyahu, and R. D. Eastman, editors, *Image registration for remote sensing*, chapter 12. Cambridge University Press, 2011.

[68] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[69] W. R. Crum, T. Hartkens, and D. L. G. Hill. Non-rigid image registration: theory and practice. *British Journal of Radiology*, 77(SPEC. ISS. 2):S140–S153, 2014.

[70] O. Cuisenaire and B. Macq. Fast Euclidean distance transformation by propagation using multiple neighborhoods. *Computer Vision and Image Understanding*, 76(2):163–172, 1999.

[71] Y. Díez, A. Oliver, M. Cabezas, S. Valverde, R. Martí, J. C. Vilanova, L. Ramió-Torrentà, À. Rovira, and X. Lladó. Intensity based methods for brain mri longitudinal registration. a study on multiple sclerosis patients. *Neuroinformatics*, 12(3):365–379, 2014.

[72] Y. Díez, A. Oliver, X. Lladó, J. Freixenet, J. Martí, J. C. Vilanova, and R. Martí. Revisiting intensity-based image registration applied to mammography. *IEEE Transactions on Information Technology in Biomedicine*, 15(5):716–725, 2011.

[73] Y. Díez, A. Oliver, X. Lladó, and R. Martí. Comparison of registration methods using mammographic images. In *Proc. IEEE International Conference on Image Processing*, pages 4421–4424, 2010.

[74] K. Doi. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31(4–5):198–211, 2007.

[75] F. Dru, M. P. Wachowiak, and T. M. Peters. An itk framework for deterministic global optimization for medical image registration. In *Proceedings of the SPIE, Medical Imaging*, volume 6144, pages 61442J–1–61442J–12, 2006.

[76] R. P. W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. De Ridder, and D. M. J. Tax. Prtools4: A matlab toolbox for pattern recognition. *Delft University of Technology*, 2004.

[77] S. Eberl, I. Kanno, R. R. Fulton, A. Ryan, B. F. Hutton, and M. J. Fulham. Automated interstudy image registration technique for SPECT and PET. *Journal of Nuclear Medicine*, 37:137–145, 1996.

[78] A. El-Zaart. Expectation–maximization technique for fibro-glandular discs detection in mammography images. *Computers in Biology and Medicine*, 40:392–401, 2010.

[79] A. Elmoufidi, K. El Fahssi, S. J. Andaloussi, and A. Sekkaki. Evaluate dynamic k-means algorithm for automatically segmented different breast regions in mammogram based on density by using seed region growing technique. *Journal of Theoretical and Applied Information Technology*, 72(2):280–288, 2015.

[80] M. Elter and A. Horsch. CADx of mammographic masses and clustered microcalcifications: a review. *Medical Physics*, 36(6):2052–2068, 2009.

[81] A. Eng, Z. Gallant, J. Shepherd, V. McCormack, J. Li, M. Dowsett, S. Vinnicombe, S. Allen, and I. dos Santos Silva. Digital mammographic density and breast cancer risk: a casecontrol study of six alternative density assessment methods. *Breast Cancer Research*, 16(5):1–12, 2014.

[82] D. R. Ericeira, A. C. Silva, A. C. De Paiva, and M. Gattass. Detection of masses based on asymmetric regions of digital bilateral mammograms using spatial description with variogram and cross-variogram functions. *Computers in Biology and Medicine*, 43(8):987–999, 2013.

[83] R. Fabbri, L. D. F. Costa, J. C. Torelli, and O. M. Bruno. 2D Euclidean distance transform algorithms: a comparative survey. *IEEE Transactions on Medical Imaging*, 40(1):2:1–2:44, 2008.

[84] A. X. Falcao, J. Stolfi, and S. de Alencar. The image foresting transform: intelligence theory, algorithms, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):19–29, 2004.

[85] J. Ferlay, E. Steliarova-Foucher, J. Lortet-Tieulent, S. Rosso, J. W. W. Coebergh, H. Comber, D. Forman, and F. Bray. Cancer incidence and mortality patterns in europe: estimates for 40 countries in 2012. *European Journal of Cancer*, 49(6):1374–1403, 2013.

[86] R. J. Ferrari, R. M. Rangayyan, R. A. Borges, and A. F. Frere. Segmentation of the fibro-glandular disc in mammograms via Gaussian mixture modelling. *Medical and Biological Engineering and Computing*, 42(3):378–387, 2004.

[87] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges, and A. F. Frère. Automatic identification of the pectoral muscle in mammograms. *IEEE Transactions on Medical Imaging*, 23(2):232–245, 2004.

[88] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, and A. F. Frère. Analysis of asymmetry in mammograms via directional filtering with gabor wavelets. *IEEE Transactions on Medical Imaging*, 20(9):953–964, 2001.

[89] N. Ferreira and P. J. F. Lucas. Probabilistic relational modelling of mammographic images. In *Proc. IEEE International Symposium on Computer-Based Medical Systems*, pages 1–8, 2009.

[90] N. Ferrreira, M. Velikova, and P. J. F. Lucas. Bayesian modelling of multi-view mammography. In *Proc. International Conference on Machine Learning / Conference on Uncertainty in Artificial Intelligence / Conference on Learning Theory*, 2008.

[91] P. Filev, L. Hadjiiski, B. Sahiner, H. P. Chan, and M. A. Helvie. Comparison of similarity measures for the task of template matching of masses on serial mammograms. *Medical Physics*, 32(2):515–529, 2005.

[92] B. Fischer and J. Modersitzki. Ill-posed medicinean introduction to image registration. *Inverse Problems*, 24(3):034008 (1–16), 2008.

[93] J. M. Fitzpatrick, D. L. G. Hill, and C. R. Maurer Jr. Image registration. In M. Sonka and J. M. Fitzpatrick, editors, *Handbook of medical imaging. Volume 2. , Medical image processing and analysis*, chapter 8. SPIE Press, 2000.

[94] T. W. Freer and M. J. Ulissey. Screening mammography with computer-aided detection: Prospective study of 12860 patients in a community breast center. *Radiology*, 220:781–786, 2001.

[95] J. Freixenet, A. Oliver, X. Lladó, R. Martí, J. Pont, E. Pérez, E.R.E. Denton, and R. Zwiggelaar. Eigendetection of masses considering false positive reduction and breast density information. *Medical Physics*, 35(5):1840–1853, 2008.

[96] Y. Freund and R. E. Schapire. Experiments with a new boosting algorithm. In *Proc. International Conference on Machine Learning / Conference on Uncertainty in Artificial Intelligence / Conference on Learning Theory*, pages 148–156, 1996.

[97] F. Georgsson. Anatomical coordinate system for bilateral registration of mammograms. In *Proc. Scandinavian Conference on Image Analysis*, pages 335–342, 2003.

[98] F. Georgsson. Differential analysis of bilateral mammograms. *International Journal of Pattern Recognition and Artificial Intelligence*, 17(7):1207–1226, 2003.

[99] M. L. Giger, H. P. Chan, and J. Boone. Anniversary paper: history and status of cad and quantitative image analysis: the role of medical physics and aapm. *Medical Physics*, 35(12):5799–5820, 2008.

[100] M. L. Giger, F. F. Yin, K. Doi, Y. Wu, C. J. Vyborny, R. A. Schmidt, and Z. Huo. Computerized detection and characterization of mass lesions in digital mammography. In *Proc. IEEE International Conference on Systems, Man and Cybernetics*, pages 1370–1372, 1992.

[101] A. Godavarty, S. Rodriguez, Y.-J. Jung, and S. Gonzalez. Optical imaging for breast cancer prescreening. *Breast Cancer: Targets and Therapy*, 7:193–209, 2015.

[102] Y. C. Gong, M. Brady, and S. Petroudi. Texture based mammogram classification and segmentation. *Lecture Notes in Computer Science*, 4046:616–625, 2006.

[103] W. F. Good, X. H. Wang, and G. S. Maitz. Feature-based differences between mammograms. In *Proceedings of the SPIE*, volume 5032, pages 919–929, 2003.

[104] M. M. Goodsitt, H. P. Chan, B. Liu, S. Guru, A. R. Morton, S. Keshavmurthy, and N. Petrick. Classification of compressed breast shapes for the design of equalization filters in x-ray mammography. *Medical Physics*, 25(6):937–948, 1998.

[105] A. A. Goshtasby. *Image Registration: Principles, Tools and Methods*. Springer Publishing Company, Incorporated, 2012.

[106] A. Gubern-Mérida, M. Kallenberg, B. Platel, R. M. Mann, R. Martí, and N. Karssemeijer. Volumetric breast density estimation from full-field digital mammograms: a validation study. *Public Library of Science*, 9(1), 2014.

[107] Y. Guo, R. Sivaramakrishna, C. C. Lu, J. Suri, and S. Laxminarayan. Breast image registration techniques: a survey. *Medical and Biological Engineering and Computing*, 44(1):15–26, 2006.

[108] S. Gupta, P. F. Chyn, and M. K. Markey. Breast cancer cadx based on BI-RADS descriptors from two mammographic views. *Medical Physics*, 33(6):1810–1817, 2006.

[109] S. Gupta and M. K. Markey. Correspondence in texture features between two mammographic views. *Medical Physics*, 32(6):1598–1606, 2005.

[110] S. Gupta, D. Zhang, M. P. Sampat, and M. K. Markey. Combining texture features from mlo and cc views for mammographic cadx. In *Proceedings of the SPIE*, volume 6144, pages 61445V.1–61445V.9, 2006.

[111] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46:389–422, 2002.

[112] E. Haber and J. Modersitzki. Numerical methods for volume preserving image registration. *Inverse Problems*, 20(5):1621–1638, 2004.

[113] M. Hachama, A. Desolneux, and F. Richard. A probabilistic approach for the simultaneous mammogram registration and abnormality detection. In *Proc. International Workshop on Digital Mammography*, pages 205–212, 2006.

[114] M. Hadjiiski, H. P. Chan, B. Sahiner, N. Petrick, and M. A. Helvie. Automated registration of breast lesions in temporal pairs of mammograms for interval change analysis–local affine transformation for improved localization. *Medical Physics*, 28(6):1070–1079, 2001.

[115] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.

[116] R. M. Haralick, K. S. Shanmugan, and I. Dunstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621, 1973.

[117] K. Hartman, R. Highnam, R. Warren, and V. Jackson. Volumetric assessment of breast tissue composition from ffdm images. In *Proc. International Workshop on Digital Mammography*, pages 33–39, 2008.

[118] A. Hasegawa, H. Neemuchwala, H. Tsunoda-Shimizu, S. Honda, K. Shimura, M. Sato, T. Koyama, M. Kikuchi, and S. Hiramatsu. A tool for temporal comparison of mammograms: Image toggling and dense-tissue-preserving registration. In *Proc. International Workshop on Digital Mammography*, pages 447–454, 2008.

[119] P. M. Hayton, M. Brady, S. M. Smith, and N. Moore. A non-rigid registration algorithm for dynamic breast mr images. *Artificial Intelligence*, 114(1):125–156, 1999.

[120] W. He, E. R. Denton, and R. Zwiggelaar. Mammographic image segmentation and risk classification using a novel texture signature based methodology. In *Proc. International Workshop on Digital Mammography*, pages 526–533, 2010.

[121] W. He, E. R. Denton, and R. Zwiggelaar. Mammographic segmentation and risk classification using a novel binary model based bayes classifier. In *Proc. International Workshop on Digital Mammography*, pages 40–47, 2012.

[122] W. He, E. R. E. Denton, K. Stafford, and R. Zwiggelaar. Mammographic image segmentation and risk classification based on mammographic parenchymal patterns and geometric moments. *Biomedical Signal Processing and Control*, 6(3):321–329, 2011.

[123] W. He, E. R. E. Denton, and R. Zwiggelaar. Mammographic segmentation based on mammographic parenchymal patterns and spatial moments. In *Proc.*

IEEE Conference on Information Technology Applications in Biomedicine, pages 1–4, 2009.

[124] W. He, A. Juette, E. R. E. Denton, and R. Zwiggelaar. Novel multiresolution mammographic density segmentation using pseudo 3D features and adaptive cluster merging. In *Proceedings of the SPIE, Medical Imaging*, volume 9413, pages 94133I–1–94133I–6, 2015.

[125] W. He, M. Kibiro, A. Juette, P. Hogg, E. R. Denton, and R. Zwiggelaar. A novel breast image preprocessing for full field digital mammographic segmentation and risk classification. In *Proc. Medical Image Understanding and Analysis*, pages 40–47, 2014.

[126] W. He, I. Muhimmah, E. R. Denton, and R. Zwiggelaar. Mammographic segmentation based on texture modelling of tabár mammographic building blocks. In *Proc. International Workshop on Digital Mammography*, pages 17–24, 2008.

[127] W. He and R. Zwiggelaar. Breast parenchymal pattern analysis in digital mammography: Associations between tabár and birads tissue compositions. In *Proc. International Conference on Computer Analysis of Images and Patterns*, pages 386–393, 2013.

[128] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. J. Kegelmeyer. The Digital Database for Screening Mammography. In *Proc. International Workshop on Digital Mammography*, pages 212–218, 2000.

[129] M. D. Heath and K. W. Bowyer. Mass detection by relative image intensity. In *Proc. International Workshop on Digital Mammography*, pages 219–225, 2000.

[130] J. J. Heine, K. Cao, D. E. Rollison, G. Tiffenberg, and J. A. Thomas. A quantitative description of the percentage of breast density measurement using full-field digital mammography. *Academic Radiology*, 18(5):556–564, 2011.

[131] J. J. Heine, M. J. Carston, C. G. Scott, K. R. Brandt, F. F. Wu, V. S. Pankratz, T. A. Sellers, and C. M. Vachon. An automated approach for esti-

mation of breast density. *Cancer Epidemiology, Biomarkers and Prevention*, 17(11):3090–3097, 2008.

[132] J. J. Heine, C. G. Scott, T. A. Sellers, K. R. Brandt, D. J. Serie, F. F. Wu, M. J. Morton, B. A. Schueler, F. J. Couch, J. E. Olson, Pankratz V.S., and Vachon C.M. A novel automated mammographic density measure and breast cancer risk. *Journal of the National Cancer Institute*, 104(13):1028–1037, 2012.

[133] J. J. Heine and R. P. Veltzuihen. A statistical methodology for mammographic density detection. *Medical Physics*, 27(12):2644–2651, 2000.

[134] R. Highnam, M. Brady, M. J. Yaffe, N. Karssemeijer, and J. Harvey. Robust breast composition measurement-volparatm. In *Proc. International Workshop on Digital Mammography*, pages 342–349, 2010.

[135] W. T. Ho and P. W. T. Lam. Clinical performance of computer-assisted detection (CAD) system in detecting carcinoma in breasts of different densities. *Clinical Radiology*, 58(2):133–136, 2003.

[136] C. K. Hoh, M. Dahlbom, G. Harris, Y. Choi, R. A. Hawkins, M. E. Phelps, and J. Maddahi. Automated iterative three-dimensional registration of positron emission tomography images. *Journal of Nuclear Medicine*, 34:2009–2018, 1993.

[137] Inc. Hologic. http://www.hologic.com/products/imaging/mammography/image-analytics. *ImageChecker CAD*, accessed June 15, 2016.

[138] C. Huang and O. Mitchell. A Euclidean distance transform using grayscale morphology decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(4):443–448, 1994.

[139] R. Hupse and N. Karssemeijer. Use of normal tissue context in computer aided detection of masses in mammograms. *IEEE Transactions on Medical Imaging*, 28(12):2033–2041, 2009.

[140] L. Ibáñez, L. Ng, J. Gee, and S. Aylward. Registration patterns: the generic framework for image registration of the insight toolkit. In *Proc. IEEE Workshop on Biomedical Image Analysis*, pages 345–348, 2002.

[141] L. Ibáñez, W. Schroeder, L. Ng, and J. Cates. *The ITK Software Guide*. Kitware, Inc. ISBN 1-930934-10-6, http://www.itk.org/ItkSoftwareGuide.pdf, first edition, 2003.

[142] iCAD Inc. http://www.icadmed.com/secondlook-digital.html. *SecondLook® Digital*, accessed June 15, 2016.

[143] A. K. Jain, Y. Zhong, and S. Lakshmanan. Object matching using deformable templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(3):267–278, 1996.

[144] G. Janssens, L. Jacques, J. Orban de Xivry, X. Geets, and B. Macq. Diffeomorphic registration of images with variable contrast enhancement. *International Journal of Biomedical Imaging*, 2011, 2011.

[145] M. Jeffreys, J. Harvey, and R. Highnam. Comparing a new volumetric breast density method (volparatm) to cumulus. In *Proc. International Workshop on Digital Mammography*, pages 408–413, 2010.

[146] G. H. John and P. Langley. Estimating continuous distributions in bayesian classifiers. In *Proc. International Conference on Machine Learning / Conference on Uncertainty in Artificial Intelligence / Conference on Learning Theory*, pages 338–345, 1995.

[147] H. J. Johnson and G. E. Christensen. Landmark and intensity-based, consistent thin-plate spline image registration. In *Proc. of the Information Processing in Medical Imaging*, pages 329–343, 2001.

[148] I. T. Jolliffe. *Principal Component Analysis*. Springer, 2nd edition, 2002.

[149] S. Kabus, T. Netsch, B. Fischer, and J. Modersitzki. B-spline registration of 3d images with levenberg-marquardt optimization. In *Proceedings of the SPIE, Medical Imaging*, pages 304–313, 2004.

[150] M. G. J. Kallenberg and N. Karssemeijer. Compression paddle tilt correction in full-field digital mammograms. *Physics in Medicine and Biology*, 57(3):703–715, 2012.

[151] M. G. J. Kallenberg, M. Lokate, C. H. Van Gils, and N. Karssemeijer. Automatic breast density segmentation: an integration of different approaches. *Physics in Medicine and Biology*, 56(9):2715–2729, 2011.

[152] M. Kallergi, G. M. Carney, and J. Gaviria. Evaluating the performance of detection algorithms in digital mammography. *Medical Physics*, 26:267–275, 1999.

[153] N. Karssemeijer. Automated classification of parenchymal patterns in mammograms. *Physics in Medicine and Biology*, 43:365–378, 1998.

[154] N. Karssemeijer and G. M. te Brake. Detection of stellate distortions in mammograms. *IEEE Transactions on Medical Imaging*, 15(5):611–619, 1996.

[155] N. Karssemeijer and G. M. te Brake. Combining single view features and asymmetry for detection of mass lesions. In *Proc. International Workshop on Digital Mammography*, pages 95–102, 1998.

[156] S. Katsuhara, H. Futamura, S. Kasai, T. Morita, and T. Endo. Computerized scheme for focal asymmetric densities on mammograms by use of geometric and texture analyses. In *Proc. International Workshop on Digital Mammography*, pages 315–322, 2008.

[157] B. Keller, D. Nathan, Y. Wang, Y. Zheng, J. Gee, E. Conant, and D. Kontos. Adaptive multi-cluster fuzzy c-means segmentation of breast parenchymal tissue in digital mammography. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 562–569, 2011.

[158] B. M. Keller, D. L. Nathan, Y. Wang, Y. Zheng, J. C. Gee, E. F. Conant, and D. Kontos. Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation. *Medical Physics*, 39(8):4903–4917, 2012.

[159] S. P. Keshavmurthy, M. M. Goodsitt, H. P. Chan, M. A. Helvie, and E. Christodoulou. Design and evaluation of an external filter technique for

exposure equalization in mammography. *Medical Physics*, 26(8):1655–1669, 1999.

[160] J. Kim and J. A. Fessler. Intensity-based image registration using robust correlation coefficients. *IEEE Transactions on Medical Imaging*, 23:1430–1444, 2004.

[161] Y. Kim, C. Kim, and J. H. Kim. Automated estimation of breast density on mammogram using combined information of histogram statistics and boundary gradients. In *Proceedings of the SPIE, Medical Imaging*, pages 76242F–76242F, 2010.

[162] Y. Kita, R. Highnam, and M. Brady. Correspondence between different view breast x-rays using a simulation of breast deformation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 700–707, 1998.

[163] Y. Kita, R. Highnam, and M. Brady. Correspondence between different view breast x-rays using curved epipolar lines. *Computer Vision and Image Understanding*, 83(1):38–56, 2001.

[164] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. Hyun Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parseya. Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *NeuroImage*, 46(3):786–802, 2009.

[165] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim. Elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 2010.

[166] S. Klein, M. Staring, and J. P. W. Pluim. Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines. *IEEE Transactions on Image Processing*, 16(12):2879–2890, 2007.

[167] S. L. Kok-Wiles, M. Brady, and R. Highman. Comparing mammogram pairs for the detection of lesions. In *Proc. International Workshop on Digital Mammography*, pages 103–110, 1998.

[168] D. Kopans. *Breast Imaging*. Lippincott-Raven, Philadelphia, 1998.

[169] K. Kriti and J. Virmani. Breast density classification using laws' mask texture features. *International Journal of Biomedical Engineering and Technology*, 19(3), 2015.

[170] C. K. Kuhl, S. Schrading, C. C. Leutner, N. Morakkabati-Spitz, E. Wardelmann, R. Fimmers, W. Kuhn, and H. H. Schild. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *Journal of Clinical Oncology*, 23(33):8469–8476, 2005.

[171] R. Kumar, J. C. Asmuth, K. Hanna, J. Bergen, C. Hulka, D. B. Kopans, R. Weisskoff, and R. H. Moore. Application of 3D registration for detecing lesions in magnetic resonance scans. In *Proceedings of the SPIE, Medical Imaging*, volume 2710, pages 646–656, 1996.

[172] C. M. Kuzmiak, G. A. Millnamow, B. Qaqish, E. D. Pisano, E. B. Cole, and M. E. Brown. Comparison of full-field digital mammography to screen-film mammography with respect to diagnostic accuracy of lesion characterization in breast tissue biopsy specimens. *Academic Radiology*, 9:1378–1382, 2002.

[173] S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, and M. T. Rickard. Automatic pectoral muscle segmentation on mediolateral oblique view mammograms. *IEEE Transactions on Medical Imaging*, 23(9):1129–1140, 2004.

[174] J. Kybic and M. Unser. Fast parametric elastic image registration. *IEEE Transactions on Medical Imaging*, 12(11):1427–1442, 2003.

[175] S. H. Lai and M. Fang. An adaptive window width/center adjustment system with online training capabilities for MR images. *Artificial Intelligence in Medicine*, 33:89–101, 2005.

[176] T. Lange, N. Papenberg, S. Heldmann, J. Modersitzki, B. Fischer, H. Lamecker, and P. M. Schlag. 3D ultrasound-CT registration of the liver using combined landmark-intensity information. *International Journal of Computer Assisted Radiology and Surgery*, 4(1):79–88, 2009.

[177] T. Lau and W. F. Bischof. Automated detection of breast tumors using the asymmetry approach. *Computers and Biomedical Research*, 24(3):273–295, 1991.

[178] K. Laws. Rapid texture identification. *SPIE, Image Processing for Missile Guidance*, 238:376–380, 1980.

[179] M. J. Ledesma-Carbayo, J. Kybic, M. Desco, A. Santos, M. Sühling, P. Hunziker, and M. Unser. Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation. *IEEE Transactions on Medical Imaging*, 24(9):1113–1126, 2005.

[180] T. M. Lehmann, C. Gönner, and K. Spitzer. Survey: Interpolation methods in medical image processing. *IEEE Transactions on Medical Imaging*, 18(11):1049–1075, 1999.

[181] J. Li, L. Szekely, L. Eriksson, B. Heddson, A. Sundbom, K. Czene, P. Hall, and K. Humphreys. High-throughput mammographic-density measurement: a tool for risk prediction of breast cancer. *Breast Cancer Research*, 14(4):1–12, 2012.

[182] X. Li, B. M. Dawant, E. B. Welch, A. B. Chakravarthy, D. Freehardt, I. Mayer, M. Kelley, I. Meszoely, J. C. Gore, and T. E. Yankeelov. A nonrigid registration algorithm for longitudinal breast mr images and the analysis of breast tumor response. *Magnetic Resonance Imaging*, 27(9):1258–1270, 2009.

[183] L. Liu, J. Wang, and K. He. Breast density classification using histogram moments of multiple resolution mammograms. In *Proc. International Conference on Biomedical Engineering and Informatics*, volume 1, pages 146–149, 2010.

[184] M. Lokate, M. G. Kallenberg, N. Karssemeijer, M. A. Van den Bosch, P. H. Peeters, and C. H. Van Gils. Volumetric breast density from full-field digital mammograms and its association with breast cancer risk factors: a comparison with a threshold method. *Cancer Epidemiology, Biomarkers and Prevention*, 19(12):3096–3105, 2010.

[185] M. Lokate, P.H.M. Peeters, L.M. Peelen, G. Haars, W.B. Veldhuis, and C.H. Gils. Mammographic density and breast cancer risk: the role of the fat surrounding the fibroglandular tissue. *Breast Cancer Research*, 13(5):1–8, 2011.

[186] L. J. W. Lu, T. K. Nishino, T. Khamapirad, J. J. Grady, M. H. Leonard Jr, and D. G. Brunder. Computing mammographic density from a multiple regression model constructed with image-acquisition parameters from a full-field digital mammographic unit. *Physics in Medicine and Biology*, 52(16):4905, 2007.

[187] Y. Lu, M. Yousefi, J. Ellenberger, R. H. Moore, D. B. Kopans, A. Krzyzak, and C. Y. Suen. 3D tomosynthesis to detect breast cancer. In C. H. Chen, editor, *Handbook of Pattern Recognition and Computer Vision*, chapter 2.10. World Scientific, 2015.

[188] F. Ma, M. Bajger, and M. J. Bottema. Temporal analysis of mammograms based on graph matching. In *Proc. International Workshop on Digital Mammography*, pages 158–165, 2008.

[189] F. Ma, M. Bajger, S. Williams, and M. J. Bottema. Improved detection of cancer in screening mammograms by temporal comparison. In *Proc. International Workshop on Digital Mammography*, pages 752–759, 2010.

[190] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[191] F. Maes, D. Vandermeulen, and P. Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.

[192] I. E. Magnin, F. Cluzeau, C. L. Odet, and A. Bremond. Mammographic texture analysis: an evaluation of risk for developing breast cancer. *Optical Engineering*, 25:780–784, 1986.

[193] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 2008.

[194] S. Malkov, J. Wang, K. Kerlikowske, S. R. Cummings, and J. A. Shepherd. Single x-ray absorptiometry method for the quantitative mammographic measure of fibroglandular tissue volume. *Medical Physics*, 36(12):5525–5536, 2009.

[195] K. Marias, C. Behrenbruch, S. Parbhoo, A. Seifalian, and M. Brady. A registration framework for the comparison of mammogram sequences. *IEEE Transactions on Medical Imaging*, 24(6):782–790, 2005.

[196] K. Marias, C.P. Behrenbruch, M. Brady, S. Parbhoo, and A. Seifalian. Multiscale landmark selection for improved registration of temporal mammograms. In *Proc. International Workshop on Digital Mammography*, pages 580–586, 2000.

[197] K. Marias, J. M. Brady, R. P. Highnam, S. Parbhoo, A. M. Seifalian, and M. Wirth. Registration and matching of temporal mammograms for detecting abnormalities. *Proc. Medical Image Understanding and Analysis*, 1999.

[198] K. Marias, M. G. Linguraru, M. Ballester, S. Petroudi, M. Tsiknakis, and M. Brady. Automatic labelling and BI-RADS characterisation of mammogram densities. In *Proc. International Conference IEEE Engineering in Medicine and Biology Society*, pages 6394–6398, 2005.

[199] R. Martí, D. Raba, A. Oliver, and R. Zwiggelaar. Mammographic registration: proposal and evaluation of a new approach. In *Proc. International Workshop on Digital Mammography*, pages 213–220, 2006.

[200] R. Martí, R. Zwiggelaar, and C. M. E. Rubin. Automatic point correspondence and registration based on linear structures. *International Journal of Pattern Recognition and Artificial Intelligence*, 16(3):331–340, 2002.

[201] K. E. Martin, M. A. Helvie, C. Zhou, M. A. Roubidoux, J. E. Bailey, C. Paramagul, C. E. Blane, K. A. Klein, S. S. Sonnad, and H. P. Chan. Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and BI-RADS categories. *Radiology*, 240(3):656–665, 2006.

[202] C. R. Maurer, R. Qi, and V. Raghavan. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):265–270, 2003.

[203] V. A. McCormack and I. dos Santos Silva. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiology, Biomarkers and Prevention*, 15(6):1159–1169, 2006.

[204] V. A. McCormack, N. Highnam, R.and Perry, and I. dos Santos Silva. Comparison of a new and existing method of mammographic density measurement: intramethod reliability and associations with known risk factors. *Cancer Epidemiology, Biomarkers and Prevention*, 16(6):1148–1154, 2007.

[205] A. J. Méndez, M. Souto, P. G. Tahoces, and J. J. Vidal. Computer aided diagnosis for breast masses detection on a telemammography system. *Computerized Medical Imaging and Graphics*, 27:497–502, 2003.

[206] A. J. Méndez, P. G. Tahoces, M. J. Lado, M. Souto, and J. J. Vidal. Automatic detection of breast border and nipple in digital mammograms. *Computer Methods and Programs in Biomedicine*, 49(3):253–262, 1996.

[207] A. J. Méndez, P. G. Tahoces, M. J. Lado, M. Souto, and J. J. Vidal. Computerized-aided diagnosis: Automatic detection of malignant masses in digitized mammograms. *Medical Physics*, 25(6):957–964, 1998.

[208] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, G. L. Lecarpentier, J. B. Fowlkes, M. A. Roubidoux, and P. L. Carson. Semiautomatic registration of volumetric ultrasound scans. *Ultrasound in Medicine & Biology*, 25(3):339–347, 1999.

[209] M. I. Miga. A new approach to elastography using mutual information and finite elements. *Physics in Medicine and Biology*, 48(4):467, 2003.

[210] P. Miller and S. M. Astley. Classification of breast tissue by texture analysis. *Image and Vision Computing*, 10:277–282, 1992.

[211] P. Miller and S. M. Astley. Automated detection of breast asymmetry using anatomical features. In *State of the Art in Digital Mammographic Image Analysis*, pages 247–261, 1994.

[212] J. Mitra, A. Oliver, R. Martí, X. Lladó, J. C. Vilanova, and F. Meriaudeau. A thin-plate spline based multimodal prostate registration with optimal correspondences. In *Proc. IEEE Conference on Signal-Image Technologies and Internet-Based System*, pages 7–11, 2010.

[213] Mohanalin, P. K. Kalra, and Kumar N. An automatic method to enhance microcalcifications using normalized tsallis entropy. *Signal Processing*, 90(3):952–958, 2010.

[214] I Muhimmah, W. He, E. R. Denton, and R. Zwiggelaar. Segmentation based on textons and mammographic building blocks. In *Proc. Medical Image Understanding and Analysis*, pages 228–232, 2007.

[215] I. Muhimmah, A. Oliver, E. R. E. Denton, J. Pont, E. Pérez, and R. Zwiggelaar. Comparison between Wolfe, Boyd, BI-RADS and Tabár based mammographic risk assessment. In *Proc. International Workshop on Digital Mammography*, pages 407–415, 2006.

[216] I. Muhimmah and R. Zwiggelaar. Mammographic density classification using multiresolution histogram information. In *Proc. IEEE Conference on Information Technology Applications in Biomedicine*, 2006.

[217] M. Mustra and M. Grgic. Dense tissue segmentation in digitized mammograms. In *Proc. IEEE International Symposium of the Croatian Society Electronics in Marine*, pages 55–58, 2013.

[218] M. Mustra and M. Grgic. Robust automatic breast and pectoral muscle segmentation from scanned mammograms. *Signal Processing*, 93(10):2817–2827, 2013.

[219] M. Mustra, M. Grgic, and K. Delac. Feature selection for automatic breast density classification. In *Proc. IEEE International Symposium of the Croatian Society Electronics in Marine*, pages 9–16, 2010.

[220] L. Nanni, A. Lumini, and S. Brahnam. Local binary patterns variants as texture descriptors for medical image analysis. *Artificial Intelligence in Medicine*, 49:117–125, 2010.

[221] J. T. Neyhart, M. Kirlakovsky, L. M. Coleman, R. Polikar, M. Tseng, and S. A. Mandayam. Automated segmentation and quantitative characterization of radiodense tissue in digitized mammograms. In *AIP Conference Proceedings*, volume 615, pages 1866–1873, 2002.

[222] C. Nickson, Y. Arzhaeva, Z. Aitken, T. Elgindy, M. Buckley, M. Li, English D. R., and A. M. Kavanagh. Autodensity: an automated method to measure mammographic breast density that predicts breast cancer risk and screening outcomes. *Breast Cancer Research*, 15(5):R80, 2013.

[223] R. M. Nishikawa. Computer-aided detection and diagnosis. In U. Bick and F. Diekmann, editors, *Digital Mammography*, chapter 6. Springer Berlin Heidelberg, 2010.

[224] L. Nyul and J. Udupa. On standardizing the MR image intensity scale. *Magnetic Resonance in Medicine*, 42:1072–1081, 1999.

[225] International Association of Cancer Registries. http://globocan.iarc.fr. *Global Cancer Statistics*, accessed June 15, 2016.

[226] U.S. Department of Health and Human Services. http://www.fda.gov. *U.S. Food and Drug Administration*, accessed June 15, 2016.

[227] American College of Radiology. *Illustrated Breast Imaging Reporting and Data System BIRADS*. American College of Radiology, 3rd edition, 1998.

[228] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative-study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29(1):51–59, 1996.

[229] A. Oliver, J. Freixenet, J. Martí, E. Pérez, J. Pont, E. R. E. Denton, and R. Zwiggelaar. A review of automatic mass detection and segmentation in mammographic images. *Medical Image Analysis*, 14(2):87–110, 2010.

[230] A. Oliver, J. Freixenet, R. Martí, J. Pont, E. Pérez, E.R.E. Denton, and R. Zwiggelaar. A novel breast tissue density classification methodology. *IEEE Transactions on Information Technology in Biomedicine*, 12(1):55–65, 2008.

[231] A. Oliver, J. Freixenet, and R. Zwiggelaar. Automatic classification of breast density. In *Proc. IEEE International Conference on Image Processing*, volume 2, pages 1258–1261, 2005.

[232] A. Oliver, X. Lladó, J. Freixenet, R. Martí, E. Pérez, J. Pont, E.R.E. Denton, and R. Zwiggelaar. Influence of using manual or automatic breast density information in a mass detection cad system. *Academic Radiology*, 17(7):877–883, 2010.

[233] A. Oliver, X. Llado, R. Martí, J. Freixenet, and R. Zwiggelaar. Classifying mammograms using texture information. In *Proc. Medical Image Understanding and Analysis*, pages 223–227, 2007.

[234] A. Oliver, X. Lladó, E. Pérez, J. Pont, E. R. Denton, J. Freixenet, and J. Martí. A statistical approach for breast density segmentation. *Journal of Digital Imaging*, 23(5):527–537, 2010.

[235] A. Oliver, A. Torrent, X. Lladó, M. Tortajada, L. Tortajada, M. Sentís, J. Freixenet, and R. Zwiggelaar. Automatic microcalcification and cluster detection in digital and digitised mammograms. *Journal of Knowledge-Based Systems*, 28:68–75, 2012.

[236] C. Olsén and A. Mukhdoomi. Automatic segmentation of fibroglandular tissue. In *Proc. Scandinavian Conference on Image Analysis*, pages 679–688, 2007.

[237] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[238] Y. Ou, S. P. Weinstein, E. F. Conant, S. Englander, X. Da, B. Gaonkar, M. K. Hsieh, M. Rosen, A. DeMichele, C. Davatzikos, and D. Kontos. Deformable registration for quantifying longitudinal tumor changes during neoadjuvant chemotherapy. *Magnetic Resonance in Medicine*, 73(6):2343–2356, 2015.

[239] J. Padayachee, M. Alport, and W. Rae. Mammographic cad: Correlation of regions in ipsilateral views–a pilot study. *South African Journal of Radiology*, 13(3):48–54, 2009.

[240] S. Paquerault, N. Petrick, H. P. Chan, B. Sahiner, and M. A. Helvie. Improvement of computerized mass detection on mammograms: Fusion of two-view information. *Medical Physics*, 29(2):238–247, 2002.

[241] L. F. Parr, A. L. Anderson, B. R. Glennon, and P. Fetherston. Quality control issues on high resolution diagnostic monitors. *Journal of Digital Imaging*, 14(2 suppl 1):22–26, 2001.

[242] O. Pawluczyk, B. J. Augustine, M. J. Yaffe, D. Rico, J. Yang, G. E. Mawdsley, and N. F. Boyd. A volumetric method for estimation of breast density on digitized screen-film mammograms. *Medical Physics*, 30(3):352–364, 2003.

[243] H. R. Peppard, B. E. Nicholson, C. M. Rochman, J. K. Merchant, R. C. Mayo III, and J. H. Harvey. Digital breast tomosynthesis in the diagnostic setting: Indications and clinical applications. *Radiographics*, 35(4):975–990, 2015.

[244] D. C. Pereira, M. Z. Nascimento, R. P. Ramos, and R. D. Dantas. Automatic detection of breast masses using two-view mammography. In *World Congress on Medical Physics and Biomedical Engineering*, pages 917–920, 2009.

[245] S. Petroudi and M. Brady. Breast density segmentation using texture. *Lecture Notes in Computer Science*, 4046:609–615, 2006.

[246] S. Petroudi, T. Kadir, and M. Brady. Automatic classification of mammographic parenchymal patterns: A statistical approach. In *Proc. International Conference IEEE Engineering in Medicine and Biology Society*, volume 1, pages 798–801, 2003.

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

[248] J. M. Peyrat, H. Delingette, M. Sermesant, C. Xu, and N. Ayache. Registration of 4D cardiac CT sequences under trajectory constraints with multichannel diffeomorphic demons. *IEEE Transactions on Medical Imaging*, 29(7):1351–1368, 2010.

[249] J. W. Pluim, J. A. Maintz, and M. Viergever. Mutual-information-based registration of medical images: A survey. *IEEE Transactions on Medical Imaging*, 22:986–1004, 2003.

[250] C. E. Priebe, R. A. Lorey, D. J. Marchette, Solka J. L., and Rogers G. W. Nonparametric spatio-temporal change point analysis for early detection in mammography. In *Proc. International Workshop on Digital Mammography*, pages 111–120, 1994.

[251] J. Pu, B. Zheng, J. K. Leader, and D. Gur. An ellipse-fitting based method for efficient registration of breast masses on two mammographic views. *Medical Physics*, 35(2):487–494, 2008.

[252] W. Qian, D. Song, M. Lei, R. Sankar, and E. Eikman. Computer-aided mass detection based on ipsilateral multiview mammograms. *Academic Radiology*, 14(5):530–538, 2007.

[253] Y. Qu, C. Shang, Q. Shen, N. M. Parthaláin, and W. Wu. Kernel-based fuzzy-rough nearest-neighbour classification for mammographic risk analysis. *International Journal of Fuzzy Systems*, 17(3):471–483, 2015.

[254] D. Raba, A. Oliver, J. Martí, M. Peracaula, and J. Espunya. Breast segmentation with pectoral muscle suppression on digital mammograms. In *Lecture Notes in Computer Science*, volume 3523, pages 471–478, 2005.

[255] R. M. Rangayyan, F. J. Ayres, and J. E. L. Desautels. A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. *Journal of the Franklin Institute*, 344(3–4):312–348, 2007.

[256] F. J. P. Richard. A new approach for the registration of images with inconsistent differences. In *Proc. International Conference on Pattern Recognition*, volume 4, pages 649–652, 2004.

[257] F. J. P. Richard and L. D. Cohen. A new image registration technique with free boundary constraints: application to mammography. *Computer Vision and Image Understanding*, 89:166–196, 2003.

[258] C. Romero, C. Varela, R. Cuena, A. Almenar, J. M. Pinto, and M. Botella. Impact of mammographic breast density on computer-assisted detection (CAD) in a breast imaging department. *Radiología*, 53(5):456–461, 2011.

[259] A. Rosenfeld and J. Pfaltz. Sequential operations in digital picture processing. *Journal of the ACM*, 13(4):471–494, 1966.

[260] D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers. Diffeomorphic registration using b-splines. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 702–709, 2006.

[261] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.

[262] P. K. Saha, J. K. Udupa, E. F. Conant, P. Chakraborty, and D. Sullivan. Breast tissue density quantification via digitized mammograms. *IEEE Transactions on Medical Imaging*, 20(8):792–803, 2001.

[263] N. Saidin, U. K. Ngah, H. Sakim, D. N. Siong, and M. K. Hoe. Density based breast segmentation for mammograms using graph cut techniques. In *Proc. IEEE Region 10 Conference*, pages 1–5, 2009.

[264] M. Sallam and K. Bowyer. Detecting abnormal densities in mammograms by comparison with previous screenings. In *Proc. International Workshop on Digital Mammography*, pages 417–420, 1996.

[265] M. Sallam and K. Bowyer. Registration and difference analysis of corresponding mammogram images. *Medical Image Analysis*, 3(2):103–118, 2001.

[266] M. Samulski and N. Karssemeijer. Matching mammographic regions in mediolateral oblique and cranio caudal views: a probabilistic approach. In *Proceedings of the SPIE*, volume 6915, page 69151M, 2008.

[267] S. Sanjay-Gopal, H. P. Chan, T. Wilson, M. Helvie, N. Petrick, and B. Sahiner. A regional registration technique for automated interval change analysis of breast lesions on mammograms. *Medical Physics*, 26(12):2669–2679, 1999.

[268] J. Santos, A. J. Chaudhari, A. A. Joshi, A. Ferrero, K. Yang, J. M. Boone, and R. D. Badawi. Non-rigid registration of serial dedicated breast ct, longitudinal dedicated breast ct and pet/ct images using the diffeomorphic demons method. *Physica Medica*, 30(6):713–717, 2014.

[269] S. E. Selvan, C. C. Xavier, N. Karssemeijer, J. Sequeira, R. A. Cherian, and B. Y. Dhala. Parameter estimation in stochastic mammogram model by heuristic optimization techniques. *IEEE Transactions on Information Technology in Biomedicine*, 10(4):685–695, 2006.

[270] J.M. Seo, E.S. Ko, B.K. Han, E.Y. Ko, J.H. Shin, and S.Y. Hahn. Automated volumetric breast density estimation: a comparison with visual assessment. *Clinical Radiology*, 68(7):690–695, 2013.

[271] J. A. Shepherd, L. Herve, J. Landau, B. Fan, K. Kerlikowske, and S. R. Cummings. Novel use of single x-ray absorptiometry for measuring breast density. *Technology in Cancer Research and Treatment*, 4(2):173–182, 2005.

[272] J. A. Shepherd, K. Kerlikowske, L. Ma, F. Duewer, B. Fan, J. Wang, S. Malkov, E. Vittinghoff, and S. R. Cummings. Volume of mammographic density and risk of breast cancer. *Cancer Epidemiology, Biomarkers and Prevention*, 20(7):1473–1482, 2011.

[273] J. A. Shepherd, K. M. Kerlikowske, R. Smith-Bindman, H. K. Genant, and S. R. Cummings. Measurement of breast density with dual x-ray absorptiometry: Feasibility 1. *Radiology*, 223(2):554–557, 2002.

[274] F. Y. C. Shih and Y. T. Wu. The efficient algorithms for achieving Euclidean distance transformation. *IEEE Transactions on Image Processing*, 13(8):1078–1091, 2004.

[275] E. A. Sickles, C. J. D'Orsi, L. W. Bassett, C. M. Appleton, W. A. Berg, and E. S. Burnside. Acr bi-rads mammography. In *ACR BI-RADS Atlas, Breast Imaging Reporting and Data System*. Reston, V. A.: American College of Radiology, 2013.

[276] E. Siegel, E. Krupinski, E. Samei, M. Flynn, K. Andriole, B. Erickson, J. Thomas, A. Badano, J. A. Seibert, and E. D. Pisano. Digital mammography image quality: Image display. *Journal of the American College of Radiology*, 3:615–627, 2006.

[277] V. Siless, P. Guevara, X. Pennec, and P. Fillard. Joint t1 and brain fiber diffeomorphic registration using the demons. In *Proc. International Workshop on Multimodal Brain Image Analysis*, pages 10–18, 2011.

[278] S. Singh and K. Bovis. An evaluation of contrast enhancement techniques for mammographic breast masses. *IEEE Transactions on Information Technology in Biomedicine*, 9(1):109–119, 2005.

[279] R. Sivaramakrishna, N. A. Obuchowski, W. A. Chilcote, and K. A. Powell. Automatic segmentation of mammographic density. *Academic Radiology*, 8(3):250–256, 2001.

[280] P. Skaane, A. I. Bandos, E. B. Eben, I. N. Jebsen, M. Krager, U. Haakenaasen, U. Ekseth, M. Izadi, S. Hofvind, and R. Gullien. Two-view digital breast tomosynthesis screening with synthetically reconstructed projection images: comparison with digital breast tomosynthesis with full-field digital mammographic images. *Radiographics*, 271(3):655–663, 2014.

[281] A. Smith. Fundamentals of breast tomosynthesis. *White Paper, Hologic Inc.*, 8, 2008.

[282] A. P. Smith. Fundamentals of digital mammograhpy: physiscs, technology and practical considerations. *Radiology Management*, 25(5):18–31, 2003.

[283] S. M. Smith, G. Kindlmann, and S. Jbabdi. Cross-subject comparison of local diffusion MRI parameters. In H. Johansen-Berg and T. E. J. Behrens, editors, *Diffusion {MRI} (Second Edition)*, chapter 10. Academic Press, 2014.

[284] P. R. Snoeren and N. Karssemeijer. Thickness correction of mammographic images by means of a global parameter model of the compressed breast. *IEEE Transactions on Medical Imaging*, 23(7):799–806, 2004.

[285] P. R. Snoeren and N. Karssemeijer. Gray-scale and geometric registration of full-field digital and film-screen mammograms. *Medical Image Analysis*, 11(2):146–156, 2007.

[286] A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging*, 32(7):1153–1190, 2013.

[287] U. Sovio, J. Li, Z. Aitken, K. Humphreys, K. Czene, S. Moss, P. Hall, V. McCormack, and I. dos Santos Silva. Comparison of fully and semi-automated area-based methods for measuring mammographic density and predicting breast cancer risk. *British Journal of Cancer*, 110(7):1908–1916, 2014.

[288] S. V. Sree, E. Y.-K. Ng, R. U. Acharya, and O. Faust. Breast imaging: A survey. *World Journal of Clinical Oncology*, 4(2):171–178, 2011.

[289] E. A. Stamatakis, A. Y. Cairns, I. W. Ricketts, C. Walker, P. E. Preece, and A. J. Thompson. A novel approach to aligning mammograms. In *Proc. International Workshop on Digital Mammography*, pages 355–364, 1994.

[290] E. A. Stamatakis, I. W. Ricketts, A. Y. Cairns, C. Walker, and P. E. Preece. Detecting abnormalities on mammograms by bilateral comparison. In *Proc. IEEE Colloquium on Digital Mammography*, pages 12/1–12/4, 1996.

[291] A. P. Stefanoyiannis, L. Costaridou, P. Sakellaropoulos, and G. Panayiotakis. A digital density equalization technique to improve visualization of breast periphery in mammography. *British Journal of Radiology*, 73(868):410–420, 2000.

[292] A. P. Stefanoyiannis, L. Costaridou, S. Skiadopoulos, and G. Panayiotakis. A digital equalisation technique improving visualisation of dense mammary gland and breast periphery in mammography. *European Journal of Radiology*, 45(2):139–149, 2003.

[293] H. Strange, E. Denton, M. Kibiro, and R. Zwiggelaar. Manifold learning for density segmentation in high risk mammograms. In *Proc. Iberian Conference on Pattern Recognition and Image Analysis*, pages 245–252, 2013.

[294] T. S. Subashini, V. Ramalingam, and S. Palanivel. Automated assessment of breast tissue density in digital mammograms. *Computer Vision and Image Understanding*, 114(1):33–43, 2010.

[295] J. Suckling, D. R. Dance, E. Moskovic, D. J. Lewis, and S. G. Blacker. Segmentation of mammograms using multiple linked self-organizing neural networks. *Medical Physics*, 22:145–152, 1995.

[296] J. Suckling, J. Parker, D. R. Dance, S. M. Astley, I. Hutt, C. R. M. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. L. Kok, P. Taylor, D. Betal, and J. Savage. The Mammographic Image Analysis Society digital mammogram database. In *Proc. International Workshop on Digital Mammography*, pages 211–221, 1994.

[297] X. J. Sun, W. Qian, and D. S. Song. Ipsilateral-mammogram computer-aided detection of breast cancer. *Computerized Medical Imaging and Graphics*, 28(3):151–158, 2004.

[298] X. J. Sun, W. Qian, D. S. Song, and R. A. Clark. Ipsilateral multi-view cad system for mass detection in digital mammography. In *IEEE Workshop on Mathematical Methods in Biomedicine and Image Analysis*, pages 19–26, 2001.

[299] L. Tabar and P. B. Dean. *Teaching atlas of mammography.* Thieme, New York, 1985.

[300] L. Tabár, T. Tot, and P. B. Dean. *Breast Cancer - The Art and Science of Early Detection with Mammography.* Georg Thieme Verlag, Stuttgart, 2005.

[301] D. Tahmoush and H. Samet. Image similarity and asymmetry to improve computer-aided detection of breast cancer. In *Proc. International Workshop on Digital Mammography*, pages 221–228, 2006.

[302] D. Tahmoush and H. Samet. Using image similarity and asymmetry to detect breast cancer. In *Proceedings of the SPIE*, volume 6144, pages 61441S.1–61441S.7, 2006.

[303] P. G. Tahoces, J. Correa, M. Soutu, L. Gómez, and J. J. Vidal. Computer-assisted diagnosis: the classification of mammographic breast parenchymal patterns. *Physics in Medicine and Biology*, 40:103–117, 1995.

[304] P. Taylor, S. Hajnal, M. H. Dilhuydy, and B. Barreau. Measuring image texture to separate "difficult" from "easy" mammograms. *British Journal of Radiology*, 67:456–463, 1994.

[305] P. Thévenaz, T. Blu, and M. Unser. Interpolation revisited [medical images application]. *IEEE Transactions on Medical Imaging*, 19(7):739–758, 2000.

[306] J. P. Thirion. Fast non-rigid matching of 3D medical images. Technical Report RR-2547, INRIA, 1995.

[307] S. Timp and N. Karssemeijer. Interval change analysis to improve computer aided detection in mammography. *Medical Image Analysis*, 10(1):82–95, 2006.

[308] S. Timp, S. van Engeland, and N. Karssemeijer. A regional registration method to find corresponding mass lesions in temporal mammogram pairs. *Medical Physics*, 32(8):2629–2638, 2005.

[309] S. Timp, S. van Engeland, and N. Karssemeijer. Temporal change analysis for characterization of mass lesions in mammography. *IEEE Transactions on Medical Imaging*, 26(7):945–953, 2007.

[310] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal. Global cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, 65(2):87–108, 2015.

[311] A. Torrent, A. Bardera, A. Oliver, J. Freixenet, I. Boada, M. Feixes, R. Martí, X. Lladó, J. Pont, E. Pérez, S. Pedraza, and J. Martí. Breast density segmentation: A comparison of clustering and region based techniques. In *Lecture Notes in Computer Science*, volume 5116, pages 9–16, 2008.

[312] G. Torres-Mejía, B. De Stavola, D. S Allen, J. J. Pérez-Gavilán, J. M. Ferreira, I. S Fentiman, and I. dos Santos Silva. Mammographic features and subsequent risk of breast cancer: A comparison of qualitative and quantitative evaluations in the guernsey prospective studies. *Cancer Epidemiology, Biomarkers and Prevention*, 14(5):1052–1059, 2005.

[313] M. Tortajada, A. Oliver, R. Martí, M. Vilagran, S. Ganau, L. Tortajada, M. Sentís, and J. Freixenet. Adapting breast density classification from digitized to full-field digital mammograms. In *Lecture Notes in Computer Science*, volume 7361, pages 561–568, 2012.

[314] C. E. Tromans, M. R. Cocker, and S. M. Brady. Quantification and normalization of x-ray mammograms. *Physics in Medicine and Biology*, 57(20):6519–6540, 2012.

[315] M. A. Turk and A. P. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[316] S. D. Tzikopoulos, M. E. Mavroforakis, H. V. Georgiou, N. Dimitropoulos, and S. Theodoridis. A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry. *Computer Methods and Programs in Biomedicine*, 102(1):47–63, 2011.

[317] M. A. Unser, A. Aldroubi, and C. R. Gerfen. Multiresolution image registration procedure using spline pyramids. In *Proceedings of the SPIE*, pages 160–170, 1993.

[318] C.M. Vachon, C.H. van Gils, T.A. Sellers, K. Ghosh, S. Pruthi, K.R. Brandt, and V.S. Pankratz. Mammographic density, breast cancer risk and risk prediction. *Breast Cancer Research*, 9(217), 2007.

[319] N. Vállez, G. Bueno, O. Déniz, J. Dorado, J. A. Seoane, A. Pazos, and C. Pastor. Breast density classification to reduce false positives in {CADe} systems. *Computer Methods and Programs in Biomedicine*, 113(2):569–584, 2014.

[320] S. van Engeland and N. Karssemeijer. Matching breast lesions in multiple mammographic views. In *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1172–1173, 2001.

[321] S. van Engeland and N. Karssemeijer. Combining two mammographic projections in a computer aided mass detection method. *Medical Physics*, 34(3):898–905, 2007.

[322] S. van Engeland, P. R. Snoeren, J. H. C. L. Hendriks, and N. Karssemeijer. A comparison of methods for mammogram registration. *IEEE Transactions on Medical Imaging*, 22(11):1436–1444, 2003.

[323] S. van Engeland, P. R. Snoeren, H. Huisman, C. Boetes, and N. Karssemeijer. Volumetric breast density estimation from full-field digital mammograms. *IEEE Transactions on Medical Imaging*, 25(3):273–282, 2006.

[324] S. van Engeland, P. R. Snoeren, N. Karssemeijer, and J. H. C. L. Hendriks. Optimized perception of lesion growth in mammograms using digital display. In *Proceedings of the SPIE, Medical Imaging*, volume 5034, pages 25–31, 2003.

[325] S. van Engeland, S. Timp, and N. Karssemeijer. Finding corresponding regions of interest in mediolateral oblique and craniocaudal mammographic views. *Medical Physics*, 33(9):3203–3212, 2006.

[326] J. Vandemeulebroucke, S. Rit, J. Kybic, P. Clarysse, and D. Sarrut. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. *Medical Physics*, 38(1):166–178, 2011.

[327] V. Vapnik. *Statistical Learning Theory*. John Wiley & Sons, New York, 1998.

[328] M. Velikova, M. Samulski, P. J. Lucas, and N. Karssemeijer. Improved mammographic cad performance using multi-view information: a bayesian network frameworks. *Physics in Medicine and Biology*, 54(5):1131–1147, 2009.

[329] N. Venugopal, B. McCurdy, A. Hnatov, and A. Dubey. A feasibility study to investigate the use of thin-plate splines to account for prostate deformation. *Physics in Medicine and Biology*, 50(12):2871, 2005.

[330] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Diffeomorphic demons using itk's finite difference solver hierarchy. In *Insight Journal – ISC/NA-MIC Workshop on Open Science at MICCAI 2007*, 2007.

[331] R. Visser, W. J. H. Veldkamp, D. Beijerinck, P. A. M. Bun, J. J. M. Deurenberg, M. W. Imhof-Tas, K. H. Schuur, M. M. Snoeren, den Heeten G. J., N. Karssemeijer, and M. J. M. Broeders. Increase in perceived case suspiciousness due to local contrast optimisation in digital screening mammography. *European Radiology*, 22(4):908–914, 2012.

[332] T. J. Vogl, W. Reith, and E. J. Rummeny. *Diagnostic and interventional radiology*. Springer, 2016.

[333] M. P. Wachowiak and T. M. Peters. High-performance medical image registration using new optimization techniques. *IEEE Transactions on Information Technology in Biomedicine*, 10(2):344–353, 2006.

[334] L. C. C. Wai and M. Brady. Curvilinear structure based mammographic registration. In *Proc. International Workshop on Computer Vision for Biomedical Image Applications*, pages 261–270, 2005.

[335] J. Wang and Y. Tan. Efficient Euclidean distance transform algorithm of binary images in arbitrary dimensions. *Pattern Recognition*, 46(1):230–242, 1999.

[336] K. Wang, Y. He, H. Qin, P. R. Fisher, and W. Zhao. Temporal registration of 2D x-ray mammogram using triangular b-splines finite element method (TBFEM). In *Proceedings of the SPIE, Medical Imaging*, pages 614436–1–614436–8, 2006.

[337] K. Wang, H. Qin, P. R. Fisher, and W. Zhao. Automatic registration of mammograms using texture-based anisotropic features. In *Proc. IEEE Inter-*

national Symposium on Biomedical Imaging: Macro to Nano, pages 864–867, 2006.

[338] X. Wang, D. Lederman, J. Tan, X. H. Wang, and B. Zheng. Computerized detection of breast tissue asymmetry depicted on bilateral mammograms: A preliminary study of breast risk stratification. *Academic Radiology*, 17(10):1234–1241, 2010.

[339] J. Wei, H. P. Chan, B. Sahiner, L. M. Hadjiiski, M. A. Helvie, M. A. Roubidoux, C. Zhou, and J. Ge. Dual system approach to computer-aided detection of breast masses on mammograms. *Medical Physics*, 33(11):4157–4168, 2006.

[340] J. Wei, H. P. Chan, B. Sahiner, C. Zhou, L. M. Hadjiiski, M. A. Roubidoux, and M. A. Helvie. Computer-aided detection of breast masses on mammograms: Dual system approach with two-view analysis. *Medical Physics*, 36(10):4451–4460, 2009.

[341] J. Wei, H. P. Chan, C. Zhou, Y. T. Wu, B. Sahiner, L. M. Hadjiiski, M. A. Roubidoux, and M. A. Helvie. Computer-aided detection of breast masses: Four-view strategy for screening mammography. *Medical Physics*, 38(4):1867–1876, 2011.

[342] K. White, K. Berbaum, and W. L. Smith. The role of previous radiographs and reports in the interpretation of current radiographs. *Investigative Radiology*, 29(3):263–265, 1994.

[343] M. B. Williams, M. J. Yaffe, A. D. A. Maidment, M. C. Martin, J. A. Seibert, and E. D. Pisano. Image quality in digital mammography: Image acquisition. *Journal of the American College of Radiology*, 3:589–608, 2006.

[344] M. Wirth, D. Nikitenko, and J. Lyon. Segmentation of the breast region in mammograms using a rule-based fuzzy reasoning algorithm. *International Journal on Graphics, Vision and Image Processing*, 5(2):45–54, 2005.

[345] M. A. Wirth, C. Choi, and A. Jennings. A nonrigid-body approach to matching mammograms. In *Proc. International Conference on Image Processing and Its Application*, pages 484–488, 1999.

[346] M. A. Wirth, J. Narhan, and D. W. S. Gray. Nonrigid mammogram registration using mutual information. In *Proceedings of the SPIE*, volume 4684, pages 562–573, 2002.

[347] I. H. Witten and E. Frank. *Data Mining Pactical Machine Learning Tools and Technique*. Morgan Kaufmann, 2nd edition, 2005.

[348] J. N. Wolfe. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer*, 37(5):2486–2492, 1976.

[349] Y.T. Wu, J. Wei, L. M. Hadjiiski, B. Sahiner, C. Zhou, J. Ge, J. Shi, Y. Zhang, and H.P. Chan. Bilateral analysis based false positive reduction for computer-aided mass detection. *Medical Physics*, 34(8):3334–3344, 2007.

[350] F. F. Yin, M. L. Giger, K. Doi, C. J. Vyborny, and R. A. Schmidt. Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images. *Medical Physics*, 18(5):955–963, 1991.

[351] F. F. Yin, M. L. Giger, K. Doi, C. J. Vyborny, and R. A. Schmidt. Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique. *Medical Physics*, 21(3):445–452, 1994.

[352] H. J. Yoon, B. Zheng, B. Sahiner, and D. P. Chakraborty. Evaluating computer-aided detection algorithms. *Medical Physics*, 34(6):2024–2038, 2007.

[353] Y. Yuan, M. L. Giger, H. Li, and C. Sennett. Correlative feature analysis on ffdm. *Medical Physics*, 35(12):5490–5500, 2008.

[354] B. Zheng. Breast cancer: computer-aided detection. In M. A. Hayat, editor, *Methods of Cancer Diagnosis, Therapy and Prognosis. Breast Carcinoma*, chapter 2. Springer, 2008.

[355] B. Zheng, Y. H. Chang, and D. Gur. Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis. *Academic Radiology*, 2(11):959–966, 1995.

[356] B. Zheng, J. K. Leader, G. S. Abrams, A. H. Lu, L. P. Wallace, G. S. Maitz, and D. Gur. Multiview-based computer-aided detection scheme for breast masses. *Medical Physics*, 33(9):3135–3143, 2006.

[357] B. Zheng, J. Tan, M.A. Ganott, D.M. Chough, and D. Gur. Matching breast masses depicted on different views a comparison of three methods. *Academic Radiology*, 16(11):1338–1347, 2009.

[358] C. Zhou, H. P. Chan, N. Petrick, M. A. Helvie, M. M. Goodsitt, B. Sahiner, and L. M. Hadjiiski. Computerized image analysis: Estimation of breast density on mammograms. *Medical Physics*, 28(6):1056–1069, 2001.

[359] B. Zitova. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

[360] W. K. Zouras, M. L. Giger, P. Lu, D. E. Wolverton, C. J. Vyborny, and K. Doi. Investigation of a temporal subtraction scheme for computerized detection of breast masses in mammograms. In *Proc. International Workshop on Digital Mammography*, pages 411–415, 1996.

[361] M. L. Zuley, B. Guo, V. J. Catullo, D. M. Chough, A. E. Kelly, A. H. Lu, G. Y. Rathfon, M. L. Spangler, J. H. Sumkin, L. P. Wallace, and A. I. Bandos. Comparison of two-dimensional synthesized mammograms versus original digital mammograms alone and in combination with tomosynthesis images. *Radiographics*, 271(3):664–671, 2014.

[362] C. S. Zuo, A. Jiang, B. L. Buff, T. G. Mahon, and T. Z. Wong. Automatic motion correction for breast mr imaging. *Radiology*, 198(3):903–906, 1996.

[363] R. Zwiggelaar. Local greylevel appearance histogram based texture segmentation. In *Proc. International Workshop on Digital Mammography*, pages 175–182, 2010.

[364] R. Zwiggelaar, L. Blot, D. Raba, and E. R. E Denton. Set-permutation-occurrence matrix based texture segmentation. In *Proc. Iberian Conference on Pattern Recognition and Image Analysis*, pages 1099–1107, 2003.

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

[366] R. Zwiggelaar and E. R. Denton. Texture based segmentation. In *Proc. International Workshop on Digital Mammography*, pages 433–440, 2006.

[367] R. Zwiggelaar, P. Planiol, J. Martí, R. Martí, L. Blot, E. R. E. Denton, and C. M. E. Rubin. Em texture segmentation of mammographic images. In *Proc. International Workshop on Digital Mammography*, pages 223–227, 2002.