Caracterización de Patrones Anormales en Mamografías

Fabián Rodrigo Narváez Espinoza

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería Mecánica y Mecatrónica
Bogotá D.C, Colombia
2016

Characterization of Breast Abnormality Patterns in Digital Mammography

Fabián Rodrigo Narváez Espinoza

Thesis submitted in partial fulfillment of the requirements for the degree of:
Doctor en Ingeniería - Ingeniería Mecánica y Mecatrónica

Advisor:
Eduardo Romero Castro, PhD

Research Field:
Applied Computing - Image Processing
Research Group:
Computer Imaging and Medical Applications Laboratory - CIM@LAB

Universidad Nacional de Colombia
Facultad de Ingeniería, Departamento de Ingeniería Mecánica y Mecatrónica
Bogotá D.C, Colombia
2016

En memoria de mi Padre, quién no logró acompañarme hasta alcanzar este sueño, pero vive siempre en mi corazón y habita en mis recuerdos...

Contents

Acknowledgments

Abstract

Computer-guided image interpretation is an extensive research area whose main purpose is to provide tools to support decision-making, for which a large number of automatic techniques have been proposed, such as, feature extraction, pattern recognition, image processing, machine learning, among others. In breast cancer, the results obtained at this area, they have led to the development of diagnostic support systems, which have even been approved by the FDA (Federal Drug Administration). However, the use of those systems is not widely extended in clinic scenarios, mainly because their performance is unstable and poorly reproducible. This is due to the high variability of the abnormal patterns associated with this neoplasia. This thesis addresses the main problem associated with the characterization and interpretation of breast masses and architectural distortion, mammographic findings directly related to the presence of breast cancer with higher variability in their form, size and location.

This document introduces the design, implementation and evaluation of strategies to characterize abnormal patterns and to improve the mammographic interpretation during the diagnosis process. The herein proposed strategies allow to characterize visual patterns of these lesions and the relationship between them to infer their clinical significance according to BI-RADS (Breast Imaging Reporting and Data System), a radiologic tool used for mammographic evaluation and reporting. The obtained results outperform some obtained by methods reported in the literature both tasks classification and interpretation of masses and architectural distortion, respectively, demonstrating the effectiveness and versatility of the proposed strategies.

Keywords: Breast Cancer, BI-RADS, Breast Mass, Architectural Distortion, Patterns Recognition, Multiresolution Image Analysis, Content Based Image Retrieval.

Resumen

La interpretación de imágenes guiada por computador es una área extensa de investigación cuyo objetivo principal es proporcionar herramientas para el soporte a la toma de decisiones, para lo cual se han usado un gran número de técnicas de extracción de características, reconocimiento de patrones, procesamiento de imágenes, aprendizaje de máquina, entre otras. En el cáncer de mama, los resultados obtenidos en esta área han dado lugar al desarrollo de sistemas de apoyo al diagnóstico que han sido incluso aprobados por la FDA (Federal Drug Administration). Sin embargo, el uso de estos sistemas no es ampliamente extendido, debido principalmente, a que su desempeño resulta inestable y poco reproducible frente a la alta variabilidad de los patrones anormales asociados a esta neoplasia. Esta tesis trata el principal problema asociado a la caracterización y análisis de masas y distorsión de la arquitectura debido a que son hallazgos directamente relacionados con la presencia de cáncer y que usualmente presentan mayor variabilidad en su forma, tamaño y localización, lo que altera los resultados diagnósticos.

Este documento introduce el diseño, implementación y evaluación de un conjunto de estrategias para caracterizar patrones anormales relacionados con este tipo de hallazgos para mejorar la interpretación y soportar el diagnóstico mediante la imagen mamaria. Los modelos aquí propuestos permiten caracterizar patrones visuales y la relación entre estos para inferir su significado clínico según el estándar BI-RADS (Breast Imaging Reporting and Data System) usado para la evaluación y reporte mamográfico.
Los resultados obtenidos han demostrado mejorar a los resultados obtenidos por los métodos reportados en la literatura en tareas como clasificación e interpretación de masas y distorsión arquitectural, demostrando la efectividad y versatilidad de las estrategia propuestas.

Palabras Claves: Cáncer de mama, Nódulos de Mama, Distorsión Arquitectural, BI-RADS, Reconocimiento de Patrones, Análisis Multiresolución, Recuperación de Imágenes basado en Contenido

List of Tables

List of Figures

1. Introduction

Despite the considerable efforts performed by health systems to reduce the burden of disease caused by breast cancer, this is still one of the most common diagnosed cancer in women worldwide [5]. The American Cancer Society estimates that 1 in 8 women (about 12%) will develop invasive breast cancer during their lifetime. In 2016 for the United States, 246,660 new cases of women invasive breast cancer are expected to be diagnosed, along with 61,0000 new cases of non-invasive (in situ) breast cancer [5]. Supportive evidence and different studies have shown a higher survival rate when this kind of cancer is diagnosed in early stages, increasing from 27% to 98% the 5-year survival rate [4, 5] .

Among the different imaging modalities used for breast cancer detection, X-ray mammography is the examination for both routine screening and symptomatic diagnosis [108]. Mammography screening has demonstrated to be effective in reducing breast cancer mortality, in percentages that vary from 30-70%, yet the screening sensitivity has been reported to be between 85 and 90% [14, 115]. Despite this evidence, mammography is still considered a method with limitations, since it only achieves an average sensitivity of 73% to 88% and a specificity between 83% and 96% [110, 20]. Different causes may influence the variability of the observed sensitivity, among others the quality of the image, the radiologist's expertise, the detail or meticulousness of the evaluation and the intrinsic difficulty for the differentiation of visual patterns in the image [6, 94]. This variability produces a high index of false positives and false negatives rates.

Several strategies have aimed to increase the diagnosis accuracy and reduce the interpretation variability. The American College of Radiology (ACR) developed the Breast Imaging-Reporting and Database System (BI-RADS), a standardization of the mammography findings by a radiologic semiology that aims to reduce the radiologist subjectivity [108]. On the other hand, while the BI-RADS lexicon is an important step ahead in reducing the variation of the mammography report, automated screening systems known as "Computer Aided Diagnosis" systems (CAD), are being developed to assist the diagnostic process [36, 98, 124]. Nowadays, CAD systems are considered as a potential support for diagnosis and screening, reducing the workload of radiologists and improving the detection of breast abnormalities [16, 117, 43]. However, evaluation of commercial CADs show a relatively low sensitivity, specially for detecting some radiological findings such as masses and architectural distortions [22], reason why the radiologist's confidence in these systems is low [87, 132, 61], a reluctance worsened by the fact that most CADs are merely designed to yield a cancer/no-cancer opinion. Overall, this is useless in terms of interpretation since no clue is given about the described radiologic signs. Provided that the radiologist workflow is mainly based upon the analysis of such signs, any diagnosis support should not only detect malignant lesions but it also should help the radiologist to improve the reading of such radiologic signs, independently of the final diagnosis. This is why the clinical use of such kind of tools has still remained very limited, among others because evidence is not conclusive and sometimes even contradictory [36,

31, 32, 58]. Researchers agree that current commercial CAD systems still need improvement for they are introduced in actual clinical and screening applications [87, 132]. Therefore, the present investigation was addressed to provide new methods of diagnosis processing of mammography, particularly for breast masses and architectural distortions of breast tissues.

1.1. Breast cancer screening based on Mammography

Mammography is a specific type of breast imaging obtained by the projection onto an image sensor of low-dose X-rays that go through several anatomical planes of the breast. Two planar images, craniocaudal (CC) and mediolateral-oblique (MLO) projections, are routinely obtained for diagnostic purposes, as illustrated in figure 1-1. In this image, internal breast structures are observed as differences between radiopaque and radiolucent zones, with radiopaque corresponding to dense tissue i.e. epithelial tissue and stroma, and radiolucent regions corresponding to adipose tissue. Normal breasts have a wide variation in the distribution of dense and adipose regions. Radiologists are trained for recognizing alterations out of the expected variability, suspicious signs such as calcifications, masses or another type of distortions, by searching specific changes on the distribution of the observed intensity values. However, diagnosis is a very difficult process since normal dense tissue and abnormal structures exhibit similar visual characteristics.
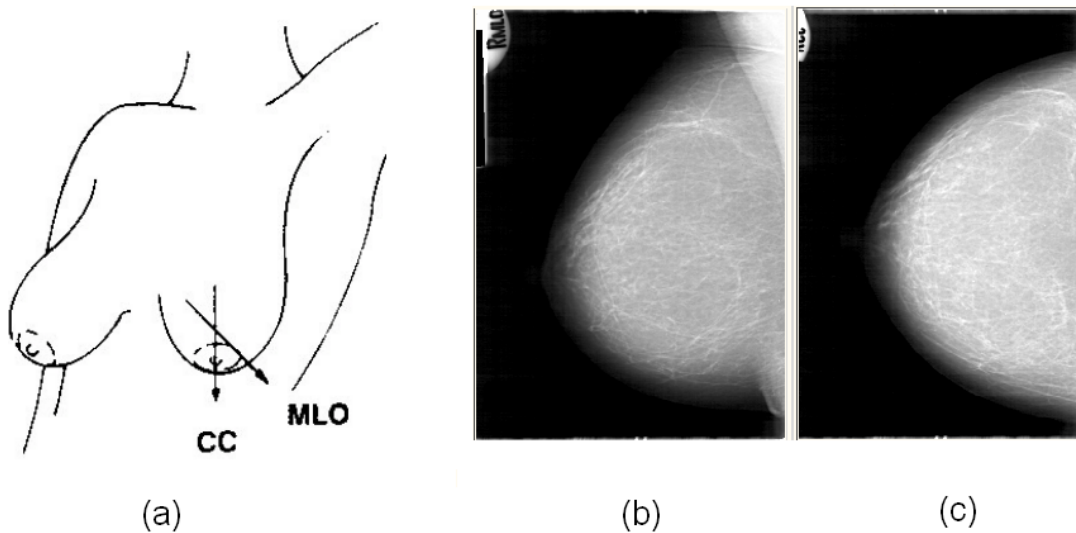
Figure 1-1.: X-ray direction to each projection, (b) Mediolateral Oblique (MLO) projection, (c) Cranio-caudal Projection [51].

1.1.1. Mammography interpretation process

The interpretation process, routinely followed at interpreting mammography, can be summarized as follows. Once the mammography is obtained, a visual inspection is carried out to establish if the capture quality is appropriate for diagnosis purposes. Images, defined as acceptable, are

subjected to visual inspection by the radiologist, who generates a report that should follows the BI-RADS system [108]. This report informs about the overall breast mass composition, description of all important findings, a final assessment and management recommendations. The final assessment consists in assigning a BI-RADS category value, between 0 and 6, that should agree with both the interpreted findings and the management recommendations, as shown in Table 1-1. One important task when reading mammography is to detect and describe relevant findings associated to cancer disease, i.e. bilateral asymmetries, calcifications, solid masses and architectural distortions. These findings exhibit a high variability, several shapes, sizes, locations and levels of radiographic densities which closely resemble the normal tissue, making the mammography interpretation an actual radiology challenge. Among them, architectural distortions and description of breast masses are the most challenging tasks.

Table 1-1.: Agreement Between BI-RADS Assessment Categories, Important Findings and Management Recommendations

Assessment	Important Findings	Management
Category 0: Incomplete - Need Additional Imaging Evaluation and Prior Mammograms for Comparison	There is a finding for which additional imaging evaluation is needed	Recall for additional imaging and comparison with prior examination(s)
Category 1: Negative	No findings reported	Routine mammography screening
Category 2: Benign	Benign findings such as rounded and circumscribed masses,were reported	Routine mammography screening
Category 3: Probably Benign	Findings having a $\leq 2\%$ likelihood of malignancy. Some typical probably benign findings are noncalcified circumscribed masses, focal asymmetries, and solitary group of punctate calcifications.	Short-interval (6-month) follow-up or continued surveillance mammography
Category 4: Suspicious Category 4A: Low suspicion for malignancy Category 4B: Moderate suspicion for malignancy Category 4C: High suspicion for malignancy	Findings that not have all features of malignancy but are suficiently suspicious such as heterogeneus microcalcifications, masses with microlobulated or irregular marging and high density, architectural distortions, among others.	Tissue diagnosis
Category 5: Highly Suggestive of Malignancy	Findings with a a very high probability ($\geq 95\%$) of malignancy. Reserved to spiculated masses and pleomorphic microcalcifications.	Tissue diagnosis
Category 6: Known Biopsy-Proven Malignancy	there are no mammographic abnormalities other than the known cancer that might need additional evaluation.	Surgical excision when clinically appropriate

1.2. Computer-Aided Diagnosis Systems

Image processing and computer vision areas have developed computational systems in different domains. In particular, in radiology there is a modern trend towards developing systems that automatically locate suspicion regions, a research area currently-known as Computer Aided Diagnosis (CAD) systems. This type of systems aims to either guide radiologist in locating and identifying possible abnormalities, known as computer aided detection (CADe) systems [50, 22], or to assist in the diagnosis process of abnormalities previously detected, known as computer-aided diagnostic (CADx) models [42]. The use of these systems reduce the subjectivity, decreasing common errors and increasing interpretation reproducibility [86, 69, 91]. Moreover, new radiologist training has been improved when CAD systems are used as teaching assistants. However, large variability among radiological findings also cause a variable detection and diagnosis performance. Recent evaluations of commercial CADs [50, 22], show that it reaches a high sensitivity in detecting calcifications (92%), but a lower figure, 75%, when detecting masses [125, 12, 55, 102] and 50% for architecture distortions [9]. Additionally, a high rate of false positives has been also reported, leading women without cancer to undergo further clinical evaluation or directly to unnecessary breast biopsies [109, 54, 73, 39].

Most of CAD systems are black boxes that may detect suspicious regions but the user has no clue about the criteria upon which such decision is based on, i.e., they are only used to classify regions as either malignant or benign. This drawback has motivated the development of interactive CAD systems (iCAD) that aid to interpret suspicious regions rather than to assist with the initial automatic detection of these regions [48, 125]. Different initiatives have been proposed, among them, content-based mammogram retrieval systems that pursuit to visually support the diagnosis process by retrieving suspicious regions of interest with similar visual content from a mammography database [52, 118, 127, 59]. This strategy aims to exploit the high-level information associated to each retrieved region and to help the clinician with the evidence that might support a particular case decision. Nevertheless, these systems ignore the importance of providing information for the final diagnosis. [52, 127, 112]. From this discussion, systems that take advantage of both pattern recognition and content-based retrieval based on image analysis could be useful in the development of new CAD systems.

On the other hand, recently some researchers proposed a new type of CAD system, which beyond a final decision about malignancy of a radiological finding, aids radiologists to perform intermediate steps in the interpretation process. Specifically, the idea of automatically describing morphological features of detected radiological findings, is actually considered relevant to reduce errors in breast cancer image analysis [3, 121, 112].

1.3. Research Challenges

As aforementioned, the idea of providing systems that aid radiologist to perform the radiological BI-RADS report, could be useful to improve the breast cancer diagnosis. Particularly, detection of architecture distortions and description of masses are not only the most challenging tasks but the most relevant for defining a final BI-RADS category. Figure 1-2 shows some examples from architectural distortions (1-2-a), masses (1-2-b) and segments of normal tissue. As observed, the

first challenge is to decide whether a RoI contains or not a lesion since normal and abnormal tissues exhibit a huge visual variability such as shape, margin and texture, furthermore, it is very difficult to define accurately the edges of structures, reason by which segmentation based methods often fail at detecting this kind of lesions.
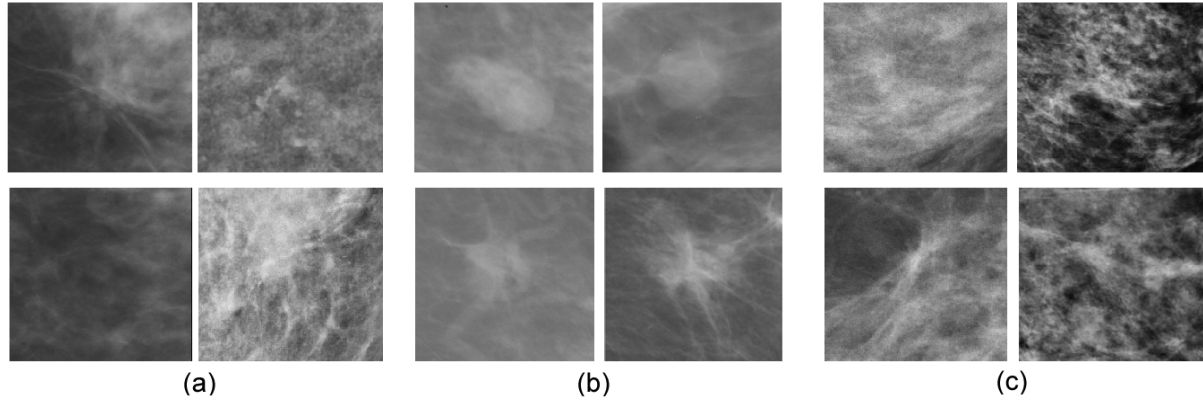
Figure 1-2.: Regions of interest containing architectural distortions (a), masses (b) and normal parenchymal tissue (c).

Detection of architectural distortions, a strong evidence of breast cancer, is crucial since the sole detection is sufficient to assign the mammography categorization. In contrast, for a breast mass, it is necessary to describe the mass features, namely shape, margin and density and with these elements it is then possible to assign the category. As observed in Figure 1-3, after the BI-RADS descriptor, the shape can be oval, round and irregular; the margin circumscribed, obscured, microlobulated, spiculated or indistinct; and the density, i.e. the attenuation of the mass with respect to the glandular tissue, can be described as High density, Equal density, Low density or Fat-containing. From a pattern recognition standpoint, this is a multi-labeling and multi-class problem, a task far more challenging than the simple binary classification, i.e. benign or malignant.

In both cases, lesion detection and mass feature description, a main challenge is to obtain a region characterization that allows to discriminate the content of a region of interest.

1.4. Research Questions

This thesis is aimed to strengthen the breast cancer diagnostic process, for which automatic recognition strategies have been proposed and used to describe lesions in terms of low-level descriptors. In order to address the main research problem, the following research question was purposed: How to reduce the diagnostic subjectivity produced by the high variability of mammographic findings?
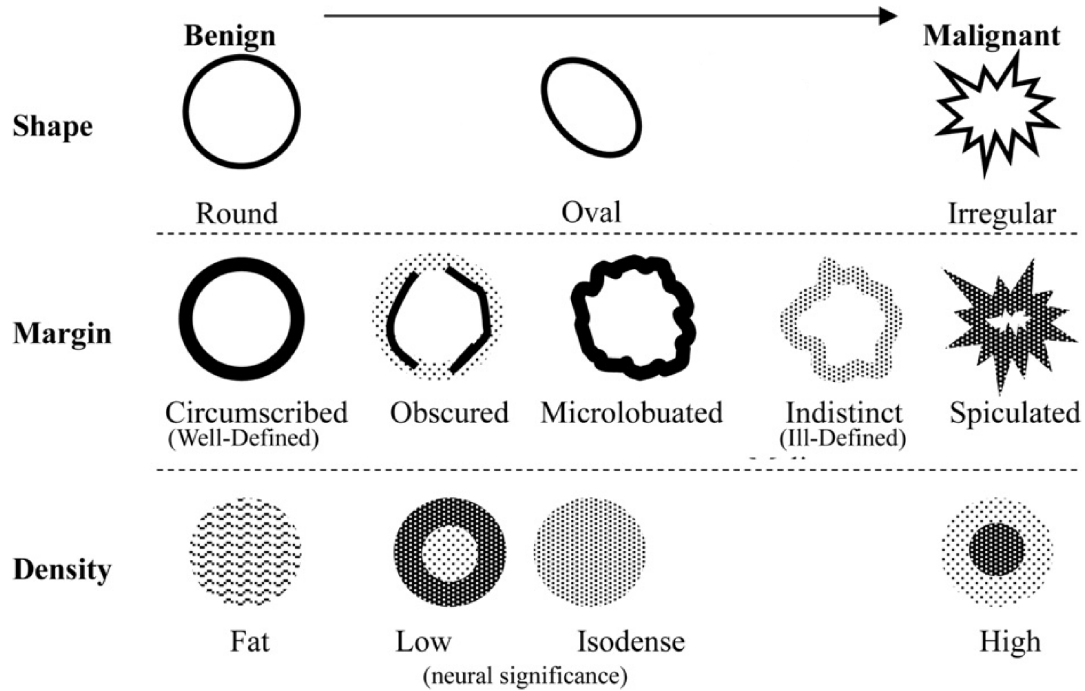
Figure 1-3.: Feature descriptors defined by the BI-RADS system for labeling a mass.

1.5. Contributions and thesis outline

This work presents several contributions for breast abnormality analysis, focused on radiologic interpretation. The different features are handcrafted but sufficiently general as to be part of a clinic application, with minimal intervention of the radiologist and with no previous segmentation process at all. Main contributions are the original description of different biological patterns by the use of two important multi-resolution image decomposition and linear saliency maps for mammographic sings analysis: such as Zernike moments, Curvelets transforms and oriented linearity patterns. The following are the main applications covered by the methods proposed in this thesis, together with the corresponding references to published works.

Characterization of Breast Masses

An accuracy classification of breast masses between benign and malignant is still a difficult task, a task usually dependent on a previous accurate segmentation of the lesion. The proposed characterization was evaluated in two contexts: Classification and Content based image retrieval for masses analysis. As a result of this thesis, the following papers have been published, which are referred in the text:

- Fabián Narváez, Gloria Díaz, Cesar Poveda and Eduardo Romero:*"An Automatic BI-RADS Interpretation of Regions of Interest with Breast Masses by Fusing Multiresolution Descriptors"*. Journal of Expert Systems With Applications. In press, 2017.

- Fabián Narváez and Eduardo Romero: *"Breast mass classification using orthogonal moments"*. Breast Imaging, Lecture Notes in Computer Science, Volume 7361, pages 64-71, 2012. ISBN: 978-3-642-31270-0 and ISSN 0302-9743

- Fabián Narváez, Gloria Díaz, Francisco Gómez and Eduardo Romero:*"A content-based retrieval of mammographic masses using the curvelet descriptor"*. Medical Imaging, Proc. SPIE Vol 8315, pages 83150A, 2012. ISBN: 9780819489647 and ISSN 1605-7422.

- Fabián Narváez, Gloria Díaz and Eduardo Romero: *"Multi-view Information Fusion for Automatic BI-RADS Description of Mammographic Masses"*.Medical Imaging 2011, Proc. SPIE, Vol 7963, 79630A, 2011. ISBN: 9780819485052 and ISSN 1605-7422

- Fabián Narváez, Gloria Díaz, Francisco Gómez and Eduardo Romero: *"Multiscale Breast Mass Analysis using the Curvelet Transform"*. MICCAI 2011 - Workshop on Breast Image Analysis. ISBN: 978-87-981270-9-3 and ISSN: 0107-8283

- Fabián Narváez, Andrea Rueda and Eduardo Romero: *"Breast Masses Classification using a Sparse Representation."*. Proceedings of MIAD 2011 - 2nd International Workshop on Medical Image Analysis and Description for Diagnosis Systems, ISBN: 978-989-8425-38-6, pages 26-33, 2011.

- Fabián Narváez, Gloria Díaz and Eduardo Romero: *" Automatic BI-RADS description of Mammographic Masses"*. 10th International Workshop on Digital Mammography. Lecture Notes in Computer Science, Volumen 6136, pages 673-68, 2010. ISBN 978-3-642-13666-5 and ISSN 0302-9743

1.5.1. Characterization of Architectural Distortion

Unlike classic approaches that characterize linear structures by transforming the image to specialized spaces, a powerful technique was proposed to search of the linear saliency of a stellate pattern in the spatial domain, aiming to strength the classification process. The main contribution consists in detecting and characterizing the Linear Saliency in the spatial domain. The performance of this strategy was published in:

- Fabián Narváez, Jorge Alvarez, Juan D. GarciaArteaga, Jonathan Tarquino and Eduardo Romero: *"Characterizing Architectural Distortion in mammograms by Linear Saliency"*. Accepted for the Journal of Medical Systems, 2016.

- Jorge Alvarez, Fabián Narváez, Cesar Poveda and Eduardo Romero:*"Characterization of architectural distortion on mammograms using a Linear Energy Detector"*. SIPAIM2013, In Proccedings on IX International Seminar on Medical Information Processing and Analysis, 2013.

- Fabián Narváez and Eduardo Romero:*"Automatic Detection of Architectural Distortion in Mammograms using Sparse Overcomplete Dictionaries of a Curvelet Descriptor"*. MICCAI 2013 - Workshop on Breast Image Analysis, In Proccedings Vol 1, pp 106-113, 2013. ISSN: 0107-8283 and ISBN: 978-87-996443-0-8

1.5.2. Thesis Outline

The remaining chapters of the thesis are organized as follows:

- **Chapter 2:** *Characterization of Breast Masses*: This chapter introduces a new automatic BI-RADS characterization of breast masses contained in a Regions of Interest (RoI). A detailed description of breast masses is achiveved in terms of their shape, margin, and density. The proposed approach learns relevant radiological characteristics from several multiscale decomposition of the visual information, the Zernike polynomials and Curvelet bases, which are optimally fused by a Multiple Kernel Learning (MKL) strategy. The method first assesses the presence of masses in the RoI by training a conventional support vector machine (SVM) classifier. Once a RoI is identified as mass-positive, it feeds a bank of SVM binary classifiers that selects the type of shape from the BI-RADS terms. The mass margin and density are then set by combining the BI-RADS labels of the five most similar shapes in the database. Accuracy and precision of the proposed approach is evaluated using two mammogram databases, a full-field digital mammography and a digitized film-screen version of mammographies.

- **Chapter 3:** *Characterization of Architectural Distortion*: This chapter presents a novel characterization of architectural distortion by representing the linear saliency in mammography Regions of Interest (ROI) as a graph composed of nodes, corresponding to locations along the ROI boundary, and edges with a weight proportional to the line intensity integral along the path connecting any pair of nodes. A set of eigenvectors from the adjacency matrix is then used to extract discriminant coefficients that represent those nodes with higher salient lines. A dimensionality reduction is further accomplished by selecting the pair of nodes with major contribution for each of the computed eigenvectors. The set of main salient lines is then assembled as a feature vector that inputs a conventional Support Vector Machine (SVM). Accuracy and precision of the proposed approach is evaluated also using two mammograms database, additionally an exhaustive comparison with other existing approaches reported in the state-of-art demonstrates its benefits.

- **Chapter 4:** *Conclusions and Perspectives*: This final chapter presents the main conclusions of the proposed work, highlighting the main contributions achieved and its impact in the research area. In addition, it depicts some of the future research directions and perspectives promoted by this thesis.

1.5.3. Contributing to related work

In the process of performing research the author has also contributed significantly to the following publication.

- Lina Pedraza, Carlos Vargas, Fabián Narváez, O. Durán, E. Muñoz and Eduardo Romero: *"An open access thyroid ultrasound image database"*. Tenth International Symposium on Medical Information Processing and Analysis, 92870W, 2015.

1.6. Research Funding

2. Characterization of breast masses

*A correct mammography evaluation demands great expertise and rigorous interpretation of some mammographic findings. Currently, the Breast Imaging Reporting and Data System (BI-RADS) is accepted as a standardized method to communicate mammographic findings. This report assigns a detailed description to breast lesions in terms of their shape, margin, and density, a process highly dependent on the observer experience. This chapter presents a new automatic BI-RADS characterization of breast masses contained in a Regions of Interest (RoI). Instead of attempting a mass segmentation, the automatic description is based on a multiresolution characterization of a region on interest (RoI). Likewise, the description problem is recast by a conventional classification which is enriched with the knowledge from a mammographic database, thereby improving the final description result. The contributions of this work is twofold: on the one hand, a novel representation of breast masses that fuses two complementary multiresolution bases, the Zernike pyramidal analysis, a complex orthogonal base, and a Curvelet transformation, a multiscale decomposition suitable for capturing complex mixes of curved lines. On the other hand, a method that integrates a conventional shape classification and a retrieval strategy which is used to assign the mass margin and density, completing thereby a description fully compatible with the BI-RADS lexicon. The complete content of this chapter has been published as a research article in the **Journal of Expert Systems With Applications** (see [84]).*

2.1. Introduction

Breast cancer is the most frequently diagnosed cancer in women and the second cause of cancer women deaths [5], a cancer usually treatable if an early diagnosis is possible. Currently, X-ray mammography is the only examination approved for both routine screening and symptomatic diagnosis [106], which has demonstrated to be effective in reducing breast cancer mortality, in percentages that vary from 30-70%, yet the screening sensitivity has been reported to be between 85 and 90% [14, 115]. Different radiologic signs such as masses, calcifications, bilateral asymmetry and architectural distortion may be detected and followed up [108]. Usually, mammographic findings exhibit a high variability, several shapes, sizes, locations and levels of radiographic densities which closely resemble the normal tissue, making the mammography interpretation an actual radiology challenge. An incorrect interpretation may lead women without cancer to undergo further clinical evaluation or directly to an unnecessary breast biopsy [73, 39].
Breast masses and micro-calcifications are important cancer markers, but masses may be much more difficult to detect or interpret than micro-calcifications [130], among others because of their poor contrast and blurred features. In fact, the complex superposition of breast tissues in mammograms misleads the radiologic findings that define a degree of malignancy (benign or malignant). Several strategies have aimed to increase the diagnosis accuracy and reduce the variable interpretation. Among them, the American College of Radiology (ACR) developed the Breast
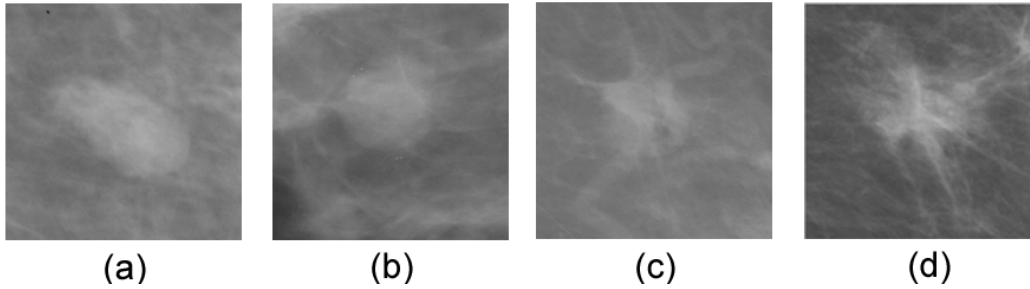
Figure 2-1.: Different mass shapes and margins described according to the fifth BI-RADS version. a) Oval-Circumscribed b) Round-Circumscribed c) Irregular-Indistinct d)Irregular-Spiculated

Imaging-Reporting and Database System (BI-RADS), which standardizes the lesion description by a radiologic semiology that categorizes specific radiologic features. These descriptions support the physicians BI-RADS decision, between 0 and 6, a value that determines the probability of malignancy and the final recommendation [108]. Specifically, breast masses are described by their shape, margin and density. In the fifth BI-RADS edition, the shape class is described as *oval*, *round* and *irregular* and the margin as *circumscribed, obscured, microlobulated, spiculated* or *indistinct*. Finally, the density class, the attenuation of the mass with respect to the glandular tissue, can be described as *High density, Equal density, Low density* or *Fat-containing* . Figure **2-1** illustrates different mass shapes and margins described with the BI-RADS lexicon. Once the radiologist describes the mass in terms of the BI-RADS lexicon, a cancer probability of the lesion is set (0-6 scale). As illustration, a mass with round or oval shape, circumscribed margin and low density has a high probability of being benign, while a mass with irregular shape, spiculated margin and high-density is cancer suspicious [39]. This value also defines the particular patient management, i.e., low values imply the patient is followed up while high values usually lead to the possibility of biopsy and pathological study.

The BI-RADS system has remarkably improved the agreement among radiologists and decreased the intra reader variability [119]. However, large variability at describing masses is still reported [67, 73, 18]. Therefore, automatic description of these features is in consequence a potential benefit for the process of supporting the final radiologist decision. Computer Aided Diagnosis (CAD) systems have been proposed for improving the reading workflow and the radiologist workload. Most of them have been developed for detecting and classifying benign/malign breast masses as second readers [36, 88, 87]. However, assigning BI-RADS descriptors is a much more complex multi-class task that has been poorly explored in the literature [105, 121, 103, 53]. As far as we know, automatic BI-RADS descriptions in terms of shape, margin, and density, has been only proposed by [64, 63].

This chapter presents a novel method that automatically describes non-calcified mammography masses following the lexicon introduced by the BI-RADS fifth edition. Instead of attempting a mass segmentation, the automatic description is based on a multiresolution characterization of a region on interest (RoI). Our previous works have shown the potential of multiresolution features as mass descriptors [83, 82] and this chapter extends these works by generalizing the representa-

tion and learning the relevant features for each descriptor. Likewise, the description problem is recast by a conventional classification which is enriched with the knowledge from the database, thereby improving the final description result.

The main contributions of this work are summarized below:

1. A novel representation of breast masses that fuses two complementary multiresolution bases, the Zernike pyramidal analysis, a complex orthogonal base, and a Curvelet transformation, a multiscale decomposition suitable for capturing complex mixes of curved lines.

2. A method that integrates a conventional shape classification and a retrieval strategy which is used to assign the mass margin and density, completing thereby a description fully compatible with the BI-RADS lexicon.

The rest of this chapter is organized as follows: in the next section a literature review of previews works is presented. In Section 2.3, the proposed method is described. Section 2.4 presents the experimental results, and the last sections are devoted to discussions, conclusions and future works.

2.2. Related works

The proposed approach deals with two breast different analyses: deciding whether a RoI contains a mass or not (mass/non mass classification) and assigning a description to the mass, according to the BI-RADS lexicon. This section presents a summary of the proposed methods related to each of them.

2.2.1. Masses against normal tissue

Several methods have aimed to distinguish control tissue from masses in mammography, either as main task [101] or as a first step of an automatic diagnosis [75, 123, 70]. Since shape and texture are relevant at characterizing masses, several feature extraction approaches have been introduced. Usually, Zernike moments capture shape features from previously delineated masses [107, 127, 128, 116]. These descriptors have shown to be more discriminant than others based on geometrical and margin features [116, 107]. Moreover, Zernike moments and sharpness measures, obtained from a segmented mass, have been also used by a Content Based-Image Retrieval (CBIR) system [128, 127]. On the other hand, texture patterns have been represented by grey-level statistics [101], local features [71, 1] and frequency-based approaches [75, 30, 95, 29, 25]. [30] evaluated the performance of two multiresolution representations, curvelets and wavelets, for distinguishing between benign, malignant and normal tissues. Best results were obtained by curvelets, obtaining an accuracy of 94.07%, while the highest performance achieved by wavelet functions was 90.05%. In 2012, the same authors evaluated the effect of introducing a feature selection step for reducing the number of coefficients [29]. Additionally, the classification task was divided into two stages, the first classifies normal versus abnormal tissues, and the second distinguishes whether the abnormal tissue is benign or malignant. The maximum accuracy obtained for the classification of normal versus abnormal tissue was 94.79% with the wavelets and 95.67% with the curvelets. Recently [1] proposed a new descriptor, the uniform local directional pattern (ULDP), which is

intensity invariant since it describes each neighborhood with edge information. An A_z of 0.91 and 0.99 were reported when this descriptor was used to distinguish between normal and abnormal tissues.

Fusion has been introduced as a promising strategy that strengths out the best properties of each descriptor. [101] evaluated different features to be used by a system that classifies RoIs as masses or normal tissues, using a SVM with RBF kernel model. After a parameter tuning, the combination of LBP and statistical features reported the best performance with an accuracy of 99.1%. [19] combined LBP operators with curvelet coefficients to discriminate normal tissue and malignant masses, and malignant from benign masses. After evaluating different classifiers, the best performance was obtained by a SVM with polynomial kernel (Accuracy of 91 and 100% for DDSM and BCDR-FMR datasets, respectively). [70] fuses texture and shape descriptors for establishing if a RoI is normal tissue, benign or malignant lesion. Zernike moments are calculated from a series of wavelet components and then used by a kernel-oriented neural network model. The best results (Accuracy of 94.1%) were obtained using multiresolution Symlet 8.

2.2.2. Automatic description of BI-RADS mass features

As aforementioned, few strategies have been proposed to support the process of evaluating a mammography mass using the lexicon defined by the BI-RADS system. The morphology of breast masses is of course the base of the diagnosis, reason by which some approaches have been proposed to characterize some mass shape, margin and density related features that are then used for benign/malign classification [41, 118, 114, 10], for BI-RADS categorization (0 to 6) [76] and for information retrieval [128] tasks. Moreover, some authors have proposed to simply use the radiologist's BI-RADS mass description for the same purposes [74, 112, 88], but they have not attempted shape, margin or density description tasks.

Although mass density has demonstrated to be an important predictor of breast cancer [129, 33, 130], only the method proposed by Kisilev et al. [64, 63] tried to classify it. Some works have addressed the problem of classifying breast density [28, 60, 3, 89, 71], a feature also relevant for the mammography report but different from the mass density, which is a description of the mass density relative to normal fibroglandular breast tissue. Shape and margin descriptions have been carried out using geometric and boundary-based features [113, 121, 103, 53], which require a previous segmentation step. Overall, methods describing the margin are limited to distinguish between spiculated and circumscribed masses [103, 53], a task simpler than the whole description including the other three categories, namely obscured, microlobulated and indistinct.

2.3. Materials and Methods

2.3.1. Method overview

The method pipeline is illustrated in figure 2-2, starting by a preprocessing phase which is different for each of the multiresolution bases and followed by their characterization.

Once the representation process is completed, a Multiple Kernel Learning (MKL) technique combines and learn the most relevant features from each of the two previous feature spaces. Using

this image representation as input, the RoI is firstly classified as breast mass or control tissue (non-mass) by a conventional support vector machine (SVM) classifier. If a RoI is mass-positive, a multiclass bank of classifiers assigns the RoI to a single shape term, namely *Oval, Round* or *Irregular*. The RoI margin and density terms are then set by a different strategy that starts by recovering the most similar mass shapes from a database using the histogram intersection as the similarity metric. A weight is assigned to each of these masses, a weight proportional to the distance of each shape to the hyperplane defined by the classifier that assigned the mass shape. The margin and density terms are finally chosen by a simple voting strategy with the set of five most similar (weighted) shapes.
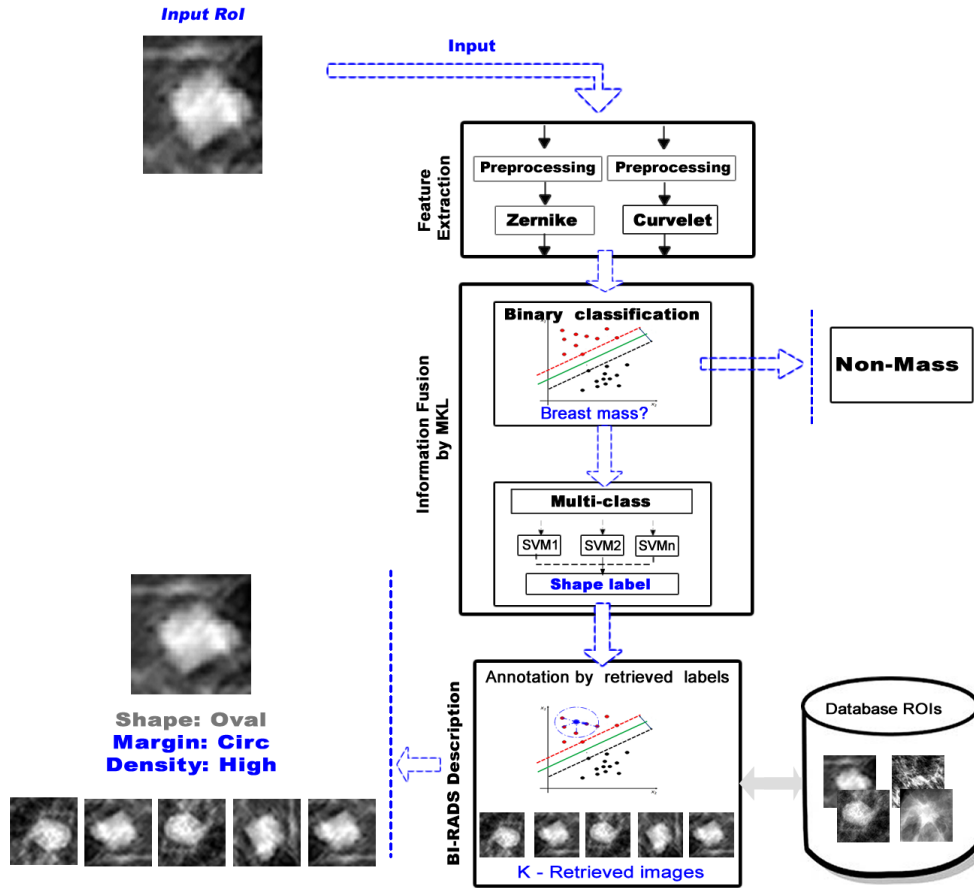
Figure 2-2.: Proposed pipeline

2.3.2. Extraction of characteristics

Recent evidence suggests radiologists integrate visual information from different scales and frequencies [13]. For the particular case of breast masses, radiologic analysis integrates information from edges and local spatial patterns at different scales and orientations. Inspired by these studies, the present investigation has explored two complementary representations that describe breast lesions from global (coarse shape) and local (texture) standpoints: a Zernike pyramidal analysis to

capture the essential shape information and a Curvelet representation that describes the local information. Both representation bases require different pre-processing schemes.

A region-based shape analysis:

Mass shape has been reported as a main radiological sign [66], while Zernike moments have been classically used as good region-based shape descriptors [107, 90, 46]. Provided that Zernike moments are insensitive to mass segmentation [127], this transformation results suitable for representing masses with blurry boundaries. In this case, the pre-processing step consisted in enhancing the RoI by using an adaptive histogram equalization [93], followed by a reduction of the dynamic range, i.e., a pixel-depth conversion strategy [2] that decreased the mammography range from 16 to 8-bits. Overall, the distribution of intensities is highly variable along a mass boundary and is usually a real challenge to determine where an edge is actually located. A reduction of the dynamic range produces the effect of homogenizing the distribution of intensities while real edges are conserved. In consequence, an additional further dimensionality reduction was performed by mapping the $[0, 255]$ usual intensity range to an interval with $[0, 12]$ gray levels. The main purpose of this nonlinear transformation was to obtain coarse mass representations that were more easily captured by the Zernike moments, as illustrated in the upper panels of table **2-1**, where a RoI with $[0, 255]$ grey levels is sequentially transformed to $[0, 128]$, $[0, 64]$ and $[0, 12]$ grey levels. This transformation preserves the Zernike representation, as demonstrated in table **2-1**, where the variances of the different p orders are shown for the sequence of images in the upper row. While the mean is approximately the same (not shown in table **2-1**), the variance (σ) of the sets of coefficients for the different p orders (10 to 50), changes very little when the number of gray levels of the RoI is adjusted from 256 to 12. As a conclusion, this drastic dimensionality reduction captures the low shape frequencies and decreases the computational cost of the Zernike basis.
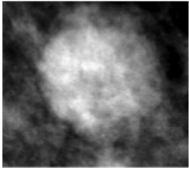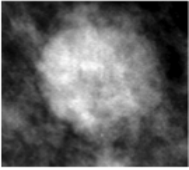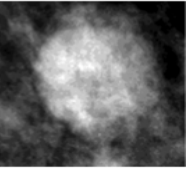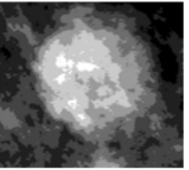
p-order	[0,255] σ	[0,128] σ	[0,64] σ	[0,12] σ
10	2.357	2.357	2.354	2.344
20	1.261	1.260	1.261	1.273
30	0.897	0.897	0.897	0.902
40	0.728	0.729	0.729	0.738
50	48.042	48.718	49.914	50.257

Table 2-1.: Grey-level reduction analysis

Once the RoI was preprocessed, a region-based shape description is reached by computing the Zernike moments as shown in [46], selecting a subset of coefficients from the Zernike pyramidal analysis. The complex 2D Zernike moments of order p and q repetitions are defined in the unitary

circle as

$$Z_{p,q} = \frac{p+1}{\pi} \int\limits_{0}^{2\pi} \int\limits_{0}^{1} [V_{p,q}(r,\theta)]^* f(rcos\theta, rsin\theta) r \, dr \, d\theta \tag{2-1}$$

where $p = 0, 1, 2, 3, ..., \infty$ and q is a positive or negative integer depending on $p - |q| = even$ and $|q| \leq p$ [126, 46]. In discrete terms, a $N \times N$ digital image is projected to the Zernike space as

$$Z_{pq} = \frac{p+1}{\lambda} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} V_{pq}^*(x,y) f(x,y) \tag{2-2}$$

where $\lambda = \delta A/\pi$ is a normalizing constant and δA is the image area when projected onto the unit circle of Zernike polynomials.

In this work, to achieve translation and scale invariance, a normalization technique is applied as described in [26] by centering and scaling the reference frame (x,y) to $(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y})$, being $(\bar{x}, \bar{y})$ the $f(x,y)$ centroid and $a = \sqrt{\beta/m_{00}}$ the scaling factor, with β the number of pixels in the mass obtained by binarizing the RoI and m_{00} the geometrical moment of order zero. A particular RoI is then mapped to the unitary circle by making the centroid of the RoI and the unitary circle centers to coincide.

Classically, shape representations take the first five Zernike orders as approximation to the optical aberrations produced by the human eye [127, 116], being ignored the high-order coefficients since, depending on the image size, a simple calculation of the first 40 moments may take one hour [46]. Provided that the computational cost was herein drastically reduced by the pre-processing phase, moments up to the fifty order can be computed in less than $5\ s$. Additionally, the subset of moments was also selected to represent the global mass characteristics (see figure 2-3). For doing so, three sets of ten masses were randomly selected, decomposed and used to determine the maximum number of orders required to obtain a peak signal-to-noise ratio (PSNR) larger than $40\ dB$ with respect to the original image. The test consisted in measuring the PSNR using increments of ten orders, i.e., reconstruction was performed with the first ten orders, then twenty and so on, finding that the PSNR reconstruction with the first fifty orders was always larger than $40\ dB$.

(a)   (b)   (c)   (d)
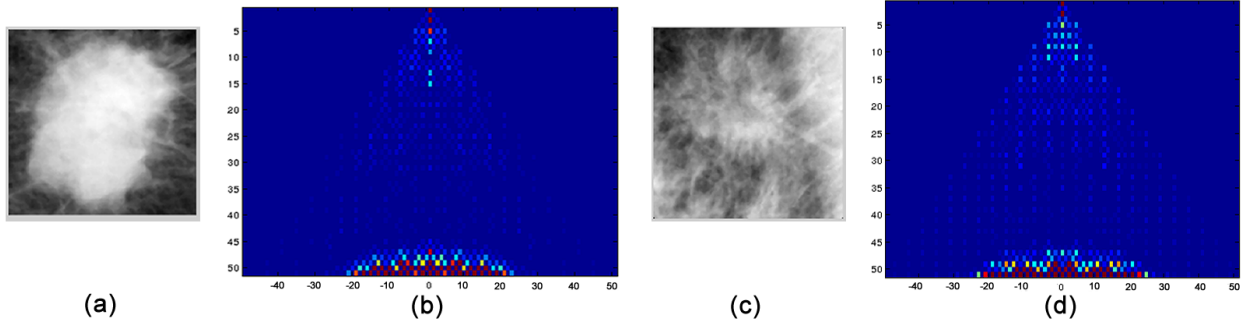
Figure 2-3.: Panels b) and d) illustrate pyramidal Zernike representations, being the x-$axis$ the repetition q and the y-$axis$ the order p. Panels a) and b) show an oval mass and its fifty Zernike moments. Panels c) and d) depict an irregular mass and its corresponding fifty Zernike moments

The group of 50 Zernike orders ($p = 50$) was further reduced by selecting the subset of moments with the larger statistical independence among the three groups of masses, namely round, oval and irregular. This subset of Zernike orders was experimentally determined as follows: ten masses per group were randomly selected and mapped to the Zernike pyramid of the first 50 orders (illustrated in figure 2-3). Exhaustive comparison of the statistical differences for the three different groups was performed by progressively assembling the Zernike orders, i.e., the first two orders, the first three orders and so on. A two-way ANOVA tested the null hypothesis that all the populations from the three groups of moments have identical means. Statistical differences ($p < 0.05$) were only demonstrated for the set of moments between 45 and 50 and the analysis was therefore performed only within this range. The vector feature is then obtained by concatenating six normalized histograms, each corresponding to an order between 45 and 50, and the similarity metrics is the histogram intersection. Interestingly, the choice of a maximal order (50) avoided any numerical instability produced when larger frequencies are included, a limitation of the Zernike polynomials produced by the discretization of these continuous bases. In this work, the numerical and geometrical errors were corrected as shown in [46], while the iterative algorithm [126] corresponds to a modification of the original method presented in [62].

The Curvelet characterization

The Curvelet transform performs a multiscale decomposition [21] using a set of statistically dependent bases. In addition to the usual scale and location informations, each of these Curvelet bases captures the associated orientation. This transformation in particular constrains each scale to follow the parabolic anisotropic scaling law: $width \approx length^2$ [21], ensuring that mass shape and texture patterns are preserved at different scales. Several investigations have demonstrated the Curvelet is a proper texture descriptor in different public databases [37], while others have shown Curvelets outperform transforms like wavelets or Gabor in classification tasks [30, 29]. In the frequency domain, a Curvelet is constructed as the product of two windows: the angular and the radial dyadic frequential coronas. The angular window represents the directional analysis and the radial dyadic window is a bandpass filter with cut frequencies following the parabolic

anisotropic scaling law [21]. Because of this anisotropic property, Curvelets adapt much better to scaled curves than other transforms and therefore result very robust to noise.
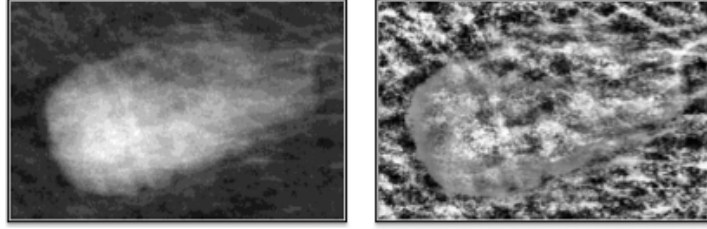
Figure 2-4.: Pre-processed RoI by using the adaptative histogram equalization and median filter

The pre-processing step aimed in this case to preserve main texture patterns and to define a comparable level of intensity. A median filter smoothed out the particular RoI while an adaptive adjustment of the histogram [65] highlighted the main textural patterns and conserved important edges, as illustrated in figure **2-4**.

The RoI was then decomposed into four scales with 32 orientations, for a total of 128 subbands. The curvelet descriptor was constructed by concatenating normalized versions of the different subbands, i.e., 32 orientations at the four scales. The similarity metrics in this case is the histogram intersection of the whole descriptor. The curvelet transform uses herein unequally-spaced fast Fourier transform, a decimated rectangular grid tilted along the main direction of each curvelet [21].

Fusion of Multiresolution Descriptors

Once the Zernike and Curvelet analyses are performed, a kernel strategy aims to choose the optimal way of combining the information obtained from these two representation sources. Specifically, the fusion tries to find optimal values for the parameters that maximize the discriminative ability. A common strategy is to weight the feature spaces and then sum them up to calculate an optimal characterization [35]:

$$\hat{S}(x) = \sum_{\phi} \omega_{\phi} S^{\phi}(x) \tag{2-3}$$

A linear combination $\hat{S}(x)$ of the different features $S^{\phi}(x)$ extracted from the two spaces, with proper weights, might increase the between class separation. The problem with this strategy is that important information from individual features may be lost when somehow linearly combining feature spaces. The proposed fusion uses then a different strategy that keeps all the information from the two feature spaces and only fuses them when a decision needs to be made. Specifically, Zernike and Curvelet spaces are used as input to a discriminant function, $D_W(x)$, that indicates to which extend a particular RoI x belongs to a particular class:

$$D_W(x) = \left\langle W, \left(\Psi_1(S^1(x)), \ldots, \Psi_{\Phi}(S^{|\Phi|}(x)) \right) \right\rangle \tag{2-4}$$

where $W \in \mathbb{R}^{|\Phi| \times M \times N}$ is the vector of parameters that indicates the relative importance of the feature vector $S^{\phi}(x)$ extracted from each space and $\Psi_{1,\dots,\phi} : [S^{\phi}]^{M \times N} \to F$ is a function that maps each descriptor $S^{\phi}(x)$ to a feature space F, herein defined by the histogram intersection metrics. The value of $D_W(x)$ is expected to satisfy $D(x) \geq 0$ if x corresponds to a specific class and $D_W(x) < 0$ if x is not part of that class. This formulation has the advantage of improving the flexibility by allowing the model to account for potentially complex non-linear interactions of the original single space values corresponding to different features and scales.

Provided that only two different spaces need to be combined, the functions Ψ are charged of mapping the descriptors to a new feature space with a dot product defined by the histogram intersection, where complex non-linear patterns become linear. This is a well known strategy used in kernel methods, the *kernel trick*, where the mappings Ψ are implicitly induced by a kernel function. A kernel is a function $k : X \times X \to \mathbb{R}$ associated to a mapping $\Psi_{1,\dots,\phi} : X \to F$ such that $\forall x, y \in X, k(x, y) = < \Psi_1(x), \Psi_1(y) >_F$, i.e., k calculates the dot product in F. Intuitively, a kernel may be seem as a function that measures the similarity between two objects of the input space, in the present case, the histogram intersection. In the proposed model, the input space may be the space of Curvelets, the Zernike maps or the fused descriptor.

The optimization problem is formulated then as the following min-max-margin discrimination problem:

$$\min_{W} \left\{ C \sum_{i=1}^{N} \max(0, 1 - y_i D_W(x_i)) + \|W\|_2^2 \right\} \tag{2-5}$$

where x_i represents a training RoI, $y_i \in \{-1, 1\}$ represents the corresponding label, N is the number of training samples, and C controls the regularization of the model. The first term of the objective function in equation 2-5 is a loss function that penalizes the wrong classification of training samples and the second term is the regularizer of the W parameter. The C parameter was set to 1000 after an exhaustive search of the optimum value under a 10-fold cross validation of the available data. The resultant weights combine the different kernels, associated to each feature vector, aiming to obtain a maximal separation between the two classes

Regarding the functions Ψ, in this work we used the histogram intersection kernel defined as:

$$\Psi = k_{hi}(s_p, s_q) = \sum_{i} \sum_{j} \min \left(s_p(i, j), s_q(i, j) \right) \tag{2-6}$$

This histogram intersection kernel requires the different orders of the Zernike moments and the several sub-bands of the Curvelet coefficients are normalized.

The proposed model fuses the different spaces by combining their respective kernels, instead of directly adding the spaces themselves, the optimization problem in equation 2-5 is treated as a Multiple Kernel Learning (MKL) problem, aiming to find a good discriminant hyperplane. Several MKL formulations have been introduced so far, the approach herein used was proposed by [122], the Generalized Multiple Kernel Learning (GMKL). In practice the fusion is implemented at the level of the different descriptors obtained by the multiresolution analysis, namely four scales and 32 orientations for the curvelet and five orders (from 45 to 50) for the Zernike base. Each

of these descriptors is basically a normalized histogram, i.e., a probability distribution function (pdf). The training phase constructs the similarity matrices by intersecting the pdfs corresponding to each descriptor, i.e. separately per scales, orientations and orders. The several obtained kernels or similarity matrices are then linearly combined by the MKL which finds out a linear combination that optimizes the separation between classes. The SVM is then trained with this optimized kernel version. Once a new case arrives, the obtained multiresolution descriptors for this case are intersected against the support vectors obtained in the SVM training phase and linearly combined using the weights computed by the MKL. The new kernel then feeds the SVM model that takes the binary decision.

2.3.3. Screening RoIs for masses

A first stage of the proposed method consists in establishing whether or not a particular RoI contains a mass. Once breast images are mapped to the three different spaces as described so far, the different representations feed SVM classical binary classifiers that use its precomputed kernels by computing the histogram intersection as distance metric, and classify the particular RoI as containing or not a mass.

2.3.4. Setting the mass shape with a bank of classifiers

Support Vector Machine (SVM) is a classical classifier that discriminates two classes [24], yet this algorithm can be adapted to a multi-class problem. A multi-class SVM classifier can be obtained by training several classifiers and combining their results with a structure known as the bank of classifiers. The bank herein implemented trains one classifier per class to discriminate between that class and the rest of classes. Each classifier assumes positive values when cases belong to the trained class and negative values otherwise. In a multiclass problem, these values are compared and the decision of the bank of independent classifiers is the value for which the discriminating function is maximal. The most common discrimination function is the signed distance between a particular case to classify and the hyperplane [47]

$$f_i = \frac{\mathbf{w}_i.\mathbf{x} + b_i}{||\mathbf{w}_i||}, \ i = 1, 2, ..., n \tag{2-7}$$

where $\mathbf{w}_i$ and $\mathbf{b}_i$ are, respectively, the normal and the bias of the hyperplane related to the ith classifier. Hence, given a RoI x, the label $\mathbf{c}$ is assigned as $c(x) = \max\limits_{i=1,2...,n} f_i(x)$. This baseline assigns thereby a unique label to the different classes and a complete description is obtained.

2.3.5. Estimating a mass description by retrieving similar shapes

One main contribution of the present proposal is the integration of prior knowledge in the database, improving that way the description quality. The process of annotating a mass consists in assigning a particular label to a mass feature. However, in this particular problem the classes are not independent and the description may be considered as hierarchical, that is to say, certain dependences are unavoidable, for instance a mass can be round and the margin circumscribed but never

spiculated. This sort of conceptual connection was herein exploited by designing a new automatic description strategy. The approach starts by classifying a particular shape since this is the most global feature and therefore less prone to be contaminated by noise. Once the evaluated RoI is assigned, the strategy retrieves RoIs in the database with the same shape and selects the k most similar ones, using the discrimination function in equation 2-7 as the similarity metric. The k number of masses was determined by an extensive tuning over the set of available masses, five for the present investigation. This set of similar RoIs assigns the missing classes, margin and density (recall only shape has been set so far), using the associated terms of the recovered RoIs. A simple majority voting decides any of the missing margin and density, but every vote is weighted by the distance to the hyperplane found by the shape classifier using the metrics defined in equation 2-7.

2.3.6. Evaluation metrics

In a typical classification task, a classifier assigns an instance to one of a finite and mutually exclusive collection of possibilities and therefore each instance is associated to a single label. In contrast, multi-label problems include cases that allow a variable number of labels, which are assigned to each instance, as in the present case. Multi-label learning can be stated as the problem of finding a model that maps the input to a binary vector, rather than to scalar outputs, as in the ordinary classification problem. Usually, this multi-label problem is transformed into a simpler binary problem for each label. This is usually found in the literature as the binary relevance method [47], which consists in independently training one binary classifier for each label so that any of these bank-binary-classifiers can then be individually applied to assign a particular label.

The classification performance for shape, margin and density classes were herein independently evaluated under a multi-label approach. In consequence, a set of binary "one-vs-all" SVM classifiers were trained and evaluated. The best set of parameters for each classifier were chosen from a conventional 10-fold cross validation. The performance of the multi-label task was measured by the next metrics: If the sets of labels are denoted as T, the true, and P, the predicted, the metrics were: precision $\frac{|T \bigcap P|}{|P|}$, recall $\frac{|T \bigcap P|}{|T|}$ and F_1 score (F-measure) as the harmonic mean of the precision and recall , $2\left(\frac{precision \times recall}{precision + recall}\right)$. In addition, Accuracy $\frac{tp+tn}{T+N}$ was computed from the confusion matrix, where each column of the matrix represents the instances in a predicted class while each row represents the instances in an actual class, namely tp (true positive), tn (true negative), T (all positive) and N (all negative). Finally, the performance of the mass recovery strategy was evaluated by the conventional precision-recall curve obtained when iteratively recovering sets of the five most similar lesions.

2.3.7. Mammographic Databases

The herein proposed strategy is evaluated using different quality levels of mammographic images, a digitized version of screen-film mammographies and full-field digital mammographies. For this, two separate datasets are used containing each one a set of RoIs extracted from two public mammographic image databases, the Digital Database for Screening Mammography (DDSM)[1] and the

---

[1]http://marathon.csee.usf.edu/Mammography/Database.html

INBreast database[2], respectively. A brief description for each dataset is underneath shown.

- The DDSM database: the Digital Database for Screening Mammography database (DDSM) [44], a mammography database widely used as an evaluation benchmark [79, 45, 131], was herein used to assess the present proposal. This open access database is constituted of digitized images of mammographic films with the corresponding technical and clinical information. The whole DDSM database contains a total of 2620 cases, each including four images obtained from Cranio-Caudal (CC) and Mediolateral-Oblique (MLO) views as well as specific BI-RADS descriptions that were annotated by expert radiologists according to BI-RADS fourth edition. In this investigation, a subset of mammographies were chosen, the inclusion criteria were: 1) similar image quality, 2) balanced number of the type of masses, 3) only a single lesion was assessed, i.e., this study excluded masses with calcifications, architectural distortion or cases with the presence of clips or pencil marks in the mass area. Thus, a subset of mammographies digitized either with Lumisys laser film scanner at 50 μm or a Howtek scanner at 43.5 μm pixel resolution and at a dynamic range of intensities of 2^{12}=4096 gray level tones, were selected. Then, a cropping process was carried out. The mass location and size, information associated to each case, were used to manually crop squared sub-images centered at the mass. This is because of image annotations, in DDSM database, include level boundary of the findings, which is not considered an accurate segmentation. In this work, the input size to RoIs was defined to n x n pixels. Specifically, RoIs were cropped to the bounding box of the lesions and rescaled to n x n pixels preserving the aspect ratio when either width or hight of the bounding box are greater than n, otherwise the lesion is centered without scaling and preserving the background tissues. In consequence, a set of 980 regions of interest were extracted from this DDSM database by including CC and MLO views, distributed as 490 regions of interest with normal tissue (non-mass) and 490 regions of interest with breast masses. The RoIs used as normal tissues, with no mass, corresponded to exactly the same dimensions but at a random locations in the glandular tissue. Figure 5 shows the distribution of the set of different masses according to BI-RADS description extracted from DDSM database.

- The INBreast database: additional evaluation was performed using the cases stored in the INBreast project [77]. The INBreast database contains mammographies in DICOM format and acquired at the Hospital de São João, Breast Centre, Porto, Portugal. This collection of cases includes masses, calcifications, asymmetries and distortions. A total of 115 cases (410 images) constitutes this database, from which 90 cases are bilaterally documented (4 images per case) and 25 are patients mastectomized (2 images per case in these cases). In consequence, a set of 108 masses were extracted as well as 108 normal RoIs. The mass location and size, associated to each case, were taken from the contours drawn by experts and provided in XML format.
Finally, the set of RoIs evaluated in the present investigation was updated to the fifth version by our expert radiologist who has reported experience reading mammographies by at least fifteen years. The RoIs with control tissues were chosen from the fibroglandular tissue, for which the expert selected a squared region of the same size of the mass. Likewise, the mass density was described by the expert, a sign not included in the original DDSM database. Figure **2-5** shows

---

[2]http://medicalresearch.inescporto.pt/breastresearch/index.php/Get_INbreast_Database
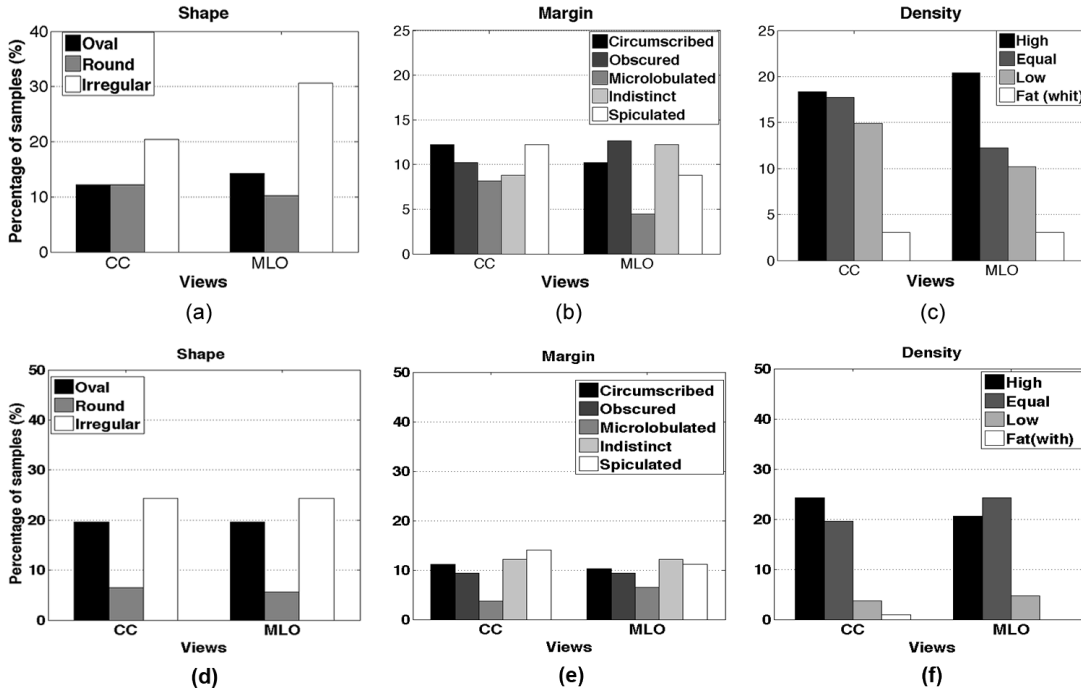
Figure 2-5.: Histogram distribution of the different BI-RADS features in percentage for both databases, the DDSM (top row) and the INBreast (bottom row). a) and d) shape distribution for features Oval, round and Irregular; b) and e) margin distribution for features circumscribed, obscured, microlobulated, indistinct and spiculated; c) and f) density distribution for features high, equal, low and fat

the distribution of masses for the different shapes, margins and densities that were herein chosen, aiming to preserve a balanced number of masses for each database.

2.4. Results

The proposed method was twofold evaluated. First, for the task of classifying an input RoI as mass or control tissue (non-mass). Second, when assessing the accuracy of assigning a particular shape and the subsequent margin and density terms using the subset of retrieved masses with a similar shape. The experimental evaluation process was implemented in MATLAB R14, running on a Linux PC with 2 Intel Quad Core i7 at 3.07 GHz and 24 GB of RAM. In all experiments, we used the Generalized Multiple Kernel Learning (GMKL) source code as the implementation of the Multiple Kernel Learning strategy [122], while the SVM classifier was implemented using the LIBSVM toolbox [23], with precomputed kernel matrices.

2.4.1. Screening RoIs for Masses

A first task with the proposed method consisted in detecting whether or not a mass was present in the experimental set of RoIs, composed of 980 RoIs extracted from the DDSM database (490 with

normal tissues and 490 with masses) and a set of 216 RoIs from the INBreast database (108 controls
and 108 with masses), respectively. This evaluation compared the performance obtained by using
Zernike (ZM), Curvelet (CT) and the proposed fused (ZM-CT) bases. The three representation
bases were used as input to a classical binary SVM classifier with precomputed kernel matrices.
The two sets were 70-30 split, a random selected 70% for training while the remaining 30% was
used for validation, a process repeated 10 times. In addition to the aforementioned evaluation
metrics, the area under the curve (A_z) of the ROC curve was also included. The results obtained
for this task are presented in Table 2-2.

Table 2-2.: Averaged performance under a $70 - 30\%$ fold validation scheme for the task of classi-
fying RoIs as normal or containing breast masses for DDSM and INBreast databases

	DDSM			INBreast		
	ZM-CT	ZM	CT	ZM-CT	ZM	CT
A_z	**0.94**	0.79	0.78	**0.98**	0.88	0.87
$Sensitivity$	**96.2**	83.3	82.2	**98.6**	90.1	88.4
$Specificity$	**93.1**	80.1	78.3	**99.1**	89.3	85.2
$Precision$	**95.3**	80.2	80.1	**99.6**	92.1	87.3
$Accuracy$	**96.8**	84.2	82.2	**99.8**	91.0	89.7
F_1-**score**	**95.7**	83.1	82.1	**99.0**	90.1	87.8

Overall, the fused descriptor (ZM-CT) outperforms the Zernike (ZM) and Curvelet (CT) bases,
reaching a sensitivity of 96.2 % for the DDSM and 98.6 % for the INBreast databases, while the
respective specificities were 93.1 % and 99.1 %. In addition, the Accuracy and A_z metrics shows
remarkable results, with 96.8 and 0.94 for DDSM database and 99.8 and 0.98 for INBreast database,
respectively. Finally, the F_1-score evidences also a higher accuracy with the fused base when
comparing with the two baselines.

2.4.2. The mass description

The description task is achieved by firstly assigning a particular shape to the mass with a bank
of classifiers and then completing the margin and density labels by integrating the information
associated to a set of similar masses retrieved from the database.

Setting the mass shape with a bank of binary classifiers

This experimentation assessed the discrimination power of the fused base to assign a shape term
to the RoI, namely *Round*, *Oval* and *Irregular*. This multi-label problem was approached by a bank
of binary SVM classifiers, trained using a conventional 10-fold cross validation scheme, under a
"one vs all" paradigm, i.e., every binary classifier associates a particular shape term to the positive
label while the others are considered as negative. The histogram intersection was the metrics used
as a precomputed kernel function for each classifier. If only a single classifier is positive, the mass
is assigned after this classifier, but if more than one classifier is positive, the decision is taken using

a distance criterion to the SVM hyperplane, i.e., the classifier with the largest distance to the SVM hyperplane (see equation 7). This choice is guided by the hypothesis that the farther the sample is w.r.t the classification hyperplane the larger the probability the sample belongs to that class, or in other words, the method selects the sample that is closer to the class centroid.

Table 2-3.: Performance of the bank of classifiers and the three descriptors at setting the RoIs with a shape term, namely Round, Oval and Irregular for each database, respectively

	DDSM			INBreast		
	ZM-CT	ZM	CT	ZM-CT	ZM	CT
Sensitivity	**92.3**	89.4	86.4	**97.7**	93.3	87.5
Specificity	**89.2**	87.4	81.4	**96.3**	89.8	83.6
Precision	**88.2**	82.3	78.9	**96.5**	90.2	85.4
Accuracy	**92.1**	88.6	80.2	**98.2**	92.3	90.6
F_1-**score**	**90.2**	85.7	82.4	**96.9**	90.1	84.6

Results are shown in table **2-3**. In general, the best performance was obtained for the INBreast database. Zernike bases show they capture low level features, with sensitivity of about 89 % and 93 % and respective specificities of 87 % and 89 % for the two databases. Accuracy, precision and F_1-measure evidence a slight superiority of the Zernike descriptor, likely indicating a better global description. The fused descriptor outperforms the other two at assigning the shape class description for the two databases

Recovering similar masses

Provided that a classification function unavoidably generates a biased discrete topology or partition of the feature space, we also evaluated the metrics for recovering a series of similar sets of masses, that is to say the aptness of the method to obtain a similar description. Unlike the classification task, the retrieval assesses the bias of the similarity function, a task that is usually evaluated with the precision-recall curve. In general, precision (positive predictive value) is the proportion of retrieved masses that is relevant in a particular search, while recall (sensitivity) is the fraction of relevant masses that are retrieved. Both precision and recall quantify the number of relevant masses.

Specifically, the test consisted in determining what percentage of the retrieved set of masses exactly matched the same description of the query. The set of evaluated queries is presented in table **2-4**, being each query a set of ten randomly selected RoIs from the database with the same description. For a particular query, the system retrieved the set of the most similar 50 RoIs from the whole database, with increments of 5 RoIs.

Unlike the DDSM dataset, the INBreast database contains a reduced number of breast masses with a similar BI-RADS description (i.e In terms of shape, margin and density). Therefore, this task was only performed for the DDSM dataset.

Six different queries were assessed and are shown in table **2-4**. An oval shape is then either circumscribed with high density or obscured with equal density. The round mass is either circumscribed

Table 2-4.: The different full BI-RADS descriptions available in the DDSM database and the num-
ber of RoIs for each. The query number stands for a particular experiment

Query	Shape	Margin	Density	Number
1	*Oval*	Circumscribed	High	50
2		Obscured	Equal	54
3	*Round*	Circumscribed	High	50
4		Obscured	Equal	103
5	*Irregular*	Spiculated	High	130
6		Indistinct	Equal	103

with high density or obscured with equal density while the irregular mass is either spiculated
with high density or Indistinct with equal density. Obviously the retrieval task requires a simi-
larity notion, which in this case corresponded to the strategy presented in section 2.3.5. Briefly,
a bank of classifiers establishes the mass shape while the margin and density terms are chosen
by the voting strategy presented in section 2.3.5. A positive result is an exact match of the three
description features, namely *shape*, *margin* and *density*.

Figure **2-6** shows the Precision-Recall curves obtained for oval masses with different margin and
density, that is to say, queries 1 and 2 in table **2-4**. Each point of this curve corresponds to the
precision-recall obtained with increments of five masses. In both curves the fused approach (blue
line) outperforms the other two representations.

(a) Oval-Circumscribed-High   (b) Oval-Obscured-Equal
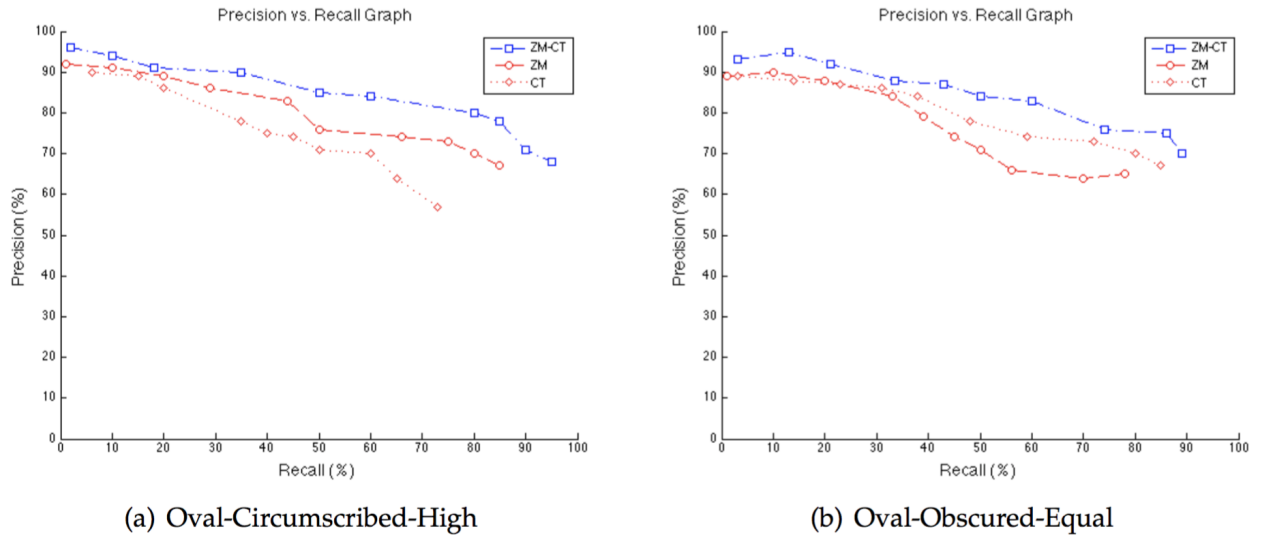
Figure 2-6.: Precision vs. Recall curves for Oval masses with (a) Query 1: Circumscribed margin
and High density, (b) Query 2: Obscured margin and Equal density

When retrieving similar oval masses with a circumscribed margin and high density, the precision
and recall averages were 93.7 % and 72.5 % with the fused descriptor, 85.72 % and 56 % with the

Zernike moments and 85.4 % and 50.9 % with the curvelets. On the other hand, for masses also Oval but margin Obscured and Equal density, the fused descriptor averaged precision and recall were 94.3 % and 71.7 %, they were 80.1 % and 65.5 % with the Zernike moments and 84.1 % and 55.3 % with the curvelets.

(a) Round-Circumscribed-High   (b) Round-Obscured-Equal
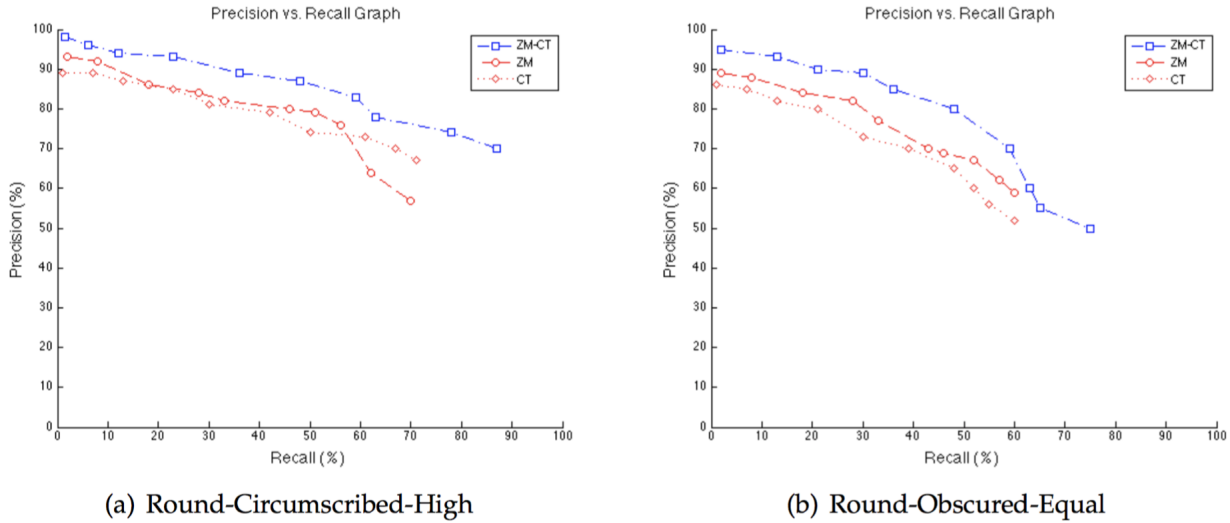
Figure 2-7.: Precision vs. Recall curves for Round masses with (a) Query 3: Circumscribed margin and High density, (b) Query 4: Obscured margin and Equal density

Figure 2-7 shows the Precision-Recall curves obtained for round masses with different margin and density. Again, both curves show how the fused approach outperforms the other two descriptors. For circumscribed margin and high density, the precision and recall averages were 90.8 % and 67.2 % with the fused descriptor, 83.1 % and 60.1 % with the Zernike moments and 81.1 % and 60.6 % with the Curvelets. On the other hand, for masses also round but margin obscured and equal density, the averaged precision and recall were 89.3 % and 61.2 % with the fused descriptor, 84.7 % and 54.5 % with the Zernike moments and 79.8 % and 52.4 % with the Curvelets.

Figure 2-8 shows the Precision-Recall curves obtained for irregular masses with different margin and density. As in previous queries, both curves show the fused approach (blue line) outperforming the other two representations. For retrieving the most similar irregular masses with a spiculated margin and high density, the precision and recall averages were 89.7 % and 64.7 % with the fused descriptor, 80.2 % and 57.2 % with the Zernike moments and 83.3 % and 65.7 % with the Curvelets. On the other hand, with masses also irregular but margin indistinct and equal density, the averaged precision and recall were 86.7 % and 55.7 % with the fused descriptor, 79.2 % and 56.2 % with the Zernike moments and 80.3 % and 53.7 % with the Curvelets.

Describing the mass with a set of similar retrieved masses

The description method consists in setting the mass shape by a bank of classifiers and then assigning the margin and density using the information associated to the set of five most similar shapes

(a) Irregular-Spiculated-High   (b) Irregular-Indistinct-High
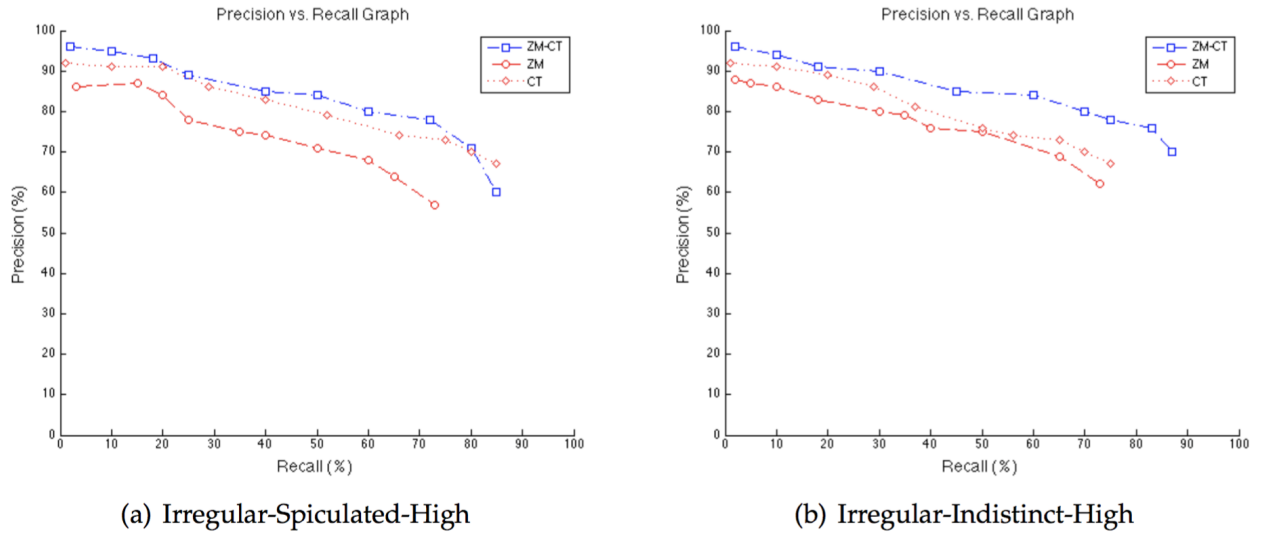
Figure 2-8.: Precision vs. Recall curves for Irregular masses with (a) Query 5: Spiculated margin and High density, (b) Query 6: Indistinct margin and High density

retrieved from a database. This approach hypothesizes that a mass shape is a global characteristic that is less sensitive to noise and uncertainty, unlike margin and density which are concepts much more related to local information. The complete description is achieved by a weight voting of the margin and density associated to these five retrieved masses. These weights correspond to the distance of each of these similar lesions, in the SVM space, to the closer hyperplane.

Table 2-5.: Classification performance for the margin and density terms for DDSM database. Recall that the shape has been already assigned at this point and the whole description requires also the margin and density terms

	ZM-CT			Retrieval Approach		
	Sensitivity	Specificity	Acc	Sensitivity	Specificity	Acc
Margin	80.3	79.1	82.6	**89.2**	**85.6**	**92.2**
Density	79.3	83.2	84.3	**87.4**	**94.4**	**93.6**

Table **2-5** presents the results of the evaluation of the complete description strategy with the same set of lesions used for the previous shape classification for DDSM database, the shape values were not included since they were already presented in Section 2.4.2. This scheme was compared against the bank of classifiers which was also trained to determine each term of the margin and density BI-RADS features. The three metrics, sensitivity, specificity and accuracy, remarkably improved when applying the retrieving strategy, for the case of margin the sensitivity improved from 80% to about 90%, the specificity increased from 80% to 85%, and the accuracy also increased from 82% to 92%,. The density feature also improved a sensitivity of about 80 % to 87 %, specificity from 83 % to 94 % and accuracy from 84 % to 93 %.

This task was also performed with the INBreast database, achieving the complete mass description under the same protocol of progressively recovering 5 more masses. Table 2-6 presents the results for the INBreast dataset. As observed for the other database, the obtained sensitivity, specificity and accuracy improved when applying the retrieving strategy, for the margin sensitivity increased from 82% to about 88%, the specificity from 80% to 87% and the accuracy from 82% to 92%. The mass density feature also raised the sensitivity from 83 % to 87 %, the specificity from 85 % to 93 % and the accuracy from 84 % to 93 %. Despite the reduced number of RoIs in the INBreast dataset, the fused descriptor proved to capture relevant information.

Table 2-6.: Classification performance for the margin and density terms for INBreast dataset. Recall that the shape has been already assigned at this point and the whole description requires also the margin and density terms

| | ZM-CT | | | **Retrieval Approach** | | |
	Sensitivity	Specificity	Acc	Sensitivity	Specificity	Acc
Margin	82.4	80.3	82.0	**88.7**	**87.2**	**91.4**
Density	83.6	85.5	84.8	**87.8**	**93.2**	**93.1**

Finally figure 2-9 illustrates the complete mass description for DDSM database, refined by the retrieval process. The two masses are shown at the left panel of top and bottom rows together with their most similar retrieved masses. The classification distance is used to weight each of the five masses which are then combined by the weighted voting strategy.
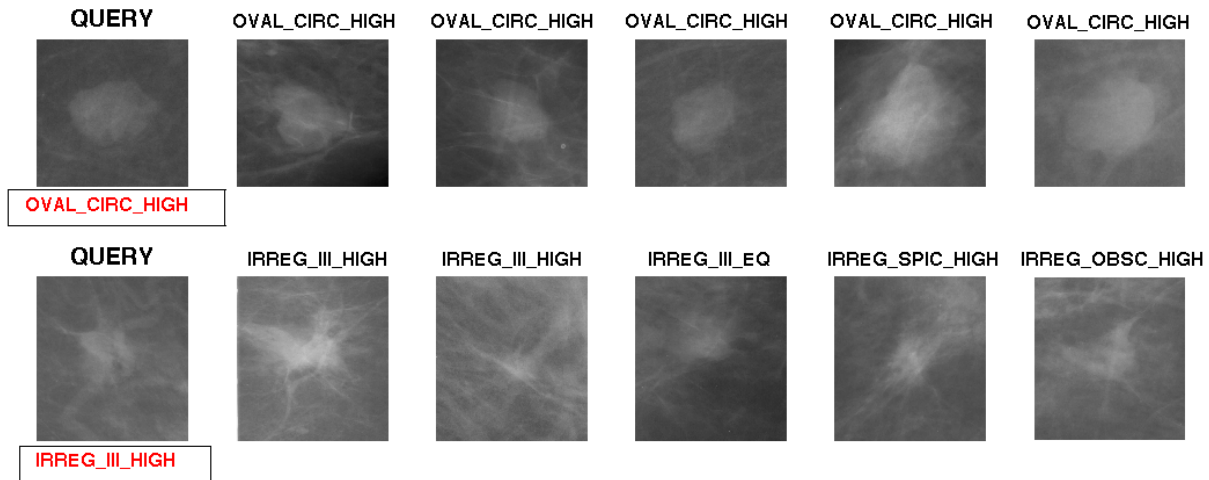
Figure 2-9.: The retrieval scheme and the whole description of two masses at the upper and bottom rows: the queried mass is shown at the first panel from left to right. The 5 most similar masses are shown by order of similarity from left to right. The annotations associated to the retrieved masses are combined by the weight voting strategy.

2.5. Discussion

Overall, there is a great difficulty at attempting to contextualize the presented results since other existing methods either use different databases or, even using the same database, the experimental setup is different. However, the presented results are competitive or outperform many methods of the state-of-the art.

Table **2-7** shows a summary of several methods that differentiate a mass from normal tissue in RoIs extracted from different databases, including the two used in the present investigation. These methods inform the area under the ROC curve A_z or the Accuracy, two metrics that can be seen as complementary. Evaluation of these methods is performed with different databases, MIAS [75, 29, 1], BCDR [19], IRMA [70] and the two databases used here (DDSM and INBreast). Overall, the present method is competitive and the evaluation is obviously more extensive, the number of DDSM RoIs was 980 while the number of ROIS with the other databases is considerably smaller. This factor is important if one considers that masses are likely the most variable lesions of the breast pathology at a radiologic level. The method presented by [1] reported a slightly larger A_z than ours, 0.99 against 0.98 for the INBreast database, but the present method reaches a higher value for the digitized database (0.94 for the DDSM against 0.91 for the MIAS). The present approach also outperforms other methods that proposed to combine different texture features [75, 101, 19], even the method proposed by [70] that fuses texture (wavelet decomposition) and shape (Zernike moments) representations.

Table **2-8** summarizes the best results presented by several methods when describe masses in terms of the BI-RADS mass features. Table **2-8** is organized by first introducing some methods that have classified the mass shape, then the mass margin and finally those works that have investigated the automatic classification of the mass shape, margin and density. As was mentioned above, there are no previous works that only classified the mass density feature. Regarding the shape description, the methods presented in [113] and [121] reported a very high accuracy but with less than the fourth part of the cases that we used here, reason by which the sources of noises of this work are considerably smaller and results can be considered more biases than those herein presented. Regarding the margin, the proposed approach clearly outperforms the other two methods [103, 53]. Moreover, since those methods differentiate only spiculated from circumscribed masses. Likewise, when comparing with the methods proposed by kisilev et al. [64, 63] that carry out a full description of the shape, margin and density, the present method reaches higher values of the accuracy metrics.

On the other hand, recent content based image retrieval investigations have been proposed to support diagnosis [27, 133], proving a substantial improvement when examining masses in mammograms [80]. This kind of systems aims to retrieve cases with similar visual content and use the associated information to improve the final diagnosis. One of the most important issues when developing a retrieval system is that the extracted features have to be discriminative enough to represent different lesion characteristics. Recent content based image retrieval investigations have been proposed to support diagnosis of breast masses. [127] devised a system that selected the best characteristics under a custom made metrics, a relevance measure that is weighted after the radiologist feedback. They evaluated the effect of both Zernike moments and margin features as

Table 2-7.: Comparison of results obtained for the BI-RADS mass description task.

Reference	Database	RoI	Feature Extraction	Classifier	Results
Moayedi(2010)	MIAS	90	Contourlets, GLCM and Morphology	SVM	*Ac* 96.6
Eltoukhy (2012)	MIAS	322	Wavelets	SVM	*Ac* 94.8
			Curvelets		*Ac* 95.7
Reyad (2014)	DDSM	512	Wavelets	SVM	*Ac* 95.7
			Contourlets		*Ac* 98.2
			Statistical features		*Ac* 98.4
			LBP		*Ac* 98.4
			LBP and Statistical		*Ac* 98.6
Abdel-Nasser (2015)	MIAS	312	ULDP	LSVM	A_z 0.93
	IN-breast	407			A_z 0.99
Bruno (2016)	DDSM	240	LBP and Curvelets	PL	A_z 0.91
	BCDR-FMR	106			A_z 1.0
	BCDR-DMR	95			A_z 1.0
de Lima (2016)	IRMA	355	Zernike-Wavelets	SVM-ELM	*Ac* 94.1
Our method	DDSM	980	fused ZM-CT	SVM	**Az 0.94, Acc 96.8**
	INBreast	216	fused ZM-CT	SVM	**Az 0.98, Acc 99.8**

low level descriptors and assessed their aptness to retrieve similar masses under a full BI-RADS description. The data set consisted of 1919 mammograms from the DDSM database. Four mass-classes were tested, namely IRR-ILL (265 masses with irregular shape and ill-defined margin), LOB-CIR (178 masses with lobulated shape and circumscribed margin), OVA-OBS (152 masses with oval shape and obscured margin), and ROU-CIR (59 masses with round shape and circum-scribed margin). Their results showed the highest precision, 70%, with the ROU-CIR class, when retrieving the first nine RoIs and 68 to 65% for the rest of the ninety retrieved RoIs. Precision fell down with the other three lesion classes, from 57% to 45%. In contrast, the presented method retrieved the first five most similar regions with a precision of 92% for IRR-ILL (masses with ir-regular shape and ill-defined margin), while the other three classes were found to have precision rates from 80% to 85%

The use of the Zernike moments has been reported for the task of classifying a breast mass as

Table 2-8.: Comparison of results obtained for the BI-RADS mass description task.

Shape

Reference	Database	RoI	Feature Extraction	Classifier	Results	
[105]	DDSM	100	Beamlet Decorated Recursive Dyadic Partition	KNN	Ac 72 - 78	
[113]	DDSM	224	Geometric shape and margin features	Neural network	Ac 97,5	
[121]	DDSM	224	Geometric shape and margin features	Decision tree	Ac 87.7	

Margin

Reference	Database	RoI	Feature Extraction	Classifier	Results	
[103]	Home-made	396	Spiculation and cicumscription measurements	Threshold	A_z 0.87 and 0.96	
[53]	DDSM	211	Spiculation index	Threshold	A_z 0.70	

Shape, margin, and density

Reference	Database	RoI	Feature Extraction	Classifier	Results	
[64]	DDSM	1546	Geometric shape, boundary sharpness, (LBP) texture	SSVM	Ac 71 (Shape) Ac 69 (Margin) Ac 73 (Density)	Ac Ac
	home-made	270			Ac 79 (Shape) Ac 78 (Margin) Ac 82 (Density)	Ac Ac
[63]	DDSM	974	–	CNN	Ac 82 (Shape) Ac 77 (Margin)	Ac
	home-made	646			Ac 88 (Shape) Ac 86 (Margin) Ac 84 (Density)	Ac Ac
Our method	DDSM	980	Fused ZM-CT	SVM	**Ac 92.1** (Shape) **Ac 92.2** (Margin) **Ac 93.6** (Density)	
	INBreast	216			**Ac 98.2** (Shape) **Ac 91.4** (Margin) **Ac 93.1** (Density)	

benign or malignant [116, 107], but not for discriminating normal tissue from masses. These approaches characterize manually delineated regions using very low (up to 10th) or low (from 10th to 17th) orders of the Zernike moments which are then used as input to some classifier, while in this investigation a mass was found and described in a RoI without any previous segmentation (a more challenging task). A main contribution of the present investigation has been the adaptation of the Zernike moments to this specific problem, resulting in a small subset of moments with little information loss, i.e., from the 45th to 50th orders, and drastically reducing the computational cost (about $5\ s$ compared with the described scale of hours). This approach is further complemented by the fusion with the curvelet base and a retrieval strategy that both mitigate what might be lost by the use of the pure Zernike moments. The exhaustive evaluation has demonstrated a remarkable performance of the present method in very different tasks, classification or retrieval, and the possibility of learning from new data. In spite of that, some limitations of the proposed method are:

- Since the proposed system has been devised to be integrated with the radiologist work flow, the input RoIs must be previously selected. This human intervention limits the method to be part of a conventional computer aided detection system (CADx) but increases the control the expert has on the final diagnoses.

- An accurate margin and density description fully depends on the initial shape classification, a task for which the method showed good results in the present investigation. However, a necessary further evaluation would require a larger database. Some errors might be observed with masses located near the pectoral muscle or within highly dense tissue distribution, case in which even for an experimented radiologist the diagnosis could be misleading.

- In addition, the margin and density description also depends on the number of available masses in the database: The smaller the mass variability, the higher the probability of misdiagnoses.

- Another point is that the method was devised and trained to describe only raw masses, that is to say, with no other abnormalities. However, many masses present mixed patterns with several lesions, i.e., a mass may contain some calcifications or architectural distortions. These patterns of course surely would require a different type of training that however was far from the scope of this investigation.

2.6. Conclusions and future works

The whole method is built upon a pair of complementary multiresolution representations, i.e., Zernike moments and Curvelet coefficients, optimally fused (combined) by an adaptable Multiple Kernel Learning framework. After the exhaustive evaluation in classification (mass/non mass) and retrieval (mass description) tasks, the method has demonstrated a high accuracy level. The CBIR strategy has integrated the knowledge stored in the database with the actual radiologist's work flow, facilitating thereby the interaction of the radiologist with cases already diagnosed in the collection.

The multiresolution analysis captures shape information with no need of a previous mass segmentation or delineation, a side advantage of the method. Evaluation demonstrates the Zernike higher order moments (45-50) discriminate slight shape variation, v.g., round and oval shape. Unlike the classic use of this representation, the method here attempts to catch relevant texture or high frequency information and uses the moments describing these frequency bands to properly classify masses, for instance defined or spiculated shapes (benigh/malignant).

The obtained results have confirmed the method adapts to different image resolutions as proved when the model trained with the DDSM database was used to classify cases from the INbreast database. This type of evaluation illustrates how the present approach results highly reliable to be used in actual clinical scenarios and adapted to any CAD strategy, either in classification tasks or in retrieval goals.

Finally, the proposed method describes the mass density using annotations from the database and therefore avoids the usual comparison the radiologist performs with the database, a very subjective process which is while error prone.

In addition, future works might be:

- The system could be more complete if a BI-RADS category is assigned to the case (between 1 and 6) by fusing the shape description and the clinical information, performing the second reading that actually helps the diagnosis process out.

- Another important point is predicting mass malignity and include this estimation into a CAD system.

- Proposed descriptor could be applied to other types of lesions like the architecture distortion or calcification or even used to describe the tissue density.

- The present analysis in addition might be extended to other modalities like ultrasound or magnetic resonance that require lesion characterization in terms of shape and margin.

- Finally, other findings in the same RoI could be also described with this strategy.

3. Characterization of architectural distortion

*Among the different mammographic findings, architectural distortion (AD) is a common cause of false-negatives in mammograms. This lesion usually consists of a central retraction of the connective tissue and a spiculated pattern radiating from it. This pattern is difficult to detect due the complex superposition of breast tissue. This chapter presents a novel AD characterization by representing the linear saliency of a stellate pattern in the spatial domain, aiming to strength the classification process. The proposed Linear Saliency Domain method (LSD) represents a RoI as an (initially) fully connected graph with nodes corresponding to the RoI boundary and edge weights calculated using line integrals of the intensities along the path connecting any pair of nodes. By doing this, the edges corresponding to the most salient lines are assigned a larger weight in the graph. A centrality measure based on the graph's adjacency matrix's eigenvectors is used to identify the most relevant nodes. A further dimensionality reduction is accomplished by selecting only the most salient edge originating from relevant nodes. The complete content of this chapter has been published as a research article in the **Journal of Medical Systems** (see [81]).*

3.1. Introduction

Despite the widespread use and proven effectiveness of mammographies, screening mammography programs showed a high intra and inter-observer variability in the early 90s of last century with undetected cancer rates of between 10% and 30% [100]. Previous studies have reported rates of between 10% and 25% [14, 115]. This high rate is partly due to the fact that identifying radiologic signs for mammographic interpretation remains a complex visual task because of the blurriness originating from the overlaying of a complex superposition of breast tissues. Among the most important mammographic findings, AD has been reported as the most commonly missed abnormality in false-negative cases [56]. An AD lesion is any change of the spatial distribution of the breast connective tissue, in general with no visible mass. In such mammographic finding, the typical retraction of tissues is variable and gradual, commonly showing radial spiculations. Focal retraction or distortion at the boundary of the parenchyma without a visible or palpable mass may also be present. The set of radiological signs of an AD has been collected and formalized within the Breast Imaging Reporting and Database Systems (BI-RADS) standard [108]. Figure **3-1** illustrates the linear patterns in both AD and normal tissue, manually selected by a radiologist.

3.1.1. Previous work

Nowadays, despite of development of Computer Assisted Diagnosis Systems (CAD) in mammography has proved to be effective in finding calcifications and masses [88], CAD systems have not been as effective in the detection of AD with adequate levels of accuracy (about 49 % of precision) [9, 78]. Several methods have been introduced to characterize AD connective tissue radiating from
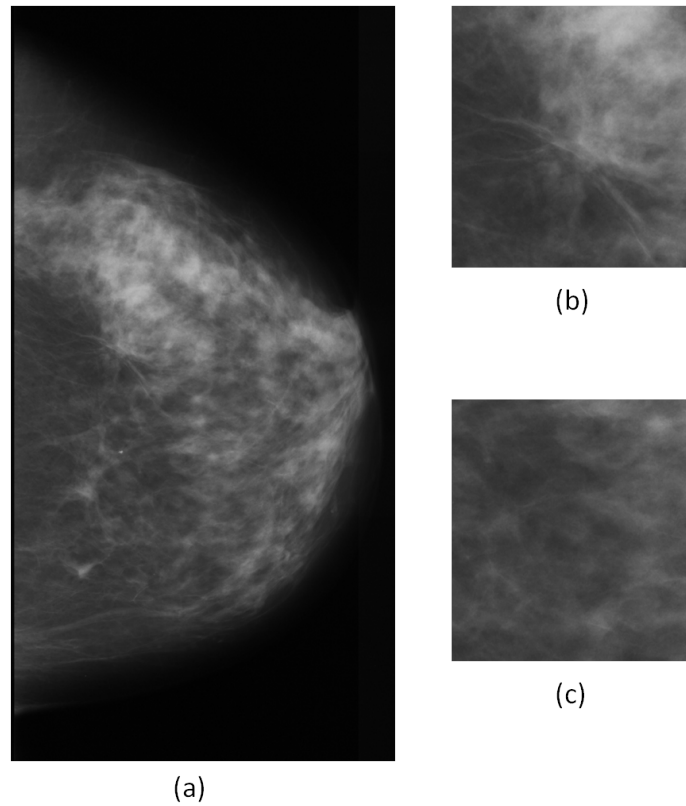
Figure 3-1.: a) Mammogram showing AD. b) A 128×128 AD ROI showing spiculations radiating from a central point. c) Normal 128×128 ROI tissue portion. The ROIs shown in (b) and (c) were both extracted from the mammogram shown in (a).

a center which have the appearance of radiolucent lines or stellate patterns. In these methods linear structures are enhanced to highlight information related to orientation, scale, or strength of lines, facilitating the characterization process. Several techniques have detected areas with potential stellate patterns, among others linear detectors [135], steerable filters [105], oriented field analysis by phase portraits maps [99], spicules enhancement by image transformation (Gabor and Radon spaces) [104, 8, 11] or fractal functions [40, 120]. Karssemeijer *et al.* [57] detected stellate patterns, including spiculated masses and AD using a multi scale method that classified the output of three-directional second order Gaussian operators and reported a sensitivity of about 90% with one false-positive per image (FP/image). Zwiggelaar *et al.* [134] evaluated four distinct methods for detecting and classifying linear structures in synthetic images. The reported area under the curve (A_z) using line operators, Orientated Bins, Gaussian Derivatives and Ridge Detectors were of 0.94, 0.91, 0.9 and 0.82, respectively. The type of linear structure detector and the parameter selection depends on prior knowledge of the lesion being imaged. Sampat *et al.* [105, 104] proposed a transformed image and radial spiculation filters in the Radon space to detect spiculated masses and AD. The reported sensitivity was of 80% and 14 FP/image for AD and the reported sensitivity was of 91% and 12 FP/image for spiculated masses. Guo *et al.* [40] applied five methods to estimate a Region of interest (ROI) of a Fractal Dimension that was used to train a

support vector machine classifier (SVM). The authors reported an A_z of 0.875 when differentiating masses and AD from the normal parenchyma. Ichikawa *et al.* [49] used a concentration index of linear structures obtained by the mean curvature of the image, resulting in a sensitivity of 68% and 3.4 FP/image. Nemoto *et al.* [85] proposed the likelihood of spiculation and a modified point convergence index weighted by the likelihood to enhance AD, establishing a sensitivity of 80% and 0.80 FP/image. Biswas *et al.* [15] introduced a generative model with a bank of filters that learned "texton" patterns using a subset of images from the Mammographic Image Analysis Society (mini-MIAS) dataset [111] (19 AD and 21 normal tissues). Evaluation was performed under a leave-one-out scheme, obtaining 3.6 FP/image and a sensitivity of 81.3% with a ROI radius of 5 mm (the minimum radius of an AD region in mini-MIAS). Kamra *et al.* [56] use a combination of Spatial Gray Level Co-occurrence Matrix (SGLCM), fractal and Fourier power spectrum based features to characterize the AD lesion, quantified in four directions $\theta = 0, 45, 90, 135$. This strategy was evaluated on a subset of ROIs from mini-MIAS [111] and the Digital Database for Screening Mammography (DDSM) [44] datasets with different size of ROI and a subset of fixed ones, the reported accuracy, sensitivity, specificity of the fixed ROI versions of DDSM was 92.9%, 93.3%, 92%, for mini-MIAS was 97.2%, 85.7%, 95.3% respectively. Matsubra *et al.* [72] applied directional and background filters to analyze the linear structure in eight directions. The detection sensitivity and the FP/image reported was 81% and 2.5 respectively. The dataset used in this study contains 174 AD and 580 normal diagnosed control cases. Several studies have used the phase portrait maps to capture the outward radiating linear structure, a strategy based on Gabor filters that highlight linear patterns at certain orientations [7, 96, 97, 11] with significant results. Banik *et al.* [11] detected the AD using the Gabor filters and phase portrait analysis. For each detected ROI, the fractal dimension, the entropy of the angular power spread and 14 Haralick's features were computed. Results showed an A_z of 0.76 with the Bayesian classifier, 0.75 with Fisher linear discriminant analysis, and 0.78 with a single-layer feed-forward neural network. These approaches face the AD detection problem as a classification task of ROIs of the entire mammogram. The major drawback is then the need of choosing a large number of parameters to construct linear detectors, Gabor filters, steerable filters, linear filters or features in both spatial and transform image domain to measure spicules. Additionally, the intended detection of a node-pattern, representing radiating-out linear structures often fails because the spicules, in general, do not fully describe a well-defined converging pattern but instead spicules mostly make an incomplete "star" shape with missing parts, making the texture orientation highly noise sensitive.

3.1.2. Contribution

Unlike previous approaches, this work is focused on the search of the linear saliency of a stellate pattern in the spatial domain, aiming to strength the classification process. The proposed Linear Saliency Domain method (LSD) represents a ROI as an (initially) fully connected graph with nodes corresponding to the ROI boundary and edge weights calculated using line integrals of the intensities along the path connecting any pair of nodes. By doing this, the edges corresponding to the most salient lines are assigned a larger weight in the graph. A centrality measure based on the graph's adjacency matrix's eigenvectors is used to identify the most relevant nodes. A further dimensionality reduction is accomplished by selecting only the most salient edge origi-

nating from relevant nodes. The set of salient lines is then assembled as a feature vector used to train an SVM classifier. Experimental results in two benchmark databases, the DDSM and mini-MIAS databases, demonstrate that LSD outperforms the baseline techniques [7, 96] in accuracy and speed with a precision rate of 89% and 87%, a sensitivity rate of 85% and 95%, a specificity rate of 93% and 84% respectively.

The rest of this chapter is organized as follows: Section 3.2 explains in detail the methodology and steps of the LSD. In Section 3.3 the datasets used to test the method are presented. In section 3.4, results of the validation experimentation, are shown. Section 3.5 presents the discussion, conclusions and possible future work directions are explored.

3.2. Method

The LSD method consists of five consecutive steps: Image Enhancement, Graph Construction, Central Node Detection, Detection of Salient Edges and Classification. The pipeline of the whole process is presented in Figure 3-2 and each step is explained in detail in the following sub-sections.
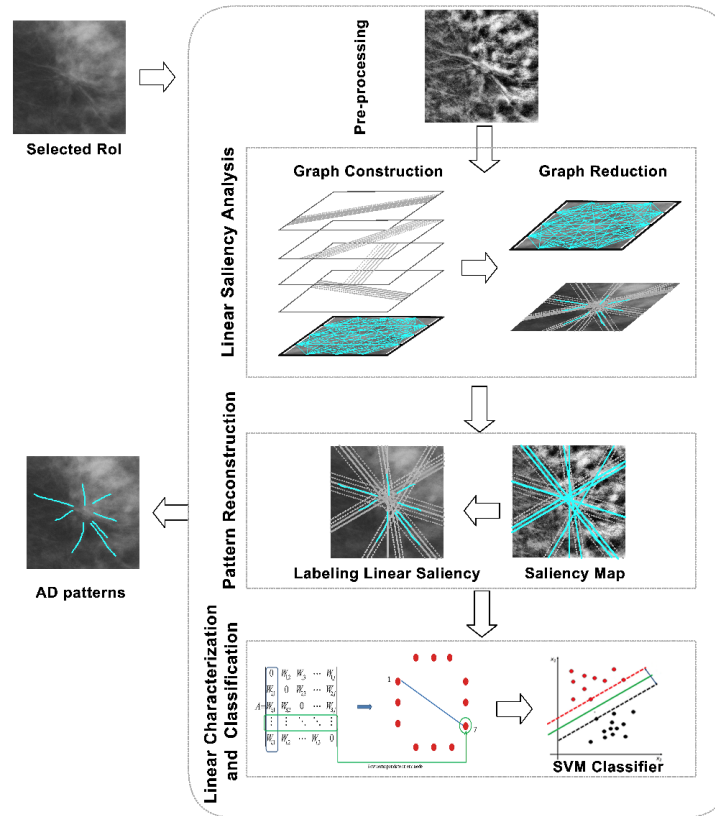
Figure 3-2.: Pipeline of LSD.

3.2.1. Image enhancement

The ROIs of the images selected as input for LSD are smoothed using a non-linear median filter [17] aiming to remove the image noise. Spicular details are enhanced by mapping each image's dynamical range to the maximum and minimum gray level values of the interval ([0, 255]), followed by a contrast limited adaptive histogram equalization approach (CLAHE) [93].

3.2.2. Graph Construction

The perimeter of a squared ROI is divided into 32 segments per border for a total of $N = 4 \times 32 = 128$ segments. The segments' width, 4 pixels in the mini-MIAS database and 32 pixels in the DDSM database, were chosen to approach the typical width of a single spicula, *i.e.* approximately $200 \, \mu m$ [8]. A graph $G(V, E)$ is built with the N segments corresponding to nodes V. The weights of the edges E are computed as the integral of the ROI intensities along the lines connecting every pair of nodes, *i.e.* the value of the integral between node V_i and node V_j corresponds to edge $E_{ij} = E_{ji}$. Notice that when nodes lie on the same border, *e.g.* two nodes on the top border, the thickness of the line and the integral result are both zero, making G is a partially connected graph.

Graph relations are summarized as an adjacency matrix A, where each position $[i, j]$ corresponds to the weight $w_{ij} = w_{ji}$ of the edge between two vertices (V_i, V_j). An illustrated example of the construction of the graph may be seen in Figure 3-3 and its adjacency matrix is shown in Equation 3-1.
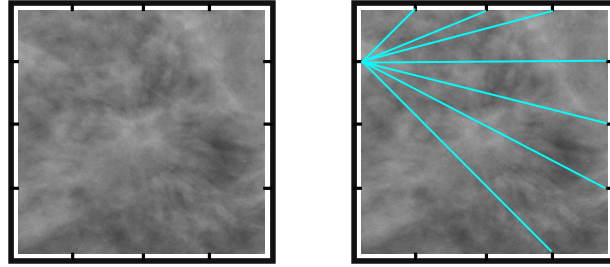
Figure 3-3.: The ROI boundary is divided into N segments as illustrated in the left panel. The center of each segment corresponds to a node V_i. The weight w of the edges connecting any pair of nodes corresponds to the value of the line integrals.

$$A = \begin{pmatrix} 0 & w_{1,2} & w_{1,3} & \cdots & w_{1,j} \\ w_{2,1} & 0 & w_{2,3} & \cdots & w_{2,j} \\ w_{3,1} & w_{3,2} & 0 & \cdots & w_{3,j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{i,1} & w_{i,2} & w_{i,3} & \vdots & 0 \end{pmatrix} \qquad (3\text{-}1)$$

3.2.3. Central Node Detection

A Markov chain may be associated to the graph $G(V, E)$ by normalizing the weights of the edges. Nodes can then be interpreted as the states of the Markovian process, while the weights are equivalent to transition probabilities. Under the appropriate assumptions, such as irreducibility due to strong connectivity, this chain tends to a stationary probability distribution that represents the frequency of node visits that a random walker would make if allowed to walk forever. In the present problem, this probability distribution would naturally be concentrated at nodes that have higher linear intensities since transitions into subgraphs containing them is more likely.

This type of representation has been shown to be equivalent to the eigenvector centrality analysis described in [34], where the node importance is calculated in a self-referential way, *i.e.* a node's importance is larger if it is adjacent to other important nodes. Different artificial vision strategies have represented the image relationships by a graph structure. Gopalakrishnan *et al.* [38], for example, have proposed that Markov random walks may represent the global or local image properties by performing the centrality analysis on either the complete or sparse k-regular graphs. From this information, the most salient foreground and background objects may be set and labeled.

The most important nodes may be extracted from the eigenvectors sorted using the eigenvalues of A. The first eigenvector represents the steady state of the associated induced Markov chain and includes the nodes with the most important connections. Subsequently, the nth eigenvector corresponds to the most important nodes except those in the mth eigenvector, for all $m < n$.

In each eigenvector the largest element $\overrightarrow{V}_k$ corresponds to the position of the most relevant node v_m :

$$v_m = max(\overrightarrow{V}_k) \quad \forall k \in [1, \ldots, N]. \tag{3-2}$$

3.2.4. Detection of salient edges

Once the central nodes have been detected, it is necessary to select the most salient edges that radiate from them. This is done by locating the highest weights in the original A adjacency matrix. As stated before, higher values correspond to the more radiolucent lines. In particular, each relevant node stands for a column of the adjacency matrix A, as illustrated in Figure 3-4.

The highest value of the column corresponds to the most important connection of that node, as shown in the Figure 3-5. This connection is then the most salient pattern for the selected central node.

ROIs of 128×128 pixels, such as the mini-MIAS database, correspond to a graph G of 32×32 nodes with at least 32 lines representing the saliency for any of the possible directions, *i.e.*, a particular side is connected with each of three possible sides of the ROI.

3.2.5. Classification

The binary AD classification problem is approached by an SVM classifier [24]. This strategy is applied to datasets with complex separation boundaries between classes and is based on maximizing
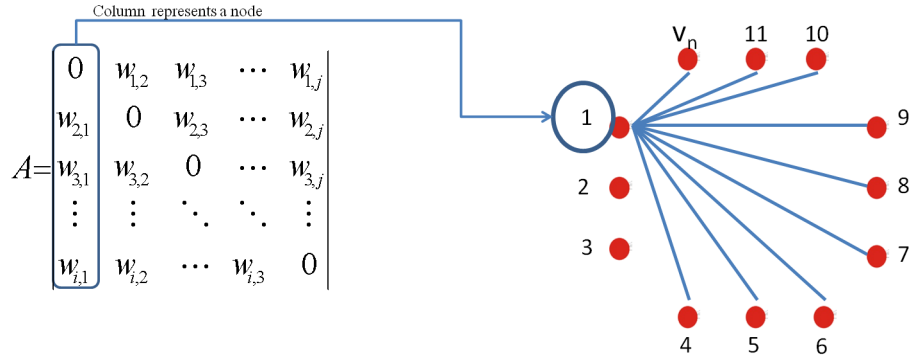
Figure 3-4.: In the adjacency matrix A each column corresponds to a node in the graph and its connections.
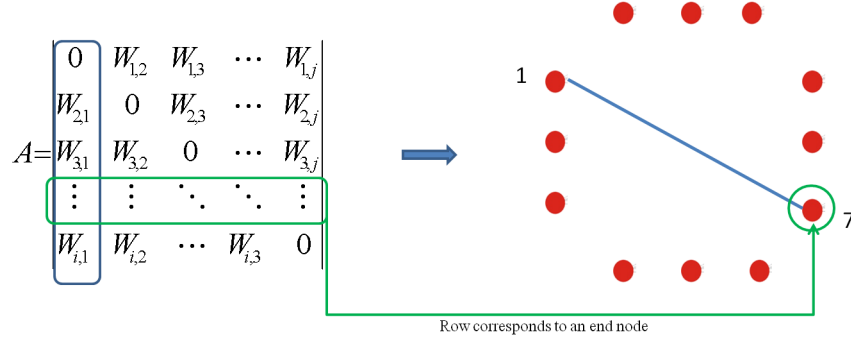
Figure 3-5.: The highest value of the nth column corresponds to the most salient edge radiating from the nth node.

the separating margin between the two classes in an alternative space in which data relationships are linear. In this case two classes are defined: normal breast tissue and ADs [40].

The proposed method trains the classifier model using radial basis functions (RBF) as a kernel. Positive values correspond to AD and negative values to control cases. The optimal hyperplane is found by solving the discriminating function:

$$f(x) = sign\left(\sum_{i=1}^{n} y_i a_j \phi^T(x_i).\phi(x_j) \right) \tag{3-3}$$

where sets of labels y_i module the Lagrange multipliers a_j and the kernel function $K(x_i, x_j)$ is a similarity measure defined by computing the inner-product between the feature vectors as

$$\phi^T(x_i).\phi(x_j) = e^{-\left(\frac{1}{2\sigma^2}\|x-x_i\|^2\right)}. \tag{3-4}$$

The kernel parameters were optimized by cross validating on the synthetic data ($\gamma = \frac{1}{5128} = 0.00019$ and $\sigma = 1$).

3.3. Datasets

3.3.1. Clinical cases

Each AD ROI corresponded to a window with the size and location annotated in the database. In contrast, the reference tissue or control case was selected by and expert radiologist who manually set a window within the normal fibroglandular tissue, with a similar size to the the marked AD regions. Two public mammography databases were used to test the method: DDSM and mini-MIAS. A ROI is the smallest square region containing the complete lesion, 128×128 for the mini-MIAS database and 1024×1024 for the DDSM collection.

- The DDSM database: the Digital Database for Screening Mammography database (DDSM) [44] is a mammography database widely used as an evaluation benchmark [79, 45, 131]. This open access database is constituted of digitized images of mammographic films with the corresponding technical and clinical information. The whole DDSM database contains a total of 2620 cases, each including four images obtained from Cranio-Caudal (CC) and Mediolateral-Oblique (MLO) views as well as specific BI-RADS descriptions that were annotated by expert radiologists according to BI-RADS fourth edition. In this work, a subset of mammographies were chosen, the inclusion criteria were: 1) similar image quality, 2) balanced number of ROIs (with normal tissues and with Architectural Distortion), 3) only a single lesion was assessed, i.e., this study excluded masses, calcifications or cases with the presence of clips or pencil marks in the lesion area. A subset of selected mammographies were digitized either with Lumisys laser film scanner at 50 μm or a Howtek scanner at 43.5 μm pixel resolution and at a dynamic range of intensities of 2^{12}= 4096 gray level tones. The AD location and size, associated to each case, were used to manually crop squared sub-images centered at the lesion. Specifically, RoIs were cropped to the bounding box of the lesions and rescaled to 1024×1024 pixels preserving the aspect ratio when either width or height were greater than 1024. Otherwise the lesion is centered without scaling and preserving the background tissues. In consequence, a set of 246 ROIs were extracted from this DDSM database, distributed as 123 ROIs with normal tissue and 123 ROIs with Architectural Distortion.

- The mini-MIAS database: a collection published by the Mammography Image Analysis Society (MIAS),[1], provides digitized mammograms from a screened population. The X-ray films in the database have been carefully selected from the United Kingdom National Breast Screening Programme and digitized with a Joyce-Lobel scanning microdensitometer to a resolution of 50 μm^2 and representing each pixel with 8-bit depth. The database contains left and right breast images from 161 patients, and radiologist's truth-markings. For each film, experienced radiologists give the type, location, scale, and other useful information. The database provides a metadata file describing the lesions as well as their location and size. In consequence, a set of 19 ROIs with Architecture Distortion were extracted as well as 19 normal RoIs. The AD location and size were taken from the metadata file. The cropping process was carried out as described for the DDSM database, except that the size was 128×128 pixels.

---

[1]http://peipa.essex.ac.uk/info/mias.html

3.3.2. Synthetic Data

A set of synthetic images was constructed to assess and fine tune the method's parameters. These digital phantoms aim to emulate actual AD and therefore must keep the radiated pattern of linear structures around a focus. The method implemented was presented by Parr *et al.* [92] and consists of finding a representative element of a series of artificially constructed radiating patterns. Briefly, the method splits a square image of 81×81 pixels into tiles of 9×9 that serve to focus radiated patterns starting from the center of each tile. For a particular centered pattern, the same 9×9 representation grid is used but this time each tile stores the average angle of every line crossing the tile. From these images, the obtained 81 radiating patterns storing the mean angle for each tile, are then vectorized and used to feed a Principal Component Analysis strategy (PCA). The main twelve eigenvectors of the covariance matrix, that explain a $90\,\%$ of the variability, are set as the statistical representative elements. Afterwards, the weights of the six vectors that approximate 80% of the variance are retrieved by simply subtracting the mean to the observation and projecting the result onto the matrix of twelve eigenvectors.

The weights of the principal directions b_i are found then by centering the data x_i and projecting onto the principal direction matrix P, as stated in Equation 3-5 and illustrated in Figure **3-6**

$$b_i = P^T(x_i - \mu) \tag{3-5}$$

$$
\begin{array}{ccc}
\text{(a)} & \text{(b)} & \text{(c)}
\end{array}
$$
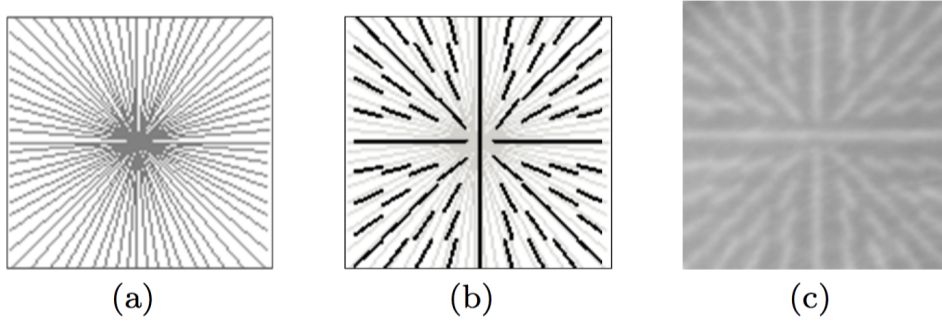
Figure 3-6.: a) Artificially generated radiating pattern. b) Angular representation of a radiating pattern. c) Final focal radiating phantom.

Different linear patterns can be obtained from this representation by simply de-centering the data by one or two standard deviations, as formulated in Equation 3-6 and illustrated in Figure **3-7**. The final phantom consists of a pattern superimposed to an actual mammogram region diagnosed as normal. The dataset is composed of 162 phantoms, namely 81 focal that are de-centered as previously explained by one standard deviation to yield the 81 non-focal patterns.

$$b_i = P^T(x_i - \mu + s) \tag{3-6}$$
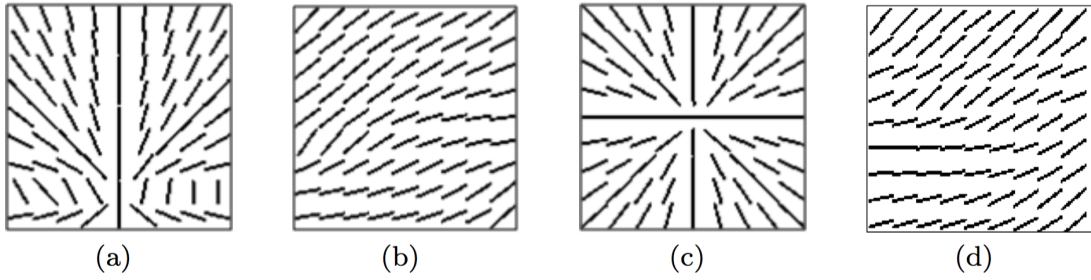
Figure **3-7**.: Effects of de-centering the focal line pattern

3.3.3. Experimental Setup

The performance of the LSD method was tested using series of normal and AD ROI either synthetically constructed or extracted from the two public datasets described above. Classification followed an SVM strategy with an RBF kernel using the LIBSVM library implementation [23]. The strategy herein implemented consists of 128 nodes that result in 32 salient lines from the graph introduced in previous section. The different metrics used to evaluate the results were: $accuracy = \frac{TP+TN}{TP+FP+FN+TN}$, $sensitivity = \frac{TP}{TP+FP}$ and $specificity = \frac{TN}{TN+FN}$, being TP the number of True Positives, TN the true negatives, FP the false positives and FN the false negatives, respectively. Performance is also evaluated using the ROC curves generated by setting thresholds as the membership.

3.4. Results

Comparison with baseline methods was performed by implementing two different strategies:

- Firstly, the classic approach proposed by Ayres and Rangayyan [7]: This seminal work was included as a baseline as it is, to the best of our knowledge, the earliest example of an AD detection method in the Gabor space.

- Secondly, the strategy proposed by Banik *et al.* [11] which analyzes both the node maps of the phase portraits and Haralick texture descriptors from the Gabor magnitude.

These two representations are based on projections to the Gabor space, but while Ayres and Rangayyan use the phase space orientation map, from which different features are extracted, Banik *et al.* include also some characteristics of the phase space magnitude. It is worth mentioning that these methods were implemented using exclusively the frequency and texture features reported by the authors to be significative.

3.4.1. Experiments with Synthetic data

As introduced in section 3.3.2, synthetic data emulates different AD focal and non-focal oriented patterns, corresponding to the pathological and control cases, superimposed to a regular distribution of normal tissue, as illustrated in Figure 3-8. This background was chosen by an expert

from a case which was randomly selected from the mini-MIAS set of control cases. The obtained database is composed of 162 ROIs, 81 focal and 81 no focal oriented patterns.
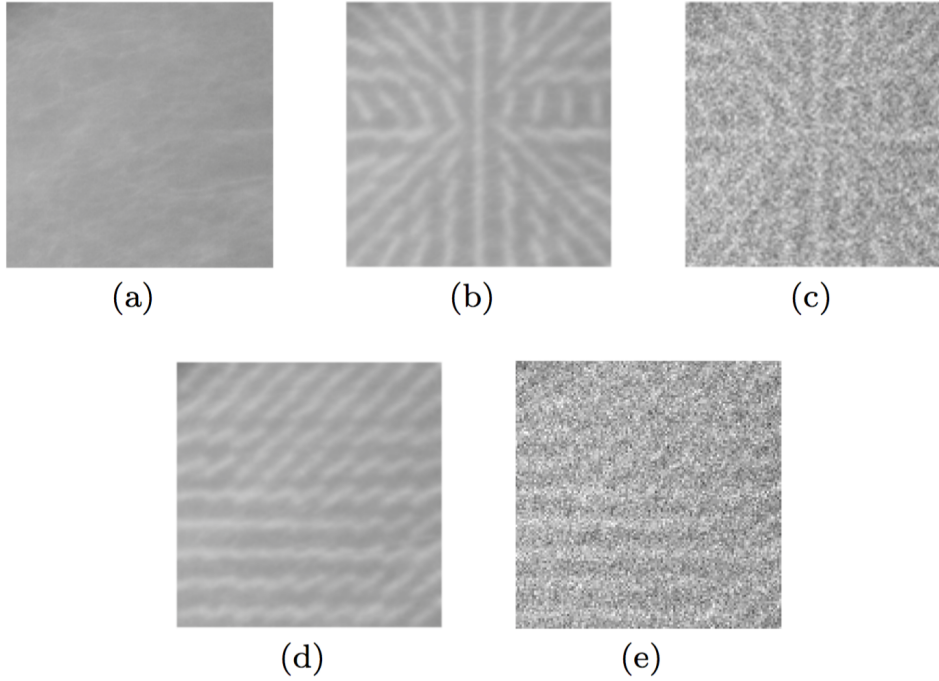
Figure 3-8.: Different phantom patterns. (a) background from a control MIAS Mammogram. (b) focal AD phantom. (c) focal AD phantom with Gaussian noise. (d) Non-focal regular AD phantom. (e) Non-focal regular AD phantom with Gaussian noise. Gaussian noise was set to $\mu = 0$ and $\sigma = 0.05$.

This evaluation tested not only the classification using pure phantoms, but also the quality of the result obtained when the phantom was contaminated with different levels of Gaussian noise. The added noise emulates the type of noise resulting from the contribution of many independent sources and associated to tomography and radiographic images [68]. The Gaussian noise was set to $\mu = 0$ and $\sigma = 0.05$, a level for which the peak signal to noise ratio (PSNR) between the original image and the degraded one was of 19.6 dB. A total of 162 phantoms (81 focal and 81 regular patterns) were synthetically generated, and each of these subsets was randomly sorted out. These two sets were then split to a, roughly, 70 to 30 training to testing cross-validation ratio with 57 of each group used for training and the remaining 24 of each group used for evaluation. A first test was made on the pure phantoms and a second one over the original patterns corrupted with the Gaussian noise. This classification task was carried out using a conventional SVM classification with an RBF kernel. Comparative results obtained for the two baseline methods and LSD, using both the original and corrupted patterns, are shown in Table **3-1**.

The baseline method proposed by Banik *et al.* [11], for the original test P shows a perfect performance for the set of generated patterns, demonstrating its effectiveness for detection of focal patterns. Likewise, LSD showed an accuracy of 98%, a sensitivity of 100% and a specificity of 96%, evidencing a comparable performance with these simulated lesions. These two methods outper-

	Acc		Sens		Spec		A_z	
	P	P_G	P	P_G	P	P_G	P	P_G
LSD	0.98	0.92	1.0	0.87	0.96	0.96	1.0	0.96
Ayres *et al.*	0.54	0.48	0.66	0.33	0.42	0.62	0.52	0.52
Banik *et al.*	1.0	0.5	1.0	0.75	1.0	0.25	1.0	0.5

Table 3-1.: Comparative results of the classification with synthetic images for each evaluated method. Column P shows the results for the initial synthetic image and P_G for the Gaussian corrupted phantom test sets.

form the implementation of the Ayres *et al.* proposal [7]. In the corrupted phantoms tests LSD outperformed the two baseline methods showing comparable figures to what was observed with the uncorrupted patterns. The baseline methods completely lose the discriminative power shown with the uncorrupted dataset.

Figure **3-9** presents the obtained results as ROC curves that depict the classification in terms of $sensitivity$ vs $1 - specificity$. Panel (a) in Figure **3-9** shows a perfect performance for the proposed method and the baseline proposed by Banik, yielding an A_z of about 1.0 in both case. In contrast the classical Ayres method shows a very low performance, with an A_z of 0.52. Interestingly, when a Gaussian noise is added to the synthetic phantoms, the proposed approach results remarkably more resistant than the baseline approaches, as illustrated in panel (b) of Figure **3-9**. The A_z of the tested methods were LSD= 0.96, Ayres= 0.52 and Banik= 0.5.

(a)   (b)
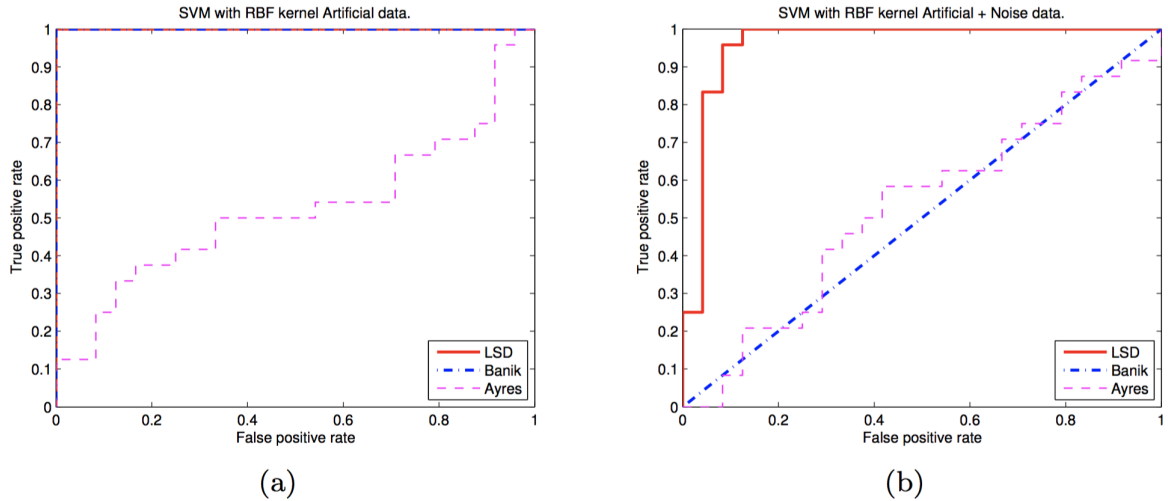
Figure 3-9.: ROC curves after the evaluation of each method for both experimental groups. Results of the original synthetic images are shown in panel (a) and the results of the corrupted ones are shown in panel (b).

3.4.2. Experiments with DDSM and mini-MIAS Databases

Given the small number of AD cases (19) available in the mini-MIAS database, cross-validation tests were carried out using a leave-one-out scheme.

The baseline method proposed by Banik *et al*. showed an accuracy of 89%, a sensitivity of 84% and a specificity of 89%, evidencing a good performance for the mini-MIAS set. The proposed method showed for the same database an accuracy of 87%, a sensitivity of 95% and a specificity of 84%, showing a performance comparable to Banik's method. These two methods strongly outperforms the Ayres *et al.* strategy. Results are shown in Table **3-2**.

	Acc		**Sens**		**Spec**	
	DDSM	*mini-MIAS*	*DDSM*	*mini-MIAS*	*DDSM*	*mini-MIAS*
LSD	0.89 ± 0.07	0.87 ± 0.09	0.85 ± 0.13	0.95 ± 0.12	0.93 ± 0.13	0.84 ± 0.1
Ayres *et al.*	0.59 ± 0.13	0.65 ± 0.13	0.33 ± 0.15	0.68 ± 0.13	0.68 ± 0.14	0.63 ± 0.15
Banik *et al.*	0.76 ± 0.11	0.89 ± 0.12	0.75 ± 0.13	0.84 ± 0.13	0.76 ± 0.15	0.89 ± 0.14

Table 3-2.: Comparative classification results with the DDSM and mini-MIAS datasets for each evaluated method. Accuracy, Sensitivity and Specificity are reported.

The number of AD cases in the DDSM dataset allowed a k-fold cross validation scheme with k=10 folds to be used. In this test, LSD showed an accuracy of 89%, a sensitivity of 85% and a specificity of 93%: a much better performance than the one observed with Banik *et al.* that yielded an accuracy of 76%, a sensitivity of 75% and a specificity of 76%. The classical method proposed by Ayres *et al.* showed an even lower performance.

Figure **3-10** presents the results of SVM classification as a ROC curves. Panel (a) shows a high performance on the DDSM for LSD, with an A_z of 0.93 while the baseline proposed by Banik *et al.* yield 0.83 and the classical method proposed by Ayres *et al.* show an A_z of 0.55.

3.5. Discussion and Conclusions

This work introduced a new form of characterizing AD on mammograms by representing ROI tissue distribution as a graph that captures the salient directions of linear structures within a particular ROI. The AD saliency is highlighted using a graph structure that stores the intensity linear information in the weight of the edges connecting the borders of a ROI. This external border undergoes a dyadic partition that is driven by the maximal width of an AD spicula, in the present investigation four pixels for the mini-MIAS database and 32 for the DDSM.

The intrinsic AD characteristics contained in the graph are quantified using the spectral decomposition of the adjacency matrix. This allows the identification of linear clusters, represented by the set of ordered eigenvectors, where each eigenvector stands for a different degree of relevance or saliency. The final descriptor is constructed by concatenating the lines associated to the most important nodes, *i.e.*, the largest value of each eigenvector.

This descriptor was assessed and compared against two baselines, outperforming both a classic and a state-of-the-art approach. The former, proposed by Ayres and Rangayyan, set a precedent by using directional and distribution information from curvilinear structures present in the breast
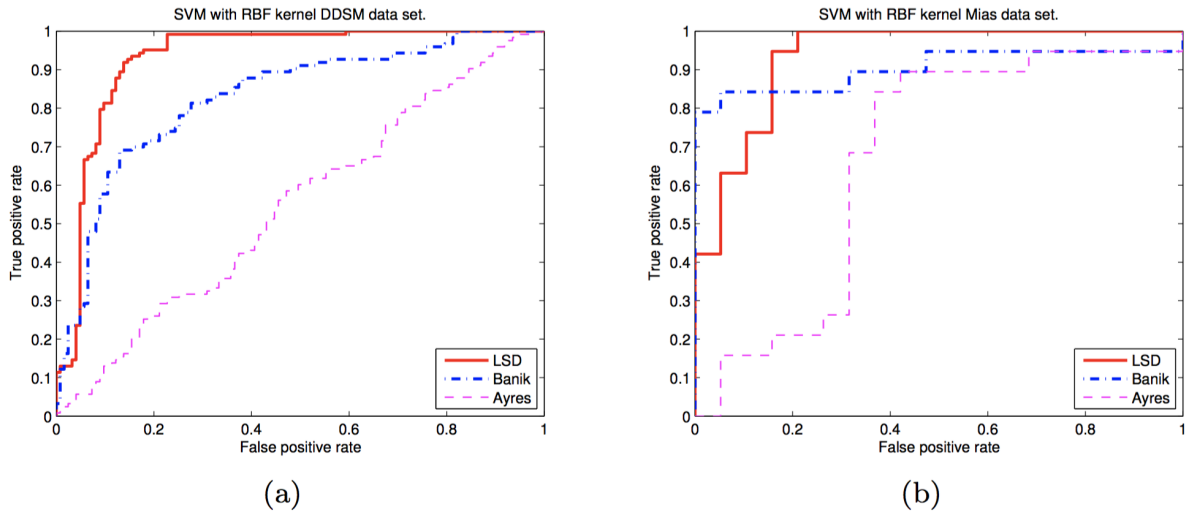
Figure 3-10.: (a) ROC curves for the DDSM and mini-MIAS databases obtained from a classical 10-fold cross validation for the set of ROIs. (b) The ROC curve corresponds to a leave one out scheme.

parenchyma. The latter is a recent improvement proposed by Banik *et al.* uses node map information combined with texture features such as Haralick's or the Fractal Dimension. The presented work was compared against this method using Haralick texture features because is a feature being commonly and widely used.

Results using synthetic images, which can be viewed as an "idealized case" of an AD, showed almost perfect classification results. This indicates that, even though it does not explicitly try to extract linear or spiculated patterns, LSD does highlight the image information permitting a consistent detection under ideal conditions. The main advantage of the method, however, is its robustness to noise. This is evidenced by its continued good performance under noisy conditions, opposed to the notable drop in performance of the method proposed by Banik. The method proposed by Ayres shows a very low discriminative power which is only slightly better than chance when noise is added (A_z=0.52).

When evaluated with open databases, such as DDSM, the robustness of LSD is evidenced by showing similar results to the synthetic datasets and outperforming both the Banik and the Ayres methods. Although the features proposed by Banik are more rapidly learned, they are unable to cope with the larger biological variability introduced in the dataset.

As expected, the results are less concluding in case of the mini-MIAS database, probably due to a much smaller population (barely nineteen cases). Still, for the mini-MIAS basically the Banik strategy and LSD are equivalent, with very similar A_z (LSD=0.93 and Banik= 0.9).

Future work includes evaluation of the whole mammography as a screening approach. Interestingly, the graph structure might be used to compare different lesions in topological terms and to establish a similarity metric between lesions.

It should be noted that the method is dependent on parameters such as the cross sectional width of a typical spicule. This is a possible source of error that might be mitigated by fine tuning this

parameter for a particular set of mammograms, *e.g.* by varying this number of pixels between nodes. This limitation is clearly common to any method that attempts to search spiculated lesions [135, 104, 8, 11, 57, 85, 56, 72].

4. Conclusions and Perspectives

This thesis has addressed the problem of characterizing breast abnormality patterns contained on regions of interest in mammographies. Specifically, we have focused on on identifying and describing the most important radiological signs associated to breast cancer, namely masses and architectural distortion, aiming to improve the diagnosis accuracy of mammographic findings, an important requirement for any support system to be part of an actual radiology workflow. The herein proposed strategies characterize visual patterns of these lesions and the relationship between them, to infer their clinical significance. Unlike conventional Computer Aided Diagnostic systems, the strategies proposed in this thesis do not intend to segment suspicious regions of interest. In contrast, they establish a detailed characterization of some mammographic findings according to standard descriptors used for the mammogram reports, defined in the BI-RADS (Breast Imaging Reporting and Data System) standard . BI-RADS is a complex and challenging interpretation task due to the high anatomical variability of these lesions when imaged with mammography.

Unlike classic approaches used in diagnosis support systems, the strategies proposed in this Thesis do not attempt a previous segmentation or delineation of mammographic findings for extracting visual features. They characterize the morphological and structural information of breast tissues from the whole region of interest by decomposing the image in different image levels and scales. Thus their performance does not depend on exact segmentation techniques, which is a difficult task due to the complex distribution of breast tissue and the presence of noise in this kind of diagnostic images. This is an important contribution of our approach in the area of mammography analysis. The exhaustive evaluation of the proposed methods has demonstrated a high accuracy level for classifying normal and abnormal patterns.

The proposed strategies constitute an innovative radiological tool in the context of mammographic evaluation: the representation of mammographic findings together with visual interpretation methods, to which machine learning techniques and content based image retrieval strategies have been added, providing efficient visual support of the mammographic findings in terms of visual features and leading to reduce the variability inter and intra-observer during the diagnosis process. In the particular case of breast mass analysis, the proposed strategies have been combined with a CBIR approach, which has allowed to include the expert's knowledge stored in different mammographies database with the actual workflow of the radiologist. This facilitates the interaction of new radiologists and their training process using retrieved cases already diagnosed in the image collection. This is an advantage during the decision-making step and it helps to reduce the rate of unnecessary biopsies in mammogram screening programs.

The strategies introduced in this thesis have led to an innovative perspective of computer aided

image interpretation for breast cancer analysis that can be considered as an important require-
ment of any traditional computer aided diagnosis system intended to be used in the actual clinical
practice. In particular, the proposed techniques have shown to be capable of supporting the mam-
mographic evaluation process, specially by identifying the shape, margin and density descriptors,
as well as oriented breast tissues. These descriptors may help to define the malignancy level of
lesions, the final BI-RADS categorization. Finally, the proposed strategies have shown remark-
able results when used on different mammography technologies, such as digital mammograms
and a digitized version of film-screen mammographies, confirming that our methods are easily
adapted to different image resolutions, a condition highly useful when used in actual mammog-
raphy screening programs and when adapted to current CAD systems.

4.0.1. Perspectives

This thesis was developed to contribute to the field of mammographic image analysis for sup-
porting the mammography evaluation process in screening programs. It aims to provide compu-
tational tools for an automatic interpretation and characterization of breast abnormality patterns
that serve as a support for breast cancer diagnosis, training of inexperienced radiologists and re-
duction of intra-observer variability. A full mammography evaluation allows to define the clinical
condition of the patient, to establish the adequate medical procedure and to improve current clin-
ical treatments, impacting directly the patient's quality-of-life. We expect that, in the future, the
proposed computational tools will contribute to improve the complete analysis workflow of the
radiologist. Meanwhile, further work must be performed to prepare these tools to be used by
radiologist experts in a daily work. Some of this work includes:

1. **Multi-view mammography analysis.** During a complete mammography evaluation in screen-
   ing programs, the radiologist correlates the visual information from different mammogra-
   phy projections (MLO and CC views), identifying the spatial location and size of the same
   breast lesion at each mammography, a process that allows to reduce the false positives rate
   and helps to establish the highest degree of malignancy of the lesion according to its most
   relevant visual properties from each views. Then, the final report is performed by suggest-
   ing some standard concepts, which are established in the BI-RADS lexicon. For instance,
   a mass is taken into account if it is visualized in at least two projections. Otherwise it is
   considered an architectural distortion. However, this associated visual information is useful
   to exactly establish the malignancy level of each detected lesion. Therefore, an extended
   approach based on our proposed strategies may be addressed by fusing the visual features
   extracted of the lesion under analysis from multi-views when a specific case is evaluated
   during a diagnosis session.

2. **Assignation of BI-RADS categorization**.

   Once the mammographic description has been carried out, the challenge is to determine a
   BI-RADS category (from 1 to 6 level), which defines the appropriate treatment, management
   and follow-up of patients during the screening programs. On the other hand, this classifi-
   cation task has been used to establish the probability level of malignancy of lesions (benign
   or malignant). However, in the particular case of breast masses, a problem emerges when

a BI-RADS category between 3 to 4 can be assigned (probably malignant or benign). We believe that the methods propose in this Thesis may be evaluated for this task, for which a more extensive mammogram database whit this specific annotation is required. This will objectively allow to support the decision-making, without the need of invasive procedures such as a biopsy.

3. **Developing a complete CAD system based on our strategies.**

   Computer-aided diagnosis systems are in permanent development for breast cancer diagnosis, constantly integrating new computational tools. However, currently, some systems lack the tools to interpret images according to the BI-RADS lexicon, breast tissue estimations, as well as the risk factors analysis (Breast density estimation). We believe that the strategies proposed in this work could be evaluated for breast density estimation, calcification description and bilateral asymmetry analysis, due to the results obtained when the tissue shape and margin were analyzed. Besides, these have been correctly adjusted to the different image technologies from real clinical cases.

4. **Characterization of lesion from multimodal information**. Finally, the proposed strategies might be extended to other image modalities such as ultrasound or magnetic resonance, that require lesion characterization in terms of shape, lesion margin and density.

A. Characterization of breast abnormalities patterns based on a sparse representation strategy

An application of sparse representations in a classification task was explored in the context of severity (benign or malign) identification for breast masses and architectural distortion. The proposed multiresolution descriptors (curvelets) were used as low-level features. Hence, a set of Regions of Interest (RoIs) were characterized by their projection onto learned malign and benign dictionaries, and the region class was identified using a decision rule algorithm. Those works were published in:

Breast Masses Classification using a Sparse Representation

Fabián Narváez, Andrea Rueda and Eduardo Romero

BioIngenium Research Group
Universidad Nacional de Colombia, Bogotá, Colombia

Abstract. Breast mass detection and classification in mammograms is considered a very difficult task in medical image analysis. In this paper, we present a novel approach for classification of masses in digital mammograms according with their severity (benign or malign). Unlike other approaches, we do not segment masses but instead, we attempt to describe entire regions of interest (RoIs) based on a sparse representation. A set of patches selected by a radiologist in a RoI are characterized by their projection onto learned dictionaries, constructed previously from classified regions. Finally, the region class was identified using a decision rule algorithm. The strategy was assessed in a set of 80 masses with different shapes extracted from the DDSM database. The classification was compared with a ground truth already provided in the data base, showing an average accuracy rate of 70%.

1 Introduction

Breast cancer is the most frequent disease in women and is considered as the largest public health problem in women population [1]. This disease is fully curable if diagnosis is achieved early and mammography is the more efficient method for visualizing abnormalities in the very early stages [17, 4]. However, mammographic interpretation is really hard and there exist studies showing that between 10% and 25% of breast cancer are not detected in mammography [3]. Abnormal lesions that are directly related to the presence of breast cancer are masses and calcifications. In clinical practices, a final diagnosis is determined by pathological analysis of abnormal lesions, an invasive procedure well known as biopsy. In order to reduce innecessary biopsies and interpretation variability between radiologists, the American College of Radiology diffused the *Breast and Imaging Report and Database System* (BI-RADS) as a classification standard to reporting breast lesions, which allows to classify different pathologies as well as their severity [2]. This standard established a basic classification for masses based on their shape, margin and density, which usually correspond to low level descriptors, and the severity level is defined a semantic interpretation of the first two features. In real clinical scenario, the radiologist identified the severity level of masses by visual features analysis, as circumbscribed margin of lesions, which are compact and lobular or circular shaped, and spiculated margin of lesions, which consist of a central mass with radiating spicules in some or many directions. Therefore , edge and shape information of mass defined a severity level (malign or benign lesion).

Actually, Computer Aided Detection (CAD) and Diagnosis for mammography has decreased innecessary biopsy practice and variability effects since the radiologist can have a support for their diagnosis [16, 18, 15], becoming a well accepted clinical practice to assist radiologists interpreting mammograms, when they search and identify micro-calcification clusters [12]. However, the relatively low performance of CAD schemes in mass detection [7] make them less accepted as mass diagnosis tools. Two main factors makes breast mass detection in mammograms a very difficult task in medical image analysis. Firstly, there is a large variation in the appearance of both normal breast tissue and cancerous tissue [6]. Secondly, CAD systems are usually based on automatic detection and segmentation of abnormal lesions, issue that increases the false positive rate. As an alternative to overcome these difficulties, interactive CAD systems have been developed [18]. Given a query lesion, these systems identify other similar mass lessions in a large database, which are eventually clinically relevant to the actual one, allowing to provide a suggestion to the specialist in diagnosis tasks. On the other hand, CBIR-based CAD schemes [16] have the potential to provide radiologist with visual aid and increase their confidence in accepting CAD-cued results in the decision making process. In a recent work, we have proposed an interactive CAD system that provides a BI-RADS mass description of a manually selected region of interest (RoI) by region-based descriptors [13].

In this paper, a new approach for breast mass classification from a set of regions of interest (RoIs) is proposed. A set of image patches are extracted from previously classified RoIs and then characterized using a multi-scale edge analysis to project them in a feature space, using a sparse representation. This process allows to identify feature clusters that corresponds to the severity of the masses (malign or benign). Finally, a new RoI can be classified by projecting some patches in the feature space and analyzing their relationships with the severity clusters. This strategy was assessed in a set of 80 masses with different shapes extracted from the DDSM database, where 30 benign and 30 malign masses were used as the training set and the remaining 20 masses were selected for testing. The classification was compared with a ground truth already available in the data base, showing an accuracy rate of 70%.

2 Methodology

The proposed method for classification of breast masses can be roughly divided in two stages: an offline learning process and an online classification procedure. At the offline learning process, the main goal is to identify the feature vectors that characterize each selected class, in this case, malign and benign masses. For doing so, three different tasks are involved in this process. First, two different sets of RoIs with benign and malign masses are selected by a radiologist and preprocessed to enhance the mass shape characteristics. As the class characterization process will be based in a sparse representation, the next step in the learning process includes to construct malign and benign severity dictionaries. Then, an image patch dictionary is constructed for each selected class by randomly sampling patches from malign and benign RoIs, which are thus characterized with a multi-scale edge analysis. Finally, a new set of relevant patches (that capture edge and background information) are manually selected at each RoI (Figure
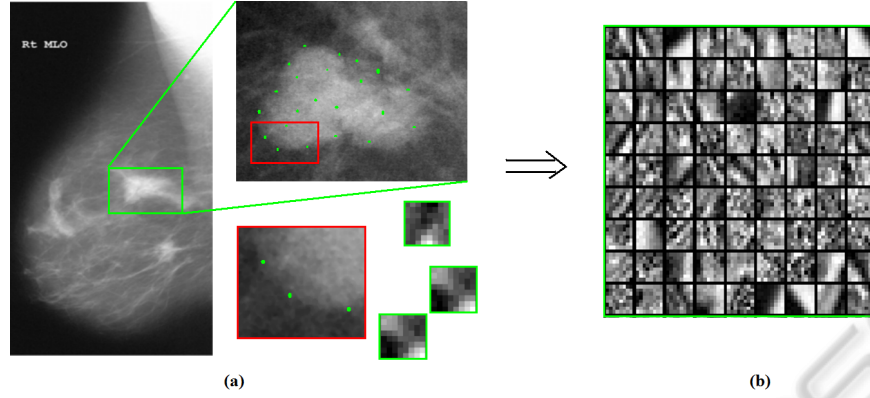
Fig. 1. Point selections by radiologist and sparse representation. (a) illustrates a manual selection of the patches as points of interest. (b) illustrates the dictionary formed by the selection of patches.

1(a)) and then characterized with a sparse representation, by using its projection onto the malign and benign severity dictionaries previously constructed (Figure 1(b)).

The online classification procedure takes place when the severity of a test RoI needs to be defined. The radiologist manually select a set of patches on the test RoI, which are individually classified using again a sparse representation. Each patch is then projected onto the malign and benign severity dictionaries, and these projections are then compared with the characterizations previously obtained for the training patches. This process, followed by a decision rule, allows to establish the membership of the entire RoI.

2.1 RoI Pre-processing

Mammography analysis generally must deal with regions difficult to interpret [6], since they are associated to hard acquisition conditions. In most cases, diagnostic characteristics, such as mass edges, are small and have low contrast with respect to the surrounding breast tissues. To improve the particular region characteristics and to highlight the grey level intensity information, a preprocessing stage was carried out on each RoI. A contrast enhancement method is then used based on mean and standard deviation information of each RoI, allowing to stretch the maximum and minimum gray levels to the interval $[0, 255]$. With this procedure shape features are improved, while preserving edge details. Finally, the whole region is smoothed using a median filter [19].

2.2 Dictionary Construction

The next step is to build dictionaries $\mathbf{D}_m$ and $\mathbf{D}_b$ for malign and benign masses, respectively, as arrays of patches (atoms). Such an approach has been successfully used for image classification [9]. We selected a set of N RoIs with different mass shapes, according to their level of severity (malign or benign) as training RoIs. First, a set of K random patches per RoI were selected and then characterized using a multi-scale

edge analysis. This analysis attempts to describe the tissues present in mammographic images in terms of edge and background information. We used the 3×3 and 5×5 Sobel kernels, applied in the horizontal and vertical directions, and concatenated as a single feature vector. Finally, this vectors are stored as columns of the matrices $\mathbf{D}_m$ and $\mathbf{D}_b$, leading to 2 different dictionaries that represent the mass severity, one for malign and one for benign masses. This process is illustrated for the benign dictionary in Figure 2.
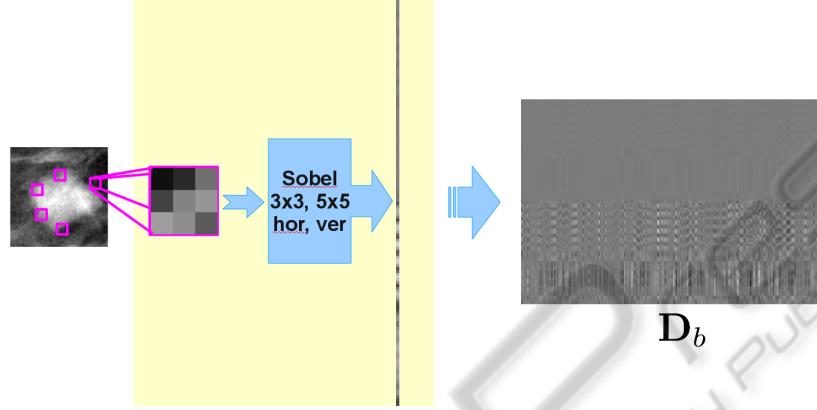
Fig. 2. Construction of a feature dictionary for benign masses.

2.3 Sparse Representation and Characterization

Once the dictionaries are built, the main goal is to identify the set of feature vectors that characterizes the benign and malign classes. Therefore, a new set of patches selected from the training images are projected onto the previously constructed dictionaries, following a sparse representation. The coefficients of the projection will be used to place each patch on a feature space, where each class will be defined as clouds of feature points.

Sparse representation techniques allows to identify the constituent parts of a scene and then, using some of them, the same scene or similar ones may be accurately reconstructed. These parts, denoted as basis functions or patches (atoms), are usually arranged in overcomplete dictionaries with a larger number of elements than the effective dimensionality of input space, thereby representing a wider range of image phenomena [14, 11]. Formally, consider a $n \times m$ matrix $\mathbf{D}$, where each column is a possible image in $\mathbb{R}^n$ (atomic images), a dictionary of patches. The projection of an image x onto the space spanned by $\mathbf{D}$ yields a weighting vector α ($x = \mathbf{D}\alpha$). Furthermore, if α is sparse (with $k_0 \ll m$ nonzeros), this produces a linear combination of k_0 patches with varying weights. To find the adequate α, we need to solve the optimization problem denoted as $\mathcal{G}_1(\mathbf{D}, x, \lambda)$, which has the form

$$\mathcal{G}_1(\mathbf{D}, x, \lambda): \min_{\alpha} \lambda \|\alpha\|_1^1 + \frac{1}{2}\|x - \mathbf{D}\alpha\|_2^2$$

The solution of this problem consist in finding the sparsest vector α that weights x as

a linear combination of patches from $\mathbf{D}$, using the norm ℓ_1 as a measure of sparsity. Different approximation methods to solve this problem have been recently proposed, detailed descriptions and references can be found in [5]

This process, applied independently to the benign and malign RoI sets, delivers a set of representation coefficients per class (a set of α vectors obtained by solving the optimization problem), which allows to characterize the entire class as a set of feature points. After a new set of k relevant patches are manually selected for capture additional mass information from the training RoIs, these are characterized using a multi-scale edge analysis and projected onto the previously constructed dictionaries, following a sparse representation, in terms of $x = \mathbf{D}\alpha$, where x is the feature vector of a RoI patch and α corresponds to the projection coefficients of x. $\mathbf{D}$ is replaced by $\mathbf{D_m}$ if the patch belongs to a malign RoI or by $\mathbf{D_b}$ if the patch comes from a benign RoI. This coefficients allow to represent each patch in a feature space, thus defining each class (malign or benign) as clouds of feature points in this space.

2.4 Classification

When a new RoI under analysis arrives, a mass classification strategy that uses the K-NN rule (K-Nearest Neighbor) was implemented. First, a set of patches are manually selected at the test RoI, and then characterized by proyecting each patch onto the severity dictionaries $\mathbf{D}_m$ and $\mathbf{D}_b$. For each patch, two different representation coefficients are obtained after applying the sparse representation framework (described in Subsection 2.3), one indicating the projection onto the benign dictionary, α_b, and the other one describing the projection onto the malign dictionary, α_m. Then, the complete set of coefficients is located as a set of points in the feature space, and each point is classified as benign or malign using the k-nearest neighbors algorithm. The algorithm used a weighted Mahalanobis distance (wd) to measure the similarity among the points in the feature space describing both the benign and malign class.

Finally, the classification of the entire RoI, S^I, is obtained by applying a decision rule [13], which uses each classified point, weighted by the distance to the nearest neighbor, to infer the corresponding class for the RoI. The decision rule can be written as follows

$$S^I = \arg\max_{S_i} |S_1, S_2|, \quad S_i = \sum_{i=1}^{K} w_d^{s_i}, \quad i = 1, 2 \tag{1}$$

where S_1 and S_2 corresponds to benign and malign classes, respectively, and $w_d = 1/d(\boldsymbol{x}, \boldsymbol{y})$ is the point weight, calculated as the Mahalanobis distance between the nearest neighbor ($\boldsymbol{y}$) and the actual point ($\boldsymbol{x}$).

3 Preliminar Results

A small set of 80 regions, extracted from the *Digital Database for Screening Mammography (DDSM)* [8], were used to preliminary evaluate the performance of the proposed approach. Each RoI was previously classified as benign or malign by a group of breast radiologists, according the BI-RADS standard. The set of RoIs was splitted into two

training sets (30 benign RoIs and 30 malign RoIs) and one testing set (20 RoIs). The training set was used for constructing the $\mathbf{D}_m$ and $\mathbf{D}_b$ dictionaries, 60 image patches (size: 3×3 pixels) were randomly sampled from each training RoI, leading to two severity dictionaries, each one containing 1800 patches. Then, to characterize each class in the feature space, 900 feature points were used per class, obtained after applying the sparse representation framework to 30 manually sampled patches per each training RoI. For the sparse representation, we have used the SparseLab[1] library that provides a set of solvers for the optimization problem (from this library we have chosen the Basis Pursuit solver).For classification of each test image, 30 manually sampled patches were Selected per RoI, and then projected onto the two severity dictionaries, leading to a set of 60 feature points. The optimal number of k for the k-nearest neighbors algorithm was estimated by a 10-fold cross validation assessment. Results showed that a minimal of 11 neighboring feature points are needed for establish optimally the corresponding severity level.

Classification performance was assessed by computing the accuracy rate from a confusion matrix of the test images, according to the ground truth provided with the DDSM mammogram database (defined by experienced radiologists). The accuracy was defined as:

$$\text{Acc} = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

where TP, TN, FP and FN stand for true positives, true negatives, false positives and false negatives, respectively. From the 20 test regions, 5 benign RoIs and 9 malign RoIs were correctly classified, leading to an accuracy rate of 70%. This results are reported in the Table 1.

Table 1. Confusion Matrix for classification of 20 test RoIs.

	Benign	Malign
Benign	5	4
Malign	2	9

4 Conclusions

In this paper a new strategy for breast mass classification from mammography images based on a sparse representation scheme was proposed, implemented and evaluated. This strategy provided a BI-RADS mass classification of a RoI as benign or malign, which was supported by a set of diagnosed images that were previously classified by expert radiologists. Instead of attempting to segment masses, we proposed a mass feature description, based on its internal structure with no explicit mass boundary detection.

The proposed approach was evaluated on a public image database (DDSM). The preliminar results have shown that this approach is successfully able to classify the severity of a RoI using learned dictionaries. Even though the proposed classification

---

[1] http://sparselab.stanford.edu/

scheme have been tested with a small dataset, the obtained accuracy of 70% seems to be promising for automatic classification of breast masses. These preliminary results have opened up new strategies for the development of computer-aided tools, based on the sparse representation framework, for mammographic diagnosis. Further work includes to perform extensive validations with bigger datasets and to include other breast mass characteristics, like shape, margin and density.

References

1. American Cancer Society: American Cancer Statistics. (2007) Updated: September 2, 2008.
2. American College of Radiology (ACR): Illustrated Breast Imaging Reporting and Data System (BI-RADS). ACR (1998)
3. R. Bird, T. Wallace, and B. Yankaskas, Analysis of cancers missed at screening mammography, Radiology 178 (1992), 234–247.
4. S. Buseman, J. Mouchawar, N. Calonge, and T. Byers., Mammography screening matters for young women with breast carcinoma., Cancer 97 (2003), 352–358.
5. A. M. Bruckstein, D. L. Donoho, and M. Elad., From Sparse Solutions of Systems of Equations to Sparse Modeling of Signals and Images., SIAM Review 51 (2009), 34–81.
6. H. D. Cheng, X. J. Shi, R. Min, L. M. Hu, X. P. Cai, H. N. Du. Approaches for automated detection and classification of masses in mammograms., Pattern Recognition 39 (2006), 646–668
7. D. Gur, J. S Stalder, L. A. Hardesty, B. Zheng, J. H. Sumkin, D. M Chough, B. E. Shindel, and H. E. Rockette, Computer-aided detection performance in mammographic examination of masses: assessment., Radiology 233 (2004), 418–423.
8. M. Heath, K. Bowyer, D. Kopans, R. Moore, and W. P. Kegelmeyer, The digital database for screening mammography, in Proceedings of the Fifth International Workshop on Digital Mammography, Medical Physics Publishing M.J. Yaffe, ed (2001), 212–218.
9. J. Herredsvela, K. Engan, T. O. Gulsrud, and K. Skretting, Detection of masses in mammograms by watershed segmentation and sparse representation using learned dictionaries. (paper in pdf-format), Proceedings NORSIG (2005), 35–40.
10. H. Kim and J. Kim, Region-based shape descriptor invariant to rotation, scale and translation., Signal Proc.: Image Communication 16 (2000), 87–93.
11. S. G. Mallat and Z. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Transactions on signal processing 41 (1993), no. 12, 3397–3415.
12. R. M. Nishikawa, Current status and future directions of computer-aided diagnosis in mammography, Computerized Medical Imaging and Graphics 31 (2007), 224–235.
13. F. Narvaez, G. Diaz, E. Romero, Automatic BI-RADS description for mammographic masses, IWDM2010 Digital Mammograpyhy, LNCS 6136 (2010), 673–681.
14. B. A. Olshausen, Principles of image representation in visual cortex, pp. 1603–1615, MIT Press, 2003.
15. N. A. Rosa, J. C. Felipe, A. J. Traina, R. M Rangayyan, and P. M. Azevedo-Marques, Using relevance feedback to reduce the semantic gap in content-based image retrieval of mammographic masses., Conf Proc IEEE Med Biol Soc (2008), 406–409.
16. Y. Tao, S. B. Lo, M. T. Freedman, and J. Xuan, A preliminary study of content-based mammographic masses retrieval, Proc SPIE 6514 (2007), 65141Z.
17. K. Verma and J. Zakos, A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques, IEEE Transactions on Information Technology in Biomedicine 16 (2002), 219–223.

18. B. Zheng, C. Mello-Thoms, X. H. Wang, G. S. Abrams, J. H. Sumkin, D. M. Chough, M. A. Ganott, A. Lu, and D. Gur, Interactive computer aided diagnosis of breast masses: Computerized selection of visually similar image sets from a reference library, Academical Radiology 14 (2007), 917–927.

19. K. Wongsritong, K. Kittayaruasiriwat, F. Cheevasuvit, K. Dejhan, A. Somboonkaew. Contrast enhancement using multipeak histogram equalization with brightness preserving., IEEE Asia-Pacific Conference on Circuits and Systems Proceedings, (1998), 455–458

Automatic Detection of Architectural Distortion in Mammograms using Sparse Overcomplete Dictionaries of a Curvelet Descriptor

Fabián Narváez, Cesar Poveda*, and Eduardo Romero

Computer Imaging and Medical Application Laboratory - Cim@Lab,
Departament of Medicine,
Universidad Nacional de Colombia, Bogotá, Colombia
* Instituto Nacional de Cancerología, Bogotá, Colombia
http://www.cimlaboratory.com

Abstract. Mammography aims at the detection of early breast disease stages, a condition for which the radiological illness signs are usually blurred. In particular, the architectural distortion, a common cause of false-negative on mammograms, arise as the change of normal oriented texture of the breast, which is usually hidden by the intricate breast anatomy. This paper presents a novel detector of Regions of Interest (RoIs) with a certain degree of architectural distortion, i.e., an associated normal control (NC) or architectural distortion (AD) membership value. The approach carries out a multiscale image decomposition using the curvelet transform and representing the marginal curvelet subband as the parameters of a Generalized Gaussian Density (GGD) that approximates the subband coefficient distribution. A set of training images serves to construct NC and AD dictionaries in that multiscale space from randomly selected RoIs. A mammogram is decomposed into overlapping patches that are projected to the multiscale space, reconstructed with the two dictionaries and assigned to any of the two classes with the smallest ℓ_2 norm of the reconstructing vector α. The membership function is computed as the Kullback−Leibler distance between any multiscale represented RoI patch and the closest basis of the dictionary, set by the largest projection coefficient in the α vector. We demonstrate the effectiveness of the proposed descriptor by classifying AD in a set of 19 RoIs (4269 patches) of the MIAS database and we obtained a sensitivity and specificity rates of 65 % and 82 %, respectively while the number of false positives per image was 0.9.

Keywords: Breast Cancer, BI-RADS, Architectural Distortion, Curvelet transform, Kull-back Leibler Divergence

1   Introduction

Breast cancer is fully tractable if diagnosis is achieved early, reason for breast cancer screening programs have been introduced in most health systems. Mammography, as part of these screening programs, is considered the most cost-effective method for visualizing abnormalities in the very early stages [21, 5].

2   F. Narváez et al.

Mammography allows detection of some abnormalities such as calcifications, bilateral asymmetry, masses and architectural distortion [5], being architectural distortion the third most common mammographic finding. This breast imaging term is associated to parenchymal asymmetry and usually amounts to any change in the usual parenchymal pattern. This includes spiculations radiating from a point and focal retraction or distortion at the edge of the parenchyma [4]. Likewise, focal retraction is considered as easier to perceive than spiculated distortion and could be categorized as malignant or benign. However, mammographic interpretation is a really difficult examination, with high intra and inter-observer variabilities. Studies have shown that between 10% and 25% of breast cancers are not detected [4]. Scarring from previous surgery, radial scar and crossing breast structures are usually misdiagnosed as architectural distortions. Architectural distortions have indefinite characteristic features, e.g. number of spicules, width and length of spicules and density. An agreement to reduce such variability resulted in the *Breast and Imaging Report and Database System* (BI-RADS), designed by the the American College of Radiology , as a standard description to report breast lesions and categorize different pathologies as well as their severity level [17].

Automatic aid diagnosis systems for mammography categorization is becoming largely used in actual clinical scenarios to complement the clinician first diagnosis impression. Computer Assisted Diagnosis Systems (CAD) have decreased variability diagnosis, becoming a well accepted clinical practice to assist radiologists interpreting mammograms when they search micro-calcification clusters [13]. However, architectural distortion is still a very open problem. Several automated methods for AD detection have been reported, all of them relying on the fact that AD is a group of aligned structures with different orientation. RoIs have been characterized [11, 3, 20, 18, 10] by segmenting the breast, sharpening the spicules, filtering in the Radon domain, extracting the mean direction, filtering with a Bank of Gabor filters, extracting potential AD with the phase portrait model or the fractal dimension and calculating the Haralick's texture measures. All these approaches use a Support Vector Machine to set an optimal classification threshold. In spite of the multiple techniques, the proposed solutions are still unsatisfactory. There is either a low sensitivity or a high number of false positives per image (FPI). Ichikawa et al. [11] reported a sensitivity for detection of only 56 % while the number of false positives per image obtained by Sampat et al. [20] was 14. The results of Rangayyan et al. [18] showed a sensitivity of 84 % but with a high false-positive rate of 7.4 per image. The commercial CAD systems [9] correctly marked a 90 % of mass and calcification cases, whereas the sensitivity of AD detection was reported to be only of 40 %. The R2 Image Checker system [7] successfully identified 49 % cases containing AD. The CADx Second Look system [1] successfully detected 33 % cases of AD with 1.27 false positives per image [19]. Additionally, only 48 - 60 % of ADs that are biopsied are found to be cancer [19].

In this paper, we proposed a new approach to detect and discriminate AD lesions in a mammogram. Given a particular mammogram, the method obtains

a normalized mammogram membership function, which is thresholded to determine to which class a RoI belongs. This function is found by firstly splitting the mammogram into a set of overlapped patches, which are projected to the curvelet space. The subband coefficients are well approximated by the GGD parameters, achieving thereby a considerable dimensionality reduction, i.e., a subband from a 64×64 patch, usually composed of 60 coefficients, is described by two parameters. A particular patch is then set to any of the two classes by the smallest ℓ_2 sparse reconstructing vector. Once the class has been assigned, the patch is characterized by the Kullback−Leibler distance of the patch to the basis dictionary with the largest sparse representation coefficient. Finally, the membership function is obtained by normalizing the distances associated to any patch of the mammogram.

The rest of this article is organized as follows: after this introduction, next section presents the methodology, then results are shown and last section discusses future work and conclusions.

2   Methodology

2.1   The method
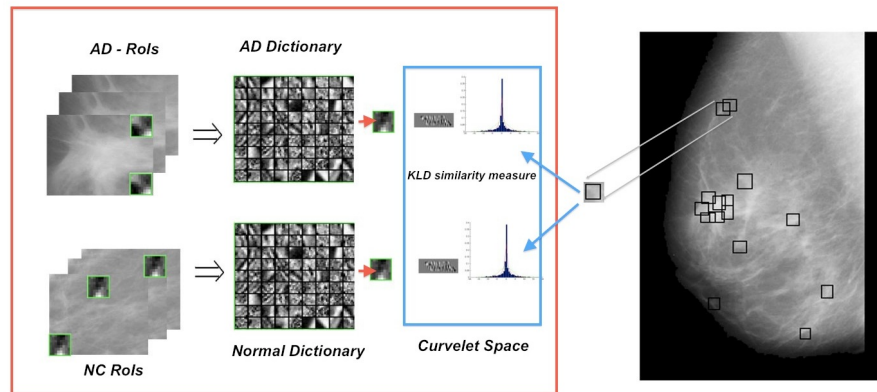
Fig. 1. Method overview: red block shows the process of construction of dictionaries (AD-NC), and automatic detection of architectural distortion based on a curvelet descriptor of different patches

The whole method searches to characterize a particular RoI by projecting the mammogram data to the curvelet space, thereby improving data sparsity. Two overcomplete dictionaries also projected to the curvelet space, representing the AD and NC classes in the curvelet space, serve as a reference conceptual

4   F. Narváez et al.

frame (see figure 1), i.e., the minimal distance to each of the two dictionaries will define the classification problem. The curvelet sparsity is then additionally used to reduce dimensionality by approximating the subband coefficients with two parameters of a Generalized Gaussian Distribution. The result is that the whole curvelet decomposition, with for instance four scales and eight directions, is represented as a vector with 32 values. Provided that each curvelet sub-band is approximated by a probability distribution function, The Kullback−Leibler divergence is used as the metrics to introduce a distance notion between a RoI and each of the spaces spanned by the dictionaries. Once the RoI is characterized by the set of couples (μ, σ) representing the subbands of the curvelet decomposition, the Kullback−Leibler distance to the closest basis of the two dictionaries defines a membership function that will define the particular region classification.

2.2   RoI Pre-processing

Every image was herein stretched to the maximum and minimum gray level values ($[0, 255]$), but adaptively equalizing the histogram so that structural details were preserved. Resultant images were smoothed out by a median filter to remove the remaining noise [16].

2.3   The curvelet transform: statistical characterization

The curvelet transform is a multiscale decomposition [6] with a set of redundant bases which optimally represent 2D curves. In addition to the usual information about scale and location, already available from a wavelet, each of these curvelet bases is able to capture information about orientation while fulfilling a parabolic anisotropic scaling law $width \approx length^2$, that assures curves at different scale levels conserve their geometrical relationships [6]. A curvelet in the frequency domain, is constructed as the product of two windows: the angular and the radial dyadic frequential coronas. The angular window corresponds to a directional analysis, i.e., a Radon transform, and the radial dyadic window is a bandpass filter with cut frequencies following the parabolic anisotropic scaling law [6].

In general, the subband curvelet coefficient distribution is characterized by a sharper peak at zero with smooth tails. This leptokurtic pattern is associated to the sparse property of this transformation, i.e., few coefficients have high probability of being used to represent. It has been shown that the parameters of a generalized Gaussian density can fully describe the marginal distribution of the subband curvelet coefficients. The GGD reads as

$$p(x; \theta, \gamma) = \frac{\gamma}{2\theta\Gamma(1/\gamma)} e^{-(|x|/\theta)^\gamma}$$

where $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$, $z > 0$ is the Gamma function, θ is the variance and γ is related to the decreasing rate of the GGD. The parameters θ and γ are estimated from the subbband data through Maximum Likelihood, as detailed in [8]. The parameters (θ, γ) may be used as descriptor of the probability density function of the energy levels inside each curvelet subband.

2.4   Similarity measure

The similarity between subband curvelets is measured using the Kullback-Leibler divergence (KLD) of the corresponding GGDs:

$$D(p(.;\theta_1,\gamma_1)||p(.;\theta_2;\gamma_2)) = \log\left(\frac{\gamma_1\theta_2\Gamma(1/\gamma_2)}{\gamma_2\theta_1\Gamma(1/\gamma_1)}\right) + \left(\frac{\theta_1}{\theta_2}\right)^{\gamma_2}\frac{\Gamma((\gamma_2+1)/\gamma_1)}{\Gamma((1/\gamma_1)} - \frac{1}{\gamma_1}$$

where (θ_1,γ_1) and (θ_2,γ_2) are the GGD parameters estimated for each subband. This metric does not require additional normalization and shows good performance in other multiscale domains [8].

2.5   Dictionary Construction

The next step is to build dictionaries $\mathbf{D}_{AD}$ and $\mathbf{D}_{NC}$ for the architectural distortion and normal controls, respectively. For doing so, it was selected a set of N RoIs with different architectural distortion lesions or normal controls. Then, a set of K 64×64 patches were randomly selected from each RoI and transformed to the curvelet space. Each subband of any patch is approximated by the couple of the GGD parameters so that the patch is characterized by a vector of couples. Finally, the characterized patches are stored as columns of the Dictionaries $\mathbf{D}_{AD}$ and $\mathbf{D}_{NC}$.

2.6   Class Characterization

Sparse representations attempt to identify the constituent parts of an object and use them for reconstructing any scene with a certain accuracy. These parts, denoted as basis functions or atoms, are usually arranged in overcomplete dictionaries with a larger number of elements than the effective dimensionality of the input space, thereby representing a wider range of image phenomena [15, 14, 12]. Formally, consider a $n\times m$ matrix $\mathbf{D}$, where each column is a possible image in $\mathbb{R}^n$ (atomic images), a dictionary of atoms. The projection of an image x onto the space spanned by $\mathbf{D}$ yields a weighting vector α ($x = \mathbf{D}\alpha$). Furthermore, if α is sparse (with $k_0 \ll m$ nonzeros), this produces a linear combination of k_0 atoms with varying weights. To find the adequate α, the optimization problem to solve is $\mathcal{G}_1(\mathbf{D}, x, \lambda)$, which has the form

$$\mathcal{G}_1(\mathbf{D}, x, \lambda): \min_{\alpha} \lambda\|\alpha\|_1^1 + \frac{1}{2}\|x - \mathbf{D}\alpha\|_2^2$$

The solution consists in finding the sparsest vector α that weights x as a linear combination of atoms from $\mathbf{D}$, using the norm ℓ_1 as a measure of sparsity.

After a random selection of patches from two different types of RoIs, namely architectural distortion and normal controls, two different dictionaries $\mathbf{D}_{AD}$ and $\mathbf{D}_{NC}$ were obtained. To identify whether a particular patch is architectural distortion or not, the patch is firstly characterized by the set of parameters of each of the curvelet subbands. The patch feature vector is then reconstructed using

6   F. Narváez et al.

both dictionaries and assigned to the class with the smallest ℓ_2 reconstructing vector. Afterward, the patch is characterized with the distance between its curvelet representation and the most important atom for the reconstruction in the curvelet space. This distance is then normalized with respect to the distances of every patch in the image and such patch classified after a simple threshold. For the sake of classification, this threshold was herein set to 0.62 and found by automatically adjusting the minimum number of False Positives in the ROC curve.

3   Experimental results

3.1   Experimental setup

A total of 38 mammograms, 19 containing AD lesions, extracted from the *MiniMias Database* [2] were used for the evaluation. Images were curvelet transformed, with 4 scales and 32 orientations, resulting in a descriptor with 66 subbands for each patch. Dictionaries were constructed by randomly sampling 250 overlapped patches from a set of ten images containing the AD lesion, each patch having 64×64 pixels. Evaluation was then performed in the remaining set of 9 mammograms.

3.2   AD detection and False Positive minimization

Once the membership function is computed for the whole mammogram, a simple threshold defines the patch classification. This threshold is automatically determined by setting a value which minimizes the number of False Positives in the ROC curve with the restriction that the area under the curve should be maximal.
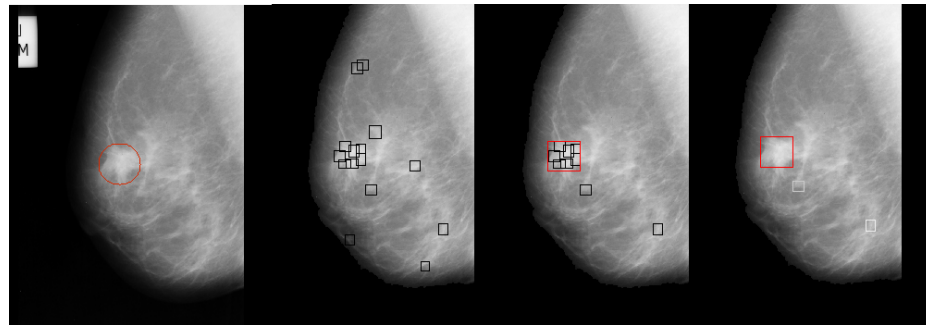

Fig. 2. Left panel shows the selected AD while the sequence of panels illustrate different numbers of False Positive after setting the binary membership function to several thresholds

Title Suppressed Due to Excessive Length   7

3.3 Sensitivity and False Positive Rate

Sensitivity, specificity and a minimal number of False Positives was determined in an independent set of 9 images

Table 1.

	Sens.	Spec.	NFP
AD	65	82	0.9
NC	86	92	0.93

4 Conclusion and future works

We have introduced a new descriptor for AD lesions, based on curvelets and a statistical model of the curvelet coefficients for RoIs extracted from mammography images. By applying the curvelet transform and adjusting the levels of energy for each subband to a generalized Gaussian model, we obtained a robust representation which captures the edge distribution at different orientation and scales. Experimental results indicate that the new feature improves detection and classification performance. Future works include improving the feature with invariance to rotation and scale and extensive experimentation in large mammograms databases.

5 Acknowledgment

References

1. (2012), `http://www.icadmed.com`
2. et al, J.S.: The mammographic image analysis society digital mammogram database exerpta medica. vol. 1069, pp. 375–378 (1994)
3. Ayres, F.J., Rangayyan, R.M.: Characterization of architectural distortion in mammograms. IEEE Enginering in Medicine and Biology Magazine pp. 59–67 (2005)
4. Bird, R., Wallace, T., Yankaskas, B.: Analysis of cancers missed at screening mammography. Radiology 178, 234–247 (1992)
5. Buseman, S., Mouchawar, J., Calonge, N., Byers., T.: Mammography screening matters for young women with breast carcinoma. Cancer 97, 352–358 (2003)
6. Candes, E., Demanet, L., Donoho, D., Ying, L.: Fast discrete curvelet transforms. Multiscale Modeling and Simulation 5(3), 861–899 (2006)
7. Checker:, T.R.T.I.: Image checker (2012), `http://www.r2tech.com`

8   F. Narváez et al.

8. Do, M., Vetterli, M.: Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance. IEEE Transactions on Image Processing 11(2), 146–158 (2002)

9. Garvican, L., Field, S.: A pilot evaluation of the r2 image checker system and users' response in the detection of interval breast cancers on previous screening films. Clinic Radiology 56(10), 833–837 (2001)

10. Guo, Q., Shao, J., Ruiz, V.: Investigation of support vector machine for the detection of architectural distortion in mammographic images. Journal of Physics: Conference Series(15) (2005)

11. Ichikawa, T., Matsubara, T., Hara, T., Fujita, H., Endo, T., Iwase, T.: Automated detection method for architectural distorion areas on mammograms based on morphological processing and surface analysis. In: Processing, I. (ed.) Medical Imaging. vol. 5370 (2004)

12. Mallat, S.G., Zhang, Z.: Matching pursuits with time-frequency dictionaries. IEEE-Transactions on signal processing 41(12), 3397–3415 (1993)

13. Nishikawa, R.M.: Current status and future directions of computer-aided diagnosis in mammography. Computerized Medical Imaging and Graphics 31, 224–235 (2007)

14. Olshausen, B.: Principles of image representation in visual cortex. No. 1603-1615, MIT (2003)

15. Olshausen, B.A., Field, D.J.: Sparse coding whit an overcomplete basis set: A strategy employed by v1? Vision Res. Pergamon 37(23), 3311–3325 (1997)

16. Petrick, N., Chan, H.P., Sahiner, B., Wei, D.: An adaptive density weighted contrast enhancement filter for mammographic breast mass detection. IEEE Trans. Med. Imaging 15 1, 59–67 (1996)

17. of Radiology (ACR), A.C.: Illustrated Breast Imaging Reporting and Data System (BI-RADS). ACR (1998)

18. Rangayyan, R.M., Prajna, S., Ayres, F.J., Desautels, J.E.L.: Detection of architectural distortion in prior screening mammograms using gabor filters, phase portraits, fractal dimension and texture analysis. Int J CARS. Springer (2008)

19. Rangayyana, R.M., Ayres, F.J., Desautelsa, J.E.L.: A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. Journal of the Franklin Institute 344, 312–348 (2007)

20. Sampat, M.P., Whitman, G.J., Markey, M.K., Bovik, A.C.: Evidence based detection of spiculated masses and architectural distortions. In: Medical Imaging. Image Processing, vol. 5747, pp. 26–37. SPIE (2005)

21. Verma, K., Zakos, J.: A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques. IEEE Transactions on Information Technology in Biomedicine 16, 219–223 (2002)

B. Multi-view information fusion for description of mammographic masses

In order to improve the diagnosis process of breast cancer, an ipsilateral multi-view information fusion was investigated for mass description in mammograms by exploiting correlative information of suspicious lesions between mammograms of the same breast. The contribution of this work was published in:

Multi-view Information Fusion for automatic BI-RADS description of mammographic masses

Fabian Narvaez[1], Gloria Diaz[1] and Eduardo Romero[1]

[1]BioIngenium Research Group, National University of Colombia, Bogota, Colombia

ABSTRACT

Most CBIR-based CAD systems (Content Based Image Retrieval systems for Computer Aided Diagnosis) identify lesions that are eventually relevant. These systems base their analysis upon a single independent view. This article presents a CBIR framework which automatically describes mammographic masses with the BI-RADS lexicon, fusing information from the two mammographic views. After an expert selects a Region of Interest (RoI) at the two views, a CBIR strategy searches similar masses in the database by automatically computing the Mahalanobis distance between shape and texture feature vectors of the mammography. The strategy was assessed in a set of 400 cases, for which the suggested descriptions were compared with the ground truth provided by the data base. Two information fusion strategies were evaluated, allowing a retrieval precision rate of 89.6% in the best scheme. Likewise, the best performance obtained for shape, margin and pathology description, using a ROC methodology, was reported as $AUC = 0.86$, $AUC = 0.72$ and $AUC = 0.85$, respectively.

Keywords: Mammography, BI-RADS, Computer Aided Diagnosis Systems, Content-based Image Retrieval, Automatic Annotation, Multi-View Analysis, Information Fusion, Breast Cancer

1. INTRODUCTION

Mammography is still the choice method for early detection of breast cancer, which is the largest public health problem in women population.[1,2] Mammography has the reputation of being the most difficult radiological examination.[3] It has been reported that radiologists fail to detect about $10 - 25\%$ visible cancers on mammograms.[4] Computer-aided detection and diagnosis (CAD) systems have been developed to assist radiologists in making more accurate diagnosis on breast cancer with mammograms, which have been studied widely to classify malignant and benign abnormalities.[5] These systems seek to increase diagnostic accuracy and to improve the efficacy and efficiency of mammographic interpretations by using some paradigms of signal processing and pattern recognition approaches such as image noise reduction, automatic segmentation, efficient characterization and classification methods.[6] On the other hand, as an attempt for reducing diagnosis variability among radiologists, the American College of Radiology has developed the *Breast and Imaging Report and Database System* (BI-RADS), a standard description for reporting breast lesions that allows to categorize different pathologies as well as their severity level.[7] The use of CAD systems and BI-RADS description has been separately studied, showing a reduction of the variability effect. However, a high incidence of false positives (FP) rates has been reported to the most CAD systems.[8]

One of the most challenging tasks, for both radiologists and CAD systems, is to accurately detect and classify mammographic masses (an important sign of breast cancer). According to the BI-RADS lexicon, masses are described as "space occupying lesions visible in at least two mammographic views".[7] So, one important principle in the radiologist's practice consists in reading, comparing and combining information from the two different views of the same breast to reach a final decision.[9,10] Figure 1 shows a typical mammographic case. Left panel corresponds to the mediolateral oblique (MLO) view, a projection angle of 45 degrees which shows part of the pectoral muscle. Right panel stands for the craniocaudal (CC) view, a top-down breast projection. Each view contains spatial information about a single breast, which amounts to a different distortion of the breast tissue. Because of the isotropic breast characterization, many kinds of image features tend to be invariant to the different breast views, i.e. a mass projected in one view will appear in another view. However, as illustrated in

---

Figure 1, in many cases these features can be blurred by overlap of internal breast tissues, making that masses may appear differently from different views. In this figure, the mass is observed in the MLO view as an irregular shape (spiculated), with blurred margin and poorly-defined edges, while in CC view it presents an irregular shape, defined margin and defined edges. In this case, the mass was finally described as irregularly shaped, poor-defined margin and malign lesion.
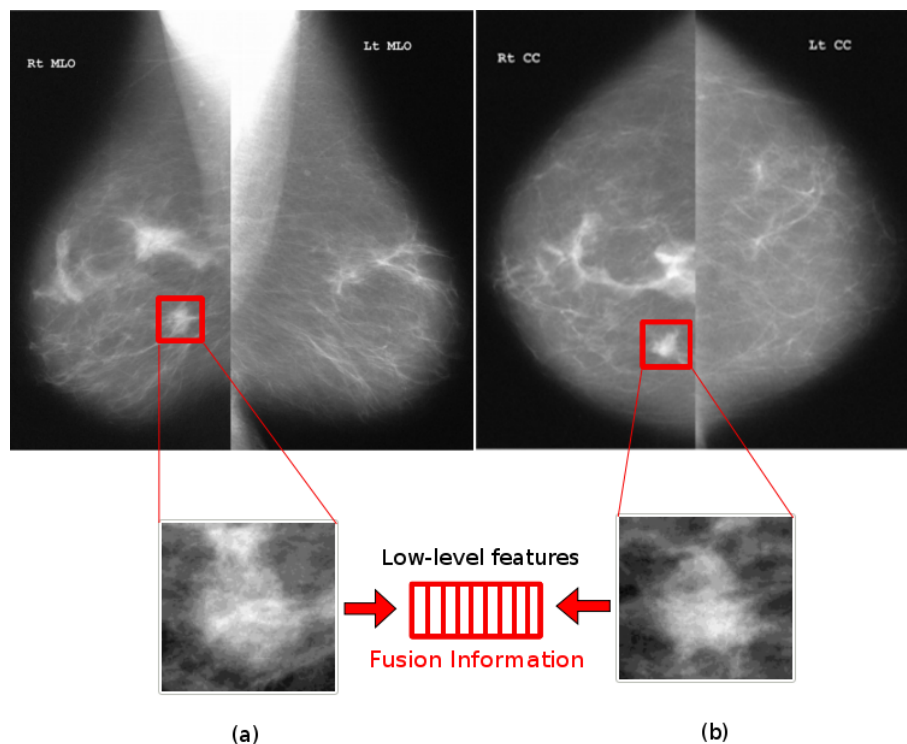

Figure 1. Typical mammographic examination composed by two projections (MLO and CC) for both right and left breasts (Top panel). In down panel a magnification of a mammographic mass, extracted from the two views, which present different visual features for each.

Most CAD systems ignore differences from mass projections and limit their analysis to single features, but each view is independently evaluated. Therefore, information between the complementary breast views, useful for the diagnosis process, is unfortunately disregarded. This fact results in inaccuracies and high incidence of false positives (FP) rates.[8] Recently, the use of information from the two mammographic views have been proposed for improving the single-view CAD performance.[11–15] Thus far, these approaches follow a two-step approach. Initially, traditional single-view CAD systems are used to identify suspicious mammographic regions, from which morphological and textural features are extracted. These features are then used for determinating the correspondence between pair of regions extracted from the two breast views, either using different discriminative classifiers such as Linear Discriminant Analysis (LDA)[12,13,16] and k-Nearest Neighbor.[14] These approaches demonstrated an ability to discriminate between true and false correspondences, reducing the false positives reported by the use of a single-view CAD system. Although the use of multi-view information has shown to be useful for improving the detection of mammographic masses, these approaches are highly dependent on accurate segmentation, a difficult challenge in masses with blurred boundaries.

In a previous work[17] we proposed an interactive CAD system * that provided a BI-RADS mass description of a manually selected region of interest (RoI). Briefly, a radiologist selects a region of interest with a mass, as query RoI, which is enhanced to improve visual details. Afterward, the method finds the most similar regions in a database, according to the BI-RADS lexicon, using the two most important diagnostic features for

---

*content based image retrieval scheme (CBIR)

describing masses: shape and texture. Shape feature was described using the Zernike moments, and texture was captured via the Neighborhood Gray Tone Difference Matrix (NGTDM). Once these basic features were computed, a further dimensionality reduction was achieved using a standard Principal Component Analysis (PCA). A multiclass retrieval algorithm, based on the k-NN rule, was constructed for the shape, margin and pathology classification, aiming at retrieving the most similar regions. Finally, the BI-RADS descriptions of these retrieved regions were used to suggest the most probable description of the selected region. In that work the multi-view and multi-region breast dependences were ignored.

In this paper, we evaluate the effect of using information extracted from both CC and MLO views using the same CBIR scheme. Given a particular mass query, extracted from a "case" (CC-MLO regions of the same breast), the method retrieves the most similar masses, and each mass is associated with some case from the database, according to the BI-RADS lexicon. Once these basic features are computed, they are combined in a unique vector that integrates multi-feature (texture and shape) and multi-view (CC and MLO) information. We evaluated two fusion approaches, in the former, features from each view were assembled into a unique vector and then reduced while in the latter, features were firstly reduced and then fused. Figure 1 illustrates the schemes of the evaluated fusion approaches.

The rest of this article is organized as follows: after this introduction, Section 2 describes the different techniques and methods proposed for retrieving and annotating mammographic masses. Experimental results are reported in Section 3. Finally, Section 4 present the conclusions and perspectives of this work.

2. METHODS

The framework used in this investigation is similar to that presented in[17] as illustrated in Figure 2. Briefly, a radiologist selects a ROI from each of the two views, considered as complementary. They are preprocessed to enhance the mass shape and texture characteristics. Afterward, the mass description is extracted and compared to the information stored in a reference database, so the most similar cases are retrieved. Finally, the BI-RADS descriptions associated to the retrieved cases, are used for assigning the most probable description to the queried mass. Further details can be found in.[17]
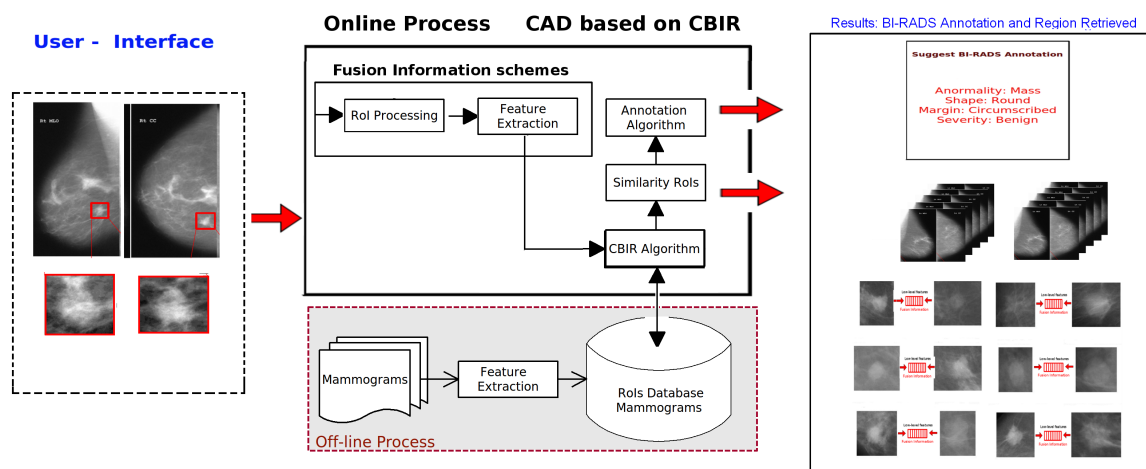
Figure 2. Scheme of the proposed CAD based on CBIR approach. Regions selected by the expert (small squares in the mammographic views) are preprocessed and their feature vectors are fused for generating a description that combine information of both CC and MLO views. This information is used for retrieving the most similar cases in a database that stores classified masses. As a result, the system shows the similar cases and indicates the most probable category for each BI-RADS descriptor.

RoI Pre-processing Each single-view RoI was pre-processed for generating two enhanced images, the former captures mass edges by stretching the maximum and minimum gray level values, the other adaptively equalizes the histogram so that structural details are preserved. In both cases, the resultant images were smoothed out by a median filter.

Feature Extraction According to the BI-RADS lexicon, shape and texture are the most important criteria for mass diagnosis. In this work, shapes were described using the Hosny's implementation[18] of Zernike moments, which constitutes a multiresolution shape representation[19] that allows reconstruction with minimal losses. The number of order moments used for generating the shape descriptor was set by minimizing the reconstruction error. So, the first 60 order moments were selected for generating a descriptor of 961 features for both CC and MLO views, independently. On the other hand, the essential texture features were captured via the Neighborhood Difference Gray Tone Matrix.[20] The average difference was stored in a histogram whose bins 1, 2, 3, 4 and 5 corresponded to the neighborhood sizes. Five histograms of 256 positions were generated and five features were calculated from each, as described in.[20]

Information Fusion Once the feature extraction is performed, multi-feature (shape and texture) and multi-view (CC and MLO) information is fused together for being used in the classification stage. In this work we evaluate two fusing schemes as illustrated in Figure 3. The description from each view is firstly obtained by concatenating the 961 shape and the 25 texture features. Then, two fusion methodologies were assessed. In the former scheme (SCA, dashed lines), a PCA analysis was performed to each view description and the resulting projected vectors were finally concatenated into a unique vector of 30 dimensions. In the second scheme (SCB, continuous line), the feature descriptors from the two mammographic views were first concatenated, generating a unique vector of 1972 features, which was then PCA projected for assembling a descriptor of 35 dimensions. In both cases, the eigenvectors that reported a variability larger than 85% were selected.
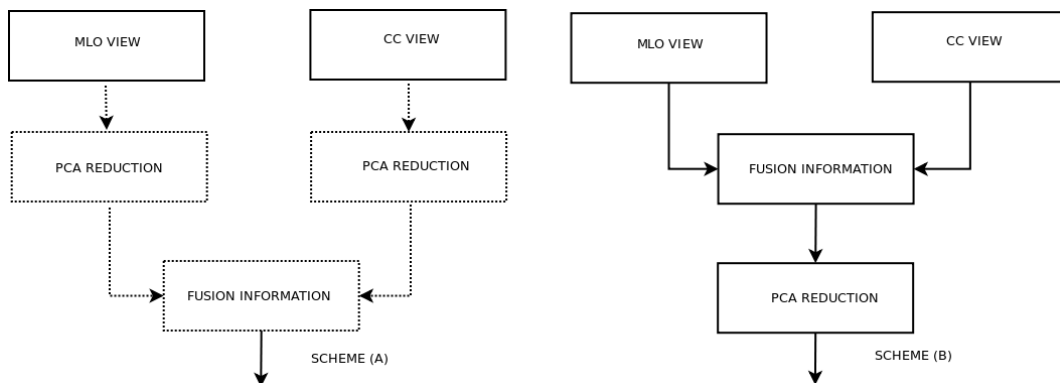
Figure 3. Schemes assessed in the information fusion stages

Automatic Mass Description A CBIR strategy that uses the K-NN rule was implemented using a weighted Mahalanobis distance to measure the similarity among the feature vectors, describing both the database and the query mass. Once a set of K cases is retrieved, each shape, margin and pathology BI-RADS description of the query mass is set using the decision rule in equation 1, where $Label(S)$ assigns the value description with the largest weight.

$$Label(S) = \arg\max_{S_i} |S_1, ..., S_n|, \quad S_i = \sum_{i=1}^{K} w_d^{s_i} \tag{1}$$

where $w_d = 1/d(\vec{x}, \vec{y})$ is the relative weight of each possible value description and S_n corresponds to number of possible labels for each shape, margin and pathology BI-RADS description.

3. EXPERIMENTAL RESULTS

The strategy was evaluated on a total of 400 cases, including pathological masses previously annotated by a group of radiologists, extracted from the *Digital Database for Screening Mammography (DDSM)*.[21] This dataset was split into training (300) and test (100) subsets. Both the CBIR scheme and the automatic annotation algorithm were independently assessed.

Content-based image retrieval evaluation   Retrieval performance was assessed by computing the relevance of the recovered images, according to the ground truth of DDSM mammogram database using a conventional Precision-Recall (P-R) graph. Precision (P) quantifies the relevance of the retrieved images ($P = \sum_{i=1}^{k} S_i/K$), whilst recall (R) amounts to the relevance of retrieved images among all the relevant images stored in the database ($R = \sum_{i=1}^{k} S_i / \sum_{i=1}^{n} S_i$). Figure 4 shows the Precision-Recall curve from evaluation of the testing set, when the 15 most similar cases were retrieved. The first point represents the average precision and recall rates for the first returned case, the second point corresponds to the precision and recall for the first two retrieved cases and so forth. As expected, as long as the recall measure increases, the precision decreases. However and interestingly, precision was higher than 70% in general. As it is observed, the second fusion scheme (SCB) reports the best performance with a precision up to 78% in all cases, whilst the first scheme (SCA) achieves precisions higher than 70%. This shows that fusion is able to preserve cross discriminating information relevant to mammographic masses that is lost when dimensionality reduction is applied for each view descriptor. As expected, both results outperform those reported by using single view mass descriptions,[17] which reached to 65%.
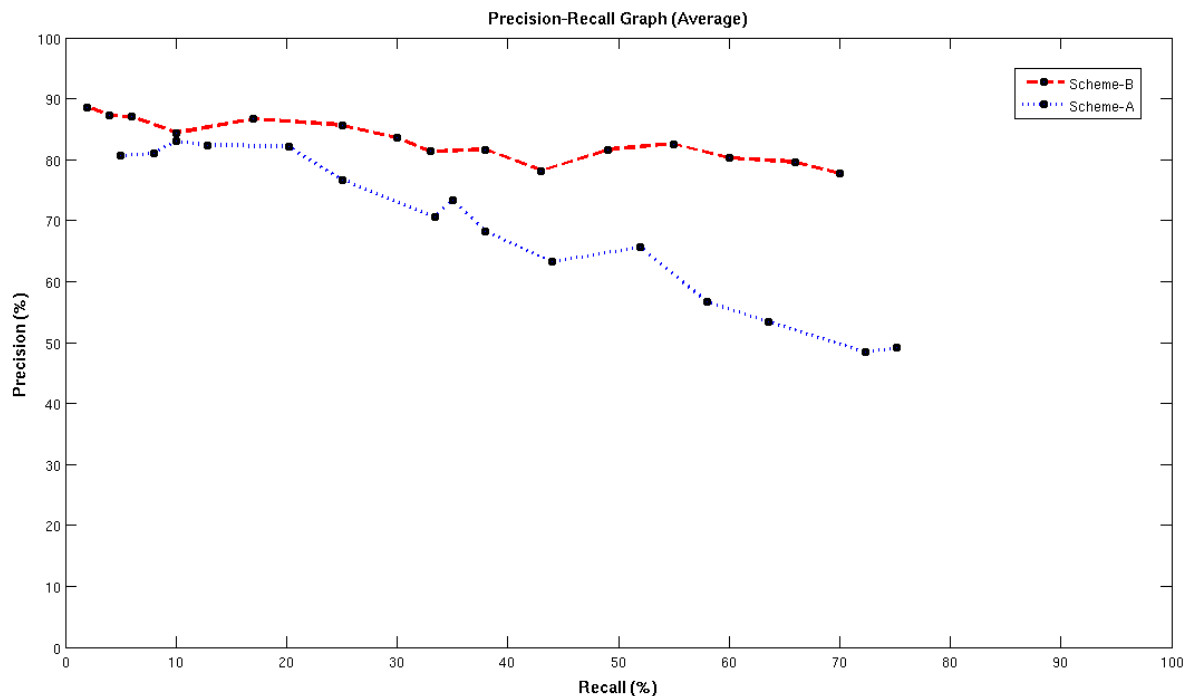
Figure 4. Precision-Recall Average Curve

Automatic annotation assessment   Each shape, margin and pathology BI-RADS description was assigned to each case query, based on equation 1. The optimal number of used cases was estimated by a 10-fold cross validation assessment. Results showed that the best shape and texture descriptions were achieved using the first 7 retrieved images, whilst best pathology description was obtained using 9 of them. Results from automatic

BI-RADS description in the testing set are shown in Table 1. As expected, performance reported by using images retrieved by the scheme B was higher than the scheme A. The three BI-RADS descriptors achieved very good results, i.e. accuracy above 75.1 and 65.1 for schemes A and B, respectively. In general, the most difficult description was the margin, probably due to the blurred boundaries presented in many mass classes. However, average performance annotation was higher than $80, 23\%$ to the best scheme, a figure very acceptable for a system that tries to assist a diagnosis task.

Table 1. Performance of Automatic BI-RADS description for the two information fusion schemes.

Label	Scheme A		Scheme B	
	Accuracy	AUC ROC (A_Z)	Accuracy	AUC ROC (A_Z)
Shape	78.6	0.76 ± 0.03	80.3	0.86 ± 0.03
Margin	65.2	0.65 ± 0.04	75.1	0.72 ± 0.04
Pathology	80.2	0.78 ± 0.04	85.3	0.85 ± 0.04

4. CONCLUSIONS

Accurate identification and description of mammographic masses is actually performed by simultaneous analysis of both craniocaudal (CC) and Medio Lateral Oblique (MLO) views of each breast, a fact that is usually ignored by the CAD systems. In this paper a new content based image retrieval and annotation strategy based on multi-feature (shape-texture) and multi-view (CC-MLO) information fusion is introduced. Given a suspicious region, the proposed approach provides two important tools that support the diagnostic decision of radiologists: a set of visually similar cases, stored in a reference database, and a textual annotation that describes the visual features of the region, according to the BI-RADS lexicon. Two fusion schemes were evaluated, the first one concatenated in a feature vector the PCA projected feature vectors that describe each single-view region. In the second one, the original single-view feature vectors were firstly concatenated in a unique feature vector that was then PCA projected. In overall cases, experimental results demonstrated that information fusion could represents accurately mass characteristics allowing outperform retrieval and description results reported by a similar strategy that limited the analysis to single views. Moreover, the second scheme reported the best results, showing that the fusion of original vectors can to preserve information that results discriminative when the two views are combined.

In the future, we plan to deal with the description of the mass density, which requires more sophisticated feature extraction process that include information of the adjacent tissues. Additionally, we plan to integrate other abnormalities such as architectural distortion and bilateral asymmetry in the BI-RADS based CBIR and description strategy that maybe will become a useful tool for radiologists training in the diagnosis of breast cancer in mammograms.

REFERENCES

[1] "Management of breast cancer in women. a national clinical guideline," tech. rep., Scottish Intercollegiate Guidelines Network (2005).
[2] "American cancer statistics," tech. rep., American Cancer Society (2007).
[3] Maggio, C. D., "State of the art of current modalities for the diagnosis of breast lesions," *Eur. J. Nucl. Med. Mol. Imaging* **31**, S59–S69 (2004).
[4] Bird, R., Wallace, T., and Yankaskas, B., "Analysis of cancers missed at screening mammography," *Radiology* **178**, 234–247 (1992).
[5] Freer, T. and Ulissey, M., "Screening mammography with computer-aided detection: prospective study of 12860 patients in a community breast center," *Radiology* **220**, 781–786 (2001).
[6] Oliver, A., Freixenet, J., Mart, J., Prez, E., Pont, J., Denton, E. R., and Zwiggelaar, R., "A review of automatic mass detection and segmentation in mammographic images," *Medical Image Analysis* **14**, 87–110 (2010).

[7] [*Illustrated Breast Imaging Reporting and Data System (BI-RADS)*], American College of Radiology (1998).

[8] Nishikawa, R. M., "Current status and future directions of computer-aided diagnosis in mammography," *Computerized Medical Imaging and Graphics* **31**, 224–235 (2007).

[9] Sun, L., Li, L., Xu, W., Liu, W., Zhang, J., and Shao, G., "A novel classification scheme for breast masses based on multi-view information fusion," *4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE)* , 1–4 (2010).

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

[11] Good, W., Zheng, B., Chang, Y., Wang, X., Maitz, G., and Gur, D., "Multi-image cad employing features derived from ipsilateral mammographic views," *Proceedings of SPIE* **Medical Imaging 3661**, 474–485 (1999).

[12] van Engeland, S., Timp, S., and Karssemeijer, N., "Finding corresponding regions of interest in mediolateral oblique and craniocaudal mammographic views," *Medical Physics* **33**, 3203–3212 (2006).

[13] van Engeland, S. and Karssemeijer, N., "Combining two mammographic projections in a computer aided mass detection method," *Medical Physics* **34(3)**, 898–905 (2007).

[14] Samulski, M. and Karssemeijer, N., "Matching mammographic regions in mediolateral oblique and cranio caudal views : A probabilistic approach," *Proceedings of SPIE, the International Society for Optical Engineering* **6915(2)**, 69151M.1–69151M.10 (2008).

[15] Velikova, M., Lucas, P. J. F., Ferreira, N., Samulski, M., and Karssemeijer, N., "A decision support system for breast cancer detection in screening programs," in [*Proceeding of the 2008 conference on ECAI 2008: 18th European Conference on Artificial Intelligence*], (2008).

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

[17] Narvaez, F., Diaz, G., and Romero, E., "Automatic bi-rads description of mammographic masses," *Digital Mammography, 10th international workshop on digital Mammography* **Volume 6136**, 673–681 (2010).

[18] Hosny, K. M., "Fast computation of accurate zernike moments," *J Real-Time Image Proc* **3**, 97–107 (2008).

[19] Belkasim, S., Hassan, E., and Obeidi, T., "Radial zernike moment invariants," *The Fourth Int. Conf. on Computer and Information Tech* **1**, 790–795 (2004).

[20] Amadasun, M. and King, R., "Textural features corresponding to textural properties," *IEEE Transactions on systems, man, and Cybernetics* **19**, 1264–1274 (September 1989).

[21] Heath, M., Bowyer, K., Kopans, D., Moore, R., and Kegelmeyer, W. P., "The digital database for screening mammography," *in Proceedings of the Fifth International Workshop on Digital Mammography, Medical Physics Publishing* **M.J. Yaffe, ed**, 212–218 (2001).

C. An open access thyroid ultrasound-image Database

In the process of performing research the author has also contributed significantly to the following work:

An open access thyroid ultrasound-image Database

Lina Pedraza [1], Carlos Vargas [1] Fabián Narváez [1], Oscar Durán [2], Emma Muñoz [2] and
Eduardo Romero [1]

[1]Cim@Lab, Faculty of Medicine, Universidad Nacional de Colombia, Bogotá, Colombia
[2]IDIME - Instituto de Diagnóstico Médico, Bogotá, Colombia

ABSTRACT

Computer aided diagnosis systems (CAD) have been developed to assist radiologists in the detection and diagnosis of abnormalities and a large number of pattern recognition techniques have been proposed to obtain a second opinion. Most of these strategies have been evaluated using different datasets making their performance incomparable. In this work, an open access database of thyroid ultrasound images is presented. The dataset consists of a set of B-mode Ultrasound images, including a complete annotation and diagnostic description of suspicious thyroid lesions by expert radiologists. Several types of lesions as thyroiditis, cystic nodules, adenomas and thyroid cancers were included while an accurate lesion delineation is provided in XML format. The diagnostic description of malignant lesions was confirmed by biopsy.The proposed new database is expected to be a resource for the community to assess different CAD systems.

1. INTRODUCTION

Thyroid disorders are the most common endocrine pathologies, among which thyroid cancer is the more frequent. According to the American Cancer Society, 62.980 new cases and 1.890 deaths will take place in the United States[1] in 2014. The incidence of palpable thyroid nodules on the adult population is about 67% and 10% of them can be malign[2].[3] Modern imaging techniques such as the computer tomography (CT) and Ultrasound (US) image modalities, are used to detect, diagnose and manage thyroid nodules. B-mode Ultrasound image is the best cost-effective method and less invasive technique. However, interpretation of ultrasound images requires a high degree of expertise and training while the reading performance is affected in any case by the noise and speckle of the ultrasound images as well as the ability of the operator to properly acquire the image.[4] An inaccurate US capture of a nodule might result in unnecesary fine-needle aspiration (biopsy).

Usually, a thyroid nodule is described as hypo-echoic, iso-echoic, or hyper-echoic, being the echogenicity its brightness when compared to the normal parenchyma surrounding this. Hypoechogenicity has been associated to thyroid malignancy.[5] Most of the thyroid nodules tend to have various internal echogenicities, making the final diagnosis very complicated. For instance, the follicular adenoma can be either iso-echoic or hyper-echoic.[6] Several ultrasound features have been found to be associated with an increased risk of thyroid cancer, including hypo-echogenicity, predominantly solid composition and calcifications.[7] Usually, thyroid nodules are heterogeneous, composed of various internal echo patterns that confuse radiologists and physicians at the moment of identifying the thyroid nodule. In order to standardize the ultrasound report that describes and evaluates thyroid lesions, an agreement, still under evaluation, is the Thyroid Imaging Reporting and Data System (TIRADS) that categorizes the nodules as benign, probably benign, borderline, suspicious for malignancy and malignant.[8]

On the other hand, computer aided diagnosis of thyroid ultrasound, systems aimed to support the diagnosis process, are clasically composed of two main stages, feature extraction and classification (between benign or malignant) stages. These systemas attempt to eliminate the weaknesses of operator dependency and to improve the diagnostic accuracy. However, an actual limitation for these systems to develop is the necessity of enlarginkg the scientific community interested in this kind of problems. Thyroid ultrasound-image databases play then a central role in developing algorithms devoted at detecting and diagnosing lesions. Such datasets facilitate comparison of results from different studies. Different opinions from the scientific community agree about the urgency of building a benchmark database of US images,[2] accessible publicly that supports the comparison

---

and evaluation of different algorithms and CAD systems. To the best of our knowledge, there are no available ultrasound-image databases with the associated diagnostic description and report of main findings.

In this work, an image database of thyroid ultrasound images is presented and made public (available in www.cimalab.unal.edu.co). The proposed database contains a set of B-mode Ultrasound images, corresponding a 389 cases of study that include a complete annotation and diagnostic description of suspicious thyroid lesions, using the TI-RADS lexicon description performed by at least two expert radiologists. Several types of lesions such as thyroiditis, cystic nodules, adenomas and thyroid cancers are included and their accurate contours are provided in XML format. The diagnostic description of malignant lesions was confirmed by a biopsy procedure. Furthermore, additional annotation of training radiology students was compare with the expert's , helping them to used this data set as a tool in the training process for the new radiologists.

2. MATERIALS AND METHODS

A set of cases with relevant thyroid disorders were selected from the IDIME Ultrasound Departament (Institute with over ten years of experience), one of the largest diagnostic imaging centers in Colombia and who performs more than 2000 thyroid ultrasounds associated to Fine Needle Aspiration per year. The selection of patients was based on TI-RADS description.[7] These selected patients signed the informed consent agreeing about using this information exclusively with scientific purposes and they underwent both the ultrasound and the biopsy in the same session. Thus, the staff of head and neck experts of IDIME collected a total number of 299 patients, 270 women and 29 men whose ages varied as 57.35 ± 16.2 years. The experts evaluated the patient individually and described the specific features filling the TIRADS requirements. They were separated in different consulting rooms, none of them knew the outcome of the evaluation. The second observer performed the biopsy following the ATA (American Thyroid Association) recommendations. The images were extracted from thyroid ultrasound video sequences captured with a TOSHIBA Nemio 30 and a TOSHIBA Nemio MX Ultrasound devices, both set to 12 MHz convex and linear transducers, containing the most relevant pathological features seen such as size, shape, margin, composition, calcifications and echogenicity for a given view, i.e. sagital or transverse, which were selected by the radiologists and their pathologies confirmed by biopsy using the BETHSEDA system.

Once, the images were acquired and saved in an uncompressed JPEG format, an image annotation tool was developed, which is designed to allows segmented nodules and annotated labels to be instantly shared via the web and to grow over time. For this, we designed an online Javascript drawing interface that works on many platforms, is easy to use, and allows to make a TI-RADS description of nodules over collected data. Figure 1, shows a snapshot of the proposed online annotation tool. The tool provides a simple drawing interface that allows users to outline the silhouettes of the nodule present in each image. When the user opens the application, a new image is displayed. The user provides an annotation by clicking along the boundary of an nodule to form a polygon. The user closes the polygon by clicking on the initial point or with a right click. After the polygon is closed, a pop-up dialog box appears querying for the nodule description according to TI-RADS category such as (1) Normal thyroid, (2) Benign, (3) No suspicious US feature, (4a) One suspicious US feature, (4b) Two suspicious US features, (4c) Three or four suspicious US features and (5) Five suspicious features.The experts and the radiology resident had an individual interface to perform the annotation, this interface allowed them evaluate different images from the same nodule. The annotation was performed without additional information.After this procedure, ultrasound properties were include as well as the location, examination date and frequency (MHz) in an alone XML file. These annotations are added to the database and becomes available for immediate download for research.

On the other hand, the whole cases were categorized by the experts radiologists among thyroiditis, cystic nodules, malignant or benign nodules. The initial diagnosis based on the ultrasound features was confirmed by the pathological report. These pathologic description, as well as the BETHESDA (System for Reporting Thyroid Cytopathology)[9] results, were also included in XML file. The expert's evaluations were compared to analize the accuracy of both radiologist as well as the variability inter observers with the BETHESDA report, the Gold Standard for thyroid cancer. The table 3, shows these results. The XML files provide discriminant information such as malignant and benign cases, each contains the number case, following of patient's information as age and sex. The nodule's information was separed into composition,size, echogenicity, margin characteristics,

Figure 1. Screenshot of the online annotation-tool

calcification presence and TIRADS score. Finally, the free hand annotation nodule coordinates are described as a code-point. Figure 2 illustrate a manual nodule segmentation by the radiologist.
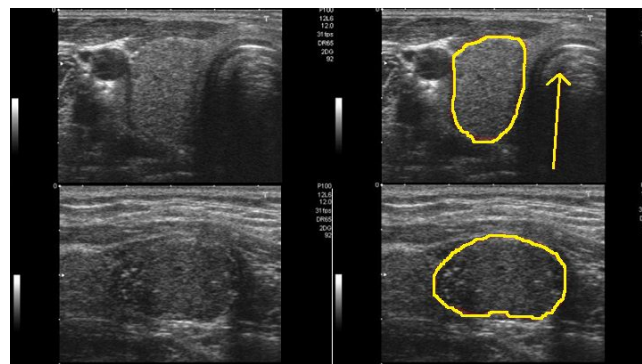
Figure 2. a) original image, b) segmented image by radiologist

Table 1. BETHESDA SYSTEM

SCORE[9]	DESCRIPTION[9]
1	Nondiagnostic or Unsatisfactory
2	Benign Consistent with a benign follicular nodule
3	Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance
4	Follicular Neoplasm or Suspicious for a Follicular Neoplasm
5	Suspicious for Malignancy
6	Malignant

3. RESULTS

The DDTI (Digital Database Thyroind Image) is a public open access dataset. It contains the analysis of 347 thyroid ultrasound images, performed by two experts in 299 patients with thyroid disorders. The patients were classified by the experts using the TIRADS system.Table 2, shows how many cases of each category the dataset contains per observer. The BETHESDA system confirmed 200 cases. Thyroiditis and goiter cases were 137 . Spongiform nodules were 18 cases. The papillar and follicular cancer was confirmed in 21 patients. 24 samples were unsatisfactory for the pathology study.

A discrepancy among the TIRADS score and the Bethesda system was analyzed for each observer. The TIRADS system has been validated to correlate the image features with the pathological diagnosis. This classification pretends avoid the unnecessary FNA and performs an accuracy diagnostic related with the thyroid disorder. One of the most interesting aspect was the behavior of the thyroid disorders in the male population. The majority of the nodules in men are related with cancer, this aspect was known by the observers and became in a bias, almost all male cases were classified with a TIRADS over 4a. Nevertheless, the Bethesda system confirmed only 4 malignant cases.

The most prevalent nodule composition in the 299 patients was solid, in multiples studies this characteristic is related with malignancy but in our dataset benign and malignant nodules presented it.

Another important feature for malignancy was the presence of calcifications, most of the nodules have micro and macro calcifications that can only be seen during the ultrasound process and after the image capture these lost,in our data set the experts performed the annotation not only of the boundaries of the nodules also of some specific features as the calcifications,veins, arteries, muscles and trachea.

On the other hand, the dataset contains 38 images without clinical information, only the definition of the boundaries is available. These images were used for training. The radiology students had the opportunity to evaluated the images and performed the annotation following the TIRADS requirments. After this process the annotation was compared with the expert's.
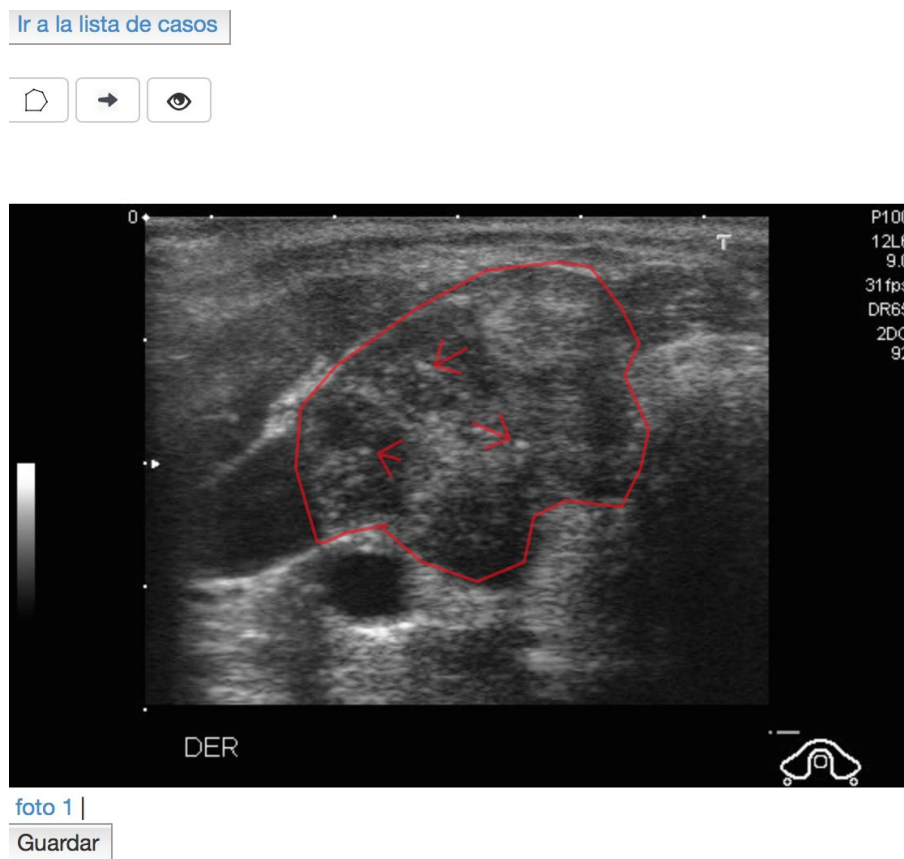

Figure 3. Screenshot of the online training annotation-tool

The original uncompressed JPEG images and XML files have been uploaded to the http://www.cimalab.unal.edu.co web site. The access to the dataset is open and the user can download the cases divided in benign and malign following the TIRADS score, as well the images for training.

TIRADS	Observer 1 No. Diagnosed Cases			Observer 2 No. Diagnosed Cases		
	Male	Female	Total	Male	Female	Total
2	3	35	38	2	18	20
3	2	12	14	3	30	33
4a	9	76	85	11	79	90
4b	5	70	75	8	93	101
4c	6	50	56	1	34	35
5	4	26	30	4	16	20

Table 2. Distribution of the TIRADS classification for each observer

BETHESDA	CASES
1	24
2	137
3	18
4	8
5	13
6	0

Table 3. PATOLOGIC DIAGNOSIS

4. DISCUSSION

The analysis of thyroid ultrasound images is a wide field in continuous expansion due to the difficulty in the ultrasound nodule detection. The clinical aspects of this study allowed us to find the most relevant features of the images defined by neck and head experts. The wealth of this database is given by the variety of diagnoses and cases, among thyroiditis, goiter, nodules and cancer. The complexity of the images and their interpretation reflect the need to develop strategies to support the diagnosis and the following.On the other hand, is relevant the possibility of having multiple datasets for the radiology community to validate retrospectively the TIRADS classification.

The DDTI main interest is based on the use of the images to apply CAD (Computer Aid Diagnosis) systems through the extraction of features from the nodule image. Each malignant feature can be detected using specific image analysis in the thyroid ultrasound to minimize the operator-dependent nature inherent in US images and make the diagnostic process reproducible.

The use of CAD systems in the ultrasound analysis, is related with the reduction of the inter-observer variability and the extraction of nodule features that can not to be obtain just with the visual component. The main objective of CAD in thyroid ultrasound is give a support to the expert to help him to improve the accuracy diagnosis and decrease the unnecessary thyroid biopsies. Nowadays, there are no available open access thyroid ultrasound-image databases with the associated diagnostic description and report of main findings.

5. FUTURE WORK

The Digital Database of Thyroid Ultrasound Images (DDTI) is a resource for the scientific community. Its open access will allow the use of multiple thyroid ultrasound images for their analysis, the application of CAD systems and algorithms.

Currently, DDTI contains 299 cases with 347 images. Each month a set of new cases will be downloaded with the expert annotation, pathological diagnosis and clinical description.More observers will be added. Currently , we have two head and neck experts from the same institution. In order to support the radiologist's diagnosis,

more institutios will be added too as well some patohology laboratories to perform the correlation among the histological and ultrasound components.

The main objective of the DDTI is developed an algorithm to support the diagnosis, pretending identify the main nodule features to reduce the variability inter-intra observer. This database will allowed to compare other CAD systems used in thyroid ultrasound images.

The DDTI will be use as a training tool for the radiology students of the Universidad Nacional de Colombia.

REFERENCES

1. N. C. Institute, "Statistics fact sheets: Thyroid cancer," *Surveillance, Epidemiology, and End Results Program Turning Cancer Data Into Discovery.*, 2010.
2. D. Koundal, S. Gupta, and S. Singh, "Computer aid diagnosis of thyroid nodule : A review," *International Journal of Computer Science and Engineering Surveys* **3**, August 2012.
3. A. G. Unnikrishnan and U. V. Menon, "Thyroid disorders in india: An epidemiological perspective," *Indian Journal of Endocrinology and Metabolism* **15**, pp. 78–81, 2011.
4. S. Choi, E. Kim, and J. Kwak, "Interobserver and intraobserver variations in ultrasound assesment of thyroid nodules," *Thyroid Radiology And Nuclear Medicine* **20**(2), 2010.
5. E. Horvath, S. Majilis, and R. Rossi, "An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management," *Journal Clinical Endocrinology and Metabolism* **90**, pp. 1748–1751, 2009.
6. W. Moon, J. S, and J. Lee, "Benign and malignt thyroid nodules: Us differentiation multicenter retrospective study," *Radiology* **247**, 2008.
7. J. Kwak, I. Jung, and et All, "Image reporting and characterization system for ultrasound features of thyroid nodules: Multricentric korean restrospective study.," *Korean Journal of Radiology* **14**, pp. 110–117, 2013.
8. J. Kwak, K. Han, and J. Yoon, "Thyroid imaging reporting and data systems for ultrasound features of nodules: A step in stablishing better stratification of cancer risk," *Radiology* **260**, September 2011.
9. E. S. Cibas and S. Z. Ali, "The bethesda system for reporting thyroid cytopathology," *American Journal for Clinical Pathology* **132**, 2009.

Bibliography

[1] ABDEL-NASSER, Mohamed ; RASHWAN, Hatem A. ; PUIG, Domenec ; MORENO, Antonio: Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern. En: *Expert Systems with Applications* 42 (2015), Nr. 24, p. 9499–9511

[2] ABUBAKER, Ayman A. ; S. QAHWAJI, Rami ; AQEL, Musbah J. ; SALEH, Mohmmad H.: Mammogram Image Size Reduction Using 16-8 bit Conversion Technique. En: *International Journal of Biomedical Sciences* 1 (2006), Nr. 3, p. 83–90

[3] AHMED M. ANTER, MOHAMED ABU ELSOUD, Aboul Ella H.: Automatic mammographic parenchyma classification according to BIRADS dictionary. En: *Computer Vision and Image Processing in Intelligent Systems and Multimedia Technologies*. 2014, p. 22–37

[4] AMENDOLIA, S R. ; BISOGNI, M G. ; BOTTIGLI, U ; CECCOPIERI, A ; DELOGU, P ; FANTACCI, M E. ; MARCHI, A ; MARZULLI, V M. ; PALMIERO, R ; STUMBO, S: The CALMA project: a CAD tool in breast radiography. En: *Nuclear Instruments and Methods in Physics Research* 460 (2001), p. 107–112

[5] AMERICAN CANCER SOCIETY: Breast Cancer / American Cancer Society. Atlanta, 2015. – Informe de Investigación. – Abstract p.. – ISBN 978–1–84996–313–8

[6] AVRIN, D ; MORIN, R ; PIRAINO, D ; ROWBERG, A ; DETORIE, N ; ZULEY, M ; J. A. SEIBERT ; PISANO., E D.: Storage, transmission, and retrieval of digital mammography, including recommendations on image compression. En: *Journal of the American College of Radiology* 3 (2006), p. 609–614

[7] AYRES, F J. ; RANGAYYAN, R M.: Characterization of architectural distortion in mammograms. En: *IEEE Enginering in Medicine and Biology Magazine* (2005), p. 59–67

[8] AYRES, F J. ; RANGAYYAN, R M.: Reduction of false positives in the detection of architectural distortion in mammograms by using a geometrically constrained phase portrait model. En: *International Journal of Computer Assisted Radiology and Surgery* 1 (2007), p. 361–369

[9] BAKER, Jay A. ; ROSEN, Eric L. ; LO, Joseph Y. ; GIMENEZ, Edgardo I. ; WALSH, Ruth ; SOO, Mary S.: Computer-Aided Detection (CAD) in Screening Mammography: Sensitivity of Commercial CAD Systems for Detecting Architectural Distortion. En: *American Journal of Roentgenology* 181 (2003), Nr. 4, p. 1083–1088

[10] BALANICA, Victor ; DUMITRACHE, Ioan ; PREZIOSI, Luigi: Breast cancer diagnosis based on spiculation feature and neural network techniques. En: *International Journal of Computers Communications & Control* 8 (2013), Nr. 3, p. 354–365

[11] BANIK, Shantanu ; M.RANGAYYAN, Rangaraj ; DESAUTELS, J E L.: Detection of Architectural Distortion in Prior Mammograms. En: *IEEE Transactions on Medical Imaging* 30 (2011), Nr. 2, p. 279–294

[12] BEHESHTI, S. M. A. ; AHMADINOUBARI, H. ; FATEMIZADEH, E. ; KHALILI, M.: An efficient fractal method for detection and diagnosis of breast masses in mammograms. En: *Journal of digital imaging* 27 (2014), oct, Nr. 5, p. 661–669. – ISSN 1618–727X

[13] BEUTEL, J ; KUNDEL, H ; METTER, R V. ; PRESS, SPIE (Ed.): *Phys. Psychophys*. Vol. 1: *Handbook of Medical Imaging*. Bellingham, WA, 2000

[14] BIRD, R. ; WALLACE, T. ; YANKASKAS, B.: Analysis of cancers missed at screening mammography. En: *Radiology* 178 (1992), p. 234–247

[15] BISWAS, Sujjoy K. ; MUKHERJEE, Dipti P.: Recognizing Architectural Distortion in Mammogram: A Multiscale Texture Modeling Approach with GMM. En: *IEEE Transactions on Biomedical Engineering* 58 (2011), Nr. 7, p. 2023–2030

[16] BLUEKENS, A M. ; HOLLAND, R ; KARSSEMEIJER, N ; BROEDERS, M J. ; DEN HEETEN, G J.: Comparison of digital screening mammography and screen-film mammography in the early detection of clinically relevant cancers: a multicenter study. En: *Radiology* 265 (2012), Nr. 3, p. 707–714

[17] BOVIK, Alan C. ; HUANG, Thomas S. ; MUNSON JR, David C.: A generalization of median filtering using linear combinations of order statistics. En: *Acoustics, Speech and Signal Processing, IEEE Transactions on* 31 (1983), Nr. 6, p. 1342–1350

[18] BOYER, Bruno ; CANALE, Sandra ; ARFI-ROUCHE, Julia ; MONZANI, Quentin ; KHALED, Wassef ; BALLEYGUIER, Corinne: Variability and errors when applying the BIRADS mammography classification. En: *European journal of radiology* 82 (2013), mar, Nr. 3, p. 388–397. – ISSN 1872–7727

[19] BRUNO, Daniel O T. ; DO NASCIMENTO, Marcelo Z. ; RAMOS, Rodrigo P. ; BATISTA, Valério R ; NEVES, Leandro A. ; MARTINS, Alessandro S.: *LBP* operators on curvelet coefficients as an algorithm to describe texture in breast cancer tissues. En: *Expert Systems with Applications* 55 (2016), p. 329–340

[20] BUSEMAN, S ; MOUCHAWAR, J ; CALONGE, N ; BYERS, T: Mammography screening matters for young women with breast carcinoma. En: *Cancer* 97 (2003), p. 352–358

[21] CANDES, Emmanuel ; DEMANET, Laurent ; DONOHO, David ; YING, Lexing: Fast Discrete Curvelet Transforms. En: *Multiscale Modeling and Simulation* 5 (2006), Nr. 3, p. 861–899

[22] CASCIO, Donato ; FAUCI, Francesco ; IACOMI, Marius ; RASO, Giuseppe ; MAGRO, Rosario ; CASTROGIOVANNI, Debora ; FILOSTO, Guido ; IENZI, Raffaele ; VASILE, Maria S.: Computer-aided diagnosis in digital mammography: comparison of two commercial systems. En: *Imagin in Medicine* 6 (2014), Nr. 1, p. 13–20

[23]   CHANG, C.C. ; LIN, C.J.: LIBSVM : a library for support vector machines. En: *ACM Transactions on Intelligent Systems and Technology* 2 (2011), Nr. 27, p. 1–27

[24]   CORTES, Corinna ; VAPNIK, Vladimir: Support-Vector Networks. En: *Mach. Learn.* 20 (1995), sep, Nr. 3, p. 273–297. – ISSN 0885–6125

[25]   DHAHBI, Sami ; BARHOUMI, Walid ; ZAGROUBA, Ezzeddine: Breast cancer diagnosis in digitized mammograms using curvelet moments. En: *Computers in biology and medicine* 64 (2015), jun, p. 79–90. – ISSN 1879–0534

[26]   DONG, Ping ; GALATSANOS, Nikolas P.: Affine transformation resistant watermarking based on image normalization. En: *Image Processing. 2002. Proceedings. 2002 International Conference on* Vol. 3 IEEE, 2002, p. 489–492

[27]   EL-NAPA, I ; YANG, Y ; GALATSANOS, N P. ; NISHIKAWA, R M. ; WERNICK, M N.: A similarity learning approach to content-based image retrieval: application to digital mammography. En: *IEEE Trans Med Imaging* 23 (2004), p. 1233–1244

[28]   ELSOUD, Mohamed A. ; ANTER, Ahmed M.: Automatic mammogram segmentation and computer aided diagnoses for breast tissue density according to BIRADS dictionary. En: *International Journal of Computer Aided Engineering and Technology* 4 (2012), mar, Nr. 2, p. 165–180

[29]   ELTOUKHY, Mohamed M. ; FAYE, Ibrahima ; BELHAOUARI, Brahim: A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation. En: *Computers in biology and medicine* 42 (2012), jan, Nr. 1, p. 123–128. – ISSN 1879–0534

[30]   ELTOUKHY, Mohamed M. ; FAYE, Ibrahima ; SAMIR, Brahim B.: A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram. En: *Computers in Biology and Medicine* 40 (2010), apr, Nr. 4, p. 384–391. – ISSN 1879–0534

[31]   FENTON, Joshua J. ; ABRAHAM, Linn ; TAPLIN, Stephen H. ; GELLER, Berta M. ; CARNEY, Patricia A. ; D'ORSI, Carl ; ELMORE, Joann G. ; BARLOW, William E.: Effectiveness of Computer-Aided Detection in Community Mammography Practice. En: *Journal of the National Cancer Institute* 103 (2011), p. 1152–1161

[32]   FENTON, Joshua J. ; TAPLIN, Stephen H. ; CARNEY, Patricia A. ; ABRAHAM, Linn ; SICKLES, Edward A. ; D'ORSI, Carl ; BERNS, Eric A. ; CUTTER, Gary ; HENDRICK, R E. ; BARLOW, William E. ; ELMORE, Joann G.: Influence of computer-aided detection on performance of screening mammography. En: *N. Engl. J. Med.* 356 (2007), p. 1399–1409

[33]   FERREIRA, Pedro ; DE CASTRO DUTRA, Inês ; FONSECA, Nuno A. ; WOODS, Ryan W. ; BURNSIDE, Elizabeth S.: Studying the Relevance of Breast Imaging Features. En: *HEALTHINF*, 2011, p. 337–342

[34]   FREEMAN, L C.: Centrality in social networks: Conceptual clarification. En: *Social Networks* 1 (1979), p. 215–239

[35] FRINTROP, S ; BACKER, G ; ROME, E: Goal-directed search with a top-down modulated computational attention system. En: *Pattern Recognit* (2005), p. 117–124

[36] GANESAN, Karthikeyan ; ACHARYA, U R. ; CHUA, Chua K. ; MIN, Lim C. ; ABRAHAM, K T. ; NG, Kwan-Hoong: Computer-aided breast cancer detection using mammograms: a review. En: *IEEE reviews in biomedical engineering* 6 (2013), jan, p. 77–98. – ISSN 1941–1189

[37] GOMEZ, Francisco ; ROMERO, Eduardo: Texture Characterization using a Curvelet Based Descriptor. En: *Proceedings of the 14th Iberoamerican Conference on Pattern Recognition: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, 2009

[38] GOPALAKRISHNAN, Viswanath ; HU, Yiqun ; RAJAN, Deepu: Random Walks on Graphs for Salient Object Detection in Images. En: *Image Processing, IEEE Transactions on* 19 (2010), p. 3232–3242

[39] GRIMM, Lars J. ; GHATE, Sujata V. ; YOON, Sora C. ; KUZMIAK, Cherie M. ; KIM, Connie ; MAZUROWSKI, Maciej A.: Predicting error in detecting mammographic masses among radiology trainees using statistical models based on BI-RADS features. En: *Medical physics* 41 (2014), mar, Nr. 3, p. 31909. – ISSN 0094–2405

[40] GUO, Q ; SHAO, J ; RUIZ, V: Investigation of support vector machine for the detection of architectural distortion in mammographic images. En: *Journal of Physics: Conference Series,2005* Vol. 15 IOP Publishing, 2005, p. 88

[41] GUPTA, Shalini ; CHYN, Priscilla F. ; MARKEY, Mia K.: Breast cancer CADx based on BI-RADS descriptors from two mammographic views. En: *Medical Physics* 33 (2006), Nr. 6, p. 1810–1817. – ISSN 00942405

[42] GUR, D ; STALDER, J S. ; HARDESTY, L A. ; ZHENG, B ; SUMKIN, J H. ; CHOUGH, D M. ; SHINDEL, B E. ; ROCKETTE, H E.: Computer-aided detection performance in mammographic examination of masses: assessment. En: *Radiology* 233 (2004), p. 418–423

[43] GUR, D ; SUMKIN, J H. ; ROCKETTE, H E. ; GANOTT, M ; HAKIM, C ; HARDESTY, L ; POLLER, W R. ; SHAH, R ; WALLACE, L: Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. En: *J Natl Cancer Inst* 96 (2004), p. 185–190

[44] HEATH, Michael ; BOWYER, Kevin ; KOPANS, Daniel ; MOORE, Richard ; KEGELMEYER, W P.: The Digital Database for Screening Mammography. En: *in Proceedings of the Fifth International Workshop on Digital Mammography, Medical Physics Publishing* M.J. Yaffe (2001), p. 212–218

[45] HORSCH, Alexander ; HAPFELMEIER, Alexander ; ELTER, Matthias: Needs assessment for next generation computer-aided mammography reference image databases and evaluation studies. En: *International Journal of Computer Assisted Radiology and Surgery* 6 (2011), Nr. 6, p. 749–767

[46] HOSNY, Khalid M.: Fast computation of accurate Zernike moments. En: *J Real-Time Image Proc* 3 (2008), p. 97–107

[47] HSU, Chih-Wei ; LIN, Chih-Jen: A comparison of methods for multiclass support vector machines. En: *IEEE Transactions on Neural Networks* 13 (2002), Nr. 2, p. 415–425

[48] HUPSE, R ; SAMULSKI, M ; LOBBES, M B. ; MANN, R M. ; MUS, R ; DEN HEETEN, G J. ; BEIJERINCK, D ; PIJNAPPEL, R M. ; BOETES, C ; KARSSEMEIJER, N: Computer-aided detection of masses at mammography: interactive decision support versus prompts. En: *Radiology* 266 (2013), Nr. 1, p. 123–129

[49] ICHIKAWA, T ; MATSUBARA, T ; HARA, T ; FUJITA, H ; ENDO, T ; IWASE, T: Automated detection method for architectural distorion areas on mammograms based on morphological processing and surface analysis. En: PROCESSING, Image (Ed.): *Medical Imaging* Vol. 5370, 2004

[50] IMAGECHECKER, R2: http://www.r2tech.com. 2008. – Informe de Investigación

[51] IMAGINIS: Mammographic screening is key to the early detection of breast cancer. / Imaginis. http://www.imaginis.com/breasthealth/screening.asp, acceded Janury 10-2016, 2016. – Technical report

[52] ISSAM EL NAQA, Yongyi Y.: The Role of Content-Based Image Retrieval in Mammography CAD. En: SUZUKI, Kenji (Ed.): *Computational Intelligence in Biomedical Imaging*. New York, NY : Springer New York, 2014, p. 33–53

[53] JIANG, Luan ; SONG, Enmin ; XU, Xiangyang ; MA, Guangzhi ; ZHENG, Bin: Automated Detection of Breast Mass Spiculation Levels and Evaluation of Scheme Performance. En: *Academic Radiology* 15 (2008), dec, Nr. 12, p. 1534–1544. – ISSN 1878–4046

[54] JIANG, Menglin ; ZHANG, Shaoting ; LI, Hongsheng ; METAXAS, Dimitris N.: Computer-aided diagnosis of mammographic masses using scalable image retrieval. En: *IEEE transactions on bio-medical engineering* 62 (2015), feb, Nr. 2, p. 783–792. – ISSN 1558–2531

[55] KACHOURI, Imene C. ; DJEMAL, Khalifa ; MAAREF, Hichem: Characterisation of mammographic masses using a new Spiculated Mass Descriptor in computer aided diagnosis systems. En: *International Journal of Signal and Imaging Systems Engineering* 5 (2012), jul, Nr. 2, p. 132—

[56] KAMRA, Amit ; JAIN, V K. ; SINGH, Sukhwinder ; MITTAL, Sunil: Characterization of Architectural Distortion in Mammograms Based on Texture Analysis Using Support Vector Machine Classifier with Clinical Evaluation. En: *Journal of Digital Imaging* (2015), p. 1–11. – ISSN 0897–1889

[57] KARSSEMEIJER, N ; TE BRAKE, G M.: Detection of stellate distortions in mammograms. En: *IEEE Transactions on Medical Imaging* 15 (1996), Nr. 5, p. 611–619

[58] KARSSEMEIJER, N ; OTTEN, J D M. ; RIJKEN, H ; HOLLAND, R: Computer aided detection of masses in mammograms as decision support. En: *The British Journal of Radiology* 79 (2006), Nr. special_issue_2, p. S123–S126

[59] KAUR, Simranjit ; SHARMA, Vipul ; SINGH, Sukhwinder ; GUPTA, Savita: A content based framework for mass retrieval in mammograms. En: AYLWARD, Stephen (Ed.) ; HADJIISKI, Lubomir M. (Ed.): *SPIE Medical Imaging*, International Society for Optics and Photonics, mar 2014, p. 90351N

[60] KAZEROUNI, I. A. ; HADDADNIA, J.: A mass classification and image retrieval model for mammograms. En: *The Imaging Science Journal* 62 (2014), sep, Nr. 7, p. 353–357. – ISSN 1368–2199

[61] KHOO, L A L. ; TAYLOR, P ; GIVEN-WILSON, R M.: Computer-aided detection in the United Kingdom national breast screening programme: prospective study. En: *Radiology.* 237(2) (2005), p. 444–449

[62] KINTNER, E C.: On the mathematical properties of the Zernike polynomials. En: *Opt. Acta* 23 (1976), Nr. 8, p. 679–680

[63] KISILEV, Pavel ; SASON, Eli ; BARKAN, Ella ; HASHOUL, Sharbell: Medical Image Description Using Multi-task-loss CNN. En: *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis* Springer, 2016, p. 121–129

[64] KISILEV, Pavel ; WALACH, Eugene ; HASHOUL, Sharbell ; BARKAN, Ella ; OPHIR, Boaz ; ALPERT, Sharon: Semantic description of medical image findings: structured learning approach. En: XIE, Xianghua (Ed.) ; JONES, Mark W. (Ed.) ; TAM, Gary K. L. (Ed.): *Proceedings of the British Machine Vision Conference (BMVC)*, BMVA Press, September 2015, p. 171.1–171.11

[65] KOM, Guillaume ; TIEDEU, Alain ; KOM, Martin: Automated detection of masses in mammograms by local adaptive thresholding. En: *Computers in Biology and Medicine* 37 (2007), p. 37–48

[66] KOPANS, D: *Breast Imaging*. 3rd. Williams & Wilkins: Philadelphia, 2007

[67] LAZARUS, Elizabeth ; MAINIERO, Martha B. ; SCHEPPS, Barbara ; KOELLIKER, Susan L. ; LIVINGSTON, Linda S.: BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value 1. En: *Radiology* 239 (2006), Nr. 2, p. 385–391

[68] LEI, Tianhu ; SEWCHAND, Wilfred: Statistical approach to X-ray CT imaging and its applications in image analysis. II. A new stochastic model-based image segmentation technique for X-ray CT image. En: *Medical Imaging, IEEE Transactions on* 11 (1992), Nr. 1, p. 62–69

[69] LI, L. ; CLARK, R.A. ; THOMAS, J.A.: Computer-aided diagnosis of masses with full-field digital mammography. En: *Academic Radiology* 9 (2002), p. 4–12

[70] DE LIMA, Sidney M. ; DA SILVA-FILHO, Abel G. ; DOS SANTOS, Wellington P.: Detection and classification of masses in mammographic images in a multi-kernel approach. En: *Computer Methods and Programs in Biomedicine* (2016)

[71] MASMOUDI, Alima D. ; AYED, Norhen Gargouri B. ; MASMOUDI, Dorra S. ; ABID, Riad:
Robust mass classification–based local binary pattern variance and shape descriptors. En:
International Journal of Signal and Imaging Systems Engineering 19 (2013), jan, p. 1–9

[72] MATSUBARA, T ; ITO, A ; TSUNOMORI, A ; HARA, T ; MURAMATSU, C ; ENDO, T ; FUJITA,
H: An automated method for detecting architectural distortions on mammograms using di-
rection analysis of linear structures. En: *Engineering in Medicine and Biology Society (EMBC),
2015 37th Annual International Conference of the IEEE*, 2015, p. 2661–2664

[73] MAZUROWSKI, Maciej A. ; BARNHART, Huiman X. ; BAKER, Jay A. ; TOURASSI, Georgia D.:
Identifying error-making patterns in assessment of mammographic BI-RADS descriptors
among radiology residents using statistical pattern recognition. En: *Academic radiology* 19
(2012), Nr. 7, p. 865–871

[74] MIRANDA, Gisele Helena B. ; FELIPE, Joaquim C.: Computer-aided diagnosis system based
on fuzzy logic for breast cancer categorization. En: *Computers in biology and medicine* 64
(2015), oct, p. 334–346. – ISSN 1879–0534

[75] MOAYEDI, Fatemeh ; AZIMIFAR, Zohreh ; BOOSTANI, Reza ; KATEBI, Serajodin: Contourlet-
based mammography mass classification using the SVM family. En: *Computers in biology
and medicine* 40 (2010), apr, Nr. 4, p. 373–383. – ISSN 1879–0534

[76] MOONA, Woo K. ; LO, Chung-Ming ; CHO, Nariya ; CHANG, Jung M. ; HUANG, Chiun-
Sheng ; CHEN, Jeon-Hor ; CHANG, Ruey-Feng: Computer-aided diagnosis of breast masses
using quantified BI-RADS findings. En: *Computer methods and programs in biomedicine* 111
(2013), jul, Nr. 1, p. 84–92. – ISSN 1872–7565

[77] MOREIRA, Ines C. ; AMARAL, Igor ; DOMINGUES, Inês ; CARDOSO, António ; CARDOSO,
Maria J. ; CARDOSO, Jaime S.: INbreast: toward a full-field digital mammographic database.
En: *Academic Radiology* 19 (2012), feb, Nr. 2, p. 236–48. – ISSN 1878–4046

[78] MORTON, Marilyn J. ; WHALEY, Dana H. ; BRANDT, Kathleen R. ; AMRAMI, Kimberly K.:
Screening Mammograms: Interpretation with Computer aided Detection Prospective Eval-
uation. En: *Radiology* 239 (2006), Nr. 2, p. 375–383

[79] MOURA, Daniel C. ; GUEVARA LÓPEZ, Miguel A.: An evaluation of image descriptors
combined with clinical data for breast cancer diagnosis. En: *International journal of computer
assisted radiology and surgery* 8 (2013), jul, Nr. 4, p. 561–574. – ISSN 1861–6429

[80] MURAMATSU, Chisako ; SCHMIDT, Robert A. ; SHIRAISHI, Junji ; LI, Qiang ; DOI, Kunio:
Presentation of Similar Images as a Reference for Distinction Between Benign and Malignant
Masses on Mammograms: Analysis of Initial Observer Study. En: *Journal of Digital Imaging*
23 (2010), Nr. 5, p. 592–602

[81] NARVÁEZ, Fabián ; ALVAREZ, Jorge ; GARCIA-ARTEAGA, Juan D. ; TARQUINO, Jonathan
; ROMERO, Eduardo: Characterizing Architectural Distortion in Mammograms by Linear
Saliency. En: *Journal of Medical Systems* 41 (2017), p. 1–26

[82] NARVÁEZ, Fabian ; DÍAZ, Gloria ; GÓMEZ, Francisco ; ROMERO, Eduardo: A content-based retrieval of mammographic masses using the curvelet descriptor. En: *Proc. SPIE* Vol. 8315, 2012, p. 83150A–83150A–7

[83] NARVAEZ, Fabian ; DIAZ, Gloria ; ROMERO, Eduardo: Automatic BI-RADS Description of Mammographic Masses. En: *Digital Mammography* Vol. 6136, 2010. – ISBN 978–3–642–13665–8, p. 673–681

[84] NARVÁEZ, Fabián ; DÍAZ, Gloris ; POVEDA, Cesar ; ROMERO, Eduardo: An automatic BI-RADS description of mammographic masses by fusing multiresolution features. En: *Expert Systems with Applications* 74 (2017), p. 82–95

[85] NEMOTO, M ; HONMURA, S ; SHIMIZU, A ; FURUKAWA, D ; KOBATAKE, H ; NAWANO, S: A pilot study of architectural distortion detection in mammograms based on characteristics of line shadows. En: *International Journal of Computer Assisted Radiology and Surgery* 4 (2009), Nr. 1, p. 27–36

[86] NISHIKAWA, R M. ; KALLERGI, M: Computer-aided detection in its present form is not an effective aid for screening mammography. En: *Med Phys* 33 (2006), p. 811–814

[87] NISHIKAWA, Robert M. ; GUR, David: CADe for Early Detection of Breast Cancer-Current Status and Why We Need to Continue to Explore New Approaches. En: *Academic Radiology* 21 (2014), Nr. 10, p. 1320–1321

[88] NITHYA, R. ; SANTHI, B.: Computer Aided Diagnosis System for Mammogram Analysis: A Survey. En: *Journal of Medical Imaging and Health Informatics* 5 (2015), aug, Nr. 4, p. 653–674. – ISSN 21567018

[89] OLIVER, Arnau ; TORTAJADA, Meritxell ; LLADÓ, Xavier ; FREIXENET, Jordi ; GANAU, Sergi ; TORTAJADA, Lidia ; VILAGRAN, Mariona ; SENTÍS, Melcior ; MARTÍ, Robert: Breast Density Analysis Using an Automatic Density Segmentation Algorithm. En: *Journal of digital imaging* (2015), feb. – ISSN 1618–727X

[90] PAPAKOSTAS, G A. ; BOUTALIS, Y S. ; KARRAS, D A. ; MERTZIOS, B G.: A new class of Zernike moments for computer vision applications. En: *Information Sciences* 177 (2007), Nr. 13, p. 2802–2819

[91] PARK, S C. ; PU, J ; ZHENG, B: Improving performance of computer-aided detection scheme by combining results from two machine learning classifiers. En: *Academic Radiology* 16 (2009), p. 266–274

[92] PARR, Tim C. ; TAYLOR, Christopher J. ; ASTLEY, Susan M. ; BOGGIS, Caroline R M.: Statistical modeling of oriented line patterns in mammograms. En: *Medical Imaging 1997* International Society for Optics and Photonics, 1997, p. 44–55

[93] PISANO, E D. ; ZONG, S ; HEMMINGER, B M. ; DELUCA, M ; JOHNSTON, R E. ; MULLER, K ; BRAEUNING, M P. ; PIZER, S M.: Contrast limited adaptive histogram equalization image

processing to improve the detection of simulated spiculations in dense mammograms. En: *Journal of Digital Imaging* 11 (1998), p. 193–200

[94] R., Highnam ; BRADY, M: *Mammographic Image Analysis*. Kluwer Academic Publisher, 1999

[95] RAMOS, Rodrigo P. ; ZANCHETTA DO NASCIMENTO, Marcelo ; PEREIRA, Danilo C.: Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms. En: *Expert Systems with Applications* 39 (2012), sep, Nr. 12, p. 11036–11047. – ISSN 09574174

[96] RANGAYYAN, Rangaraj M. ; CHAKRABORTY, Jayasree ; BANIK, Shantanu ; MUKHOPADHYAY, Sudipta ; DESAUTELS, J E L.: Detection of Architectural Distortion Using Coherence in Relation to the Expected Orientation of Breast Tissue. En: *IEEE Transactions on CBMS* 2 (2012), p. 1–4

[97] RANGAYYAN, RangarajM. ; BANIK, Shantanu ; DESAUTELS, J.E.Leo: Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer. En: *Journal of Digital Imaging* 23 (2010), Nr. 5, p. 611–631. – ISSN 0897–1889

[98] RANGAYYANA, M R. ; AYRES, Fabio J. ; DESAUTELSA, J E L.: A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. En: *Journal of the Franklin Institute* 344 (2007), p. 312–348

[99] RAO, A.R. ; JAIN, R.C.: Computerized flow field analysis: Oriented texture fields. En: *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (1992), Nr. 7, p. 693–709

[100] REDONDO, A. ; COMAS, M. ; MACIÀ, F. ; FERRER, F. ; MURTA-NASCIMENTO, C. ; MARISTANY, M.T. ; MOLINS, E. ; SALA, M. ; CASTELLS, X.: Inter- and intraradiologist variability in the BI-RADS assessment and breast density categories for screening mammograms. En: *The British journal of radiology* 85 (2012), nov, Nr. 1019, p. 1465–1470. – ISSN 1748–880X

[101] REYAD, Yasser A. ; BERBAR, Mohamed A. ; HUSSAIN, Muhammad: Comparison of statistical, LBP, and multi-resolution analysis features for breast mass classification. En: *Journal of medical systems* 38 (2014), Nr. 9, p. 1–15

[102] ROUHI, Rahimeh ; JAFARI, Mehdi ; KASAEI, Shohreh ; KESHAVARZIAN, Peiman: Benign and malignant breast tumors classification based on region growing and CNN segmentation. En: *Expert Systems with Applications* 42 (2015), feb, Nr. 3, p. 990–1002. – ISSN 09574174

[103] SAHINER, Berkman ; HADJIISKI, Lubomir M. ; CHAN, Heang-Ping ; PARAMAGUL, Chintana ; NEES, Alexis ; HELVIE, Mark ; SHI, Jiazheng: Concordance of computer-extracted image features with BI-RADS descriptors for mammographic mass margin. En: GIGER, Maryellen L. (Ed.) ; KARSSEMEIJER, Nico (Ed.): *Medical Imaging*, International Society for Optics and Photonics, mar 2008, p. 69151N

[104] SAMPAT, M P. ; MARKEY, M K. ; BOVIK, A C.: Measurement and detection of spiculated lesions. En: *Image Analysis and Interpretation, 2006*, 2006, p. 105–109

[105] SAMPAT, Mehul P. ; BOVIK, Alan C. ; MARKEY, Mia K.: Classification of mammographic lesions into BI-RADS shape categories using the beamlet transform. En: FITZPATRICK, J. M. (Ed.) ; REINHARDT, Joseph M. (Ed.): *SPIE Medical Imaging*, International Society for Optics and Photonics, apr 2005, p. 16–25

[106] SÉRADOUR, Brigitte ; HEID, Patrice ; ESTÈVE, Jacques: Comparison of direct digital mammography, computed radiography, and film-screen in the French national breast cancer screening program. En: *AJR. American journal of roentgenology* 202 (2014), jan, Nr. 1, p. 229–236. – ISSN 1546–3141

[107] SHARMA, Shubhi ; KHANNA, Pritee: Computer-aided diagnosis of malignant mammograms using Zernike moments and SVM. En: *Journal of digital imaging* 28 (2014), feb, Nr. 1, p. 77–90. – ISSN 1618–727X

[108] SICKLES, E A. ; DÓRSI, C J. ; BASSETT, L W. ; AL., Et: *ACR BI-RADS Mammography. In: ACR BI-RADS Atlas, Breast Imaging Reporting and Data System.* Fifth Edit. American College of Radiology, Reston, VA, 2013

[109] SOARES SÉRVULO DE OLIVEIRA, Fernando ; OSEAS DE CARVALHO FILHO, Antonio ; CORRÊA SILVA, Aristófanes ; CARDOSO DE PAIVA, Anselmo ; GATTASS, Marcelo: Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM. En: *Computers in biology and medicine* 57 (2015), feb, p. 42–53. – ISSN 1879–0534

[110] SOBTI, A ; SOBTI, P ; KEITH, L G.: Screening and diagnostic mammography: Why the gold standard does not shine more brightly? En: *Int. J Fert. Womens Med.* 50 (2005), p. 199–206

[111] SUCKLING, J. ; PARKER, J. ; DANCE, D. ; ASTLEY, S. ; HUTT, I. ; BOGGIS, C. ; RICKETTS, I. ; STAMATAKIS, E. ; CERNEAZ, N. ; KOK, S. ; TAYLOR, P. ; BETAL, D. ; SAVAGE, J.: The mammographic image analysis society digital mammogram database. En: *in Second International Workshop on Digital Mammography* Vol. 1069, 1994, p. 375–378

[112] SURENDIRAN, B. ; RAMANATHAN, P. ; VADIVEL, A.: Effect of BIRADS shape descriptors on breast cancer analysis. En: *International Journal of Medical Engineering and Informatics* 7 (2014), dec, Nr. 1, p. 65–79

[113] SURENDIRAN, B ; VADIVEL, A: Classifying Mammographic Masses into BI–RADS shape categories using various geometric shape and margin features. En: *International Journal of Biomedical Signal Processing, International Science Press* 2 (2011), Nr. 1, p. 43–47

[114] SURENDIRAN, B. ; VADIVEL, A.: Mammogram mass classification using various geometric shape and margin features for early detection of breast cancer. En: *International Journal of Medical Engineering and Informatics* 4 (2012), feb, Nr. 1, p. 36–54

[115] TABAR, L ; YENG, M F. ; VITAK, B ; CHENG, H H T. ; SMITH, R A. ; DUFFY, S W.: Mammography service screening and mortality in breast cancer patients: 20-year follow\-up before and after introduction of screening. En: *Lacent* 361 (2003), p. 1405–1410

[116] TAHMASBI, Amir ; SAKI, Fatemeh ; SHOKOUHI, Shahriar B.: Classification of benign and malignant masses based on Zernike moments. En: *Computers in biology and medicine* 41 (2011), aug, Nr. 8, p. 726–35. – ISSN 1879–0534

[117] TANG, J ; RANGAYYAN, R M. ; XU, J ; NAQA, I E. ; YANG, Y: Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. En: *IEEE Trans. Inform. Technol. Biomed* 13 (2009), p. 236–251

[118] TAO, Yimo ; LO, Shih-Chung B. ; HADJISKI, Lubomir ; CHAN, Heang-Ping ; FREEDMAN, Matthew T.: BI-RADS guided mammographic mass retrieval. En: SUMMERS, Ronald M. (Ed.) ; VAN GINNEKEN, Bram (Ed.): *SPIE Medical Imaging*, International Society for Optics and Photonics, mar 2011, p. 79632H–79632H–6

[119] TIMMERS, J M H. ; VAN DOORNE-NAGTEGAAL, H J. ; VERBEEK, A L M. ; DEN HEETEN, G J. ; BROEDERS, M J M.: A dedicated BI-RADS training programme: effect on the inter-observer variation among screening radiologists. En: *European journal of radiology* 81 (2012), sep, Nr. 9, p. 2184–2188. – ISSN 1872–7727

[120] TOURASSI, G D. ; DELONG, D M. ; FLOYD, C E.: A study on the computerized fractal analysis of architectural distortion in screening mammograms. En: *Physics in Medicine and Biology* 51 (2006), Nr. 5, p. 1299–1312

[121] VADIVEL, A ; SURENDIRAN, B: A fuzzy rule-based approach for characterization of mammogram masses into BI-RADS shape categories. En: *Computers in biology and medicine* 43 (2013), may, Nr. 4, p. 259–267. – ISSN 1879–0534

[122] VARMA, M ; BABU, B: More generality in efficient multiple kernel learning. En: *ACM 26th Annu. Int. Conf. Mach. Learn*, 2009, p. 1065–1072

[123] VIBHA, L ; HARSHAVARDHAN, GM ; PRANAW, K ; SHENOY, P D. ; VENUGOPAL, KR ; PATNAIK, Lalit M.: Classification of mammograms using decision trees. En: *2006 10th International Database Engineering and Applications Symposium (IDEAS'06)* IEEE, 2006, p. 263–266

[124] VYBORNY, C J. ; GIGER, M L. ; NISHIKAWA, R M.: Computer-aided detection and diagnosis of breast cancer. En: *Radiology Clinic North America* 38 (2000), p. 725–740

[125] WANG, Xingwei ; LI, Lihua ; LIU, Wei ; XU, Weidong ; LEDERMAN, Dror ; ZHENG, Bin: An interactive system for computer-aided diagnosis of breast masses. En: *Journal of Digital Imaging* 25 (2012), oct, Nr. 5, p. 570–579. – ISSN 1618–727X

[126] WEE, Chong-Yaw Y. ; PARAMESRAN, Raveendran: On the computational aspects of Zernike moments. En: *Image and Vision Computing* 25 (2007), Nr. 6, p. 967–980

[127] WEI, Chia-Hung ; CHEN, Sherry Y. ; LIU, Xiaohui: Mammogram retrieval on similar mass lesions. En: *Computer Methods and programs in biomedicine* 106 (2012), jun, Nr. 3, p. 234–248. – ISSN 1872–7565

[128] WEI, Chia-Hung ; LI, Yue ; HUANG, Pai J.: Mammogram retrieval through machine learning within BI-RADS standards. En: *Journal of Biomedical Informatics* 44 (2011), aug, Nr. 4, p. 607–614. – ISSN 1532–0480

[129] WOODS, Ryan W. ; SISNEY, Gale S. ; SALKOWSKI, Lonie R. ; SHINKI, Kazuhiko ; LIN, Yunzhi ; BURNSIDE, Elizabeth S.: The mammographic density of a mass is a significant predictor of breast cancer. En: *Radiology* 258 (2011), feb, Nr. 2, p. 417–25. – ISSN 1527–1315

[130] WU, Yirong ; ALAGOZ, Oguzhan ; AYVACI, Mehmet U S. ; MUNOZ DEL RIO, Alejandro ; VANNESS, David J. ; WOODS, Ryan ; BURNSIDE, Elizabeth S.: A comprehensive methodology for determining the most informative mammographic features. En: *Journal of digital imaging* 26 (2013), oct, Nr. 5, p. 941–7. – ISSN 1618–727X

[131] ZHANG, Xinsheng: A New Ensemble Learning Approach for Microcalcification Clusters Detection. En: *Journal of Software* 4 (2009), Nr. 9, p. 1014–2021

[132] ZHENG, B ; SUMKIN, J H. ; ZULEY, M L. ; LEDERMAN, D ; WANG, X ; GUR, D: Computer-aided detection of breast masses depicted on full-field digital mammograms: a performance assessment. En: *The British Journal of Radiology* 84 (2012), Nr. 1014, p. e153–e161

[133] ZHENG, Bin ; MELLO-THOMS, Claudia ; WANG, Xiao-Hui ; ABRAMS, Gordon S. ; SUMKIN, Jules H. ; CHOUGH, Denise M. ; GANOTT, Marie A. ; LU, Amy ; GUR, David: Interactive computer aided diagnosis of breast masses: computerized selection of visually similar image sets from a reference library. En: *Academical Radiology* 14 (2007), aug, Nr. 8, p. 917–927. – ISSN 1076–6332

[134] ZWIGGELAAR, Reyer ; ASTLEY, Susan M. ; BOGGIS, Caroline R M. ; TAYLOR, Christopher J.: Linear Structures in Mammographic Images: Detection and Classification. En: *IEEE Transactions on Medical Imaging* 23 (2004), Nr. 9, p. 1077–1086

[135] ZWIGGELAAR, Reyes ; PARR, Timothy C. ; SCHUMM, James E. ; HUTT, Ian W. ; TAYLOR, Christopher J. ; ASTLEY, Susan M. ; BOGGIS, Caroline R M.: Model-based detection of spiculated lesions in mammograms. En: *Medical Image Analysis* 3 (1999), Nr. 1, p. 39–62