Computerized Analysis of Mammographic Images for Detection and Characterization of Breast Cancer

Synthesis Lectures on Biomedical Engineering

Lectures in Biomedical Engineering will be comprised of 75- to 150-page publications on advanced and state-of-the-art topics that span the field of biomedical engineering, from the atom and molecule to large diagnostic equipment. Each lecture covers, for that topic, the fundamental principles in a unified manner, develops underlying concepts needed for sequential material, and progresses to more advanced topics. Computer software and multimedia, when appropriate and available, are included for simulation, computation, visualization and design. The authors selected to write the lectures are leading experts on the subject who have extensive background in theory, application and design.
The series is designed to meet the demands of the 21st century technology and the rapid advancements in the all-encompassing field of biomedical engineering that includes biochemical processes, biomaterials, biomechanics, bioinstrumentation, physiological modeling, biosignal processing, bioinformatics, biocomplexity, medical and molecular imaging, rehabilitation engineering, biomimetic nano-electrokinetics, biosensors, biotechnology, clinical engineering, biomedical devices, drug discovery and delivery systems, tissue engineering, proteomics, functional genomics, and molecular and cellular engineering.

Digital Image Processing for Ophthalmology: Detection and Modeling of Retinal Vascular Architecture
Faraz Oloumi, Rangaraj M. Rangayyan, and Anna L. Ells
2014

Biomedical Signals and Systems
Joseph V. Tranquillo
2013

Health Care Engineering, Part II: Research and Development in the Health Care Environment
Monique Frize
2013

Health Care Engineering, Part I: Clinical Engineering and Technology Management
Monique Frize
2013

Computer-aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer
Shantanu Banik, Rangaraj M. Rangayyan, and J.E. Leo Desautels
2013

Content-based Retrieval of Medical Images: Landmarking, Indexing, and Relevance Feedback
Paulo Mazzoncini de Azevedo-Marques and Rangaraj Mandayam Rangayyan
2013

Chronobioengineering: Introduction to Biological Rhythms with Applications, Volume 1
Donald McEachron
2012

Medical Equipment Maintenance: Management and Oversight
Binseng Wang
2012

Fractal Analysis of Breast Masses in Mammograms
Thanh M. Cabral and Rangaraj M. Rangayyan
2012

Capstone Design Courses, Part II: Preparing Biomedical Engineers for the Real World
Jay R. Goldberg
2012

Ethics for Bioengineers
Monique Frize
2011

Computational Genomic Signatures
Ozkan Ufuk Nalbantoglu and Khalid Sayood
2011

Digital Image Processing for Ophthalmology: Detection of the Optic Nerve Head
Xiaolu Zhu, Rangaraj M. Rangayyan, and Anna L. Ells
2011

Modeling and Analysis of Shape with Applications in Computer-Aided Diagnosis of Breast Cancer
Denise Guliato and Rangaraj M. Rangayyan
2011

Analysis of Oriented Texture with Applications to the Detection of Architectural Distortion in Mammograms
Fábio J. Ayres, Rangaraj M. Rangayyan, and J. E. Leo Desautels
2010

Fundamentals of Biomedical Transport Processes
Gerald E. Miller
2010

Models of Horizontal Eye Movements, Part II: A 3rd Order Linear Saccade Model
John D. Enderle and Wei Zhou
2010

Models of Horizontal Eye Movements, Part I: Early Models of Saccades and Smooth Pursuit
John D. Enderle
2010

The Graph Theoretical Approach in Brain Functional Networks: Theory and Applications
Fabrizio De Vico Fallani and Fabio Babiloni
2010

Biomedical Technology Assessment: The 3Q Method
Phillip Weinfurt
2010

Strategic Health Technology Incorporation
Binseng Wang
2009

Phonocardiography Signal Processing
Abbas K. Abbas and Rasha Bassam
2009

Introduction to Biomedical Engineering: Biomechanics and Bioelectricity - Part II
Douglas A. Christensen
2009

Introduction to Biomedical Engineering: Biomechanics and Bioelectricity - Part I
Douglas A. Christensen
2009

Landmarking and Segmentation of 3D CT Images
Shantanu Banik, Rangaraj M. Rangayyan, and Graham S. Boag
2009

Basic Feedback Controls in Biomedicine
Charles S. Lessard
2009

Understanding Atrial Fibrillation: The Signal Processing Contribution, Part I
Luca Mainardi, Leif Sörnmo, and Sergio Cerutti
2008

Understanding Atrial Fibrillation: The Signal Processing Contribution, Part II
Luca Mainardi, Leif Sörnmo, and Sergio Cerutti
2008

Introductory Medical Imaging
A. A. Bharath
2008

Lung Sounds: An Advanced Signal Processing Perspective
Leontios J. Hadjileontiadis
2008

An Outline of Informational Genetics
Gérard Battail
2008

Neural Interfacing: Forging the Human-Machine Connection
Susanne D. Coates
2008

Quantitative Neurophysiology
Joseph V. Tranquillo
2008

Tremor: From Pathogenesis to Treatment
Giuliana Grimaldi and Mario Manto
2008

Introduction to Continuum Biomechanics
Kyriacos A. Athanasiou and Roman M. Natoli
2008

The Effects of Hypergravity and Microgravity on Biomedical Experiments
Thais Russomano, Gustavo Dalmarco, and Felipe Prehn Falcão
2008

A Biosystems Approach to Industrial Patient Monitoring and Diagnostic Devices
Gail Baura
2008

Multimodal Imaging in Neurology: Special Focus on MRI Applications and MEG
Hans-Peter Müller and Jan Kassubek
2007

Estimation of Cortical Connectivity in Humans: Advanced Signal Processing Techniques
Laura Astolfi and Fabio Babiloni
2007

Brain-Machine Interface Engineering
Justin C. Sanchez and José C. Principe
2007

Introduction to Statistics for Biomedical Engineers
Kristina M. Ropella
2007

Capstone Design Courses: Producing Industry-Ready Biomedical Engineers
Jay R. Goldberg
2007

BioNanotechnology
Elisabeth S. Papazoglou and Aravind Parthasarathy
2007

Bioinstrumentation
John D. Enderle
2006

Fundamentals of Respiratory Sounds and Analysis
Zahra Moussavi
2006

Advanced Probability Theory for Biomedical Engineers
John D. Enderle, David C. Farden, and Daniel J. Krause
2006

Intermediate Probability Theory for Biomedical Engineers
John D. Enderle, David C. Farden, and Daniel J. Krause
2006

Basic Probability Theory for Biomedical Engineers
John D. Enderle, David C. Farden, and Daniel J. Krause
2006

Sensory Organ Replacement and Repair
Gerald E. Miller
2006

Artificial Organs
Gerald E. Miller
2006

Signal Processing of Random Physiological Signals
Charles S. Lessard
2006

Image and Signal Processing for Networked E-Health Applications
Ilias G. Maglogiannis, Kostas Karpouzis, and Manolis Wallace
2006

Computerized Analysis of Mammographic Images for Detection and Characterization of Breast Cancer

Paola Casti, Arianna Mencattini, Marcello Salmeri, and Rangaraj M. Rangayyan

www.morganclaypool.com

ISBN: 9781681731568 paperback
ISBN: 9781681731575 ebook

DOI 10.2200/S00776ED1V01Y201704BME056

A Publication in the Morgan & Claypool Publishers series
SYNTHESIS LECTURES ON BIOMEDICAL ENGINEERING

Lecture #56
Series Editor: John D. Enderle, *University of Connecticut*
Series ISSN
Print 1930-0328 Electronic 1930-0336

Computerized Analysis of Mammographic Images for Detection and Characterization of Breast Cancer

Paola Casti, Arianna Mencattini, and Marcello Salmeri
University of Rome Tor Vergata, Rome, Italy

Rangaraj M. Rangayyan
University of Calgary, Calgary, Alberta, Canada

SYNTHESIS LECTURES ON BIOMEDICAL ENGINEERING #56

MORGAN & CLAYPOOL PUBLISHERS

ABSTRACT

The identification and interpretation of the signs of breast cancer in mammographic images from screening programs can be very difficult due to the subtle and diversified appearance of breast disease. This book presents new image processing and pattern recognition techniques for computer-aided detection and diagnosis of breast cancer in its various forms. The main goals are: (1) the identification of bilateral asymmetry as an early sign of breast disease which is not detectable by other existing approaches; and (2) the detection and classification of masses and regions of architectural distortion, as benign lesions or malignant tumors, in a unified framework that does not require accurate extraction of the contours of the lesions. The innovative aspects of the work include the design and validation of landmarking algorithms, automatic Tabár masking procedures, and various feature descriptors for quantification of similarity and for contour-independent classification of mammographic lesions. Characterization of breast tissue patterns is achieved by means of multidirectional Gabor filters. For the classification tasks, pattern recognition strategies, including Fisher linear discriminant analysis, Bayesian classifiers, support vector machines, and neural networks are applied using automatic selection of features and cross-validation techniques. Computer-aided detection of bilateral asymmetry resulted in accuracy up to 0.94, with sensitivity and specificity of 1 and 0.88, respectively. Computer-aided diagnosis of automatically detected lesions provided sensitivity of detection of malignant tumors in the range of [0.70, 0.81] at a range of falsely detected tumors of [0.82, 3.47] per image. The techniques presented in this work are effective in detecting and characterizing various mammographic signs of breast disease.

KEYWORDS

angular stationarity, architectural distortion, bilateral asymmetry, breast cancer, computer-aided diagnosis (CAD), correlation analysis, feature selection, Gabor filters, Gaussian curvature, landmarking, mammography, masses, pattern recognition, radial stationarity, structural similarity, semivariogram analysis, Tabár masking

Paola Casti dedicates this book to her son Matteo

Contents

Preface

This work presents new image processing and pattern recognition strategies for detection and diagnosis of breast cancer with mammography. The first part of the book focuses on a computerized system for the identification of bilateral asymmetry as an early sign of tumor which is not detectable by other existing approaches. With the purpose of detecting bilateral asymmetry, novel landmarking procedures are presented to identify anatomical structures on the mammogram, including the nipple and the breast contour. Together with the pectoral muscle, they serve as landmarks for bilateral matching and, in addition, as boundary structures to confine the breast region where the abnormalities are localized. Following radiologists' criteria in interpreting mammograms, computerized Tabár masking procedures are described as a means to derive corresponding regions of the left and right mammograms for comparison and analysis. The extraction of the oriented patterns is performed to characterize the breast tissue patterns. Structural similarity or dissimilarity between paired regions on the mammograms are then quantified by means of various specifically designed measures of similarity. The feature descriptors developed for the application consist of: (1) a novel application of Moran's index to measure the angular covariance between rose diagrams related to the phase and magnitude responses of multidirectional Gabor filters; (2) features for the analysis of spatial correlation of pixel values with respect to the nipple position; and (3) spherical semivariogram descriptors and new correlation-based structural similarity indices in the spatial and complex wavelet domains.

In the second part of the book, the development of a unified and comprehensive computerized system for detection and diagnosis of breast cancer is presented, in which contour-independent diagnosis of malignant tumors, including masses and regions of architectural distortion, is performed on automatically detected suspicious focal areas. Analysis of the gradient vector field via the eigenvalues of the Hessian is performed to identify the focal areas, while a differential approach is implemented to derive features for detection of lesions. New feature descriptors are designed for quantification of 2D spatial correlation and trends over the radial and angular directions of circular regions including a lesion. They serve the purpose of contour-independent classification of candidates as benign lesions or malignant tumors. Finally, a 3D free-response receiver operating characteristic framework is introduced for evaluation of two binary categorization problems in series.

Sequential forward/backward selection and stepwise logistic regression are used for automatic selection of the various extracted features. Pattern recognition techniques, including Fisher linear discriminant analysis, Bayesian classifiers, support vector machines, and neural networks, are applied to experimental training data and the obtained models are used for the automatic classification of test data. The effectiveness of the developed systems is demonstrated

through cross-validation techniques such as leave-one-patient-out and k-fold analysis. Two public databases, Digital Database for Screening Mammography (DDSM) and Mammographic Image Analysis Society (MIAS) database, together with a private database of full-field digital mammograms (FFDMs) from San Paolo Hospital of Bari, Italy, are used for the analysis. Multiple comparisons of the results achieved in this work with the results reported in previous research work are reported.

Computer-aided detection (CADe) of bilateral asymmetry has not been studied adequately. There is increasing interest in this area, as indicated by the appearance of publications addressing the problem. The performance of the methods developed in this work for the detection of bilateral asymmetry resulted in accuracy up to 0.94, with sensitivity and specificity of 1 and 0.88, respectively; the obtained results are better than the results reported in other works in the scientific literature and are expected to improve the performance of techniques for mammography and breast cancer.

Computer-aided diagnosis (CADx) of mammographic lesions, in particular of automatically detected masses and regions of architectural distortion, is an important but yet-to-be addressed task that can facilitate accurate interpretation of mammograms and reduce unnecessary breast biopsies. Some of the related subproblems have been addressed by researchers as independent tasks. The integration of the various aspects of detection and classification of mammographic lesions pose a new challenge to be addressed, which demands the design of a unified CADe/CADx system. Moreover, the presence of tumors with obscured or ill-defined margins, for which the existing approaches based on accurate segmentation of the lesions are prone to fail, has motivated the contour-independent approach presented in this work. The results obtained with the CADe/CADx system indicate sensitivity of detection of malignant tumors in the range of [0.70–0.81] at a range of falsely detected tumors of [0.82–3.47] per image. The results obtained with FFDMs, in particular, compared favorably with the performance of the existing commercial systems for the automatic detection of masses, which is a simpler problem than integrating detection and classification of lesions. The methods presented in this work are expected to improve the scope and performance of CADe/CADx systems for breast cancer.

Acknowledgments

The research work presented in this book would not have been possible without the valuable support of many individuals and organizations.

We thank the Diagnostic Radiology Unit, San Paolo Hospital of Bari, Italy, in particular Dr. Antonietta Ancona, Dr. Fabio F. Mangieri, and Dr. Maria Luisa Pepe from the Integrated Operational Unit of Diagnostic Imaging, ASL, Taranto, Italy, and Prof. Maria Grazia Raguso from the Department of Mathematics, University of Bari, Italy, for providing the digital mammograms used in this work, the ground truth annotations, and the support needed for validation of results.

We thank J. Pont and E. Pérez from the Department of Radiology of the Girona University Hospital "Dr. Josep Trueta" (Spain) and E.R.E. Denton from the Department of Breast Imaging of the Norwich and Norfolk University Hospital (UK) for providing the BI-RADS classification of the MIAS database used in this work. We also thank Dr. Ricardo J. Ferrari from Departamento de Computação, Universidade Federal de São Carlos, São Paulo, Brasil, and his colleagues for having made available additional ground-truth contours of the MIAS database.

We gratefully thank the Department of Electronics Engineering of the University of Rome Tor Vergata, Italy, including the affiliated professors, researchers, technical and administrative personnel, and students who supported our work and enriched it with their unique contributions.

Some of the materials and illustrations have been reproduced, with permission, from the associated organizations, from our publications listed below.

1. P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangieri, M.L. Pepe, and R.M. Rangayyan, Contour-independent detection and classification of mammographic lesions, *Biomed. Signal Process. Control*, 25:165–177, 2016. © Elsevier.

2. P. Casti, A. Mencattini, M. Salmeri, and R.M. Rangayyan, Analysis of structural similarity in mammograms for detection of bilateral asymmetry, *IEEE Trans. Med. Imag.*, 34(2), 662–671, 2015. © IEEE.

3. P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangieri, M.L. Pepe, and R.M. Rangayyan, Estimation of the breast skin-line in mammograms using multidirectional Gabor filters, *Comput. Biol. Med.*, 43(11), 1870–1881, 2013. © Elsevier.

4. P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangieri, M.L. Pepe, and R.M. Rangayyan, Automatic detection of the nipple in screen-film and full-field digital mam-

mograms using a novel Hessian-based method, *J. Dig. Imag.*, 26(5), 948–957, 2013. © Springer.

5. P. Casti, A. Mencattini, M. Salmeri, and R.M. Rangayyan, Masking procedures and measures of angular similarity for detection of bilateral asymmetry in mammograms, 4*th IEEE Conference on e-Health and Bioengineering (EHB)*, Iasi, Romania, 2013. © IEEE.

6. P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangieri, M.L. Pepe, and R.M. Rangayyan, Design and analysis of contour-independent features for classification of mammographic lesions, *Proc. of the* 4*th IEEE Conference on e-Health and Bioengineering (EHB)*, Iasi, Romania, 2013. © IEEE.

7. P. Casti, A. Mencattini, M. Salmeri, F. Mangieri, and R.M. Rangayyan, Measures of radial correlation and radial trends for classification of breast masses in mammograms, *Proc. of the* 35*th IEEE Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 2013. © IEEE.

8. A. Mencattini, M. Salmeri, P. Casti, and M.L. Pepe, Local active contour models and Gabor wavelets for an optimal breast region segmentation, *Proc. of the* 26*th International Congress and Exhibition: Computer Assisted Radiology and Surgery (CARS)*, Pisa, Italy, 2012. © Springer.

We thank our families, relatives, and friends for their loving support, encouragement, and understanding, and for making our lives meaningful.

Paola Casti, Arianna Mencattini, Marcello Salmeri, and Rangaraj M. Rangayyan
May 2017

CHAPTER 1

Introduction

1.1 BREAST CANCER AND MAMMOGRAPHY

1.1.1 BREAST CANCER STATISTICS

With 1.67 million new cases in 2012, breast cancer represents 25% of all diagnoses of cancer cases worldwide. The estimates for 2015 indicate 1.70 million of new cases [64]. Among women, breast cancer is the most frequent form of cancer and it is second only to lung cancer as the most frequent cause of cancer death worldwide. According to the ISTAT [66], breast cancer is both the most frequent form of cancer and the most common cause of cancer death in all age groups of Italian women: it accounts for 28% of deaths among young women, 21% among adults, and 14% among women aged more than 70 years.

When breast cancer is diagnosed at an early stage, the prognosis for the patient is favorable and surgery can be resolutive even if limited to the lesion and its surrounding tissue. At higher stages of the disease, the surgical procedure may need to be followed by radiation therapy or, in some cases, chemotherapy, which have the purpose of protecting the remaining glandular tissue from the risk of local recurrence of cancer [137]. However, 25% of women with breast cancer present are diagnosed with advanced forms of the disease that need to be treated with more aggressive treatments consisting of the removal of the entire breast, named radical mastectomy, and involving chemotherapy before surgery, followed by additional chemotherapy and radiation; in such cases, the chances of survival are drastically reduced [84].

1.1.2 MAMMOGRAPHY SCREENING PROGRAMS

Through the years, researchers have learned that early diagnosis is critical to cure breast cancer and that access to screening tools by women facilitates efficient treatments for most patients [40]. If the cancer is detected at an early stage, in fact, more treatment options are available and the patient's life can be saved. At present, the best radiographic method for detecting breast cancer is mammography. Therefore, cancer screening is best provided by the combination of mammography and clinical breast examination, performed at standard intervals. The majority of European countries adopted national or regional mammography screening programs, which consist of periodical mammographic examinations of asymptomatic women aged 50–69 years with a 2-year screening interval [112]. The accumulated evidence [144] suggests that a reduction of 20% in breast cancer mortality can be achieved by organized mammography programs. A recent study of the IMPACT Working Group [51] has evaluated the effects of the national

mammographic screening programs on the incidence of breast cancers at advanced stages in Italy, showing a significant and stable reduction in breast cancer-related specific mortality in the range of 20–30%.

1.1.3 THE MAMMOGRAPHIC EXAMINATION

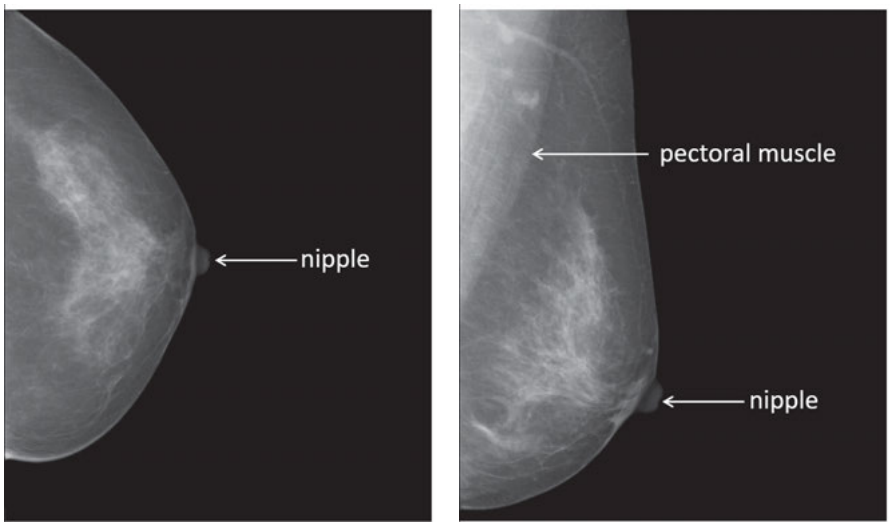
The beginning of mammographic examinations dates back to 1913, when the German surgeon Albert Salomon obtained the first X-ray image of breast tissue removed from cancer patients [115]. Almost 20 years later, the American radiologist Stafford L. Warren performed the first in vivo mammography by means of an apparatus for conventional X-ray imaging. Since then, physicians all over the world have contributed to improvements of the technique, and today, mammography has become the recognized technique for detection of breast cancer worldwide [115].

Mammography consists of an X-ray examination of the breast made with a specific X-ray equipment that is capable of finding tumors too small to be palpable. The output is a shadowgram of the breast which is recorded by an image receptor and that results from the attenuation of X-rays along paths passing through the glandular structures. The projected structures are magnified onto the image receptor due to the spreading of the X-rays from the source, while the differential X-ray attenuation among the various tissue structures is responsible for the image contrast. The magnification of the internal structures of the breast and the obtained contrast enable, in the presence of pathological processes, the identification of alteration of the mammary gland. With this purpose, the spatial resolution of the mammographic system can be as small as 10 μm in order to discriminate fine details and low-energy X-rays in the range of 24–32 keV are used to obtain good contrast. Limits to the image quality are posed by the restrictions on the radiation dose absorbed by the patient, which should be less than 3 mGy, and by an overall random fluctuation, referred to as mottle or noise. Such limitations, together with the superimposition of tissue during projection, contribute to make the mammographic interpretation difficult even by experienced radiologists [14].

Screen-film mammography (SFM) and full-field digital mammography (FFDM) are the two modalities for performing mammographic examinations nowadays. In SFM, which is the conventional analog system, the X-ray photons are converted to light by a phosphorescent screen; the light image is captured by a film, which is then developed for interpretation by the radiologist. Digital images can be produced by digitization after the film is processed. In FFDM, the X-ray photons are converted by solid-state detectors into electronic signals that can be displayed directly on a high-resolution monitor. SFM had been the modality of choice for screening programs; however, FFDM is replacing conventional film imaging systems, mainly due to increased quality of images and a wider dynamic range, in addition to the benefits of digital technology in data transmission, retrieval, display, and storage [62].

During screening mammography, the left and right breasts of a woman are imaged separately. Two views are obtained by compressing each breast along different directions of projec-

tions to cover the majority of fibroglandular structures. The craniocaudal (CC) view is acquired by a vertical projection of the breast, which is compressed in the head-to-toe direction. The mediolateral-oblique (MLO) view is taken by compressing the breast from the middle of the chest to the outside of the body in an oblique direction. In Figs. 1.1a and b, two normal mammograms of the right breast of a patient in the CC and MLO projections, respectively, are shown. The brighter areas at the center of the images correspond to the fibroglandular tissue. The nipple is in profile and visible in both views, while the pectoral muscle is visible only in the MLO view. The mammograms were acquired at the San Paolo Hospital of Bari, Italy. More information on the related database and how the images were obtained are provided in Section 2.1.1.



(a) (b)

Figure 1.1: (a) The craniocaudal (CC) view and (b) the mediolateral-oblique (MLO) view of the right breast of a patient. The mammograms show normal breast parenchyma with normal radiographic density. The FFDM images were acquired at the San Paolo Hospital of Bari, Italy (see Section 2.1.1 for more details).

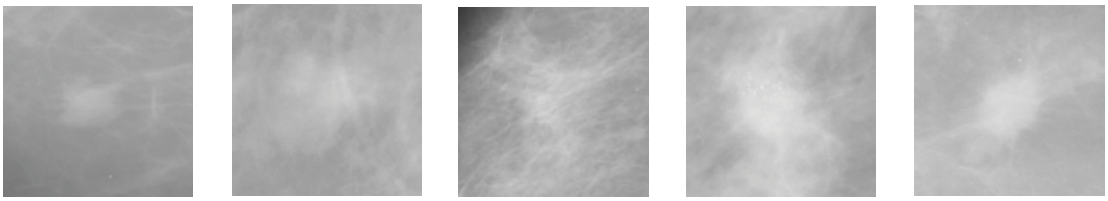
1.1.4 MAMMOGRAPHIC SIGNS OF BREAST DISEASE

The Breast Imaging Reporting and Data System (BI-RADS) of the American College of Radiology (ACR) [39], which is the standard reference for reporting mammographic results, describes four categories of breast abnormalities that can indicate breast cancer: masses, calcifications, architectural distortions, and bilateral asymmetry.

Masses

A mass is defined as a space occupying lesion seen in two different projections [39]. When a potential mass is seen in a single view it is named "density" until its nature is confirmed. Masses usually appear as areas brighter than the surrounding tissue due to increased attenuation of X-rays if they possess higher density, but they can also result in equal density (isodense) or less density (hypodense) regions on the mammogram. There are also fat containing masses which appear as radiolucent regions. The margins of masses are classified as circumscribed, microlobulated, obscured, indistinct, or spiculated [39]. Examples of masses from each category of margins are illustrated in Fig. 1.2. The regions of interest (ROIs) are from the Digital Database for Screening Mammography (DDSM) [61], which will be described in more detail in Section 2.1.3. Different shapes of masses are also possible: round, oval, lobular, or irregular [39].

When a suspicious area is detected on the mammogram by the radiologist, its nature is further investigated by means of a histological examination of a biopsy sample of the lesion in order to determine whether the area corresponds to a malignant tumor or a benign lesion. Tumors as small as 2 mm in diameter can be detected on the mammogram but are also the most difficult to identify, especially in the presence of dense fibroglandular tissue [69].



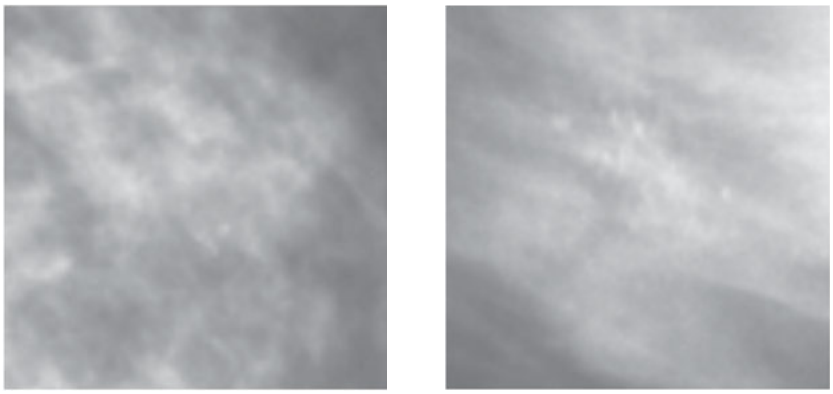
(a) (b) (c) (d) (e)

Figure 1.2: Examples of regions of interest (ROIs) including a mass with (a) circumscribed, (b) microlobulated, (c) obscured, (d) indistinct, or (e) spiculated margins. (a,c) Benign lesions. (b,d,e) Malignant Tumors. The ROIs are extracted from the Digital Database for Screening Mammography (DDSM) [61]; see Section 2.1.3 for more details.

Calcifications

Calcifications are small and bright spots on the mammogram due to the deposition of calcium within the breast parenchyma. They are characterized in terms of size, morphology, number, and distribution [39]. The presence of a cluster of calcifications is associated with an increased risk of malignancy, but benign clusters may also occur. The detection of calcifications is limited by the signal-to-noise ratio. The size varies from 100 μm to 1 mm. However, the typical size of calcifications that is detectable by conventional mammography is around 200 μm. The detection of smaller calcifications requires geometric magnification techniques, which are performed as part of a diagnostic mammogram workup for patients suspected to have breast disease but not

during screening mammography [5]. Examples of calcifications from the DDSM are shown in Fig. 1.3.



(a) (b)

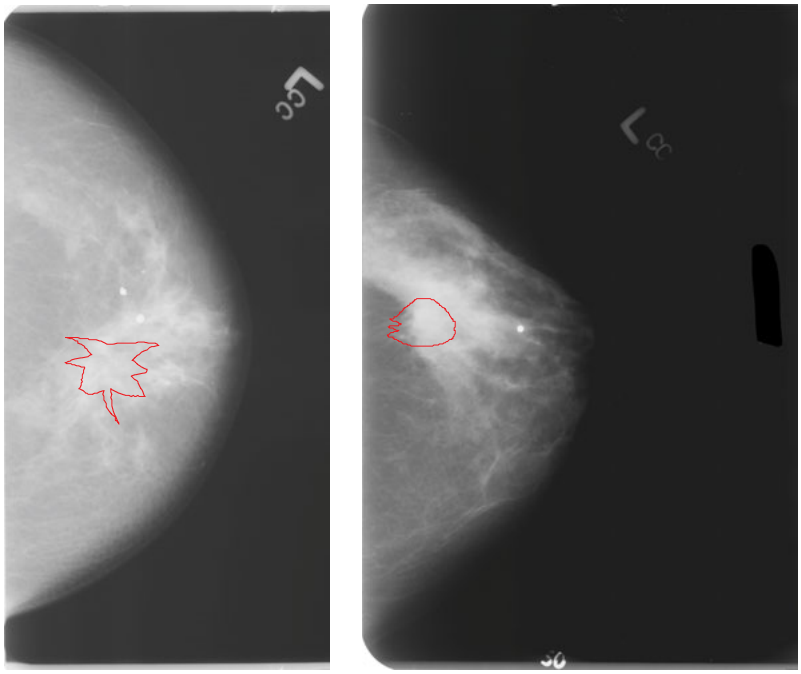
Figure 1.3: Examples of (a) benign and (b) malignant clusters of calcifications from the DDSM [61].

Architectural Distortion

When the normal architecture of the breast parenchyma is distorted in the absence of visible masses, it is defined as architectural distortion. It is usually characterized by the presence of spiculations radiating from a point, but it can have a more subtle appearance manifested by focal retraction or distortion at the edge of the parenchyma [39]. Two examples of regions of architectural distortion from the DDSM [61] are illustrated in Figs. 1.4a and b.

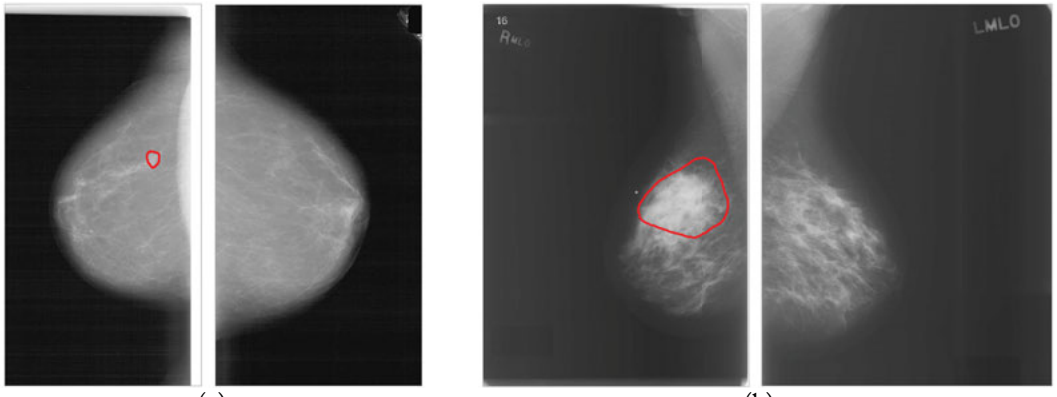
Bilateral Asymmetry

Radiologists perform comparative studies of the left and right mammograms of a given patient to prevent missing signs of breast disease. When a greater area of tissue with fibroglandular density is detected in a mammogram relative to the corresponding region in the contralateral breast, it is reported as an asymmetric finding, either local or global [39]. The presence of a greater area of tissue with fibroglandular density when judged relative to the contralateral breast defines a class of mammographic lesions denoted as bilateral asymmetry [39]. The condition of asymmetry is reported as global if the observed differences in areas of fibroglandular tissue are extensive, or focal if the difference in fibroglandular tissue density is confined to a small region but lacks the conspicuity of a mass [39]. Examples of two pairs of focal and global bilateral asymmetry from the DDSM [61] are illustrated in Fig. 1.5a and b, respectively. The identification of all asymmetric findings in a given pair of mammograms is important, because they may be the only clue to breast disease that is detectable on standard mammographic projections, especially when



(a) (b)

Figure 1.4: Examples of mammograms with regions of architectural distortion from the DDSM [61]. The red contours indicate the regions outlined by the radiologist.



(a) (b)

Figure 1.5: Examples of two pairs of (a) focal and (b) global bilateral asymmetry from the DDSM [61].

masses, microcalcifications, and/or architectural distortion are not visible [134]. Bilateral asymmetry has proved to be an indicator of increased risk of developing breast cancer [59, 133, 170], stressing the importance of special surveillance and follow-up observations of the patients to establish the nature of the asymmetry present. Asymmetric findings on mammograms may indicate a developing or underlying mass. They can be subtle in presentation and hence overlooked or misinterpreted by radiologists. The difficulty with the detection of asymmetry arises because the bilateral anomalies caused by a developing or underlying pathological process need to be differentiated from the physiological differences between the two breasts and distortions due to projection artefacts. These confounding factors and subtlety in presentation can cause overlooking or misinterpretation, even by experienced radiologists [88]. Clinical studies have reported that asymmetry accounts for 3–9% of breast cancer cases incorrectly reported by radiologists as showing no evidence of a tumor [15]. Evidence also suggests that asymmetric distribution of fibroglandular density is a common source of false-positive (FP) diagnosis [154].

1.1.5 BI-RADS MAMMOGRAPHIC DENSITY CATEGORIES

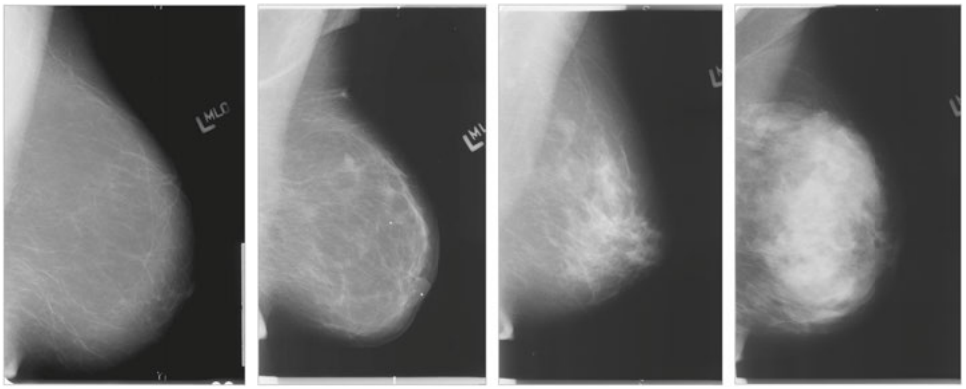
Mammograms exhibit differences in terms of the type of breast tissue composition. Due to the variable proportion of fatty and fibroglandular tissues in the breast composition mammographic images are difficult to interpret by radiologists. Fibroglandular tissue is composed by the stroma that forms the connective components of the breast, by the glandular component that represents its functional part, and by the breast ducts that correspond to the mammary parenchyma. Fatty and fibroglandular tissues have different X-ray attenuation coefficients. In particular, regions of fat appear darker than the fibroglandular components on the mammogram. Fibroglandular regions appear as brighter regions and are referred to as "mammographic density." It is well known that a strong correlation exists between the presence of relatively large regions of density in mammograms and the risk of developing breast cancer in the near term [161]. Detecting breast cancer in mammograms of dense breasts is more difficult due to the superimposition of projected tissues that may obscure small tumors. As a consequence, the accuracy of mammography is inversely correlated with density. It is then important to determine the category of density of a subject in order to have an indication of the detection capability of the examination [82]. The BI-RADS [39] lexicon defines four density classes (B-I to B-IV) and establishes the corresponding effects on the diagnostic accuracy. The accuracy of mammography to detect suspicious lesions decreases for types III and IV as follows.

B-I: The breast is almost entirely fat and the accuracy of mammography is very high.

B-II: The breast has scattered fibroglandular densities and the accuracy of mammography is high.

B-III: The breast is heterogeneously dense and the accuracy of mammography is limited.

B-IV: The breast is extremely dense and the accuracy of mammography is limited.



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

Figure 1.6: Examples of mammograms from the DDSM [61] for the BI-RADS mammographic density categories (a) B-I, (b) B-II, (c) B-III, and (d) B-IV.

1.1.6 TABÁR MASKING

The perception of subtle radiographic abnormalities in breast cancer screening can be improved by the use of a systematic approach to the analysis of mammograms, aimed at reducing false-positive rates (FPRs) and maintaining high levels of sensitivity. A complete mammographic study requires side-by-side viewing of corresponding areas of both breasts, whose practical realization should be strengthened by the technique of masking, as described by Tabár [145], ensuring that all regions of the breasts are viewed and compared in detail with the contralateral regions.

An exhaustive Tabár masking process would require at least four different types of analysis for each pair of views, performed with stepwise movements: horizontal and oblique masking of the MLO views, both in the cranial and caudal directions; horizontal and vertical masking of the CC views, the former in the medial and lateral directions and the latter in the proximal and distal directions. Particular attention should be given to the so-called "forbidden areas," specific areas where the majority of breast cancers are found in the early phase [145]:

(a) *medial-half area*: the medial half of the breast on CC projections;

(b) *retroglandular area*: the retroglandular space on CC projections;

(c) *milky area*: the region parallel with the edge of the pectoral muscle on MLO projections; and

(d) *retroareolar area*: the retroareolar region on MLO projections.

The "forbidden areas" on mammograms are illustrated in Figs. 1.7 and 1.8 together with possible masking procedures. During the masking procedures, different regions of the mammograms are compared by a radiologist with step-by-step movements. At each step, the areas under investigation can be matched singularly, changing stepwise the analyzed regions (*Stepwise Tabár masking*), or gradually, increasing the size of the paired observation windows (*Incremental Tabár masking*). The first approach enhances the perception of focal anomalies, whereas the latter allows a better understanding of global changes in the breast parenchyma.

1.1.7 DRAWBACKS AND LIMITATIONS OF MAMMOGRAPHY

Although mammography is the most widely used screening modality with solid evidence of benefit for women, some authors have been stressing its limitations, including false-negative (FN) and FP outcomes, overdiagnosis, and overtreatment of patients [2, 97]. Mammographic interpretation is a difficult task: signs of breast cancer can be very subtle and are often obscured by normal fibroglandular breast tissue with which these signs have many features in common, making their visual detection and analysis difficult. The accuracy of interpretation of screening mammograms, in particular, is affected by several factors, such as image quality and the radiologists' level of expertise. The rate of malignant cases missed by radiologists in the past few years has been reported to be 10–30% [88]. Additional reasons include the low prevalence of the disease in a screening population and the large number of mammograms that radiologists need to assess every day. To overcome such limitations and to improve radiologists' performance in interpreting mammograms, double reading, which consists of having two radiologists interpreting each case independently, has been introduced [54]. The alternative to double reading in the current screening practice consists of the use of computerized systems as second readers.

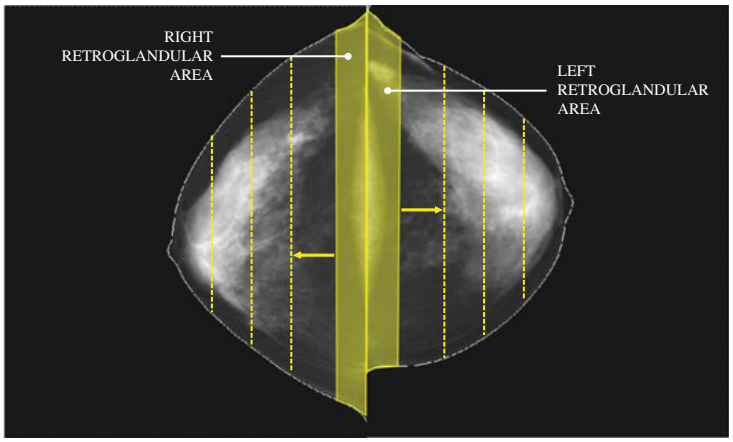
1.2 COMPUTER-AIDED DETECTION AND DIAGNOSIS WITH MAMMOGRAPHY

1.2.1 THE ROLE OF CAD AS A SECOND READER

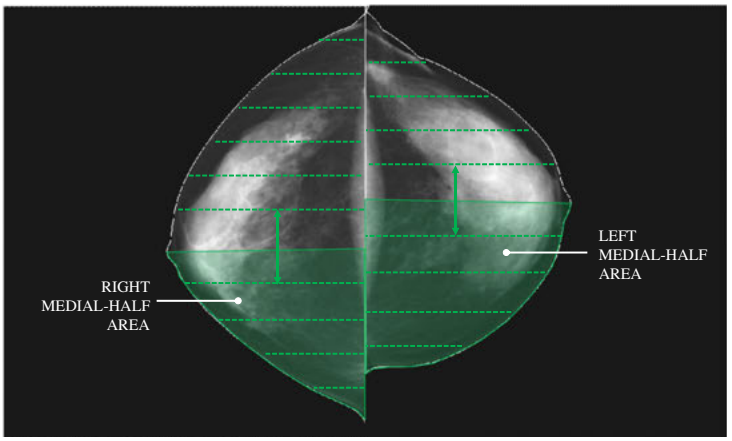
Computer-aided detection and diagnosis (CAD) techniques and systems involve the use of computer algorithms to detect patterns in images associated with signs of disease. In mammography, they can support radiologists in the role of a second reader, prompting the radiologists to review areas in a mammogram deemed to be suspicious (computer-aided detection, or CADe) and distinguishing between a lesion that is decidedly negative on a mammogram as opposed to one that needs regular monitoring or requires a biopsy (computer-aided diagnosis, or CADx).

For an understanding of the limits and potential of CAD of breast cancer, it is of interest to report what Alan Turing, the father of theoretical computer science and artificial intelligence, said about computing machinery and intelligence:

> *"I would say that fair play must be given to the machine. Instead of it giving no answer we could arrange that it gives occasional wrong answers. But the human mathematician*

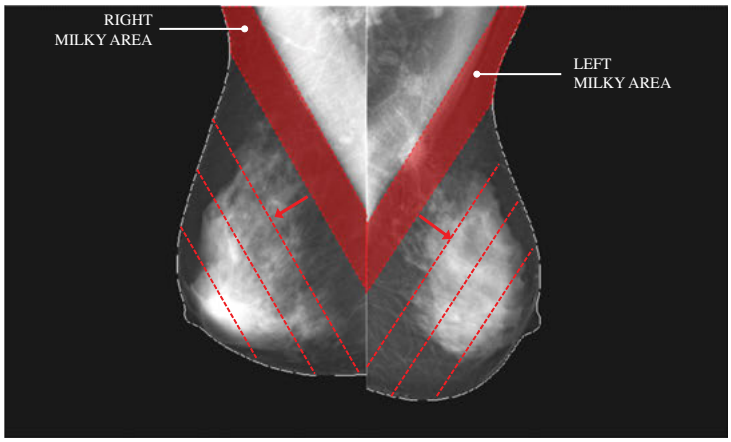


(a)

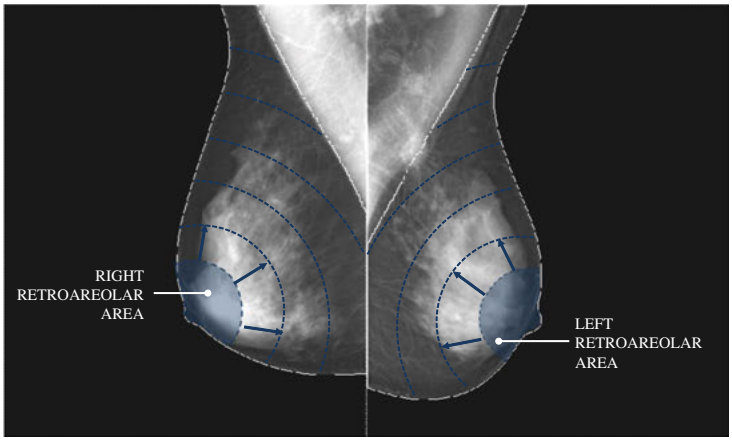


(b)

Figure 1.7: Masking procedures and "forbidden areas" (shown with labels) on mammograms in CC projections.



(a)



(b)

Figure 1.8: Masking procedures and "forbidden areas" (shown with labels) on mammograms in MLO projections.

would likewise make blunders when trying out new techniques. In other words then, if a machine is expected to be infallible, it cannot also be intelligent. There are several mathematical theorems which say almost exactly that. But these theorems say nothing about how much intelligence may be displayed if a machine makes no pretence at infallibility."

Alan Turing 1912–1954

Beside the potentials of artificial intelligence, Turing stressed the importance of its limits. Such limits, today, are at the basis of a correct fruition of computerized systems. Accurate quantification of the performance of CAD systems in solving real problems is crucial for efficient physician computer interaction. However, given the limits and benefits of both computer and human vision, the key for improved levels of sensitivity and specificity of diagnostic tests lies in the integration of automated approaches for quantitative analysis with human intuition. The final performance of a CAD system in mammographic reading should correspond to the performance achieved by the radiologist when interpreting mammograms by using the computer output as a second opinion, not by the performance of the CAD system itself. In this way, even if the accuracy levels of CAD systems do not surpass the levels achievable by radiologists, it is their interaction with the radiologist that determines the final benefits of CAD.

1.2.2 CLINICAL UTILITY OF CAD SYSTEMS

There are commercially available CAD systems for mammography whose benefits in a screening or diagnostic environment have been evaluated. Two of the most widely available are the ImageChecker system (Hologic, Inc., Bedford, MA) and the SecondLook system (iCAD, Inc., Nashua, NH).

The first large reported study on the effects of CAD systems in a screening environment was performed by Freer and Ulissey [52]; over a 12-month period the radiologists' performance using CAD was analyzed with respect to the radiologists' performance without CAD, showing an increase of 19.5% in the detected cancers. The study also reported a clinically significant ability of CAD systems in detecting tumors at an early stage (0 and I), which is critically important for saving women's lives. A consequent drawback was an increase of 19% in the number of recalled patients. However, the overall positive-predictive value for biopsy remained unchanged. Subsequent studies [11, 76, 101] confirmed the benefits in breast cancer detection due to the use of CAD systems for mammography, even if the reported increases were lower and ranged from 4.7–16.1%.

The next level of evaluation of CAD systems consists of determining the performance of radiologists using CAD in order to evaluate the effectiveness of CAD as an alternative to double reading. The interpretation of each case by two radiologists separately has been shown to increase the rate of cancer detection by 4–14%, but there are situations where double reading is not practiced [54]. The study by Ciatto et al. [34] estimated an increase of sensitivity of detection of interval cancers with prior negative screening mammograms of around 0.9% by introducing CAD-assisted reading with respect to double reading. A subsequent large study by

Gilbert et al. [54] found evidence on the equivalence of single reading with CAD and double reading. The economic and social benefits of the two alternative screening approaches have been analyzed in a recent study by Sato et al. [132], which found the use of CAD-assisted reading more cost-effective than double reading. However, the benefits due to the use of CAD systems in mammography are still a matter of debate and need further investigation. The same authors as cited above [11, 34, 52, 76, 101], in fact, have also reported the additional recall in 8–35% of cases produced by CAD and some results are still controversial. The study by Fenton et al. [46], for example, concluded that the increased rate of biopsy caused by CAD technology is not correlated with improved detection of breast cancer. The results are seen as the major limitation of CAD systems for mammography because of the increased costs of unnecessary followup examinations, resulting in additional anxiety for patients. Moreover, although several CAD techniques are effective in detecting masses and calcifications, the results are still less favorable for bilateral asymmetry, architectural distortion, and masses with obscured or ill-defined margins. Research is still in progress to overcome such limitations and also to move from CADe to CADx. New solutions can be found via image processing and computer vision techniques to be integrated into routine clinical practice and to improve the existing systems.

1.2.3 STATISTICAL EVALUATION OF DIAGNOSTIC PERFORMANCE

The diagnostic performance of any computerized system for mammography is assessed by comparing the result of the prediction test with a gold standard test. In fact, the gold standard test defines unequivocally the presence or absence of disease, as does the radiologist by reporting the presence of abnormalities or the biopsy result by indicating the benign or malignant nature of a lesion. The result of a prediction test for a two-class problem, as a diagnostic test, is negative, if it does not indicate the presence of the disease, or it is positive, when it indicates instead that the disease is present. For a given instance, the CAD system can give four possible outcomes with reference to the gold standard test.

TP (True Positive): the CAD system makes a positive prediction and the instance is actually positive.

TN (True Negative): the CAD system makes a negative prediction and the instance is actually negative.

FP (False Positive): the CAD system makes a positive prediction but the instance is actually negative (type I error).

FN (False Negative): the CAD system makes a negative prediction but the instance is actually positive (type II error).

The relative frequencies of the correct results obtained on the set of all instances define the sensitivity, or TPR, and the specificity, or true-negative rate (TNR), of the system as:

$$TPR = \frac{\#TP}{(\#TP + \#FN)} \tag{1.1}$$

and

$$TNR = \frac{\#TN}{(\#TN + \#FP)}, \tag{1.2}$$

which give an indication on how reliable the system is in making positive and negative identifications, respectively. In addition, the relative frequencies of the obtained incorrect results are quantified by means of the false-positive rate (FPR) and false-negative rate (FNR) of the system as:

$$FPR = \frac{\#FP}{(\#FP + \#TN)} \tag{1.3}$$

and

$$FNR = \frac{\#FN}{(\#FN + \#TP)}. \tag{1.4}$$

A global index of reliability is given by the accuracy of the system, which is defined as

$$Accuracy = \frac{\#TN + \#TP}{(\#TN + \#FP + \#TP + \#FN)}. \tag{1.5}$$

However, the values of TPR, TNR, FPR, FNR, and accuracy provide a static representation of the diagnostic test that does not consider the strength with which each instance belongs to one of the two classes. In fact, for every test, the calculated values of the described performance indices vary based on the particular cutoff value chosen to distinguish normal and abnormal results. This aspect is shown in Fig. 1.9, where increasing the cutoff level would make the test more specific but less sensitive: decreasing the number of FP instances and increasing the number of FN instances. Similarly, lowering the cutoff value would increase sensitivity while decreasing specificity: lowering the number of FP instances and decreasing the FN instances.

The decision about what cutoff to use in calling a test abnormal corresponds to a decision about whether it is better to tolerate FP instances (people without the disease inappropriately classified as diseased) or FN instances (missed cases with the disease). The choice of the cutoff depends on the disease in question and on the purpose of test. A good screening system requires a high TPR to detect all the possible lesions, giving the possibility of treating them early, and a high TNR to avoid useless and expensive further investigation of cases without the disease. This may not be achievable in reality. For the above-mentioned motivations, the best way to characterize a test is by extending the analysis of the diagnostic performance of a test by varying the cutoff, or threshold. The typical way to show this relationship is to plot the pairs $(TPR, 1 - TNR)$ for the entire range of possible threshold values. The resulting curve, known as the receiver operating characteristics (ROC) curve, which was originally described by researchers investigating

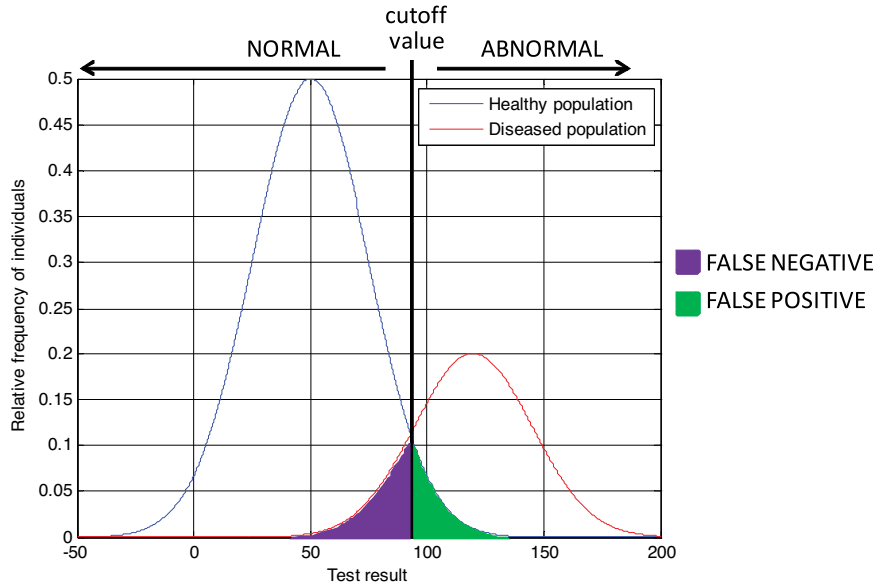


Figure 1.9: Example of a frequency distribution of a diagnostic test result in healthy and diseased individuals.

methods of electromagnetic-signal detection [113] during World War II and later applied to the field of psychology, facilitates improved analysis of the classification performance of a diagnostic method. A perfect diagnostic system has a ROC curve defined by the pairs $(0, 0)$, $(0, 1)$, and $(1, 1)$, whereas a random guess system corresponds to the diagonal line that goes from $(0, 0)$ to $(1, 1)$ [121]. The area under the ROC curve (AUC), also named A_z when estimated with a binormal model [140], ranges from zero to one and provides a measure of the system as ability to discriminate between actual positive cases and actual negative ones. AUC $= 0.5$ corresponds to random guess and AUC $= 1$ indicates an ideal diagnostic test with perfect separation between the positive and the negative classes. An example ROC curve with AUC $= 0.79$ is illustrated in Fig. 1.10.

When discrete and countable abnormalities (such as masses) need to be detected, it is important to establish the values of TPR obtained by the diagnostic system against the number of FPs per image (FPpI). This extends the ROC curve to a form known as the free-response receiver operating characteristic curve (FROC) [121].

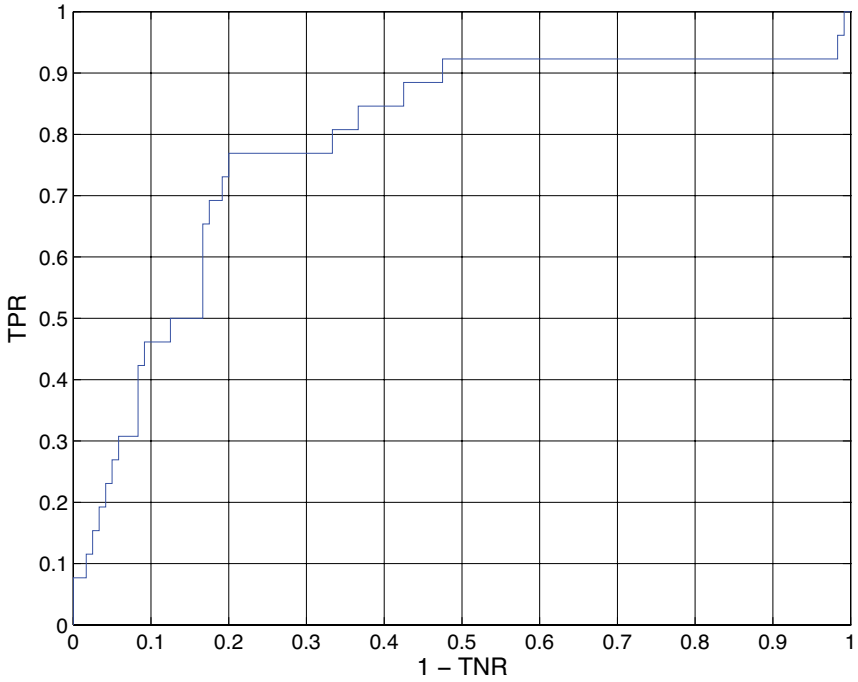


Figure 1.10: Receiver operating characteristic curve for a hypothetical diagnostic test with AUC = 0.79.

1.3 SCOPE AND ORGANIZATION OF THE BOOK

1.3.1 AIMS OF THE WORK

In the work presented in this book, four main tasks within the context of mammography and breast cancer have been addressed by means of image processing and pattern recognition techniques.

1. Design of novel landmarking algorithms for the extraction of the breast-skin line and the detection of the nipple. Reference anatomical structures, or landmarks, on mammograms, i.e., the nipple, breast-skin line, and pectoral muscle, are used in this work for effective matching between corresponding regions of the left and right breasts of a patient with the aim of performing automatic Tabár masking procedures and quantitative bilateral comparisons between pairs of mammograms.

2. Development and validation of a CAD system for the identification of bilateral asymmetry in mammograms as an early sign of breast cancer. The procedure is based on the analysis of the structural similarity or dissimilarity between paired mammographic regions. The

methods should improve the diagnostic sensitivity of CAD systems for mammography by providing clues about the presence of breast cancer which are not detected by other existing approaches.

3. Design of features for classification of mammographic regions as benign lesions or malignant tumors without relying on accurate extraction of the contours of the lesions. This is important, especially for masses or regions of architectural distortion with obscured or ill-defined margins, for which the other available approaches based on the segmentation of the lesions are prone to fail.

4. Development and validation of a novel comprehensive and multistage CADe/CADx system for automatic detection and diagnosis of malignant tumors. The task is addressed in a realistic scenario of a three-class environment, i.e., in the presence of normal parenchymal tissue, benign lesions, and malignant tumors, including masses and regions of architectural distortion. Integrated systems for mammography are expected to reduce the recall rate in screening mammography.

1.3.2 OVERVIEW

Chapter 2 documents the characteristics of the three databases of mammograms used in this work, including SFMs and FFDMs, as well as the validation strategy implemented.

Chapter 3 presents a technique for the extraction of the directional components of the breast used as a fundamental step in the analysis of mammograms.

Chapter 4 presents algorithms for automatic detection of anatomical reference structures, or landmarks, on the mammogram, including an algorithm for detection of the pectoral muscle and new methods for extraction of the breast-skin line and for detection of the nipple.

Chapter 5 contains novel methods for quantification of structural similarity or dissimilarity between the right and left mammograms of a patient as part of a CAD system for detection of bilateral asymmetry.

Chapter 6 provides details on the design process of contour-independent features for classification of masses as benign lesions or malignant tumors.

Chapter 7 focuses on the various original steps of analysis developed to design an integrated CADe/CADx system for detection and diagnosis of malignant tumors, including masses and regions of architectural distortion.