Pattern Recognition 42 (2009) 1138 – 1148

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Toward breast cancer diagnosis based on automated segmentation of masses in mammograms

Alfonso Rojas Domínguez[1], Asoke K. Nandi*

Department of Electrical Engineering and Electronics, The University of Liverpool, Brownlow Hill, Liverpool L69 3GJ, UK

ARTICLE INFO

Article history:
Received 29 November 2007
Received in revised form 7 June 2008
Accepted 4 August 2008

Keywords:
Breast cancer
Breast masses
Mammography
Image analysis

ABSTRACT

This work explores the use of characterization features extracted based on breast-mass contours obtained by automated segmentation methods, for the classification of masses in mammograms according to their diagnosis (benign or malignant). Two sets of mass contours were obtained via two segmentation methods (a dynamic-programming-based method and a constrained region-growing method), and simplified versions of these contours (modeling the contours as ellipses) were employed to extract a set of six features designed for characterization of mass margins (contrast between foreground region and background region, coefficient of variation of edge strength, two measures of the fuzziness of mass margins, a measure of spiculation based on relative gradient orientation, and a measure of spiculation based on edge-signature information). Three popular classifiers (Bayesian classifier, Fisher's linear discriminant, and a support vector machine) were then used to predict the diagnosis of a set of 349 masses based on each of said features and some combinations of these. The systems (each system consists of a segmentation method, a featureset, and a classifier) were compared with each other in terms of their performance on the diagnosis of the set of breast masses. It was found that, although there was a percent difference of about 14% in the average segmentation quality between methods, this was translated into an average percent difference of only 4% in the classification performance. It was also observed that the spiculation feature based on edge-signature information was distinctly better than the rest of the features, although it is not very robust to changes in the quality of the segmentation. All systems were more efficient in predicting the diagnosis of benign masses than that of the malignant masses, resulting in low sensitivity and high specificity values (e.g. 0.6 and 0.8, respectively) since the positive class in the classification experiments is the set of malignant masses. It was concluded that features extracted from automated contours can contribute to the diagnosis of breast masses in screening programs by correctly identifying a majority of benign masses.

1. Introduction

Breast cancer is the most common form of cancer in the female population, affecting one of approximately 11 women at some stage of their life in the Western world [1,2]. As with any form of cancer, early detection and diagnosis of breast cancer is one of the most important factors affecting the possibility of recovery from the disease. Early diagnosis of breast cancer can be achieved through mammography screening programs assisted by computers [3]. Over the past one and a half decades, several researchers have studied and proposed methods for computer-aided diagnosis (CADx) of abnormalities related to breast cancer in mammograms. The objective of CADx systems is to help radiologists in making a recommendation for patient management. CADx systems are used after a positive detection of a breast abnormality has occurred. If the abnormality is suspected to be malignant, a biopsy must be performed to confirm or reject this suspicion. Automated methods for diagnosis of breast masses are still under development. The large variability in the appearance of breast masses, added to the significant overlap in the appearance of malignant and benign masses, and the fact that abnormalities (masses and others) are often occluded or hidden in dense breast tissue, make both their detection and diagnosis very difficult.

The ultimate objective of automated methods for classification of masses is to provide a tentative diagnosis (the final decision is produced by a human expert) of individual masses, based on their physical attributes. These methods are incarnations of a generic model of supervised pattern classification systems. According to this model, a classifier is presented with features obtained from a selection of

* Corresponding author. Tel.: +44 1517944525; fax: +44 1517944540.
E-mail addresses: Alfonso.Rojas-Dominguez@liv.ac.uk (A. Rojas Domínguez), A.Nandi@liv.ac.uk (A.K. Nandi).
[1] Supported in part by the National Council of Science and Technology (CONACYT) of Mexico.

0031-3203/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2008.08.006

the objects that are to be classified, in a process known as training. The trained classifier can later label objects which were not used in its training, an ability known as generalization. The performance of classification methods depends on the type and quality of the features employed to train the classifier. Segmentation of the masses is almost always necessary for feature extraction, and may have an impact on the quality of the features obtained. This is because some of the most important information for the classification of breast masses is found in the transition region between a mass and the background tissue surrounding it. If the segmentation of a mass is of poor quality, in the sense that the boundary between mass and background is not correctly identified, the classification features may not be able to capture the important information contained in the transition region.

The general methodology for diagnosis of breast masses consists of the following main steps:

(1) Input: region of interest (ROI, portion of a mammogram containing a breast mass).
(2) Segmentation of the breast mass.
(3) Feature extraction.
(4) Classification of the breast mass.
(5) Output: predicted diagnosis of the breast mass.

Each of the steps from (2) to (4) corresponds to what can be identified as one main component of a CADx system. In other words, a diagnosis system is composed of a segmentation method, a set of features for classification, and a classifier. For each of these main components researchers have developed a number of options from which it is nowadays possible to select particular components and use these in a modular fashion to build a complete diagnosis system. However, even when breast cancer classification can be semi-automated, the involvement of experts in drawing accurate boundaries around masses is extremely time consuming and is not helpful in the endeavor to fully automate breast cancer diagnosis. In this context, the only way to make a significant advance in computer-aided breast cancer diagnosis may be to automate the mass-boundary drawing as well as the classification. Whilst there is a lot of research work published in automating classification, very little is available in the way of generating complex mass boundaries automatically. Attempts to generate complicated mass boundaries automatically have not been successful.

The approach adopted in this paper to the problem of automated diagnosis of masses is not via measures that may strongly depend on the accuracy of the mass contours (since these require the presence of an operator to produce such contours), but through the design of robust features that directly measure appropriate characteristics of the masses. This is not to say that mass contours are to be forgotten as a whole, but rather, that contours should only be used as pointers to the image regions from which effective features can be obtained. This constitutes a paradigm shift that shows encouraging results.

Spiculation is perhaps the single main indication of the malignancy of a breast mass. Another important indication is the roughness, raggedness, or fuzziness of the mass margins. The objective of this study is to explore the possibility of employing features designed to measure these characteristics, and extracted from contours obtained in an automated fashion (as opposed to features extracted from manually obtained contours), for automated analysis of breast masses in screening mammography. Another objective is to observe the effect that different segmentation methods (in terms of the segmentation quality) and different classifiers have on the systems and on the effectiveness of the classification features. Identification of the components that have the largest influence on the systems can indicate what are the areas that require the most improvement.

In this paper, six features designed to show some degree of robustness to errors in the location of the mass contours have been employed in classification experiments. The features were tested individually and in a number of combinations.

Two segmentation methods were considered in this study. One is based on dynamic programming (DP), and it is identified by the label ID^{2}PBT [4], the other method is known as constrained region growing [5,6], and it is identified by the label CRG. The ID^{2}PBT segmentation algorithm is completely automated. The CRG method used in this work requires the participation of a user to indicate the size of the constraint function. This is discussed in Section 3.1.

Three popular classifiers have been employed in the classification experiments. These classifiers are briefly described in Section 5. A detailed exposition of these and other classifiers can be found in Refs. [7–9]. A recent survey of detection and classification of masses in mammograms, including applications of classifiers can be found in Ref. [10].

2. Image database

In this work, a set of images selected from the mini-MIAS database [11] and from the digital database for screening mammography (DDSM) of the University of South Florida [12] are used for validation of the results. Our test set consists of 349 ROIs (43 from the mini-MIAS database and 306 from the DDSM) containing circumscribed masses—circumscribed masses are most probably benign—(150 masses) as well as spiculated and microlobulated masses—which are considered most probably malignant—(199 masses). The selection of cases included lesions with different degrees of subtlety and different sizes (see Fig. 1), from images with different breast-tissue densities (see Fig. 2), and lesions of benign and malignant diagnoses. A total of 207 masses have a benign pathology; the other 142 masses have a malignant pathology. The ROIs were all adjusted to be 256×256 pixels at 200 μm per pixel and 8 bpp (256 gray levels); bicubic interpolation was used when re-scaling was necessary to preserve uniformity in the images spatial resolution, but no masses larger than the area of 256×256 pixels were included in the test set. The choice of the ROIs was based on several factors: very large masses were not included, ROIs with two or more overlapping masses were not included, masses with calcifications were not included, etc. Since the effect of the segmentation algorithms is also of interest, criteria related to the application of these algorithms were also considered: that the boundaries of the

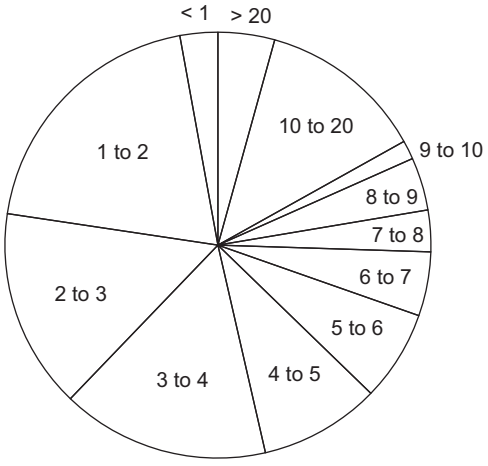

Fig. 1. Mass size distribution. Units are in thousands of pixels. The spatial resolution is 200 μm per pixel, therefore, a mass with 1000 pixels has an area approximately equivalent to that of a circle with a radius of 3.6 mm.

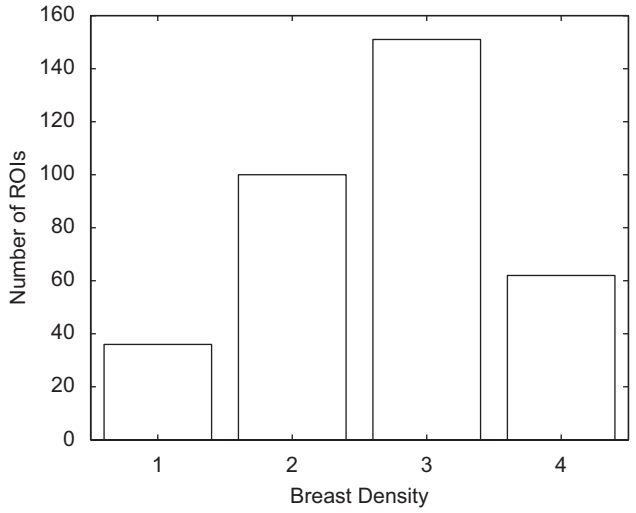

Fig. 2. Breast tissue-density distribution. 1. Entirely fatty, 2. scattered fibroglandular, 3. heterogeneously dense, and 4. extremely dense. Note: the MIAS database considers only the first three tissue-density categories, the DDSM database considers all four. The categories are those of the BI-RADS published by the American College of Radiology [46].

masses were not obstructed (for instance, by the pectoral muscle), that the boundaries appeared complete (sometimes only a portion of the mass can be seen in the mammograms), etc. A complete set of manual segmentations was produced based on the intensity and gradient magnitude of the mammograms and the annotations provided by the databases. These manually segmented regions were regarded as ground truth of the boundary of the masses.

3. Segmentation methods

3.1. Constrained region growing

Region growing consists in adding pixels to an initial set, known as seed, based on predefined criteria for growth [13]. Neighboring pixels that share similar properties with the seed region are appended to it, thus creating larger regions. The process is useful for segmentation of irregularly shaped objects.

While conventional region growing defines the lesion segmentation based solely on gray-level information (as long as connectivity is satisfied), researchers have proposed to introduce geometrical information to improve the segmentation results. In conventional region growing, the shape of the region being grown is not constrained in any way. Thus, the regions being grown can sometimes move into adjacent structures to the one of the seed location, producing wrong partitions. The geometrical information, based on prior knowledge of the general shapes of the objects to be segmented, takes the form of a constraint function used to prevent the inclusion of pixels that differ significantly from the reference shape, or that are located away from the compact region of the object of interest. Pixels with the same gray level but different locations are modified by an appropriate constraint function so that those pixels closer to the seed location will now differ from those pixels distant from the seed. The introduced dissimilarity between close and distant pixels effectively prevents the region from growing excessively. The resulting segmentation method is known as CRG.

The form and complexity of the constraint function will obviously have an effect on the effectiveness of the CRG method. In the literature of segmentation of breast masses, a popular constraint function is an isotropic Gaussian function centred on the location of the seed point [5,6]. The choice of this function derives from the prior knowledge (or assumption) that most mammographic lesions are approximately circular in shape. The Gaussian function is multiplied with the ROI before the application of region growing. The effect of the constraint function is controlled by the variance of the Gaussian function, σ_c^2, which is directly related to the size of the objects to be segmented.

In this paper, data computed beforehand was provided to the CRG algorithm, so that it could adjust the size of the constraint function to that required to segment each individual mass in the dataset. This was done in order to simplify the operation of the segmentation algorithm and to reduce its processing time. As a consequence, the CRG algorithm also received some advantage over the ID^{2}PBT algorithm, which is fully automated.

3.2. DP-based boundary tracing

The technique of DP was first employed for delineation of noisy contours by Montanari [14], Gerbrands [15], and others. Applications of DP to the segmentation of medical images soon followed and continue to appear in the literature [16–21]. At the heart of DP-based segmentation algorithms is the so-called cost function, which is used to find the path that most efficiently represents the contour that is being delineated. The cost function should incorporate all the information that characterizes the desired contour; it typically includes three components related to gradient (edge strength), intensity (gray level), and shape (and size) of the objects to be segmented [17,18,20].

A geometrical transformation of the ROI to polar-variable representation can be employed to rectify the image matrix for reduction of the number of DP iterations [15]. In this case, the DP-based method first constructs a local cost function that assigns a cost to each pixel in the polar-variable representation of a ROI. From the local cost image, a cumulative cost matrix is obtained where the cumulative cost of each path from the first column to the last column of the cost image is stored. The contour of the lesion is defined in the cumulative cost matrix by those pixels that linked together form the path with the lowest cumulative cost.

In the experiments reported in this paper, a particular instance of the DP-based segmentation algorithm, labeled ID^{2}PBT has been employed. The details of the implementation can be found elsewhere [4]. In summary, the components of the ID^{2}PBT algorithm are obtained as follows:

Edge strength component: pixels with strong edge content are assigned a low cost, and vice versa. In order to determine edge strength, first the local standard deviation of the intensity of pixels is obtained, producing an image where edges are highlighted. This is then subjected to a refinement procedure by computation of a feature known as the intrinsic coherence [22,23].

Gray-level component: pixels with gray-level values similar to a preferred gray level (which corresponds to the boundary of masses) are given a low cost. The preferred gray level used in the segmentation of a particular mass is calculated by the algorithm based on estimates of the average gray-level values of the mass and the background, and on an estimate of the size of the mass.

Shape component: the shape of each particular mass to be segmented is modeled by an ellipse. The parameters of the ellipse (axes and orientation) are computed by the algorithm based on an estimate of the boundary of the mass. The pixels on the coordinates given by such a shape estimate (the particular ellipse found) are given a low cost. The rest of the pixels are given a higher cost, in proportion to the deviation of their location from the shape estimate.

The three components above are combined into a total cost function through a linear combination with weights that are dynamically adjusted by the ID^{2}PBT algorithm based on the relative agreement of the components. From this total cost function, the cumulative cost matrix is computed, and DP is applied to trace the boundary of the

mass (the path with the lowest cost across the cumulative cost matrix). The operation of the ID^{2}PBT algorithm is fully automated.

4. Features

The features included in this study were all designed or modified to use as little information as possible from the contour of the masses, so that an approximate contour (such as one extracted via an automated segmentation algorithm) can be employed in their computation. In this way, the features are expected to show some degree of robustness to errors in the location of the boundary of the masses (the quality of the segmentation).

The features are:

- Contrast between foreground region and background region (Co).
- Coefficient of variation of edge strength (CV_{ES}).
- Measures of fuzziness of mass margins (Fz_1 and Fz_2).
- Measure of spiculation based on relative gradient orientation (Sp_{GO}).
- Measure of spiculation based on edge-signature information (Sp_{SI}).

4.1. Selection of the approximate mass-boundary region

In order to obtain the features described in this work, a set of simplified contours of the breast masses is first produced. These simplified contours are ellipses that approximate the shape of the contours of masses, and are obtained by least-squares curve fitting of the points of each mass contour [24]. The ellipses are used to guide the selection of a band of pixels. The selection includes the boundary of the mass, and may also be extended to include portions from the interior of the mass and from the background region that surrounds the mass. The size of the band, and the regions included in the selected band depend on the particular feature that is to be extracted. This is further discussed in the sections describing each of the features. In some cases, also depending on the particular feature desired, the band of pixels selected is converted into a rectangle via the rubber band straightening transform (RBST) [25]. This is done to simplify the application of some operations that work on rectangular arrays. The steps described above are illustrated schematically in Fig. 3. In this paper the width of the band across the guiding mass contour was kept fixed. However, a band with a variable width, proportional to the diameter of the mass, could be more appropriate and have an effect on the system performance. This may be explored in a future study.

4.2. Coefficient of variation of edge strength

Mudigonda et al. [26] proposed to use the coefficient of variation of edge strength (CV_{ES}) as a feature for diagnosis of masses. They defined edge strength based on the variance (σ_w^2) of the intensity values of pixels in a line perpendicular to the boundary of a mass

$$\sigma_w^2 = \frac{1}{M}\sum_{n=\lfloor -M/2\rfloor}^{\lfloor M/2\rfloor}[f_i(n)-\mu_w]^2, \tag{1}$$

$$\mu_w = \frac{1}{M}\sum_{n=\lfloor -M/2\rfloor}^{\lfloor M/2\rfloor} f_i(n), \tag{2}$$

where M is an odd number of pixels, $f_i(n)$ are the values of the pixels considered at the i th boundary point in the perpendicular direction. The maximum of the variance values is chosen as the edge strength value of the boundary point being processed. The coefficient of variation of the edge strength is then computed over all the boundary points. In this study, the feature was computed from the RBST of a band of 20 pixels to either side of the guiding contour.

4.3. Contrast

Let μ_f and μ_b be the mean intensity values of pixels in two separate regions labeled *foreground* and *background*, respectively. One definition of the contrast between foreground region and background region is

$$Co = \frac{\mu_f - \mu_b}{\mu_f + \mu_b}. \tag{3}$$

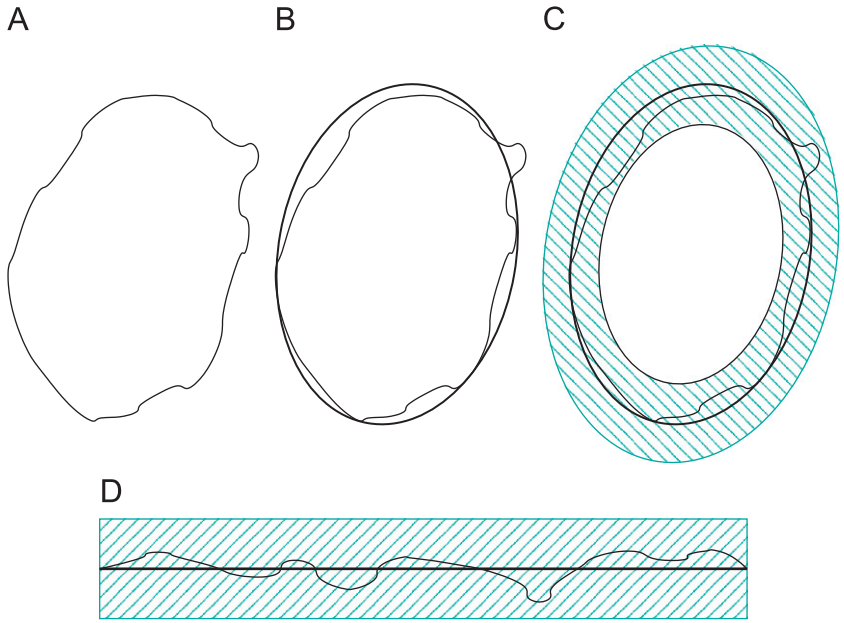

Fig. 3. (A) Contour of a mass obtained by automated segmentation method. (B) Ellipse fitted to the mass, to be used a simplified guiding contour. (C) Band or ribbon of pixels selected around the guiding contour. (D) Rubber-band straightening transform of the band of pixels selected in (C).

In the particular case of a contrast measure for classification of breast masses, the foreground region corresponds to a mass, and the background region to the tissue surrounding the mass. Only the pixels in these regions, that also overlap the boundary ribbon, are included in the computation of the contrast feature. This features was extracted from a band of 40 pixels around the guiding contour (20 pixels at either side of the contour).

4.4. Radial to tangential signature information

This feature is a measure of similarity between three different edge signatures of a breast mass. One of the signatures, the omnidirectional edge signature, contains all of the edge content in the mammographic image. The other two signatures, the radial edge signature and the tangential edge signature, are designed to retain only the edge content that appears oriented in a certain direction. In the radial edge signature, only the structures that possess some edge content oriented in the radial direction are represented. In the tangential signature, only the structures that possess edge content oriented in the tangential direction are represented. By comparing (in terms of a similarity measure) the edge signatures of a mass in a mammographic image it is possible to obtain an indication of the characteristics of the mass. This is the rationale behind the Sp_{SI} feature, which is described as follows.

Consider a log-Gabor function [27,28] (see also Ref. [29, Chapter 8]) with a transfer function $\mathscr{G}$. A bank of oriented filters, $F_{n,k}$ can be constructed using the following equations:

$$\mathscr{G}(w, w_n) = \exp(-(\log(w/w_n)^2)/2(\log(\sigma_w))), \tag{4}$$

$$G(\theta, \theta_k) = \exp(-(\theta - \theta_k)^2/2\sigma_\theta^2), \tag{5}$$

$$F_{n,k} = \mathscr{G}(w, w_n)G(\theta, \theta_k), \tag{6}$$

where w_n is the center frequency of the filters, σ_w is a term used to obtain constant-shape ratio filters, $G(\theta, \theta_k)$ is a Gaussian spreading function which acts as an angular envelope oriented in direction θ_k, σ_θ is the standard deviation of $G(\theta, \theta_k)$ in the angular direction (used to control the spread of the envelope). The subscripts n and k indicate a particular scale (n) and orientation (k) of a filter in the filter bank.

Let I be the 2-D DFT of a ROI containing a breast mass, with the mass centred on the ROI, and let $\mathscr{F}^{-1}[\cdot]$ represent the inverse 2-D DFT. Then the radial response $r_{n,k}$, the tangential response $t_{n,k}$, the radial edge signature $R(i,j)$, the tangential edge signature $T(i,j)$ and the omnidirectional edge signature $O(i,j)$ of the ROI are given by

$$E_{n,k} = \mathscr{F}^{-1}[IF_{n,k}], \tag{7}$$

$$r_{n,k} = |E_{n,k}|G(\theta, \theta_k), \tag{8}$$

$$t_{n,k} = |E_{n,k\perp}|G(\theta, \theta_k), \tag{9}$$

$$R(i,j) = \sum_n \sum_k r_{n,k}(i,j), \tag{10}$$

$$T(i,j) = \sum_n \sum_k t_{n,k}(i,j), \tag{11}$$

$$O(i,j) = \max_{n,k} |E_{n,k}(i,j)|, \tag{12}$$

where $E_{n,k}$ represents the complex response of one filter for a certain scale n and orientation k, and $k \perp$ is used to indicate the orientation orthogonal to orientation k. The spatial horizontal and vertical variables for the location of a pixel are represented by i and j, respectively.

Once the edge signatures have been computed, the spatial portions of the signatures that roughly correspond to the mass boundary are selected using the simplified contour of the masses (ellipses) to guide the RBST transform. The width of the selected band of pixels was chosen to be 20 pixels to either side of the guiding contour.

The 2-D mutual information [30,31], $M(X;Y)$, between images X and Y is defined in terms of the marginal entropies $H(X)$ and $H(Y)$, and the joint entropy $H(X,Y)$ as $M(X;Y) = H(X) + H(Y) - H(X,Y)$. Let $R_{\hat{B}}$, $T_{\hat{B}}$, and $O_{\hat{B}}$, represent the selected portions of the radial, tangential and omnidirectional signatures, respectively (we use the subscript $\hat{B}$ to indicate that the selected portions correspond to estimates of the boundary location given by the simplified contours). Then, the following rule can be used to discriminate spiculated masses: If $M(O_{\hat{B}}, R_{\hat{B}}) > M(O_{\hat{B}}, T_{\hat{B}})$, then the mass is most probably spiculated. In this way, the two measures of mutual information can be used as features in a classifier. Alternatively, instead of having two features it is possible to combine them into one dimensionless feature. Notice that, since the edge signatures of a mammographic mass are not independent, none of the mutual information measures can be zero. Thus, the ratio of the two mutual information quantities can be used as a combined feature. Let this feature be called spiculation measure based on signature information:

$$Sp_{SI} = \frac{M(O_{\hat{B}}, R_{\hat{B}})}{M(O_{\hat{B}}, T_{\hat{B}})}. \tag{13}$$

The larger the value of Sp_{SI} for a given mammographic mass, the most likely it is that the mass is a spiculated mass.

4.5. Spiculation measure based on relative gradient orientation

This feature is a spiculation measure based on the relative gradient orientation of pixels on spiculations, Sp_{GO}. Spiculations appear as linear structures with a positive image contrast. As a result of their linearity, the gradient directions at image pixels on or close to the spicules have approximately the same orientation relative to the spicules. Spicules develop in an approximately radial direction to the mass (normal to the mass contour), and the gradient at the pixels on a spicule is orthogonal to the direction of the spicule (i.e. tangential to the mass contour). This is schematically illustrated in Fig. 4A, where the contour of the mass is represented by a circle, the line containing the segment $\overline{PP'}$ represents a spicule, and the vector g represents the gradient direction at point P' on the spicule. As a result of the characteristics of the spicules, the angle θ between $\overline{PP'}$ and the gradient vector g is close to $\pi/2$.

The Sp_{GO} feature exploits the relationship between the image gradient direction and the spicules direction to discriminate between spiculated and non-spiculated masses (a similar approach has been applied in the past for detection of spiculated lesions, see Refs. [32,33]). However, instead of measuring the angle θ, we measure another angle, β, between the image gradient vector g, and the line joining the centroid of the mass region with the point P'. If the masses were perfectly circular, the two angles β and θ would be equal. In reality, since the masses are not circular, the angles are not the same. The angle β can be considered an approximation of the angle θ, with the advantage that the boundary point P is not needed to measure the angle β. This is illustrated in Fig. 4B.

The Sp_{GO} feature is defined as the average value of the sine function of the angle β, computed over a selection of pixels around the mass. In order to select the pixels for the computation of Sp_{GO} a feature known as the phase congruence (PC) [27,34] is employed. The PC is a dimensionless measure of edge content which is invariant to image illumination and contrast. The PC range is [0–1], and thus, it is possible to use a threshold on the PC image to pick up increasingly (or decreasingly) significant image features. PC is a very sensitive feature and can be affected by noise. For this reason we apply it on the gradient magnitude of the mammographic images instead of applying it directly to the mammograms. Used in this way, PC allows the selection of pixels which possess a high probability of belonging to either the spicules of a spiculated mass, or the contour of a circumscribed mass.

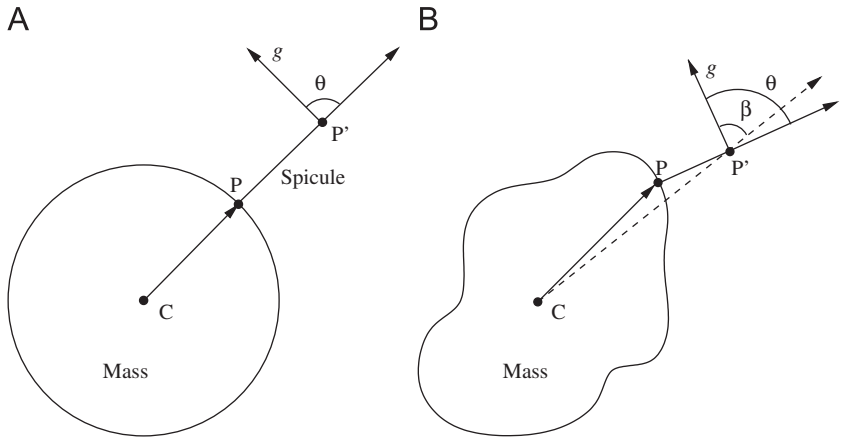

Fig. 4. (A) In the ideal case of a perfectly circular mass, the angle θ between the gradient g and the line $\overline{PP'}$ is the same as the angle between g and the line $\overline{CP'}$. (B) In a real case the mass is not circular. Here the angle θ between g and $\overline{PP'}$ is different than the angle β between g and $\overline{CP'}$.

A second restriction is imposed on the location of the pixels selected for the computation of Sp_{GO}. This restriction consists in limiting the selection to those pixels that belong to a region A around the mass. The region was defined as a ribbon of thirty pixels to the exterior of the guiding contour of each mass.

Let $\hat{g}_i$ represent the unitary gradient vector at a pixel location P'_i, and $\hat{c}_i$ represent the unitary vector with the direction of the line $\overline{CP'_i}$ joining the centroid of the mass, C, and the selected pixel P'_i. Then, the Sp_{GO} is mathematically expressed as

$$Sp_{GO} = \frac{1}{N}\sum_i \sin(\arccos(\hat{g}_i \cdot \hat{c}_i)),$$
$$i \in \{A \cap (PC(x,y) > \alpha)\}, \quad i = 1,2,\ldots,N, \tag{14}$$

where N is the total number of selected pixels, A represents the selection-allowed region (see Fig. 3C), $PC(x,y)$ is the PC value of the pixel at location (x,y), and $\alpha \in (0{-}1)$ is a threshold on the PC that becomes a parameter of the Sp_{GO} feature.

4.6. Measures of the fuzziness of mass margins

These features are measures of the local fuzziness of the mass margins. A feature similar to the Fz_k features described below was proposed by Varela et al. [35], but in that study the computation of the feature was strongly dependent on the boundary of the masses. The method to compute these is more easily implemented and described on the polar-variable representation of the ROIs. In this representation, $(x,y) \mapsto (r,\theta)$, with the centroid of the mass as the origin of the transformation, $r = 1,2,\ldots,128$ (maximum valid radius in ROIs of size 256×256), and $\theta = 1,2,\ldots,360$. Let (i,j) represent the indexes in the polar-variable representation of the ROIs.

First the derivative (first or second order) in the radial direction is computed. Next, the pair of points, separated a fixed distance apart along the radial direction, δ_i, for which the difference between derivative values is maximum are found. This is done for each column of the image matrix (i.e. along the angular variable). The location halfway between these points is the location of the strongest edge in the radial direction, which is used to characterize the mass margins. The location of the strongest-edge is restricted to a band of forty pixels around the guiding contour of each mass in order to reduce possible noise being included in the computation. When the mass margins are well defined and smooth around the mass (such as with circumscribed masses), the difference between the strongest-edge locations is small. If the mass-margins are irregular or jagged (a spiculated or micro-lobulated mass), the difference between the strongest-edge locations is large. Thus, if by computing the average of the radial difference between strongest-edge locations along the tangential direction, one obtains an indicator or measure of the fuzziness of the mass margins. Disregarding the treatment of index values for which the derivatives are undefined, the above procedure is mathematically expressed as

$$d_k(i,j) = |p_k(i-\delta_i,j) - p_k(i+\delta_i,j)|, \tag{15}$$
$$m_k(j) = arg\max_i d_k(i,j), \tag{16}$$
$$Fz_k = \frac{1}{J}\sum_j |m_k(j) - m_k(j+1)|, \tag{17}$$

where $p_k(i,j)$ represents the k-th order difference in the radial direction, $k = \{1,2\}$, and J represents the maximum of index j.

5. Classifiers

Three popular classifiers have been employed in the classification experiments. These classifiers are briefly described below. A detailed exposition of these and other classifiers can be found in Refs. [7–9]. A recent survey of detection and classification of masses in mammograms, including applications of classifiers can be found in Ref. [10].

5.1. Bayesian classifier

A Bayesian classifier minimizes the classification error probability by application of the *Bayes classification rule*. Let $\mathbf{x}$ represent a feature vector, and ω_i, with $i = 1,2$, represent the classes in a two-class problem. In this case the Bayes classification rule can be stated as [9]: If $p(\mathbf{x}|\omega_1)P(\omega_1) > p(\mathbf{x}|\omega_2)P(\omega_2)$, $\mathbf{x}$ is classified to ω_1, otherwise, $\mathbf{x}$ is classified to ω_2. The *a priori* class probabilities $P(\omega_i)$, and the class-conditional probability density functions $p(\mathbf{x}|\omega_i)$ can all be estimated from the available training feature vectors. A common choice for the class-conditional density functions is the Gaussian or normal density function, due to its computational tractability and the fact that, in general, it models adequately a large number of cases found in practice. However, not all classification problems are well (or easily) solved by minimization of the classification error. For example, in some cases it is not easy to estimate the probability distribution functions involved. It may be preferable to compute decision surfaces directly by means of alternative cost functions or criteria. The next classifier discussed is an example of this approach.

5.2. Fisher's linear discriminant (FLD)

Instead of minimizing the classification error or classification risk, the FLD maximizes a separability criterion [36]. The criterion adopts large values if the data is organized in compact clusters corresponding to each class, and the clusters of different classes are well separated from each other. The approach undertaken by the FLD is to transform the m-dimensional feature vector $\mathbf{x}$ (via a linear transformation) into another l-dimensional vector $\mathbf{y}$ so that the separability criterion is maximized. FLD does not make assumptions about the data distribution such as normally distributed classes or equal class covariances. In practice, the class means and covariances needed to compute the separability criterion are not known, but can be estimated from the training set. However, although the estimates of the covariance may be considered optimal in some sense, this does not mean that the resulting discriminant obtained by substituting these values will be optimal. Also, in many practical cases linear discriminants are not suitable. The FLD can be extended for use in nonlinear classification via a technique known as *the kernel trick*. This extension is known as the *kernel Fisher discriminant*. We shall consider another family of classifiers that also employs the kernel trick, known as support vector machines (SVMs).

5.3. Support vector machines

Consider a two-class problem with linearly separable data. It is easy to see that the optimal discriminant function takes the form of a hyperplane located somewhere between the clusters of data corresponding to each of the classes. The optimal position of this hyperplane will be dictated by the (relatively) small number of data samples from each class that are closer to samples from the other class. Because of this, these samples are called *support vectors*, and the classifiers that determine the optimum hyperplane are called SVMs [37]. When the problem is expanded to multi-class problems and non-separable data, the methodology to obtain a solution to the SVM becomes quite complex. For the purposes of this brief introduction, it suffices to say that there are a number of properties and constraints in the mathematical formulation of SVMs that permit the use of optimization techniques which ultimately return a solution. One of such properties is exploited to allow the application of SVMs to nonlinear classification problems via an implicit mapping of the original feature vector into a higher dimensional space where linear classification is effective (linear classification in the higher dimensional space becomes equivalent to nonlinear classification in the original space). The functions used in the mapping are called *Mercer kernels*, and the operation is known as the *kernel trick* [38]. Typical kernel functions used in pattern recognition applications include: polynomials, radial basis functions, and hyperbolic functions. Once an appropriate kernel has been adopted, the solution to the SVM problem can proceed as in the linear case.

6. Experiments

Three different featuresets were employed in this study. One of these sets is used to train the classifiers, and is referred to as the training set. The training set was produced using guiding contours based on the ground-truth contours. Two more sets were used to evaluate the performance of automated diagnosis methods (i.e. combinations of segmentation methods, features, and classifiers). One test set was obtained using the contours produced by the DP-based segmentation method (ID^{2}PBT). The other test set was obtained using the contours produced by the CRG segmentation method. The whole set of 349 masses was employed in the tests under a leave-one-out classification framework. In other words, in order to obtain the results presented, each classifier was trained with 348 masses and using the feature values from the training set. Then the classifier was tested on the mass that was "left out" and using the feature values from each of the test sets. This procedure was repeated, leaving out a different mass, until the classifiers were tested on all of the masses in the dataset. The average area overlap measure (AOM, a measure of agreement between two regions [5,6,39]) between ground truth and the set of contours of each segmentation method was 0.72 and 0.83 for ID^{2}PBT and CRG, respectively (percent difference of 14.19).

Each diagnosis system consists of one segmentation method (ID^{2}PBT or CRG), one feature set (including one or more features) with features based on either the ID^{2}PBT or on the CRG segmentation methods, and one of three classifiers (Bayesian classifier, FLD, or SVM). None of the operational parameters of the classifiers were optimized for the particular dataset. The features were tested individually and in combination. The main results of the experiments are reported in Table 1 and Figs. 5–9, which report the classification success as a percentage of the total number of masses and also give the percent differences between systems that employ ID^{2}PBT and those that employ CRG; and Tables 2–4 which contain the sensitivity (SE), specificity (SP), positive predictive value (PPV), and negative predictive value (NPV) of each system, rounded to the nearest half tenth. Recall that SE is computed as the number of masses correctly classified as malignant divided by the total number of malignant masses in the database; similarly, SP is defined as the number of masses correctly classified as benign divided by the total number of benign masses in the database. For the results in Tables 2–4 it

Table 1
Performance-classification results (expressed as % of success) of different diagnosis systems

Featureset	ID^{2}PBT			CRG			Percent diff. (%)		
	FLD	Bay.	SVM	FLD	Bay.	SVM	FLD	Bay.	SVM
6	63.64	70.09	69.50	71.88	73.33	73.04	12.16	4.52	4.97
7	60.41	59.82	58.06	58.26	60.58	57.97	3.62	1.26	0.16
8	61.03	62.75	63.04	62.46	61.89	62.46	2.32	1.38	0.92
9	64.81	70.38	70.09	71.88	73.04	74.78	10.34	3.71	6.47
10	63.64	61.88	64.81	63.48	62.90	66.38	0.25	1.63	2.39
11	69.50	68.04	69.21	72.46	72.46	72.17	4.17	6.29	4.19
12	66.28	70.97	72.73	71.59	72.46	73.33	7.70	2.08	0.82
13	69.50	67.45	71.55	71.30	69.57	72.46	2.56	3.09	1.26

The percent difference is computed between systems that use ID^{2}PBT and those that use CRG.

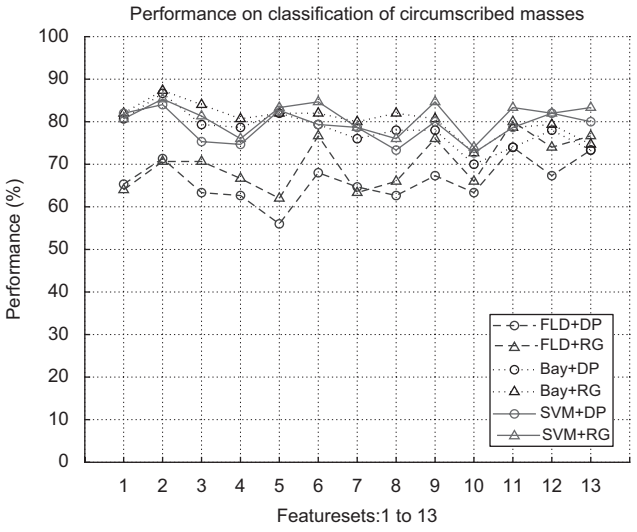

Fig. 5. Performance of systems in classification of circumscribed masses.

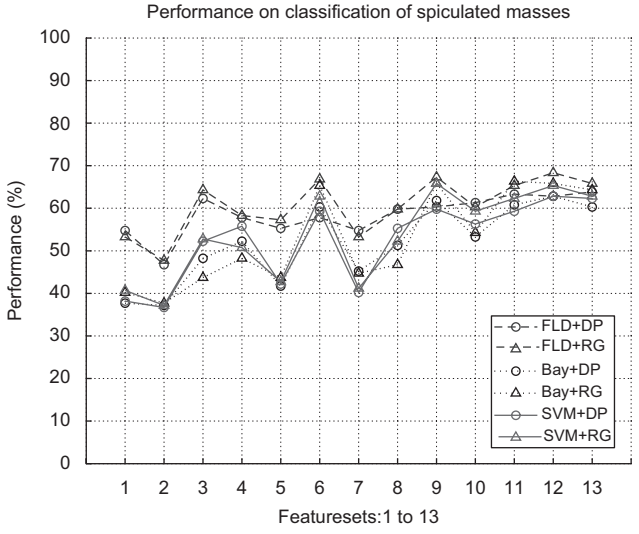

Fig. 6. Performance of systems in classification of non-circumscribed masses.

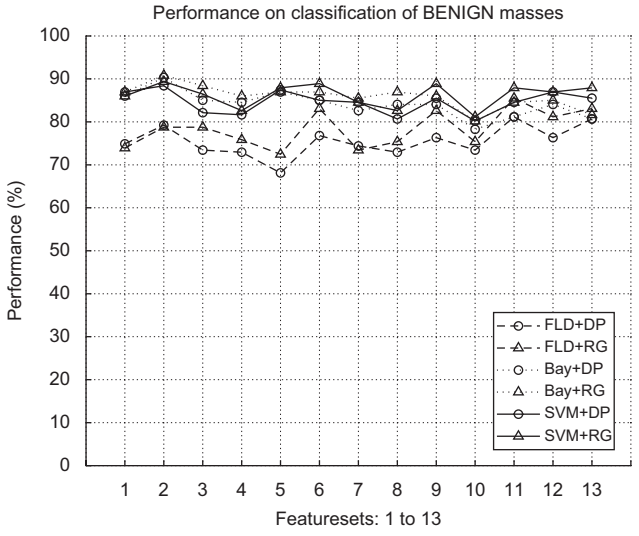

Fig. 7. Performance of systems in classification of benign masses.

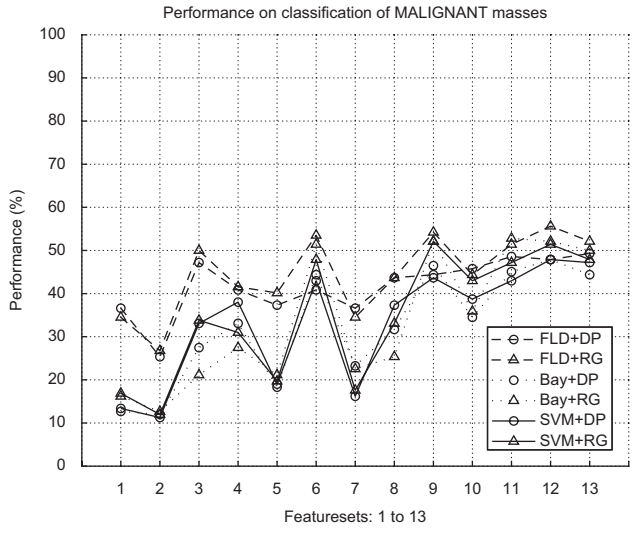

Fig. 8. Performance of systems in classification of malignant masses.

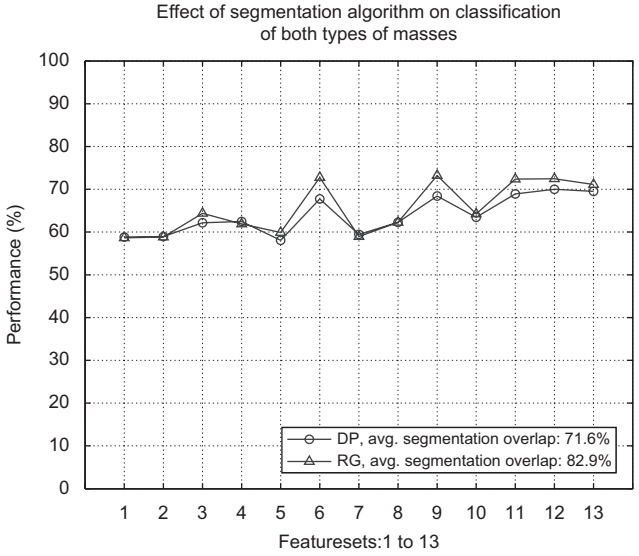

Fig. 9. Performance of systems in classification of all masses.

Table 2
Characterization measures (sensitivity, specificity, positive and negative predictive values) of systems with FLD

Featureset	SE	SP	PPV	NPV
6	0.65, 0.60	0.65, 0.80	0.55, 0.65	0.70, 0.75
7	0.45, 0.45	0.70, 0.65	0.50, 0.50	0.65, 0.65
8	0.60, 0.60	0.60, 0.65	0.50, 0.55	0.70, 0.70
9	0.65, 0.65	0.65, 0.80	0.55 0.65	0.75, 0.75
10	0.65, 0.65	0.60, 0.65	0.55, 0.55	0.75, 0.70
11	0.65, 0.60	0.75, 0.80	0.60, 0.65	0.75, 0.75
12	0.65, 0.65	0.65, 0.75	0.55, 0.65	0.75, 0.75
13	0.65, 0.65	0.70, 0.75	0.60, 0.65	0.75, 0.75

In each case the values before the comma correspond to systems with ID^{2}PBT, and after the comma to systems with CRG.

Table 3
Characterization measures (sensitivity, specificity, positive and negative predictive values) of systems with Bayesian classifier

Featureset	SE	SP	PPV	NPV
6	0.45, 0.45	0.85, 0.95	0.70, 0.80	0.70, 0.70
7	0.25, 0.25	0.85, 0.85	0.50, 0.50	0.60, 0.60
8	0.35, 0.25	0.80, 0.85	0.55, 0.55	0.65, 0.65
9	0.45, 0.45	0.85, 0.90	0.70, 0.80	0.70, 0.70
10	0.55, 0.55	0.70, 0.70	0.55, 0.55	0.70, 0.70
11	0.55, 0.55	0.80, 0.85	0.60, 0.70	0.70, 0.70
12	0.50, 0.50	0.85, 0.85	0.70, 0.75	0.70, 0.70
13	0.55, 0.55	0.75, 0.80	0.60, 0.65	0.70, 0.70

In each case the values before the comma correspond to systems with ID^{2}PBT, and after the comma to systems with CRG.

has been estimated that the differences reported between systems are statistically significant at the 5% level of confidence. The feature sets used are listed below:

(1) Contrast (Co).
(2) Coefficient of variation of edge strength (CV_{ES}).
(3) Measure of the fuzziness of mass margin of first order (Fz_1).
(4) Measure of the fuzziness of mass margin of second order (Fz_2).
(5) Measure of spiculation based on gradient orientation (Sp_{GO}).
(6) Measure of spiculation based on signature information (Sp_{SI}).
(7) $\{Co, CV_{ES}\}$.

Table 4
Characterization measures (sensitivity, specificity, positive and negative predictive values) of systems with SVM

Featureset	SE	SP	PPV	NPV
6	0.45, 0.40	0.85, 0.95	0.70, 0.85	0.70, 0.70
7	0.15, 0.15	0.90, 0.85	0.45, 0.45	0.60, 0.60
8	0.50, 0.40	0.75, 0.80	0.55, 0.55	0.65, 0.65
9	0.45, 0.45	0.85, 0.95	0.70, 0.85	0.70, 0.70
10	0.50, 0.50	0.75, 0.80	0.55, 0.60	0.70, 0.70
11	0.50, 0.50	0.85, 0.90	0.65, 0.75	0.70, 0.70
12	0.55, 0.50	0.85, 0.90	0.70, 0.75	0.75, 0.70
13	0.55, 0.55	0.80, 0.85	0.70, 0.70	0.75, 0.70

In each case the values before the comma correspond to systems with ID^{2}PBT, and after the comma to systems with CRG.

(8) $\{Fz_1, Fz_2\}$.
(9) $\{Sp_{GO}, Sp_{SI}\}$.
(10) $\{Co, CV_{ES}, Fz_1, Fz_2\}$.
(11) $\{Co, CV_{ES}, Sp_{GO}, Sp_{SI}\}$.
(12) $\{Fz_1, Fz_2, Sp_{GO}, Sp_{SI}\}$.
(13) All features.

Of the six individual features, Sp_{SI} was found to be the more successful, and the only one comparable to the combinations of features. Because of this, and in order to save some space, the numerical results of the individual features other than Sp_{SI} were not included in the tables.

7. Discussion

The baseline SE of CADx systems for screening mammography is yet to be defined (CADx systems are not currently used in screening programs). Some of the best results that have been published in the literature (see for example Refs. [26,40,41]) correspond to systems that are still under development and that have employed manually segmented mass contours in feature extraction. Other studies have reported on the use of automated segmentation methods [20,39,42]. Some of the best results obtained by the systems in this work overlap with those in the latter type of studies. For instance, some of the systems studied in this work that employ the FLD as a classifier produced a SE of approximately 0.65, with a SP in the range of 0.65–0.8. The systems that employ the Bayesian classifier or the SVM achieve a SE in the range of approximately 0.50–0.55, with a SP in the range of approximately 0.7–0.9.

Screening mammography has a very low PPV. The specific value varies according to factors such as age, race, family history of breast cancer, and others. According to the literature [43–45], the PPV of mammography is in the range of 0.03–0.35 (this includes all types of abnormalities, not only masses). This means that on average for every 100 biopsies (or similar procedure ordered to confirm a diagnosis of malignancy based on mammography), only 35 (in the best case) will confirm the diagnosis; the other 65–97 procedures would have been performed, in retrospect, unnecessarily. Of the systems studied in this work, the ones with the highest SE show a PPV in the range 0.60–0.75, which is roughly the double of the PPV of screening mammography.

In general, the features studied showed a much higher ability of classifying the benign masses in the dataset (78–98% correctly classified) than the malignant masses (0–51%), compare Fig. 7 with Fig. 8. Since circumscribed masses are in most cases benign, the classification success of circumscribed masses was also higher (76–85%) than the classification success of spiculated and microlobulated masses (42–67%), compare Fig. 5 with Fig. 6.

When examining individual features (see Figs. 7 and 8), it was observed that features Co, CV_{ES}, and Sp_{GO} produce the largest percentage of benign masses correctly classified, together with the smallest percentage of malignant masses correctly classified. However, this behavior is observed only with the use of the Bayesian classifier or the SVM. In the case of the FLD, the percentages of correct classification are much more balanced between the two classes. This indicates that there exists an evident overlap between classes in each of these features that the nonlinear classifiers resolve in favor of one class (the one with more instances in the training set) and in detriment of the other. The FLD, in contrast, compromises between the two classes and accepts a similar error in both classes. Nevertheless, the overall classification success (considering both benign and malignant masses) is very similar amongst classifiers. The choice of segmentation method had very little effect on the results of the systems. These features can be considered robust to differences in the (average) quality of the segmentation.

In the case of the features Fz_1 and Fz_2, the success rate of classification was also in favor of the benign class. However, in this case the tendency was not as strong as with the features discussed above. The choice of a segmentation method had a noticeable effect on the classification of circumscribed masses and of spiculated and microlobulated masses (see Figs. 5 and 6), but in the overall classification the effect was only moderate and affecting feature Fz_1 more than Fz_2. These features are relatively robust to differences in the quality of segmentation, specially when they are combined with each other.

The feature Sp_{SI} demonstrated the highest discriminant ability of any single feature (approximately 73% with CRG, and approximately 68% with ID^{2}PBT). The number of benign masses correctly classified and the number of malignant masses correctly classified were also more similar than in the cases of the features discussed above. Roughly, the proportion was 89% of benign masses correctly classified, and 49% of malignant masses correctly classified, when either the Bayesian classifier or the SVM was used. A moderate decrease in these percentages (around 5%) was observed between the systems with CRG and the ones with ID^{2}PBT. If the FLD was used, the proportion was 80% and 60% with CRG segmentation, but decreased to roughly 64% for both classes if the segmentation method was ID^{2}PBT. This indicates that the discrimination ability of this feature is affected by the quality of the segmentation, in particular when a relatively simple classifier is employed. The strength of this feature is its ability of classifying, relatively well, both the class of circumscribed masses as well as the class of spiculated and microlobulated masses (overall average of approximately 79% circumscribed, 64% spiculated and microlobulated, see Figs. 5 and 6).

The combination $\{Sp_{SI}, Sp_{GO}\}$ with CRG segmentation achieved an average of approximately 73% masses correctly classified (approximately 88% benign masses and 51% malignant masses correctly classified). With ID^{2}PBT the average performance was 68.4% (80% benign, 52% malignant). These quantities indicate that the performance of the combined features are roughly the same as the performance of the Sp_{SI} feature alone. The similarity is also observed in terms of the PPV, which is practically the same for the combined featureset and the Sp_{SI} feature. See Fig. 9 and compare the PPV values of featuresets 6 and 9 in Tables 2–4.

The combination $\{Co, CV_{ES}, Fz_1, Fz_2\}$ (featureset 10) achieved an average classification success of approximately 64%, which is only slightly higher (approximately 1.6% higher) than the performance achieved by the combination $\{Fz_1, Fz_2\}$ (featureset 8). In general terms, the choice of the segmentation method did not have an effect on the performance of these features (see Fig. 9). The PPV of featureset 10 was the same as that of featureset 8. All this indicates that Co and CV_{ES} do not have any positive influence when combined with Fz_1 and Fz_2.

In the case of featuresets 11–13, the average correct classification values of systems with CRG were always slightly higher than those of systems with ID^{2}PBT (average of 2.5% higher). The performance

achieved by these featuresets was around 72% (Fig. 9). The PPV was almost the same between featuresets, and the values were slightly lower with the FLD than with the Bayesian classifier and the SVM (0.62, 0.67, and 0.71, respectively).

It can be concluded that in terms of the SP, PPV, and average classification performance, the feature Sp_{SI} is the most effective in predicting the diagnosis of the masses in the dataset. The combination of this feature with some of the other features discussed can produce a slightly higher SE and in some cases (e.g. featureset 11) reduce the difference in performance between systems with the same segmentation method.

8. Conclusions

Ultimately, the objective of breast mass analysis from mammograms is to be able to predict the nature of masses (i.e. to diagnose). However, it cannot be said that a spiculated mass is definitively malignant, or that a non-spiculated mass is definitively benign. This poses a problem for researchers because, from a theoretical point of view, one can only derive conclusions about the physical characteristics of the masses that are perceivable from the mammographic images; thus, diagnosis corresponds to a second degree of analysis where other types of information (such as the clinical profile and background of each patient) should be included. In other words, the information available at the mammographic level only permits the characterization of masses with some amount of certainty, but, strictly speaking, not their diagnosis.

The features presented possess, by design, a limited discrimination power. This is because a number of necessary approximations were introduced to compensate for the uncertainty of mass-boundary information. For example, the design implicitly considers all masses to be approximately circular in shape, which is certainly not accurate. If there were a guarantee that boundaries extracted by automated segmentation algorithms would comply to a certain degree of precision with the true boundaries of all masses, then the design of the features could be adapted to follow such contours. Without such guarantee, the trade-off between performance and feasibility of implementation will continue to be a factor to consider while designing features for automated analysis of breast masses.

Given the characterization measures reported in the previous section, it can be concluded that some of the systems studied, and in particular, those utilizing the feature Sp_{SI}, are useful in the diagnosis of benign masses. This is indicated by the specificity and the NPV of these systems. A high sp indicates that a system is able to recognize a majority of the negative cases. Furthermore, it is well known that, due to the prevalence rates of breast cancer, the number of benign breast masses in a given population of screening mammograms with abnormal findings is predictably much larger than the number of cancerous masses in the same population. Approximately 80% of the total of benign masses can be diagnosed correctly by some of the systems; at the same time, about 50% of the malignant masses will be misdiagnosed (because of the low SE of the systems). However, due to the comparatively small fraction of malignant masses, these numbers indicate that a vast majority of the negative-diagnoses returned by the systems will be correct. If, for example, the fraction of malignant masses is 0.1, then 90% of the negative-diagnoses returned by the systems would be correct. With a lower fraction (closer to the actual scenario) this percentage would be higher.

Regarding the effect of the segmentation method used in CAD systems, if it is assumed that the systems are to be completely automated, then the results obtained by systems with ID^{2}PBT segmentation are to be considered as the baseline performance. The results of the systems with CRG can then be used to obtain some conclusions as to what would happen if there is an improvement in the performance of the segmentation method. Due to the use of the simplified contours in the feature extraction procedure, a 14% improvement in the quality of the segmentation is translated into only about one-third of that improvement in the final task of diagnosis of the masses (compare the percent difference values in Table 1 with the percent difference of the segmentation task between segmentation algorithms, which is 14.19%). An average segmentation improvement of 14% (absolute difference of 0.11 in the AOM) is hard to achieve. This improvement would not be sufficient to justify a feature extraction procedure based directly on the contours extracted in an automated fashion (a larger improvement is needed). Thus, one can conclude that given a scenario of limited resources, these would be well employed in the design of classification features with more discriminative power and that are more robust to inaccuracies of the contours extracted by automated methods (since it is probable that these inaccuracies will still be significant even if a very efficient segmentation method is used).

Regarding the operation of the classifiers, the distribution of the classification error indicates that in order to increase the sensitivity of the systems, perhaps the classifiers could be trained using a scheme that penalizes the misclassification of the instances in the malignant class or otherwise accepts a very small classification error in this class, independently of the overall error. Another possibility is to adopt more advanced classifiers that those employed in this study. For instance, Mu et al. [40] have proposed a nonlinear pairwise Rayleigh quotient classifier that is able to improve on the performance of some types of SVMs. Similarly, Nandi et al. [41] have applied genetic programming combined with feature selection to the classification of breast masses, with very promising results.

As a concluding remark it should be pointed out that, out of the features studied, the most effective feature (Sp_{SI}) is different from the other features in that it does not attempt to compute an absolute measure (of spiculation), but rather, it compares the evidence indicating that a particular mass may be spiculated with the evidence indicating that it may be not. In a sense, the feature is the result of a classification performed based on the information contained in a single mass. Perhaps this is what makes Sp_{SI} distinctly better than the other features presented, and it may be a possible line of research to follow in future work regarding the design of features for characterization of breast masses.

References

[1] Breast cancer factsheet—February 2004, Cancer Research UK, February 2004, Available online at: ⟨http://info.cancerresearchuk.org/ourpublications/⟩.
[2] Canadian cancer statistics 2006, National Cancer Institute of Canada, 2006. Available online at: ⟨http://www.cancer.ca⟩.
[3] C. Di Maggio, State of the art of current modalities for the diagnosis of breast lesions, Eur. J. Nucl. Med. Mol. Imaging 31 (Supplement 1) (2004) S56–S69.
[4] A. Rojas, A.K. Nandi, Improved dynamic-programming-based algorithms for segmentation of masses in mammograms, Med. Phys. 34 (2007) 4256–4269.
[5] M.A. Kupinski, M.L. Giger, Automated seeded lesion segmentation on digital mammograms, IEEE Trans. Med. Imaging 17 (4) (1998) 510–517.
[6] G.M. te Brake, N. Karssemeijer, Segmentation of suspicious densities in digital mammograms, Med. Phys. 28 (2) (2001) 259–266.
[7] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, Wiley-Interscience, New York, USA, 2001.
[8] A. Webb, Statistical Pattern Recognition, second ed., Wiley, West Sussex, England, 2002.
[9] S. Theodoridis, K. Koutroumbas, Pattern Recognition, 3rd ed., Academic Press, San Diego, California, USA, 2006.
[10] 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 (4) (2006) 646–668.
[11] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S. Kok, P. Taylor, D. Betal, J. Savage, The Mammographic Image Analysis Society digital mammogram database, Exerpta Medica, International Congress Series, vol. 1069, 1994, pp. 375–378, Available online at: ⟨http://peipa.essex.ac.uk/info/mias.html⟩.

[12] M. Heath, K. Bowyer, D. Kopans, P. Kegelmeyer Jr., R. Moore, K. Chang, S. Munishkumaran, Current status of the digital database for screening mammography, in: N. Karssemeijer, M. Thijssen, J. Hendriks, L. van Erning (Eds.), Digital Mammography, Kluwer Academic Publishers, Boston, Massachusetts, USA, 1998, pp. 457–460, Available online at: ⟨http://marathon.csee.usf.edu/⟩.
[13] R.C. Gonzalez, R.E. Woods, S.L. Eddins, Digital Image Processing Using MATLAB, Prentice-Hall, NJ, USA, 2004.
[14] U. Montanari, On the optimal detection of curves in noisy pictures, Commun. ACM 14 (5) (1971) 335–345.
[15] J.J. Gerbrands, Segmentation of noisy images, Ph.D. dissertation, Delft University of Technology, The Netherlands, 1988.
[16] J.H.C. Reiber, S.P. Lie, M.L. Simoons, C. Hoek, J.J. Gerbrands, W. Wijns, W.H. Bakker, P.P.M. Kooij, Clinical validation of fully automated computation of ejection fraction from gated equilibrium blood-pool scintigrams, J. Nucl. Med. 24 (12) (1983) 1099–1107.
[17] J. Nuyts, P. Suetens, A. Oosterlinck, M. De Roo, L. Mortelmans, Delineation of ECT images using global constraints and dynamic programming, IEEE Trans. Med. Imaging 10 (4) (1991) 489–498.
[18] L. Maes, B. Bijnens, P. Suetens, F.V. de Werf, Automated contour detection of the left ventricle in short axis view in 2D echocardiograms, Mach. Vision Appl. 6 (1) (1993) 1–9.
[19] P. de Bruin, F. Vos, A. Vossepoel, F. Post, S. de Blok, A. Aarden, Supporting hysteroscopic surgery by 3D imaging, modelling and visualization, in: M. Boasson, J. Kaandorp, J. Tonino, M. Vosselman (Eds.), Proceedings of the Fifth Annual Conference of the Advanced School for Computing and Imaging, June 1999, pp. 46–51.
[20] S. Timp, N. Karssemeijer, A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography, Med. Phys. 31 (5) (2004) 958–971.
[21] J.G. Bosch, Automated contour detection in echocardiographic images, Ph.D. Dissertation, Leiden University, Leiden, The Netherlands, 2006.
[22] N.R. Mudigonda, R.M. Rangayyan, J.E.L. Desautels, Detection of breast masses in mammograms by density slicing and texture flow-field analysis, IEEE Trans. Med. Imaging 20 (12) (2001) 1215–1227.
[23] A.R. Rao, B.G. Schunck, Computing oriented texture fields, CVGIP: Graphical Models Image Process. 53 (1991) 157–185.
[24] R. Halif, J. Flusser, Numerically stable direct least squares fitting of ellipses," Department of Software Engineering, Charles University, Czech Republic, Technical Report, 2000.
[25] B. Sahiner, H. Chan, N. Petrick, M. Helvie, M.M. Goodsitt, Computerized characterization of masses on mammograms: the rubber band straightening transform and texture analysis, Med. Phys. 20 (4) (1998) 516–526.
[26] N.R. Mudigonda, R.M. Rangayyan, J.E.L. Desautels, Gradient and texture analysis for the classification of mammographic masses, IEEE Trans. Med. Imaging 19 (10) (2000) 1032–1043.
[27] P. Kovesi, Image features from phase congruency, Videre: J. Comput. Vision Research 1 (3) (1999) 1–26.
[28] Z. Wang, M. Jenkin, Using complex gabor filters to detect and localize edges and bars, in: C. Archibald, E. Petriu (Eds.), Advances in Machine Vision: Strategies and Applications, World Scientific Publishing Co., Inc. River Edge, NJ, USA, 1992, pp. 151–170.
[29] R.M. Rangayyan, Biomedical Image Analysis, CRC Press, Florida, USA, 2005.
[30] T.M. Cover, J.A. Thomas, Elements of Information Theory, second ed., Wiley-Interscience, New Jersey, USA, 2006.
[31] J.P.W. Pluim, J.B.A. Maintz, M.A. Viergever, Mutual-information-based registration of medical images: a survey, IEEE Trans. Med. Imaging 22 (8) (2003) 986–1004.
[32] H. Kobatake, M. Murakami, H. Takeo, S. Nawano, Computerized detection of malignant tumors on digital mammograms, IEEE Trans. Med. Imaging 18 (5) (1999) 369–378.
[33] C. Varela, P.G. Tahoces, A.J. Mendez, M. Souto, J.J. Vidal, Computerized detection of breast masses in digitized mammograms, Comput. Biol. Med. 37 (2) (2007) 214–226.
[34] P. Kovesi, Phase congruency detects corners and edges, in: Proceedings of the Australian Pattern Recognition Society Conference, Sydney, December 2003, pp. 309–318.
[35] C. Varela, S. Timp, N. Karssemeijer, Use of border information in the classification of mammographic masses, Phys. Med. Biol. 51 (2) (2006) 425–441.
[36] R.A. Fisher, The use of multiple measurements in taxonomic problems, Ann. Eugen. 7 (1936) 179–188.
[37] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (1995) 273–297.
[38] S. Abe, Support Vector Machines for Pattern Classification (Advances in Pattern Recognition), first ed., Springer-Verlag London Limited, London, UK, 2005.
[39] B. Sahiner, N. Petrick, H. Chan, L.M. Hadjiiski, C. Paramagul, M.A. Helvie, M.N. Gurcan, Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization, IEEE Trans. Med. Imaging 20 (12) (2001) 1275–1284.
[40] T. Mu, A.K. Nandi, R.M. Rangayyan, Analysis of breast tumors in mammograms using the pairwise rayleigh quotient classifier, J. Electron. Imaging 16 (2007) 043004-1–043004-11.
[41] R.J. Nandi, A.K. Nandi, R.M. Rangayyan, D. Scutt, Classification of breast masses in mammograms using genetic programming and feature selection, Med. Biol. Eng. Comput. 44 (8) (2006) 683–694.
[42] P. Delogu, M.E. Fantacci, P. Kasae, A. Retico, Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier, Comput. Biol. Med. 37 (10) (2007) 1479–1491.
[43] D.B. Kopans, The positive predictive value of mammography, Am. J. Roentgenol. 158 (1992) 521–526.
[44] M.L. Brown, F. Houn, E.A. Sickles, L.G. Kessler, Screening mammography in community practice: positive predictive value of abnormal findings and yield of follow-up diagnositic procedures, Am. J. Roentgenol. 165 (1995) 1373–1377.
[45] K. Kerlikowske, D. Grady, J. Barclay, E.A. Sickles, A. Eaton, V. Ernster, Positive predictive value of screening mammography by age and family history of breast cancer, J. Am. Med. Assoc. 270 (1993) 2444–2450.
[46] American College of Radiology (ACR), Breast Imaging Reporting and Data System (BI-RADS), 4th ed., American College of Radiology, Reston, Virginia, USA, 2003.

About the Author—ALFONSO ROJAS DOMÍNGUEZ was born in Mexico City, Mexico, in 1978. He received the degree of B.Sc. in Telecommunications Engineering from the National Autonomous University of Mexico (UNAM) in 2002, the degree of M.Sc. with distinction in Intelligence Engineering from the University of Liverpool, UK, in 2003, and the degree of Ph.D. in Electrical Engineering and Electronics, from the same university, in 2007. His research interests include automated image processing and analysis, machine learning, pattern recognition, and genetic programming.

About the Author—ASOKE K. NANDI received the degree of Ph.D. from the University of Cambridge (Trinity College), Cambridge, UK, in 1979. He held positions in Rutherford Appleton Laboratory (UK), European Organization for Nuclear Research (Switzerland), Queen Mary College (London, UK), University of Oxford (UK), Imperial College (London), and University of Strathclyde (Glasgow, UK). Since March 1999 he holds the David Jardine Chair of Signal Processing at the University of Liverpool. In 1983 he was a member of the UA1 team at CERN that discovered the three fundamental particles known as W^+, W^- and Z^0 providing the evidence for the unification of the electromagnetic and weak forces, which was recognized by the Nobel Committee for Physics in 1984. Currently, he is interested in the areas of bio-medical signal processing, communications, and machine learning research. He has authored or co-authored over 350 technical publications. He is a Fellow of the Cambridge Philosophical Society, the Institution of Electrical Engineers, the Institute of Mathematics and its applications, the Institute of Physics, the Institution of Mechanical Engineers, the British Computer Society, and the Royal Society for Arts.