



LBP operators on curvelet coefficients as an algorithm to describe texture in breast cancer tissues



Daniel O. Tambasco Bruno^{a,1}, Marcelo Z. do Nascimento^{a,b,*}, Rodrigo P. Ramos^{c,2}, Valério R. Batista^{a,1}, Leandro A. Neves^{d,3}, Alessandro S. Martins^{e,4}

^a UFABC - CMCC, av. dos Estados 5001, Bl.B, 09210-580 St. André, SP, Brazil

^b UFU - FACOM, av. João Neves de Ávila 2121, Bl.B, 38400-902 Uberlândia, MG, Brazil

^c UNIVASF - CENEL, av. Antônio C. Magalhães 510, 48902-300, Juazeiro, BA, Brazil

^d UNESP - DCCE, r. Cristóvão Colombo 2265, 15054-000, S. J. Rio Preto, SP, Brazil

^e IFMT, r. Belarmino Vilela Junqueira S/N, 38305-200, Ituiutaba, MG, Brazil

ARTICLE INFO

Keywords:

Breast cancer tissues
Texture analysis
Local binary pattern
Curvelet transform
Computer aided diagnosis
Polynomial classifier

ABSTRACT

In computer-aided diagnosis one of the crucial steps to classify suspicious lesions is the extraction of features. Texture analysis methods have been used in the analysis and interpretation of medical images. In this work we present a method based on the association among curvelet transform, local binary patterns, feature selection by statistical analysis and distinct classification methods, in order to support the development of computer aided diagnosis system. The similar features were removed by the statistical analysis of variance (ANOVA). The understanding of the features was evaluated by applying the decision tree, random forest, support vector machine and polynomial (PL) classifiers, considering the metrics accuracy (AC) and area under the ROC curve (AUC): the rates were calculated on images of breast tissues with different physical properties (commonly observed in clinical practice). The datasets were the Digital Database for Screening Mammography, Breast Cancer Digital Repository and UCSB biosegmentation benchmark. The investigated groups were normal-abnormal and benign-malignant. The association of curvelet transform, local binary pattern and ANOVA with the PL classifier achieved higher AUC and AC values for all cases: the obtained rates were among 91% and 100%. These results are relevant, specially when we consider the difficulties of clinical practice in distinguishing the studied groups. The proposed association is useful as an automated protocol for the diagnosis of breast tissues and may contribute to the diagnosis of breast tissues (mammographic and histopathological images).

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Routine physical breast examination should always be complemented with further analyses through clinical equipment techniques. These are screen-film mammography, ultrasonography, magnetic resonance and tomosynthesis. Regarding accessibility combined with cost-benefit for early detection of lesions, the

screen-film mammography still prevails as the best one (Hussain, 2013; Nascimento et al., 2013). As a complementary protocol, histological images of breast cancers are also evaluated to confirm a preliminary diagnosis. This procedure is commonly applied in clinical practice (Cireşan, Giusti, Gambardella, & Schmidhuber, 2013; Malhotra, Zhao, Band, & Band, 2010; Tashk, Helfroush, Danyali, & Akbarzadeh-jahromi, 2015; Veta et al., 2015; Veta, Pluim, van Diest, Viergever et al., 2014).

Precision and correctness are two requirements always expected on a diagnosis, specially because studies show that cancer cases have been increasing over the years. The 2012 report of the World Health Organization's International Agency for Research on Cancer indicates that 8.2 million obits were caused by cancer in that year. This means a rise of 8% compared to the previous 2008 report. Specifically in respect to breast cancer there has been a rise of 14%, with an amount of 522,000 women in 2012 (Ferlay et al., 2012).

* Corresponding author at: UFU - FACOM, av. João Neves de Ávila 2121, Bl.B, 38400-902 Uberlândia, MG, Brazil. Tel.: +55 34 3239 4531, +55 11 4996 0077; fax: +55 34 3239 4144.

E-mail addresses: danielbruno@gmail.com (D.O. Tambasco Bruno), marcelo.zanchetta@gmail.com (M.Z. do Nascimento), rodrigo.ramos@univasf.edu.br (R.P. Ramos), vramos1970@gmail.com (V.R. Batista), neves.leandro@gmail.com (L.A. Neves), alessandros2006@yahoo.com.br (A.S. Martins).

¹ Tel.: +55-11-4996-0077.

² Tel.: +55-74-2102-7630.

³ Tel.: +55-17-3221-2211.

⁴ Tel.: +55-34-3271-4000.

As we have mentioned, clinical equipment techniques help hinder these rates. The cutting-edge one is tomosynthesis but very few women can afford it in developing countries. As an example, in Brazil the women population is approximately 100 million, where the cases of breast cancer for 2014 were predicted to be 57,120 with a fatal rate of 23.4% (INCA, 2014).

In spite of being the most accessible technique, mammographic images are difficult to analyse. This occurs because the breast can have some adjacent tissues similar to lesions. Such false positive lesions must be ruled out by the analysis of radiologists (Thurfjell, Lernevall, & Taube, 1994; Warren & Duffy, 1995). However, nowadays there is a high demand for diagnoses, which increases both the operational costs and the workload of these professionals. This leads to more frequent human error resulting in incorrect diagnoses.

Since early 1990, computer aided diagnosis (CAD) has been an alternative tool to supplementary analysis that helps radiologists achieve more precise diagnoses. The CAD system can be divided into two parts: computer-aided detection (CADe) and computer-aided diagnosis (CADx). CADe schemes are systems that locate suspicious lesions in mammograms automatically. In Oliver et al. (2010) there is a good survey on the state of the art of CADe systems.

CADx systems are devoted to classify lesions. They complement the CADe analysis with the characterisation of regions and the computation of probabilities of lesion malignancy (Elter & Horsch, 2009). In CADx schemes, one of the crucial steps to classify suspicious lesions is the feature extraction. Typically, two classes of characteristics are extracted from mammograms, namely morphological and non-morphological features.

Morphological features give information about size and shape of the lesion. Non-morphological features work on grey level properties. For instance, a consistent variation of grey levels in the image is detected as a pattern. Patterns are then grouped to form a set called *image texture*.

Texture is a general concept that also applies to colour images. In this case patterns are identified by colour gradients along the image (Gonzalez & Richard, 2002). Texture involves basic grey levels or colour texture primitives that form elements called *textons* (Zhu, Guo, Wang, & Xu, 2005) or *texels* (Haralick, 1979). Several methods have been developed for texture analysis, being classified as statistical, model-based and frequency-based (Tang & He, 2013).

Statistical methods are based on the distribution of the grey scale. Some of them use the grey scale histogram as a measure of texture (Gupta & Markey, 2005; Souza Jacomini, Nascimento, Dantas, & Ramos, 2012). The grey scale distribution can be described by either first or second order statistics and summarised as a co-occurrence matrix (Nanni, Brahnam, Ghidoni, & Menegatti, 2014; Nanni, Brahnam, Ghidoni, Menegatti, & Barrier, 2013). The metric considers only pixels individually, but this causes measures to be more sensitive to image pixel changes. Finally, in Schwartz, Roberti de Siqueira, and Pedrini (2012) the authors use a statistical method that works on transitions between grey levels of pixels.

Model-based approaches aim at establishing stochastic properties that can define texture. Some techniques used for this purpose are Markov random fields (Suliga, Deklerck, & Nyssen, 2008; Yu & Huang, 2010), fractal features (Neves, do Nascimento, & de Godoy, 2014) and autoregressive model (Mayerhoefer et al., 2010). They can describe texture through some few parameters of microtextures, which however give macrotexture information in the end. Therefore, either little is known about the texture or there exist more than one possible texture.

Frequency-based models produce information derived from local operators and statistical attributes of images in the frequency domain. Some examples of these models are the wavelet transform (Dheeba & Tamil Selvi, 2012; Jacomini, Nascimento, Dantas,

& Ramos, 2013; Nascimento et al., 2013), the ridgelet transform (Ramos, Nascimento, & Pereira, 2012) and the curvelet transform (Eltoukhy, Faye, & Belhaouari Samir, 2010a). They decompose the original image into subbands that preserve high and low frequency information. These are the most adequate techniques to extract texture, for they enable the image to be represented by multiple scales in a way that is quite close to what is done by the human eyes (Tang & He, 2013).

Several recent works have been developed for the diagnosis of suspicious regions in mammograms, mainly focused on the understanding of feature extraction from images of breast lesions. Thus, a relevant problem in CAD is the study of normal, benign and malignant lesions in images with suspicion of breast cancer. For instance, in Liu and Tang (2014) the authors extract texture features based on a grey level co-occurrence matrix (GLCM). In Rouhi, Jafari, Kasaei, and Keshavarzian (2015) the authors present approaches based on artificial neural network, cellular neural network and genetic algorithm. In Liu and Zeng (2015) the authors develop a method considering an adaptive region growing method to locate suspicious regions. Afterwards, they extract geometrical and texture features (GLCM and completed local binary pattern (CLBP)) that are then applied to classify regions of interest (ROIs). The ROIs are classified by means of support vector machine (SVM), with supervision provided by the diagnosis of a radiologist. In Abdel-Nasser, Rashwan, Puig, and Moreno (2015) the authors present a system with segmentation of the ROIs, feature extraction based on uniform local directional pattern and classifications using the SVM. In Rouhi and Jafari (2016) the authors consider texture feature computed from GLCM and CLBP, with classification obtained via SVM and supervised by radiologists. In these works the authors conclude that these approaches can describe breast tissues properly and with promising results. However, as mentioned in Abdel-Nasser et al. (2015), it is not possible to elect just one of these feature descriptors as optimal to quantify breast tissues.

Also, there are many techniques to identify patterns, such as local and mid-level features (Kong, Jiang, & Yang, 2015). For instance, local binary patterns (LBP) have been relevant to texture analysis in several applications (Guo, Zhao, & Pietikäinen, 2014; Nanni, Lumini, & Brahnam, 2010; Song, Yan, Chen, & Zhang, 2013). The LBP is an effective texture description operator with many significant advantages, for it generates histograms that are very useful to represent texture features that are invariant by rotations and brightness levels. LBP labels can be regarded as local primitives such as curved edges, spots and flat areas, among others. They have been largely applied to classify abnormalities detected in images (Choi, Kim, & Ro, 2012; Llad, Oliver, Freixenet, Mart, & Mart, 2009).

In fact, there is not a universal texture descriptor that always gives the best quantification for all different kinds of images. However, curvelet transform and LBP appear to be more effective for texture analysis. Eltoukhy, Faye, and Belhaouari Samir (2010a) used the 100 greatest coefficients from each decomposition level of the curvelet transform in order to identify lesions in mammograms. Zhang and Pham (2011) applied curvelet transform with Haralick to describe microscopy images. Zhou Li-Jian (2012) also applied curvelet transform and LBP to solve problems in face recognition when illumination is variable. In their approach the LBP technique was applied only over subband level 1. Nagaraja, Prabhakar, and Praveen Kumar (2013) proposed a technique for facial expression representation based on a combination that uses curvelet transform to obtain the several approximation subbands, of which the ones with highest energy are selected and then submitted to CLBP. Al-Hammadi, Muhammad, Hussain, and Bebis (2013) used the LBP technique over the curvelet transform coefficients and SVM classifier to describe information from chrominance components for image forgery detection. Dhahbi, Barhoumi, and Zagrouba (2015) investigated the moment theory to characterise the distribution

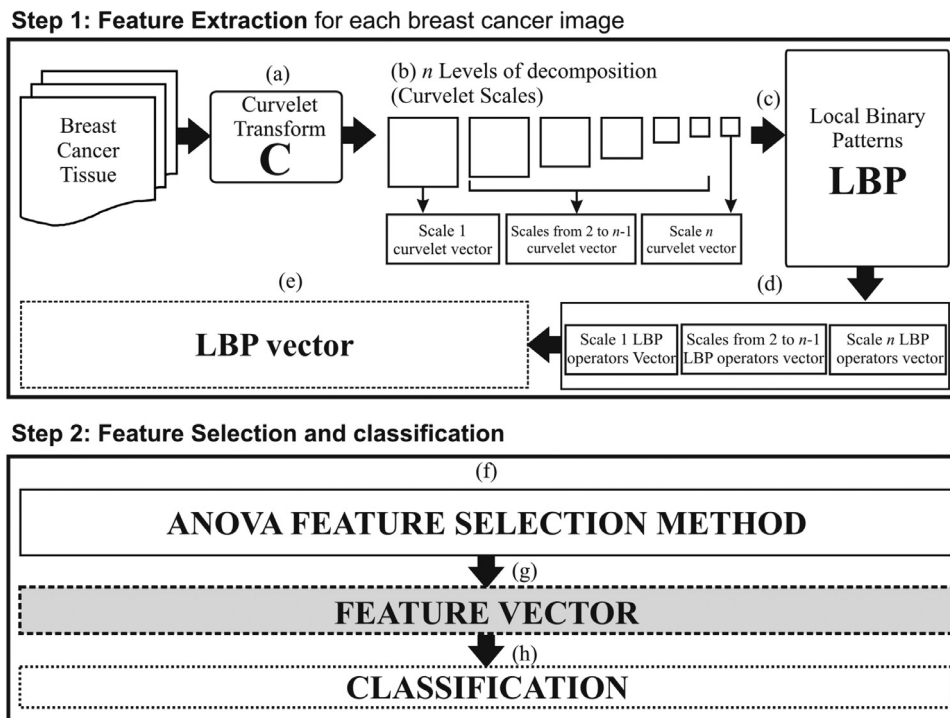


Fig. 1. Block diagram of the main steps for processing images; (a) application of the curvelet transform to the image; (b) extraction of the curvelet coefficients to construct scale vectors for each decomposition; (c) application of the LBP to the coefficients; (d) construction of the LBP-vectors; (e) aggregation of the LBP-vectors; (f) reapplication of the previous steps to all the images; (g) ANOVA application; (h) vector of features; (i) classification.

of curvelet coefficients on mammograms. The combination of LBP and curvelet transform was also studied by considering part of the curvelet subbands or coefficients (Eltoukhy, Faye, & Belhaouari Samir (2010a); Nagaraja et al., 2013; Zhou Li-Jian, 2012), by applying different descriptors over curvelet transform coefficients (Dhahbi et al., 2015; Zhang & Pham, 2011) or even by associating curvelet transform, LBP and only one classification method, namely the SVM algorithm (Al-Hammadi et al., 2013).

In this context, a method that combines curvelet transform, LBP and feature selection by statistical analysis is a relevant and unexplored combination, specially when we consider different image datasets commonly studied by specialists in clinical practice. In this paper we present and detail such a method, which extracts features from images of breast cancer tissues, and also supports the development of CAD systems. Another innovative characteristic of our method is that it defines the best association after applying well known algorithms of artificial intelligence, such as decision tree (DT), random forest (RaF), SVM and polynomial (PL) (Moschidis, Chen, Taylor, & Astley, 2014; Nascimento et al., 2013; de Nazar Silva, de Carvalho Filho, Corra Silva, Cardoso de Paiva, & Gattass, 2015; Ramos et al., 2012). The results were evaluated by considering both the rates accuracy (AC) and area under the ROC curve (AUC), which were computed for images with different physical properties. The investigated groups were normal, benign and malignant (Nascimento et al., 2013).

The contributions of this paper can be summarised as follows:

- An original combination of curvelet transform, local binary patterns and feature selection with statistical analysis that, together with distinct classification methods (SVM, DT, RaF and PL), was tested in order to support the development of CAD systems;
- The evaluation of ANOVA as the statistical method applied to exclude similar features after associating the curvelet transform with the LBP descriptor, in order to enhance the classification step with well known algorithms of artificial intelligence;

- The identification of the PL classifier as the one that, in some cases, builds an ideal system when aggregated to our proposed combination of curvelet transform, LBP and ANOVA. Even in the non-ideal cases the PL classifier always achieved the highest AUC and AC values after tests with different datasets of breast images, such as mammograms (digital and analogical) and histology.

2. Methodology

The steps involved on the feature extraction algorithm are schematised in Fig. 1. They are based on the application of LBP to curvelet transforms with the purpose of differentiating normal, benign and malignant in breast tissue images. The features are extracted from the breast tissue images and grouped into three classes. After obtaining the desired features we use the ANOVA technique to select features that are statistically different. Finally, we apply the PL classifier for distinguishing among the different groups. In this study there is a comparison with other classification methods (DT, RaF and SVM). Each step is detailed in the following subsections.

2.1. DataSets

We evaluated our proposed approach with four distinct classifiers, two metrics and four breast cancer datasets obtained with different clinical equipment techniques. The datasets used in our evaluation were:

- The *digital database for screening mammography (DDSM)*: this dataset (Heath, Bowyer, Kopans, Moore, & Kegelmeyer, 2001) contains 2500 cases, and each of them is composed of 4 images: CC-left, MLO-left, CC-right and MLO-right. We have worked with cases in which the number of masses was at most one. More specifically, we have worked with 240 ROIs cut out

at a standard size of 512×512 pixels of the CC-view. Our reference point to locate the centre of this frame was given by the coordinates of the region of interest (ROI) in the cases malignant and benign. For the normal case the ROI was taken at random. Despite the random selection there was special care to guarantee that each ROI had a resolution of 12 bits per pixel and a quantisation of $50 \mu\text{m}$.

- *Breast cancer digital repository (BCDR)*: this database (López et al., 2012; Moura & López, 2013) is subdivided in two different datasets: the film mammography repository (BCDR-FMR) and the full field digital mammography repository (BCDR-DMR). In the BCDR-FMR the images have grey level with a bit depth of 8 bits per pixel and in the BCDR-DMR they have a bit depth of 14 bits per pixel. For BCDR-FMR we have worked with 106 ROIs cut out at a standard size of 512×512 pixels. These are mammogram ROIs obtained from the CC-view of the types malignant, benign and normal. In the case of BCDR-DMR we used 95 ROIs of images of that three types as well. The databanks include in each mammogram the coordinates of the ROI for malignant and benign. In the normal case the ROI was taken at random.
- The *UCSB biosegmentation benchmark (UCSB-BB)*: this dataset (Gelasca, Obara, Fedorov, Kvilekval, & Manjunath, 2009) belongs to David Rimm's Laboratory at Yale. It contains 58 H&E stained histopathology images that were used in order to detect breast cancer cells. We have worked with all the 58 images, half of them show the malignant case and the other half the benign. They are colour images in the RGB system with 896×768 pixels of resolution and 24 bits of quantisation. In this study we also converted them into grey level images.

2.2. Extraction of descriptors

Textural features were used to represent medical images. The vectors of features were obtained by applying LBP operators on curvelet transform coefficients obtained from the breast tissue images. These two techniques and the way they were used will be resumed next.

2.2.1. The fast discrete curvelet transform

In Candes, Demanet, Donoho, and Ying (2006) the authors introduced a variation of the curvelet transform called *fast discrete curvelet transform* (FDCT). Instead of the complex ridgelet, the FDCT uses *fast Fourier transform* (FFT). The result is a considerable simplification in the computational algorithm. In that work the authors explain two ways of implementing FDCT, namely curvelet transform via unequally spaced fast Fourier transform and curvelet transform via wrapping function. In the present work we have chosen the wrapping method because it is faster and reduces redundant information (Candes et al., 2006).

Wrapping method is based on the translation and wrapping of specially selected Fourier samples. It starts by taking a 2D-image as a cartesian array $f[m, n]$, where $0 \leq m < M$, $0 \leq n < N$ and $M \times N$ are the dimensions of the array. This image is scanned by wedge data and the output will be a subset of curvelet coefficients $c^D(j, \ell, k_1, k_2)$ indexed by a scale j , an orientation ℓ and spatial location parameters k_1 and k_2 (AlZubi, Islam, & Abbod, 2011). They consist of parallelograms containing a *digital curvelet waveform* $\varphi_{j\ell k_1 k_2}$ depicted as an oval. Several copies of it are represented in Fig. 2.

The wrapping consists of collecting information from the parallelograms into the rectangle in Fig. 2. The rectangle is centred at the origin, and its width and height are $2^{j/2}$ and 2^j , respectively.

The corresponding curvelet coefficient is then computed via

$$c^D(j, \ell, k_1, k_2) = \sum_{n=1}^{N_j} \sum_{m=1}^M f[m, n] \varphi_{j\ell k_1 k_2}^D[m, n]. \quad (1)$$

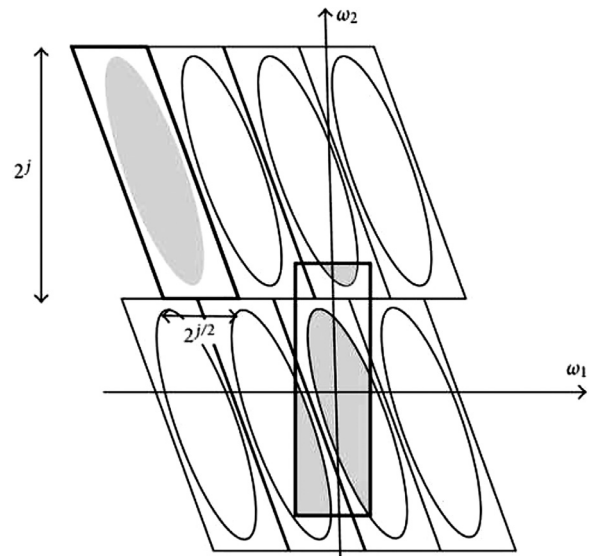


Fig. 2. Wrapping an image by wedge data (AlZubi et al., 2011, Fig. 10).

The better $f[m, n]$ fits $\varphi_{j\ell k_1 k_2}$, the higher is $c^D(j, \ell, k_1, k_2)$. In fact, j and ℓ can be viewed in the frequency domain as a rectangular tiling, as depicted in Fig. 3. It can be seen that curvelets become thinner and smaller in the spatial domain, where 2^j and $2^{j/2}$ are reverted to 2^{-j} and $2^{-j/2}$. They are optimal at capturing curves in an image, and one just need a few coefficients to identify curved singularities.

Essentially, by scanning the image, curves in it are represented by gaps along neighbouring grey levels. The curvelets are numerical data that translate these curves into coefficients. The sharper the gaps, the higher the coefficients. The reader can refer to AlZubi et al. (2011); Candes et al. (2006); Candes and Donoho (2000) for more detailed information.

In this study we used FDCT via wrapping method to extract textural features of each breast tissue image. The curvelet coefficients $c^D(j, \ell, k_1, k_2)$ at different scales and orientations were derived by Eq. (1). We have used orientation ℓ of 16 degrees and scale of 4 based on Eltoukhy, Faye, and Samir (2010b). In the experiments the parameters ℓ and j were 16 degrees and 4 levels, respectively. Based on the subband division, $50 = 1 + 16 + 32 + 1$ subbands of curvelet coefficients were obtained.

2.2.2. Local binary patterns

In 1994 Ojala introduced the technique of local binary patterns (LBP). It was later resumed by Pietikäinen (Ojala, Pietikäinen, & Maenpää, 2002; Pietikäinen, Zhao, Hadid, & Ahonen, 2011). By scanning the image, among others we get two pieces of information: the *spatial structure* and the *contrast*. The LBP then uses both in order to derive texture.

After the decomposition of each breast tissue image into a series of curvelet subbands, the LBP codes of these subbands were calculated by using LBP operator. Then, given a central pixel c as threshold value, the LBP operator is computed by comparing its value with those of its neighbours p according to the following equation:

$$LBP_{p,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, \quad (2)$$

where $s(x)$ is 1 for $x \geq 0$ and 0 for $x < 0$, g_c is the intensity of the central pixel, g_p is the one of each neighbouring pixel p , P is the number of neighbours and R is the radius of the neighbourhood.

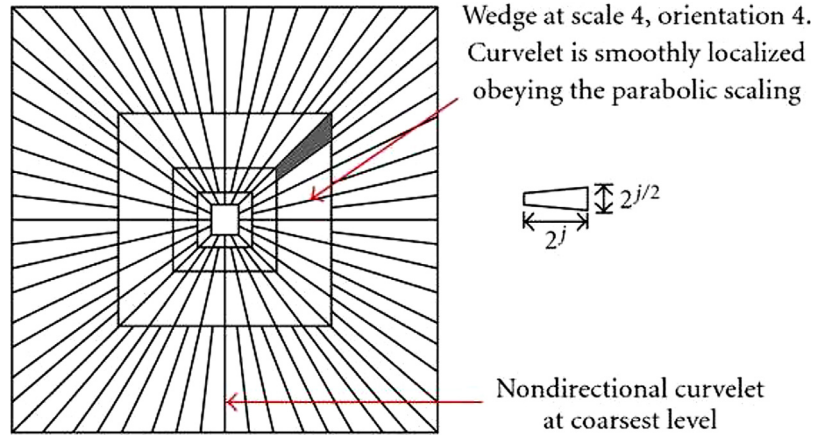
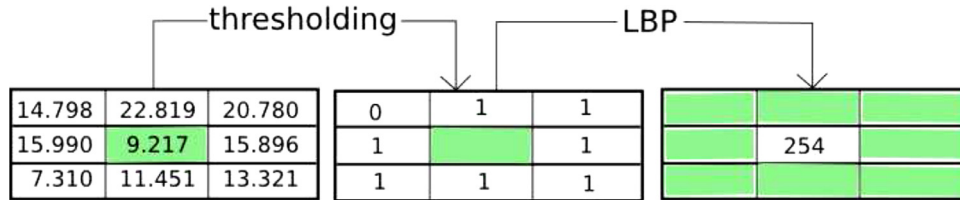


Fig. 3. Digital image tiling in the frequency domain (AlZubi et al., 2011, Fig. 9).



$$\mathbf{LBP} = 0 + 2 + 4 + 8 + 16 + 32 + 64 + 128 = 254$$

Fig. 4. LBP operator performed onto a 3 × 3 neighbourhood.

Fig. 4 shows an example of circularly symmetric neighbour sets used to compute the basic LBP operator for parameters $P = 8$ and $R = 1$.

In general, for an $N \times M$ input matrix, after LBP converts it into a logical matrix (0 and 1 entries), the weights 2^p of each pixel (i, j) are represented by an histogram vector \mathbf{h} of length K , whose bins are given by:

$$H(k) = \sum_{i=1}^N \sum_{j=1}^M f(LBP_{P,R}(i, j), k), \quad k \in [0, K] \quad (3)$$

where K is the maximal LBP pattern and $f(x, y)$ coincides with the threshold σ given by:

$$\sigma(x) = \begin{cases} 1 & \text{if } x = y; \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

In Ojala et al. (2002) the authors introduce an improvement to the non-uniform LBP, namely

$$LBP_{P,R}^u = \begin{cases} \sum_{p=1}^{P-1} s & \text{if } U(LBP_{P,R}) \leq 2; \\ P + 1 & \text{otherwise,} \end{cases} \quad (5)$$

where

$$U(LBP_{P,R}) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|. \quad (6)$$

The application of LBP operators on curvelet coefficients is similar to LBP operation on grey scale images. The only difference is that the LBP kernel is moving along all the curvelet coefficients in each subband. In the decomposition stage, the subbands contain both positive and negative coefficients. Therefore, we used the absolute value of the coefficients to compute LBP, which contain the histograms \mathbf{h} for each subband. We chose the parameters $P = 8$ and $R = 2$ by following (Nanni, Brahmam, & Lumini, 2011).

2.3. Reduction of attributes

After the computation of the texture histograms for each subband, we have applied analysis of variance (ANOVA) in order to select the relevant features. By applying ANOVA to our large set of descriptors they are grouped into statistically similar subgroups, except for isolated descriptors discarded as noise (Surendiran & Vadivel, 2010).

In the present work, ANOVA technique compared breast tissue images that showed either lesion or normal states. The texture histograms were evaluated for each attribute and when two attributes were statistically similar, both were discarded. Moreover, the confidence interval was taken to be within 5% in our case (Jacomini et al., 2013). The following hypotheses were considered:

- The descriptors are all random and statistically independent variables;
- They are all obtained from the breast tissues described in Section 2.1;
- Each subgroup consists of breast tissues that have the same variance.

2.4. Performance of the proposed algorithms

In this paper we study each dataset by subdividing it into two groups. The first group is composed of tissues stated as normal and malignant, and the second group is formed by benign and malignant breast tissues. The purpose of the classifier is to identify in which class each tissue is likely to be, considering the two groups separately. Four classification algorithms, PL, DT, RaF and SVM, were used to distinguish the two classes for each group. They are briefly described in this subsection.

2.4.1. Polynomial classifier

For the classification of different groups of breast tissue images, the feature vectors were passed as inputs defined by

$\mathbf{x} = [x_1 \dots x_d]^T$, and they were expanded in terms of the vector-based polynomial $\mathbf{p}_n(\mathbf{x})$. This process make it possible to map a d -dimensional feature vector into an L -dimensional vector. For example, given a two-dimensional input vector $\mathbf{x} = [x_1 \ x_2]^T$, the elements of $\mathbf{p}_2(\mathbf{x})$ result in parameters similar to those shown in Campbell, Assaleh, and Broun (2002):

$$\mathbf{p}_2(\mathbf{x}) = [1 \ x_1 \ x_2 \ x_1^2 \ x_1x_2 \ x_2^2]^T. \quad (7)$$

Therefore, the dimensionality of the expanded vector $\mathbf{p}_n(\mathbf{x})$ can be expressed in terms of the polynomial order and the dimensionality of the input vector \mathbf{x} . Finally, the classifier output y_i was obtained after a linear combination of the expanded vector $\mathbf{p}_n(\mathbf{x})$ as

$$y_i = \mathbf{a}_i^T \mathbf{p}_n(\mathbf{x}), \quad (8)$$

where \mathbf{a}_i is a coefficient vector of the polynomial discriminant function for the class ω_i . The model was divided into two stages, namely training and testing, which are detailed in Nascimento et al. (2013).

2.4.2. Decision tree

Decision tree (DT) is a statistical method composed of nodes connected by edges. This structure is established from the information gain, as for instance the difference in entropy: the descriptor with the highest information gain is the splitting criterion. Therefore, a node indicates that samples were divided into enriched subsets. The splitting process is tested on each subset in a recursive manner, resulting in a hierarchical structure. For a DT linearised into decision rules, the rules are defined by the conditions along the path and the outcome is the contents of the leaf node (Chau, Li, & Yu, 2013; Freund & Mason, 1999; Quinlan, 1993; Safavian & Landgrebe, 1991; Snousy, El-Deeb, Badran, & Khilil, 2011).

2.4.3. Random forest

The random forest (RaF) is an algorithm that combines many decision trees to make predictions. It is a concept of regression trees induced by bootstrap samples of a training data set, with random features selected in the induction tree process (Pang-Ning, Steinbach, & Kumar, 2006)[PP.290–294] and (Nisbet, Elder IV, & Miner, 2009)[PP.248–250]. The proposed approach produces a large set of features that we call ν . In its turn, ν is submitted to the RaF algorithm, which starts by taking a random training set $I \subset \nu$.

After getting I , RaF makes each tree grow randomly from different trunks $I' \subset I$. Each node of the tree is labelled by a number n , and the corresponding $I_n \subset I$ is split into left and right subsets I_ℓ and I_r , respectively. Notice that $I_1 = I'$. The splitting is determined by a threshold t and a function $f(v_i)$, $v_i \in \nu$ and $i \in I_n$. Once these decision trees are obtained, RaF combines them in order to construct a new one called I^* .

2.4.4. Support vector machines

In the 1990s, Vapnik introduced a technique called support vector machine (SVM) (Vapnik, 1999). Its purpose was to minimise and balance two types of error: training set and test set. To be more specific, consider a training set

$$\{(z_i, y_i) : z_i \in \mathbb{R}^{n_f}, y_i \in \{-1, +1\}\}_{i=1}^{n_s}, \quad (9)$$

where n_f is the number of characteristics and n_s is the number of samples. Hence, in order to establish the two classes normal and abnormal (malignant or benign) we can use the dual form to find the Lagrange multipliers $\{\alpha_i\}_{i=1}^{n_s}$ that minimise the objective function

$$Q(\alpha) = \frac{1}{2} - \sum_{i=1}^{n_s} \alpha_i + \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} \alpha_i \alpha_j y_i y_j K(z_i, z_j) \quad (10)$$

constrained to

$$\sum_{i=1}^{n_s} \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C, \quad \forall i = 1, 2, \dots, n_s, \quad (11)$$

where C is a positive constant given by the user, and $K(\cdot, \cdot)$ is a partially defined positive kernel function.

Let $\{\alpha_i\}_{i=1}^{n_s}$ be an optimal solution of (10) and (11). The decision function is given by

$$\beta_{SVM}(z) = b + \sum_{i=1}^{n_s} \alpha_i y_i K(z_i, z), \quad (12)$$

where b is some bias value. Here, we have used the polynomial kernel function given by

$$K(x_i, x_j) = (x_i x_j + 1)^2. \quad (13)$$

2.4.5. Cross-validation

We trained and tested the proposed approach using an n -fold cross-validation method. The samples are distributed into n subsets. Both the distribution and the number of samples in each subset are as uniform as possible. Afterwards, we take one of the n subsets as test and the remaining $n - 1$ partitions as the training set. The accuracy of the model is then computed for this $1 \times (n - 1)$ separation.

This procedure is repeated for each one of the n subsets, and so we end up with n values of accuracy. The final accuracy is computed as an average of that n values. A final accuracy close to zero means that data generation is well-represented by the model. Fig. 5 shows how the sets were distributed in order to apply the cross-validation method to the classifiers.

2.5. Evaluation of performance

The feature vectors were combined and compared in order to evaluate the proposed approach, where the metrics accuracy (AC) and area under the ROC curve (AUC) were considered (Dua, Singh, & Thompson, 2009; Fawcett, 2006). All tests were evaluated with these metrics because they are the most significant measures to evaluate the classification algorithms (Huang & Ling, 2005).

The AC measure is defined as the proportion of correct predictions related to the total number of evaluated cases:

$$AC = \frac{TP + TN}{P + N}, \quad (14)$$

where TP is the rate of true positive cases, TN is the rate of true negative cases, both calculated in relation to all positive (P) and all negatives (N) cases.

The AUC metric relates the success rate to the failure rate of groups (Beck & Shultz, 1986). Both rates range from zero to one. The ROC curve is plotted in the xOy -plane, in which the sensitivity varies along Oy , whereas the specificity varies along Ox . One takes the AUC as an evaluation of the classifier. A classification test is called *perfect* when its AUC equals 1.0, and it is called *perfectly wrong* when its AUC equals 0.0.

The algorithms were implemented in MATLAB R2012b with the help of WEKA platform 3.6.6 to classify features. We have used MATLAB packages *curvelet* (Demanet, 2007) to compute curvelet coefficients and the function *anova1* to select features with ANOVA technique.

3. Results

The quantification and distinction of the several images of breast tissues were performed by applying the techniques described previously. Examples of breast tissues are shown in Figs. 6, 7 and 8. The next Section 3.1 and 3.2 explain our procedures in a

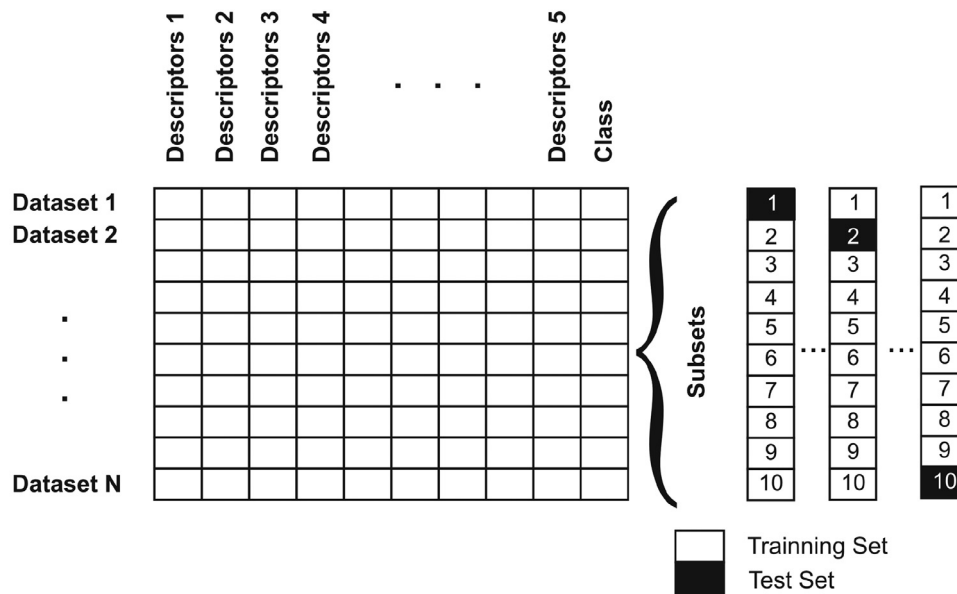


Fig. 5. Relation between training sets and test sets.

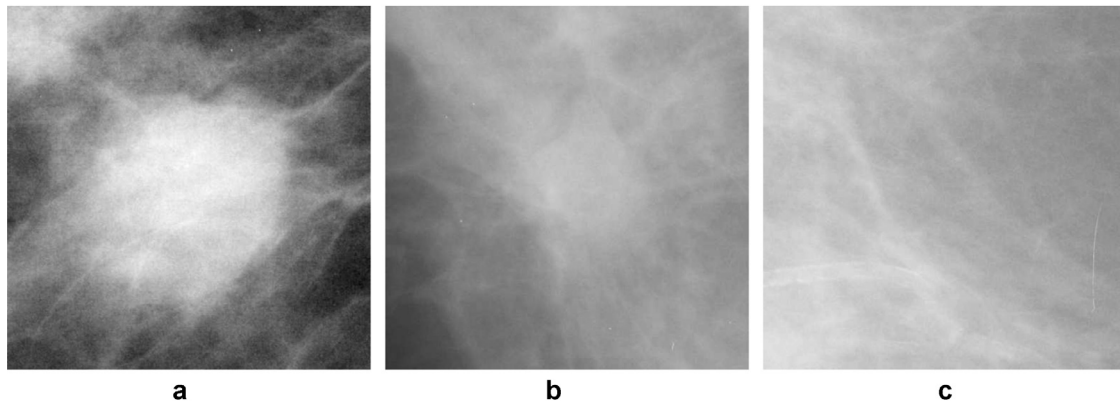


Fig. 6. ROIs taken from DDSM, (a) C-0009-1-Right (malignant tissue); (b) B-3128-1-Right (benign tissue); (c) B-3013-1-Left (normal tissue).

way that will later make it easier to understand and discuss our final results. Now we are going to consider the steps of Extraction of Descriptors and Reduction of Attributes, both already described in Sections 2.2 and 2.3, respectively.

3.1. Descriptors

In Section 2.2.1 we applied the curvelet transform to images of breast cancer tissues. These images belong to the dataset described in Section 2.1. The curvelet transform gave a total of 740,897 coefficients for each image composed of 512×512 pixels. For each image that consisted of 896×768 pixels we got far more coefficients, namely 1,111,113. Some examples of coefficients obtained with the Curvelet Transform are shown in Fig. 9. The RGB model was used on images of Fig. 9 to show the difference between values of the coefficients in each scale. We randomly chose a subband of the scale 2 and of the scale 3. Afterwards, the LBP technique explained in Section 2.2.2 was applied and the number of coefficients for each breast cancer image of the dataset was reduced to 6,656.

3.2. Reduction and classification

The attribute reduction technique ANOVA explained in Section 2.3 was then applied to that 6656 coefficients of each image. By means of ANOVA, the samples obtained via our data

mining technique proved it to be successful. The total values are in accordance with existing data in nine-folds of the training step. For instance, by considering images in size 512×512 pixels the ANOVA technique gave an average of 961 features for the normal-malignant group. This average was 1889 for the benign-malignant group of nine-folds used in the training step with DDSM.

But for same size images in the datasets BCDR-FMR and BCDR-DMR, the results of the normal-malignant group had averages 626 and 376, respectively. For the benign-malignant group the averages were 331 and 444, respectively. Fig. 10 summarises these results, which consider the amount of curvelet coefficients, LBP operators and finally their reduction by ANOVA in accordance with the nine-folds of the training step.

After the attribute reduction, we applied the PL, DT, RaF and SVM classifiers mentioned in Section 2.4. The results are shown in Tables 1, 2, 3 and 4, which consider the datasets DDSM, BCDR-FMR, BCDR-DMR and UCSB-BB, respectively. Discriminative performances were evaluated by the metrics AC and AUC.

From Tables 1–4 the association of curvelet transform, LBP and ANOVA with a PL classifier achieved higher AUC and AC values for all cases compared with other tested combinations, namely the ones in which algorithms DT, RaF and SVM were used. The obtained rates are between 91% and 100%.

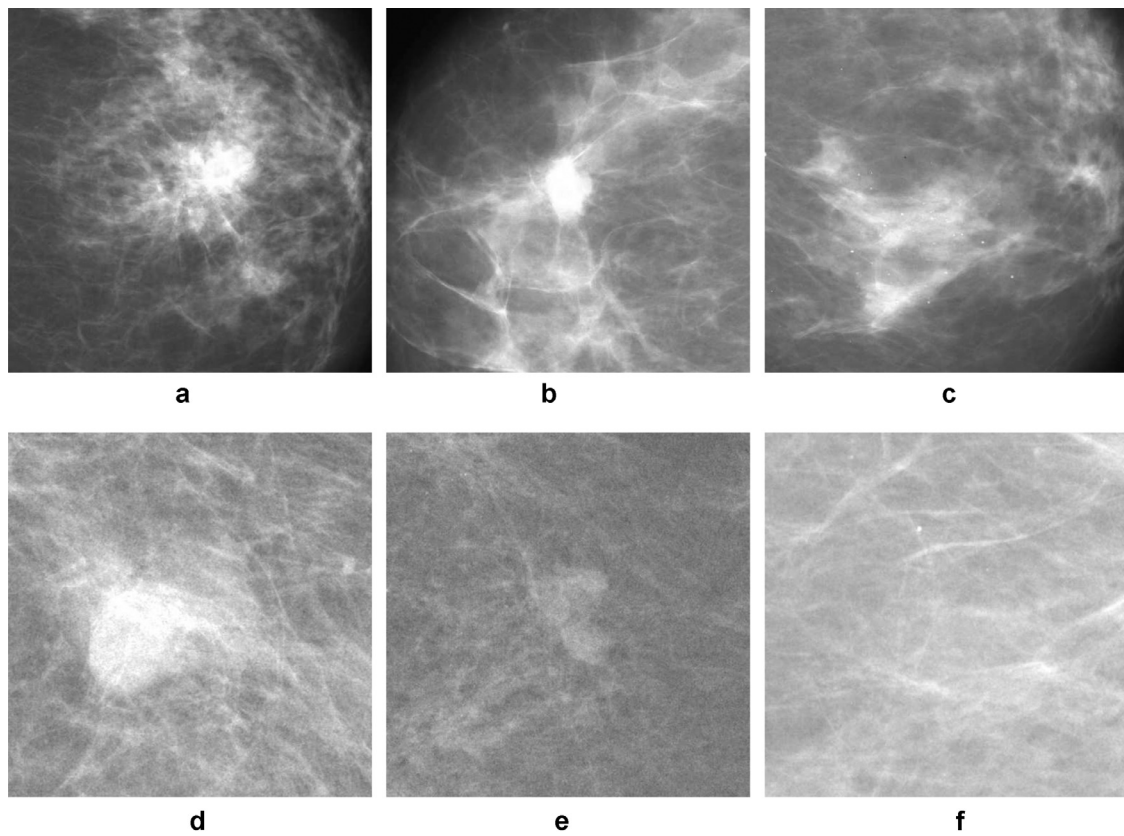


Fig. 7. ROIs taken from images of the BCDR datasets, BCDR-FMR: (a) img-21-22-1-Left (malignant tissue); (b) img-269-279-1-Right (benign tissue); (c) img-13-14-2-Left (normal tissue). BCDR-DMR: (d) img-3-4-1-Left (malignant tissue); (e) B-3128-1-Right (benign tissue); (f) B-3013-1-Left (normal tissue).

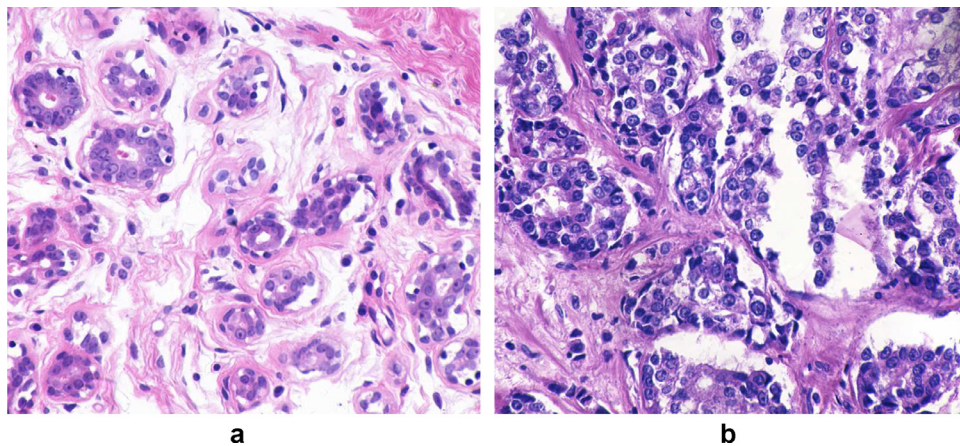


Fig. 8. Images taken from UCSB-BB, (a) ytma23-022103-malignant3-ccd; (b) ytma10-010704-benign1-ccd.

Table 1

Performance of the PL, DT, RaF and SVM algorithms for tissues of the DDSM dataset.

Classification Algorithm	Normal-malignant		Benign-malignant	
	AC	AUC	AC	AUC
PL	91	0.91	94	0.94
DT	64	0.67	72	0.72
RaF	66	0.75	79	0.89
SVM	63	0.63	85	0.85

Table 2

Results obtained with classification algorithms for breast tissue images of the BCDR-FMR dataset.

Classification Algorithm	Normal-malignant		Benign-malignant	
	AC	AUC	AC	AUC
PL	100	1.00	100	1.00
DT	45	0.45	52	0.52
RaF	53	0.53	45	0.47
SVM	80	0.80	70	0.70

When considering mammographic images the proposed approach achieved the best value $AUC = 1$ for the datasets BCDR-FMR and BCDR-DMR. Regarding lower AUC values DDSM presented

the worst performance for the proposed association. However, one should notice how the proposed approach improves performance. When the Polynomial classifier was applied to the descriptors, in

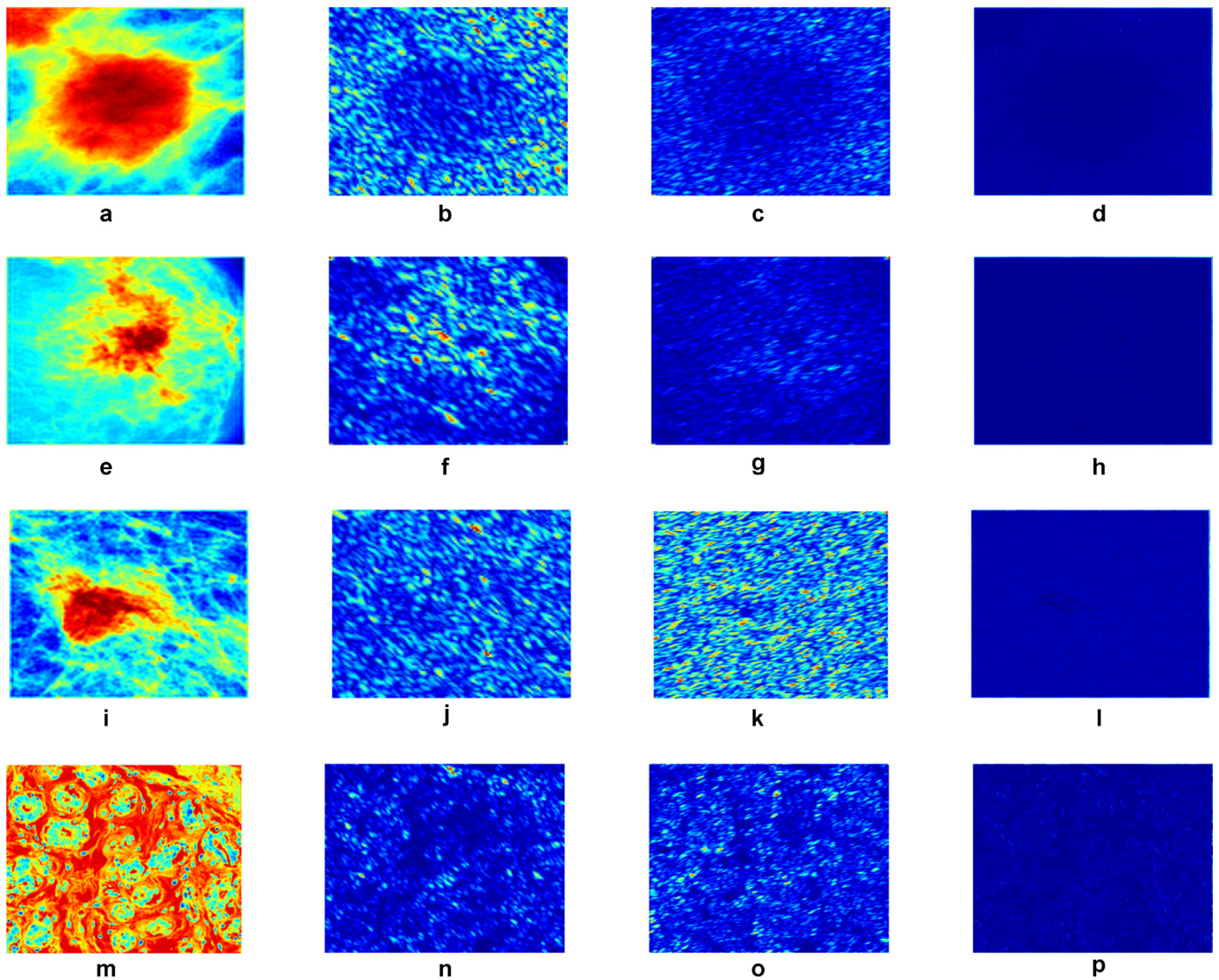


Fig. 9. Examples of curvelet coefficients of the four scales resulting in the decomposition process of the image. Case C-0009-1-Right: (a) scale 1; (b) scale 2; (c) scale 3 and (d) scale 4. Case img-21-22-1-Left: (e) scale 1; (f) scale 2; (g) scale 3 and (h) scale 4. Case img-3-4-1-Left: (i) scale 1; (j) scale 2; (k) scale 3 and (l) scale 4. Case ytna23-022103-malignant3-ccd: (m) scale 1; (n) scale 2; (o) scale 3 and (p) scale 4.

Table 3

Results obtained with PL, DT, RaF and SVM algorithms for tissues of the BCDR-DMR dataset.

Classification Algorithm	Normal-malignant		Benign-malignant	
	AC	AUC	AC	AUC
PL	100	1.00	100	1.00
DT	55	0.59	63	0.62
RaF	46	0.47	43	0.46
SVM	61	0.61	53	0.53

the normal-malignant group of DDSM we achieved $AUC = 0.91$. By applying DT, RaF and SVM the AUC values were 0.72, 0.89 and 0.85, respectively.

For DDSM the proposed approach, which uses the Polynomial classifier, led to an improvement greater than 24%, 16%, 28% compared with the other approaches that use DT, RaF and SVM, respectively. In these tests the best performance $AUC = 1$ was achieved with the Folds 4 (benign-malignant) and 5 (normal-malignant). These results can indicate a presence of noise or artefacts influ-

encing the other Folds. Regarding UCSB-BB in Table 4, which is the only dataset with colour images, although they were also analysed with separate RGB colour channels and grey scale, it again achieved the best result $AUC = 1$. This indicates that colour channels are irrelevant for classification. Comparing the performance in this dataset, the association among curvelet transform, LBP, ANOVA and polynomial classifier was also the best one. Considering the results with DT, RaF and SVM classifiers, the association with the polynomial method presented differences greater than 12% with respect to RaF with the Red channel, and the peak of 40% with respect to DT with the Blue channel.

Several methods have been developed to study breast tissues. Among others we cite (Dhahbi et al., 2015; Eltoukhy, Faye, & Belhaouari Samir, 2010a). But none of them has used our proposed combination. However, an illustrative overview is now propitious to show the good quality of our method. Table 5 lists several values of AC and AUC in the literature, including those presented in this work. One sees that the proposed combination leads to AUC values similar to the best rates found by other authors. Of course, each author chooses a strategic methodology that looks promising, although its consistency will only be evidenced by the final

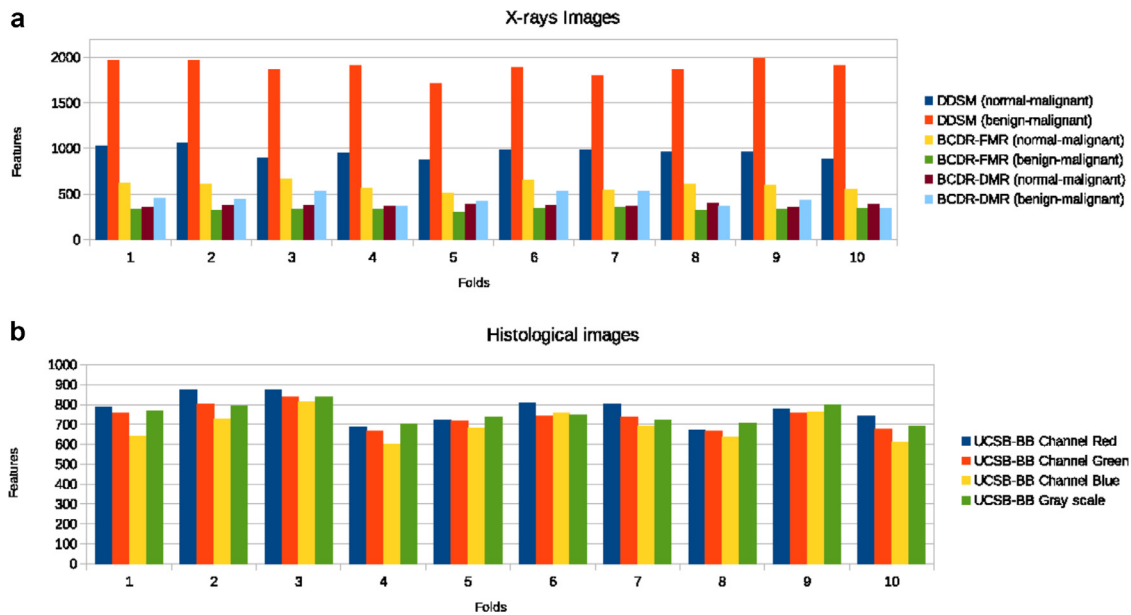


Fig. 10. Number of features after the application of the proposed method for each fold used during the training step, (a) X-rays images; (b) histological images.

Table 4
Evaluation of the benign and malignant tissues of the dataset UCSB-BB using the classification algorithms.

Classification Algorithm	Channel red		Channel green		Channel blue		Grey scale	
	AC	AUC	AC	AUC	AC	AUC	AC	AUC
PL	100	1.00	100	1.00	100	1.00	100	1.00
DT	67	0.70	65	0.65	60	0.60	67	0.68
RaF	67	0.88	82	0.87	75	0.85	72	0.72
SVM	72	0.72	750	0.75	82	0.82	72	0.72

Table 5
Comparison of AUC values from different approaches for benign-malignant group.

Reference	Dataset	ROIs	Feature extraction	Classifier	AC	AUC
Eltoukhy, Faye, and Samir (2010b)			Curvelet transform and Haralick features	Euclidean distance	91.68	
Eltoukhy, Faye, and Belhaouari Samir (2010a)	MIAS	322	Greatest coefficients of curvelet transform		98.59	–
Eltoukhy, Faye, and Belhaouari Samir (2012)			Features in curvelet coefficients	SVM	97.30	
Tai, Chen, and Tsai (2014)	DDSM	358	Co-occurrence matrix and Haralick features	LDA	–	0.98
Görgel, Sertbas, and Uan (2015)	MIAS	60	Spherical wavelet transform		83.3	–
Dhahbi et al. (2015)		252	Curvelet transform and moment theory	k-NN	91.27	–
Our method	DDSM	240	Curvelet transform, LBP and ANOVA	PL	94.00	0.94
	BCDR-FMR	106				
	BCDR-DMR	95				
	UCSB-BB	58			100.00	1.00

results. Most of the methods in Table 5 lead to an “almost” ideal system. However, we could not compare them in order to elect the best one, for this would be a difficult task, not to say senseless. In fact, the different methodologies are rather complementary than ratable. Therefore, we believe that our method is reliable for applications in any other similar image processing. Moreover, we contribute to future works, in which one could profit from valuable information contained in each of these references.

4. Conclusions

We have presented a new method based on the association among curvelet transform, LBP, feature selection by statistical analysis and distinct classification techniques, in order to support the development of CAD systems for cancer breast tissues. The studied groups were normal-malignant and benign-malignant from the different datasets: DDSM, BCDR-FMR, BCDR-DMR and UCSB-BB. The

best association was curvelet transform, LBP and ANOVA with a PL classifier, when the AUC and AC values are compared with the others achieved by the algorithms DT, RaF and SVM. These algorithms have been largely used in the literature, as explained before. The obtained rates lie between 91% and 100%, which are very competitive results when compared to the described in state-of-the-art approaches. Another advantage is in tested results on images with different physical properties commonly explored in clinical practice. Of course, even if our proposed method outperforms the others, once again we emphasize that it is senseless to rate our results with the others, because the different methodologies are in fact complementary. However, to the best of our knowledge, the proposed methods in our work are *new*. Also, another innovation was the exclusion of similar features by ANOVA, which provided relevant rates. For instance, by considering images in size 512×512 pixels, the reduction rate was approximately 85% for the normal-malignant group and 71% for the benign-malignant group of nine-folds used in the training step with the DDSM dataset. For the BCDR-FMR and BCDR-DMR datasets the rates were 90% and 94%, respectively. These values endorse the proposed association as an improvement of the classification step towards exactness with well known techniques of artificial intelligence and the chosen best combination.

According to the analyses presented previously, our contributions are: (a) The presented values with the proposed association are clinically significant, specially when one considers the difficulties in clinical practice to distinguish the two groups normal versus malignant and benign versus malignant. These groups were well represented by the images used in this work, namely from the datasets DDSM, BCDR-FMR, BCDR-DMR and UCSB-BB; (b) The proposed association (curvelet transform, LBP, feature selection by ANOVA and PL classifier) is useful as an automated protocol for the diagnosis of breast tissues: a reference to pathologists and specialists; (c) Considering the fact that cancer cases have been increasing over the years (Ferlay et al., 2012) and have been causing a considerable loss in the patients' quality of life, the presented association may contribute to the correct diagnosis of breast tissues (mammographic and histopathological images) and to the efficient development of CAD systems.

As a variation of this present method, in future works we intend to investigate local binary pattern operators (CLBP, LTP, etc.) applied to different levels of subbands and sub-images obtained via the curvelet transform. Of course, this study can be complemented and improved with the following strategies: (1) apply other statistical methods in the feature selection to understand their influence on the classification step; (2) identify the most relevant features for distinction among groups with the polynomial classifier; (3) test other classification methods; (4) evaluate our proposed method applied to many more datasets other than clinical images. In Computer Vision there are many problems that can be handled with our method. We intend to assess its generality and its quality when applied to these problems.

Acknowledgements

We are grateful to the following persons for having conceded the use of image datasets: Kopans and Moore of Massachusetts General Hospital, Bowyer of the University of South Florida and Kegelmeyer of Sandia National Laboratories for providing the DDSM dataset; López and the Breast Research Group of the University of Aveiro for providing the BCDR datasets; Manjunath and the Bio-image Informatics Research Group for images from the UCSB-BB dataset.

References

- Abdel-Nasser, M., Rashwan, H. A., Puig, D., & Moreno, A. (2015). Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern. *Expert Systems With Applications*, 42(24), 9499–9511.
- Al-Hammadi, M. H., Muhammad, G., Hussain, M., & Bebis, G. (2013). Curvelet transform and local texture based image forgery detection. In *Advances in visual computing* (pp. 503–512). Springer.
- AlZubi, S., Islam, N., & Abbod, M. (2011). Multiresolution analysis using wavelet, ridgelet, and curvelet transforms for medical image segmentation. *Journal of Biomedical Imaging*, 2011, 4.
- Beck, J. R., & Shultz, E. K. (1986). The use of relative operating characteristic (ROC) curves in test performance evaluation. *Archives of Pathology & Laboratory Medicine*, 110(1), 13–20.
- Campbell, W. M., Assaleh, K. T., & Broun, C. C. (2002). Speaker recognition with polynomial classifiers. *IEEE Transactions on Speech and Audio Processing*, 10(4), 205–212.
- Candes, E., Demanet, L., Donoho, D., & Ying, L. (2006). Fast discrete curvelet transforms. *Multiscale Modeling & Simulation*, 5(3), 861–899.
- Candes, E. J., & Donoho, D. L. (2000). Curvelets: A surprisingly effective nonadaptive representation for objects with edges. *Technical Report*. DTIC Document.
- Chau, A. L., Li, X., & Yu, W. (2013). Support vector machine classification for large datasets using decision tree and fisher linear discriminant. *Future Generation Computer Systems*. doi:10.1016/j.future.2013.06.021.
- Choi, J. Y., Kim, D. H., & Ro, Y. M. (2012). Combining multiresolution local binary pattern texture analysis and variable selection strategy applied to computer-aided detection of breast masses on mammograms. In *Proceedings of the IEEE-EMBS international conference on biomedical and health informatics (BHI)* (pp. 495–498). IEEE.
- Cireşan, D. C., Giusti, A., Gambardella, L. M., & Schmidhuber, J. (2013). Mitosis detection in breast cancer histology images with deep neural networks. In *Proceedings of the medical image computing and computer-assisted intervention—MICCAI* (pp. 411–418). Springer.
- Demanet, L. (2007). The curvelet organization. <http://www.curvelet.org/software.html>. Accessed 07.02.16.
- Dhabhi, S., Barhoumi, W., & Zagrouba, E. (2015). Breast cancer diagnosis in digitized mammograms using curvelet moments. *Computers in Biology and Medicine*, 64, 79–90. doi:10.1016/j.compbiomed.2015.06.012.
- Dheeba, J., & Tamil Selvi, S. (2012). An improved decision support system for detection of lesions in mammograms using differential evolution optimized wavelet neural network. *Journal of Medical Systems*, 36(5), 3223–3232. Accessed 07.02.16.
- Dua, S., Singh, H., & Thompson, H. W. (2009). Associative classification of mammograms using weighted rules. *Expert Systems with Applications*, 36(5), 9250–9259.
- Elter, M., & Horsch, A. (2009). Cadx of mammographic masses and clustered microcalcifications: A review. *Medical Physics*, 36(6), 2052–2068.
- Eltoukhy, M. M., Faye, I., & Belhaouari Samir, B. (2010a). A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram. *Computers in Biology and Medicine*, 40(4), 384–391.
- Eltoukhy, M. M., Faye, I., & Belhaouari Samir, B. (2012). A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation. *Computers in Biology and Medicine*, 42(1), 123–128. Accessed 07.02.16.
- Eltoukhy, M. M., Faye, I., & Samir, B. B. (2010b). Curvelet based feature extraction method for breast cancer diagnosis in digital mammogram. In *Proceedings of the international conference on intelligent and advanced systems (ICIAS)* (pp. 1–5). IEEE.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. doi:10.1016/j.patrec.2005.10.010.
- Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., et al. (2012). GLOBOCAN 2012 v1.0, cancer incidence and mortality worldwide: IARC Cancer-Base No. 11. [01/01/2015]. *Technical Report*. Lyon, France: International Agency for Research on Cancer.
- Freund, Y., & Mason, L. (1999). The alternating decision tree learning algorithm. In *Proceedings of the ICML: Vol. 99* (pp. 124–133).
- Gelasca, E. D., Obara, B., Fedorov, D., Kvilekval, K., & Manjunath, B. (2009). A bio-segmentation benchmark for evaluation of bioimage analysis methods. *BMC Bioinformatics*, 10(1), 368.
- Gonzalez, R. C., & Richard, E. (2002). *Digital image processing*. Prentice Hall Press. ISBN 0-201-18075-8.
- Görgel, P., Sertbas, A., & Uan, O. N. (2015). Computer-aided classification of breast masses in mammogram images based on spherical wavelet transform and support vector machines. *Expert Systems*, 32(1), 155–164. doi:10.1111/exsy.12073.
- Guo, Y., Zhao, G., & Pietikäinen, M. (2014). Local configuration features and discriminative learnt features for texture description. In *Local binary patterns: New variants and applications* (pp. 113–129). Springer.
- Gupta, S., & Markey, M. K. (2005). Correspondence in texture features between two mammographic views. *Medical Physics*, 32(6), 1598–1606.
- Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), 786–804. doi:10.1109/proc.1979.11328.
- Heath, M., Bowyer, K., Kopans, D., Moore, R., & Kegelmeyer, P. (2001). The digital database for screening mammography. In *Proceedings of the fifth international workshop on digital mammography* (pp. 212–218). Medical Physics Publishing. <http://marathon.csee.usf.edu/Mammography/Database.html>.

- Huang, J., & Ling, C. X. (2005). Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 17(3), 299–310. doi:10.1109/TKDE.2005.50.
- Hussain, M. (2013). False-positive reduction in mammography using multiscale spatial weber law descriptor and support vector machines. *Neural Computing and Applications*, 1–11. doi:10.1007/s00521-013-1450-7.
- INCA (2014). Estimativa 2014: Indicência de Câncer no Brasil. *Technical Report*. Rio de Janeiro, Brazil: Instituto Nacional de Câncer José Alencar Gomes da Silva. Coordenação de Prevenção e Vigilância.
- Jacomini, R., Nascimento, M., Dantas, R., & Ramos, R. (2013). Classification of mass in two views mammograms: Use of analysis of variance (ANOVA) for reduction of the features. *Recent Patents on Medical Imaging*, 3(1), 80–88. Accessed 07.02.16.
- Kong, S., Jiang, Z., & Yang, Q. (2015). Modeling neuron selectivity over simple mid-level features for image classification. *IEEE Transactions on Image Processing*, 24(8), 2404–2414.
- Liu, X., & Tang, J. (2014). Mass classification in mammograms using selected geometry and texture features, and a new SVM-based feature selection method. *IEEE Systems Journal*, 8(3), 910–920.
- Liu, X., & Zeng, Z. (2015). A new automatic mass detection method for breast cancer with false positive reduction. *Neurocomputing*, 152, 388–402.
- Llad, X., Oliver, A., Freixenet, J., Mart, R., & Mart, J. (2009). A textural approach for mass false positive reduction in mammography. *Computerized Medical Imaging and Graphics*, 33(6), 415–422. doi:10.1016/j.compmedimag.2009.03.007.
- López, M. A. G., Posada, N. G., Moura, D. C., Polln, R. R., Valiente, J. M. F., Ortega, C. S., et al. (2012). BCDR: A breast cancer digital repository. In *Proceedings of the ICMI15: 15th international conference on experimental mechanics* (pp. 22–27).
- Malhotra, G. K., Zhao, X., Band, H., & Band, V. (2010). Histological, molecular and functional subtypes of breast cancers. *Cancer Biology & Therapy*, 10(10), 955–960.
- Mayerhoefer, M., Perry, N., Milner, S., Mokbel, K., Duffy, S., & Pinker, K. (2010). Texture analysis applied to full-field digital mammography: ability to discriminate between invasive ductal and invasive lobular breast cancer - preliminary results. *Breast Cancer Research*, 12(Suppl 3), P12. doi:10.1186/bcr2665.
- Moschidis, E., Chen, X., Taylor, C., & Astley, S. (2014). Texture-based breast cancer prediction in full-field digital mammograms using the dual-tree complex wavelet transform and random forest classification. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8539 LNCS, 209–216. Accessed 07.02.16.
- Moura, D. C., & López, M. A. G. (2013). An evaluation of image descriptors combined with clinical data for breast cancer diagnosis. *International Journal of Computer Assisted Radiology and Surgery*, 8(4), 561–574.
- Nagaraja, S., Prabhakar, C., & Praveen Kumar, P. (2013). Complete local binary pattern for representation of facial expression based on curvelet transform. In *Proceedings of the International conference on multimedia processing, communication and information technology (MPCIT), published in ACEEE* (pp. 48–56).
- Nanni, L., Brahnam, S., Ghidoni, S., & Menegatti, E. (2014). A survey of region-based approaches and descriptors extracted from the co-occurrence matrix. *International Journal of Latest Research in Science and Technology*, 3(6), 192–200.
- Nanni, L., Brahnam, S., Ghidoni, S., Menegatti, E., & Barrier, T. (2013). Different approaches for extracting information from the co-occurrence matrix. *PLOS one*, 8(12).
- Nanni, L., Brahnam, S., & Lumini, A. (2011). A very high performing system to discriminate tissues in mammograms as benign and malignant. *Expert Systems With Applications*, 39(2), 1968–1971.
- Nanni, L., Lumini, A., & Brahnam, S. (2010). Local binary patterns variants as texture descriptors for medical image analysis. *Artificial Intelligence in Medicine*, 49(2), 117–125. doi:10.1016/j.artmed.2010.02.006.
- Nascimento, M., Martins, A., Neves, L., Ramos, R., Flores, E., & Carrijo, G. (2013). Classification of masses in mammographic image using wavelet domain features and polynomial classifier. *Expert Systems with Applications*, 40(15), 6213–6221. Accessed 07.02.16.
- de Nazar Silva, J., de Carvalho Filho, A., Corra Silva, A., Cardoso de Paiva, A., & Gattass, M. (2015). Automatic detection of masses in mammograms using quality threshold clustering, correlogram function, and SVM. *Journal of Digital Imaging*, 28(3), 323–337. doi:10.1007/s10278-014-9739-3.
- Neves, L. A., do Nascimento, M. Z., & de Godoy, M. F. (2014). Multiscale fractal descriptors to quantify behaviours of healthy and diseased tissues in mammographic images. *Cancer - Cares, Treatments and Preventions. 1ed.: iConcept Press Ltd*, 1, 234–254.
- Nisbet, R., Elder IV, J., & Miner, G. (2009). *Handbook of statistical analysis and data mining applications*. Academic Press.
- Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987.
- Oliver, A., Freixenet, J., Mart, J., Prez, E., Pont, J., Denton, E. R., et al. (2010). A review of automatic mass detection and segmentation in mammographic images. *Medical Image Analysis*, 14(2), 87–110. doi:10.1016/j.media.2009.12.005.
- Pang-Ning, T., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining. In Library of congress*.
- Pietikäinen, M., Zhao, G., Hadid, A., & Ahonen, T. (2011). *Computer Vision Using Local Binary Patterns. No. 40 in Computational Imaging and Vision*. Springer.
- Quinlan, J. R. (1993). *C4. 5: programs for machine learning: Vol. 1*. Morgan kaufmann.
- Ramos, R. P., Nascimento, M. Z., & Pereira, D. C. (2012). Texture extraction: An evaluation of ridgelet, wavelet and co-occurrence based methods applied to mammograms. *Expert Systems with Applications*, 39(12), 11036–11047.
- Rouhi, R., & Jafari, M. (2016). Classification of benign and malignant breast tumors based on hybrid level set segmentation. *Expert Systems with Applications*, 46, 45–59.
- Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Systems with Applications*, 42(3), 990–1002.
- Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man and Cybernetics*, 21(3), 660–674.
- Schwartz, W. R., Roberti de Siqueira, F., & Pedrini, H. (2012). Evaluation of feature descriptors for texture classification. *Journal of Electronic Imaging*, 21(2). doi:10.1117/1.JEI.21.2.023016.023016-1-023016-17.
- Snousy, M. B. A., El-Deeb, H. M., Badran, K., & Khilil, I. A. A. (2011). Suite of decision tree-based classification algorithms on cancer gene expression data. *Egyptian Informatics Journal*, 12(2), 73–82.
- Song, K.-C., Yan, Y.-H., Chen, W.-H., & Zhang, X. (2013). Research and perspective on local binary pattern. *Acta Automatica Sinica*, 39(6), 730–744.
- Souza Jacomini, R., Nascimento, M., Dantas, R. D., & Ramos, R. (2012). Comparison of PCA and ANOVA for information selection of CC and MLO views in classification of mammograms. In H. Yin, J. Costa, & G. Barreto (Eds.), *Intelligent data engineering and automated learning - ideal. In Lecture Notes in Computer Science: Vol. 7435* (pp. 117–126). Springer Berlin Heidelberg. doi:10.1007/978-3-642-32639-4_15.
- Suliga, M., Deklerck, R., & Nyssen, E. (2008). Markov random field-based clustering applied to the segmentation of masses in digital mammograms. *Computerized Medical Imaging and Graphics*, 32(6), 502–512. Accessed 07.02.16.
- Surendiran, B., & Vadivel, A. (2010). Feature selection using stepwise ANOVA, discriminant analysis for mammogram mass classification. *International J. of Recent Trends in Engineering and Technology*, 3(2), 55–57.
- Tai, S.-C., Chen, Z.-S., & Tsai, W.-T. (2014). An automatic mass detection system in mammograms based on complex texture features. *IEEE Journal of Biomedical and Health Informatics*, 18(2), 618–627. doi:10.1109/JBHI.2013.2279097.
- Tang, L., & He, C. (2013). Multiscale texture extraction with hierarchical (bv, g p, l 2) decomposition. *Journal of Mathematical Imaging and Vision*, 45(2), 148–163.
- Tashk, A., Helfroush, M. S., Danyali, H., & Akbarzadeh-jahromi, M. (2015). Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features. *Applied Mathematical Modelling*, 39(20), 6165–6182.
- Thurfjell, E. L., Lernevall, K. A., & Taube, A. (1994). Benefit of independent double readings in a population-based mammography screening program. *Radiology*, 191(1), 241–244.
- Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, 10(5), 988–999.
- Veta, M., van Diest, P. J., Willems, S. M., Wang, H., Madabhushi, A., Cruz-Roa, A., et al. (2015). Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Medical image analysis*, 20(1), 237–248.
- Veta, M., Pluim, J. P. W., van Diest, P. J., & Viergever, M. (2014). Breast cancer histopathology image analysis: A review. *IEEE Transactions on Biomedical Engineering*, 61(5), 1400–1411.
- Warren, R., & Duffy, W. (1995). Comparison of single reading with double reading of mammograms, and change in effectiveness with experience. *British journal of radiology*, 68(813), 958–962.
- Yu, S.-N., & Huang, Y.-K. (2010). Detection of microcalcifications in digital mammograms using combined model-based and statistical textural features. *Expert Systems with Applications*, 37(7), 5461–5469. Accessed 07.02.16.
- Zhang, B., & Pham, T. D. (2011). Phenotype recognition with combined features and random subspace classifier ensemble. *BMC Bioinformatics*, 12(1), 128.
- Zhou Li-Jian, & Liu Mo-Quan, S. J. (2012). Variable illumination face recognition based on curvelet and LBP. *Computer Engineering*, 38(22), 176. doi:10.3969/j.issn.1000-3428.2012.22.043.
- Zhu, S.-C., Guo, C.-E., Wang, Y., & Xu, Z. (2005). What are textons? *International Journal of Computer Vision*, 62(1–2), 121–143. doi:10.1023/b:visi.0000046592.70770.61.