

# A ranklet-based image representation for mass classification in digital mammograms

Matteo Masotti<sup>a)</sup>

Department of Physics, University of Bologna, Viale Bertini-Pichat 6/2, 40127, Bologna, Italy

(Received 16 December 2005; revised 29 June 2006; accepted for publication 15 August 2006; published 27 September 2006)

Regions of interest (ROIs) found on breast radiographic images are classified as either tumoral mass or normal tissue by means of a support vector machine classifier. Classification features are the coefficients resulting from the specific image representation used to encode each ROI. Pixel and wavelet image representations have already been discussed in one of our previous works. To investigate the possibility of improving classification performances, a novel nonparametric, orientation-selective, and multiresolution image representation is developed and evaluated, namely a ranklet image representation. A dataset consisting of 1000 ROIs representing biopsy-proven tumoral masses (either benign or malignant) and 5000 ROIs representing normal breast tissue is used. ROIs are extracted from the digital database for screening mammography collected by the University of South Florida. Classification performances are evaluated using the area  $A_z$  under the receiver operating characteristic curve. By achieving  $A_z$  values of  $0.978 \pm 0.003$  and 90% sensitivity with a false positive fraction value of 4.5%, experiments demonstrate classification results higher than those reached by the previous image representations. In particular, the improvement on the  $A_z$  value over that achieved by the wavelet image representation is statistically relevant with the two-tailed  $p$  value  $< 0.0001$ . Besides, owing to the tolerance that the ranklet image representation reveals to variations in the ROIs' gray-level intensity histogram, this approach discloses to be robust also when tested on radiographic images having gray-level intensity histogram remarkably different from that used for training. © 2006 American Association of Physicists in Medicine. [DOI: 10.1118/1.2351953]

Key words: ranklets, wavelets, support vector machine, pattern classification, computer-aided detection, mammography

## I. INTRODUCTION

Breast cancer is one of the most devastating causes of death among women from all over the world.<sup>1</sup> The detection of this disease at early stage is critical, since primary prevention is thus far impossible. To this purpose, screening mammography is undoubtedly the most effective tool, even though radiologists still miss up to 10%–30% of breast cancers.<sup>2</sup> In the last years, computer-aided detection (CAD) systems have been expressly introduced in order to help them in the interpretation of radiographic images.<sup>3–5</sup> By detecting suspicious regions independently from the human reader, those systems provide the radiologist with a second opinion that serves as a cue, yet leaving the final decision between follow-up or biopsy strictly to him. Recent studies illustrate that, when used in this way, CAD systems improve the radiologists' efficiency by as much as 5%–15%.<sup>6–8</sup>

Together with microcalcifications, masses are the most common lesions associated with the presence of breast cancer. In radiographic images, they appear as thickenings of breast tissue with size ranging from 3 to 30 mm.<sup>9</sup> In order to detect them, a large number of the computer-aided mass detection algorithms so far developed follow the scheme depicted in Fig. 1(a). Regions of interest (ROIs) are first encoded by means of a specific image representation, e.g., a pixel image representation, a wavelet image representation,

etc. Features are extracted from ROIs as they appear in the image representation domain, then submitted to the classifier. Typical features refer explicitly to shape, border, size, contrast, and texture of the regions under analysis. As an example, texture features in combination with linear discriminant analysis have been used by Wei *et al.*,<sup>10</sup> thus achieving 90% sensitivity with a false positive fraction (FPF) value of approximately 35% ( $A_z=0.86$ ). On the same database, texture features in combination with convolution neural networks yielded 90% sensitivity with an FPF value of 31% ( $A_z=0.87$ ), as described by Sahiner *et al.*<sup>11</sup> Features based on the geometry, intensity, gray levels, and gradients of ROIs have been investigated by Kupinski *et al.*,<sup>12</sup> in combination with neural networks ( $A_z=0.97$ ). Te Brake *et al.* designed a number of features aimed at capturing ROIs characteristics like intensity, isodensity, location, and contrast.<sup>13</sup> By measuring the suspiciousness of each ROI using an artificial neural network, 75% sensitivity has been achieved with 0.1 false positive per image.

Unfortunately, masses vary considerably in shape, border, size, contrast, and texture. It turns out that identifying a common set of features effective for every kind of mass is extremely difficult. In some recent works we proposed an alternative approach to mass detection which does not refer explicitly to the shape, border, size, contrast, or texture of ROIs.<sup>14–17</sup> As evident from Fig. 1(b), ROIs are simply en-

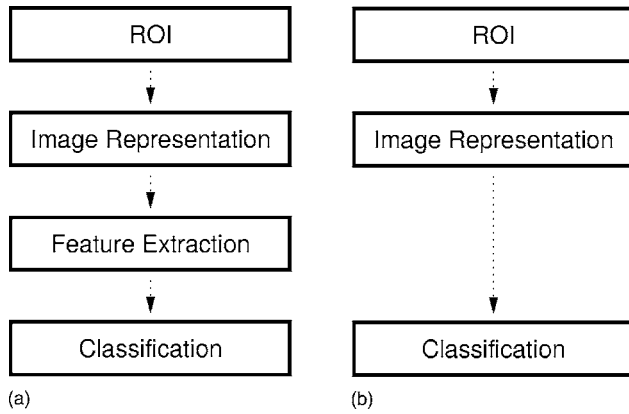


FIG. 1. Different approaches to computer-aided mass detection. The traditional approach based on (1) ROIs encoding by means of a specific image representation, (2) feature extraction from ROIs as they appear in the image representation domain, and (3) classification is shown in (a). The alternative approach adopted in this work and based on the direct classification of encoded ROIs is shown in (b).

coded by means of a specific image representation. Encoded ROIs are thus submitted to the classifier as they appear in the image representation domain. More specifically, they are directly classified by means of a support vector machine (SVM) classifier, an artificial intelligence technique based on Vapnik's statistical learning theory.<sup>18,19</sup> To investigate in detail the performances of this approach, several image representations have already been evaluated in one of our recent works,<sup>20</sup> namely a pixel image representation, a discrete wavelet image representation, and an overcomplete wavelet image representation.

With the purpose of improving classification performances, a novel image representation is proposed and evaluated in this paper. Being based on a recently developed transform known as ranklet transform,<sup>21</sup> in the following it will be referred to as ranklet image representation. Recently, Smeraldi *et al.* showed the feasibility of using the nonparametric, orientation-selective, and multiresolution properties of this transform for image classification tasks.<sup>22,23</sup> They used this image representation to encode the appearance of image frames in face identification problems. However, in order to emphasize the descriptive power of ranklet features, they addressed the classification task by using particularly simple classifiers, namely density estimation and hypothesis testing schemes. Conversely, for the study presented herein, we propose to combine the ranklet properties together with a more sophisticated classifier such as SVM, which, in turn, is much more suited for handling very high dimensional feature spaces and achieving better classification performances. Our main objective in the present study is hence to investigate this innovative approach thoroughly in mammography.

## II. MATERIALS AND METHODS

### A. Image databases

Evaluation and optimization of the ranklet image representation in combination with SVM is performed using ROIs extracted from the digital database for screening mammog-

raphy (DDSM) collected by the University of South Florida.<sup>24</sup> From the DDSM benign and malignant cases, a total of 1000 diagnosed masses are extracted using the provided ground truth annotations: 4% of them with subtlety rating 1, 12% with 2, 20% with 3, 24% with 4, and 40% with 5; here, rating 1 corresponds to subtle mass, whereas rating 5 to obvious mass. A square crop centered on the location of each annotated mass is selected. The size is chosen so that 70% of its area is occupied by the annotated mass and the remaining 30% by background. This choice is rather arbitrary, nevertheless, initial investigations demonstrated that this factor is not particularly relevant for the proposed system. Since SVM deals exclusively with dimensionally homogeneous vectors, all the crops containing a mass are bilinearly resized to an arbitrarily prefixed size of  $64 \times 64$  pixels. For our CAD system, this choice has already been demonstrated to offer a good tradeoff between the cropped image's detail preservation and computational complexity.<sup>15</sup> Similarly, for the nonmass class, a total of 5000 square crops are extracted from the DDSM normal cases. The whole dataset is therefore composed of 6000 regions with pixel size  $64 \times 64$ . To allow a direct comparison, the crops used in this paper are exactly the same as those used in our previous work<sup>20</sup> to evaluate different image representations.

Once the ranklet image representation's and classifier's parameters are tuned, a test of the full CAD system is performed on a set of digital radiographic images coming out from an IMS Giotto Image MD full field digital mammography (FFDM) system. This dataset includes 78 radiographic images with one tumoral mass and 100 normal. They have been collected both in the course of the clinical evaluation of the FFDM system and subsequently during the regular clinical examinations. Locations of masses have been marked by expert radiologists and collected together with radiographic images. Unfortunately, subtlety is not provided.

### B. The ranklet image representation

The nonparametric property of the ranklet transform derives from its being based on the relative rank of pixels rather than on their intensity values. In particular, being based on nonparametric techniques such as the rank transform, the Wilcoxon test, and the Mann-Whitney test, some details will be given on that. Its orientation-selective property, instead, derives from its being mainly modeled on Haar wavelets. In analogy to the wavelet transform, in fact, ranklet coefficients can be computed at different orientations by applying vertical, horizontal, and diagonal Haar wavelet supports to each image under analysis. To this purpose, Haar wavelet supports will be introduced and their relationship with the ranklet transform will be clarified. Similarly, its multiresolution property derives from the feasibility of calculating the ranklet coefficients at different resolutions by stretching and shifting the aforementioned Haar wavelet supports. To this purpose, a suitable stretch and shift scheme will be proposed.

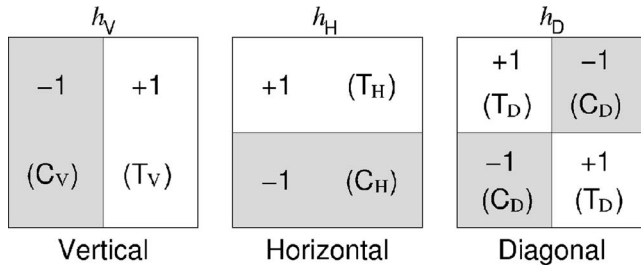


FIG. 2. The three Haar wavelet supports  $h_V$ ,  $h_H$ , and  $h_D$ . From left to right, the vertical, horizontal, and diagonal Haar wavelet supports.

### 1. Nonparametric and orientation-selective properties

Suppose a set  $Z$  of  $p_1, \dots, p_N$  pixels is given. The so-called rank transform  $\pi$  substitutes each pixel intensity value with its relative order among all the other pixels, namely with its rank.<sup>25</sup> As an example

$$Z = \begin{pmatrix} 55 & 99 & 25 & 153 \\ 26 & 75 & 92 & 200 \\ 21 & 64 & 88 & 154 \end{pmatrix} \Rightarrow \pi^Z = \begin{pmatrix} 4 & 9 & 2 & 10 \\ 3 & 6 & 8 & 12 \\ 1 & 5 & 7 & 11 \end{pmatrix}.$$

In case the set  $Z$  contains pixels with equal intensity values, mid ranks are introduced. They are computed assigning to each group of pixels with an equal intensity value the average of the ranks they occupy. The rank transform is closely related to the so-called Wilcoxon test. Given the same set  $Z$  of  $p_1, \dots, p_N$  pixels, suppose it is split into two subsets  $T$  and  $C$ , with  $n$  and  $m$  pixels each, so that  $n+m=N$ . To state whether the  $n$  pixels in  $T$  have intensity values significantly higher than the  $m$  pixels in  $C$ , the entire set of the  $N$  pixels is ranked by applying the rank transform  $\pi$  on  $Z$ . The Wilcoxon test  $W_S$  is then defined as the sum of the  $n$  ranks in  $T$ , namely, as  $W_S = \sum_{p \in T} \pi^Z(p)$ .<sup>26</sup> The  $n$  pixels in  $T$  are thus judged to have intensity values significantly higher than the  $m$  pixels in  $C$  if the Wilcoxon test is above a critical value  $\tau$ . In order to deal with something equivalent to the Wilcoxon test, but with an immediate interpretation in terms of pixels comparison, the Mann-Whitney test  $W_{XY}$  is defined as  $W_{XY} = W_S - n(n+1)/2$ .<sup>26</sup> As can be easily demonstrated, the value of the Mann-Whitney test  $W_{XY}$  is equal to the number of pixel pairs  $(p_i, p_j)$ , with  $p_i \in T$  and  $p_j \in C$ , such that the intensity value of  $p_i$  is higher than  $p_j$ . Its values range therefore from 0 to the number of pairs  $(p_i, p_j) \in T \times C$ , namely  $nm$ .

After having provided the basis for the nonparametric property of the ranklet transform, suppose that an image is constituted by the aforementioned set  $Z$  of  $p_1, \dots, p_N$  pixels. In order to compute the Mann-Whitney test, a possible choice in dividing the  $N$  pixels is to split them into two subsets  $T$  and  $C$ , of size  $n=m=N/2$ , thus assigning half of the pixels to the subset  $T$  and half to the subset  $C$ . With this in mind, it is possible to define the two subsets  $T$  and  $C$  being inspired by the three Haar wavelet supports depicted in Fig. 2. In particular, for the vertical Haar wavelet support ( $h_V$ ), the two subsets  $T_V$  and  $C_V$  are defined; similarly for the horizontal and diagonal ones. It is worth noting that the ar-

bitrariness that characterizes the selection of the two subsets  $T$  and  $C$  is crucial in order to freely choose the two subsets based on the Haar wavelet supports. This forms the basis for the orientation-selective property of the ranklet transform.

The definition of the ranklet transform is now almost straightforward. Given an image constituted by the set  $Z$  of  $p_1, \dots, p_N$  pixels, the horizontal, vertical, and diagonal ranklet coefficients can be computed as

$$R_j = \frac{W_{XY}^j}{nm/2} - 1, \quad j = V, H, D, \quad \text{where } -1 \leq R_j \leq +1. \quad (1)$$

Here,  $W_{XY}^j$  is computed by splitting the  $N$  pixels into the two subsets  $T_j$  and  $C_j$  differently for each  $j=V, H, D$ , as discussed previously for the Haar wavelet supports. As depicted in Fig. 3, the geometric interpretation of the ranklet coefficients  $R_j$  is thus quite simple. Suppose that the image we are dealing with is characterized by a vertical edge with the darker side on the left, where  $C_V$  is located, and the brighter side on the right, where  $T_V$  is located.  $R_V$  is then close to  $+1$ , as many pixels in  $T_V$  have higher intensity values than the pixels in  $C_V$ . Conversely,  $R_V$  is close to  $-1$  if the dark and bright sides are reversed. Note also that horizontal edges or other patterns with no global left-right variation of intensity give a value close to 0. Similar considerations can be drawn for the other ranklet coefficients, namely  $R_H$  and  $R_D$ .

### 2. Multiresolution property

The close correspondence between the Haar wavelet transform and the ranklet transform leads directly to the extension of the latter to its multiresolution formulation. Similarly to what is done for the bidimensional Haar wavelet transform, the ranklet coefficients can be computed at different resolutions by simply stretching and shifting the Haar wavelet supports. The multiresolution ranklet transform of an image is thus a set of triplets of vertical, horizontal, and diagonal ranklet coefficients, each one corresponding to a specific stretch and shift of the Haar wavelet supports. As an example, suppose that the multiresolution ranklet transform of an image with pixel size  $8 \times 8$  is performed at resolutions  $\{8, 4, 2\}$ , namely using Haar wavelet supports with pixel size  $8 \times 8$ ,  $4 \times 4$ , and  $2 \times 2$ , as shown in Fig. 4. Suppose also that the horizontal and vertical shifts of the Haar wavelet supports along the horizontal and vertical dimensions of the image are of 1 pixel. The multiresolution ranklet transform of the image is then composed by 1 triplet  $R_{V,H,D}$  of ranklet coefficients deriving from the ranklet transform at resolution 8, 25 triplets from that at resolution 4, and 49 triplets from that at 2. To generalize the calculation to arbitrary image sizes and resolutions, the number  $n_T$  of triplets  $R_{V,H,D}$  at each resolution is thus computed as  $n_T = (I+1-S)^2$ , where  $I$  and  $S$  represent the linear dimension of the image and that of the Haar wavelet support, respectively.

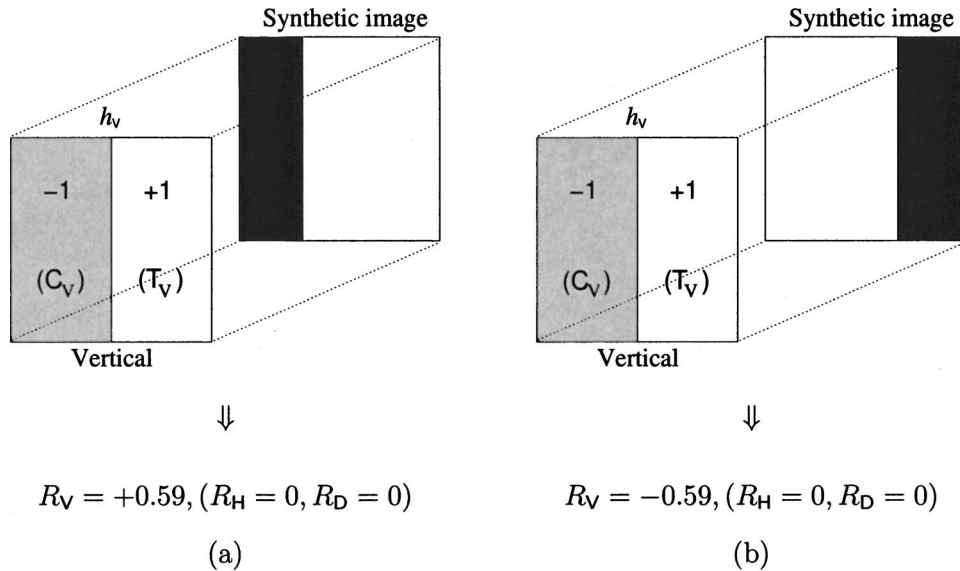


FIG. 3. Ranklet transform applied to two synthetic images. The application of the vertical Haar wavelet support ( $h_v$ ) results in  $R_V = +0.59$  and  $R_V = -0.59$  for the synthetic images shown in (a) and (b), respectively. Conversely, due to symmetry reasons, the application of the horizontal and diagonal Haar wavelet supports ( $h_H, h_D$ ) results in  $R_H = 0$  and  $R_D = 0$ , regardless of the synthetic image analyzed.

**C. SVM classification**

One intrinsic drawback of the proposed approach is that the number of classification features is very high. It is thereby of fundamental importance to use a classifier able to deal with such high dimensional feature spaces. To this purpose, SVM is chosen.

SVM constructs a binary classifier from a set of  $l$  training samples that consists of labeled patterns  $(\mathbf{x}_i, y_i) \in \mathbb{R}^N \times \{\pm 1\}$ , with  $i = 1, \dots, l$ .<sup>18,19</sup> In order to separate the two classes, it selects the maximal margin hyperplane, namely the hyperplane which causes the largest separation in the feature space between itself and the borderline training samples of the two classes. The decision function is thus computed as  $f(\mathbf{x}) = \text{sgn}(\sum_{i=1}^l y_i \alpha_i (\mathbf{x} \cdot \mathbf{x}_i) + b)$ , where the coefficients  $\alpha_i$  and  $b$  are calculated by solving a quadratic programming problem. Note that, since the coefficients  $\alpha_i$  are sparse, as known from SVM theory,<sup>18,19</sup> the number of dot products  $\mathbf{x} \cdot \mathbf{x}_i$  that must

be actually computed is sensibly smaller than  $l$ . As a result, SVM is particularly well suited for classification problems with very high dimensional feature spaces. When samples are not linearly separable in the feature space, a nonlinear transformation  $\phi(\mathbf{x})$  is used.<sup>27</sup> In order to implement this mapping, the dot products  $\mathbf{x} \cdot \mathbf{x}_i$  are substituted by  $\phi(\mathbf{x}) \cdot \phi(\mathbf{x}_i) \equiv K(\mathbf{x}, \mathbf{x}_i)$ , commonly referred to as kernel functions. Admissible and typical kernels are the linear kernel  $K(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ , the polynomial kernel  $K(\mathbf{x}, \mathbf{y}) = (\gamma \mathbf{x} \cdot \mathbf{y} + r)^d$ , etc.

To evaluate and optimize the classification performances of the ranklet image representation in combination with SVM, a 10-fold cross-validation procedure<sup>28</sup> is adopted while using DDSM crops. The dataset is partitioned into ten distinct and homogeneous random partitions. Nine of the partitions are used for training, whereas the remaining for test. Training and test are repeated ten times by permuting the test partition in a round-robin manner. For each permu-

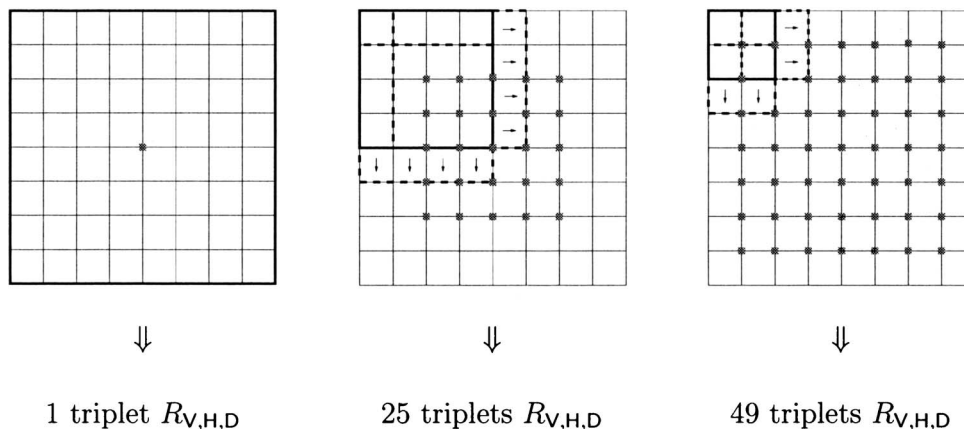


FIG. 4. Multiresolution ranklet transform of an image with pixel size  $8 \times 8$  at resolutions  $\{8, 4, 2\}$ , namely, using Haar wavelet supports with pixel size  $8 \times 8, 4 \times 4$ , and  $2 \times 2$ .

tation, SVM is trained with 900 mass crops and 4500 non-mass crops, then tested, respectively, on 100 and 500. Classification results are presented by performing the receiver operating characteristic (ROC) curve analysis of the system.<sup>29</sup> The ROC curves and their associated areas  $A_z$  are estimated using the ROCKIT software by Metz *et al.*<sup>30</sup> On the other hand, performances of the full CAD system are estimated on FFDM radiographic images and presented in terms of free-response operating characteristic (FROC) curves.<sup>31</sup>

### III. TESTS AND RESULTS

Four main tests are carried out. The first three are intended to evaluate and optimize the ranklet image representation in combination with SVM. To this purpose, DDSM crops are considered. The highest classification result achieved will be compared with those reached by the three best-performing image representations yet developed and evaluated, namely a pixel (**PixHRS**), a discrete wavelet (**DwtHS3**), and an overcomplete wavelet (**OwtS2**) image representation.<sup>20</sup> For the pixel image representation, each ROI is encoded by concatenating its gray-level intensity values. As far as the discrete wavelet and overcomplete wavelet image representations are concerned, instead, classification features are represented by the wavelet coefficients resulting from the application of the discrete wavelet and the overcomplete wavelet transform to each ROI, respectively. Also, in the fourth test, the best ranklet image representation's and classifier's configuration will be implemented on the full CAD system for testing on FFDM radiographic images.

The ranklet image representations will be indicated as **RankletS**<sub>{RES1,RES2,...}</sub>, in the following. The prefix **Ranklet** stands for its being a ranklet image representation, whereas the post-fix **S** for its having correspondent features (i.e., ranklet coefficients) scaled in the interval  $[-1, 1]$ . This last property, in particular, is automatically assured by the definition of ranklet coefficients given in Eq. (1). An eventual integer number on the right side of the post-fix **S** indicates the degree of the polynomial SVM kernel used for classification. This corresponds to the parameter  $d$  discussed in Sec. II C. When none, a linear SVM kernel is used instead. The subscript {RES1,RES2,...} indicates the resolutions at which the multiresolution ranklet transform is performed. In particular, for the sake of conciseness, the ranklet resolutions {16, 14} are broadly categorized as low resolutions, the ranklet resolutions {12, 10, 8, 6} as intermediate resolutions, and the ranklet resolutions {4, 2} as high resolutions.

The multiresolution ranklet transform of a  $64 \times 64$  crop results in a huge amount of ranklet coefficients, specifically when high resolutions are considered. The reason for that relies in the square law that links the number  $n_T$  of triplets  $R_{V,H,D}$  to each resolution, i.e.,  $n_T = (I+1-S)^2$ . The 6000 crops of the image database are thus required to be further resized from their original  $64 \times 64$  pixel size. After initial experimentation, the new pixel size is chosen to be  $16 \times 16$ . See Table I, just to get an idea of the number of features hence involved. Here, the correspondence between some of the

TABLE I. Multiresolution ranklet transform of an image with pixel size  $16 \times 16$ . Number of resulting ranklet coefficients for each different combination of resolutions.

Resolutions	# Ranklet coefficients
{16, 14, 12, 10, 8, 6, 4, 2}	2040
{16, 8, 4, 2}	1428
{16, 8, 2}	921
{16, 2}	678
{16, 4}	510
{16, 14, 12, 10}	252
{16, 8}	246

resolutions at which the multiresolution ranklet transform can be performed on  $16 \times 16$  crops and the relative number of ranklet coefficients computed is shown. As regards the other parameters of the polynomial SVM kernel (i.e.,  $\gamma, r$ ) and the regularization parameter  $C$  of SVM (see Vapnik *et al.*<sup>18,19</sup>), they are arbitrarily set equal to unity, since classification performances are found to be almost completely unaffected by any change on them.

#### A. SVM kernels

The first test performed is intended to understand the influence of different SVM kernels on classification performances. To this aim, original crops are resized from  $64 \times 64$  pixel size to  $16 \times 16$  by means of bilinear resizing. The multiresolution ranklet transform is then applied at resolutions {16, 8, 4, 2} to the resized crops, thus producing 1428 classification features for each ROI. This choice is arbitrary, nevertheless it is a reasonable starting point, since it spans over the entire range of resolutions, from coarsest to finest. Using as an image representation the resulting ranklet coefficients, several SVM kernels are hence evaluated, namely linear, polynomial with degree 2 and 3.

Figure 5 shows the results achieved using this procedure. Classification performances are improved when the polynomial degree of the SVM kernel is increased. More specifically, the linear SVM kernel (**RankletS**<sub>{16,8,4,2}</sub>) achieves average  $A_z$  value of  $0.946 \pm 0.004$ , the polynomial SVM kernel with degree 2 (**RankletS2**<sub>{16,8,4,2}</sub>) of  $0.974 \pm 0.003$ , and the polynomial SVM kernel with degree 3 (**RankletS3**<sub>{16,8,4,2}</sub>) of  $0.978 \pm 0.003$ . In particular, the difference in the  $A_z$  values between the polynomial SVM kernels with degree 2 or 3 and the linear SVM kernel is statistically significant (two-tailed  $p$  value  $< 0.0001$ ).

#### B. Ranklet resolutions

The second test performed is intended to comprehend the effects of the multiresolution property on classification performances. As for the previous test, original crops are resized from  $64 \times 64$  pixel size to  $16 \times 16$  by means of bilinear resizing. The multiresolution ranklet transform is then applied to the resized crops by using several combinations of different resolutions, namely some of those shown in Table I. The number of resulting classification features hence varies ac-

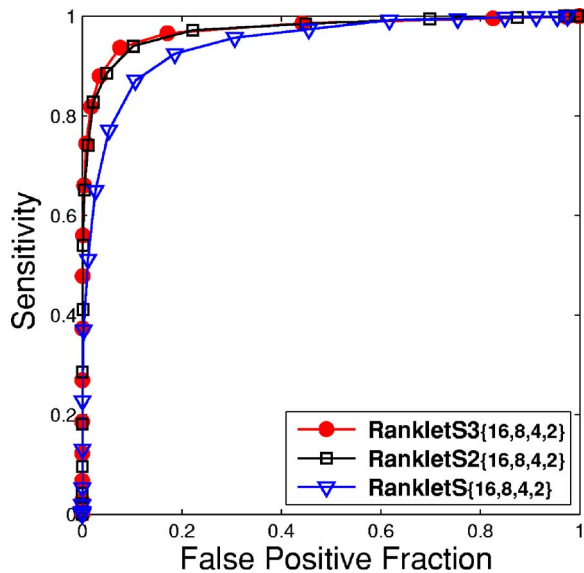


FIG. 5. ROC curves obtained varying SVM kernels. **RankletS**<sub>{16,8,4,2}</sub> ( $A_z$  values of  $0.946 \pm 0.004$ ) uses a linear SVM kernel. **RankletS**<sub>2{16,8,4,2}</sub> ( $0.974 \pm 0.003$ ) and **RankletS**<sub>3{16,8,4,2}</sub> ( $0.978 \pm 0.003$ ) use a polynomial SVM kernel with degree 2 and 3, respectively. The difference in the average  $A_z$  value of **RankletS**<sub>2{16,8,4,2}</sub> or **RankletS**<sub>3{16,8,4,2}</sub> and that of **RankletS**<sub>{16,8,4,2}</sub> achieves statistical significance with two-tailed  $p$  value  $< 0.0001$ .

according to the resolutions at which the analysis is performed. Owing to the results achieved in Sec. III A, a polynomial SVM kernel with degree 3 is used.

Figure 6 shows, for instance, the results achieved by classifying the ranklet coefficients of resolutions  $\{16, 14, 12, 10\}$ ,

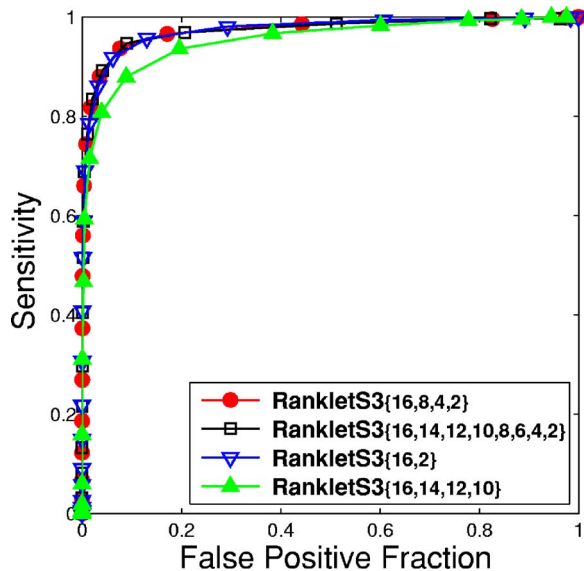


FIG. 6. ROC curves obtained varying ranklet resolutions. **RankletS**<sub>3{16,14,12,10,8,6,4,2}</sub> ( $A_z$  values of  $0.976 \pm 0.003$ ) and **RankletS**<sub>3{16,8,4,2}</sub> ( $0.978 \pm 0.003$ ) account for low, intermediate, and high ranklet resolutions. **RankletS**<sub>3{16,2}</sub> ( $0.977 \pm 0.002$ ) for low and high ranklet resolutions. **RankletS**<sub>3{16,14,12,10}</sub> ( $0.957 \pm 0.004$ ) for low ranklet resolutions. The difference in the average  $A_z$  value of **RankletS**<sub>3{16,14,12,10,8,6,4,2}</sub>, **RankletS**<sub>3{16,8,4,2}</sub> or **RankletS**<sub>3{16,2}</sub> and that of **RankletS**<sub>3{16,14,12,10}</sub> achieves statistical significance with two-tailed  $p$  value  $< 0.0001$ .

$\{8, 6, 4, 2\}$  and  $\{16, 8, 4, 2\}$ . In other words, the whole range of resolutions is considered, as for the  $\{16, 14, 12, 10, 8, 6, 4, 2\}$  case. At least, a sampled version of them is taken into account, as for the  $\{16, 8, 4, 2\}$  case. In either ways, low, intermediate, and high resolutions are all contemplated. As evident from the ROC curve analysis, both those image representations (**RankletS**<sub>3{16,14,12,10,8,6,4,2}</sub> and **RankletS**<sub>3{16,8,4,2}</sub>) perform quite similarly, the former achieving average  $A_z$  values of  $0.976 \pm 0.003$ , the latter of  $0.978 \pm 0.003$ . The difference between their areas does not reach statistical significance.

Figure 6 shows also the results achieved by classifying the ranklet coefficients of resolutions  $\{16, 2\}$ , thus ignoring all intermediate resolutions. Looking at the classification performances, it is evident that intermediate resolutions are not essential for classification purposes. The results obtained by **RankletS**<sub>3{16,2}</sub> (i.e., average  $A_z$  values of  $0.977 \pm 0.002$ ), in fact, are not significantly different from those obtained by **RankletS**<sub>3{16,8,4,2}</sub> (i.e., average  $A_z$  values of  $0.978 \pm 0.003$ ).

In almost the same way, Fig. 6 shows the results achieved by classifying the ranklet coefficients of resolutions  $\{16, 14, 12, 10\}$ , thus ignoring all high resolutions. In this case, it is evident that high resolutions are quite important for classification purposes. The results achieved by **RankletS**<sub>3{16,14,12,10}</sub> (i.e., average  $A_z$  values of  $0.957 \pm 0.004$ ), in fact, perform significantly worse than those achieved by **RankletS**<sub>3{16,8,4,2}</sub> (i.e., average  $A_z$  values of  $0.978 \pm 0.003$ ), with two-tailed  $p$  value  $< 0.0001$ . Similar results are found when low resolutions are ignored.

### C. Histogram equalization

The third test performed is intended to investigate the influence of histogram equalization on classification performances: previous investigations on pixel and wavelet image representations, in fact, demonstrated that the application of histogram equalization on ROIs affects rather sensibly classification performances.<sup>20</sup> Histogram equalization is a well-known image processing technique whose aim is to spread the histogram of the processed image over the entire range of possible gray levels.<sup>32</sup> As a result, the processed image shows higher contrast. This aspect is explored for the ranklet approach by first processing the original  $64 \times 64$  pixels size crops by means of histogram equalization. The resulting crops are then resized to  $16 \times 16$ , transformed by means of the multiresolution ranklet transform at resolutions  $\{16, 8, 4, 2\}$  and, finally, classified by SVM. As for the first test, the number of resulting classification features is 1428. Again, owing to the results achieved in Sec. III A, a polynomial SVM kernel with degree 3 is used.

Figure 7 shows the results achieved by the above discussed image representation (**RankletHS**<sub>3{16,8,4,2}</sub>, the postfix **H** indicating crops submitted to histogram equalization). In this case, average  $A_z$  values of  $0.977 \pm 0.003$  are reached. This image representation is compared to the previously discussed **RankletS**<sub>3{16,8,4,2}</sub> ( $A_z = 0.978 \pm 0.003$ ), namely an identical image representation with the only difference that crops are not submitted to histogram equalization. Due to the

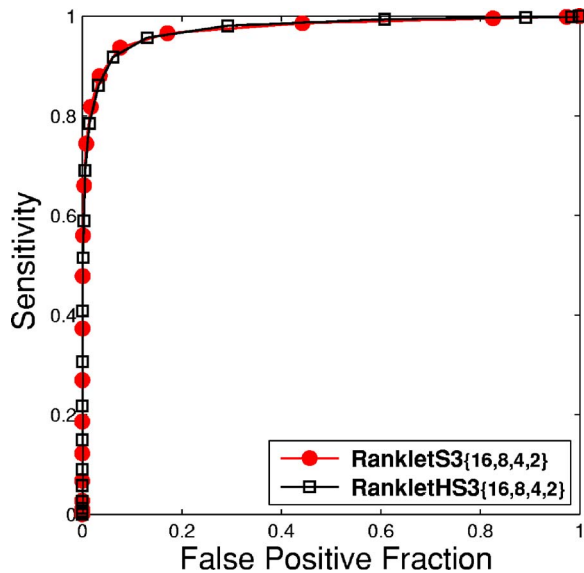


FIG. 7. ROC curves obtained applying histogram equalization. **RankletHS3**<sub>{16,8,4,2}</sub> ( $A_z$  values of  $0.977 \pm 0.003$ ) deals with crops submitted to histogram equalization before the ranklet transform is performed. The difference in its average  $A_z$  value and that of **RankletS3**<sub>{16,8,4,2}</sub> ( $0.978 \pm 0.003$ ) does not achieve statistical significance.

high  $p$  value (i.e., 0.70) for the difference in their  $A_z$  values, the results achieved in either cases are almost the same.

#### D. Ranklet image representation implemented on the full CAD system

The results described in Sec. III C seem to indicate that the ranklet image representation is quite insensitive to variations on the gray-level intensity histogram of the crops. The fourth and last test performed is thus intended to investigate the actual beneficial effects of this property, specifically when classification of ROIs by means of the ranklet image representation is settled in the complete framework of the CAD system. Of course, the system also relies on several other steps, such as, for example, the segmentation of radiographic images, detection of ROIs, and false positives reduction.<sup>14–17</sup> However, a detailed description of these steps is out of the scope of this study.

The SVM classifier trained by means of the DDSM crops encoded using the ranklet image representation **RankletS3**<sub>{16,8,4,2}</sub> is considered. Also, the SVM classifier trained by means of the same crops, but encoded using the overcomplete wavelet image representation **OwtS2**, is considered. In this way, they are both trained using digitized screen film radiographic images. In other words, they are trained by means of crops whose gray-level intensity histograms have some peculiar specifics, for instance, having gray-level intensity values centered on 2400 and ranging approximately between 2000 and 2800, as shown in Fig. 9(a). A test set **T<sub>DIGITAL</sub>**, consisting of the FFDM digital radiographic images described in Sec. II A, is considered as well. As shown in Fig. 9(b), their gray-level intensity histogram is sensibly different from that of digitized DDSM radiographic images, for instance, having gray-level intensity values cen-

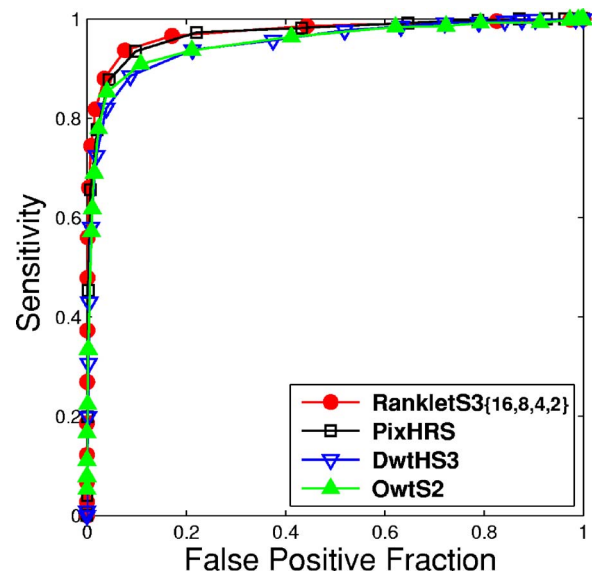


FIG. 8. ROC curves comparison. **RankletS3**<sub>{16,8,4,2}</sub> ( $A_z$  value of  $0.978 \pm 0.003$ ) is one of the best ranklet image representations evaluated. **PixHRS** ( $0.973 \pm 0.002$ ), **DwtHS3** ( $0.948 \pm 0.004$ ), and **OwtS2** ( $0.956 \pm 0.003$ ) are the best pixel, discrete wavelet, and overcomplete wavelet image representations previously developed and evaluated (Ref. 20). The difference in the  $A_z$  value of **RankletS3**<sub>{16,8,4,2}</sub> and that of **PixHRS** does not achieve statistical significance. Conversely, the difference in the  $A_z$  value between **RankletS3**<sub>{16,8,4,2}</sub> and **DwtHS3** or **OwtS2** is statistically relevant with two-tailed  $p$  value  $< 0.0001$ .

tered on 32 100 and ranging approximately between 32 000 and 32 200. Nevertheless, by applying a sigmoidal look-up table (LUT),<sup>16</sup> the gray-level intensity histogram of digital radiographic images can be mapped into a histogram very close to that of digitized ones. These resulting radiographic images are collected in an independent test set referred to as **T<sub>DIGITAL+LUT</sub>**. By testing our CAD system with the **RankletS3**<sub>{16,8,4,2}</sub> SVM classifier both on **T<sub>DIGITAL</sub>** and **T<sub>DIGITAL+LUT</sub>**, the typical result achieved is that depicted in Figs. 10(a) and 10(b). In other words, even though the training procedure is performed using digitized crops, the classification performances are almost identical on both digital radiographic images and digital radiographic images processed with LUT. In particular, approximately 70% sensitivity per image is achieved with 0.7 false positives on both **T<sub>DIGITAL</sub>** and **T<sub>DIGITAL+LUT</sub>**; see the two correspondent FROC curves in Fig. 11. Conversely, when testing our CAD system with the **OwtS2** SVM classifier, the typical situation is that depicted in Figs. 10(c) and 10(d). In other words, the wavelet image representation achieves dramatically low classification performances when tested on digital images, whereas higher classification performances when tested on digital images processed with LUT. In particular, approximately zero sensitivity per image (e.g., 2.6%) and zero false positives (e.g., 0.02) are constantly achieved on **T<sub>DIGITAL</sub>**, whereas approximately 58% sensitivity per image is achieved with 0.7% false positives on **T<sub>DIGITAL+LUT</sub>**; see the two correspondent FROC curves on Fig. 11.

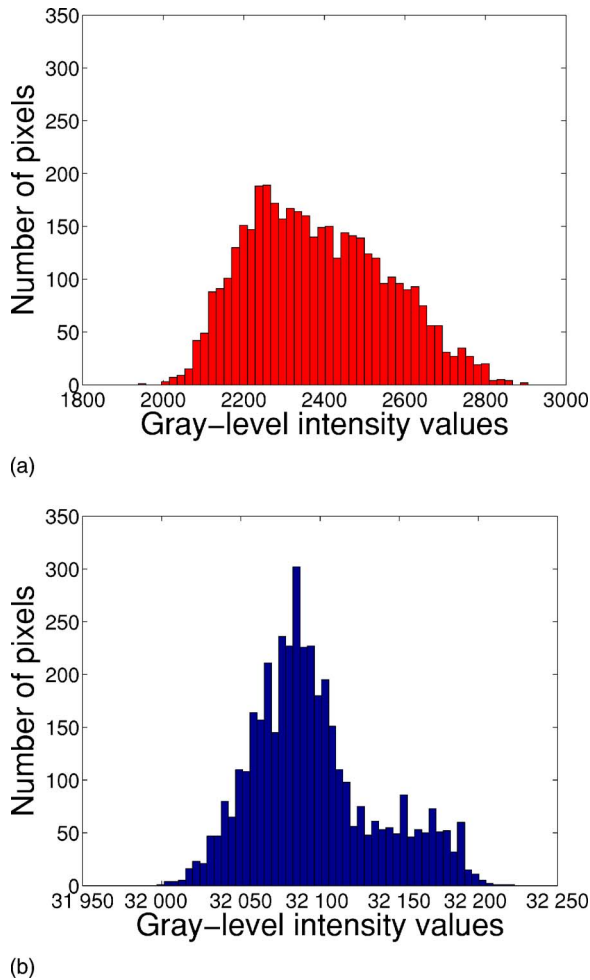


FIG. 9. Gray-level intensity histograms of two crops extracted, as an example, from two different images. Digitized DDSM radiographic image (a). Digital FFDM radiographic image (b).

#### IV. DISCUSSION

The performances presented show that **RankletS3**<sub>{16,8,4,2}</sub> provides the best classification results among all the ranklet image representations, namely an average  $A_z$  value of  $0.978 \pm 0.003$  and 90% sensitivity with an FPF value of 4.5%. See Table II for a summary of the results reached by the different ranklet image representations evaluated. Also, for comparison purposes, the classification results achieved by the previously developed image representations are reported in Table III. By pairwise comparison, it turns out that the improvement in the  $A_z$  value with **RankletS3**<sub>{16,8,4,2}</sub> over that of **PixHRS** ( $A_z = 0.973 \pm 0.002$ , 90% sensitivity with an FPF value of 6%) does not reach statistical significance. Conversely, by showing a two-tailed  $p$  value  $< 0.0001$ , the improvement over that of **DwtHS3** ( $A_z = 0.948 \pm 0.004$ , 90% sensitivity with an FPF value of 11%) and **OwtS2** ( $A_z = 0.956 \pm 0.003$ , 90% sensitivity with an FPF value of 7%) is statistically significant. The ROC curves for classification in the **RankletS3**<sub>{16,8,4,2}</sub> feature space and in the previously developed ones are shown in Fig. 8. Also, Table II and Table III show how the number of features affects the calculation times required by each image representation for the analysis

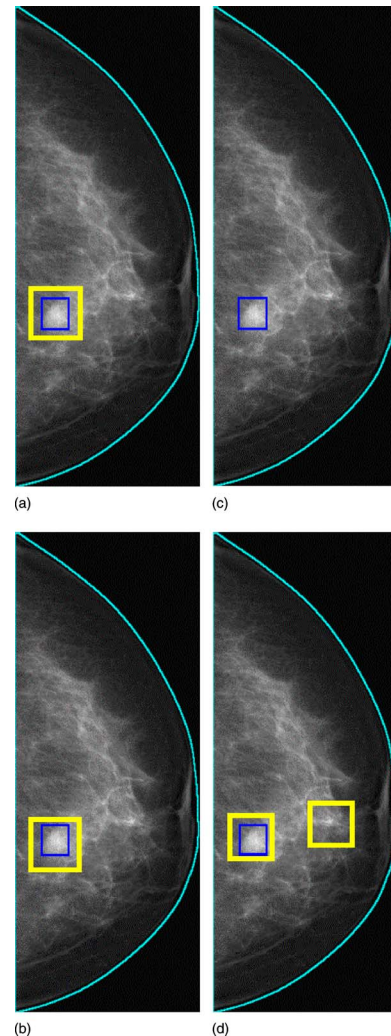


FIG. 10. Ranklet image representation and overcomplete wavelet image representation in action. **RankletS3**<sub>{16,8,4,2}</sub> SVM classifier tested on  $T_{\text{DIGITAL}}$  (a) and  $T_{\text{DIGITAL+LUT}}$  (b). **OwtS2** SVM classifier tested on  $T_{\text{DIGITAL}}$  (c) and  $T_{\text{DIGITAL+LUT}}$  (d). The small square marks represent the radiologist's interpretation, whereas the large ones represent the automatic CAD's analysis.

of an entire radiographic image. Tests are performed on a dual Intel Xeon 2.6 GHz PC. Even though the achieved calculation times seem to indicate that all the image representations considered are quite reliable for real-time CAD implementations, feature reduction techniques are currently under investigation.

#### A. SVM kernels

From the results presented in Sec. III A it emerges that, when dealing with the ranklet image representation, polynomial SVM kernels perform better than linear SVM kernels.

Being computed as  $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})$ , the inner products of linear SVM kernels put emphasis specifically on correlations among correspondent features. This is particularly useful when dealing with monodimensional image representations, such as, for example, **PixHRS**. In this case, in fact, correlations among correspondent pixels are much more reliable as features than correlations among distant pixels.<sup>33</sup> On the con-



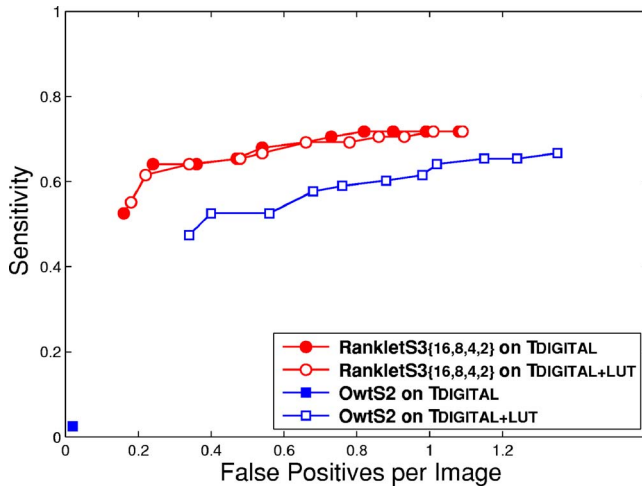


FIG. 11. FROC curves. **RankletS3**<sub>{16,8,4,2}</sub> SVM classifier tested on  $T_{\text{DIGITAL}}$  and  $T_{\text{DIGITAL+LUT}}$ . **OwtS2** SVM classifier tested on  $T_{\text{DIGITAL}}$  and  $T_{\text{DIGITAL+LUT}}$ .

rary, when dealing with multiresolution image representations, such as, for example, **DwtHS3**, **OwtS2**, and **RankletS3**<sub>{16,8,4,2}</sub>, correlations among distant features are important as well as correlations among correspondent features. In those cases, in fact, each pixel of the original suspicious region is represented many times in the feature vector, namely by a number of wavelet or ranklet coefficients, which is proportional to the number of multiresolution levels at which the analysis is performed. Being computed as  $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})^2$  or  $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y})^3$ , polynomial SVM kernels with degree 2 or 3 are thus in the condition of taking this aspect into account, hence to provide better classification performances.

## B. Ranklet resolutions

By looking at the results presented in Sec. III B, it emerges that low and high resolutions are quite important in order to achieve high classification performances. Conversely, intermediate resolutions appear less important, since they can be eliminated without sensibly affecting the classification results.

A practical explanation for that relies in the ability of low and high resolutions in encoding characteristics which are

TABLE II. Results summary. Ranklet image representations: average  $A_z$  values, FPF values at 90% sensitivity, and calculation times for the analysis of an entire radiographic image.

Image representation	# Features	$A_z$	FPF (%)	Time (s)
<b>RankletS</b> <sub>{16,8,4,2}</sub>	1428	0.946±0.004	15	30
<b>RankletS2</b> <sub>{16,8,4,2}</sub>	1428	0.974±0.003	6	30
<b>RankletS3</b> <sub>{16,8,4,2}</sub>	1428	0.978±0.003	4.5	30
<b>RankletS3</b> <sub>{16,14,12,10,8,6,4,2}</sub>	2040	0.976±0.003	4.5	45
<b>RankletS3</b> <sub>{16,2}</sub>	678	0.977±0.002	4.5	15
<b>RankletS3</b> <sub>{16,14,12,10}</sub>	252	0.957±0.004	13	5
<b>RankletHS3</b> <sub>{16,8,4,2}</sub>	1428	0.977±0.003	4.5	30

TABLE III. Results comparison. Best performing ranklet, pixel, discrete wavelet, and overcomplete wavelet image representations: average  $A_z$  values, FPF values of 90% sensitivity, and calculation times for the analysis of an entire radiographic image.

Image representation	# Features	$A_z$	FPF (%)	Time (s)
<b>RankletS3</b> <sub>{16,8,4,2}</sub>	1428	0.978±0.003	4.5	30
<b>PixHRS</b>	256	0.973±0.002	6	5
<b>DwtHS3</b>	4032	0.948±0.004	11	90
<b>OwtS2</b>	2955	0.956±0.003	7	60

particularly useful to discriminate ROIs containing masses from those representing normal tissue. Low resolutions encode informations concerning the symmetry of the gray-level intensity distribution that characterizes the image when analyzed at coarse resolutions. Owing to that, they are particularly suited to encode informations concerning the presence or absence in the suspicious region of a bright centered nucleus surrounded by more heterogeneous structures characterizing normal tissue. A typical signature indicating the presence of a mass is thus represented by the existence, in the feature vector, of vertical, horizontal, and diagonal low-resolution ranklet coefficients with high absolute values, namely approximately  $\pm 1$ . Conversely, high resolutions encode informations concerning the symmetry of the gray-level intensity distribution that characterizes the image when analyzed at fine resolutions. Differently from low resolutions, high resolutions are hence particularly suited to encode informations concerning the presence or absence in the suspicious region of a boundary delimiting the bright centered nucleus. At those fine resolutions, a signature indicating the presence of a mass is thus represented by vertical, horizontal, and diagonal high-resolution ranklet coefficients (in particular, those arising from the analysis at the borders of the image) with high absolute values.

## C. Histogram equalization

As far as the results discussed in Sec. III C are concerned, they show that histogram equalization is ineffective in improving classification performances.

This result is consistent with the fact that the ranklet transform deals with the relative rank of pixels rather than with their intensity values. This provides an image representation that is very robust to gray-level intensity variations, such as those provided by the application of histogram equalization on suspicious regions. In particular, the application of the ranklet transform to either a crop or an equalized crop results in two very similar feature vectors. Owing to that, the classification results achieved by either applying or not histogram equalization are almost identical.

## D. Ranklet image representation implemented on the full CAD system

According to the results presented in Sec. III D, when using the ranklet image representation, the CAD system's classifier must not be necessarily trained with crops extracted

from radiographic images having the same intensity characteristics as those used in test. On the contrary, it can be trained with a set of radiographic images collected by a specific mammographic detector, then tested on a set of radiographic images collected by a different mammographic detector; this, without worsening its classification performances.

Evidently, the reason for this important result relies again in the tolerance of the ranklet image representation to variations in the ROIs' gray-level intensity histogram. Conversely, being based on sums and differences in gray-level intensity content, the wavelet image representation is much less tolerant to histogram variations, and, in particular, to histogram's spread variations. Owing to that, being trained with crops whose gray-level intensity histogram is centered on 2400 and spread over the values 2000–2800, it is quite reasonable that no marks are drawn on digital images whose gray-level intensity histogram is centered on 32 100 and spread over the values 32 000–32 200. Note, by the way, that identical considerations could be drawn for **PixHRS** and **DwtHS3** as well. The results presented herein have a very important practical implication. By using the ranklet image representation, in fact, the CAD system does not have to be repeatedly tuned when radiographic images coming from different detectors are considered. On the contrary, the same mass detection scheme can be applied to radiographic images coming from different detectors or, eventually, acquired in different exposure conditions. As a further result, this allows us to exploit freely available mammographic image databases (such as the DDSM) for training the system, then to test it on sets of digital radiographic images collected in collaboration with local hospitals.

## V. CONCLUSION

The classification of breast tumoral masses and normal tissue is targeted. Classification features are represented by the coefficients that results from the specific image representation used in order to encode each ROI. Owing to its ability in handling very high dimensional feature spaces, an SVM classifier is chosen for the classification task.

To investigate whether better classification performances can be achieved with respect to a pixel, a discrete wavelet, and an overcomplete wavelet image representation previously presented,<sup>20</sup> a novel image representation is developed and evaluated, namely a ranklet image representation. Results demonstrate a great attitude. The best result achieved shows an average  $A_z$  value of  $0.978 \pm 0.003$  and 90% sensitivity with an FPF value of 4.5%. The improvement over the average  $A_z$  value of the pixel image representation ( $A_z = 0.973 \pm 0.002$ , 90% sensitivity with an FPF value of 6%) does not reach statistical significance. Conversely, the improvement over the average  $A_z$  value of the discrete wavelet image representation ( $A_z = 0.948 \pm 0.004$ , 90% sensitivity with an FPF value of 11%) and that of the overcomplete wavelet image representation ( $A_z = 0.956 \pm 0.003$ , 90% sensitivity with an FPF value of 7%) is statistically significant with two-tailed  $p$  value  $< 0.0001$ .

Those high classification results acquire even more importance when the robustness of this image representation to variations in the ROIs' gray-level intensity histogram is considered. As discussed in this work, when using the ranklet image representation, the CAD system's SVM classifier is (also) well suited to be trained on radiographic images collected by a specific mammographic detector and tested on radiographic images collected by a different one. As an extreme example, the train set can be comprised of digitized radiographic images, whereas the test set of digital radiographic images. Owing to the different conditions that may occur during the image acquisition procedure, and due to the difficulties of collecting really homogeneous image databases, this characteristic can be very useful in CAD systems.

## ACKNOWLEDGMENTS

The author would like to thank Prof. Charles E. Metz for providing the ROCKIT program. The author is grateful also to Dr. Med. Gianni Saguatti (Maggiore Hospital, Bologna, Italy) and Dr. Med. Claude Mair-Noack (Triemli Hospital, Zürich, Switzerland) for providing FFDM digital radiographic images. Finally, many thanks to Dr. Nico Lanconelli and Prof. Renato Campanini for helpful comments and critical revision of the manuscript.

<sup>a)</sup>Corresponding author. Electronic mail: masotti@bo.infn.it. Telephone: +39 051 2095136. Fax: +39 051 2095047.

Web: <http://www.bo.infn.it/~masotti/>

<sup>1</sup>A. Jemal, T. Murray, E. Ward, A. Samuels, R. C. Tiwari, A. Ghafoor, E. J. Feuer, and M. J. Thun, *Ca-Cancer J. Clin.* **55**(1), 10–30 (2005).

<sup>2</sup>H. C. Zuckerman, "The role of mammography in the diagnosis of breast cancer," in *Breast Cancer, Diagnosis and Treatment* (McGraw-Hill, New York, 1987), pp. 152–172.

<sup>3</sup>M. L. Giger, K. Doi, H. MacMahon, R. M. Nishikawa, K. R. Hoffmann, C. J. Vyborny, R. A. Schmidt, H. Jia, K. Abe, and X. Chen, "An intelligent workstation for computer-aided diagnosis," *Radiographics* **13**, 647–656 (1993).

<sup>4</sup>N. Ibrahim, H. Fujita, T. Hara, and T. Endo, "Automated detection of clustered microcalcifications on mammograms: CAD system application to MIAS database," *Phys. Med. Biol.* **42**, 2577–2589 (1997).

<sup>5</sup>H. P. Chan, B. Sahiner, M. A. Helvie, N. Petrick, M. A. Roubidoux, T. E. Wilson, D. D. Adler, C. Paramagul, J. S. Newman, and S. Sanjay-Gopal, "Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: An ROC study," *Radiology* **212**, 817–827 (1999).

<sup>6</sup>M. L. Giger, N. Karssemeijer, and S. G. Armato III, "Computer-aided diagnosis in medical imaging," *IEEE Trans. Med. Imaging* **20**, 1205–1208 (2001).

<sup>7</sup>M. L. Giger, "Computer-aided diagnosis of breast lesions in medical images," *Comput. Sci. Eng.* **2**, 39–45 (2000).

<sup>8</sup>K. Doi, H. MacMahon, S. Katsuragawa, R. M. Nishikawa, and Y. Jiang, "Computer-aided diagnosis in radiology: potential and pitfalls," *Eur. J. Radiol.* **31**, 97–109 (1999).

<sup>9</sup>E. Shaw de Paredes, "Radiographic breast anatomy: Radiologic signs of breast cancer," *RSNA Categorical Course in Physics*, 1993, pp. 35–46.

<sup>10</sup>D. Wei, H. P. Chan, M. A. Helvie, B. Sahiner, N. Petrick, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis," *Med. Phys.* **22**, 1501–1513 (1995).

<sup>11</sup>B. Sahiner, H. P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images," *IEEE Trans. Med. Imaging* **15**, 598–610 (1996).

<sup>12</sup>M. A. Kupinski and M. L. Giger, "Investigation of regularized neural networks for the computerized detection of mass lesions in digital mammograms," in *Proceedings of the IEEE International Engineering Medi-*

- cine and Biology Conference*, 1997, pp. 1336–1339.
- <sup>13</sup>G. M. te Brake, N. Karssemeijer, and J. H. Hendriks, “An automatic method to discriminate malignant masses from normal tissue in digital mammograms,” *Phys. Med. Biol.* **45**, 2843–2857 (2000).
- <sup>14</sup>R. Campanini, A. Bazzani, A. Bevilacqua, D. Bollini, D. N. Dongiovanni, E. Iampieri, N. Lanconelli, A. Riccardi, M. Roffilli, and R. Tazzoli, “A novel approach to mass detection in digital mammography based on Support Vector Machines (SVM),” in *Proceedings of the International Workshop on Digital Mammography*, 2002, pp. 399–401.
- <sup>15</sup>R. Campanini, D. Dongiovanni, E. Iampieri, N. Lanconelli, M. Masotti, G. Palermo, A. Riccardi, and M. Roffilli, “A novel featureless approach to mass detection in digital mammograms based on Support Vector Machines,” *Phys. Med. Biol.* **49**, 961–975 (2004).
- <sup>16</sup>R. Campanini, E. Angelini, D. N. Dongiovanni, E. Iampieri, N. Lanconelli, C. Mair-Noack, M. Masotti, G. Palermo, M. Roffilli, G. Saguatti, and O. Schiaratura, “Preliminary results of a featureless CAD system on FFDM images,” in *Proceedings of the International Workshop on Digital Mammography*, 2004, in press.
- <sup>17</sup>R. Campanini, E. Angelini, E. Iampieri, N. Lanconelli, M. Masotti, M. Roffilli, O. Schiaratura, and M. Zanoni, “A fast algorithm for intra-breast segmentation of digital mammograms for CAD systems,” in *Proceedings of the International Workshop on Digital Mammography*, 2004, in press.
- <sup>18</sup>V. Vapnik, *The Nature of Statistical Learning Theory* (Springer-Verlag, New York, 1995).
- <sup>19</sup>V. Vapnik, *Statistical Learning Theory* (Wiley, New York, 1998).
- <sup>20</sup>E. Angelini, R. Campanini, E. Iampieri, N. Lanconelli, M. Masotti, and M. Roffilli, “Testing the performances of different image representations for mass classification in digital mammograms,” *Int. J. Mod. Phys. C* **17**, 113–131 (2006).
- <sup>21</sup>F. Smeraldi, “Ranklets: orientation selective non-parametric features applied to face detection,” in *Proceedings of the International Conference on Pattern Recognition*, 2002, Vol. 3, pp. 379–382.
- <sup>22</sup>F. Smeraldi, “A nonparametric approach to face detection using ranklets,” in *Proceedings of the International Conference on Audio and Video-Based Biometric Person Authentication*, 2003, pp. 351–359.
- <sup>23</sup>E. Franceschi, F. Odone, F. Smeraldi, and A. Verri, “Finding objects with hypothesis testing,” in *Proceedings of the Workshop on Learning for Adaptable Visual Systems*, 2004.
- <sup>24</sup>M. Heath, K. W. Bowyer, D. Copans, R. Moore, and P. Kegelmeyer, “The digital database for screening mammography,” in *Proceedings of the International Workshop on Digital Mammography*, 2000, pp. 212–218. URL <http://marathon.csee.usf.edu/Mammography/Database.html>
- <sup>25</sup>R. Zabih and J. Woodfill, “Non-parametric local transforms for computing visual correspondence,” in *Proceedings of the European Conference on Computer Vision*, 1994, Vol. 2, pp. 151–158.
- <sup>26</sup>E. L. Lehmann, *Nonparametrics: Statistical Methods Based on Ranks* (Holden-Day, San Francisco, 1975).
- <sup>27</sup>K. R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, “An introduction to kernel-based learning algorithms,” *IEEE Trans. Neural Netw.* **12**, 181–201 (2001).
- <sup>28</sup>R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Proceedings of the International Joint Conference on Artificial Intelligence*, 1995, pp. 1137–1145.
- <sup>29</sup>C. E. Metz, “ROC methodology in radiologic imaging,” *Invest. Radiol.* **21**, 720–733 (1986).
- <sup>30</sup>C. E. Metz, B. A. Herman, and J. H. Shen, “Maximum-likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data,” *Stat. Med.* **17**, 1033–1053 (1998).
- <sup>31</sup>P. C. Bunch, J. F. Hamilton, G. K. Sanderson, and A. H. Simmons, “A free response approach to the measurement and characterization of radiographic observer performance,” *J. Appl. Photogr. Eng.* **4**, 166–171 (1978).
- <sup>32</sup>R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. (Addison-Wesley, Reading, MA, 1992).
- <sup>33</sup>B. Schölkopf, P. Y. Simard, A. J. Smola, and V. N. Vapnik, “Prior knowledge in support vector kernels,” in *Proceedings of the Advances in Neural Information Processing Systems Conference*, 1998, Vol. 10, pp. 640–646.