A ranklet–based image representation for mass classification in digital mammograms

Matteo Masotti *

Department of Physics, University of Bologna

Abstract

Ranklets are non–parametric, multi–resolution and orientation selective features modelled on Haar wavelets. A ranklet–based image representation is proposed in this paper in order to solve a two–class classification problem. The first class is constituted by masses, breast tumors with size ranging from 3 mm to 30 mm, whereas the second class is constituted by non–masses. Masses and non–masses are both extracted from the University of South Florida (USF) mammographic image database, submitted to the ranklet transform and finally classified by means of a Support Vector Machine (SVM). Experiments demonstrate that the proposed image representation solves succesfully the two–class classification problem. Furthermore, it achieves an improvement over the pixel–based and wavelet–based representations tested on the same dataset by one of our previous works.

Key words: Ranklets, Wavelets, Support Vector Machine, Pattern Classification, Computer Aided Detection

* Corresponding Author, *E–mail: masotti@bo.infn.it*, *Phone: +39 051 2095136*, *Fax: +39 051 2095047*, *Address: Viale Berti–Pichat 6/2, 40127, Bologna, Italy*, *Web: http://www.bo.infn.it/∼masotti/*

1 Introduction

Breast cancer is one of the most common causes of death among women from all over the world. Its detection and diagnosis at early stage is critical, since primary prevention of cancers is thus far impossible. However, this task still proves really arduous for the radiologists, due to the complexity of breast tissues and similarity between tumoral and normal tissues. Computer Aided Detection (CAD) systems have been expressly introduced in the last years in order to aid the radiologists in the interpretation of mammograms, namely the images produced on film by the X–ray analysis of a woman's breast. CAD systems work as an objective second reader that, by means of the automatic detection of the regions suspected to be tumors, gives a further suggestion to the radiologists. As an example, in Fig. 1 the CAD's mark individuating a suspected region is shown.

The most common lesions associated with the presence of breast tumor are masses. In the mammogram, they appear as thickenings of the breast tissue with size ranging from 3 mm to 30 mm. In order to detect them, the entire mammographic image is scanned at different scales by the CAD system with a resizeable window. Each sub–image scanned by the window—also known as *crop*—is then resized to an image with pixel size 64×64. From each resized crop, some relevant features are extracted, then by means of these features the crop is classified as belonging to the mass class or to the non–mass class by a trained learning machine, namely a Support Vector Machine (SVM). For more details on the whole scanning scheme and on the application of SVM to CAD systems for mammography, see respectively our previous works [1,2].

It is quite clear that, one of the most important steps for the classification problem is the extraction from each crop of a set of suitable features able to distinguish between the two classes. This actually means choosing the image representation that gives the best classification performances. Several works in the literature have addressed this task focusing on the evaluation of texture–based, histogram–based, pixel–based or wavelet–based features. Specifically, in one of our most recent paper [3], some pixel–based and wavelet–based image representations have been evaluated on the same dataset used here. Instead, in this paper, a ranklet–based image representation is proposed. Ranklets have been introduced and applied to face detection for the first time in some recent works [4–6]. They are usually defined as non–parametric, multi–resolution and orientation selective features modelled on Haar wavelets. The non–parametric properties derive from the fact that the ranklet transform is based on the rank transform, a transform that, given $(x_1, x_2, \ldots, x_N)$ pixels, replaces the value of each x_i with the value of its order among all the other pixels. The multi–resolution and orientation selective properties derive from the fact that the ranklet transform is mainly modelled on the bi–dimensional Haar wavelets. This means that, as for the wavelet trasform, it is possible to compute the ranklet transform of each crop at different resolutions by means of a suitable stretch and shift of the Haar supports. At the same time, for each resolution, it is possible to compute the vertical, the horizontal and the diagonal ranklet coefficients.

Several experiments have been carried out, in this paper, using ranklets as features and SVM as classifier. In particular, several SVM's kernels have been evaluated, together with different combinations of resolutions and together with the application of some pre–processing techniques to the original crops, such as histogram equalization. Experiments demonstrate that the proposed

representation is quite effective in solving the two–class classification problem. Furthermore, it achieves an improvement over the pixel–based and wavelet–based image representations evaluated on the same dataset in [3].
The rest of the paper is organized as follows. In Section 2 an overview of the ranklet transform is given. Section 3 provides detailed informations about the dataset used, the features extracted and the classification method adopted. The experiments performed and the results achieved are discussed in Section 4. In Section 5 a critical discussion of the results is carried out. Conclusions are drawn in Section 6.

2 Overview of the ranklet transform

In this Section, an overview of the ranklet transform is given. First, the rank transform, the Wilcoxon test and the Mann–Whitney test are introduced, since they are at the basis of the ranklet transform and are responsible of its non–parametric properties. Second, the ranklet transform and the computation of its orientation selective coefficients is discussed. Finally, the extension of the ranklet transform to the multi–resolution case is described.

2.1 Introduction to some non–parametric statistics

2.1.1 Rank transform

Given a set of $(x_1, x_2, \ldots, x_N)$ pixels, the rank transform $\pi(x_1, x_2, \ldots, x_N)$ substitutes each pixel's intensity value with its relative order (rank) among all the other pixels [7] . Here follows an example:

$$\pi \begin{pmatrix} 55 & 99 & 25 & 153 \\ 26 & 75 & 92 & 200 \\ 21 & 64 & 88 & 154 \\ 101 & 190 & 199 & 222 \end{pmatrix} = \begin{pmatrix} 4 & 9 & 2 & 11 \\ 3 & 6 & 8 & 15 \\ 1 & 5 & 7 & 12 \\ 10 & 13 & 14 & 16 \end{pmatrix} \tag{1}$$

In case the set of $(x_1, x_2, \ldots, x_N)$ pixels contains pixels with equal intensity values, midranks are introduced. Midranks are computed assigning to each group of pixels with equal intensity values the average of the ranks they occupy. For example:

$$\pi \begin{pmatrix} 55 & 99 & \mathbf{25} & 153 \\ \mathbf{25} & \mathbf{64} & 92 & 200 \\ 21 & \mathbf{64} & \mathbf{64} & 154 \\ 101 & 190 & 199 & 222 \end{pmatrix} = \begin{pmatrix} 4 & 9 & \mathbf{2.5} & 11 \\ \mathbf{2.5} & \mathbf{6} & 8 & 15 \\ 1 & \mathbf{6} & \mathbf{6} & 12 \\ 10 & 13 & 14 & 16 \end{pmatrix} \tag{2}$$

2.1.2 Wilcoxon test

The rank transform is closely related to the Wilcoxon test. Given a set of $(x_1, x_2, \ldots, x_N)$ pixels, they are split into the two subsets T and C, with n and m pixels each, so that $n + m = N$. In order to state whether the n pixels in T have significantly higher intensity values than the m pixels in C, the Wilcoxon test W_S is introduced [8] and defined as the sum of the n ranks:

$$W_S = \sum_{i=1}^{n} \pi(x_i) \tag{3}$$

The n pixels in T are then judged to have significantly higher intensity values than the m pixels in C if the Wilcoxon test is above a critical value τ, in other words $W_S > \tau$. The value of τ determines the confidence level of the test.

2.1.3 Mann–Whitney test

In order to introduce a test equivalent to the Wilcoxon test, but with an immediate interpretation in terms of pixels comparison, the Mann–Whitney test W_{XY} is introduced [8]:

$$W_{XY} = W_S - \frac{n(n+1)}{2} \tag{4}$$

As can be easily demonstrated, the value of the Mann–Whitney test W_{XY} is equal to the number of pixel pairs $(\vec{x}_p, \vec{y}_q)$, with $\vec{x}_p \in T$ and $\vec{y}_q \in C$, such that the intensity value of $\vec{x}_p$ is higher than the intensity value of $\vec{y}_q$. Therefore, its values range from 0 to the number of pairs $(\vec{x}_p, \vec{y}_q) \in T \times C$, which is mn. Notice, however, that in order to compute the value of W_{XY}, these pairwise comparisons are never carried out explicitly. This, in fact, would results in a huge computational time. Instead, its value is obtained by the application of the rank transform to the set of pixels $(x_1, x_2, \ldots, x_N)$, thus leading to only $NLogN$ operations.

2.2 The ranklet transform

2.2.1 Haar wavelet supports

As already discussed in Section 1, the non–parametric properties of the ranklet transform derive from the fact that it is based on non–parametric statistics such as the rank transform, the Wilcoxon test and the Mann–Whitney test. Similarly, the orientation selective properties of the ranklet transform derive from the fact that it is manly modelled on Haar wavelets. Thus, in order to arrive at the ranklet transform definition, the first step consists in introducing the Haar wavelet supports.

Suppose that an image constituted by a set of $(x_1, x_2, \ldots, x_N)$ pixels is given. A possible choice in splitting the N pixels, in order to compute the Mann–Whitney test, 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, as shown in Fig. 2. In particular, for the vertical Haar wavelet support, represented by the image h_V, the two subsets T_V and C_V are defined. Similarly, for the horizontal Haar wavelet support h_H, the two subsets T_H and C_H are defined, whereas for the diagonal Haar wavelet support h_D, the two subsets T_D and C_D are defined.

Notice that, the arbitrariness that characterize the selection of the two subsets T and C, is fundamental in order to be able to freely choose the two subsets based on the Haar wavelet supports. In other words, this arbitrariness is at the basis of the orientation selective properties of the ranklet transform.

2.2.2 Ranklet coefficients

Once the rank transform, the Wilcoxon test, the Mann–Whitney test and the Haar wavelet supports have been introduced, the definition of the ranklet transform is straightforward. In fact, given an image constituted by a set of $(x_1, x_2, \ldots, x_N)$ pixels, it is possible to compute the horizontal, vertical and diagonal ranklet coefficients in the following way:

$$R_j = \frac{W_{XY}^j}{mn/2} - 1, \quad j = V, H, D \tag{5}$$

where 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 for the Haar wavelet supports.

The geometric interpretation of the ranklet coefficients R_j, with $j = V, H, D$, is quite simple, see Fig. 3. 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. Then R_V will be close to +1, as many pixels in T_V will have higher intensity values than the pixels in C_V. Conversely, R_V will be close to -1 if the dark and brigth side are reversed. At the same time, horizontal edges or other patterns with no global left–right variation of intensity will give a value close to 0. Analogous considerations could be drawn for the other ranklet coefficients, R_H and R_D.

2.3 The multi–resolution ranklet transform

The close correspondence between the Haar wavelet transform and the ranklet transform leads directly to the extension of the latter to its multi–resolution

formulation. Similarly to what is usually done for the Haar wavelet transform, the ranklet coefficients at different resolutions are computed simply stretching and shifting the Haar wavelet supports. This means that the multi–resolution ranklet transform of an image is a set of triplets of vertical, horizontal and diagonal ranklet coefficients, each one corresponding to a specific resolution and shift of the Haar wavelet supports.

For example, suppose that the multi–resolution ranklet transform of an image with pixel size 16×16 is performed at resolutions 16, 4 and 2 pixels, as shown in Fig. 4. This actually means that the ranklet transform of the image is computed at resolution 16 pixels, by shifting the Haar wavelet support with linear dimensions 16 pixels, at resolution 4 pixels, by shifting that with linear dimensions 4 pixels and at resolution 2 pixels, by shifting that with linear dimensions 2 pixels. 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. Then the multi–resolution ranklet transform of the image is composed by 1 triplet $R_{V,H,D}$ of ranklet coefficients deriving from the ranklet transform at resolution 16 pixels, 25 triplets $R_{V,H,D}$ from that at resolution 4 pixels and 49 triplets $R_{V,H,D}$ from that at resolution 2 pixels. Notice that, the number nT of triplets $R_{V,H,D}$ at each resolution is computed as:

$$nT = (I + 1 - S)^2 \tag{6}$$

where I and S represent the linear dimensions respectively of the image and of the Haar wavelet support, as shown in Fig. 5.

3 The implemented method

In this Section, some informations about the materials and methods adopted to test the ranklet–based image representation are given. First, the dataset used is described. Second, the usage of ranklets as classification features is discussed. Third, an overview of SVM together with some details concerning the classification strategy adopted—a 10–fold cross validation procedure—are given.

3.1 Dataset

The crops used to evaluate the ranklet–based image representation have been extracted—and then resized to 64×64—from the mammographic images of the Digital Database for Screening Mammography (DDSM). This database has been collected by the University of South Florida (USF) and is composed of images digitized with Lumisys scanner at 50 μm or Howtek scanner at 43.5 μm pixel size, with a 12–bit gray–level resolution. For detailed informations about the DDSM mammographic image database, see [9].

The total number of crops used amounts to 6000 and is partitioned in 1000 crops representing the mass class and 5000 crops representing the non–mass class. Notice that the crops used in this paper are exactly the same used to evaluate the pixel–based and wavelet–based image representations in our previous work [3]. This is important in order to be able to directly compare the ranklet–based image representation performances to those obtained, on the same dataset, by means of the pixel–based and wavelet–based image representations.

3.2 Ranklet coefficients as classification features

As already discussed, the main purpose of this paper is to understand whether the non–parametric, multi–resolution and orientation selective properties of the ranklet transform could be exploited in order to improve the performance obtained for this two–class classification problem. The basic idea is to use the ranklet coefficients, derived from the application of the ranklet transform to the mass crops and to the non–mass crops, as classification features. To this purpose, first the multi–resolution ranklet transform of each crop is performed at different resolutions by stretching and shifting the Haar wavelet supports. Then, each crop is presented to the classifier—an SVM—as a collection of several ranklet triplets $R_{V,H,D}$, each one corresponding to a specific stretch and shift of the Haar wavelet supports.

In order to speed up the computational time of the multi–resolution ranklet transform, the 6000 crops constituting the dataset are required to be resized from a 64×64 to a 16×16 pixel size by means of a bilinear resizing, as shown in Fig. 6. In this way, it is possible to compute the multi–resolution ranklet transform of a crop at several resolutions, up to the highest ones, in a reasonable time. Just to give an idea of the number of classification features involved in the problem, Tab. 1 shows the correspondence between the resolutions at which the multi–resolution ranklet transform is performed and the number of ranklet coefficients computed. For example, the multi–resolution ranklet transform of a crop with pixel size 16×16 at resolutions [16,8,4,2] pixels results in 1 triplet $R_{V,H,D}$ from the resolution at 16 pixels, 81 triplets $R_{V,H,D}$ from the resolution at 8 pixels, 169 triplets $R_{V,H,D}$ from the resolution at 4 pixels and 225 triplets $R_{V,H,D}$ from the resolution at 2 pixels,

thus for a total of $3 \times (1 + 81 + 169 + 225) = 1428$ ranklet coefficients. Notice that, the lower is the linear dimension of the Haar wavelet support, the higher is the resolution at which the multi–resolution ranklet transform is performed, the higher is the number of ranklet coefficients produced. And viceversa. This is consistent with the expression discussed in (6).

3.3 Classification

3.3.1 Support Vector Machine

As anticipated in the rest of the paper, an SVM has been chosen as classifier. SVM constructs a binary classifier from a set of l training examples, consisting of labeled patterns $(\mathbf{x_i}, y_i) \in \mathbf{R}^N \times \{\pm 1\}, i = 1, \ldots, l$, see [10,11]. The classifier aims to estimate a function $f : \mathbf{R}^N \rightarrow \pm 1$, from a given class of functions, such that f will correctly classify unseen test examples $(\mathbf{x}, y)$. An example is assigned to the class $+1$ if $f(x) \geq 0$ and to the class -1 otherwise.

SVM selects hyperplanes in order to separate the two classes. Among all the separating hyperplanes, SVM finds the one that causes the largest separation among the decision function values for the borderline examples of the two classes. The Maximal Margin Hyperplane (MMH) is computed as a decision surface of the form:

$$f(\mathbf{x}) = sgn\left(\sum_{i=1}^{l} y_i \alpha_i (\mathbf{x} \cdot \mathbf{x}_i) + b\right) \tag{7}$$

where the coefficients α_i and b are calculated by solving the following quadratic programming problem:

$$\begin{cases} \text{maximize} & \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j) y_i y_j \\ \text{with} & \sum_{i=1}^{l} \alpha_i y_i = 0 \qquad\qquad 0 \leq \alpha_i \leq C \end{cases} \tag{8}$$

C is a regularization parameter, selected by the user. The classification of a pattern $\mathbf{x}$ is therefore achieved according to the values of $f(\mathbf{x})$ in (7). It is worth mentioning that in a typical classification problem the hyperplane (7) is determined only by a small fraction of training examples. These vectors, named *support vectors*, are those with a distance from the MMH equal to half the margin.

In the more general case in which the data are not linearly separable in the input space, a non–linear transformation $\phi(\mathbf{x})$ is used to map the input vectors into a high–dimensional space. The product $K(\mathbf{x_i}, \mathbf{x_j}) \equiv \phi(\mathbf{x_i}) \cdot \phi(\mathbf{x_j})$ is called kernel function. Admissible and typical kernels are:

$$\begin{cases} K(\mathbf{x_i}, \mathbf{x_j}) = \mathbf{x_i}^T \mathbf{x_j} & \text{Linear Kernel} \\ K(\mathbf{x_i}, \mathbf{x_j}) = (\gamma \mathbf{x_i}^T \mathbf{x_j} + r)^d, \gamma > 0 & \text{Polynomial Kernel} \\ K(\mathbf{x_i}, \mathbf{x_j}) = \exp(-\gamma \|\mathbf{x_i} - \mathbf{x_j}\|^2), \gamma > 0 & \text{RBF Kernel} \\ K(\mathbf{x_i}, \mathbf{x_j}) = \tanh(\gamma \mathbf{x_i}^T \mathbf{x_j} + r) & \text{Sigmoid Kernel} \end{cases} \tag{9}$$

where γ, r and d are kernel parameters.

3.3.2 Cross–validation

One of the most common problems one has to face, when dealing with a two–class classification problem, is the lack of samples to train and test the classifier. Cross–validation is a common procedure used to train and test a classifier when the dimensionality of the dataset is limited [12]. Given a n–dimensional dataset D, first the entire dataset is divided in f homogeneus sub–datasets, also known as *folds*, $F_1, F_2, \ldots, F_f$. Second, the classifier is trained with the collection of the first $f-1$ folds, $F_1, F_2, \ldots, F_{f-1}$, then is tested on F_f, the fold left over. The procedure is then permuted for each F_i, $i = 1, \ldots, f-1$. As discussed in Section 3.1, the dataset used in this work is composed of 1000 crops representing the mass class and 5000 crops representing the non–mass class. Due to the restricted number of crops, a 10–folds cross–validation procedure is implemented, therefore the dataset is divided into 10 folds, each one containing 100 mass crops and 500 non–mass crops. In this way, for each permutation of the cross–validation procedure, SVM is trained with 900 mass crops and 4500 non–mass crops, then is tested on 100 mass crops and 500 non–mass crops.

4 Experiments and results

In this Section, details concerning the experiments performed and the results obtained are given. First, a concise introduction to the Receiver Operating Characteristic (ROC) curve analysis—and how it is used in this paper in order to give the classification results—is delineated. Then, the three main experiments performed are discussed.

4.1 Receiver operating characteristic curve

The ROC curve analysis is a widely employed method in order to evaluate the performance of a classifier used to separate two classes, as discussed in [13]. It is a plot of the classifier's True Positive Fraction (TPF) versus its False Positive Fraction (FPF). The quantity TPF is generally known as the system *sensitivity*, whereas the quantity 1 – FPF as the system *specificity*.
In this paper, the quantity FPF is represented by the fraction of non–masses incorrectly classified as belonging to the mass class, whereas the quantity TPF by the fraction of masses correctly classified as belonging to the mass class. The performances are compared using ROC curves generated by moving the hyperplane of the SVM solution. This is achieved by changing the threshold b introduced in (7). The fraction of true positives and false negatives for each choice of b are computed. Each single point of the ROC curves is then obtained by averaging the results of a 10–folds cross–validation technique applied to the entire dataset.

4.2 The tests performed

In order to test the performances of the ranklet–based image representation, three experiments are carried out. First, using as image representation the ranklet coefficients resulting from the multi–resolution ranklet transform at resolutions [16,8,4,2] pixels, several SVM kernels are varied. Second, using an SVM polynomial kernel with degree 3, different ranklet–based representations are tested by varying the resolutions at which the multi–resolution ranklet

transform is performed. Third, the application of histogram equalization to the crops, before the ranklet transform is performed, is evaluated.

The results obtained by means of the ranklet–based image representation are compared with those obtained by the two most performing representations evaluated and discussed in our previous work [3]. The first one is a pixel–based representation in combination with histogram equalization, resizing and scaling techniques. This means that the original crops with pixel size 64×64 are first treated with histogram equalization, then resized to 16×16 and finally their correspondent pixels are scaled in the range [0,1]. The second one is a wavelet–based image representation and—specifically—it is based on the Overcomplete Wavelet Transform (OWT), see [14]. In this case, the original crops with pixel size 64×64 are decomposed by OWT using the Haar wavelet filters and retaining only the wavelet coefficients corresponding to the levels 4 and 6. For the sake of simplicity and in order to use a notation coherent with that used in [3], these two image representations will be respectively reffered to as *PixHRS* and *Owt2*. Notice that, in the first case the number of classification features is equal to the number of pixels, namely $16 \times 16 = 256$, whereas in the second case is equal to the number of wavelet coefficients corresponding to the levels 4 and 6, thus about 3000.

4.2.1 Tests varying kernels

The first test is intended to understand the influence of the SVM kernel on the classification performances. To this purpose, the results obtained with the image representations *PixHRS* and *Owt2*, discussed in [3], are compared to those obtained using as classification features the ranklet coefficients produced by the multi–resolution ranklet transform at resolutions [16,8,4,2] pixels. In

particular, fixed the ranklet–based image representation in this way, different SVM kernels are evaluated, namely the polynomial kernels with degree 1, 2, 3 and 4.

Looking at Fig. 7, where the comparison is reported, the ranklet–based image representation seems to improve its performances with increasing values of the polynomial degree. In particular, while the linear SVM kernel has really poor performances, the polynomial SVM kernels with degrees 2, 3 and 4 give better results. What is particularly worth mentioning is that the results obtained by the ranklet coefficients at resolutions [16,8,4,2] pixels, with polynomial degrees 2, 3 and 4, perform better than *PixHRS* and *Owt2*, that were the most performing image representations found in [3].

4.2.2 Tests varying resolutions

The second test is intended to comprehend the effect of the multi–resolution properties of the ranklet transform on the classification performances. In order to investigate this aspect, an SVM polynomial kernel with degree 3 is used, since in the previous test it demonstrates to ensure interesting performances. Then, fixed the SVM kernel, several combinations of different resolutions are evaluated, namely those shown in Tab. 1.

Fig.8 shows the results obtained employing as image representation the ranklet coefficients resulting from the multi–resolution ranklet transform at resolutions [16,14,12,10,8,6,4,2], [16,8,4,2] and [16,8,2] pixels. It is evident from the ROC curve analysis that all these combinations perform quite similarly and that they all perfom better than *PixHRS* and *Owt2*. This result is quite important, since it demonstrates that it is possible to obtain similar performances using 921 classification features, as for the case [16,8,2], as well as 2040 classification

features, as for the case [16,14,12,10,8,6,4,2]. This means that, using the combination of resolutions [16,8,2], it is possible to obtain results rather identical to those obtained—with twice the number of features—by the combination of resolutions [16,14,12,10,8,6,4,2], thus saving a lot of computational time.

Notice that, in the tests discussed above, all the resolutions have been taken into account, as for the [16,14,12,10,8,6,4,2] case, or at least a sampled version of them has been considered, as for the [16,8,4,2] and [16,8,2] cases. In other words, low, intermediate and high resolutions have been all contemplated. Fig.9 shows instead the results obtained employing as image representation the ranklet coefficients resulting from the multi–resolution ranklet transform at resolutions [16,4] and [16,2] pixels, thus ignoring the intermediate resolutions. Looking at the performances, it is evident that they are not essential for classification purposes. In fact, the results obtained for the [16,4] and [16,2] cases are only slightly different from those obtained for the [16,8,4,2] case and they all perform better than the *PixHRS* and *Owt2* image representations. As for the tests discussed above, this result demonstrates that it is possible to obtain quite identical performances using 510 classification features, as for the case [16,4], as well as 1428 classification features, as for the case [16,8,4,2]. As discussed above, this result is worthy, since it means avoiding unnecessary waste of time.

Finally, in Fig.10 the results obtained by using as image representation the ranklet coefficients resulting from the multi–resolution ranklet transform at resolutions [16,14,12,10] and [16,8] pixels are shown. In this case, the high resolutions are ignored. Looking at the performances, it is evident that they are important for classification purposes. In fact, the results achieved by the [16,14,12,10] and [16,8] cases perform worse than those achieved by the [16,8,4,2] case and by *PixHRS* and *Owt2*.

4.2.3 Tests applying histogram equalization

The last test is intended to investigate the influence of histogram equalization on the classification performances. In order to investigate this aspect, the SVM kernel used is a polynomial kernel with degree 3 and the resolutions at which the multi–resolution ranklet transform is performed are [16,8,4,2] pixels, since in the previous tests this configuration demonstrates to be the most stable and to ensure one of the best result. This configuration is compared to an identical configuration, with the only difference that, in the latter, the crops are submitted to histogram equalization before the multi–resolution ranklet transform is applied.

In Fig. 11 the comparison between the two configurations is shown. It is quite evident that the performances achieved by both the cases are almost the same. This result is really important, since it demonstrates that a computational expensive procedure as histogram equalization is generally uneffective—when dealing with ranklet coefficients—in order to improve the classification results.

5 Discussion

The results presented in Section 4.2.1 demonstrate that—when dealing with the ranklet–based image representation—the SVM polynomial kernels with higher degrees are the best performing ones. Namely, the polynomial kernels with degree 2, 3 and 4 achieve quite identical performances, but due to the higher computational time of the latter and to the slightly worse results of the former, that with degree 3 is preferable.

The results discussed in Section 4.2.2 prove that the low and high resolutions at which the multi–resolution ranklet transform is performed are important to

achieve good performances, whereas intermediate resolutions can be ignored without affecting the classification results. These considerations suggest to perform the multi–resolution ranklet transform at resolutions [16,4] or [16,2] pixels, thus ignoring the intermediate resolutions. Or at least to perform it at resolutions [16,8,4,2] or [16,8,2] pixels, thus using a sampled version of all the resolutions. In both cases the main idea is to use a reduced number of ranklet coefficients, namely only those influencing the classification performances.

Finally, the results discussed in Section 4.2.3 show that histogram equalization is quite uneffective in order to improve the classification performances. This result is consistent with the fact that the ranklet transform is based on non–parametric statistics and could prove really useful when dealing with crops whose intensity histograms are highly variable.

In order to give also some quantitative results, other than the ROC curve analysis, the TPF values of some remarkable image representations, obtained for FPF values close to 0.01, 0.02, 0.03, 0.04 and 0.05, are shown in Tab. 2. In particular, the results obtained using *PixHRS* with an SVM linear kernel and by *Owt2* using an SVM polynomial kernel with degree 2 are reported. They are compared to the best results obtained using a ranklet–based image representation, namely using as classification features the ranklet coefficients produced by the multi–resolution ranklet transform at resolutions [16,8,4,2] pixels and an SVM polynomial kernel with degree 3.

6 Conclusions

In this paper, a ranklet–based image representation is first proposed and then compared to *PixHRS* and *Owt2*, respectively, the most performing pixel–based

and wavelet–based image representations previously evaluated in [3] to solve the same two–class classification problem, namely mass classification in digital mammograms. The results obtained demonstrate that the proposed image representation solves succesfully the classification problem. Furthermore, with an accurate choice of the SVM kernel and of the resolutions at which the multi–resolution ranklet transform is performed, it achieves an improvement over the pixel–based and wavelet–based image representations previously tested.

References

[1] R. Campanini, D. Dongiovanni, E. Iampieri, N. Lanconelli, M. Masotti, G. Palermo, A. Riccardi, M. Roffilli, A novel featureless approach to mass detection in digital mammograms based on support vector machines, Physics in Medicine and Biology 49 (6) (2004) 961–975.

[2] A. Bazzani, A. Bevilacqua, D. Bollini, R. Brancaccio, R. Campanini, N. Lanconelli, A. Riccardi, D. Romani, An svm classifier to separate false signals from microcalcifications in digital mammograms, Physics in Medicine and Biology 46 (6) (2001) 1651–1663.

[3] E. Angelini, R. Campanini, E. Iampieri, N. Lanconelli, M. Masotti, M. Roffilli, Testing the performances of image representations for mass classification in digital mammograms, Pattern Recognition, Submitted.

[4] F. Smeraldi, Ranklets: orientation selective non-parametric features applied to face detection, in: Proceedings of the 16th International Conference on Pattern Recognition, Quebec, QC, Vol. 3, 2002, pp. 379–382.

[5] F. Smeraldi, A nonparametric approach to face detection using ranklets, in: Proceedings of the 4th International Conference on Audio and Video-based

Biometric Person Authentication, Guildford, UK, 2003, pp. 351–359.

[6] F. Smeraldi, Ranklets: a complete family of multiscale, orientation selective rank features, Tech. Rep. RR0309-01, Department of Computer Science, Queen Mary, University of London, Mile End Road, London E1 4NS, UK (September 2003).

[7] R. Zabih, J. Woodfill, Non-parametric local transforms for computing visual correspondence, in: Proceedings Of The Third European Conference On Computer Vision (Vol. II), Springer-Verlag New York, Inc., 1994, pp. 151–158.

[8] E. L. Lehmann, Nonparametrics: statistical methods based on ranks, Holden–Day, 1995.

[9] M. Heath, K. W. Bowyer, D. Copans, R. Moore, P. Kegelmeyer, The digital database for screening mammography, Digital Mammography: IWDW2000 5th International Workshop on Digital Mammography (2000) 212–218.
URL `http://marathon.csee.usf.edu/Mammography/Database.html`

[10] V. Vapnik, The Nature of Statistical Learning Theory, Springer Verlag, 1995.

[11] V. Vapnik, Statistical Learning Theory, J. Wiley, 1998.

[12] L. Tarassenko, Guide to Neural Computing Applications, Butterwoth–Heinemann, 1998.

[13] C. E. Metz, Roc methodology in radiologic imaging, Investigative Radiology 21 (1986) 720–733.

[14] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, D. J. Heeger, Shiftable multi–scale transforms, IEEE Transactions on Information Theory 38 (1992) 587–607.

About The Author—Matteo Masotti received the Laurea degree cum laude in Physics from the University of Bologna in 2001. He is going to receive the Ph.D. degree in March 2005, working on the optimization of wavelet techniques for mass detection in mammography. His interests include wavelets, pattern recognition and image registration.

List of Figures

Table 1

Number of ranklet coefficients for each different combination of resolutions.

Resolutions	Number of 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

Table 2

Classification results comparison. TPF values obtained for FPF values close to 0.01, 0.02, 0.03, 0.04 and 0.05, are shown.

	FPF ~ 0.01	FPF ~ 0.02	FPF ~ 0.03	FPF ~ 0.04	FPF ~ 0.05
PixRHS	$.70 \pm .06$	$.77 \pm .07$	$.84 \pm .05$	$.86 \pm .05$	$.89 \pm .03$
Owt2	-	$.75 \pm .05$	$.82 \pm .05$	$.85 \pm .05$	$.87 \pm .05$
Ranklets	$.76 \pm .05$	$.82 \pm .05$	$.87 \pm .05$	$.89 \pm .05$	$.91 \pm .04$

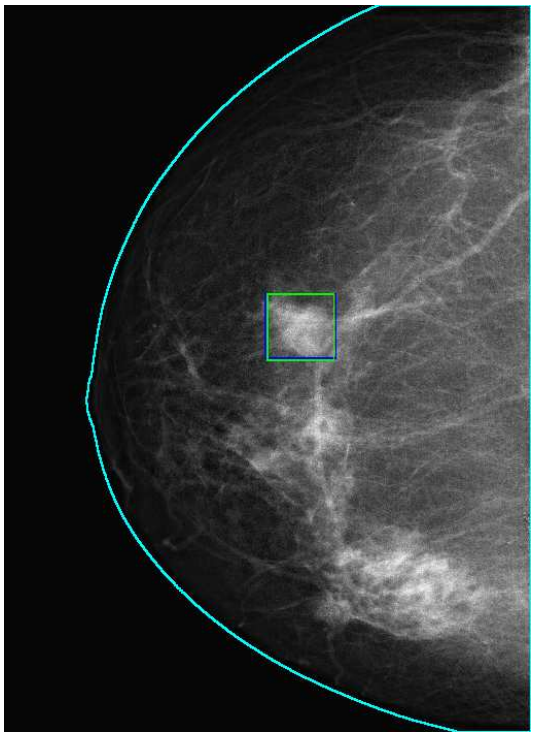


Fig. 1. A mammogram. The square mark is the CAD's automatic detection of a suspected region.

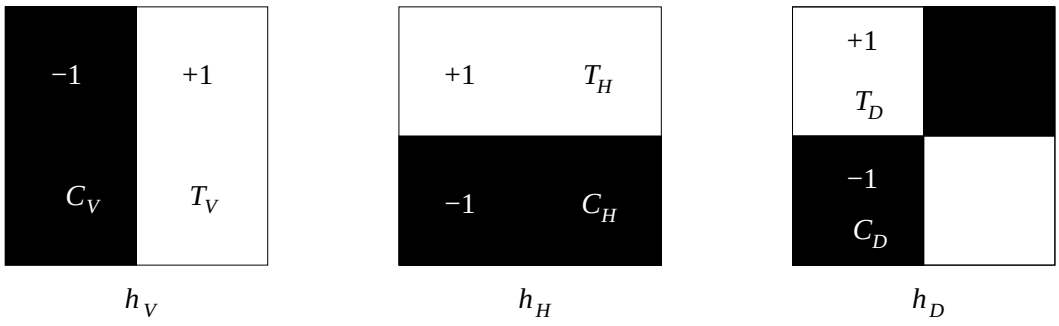
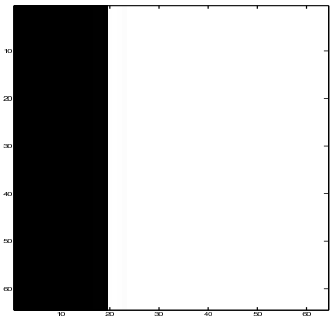
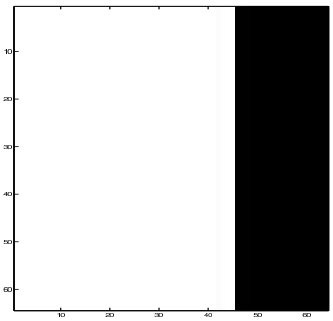


Fig. 2. The three Haar wavelet supports h_V, h_H and h_D. From left to right, the vertical, horizontal and diagonal Haar wavelet supports.

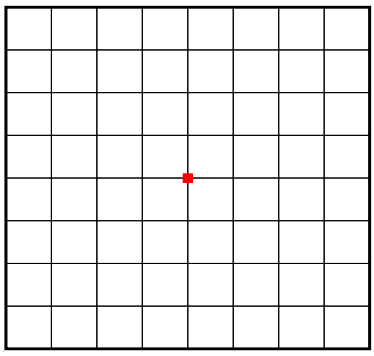


$$\Rightarrow \qquad R_{V,H,D} = [+0.5938, 0, 0]$$

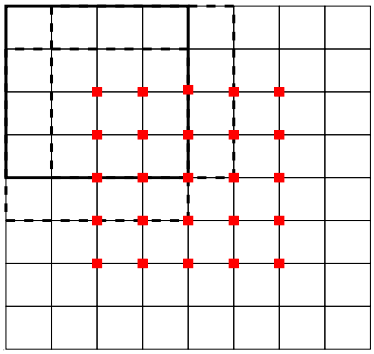


$$\Rightarrow \qquad R_{V,H,D} = [-0.5938, 0, 0]$$

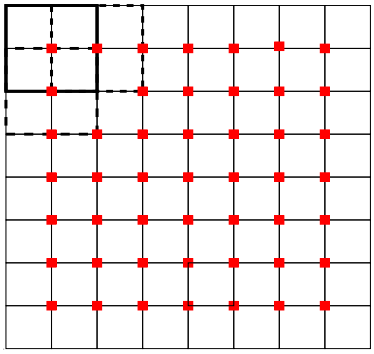
Fig. 3. Ranklet transform applied to some trivial examples.



$\Rightarrow$ 1 triplet $R_{V,H,D}$



$\Rightarrow$ 25 triplets $R_{V,H,D}$



$\Rightarrow$ 49 triplets $R_{V,H,D}$

Fig. 4. Multi–resolution ranklet transform of an image with pixel size 16×16, at resolutions 16, 4 and 2 pixels.

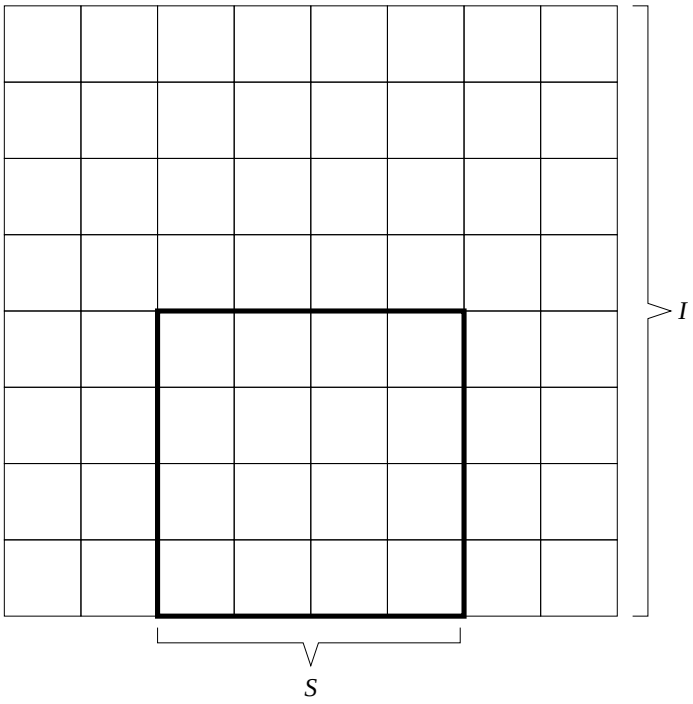


Fig. 5. Linear dimensions I and S respectively of the image and of the Haar wavelet support.

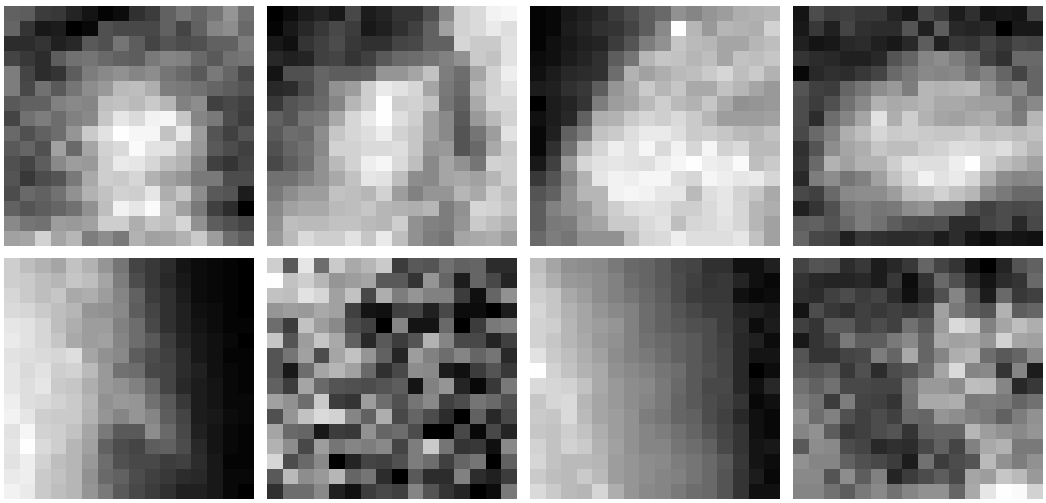


Fig. 6. The two classes after the bilinear resizing from a 64×64 to a 16×16 pixel size. Mass class (top) vs. non–mass class (down).

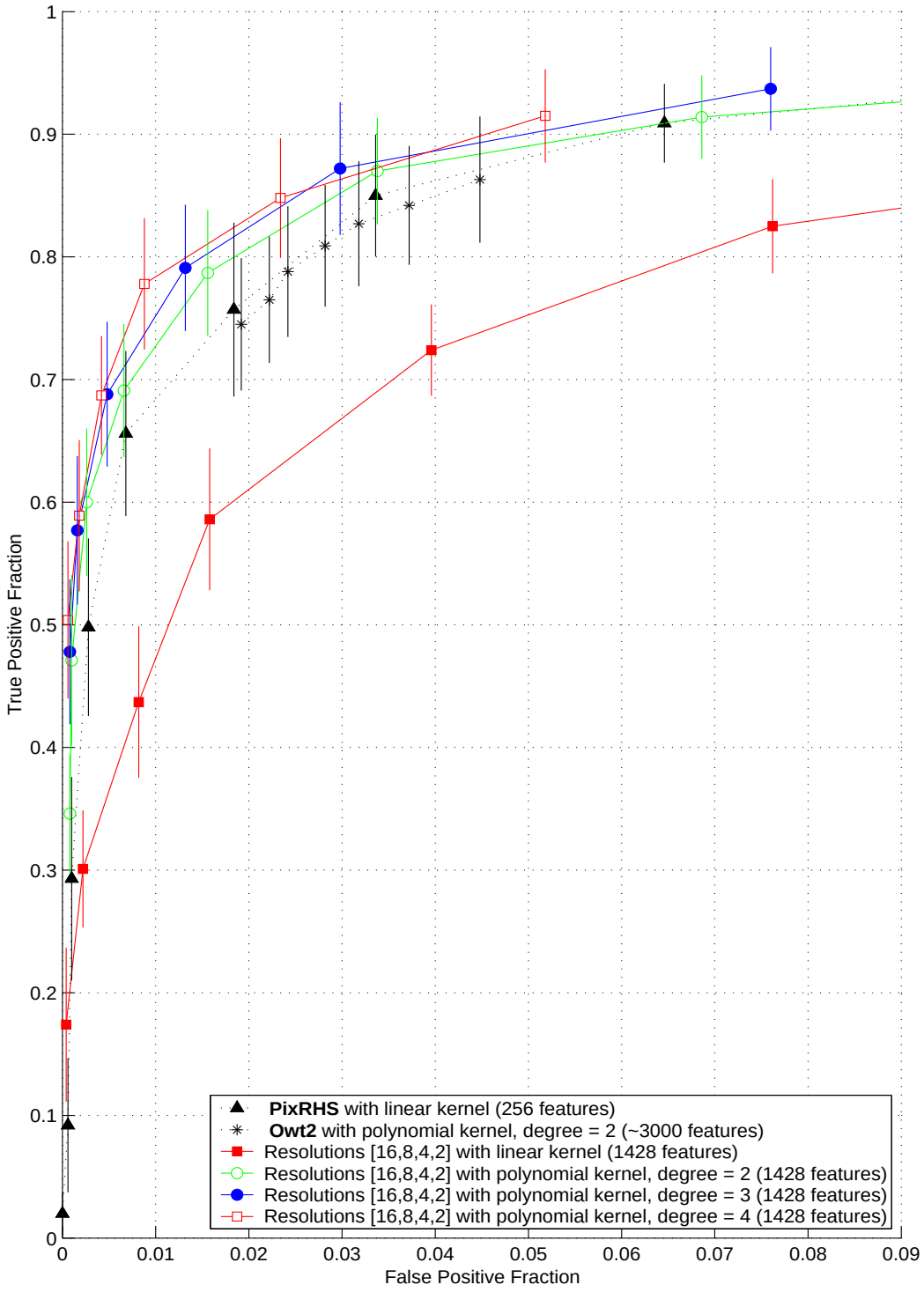


Fig. 7. ROC curves obtained varying SVM kernels.

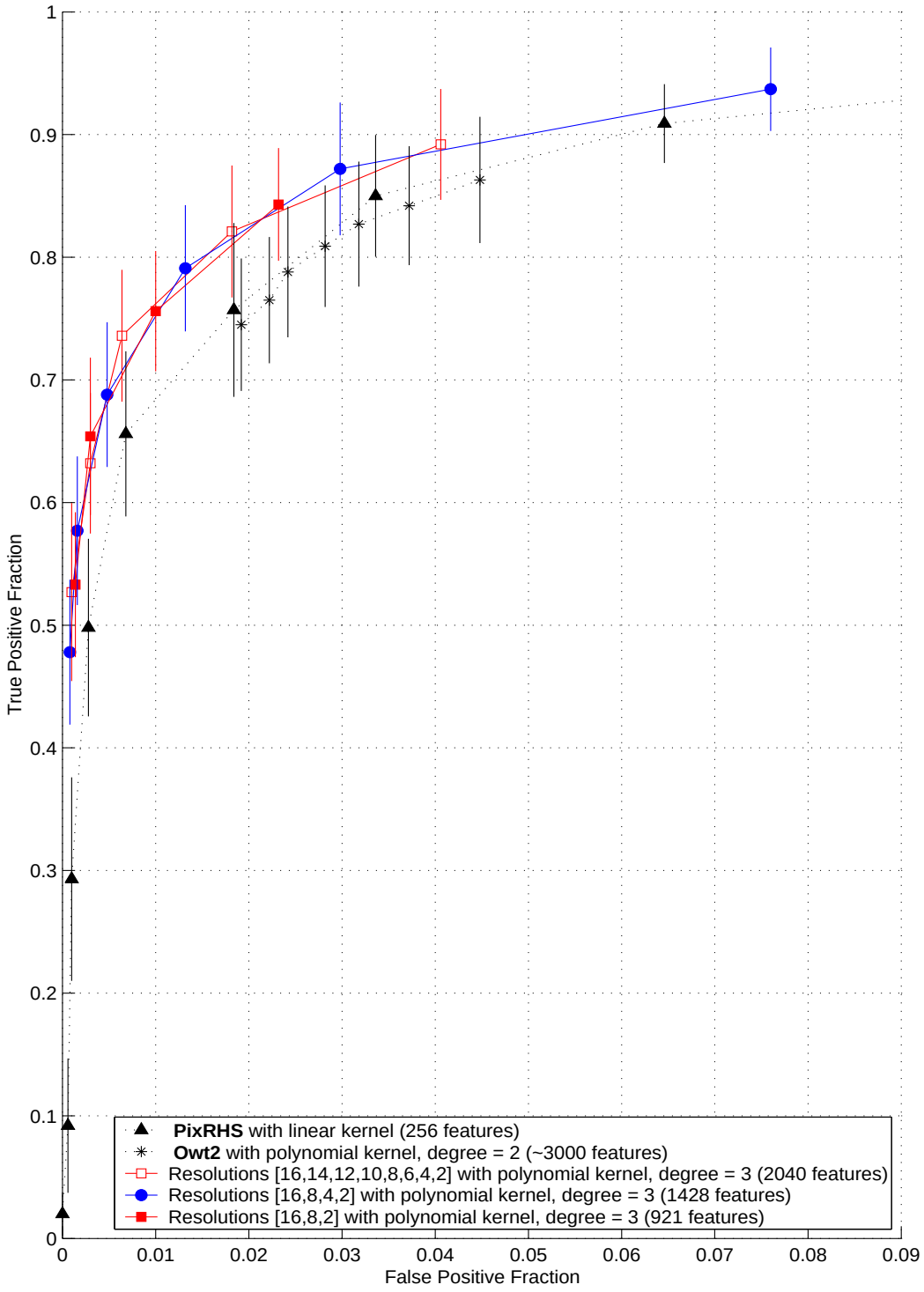


Fig. 8. ROC curves obtained varying the resolutions at which the multi–resolution ranklet transform is performed. Low, intermediate and high resolutions are taken into account.

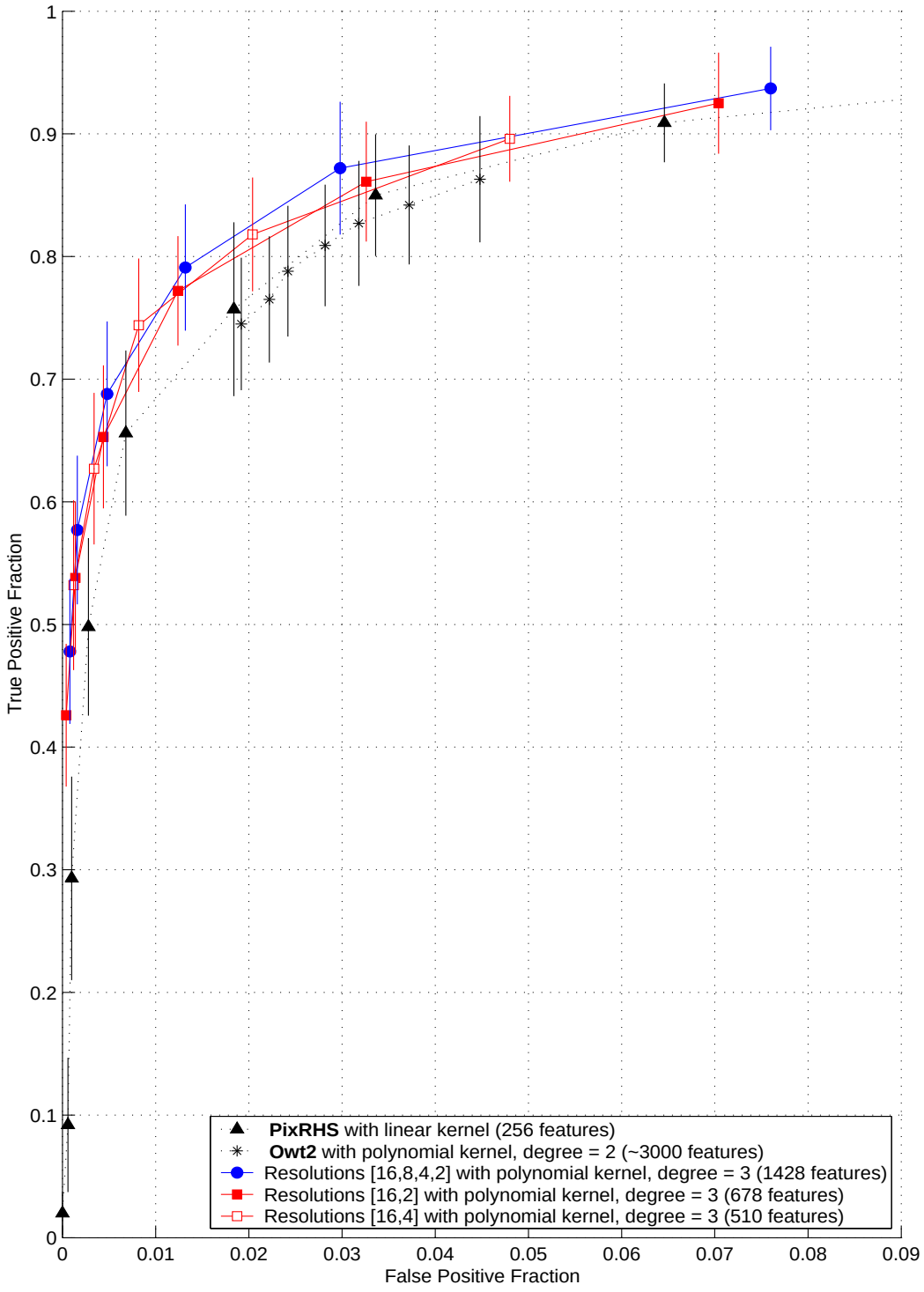


Fig. 9. ROC curves obtained varying the resolutions at which the multi–resolution ranklet transform is performed. Low and high resolutions are taken into account. Intermediate resolutions are ignored.

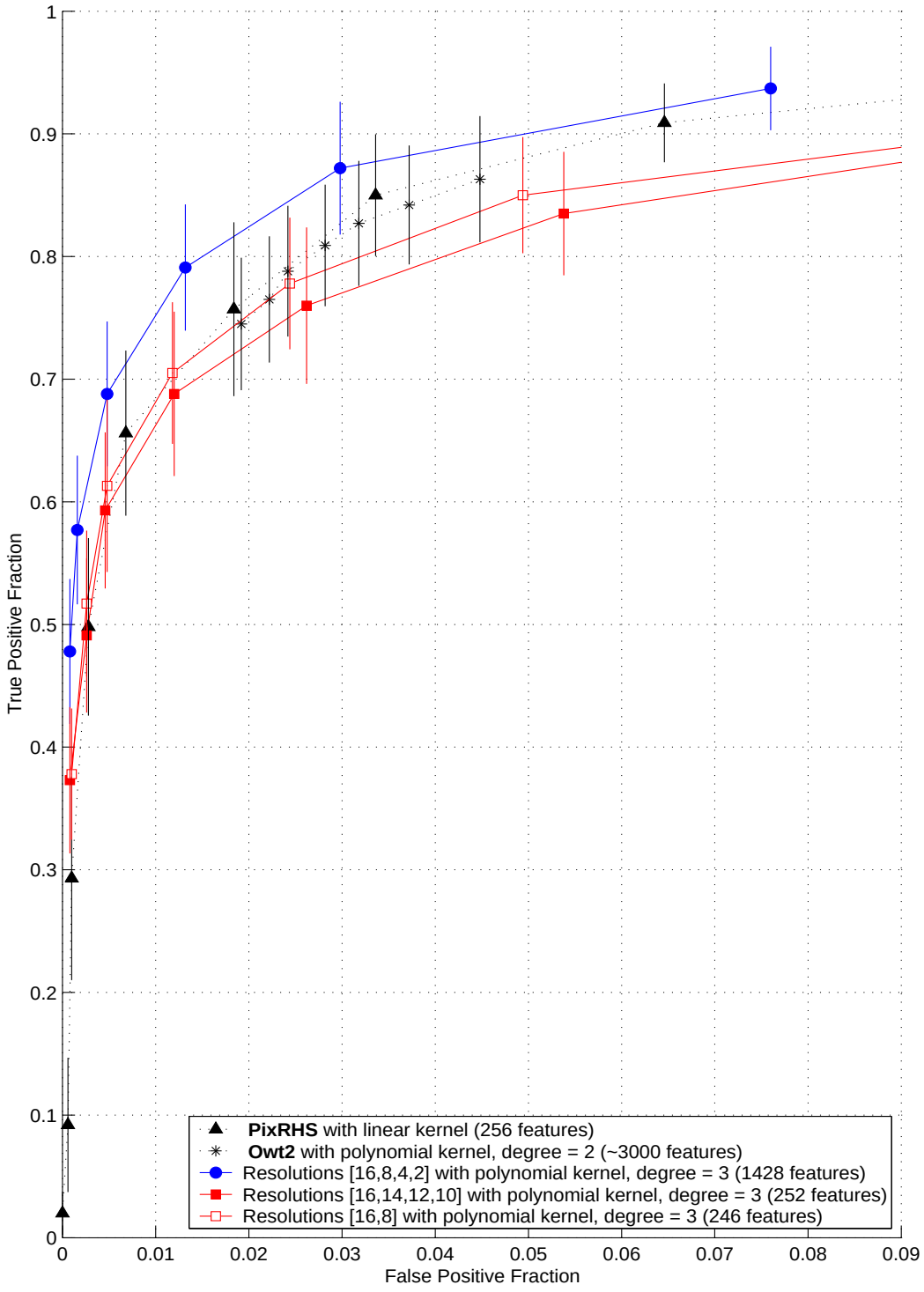


Fig. 10. ROC curves obtained varying the resolutions at which the multi–resolution ranklet transform is performed. Low and intermediate resolutions are taken into account. High resolutions are ignored.

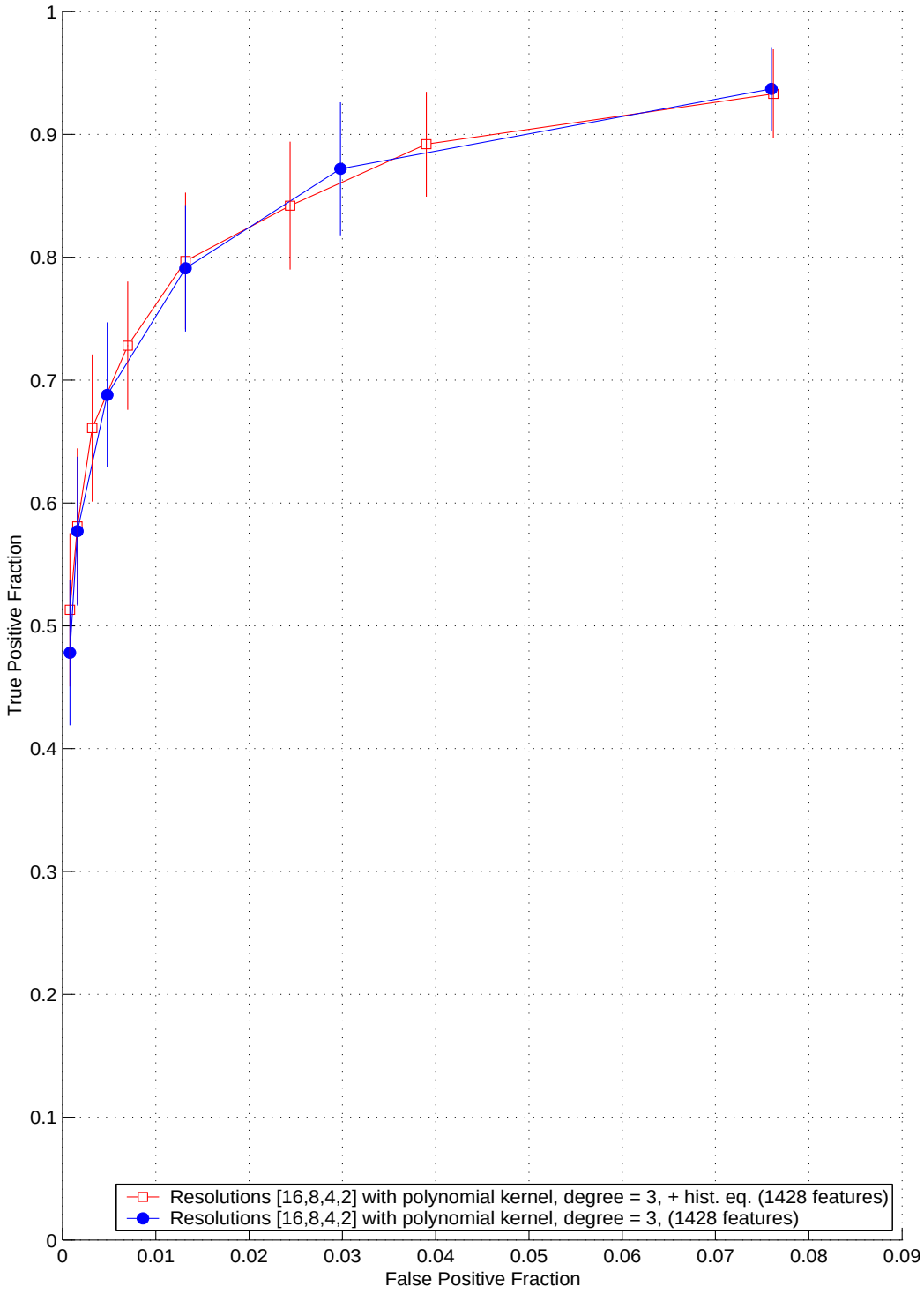


Fig. 11. ROC curves obtained applying histogram equalization.