Discrimination between tumour epithelium and stroma via perception-based features

Francesco Bianconi[a,*], Alberto Álvarez-Larrán[b], Antonio Fernández[c]

[a]*Department of Engineering, Università degli Studi di Perugia, Via G. Duranti 93, 06125 Perugia, Italy*
[b]*Haematology Department, IMIM-Hospital del Mar, Universitat Autònoma de Barcelona, Passeig Marítim 25-29, Barcelona 08003, Spain*
[c]*School of Industrial Engineering, Universidade de Vigo, Campus Universitario Lagoas-Marcosende, 36310 Vigo, Spain*

Abstract

In this work we propose the use of image features based on visual perception for discriminating epithelium and stroma in histological images. In particular, we assess the capability of the following five visual features to correctly discriminate epithelium from stroma in digitised tissue micro-arrays of colorectal cancer: coarseness, contrast, directionality, line-likeliness and roughness. The use of features directly related to human perception makes it possible to evaluate the tissue's appearance on the basis of a set of meaningful parameters; moreover, the number of features used to discriminate epithelium from stroma is very small. In the experiments we used histologically-verified, well-defined images of epithelium and stroma to train three classifiers based on Support Vector Machines (SVM), Nearest Neighbour rule (1-NN) and Naïve Bayes rule (NB.) We optimised SVM's parameters on a validation set, and estimated the accuracy of the three classifiers on a independent test set. The experiments demonstrate that the proposed features can correctly discriminate epithelium from stroma with state-of-the-art accuracy.

Keywords: Perceptual features, Image analysis, Colorectal cancer, Epithelium, Stroma

1. Introduction

Tumour-stroma ratio (TSR) has been recognised as an independent prognostic factor for a number of oncologic diseases. In patients with invasive breast cancer, a high tumor-stroma ratio was shown to correlate with increased hazard for disease relapse

*Corresponding author: Tel.: +39-075-5853706; fax: +39-075-5853703
Email addresses: bianco@ieee.org (Francesco Bianconi), 95967@parcdesalutmar.cat (Alberto Álvarez-Larrán), antfdez@uvigo.es (Antonio Fernández)
URL: http://dismac.dii.unipg.it/bianco (Francesco Bianconi), http://webs.uvigo.es/antfdez (Antonio Fernández)

[1]. In early cervical carcinoma, the disease-free and overall survival were found significantly better in the stroma-poor than in the stroma-rich group [2]. Similar findings have been described in oesophageal squamous cell carcinoma, where stroma-rich tumors were associated with poor prognosis and an increased risk of relapse [3]. Likewise, in non-small cell lung cancer, survival analysis showed that tumour-stroma ratio was significantly correlated with survival [4]. Reliable assessment of tumor-stroma ratio is therefore a key-point to patient stratification and follow-up. Courrech Staal *et al*. [5] investigated the intra- and inter-observer reproducibility of TSR assesment from oesophageal adenocarcinoma biopsies using optical microscopy. In their study they found inter-observer agreement ranging from 81% to 98% when TSR was quantised in two classes (<50% or $\geq$ 50%), but the figures dropped drastically (agreement from 51% to 72%) when TSR was quantised into four classes (<25%, $\geq$ 25% to < 25%, $\geq$ 50% to < 75% or $\geq$ 75).

Computer-assisted classification of tumour epithelium and stroma through digital image processing could be a real possibility to eliminate – or at least reduce – the variability observed among human experts. During the last years, computer-assisted analysis of tissue images has benefited from the steady improvement in imaging technologies as well as from the development of new image descriptors [6, 7, 8]. Among them, Local Binary Patterns (LBP) and variants have received a great deal of attention due to their high discrimination capability, ease of implementation and low computational cost [9, 10, 11]. Linder *et al*. [12] recently proposed a combination of Local Binary Patterns + contrast measure (LBP/C) and linear support vector machine (SVM) for automated identification of tumour epithelium and stroma obtaining strong agreement ($\approx$97%) between the human observer and the computerised approach. A potential drawback of the method, however, is that LBP features are quite difficult to interpret in terms of high-level visual cues, and rather unrelated to the way pathologists perceive and interpret human tissue. Though local binary patterns were originally believed to be related to image micro-structures such as edges, corners, and spots [13] – other studies suggested that this link could be rather weak [14]. As a consequence, LBP-based classification works, to the eye of the physician, as a 'black-box' approach.

In this paper we propose an alternative strategy based on a compact set of perception-based features: *coarseness*, *contrast*, *directionality*, *line-likeliness* and *roughness*. Our approach is inspired on the work of Tamura *et al*. [15], but also represents a significant improvement on their contribution, for we solve some substantial implementation and normalisation issues that are not addressed in the original reference. In addition, we investigate the discrimination power of each feature and the degree of correlation between couples of features. The advantages of the perceptual feature space proposed here are basically two: first, the use of features directly related to human perception makes it possible to assess the tissue's appearance on the basis of a set of values which the pathologist can interpret in a meaningful way; second, the number of required features is very small – we show that state-of-the-art accuracy can be obtained with as few as five features. Consequently, the resulting model provides a very compact description of the phenomenon, reduces the computational complexity of the whole procedure, avoids any potential problems related to the 'curse of dimensionality' [16] and helps the user understand how the model behaves and which features are important (see, for example, Ref. [17] for a discussion on this topic).

Table 1: Number of images and resolution range (in pixels).

	Resolution range	No. of images
Train		
Epithelium	230×230 – 1991×1991	41
Stroma	212×212 – 1355×1355	39
Validation		
Epithelium	172×172 – 1809×1809	395
Stroma	162×161 – 1092×1092	217
Test		
Epithelium	219×219 – 2372×2373	425
Stroma	168×168 – 1690×1691	295

In the remainder of the paper we first present the materials used in the study (Sec. 2) then provide a detailed description of the methods for feature extraction and classification (Sec. 3). The experimental activity is discussed in Sec. 4, followed by the results (Sec. 5) and some concluding considerations (Sec. 6).

2. Materials

This study is based on an image database including 1376 images of tissue samples from patients with colorectal cancer. The dataset is available within the WebMicroscope virtual platform [18] and its use for research, scientific and/or information purposes is expressly permitted[1].

The whole dataset is composed of three groups: train, validation and test. Each contains images representing regions of interest belonging either to tumour epithelium or stroma (see Fig. 1). The proportion of epithelium/stroma samples is 41/39, 395/217 and 425/295 in each of the train, validation and test group, respectively. Image resolution varies from 162×161 to 2372×2373 pixels (see Tab. 1). The tissue samples come from a series of 643 patients with histologically-verified colorectal cancer; further information on the clinico-pathological features of the patients, as well as details about the preparation and digitisation of the tissue microarrays are available in Ref. [19].

3. Methods

3.1. Features

The use of image features corresponding to visual perception was originally proposed by Tamura *et. al* [15]. Based on a set of psychological experiments, they came to define, in their seminal study, six basic textural features, namely: *coarseness*, *contrast*,

[1] http://www.webmicroscope.net/about/disclaimer.asp

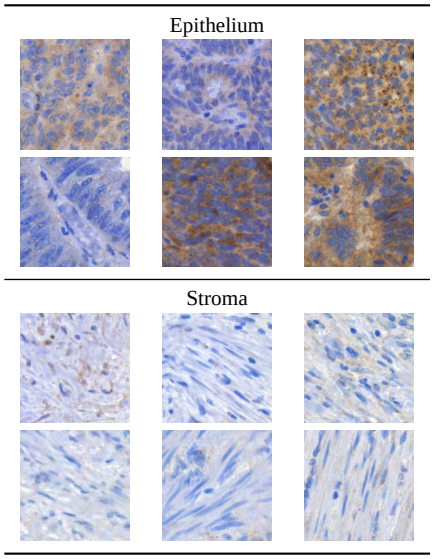


Figure 1: Sample images of tumour epithelium and stroma from Ref. [20].

directionality, *line-likeness*, *regularity* and *roughness*. The practical computation of these features, however, is not completely straightforward: Ref. [15] in fact provides just an outline of how to implement the perceptual features, but leaves many important details to the user. Nor is the matter solved in posterior works [21, 22]. Moreover, in the definitions given in Ref. [15] the output range differs from one feature to another, a condition that is likely to impair the results of any classification strategy based on such features. In the following subsections we discuss our approach to the calculation of each feature. In some cases our implementation departs significantly from the original one. In all cases – and differently from the original definition – our algorithms guarantee that each perceptual feature is represented by a real number in the [0,1] interval. As a result, all the features have the same weight in the classification phase. In the remainder we assume the origin of the image coordinate system is the upper-left pixel with the x and y axes pointing downwards and rightwards, respectively.

3.1.1. Coarseness

The concept of coarseness is related to the intrinsic size of the texture elements: the higher the size, the coarser the texture and vice versa. The computation of this features proceeds as follows. We first apply a set of mean filters to the input image, each filter being defined by a square window of dimension $2^k \times 2^k$, where $k \in \{1, \ldots, K\}$. Values outside the bounds of the image are circularly repeated by implicitly assuming the input is periodic (circular scanning). The selection of a suitable value of K is left to the user – herein we set K = 4. Let $\mathbf{A}_k$ indicate the k-th transformed image resulting from this step.

In the second step we apply, to each $\mathbf{A}_k$, a vertical and a horizontal difference mask

which assign, to each pixel, the difference between the values of the two symmetric pixels that lie vertically or horizontally at distance $2^{(k-1)}$ from the given one. Let the resulting matrices corresponding to the horizontal and vertical directions for each value of k be $\mathbf{E}_{k,h}$ and $\mathbf{E}_{k,v}$, respectively. We now search the value of k that minimises, in each point, the value of $\mathbf{E}(x,y)$ in either directions, i.e.:

$$\bar{k}(x,y) = \arg\max_{k\in\{1,\ldots,K\}} \{\mathbf{E}(x,y)_{k,h}, \mathbf{E}(x,y)_{k,v}\} \tag{1}$$

Finally, we take as coarseness the average windows size that in each point maximises the value of $\mathbf{E}(x,y)$ in either directions:

$$F_{\text{crs}} = \frac{1}{2^K}\frac{1}{WH}\sum_{x=1}^{H}\sum_{y=1}^{W} 2^{\bar{k}(x,y)} \tag{2}$$

where W and H are the dimensions of the input image. Factor $1/2^K$ in Eq. 2 normalises the output in [0,1].

3.1.2. Contrast

According to Tamura *et al*. the concept of 'contrast' depends on the distribution (histogram) of grey-levels, the sharpness of edges and the period of repeating patterns [15]. Following the approach proposed in the cited reference, we estimated this parameter through the following expression:

$$F_{\text{con}} = 2\frac{\sigma}{{\alpha_4}^n} \tag{3}$$

where σ and α_4, are, respectively, the standard deviation and the kurtosis of the distribution of the grey levels. The first reflects the 'dispersion' of the distribution; the second, for many distributions encountered in practice, their 'peakedness' (see Ref. [23, p. 53]). Parameter n is a positive number which we set to $1/4$ as suggested in Ref. [15]; factor 2 normalises the output in [0,1].

3.1.3. Directionality

Directionality is related to the probability that the variation of the pixels' intensities occurs along certain predefined orientations. An image mainly composed by parallel lines will have 'strong' directionality; one made up of almost randomly scattered points will have 'weak' directionality. To estimate this parameter, we first apply a vertical and horizontal 3×3 Sobel filter to compute the image gradient at each point; let us indicate these as $\mathbf{G}_x$ and $\mathbf{G}_y$. Then we compute the gradient orientation at each pixel $\theta(x,y)$ through Eq. 4, discarding those pixels where the magnitude of the gradient is zero:

$$\theta(x,y) = \text{atan}\left[\frac{\mathbf{G}_y(x,y)}{\mathbf{G}_x(x,y)}\right] \tag{4}$$

We quantise the value of θ into N_o discrete values each representing an angular interval of span $2\pi/N_o$. This way we approximate the probability distribution of the gradient orientation through a histogram of N_o bins. An almost 'flat' distribution will be suggestive of a weakly oriented image, whereas a distribution concentrated on just

one or few peaks will indicate a strongly oriented image. To quantify this value we compute the Euclidean distance between the estimated distribution and the uniform distribution, which ideally represents a completely non-directional image: the farther the distance the higher the directionality and vice-versa. In formulas we have:

$$F_{\text{dir}} = \sqrt{\frac{N_o}{N_o - 1} \sum_{i=1}^{N_o} \left[p(\theta_i) - \frac{1}{N_o} \right]^2} \tag{5}$$

where $p(\theta_i)$ is the probability of each orientation θ_i. Factor $N_o/(N_o - 1)$ is the reciprocal of the maximum theoretical distance to the uniform distribution, it therefore normalises F_{dir} in the [0,1] interval. In the experiments we set $N_o = 12$, which correspond to an angular span of 15°. The reader will certainly note the close relationship that directionality bears to the well-known histograms of oriented gradients [24].

3.1.4. Line-likeliness

Line-likeliness reflects the extent to which an image is perceived as composed by lines. This parameter is related to the probability that the gradient maintains the same direction across neighbouring pixels: when the direction remains nearly constant we regard such group of pixels as a line [15]. Therefore, line-likeliness can be evaluated through the joint probability of the gradient directions computed at pixels separated by a given displacement vector: the higher the probability that the two gradient vectors be collinear, the higher the line-likeliness and vice-versa.

Operatively, we can estimate line-likeliness, by computing, as a first step, the image gradient. As a second step we proceed to quantise the gradient direction into a set of discrete levels, just as we do to compute directionality (Sec. 3.1.3). Though in principle it is not required that the number of levels used to compute directionality is the same for line-likeliness, in practice this constraint makes sense, for it guarantees that the two features are computed over the same set of angles. Of the discretised gradient orientations we compute a co-occurrence matrix for a given displacement vector. A matrix with values concentrated on the main diagonal will indicate a high tendency of the gradient to maintain its direction – therefore high line-likeliness; by contrast, a matrix with values concentrated on the corners of the anti-diagonal will be suggestive of low line-likeliness. Numerically, we can express this feature as follows:

$$F_{\text{lin},\mathbf{d}} = \frac{\sum_{i=1}^{N_o} \sum_{j=1}^{N_o} \left| \cos\left[(i-j) \frac{2\pi}{N_o} \right] \right| \mathbf{CM}_{\theta,\mathbf{d}}}{\sum_{i=1}^{N_o} \sum_{j=1}^{N_o} \mathbf{CM}_{\theta,\mathbf{d}}} \tag{6}$$

where $\mathbf{CM}_{\theta,\mathbf{d}}$ is the co-occurrence matrix of the gradient direction for a given displacement vector $\mathbf{d}$. The cosine in Eq. 6 acts as weight which takes value one when the two gradient vectors are collinear, zero when they are orthogonal. To achieve invariance against rotation, $F_{\text{lin},\mathbf{d}}$ is computed at four circularly-oriented vectors and the results averaged:

$$F_{\text{lin}} = \frac{1}{4} \sum_{l=1}^{4} F_{\text{lin},\mathbf{d}_l} \tag{7}$$

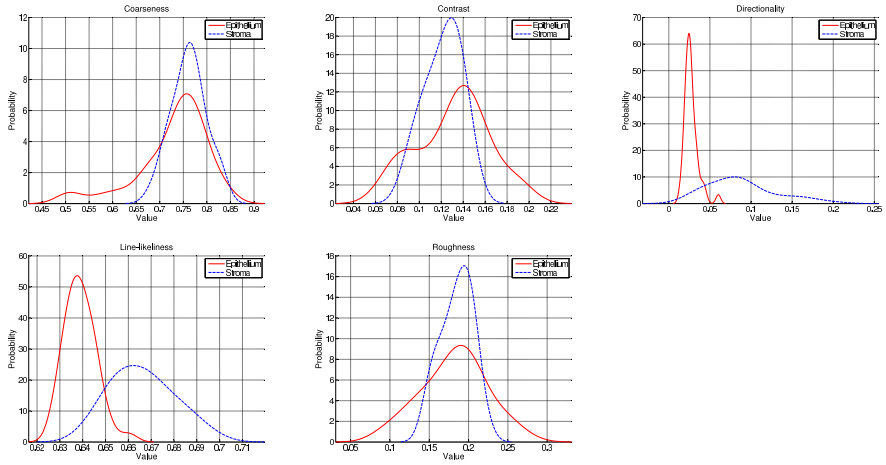


Figure 2: Probability density function of *coarseness*, *contrast*, *directionality*, *line-likeliness* and *roughness* estimated from the train set.

where: $\mathbf{d}_1 = (0, d)$; $\mathbf{d}_2 = (-d, d)$, $\mathbf{d}_3 = (-d, 0)$ and $\mathbf{d}_4 = (-d, -d)$. In the experiments we set $d = 4$, as suggested in Ref. [15]. Note that the implementation proposed here takes the module of the cosine of the angle in Eq. 6; this guarantees that the feature value is in the [0,1] interval.

3.1.5. Roughness

Roughness is generally considered a haptic concept related to the finishing of a surface. With the term 'rough' we mean a surface marked by protuberances, ridges and valleys – as opposed to a 'smooth' one, which is free from such irregularities. When it comes to images, this concept indicates the degree and the rate of variability of the intensity in a neighbourhood: a rough surface will show marked and abrupt changes in intensity; a smooth one will appear mostly even and fine. Tamura *et al.* [15] approximate roughness through the sum of coarseness and contrast as defined in Secs. 3.1.1 and 3.1.2. Such a definition, however, appears weak and rather unconvincing. Moreover, being a mere combination of other features, we expect its independent description power to be rather limited. We therefore adopted a different approach: our definition of roughness is a direct translation into the image domain of the concept of 'surface roughness' as defined by ISO standard 25178-2 [25]. More specifically, the roughness parameter used here is the equivalent of the 'Root mean square height of the scale-limited surface' – S_q. In formulas we have:

$$F_{\text{rgh}} = 2\sqrt{\frac{1}{WH}\sum_{x=1}^{H}\sum_{y=1}^{W}\left[\mathbf{I}(x,y) - R\right]^2} \tag{8}$$

where $\mathbf{I}$ is the input grey-scale image, and R a correction factor that here we set to the average intensity value of the image. Factor 2 normalises the output in the [0,1] range.

3.2. Classification

For classifying each image as either epithelium or stroma we adopted a supervised strategy based on Support Vector Machines (SVM). In a two-class problem – like the one studied herein – SVM determines, in the training step, the hyperplane that best separates the two classes. In the classification step, unlabelled patterns are assigned to one of the two classes by determining which side of the hyperplane they are on. Preliminarily, the training samples are mapped into a higher-dimensional space through a suitable kernel function; this step makes it possible to treat even those cases where the patterns to classify are not linearly separable in the original feature space. Mathematically, the problem can be stated as a minimisation one:

$$\min_{\mathbf{w}, b, \xi} \left(\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{L} \xi_i \right) \tag{9}$$

subject to:

$$\begin{aligned} y_i \left[\mathbf{w}^T \phi(\mathbf{x})_i + b \right] &\geq 1 - \xi_i \\ \xi_i &\geq 0 \end{aligned} \tag{10}$$

where $(\mathbf{x}_i, y_i)$, $i \in \{1, \ldots, L\}$, is a set of labelled patterns, being $\mathbf{x}_i$ the feature vector of the i-th pattern and $y_i \in \{1, -1\}$ the corresponding label. Function $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ is called the kernel function. In this case we used the radial basis formulation:

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2} \tag{11}$$

For the classifier to be completely defined we therefore need to assign suitable values to the parameters C (Eq. 9) and γ (Eq. 11). The first is usually referred to as 'penalisation error', and represents the weight given to misclassified samples in the training step: high values of C result in hyperplanes that strongly avoid misclassification errors; low values in hyperplanes that tolerate misclassification errors. As a consequence C can be viewed as a trade-off between generalising the model and reducing the misclassification errors [26]. The second determines the 'width' of the kernel: the smaller γ, the 'wider' the kernel and vice versa. A wider kernel will result in a smoother decision surface and more regular decision boundary; an excessively narrow kernel may lead to an overfitted classifier. The practical selection of C and γ is discussed in Sec. 4.2. In the remainder we refer to this classification strategy as 'SVM-rbf'. For calibration purposes, we also included, in the experiments, two parameter-free, non-tunable classifiers: nearest neighbourhood (1-NN) with L_2 distance and Naïve Bayes with Gaussian kernel, which in the remainder we indicate as '1-NN' and 'NB', respectively. Further details about these methods can be found in Ref. [27].

4. Experiments

4.1. Feature evaluation and ranking

The objective of this step is to estimate the capability of each perception-based feature to discriminate epithelium from stroma. For an overall appraisal Fig. 2 shows

Table 2: Discrimination properties of each feature.

Feature	FDR	AUC (Linear class.)
Line-Likeliness	3,128	0.953
Directionality	2,031	0.925
Coarseness	0,156	0.579
Contrast	0,067	0.396
Roughness	0,005	0.487

the probability density functions of each feature for the two classes. The densities have been estimated from the train samples using a Gaussian kernel. As class separability measures, we considered Fisher's discriminant ratio (FDR) and the area-under-curve (AUC) returned by a linear classifier (see Tab. 2). In a two-class problem Fisher's discriminant ratio takes the following compact form [28, p. 231]:

$$FDR = \frac{(\mu_e - \mu_s)^2}{\sigma_e^2 + \sigma_e^2} \tag{12}$$

where, μ_e, μ_s, σ_e and σ_s represent the average and standard deviation of any feature for the epithelium and stroma classes, respectively. As for the area-under-curve, this was determined through 100-fold split-half validation with stratified sampling on the train set.

Figure 2 shows that directionality and line-likeliness provide the best separability between epithelium and stroma, a finding confirmed by the values reported in Tab. 2. The probability distributions of the two features seem to correlate rather well with the human perception of the two classes (see Fig. 2). Table 3 shows samples of epithelium and stroma along with the corresponding values of coarseness, contrast, directionality, line-likeliness and roughness. Stroma clearly tends to cluster along lines more than epithelium does. The pdfs accordingly show that higher values of line-likeliness are more likely to occur in stroma than in epithelium. Regarding directionality, the pdfs suggest that high values are more likely with stroma than with epithelium. The appearance of the two classes again confirms this result: stroma's patterns tend to cluster along dominant directions, even though such directions change from image to image and even within the same image. Conversely, epithelium looks more isotropic and shows no preferred orientations. As for coarseness, though the probability distributions of epithelium and stroma largely overlap, the first exhibits a tail towards low values: the examples in Tab. 3 shows the difference between the two classes. Finally, contrast and roughness show probability distributions that largely overlap, a finding that account for the lower discriminative power of these two features (see Tab. 2.)

To better understand how the feature works, Fig. 3 provides the scatter plot and the corresponding correlation value for each couple of features. The plots and the correlation coefficients were computed on the images of the train set. The plots help to interpret the degree of correlation among the features and give an idea of the separation power of each couple of features. We can see a rather strong correlation between

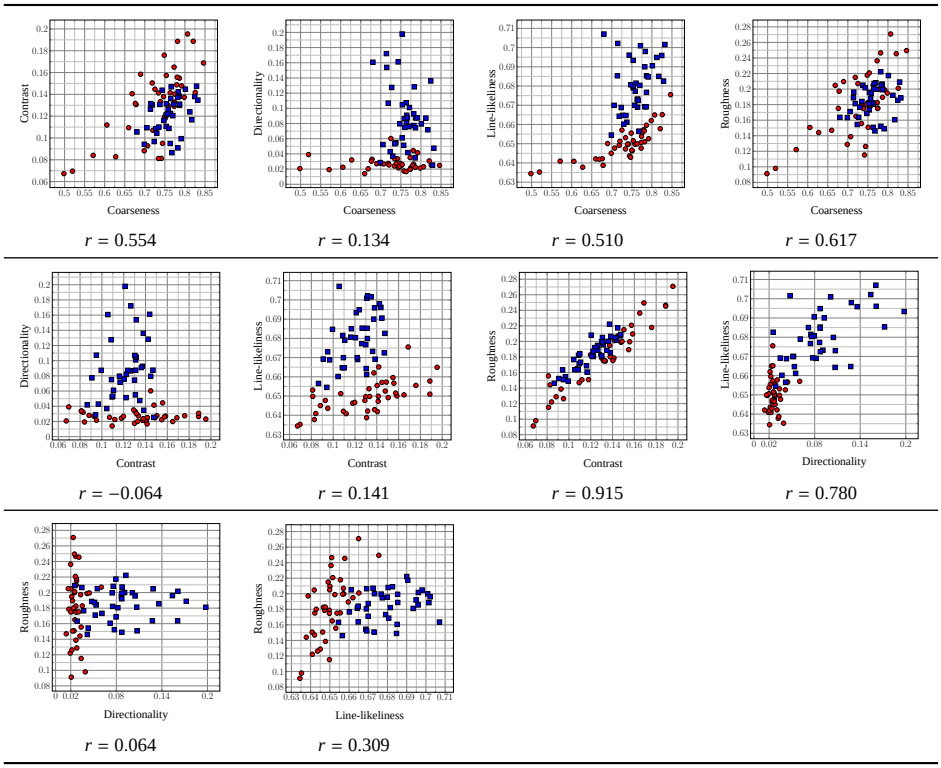


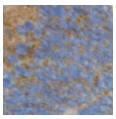
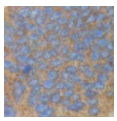
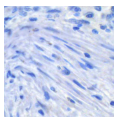
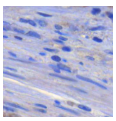
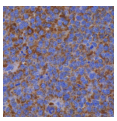
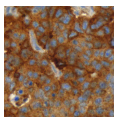
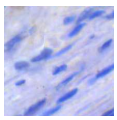
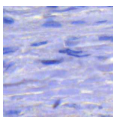
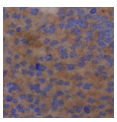
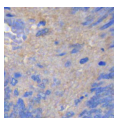
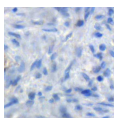
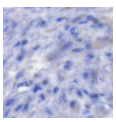
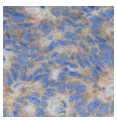
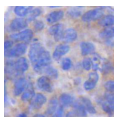
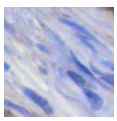
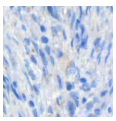
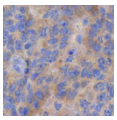
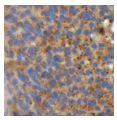
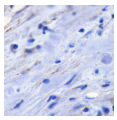
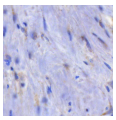
Figure 3: Correlation matrix with scatter plots. Red circles represent samples of epithelium, blue squares samples of stroma; r indicates the correlation coefficient between each couple of features.

contrast and roughness ($r = 0.915$), and a somewhat weaker one between directionality and line-likeliness ($r = 0.780$). Qualitatively we can see that some couples of features, for instance directionality/line-likeliness, coarseness/line-likeliness and contrast/line-likeliness, can actually separate the two classes rather well, a finding that confirms the soundness of the proposed approach.

4.2. Train, validation and test

In the training phase we built one classifier using each of the three classification strategies described in Sec. 3.2, namely: SVM-rbf, 1-NN and NB. In this step we used the same set of 80 images of epithelium (n = 41) and stroma (n = 39) employed in Ref. [12]. In the validation step we determined the optimal values of C and γ for the SVM-rbf classifier. To this end we performed, as suggested in Ref. [29], a grid search with cross-validation over the following set of values: $C = \{2^4, 2^5, \ldots, 2^{15}\}$ and $\gamma = \{2^{-8}, 2^{-7}, \ldots, 2^2\}$. The search returned $C = 2^{11}$ and $\gamma = 2^{-1}$ as the optimal values. Validation was based on the same 576 images of epithelium (n = 359) and stroma (n = 217) provided in Ref. [12]. Note that no validation step was required for the 1-NN and NB classifiers, for these are parameter-free methods. Finally, the effectiveness of

Table 3: Sample images of tumour epithelium and stroma along with the corresponding feature values.

	Epithelium		Stroma	
				
Line-likeliness	0.6353	0.6344	0.7010	0.7022
				
Directionality	0.0140	0.0167	0.1539	0.1979
				
Coarseness	0.6868	0.6340	0.8291	0.8324
				
Contrast	0.1413	0.1403	0.1438	0.1469
				
Roughness	0.1508	0.1787	0.2080	0.1523

each classifier was tested on an independent set of 720 images of epithelium (n = 425) and stroma (n = 295) [12]. As a measure of accuracy we considered the ratio between the number of images of the test set correctly classified and the total number of images of the test set. This parameter can as well be interpreted as the agreement between the classifier and the human observer. In order to evaluate the contribution of each of the five features presented in Sec. 3.1, we started by considering, in the first instance, only the most discriminating feature according to the FDR criterion (see Tab. 2); then we added each feature at a time in descending order of FDR (see Tab. 4).

4.3. Reproducible research

For reproducible research purposes the implementation of the perceptual features and the settings used in the experiments are available online [30][2].

[2]To access the page: user = `perceptual`, psw = `features`

Table 4: Classification accuracy.

Features	No. of feat.	SVM-rbf	1-NN	NB
Line-likeliness	1	85.97	80.56	86.11
Line-likeliness + Directionality	2	93.06	88.89	93.61
Line-likeliness + Directionality + Coarseness	3	96.11	93.33	93.33
Line-likeliness + Directionality + Coarseness + Contrast	4	96.39	92.08	92.92
All	5	**96.94**	92.36	92.50
State-of-the-art (Ref. [12])	224	96.81		

Table 5: Computing time (in seconds).

Features	Extraction	Classification		
		SVM-rbf	1-NN	NB
Line-likeliness	239.9	8.8	< 0.05	< 0.05
Line-likeliness + Directionality	277.1	9.0	< 0.05	< 0.05
Line-likeliness + Directionality + Coarseness	569.1	9.4	< 0.05	< 0.05
Line-likeliness + Directionality + Coarseness + Contrast	624.3	9.7	< 0.05	< 0.05
All	631.6	9.8	< 0.05	< 0.05
State-of-the-art (Ref. [12])	971.4	9.8	0.06	0.06

NOTE: values less that 0.05 seconds have been omitted (they are scarcely significant because mostly related to overhead.)

5. Results and discussion

The results reported in Tab. 4 show that the perception-based features proposed in this work can correctly differentiate with ≈97% accuracy. Of the three classification strategies considered in the experiments, Support Vector Machines with radial basis kernel provided the best results. Yet it is remarkable that also the simple, non-optimised and parameter-free 1-NN and Naïve Bayes gave rather good results, with an overall accuracy approaching 93%. A direct comparison with the results available in the literature [12] shows that the method herein proposed achieves state-of-the-art accuracy with significantly fewer features (i.e.: five vs. 224). Likewise, the mismatch errors reported in (Tab. 6) largely overlap the results reported in Ref. [12, Tab. 1].

Table 5 reports the feature extraction and classification time. The former includes the time required for extracting features from all the images of the train, validation and test sets; the latter that required for training, validating and testing each classifier. All the functions were coded in MATLAB®, and the experiments performed on a laptop PC equipped with INTEL® CORE™ i5- 3230M, 8GB RAM, and WINDOWS™ 7 Professional – 64 bits SP 1. The results show that the whole process is largely dominated by feature extraction, whereas classification accounts only for a small fraction of the total computing time. It is also clear that the method proposed in the manuscript runs markedly faster than the approach described in Ref. [12]. As for classification, we notice there is little difference in the computing time among the various feature sets. This find-

Table 6: Contingency table resulting from using all features and SVM-rbf classifier compared with the results obtained in Ref [12].

		Predicted (our approach)		
		Epithelium	Stroma	
Target	Epithelium	417	8	425
	Stroma	14	281	295
		431	289	720
		Predicted (Ref. [12])		
		Epithelium	Stroma	
Target	Epithelium	406	19	425
	Stroma	4	291	295
		410	310	720

ing most probably depends on the very limited number of classes involved (only two): in problems involving more classes the benefit derived from using a low-dimensional feature-space should become much more noticeable.

6. Conclusions and future work

In this work we evaluated the effectiveness of image features based on visual perception for discriminating epithelium and stroma in colorectal cancer. The results consistently showed high classification accuracy. Compared with the state-of-the-art, our approach offers two major advantages: 1) enables working with a very low-dimensional feature vector, and 2) permits a meaningful interpretation of the feature values in terms of common-sense visual properties. The experiments presented in this paper have been carried out using a set of publicly available images [20] representing well-defined samples of epithelium and stroma from histologically-verified cases of colorectal cancer. The extension of the results of this study to other types of cancer is a possible direction for future research.

Conflict of interest

The authors declare that there is no conflict of interest of any type regarding the material discussed in this paper.

References

[1] T. J. A. Dekker, C. J. H. van de Velde, G. W. van Pelt, J. R. Kroep, J.-P. Julien, V. T. H. B. M. Smit, R. A. E. M. Tollenaar, W. E. Mesker, Prognostic significance of the tumor-stroma ratio: Validation study in node-negative premenopausal breast cancer patients from the EORTC perioperative chemotherapy (pop) trial (10854), Breast Cancer Research and Treatment 139 (2) (2013) 371–379.

[2] J. Liu, J. Liu, J. Li, Y. Chen, X. Guan, X. Wu, C. Hao, Y. Sun, Y. Wang, X. Wang, Tumor-stroma ratio is an independent predictor for survival in early cervical carcinoma, Gynecologic Oncology 132 (1), 81-86.

[3] K. Wang, W. Ma, J. Wang, L. Yu, X. Zhang, Z. Wang, B. Tan, N. Wang, B. Bai, S. Yang, H. Liu, S. Zhu, Y. Cheng, Tumor-stroma ratio is an independent predictor for survival in esophageal squamous cell carcinoma, Journal of Thoracic Oncology 7 (9) (2012) 1457–1461.

[4] Z. Wang, H. Liu, R. Zhao, H. Zhang, C. Liu, Y. Song, Tumor-stroma ratio is an independent prognostic factor of non-small cell lung cancer, Chinese Journal of Lung Cancer 16 (4) (2013) 191–196.

[5] E. F. W. Courrech Staal, V. T. H. B. M. Smit, M.-L. F. van Velthuysen, J. M. J. Spitzer-Naaykens, M. W. J. M. Wouters, W. E. Mesker, R. A. E. M. Tollenaar, J. W. van Sandick, Reproducibility and validation of tumour stroma ratio scoring on oesophageal adenocarcinoma biopsies, European Journal of Cancer 47 (3) (2011) 375–382.

[6] G. Wiselin Jiji, Colour texture classification for human tissue images, Applied Soft Computing 11 (2) (2011) 1623–1630.

[7] V. Bevilacqua, Three-dimensional virtual colonoscopy for automatic polyps detection by artificial neural network approach: New tests on an enlarged cohort of polyps, Neurocomputing 116 (2013) 62–75.

[8] F. Roberti de Siqueira, W. Schwartz, H. Pedrini, Multi-scale gray level co-occurrence matrices for texture description, Neurocomputing 120 (2013) 336–345.

[9] L. Nanni, A. Lumini, S. Brahnam, Local binary patterns variants as texture descriptors for medical image analysis, Artificial Intelligence in Medicine 49 (2) (2010) 117–125.

[10] G. Kylberg, I.-M. Sintorn, Evaluation of noise robustness for local binary pattern descriptors in texture classification, EURASIP Journal on Image and Video Processing 2013 (17).

[11] S. Brahnam, L. Jain, L. Nanni, A. Lumini (Eds.), Local binary patterns: new variants and applications, Vol. 506 of Studies in computational intelligence, Springer, 2014.

[12] N. Linder, J. Konsti, R. Turkki, E. Rahtu, M. Lundin, S. Nordling, C. Haglund, T. Ahonen, M. Pietikäinen, J. Lundin, Identification of tumor epithelium and stroma in tissue microarrays using texture analysis, Diagnostic Pathology 7 (22) (2012) 1–11.

[13] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (7) (2002) 971–987.

[14] F. Bianconi, A. Fernández, On the occurrence probability of local binary patterns: a theoretical study, Journal of Mathematical Imaging and Vision 40 (3) (2011) 259–268. doi:10.1007/s10851-011-0261-7.

[15] H. Tamura, T. Mori, T. Yamawaki, Textural features corresponding to visual perception, IEEE Transactions on Systems, Man, and Cybernetics 8 (1978) 460–473.

[16] C. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

[17] V. Gómez-Verdejo, M. Verleysen, J. Fleury, Information-theoretic feature selection for functional data classification, Neurocomputing 72 (16-18) (2009) 3580–3589.

[18] Webmicroscope – web-based virtual microscopy, available online at http://www.webmicroscope.net/. Last accessed on May 16, 2014 (2014).

[19] N. Linder, E. Martelin, M. Lundin, J. Louhimo, S. Nordling, C. Haglund, J. Lundin, Xanthine oxidoreductase - clinical significance in colorectal cancer and in vitro expression of the protein in human colon cancer cells, European Journal of Cancer 45 (4) (2009) 648–655.

[20] egfr colon stroma classification - WebMicroscope, available online at http://fimm.webmicroscope.net/supplements/epistroma. Last accessed on October 28, 2013 (2012).

[21] T. Deselaers, D. Keysers, H. Ney, Features for image retrieval: An experimental comparison, Information Retrieval 11 (2) (2008) 77–107.

[22] P. Howarth, S. Rüger, Robust texture features for still-image retrieval, IEE Proceedings: Vision, Image and Signal Processing 152 (6) (2005) 868–874.

[23] J. F. Kenney, E. S. Keeping, Mathematics of Statistics. Part two., Van Nostrand, Princeton, USA, 1951.

[24] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1, 2005, pp. 886–893. doi:http://dx.doi.org/10.1109/CVPR.2005.177.

[25] ISO 25178-2: Geometrical product specifications (GPS) – surface texture: Areal – part 2: Terms, definitions and surface texture parameters, International Organization for Standardization (2012).

[26] P. Gaspar, J. Carbonell, J. L. Oliveira, On the parameter optimization of Support Vector Machines for binary classification, Journal of Integrative Bioinformatics 9 (3) (2012) 201–211.

[27] R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification, 2nd Edition, Wiley-Interscience, 2001.

[28] S. Theodoridis, K. Koutroumbas, Pattern Recognition. Third Edition, Academic Press, 2006.

[29] C. Hsu, C. Chang, C. Lin, A practical guide to support vector classification, available online at `http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf`. Last accessed on December 27, 2013 (2010).

[30] Perceptual features, code and related to this paper. Available online at `http://dismac.dii.unipg.it/pbfeatures` (2014).