Tissue and Cell 43 (2011) 318–330

Contents lists available at ScienceDirect

Tissue and Cell

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

Textural characterization of histopathological images for oral sub-mucous fibrosis detection

M. Muthu Rama Krishnan[a], Pratik Shah[b], Anirudh Choudhary[c], Chandan Chakraborty[a,*], Ranjan Rashmi Paul[d], Ajoy K. Ray[e]

[a] School of Medical Science and Technology, Indian Institute of Technology Kharagpur, West Bengal, 721302, India
[b] JFWTC, GE, Research, Bangalore, India
[c] Department of Electrical Engineering, IIT Kharagpur, India
[d] Department of Oral and Maxillofacial Pathology, Guru Nanak Institute of Dental Science and Research, Kolkata, India
[e] Department of Electronics and Electrical Communication Engineering, IIT Kharagpur, India

ARTICLE INFO

Article history:
Received 17 February 2011
Received in revised form 22 June 2011
Accepted 27 June 2011
Available online 6 August 2011

Keywords:
Texture
Wavelet family
Gabor-wavelet
Brownian motion curve
Local binary pattern
Support vector machine
Histopathology
Oral sub-mucous fibrosis

ABSTRACT

In the field of quantitative microscopy, textural information plays a significant role very often in tissue characterization and diagnosis, in addition to morphology and intensity. The aim of this work is to improve the classification accuracy based on textural features for the development of a computer assisted screening of oral sub-mucous fibrosis (OSF). In fact, a systematic approach is introduced in order to grade the histopathological tissue sections into normal, OSF without dysplasia and OSF with dysplasia, which would help the oral onco-pathologists to screen the subjects rapidly. In totality, 71 textural features are extracted from epithelial region of the tissue sections using various wavelet families, Gabor-wavelet, local binary pattern, fractal dimension and Brownian motion curve, followed by preprocessing and segmentation. Wavelet families contribute a common set of 9 features, out of which 8 are significant and other 61 out of 62 obtained from the rest of the extractors are also statistically significant ($p < 0.05$) in discriminating the three stages. Based on mean distance criteria, the best wavelet family (i.e., biorthogonal3.1 (bior3.1)) is selected for classifier design. support vector machine (SVM) is trained by 146 samples based on 69 textural features and its classification accuracy is computed for each of the combinations of wavelet family and rest of the extractors. Finally, it has been investigated that bior3.1 wavelet coefficients leads to higher accuracy (88.38%) in combination with LBP and Gabor wavelet features through three-fold cross validation. Results are shown and discussed in detail. It is shown that combining more than one texture measure instead of using just one might improve the overall accuracy.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years there has been an increase in the cancers of the oral cavity and each year more than 0.3 million new cases (Jadhav et al., 2006) of oral cancer are reported. A high incidence of oral cancer (Mukherjee et al., 2006) is mainly due to late diagnosis of potential precancerous lesions and conditions. Oral cancer develops from preexisting precancerous oral lesions. The common oral precancerous lesions are leukoplakia, erythroplakia, oral sub-mucous fibrosis, etc. The precancerous status is judged on the basis of light microscopic histopathological features of oral epithelial dysplasia (OED) and/or cellular atypia which have different grades according to involvement of the epithelial region. Moreover, dysplasia depicts associated epithelial hyperplasia or even rarely epithelial atrophy (Muthu Rama Krishnan et al., 2009; Prabhu et al., 1992; Chang et al., 2004; Muthu Rama Krishnan et al., 2010c). Interestingly epithelial hyperplasia may be noted in early stage of OSF while epithelial atrophy is consistent with advance stage (Tilakaratne et al., 1997). This hyperplasic or atrophic status may also change architecture of the epithelium which, leads to change the texture of the epithelium.

A means of inspecting histopathological characteristics at a molecular or cellular level is the motivation for the use of microscopic imaging. Although it is an invasive procedure, this modality has the advantage of providing high resolution images exposing the richness or denseness of the examined underlying texture as compared to other non-invasive imaging modalities. It assists in

Abbreviations: OSF, oral submucous fibrosis; LBP, local binary pattern; SVM, support vector machine; OED, oral epithelial dysplasia; FD, fractal dimension; BMC, Brownian motion curve; DWT, discrete wavelet transform; OSFWD, OSF without dysplasia; OSFD, OSF with dysplasia; HE, haematoxylin and eosin; CAD, computer aided diagnostic; AFP, area under first peak; NP, total number of peaks; MDP, mean distance between peaks; ANOVA, analysis of variance.

* Corresponding author. Tel.: +91 3222 283570; fax: +91 3222 28881.
E-mail address: chandanc@smst.iitkgp.ernet.in (C. Chakraborty).

0040-8166/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tice.2011.06.005

giving a better interpretation to histopathological images, by studying the effect of disease on the cellular characteristics of the body tissue. This is done by staining the extracted tissue biopsies with dyes for visual contrast improvement, which will then facilitate the delineation of cell nuclei, giving a better tissue characterization. Despite this modality being invasive, which is unpleasant for patients, physicians usually require a biopsy for a definite answer if they are suspicious about a certain abnormality in an image acquired by a non-invasive imaging modality, and a closer view of the histopathological specimens can assist in verifying the tumour type (Al-Kadi, 2010).

Pathologists have been using microscopic images to study tissue biopsies for a long time, relying on their personal experience on giving decisions about the healthiness state of the examined biopsy. This includes distinguishing normal from abnormal (i.e., cancerous) tissue, benign versus malignant tumours and identifying the level of tumour malignancy. Nevertheless, variability in the reported diagnosis may still occur (Gilles et al., 2008; Grootscholten et al., 2008; Shuttleworth et al., 2005), which could be due to, but not limited to the heterogeneous nature of the diseases; ambiguity caused by nuclei overlapping; noise arising from the staining process of the tissue samples; intra-observer variability, i.e., pathologists are not able to give the same reading of the same image at more than one occasion; and inter-observer variability, i.e., increase in classification variation among pathologists. Therefore, over the past three decades, quantitative techniques have been developed for computer-aided diagnosis, which aims to avoid unnecessary biopsies and assist pathologists in the process of cancer diagnosis (Duncan and Ayache, 2000). Currently, the challenge remains in developing a value addition diagnostic technique that not only automates the conventional procedure, but also characterizes the texture in understanding the underlying pathophysiological changes from normal to OSF.

The biopsy sections are stained with H&E. The optical density of the pixels in the light microscopic images are recorded and represented as matrix quantized as integers from 0 to 255 for each fundamental color (Red, Green, Blue), resulting in a $M \times N \times 3$ matrix of integers. Depending on either normal or OSF condition, the image has various granular structures which are self similar patterns at different scale termed "texture". It refers to the properties in respect to the smoothness, roughness and regularity of any structure (Gonzalez and Woods, 2002; Muthu Rama Krishnan et al., 2011a).

A number of research studies have been done in analyzing histopathological images for cancer detection. Some attempts for texture quantification are based on discrete wavelet transform (DWT) (Lessmann et al., 2007; Qian et al., 2007; Sertel et al., 2008) for early detection of lung cancer and neuroblastoma. In (Ferrari et al., 2001; Wu et al., 1992), Gabor filter based texture analysis is performed on mammography images and liver ultrasound images. Other measures viz., fractal dimension, gray-level co-occurrence matrix (Marghani et al., 2003; Alexandratou et al., 2008) are provided for textural classification. Using more than one measure for classification is applied as well, such as using spatial and frequency texture features for classification by regression trees analysis (Wiltgen et al., 2007). Some used morphological characteristics for feature extraction (Wittke et al., 2007; Thiran and Macq, 1996) and others focused more on classifier improvement (Seker et al., 2003; Estevez et al., 2005). Al-Kadi (2010) proposed a method for texture combination to achieve higher accuracy for meningioma classification of histological images. In (Muthu Rama Krishnan et al., 2010a), fractal dimension is used to show the textural variation over the normal and OSF without dysplasia groups. In (Muthu Rama Krishnan et al., 2010b, 2011a), Brownian motion curve is used to show the textural variation of collagen fibres over normal and OSF groups.

It is observed in the literatures that very few attempts has been made so far in developing textural classifier based on epithelial regions towards the detection of OSF (Muthu Rama Krishnan et al., 2011a). It is clinically understood that the cancerous condition at first progresses over epithelial region, which is like to affect tissue architecture leading to textural changes (Muthu Rama Krishnan et al., 2010a). In view of this, a systematic approach combining multi-scale (Wavelet, Gabor-wavelet and local binary pattern (LBP)), statistical (fractal dimension (FD) and Brownian motion curve (BMC)) techniques are devised in order to provide better understanding of textural changes over normal and OSF tissue sections. The best wavelet family i.e., biorthogonal3.1 is selected based on mean distance criteria and it is combined with other textural features for achieving the highest classification accuracy. Support vector machine (SVM) is trained and validated using *k*-fold cross-validation approach.

2. Materials and methods

Fig. 1 shows the block diagram of the proposed OSF screening system. In general, computer aided diagnosis systems can be constructed with a feature extraction subsystem and a classification subsystem. In this work, we have extracted wavelet, FD, BMC, LBP and Gabor wavelet based features. The significance of the extracted features is evaluated using statistical analysis. The significant features are fed to the SVM classifier. The role of each individual component in the block diagram is described in this section.

2.1. Selection of study subjects

Initially 45 study subjects are clinically diagnosed as OSF. All these patients are properly evaluated from medical and surgical view point and subsequently incisional biopsy are performed under local anesthesia from the buccal mucosa under their informed consent at the Department of Oral and Maxillofacial Pathology, Guru Nanak Institute of Dental Sciences and Research, Kolkata, India. Normal study samples are also collected from the buccal mucosa of 10 healthy volunteers without having any oral habits or any other known systemic diseases with prior written consent. All the above study subjects are of similar age (21–40 years) and food habits. This study is duly approved by the ethics review committee of Guru Nanak Institute of Dental Sciences and Research, Kolkata. All the biopsy samples are processed for histopathological examination and paraffin embedded tissue sections of 5 μm thickness are prepared and then stained by haematoxylin and eosin (H&E) and are evaluated subsequently. Only 24 study subjects from OSF group revealed various grades of epithelial dysplasia as well as without dysplasia.

2.2. Microscopic imaging and acquisition

90, 42, and 26 images of surface epithelium from normal oral mucosa, OSF without dysplasia (OSFWD) and OSF with dysplasia (OSFD) are optically grabbed by Zeiss Observer.Z1 Microscope (Carl Zeiss, Germany) using H&E stained histological sections under 10× objective (N.A 0.25). At a resolution used of 1.33 μm and the pixel size of 0.63 μm, the grabbed digital Images 1388 × 1040 have been stored in PACS.

2.3. Texture based epithelial layer segmentation

The second step of the OSF screening system is epithelial layer segmentation. In this section, we have proposed texture based segmentation. This segmentation combines proper partitioning of pixels within the regions, discontinuities in gray-level, color and texture at the interface as well as the intra-region similarity. The

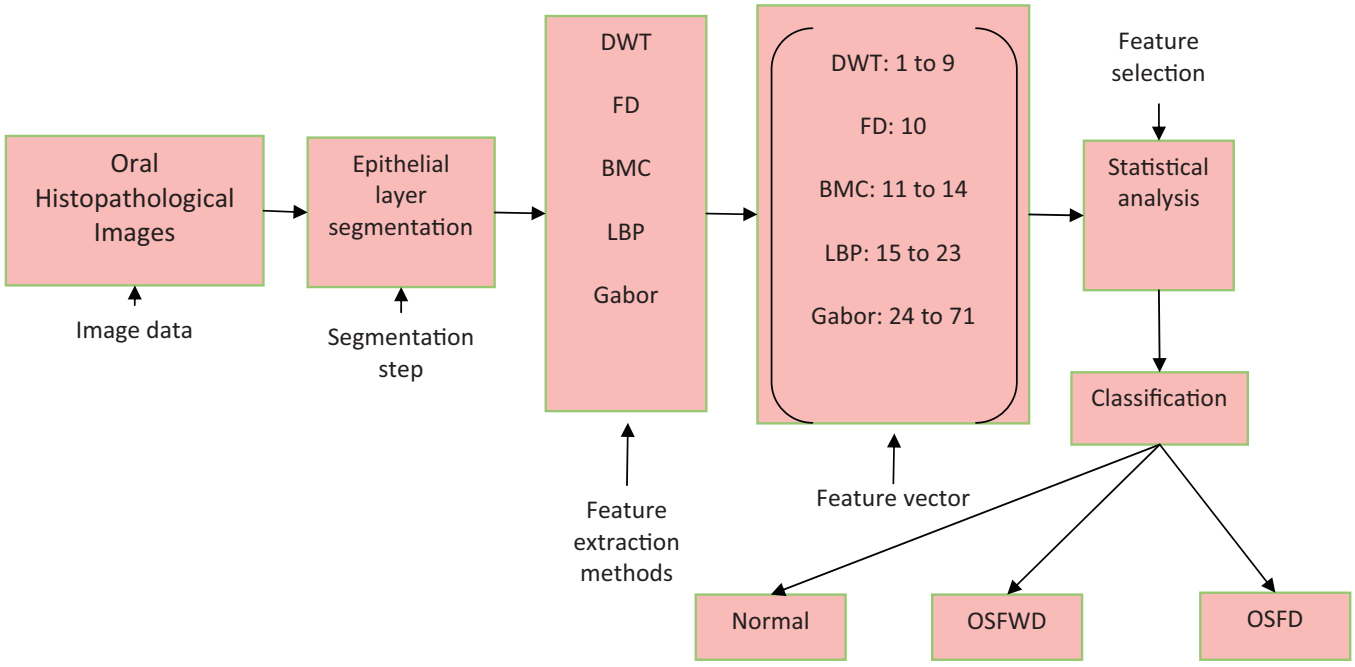

Fig. 1. The overall block diagram of OSF screening system.

watershed transformation combines elements from both the discontinuity and similarity methods. To incorporate the textural information contained within the epithelial layer (because of the cell arrangement pattern) and to compensate for the smoothing done for grayscale image gradient extraction, texture based watershed segmentation technique is proposed in this paper. In this technique, Gabor Filters (Jain and Farrokhnia, 1991) are used to compute several texture channels containing information at various frequencies and orientations from the pre-processed grayscale epithelial image. The vector gradient is computed for the multichannel texture matrix and combined with the intensity gradient yielding the final gradient image. After an automatic removal of insignificant minima from the gradient image, watershed transform (Sijbers et al., 1997) is applied on the image. Finally, a two-step region merging using RGB color information, histogram difference and texture based distance criterion are considered to obtain the final segmentation (Muthu Rama Krishnan et al., 2011b). The complete segmentation procedure consists on the following steps, shown in Fig. 2.

- *Pre-processing*: Wiener filtering, Contrast Enhancement and Shading correction are done to improve the image quality.
- *Texture analysis*: a set of *n* channels are obtained from the original image by computing *n* texture features on every pixel.
- *Gradient computation*: considering each point of the multichannel image as an *n*-valued vector, a gradient image is obtained by computing the gradient of the vector field and is combined with intensity gradient.
- *Minima selection*: dynamics is used for local minima selection from which the flooding begins by extending their influence zone in higher gray levels.
- *Region merging*: watershed regions are iteratively merged, according to a distance criterion computed using RGB color values, intensity distribution and texture to obtain the final segmentation.
- *Post processing*: edge enhancement and removal of unwanted artefacts connected to the epithelial layer edge.

2.3.1. Pre-processing

The presence of noise and staining variations within the epithelial layer (Fig. 3(a)) necessitates pre-processing. The RGB epithelial image is converted to Lab color space and the luminance channel, L is subjected to (a) wiener filtering using a 5×5 filter, (b) gray level shading correction using low pass filtering (Gonzalez and Woods, 2002), (c) contrast enhancement by mapping intensity values from 0 to 255. The processed L channel is then combined with the chrominance channels (a and b) and converted back to RGB color space. Further, the RGB image is converted in to gray scale image by forming a weighted sum of the R, G, and B components. Moreover, gray level shading correction helps to remove the erroneous staining lines present within the epithelial layer.

The texture based segmentation results are shown in Fig. 3 for all three classes viz., normal, OSF without dysplasia and OSF with dysplasia.

2.4. Feature extraction

Feature extraction is one of the most important steps in automated CAD systems, because this step extracts relevant and representative features from measurement data such as images and signals. In this work, wavelet, FD, BMC, LBP and Gabor wavelet features are extracted from normal, OSFWD and OSFD images is shown in Fig. 3(c).

2.4.1. Multi-scale texture feature extraction methods

2.4.1.1. Wavelet and sub-band decomposition based techniques. In this work, we used two dimensional (2D) DWT and Frobenius norm for feature extraction. We start the discussion of these methods by introducing the DWT, which analyzes one-dimensional (1D) signals. Later, the concept is extended to 2D signals (Chih-Chin. et al., 2010; Hui et al., 2011). This is done by considering all degrees of freedom such 2D signals offer, we arrive at the definition of 2D DWT. Finally, we define the Frobenius norm for the 2D DWT results which yield the feature vector elements.

The DWT transform of a signal x is evaluated by sending it through a sequence of down-sampling high and low pass filters.

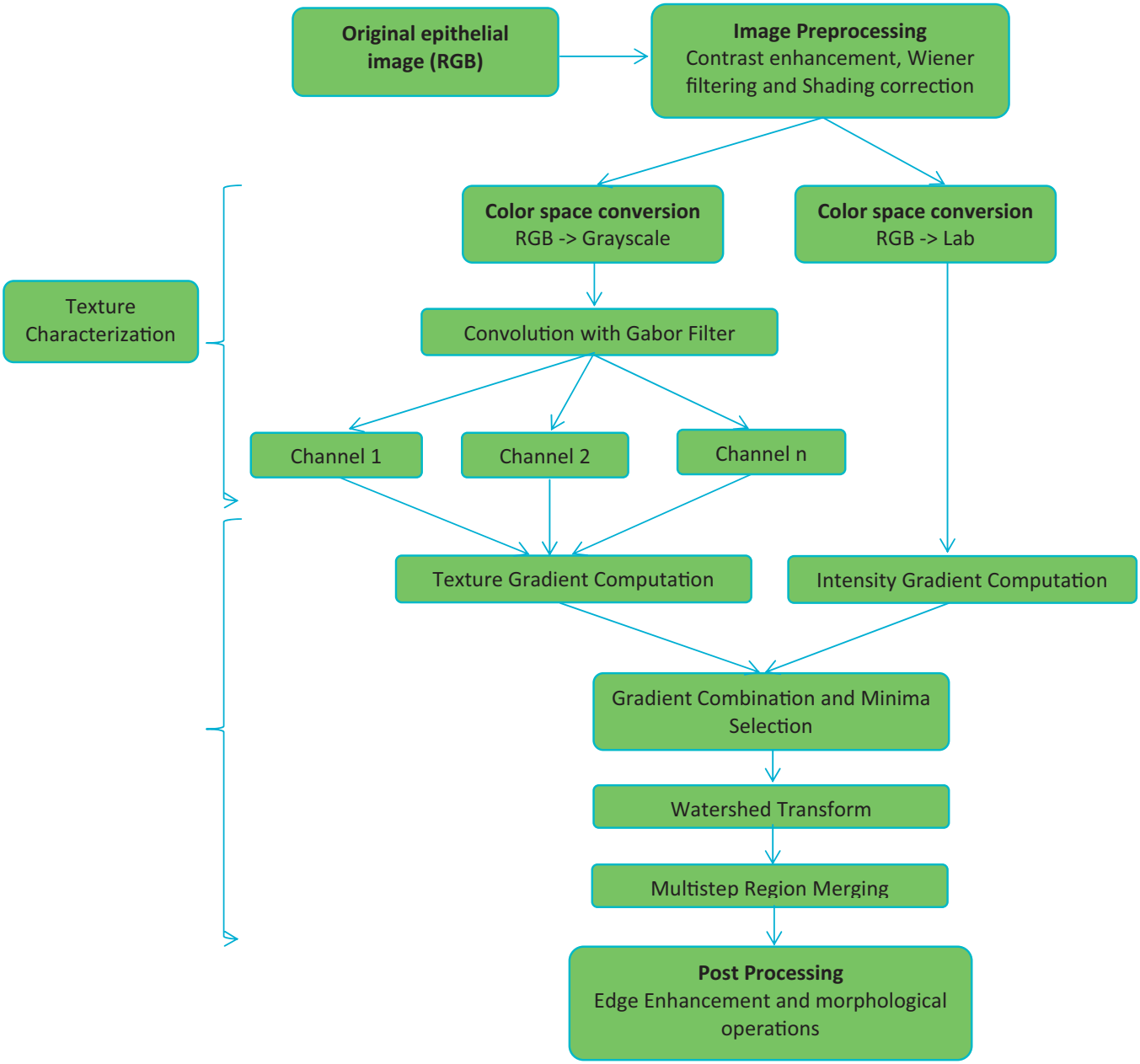

Fig. 2. Overall block diagram of texture gradient based epithelial layer segmentation (Muthu Rama Krishnan et al., 2011b).

The low pass filter is defined by the transfer function $g[n]$ and the high pass filter is defined by the transfer function $h[n]$. The output of the high pass filter $D[n]$ is known as the detail coefficients. The following equation shows how these coefficients are obtained:

$$D[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k] \quad (1)$$

The output of the low pass filter is known as the approximation coefficients. These coefficients are found by using the following equation:

$$A[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k] \quad (2)$$

The frequency resolution is further increased by cascading the two basic filter operations. To be specific, the output of the first level low pass filter is fed into the same low and high pass filter combination. The detailed coefficients are output at each level and they form the level coefficients. In general, each level halves the sample number and doubles the frequency resolution. Consequently, in the final level, both detail and approximation coefficients are obtained as level coefficients.

For 2D signals, the 2D DWT can be used. Our discussion focuses on Wavelet packets (WP) for images. These images are represented as an $m \times n$ gray scale matrix $I[i,j]$ where each element of the matrix represents the intensity of one pixel. All non-border pixels in $I[i,j]$, where $i \notin \{0, m\}$ and $j \notin \{0, n\}$, have eight immediate neighboring pixels. These eight neighbors can be used to traverse through the matrix. However, changing the direction with which the matrix is traversed just inverts the sequence of pixels and the 2D DWT coefficients are the same. For example, the WP result is the same when the matrix is traversed from left to right as from right to left. Therefore, we are left with four possible directions, which are known as decomposition corresponding to 0° (horizontal, H), 90° (vertical, V) and 45° or 135° (diagonal, D) orientations. The implementation

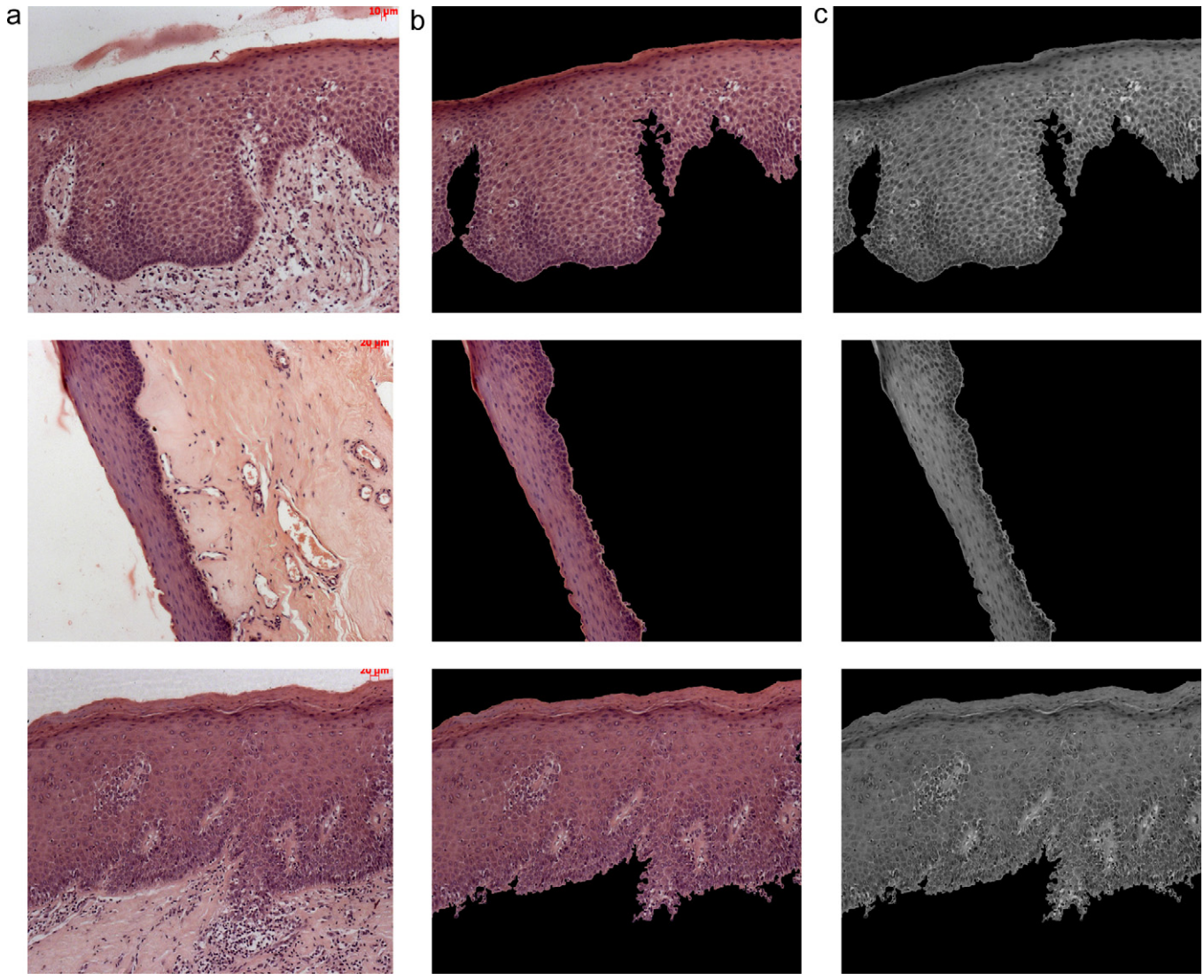

Fig. 3. Results of the texture based segmentation technique: (a) original normal, OSF without dysplasia and OSF with dysplasia images respectively from top to bottom; (b) segmented images using the texture based segmentation technique; (c) grayscale images of (a).

of this algorithm follows the block diagram shown in Fig. 4. The diagram shows the $N \times M$ dimensional input image $I[i, j]$ and the results for level 1. The results from level 1 were sufficient to obtain significant features.

In this work, we have evaluated 54 wavelet functions. Each of these wavelet functions has both a unique low pass filter transfer function $g[n]$ and a unique high pass filter transfer function $h[n]$. Among these 54 wavelet functions Biorthogonal 3.1 (*bior3.1*) performs well. The extracted features from each wavelet family are described as follows. The denseness and the degree of disorder in an image are measured by energy. The energy has been computed as a percentage of information present in each sub band. Mathematically,

$$En_k = \frac{\sum_i \sum_j C_k}{\sum_k \sum_i \sum_j C_k} \times 100 \tag{3}$$

where C_k is either horizontal, diagonal, vertical or approximation wavelet coefficients, i, j represents the rows and columns of the matrix and k represents the classes. It yields four features energy in horizontal coefficient (HE), energy in diagonal coefficient (DE), energy in vertical coefficient (VE), and energy in approximation coefficient (AE). Further, the Frobenius norm has been computed for H_1, V_1, D_1 and A_1 and denoted as $|\cdot|_F$. The Frobenius norm, sometimes also called the Euclidean norm, is matrix norm of an $m \times n$ matrix A defined as the square root of the sum of the absolute squares of its elements

The Frobenious norm of $A \in C^{m \times n}$ is defined as

$$||A||_F^2 = \sum_{i,j} |a_{ij}|^2 = \sum_i ||A_{i*}||_2^2 = \sum_j ||A_{j*}||_2^2 = trace(A * A) \tag{4}$$

The element of the feature vector (F) is the Frobenious norm of H_1, V_1, D_1 and A_1 $F = [||H_1||_F ||V_1||_F ||D_1||_F K ||A_1||_F]^T$, $i = 1, 2, 3, 4$ where K is set as 0.001.

2.4.1.2. Gabor wavelet. A two dimensional Gabor function (Manjunath and Ma, 1996) is a Gaussian modulated complex sinusoid and it is written as

$$\psi_{i,k}(m, n) = \frac{1}{2\pi\sigma_m\sigma_n} \exp\left(-\frac{1}{2}\left(\frac{m^2}{\sigma_m^2} + \frac{n^2}{\sigma_n^2}\right) + 2\pi j\omega m\right) \tag{5}$$

Here, ω is frequency of the sinusoid and σ_m and σ_n are the standard deviations of the Gaussian envelopes. 2D Gabor wavelets

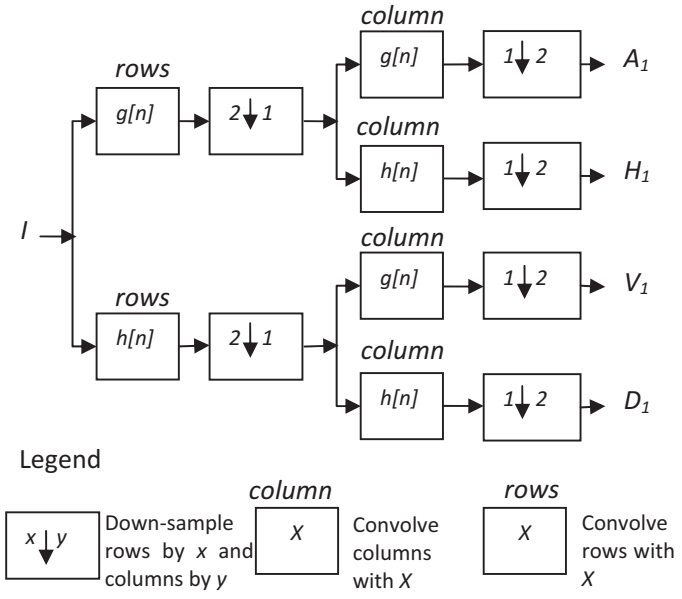

Fig. 4. Discrete wavelet transform (DWT) decomposition.

are obtained by dilation and rotation of the mother Gabor wavelet $\psi(m, n)$ using

$$\psi_{i,k}(m,n) = a^{-l}\psi[a^{-l}(m\cos\theta + n\sin\theta), a^{-l}(-m\sin\theta + n\cos\theta)], a > 1 \quad (6)$$

where a^{-l} is a scale factor, l and k are integer, the orientation θ is given by $\theta = k\pi/K$ and K is the number of orientations. The parameters σ_m and σ_n are calculated according to the design strategy proposed by (Manjunath and Ma, 1996). Given an image $I(m,n)$, its Gabor wavelet transform is obtained as

$$x_{l,k}(m,n) = I(m.n) * \psi_{l,k}(m,n) \text{ for } l = 1, 2, \ldots, S \text{ and } k = 1, 2, \ldots, K \quad (7)$$

where * denotes a convolution operator. The parameters K and S are number of orientation and number of scales, respectively. The mean and the standard deviation are used as features and given by

$$\mu_{l,k}(m,n) = \frac{1}{M\times N}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|x_{l,k}(m,n)\right| \quad \text{and} \quad \sigma_{l,k} = \left(\frac{1}{M\times N}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(\left|x_{l,k}(m,n)\right| - \mu_{l,k}\right)^2\right)^{\frac{1}{2}} \quad (8)$$

The feature vector is then constructed using $\mu_{l,k}$ and $\sigma_{l,k}$ as feature components, for $K = 6$ orientations and $S = 4$ scales, resulting in a feature vector of length 48, given by

$$f = \{\mu_{11}, \sigma_{11}, \ldots, \mu_{48}, \sigma_{48}\} \quad (9)$$

2.4.1.3. Local binary pattern. The local binary pattern (Ojala et al., 2002) is a simple and fast method for multi-scale local texture analysis. Since the epithelium tissues are at different orientations within a particular cancer case, rotation invariant texture measure is needed to describe the microstructure following the uniformity criteria. The spatial structure information is combined with contrast which is a measure of the amount of local texture.

A circular neighborhood is considered around a pixel. P points are chosen on the circumference of the circle with radius R such that they are all equidistant from the centre pixel. The gray values at points on the circular neighborhood that do not coincide exactly with pixel locations are estimated by interpolation. Let g_c be the gray value of the centre pixel and g_p, $p = 0, \ldots, P-1$, corresponds to the gray values of the P points. These P points are converted into a circular bit-stream of 0s and 1s according to whether the gray value of the pixel is less than or greater than the gray value of the centre pixel. Fig. 5 depicts circularly symmetric neighbor sets for different values of P and R.

The texture T for this pixel is given as

$$T = t(g_c, g_0, \ldots, g_{P-1}) \quad (10)$$

where $g_p(p = 0, \ldots, P-1)$ correspond to the gray values of P equally spaced pixels on a circle of radius R. The LBP code for the centre pixel is given as

$$LBP_{P,R} = \sum_{p=0}^{p-1} s(g_p - g_c)2^p \quad (11)$$

where $s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$

Rotation invariance is achieved by rotating the neighbor set clockwise such that maximal number of most significant bits is zero in the LBP code.

$$LBP^{ri}_{P,R} = \min\{ROR(LBP_{P,R,i})|i = 0, 1, \ldots, P-1\} \quad (12)$$

To improve the discrimination power of LBP^{ri}, LBP^{riu} using a uniformity measure (U) is calculated based on the no. of transitions in the neighborhood pattern. Only patterns with $U \leq 2$ are assigned the LBP code.

$$LBP^{riu2}_{P,R} = \begin{cases} \sum_{p=0}^{p-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \\ P+1 & \text{otherwise,} \end{cases} \quad (13)$$

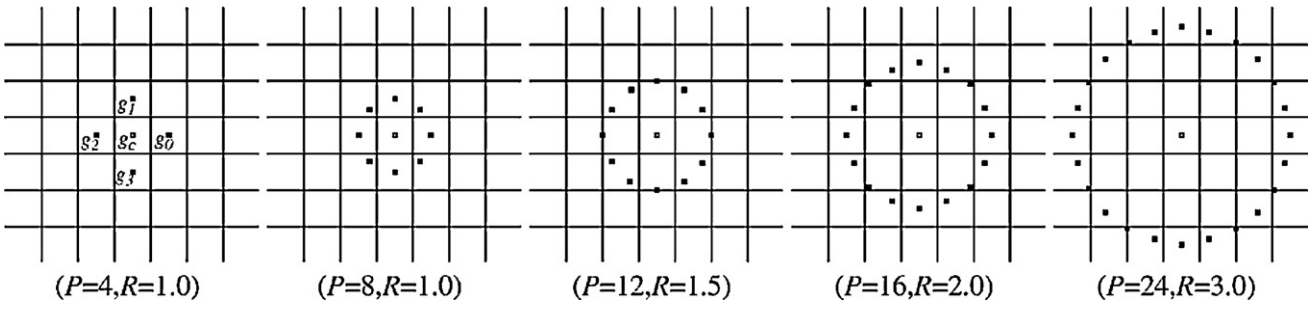

Fig. 5. Circularly symmetric neighbor sets for different P and R (Ojala et al., 2002).

To include the local image texture contrast, we define a rotation invariant measure of local variance given by:

$$VAR_{P,R} = \frac{1}{P}\sum_{p=0}^{p-1}(g_p - \mu)^2, \text{ where } \mu = \frac{1}{P}\sum_{p=0}^{p-1} g_p \quad (14)$$

The LBP^{riu2} code and local contrast values $VAR_{P,R}$ are computed for 3 radius values; $R = 8$, 16 and 24 with the corresponding pixel count P being 8, 16 and 24, respectively. The mean and variance of each of the LBP output image is calculated which is combined with the mean local contrast of the image to yield 9 features corresponding to each pixel of the epithelium tissue.

2.4.2. Statistical texture features based methods

2.4.2.1. Fractal dimension. Fractal dimension is one of the methods to perform texture analysis. Surface epithelium can be viewed as 3D object having intensity variation as covering surface on the 2D spatial plane. Any surface A in Euclidean n-space is self-similar if A is the union of N_r distinct (non-overlapping) copies of itself scaled up or down by a factor of r. Mathematically, FD is computed (Mandelbrot, 1982) using the following formula:

$$D = \frac{\log N_r}{\log\left(1/r\right)} \quad (15)$$

where D is the fractal dimension. Here we consider modified differential box counting with sequential algorithm. The input of sequential algorithm is gray-scale image where the grid size is in the power of 2 for efficient computation (Biswas et al., 1998). Maximum and minimum intensity for each box (2×2) are obtained to sum their difference, which gives the N and r by

$$r = \frac{s}{M} \quad M = \text{minimum}(R \oplus C) \quad (16)$$

where s denotes scale factor, R and C denote the number of rows and number of columns respectively when the grid size gets doubled, R and C reduces to half of its original value and above procedure is repeated iteratively untill $maximum(R \oplus C)$ is greater than 2. Linear regression model uses to fit the line from plot $\log(N)$ vs. $\log(1/r)$ and the slope gives the FD as:

$$\log N_r = D \log\left(\frac{1}{r}\right) \quad (17)$$

2.4.2.2. Brownian motion curve. Another approach to estimate texture of surface epithelium is based on BMC (Chen et al., 1998). Suface epithelium in OSF and normal mucosa have definite textural variation. Brownian motion for an image can be described in terms of intensity and distance map. Brownian motion is defined in (Hurst et al., 1965; Chen et al., 1998) as expected value of this intensity difference is proportional to $\left|\sqrt[2]{(x_2 - x_1)^2 + (y_1 - y_2)^2}\right|^H$, where H is Hurst coefficient ($0 < H < 1$).

$$E\left|I(x_2, y_2) - I(x_1, y_1)\right| = C\left|\sqrt[2]{(x_2 - x_1)^2 + (y_1 - y_2)^2}\right|^H \quad (18)$$

$$\therefore \log(E\left|I(x_2, y_2) - I(x_1, y_1)\right|) = H\log(\left|\sqrt[2]{(x_2 - x_1)^2 + (y_{1-} - y_2)^2}\right|) + \log C \quad (19)$$

where E indicates expected value of intensity difference between two pixels

Thus H varies from 0 to 1 to satisfy this equation across all types of images and hence describes texture of an image. This textural analysis asks exceptionally higher computational tasks, especially

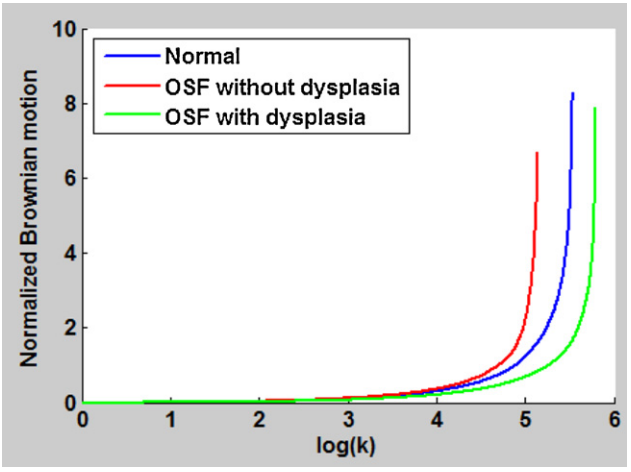

Fig. 6. BMC till first peak shows different motion for normal and OSF.

distance calculation. The problem can be overcome by considering Chebyshev distance with some approximation. Now, given an $M \times N$ image, an intensity difference vector of the image is defined as

$$IDV = [ID(1),\ ID(2),\ \ldots ID(k),\ \ldots ID(s)] \quad (20)$$

where s is the maximum possible scale and $ID\ (k)$ is defined as follows:

$$Id_1(k) = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N-k}\left|I(x,y) - I(x,y+k)\right|}{M*(N-k) - (\text{count of } I(x,y) = 0)}, \forall I(x,y) \neq 0 \quad (21)$$

$$Id_2(k) = \frac{\sum_{x=1}^{M-k}\sum_{y=1}^{N}\left|I(x,y) - I(x+k,y)\right|}{N*(M-k) - (\text{count of } I(x,y) = 0)}, \forall I(x,y) \neq 0 \quad (22)$$

$$Id_3(k) = \frac{\sum_{x=1}^{M-k}\sum_{y=1}^{N-k}\left|I(x,y) - (x+k, y+k)\right|}{(M-k)*(N-k) - (\text{count of } I(x,y) = 0)}, \forall I(x,y) \neq 0 \quad (23)$$

$$Id_4(k) = \frac{\sum_{x=1}^{M-k}\sum_{y=1}^{N-k}\left|I(x, n-y+1) - I(x+k, n+1-(y+k))\right|}{(M-k)*(N-k) - (\text{count of } I(x,y) = 0)}, \quad \forall I(x,y) \neq 0 \quad (24)$$

$$ID(k) = \frac{\sum_{j=1}^{4} Id_j(k)}{4}, \forall k = 0,\ 1,\ \ldots (m \times n) - 1 \quad (25)$$

here all pixel pairs calculated in the absolute difference ($Id_i(k)$ where $i = 1, 2, 3, 4$) are along horizontal, vertical, diagonal, and asymmetric-diagonal directions, respectively. We define function $f(k)$, normalized intensity difference vector as normalized to neighbor of the each pixel to compensate the staining variation. Mathematically,

$$f(k) = \log_{10}\left|ID(k)\right| - \log\left|ID(1)\right|, k > 1 \quad (26)$$

Thus from Eq. (17)

$$f(k) = H\ \log(k), \quad k = 1,\ 2,\ \ldots,\ \min(M, N)$$

The value of H can be estimated using least square estimation (Chen et al., 1998) from the BMC ($f(k)$ vs. $\log(k)$). Apart from this H value, three more features are extracted by the following method. The BMC is shown in Fig. 6 for normal and OSF. It can be observed from the graph there are distinct peaks between normal, OSFWD and OSFD, which will discriminate the three classes.

2.4.2.2.1. Proposed features. The BMC is used to extract the following features for classification. Here, the approach considered is to find area under the curve.

(i) The area under $f(x)$ is given by

$$A = \int_0^\infty f(x)dx \tag{27}$$

(ii) Area till first peak (AFP) is defined for discrete value of x as follows:

$$AFP = \sum_{x=0}^{X_F} f(x) \tag{28}$$

where X_F is the location of the first peak. Further analysis of the curve (after first peak) is carried out by finding the number of the peaks (NP). Apart from NP, the distance between peaks can be considered as one of the feature because it gives the rough estimation about object and background locations. Hence, mean distance between peaks (MDP) for peak location X_i, $i = 1, \ldots,$ NP is defined as

$$MDP = \sum_{i=2}^{NP} \frac{\left|X_i - X_{i-1}\right|}{NP - 1} \tag{29}$$

3. Statistical analysis

Prior to classification, it is meaningful to verify whether a feature or a set of features has the discriminating capability among the labeled classes or not. In doing so, classical statistical inference provides one of the well-established statistical tests viz., One-way ANOVA (Gun et al., 2005), which is used for comparing more than two population means. That is, to determine whether the groups are actually different in the measured characteristic. A large difference between the three sample means should lead us to reject the null hypothesis $H_0 : \mu_1 = \ldots = \mu_k$. The ANOVA test makes the following assumptions about the data

(a) All sample populations are normally distributed.
(b) All sample populations have equal variance.
(c) All observations are mutually independent.

4. Classification

In this work, we used support vector machine classifier. It is briefly described in this section.

4.1. Support vector machine approach

The extracted texture features after test of significance are fed to support vector machine (SVM) (Huang et al., 2008; Burges and Christopher, 1998) to classify healthy and OSF classes. It is a set of related supervised learning methods used for classification and regression. Let us denote a feature vector by $\underline{x} = (x_1, x_2, \ldots x_n)$ and its class label by y such that $y = \{+1, -1\}$. Therefore, consider the problem of separating n training patterns belonging to two classes as $(\underline{x}_i,\ y_i),\ \ldots \underline{x}_i \in R^n, y_i = \{+1, -1\}, i = 1, 2, \ldots, n$. A decision or discriminant function $g(x)$ that can correctly classify an input pattern x that is not necessarily from the training set.

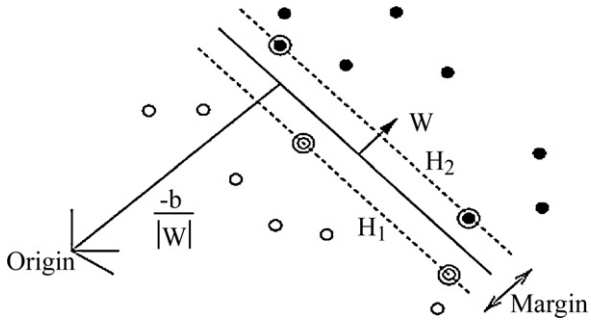

Fig. 7. Optimal separating hyperplanes.

4.1.1. Linearly separable data

A linear SVM (Vapnik, 1998; Huang et al., 2008; El-Naqa et al., 2002) is used to classify data sets which are linearly separable. In case of linear SVM, the discriminant function is of the form:

$$g(\underline{x}) = \underline{w}^T\underline{x} + b \tag{30}$$

such that $g(\underline{x}_i) \geq 0$ for $y_i = +1$ and $g(\underline{x}_i) < 0$ for $y_i = -1$. Means, hyperplane $g(\underline{x}) = \underline{w}^T\underline{x} + b = 0$ puts the two classes on either side of the plane during training. The key issue is to find the hyperplane that causes the largest separation between the decision function values and the two classes. Hence, we have to maximize the total width between two margins $(2/\underline{w}^T\underline{w})$. Mathematically, this hyperplane can be found by minimizing the following cost function:

$$J(\underline{w}) = \frac{1}{2}\underline{w}^T\underline{w} \tag{31}$$

Subject to separability constraints

$$\begin{array}{ll} \underline{w}^T\underline{x}_i + b \geq +1 & \text{for } y_i = +1 \\ \text{or } \underline{w}^T\underline{x}_i + b \leq -1 & \text{for } y_i = -1,\ i = 1, 2, \ldots n. \end{array} \tag{32}$$

Equivalently, these constraints can be re-written more compactly as

$$y_i(\underline{w}^T\underline{x}_i + b) \geq 1; \quad i = 1, 2, \ldots, n \tag{33}$$

To solve this quadratic optimization problem one must find the saddle point of the Lagrange function

$$L_p(\underline{w}, b, \alpha) = \frac{1}{2}||\underline{w}|| - \sum_{i=1}^{n}\alpha_i\left\{y_i(\underline{w}^T\underline{x}_i + b) - 1\right\} \tag{34}$$

where the α_i denotes Lagrange multipliers, hence $\alpha_i \geq 0$. The search for an optimal saddle point is necessary because L_p must be minimized with respect to the primal variables $\underline{w}$ and b and maximized with respect to the dual variable α_i. By differentiating with respect to $\underline{w}$ and b, and introducing the Karush–Kuhn–Tucker (Gunn Steve, 1998) conditions for the optimum constrained function, then is transformed to the dual Lagrange $L_D(\alpha)$

$$\underset{\alpha}{Max}\quad L_D(\alpha) = \left[\sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_i\sum_j\alpha_i\alpha_j y_i y_j \underline{x}_i^T\underline{x}_j\right] \quad \text{subject to}: \alpha_i \geq 0,\ i = 1, 2, \ldots, n \text{ and } \sum_{i=1}^{n}\alpha_i y_i = 0 \tag{35}$$

To find the optimal hyperplane (Fig. 7), a dual Lagrange L_D must be maximized with respect to non-negative. The solution α_i for the dual optimization problem determines the parameters $\underline{w}$ and b of the optimal hyperplane. The solution α_i for the dual optimization problem determines the parameters $\underline{w}^*$ and b^* of the optimal

hyperplane. Thus the optimal hyperplane decision function can be written as

$$g(\underline{x}) = sgn\left\{\sum_{i=1}^{n} y_i \alpha_i^* \underline{x}_i^T \underline{x} + b^*\right\} \quad (36)$$

4.1.2. Non-linear SVM

Higher dimensional nonlinear mapping (kernel function) of the input vector x generates the non-linearly separate input into linearly separable vector (Gunn Steve, 1998). By choosing a proper kernel function, the SVM constructs an optimal separating hyperplane in this higher dimensional space. Suppose the data is mapped to some other (possibly infinite dimensional) Euclidean space H, using a mapping which is defined by φ:

$$\varphi(.) : R^n \rightarrow R^{nh} \quad (37)$$

In this case, optimal function (35) becomes (38) with the same constraints

$$\underset{\alpha}{Max} \quad L_D(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K(\underline{x}_i^T \underline{x}_j) \quad (38)$$

where

$$K(\underline{x}_i^T \underline{x}_j) = \left\{\varphi(\underline{x}_i^T).\ \varphi(\underline{x}_j)\right\} \quad (39)$$

is the kernel function performing the non-linear mapping into feature space. The kernel function may be any of the symmetric functions that satisfy the Mercel conditions. The most commonly used are the Gaussian radial basis function and the polynomial function. Their formulas are shown in (40) and (41), respectively.

$$K(\underline{x}_i^T, \underline{x}_j) = \exp\left(\frac{||\underline{x}_i^T - \underline{x}_j||}{2\sigma^2}\right) \quad (40)$$

$$K(\underline{x}_i^T, \underline{x}_j) = (\underline{x}_i^T, x_j + 1)^d \quad (41)$$

where the parameters variance σ and degree d in (40) and (41), respectively must be preset.

There are several algorithms that extend the basic binary SVM classifier to be a multi-class classifier. The examples are one-against-one SVM, one-against-all SVM (Hsu and Lin, 2002; Weston and Watkins, 1998), half against half SVM (Lei and Govindaraju, 2005) and Directed Acyclic Graph SVM (DAGSVM) (Platt et al., 2000). In this experiment, we use the one-against-one SVM algorithm as the classification of the input among the three (normal, OSFWD and OSFD images) classes.

5. Results and discussion

This technique exploits the physiological structure of the epithelium with five different texture measures, and attempts to find the metrics of different texture measures. OSFD is more heterogeneous than OSFWD and normal. In general, epithelial dysplasias in OSF are considered as potential biological alterations towards progression to malignancy, while severe dysplasia is considered as carcinoma in situ. Therefore, it may be assumed that dysplastic conditions display some texture signatures of the ongoing process of malignant transformation as compared to the non-dysplastic state (Das et al., 2010). Dysplasia, which primarily signifies the abnormality in maturation of cells pertaining to neoplasia, within a tissue is aptly reflected by texture. In this study we have considered preneoplasia and OSF without dysplasia (epithelial atrophy) cases. It can be observed from the scatter plot of fractal dimension (Fig. 8) indicates a clear separation between the three groups. It suggests that these features (see Figs. 4 and 8) are significant and can be used to classify normal and OSF.

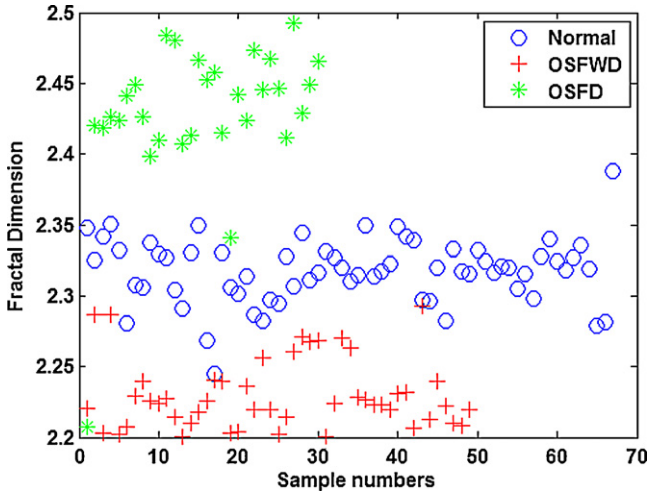

Fig. 8. Scatter plot of feature fractal dimension for normal, OSFWD and OSFD groups.

5.1. Selection of wavelet family

Initially, measurement has been made to find best wavelet family for OSF classification. It is followed by extraction of significant textural features and classification using SVM. Wavelet family selection is performed based on mean distance. For each wavelet, we found mean and standard deviation. If the standard deviation remains equal for each wavelet family as this is case for underlined problem, the criteria for best wavelet would be high interclass mean distance.

$$d_i^2 = \sum_{j=1}^{nf} (M_1(i,j) - M_2(i,j))^2 + (M_2(i,j) - M_3(i,j))^2 + (M_3(i,j) - M_1(i,j))^2,\ i = 1,\ 2,\ \ldots nw \quad (42)$$

nw and nf indicate number of wavelet family and number of features, respectively. Fig. 9 shows the distance map of all wavelets for wavelet derived features. There are 17 wavelet families as shown in the second column of Table 1. For every wavelet family there are 9 features are extracted as shown in the third column of Table 2. It is noted from Figs. 9 and 10 that Bior3.1 features have highest distance between classes.

Table 1
Wavelet family and extracted wavelet features.

Sl no	Wavelet family	Extracted wavelet features from each family
1	Bior1.3	• Norm of approximation coefficient
2	Bior1.5	• Norm of horizontal (H) coefficient
3	Bior3.1	• Norm of vertical (V) coefficient
4	Bior5.5	• Norm of diagonal (D) coefficient
5	Coif1	• Energy of approximation coefficient
6	Coif3	• Average energy of H, V, D coefficient in 1st level decomposition
7	Coif5	• Average energy of H, V, D coefficient in 2nd level decomposition
8	Db2	• Average energy of H, V, D coefficient in 3rd level decomposition
9	Db4	• Average energy of H, V, D coefficient in 4th level decomposition
10	Db6	
11	Db8	
12	Haar	
13	Sym2	
14	Sym4	
15	Sym6	
16	Sym8	
17	Sym10	

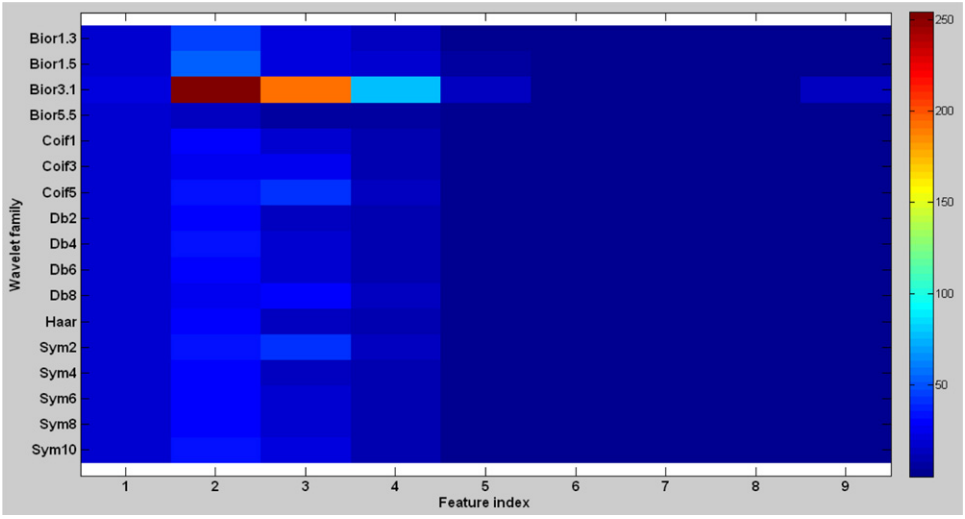

Fig. 9. Distance map for different wavelet family across wavelet features.

Table 2
Selected features between normal, OSF without dysplasia and OSF with dysplasia.

Feature index	Texture features
1–4	Wavelet coefficients*
5, 6, 7, 9	Wavelet energy*
10	Fractal dimension*
11, 12, 14	Fractional Brownian motion*
15–23	Local Binary Pattern*
24–71	Gabor wavelet*

* $p < 0.05$: statistical significance.

To investigate the significance of the results, all wavelet families have been used as an input to classifier to determine the significance of the measurement. The combination of bior3.1 wavelet with other texture measure provides higher accuracy. Fig. 11 shows the classification result for some of them.

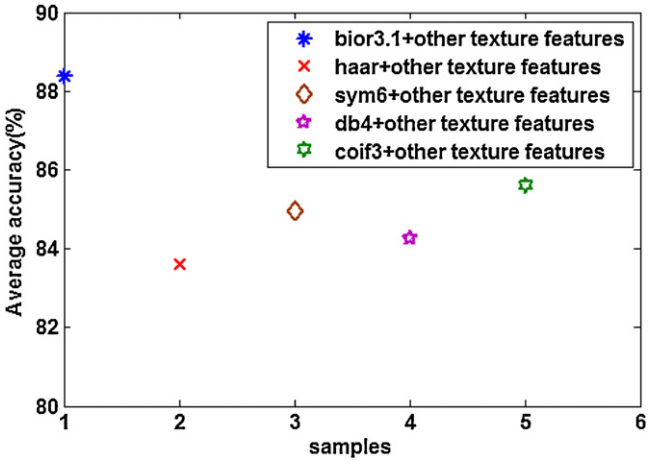

Fig. 11. Optimal combination of texture features and its classification accuracy.

5.2. Feature analysis and classification

The ranges of the 69 parameters used to feed as input to the SVM classifier. We have extracted 71 texture features using five methods. From them we chose 69 features, which are clinically significant represented in Table 2. These features are subjected to the analysis of variance (ANOVA) test to obtain the '*p*-value' (see Fig. 12). ANOVA (Gun et al., 2008) uses variances to decide whether the means are different. This test uses the variation (variance) within the groups and translates into variation (i.e., differences) between the groups, taking into account number of subjects present in the groups. If the observed differences are high then it is consid-

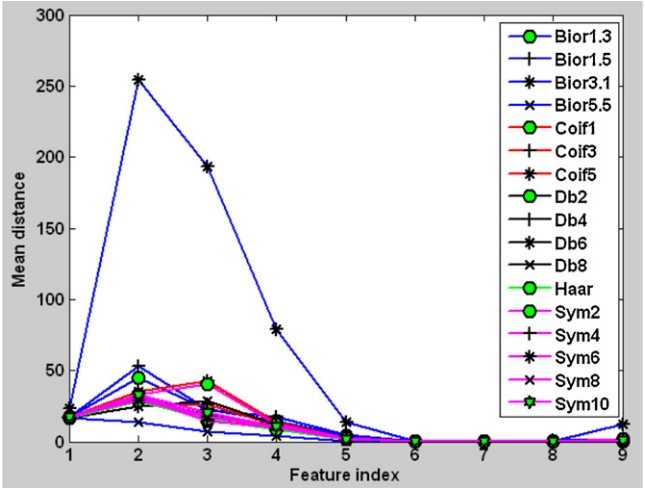

Fig. 10. Distance plot for different wavelet family across wavelet features.

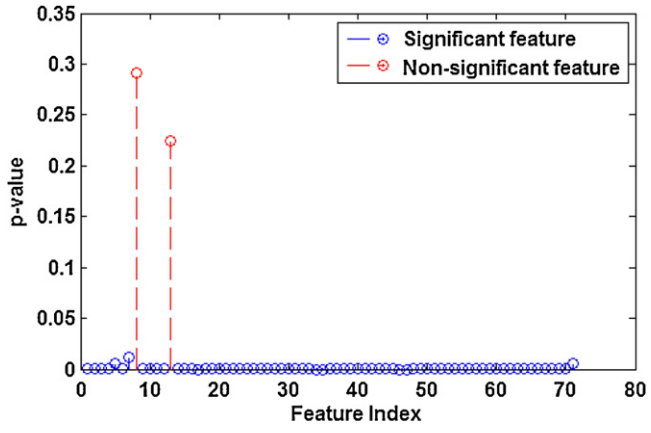

Fig. 12. *p*-Value for showing statistical significance of textural features.

Table 3
Individual texture features testing classification accuracy.

Texture feature extraction methods	Normal	OSFWD	OSFD	Over all accuracy (%)
Wavelet	43.75	63.26531	73.46939	60.16156
FD	25	24.4898	22.44898	23.97959
BMC	58.33333	51.02041	57.14286	55.49887
LBP	60.41667	73.46939	79.59184	71.1593
Gabor wavelet	**77.08333**	**77.55102**	**73.46939**	**76.03458**

ered to be statistical significant. In our work, we have obtained p value of less than 0.005, indicating that, it is statistically significant.

Considering the small size of the dataset used in this work, three-fold stratified cross validation [http://www.cs.cmu.edu/~schneide/tut5/node42.html] resampling technique is employed to test the classifiers with the sixty-nine texture features extracted using wavelet, FD, BMC, LBP and Gabor wavelet methods. That is, the whole dataset is divided into three parts such that each part contains approximately the same proportion of class samples as the original dataset. Two parts of the data (training set) are used for classifier development and the built classifier is evaluated using the remaining one part (test set). This procedure is repeated three times using a different part for testing in each case. The three test classification performance metrics are then averaged, and the average test performance is declared as the estimate of the true generalization performance. The performance metric used in this work is overall classification accuracy.

5.3. Individual and combined classification accuracies

Testing classification accuracies for each of the individual textures and in different combinations are as shown in Tables 3 and 4, respectively. These results represent the histopathological images of the normal, OSFWD and OSFD groups. When selectively combining certain texture features, the classification accuracy would increase above the highest achieved if an individual texture features method is used alone. For example, Table 3 shows the overall classification accuracies if the extracted texture features would be used individually (i.e., without combining them with each other). The Gabor wavelet texture feature achieved the highest overall accuracy by 76.03%. Yet, when fusing the texture features with each other, and in all possible combinations, some combination improved the overall accuracy up to 88.38% as in the Wavelet & LBP & Gabor wavelet paired features (as shown in Tables 4 and 5). By taking the Gabor wavelet classification accuracy from Table 3 and setting it as a threshold as it achieved the highest in case if each texture feature is used individually then we can see that the first five rows in Table 4 for the combined texture feature improved the accuracy. To investigate the significance of the results, a one way ANOVA test is applied to determine the significance between the texture measure combinations that improved the overall accuracy and the individual approaches. The test shows there is statistically significant difference on a significance level of $p < 0.05$. The optimum texture features for each of the texture features combinations that improved the classification accuracy are listed in Tables 4 and 5 with bold font.

Medical imaging plays an important role in the early detection and treatment of cancer. It provides physicians with information essential for efficient, effective and early diagnosis of various diseases. Over the last two decades, a lot of research has been (Demir and Yener, 2005) conducted on automated cancer diagnosis. This is partly because automated cancer diagnosis holds great promise for cost effective and large-scale use in the screening and management of cancer treatment.

In the present day the gold standard for qualitative method of assessment of malignant potentiality of oral precancerous lesions and conditions are primarily based on light microscopic features of the dysplastic surface epithelium, especially overall architecture of the epithelium. The general procedure is to extract various features from the epithelial layer and to allow a classifier to make the diagnostic prediction based on the fed features. One such study by Landini and Othman (2004) used the spatial arrangement of cells in neighborhoods of two sizes is characterized by constructing graph networks. They studied how the features are varying from normal to abnormal but they have not done any automated classification. Another study Muthu Rama Krishnan et al. (2010) proposed the fractal features for texture quantification of epithelium. They suggested automated classifiers for texture classification of epithelium.

In addition to histoapthological image analysis optical spectroscopy has been shown to discriminate early cancers by detecting alterations in the optical properties of human epithelial tissues, where 80% of cancers originate. Nieman et al. (2009) demonstrated that beveled multi-fibre probes are an effective tool to perform depth-resolved spectroscopy, which has the potential to improve precancer detection and monitoring. Moreover, nanotechnology offers unique opportunities for cancer detection, therapy and the

Table 4
Classification accuracy of extracted texture features in different combinations ranked in descending order.

Texture feature extraction methods	Normal	OSFWD	OSFD	Over all accuracy (%)
Wavelet & LBP & Gabor wavelet	**91.67**	**85.71**	**87.76**	**88.38**
Wavelet & FD & BMC & LBP & Gabor wavelet	**85.42**	**83.67**	**93.88**	**87.66**
LBP & Gabor wavelet	**81.25**	**89.79**	**89.79**	**86.95**
BMC & LBP	**87.50**	**89.79**	**79.59**	**85.63**
Wavelet & BMC & LBP & Gabor wavelet	**83.33**	**85.71**	**87.76**	**85.60**
FD & BMC & LBP	89.58	89.79	71.43	83.60
Wavelet & LBP	95.45	93.75	54.55	81.25
Wavelet & BMC & Gabor wavelet	79.17	79.59	81.63	80.13
BMC + Gabor wavelet	89.58	79.59	69.39	79.52
Wavelet & BMC	77.08	83.67	67.35	76.03
Wavelet & Gabor wavelet	77.27	75.00	72.73	75.00
FD & Gabor wavelet	68.75	73.47	79.59	73.94
FD & LBP	70.83	81.63	67.35	73.27
Wavelet & FD & BMC	68.75	73.47	77.55	73.26
Wavelet & FD	62.50	73.47	51.02	62.33
FD & BMC	56.25	55.10	51.02	54.12

Table 5
Optimum accuracy obtained using the texture combination.

Classes	No. of data training	No. of data testing	Percentage of correct classification (%) Wavelet, LBP and Gabor wavelet
Normal	60	30	91.67
OSFWD	28	14	85.71
OSFD	18	8	87.76
Average			**88.38**

ability to monitor therapeutic interventions. Sokolov et al. (2009) presented a case study of plasmonic nanoparticles in cancer imaging and therapy. Multispectral wide field optical imaging has the potential to improve early detection of oral cancer. The appropriate selection of illumination and collection conditions is required to maximize diagnostic ability. Roblyer et al. (2010) extracted image features and used to train and evaluate classification algorithms to discriminate tissue as non-neoplastic, dysplastic, or cancer; results were compared to histologic diagnosis. Autofluorescence imaging at 405-nm excitation provided the greatest image contrast, and the ratio of red-to-green fluorescence intensity computed from these images provided the best classification of dysplasia/cancer versus non-neoplastic tissue. A sensitivity of 100% and a specificity of 85% were achieved in the validation set. Multispectral widefield images can accurately distinguish neoplastic and non-neoplastic tissue; however, the ability to separate precancerous lesions from cancers with this technique was limited. A recent study from Roblyer and Kortum (2010) which included 124 subjects (60 patients with precancerous or cancerous lesions and 64 normal volunteers), demonstrated that spectroscopy could sucessfully identify 100% of the normal or benign regions, relative to the gold standard of histopathology.

In reality, the normal, OSFWD and OSFD groups do not necessarily have an identical structure or the same number of cells in the surface epithelium in premalignant and malignant condition. Thus, it is interesting to know how the applied texture measures cope with this situation and how their performance is affected. In this work, it is shown that combining more than one texture measure instead of using just one might improve the overall accuracy. Different texture measure tends to extract different features each capturing alternative characteristics of the examined structure. The multi-scale (Wavelet, Gabor wavelet and LBP) texture measures improved the overall accuracy up to 88.38% with none of the classified OSF subtypes achieving below 85.71%.

Since we have unequal samples (90 normal, 42 OSFWD and 26 OSFD), the accuracy is shown in Table 4. Due to unequal samples, there is a bias in the accuracy. We have used 3-fold cross validation for testing and training. Therefore, almost 8 samples in OSFD case and 30 normal samples. If approximately 10% of OSFD samples are misclassified (and other normal are classified properly) the accuracy is 97.37%. If 10% of normal samples are misclassified (and other OSFD are classified properly) the accuracy is 89.47%. The former one provides good accuracy which is misleading in the diagnosis. In the present work, the mentioned bias in performance is not considered in the accuracy. The methodology is suggested for equal samples (normal, OSFWD and OSFD), but due to the limitation in obtaining the samples we have demonstrated the methodology with this limited samples. Moreover, the accuracy can be increased by features extracted from the dysplastic region alone. Moreover, automatic detection using classifiers and good features is a promising area of research that has more opportunity for improvement.

6. Conclusion

A technique for histopathological OSF classification based on texture measures combination, which aims to overcome intra and inter-observer variability, has been proposed in this study. The texture gradient based epithelial layer segmentation is proposed in this work, and then feature extraction is performed by two statistical and three multi-scale texture measures for discrimination using a SVM classifier. The pre-processing phase represented by the appropriate texture gradient computation and minima selection proved to be necessary for increasing texture feature separability, and hence can improve classification accuracy. It can be concluded that certain selected texture measures play a complementary role to each other in the process of quantitative texture characterization. In other words, a certain texture measure can represent a pattern better than another depending on the region of interest frequency of occurrence and noise in the examined structure. This also applies to certain combinations which might outperform other texture measure fusions. However, combining more than two texture measures would not necessarily give a better accuracy even with the removal of highly correlated features. This will increase feature complexity, hence having a negative effect on the classifier's performance. It is found that the combination of the Wavelet, LBP and Gabor wavelet texture measures are the best for characterizing OSF subtypes, these three measures outperformed other measures in the study individually and combined. Furthermore, it would be interesting to test the compatibility of the suggested OSF classification approach to discern in-between subtypes and/or grades of other similar histopathological diseases.

Conflicts of interest

The authors report no conflicts of interest.

Acknowledgements

The authors would like to thank Dr. M. Pal, GNIDSR, Kolkata, India, and Dr. J. Chatterjee, SMST, IIT Kharagpur for their clinical support and valuable advices.

References

Alexandratou, E., Yova, D., Gorpas, D., Maragos, P., Agrogiannis, G., Kavantzas, N., 2008. Texture analysis of tissues in Gleason grading of prostate cancer. Imaging, Manipulation, and Analysis of Biomolecules Cells, and Tissues, 85904.

Al-Kadi, O.S., 2010. Texture measures combination for improved meningioma classification of histological images. Pattern Recognition 43, 2043–2053.

Biswas, M.K., Ghose, T., Guha, S., Biswas, P.K., 1998. Fractal dimension estimation for texture images: a parallel approach. Pattern Recognition Letters 19 (3–4), 309–313.

Burges, J., Christopher, C., 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 121–167.

Chang, R.F., Chen, C.J., Ho, M.F., 2004. Breast ultrasound image classification using fractal analysis. Proceedings of the Fourth IEEE Symposium on Bioinformatics and Bioengineering (BIBE'04).

Chen, E.L., Chung, P.C., Chen, C.L., Tsai, H.M., Chang, C.I., 1998. An automatic diagnostic system for CT liver image classification. IEEE Transactions on Biomedical Engineering 45 (6), 783–794.

Chih-Chin, L., Cheng-Chih, T., 2010. Digital image watermarking using discrete wavelet transform and singular value decomposition. IEEE Transactions on Instrumentation and Measurement 59 (11), 3060–3063.

Das, R.K., Pal, M., Barui, A., Paul, R.R., Chakraborty, C., Ray, A.K., Sengupta, S., Chatterjee, J., 2010. Assessment of malignant potential of oral submucous fibrosis through evaluation of p63, E-cadherin and CD105 expression. Journal of Clinical Pathology 63 (10), 894–899.

Demir, C., Yener, B., 2005. Automated cancer diagnosis based on histopathological images: a systematic survey. Rensselaer Polytechnic Institute, Department of Computer Science.

Duncan, J.S., Ayache, N., 2000. Medical image analysis: progress over two decades and the challenges ahead. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 85–106.

Estevez, J., Alayon, S., Moreno, L., Sigut, J., Aguilar, R., 2005. Cytological image analysis with a genetic fuzzy finite state machine. Computer Methods and Programs in Biomedicine 80, 3–15.

El-Naqa, I., Yang, Y., Wernick, M.N., Galatsanos, N.P., Nishikawa, M.R., 2002. A support vector machine approach for detection of microcalcifications. IEEE Transactions on medical imaging 21, 1552–1563.

Ferrari, R.J., Rangayyan, R.M., Desautels, J.E.L., Frere, A.F., 2001. Analysis of asymmetry in mammograms via directional filtering with Gabor wavelets. IEEE Transactions on Medical Imaging 20 (9), 953–964.

Gonzalez, R.C., Woods, R.E., 2002. Digital Image Processing, 2nd edition. Prentice Hall.

Gun, A.M., Gupta, M.K., Dasgupta, B., 2005. Fundamentals of Statistics, vol. 1, 5th edition. The World Press Pvt. Ltd.

Gun, A.M., Gupta, M.K., Dasgupta, B., 2008. Fundamentals of Statistics (Vol I & II), 4th edition. World Press Private Ltd.

Gunn Steve, R., 1998. Support Vector Machines for Classification and Regression, Technical Report, 1–66.

Gilles, F.H., Tavare, C.J., Becker, L.E., Burger, P.C., Yates, A.J., Pollack, I.F., Finlay, J.L., 2008. Pathologist interobserver variability of histologic features in childhood brain tumors: results from the CCG-945 study. Pediatric and Developmental Pathology 11, 08–117.

Grootscholten, C., Bajema, I.M., Florquin, S., Steenbergen, E.J., Peutz-Kootstra, C.J., Goldschmeding, R., Bijl, M., Hagen, E.C., Van, H.C., Houwelingen, D.R., Berden, J.H.M., 2008. Interobserver agreement of scoring of histopathological characteristics and classification of lupus nephritis. Nephrology Dialysis Transplantation 23, 223–230.

Hsu, C.W., Lin, C.J., 2002. A comparison of methods for multi class support vector machines. IEEE Transactions on Neural Networks 13, 415–425.

Huang, C.L., Liao, H.C., Chen, M.C., 2008. Prediction model building and feature selection with support vector machines in breast cancer diagnosis. Expert Systems with Applications 34, 78–587.

Hui, Y., Jie, G., Hua, G.Z., Chuan, L., 2011. An efficient method to process the quantized acoustoelectric current: wavelet transform. IEEE Transactions on Instrumentation and Measurement 60 (3), 696–702.

Hurst, H.E., Black, R.P., Simaika, Y.M., 1965. Long-Term Storage, An Experimental Study. Constable, London.

Jadhav, A.S., Banerjee, S., Dutta, P.K., Paul, R.R., Pal, M., Banerjee, P., Chaudhuri, K., Chatterjee, J., 2006. Quantitative analysis of histopathological features of precancerous lesion and condition using image processing technique. In: 19th IEEE Int. Symposium on Computer-Based Medical Systems, pp. 231–236.

Jain, A.K., Farrokhnia, F., 1991. Unsupervised texture segmentation using Gabor filters. Pattern Recognition 24 (12), 1167–1186.

Landini, G., Othman, I.E., 2004. Architectural analysis of oral cancer, dysplastic and normal epithelia. Cytometry A 61A, 45–55.

Lessmann, B., Nattkemper, T.W., Hans, V.H., Degenhard, A., 2007. A method for linking computed image features to histological semantics in neuropathology. Journal of Biomedical Informatics 40, 631–641.

Lei, H., Govindaraju, V., 2005. Half-against-half multi-class support vector machines. In: Proceeding Sixth International Work shop on Multiple Classifier Systems, Springer, Berlin, pp. 156–492.

Mandelbrot, B.B., 1982. The Fractal Geometry of Nature. WH Freeman Ed., New York.

Manjunath, B.S., Ma, W.Y., 1996. Texture features for browsing and retrieval of image data. IEEE Transaction in Pattern Analysis and Machine Intelligence 18, 837–842.

Marghani, K.A., Dlay, S.S., Sharif, B.S., Sims, A., 2003. Morphological and texture features for cancers tissues microscopic images. Medical Imaging and Image Processing 5032, 1757–1764.

Mukherjee, A., Paul, R.R., Chaudhuri, K., Chatterjee, J., Pal, M., Banerjee, P., 2006. Performance analysis of different wavelet feature vectors in quantification of oral precancerous condition. Oral Oncology 42, 914–928.

Muthu Rama Krishnan, M., Pal, M., Bomminayuni, S.K., Chakraborty, C., Paul, R.R., Chatterjee, J., Ray, A.K., 2009. Automated classification of cells in sub-epithelial connective tissue of oral sub-mucous fibrosis: an SVM based approach. Computers in Biology and Medicine 39 (12), 1096–1104.

Muthu Rama Krishnan, M., Shah, P., Pal, M., Chakraborty, C., Paul, R.R., Chatterjee, J., Ray, A.K., 2010a. Structural markers for normal oral mucosa and oral sub-mucous fibrosis. Micron 41 (4), 312–320.

Muthu Rama Krishnan, M., Shah, P., Chakraborty, C., Ray, A.K., 2010b. Statistical Analysis of Textural Features for Improved Classification of Oral Histopathological Images, doi:10.1007/s10916-010-9550-8.

Muthu Rama Krishnan, M., Pal, M., Chakraborty, C., Paul, R.R., Chatterjee, J., Ray, A.K., 2010c. Computer vision approach to morphometric feature analysis of basal cell nuclei for evaluating malignant potentiality of oral submucous fibrosis. Journal of Medical Systems 3 (4), 261–270.

Muthu Rama Krishnan, M., Shah, P., Chakraborty, C., Ray, A.K., 2011a. Brownian motion curve based textural classification and its application towards cancer diagnosis. Analytical Quantitative Cytology and Histology 33 (3), 1–11.

Muthu Rama Krishnan, M., Choudhary, A., Chakraborty, C., Ray, A.K., Paul, R.R., 2011b. Texture based segmentation of epithelial layer from oral histological images. Micron 42 (6), 632–641.

Nieman, L.T., Jakovljevic, M., Sokolov, K., 2009. Compact beveled fiber optic probe design for enhanced depth discrimination in epithelial tissues. Optics Express 17 (4), 2780–2796.

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

Platt, C.J., Chrisianini, N., Shawe-Taylor, J., 2000. Large Margin DAGs for multiclass classification. Advanced Neural Information Processing Systems 12, 547–553.

Prabhu, S.R., Wilson, D.F., Daftary, D.K., Jhonson, N.W., 1992. Oral Diseases in the Tropics. Oxford University Press.

Qian, W., Zhukov, T., Song, D.S., Tockman, M.S., 2007. Computerized analysis of cellular features and biomarkers for cytologic diagnosis of early lung cancer. Analytical and Quantitative Cytology and Histology 29, 103–111.

Roblyer, D., Kurachi, C., Stepanek, V., Schwarz, R.A., Williams, M.D., El-Naggar, A.K.J., Lee, J., Gillenwater, A.M., Kortum, R.R., 2010. Comparison of multispectral wide-field optical imaging modalities to maximize image contrast for objective discrimination of oral neoplasia. Journal of Biomedical Optics 15, 6.

Roblyer, D., Kortum, R.R., 2010. Optical diagnostics for early detection of oral cancer. ADHA Access, 22–25.

Seker, H., Odetayo, M.O., Petrovic, D., Naguib, R.N.G., 2003. A fuzzy logic basedmethod for prognostic decision making in breast and prostate cancers. IEEE Transactions on Information Technology in Biomedicine 7, 114–122.

Sertel, O., Kong, J., Shimada, H., Catalyurek, U., Saltz, J.H., Gurcan, M., 2008. Computer-aided prognosis of neuroblastoma: classification of stromal development on whole-slide Images – art. no. 69150P, Medical Imaging: Computer-Aided Diagnosis, Pts 1 and 2, 6915, 9150–9150.

Shuttleworth, J., Todman, A., Norrish, M., Bennett, M., 2005. Learning histopathological microscopy. Pattern Recognition and Image Analysis, Pt 2, Proceedings 3687, 764–772.

Sijbers, J., Scheunders, P., Verhoye, M., Linden, A.V.D., Dyck, D.V., Raman, E., 1997. Watershed segmentation of 3D MR data for volume quantization. Magnetic Resonance Imaging 15 (6), 679–688.

Sokolov, K., Tam, J., Tam, J., Travis, K., Larson, T., Aaron, J., Harrison, N., Emelianov, S., Johnston, K., 2009. Cancer imaging and therapy with metal nanoparticles. In: 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Engineering the Future of Biomedicine, EMBC.

Thiran, J.P., Macq, B., 1996. Morphological feature extraction for the classification of digital images of cancerous tissues. IEEE Transactions on Biomedical Engineering 43, 1011–1020.

Tilakaratne, W., Klinikowski, M., Saku, T., Peters, T., Warnakulasuriya, S., 1997. Oral submucous fibrosis: review on aetiology and pathogenesis. Oral Oncology 42, 561–568.

Vapnik, V., 1998. Statistical Learning Theory, 2nd edition. Wiley, New York.

Weston, J., and Watkins, C., 1998. Multi-class support vector machines Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, Egham.

Wiltgen, M., Gerger, A., Wagner, C., Bergthaler, P., Smolle, J., 2007. Evaluation of texture features in spatial and frequency domain for automatic discrimination of histologic tissue. Analytical and Quantitative Cytology and Histology 29, 251–263.

Wittke, C., Mayer, J., Schweiggert, F., 2007. On the classification of prostate carcinoma with methods from spatial statistics. IEEE Transactions on Information Technology in Biomedicine 11, 406–414.

Wu, C.M., Chen, Y.C., Hsieh, K.S., 1992. Texture features for classification of ultrasonic liver images, 11, 2, 141–152. http://www.cs.cmu.edu/~schneide/tut5/node42.html last accessed September 2010.