The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Small Bowel Tumor Detection for Wireless Capsule Endoscopy Images Using Textural Features and Support Vector Machine

Baopu Li, Max Q.-H. Meng, *Fellow IEEE*

Abstract—**Wireless capsule endoscopy (WCE) has been gradually applied in hospitals due to its great advantage that it can directly view the entire small bowel in human body compared with traditional endoscopies and other imaging techniques for gastrointestinal diseases. However, a challenging problem with this new technology is that too many images produced by WCE causes a tough task to doctors, so it is very significant to help and relief the clinicians if we can develop computer based automatic detection system to prescreen the collected large amount of images and identify the images with potential problems. In this paper, we propose a new scheme aimed for small bowel tumor detection of WCE images. This new scheme utilizes texture feature, also a powerful clue used by physicians, to detect tumor images with support vector machine. We put forward a new idea of wavelet based local binary pattern as the textural features to discriminate tumor regions from normal regions, which take advantage of wavelet transform and uniform local binary pattern. With support vector machine as the classifier, three-fold cross validation experiments on our present image data verify that it is promising to employ the proposed texture features to recognize the small bowel tumor regions.**

I. Introduction

DISEASES related to gastrointestinal (GI) tract greatly threaten human's health now. It has been reported that colorectal cancer has been the second leading cause of cancer-related deaths in U.S. [1]. Many diseases in the GI tract can be prevented and cured if early detection is possible. The traditional detection methods such as endoscopy, ultrasound, and computed tomography (CT) scan have demonstrated great values in diagnosing diseases of the digest tract. However, the main body of the GI tract, small intestine, cannot be reached by the traditional endoscopies due to their limitations. Moreover, they may have other drawbacks such as inconvenience, invasiveness, illegibility, and so on. In 2000, a new kind of GI endoscopy, i.e. wireless capsule endoscopy (WCE), was created. This new technology of endoscopy, developed by Given Imaging corporation in Israel, almost revolutionize the diagnosis methodology for the digestive tract since this small device can directly view the entire small intestine without pain, sedation, or air insufflation for the first time, and these significant breakthroughs make it rapidly used in most hospitals to detect the status of the GI tract.

As shown in Fig.1 and Fig.2, wireless capsule endoscopy, measuring $26\text{mm} \times 11\text{mm}$, is a pill-shaped device which consists of a short-focal-length CMOS camera, light source, battery and radio transmitter. We first introduce how it works briefly. After a WCE is swallowed by a patient who has a diet for about 12 hours, this little device propelled by peristalsis starts to work and record the images while moving forward along the digestive tract. Meanwhile, the images recorded by the camera are sent out wirelessly to a special recorder attached to the waist. This process continues for about eight hours until the WCE battery ends. Finally, all the image data in the special recorder are downloaded into a personal computer or a computer workstation, and physicians can view the images and analyze potential sources of different diseases in the GI tract. It should be noted that the diagnosis process exerted by physicians is very time-consuming due to the large amount of video, so the diagnosis is not a real-time process. This situation paves a potential way for off-line post processing and computer aided diagnosis. The WCE was approved by U.S. Food and Drug Administration (FDA) in 2001, and it has been reported that this new technology shows great value in evaluating gastrointestinal bleeding, Crohn's disease, ulcer and other diseases existed in the digestive tract [2].

Fig.1 Wireless capsule endoscopy

There still remains some improvement for the WCE, although this new technology shows great advantages over traditional examination techniques. One problem associated with this technology is that it would take a long period of time for physicians to inspect the large number of images it produced. There are about 50,000 images in total per examination for one patient, and it costs an experienced

The authors are with the Department of Electronic Engineering at The Chinese University of Hong Kong, Hong Kong, SAR, China, email: {bpli & max}@ee.cuhk.edu.hk This research is supported by SHIAE project #8115021 of the Shun Hing Institute of Advanced Engineering of The Chinese University of Hong Kong, awarded to Max Meng.

clinician about two hours on average to review and analyze all the video data [2]. Besides, abnormalities in the GI tract may be present in only one or two frames of the video, so they might be missed by physicians due to oversight sometimes. Moreover, there may be some abnormalities that cannot be detected by the naked eyes because of their size, color, texture and distribution. Furthermore, different clinicians may have different findings when come to the same image data. All these problems motivate the researchers to develop reliable and uniform assisting approaches to reduce the great burden of the physicians. However, it should be admitted that this goal is very challenging because the true features associated with the diseases are not exactly known. Moreover, different diseases have totally different symptoms in the digestive tract. Even the same disease shows great variations in color and shape.

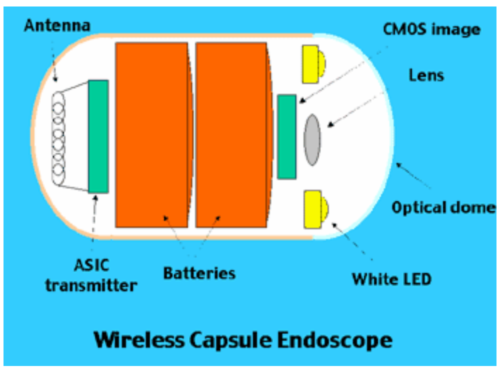


Fig.2. Components diagram of a WCE

Because of its gradually wide application, some studies have been investigated towards the direction of partially automating inspection of the WCE images so as to decrease the burden of doctors. The manufacturer itself also provides a software tool to detect the bleeding region, however, the accuracy of this system was reported to be very low [3]. Using expectation maximization clustering, the authors introduced a new technique to do bleeding detection with an encouraging result as stated in [4]. The authors in [5] proposed a method using color distribution to discriminate stomach, intestine and colon tissue. An interesting way of selecting the MPEG-7 visual descriptors as the feature extractor to do detection for several diseases such as ulcers and bleeding in the gastrointestinal tract was advanced in [6]. In paper [7], the authors proposed an algorithm to reject those invalid images through automatic detection of intestinal juices. We have investigated bleeding region detection and ulcer region detection for WCE images in [8] and [9], respectively, the preliminary experimental results show that our proposed scheme works fine for WCE images.

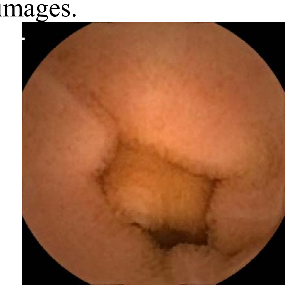
As small bowel tumor is one of the most common diseases in GI tract, we focus on small bowel tumor region recognition in this paper. To achieve this goal, we propose a new scheme that exploits texture features and support vector machine (SVM). The proposed new textural feature extraction method combines wavelet transform and uniform local binary pattern to differentiate between normal image and tumor image, leading to a better discrimination ability of the textural features. Experimental results on our present data show that this new scheme achieves a satisfying performance of tumor detection when using SVM as the classifier.

The remainder of this paper is designed as follows. The new method of texture feature extraction using wavelet transform based local binary pattern is discussed in the following section. In section III, the classifier used to implement verification of the new features by means of SVM classification is presented. Section IV gives the experimental results, and we draw some conclusions and make some discussions at the end of this paper.

II. TEXTURE ANALYSIS

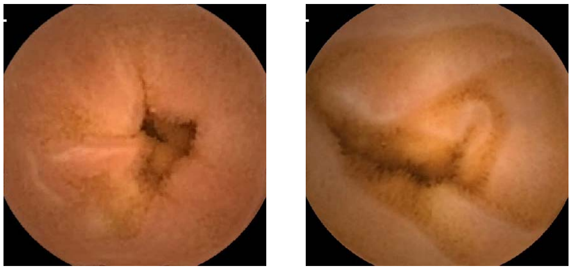
Image texture, defined as a function of the spatial variation in pixel intensity, is very useful in a variety of applications and has been intensively investigated in the literature. Generally speaking, texture analysis methods can be roughly categorized into statistical, geometrical, model-based and signal processing. Early works mainly concentrated on the analysis of statistical properties of the texture which deals with the spatial distribution of gray values. Some typical examples are co-occurrence matrix and autocorrelation function. Geometrical approaches consider texture as the combination of the texture primitives. A few stochastic models have been proposed for textural model such as Gaussian Markov random fields. The signal processing methods are based on texture filtering for extracting the features either in spatial domain or in frequency domain. Typical examples are Gabor filter and wavelet based texture extraction emerged in the last decade. Filter bank rather than a single filter has also been proposed such as Gabor filter. The major disadvantage of the Gabor transform is that its output are not mutually orthogonal, resulting into a significant correlation between texture features.

The normal regions and abnormal regions in WCE images can also be differentiated as texture information is one of the primary features analyzed by the clinicians. As illustrated in Fig.3 and Fig.4, respectively. The normal small bowel images and the small bowel images with tumor show different textural characteristics on its mucosa surface. This interesting property encourages us to investigate the color texture features of these images. We first discuss the texture feature analysis on gray images and extend the proposed feature extraction scheme to color space since WCE images are color images.



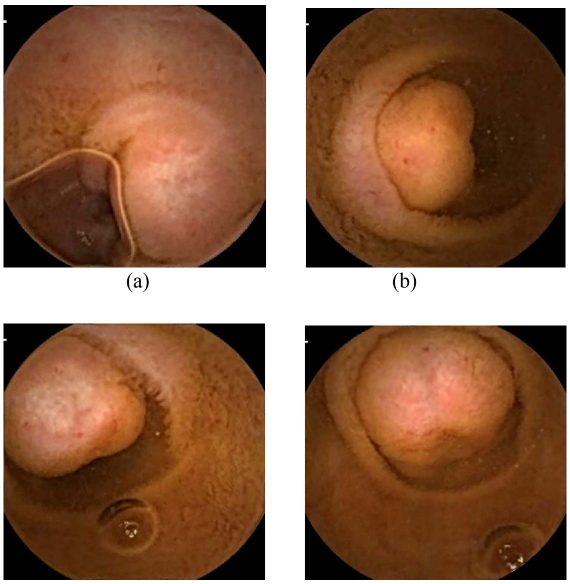
(a)

(b)



(c) (d)

Fig.3 Some representative normal small bowel WCE images



(a) (b)

(c) (d)

Fig.4 Some representative small bowel WCE images with tumor

A. Local Binary Pattern

Ojala et al.[10] proposed the famous local binary pattern (LBP) texture operator, which is invariant against any monotonic grey scale transformation and is computationally simple. This technique is based on the two level version of the texture spectrum method, and it describes the spatial structure of the local image texture. The image pixels are first labeled by thresholding the difference between the central pixel and its neighbors using the step function. Then the values of the pixels in the thresholded neighborhood are multiplied by the binomial weights given to the corresponding pixels. Finally, values of the products are summed up to obtain an LBP number of this neighborhood. This process can be better explained in Fig.5. The LBP of a 3×3 neighborhood produces up to $2^8 = 256$ local texture patterns, and the 256-bin occurrence LBP histogram computed over a region is then employed for texture description.

They continued to introduce a simple yet efficient multi-resolution approach to gray-scale and rotation invariant texture based on local binary pattern [11]. In this paper, they found that some local binary patterns are fundamental, and these fundamental patterns are called 'Uniform'. The

6	1	2
7	5	4
8	9	7

Example

1	0	0
1		0
1	1	1

Thresholded

1	2	4
128		8
64	32	16

Weights

LBP=1+16+32+64+128=241

Fig.5 Calculation of LBP

uniform patterns have circular structure that contains few transitions from 0 to 1. In order to formally define the 'Uniform' patterns, a uniformity measure U was introduced, which corresponds to the number of spatial transitions (bitwise $0/1$ changes) in the pattern. For instance, both pattern 00000000 and pattern 11111111 have U values of 0, and patterns that have U values of at most 2 are designated as "uniform". Based on the above discussions, a new rotation invariant operator is defined:

$$LBP^{riu2}_{P,R} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & \text{if } U(LBP_{P,R}) \le 2 \\ P+1 & \text{otherwise} \end{cases} \quad (1)$$

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

and $s(x)$ is the sign function. This operator is an excellent measure of local image texture, and it is robust to monotonic transformation of gray scale and also simple to implement.

In the imaging process of WCE, the images suffer from illumination variation due to the specific imaging circumstances such as motion of the camera, the rather limited range of the illumination in the digestive tract. As such, it is necessary to consider illumination variation effects on textures of WCE images because texture features are not constant to illumination variation. Uniform local binary pattern demonstrates rather robust performance to illumination change. Moreover, the uniform LBP has shown its great success in classifying textures. Due to the above reasons, we choose LBP as the basis for texture feature extraction.

B. Discrete Wavelet Transform

In the past few decades, wavelet theory has been widely used for texture analysis because wavelet provides a powerful tool for multi-resolution analysis of the image. A lot of studies have investigated the discrimination ability of the wavelet-based features in various applications such as ultrasound image [12] and mammogram [13]. As has already been shown in Fig.4, the tumors in WCE images show great variations in size. This characteristic naturally motivates us turn to wavelet transform since it can efficiently overcome this problem. Since wavelet transform for image can be achieved with discrete wavelet transform (DWT), we briefly introduce it in this section.

The DWT is similar to a hierarchical sub-band system where the sub-bands are spaced in frequency domain. For 2D images, DWT is implemented with a separable filter-bank to the image [14], and the image is convoluted with a lowpass filter H and highpass filter G recursively. Due to the decomposition of an image using DWT, the image is transformed into four sub-images which are generally denoted as LL, LH HL and HH. The LL sub-image comes from low pass filtering in both directions and it is the most like original picture, so it is called the approximation component. The remaining sub-images are called detailed components. The HL is derived from low pass filtering along the vertical direction and high pass filtering along the horizontal direction and so has the label HL. The other two sub-images LH and HH have similar explanations. In our study, we apply three levels DWT to each channel of the WCE image, and Fig.6 illustrates such a representation of one color channel for this transformation.

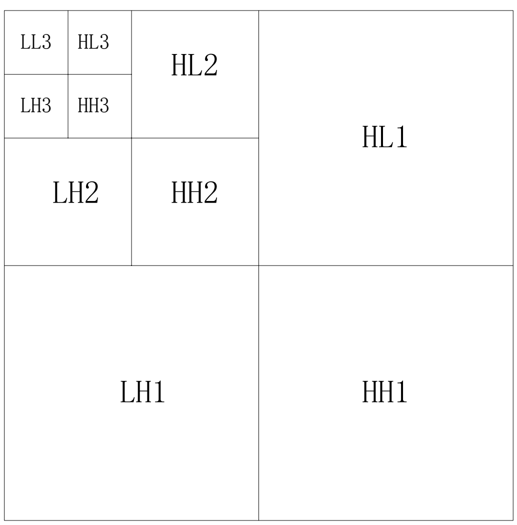


Fig. 6 Three-Level image decomposed using DWT for one channel

As has been demonstrated that the textural features are better encoded in the middle wavelet detailed sub-images [15], we choose the middle level sub-images, i.e., HL2, LH2 and HH2, as the basis for the textural feature analysis. For color images, three channels will lead to nine such sub-images. In this study, we further apply the aforementioned LBP to each sub-image to describe the color textural features for WCE images. Using the uniform LBP histogram of each sub-image in each channel, we can further obtain six statistical measurement of the histogram as the features of the texture [16] in order to reduce the number of features. Hence, each WCE image can be characterized with a feature vector with 54 ($6\times3\times3$) elements.

Concerning the color space, we will investigate the proposed feature extraction methods in RGB space and HSI space respectively. The reason why we choose them is because RGB color space is the most convenient color space used in industry and literature, and HSI color space is a representative space that separates color information into chromaticity and intensity.

III. SUPPORT VECTOR MACHINE

Support vector machine (SVM) is one kind of state-of-the-art classifiers used in a lot of applications because they have many advantages such as minimum classification structural error, good ability to handle the high dimensional problems, and so on. Since proposed in 1990's by Vapnik [17], this powerful classification technique has been a hot topic in the field of machine learning.

An SVM constructs a binary classifier from a set of labeled patterns called training sets. Let $(x_i, y_i) \in R^N \times \{-1,+1\}, i=1...m$ be such a set. The purpose here is to select a function $f:\ R^N \rightarrow \{\pm 1\}$ from a given class of functions such that f correctly classifies the test data (x, y).

Based on the above discussions, the SVM algorithm is able to find a hyper-plane defined by the equation

$$\omega\,\Phi(x)+b=0 \tag{2}$$

such that the margin of separation is maximized, where $\Phi(x)$ is a nonlinear mapping from the input space to the feature space. It was shown that [17][18] for the maximal margin hyper-plane,

$$\omega=\sum_{i=1}^{N}\lambda_i y_i \Phi^T(x) \tag{3}$$

where λ_i are the Lagrange multipliers that can be estimated through the maximization of

$$L_D=\sum_{i=1}^{N}\lambda_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j K(x_i, x_j) \tag{4}$$

with respect to λ_i, where the following constraints hold at the same time: $\sum_{i=1}^{N}\lambda_i y_i=0$ and $0\le\lambda_i\le c$. In (4), $K(x_i, x_j)$ is called a kernel function and it is defined as the inner product $K(x_i, x_j)=\Phi^T(x_i)\Phi(x_j)$. Linear, polynomial, radial basis function (RBF) and sigmoid are among the most common functions used as SVM kernels, and the Gaussian RBF usually performs better than other kernels.

The hyper-plane separating the two classes is found to be given by

$$\sum_{i=1}^{N}\lambda_i y_i K(x_i, x)+\omega_0=0 \tag{5}$$

Given a test input vector x , the trained SVM yields an output which corresponds to the label of the class it belongs to, namely,

$$s=sign\left\{\sum_{i=1}^{N}\lambda_i y_i K(x_i, x)+\omega_0\right\} \tag{6}$$

where sign is a function that returns +1 for positive and -1 for negative inputs.

IV. EXPERIMENT RESULTS

Experts of GI tract selected a data set composed of 300 representative abnormal (150) and normal (150) images from 2 patients' video data. These images are obtained from the second generation of WCE, i.e., PillCam SB2, and the effective resolution of these images is 512×512 . The original images are manually labelled to provide the ground truth. The image containing any abnormal region is labelled as a positive sample; otherwise, it is labelled as a negative sample. In order to prevent over-fitting of the classification results, we exploited three-fold cross-validation for all our classification experiments.

To demonstrate the performance of the proposed algorithm, we compare the proposed scheme with pure local binary pattern features extracted from RGB color space and HSI color space, and the color wavelet covariance (CWC) features used in [19], respectively. CWC features are very new techniques to describe the color features that are built upon the covariance of 2nd-order textural measures in the wavelet domain of the color channels of the images.

The success of classification of WCE image status using SVM is measured by accuracy, specificity, and sensitivity, respectively, which are widely employed by colleagues to assess the performance of classification. Here is the definition of them:

$$\text{Accuracy}=\frac{Number\ of\ Correct\ \ predictions}{Number\ of\ \ Positives + Number\ of\ \ Negatives} \tag{7}$$

$$\text{Specificity}=\frac{Number\ of\ Correct\ \ Negative\ \ predictions}{Number\ of\ \ Negatives} \tag{8}$$

$$\text{Sensitivity}=\frac{Number\ of\ Correct\ Positive\ \ predictions}{Number\ of\ \ Positives} \tag{9}$$

Concerning the DWT implementation, we make use of Harr wavelet for its superior discriminating power which is demonstrated in [20]. For the implementation of SVM, we referred to the work of Chang and Lin [21], and RBF was found to be the kernel function that yielded the best classification performance in our experiments. Seven sets of different parameters while tuning the SVM were experimented, the best classification results among these seven sets of parameters were considered as the performance of the classification for SVM, and we recorded the average performance of the three-fold cross-validation experimental results in RGB color space and HSI color space in table 1 and table 2, respectively. Table 1 illustrates the classification result for the proposed texture analysis method, i.e., wavelet based local binary pattern. Table 2 shows the corresponding recognition results when using the local binary pattern as the textural features. The recognition rate using CWC method is shown in Table 3.

From these three tables, we can conclude that the features extracted from the proposed method show superior performance to local binary pattern with an accuracy improvement of 19% and 7% in RGB and HSI color space, respectively; while a margin of 18.83% and 22.66% in RGB and HSI compared to those of CWC. Moreover, the accuracy of the proposed scheme in HSI color space reaches up to 96.67%, together with a specificity of 97.33% and a sensitivity of 96%. This illustrates that wavelet based local binary pattern has a better performance for tumor recognition in WCE images. Compared to LBP, the reason why the proposed features show better recognition rate for tumor detection in WCE images may be due to the fact that multi-scale based texture features can further improve the discrimination ability of the textural features compared to the single scale based texture. Compared to CWC, the reason may be ascribed to that it is robust to illumination changes for different images obtained under unstable circumstances in the GI tract.

TABLE1 CLASSIFICATION RESULTS USING WAVLET BASED LOCAL BINARY PATTERN (%)

	RGB	HSI
Sensitivity	91.33	97.33
Specificity	97.67	96.00
Accuracy	93.67	96.67

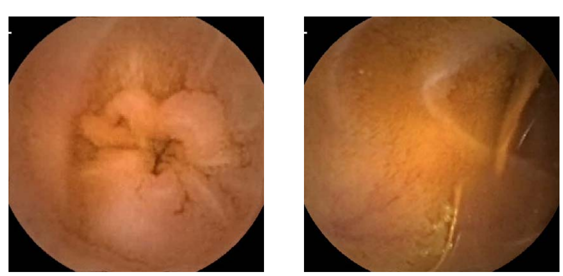
TABLE2 CLASSIFICATION RESULTS USING LOCAL BINARY PATTERN (%)

	RGB	HSI
Sensitivity	77.67	84.00
Specificity	80.67	95.33
Accuracy	74.67	89.67

TABLE3 CLASSIFICATION RESULTS USING CWC (%)

	RGB	HSI
Sensitivity	72.50	74.67
Specificity	65.67	79.67
Accuracy	78.33	69.67

Although our new scheme shows very promising performance of tumor detection for WCE images, some tough cases are still hard to judge correctly. Fig.7 illustrates two normal cases that the proposed scheme fails to judge them as tumor images (false positive detection).



(a) (b)

Fig. 7 Two normal cases failed to judge

The reason why the proposed scheme fail to recognize Fig.7(a) may be due to the subtle pump regions in the center of the WCE image. As for Fig.7 (b), it may be because the bubble region in the right corner of the region. The false positive cases motivate us to further refine the proposed scheme so as to reduce the false positive number during detection.

V. CONCLUSION

A new scheme of using multi-scale texture features and support vector machine to detect small bowel tumor for WCE images has been proposed in this paper. The novel textural features combine the advantages of wavelet transform and local binary pattern, leading to more discriminative ability for tumor detection in WCE images. Experiments on our present small bowel WCE images show that this method is promising in detecting tumor images. Future research will be directed to collect more patients' data so as to test the robustness of the proposed scheme. Moreover, there is still some room for detection accuracy improvement for the proposed scheme, so another possible avenue of our future work is to investigate specified classifiers that show better recognition results.

ACKNOWLEDGMENT

We would like to show our sincere thanks to James Lau, a professor in Prince of Wales Hospital in Hong Kong, for providing us the wireless capsule endoscopy image data mentioned in this paper.

REFERENCES

[1] S.Parker, T. Tong, S.Bolden, and P.Wingo, "Cancer statistics 1997," CA Cancer J. Clinicians, vol.47, pp.5-27,1997.

[2] D. G. Adeler, C. J.Gostout, "Wireless capsule endoscopy," Hospital Physican , pp.14-22, May, 2003.

[3] R.D.Francis et al, "Sensitivity and specificity of the red blood identification (RBIS) in video capsule endoscopy," 3Rd, Int. Conf. on Capsule Endoscopy, Feb.2004, Miami, FL, USA.

[4] Hwang, J.H. Oh, J. Cox, et al, "Blood detection in wireless capsule endoscopy using expectation maximization clustering," Proceedings of SPIE, on Medical Imaging, Volume 6144,2006, pp. 577-587

[5] J. Berens, M. Mackiewicz, and D. Bell. "Stomach, intestine and colon tissue discriminators for wireless capsule endoscopy images," Proc. of SPIE. Conference on Medical Imaging, volume 5747, Bellingham, WA, 2005, pp. 283 – 290

[6] M.T. Coimbra. J.P.S. Cunha, "MPEG-7 visual descriptors — contributions for automated feature extraction in capsule endoscopy," IEEE Transactions on Circuits and Systems for Video Technology, Volume 16, Issue 5, pp.628 – 637, 2006

[7] F.Vilarino, et al, "Automatic detection of intestinal juices in wireless capsule video endoscopy," 18th International Conference on Pattern Recognition, Volume: 4, pp. 719-722, 2006

[8]B. Li, Max Q.-H. Meng, "Computer aided detection of bleeding regions in capsule endoscopy images," IEEE Transactions on Biomedical Engineering, Vol. 56, No. 4, April, 2009, pp.1032-1039;

[9] B. Li, Max Q.-H. Meng, "Texture analysis for ulcer detection in capsule endoscopy images," Image and Vision Computing, Vol. 27, No. 9, August, 2009, pp.1336-1342;

[10] T. Ojala, M. Pietikainen, D.Harwood "A comparative study of texture measures with classification based on feature distributions," pattern recognition, vol.29, pp.51-59.

[11]T. Ojala, M. Pietikainen and T. Maenpaa "Multi-resolution gray-scale and rotation invariant texture classification with local binary pattern" IEEE Trans. on PAMI, vol.24, No.7, pp971-987,2002

[12] W.L. Lee, Y.C. Chen, K.S. H. "Ultrasonic liver tissues classification by fractal feature vector based on M-band wavelet transform," IEEE Transactions on Medical Imaging, vol.22, no.3, 2003, pp.382-392

[13] T.C. Wang, N.B. Karayiannis, "Detection of micro-calcifications in digital mammograms using wavelets " IEEE Transactions on Medical Imaging, vol.17, no.4, 1998, pp.498-509

[14]S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Machine Intell., vol.11, pp. 674–693, 1989.

[15]K. W. Abyoto, S. J. Wirdjosoedirdjo, and T. Watanabe, "Unsupervised texture segmentation using multiresolution analysis for feature extraction," J. Tokyo Univ. Inform. Sci., vol. 2, pp. 49–61, Jan. 1998

[16]M.Boulougoura, E.Wadge, "Intelligent systems for computer-assisted clinical endoscopic images analysis," proceedings of Second International Conference on Biomedical Engineering, Austria, Feb.16-18, 2004, pp. 405- 408.

[17]V. Vapnik, The Nature of Statistical Learning Theory. Springer Verlag, New York, 1995.

[18]C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, Kluwer Academic Publishers, Boston, 1998

[19] S.A. Karkanis et.al., "Computer-Aided tumor detection in endoscopic video using color wavelet features," IEEE Trans. Info. Tech. in Biomedicine. vol. 7, no.3, pp.141-152, Sep. 2003

[20]L. Semler, L. Dettori, J. Furst, "Wavelet-Based Texture Classification of Tissues in Computed Tomography," The 18th IEEE International Symposium on Computer-Based Medical Systems (CBMS'05), Dublin, Ireland, 2005, 265-270

[21] C.-C. Chang, C.-J. Lin. LIBSVM: a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/cjlin/libsvm.