# A Computer Aided Method to Detect Bleeding, Tumors, and Disease regions in Wireless Capsule Endoscopy

Omid Haji Maghsoudi
Department of Bioengineering
Temple University
Philadelphia, USA
o.maghsoudi@temple.edu

Mahdi Alizadeh
Department of Bioengineering
Temple University
Philadelphia, USA

Mojdeh Mirmomen
Radiology and Imaging Department
National institute of health
Bethesda, Maryland

***Abstract*— Wireless Capsule Endoscopy (WCE) is a relatively new technology to record the entire gastrointestinal (GI) tract, in vivo. A large amount of images (frames) are captured during the WCE examination. Reviewing this number of images by a gastroenterologist would be time consuming and prone to human error. Therefore, a diagnostic computer-aided technique is essential to detect and segment regions of abnormalities. In this study, a novel method based on textural features (such as Gabor filters, local binary pattern, and Haralick) in HSV color space, Fisher score test, and neural networks is presented to detect and differentiate regions such as bleeding, tumor, and other types of gastric diseases including Crohn's, Lymphangectasia, Stenosis, Lymphoid Hyperslasia and Xanathoma. The experimental results indicate that this method is able to classify a lesion from a normal region in every single frame and group them into normal and abnormal frames to be considered for surgery/treatment planning by an expert.**

***Keywords—Wireless capsule endoscopy; image segmentation, Gabor filter, neural network.***

## I. INTRODUCTION

Wireless capsule endoscopy (WCE) is a new device which is able to investigate the entire gastrointestinal (GI) tract without any pain and has been used to detect small bowel abnormalities such as tumors, bleeding, ulcers, Crohn's disease, and Celiac. A WCE examination records around 55,000 frames in approximately eight hours [1,2]. Considering the large number of frames captured, it is essential to use an intelligent software or automatic method for reviewing these frames. The majority of studies are limited to only finding bleeding regions in the frames but physicians also feel the need for the studies to also detect diseases.

For instance, in [3], the authors introduced a method for bleeding detection based on color similarity measurements. They used two color vector similarity coefficients to measure the degree of similarity in the RGB color space. Based on the similarity coefficients from the color vector, adequate classifiers were built and applied. This classifier was combined by the seeded region growing to implement an intelligent algorithm of bleeding detection. The sensitivity and the specificity were reported 97% and 90% respectively for detection of frames contaminated with bleeding, and 92% and 88% for detection of pixels that included bleeding. In another study [4], an intelligent technique for detection of bleeding was presented based on probabilistic neural network trained by features extracted from the RGB and the HSI color spaces to recognize bleeding pixels from other pixels. The sensitivity and specificity were 93% and 84%, respectively.

Baopu and Meng [5] presented a method based on chrominance moment as the color texture features. These color texture features were combined with uniform local binary pattern (LBP). A multilayer perceptron neural network was used to detect the bleeding regions in WCE images with a reported sensitivity of 90%.

Kumar introduced a technique [6] using edge features in four angles, color features based on the LUV color space, and texture features using Gabor filters. After extracting the features, a support vector machine (SVM) was used to classify the diseased regions. The accuracy, precision and specificity were 93%, 93%, and 97%, respectively. In addition, Baopu and Meng [7] presented a method using uniform LBP and wavelet transform. An SVM was applied to select tumor region in WCE frames. The accuracy and sensitivity of this method was reported respectively 92.4% and 88.6%.

In [8], Karargyris introduced a method to detect ulcers and polyps based on Log Gabor filters. In addition to the Gabor filter, Susan edge detector was used to extract texture features such as center of curvature (type of geometric feature to show the shape characteristic of the segmented object) for polyps. SVM classifier was then used to classify tissue in a frame using all the extracted features.

A method was presented for detection of small bowel tumors using level set method [9]. An adaptive gamma correction method was used to increase the contrast in the frames. The level set method and the snake method were then compared. The results showed that the sensitivity increased to 88% by a greater value for gamma correction while sensitivity decreased to 22%.

The previously mentioned studies, were limited to the detection of only one or two gastric diseases, but there were a whole array of diseases that were left undetected. In this work, we have presented a method for the detection of informative regions in WCE frames [10]. The method is based on chromatic and achromatic texture features by using neural

networks to differentiate between uninformative (like intestinal juice, bubbles, and dark regions in the frames) and informative regions.

In addition, detection of abnormalities using fuzzy local binary pattern was investigated in [11]. Fuzzy local binary pattern and local binary pattern were compared for segmentation of diseased regions in WCE for four color channels (red, green, gray, and hue).

The subtraction of the background can be helpful for the segmentation of objects. To achieve this goal, the methods proposed using sparse decomposition [12-14] would be useful on WCE frames.

The method presented here is able to distinguish among normal, bleeding, diseases (such as Lymphangectasia, Xanathoma, Stenosis, Crohn's, and Lymphoid hyperplasia), and tumors in the GI tract.

## II. MATERIAL AND METHOD

### A. Bleeding, Diseases and Tumors

We limited our study to three major objects in the frames: tumors (especially sub-mucosal tumors), bleeding, and diseases. The disease group consists of five objects which are Crohn's, Lymphangectasia, Lymphoid hyperplasia, Xanathoma, and Stenosis [14].

### B. Texture features

#### 1) Gabor filter bank

The Gabor filter is the same as a sinusoidal plane of particular frequency and orientation, modulated by a Gaussian envelope before convolving with images. This filter has good localization properties in both spatial and frequency domains. The impulse response is:

$$X = x \times cos(Theta) + y \times sin(Theta)$$
$$Y = -x \times sin(Theta) + y \times cos(Theta)$$
$$Sinusoidal = e^{(i\times(2\pi f \times X + Phase))}$$
$$Guassian = \frac{1}{2\pi\sigma_x\sigma_y} \times e^{\frac{-\left(\left(\frac{X}{\sigma_x}\right)^2+\left(\frac{Y}{\sigma_y}\right)^2\right)}{2}}$$
$$Gabor\ filter = Guassian \times Sinusoidal \quad (1)$$

where $x$ and $y$ are the filter ranges, Theta is the rotation degree in scale to $x$ and $y$ axis, $\sigma_x$ and $\sigma_y$ are sigma factors which determine Gaussian envelope parameters, $f$ is frequency repetition of sinusoidal part and Phase is the parameter which determine the symmetry to the center of sinusoidal [15]. In this paper, eight different parameters were used for the Gabor filter to create a filter bank as follow:

$$f = [.2\ .2\ .2\ .2\ .2\ .2\ .2\ .2];Theta = [0\ 45\ 90\ 135\ 0\ 45\ 90\ 135]$$
$$Phase = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0];Range = [20\ 20\ 20\ 20\ 20\ 20\ 20\ 20]$$
$$\sigma_x = [1.5\ 1.5\ 1.5\ 1.5\ .5\ .5\ .5\ .5];\sigma_y = [.5\ .5\ .5\ .5\ 1.5\ 1.5\ 1.5\ 1.5] \quad (2)$$

#### 2) Haralick features

A total of 18 features from Haralick gray level co-occurrence matrix (GLCM) were extracted. These features are included: autocorrelation, cluster prominence, energy, cluster shade, dissimilarity, contrast, entropy, homogeneity, maximum probability, correlation, sum of variance squares, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation and inverse difference momentum [16 and 17].

#### 3) Local binary pattern (LBP)

LBP is one of the efficient methods for distinguishing among different textures. It has been used frequently to show the differences in WCE frames [5,18].

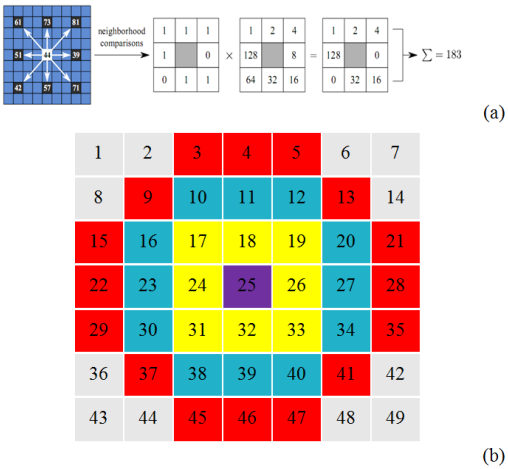


Figure 1. Image (a) shows LBP value calculation for a sample block and Image (b) demonstrates the three blocks in the equation (8).

The function of T is defined as texture in a local neighborhood of the grayscale image with the gray levels P+1,

$$T = t(g_c, g_0, \ldots, g_{P-1}), \quad (3)$$

where $g_c$ shows the center pixel value in each local neighborhood, and $g_p$(p = 0, . . ., P-1) is the gray value of pixels on a circle with radius of R which are symmetric sets of neighbors. The coordinates of the neighbors are given by ($x_c$ + R cos(2πp/P),$y_c$ - R sin(2πp/P)) in which $x_c$, $y_c$ are the coordinates of the center pixel. These circles are shown in figure 1 for three amounts of R and P. If the value of the center pixel is subtracted from the values of all neighbors, located in the circle (radius of R), then the texture function can be written as:

$$T = t(g_c, g_0 - g_c, \ldots, g_{P-1} - g_c). \quad (4)$$

Because the texture function values cover a large range (between -255 to 255), the function is redefined as follows:

$$T = t(s(g_0 - g_c), \ldots, s(g_{P-1} - g_c)), \quad (5)$$

where:

$$s = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6)$$

Now, the LBP value characterizes around $(x_c,y_c)$ based on the following equation [17 and 18]:

$$LBPP,R\ (xc,yc) = \sum_{p=0}^{P-1} s(g_{p-1} - g_c)2^p \quad (7)$$

In this paper, LBP is applied on frames by these radiuses:

$$R = \{1, 2, 3\}$$

$$\textit{Pixel numbers in a block (number of neighbors)} = \{8, 12, 16\} \quad (8)$$

Figure 1 (a) demonstrates how LBP calculates in a block with a radius of 1 and for 8 neighbors. In Figure 1 (b), the proposed block and its neighbors are shown.

### *D. Texture features in color space*

Texture features were extracted in two common and widely used color spaces (RGB and HSV color spaces) as well. RGB is an additive and subtractive model whereas HSV color space encapsulates information which is close by human interpretation [21].

Haralick features (Mean, Variance, Skewness, and Kurtosis) are extracted from 5 channels in color space (Red, Green, Blue, Hue, and Saturation) as well as the gray scaled channel. Also, the Green and gray scaled channels were used to extract the LBP features.

## III. THE PROPOSED ALGORITHM AND ITS RESULTS

Each frame was divided to 256 sub-images by resolution of 32×32 (Figure 2) to prepare the image for feature extraction.

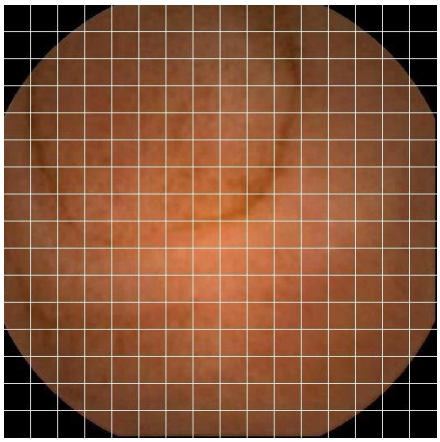

Figure 2. Creating 32×32 sub-images.

To extract features, first, LBP was applied on each sub-image. LBP was extracted for three radii (1, 2, and 3). The total of 26×26×3 = 2028 LBP values for one sub-image (three radii and 26 pixels center in each LBP block) were created. The LBP values were counted based on the ranges mentioned in the three following histograms as the function of the radius of LBP:

H1 (16n) = [1,3,7,15,31,63,127,255,511,1023,2047,4095, 8191,16383,20000,26000,32768]

H2 (12n) = [1,3,7,15,31,63,127,255,511,1023,1500,2047]

H3 (8n) = [1,3,7,15,31,63,127,255]

Using these vectors and counting the number of LBP values located between two numbers, 18 features for the radius of 3; 12 features for radius of 2; and 7 features for radius of 1 were calculated. Therefore, in total, 74 features from the gray scaled image and the green channel for each sub-image were computed. The green channel was used because the intensity values of the gray scaled image is similar to the red channel. Therefore, the extracted features from the green channel contained different information.

In addition, other features from the LBP method were obtained. We clarified which pixels had larger a value than the center for each radius. This means that another histogram was used to distinguish which neighbor was more important and also which one was repeated more than others. This is a new feature that can be extracted from the LBP values. In order to achieve this feature, neighbors were labeled by the following series:

C1 = [3 4 5 13 21 28 35 41 47 46 45 37 29 22 15 9]

C2 = [10 11 12 20 27 34 40 39 38 30 23 16]

C3 = [17 18 19 26 33 32 31 24]

and the orders were:

O1 = [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16]

O2 = [1 2 3 4 5 6 7 8 9 10 11 12]

O3 = [1 2 3 4 5 6 7 8].

We then counted repeating numbers in each sub-image (26×26 = 676 times). Using this series, the number of features increased to 110.

Also, for the gray scaled sub-image, Haralick features were extracted based on co-occurrence matrix that incorporated four different angles (0, 45, 90, and 135). Therefore, the total features increased to 110 + 92 = 202.

Two texture features remained: features extracted using the Gabor filters and the colored ones. We created the Gabor filter bank with 8 different parameters (two different frequencies and four angles). By applying each Gabor filter on the sub-image, four features (Mean, Skewness, Kurtosis, and Entropy) were extracted. The outcome of the Gabor filter in four different angles created a summation image (the average of the four filtered images), and features from the generated image were extracted. So, the features derived using the Gabor filter totaled 50.

As mentioned above (in the colored feature section), four features (Mean, Variance, Skewness, and Kurtosis) were extracted from five colored channels (Red, Green, Blue, Hue, Saturation) and the gray scaled image for each sub-image. 202 (Haralick and LBP) + 50 (Gabor) + 24 (colored features) = 276 features were extracted for each sub-image.

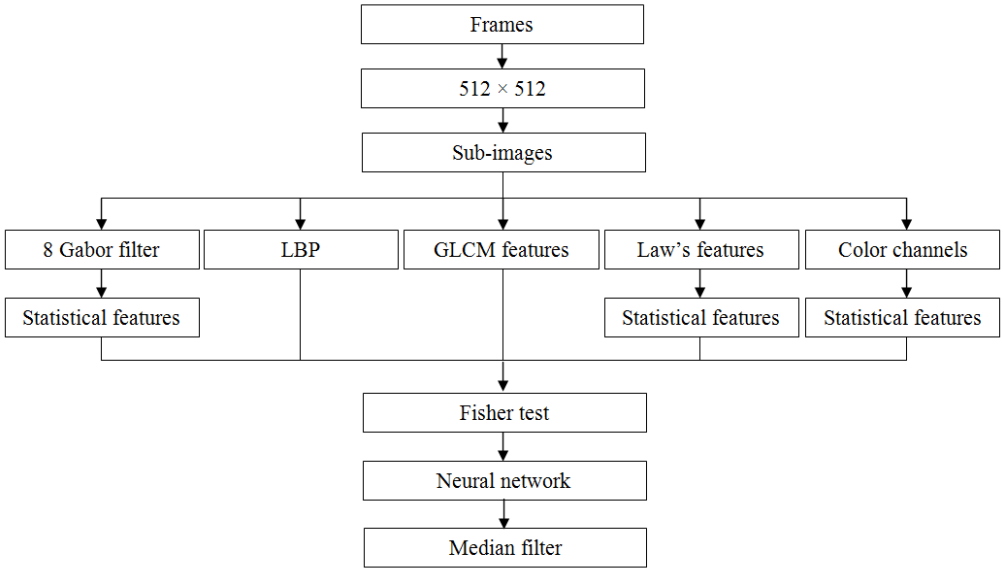


Figure 3. Block digaram for the method used in this study.

The Fisher scoring was applied to select features carrying higher information of a typical image [22]. Using Fisher scoring, 276 features were reduced to 30 providing the required information for classifiers. Finally, Multi-Layer Perceptron (MLP) neural network [23] was used to classify frames. Three networks were designed (three neural networks to distinguish between (1) normal and tumor regions; (2) normal and disease regions; and (3) normal and bleeding regions in the frames) with three hidden layers. The input layer, the first hidden layer, the second layer, the third layer, and the output layer were designed respectively by 30 neurons (number of features), 20 neurons, 10 neurons, 10 neurons, and two neurons. Figure 3 illustrates how the features discussed are extracted and these features are used to train the neural network.

## IV. RESULTS

In this study, the Given imaging videos were used [24]. Each video time was 20 seconds and we digitized them to 100 frames. Our aim was to distinguish between normal tissue and the abnormal regions in the frames.

73 tumor frames (from 12 videos), 44 completely normal frames containing no suspected objects (from 16 videos), 33 frames contained bleeding (from 18 videos), and 73 disease frames (from 39 videos) were selected. In other words, as the detection rate was not the frames but the subimages, tumor, bleeding, normal, and disease sub-images used for this study were respectively 18820, 16580, 33300, and 20195 (Crohn's 4640, Lymphangectasia 1219, Stenosis 9733, Xanathoma 3423, and Lymphoid hyperplasia 1180).


Figure 4. The two rows that are labeled as 1 show the original frame and its final segmented part is demonstrated in row 2.

Table I. The selected features for the training of each neural network.

| | LBP | Co-occurrence | Gabor | Color Channels |
|---|---|---|---|---|
| Total Extracted Features | 110 | 92 | 50 | 24 |
| Selected Features for Bleeding NN | 3 | 12 | 0 | 15 |
| Selected Features for Tomur NN | 12 | 6 | 7 | 5 |
| Selected Features for Diseases NN | 10 | 10 | 3 | 7 |

Table II. The quantitative measurements for the proposed method.

| | Tumor | Bleeding | Disease |
|---|---|---|---|
| Accuracy | 0.91176 | 0.97468 | 0.94901 |
| Precision | 0.8442 | 0.95162 | 0.90449 |
| Sensitivity | 0.9273 | 0.97331 | 0.96705 |
| Specificity | 0.90295 | 0.97536 | 0.93807 |

As discussed, 110 features were calculated respectively using H1, H2, H3, O1, O2, and O3 histograms from LBP. A total of 92 and 24 features were extracted from co-occurrence matrix. The following features selected after using Fisher scoring for each of the three neural networks are listed in Table I.

20% and 80% of the frames (subimages extracted from that frame) were selected for testing and training of neural networks using a random function in Matlab 2015. These neural networks are called: tumor neural network, disease neural network, and bleeding neural network. The performance of these neural networks were respectively 0.1028, 0.0548, and 0.0247.

After these networks, each frame was smoothed using the median filter. The median filter was applied on an image with 25 pixels (a window size of 25 pixels). Figure 4 shows the output of the method for some sample frames.

We estimated our method by measuring sensitivity, specificity, accuracy, and precision [25, 26]. Table II shows the measures for the test samples. The results are illustrated in Figure 5.

## V. DISSCUSSION

In this study, we presented a method to distinguish between normal and abnormal regions in WCE frames.

Table I shows the importance of the features for different applications. The Gabor features played no role in the detection of bleeding regions, nor in the three final selected features from LBP. The color features from the chromatic channels were the most dominant features behind the co-occurrence matrix. This fact proves the importance of color information for detection of the bleeding regions.

The role of color features mentioned above were less for the detection of tumors and diseases. However, all types of features were among the final selected features. The LBP features were more important in these two cases.

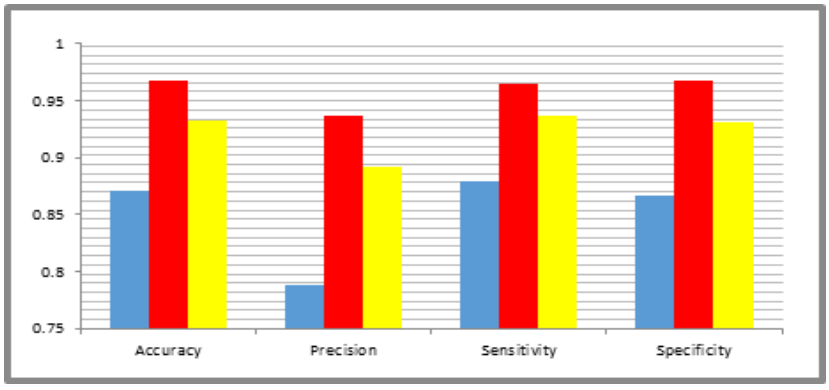


Figure 5. The quantitative measurements for detection of tumors (blue bars), bleeding (red bars), diseases (yellow bars).

Notice that a global median filter was undertaken to assign a value for undetected diseased pixels from neural networks. In other words, it smooths the segmented region and reduces the error rate of neural network. Each output pixel contains the median window size of 25×25 around the corresponding pixel (figure 4). Also, the performance of neural network for different hidden layers were investigated and our findings show that the best results were acquired by using only two hidden layers and it remained almost the same by increasing them.

The main aspect of this study that needs to be improved is the number of subjects and frames to achieve more reliable results, especially for diseases. In this study, we examined the method on available online data set [24]; however, this online data, as discussed, has so many samples for each of the groups studied here. We will present the results on a more comprehensive study using the clinical data available in hospitals.

In comparison to the available methods for detection of bleeding, the proposed method achieved respectively the sensitivity and specificity of 97% and 97%. While these two measures were reported respectively 88% and 92% for pixel detection (the method proposed here is for pixel detection) in [3]; 90% and 97% for frame detection in [3]; 84% and 93% for pixel detection in [4]; and a sensitivity of 90% in [5].

In the case of disease detection, the sensitivity and specificity were respectively 93% and 97% in [6] for detection of Crohn's disease. The reported sensitivity for tumor detection was less than 90% in [7] and [11]. Our proposed method showed a great improvement in the results as presented in these studies.

In the future, we will present a method for real time segmentation of gastric disease regions. A real time process can assist the manufacturer to add biopsy and treatments to WCE. To achieve this, we will try to use the information from different organs [27], sparse decomposition [28], and scattering transform [29].

## REFERENCES


[1] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature*, vol. 405, p. 417, 2000.

[2] P. Fritscher-Ravensand and P. Swain, "The wireless capsule: New light in the darkness," Dig. Dis., vol. 20, pp. 127–133, 2002.

[3] G. Pan, F. Xu, J. Chen, "A Novel Algorithm for Color Similarity Measurement and the Application for Bleeding Detection in WCE," Journal of Image, Graphics and Signal Processing, 2011.

[4] G. PAN, G. YAN, X. QIU, and J. CUI, "Bleeding Detection in Wireless Capsule Endoscopy Based on Probabilistic Neural Network," Journal of Medical Systems, 2010.

[5] B. Li, Max Q. H. Meng, "Computer-Aided Detection of Bleeding Regions for Capsule Endoscopy Images," IEEE Transactions on biomedical engineering, vol. 56, no. 4, April 2009.

[6] R. Kumar, Q. Zhao, S. Seshamani, G. Mullin, G. Hager, and T. Dassopoulos "Assessment of Crohn's Disease Lesions in Wireless Capsule Endoscopy Images," IEEE Transactions on biomedical engineering, Vol. 59, No. 2, February 2012.

[7] B. Li, Max Q. H. Meng, "Tumor Recognition in Wireless Capsule Endoscopy Images using Textural Features and SVM-based Feature Selection," IEEE Transaction on Information Technology in biomedicine, January 2012.

[8] A. Karargyris and N. Bourbakis, "Detection of Small Bowel Polyps and Ulcers in Wireless Capsule Endoscopy Videos," IEEE Transactions on biomedical engineering, vol. 58, no. 10, February 2011.

[9] M.Alizadeh, H.Soltanian Zadeh, O. Haji Maghsoudi, "Segmentation of Small Bowel Tumors in Wireless Capsule Endoscopy Using Level Set Method", IEEE, 2014 IEEE 27th International Symposium on Computer-Based Medical Systems, pp. 55-556, New York City, 2014.

[10] O. Haji Maghsoudi, A. Talebpour, H. Soltanian-Zadeh, M. Alizadeh, H. Soleimani Asl, "Informative and Uninformative Regions Detection in WCE Frames", Journal of Advanced Computing, vol. 3, issue 1, pp. 12-34, 2014.

[11] O. Haji Maghsoudi and H. Soltanian-Zadeh, "Detection of Abnormalities in Wireless Capsule Endoscopy Frames using Local Fuzzy Patterns", 20th Iranian conference in Biomedical Engineering (ICBME), pp. 286-291, 2013, Iran.

[12] Shervin Minaee, Yao Wang, "Screen Content Image Segmentation Using Robust Regression and Sparse Decomposition", IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2016.

[13] Shervin Minaee, Yao Wang, "Screen Content Image Segmentation Using Sparse Decomposition and Total Variation Minimization", IEEE International Conference on Image Processing, 2016.

[14] R. E. Kleinman, O. J. Goulet, G. Mieli-Vergani, "Walker's Pediatric Gastrointestinal Disease: Physiology, Diagnosis, Management". 5th ed., Hamilton, Ontario, Canada: BC Decker Inc, 2008.

[15] A. Jain, N. Ratha, and S. Lakshmanan, "Object detection using Gabor filters," Pattern Recognition, 30:295–309, 1997.

[16] R. M. Haralick, K. Shanmugam, I. Dinstein, "Textural Features for Image Classification," IEEE Trans. on systems, man and cybernetics, November 1973.

[17] http://murphylab.web.cmu.edu/publications/boland/, October

[18] B. Li, Max Q. H. Meng, "Small Bowel Tumor Detection for Wireless Capsule Endoscopy Images Using Textural Features and Support Vector Machine," The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems October 11-15, 2009 St. Louis, USA.

[19] A. Jarc, J. Pers, S.Kovacic, "A multi-stage registration method using texture features", Journal of Digital Imaging.

[20] T. Ojala, M. Pietik¨ainen and D. Harwood. Pattern Recognition 29, 51 (1996).

[21] S. K. Naik and C. A. Murthy, "Hue-preserving color image enhancement without gamut problem," IEEE Trans. on Image Processing, 12(12), 2003, 1591-1598.

[22] A. Fisher, "The Mathematical Theory of Probabilities," New York: Macmillan, 1923, vol. 1.

[23] C. T. Leondes, Algorithms and Architectures Neural Network Systems Techniques and applications, 1998.

[24] Given Imaging Ltd. Available: http://www.capsuleendoscopy.org/Pages/Home.aspx.

[25] http://en.wikipedia.org/.

[26] D. G. Altman and J. M. Bland, "Diagnostic tests. 1: Sensitivity and specificity," *BMJ*, vol. 308, (no. 6943), pp. 1552, 1994.

[27] Omid Haji-Maghsoudi, Alireza Talebpour, Hamid Soltanian-Zadeh, Navid Haji-maghsoodi "Automatic organs' detection in WCE", 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP), 2012.

[28] Shervin Minaee, Amirali Abdolrashidi, Yao Wang, "Screen Content Image Segmentation Using Sparse-Smooth Decomposition", Proceedings of the Asilomar Conference on Signals, Systems, and Computers 2015.

[29] Shervin Minaee, AmirAli Abdolrashidi, Yao Wang, "Iris Recognition Using Scattering Transform and Textural Features", IEEE Signal Processing Workshop 2015.