EMPIRICAL COMPARISON OF VISUAL DESCRIPTORS FOR ULCER RECOGNITION IN WIRELESS CAPSULE ENDOSCOPY VIDEO

Ouiem Bchir[1], Mohamed Maher Ben Ismail[1] and Nourah AL_Aseem[1,2]

[1]Computer Science Department, College of Computer and Information Sciences, King Saud University
[1,2]Computer Science Department,
College of Engineering and Computer Sciences,
Prince Sattam Bin Abdulaziz University

ABSTRACT

In this work, we empirically compare the performance of various visual descriptors for ulcer detection using real Wireless Capsule Endoscopy WCE video frames. This comparison is intended to determine which visual descriptor represents better WCE frames, and yields more accurate gastrointestinal ulcer detection. The extracted visual descriptors are fed to the ulcer recognition system which relies on Support Vector Machine (SVM) classification to categorize WCE frames as "ulcer" or "non-ulcer".

KEYWORDS

Visual descriptors, Ulcer detection, Wireless Capsule Endoscopy.

1. INTRODUCTION

A disease can be defined as a particular abnormal case, a disorder of structure or function, which impacts a specific side or all of an organism [1].Wireless Capsule Endoscopy (WCE) is an advanced technology which is used to recognize internal diseases[2], especially ulcer of the digestive tract. Its main advantages are flexibility, accuracy, pain free, and reasonable cost [3]. WCE system consists of three components: an electronic capsule sensing system, a recording device, and a computer for image review and interpretation[4].More specifically, WCE is a small capsule containing a miniature camera that is swallowed by the patient, and captures around 50,000 frames of the gastro-intestinal track that are transmitted to the receiver in a real time manner. The video is then uploaded to a computer for examination by the physician in order to identify gastrointestinal diseases. However, the recorded video is too long which makes its review awkward and time-consuming for physicians. Thus, reducing the examination time is required to make the analysis of the video less tedious[4].

David C. Wyld et al. (Eds) : IPPR, SOENG, DaMi, CSIT, AIS, CSE, CSIP, CCNET - 2018
pp. 01–09, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80501

Various types of ulcers which may affect the digestive tract can be detected using WCE. Namely, these types include: Peptic [5], Gastric[6], Duodenal[7], and Esophageal[8]. A peptic ulcer is the degradation of a tissue area by gastric juices that are produced by the stomach and the intestines to digest food. When the immunity system is weak, gastric juices attack the envelope of the gastro-intestinal track and results on peptic ulcers [9]. It is defined as arupture in the mucosal lining of the stomach [3]. For the case of Gastric ulcer, the tear is in the stomach[6]. Duodenal ulcer appears at the beginning of the small intestine [7] as an opening in the duodenum. A complicated case of acid reflux may lead Esophageal ulcer that appears at the extreme end of the esophagus[8].

Low-level feature extraction is one of the main components of any image analysis system. Their role is to convey the visual properties of an image to the recognition phase. However, choosing the appropriate visual descriptor for a specific recognition problem remains a challenging task for pattern recognition researchers. In particular, for ulcer detection, low-level visual descriptors extracted from the WCE frames do not encode and convey the same visual information to the recognition system. Thus, they do not contribute equally in the recognition power of the system. The keystone is then to identify the most discriminating visual descriptors. In other words, the answer to the question “Which visual descriptor yields the most accurate ulcer detection in WCE video frames” should be answered[6, 10]. Despite the researchers' efforts[11-12]to propose specific visual descriptors able to recognize ulcer in WCE video, no objective answer has been given to the question “What is the best descriptor to detect ulcer in WCE video?”. Moreover, none of the existing research compares the discrimination performance of these visual descriptors. In this research a comparison between the visual descriptors, used for ulcer detection WCE video frames, is conducted in order to determine which one discriminates the best between ulcer and ulcer-free frames.

The rest of this report is organized as follows. In section 2 we outline existing ulcer recognition techniques using WCE data. The Empirical comparison of visual descriptors for ulcer recognition is reported and analyzed in section 3. Finally, section 3 concludes this work and outlines potential future works.

2. LITERATURE REVIEW

Detecting and recognizing digestive ulcer using digital images as data modality is an active research field which has been promoted to assist physicians. This support aims at detecting ulcer using digestive endoscopy images [13]. During the last decade, various digestive ulcer detection techniques have been developed, and Wireless Capsule Endoscopy (WCE) emerged as an effective diagnostic tool. This technique enables doctors to gather much more gastric images. After obtaining the images, they are pre-processed. Then, visual descriptor extraction techniques are applied to encode the visual properties of the images. Finally, a supervised learning technique is launched to automatically detect ulcer frames.

In[14, 15], the researchers proposed texture visual descriptors to distinguish ulcer regions from normal regions. They used wireless capsule to take images from the inside of stomach and save them in a database. Then, new images are compared with this database in order to detect the place of the ulcer. In[14], the texture visual descriptors used to describe the images patterns are Curvelet transform[16] and local binary pattern[17]. The researchers used Neural Network [18]to classify the extracted visual descriptors and determine if there is an ulcer or not. In[15], the researchers used Curvelet transform, local binary pattern and YCbCr color[17]along with Neural Network and Support Vector Machines (SVM) to classify the images and determine if there is an ulcer or not[18, 19]. The visual descriptor used in [15] yield better accuracy than those in[14].

The researchers in [33]intended to recognize the gastrointestinal tract (GI) ulcer using digital image processing techniques. They used LM-LBP filter bank and local binary pattern as texture visual descriptor to recognize ulcer area[21]. Also, they used Image Block Dictionary (IBD) classifier and K-means clustering algorithm for training[21]. They extracted a significant number of image blocks (of size 128 by 128 pixels) of abnormal and normal textures from WCE and colonoscopy images. Then, the extracted image blocks were verified using the domain experts. Also, they used fixed size image blocks instead of regions obtained using segmentation algorithm, because it is simple to implement and fast to compute. In the detection phase, they evaluated unseen image in order to estimate whether it is abnormal or normal. After that, the image is scanned row and column wise using image blocks having a predefined overlap with the previous scan. A set of images was extracted from five real WCE videos for each abnormal and normal textures.

In[22], the authors studied the detection of gastrointestinal tract (GI) ulcer using color pattern as low-level visual descriptor. They used CIE-lab color visual descriptor[23]to represent WCE frames. The researchers used a total of 1370 representative frames captured during 252 WCE procedures with MiroCam [24]at the Royal Infirmary of Edinburgh. They used the MiroCam capsule endoscope[24] which has a frame rate of 3 frames per second, and an image resolution of 320x320 pixels. Also, they used Support Vector Machine (SVM) to classify the WCE frames and discriminate between pathology and normal frames. Another gastrointestinal ulcer recognition is proposed in[25]. The researchers used texture and color visual descriptors along with an SVM model to classify ulcer frames. The researchers concluded that a combination of texture and color descriptors enhances ulcer detection in WCE video. The authors in[26]aimed to recognize bleeding ulcer. In their study, chromaticity moments [30] were extracted as color visual descriptor, and a Multi-layer perceptron (MLP) Neural Network[28] was used to classify WCE frames. In[29], the Color Coherence Vector (CCV) visual descriptor[30] was used to classify bleeding ulcer using SVM. The researchers considered 220 images of bleeding, 159 images of ulcers, and 228 images of non-bleeding/ulcers. Thiers images were captured in the small intestine by using the PillCam SB WCE[31]. Similarly, the authors in[11] worked on the recognition of peptic ulcer in WCE video. They used HSV color and texture visual descriptor as visual descriptors. Also, they used Support Vector Machines (SVMs) to classify the images the ulcer frames. The researchers managed to run their algorithm on 20 frames with ulcer cases. Besides that, they used 10 extra frames with ulcer cases to build the SVM model. The work in[32] focused on ulcer recognition using bag-of-words, LBP and SIFT [33]. The authors proposed a visual descriptor fusion technique to aggregate the different low-level visual descriptors, and represent the content of each frame using one single vector. Finally, they used Support Vector Machines (SVMs) to classify the frame instances. In [34], the researchers used MPEG-7 Visual Descriptors[35] to detect specific anomalies such as blood and ulcers. The result showed that the Scalable Color and Homogenous texture descriptors yield the better performance measures.

3. EMPIRICAL COMPARISON OF VISUAL DESCRIPTORS FOR ULCER RECOGNITION

We study several visual descriptors in order to find the most discriminating ones in terms of ulcer detection performance. First, we collect real WCE image to assess visual descriptors. The ground truth is provided for each frame. It consists of a label encoding if the frame contains an ulcer or not. In our experiments, we run a *k-fold* cross validation with *k*=10. We used three performance measures. Namely, sensitivity, accuracy and Specificity. These performance measures aim at evaluating the discriminating power of the visual descriptors with respect to ulcer and non-ulcer frames. In other words, the visual descriptors witch yields the best performance measures with be considered as the most discriminating ones and would be recommended for automatic ulcer detection in WCE frames.

3.1 Data Set

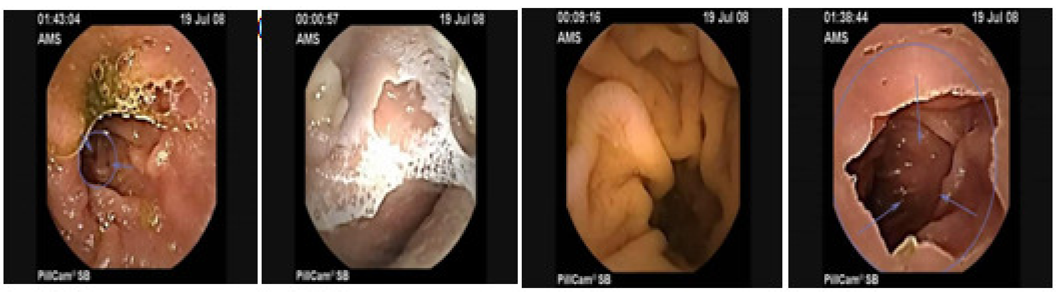

Figure 1.Sample ulcer images from WCE video.(a) Bleeding ulcer, (b)Esophageal ulcer, (c) Gastric ulcer, (d) Peptic ulcer.

We use a real WCE frames to run an experiments [36]. The dataset consists of 4274 images taken using WCE video [36], where 4024 are ulcer images, and 350 are normal. Figure 1 shows the sample ulcer images from the data set. These images taken from different parts of the digestive system and describes the appearance of different types of ulcer disease.

3.2 Experiment Description

We use various visual descriptors along with SVM classifier to find the most accurate one, and recommend it for detecting ulcer frames in WCE video. Namely, we consider Local Binary Pattern [37], Curvelet Transform [38], Chromaticity Moments Color [39], Color Coherence Vector [27], Homogenous Texture Descriptor [40], Scalable Color Descriptor [41], Ycbcr Color Histogram [42], the CIE_Lab Color Histogram [43], and the HSV Color Histogram [44]. The visual descriptors extracted from WCE frames are provided as an input to the SVM classifier, and the performance of the ulcer detection process is assessed with respect to each visual descriptor. When a descriptor yields low classification performance, it can be concluded that it was not able to convey the appropriate relevant information to the SVM classifier. Moreover, it can be claimed that the visual descriptor space of this visual descriptor does not represent efficiently "ulcer" and "non-ulcer" classes. Frame classification using SVM consists in determining the optimal hyperplane which separates data instances from both classes in the considered visual descriptor space. The optimal hyperplane corresponds to the widest margin between the two categories.

3.3 Experiments Results

The visual descriptors assessment results are summarized in Table 1. As one can notice in Table 1, the sensitivity level achieved by all visual descriptors is relatively high, and does not allow to objectively compare the obtained performances. On the other hand, the accuracy and the specificity show relevant variation, and reflect different performance levels of the different descriptors. In particular, the curvelet transform, the Homogeneous Texture Descriptor and the HSV color histograms achieved the poorest ulcer detection performance in terms of accuracy, and were drastically outperformed by the rest of the low-level visual descriptors. Also, these three descriptors along with the Chromaticity Moments Color, the Scalable Color Descriptor, the Color Coherence Vector, and the Ycbcr Color Histogram attain relatively low specificity levels. This means that these descriptors miss-classify a high portion of the “non-ulcer” frames. On the other hand, the Local Binary Pattern and the CIE_Lab Color Histogram are the two visual descriptors that were able to discriminate the best between “ulcer” and “non-ulcer” WCE video frames. In other words, they were able to represent the WCE video frames in their corresponding visual descriptor spaces in a away, they allow the SVM classifier to accurately separate between the two classes. Figure 2 shows the corresponding ROC curves.

Table 1.Accuracy, Sensitivity, and Specificity obtained for all visual descriptors using SVM classifier.

Visual descriptor	Accuracy	Sensitivity	Specificity
Local Binary Pattern	98.85%	99.4%	90.03%
Curvelet Transform	47.82%	99.22%	9.62%
Chromaticity Moments Color	77.42%	99.54%	13.79%
Color Coherence Vector	82.87%	99.93%	25.35%
Homogenous Texture Descriptor	48.50%	99.40%	09.83%
Scalable Color Descriptor	72.46%	99.79%	17.24%
Ycbcr Color Histogram	65.47%	99.96%	14.44%
CIE_Lab Color Histogram	98.95%	99.06%	96.80%
HSV Color Histogram	53.74%	96.50%	8.34%

As it can be seen, the results in Figure 2 confirms the results in Table 1. More specifically, the Local Binary Pattern and the CIE-Lab color moments outperform the other visual descriptors in terms of accuracy and sensitivity. This performance can be attributed to the fact that they are not sensitive to the monotonic color level variations that is caused by the high illumination variance of the WCE video frames that were captured at different locations of the gastrointestinal tract. Moreover, these results confirm that they visual descriptors satisfy the perceptual uniformity principle, and ensure that the difference between two patterns, as perceived by the human eye, is proportional to the distance measured within these two visual descriptor spaces.

CIE-Lab color descriptor outperforms the others colors descriptors because it is designed to model letter the human perception. Also, by definition in CIE-Lab, a color is either red or green, blue or yellow. This is appropriate to WCE frame characteristics where only red and yellow colors are present while no green or blue colors are included.

On the other hand, LBP texture descriptor beats the other texture descriptors because it encodes a combined structural and statistic of the texture pattern. Moreover, LBP provides a local information about the texture that can capture the visual properties of the WCE video frames containing ulcer symptoms.

While these two visual descriptors attain almost the same accuracy and sensitivity values, CIE-Lab color histogram achieves a higher specificity rate. Thus, it mis-classified less normal frames as "ulcer" than LBP based classification. Practically, this means that it would reduce the risk of further examination of healthy patient that were detected as "ulcer" cases.

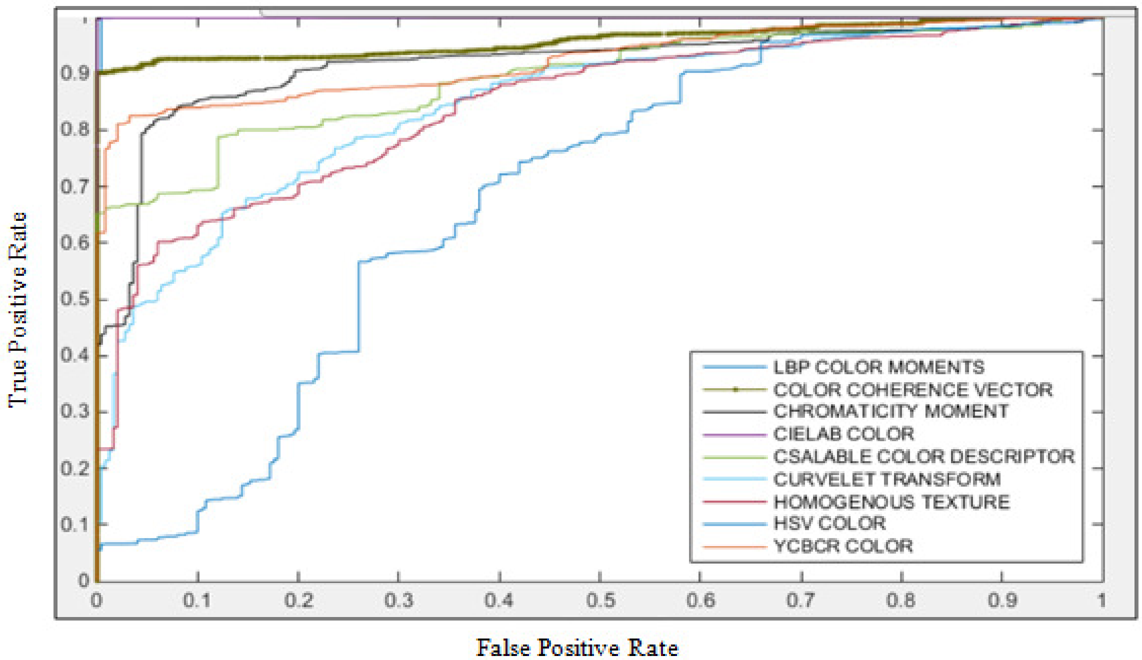

Figure 2. ROC curve obtained using the different visual descriptors and SVM classifier

4. Conclusions and Potential Future Works

Wireless Capsule Endoscopy (WCE) is the latest technology able to screen intestinal pathologies at an early stage. Despite its convenience to patients and its effectiveness to show small intestinal details, the physician involvement remains tedious and time consuming. The pattern recognition system can assist the physician by automatically detecting ulcer in WCE videos. However, the performance of such systems is sensitive to the choice of the visual descriptors. In fact, each feature encodes specific visual properties, and provides different discriminative information to the detection algorithm. The keystone is then to determine the feature(s) which yield(s) better recognition of the Ulcer pattern.

Our results showed that LBP and CIE-Lab color histogram outperform the other visual descriptors and achieved the best performance. The robustness of these two low-level features to monotonic color level variations caused by the high illumination variance in the gastrointestinal tract, along with their satisfaction of the perceptual uniformity principle, are the main properties that yield the obtained results.

Thus, we can proclaim that LBP and CIE-Lab color histogram are the most discriminating visual descriptors between "ulcer" and "non ulcer" frames in WCE video.

As potential future work, we plan to investigate visual descriptor aggregation in order to enhance the ulcer detection accuracy. This approach involves fusion techniques to optimize the combination of descriptors to best represent the visual content of the WCE video frames. For instance, efficient aggregation of LPB and CIE-Lab color histogram may enhance the overall ulcer detection accuracy.

REFERENCES

[1] A. S. Levey, K.-U. Eckardt, Y. Tsukamoto, A. Levin, J. Coresh, J. Rossert, D. d. Zeeuw, T. H. Hostetter, N. Lameire, and G. Eknoyan, (2005) "Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO)," Kidney international, vol. 67, pp. 2089-2398.

[2] C. Jingfeng, "Medicine in China," Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures, pp. 1529-1534, 2008.

[3] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, (2000) "Wireless capsule endoscopy," Nature, vol. 405, pp. 417-417.

[4] G. D. Finlayson, S. D. Hordley, and I. Tastl, (2006) "Gamut constrained illuminant estimation," International Journal of Computer Vision, vol. 67, pp. 93-109.

[5] Z. Fireman, A. Glukhovsky, H. Jacob, A. Lavy, S. Lewkowicz, and E. Scapa, (2002) "Wireless capsule endoscopy," IMAJ-RAMAT GAN-, vol. 4, pp. 717-719.

[6] T. Gevers and A. W. Smeulders, "Color-based object recognition, (1999) " Pattern recognition, vol. 32, pp. 453-464.

[7] S. A. Shafer, "Using color to separate reflection components, (1985)" Color Research & Application, vol. 10, pp. 210-218.

[8] T. Ojala and M. Pietikäinen, (1999)"Unsupervised texture segmentation using feature distributions," Pattern recognition, vol. 32, pp. 477-486.

[9] peptic ulcer. Available: www.health.harvard.edu/digestive.../peptic-ulcer

[10] J. Oh, (2013) "Detection of temporal events and abnormal images for quality analysis in endoscopy videos," UNIVERSITY OF NORTH TEXAS.

[11] A. Karargyris and N. Bourbakis, (2009) "Identification of ulcers in wireless capsule endoscopy videos," in Biomedical Imaging: From Nano to Macro, pp. 554-557.

[12] J.-Y. Yeh, T.-H. Wu, and W.-J. Tsai, (2014) "Bleeding and ulcer detection using wireless capsule endoscopy images," Journal of Software Engineering and Applications, vol. 7, p. 422.

[13] B. Marshall, J. R. Warren, E. Blincow, M. Phillips, C. S. Goodwin, R. Murray, S. Blackbourn, T. Waters, and C. Sanderson, (1988) "Prospective double-blind trial of duodenal ulcer relapse after eradication of Campylobacter pylori," The Lancet, vol. 332, pp. 1437-1442.

[14] B. Li and M.-H. Meng, (2008) "Ulcer recognition in capsule endoscopy images by texture features, WCICA 2008, pp. 234-239.

[15] B. Li and M. Q.-H. Meng, (2009) "Texture analysis for ulcer detection in capsule endoscopy images," Image and Vision computing, vol. 27, pp. 1336-1342.

[16] E. J. Candes and D. L. Donoho, (2000) "Curvelets, multiresolution representation, and scaling laws," in Proc. SPIE, pp. 1-12.

[17] T. Ojala, M. Pietikäinen, and D. Harwood, (1996) "A comparative study of texture measures with classification based on featured distributions," Pattern recognition, vol. 29, pp. 51-59.

[18] P. Wilding, M. A. Morgan, A. E. Grygotis, M. A. Shoffner, and E. F. Rosato, (1994) "Application of backpropagation neural networks to diagnosis of breast and ovarian cancer," Cancer Letters, vol. 77, pp. 145-153.

[19] J. A. Suykens and J. Vandewalle, (1999) "Least squares support vector machine classifiers," Neural processing letters, vol. 9, pp. 293-300.

[20] R. Nawarathna, J. Oh, J. Muthukudage, W. Tavanapong, J. Wong, P. C. De Groen, and S. J. Tang, (2014) "Abnormal image detection in endoscopy videos using a filter bank and local binary patterns," Neurocomputing, vol. 144, pp. 70-91.

[21] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons, (2001) " International Journal of Computer Vision, vol. 43, pp. 29-44.

[22] D. K. Iakovidis and A. Koulaouzidis, (2014) "Automatic lesion detection in capsule endoscopy based on color saliency: closer to an essential adjunct for reviewing software," Gastrointestinal endoscopy, vol. 80, pp. 877-883.

[23] G. Wyzecki and W. Stiles, (1982) "Color science: concepts and methods, quantitative data and formulae," New York, London, Sidney.

[24] L. Korman, M. Delvaux, G. Gay, F. Hagenmuller, M. Keuchel, S. Friedman, M. Weinstein, M. Shetzline, D. Cave, and R. de Franchis, (2005) "Capsule endoscopy structured terminology (CEST): proposal of a standardized and structured terminology for reporting capsule endoscopy procedures," Endoscopy, vol. 37, pp. 951-959.

[25] P. Szczypiński, A. Klepaczko, M. Pazurek, and P. Daniel, (2014) "Texture and color based image segmentation and pathology detection in capsule endoscopy videos," Computer methods and programs in biomedicine, vol. 113, pp. 396-411.

[26] B. Li and M. Q.-H. Meng, (2009) "Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments," Computers in Biology and Medicine, vol. 39, pp. 141-147.

[27] J.-M. Geusebroek, R. Van den Boomgaard, A. W. M. Smeulders, and H. Geerts, (2001) "Color invariance," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 23, pp. 1338-1350.

[28] S. Haykin, (1996), "Neural Networks: A Comprehensive Foundation", second edition ed. NewJersey: Prentice-Hall.

[29] J.-Y. Yeh, T.-H. Wu, and W.-J. Tsai, (2014) "Bleeding and Ulcer Detection Using Wireless Capsule Endoscopy Images," Journal of Software Engineering and Applications, vol. 7, pp. 422-432.

[30] G. Pass, R. Zabih, and J. Miller, (1997) "Comparing images using color coherence vectors," in Proceedings of the fourth ACM international conference on Multimedia, pp. 65-73.

[31] A. Moglia, A. Menciassi, and P. Dario, (2008) "Recent patents on wireless capsule endoscopy," Recent Patents on Biomedical Engineering, vol. 1, pp. 24-33.

[32] L. Yu, P. C. Yuen, and J. Lai, (2012) "Ulcer detection in wireless capsule endoscopy images," ICPR 2012, pp. 45-48.

[33] T.-M. Tu, P. S. Huang, C.-L. Hung, and C.-P. Chang, (2004) "A fast intensity-hue-saturation fusion technique with spectral adjustment for IKONOS imagery," Geoscience and Remote Sensing Letters, IEEE, vol. 1, pp. 309-312.

[34] M. T. Coimbra and J. S. Cunha, (2006) "MPEG-7 visual descriptors—contributions for automated feature extraction in capsule endoscopy," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 16, pp. 628-637.

[35] S.-F. Chang, T. Sikora, and A. Purl, (2001) "Overview of the MPEG-7 standard," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 11, pp. 688-695.

[36] DR.Khuroos. (16/12/2015,4:30 PM). http://drkhuroo.in/index.php

[37] E. Candes, L. Demanet, D. Donoho, and L. Ying, (2006) "Fast discrete curvelet transforms," Multiscale Modeling & Simulation, vol. 5, pp. 861-899.

[38] M. Choi, R. Y. Kim, and M.-G. Kim, (2004) "The curvelet transform for image fusion," International Society for Photogrammetry and Remote Sensing, ISPRS 2004, vol. 35, pp. 59-64.

[39] Paschos, "Fast color texture recognition using chromaticity moments, (2000) " Pattern Recognition Letters, vol. 21, pp. 837-841, 2000.

[40] Y. M. Ro, M. Kim, H. K. Kang, B. Manjunath, and J. Kim, (2001) "MPEG-7 homogeneous texture descriptor," ETRI journal, vol. 23, pp. 41-51.

[41] B. S. Manjunath, P. Salembier, and T. Sikora,(2002) "Introduction to MPEG-7: multimedia content description interface", vol. 1: John Wiley & Sons.

[42] S. Sural, G. Qian, and S. Pramanik, (2002) "Segmentation and histogram generation using the HSV color space for image retrieval," in Image Processing.

[43] G. J. Braun, M. D. Fairchild, and F. Ebner, (1998) "Color gamut mapping in a hue-linearized CIELAB color space," in Color and Imaging Conference, pp. 163-168.

[44] K. Cantrell, M. Erenas, I. de Orbe-Paya, and L. Capitán-Vallvey, (2009) "Use of the hue parameter of the hue, saturation, value color space as a quantitative analytical parameter for bitonal optical sensors," Analytical chemistry, vol. 82, pp. 531-542.