Automatic Polyp Detection in Wireless Capsule Endoscopy Images using Hybrid Patch Extraction and Supervised Classification

Sindhu C P[1], Vysak Valsan[2]

[1]PG Scholar, Department of ECE, Jawaharlal College of Engineering & Technology, Palakkad, India

[2]Assistant Professor, Department of ECE, Jawaharlal College of Engineering & Technology, Palakkad, India

Abstract: *Wireless Capsule Endoscopy (WCE) is a non-invasive and painless novel technique for diagnosing gastrointestinal disease. Computer Aided Diagnosis system is a solution to reduce the great burden of a doctor to examine images frame by frame to locate abnormalities. In this paper an automated computer aided detection system with image analysis and supervised learning method is proposed to detect polyp from WCE images. Textural features are extracted not only from single key point, by utilizing single scale-invariant feature also from neighborhood key points. Haralick texture features are extracted from each of patch size of 16*16 around the key points. After acquiring the different texture features, we develop a strategy to integrate these features with Key point descriptor. For the best classification performance, SIFT is integrated with 22 Haralick textural features. The supervised classification is performed using a Multilayer perceptron Neural Network. In our proposed method, Feature based classification performance is better than other classifiers and correctly identifies polyp images with accuracy about 98.8%.*

Keywords: Inflammatory bowel diseases (IBD), ulcer, polyp detection, SIFT, Haralick texture features, multilayer perceptron neural network, wireless capsule endoscopy images

1. Introduction

Wireless Capsule Endoscopy (WCE) is an omnipotent modern diagnostic tool used for direct visualization and non-invasive examination of patient's gastrointestinal tract. In 2001, by Given Imaging Inc. a novel technique of endoscopy, WCE [1] was introduced. Now, it has evolved into an imperative diagnostic method for inflammatory bowel diseases (IBD) [2] such as crohn's, ulcer, and diseases like gastrointestinal (GI) bleeding, polyp and tumour. The WCE is similar to a pill in its size, has length of 26mm and diameter of 11mm as shown in Fig.1 (a). It is a method of recording images and consists of a tiny camera, protecting dome lens, illuminating LED, batteries, radio frequency transmitter and an antenna. In the examination procedure, after capsule being swallowed by a patient it moves along digestive tract and camera records images for doctors to examine and provide accurate diagnostic decision [3]. Polyp is one among the common disease in intestinal mucosal layer as a result of growing protrusions of mucosa by disproportionate tissue growth in the gastro intestinal layers such as the regions of the Stomach, colon, and urinary bladder as shown in Fig.1. (b).polyps can be either beginin or non beginin. Even though
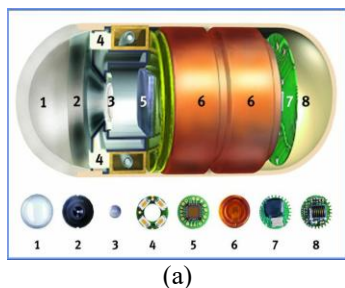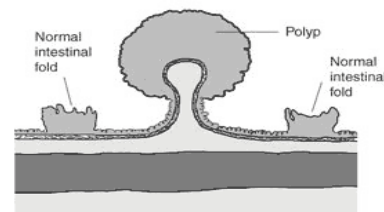
(b)

Figure 1: (a) Wireless capsule Endoscopy (b) Graphical representation of a polyp in human colon

Some polyps are benign, virtually all colon and rectal cancer starts from them. Neoplastic polyp is early stage of colorectal cancer and hence early detection of polyp is important [5].The WCE is a least invasive and least harmful modern endoscopy technology for identifying even polyps smaller than 1 cm in diameter and also provide an in-depth view of irregularly shaped polyps in digestive tract. Polyps appear in the shape of elliptical (Fig.2.(b) (c)), round (Fig.2.(c) (d)), semicircle (Fig.2.(a)) and in different sizes (Fig.2(a)(b)) as shown in the Fig.2.
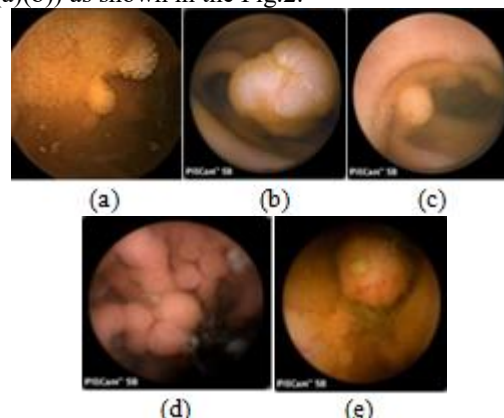

Figure 2: Polyps in WCE video

(a)

The main problem is that images capturing whole digestive tract by WCE exceed over exceed over 50,000 images for examining one patient. The abnormal images of digestive tract among entire WCE images recorded will be only a slight portion. By human perception it is not easy to detect polyp. The automatic disease detection is a solution to scale down the vast duty of a doctor to examine images frame by frame to locate abnormalities.

Here we propose an automatic method to detect polyp from WCE images. Further portions of the paper are formulated as pursued. Section II presents the related works. Section III describes methodology. Section IV describes the experimental results and discussion. Section V concludes the paper.

2. Related Works

Several computational methodologies and efforts carried out for automatic polyp detection in WCE images is being mentioned in literature [6],[7],[8]. Karargyris and Bourbakis performed Lo-Gabor filter based segmentation along with SUSAN edge detector, curvature clusters, active contour segmentation to identify polyp. In [7] authors applied a region based active contour method (ACM) along with a new technique based on geometric feature of polyp for automatic polyp detection. Li.et.al [8] based on textural features to integrate the better features of wavelet transform and uniform binary pattern. Wavelet transform is used for multi resolution analysis which mainly helps in discrimination of textural features. In [9] authors used an improved bag of feature methodology for automatic detection of polyp.

3. Methodology

Polyps are identified based on texture information in mucosa surface in WCE images. We extract traditional SIFT feature and GLCM texture features. Then we concatenate SIFT feature and texture features together to characterize the

Patch features. Multilayer perceptron neural network are used to perform supervised classification. The impact of training and testing is also studied to improve results.

The entire detection process comprises of six steps: Video to Frame Conversion, ROI Extraction, Patch Extraction, Feature Extraction, Feature Integration and Classification Method. Each step is further discussed in the following.

3.1 Video to Frame Conversion

WCE videos are often about eight to twelve hours duration. It consists of number of frames, a set of frames will form the shots and the group of shots will produce the scenes, the combination of different scenes will form the video. In order to analyses any video, first we need to study then characteristics of the frames and to analyze the properties of the video. This can be done by the video to frame conversion using MATLAB [13].

3.2 ROI Extraction

As given in the Fig.3 WCE images will be ambiguous with their black in color and visible boundaries. Hence by extracting features from the images will mirror those apparent visible defilements in each of the image. For that Maximum Square is marked within circular shaped image can be considered as required region of interest (ROI) with no significant image information's loss. Extracted ROIs are satisfactory enough to describe about significant image features and it also provides detailed description and characterization of WCE images. It makes feature extraction procedure easier and ROI images are used instead of original images with further processing such as classification after image feature extraction [12].
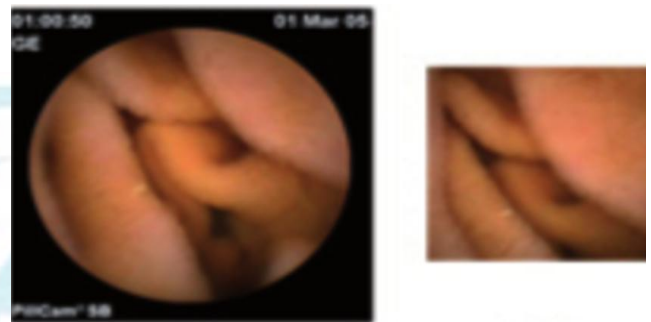

Figure 3: Illustration of ROI Extraction

3.3 Patch Extraction

In our proposed method we are combining features from neighborhood key points from WCE image which is different from describing an entire image. It is a technique for detecting

Salient, stable feature points in an image. The other key point detectors includes Laplacian of Gaussian(LOG), Difference of Gaussian(DOG), Harris Laplace(HL), FAST detector, Harris Affine(HA) and SIFT[10].The proposed method uses a SIFT key point detector for scale and rotation invariant key point detection. SIFT algorithm [11]proposed by David Lowe in 1999then developed by the same author in 2004 [12] is used in computer vision tasks for local feature detection and description in images. SIFT key points extraction consists of two steps such as key point's detection and description. The first step of SIFT key points extraction is to detect key points of an image. Firstly, Gaussian filters at variant scales are convolved with the target image. Secondly, difference of successive Gaussian blurred images is computed. Finally, through comparing their numerical values, the maximal or minimal points of Difference of Gaussians (DoG) will appear at the various scales are selected as key points, as shown in Fig 4.
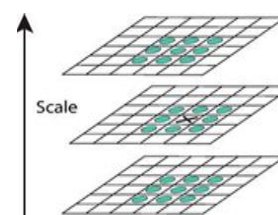

Figure 4: Key Point detection

The second step is to describe the key point with 128 spatial orientation bins. The description of a key point is derived through several steps. Firstly, in the region that surrounds the key point, the orientation and gradient magnitude at each sample point of the image are computed, which are weighted by a Gaussian-based window and denoted by an overlaid circle. Secondly, the computation results are accumulated in orientation histograms which sum up the contents through 4×4 sub regions and the length of every arrow indicates the sum of the gradient magnitudes within the region for that direction[11][12] is demonstrated with Fig. 5.
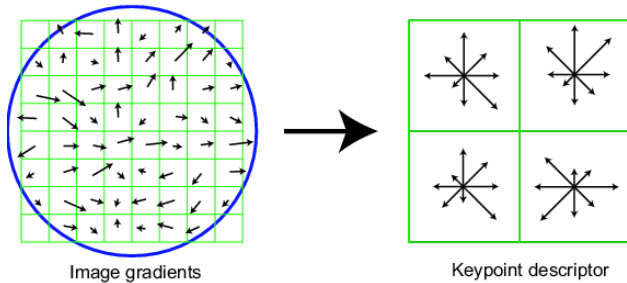

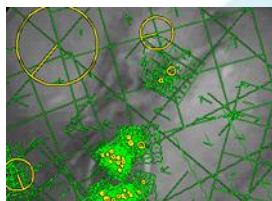**Figure 5:** Key Point Extraction

Figure 6: Illustration of Key Point Extraction

To describe WCE images different patch sizes, as shown in Fig. 6, are selected around the key points. In this work, selection of patch sizes 4×4, 8×8, 16×16 based on polyp size, Area. It is done fortesting the impact on patch size for the classification of WCE images. Normally in image processing field the size of patch selected as   to obtain SIFT related features.

3.4 Feature Extraction

The co-occurrence matrix is solitary among the universally accepted methods to excerpt texture feature because of its richness in texture information. The co-occurrence matrix calculates the arrival feasibility of two values of the pixels located at any range in the image. An assorted data can be obtained from the co-occurrence matrix. In proposed task, we extract 22 Haralick features from each of the GLCM matrix generated.

Textural features are defined based on arrangement of pixel intensities. Practically, the pixel intensities are defined by a typical distance-angular connection by a Grey Level Co-occurrence Matrix (GLCM) suggested by Haralick et al. [14]. GLCM is a second-order statistics in order to describe connection among recognizable tonal intensities by calculating the frequency along which they appear is composed at particular directions (θ) and distances d, and comes under category of statistical textural classification methods. This matrix is square with dimension   , where   represents number of gray levels in an image. Grey Level Co-occurrence Matrix feature approach itself is again considered to describe typical grey level interdependence of

WCE images. The set of matrices are described in four directions: 0, 45, 90 and 135 as in Fig.7, at a distance d and is represented as $P(i, j, \theta, d)$. Subsets of 22 Haralick features are determined from the GLCM to illustrate the textural characteristics of WCE images.
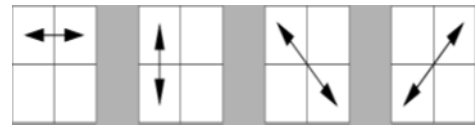

Figure 7: Illustration of Four Directions of Adjacency for Calculating Haralick Texture Features

3.5 Feature Integration

By subsequently acquiring the distinct textural features, anew approachis evolved to concatenate these features as a group to define polyps in detail. We consider $p_i(x, y)$as the key point identified by SIFT detector. Here $(x, y)$is considered as position of pixel$p_i$within theactual image. Along the given patch size a region has to be considered in such a way that p_i is the center. Then the SIFT+ Haralick feature descriptor combination is used. For describing key point p_i indicated as in the image, a 128 dimensional $(SIFT)_i$ descriptor is used. Then by selecting a particular patch size around the key pointp_i and computed textural features. Then $(SIFT)_i$ and textural features are concatenated as a group to create an array having a dimension of 150 to describe the entire patch. The normalization has to be performed before concatenation; SIFT+ Haralick textural feature=Concatenation [(SIFT)$_{i,}$ Haralick textural feature].Thus evaluation of the classification performance is done.

3.6 Classification Method

We test our proposed method using multilayer perceptron neural network [15]. It is relatively an advanced machine learning technique established on the basis of statistical learning theory. Multilayer perceptron's is based on a back propagation algorithm, one among the classic algorithm for a Supervised learning pattern identification procedure and for a Continuous study in computational neuroscience and parallel distributed processing. In our work, for classifying database we need to train the model. That is creating a model by running the Levin's berg back propagation algorithm on training data. After testing if accuracy is low, regenerate the model. After that recognize the class label of recently arrived data. This method can be used to classify unknown tuples from the database. Finally for analysis the forecasted labels are compared to actual labels of test images to calculate the classification performance.

4. Experimental Results and Discussion

4.1 Image acquisition and experimental setup

In this experiment, we have done experiment with a dataset of 406 WCE images that consist of 100polyps, 306 normal frames. These images are normally extracted and manually examined by Gastroenterologists from patient's videos. The images are obtained from Pillcam^{R}SB, device from Given imaging team by a resolution of 576×576.Among 100 polyp

samples, 70 samples is used as training set,15 samples were used as testing set and 15 used as validation set. These techniques were done replicated for many times and then mean performances were considered for assessment of classification. Several experiments were also carried out to attain best variables for polyp and normal frame classification.

4.2 Experiment Results for Polyp Frame Detection

1) Selection of Parameters

The initial technique is performed to analyze the influence of different feature combination and then to select the suitable parameters for classifying WCE images. After key point extraction, features from different patch sizes surrounding key points are extracted, and the feature concatenation is done. The evaluate classification performance; main parameter to be considered is patch size. In our experiment we use different patch sizes such as {4×4, 8×8, 16×16}.We found that as the patch size is small scale (4×4), local features extraction are not able to describe key points accurately. Thus its discrimination capability is decreased. For accurately describing key points and for getting outstanding performance of polyp detection patch size of 16×16 is selected that surrounding key points for carrying out experiments.

2) Patch and Feature Analysis

The proposed method is performed to judge the importance of feature consolidation and classification methods of gastrointestinal disease detection method. After obtaining ROI, here we need to apply key point extraction method. To show the effectiveness of our features, we need to have an analysis in feature extraction and selection process. For analyzing impact of each of the patch size around the key points and to obtain 128 SIFT related features, we apply an SIFT detector based computer vision algorithm for detecting and characterizing local features in WCE image. SIFT key points are at first obtained from the group of reference images and stored in database. It provides a reliable recognition, features extracted from training images be visible even after image scaling, noise, illumination. Those key points will be in immense contrast areas of an image that is boundaries of WCE images. It also reduces additional errors induced by regional variations and can also robustly identify key points even among clutters. By comparing with other key point extraction methods, SIFT is much accurate.

We further analyzed patch size in relation with the gastrointestinal frame classification through our experiment. The result indicates that for a small patch size (4×4) the extracted local features are not enough to give information about key points. For achieving best performance for disease detection, choose a patch size 16×16 surrounding the key points for implementing the proposed method. To characterize patch features SIFT features are concatenated with 22 Haralick textural features. Local features around every key point with the given training set were enumerated for constructing high dimensional descriptors. This method based on local textural features, could scale down the impact of redundant information and redundant information.

3) Classifier Analysis

Multilayer perceptron neural network is a nonparametric classifier which requires no prior knowledge of statistical distribution of class and it can be also successfully used for classifying the recently appeared inputs which are no more included at the time of training. More over in mentioned classification method, training time is slower but testing time is much faster. Training algorithm used is back propagation algorithm with two recognizable paths that is forward path pursued by the backward path over the layers of neural network. While forward passes computations of outputs of all neuron in network is done and in backward pass propagation of weights along with an adjustment of weight is also done. Specifications to be studied to frame back propagation algorithm is initial weight range, number of hidden layers, nodes in hidden layers, epochs, step size for gradient descent, error tolerance, critical error.

In our experiment, we need to classify polyp and normal frames from the whole WCE images. Here similarities in textural features of WCE images are used for the analysis of feature based classification accuracy. Normally growing Protrusions are polyp in mucosal surface and as the size of polyp becomes larger it leads to tumors. Initially we train a Classifier to separate between polyp and tumors from the image dataset. Out of 406frames for polyp prediction 98.8% predictions were correct and only 1.2% of predictions were wrong. After training and validation using classifiers, we analyses the parameters like performance, training state, error histogram, confusion matrix, Receiver Operating Characteristics (ROC). Outstanding validation performance is accomplished at the 4th epoch where the mean square error (MSE) is minimum .MSE is for stopping training network and if actual error less than or equal to this error then training is desired to be aborted. Then progress of training in our system is analyzed using epoch, time, performance, gradient, validation, and validation check. Number of epochs is about 10 iterations but increasing number of epoch's accuracy of this model can be improved. Training time required is about .27 only. Here Gradient is about 0.687,it is the multiplication factor for miscalculation rectification at the time of back propagation. It has a low value but is slow but steady learning. Error tolerance is also between 0 and 1.From the confusion matrix sensitivity, specificity, accuracy can be also analyzed. Here we have both sensitivity and specificity high.ROC determine whether diagnostic classification good or not. In our experiment, most of the points in ROC curve are closer to the ideal co-ordinate thus result is more accurate. As faster the curve approaches to ideal point the test results are more accurate.

The Fig.8 represents ROC characteristics, where class 1 to class2in this figure refers to polypand normal frames respectively. The overall accuracy of our proposed system is about 98.8%.It has found that with increase in the number of WCE images, the overall accuracy can be also improved. Table I represents the overall performance analysis of the proposed system.

4.3 Scope for Future Work

Our proposed method has still room for improvement. For practical implementation of proposed technique in hospitals,

more tests by large number of datasets can be performed. In our experiment, entire processing where carried out in MATLAB at core i3, windows 7 workstation with 4 GB memory. Hence further test using larger dataset are difficult for justifying the robustness and effectiveness of our suggested classification approach. Thus accuracy of the proposed classification strategy can be improved by processing with high performance systems. The alternative key point selection method and more feature integration methods could be used to achieve better result. Thus polyp frames can be more precisely detected from the set of whole WCE images.

Table 1: Comparison of State of Art Method with the Proposed System

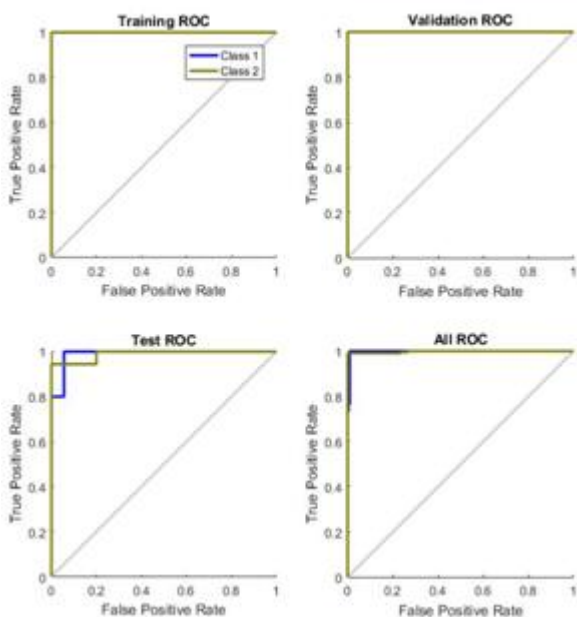| State of Art Method	Accuracy	Sensitivity	Specificity
Li's Method	88.56%	72.00%	92.70%
Yuan, Li, Meng	93.20%	90.88%	94.54%
Proposed Method	98.8%	96%	92%

Figure 8: Receiver operating Characteristics of our Proposed System.

5. Conclusion

Here we develop an automatic method to detect polyp from WCE images. A key point extraction method along with 22 Haralick textural features is used to the extract features from WCE images. The supervised classification is performed using a multilayer perceptron neural network to increase the accuracy of polyp detection. Experimental results verify that the proposed method detects polyp from the WCE images with higher accuracy and thus it is quite suitable for real-time applications.

References

[1] G.Iddan, G.Meron.A, Glukhorsky and P.S. Wain, "Wireless Capsule Endoscopy", Nature Vol.405, No.6785, 2000

[2] B.Lewis,"Expanding Role of Capsule Endoscopy in Inflammatory Bowel Disease", World Gastroenterology, Vol 14,no 26,pp 4137-4141,2008

[3] N.M Lee and G.M Eisen,"10 Years of Capsule Endoscopy: an Update, Expert Review of Gastroenterology and Hepatology, Vol 4, No.4, pp.503-512, Aug 2010

[4] A.kargyris and N.Bourbakis,"Detection of Small Bowel Polyps and Ulcers in Wireless Capsule Endoscopy Videos", IEEE Transaction Biomedical Eng., Vol.58 no 10.pp 2777-2786, Oct. 2011

[5] M.S Cappel, "The Pathophysiology Clinical Presentation and Diagnosis of Colon Cancer and Adenomas Polyp", Med Clinics North America, Vol 89, pp 1-42, 2005

[6] A. Karargyris and N. Bourbakis, "Detection of small bowel polyps and ulcers in wireless capsule endoscopy videos," IEEE Trans. Biomed. Eng., vol. 58, no. 10, pp. 2777–2786, Oct. 2011.

[7] H.Eskandari, A. Talebpour, M. Alizadeh, and H. Soltanian-Zadeh, "Polyp detection in wireless capsule endoscopy images by using region-based active contour model," in Proc. 19th Iranian Conf. BiomedEng. (ICBME), 2012, pp. 305–308.

[8] Y.G. Jiang, C.W.Ngoand, j.Yang,"Towards Optimal Bag Of features for object categorization and semantic video retrieval", in proc 6 th ACM int. conf. Image Video Retrieval,2007, pp 494-501.

[9] Lowe David G,"Object Recognition from Local scale invariant Features", Computer Vision 1999, The proceedings of seventh IEEE international Conference on 1999, Vol 2, pp 1150-1157.

[10] Lowe David G,"Distinctive Image Features from Scale Invariant Key Points", International Journal Computer vision, 2004, Vol 60,pp 91-110

[11] Punith Kumar MB, Dr.P.S.Puttaswamy,"Video to Frame Conversion of TV News Video by using MATLAB", International Journal of Advance Research in Science and Engineering, March 2014

[12] Lee, S K et.al,"A Computer Aided Design Mammography Screening System for Detection and Classification of Micro Classification", International Journal of Medical Informatics 29-57, 2000.

[13] Amit Gantara, Y.P kosta, Gaurang Pandael, Chintan Gajjar, "Initial Classification through Back Propagation in a Neural Network following Optimization through GA to Evaluate Fitness of an Algorithm", International Journal of computer science and information technology (IJCSIT), 2011