Automatic Classification of Images from WCE Videos based on SURF and Locality Constrained Linear Coding

Amritha P
M. Tech student in Applied Electronics and Communication System
Nehru College of Engineering and Research Centre
Thrissur, India

P. Rajkumar
Professor- Electronics and Communication Department
Nehru College of Engineering and Research Centre
Thrissur, India

Abstract—*Wireless Capsule Endoscopy (WCE) has become a mostly used diagnostic technique for the gastrointestinal tract at the cost of a huge quantity of data that wants to be analysed. To find a solution to this problem a new computer aided system using novel features is proposed in this paper to classify images from WCE videos automatically. In the feature learning stage to obtain the representative visual words, the training images of polyp, ulcer, bleeding, and normal images are sampled and represented by Speeded up Robust Features (SURF) descriptor, and is constructed by K-means clustering algorithm. These four types of visual words are combined to composite the representative visual words for classifying the WCE images. In the feature coding stage we propose a locality constrained linear coding (LLC) algorithm to encode the images. LLC uses the locality constraints to project each descriptor into its local-coordinate system, and the projected coordinates are concatenated by max pooling to create the final representation, and classified using SVM The experimental results exhibit a higher accuracy, sensitivity, specificity and lower processing time, validates the acceptance of the proposed method.*

Keywords: *SURF (Speeded Up Robust Feature), Wireless Capsule Endoscopy (WCE) image classification, Locality Constrained Linear Coding (LLC) algorithm, Coding bases.*

I.INTRODUCTION

WIRELESS CAPSULE ENDOSCOPY (WCE) [1][Fig. 1(a)] which was first invented by GivenImaging Incorporation in 2000 and it involves transmission of images wirelessly from the inner sides of intestinal tract to the outside environment.

WCE has been adopted instead of the traditional ones ait makes possible, painless, non-invasive, disposable [2] and effective diagnostic technique which lowers the amount of discomfort and can be indulged by the patients for the observation of small bowel [3], [4]. A normal WCE capsule measures 11 mm in diameter and 25 mm in length. WCE system which is a pill-shaped device that contains a tiny camera, light sources, a battery, radio transmitter and some other miniaturized element. After a WCE is consumed by a patient it passes through the gastrointestinal (GI) tract and starts to capture color images of the tract for an average period of 8 hrs. Simultaneously these images are then sent wirelessly to a data-recorder attached to the patient's waist and the small device is propelled through peristalsis. At last the physicians review the images that are downloaded to a computer, and diagnose various abnormalities in the gastrointestinal tract [5], [6]. An experienced clinician spends an average of 2 hours for the video reviewing and analysing of approximately 60,000 images that are captured per examination of one patient [7]. Therefore, it is essential to design a computer-aided diagnosis system for processing WCE images to assist the physicians to analyse abnormal images.

II. RELATED WORK

Many works have beenmade in the literature forautomatic abnormal image detection from the WCE videos [8]-[14].The common diseases in the intestinal tract are typically bleeding, polyp, and ulcer, as shown in Fig. 1(b)–(d).Yeh et al. proposed [8] a novel method for detecting bleeding and ulcers in WCE images by using color features to determine the status of the small intestine.Fu et al. [9] introduced rapid bleeding detection in the WCE videos. They segmented images through
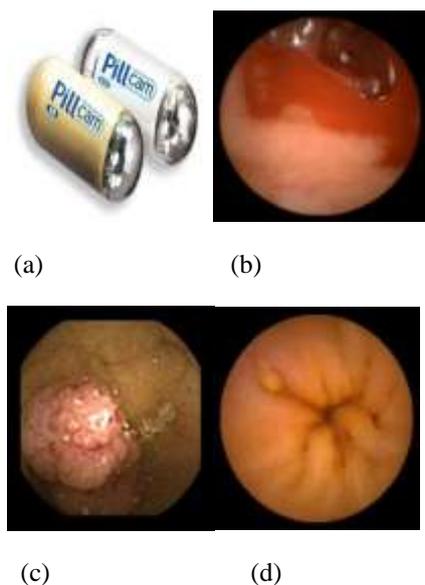
(a)   (b)

(c)   (d)

Fig.1. WCE and captured images. (a) Typical WCE capsule. Example images of (b) bleeding, (c) polyps, and (d) ulcers.

super pixel segmentation and feature of each pixel is extracted using the red color ratio in RGB color space. Yuan and Meng [10] proposed a novel texture feature that integrates the advantages of both Gabor filter and monogenic local binary pattern (LBP) methods. Followed by the linear support vector machine (SVM) classifier it detects polyp tissues from the normal images. Charisis et al. [11] introduces a Discrete Curvelet Transform (DCT), method whichcalculates the lacunarity index of DCT sub bands of the WCE imagesfor acquiring the textural information'sto detect the ulcer images.Meng and Yuan [12] proposed an algorithm based on the combination of Bag Of Features (BoF) method and saliency map to detect the polyp from WCE images. Gong et al. [13] introduces an extension of the SPM method where sparse coding was used instead of vector quantization (VQ). Yuan and Baopu [14] introduces Saliency and Locality Constrained Linear Coding (SALLC) algorithm to detect multi abnormalities in the WCE images. It consists of two stages feature learning stage and feature coding stage, in the first stage they calculates the Scale Invariant Feature (SIFT) transform of the WCE images and applies the K mean clustering to obtain the visual words.in the feature coding stage SALLC algorithm is used to encode the images.

III. PROPOSED APPROACH

In this paper, a novel method for the automatic classification of WCE images from the WCE videos based on extracting the SURF features [15] from each frames and applying locality constrained linear coding algorithm to encode the images and used for detecting bleeding and ulcers in WCE images is proposed. In this work there are two stages feature learning stage and feature coding stage. In the feature learning stage the WCE video frames are converted to WCE images and for each frame SURF is extracted and is constructed by k means clustering to generate the visual words. In the feature coding stage LLC algorithm [16] is used to encode the images.LLC uses the locality constraints to project each descriptor into its local-coordinate system, and the projected coordinates are concatenated by max pooling to create the final representation.Finally, these feature vectors are feed to a linear SVM [17] classifier to decide whether the image is bleeding, polyp, ulcerFig. 2 shows the block diagram of the proposed scheme. Each stages of the block diagram are explained below.

A DATABASE

It is a collection of videos that we used for WCE multi abnormality detection and classification.

B. VIDEO FRAMING

In image processing we use image as the input. Here our database is about video of WCE and we have to convert the video to frames for processing section.

C. SURF EXTRACTION

One of the main advantage of SURF is it is invariant to common image variations like scale changes, image rotation, illumination changes and more over it is simple, faster and more robust than SIFT.After framing the WCE videos in to images, we extract the corresponding surf points for each images. The WCE image classification shows good performance as the feature is demonstrated through the SURF points. The color information is an important factor for image characterization than the texture feature. Thus we apply the V-SURF by computing SURF descriptor over the value channel of the HSV color model to represent the SURF points. This yields a feature of 1x64 dimension for each descriptor and each mage could be represented as $X = [x1, x2, \ldots \ldots xN] \in R^{64XN}$.,where N is the number of SURF points for a given image.

D. VISUAL WORD GENERATION

To obtain the representative visual words, we use K-means clustering method on the SURF features of each type of image (bleeding, polyp, ulcer, and the normal) to obtain the visual words. Let K/4 bethe size of each type of visual word, the number of visual words ($B = [b1, b2, \ldots . bK] \in R^{64XK}$) is K.
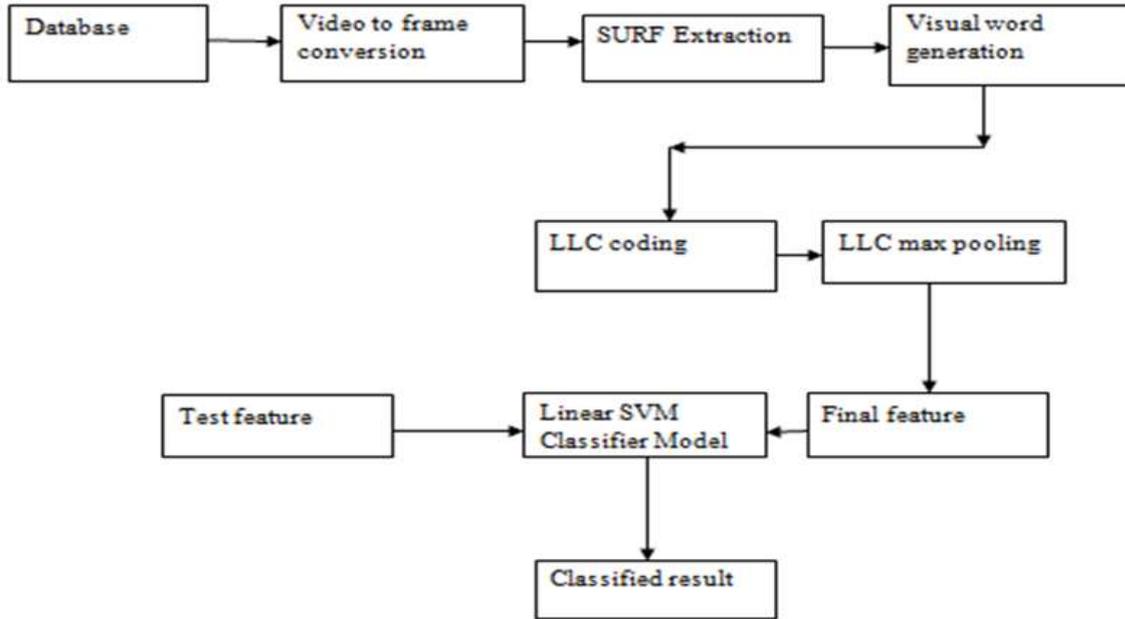
Fig. 2 Block diagram of the proposed scheme

Too small or too large vocabulary size of K leads to decrease in the discrimination performance and large computational time. In an effort to tackle this problem, we change the size of visual words from 80 to 480 with an increment size of 40.We cluster data according to the k value we specified. The centre of each cluster is selected as each visual word.

E. LLC CODING

Locality constrained linear coding used for encoding the feature data to small set of feature for fast processing. LLC utilizes the locality constraints to project each descriptor into its local-coordinate system, instead of the sparsity constraint.

Let X be a set of N-dimensional SURF descriptors extracted from the WCE images,i.e.$X = [x1, x2, \ldots \ldots . xN] \in R^{64XN}$. Given visual words of each classes with K entries,$B = [b1, b2, \ldots . bK] \in R^{64XK}$.and, $C = [c1, c2, \ldots \ldots cN] \in R^{NXK}$ is the coded feature. In order to integrate the coding bases and the feature importance to the coding process the LLC code uses the following objective function:

$$\min_{C} \sum_{i=1}^{N} ||\mathbf{x}_i - \mathbf{Bc}_i||^2 + \lambda||\mathbf{d}_i \odot \mathbf{c}_i||^2$$
$$s.t.\ \mathbf{1}^{\top}\mathbf{c}_i = 1,\ \forall i \qquad (1)$$

where, λ is a weight parameter for adjusting the locality adaptor weights and Θ denotes element wise multiplication. In our method di is the local distancebetween xi and B, $di \in R^K$ is the locality

adaptor defined as

$$di = \exp\left(\frac{dist\ (xi,B)}{\sigma}\right) \qquad (2)$$

where$dist(xi, B) = [dist(xi, b1), \ldots dist(xi, bk)]^T$ and $dist(xi, B)$ is the Euclidean distance between xi and bj. The weight decay speed for the locality adaptor is adjusted by using the parameter σ. The item $\sum_{i=1}^{N} \lambda||di\Theta ci||^2$ in (1) defines the locality constraint. In LLC coding each SURF description is projected to the fixed point.

The solution of the LLC [20] can be derived analytically. By optimizing the object function specified in (1) is as follows:

$$\min_{ci}||xi - Bci||^2 + \lambda||di\ \Theta\ ci||^2$$
$$s.t\ 1^T ci = 1\ Vi \qquad (3)$$

where $di = exp⁣(dist(xi, B)/\sigma), bj \in$ fixed q. This problem can be solved by minimizing the following Lagrange function:

$$L(ci, \beta) = ||xi - Bci||^2 + \lambda||di\Theta ci||^2 + \beta(1^T ci - 1). \qquad (4)$$

By letting the partial derivative of this Lagrange function to be zero, obtain the analytical solution of this LLC:

$$\tilde{ci} = \left(Ci + \lambda\ diag(d)\right) \qquad (5)$$

$$ci = \tilde{ci}/1^T \tilde{ci} \qquad (6)$$

where $,C_i = (B - 1x_i^T)(B - 1x_i^T)^T$ defines the data covariance matrix.

F. LLC MAX POOLING

To find the maximum dissimilar feature from the set of SURF points. To generate a WCE image representation each SURF descriptor(x_i), the codes (c_i) are pooled together by the max pooling strategy after coding. Compared with the other pooling strategies like average pooling and sum pooling [19] max pooling shows better performance.

G. FINAL FEATURE

Suppose a frame contains N SURF points and the visual word size is K, then after max pooling, this frame feature $g = [g_1, g_2, ... g_j, ... g_k]$ will be represented by a K-dimensional vector and g_j is calculated by

$$g_j = \max\{c_{1j}, c_{2j}, ... c_{Nj}\} \qquad (7)$$

The proposed method can characterize the WCE images better than the other representations because the final image representation is obtained by integrating the visual words, highlighting the SURF descriptor features.

H. TEST FEATURE

It is same as the feature extraction and LLC coding output for a single test image and sequence of videos. Test feature is given to model for classifying data.

I. CLASSIFIER MODEL

In this stage we used linear SVM for model the classifier. All featurefrom database will used to learn the machine with the help of hyper plane. WCE images which are represented by the LLC features are then given to the linear SVM classifier with a Gaussian radial basis function kernel to carry out the multi abnormal image detection task.

J. CLASSIFIED RESULT

In order to evaluate the performance of the multi abnormality detection in the WCE image videos,the classification results are expressed in terms of accuracy, sensitivity and specificity measures.

IV. EXPERIMENTAL RESULTS

The WCE videos used for the evaluation of the proposed scheme were extracted from 45 patients. There are (4 normal videos, 11 ulcer videos, 10 bleeding videos and 5 polyp videos). There are total 901 images composed in the dataset of WCE images extracted from 201 polyp images, 201 bleeding images, 201 normal images and 301 ulcer images for our experiment. There are two categories to which the set of images in each class is divided they are training set and testing set. This procedure was repeated 10 times by randomly choosing training and testing sets of images, then the average value of accuracy, sensitivity and specificity were calculated, which are defined as follows:

$$Accuracy = \frac{No: of\ correct\ predictions}{Total\ samples} \qquad (8)$$

$$Sensitivity = \frac{No: of\ correct\ positive\ predictions}{No: of\ positives} \qquad (9)$$

$$Specificity = \frac{No: of\ correct\ negative\ predictions}{No: of\ negatives} \qquad (10)$$

The fig. 3 shows the confusion matrix of our method for WCE image classification.. In the matrix class 1 represents the ulcer images, class 2 represents the polyp images, class 3 represents the bleeding images and finally class 4 represents the normal images. The accuracies for detecting the ulcer images are 96.9%, the polyp images are 98.2%, the bleeding images are 98.4% and finally for detecting the normal images are 95.8%.

TABLE I

COMPUTATIONAL PROCESSING TIME OF DIFFERENT CODING SCHEMES FOR WCE CLASSIFICATION

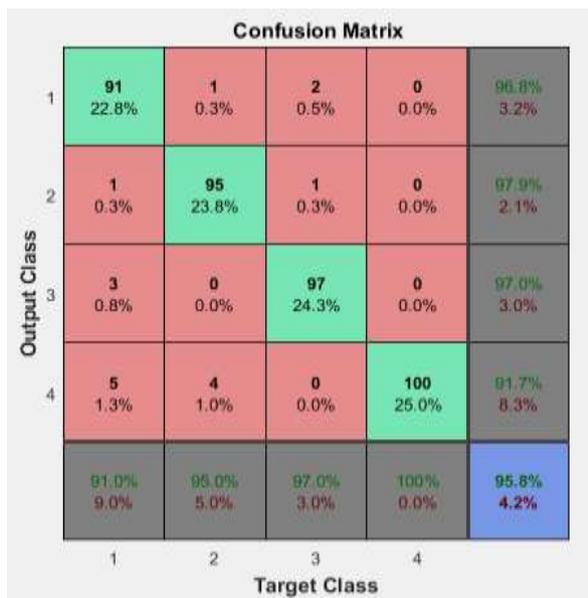|	VQ (s)	SC (s)	SALLC (s)	OURS (s)
Processing time	186.54	326.23	210.34	164.55

Fig.3 Confusion matrix for WCE image classification

To further illustrate the computational efficiency of our proposed scheme we tabulated the processing times of traditional feature coding schemes and our method in Table I. The experiment result shows that our method is faster than VQ, SC and SALLC. As for the computation time, our method is faster than the existing WCE abnormality detection schemes.

TABLE II

ACCURACY COMPARISON OF EXISTING ABNORMALITY CLASSIFICATION IN WCE IMAGES

	Normal Accuracy. (%)	Bleeding Accuracy. (%)	Polyp Accuracy. (%)	Ulcer Accuracy. (%)
Method[14]	96.18	95.02	96.58	90.12
Proposed	97.70	98.40	98.20	96.90

TABLE III

SPECIFICITY COMPARISON OF EXISTING ABNORMALITY CLASSIFICATION IN WCE IMAGES

	Normal Specificity (%)	Bleeding Specificity (%)	Polyp Specificity (%)	Ulcer Specificity (%)
Method [14]	94.35	93.25	97.17	96.36
Proposed	96.90	98.90	99.30	98.90

TABLE IV

SENSITIVITY COMPARISON OF EXISTING ABNORMALITY CLASSIFICATION IN WCE IMAGES

	Normal Sensitivity (%)	Bleeding Sensitivity (%)	Polyp sensitivity (%)	Ulcer Sensitivity (%)
Method[14]	95.63	93.10	94.00	86.35
Proposed	98.36	97.00	95.00	91.00

Table II,III, and IV shows the accuracies, sensitivities and specificities for each type of WCE image with the existing schemes. Thus our proposed method shows superior classification performance and less processing time than the existing multi abnormality WCE image classification schemes.

V. CONCLUSION

In this paper, we have presented an automatic classification of images from the WCE videos based on extracting SURF points from the images, in the feature coding stage we uses the LLC algorithm. Different from the existing schemes that classifies only single abnormality, our method classifies multi abnormalities in the WCE images.LLC uses the locality constraints to project each descriptor into its local-coordinate system, and the projected coordinates are concatenated by max pooling to create the final representation.

The experimental results acquires a better performance and the corresponding accuracies for ulcer, polyp and bleeding image detections are 96.9%, 98.20%, and 98.40%, respectively, ensures that proposed scheme is highly capable of automatic classification of WCE multi abnormality images. We can extend our work to image classification of other areas, and also by using with other bigger data sets.

References

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

[2] J. L. Toennies, G. Tortora, M. Simi, P. Valdastri, and R. J. Webster, "Swallowable medical devices for diagnosis and surgery: The state of the art," Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci., vol. 224, no. 7, pp. 1397–1414, 2010.

[3] B. Upchurch and J. Vargo, "Small bowel enteroscopy," Rev. Gastroenterol. Disorders, vol. 8, no. 3, pp. 169–177, 2007.

[4] M. Manno, R. Manta, and R. Conigliaro, "Single-balloon enteroscopy," in Ileoscopy, A. Trecca, Ed. Milan, Italy: Springer, 2012, pp. 79–85.

[5] N. M. Lee and G. M. Eisen, "10 years of capsule endoscopy: An update,"Expert Rev. Gastroenterol. Hepatol., vol. 4, no. 4, pp. 503–512, Aug. 2010.

[6] Y. Yuan, J. Wang, B. Li, and M. Q.-H. Meng, "Saliency based ulcer detection for wireless capsule endoscopy diagnosis," IEEE Trans. Med. Imag., vol. 34, no. 10, pp. 2046–2057, Oct. 2015.

[7] D. K. Iakovidis and A. Koulaouzidis, "Software for enhanced video capsule endoscopy: Challenges for essential progress," Nature Rev. Gastroenterol. Hepatol., vol. 12, no. 3, pp. 172–186, 2015.

[8] J.-Y. Yeh,T.-H. Wu, and w.-J. Tsai,"Bleeding and ulcer detection using wireless capsule endoscopy images", Journal of Software Engineering and Applications, vol. 7.no. 5,p. 422, 2014

[9] Y. Fu, W. Zhang, M. Mandal, and M. Q.-H. Meng, "Computer-aided bleeding detection in WCE video," IEEE J. Biomed. Health Informat., vol. 18, no. 2, pp. 636–642, Mar. 2014

[10] Y. Fu, W. Zhang, M. Mandal, and M. Q.-H. Meng, "Computer-aided bleeding detection in WCE video," IEEE J. Biomed. Health Informat., vol. 18, no. 2, pp. 636–642, Mar. 2014

[11] A. Eid, V. S. Charisis, L. J. Hadjileontiadis. And G. D. Sergiadis, "A curvelet- based lacunarity approach for ulcer detection from wireless capsule endoscopy images", in Proceedings of the 26th IEEE International Symposium On Computer-Based Medical Systems. IEEE,2013, pp. 273-278

[12] Y. Yuan and M. Q.-H. Meng, "Polyp classification based on bag of features and saliency in wireless capsule endoscopy," in Proc. IEEE Int. Conf. Robot. Autom., May/Jun. 2014, pp. 3930–3935.

[13] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., Jun. 2009, pp. 1794–1801.

[14] Yixuan Yuan, IEEE, Baopu Li, and Max Q.-H. Meng, "WCE Abnormality Detection Based on Saliency and Adaptive Locality-Constrained Linear Coding", IEEE transactions on automation science and engineering, vol. 14, no. 1, january 2017

[15] H. Bay, A. Ess, T. Tuytelaars and L.Van Gool. SURF: Speeded Up Robust Features. Computer Vision and Image Understanding, vol.110, no.3, pp.346-359, 2008

[16] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality constrained linear coding for image classification," in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 3360–3367.

[17] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," ACM Trans. Intell. Syst. Technol., vol. 2, no. 3, 2011, Art. no. 27.

[18] Y.-L. Boureau, F. Bach, Y. LeCun, and J. Ponce, "Learning mid-level features for recognition," in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 2559–2566

[19] Y.-L. Boureau, F. Bach, Y. LeCun, and J. Ponce, "Learning mid-levelfeatures for recognition," in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., Jun. 2010, pp. 2559–2566.