iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** V **Month of publication:** May 2018

DOI: http://doi.org/10.22214/ijraset.2018.5392

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue V, May 2018- Available at www.ijraset.com

Saliency Aware Locality Constrained Linear Coding for WCE Multi Abnormality Detection based on SURF

Amritha P[1], Rajkumar P[2]
[1,2]*Electronics and Communication Department, Nehru College of Engineering and Research Centre, India*

Abstract*: *Wireless Capsule Endoscopy (WCE) has become a widely 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 automatically detect the multi abnormalities in GI tract from WCE videos. 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. Moreover, it includes the patch saliency constrained on the feature coding stage to highlight the important information in 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*: *Patch saliency, Locality Constrained Linear Coding (LLC), Speeded up Robust Feature (SURF), Wireless Capsule Endoscopy (WCE), Coding bases.

I. INTRODUCTION

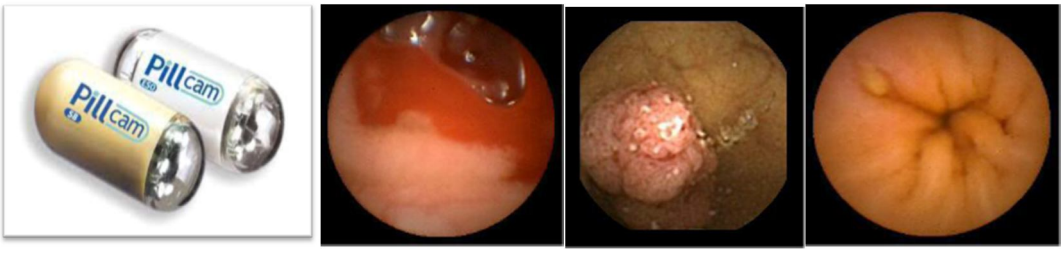
Wireless Capsule Endoscopy (WCE) [1] [Fig. 1(a)] which was first introduced by Given Imaging Incorporation in 2000 and it involves wireless transmission of images from the inner sides of GI tract to the outside environment. WCE has been accepted instead of the traditional ones as it makes possible, painless, non-invasive, disposable [2] and effective diagnostic technique which reduces 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 clinicians to analyse abnormal images.

II. RELATED WORK

Many works have been made in the literature for automatic abnormal image detection from the WCE videos [8]-[14]. The common diseases seen in the GI tract are typically bleeding, polyp, and ulcer, as shown in Fig. 1(b)–(d). Fu et al. [9] introduced rapid bleeding detection in the WCE videos. They segmented images through super pixel segmentation and feature of each pixel is extracted using the red color ratio in RGB color space 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. Charisis et al. [11] introduces a Discrete Curvelet Transform (DCT), method which calculates the lacunarity index of DCT sub bands of the WCE images for acquiring the textural information's to detect the ulcer images. 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. Gong et al. [13] introduces an extension of the SPM method where sparse coding was

©IJRASET: All Rights are Reserved

used instead of vector quantization (VQ). 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. 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.



a b c d

Fig. 1 WCE and captured images. (a) Typical WCE capsule. Example images of (b) bleeding (c) polyp and (d) ulcer.

III.PROPOSED APPROACH

In this paper, a novel method for the automatic detection of multi abnormalities in the GI tract from the WCE videos based on extracting the SURF features [15] from each patches and applying locality constrained linear coding algorithm to encode the images and used for detecting bleeding, polyp 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 each frame is segmented into different patches, we extract the corresponding patch features based on SURF and are constructed by k means clustering to generate the visual words. In the feature coding stage LLC algorithm [16] is used to encode the image patch feature. The patch saliency is explored to emphasize the important features in the coding strategy. 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, and ulcer Fig. 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, each image is segmented into different patches; we extract the corresponding patch features based on SURF. 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 $X = [x1, x2, \ldots . . xN] \in R^{64XN}$.,where N is the number of SURF points for a given image patch.

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 be the size of each type of visual word, the number of visual words ($B = [b1, b2, \ldots . . bK] \in R^{64XK}$) is K. 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

©IJRASET: All Rights are Reserved

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.

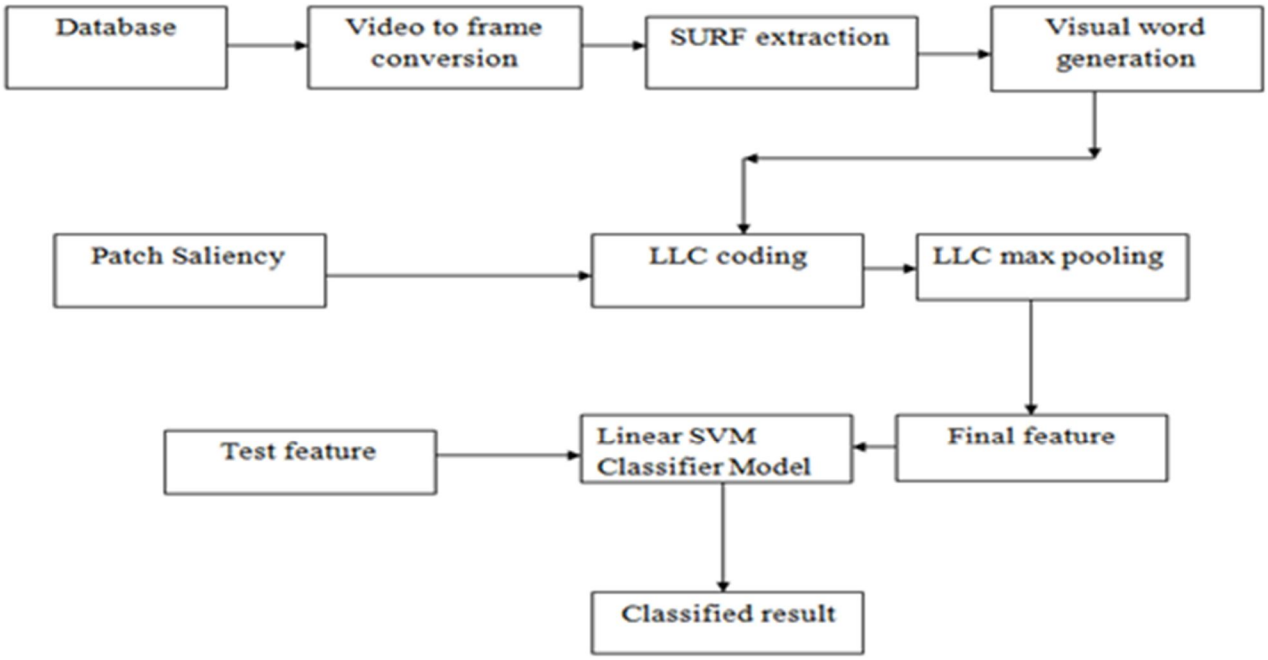


Fig. 2 Block diagram of proposed system

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 is a set of N-dimensional SURF descriptors extracted from the WCE images, i.e. $X = [x1, x2, \dots \dots xN] \in R^{64XN}$. Given visual words of each classes with K entries, $B = [b1, b2, \dots . . bK] \in R^{64XK}$.and, $H = [h1, h2, \dots \dots hN] \in R^{NXK}$ 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:

$$\min_{H} \sum_{i=1}^{N} wi\left||xi - Bhi\right||^2 + \lambda\left||di \odot hi\right||^2$$
$$s.t\ 1^T hi = 1, \forall i \qquad (1)$$

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 distance between 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 $dist(xi, B) = [dist(xi, b1), .. dist(xi, bk)]^T$ 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 $\sum_{i=1}^{N} \lambda\left||di \odot ci\right||^2$ in (1) defines the locality constraint. In LLC coding each SURF description is projected to the fixed point. The solution of the LLC [16] can be derived analytically. By optimizing the object function specified in (1) is as follows:

$$\min_{hi} wi\left||xi - Bhi\right||^2 + \lambda\left||di \odot hi\right||^2 \qquad (3)$$

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

$$L(hi, \beta) = wi\left||xi - Bhi\right||^2 + \lambda\left||di \odot hi\right||^2 + \beta(1^T hi - 1). \qquad (4)$$

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

©IJRASET: All Rights are Reserved

$$\widetilde{hi} = \left(Hi + \lambda\, diag(d)\right)$$

(5)

$$hi = \widetilde{hi}/1^T\widetilde{hi} \quad (6)$$

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

F. Patch saliency

The item $\sum_{i=1}^{N} wi\left||xi - Bhi|\right|^2$ in (1) is the saliency constraint which highlights important information about the patch features.
The patch saliency [18] is defined as the average weighted dissimilarity between a centre patch and its surrounding patches. The saliency wi of the $i\ th$ patch is defined as

$$wi = \sum_{j=1, j\neq i}^{N} \frac{df(i,j)}{1+dlocation(i,j)} \quad (7)$$

Where $df(i,j)$ is the Euclidean distance between the patch feature vectors xi and xj of the i th and j th patches. $dlocation(i,j)$ is the spatial Euclidean distance between the centre of patch i and the centre of patch j . N is the number of patches for the WCE images. In (1), if the saliency of a given patch (wi) is high, then this patch makes more contribution to the feature coding.

G. Llc max pooling

To find the maximum dissimilar feature from the set of SURF points. To generate a WCE image representation each SURF descriptor (xi), the codes (hi) 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.

H. Final Feature

Suppose a frame contains N SURF points and the visual word size is K, then after max pooling, this frame feature $T = [t1, t2, \ldots\ldots tj \ldots\ldots tK]$ will be represented by a K-dimensional vector and tj is calculated by

$$tj = \max\{h1j, h2j, \ldots\ldots hNj\} \quad (8)$$

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.

I. 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.

J. Classifier Model

In this stage we used linear SVM for model the classifier. All features from 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.

K. 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 50 patients. There are (5 normal videos, 15 ulcer videos, 20 bleeding videos and 10 polyp videos). There are total 950 images composed in the dataset of WCE images extracted from 310 polyp images, 210 bleeding images, 210 normal images and 220 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:

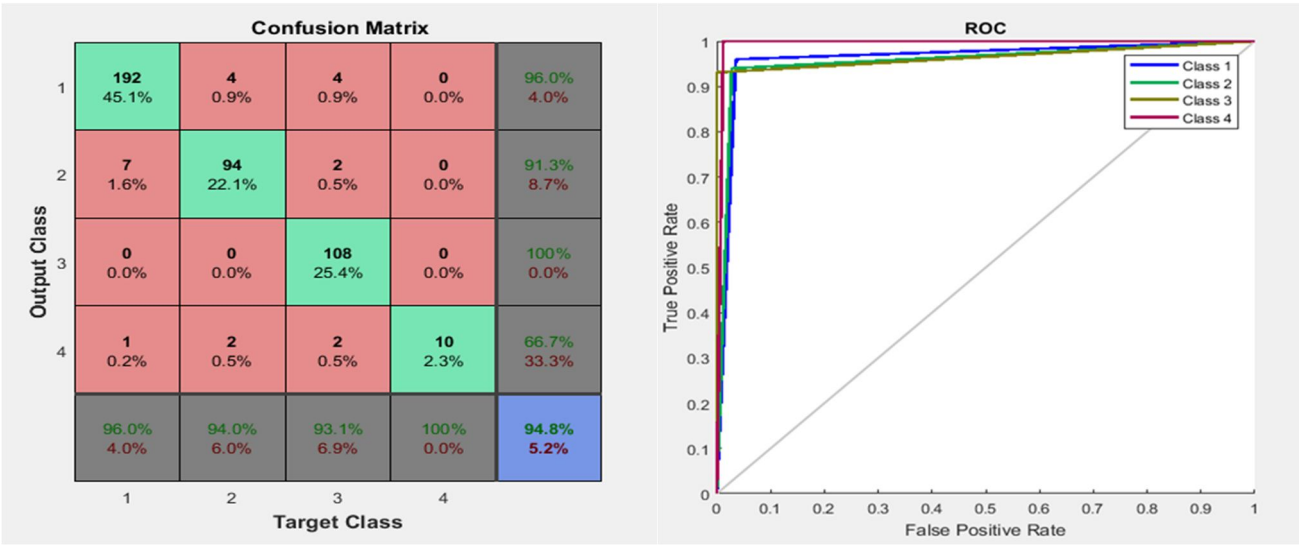
©IJRASET: All Rights are Reserved

$$Accuracy = \frac{No:of correct\ predictions}{Total\ samples} \quad (9)$$

$$Sensitivity = \frac{No:of\ correct\ positive\ predictions}{No:of\ positives} \quad (10)$$

$$Specificity = \frac{No:of\ correct\ negative\ predictions}{No:of\ negatives} \quad (11)$$

The fig. 3(a) 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 96.0%, the polyp images are 94.0%, the bleeding images are 93.1% and finally for detecting the normal images are 95.5%. Receiver Operating Characteristics (ROC) Curve is shown in fig 3 (b) in this curve x-axis with true positive rate and y-axis with false positive rate. The ROC curve is created by plotting the TPR (True Positive Rate) against the FPR (False Positive Rate) at various threshold settings. The TPR is also known as sensitivity and the FPR is specificity.



a b

Fig. 3 (a) Confusion matrix for WCE image classification, (b) ROC Curve of WCE image classification

TABLE I
COMPUTATIONAL PROCESSING TIME OF DIFFERENT CODING SCHEMES FOR WCE CLASSIFICATION

	VQ	SC	SALLC	OURS
Processing Time	186.54 s	326.23 s	210.34 s	164.55 s

TABLE II
PERFORMANCE STATISTICS

	Normal	Bleeding	Polyp	Ulcer
Accuracy	96.18%	95.20%	96.58%	90.13%
Sensitivity	95.23%	93.15%	94.12%	87.14%
Specificity	94.31%	92.52%	96.13%	95.32%

©IJRASET: All Rights are Reserved

The above table I & II describe the performance statistics and computational processing time of this work. 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.

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 detection of multi abnormalities in the GI tract from the WCE videos based on extracting SURF points from the image patches, in the feature coding stage we uses the LLC algorithm. Different from the existing schemes that detects only single abnormality; our method detects multi abnormalities from the WCE images. The patch saliency is explored to emphasize the important features in the coding strategy. 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 96.0%, 94.20%, and 93.10%, 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 set

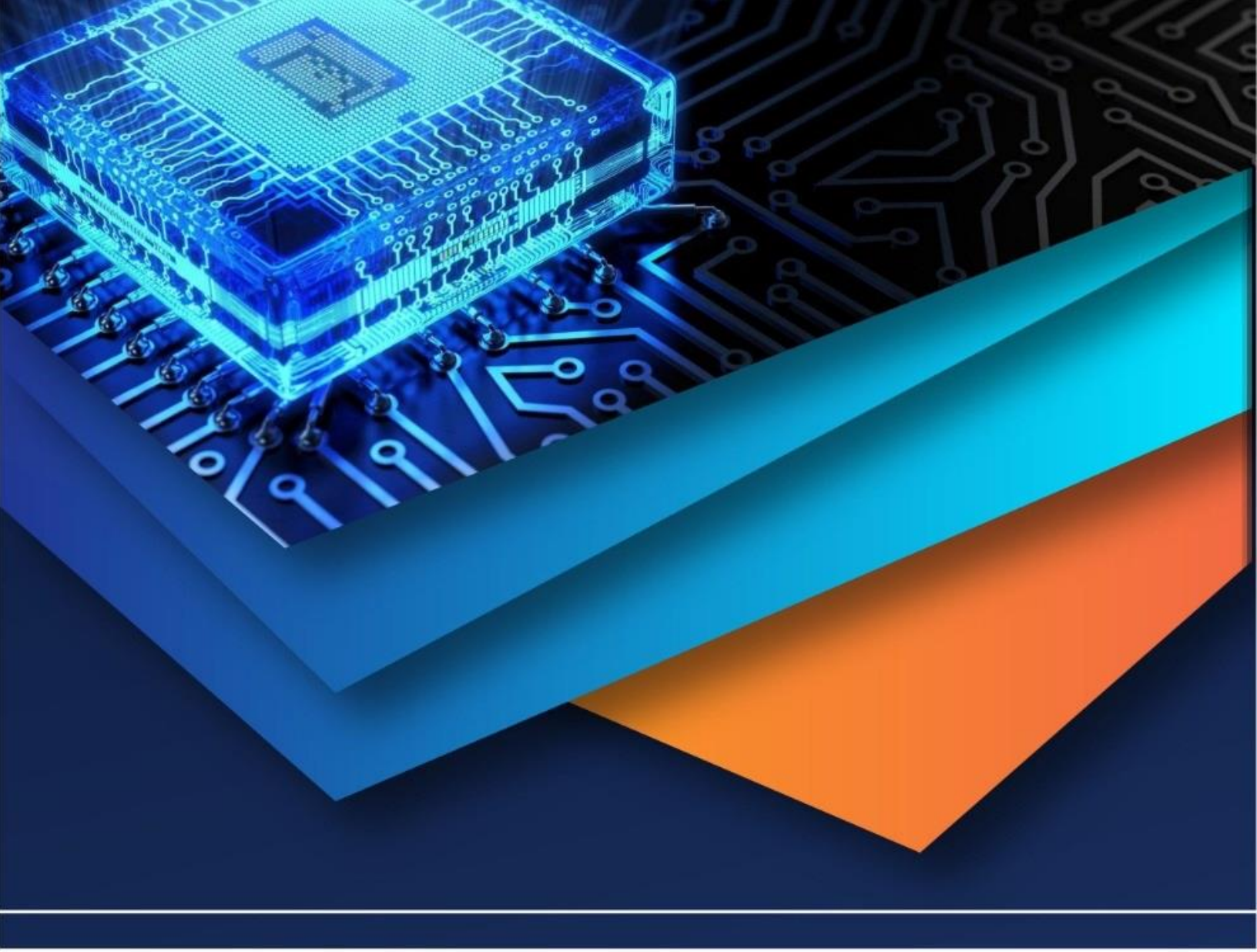
VI. ACKNOWLEDGMENT

The authors thank the head of the ECE department, Mr. S. Rajkumar for his guidance and also are greatful to Nehru College of Engineering and Research centre for the faculties provided.

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] 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
[9] 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
[10] 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
[11] 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
[12] 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.
[13] 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.
[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 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] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254–1259, Nov. 1998.
[19] 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

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)