Computer-aided diagnosis system for ulcer detection in wireless capsule endoscopy images

Said Charfi[1]*, Mohamed El Ansari[1], Ilangko Balasingham[2]

[1] *LabSIV, Department of Computer Science, Faculty of Science, Ibn Zohr University, BP 8106, 80000 Agadir, Morocco.*
[2] *Oslo University Hospital and Department of Electronic Systems, Norwegian University of Science and Technology, Norway*
* *E-mail: charfisaid@gmail.com*

Abstract: Wireless capsule endoscopy (WCE) has revolutionized the diagnosis and treatment of gastrointestinal tract, especially the small intestine which is unreachable by traditional endoscopies. The drawback of the WCE is that it produces a large amount of images to be inspected by the clinicians. Hence, the design of a computer-aided diagnosis (CAD) system will have a great potential to help reducing the diagnosis time and improve the detection accuracy. To address this problem, we propose a CAD system for automatic detection of ulcer in WCE images. Firstly, we enhance the input images to be better exploited in the main steps of the proposed method. Afterwards, segmentation using saliency map based texture and colour is applied to the WCE images in order to highlight ulcerous regions. Then, inspired by the existing feature extraction approaches, a new one has been proposed for the recognition of the segmented regions. Finally, a new recognition scheme is proposed based on hidden markov model using the classification scores of the conventional methods (support vector machine, multilayer perceptron and random forest) as observations. Experimental results with two different datasets show that the proposed method gives promising results.

1 Introduction

Wireless capsule endoscopy (WCE) is such a new technology for non-invasive examination of the gastrointestinal (GI) tract [20]. More information about the current state of things in commercially available capsule endoscopy (CE) device can be found in [30]. WCE produces around 55,000 images per patient which may take a physician around 60 to 90 min for their inspection [38]. A large amount of data makes the task time-consuming and burdensome for a thorough diagnosis. In addition to this, in many cases, it is very difficult to identify some small bleeding regions with naked eyes. These are some of the reasons why a number of research works are being carried out to reduce reading times through automatic detection of images containing abnormalities or other regions of interests (ROIs). Given Imaging (Yoqneam, Israel), one of the manufacturers of WCE, provides a tool called Suspected Bleeding Indicator (SBI) for automatic detection of frames containing possible bleeding areas. However, studies have shown that SBI is not sufficient to screen all types of diseases in the GI tract [7]. This has paved the way to researchers in the development of approaches for automatic detection of other types of abnormalities in CE images with higher accuracy.

In this paper, an automatic approach for ulcer detection is presented. The proposed method can reduce the burden of physicians in investigating WCE video to detect ulcerous frames with a high level of accuracy. Firstly, we pre-process the WCE image for enhancement. Secondly, colour and texture based segmentation using the saliency map is performed for the detection of the ulcerous regions. Clinicians discriminate the ulcer from a WCE image mainly based on some salient information. Here, texture and colour saliency maps are introduced to imitate the diagnosis of physicians. Thirdly, the segmentation process may fail in detecting the exact location of the ulcer (false positives) or segments only a small part of it. Hence, automatic classification phase is highly required to recognize the segmented regions and will be helpful in reducing the review time for physicians. Therefore, several feature descriptors are tested and combined to characterize the segmented regions. Finally, support vector machine (SVM), multilayer perceptron (MLP) and random forest classifiers are used to classify the WCE images and their scores are given as observations to a Hidden Markov Model (HMM) [19] for the recognition. The contributions of this paper can be summarized as follows.

• A new saliency map segmentation method is proposed. Different from the state-of-the art methods in which only the contrast information [1] or the colour information [39] are used to construct the saliency map. In this work, we propose to construct the saliency map using both the texture and colour information.
• Several feature descriptors are used. In the first stage, we test each of the feature descriptors independently. In the next stage, we fuse them to decide which combination is more suitable for characterizing the ulcer in WCE images.
• Different from most state-of-the art methods which use a single classifier to recognize the ulcerous images. A new and more confident recognition scheme is proposed, in which we apply multiple classifiers. The scores of these classifiers are given as observations to the HMM classifier. The constructed scheme overcomes the limitations of the usage of a single classifier and can provide the most probable and confident classification in the cases where a single classifier may fail.

The remainder of the paper is organized as follows. The related work is introduced in Section 2. The methodology used to built the proposed approach is detailed in Section 3. Experimental results are given in Section 4. Section 5 concludes the paper.

2 Related work

Many efforts have been made in the literature of automatic ulcer detection. These methods cqn be classified into three categories based on the type of the features they use for ulcer detection, i.e., colour, texture or colour and texture features.
In [27, 55], colour based methods were developed. Yuan et al. [55] computed the color scale invariant feature transform (SIFT) from the bleeding, polyp, ulcer, and normal WCE image samples separately and then applied K-means clustering on these features to obtain visual words. Then, proposed a novel approach for feature coding based on the Locality-Constrained Linear Coding method. Finally, they utilised the max pooling strategy prior to classification. Kundu et al. [26] proposed to differentiate ulcerous parts based on the brightness of colour, hence, they used the histogram of greyscale image to investigate the distribution of the greyscale values throughout the histogram bins. An other method utilising the histogram in Y plane of YIQ color space was presented in [27]. The second category

of approaches rely on texture features in order to detect ulcer from WCE image [10, 28, 29, 33, 44]. Baopu et al. [28, 29] proposed two methods inspired by wavelet and curvelet based local binary pattern (LBP), respectively, to distinguish ulcer regions from normal ones in patches selected from CE images. Furthermore, Charisis et al. [10] proposed to analyis the texture information in CE images in different colour models. This approach focuses on colour texture features in order to investigate how the structure information of healthy and abnormal tissues is distributed on red, green, and blue (RGB), hue, saturation, value (HSV) and CIELAB colour spaces. Bidimensional ensemble empirical mode decomposition was used as a pre-processing step in order to facilitate differential lacunarity analysis to extract the texture patterns of normal and ulcerous regions in WCE images. LBP and combination of Leung and Malik (LM) filter bank were investigated in [33] for detection of different abnormalities in endoscopy images. SIFT and texture features were extracted from the WCE patches and integrated then K-means algorithm was used to form visual words for automatic ulcer detection in [57]. In [44], a multi-scale approach based on LBP and Laplacian pyramid transform was proposed for ulcer detection. The last category is for methods based on both colour and texture [9, 11, 48, 52]. In [11], aiming at boosting the chromatic features of ulcerous WCE images, the former ones were color rotated. Afterwards, LBP descriptor were utilised for texture information extraction. Colour and textural features were again proposed in [52] to determine the status of the small intestine and detect ulcer and bleeding in WCE images. In this approach authors investigated the RGB, HSV and colour coherence vector for color features extraction then used the gray-level co-occurrence matrix (GLCM) to derive textural ones. In addition, methods such as One Rule (OneR), SVM-recursive feature elimination and ReliefF were utilised for feature selection. Three-stage framework for segmentation and detection of pathology was proposed in [48]. In the aforementioned work, pixels are classified based on the texture features of their neighborhood. Then, classification in performed using a new vector supported convex hull algorithm. In a related work, complete LBP and global-local oriented edge magnitude pattern descriptors were combined to detect texture and colour features of ulcer in WCE images [9]. In a different approach for ulcer detection using WCE images, the geometry features of ulcer were exploited in [25]. This approach is based on Contourlet transform and Log-Gabor filter to distinguish ulcer regions from normal ones. Table 1 displays a comparison of different ulcer detection methods. Some other interesting approaches aiming at detecting several abnormalities using WCE images were also proposed in the literature. Method for localization of bleeding in WCE videos using a combination of edge, intensity and colour saliency maps was proposed in [12]. Bag of feature for automatic polyp detection and method for bleeding detection in WCE images were stated in [46, 53, 54, 56]. Yuan et al. developed two approaches for polyp detection in WCE images. In the first approach, the visual words of all images are computed using SIFT feature vectors and K-means clustering. Then, they calculated the saliency and non-saliency maps of the WCE images. Finally, the histogram of the visual words of each image is computed to represent the WCE image. In the second approach, they improved their first one by combining both SIFT and complete LBP features instead of using only SIFT to compute the visual words. The same authors proposed a wordbased colour histogram

extraction method by applying K-means clustering on the pixel represented images to obtain the cluster centres [53]. A method based on the extraction of geometric features around SIFT keypoints and their concatenation with texture features for polyp detection was presented in [13]. Usman et al. [49], proposed pixel-based bleeding detection method in WCE videos using SVM. A new method for bleeding detection using clustering based features was stated in [16]. In [18], a new feature descriptor that extracts texture feature using normalised GLCM (NGLCM) of the magnitude spectrum was proposed for detecting GI hemorrhage in WCE videos. Yuji Iwahori et al. [22] proposed Hessian filter and histogram of gradients (HOG) features to distinguish between polyp and non-polyp regions, then used K-means++ for classification phase. Moreover, a method based on Weibull features extracted from subbands of dual-tree complex wavelet, Gabor wavelet and shearlet transfroms and fed to K-nn classifier was stated in [50]. Abnormalities detection in WCE images using LBPV descriptor based on the the discrete wavelet transform DWT image anlysis was presented in [8]. A method using superpixel segmentation and naive Bayes classifier for bleeding frames detection was recently proposed in [43]. In [23, 41, 51], systems for small intestine motility characterization and bleeding detection, based on deep convolutional neural networks were introduced. The CE scores to assess small-bowel inflammatory activity in Crohn's disease were evaluated in [37]. In [24], CE and deep enteroscopy in irritable bowel disease were investigated. Study of the effectiveness of colon CE in detecting colon polyps were presented in [40]. The paper [31] summarizes the role of the WCE in pursuing the inflammatory disease in the bowel.

3 Proposed approach

In this section, the new approach for detection and recognition of ulcer in WCE images is presented. As shown in the Fig. 1, in the first step of the proposed scheme, we pre-process the input WCE image using 3D median filter for noise removal and illumination correction. The second step involves colour and texture saliency maps generation and segmentation of the combined saliency map using thresholding. Then, we extract features from the WCE segmented regions. In the last step, the feature vectors are fed to the classification scheme to decide whether the input WCE image is normal or contains an ulcerous region. In the following subsections, we detail all steps involved in the proposed ulcer detection approach.

3.1 Pre-processing

WCE image are obtained from different patients under unstable circumstances and illumination changes in the GI tract. Hence, a pre-processing step is highly required. For these reasons, in this work the capsule images are filtered using 3-D median filtering in a $3 \times 3 \times 3$ neighborhood around each pixel for noise removal. Particularly, we apply an automatic illumination correction scheme [58], for reducing the effect of illumination.

Table 1 Ulcer detection state-of-the-art methods comparison (%).

Method	# of ulcer frames	# of normal frames	accuracy	sensitivity	specificity
[15]	110	110	87.27	88.64	85.75
[25]	65	72	94.16	96.92	91.67
[26]	75	75	87.23	85.13	90.42
[27]	75	75	87.49	83.68	91.08
[45]	133	79	95.11	95.64	94.83
[47]	260	650	92.01	89.90	91.32
[48]	36	500	-	97.00	66.00
[52]	190	258	85.71	87.90	84.11
[57]	170	170	92.65	94.12	91.18
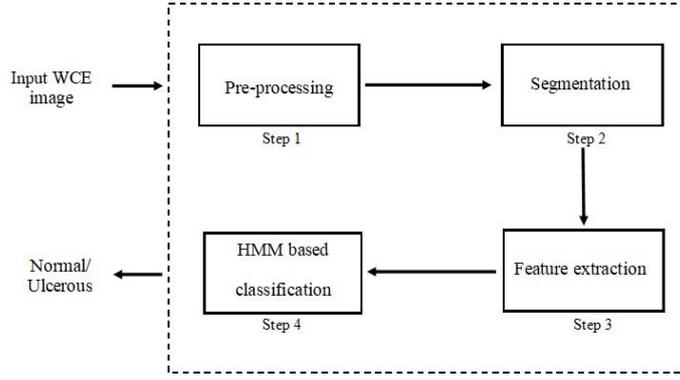
Fig. 1: Framework of the proposed scheme
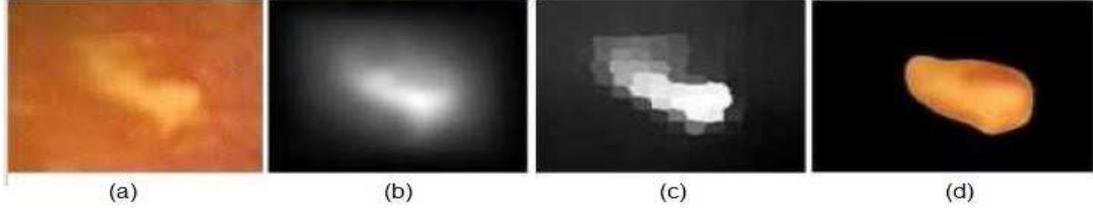
Fig. 2: Segmentation results when applied to a WCE image from the first dataset (a) Original WCE image, (b) Texture saliency map, (c) Colour saliency map, (d) Final saliency map after thresholding

3.2 Segmentation

The presented method considers the superpixel as the element of saliency estimation. First, we over-segment a given image into super-pixels using the Simple linear iterative clustering (SLIC) method [2]. Then, we compute saliency maps based on both texture and colour. Finally, we fuse the generated saliency maps. In the following, segmentation step is detailed.

3.2.1 Colour and texture based saliency map: Colour saliency map::

Colour Saliency Map: For a sub-region r_i produced by the SLIC method, we compute colour feature map using the following formula:

$$S_{col}(r_i) = \sum_{k=1(k \neq i)}^{N} w(r_i, r_k)||I_i - I_k||^2 \quad (1)$$

where

$$w(r_i, r_k) = \exp(-D_e(r_i, r_k))/C_i \quad (2)$$

where I_i are I_k are the mean colour of all pixels in sub-regions r_i and r_k, respectively. $D_e(r_i, r_k)$ is the Euclidean distance between sub-regions r_i and r_k, N is the number of the sub-regions and C_i is scale factor that make sure $\sum_{k=1(k \neq i)}^{N} w(r_i, r_k) = 1$.

Texture saliency map:: We calculate the texture saliency map based on the method presented in [39]. Nevertheless, in this work the texture features of superpixel regions are derived using LBP [34]. The saliency value of each region is given as follows:

$$S_{tex}(r_i) = \sum_{k=1(k \neq i)}^{N} w(r_i, r_k) D_t(r_i, r_k) \quad (3)$$

where $D_t(r_i, r_k)$ refers to the Euclidean distance on texture feature value between the two subregions i and k. It is given by:

$$D_t(r_i, r_k) = \sum_{m}^{l_m} \sum_{n}^{l_n} f(t_{i,m}) f(t_{k,n}) D_e(t_{i,m}, t_{k,n}) \quad (4)$$

where l_m and l_n represent the number of texture type in region r_i and r_k. $f(t_{i,m})$ is the frequency of the mth texture feature among all texture features in region r_i. $f(t_{k,n})$ is the frequency of the nth texture feature among all texture features in region r_k. The frequency of a texture feature reflect the differences between textures, which can be used to measure the weight of this texture feature.

Fusion:: The colour and texture saliency maps are combined using a regularisation constant. The final saliency value $S(r_i)$ of region r_i is given by:

$$S(r_i) = \lambda S_{col}(r_i) + (\lambda - 1)S_{tex}(r_i) \quad (5)$$

where λ is a regularisation constant for colour and texture saliency maps. It is set to 0.5 in the experiments.

3.2.2 Thresholding of saliency map::
In the proposed method, S is generated and used for foreground extraction. From Fig. 2, it can be observed that foreground has higher saliency value than background. Hence, we use thresholding to extract the foreground and the threshold T is defined as:

$$T = \sum_{i=1}^{N} S_i/N \quad (6)$$

where S_i is the saliency value of the region r_i and N is the number of the sub-regions. If $S_i \geq T$ then r_i is a salient region.

3.3 Feature extraction

According to the analysis of the approaches found in the literature, it can be noticed that feature extraction techniques differ to some

extent. Normally, the features used to characterize the ulcer affect surely the performance of any proposed system. In this work, multiple state-of-the-art feature extraction techniques have been tested and combined to verify which ones are more robust and suitable for the proposed system.
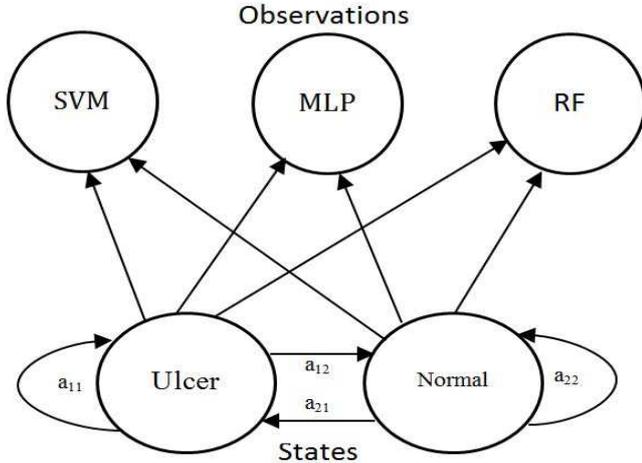

Fig. 3: First-order discrete HMM

3.3.1 *Colour LBP (CLBP):*
The grayscale-LBP descriptor [32, 34] is defined as the LBP descriptor applied to the grayscale image. Colour LBP descriptor in a given colour space is derived by individually computing the LBP descriptor on each of the three colour components. In this paper, we apply the colour LBP in the RGB colour space, especially, we use the LBP_{riu} in each color component. This produces a 30-dimensional descriptor that is formed from concatenating the 10-dimensional vectors of the three channels.

3.3.2 *Pyramid of Histograms of Orientation Gradients (PHOG):*
The PHOG descriptor [6] is able to represent an image by its local shape and the spatial layout of the shape. The local shape is captured by the distribution over edge orientations within a region, and the spatial layout by tiling the image into regions at multiple resolutions. In this paper, we extract ulcer shape information using PHOG [6]. PHOG is a spatial pyramid extension of the HOG descriptors. The HOG descriptor calculates the occurrences of gradient orientation in localized parts of an image. In the first, the PHOG descriptor detects canny edges. Then the ulcer image is divided into spatial grids at all pyramid levels. After that, orientation gradients are calculated by the Sobel mask. At last, the gradients of each grid are linked together at each pyramid level. In our experiments, we set the number of pyramids $L=3$ and the bin size $N=8$, the orientation range is 0–360.

3.3.3 *Bag of visual words:*
This algorithm encloses three main steps. The first step, feature detection, abstracts the image as several patches. The second step is feature description, deals with how to represent the patches as numerical feature descriptors. The last step, named codebook generation, handles the procedure of converting the feature descriptors into codebook which is similar to a word dictionary. In the experiments, we set the number of visual words to 500 for the performance.

3.4 *Recognition-based HMM*

Different from the conventional methods, where the features are fed to a single classifier for decision. In this study each feature vector extracted from the segmented regions is separately classified by a SVM using a radial basis function kernel, MLP and Random RF classifiers. Rather than considering the most likely class that an observation should belongs to (label), we make use of the outputted scores of the aforementioned classifiers which mean the probability

that an observation fit to a particular class. The scores of these classifiers are given as observations to a first-order discrete HMM (Fig. 3). We define the HMM scheme for CE recognition, shown in Fig. 3 as follows:

1. S = {Ulcer, Normal} denotes the interconnected states (N = 2), shown in Fig. 3 and defined as follows:
• Ulcer (s1) A CE image labeled as 'ulcer'.
• Normal (s2) A CE image not labeled as the positive case.
We denote a fixed state sequence of length T as $Q = q_1, q_2, ..., q_T$.
2. V = {SVM, MLP, RF} denotes the observations which are the scores of the classifiers. Each image is classified and a score is given. The observation sequence corresponding to Q is $O = o_1, o_2, ..., o_T$.
3. $A = a_{ij}, 1 \leq i, j \leq N, a_{ij} = P(q_{t+1} = s_j | q_t = s_i)$ denotes the state transition matrix that characterizes the probability of transition between the states for first-order HMM.
4. $B = b_j(k), b_j(k) = P(o_t = v_k | q_t = s_j)$ denotes the observation distribution matrix and represents the probability of a specific observation generated by a state. In our application, it is the probability of classification of an image to a particular class by the classifiers.
5. π denotes the initial state distribution $\pi = \pi_i, \pi_i = P(q_1 = s_i)$ and means the probability of each hidden state occurring at the beginning in any sequence.

The first-order discrete HMM used in this work performs classification in two stages, i.e., training and evaluation. For model learning, two HMMs (one model for each class) are trained using the current image sequence dataset. Feature vectors extracted from the ROI are fed to SVM, MLP and RF classifiers, and then the classifiers score vectors are given as observations to the HMM. The two HMM models are trained using the standard Baum–Welch algorithm [5] (forward procedure) which determines the parameters $\lambda = (A, B, \pi)$ that maximize the probability $P(O_i | \lambda)$.

$$\lambda^* = argmax P(O; A, B, \pi), \tag{7}$$

$$P(O; A, B, \pi) = \sum_Q P(O|Q; A, B, \pi) P(Q; A, B, \pi)$$
$$= \sum_Q (\prod_{t=1}^{T} P(o_t|q_t; B))(\prod_{t=1}^{T} P(q_t|q_{t-1}; A)) \tag{8}$$
$$= \sum_Q (\prod_{t=1}^{T} B_{q_t o_t})(\prod_{t=1}^{T} A_{q_{t-1} q_t})$$

The probability of state occupation is calculated efficiently by means of the expectation maximization algorithm. In the testing phase, we compare the likelihood given by the two HMM models for a given observation. More details about how HMM can be used in (state) classification in a supervised learning context can be found in [36].

4 Experimental results

This section presents the experimental results of the proposed scheme along with its comparison with state-of-the-art methods and discussion.

4.1 *Datasets*

Using a small number of sample frames reduces the generalization ability and can easily lead to over-tuning. The algorithms tested on small datasets can easily fail in realistic conditions as they are based on illusion of good performance. The datasets for WCE pathological abnormality detection are relatively small, which are far to meet the requirement of real applications. For example, the largest datasets so far are [48] and [42]. Although about 50 videos are used
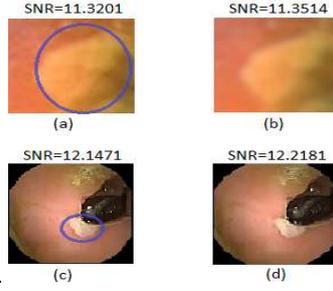
Fig. 4: Samples of ulcerous WCE images from the (a) First dataset, (b) Pre-processed first dataset, (c) Second dataset, (d) Pre-processed second dataset
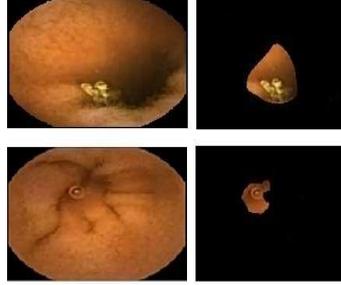
Fig. 5: Samples of false positive segmentation results

for evaluation in [48], there are only a small number of images (around 600 images). In [48] segmentation of intestine content is proposed, in which 95,148 images are used including commonly appeared bubble and turbid that are easier to obtain. In this paper, we use two different datasets for generalization. The first one [35] contains 446 WCE images in which 159 are normal and 287 with ulcer regions. The second one is publicly available at [14] and includes 17 patients with around 2,170 WCE images in which 570 are normal and the remaining contain ulcerous regions. To the best of our knowledge, all the previous works evaluate their performances on a single dataset for automatic ulcer detection for WCE images. The first dataset were divided into 211 and 235 training and testing WCE images, respectively. The resolutions of the images are 256×256 and 424×243 for first and second datastes, respectively. The second one is composed of 2170 WCE were divided into 1050 WCE images for training and 1120 for testing. The ground truth is provided by labeling the original images manually. Normal images are identified as a negative samples and abnormal images as a positive samples. In order to avoid over-fitting of the classification, three-fold cross-validation is applied for all the classification experiments. The classification results are expressed in terms of accuracy, specificity and sensitivity measures, which are defined as follows:

$$Sensitivity = \frac{Number\ of\ correct\ positive\ predictions}{Number\ of\ positives} \quad (9)$$

$$Specificity = \frac{Number\ of\ correct\ negative\ predictions}{Number\ of\ negatives} \quad (10)$$

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ samples} \quad (11)$$

4.2 Pre-processing

This section presents the results of the pre-processing step. Fig. 4 shows two WCE images from the different datasets and their corresponding pre-processed images along with signal-to-noise ratio of each image before and after pre-processing.

4.3 Segmentation

After pre-processing the WCE images, we apply the segmentation method. The segmentation strategy is illustrated using ulcerous WCE image from first dataset as shown in Fig. 2. Fig. 2 (a) shows input WCE images. Fig. 2 (b) shows the texture saliency map generated from the ulcer WCE images. The saliency map generated based of on the colour information is shown in Fig. 2 (c). Then, we fused these two colour and texture saliency maps into one saliency map. The final saliency map of the original image is thresholded to get the most salient region Fig. 2 (d). The proposed segmentation method fails in separating the ulcer regions in some images where stool is present hence resulting in false positives as shown in Fig. 5. Visual comparison of saliency maps (Graph-Based Visual Saliency (GBVS) [17], Itti et al. [21]) with the proposed saliency map is presented in Fig. 6. Table 2 depicts the performance of colour, texture and the combined colour and texture saliency maps.

4.4 Feature extraction

To justify the choice of feature descriptor we have compared their performance on the first dataset. From Table 3, we can see that the performance of CLBP + PHOG + Bag of Words (BoW) features is better than the other feature descriptors. Figs. 7 and 8 illustrate abnormal and normal WCE sample images along with their corresponding CLBP + PHOG + BoW feature vector representation, respectively. The CLBP, PHOG and BoW produce 30, 680 and 500-dimensional descriptor, respectively, given 1210 feature vector dimension when combined.

4.5 Comparison

The performance of the proposed method is compared to those described in [27, 45, 47]. These methods were implemented and applied to the same datasets for a fair comparison. Discussion of the obtained results is presented in the Section 4.6. Tables 4 and 5.
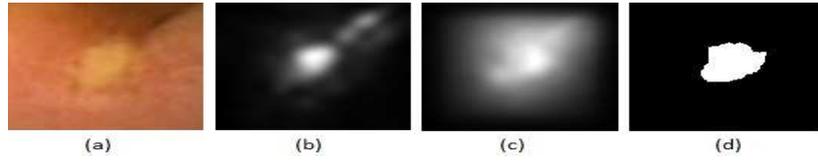
Fig. 6: Saliency maps comparison (a) The original WCE image, (b) GBVS saliency map, (c) Itti's saliency map, (d) The proposed saliency map

Table 2 Comparison between different saliency maps on the first dataset (%).

Saliency	Accuracy	Sensitivity	Specificity
colour	93.7	89.58	93.7
texture	94.9	71.4	96.1
fused	95.3	96.2	96.2

Table 4 Comparison of classification results obtained from the first dataset (%).

	Method [47]	Method [27]	Method [45]	Proposed Method
Acc	91.3	87.5	93.8	95.3
Sens	83.9	85.6	97.2	96.2
Spec	92.1	88.9	87.2	96.2

Table 5 Comparison of classification results obtained from the second dataset (%).

	Method [47]	Method [27]	Method [45]	Proposed Method
Acc	91.3	87.4	88.5	94.8
Sens	83.9	83.6	72.0	96.2
Spec	92.1	91.0	92.7	95.5

Table 3 Feature descriptors performance on the first dataset (%).

Descriptor	Accuracy	Sensitivity	Specificity
CLBP	86.3	90.6	87.1
PHOG	65.4	68.1	84.1
BoW	76.3	76.6	89.4
CLBP + PHOG	81.5	80.0	93.9
CLBP + BoW	93.8	94.1	96.2
PHOG + BoW	91.0	91.9	93.9
CLBP + PHOG + BoW	95.3	96.2	96.2

4.6 Discussion

Aiming at describing WCE image, we proposed a novel saliency map segmentation method that uses both texture and colour information to estimate two saliency maps, since ulcer regions show different colour and texture characteristics compared to their surrounding ones. The generated saliency maps are then fused to construct a better saliency map which better highlights the ulcerous regions in WCE images. The proposed saliency map segmentation method can be considered as an alternative to the classical WCE image segmentation methods since it can outline abnormal region correctly. Subsequently, we apply thresholding method to extract the foreground from the fused saliency map. The experimental results show the superiority of our model in comparison with the existing saliency models in terms of visual effect (shown in Fig. 6). We tested many feature extraction method (i.e., LBP [34], uniform LBP [4], colour LBP [3], PHOG and BoW) to describe the segmented regions and found that when we combine uniform colour LBP + PHOG + BoW feature extractor could obtain better result than the others (Table 3). This performance could be explained by the multi-scale, texture and multi-orientation nature of the ulcer which are better detected by uniform colour LBP + PHOG + BoW descriptors. The task of recognizing ulcer dependent on the segmentation results and the features extracted from the segmented ulcer regions. The presented scheme has proven its capability in detecting and recognizing ulcer disease from WCE images with 94.8% accuracy, 96.2% sensitivity and 95.5% specificity. Compared with the state-of-art methods in the first dataset. The proposed method outperforms the method in [47] with 3.5, 12.3 and 3.4% and the method in [27] with 7.8, 10.6 and 7.3% in terms of accuracy, sensitivity and specificity, respectively. In addition, the proposed method shows ameliorations of 1.5 and 9.0% in terms of accuracy and specificity, respectively, compared to method in [45]. Regarding the second dataset, the proposed method outperforms the method in [47] with 3.5, 12.3 and 3.4%, the method in [27] with 7.4, 12.6 and 4.5% and the method in [45] with 6.3, 24.0 and 2.8% in terms of accuracy, sensitivity and specificity, respectively. This outperformance is due to the segmentation method that highlights well the ulcer regions based on their colour and texture while in [47] they segment the ulcer using the masking method in the RGB bands separately. These masks may not be suitable for all images as the ulcer exhibits different appearances from one WCE image to
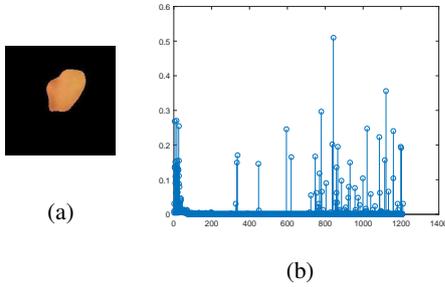
Fig. 7: Segmented ulcer WCE image from the first dataset with its corresponding feature representation (a) Segmented ulcerous WCE image from first dataset, (b) Feature representation
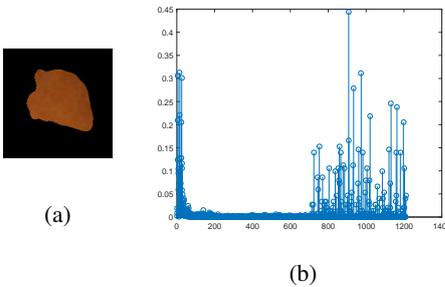
Fig. 8: Segmented normal WCE image from the first dataset with its corresponding feature representation (a) Segmented normal WCE image from the first dataset, (b) Feature representation

another. While in [27, 45], the authors extract colour or texture features directly from the input image without segmenting the ulcerous regions. Besides, the proposed recognition scheme makes use of an ensemble of classifiers that have been tested in most state-of-the art researches to ensure more confidence in the classification results.

Speed efficiency is another important factor to evaluate the performance for potential real-time clinical applications, especially for tasks with large-scale data. Even though the proposed method involves many steps i.e. images enhancement, segmentation, feature extraction and classification using different algorithms still it is computationally efficient. This efficiency in processing time comes from the proposed step of saliency map computation as well as the computational simplicity of the LBP, PHOG and BoW feature extractors. In addition, the classification phase is not timeconsuming as the training phase, that requires a huge number of images, is conducted offline only once. In this work, the training and testing processes are carried out on a computer with 2.0 GHz Intel(R) Core(TM) i3-5005U CPU and 4.0 GB memory. Most of our codes are implemented by Matlab, the time costs for preprocessing, segmentation, feature extraction, classification and total average time spent per image are 0.1262, 0.4331, 0.2408, 0.0033 and 0.8035 s, respectively.

The experimental results shown in section 4 demonstrates the viability of the proposed method and also supports the discussion arguments. For further improvements of the proposed method investigation in other directions are still possible to make the proposed scheme more robust. The number of images in the datasets used for WCE analysis is still one of the biggest limitations. Hence, there is a need of a standard datasets with large number of images in order to effectively validate the performance of the proposed method.

5 Conclusion

In this paper, a novel method for ulcer detection and recognition in WCE images is developed. In the first step, we have pre-processed the input WCE images. Then, we proposed to integrate colour and texture saliency maps based on we segment the resulted map using thresholding. For feature extraction, we found that CLBP + PHOG + BoW give the best result. These features are fed to a new classification scheme based on HMM model and the conventional classifiers. Experimental evidences have proved that the proposed approach remarkably improves the classification results, hence this method is effective in detecting ulcerous images. However, our future work will be devoted to investigate other directions for more improvements of the proposed approach.

6 Acknowledgments

7 References

1 R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk. *Salient Region Detection and Segmentation*, pages 66–75. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

2 R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. Slic superpixels. Technical report, 2010.

3 S. Banerji, A. Verma, and C. Liu. Novel color lbp descriptors for scene and image texture classification. In *15th International Conference on Image Processing, Computer Vision, and Pattern Recognition, Las Vegas, Nevada*, pages 537–543, 2011.

4 O. Barkan, J. Weill, L. Wolf, and H. Aronowitz. Fast high dimensional vector multiplication face recognition. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1960–1967, 2013.

5 J. S. Boreczky and L. D. Wilcox. A hidden markov model framework for video segmentation using audio and image features. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 6, pages 3741–3744. IEEE, 1998.

6 A. Bosch, A. Zisserman, and X. Munoz. Representing shape with a spatial pyramid kernel. In *Proceedings of the 6th ACM International Conference on Image and Video Retrieval*, CIVR '07, pages 401–408, New York, NY, USA, 2007. ACM.

7 J. M. Buscaglia, S. A. Giday, S. V. Kantsevoy, J. O. Clarke, P. Magno, E. Yong, and G. E. Mullin. Performance characteristics of the suspected blood indicator feature in capsule endoscopy according to indication for study. *Clinical gastroenterology and hepatology*, 6(3):298–301, 2008.

8 S. Charfi and M. El Ansari. Computer-aided diagnosis system for colon abnormalities detection in wireless capsule endoscopy images. *Multimedia Tools and Applications*, 77:1–18, 2017.

9 S. Charfi and M. El Ansari. Computer-aided diagnosis system for ulcer detection in wireless capsule endoscopy videos. In *Advanced Technologies for Signal and Image Processing (ATSIP), 2017 International Conference on*, pages 1–5. IEEE, 2017.

10 V. S. Charisis, L. J. Hadjileontiadis, C. N. Liatsos, C. C. Mavrogiannis, and G. D. Sergiadis. Capsule endoscopy image analysis using texture information from various colour models. *Comput. Meth. Prog. Bio.*, 107(1):61–74, 2012.

11 V. S. Charisis, C. Katsimerou, L. J. Hadjileontiadis, C. N. Liatsos, and G. D. Sergiadis. Computer-aided capsule endoscopy images evaluation based on color rotation and texture features: An educational tool to physicians. In P. P. Rodrigues, M. Pechenizkiy, J. Gama, R. Cruz-Correia, J. Liu, A. J. M. Traina, P. J. F. Lucas, and P. Soda, editors, *CBMS*, pages 203–208. IEEE Comput. Soc., 2013.

12 H. Chen, S. Wang, Y. Ding, and D. Qian. Saliency-based bleeding localization for wireless capsule endoscopy diagnosis. *International journal of biomedical imaging*, 2017, 2017.

13 M. El Ansari and S. Charfi. Computer-aided system for polyp detection in wireless capsule endoscopy images. In *Wireless Networks and Mobile Communications (WINCOM), 2017 International Conference on*, pages 1–6. IEEE, 2017.

14 C. Endoscopy. Capsule endoscopy products, 2018.

15 T. Ghosh, A. Das, and R. Sayed. Automatic small intestinal ulcer detection in capsule endoscopy images. *International Journal of Scientific and Engineering Research*, 7(10):737–741, October 2016.

16 T. Ghosh, S. A. Fattah, K. A. Wahid, W.-P. Zhu, and M. O. Ahmad. Cluster based statistical feature extraction method for automatic bleeding detection in wireless capsule endoscopy video. *Computers in Biology and Medicine*, 2018.

17 J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. In *Advances in neural information processing systems*, pages 545–552, 2007.

18 A. R. Hassan and M. A. Haque. Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos. *Computer methods and programs in biomedicine*, 122(3):341–353, 2015.

19 M. Hassan, M. Amin, I. Murtza, A. Khan, and A. Chaudhry. Robust hidden markov model based intelligent blood vessel detection of fundus images. *Computer Methods and Programs in Biomedicine*, 151:193–201, 2017.

20 G. Iddan, G. Meron, A. Glukhovsky, and P. Swain. Wireless capsule endoscopy. *Nature*, 405(6785):405–417, 2000.

21 L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 20(11):1254–1259, 1998.

22 Y. Iwahori, A. Hattori, Y. Adachi, M. K. Bhuyan, R. J. Woodham, and K. Kasugai. Automatic detection of polyp using hessian filter and hog features. In *KES*, volume 60, pages 730–739, 2015.

23 X. Jia and M. Q. Meng. A deep convolutional neural network for bleeding detection in wireless capsule endoscopy images. In *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2016, Orlando, FL, USA, August 16-20, 2016*, pages 639–642, 2016.

24 U. Kopylov, D. Carter, and A. R. Eliakim. Capsule endoscopy and deep enteroscopy in irritable bowel disease. *Gastrointest. Endosc. Clin. N. Am.*, 26(4):611–627, 2016.

25 N. E. Koshy and V. P. Gopi. A new method for ulcer detection in endoscopic images. In *Electronics and Communication Systems (ICECS), 2015 2nd International Conference on*, pages 1725–1729. IEEE, 2015.

26 A. Kundu, A. Bhattacharjee, S. Fattah, and C. Shahnaz. Automatic ulcer detection scheme using gray scale histogram from wireless capsule endoscopy. In *Electrical and Computer Engineering (WIECON-ECE), 2016 IEEE International WIE Conference on*, pages 242–245. IEEE, 2016.

27 A. Kundu, A. Bhattacharjee, S. Fattah, and C. Shahnaz. An automatic ulcer detection scheme using histogram in yiq domain from wireless capsule endoscopy images. In *Region 10 Conference, TENCON 2017-2017 IEEE*, pages 1300–1303. IEEE, 2017.

28 B. Li and M. Q. H. Meng. Texture analysis for ulcer detection in capsule endoscopy images. *Image Vision Comput.*, 27(9):1336–1342, Aug. 2009.

29 B. Li, M. Q.-H. Meng, and J. Y. W. Lau. Computer-aided small bowel tumor detection for capsule endoscopy. *Artif. Intell. Med.*, 52(1):11–16, 2011.

30 Z. Li, D. Carter, R. Eliakim, W. Zou, H. Wu, Z. Liao, Z. Gong, J. Wang, J. W. Chung, S. Y. Song, et al. The current main types of capsule endoscopy. pages 5–45, 2014.

31 I. V. Mitselos, D. K. Christodoulou, K. H. Katsanos, and E. V. Tsianos. Role of wireless capsule endoscopy in the follow-up of inflammatory bowel disease. *world journal of gastrointestinal endoscopy*, 7:643–651, 2015.

32 V. Murugappan and R. Sabeenian. Texture based medical image classification by using multi-scale gabor rotation-invariant local binary pattern (mgrlbp). *Cluster Computing*, pages 1–14, 2017.

33 R. Nawarathna, J. Oh, J. Muthukudage, W. Tavanapong, J. Wong, P. C. De Groen, and S. J. Tang. Abnormal image detection in endoscopy videos using a filter bank and local binary patterns. *Neurocomputing*, 144:70–91, 2014.

34 T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987, July 2002.

35 W. E. Organization. Weo clinical endoscopy atlas, 1962.

36 J. Pohle, R. Langrock, F. M. van Beest, and N. M. Schmidt. Selecting the number of states in hidden markov models: pragmatic solutions illustrated using animal movement. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3):270–293, 2017.

37 A. Ponte, R. Pinho, A. Rodrigues, J. Silva, J. Rodrigues, M. Sousa, and J. Carvalho. Evaluation and comparison of capsule endoscopy scores for assessment of inflammatory activity of small-bowel in crohn's disease. *Gastroenterologia y hepatologia*, 2017.

38 W. A. Qureshi. Current and future applications of the capsule camera. *Nature reviews drug discovery*, 3(5):447–450, 2004.

39 Y.-F. Ren and Z.-C. Mu. Salient object detection based on global contrast on texture and color. In *Machine Learning and Cybernetics (ICMLC), 2014 International Conference on*, volume 1, pages 7–12. IEEE, 2014.

40 T. Rokkas, K. Papaxoinis, K. Triantafyllou, and S. D. Ladas. A meta-analysis evaluating the accuracy of colon capsule endoscopy in detecting colon polyps. *Gastrointest. Endosc.*, 71(4):792–798, 2010.

41 S. Seguí, M. Drozdzal, G. Pascual, P. Radeva, C. Malagelada, F. Azpiroz, and J. Vitrià. Generic feature learning for wireless capsule endoscopy analysis. *Computers in biology and medicine*, 79:163–172, 2016.

42 S. Seguí, M. Drozdzal, F. Vilariño, C. Malagelada, F. Azpiroz, P. Radeva, and J. Vitrià. Categorization and segmentation of intestinal content frames for wireless capsule endoscopy. *IEEE Trans. Information Technology in Biomedicine*, 16(6):1341–1352, 2012.

43 P. Sivakumar and B. M. Kumar. A novel method to detect bleeding frame and region in wireless capsule endoscopy video. *Cluster Computing*, pages 1–7, 2018.

44 M. Souaidi, A. A. Abdelouahad, and M. El Ansari. A fully automated ulcer detection system for wireless capsule endoscopy images. In *Advanced Technologies for Signal and Image Processing (ATSIP), 2017 International Conference on*, pages 1–6. IEEE, 2017.

45 M. Souaidi, A. A. Abdelouahed, and M. El Ansari. Multi-scale completed local binary patterns for ulcer detection in wireless capsule endoscopy images. *Multimedia Tools and Applications*, pages 1–18, 2018.

46 M. Souaidi, S. Charfi, A. A. Abdelouahad, and M. El Ansari. New features for wireless capsule endoscopy polyp detection. In *2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, pages 1–6. IEEE, 2018.

47 S. Suman, F. A. Hussin, W. Nicolas, and A. S. Malik. Ulcer detection and classification of wireless capsule endoscopy images using rgb masking. *Advanced Science Letters*, 22(10):2764–2768, Obtober 2016.

48 P. Szczypiński, A. Klepaczko, M. Pazurek, and P. Daniel. Texture and color based image segmentation and pathology detection in capsule endoscopy videos. *Computer methods and programs in biomedicine*, 113(1):396–411, 2014.

49 M. A. Usman, G. B. Satrya, M. R. Usman, and S. Y. Shin. Detection of small colon bleeding in wireless capsule endoscopy videos. *Computerized Medical Imaging and Graphics*, 54:16–26, 2016.

50 G. Wimmer, T. Tamaki, J. J. W. Tischendorf, M. Häfner, S. Yoshida, S. Tanaka, and A. Uhl. Directional wavelet based features for colonic polyp classification. *Med. Image Anal.*, 31:16–36, 2016.

51 K.-j. Xia, H.-s. Yin, and J.-q. Wang. A novel improved deep convolutional neural network model for medical image fusion. *Cluster Computing*, pages 1–13.

52 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*, 7(5):422, 2014.

53 Y. Yuan, B. Li, and M. Q. Meng. Bleeding frame and region detection in the wireless capsule endoscopy video. *IEEE J. Biomed. Health Inform.*, 20(2):624–630, 2016.

54 Y. Yuan, B. Li, and M. Q. Meng. Improved bag of feature for automatic polyp detection in wireless capsule endoscopy images. *IEEE Trans. Autom. Sci. Eng.*, 13(2):529–535, 2016.

55 Y. Yuan, B. Li, and M. Q.-H. Meng. Wce abnormality detection based on saliency and adaptive locality-constrained linear coding. *IEEE Trans. Autom. Sci. Eng.*, 7(10):737–741, October 2016.

56 Y. Yuan and M. Q.-H. Meng. Polyp classification based on bag of features and saliency in wireless capsule endoscopy. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 3930–3935. IEEE, 2014.

57 Y. Yuan, J. Wang, B. Li, and M. Q.-H. Meng. Saliency based ulcer detection for wireless capsule endoscopy diagnosis. *IEEE Trans. Med. Imag.*, 34(10):2046–2057, 2015.

58 Y. Zheng, J. Yu, S. B. Kang, S. Lin, and C. Kambhamettu. Single-image vignetting correction using radial gradient symmetry. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.