Biomedical Signal Processing and Control 71 (2022) 103219

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Feature extraction for abnormality detection in capsule endoscopy images

Zahra Amiri [a], Hamid Hassanpour [a,*], Azeddine Beghdadi [b]

[a] Image Processing & Data Mining Lab, Shahrood University of Technology, Shahrood, Iran
[b] Department of Computer Science & Engineering, University Sorbonne Paris Nord, Villetaneuse, France

ARTICLE INFO

Keywords:
Abnormality detection
Angiodysplasia
Bleeding
Capsule endoscopy

ABSTRACT

Capsule endoscopy is a non-invasive method for diagnosing gastrointestinal diseases. This new technology has many advantages over conventional endoscopy. However, investigating endoscopic video frames in search of diseases is a tedious task for physicians. Hence, a system is required to automatically detect suspicious frames for further medical examination. Different abnormalities may exist in capsule endoscopy images. In this paper, a novel method is proposed to investigate capsule endoscopy images for abnormalities such as bleeding and angiodysplasia lesions. The proposed method identifies potential regions of interest by using an expectation maximization based image segmentation algorithm, and extracts features from them using a combination of color histogram analysis and statistical features to classify frames into normal and abnormal classes. The results show that the proposed method can distinguish the two objective classes with an approximate precision and recall of 96.5% and 95.9%, respectively.

1. Introduction

Wireless Capsule Endoscopy (WCE) is a technology developed to record video frames from the gastrointestinal tract for endoscopic examination. WCE comes in the form of a pill containing a miniaturized camera consisting of CCD or CMOS sensors, batteries and a radio frequency transmitter [1]. The whole acquisition and transmission process takes approximately 8 h generating around 50,000 frames per patient [2]. WCE has many applications, especially investigating abnormalities in the small bowel because this organ is very difficult to explore with the other conventional endoscopic examination techniques.

The analysis of such huge volume data is time-consuming and very tedious for a physician. It usually takes between one to three hours to inspect all the acquired video frames [3]. Therefore, it is desired to design a computer-aided diagnosis system to detect abnormal frames in order to help the physician for subsequent analysis and ultimate verification. It was reported that bleeding in the gastrointestinal tract is the most common abnormality witnessed in WCE frames. Besides, the most common cause of life-threatening overt bleeding is due to angiodysplasia (AD) lesion which is a small vascular malformation of the gut [4]. Hence in this paper, a novel technique is proposed for automatically detecting AD and bleeding in WCE frames. Several samples of WCE images are shown in Fig. 1.

Several challenges and limitations exist in WCE abnormality detection. Quality of WCE images is not high due to the power and volume limitations. WCE images also suffer from noise and blurriness [1]. Another issue that complicates the abnormality detection process is the high similarity between normal and abnormal frames in some cases (see Fig. 2). The presence of bubbles and turbid (food residues in the gastrointestinal tract) in some frames makes the detection process difficult [3]. Despite all the existing limitations, we will present a method that is able to detect AD and bleeding frames with high accuracy.

In our previous work [5], we introduced a method to accurately detect bleeding in WCE images. However, this method fails to diagnose the abnormality if both bleeding and AD coexist in the gastrointestinal tract. Hence, we extend our earlier work in this study, and develop a new solution to identify the existence of any of the two lesions in WCE images.

The proposed method operates in two steps. First, potential regions of interest (ROI) are automatically detected from the images by using the Expectation Maximization (EM) based image segmentation algorithm. Then, the Histogram Equalization (HE) based contrast enhancement is applied on each image in order to reduce changes in illumination and improve the visibility of the salient features in the frames. There are of course more effective but rather complex contrast enhancement methods. But given the huge amount of data to be processed, the classical HE method is a good choice because of its simplicity and the low

* Corresponding author.
E-mail address: h.hassanpour@shahroodut.ac.ir (H. Hassanpour).

https://doi.org/10.1016/j.bspc.2021.103219
Received 12 September 2020; Received in revised form 11 September 2021; Accepted 30 September 2021
Available online 9 October 2021

cost of computing time. Afterward, by using a combination of color histogram analysis and some statistical features, a set of relevant features is extracted from the ROIs. More precisely, we extend the set of features used in our previous work [5] by identifying effective color channels in the detection of considered lesions, and adding a new and effective feature using color histogram analysis. Indeed, the color histogram analysis can reveal the existence of distinctive color due to lesion in the background. Finally, the features are fed into a multi-layer perceptron (MLP) neural network to categorize the frames into normal and abnormal classes. The main contributions and novelties of the proposed study are the following:

- Detecting any AD or bleeding lesions in WCE images by using two set of novel features.
- Proposing a novel and new feature set that can reveal existence of sudden color change due to lesion in background by analyzing histogram.
- Providing simple and efficient method for identifying effective color channels in different color space for feature extraction from lesion areas.

This paper is organized as follows. The related works are reviewed in Section 2. The proposed method is explained in Section 3. Section 4 involves experiments and results. Finally, conclusion and future works are described in Section 5.

2. Related works

Detection of abnormal frames in capsule endoscopy images has been investigated in several studies [6–12], including detection of tumor, Crohn's disease, polyp, bleeding, ulcer, lymphangiectasia and other intestinal lesions. Existing methods usually use feature extraction followed by a detection method. The detection methods either differentiate between frames containing lesion and normal ones [6], or employ multi-label approaches to detect and classify different types of abnormality [7]. In these approaches features are commonly extracted from the color descriptors, statistical features analysis, and texture and morphological analysis in the images. These methods are either pixel-based [6,7] or region-based [8,9] approaches. Support Vector Machine (SVM) [10] and Artificial Neural Networks (ANNs) [11,12] are the two frequently used methods for classification in the literature.

Several methods have been developed in literature for hemorrhage detection. Deeba *et al.* in [13] proposed a two-stage unsupervised abnormality detection algorithm. In the first stage, a Retinex-theory based algorithm is applied to the images for adaptive color contrast enhancement. In the second stage, clinically significant regions are detected. The method yielded poor results. Two reasons could lead to this results: inappropriate evaluation criteria and very small dataset that contained totally 77 images for seven different abnormalities. The percentage of ROI detected by this method was considered as evaluation criteria which was not appropriate for evaluating segmentation method. Because this criterion did not consider false positives rate. Therefore, in the article, it can be seen that the value of specificity was low and this was the result of the existence of many false positives in the proposed area.

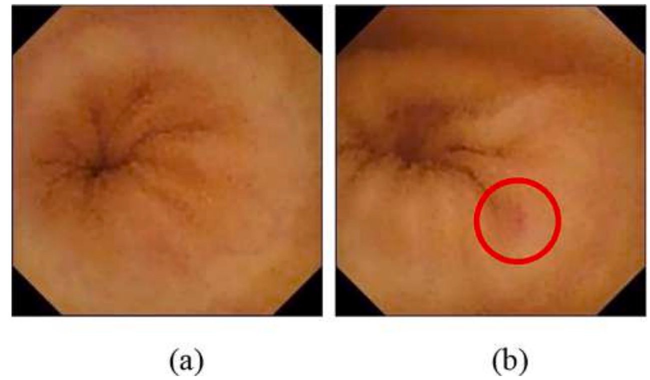


Fig. 2. High similarity between the two frames in different classes: a) normal, b) AD. The AD lesion is marked with a circle.

Supervised abnormality detection methods are more common. This is the case of the pixel-wise classification [11,14] methods. However, these approaches have high computational complexity. Hence, they are not a good choice for processing a large number of video frames which is the case in the context of WCE.

Extracting features from the image patches [12,15] or the whole image [16,17] is considered as a solution to speed up investigating WCE frames for abnormalities. However, these methods, especially feature extraction from the whole image, have a low sensitivity at detecting small bleeding regions [18]. In Yuan's method [17], the colors existing in WCE images are clustered by using *k*-means algorithm and the centers of clusters are considered as words for the computation of a histogram. Then each pixel in the image is mapped to the nearest center. So after this step, there is a set of pixels for each center. The centers, along with the number of pixels assigned to them, form the final histogram. This histogram for each image are considered as a feature set. Then SVM is used to classify images into normal or bleeding ones. Recently, deep learning methods have been considered for bleeding diagnosis [19–21]. However, deep learning methods require intensive and time-consuming training on a very large training dataset.

Our investigations indicate that AD detection has been considered in few studies. In some of the earlier studies, one of the main drawbacks is that they used databases with a limited number of lesion instances. In [22], for instance, the authors used an AD image dataset with only 27 samples to segment AD in WCE images. The authors have claimed that their method is the first regarding automatic AD lesions detection. This work proposed an automatic selection of ROI using an image segmentation method based on the EM algorithm. In some other researches,

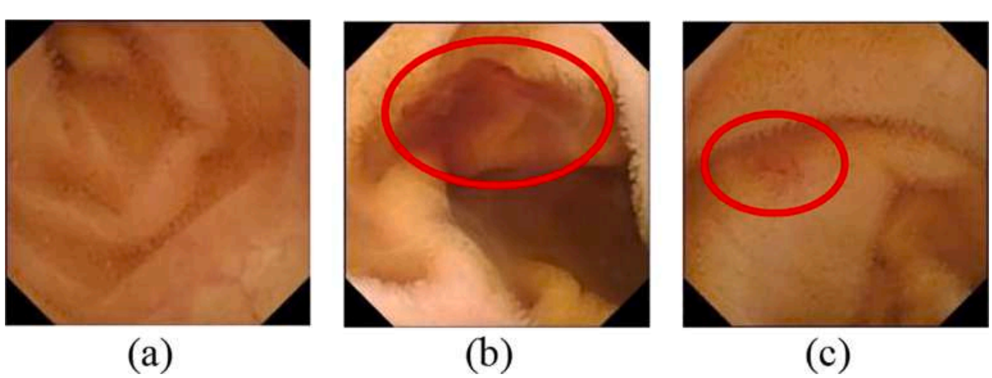


Fig. 1. Sample WCE images: a) normal frame, b) frame with bleeding, c) AD lesion. The lesion area is specified in an ellipse.

such as [8], the number of lesion instances is considerable but the methods do not have high accuracy. In [8], ROIs are extracted from images and are classified using a combination of texture, color and statistical and morphological features. The authors reported a sensitivity of 89.51% and a specificity of 96.8%. But these results are merely from detection of potential ROIs. However, the overall accuracy of the method drops when it attempts to distinguish normal and abnormal images as several potential ROIs may exist in each image. Therefore, the classification of the whole image into normal and abnormal is less accurate.

Therefore, the introduced methods have various problems such as inadequate accuracy, inappropriate evaluation criteria, evaluation on a dataset with insufficient number of images or focusing on a specific lesion. In the following, we proposed a method that has the ability to detect two lesions together with high accuracy. To evaluate the proposed method proper evaluation criteria on a large dataset have been used.

3. Material and method

In this paper, a novel method for abnormality detection in capsule endoscopy frames is proposed. The ROIs corresponding to lesion candidates are initially selected from the frames. Then suitable features are extracted from the ROIs for abnormality detection.

The block diagram of the proposed method can be seen in Fig. 3. The proposed method has three main stages including ROI extraction, feature extraction and classification. In the first stage, ROIs are extracted from images by using the EM segmentation method on the '*a*' component of the CIELab color space. Then, the winner segment is selected and a morphology filter is applied to it. In the next stage, initially histogram equalization is applied to the image. Then, two sets of features containing statistical features and histogram-based features are extracted from the ROI. Statistical features are extracted from four different components of various color spaces, and histogram-based features are extracted from the '*a*' component of the CIELab color space. Finally, the Multi-Layer Perceptron (MLP) [23] classifier is used to determine the class of images.

3.1. ROI selection

Segmentation of images into suspect and non-suspect regions is the main aim of this section. Since angiodysplasia can be a very small lesion compared to the whole image, we initially extract the suspect region for further investigation. ROI extraction helps to generate features from lesion or non-lesion area more precisely [24]. A lesion is an integrated area in a background. Although the area is distinctive, some of their pixels may have the same color as the background. However, the distribution of pixels from the lesion is different from those in the background. In addition, our investigations show that bleeding and AD lesions in component 'a' of CIELab color space distinctively appear with higher values compared to the background (see Fig. 4). Therefore, the difference in the distribution of lesion pixels from the background is more obvious in this channel. In Fig. 5 the normal distribution for the lesion and non-lesion (background) areas is shown on three images for the 'a' component. As can be seen, the normal distribution of the lesion area has a higher mean and lower variance values compared to the non-lesion area in the three images. To further investigate, we randomly selected 200 images from the dataset and performed the same experiment on them. The averages of mean and variance of normal distributions in lesion area are respectively higher and lower than non-lesion areas as reported in Table 1.

Image pixels can be considered as a combination of several different Gaussian distributions (k distributions). Hence, image pixels can be modeled using a Gaussian Mixture Model (GMM) [25]. Indeed, images containing abnormality can be segmented considering these different Gaussian distributions. EM algorithm is a good choice for estimating parameters of the k Gaussian distributions.

A mixture of k distinct normal distributions is considered for N pixels in the image $\{x^{(1)}, x^{(2)}, \cdots, x^{(N)}\}$. After segmentation, the brightest segment is suspected to contain the lesion. So, modeling the data by specifying a joint distribution is desired using Eq. (1).

$$p\left(x^{(i)}, z^{(i)}\right) = p\left(x^{(i)}|z^{(i)}\right)p\left(z^{(i)}\right) \tag{1}$$

where

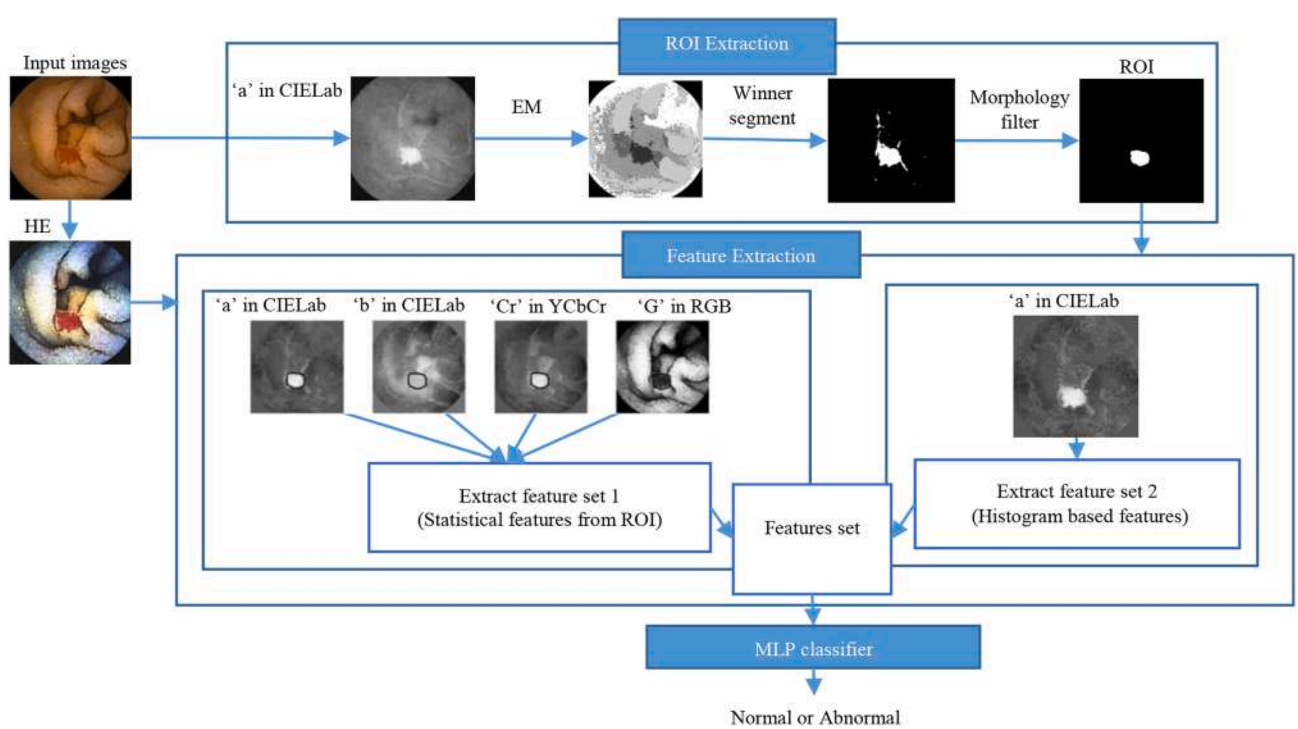


Fig. 3. Block diagram of the proposed method.

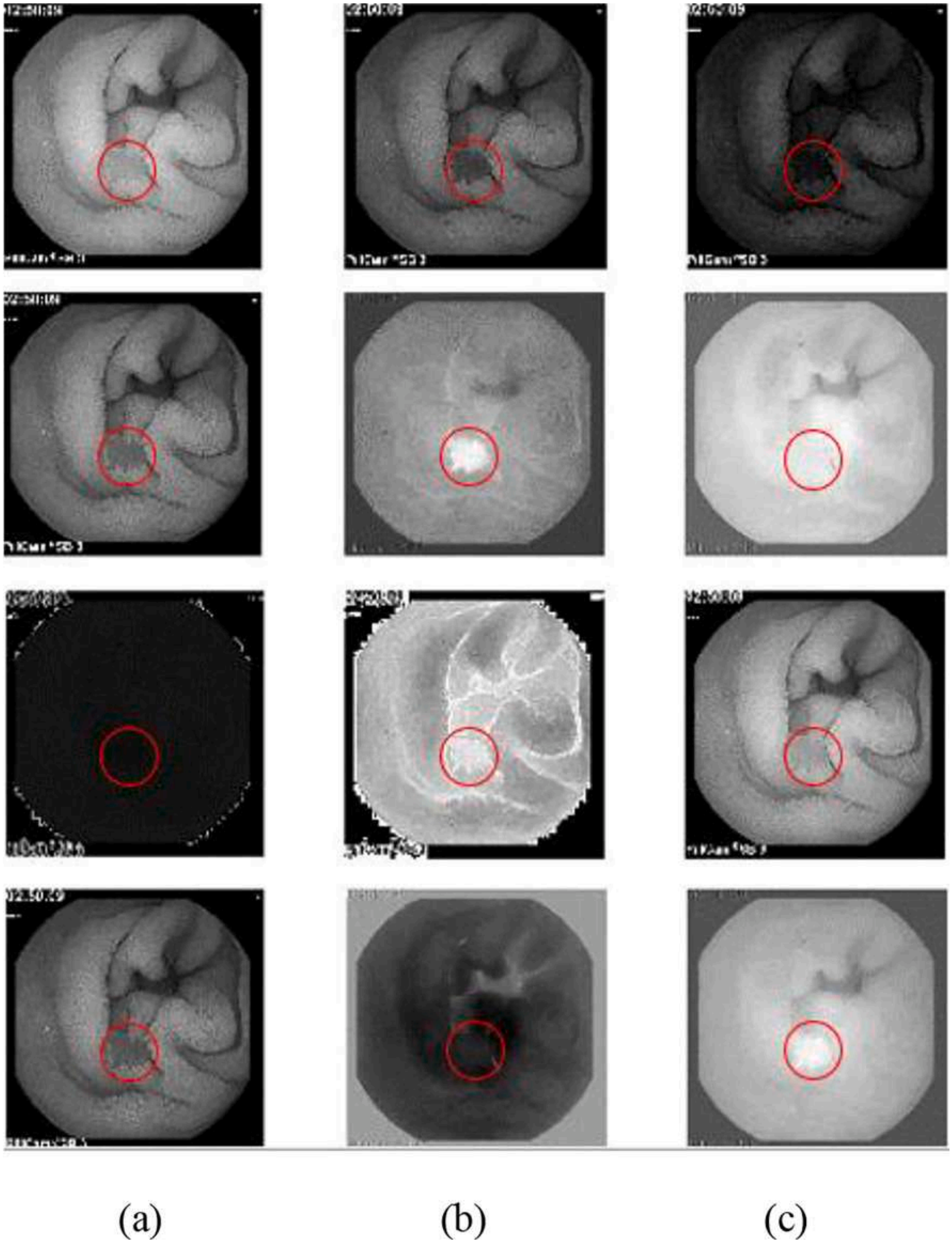


Fig. 4. AD lesion in twelve color channels of different color spaces. Each row shows the image in different channels of a specific color space. The rows are in order from top to bottom, RGB, CIELAB, HSV, and YCbCr color space. The columns are respectively: a) the first channel, b) the second channel, c) the third channel in the related color space. Lesions areas are covered with a red circle.

$$\begin{cases} z^{(i)} \sim mutinomial(\varphi), \varphi_j \geq 1, \sum_{j=1}^{k} \varphi_j = 1 \\ x^{(i)} | z^{(i)} = j \sim N(\mu_j, \sigma_j) \end{cases}$$

In Equation (1), $z^{(i)}$is a latent variable and can take k values. Also, φ_j gives the probability that $z^{(i)}$ equals to the jth class. μ_j and σ_j are the mean and standard deviation of the jth considered normal distribution. To estimate the parameters, the Maximum Likelihood (ML) criterion can be used according to Equation (2).

$$\ell(\varphi, \mu, \sigma) = \sum_{i=1}^{N} logp(x^{(i)}; \varphi, \mu, \sigma) = \sum_{i=1}^{N} log \sum_{z^{(i)}=1}^{k} p(z^{(i)}; \varphi) p(x^{(i)} | z^{(i)}; \mu, \sigma) \quad (2)$$

Maximization of Eq. (2) leads to obtaining the values of class coefficient, mean and variance, respectively in Eqs. (3)–(5).

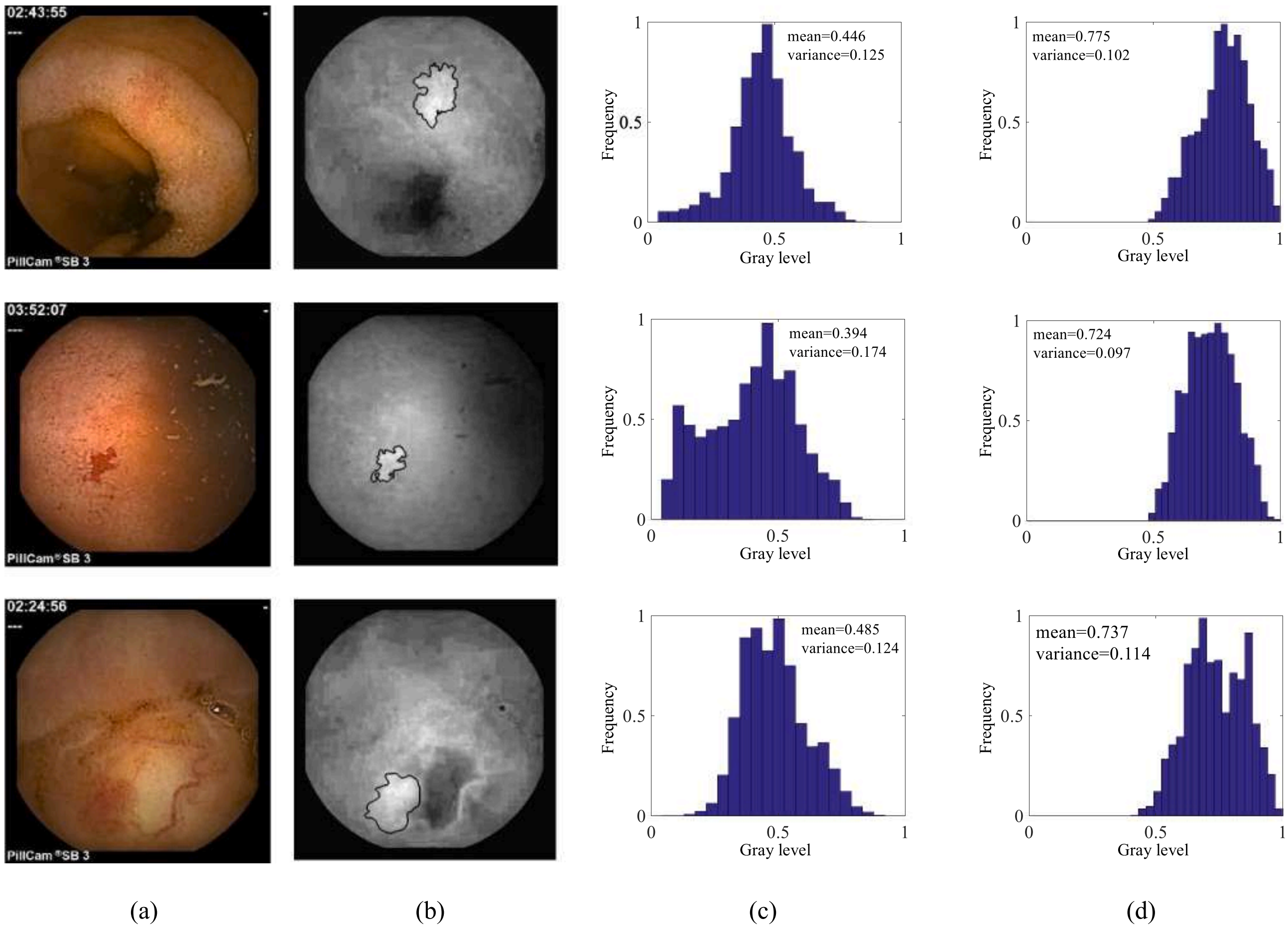


Fig. 5. Normal distributions of pixels in lesion area background (non-lesion area): a) main image, b) 'a' component in CIELab color space where the lesion area is shown using a black-line curves, c) normal distribution of non-lesion area, d) normal distribution of lesion area.

Table 1
Parameters of normal distributions for pixels in lesion and non-lesion areas on 200 images.

	Average of Mean	Average of Variance
Lesion area	0.715	0.104
Non-lesion Area	0.411	0.165

$$\varphi_j = \frac{1}{N}\sum_{i=1}^{N} p(z^{(i)} = j|x^{(i)};\varphi,\mu,\sigma) \tag{3}$$

$$\mu_j = \frac{\sum_{i=1}^{N} p(z^{(i)} = j|x^{(i)};\varphi,\mu,\sigma)x^{(i)}}{\sum_{i=1}^{N} p(z^{(i)} = j|x^{(i)};\varphi,\mu,\sigma)} \tag{4}$$

$$\sigma_j = \frac{\sum_{i=1}^{N} p\left(z^{(i)} = j\,|x^{(i)};\varphi,\mu,\sigma\right)(x^{(i)} - \mu_j)\left(x^{(i)} - \mu_j\right)^T}{\sum_{i=1}^{N} p(z^{(i)} = j\,|x^{(i)};\varphi,\mu,\sigma)} \tag{5}$$

The best values for the parameters can be calculated iteratively by the EM algorithm as follow:

1- Initializing parameters $(\varphi,\mu,\sigma)$by using the Iterative Self Organizing Data Analysis Technique Algorithm (ISODATA) [26]. The ISODATA is an iterative method for multi-threshold selection to segment an image into k classes. The algorithm of the method is shown in Algorithm 1.

2- E-step: The probability of each pixel belonging to a class is calculated as Equation (6).

$$w_j^{(i)} = p(z^{(i)} = j|x^{(i)};\varphi,\mu,\sigma) \tag{6}$$

1. M−step: updating the parameters (Equations (3–5)) with respect to new acquired values of *w(i)*.
2. Go to step 2 if convergence is not achieved.

After convergence of algorithm, each pixel is assigned to the most probable distribution. To find the brightest segment, the mean of pixel values in each distribution is calculated and the segment with the highest mean value is considered as the ROI. Results of applying the introduced method on two normal and abnormal frames are illustrated in Fig. 6. As can be seen, in the second row (normal frames), no suspicious region has been recognized by the physician. However, at least one area is extracted from each image in the introduced method. The ROI is further investigated for the existence of the abnormality rather than the whole image.

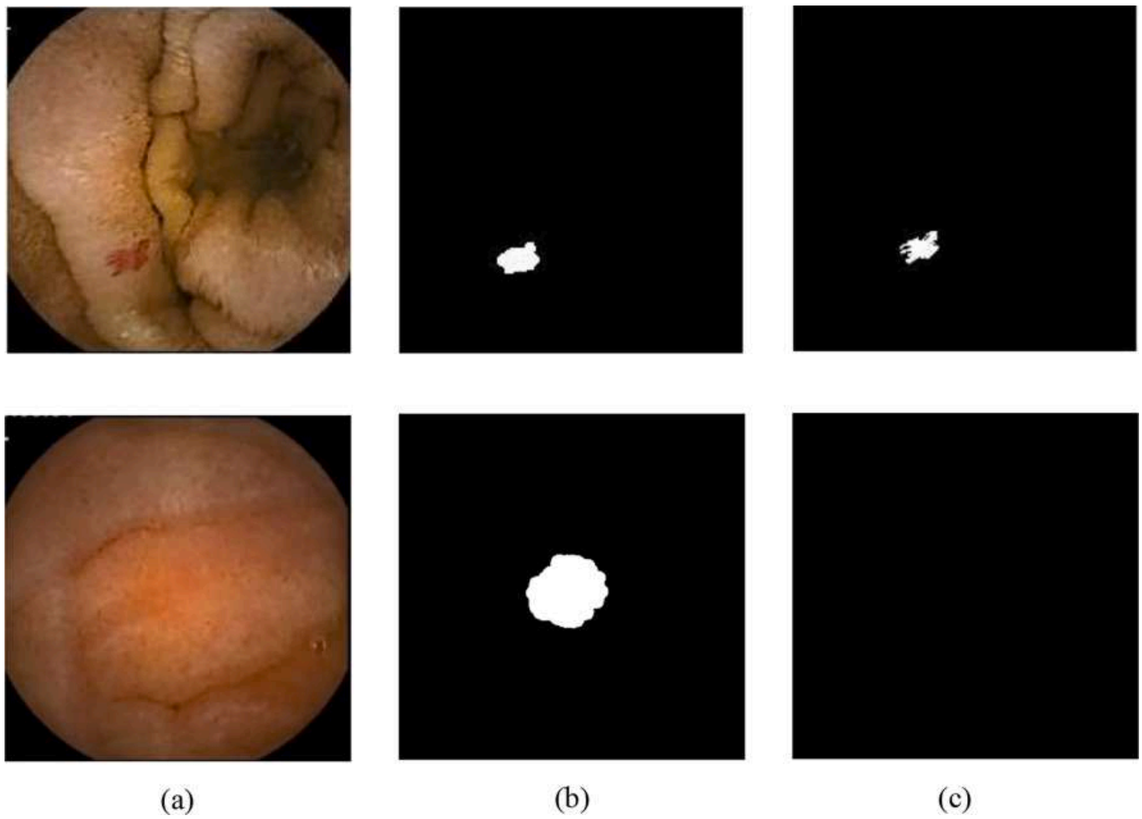


Fig. 6. Segmentation of WCE frames into suspect and non-suspect areas. The columns are respectively: a) the original image, b) the suspicious region using the EM algorithm, c) the lesion selected area from the database. The rows are an abnormal image and a normal image from top to bottom.

Input: image, number of class (c)

1. Calculate the histogram of image:
   h = histogram(image)
2. Initialize $T_0, T_1, \cdots, T_c$ such that $T_0 \leq T_1 \leq \cdots \leq T_c$, T_0 and T_c are the minimum and maximum gray level in the image
3. Calculate the mean values using Equation (7)

$$\mu_k = \frac{\sum_{i=T_{k-1}}^{T_k} i \times h(i)}{\sum_{i=T_{k-1}}^{T_k} h(i)} \quad (7)$$

4. Update thresholds as: $T_k^{new} = \left[\frac{\mu_k + \mu_{k+1} + 1}{2}\right]$
5. Update the mean values using Equation (7)
6. If $T_k^{new} \neq T_k$ go to Step 3.
7. Calculate σandφ:

$$\sigma_k = \frac{\sum_{i=T_{k-1}}^{T_k} h(i) \times (i - \mu_k)^2}{\sum_{i=T_{k-1}}^{T_k} h(i)} \quad (8)$$

$$\varphi_k = \frac{\sum_{i=T_{k-1}}^{T_k} h(i)}{\sum_{i=T_0}^{T_c} h(i)} \quad (9)$$

Algorithm 1- The ISODATA algorithm.

3.2. Abnormality detection

The proposed method for abnormality detection involves three main stages, including preprocessing, feature extraction and classifying the images into one of the two normal and abnormal classes. WCE images often suffer from low contrast. Contrast enhancement can be considered to detect lesions more accurately. Hence, the first stage is assigned to HE for contrast enhancement. The number of bins in the HE algorithm is considered 64. This preprocessing assists the next stage to extract more appropriate features for abnormality detection. The results of applying this technique on two sample WCE frames are shown in Fig. 7.

The second stage is for feature extraction from ROI. It was expected that abnormalities do not have the same effect on various color spaces as shown in Fig. 4. In addition, the worthiness of various channels may not be the same in representing the abnormality. We computed the Normalized Cumulative Histogram (NCH) of a WCE image containing abnormality for the three channels of CIELab color spaces. Then the abnormal region was removed from the image (abnormal region deleted image) and the NCH was computed for all of the three channels again. In Fig. 8, the NCH values for abnormal image and abnormal region deleted image were plotted in two distinct colors. As can be seen, the changes of NCH for channel '*a*' is significantly more obvious compared to the other channels. This result shows that the abnormal region has a greater impact on channel '*a*' in CIELab color space. We performed this experiment on 300 WCE frames containing abnormality randomly selected from the dataset for all channels of the images in RGB, HSV, YCbCr and CIELab color spaces. Fig. 9 shows average of the Mean of Absolute Difference (MAD) of the two NCHs associated with the abnormal image and abnormal region deleted image on all of the 300 images in various channels of the four color spaces. This figure indicates that the four highest effective channels include '*a*' and '*b*' in CIELab, 'Cr' in YCbCr, and 'G' in RGB color space. Hence, we consider these four channels to extract suitable features for discriminating between normal ROIs and those containing abnormality.

We compute eight parameters including the mean, median, variance, maximum, minimum, mode, entropy and contrast from the ROI of the four selected channels. These eight parameters have been employed in several studies for abnormality detection in WCE [8,24]. However, our investigation indicates that merely leaning on this feature set cannot provide an accurate abnormality detection. The problem is that bleeding and AD may appear in various colors other than red, (see Fig. 10). As the figure shows, the color of the lesions, regardless of their color, are distinctive with respect to the background. We propose a novel method in this research which reveals unexpected color changes in ROIs with respect to the background. The following steps are taken to extract the

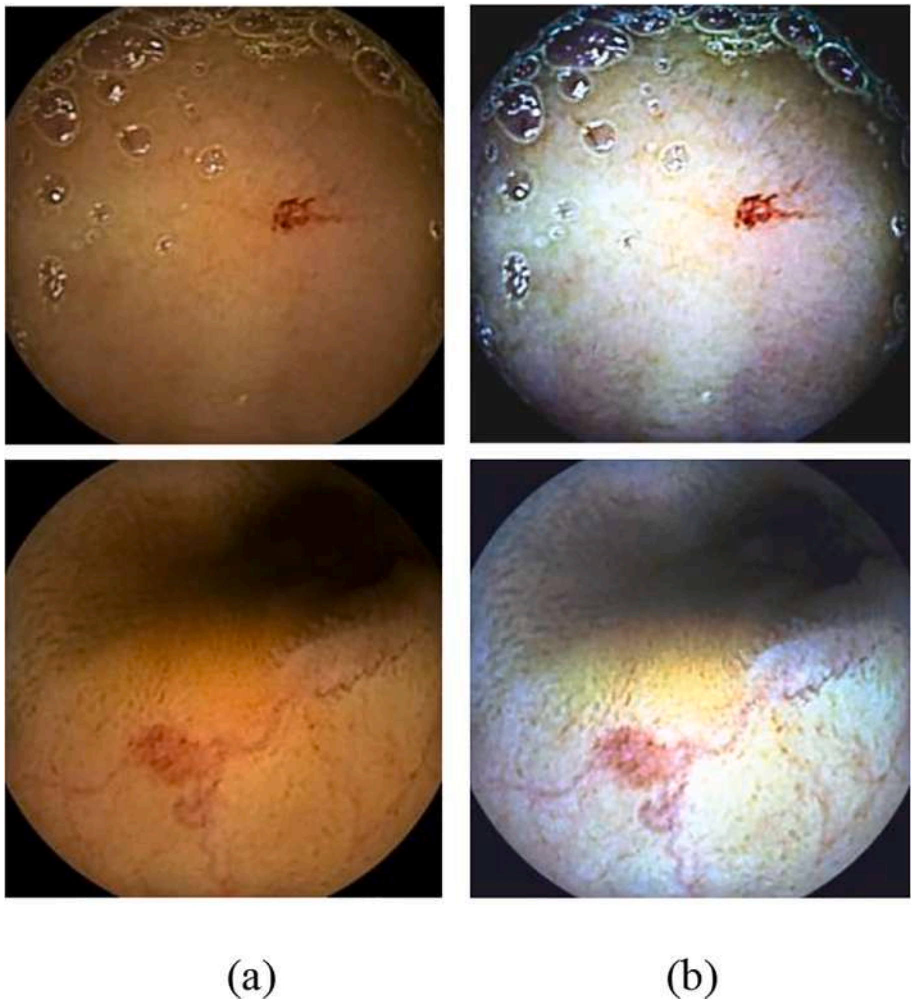


Fig. 7. Two samples of WCE frames: a) original frames, b) the frames after preprocessing using HE technique.

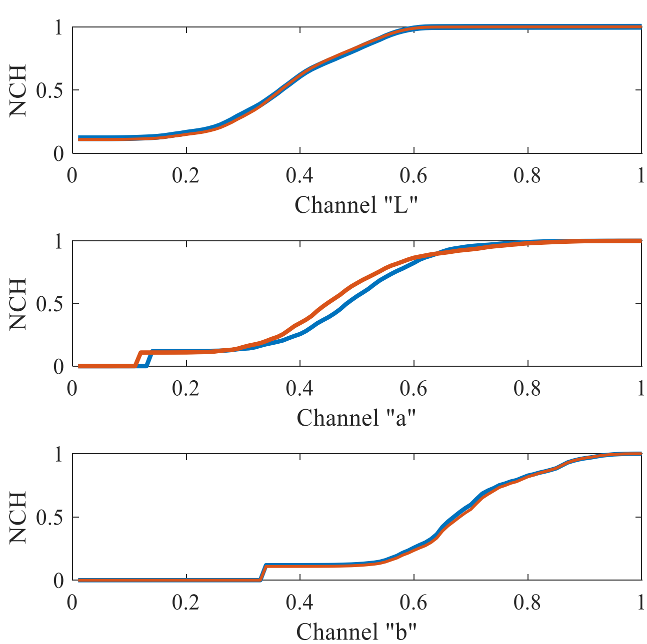


Fig. 8. NCH values for the three channels in CIELab color space for two abnormal image (red graph) and abnormal region deleted image (blue graph).

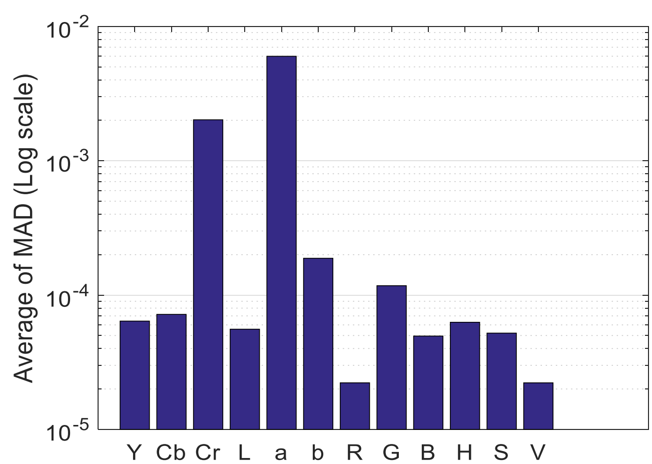
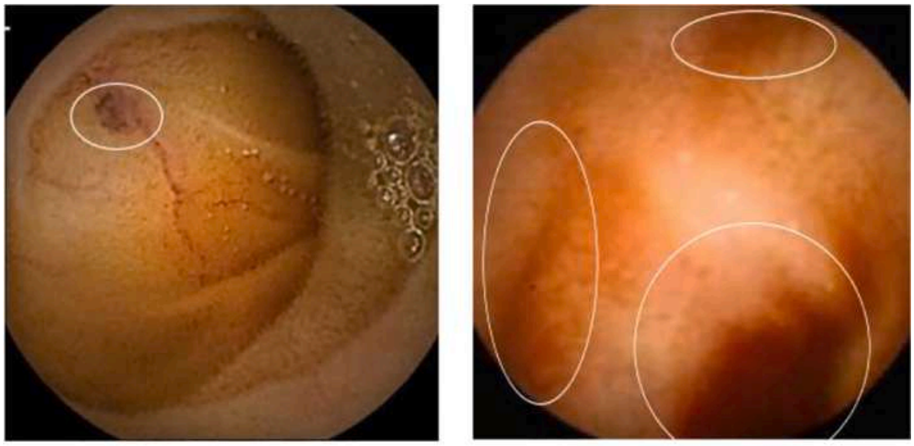


Fig. 9. Average of MAD in all channels of YCbCr, CIELab, RGB, HSV.

second feature set:

1. A small part of the frame containing the ROI, namely ROI+, is considered for further analysis to detect the presence of abnormality. Indeed, the ROI + contains the ROI in a background. To obtain ROI+, a morphological dilation operator with a 20 ×20 square structuring



(a) (b)

Fig. 10. Bleeding (a) and AD (b) lesion in various colors other than red. Lesions areas are covered with white ellipse.

element is applied on the ROI. The size of the structuring element was obtained experimentally.

2. The NCH for the ROI + is computed from channel 'a' in CIELab color space. As stated earlier, the effect of abnormality is highly reflected in this color space. Our investigations indicate that a sudden change in the NCH represents the existence of abnormality in ROI+ (see Fig. 11).
3. To detect sudden changes in the histogram, the NCH values are quantized in 10 levels. The number of NCH values in each level is considered as the second feature set.

In general, two sets of features are used in the feature extraction stage. The length of the combined feature set is 42, in which the first and the second feature set each contain 32 and 10 features respectively. The feature set is fed to an MLP neural network for classifying the ROIs into normal and abnormal classes. Structure of the MLP consists of input layers with 42 neurons, a hidden layer with 22 neurons, and an output layer with two neurons. A sigmoid function is used as the activation function. The learning rate and momentum rate are set to 0.2 and 0.1, respectively.

4. Experimental results and discussion

In this section, at first, the datasets used in this study are introduced. Then the evaluation criteria and results are discussed.

4.1. Datasets

The image sets used were obtained from two datasets, the first of which was provided by the Gastrointestinal Image Analysis Challenge 2017 [27]. This dataset consists of 600 normal and 600 abnormal images with AD lesion. There exists one binary ground truth mask for each image as shown in Fig. 12-a. The size of these images is 704 × 704 pixels. The second database consists of 3895 images in 320 × 320 pixels, containing normal frames and those with annotated bleeding [28]. See Table 2 for more details about the second dataset. From this dataset, 600 images containing bleeding from set 1 were randomly selected for abnormality detection in WCE frames. Fig. 12-b shows one bleeding image from this dataset with the related binary ground truth mask. In the proposed method we resize all images into 512 × 512 pixels.

4.2. Evaluation criteria

Several performance criteria which are widely used in medical image classification and image segmentation techniques are employed in this study to evaluate the proposed method. These evaluation criteria include: Accuracy (AC), False Positive Rate (FPR), Precision, Recall (Sensitivity), Specificity, F-Measure, Matthews Correlation Coefficient (MCC) and Dice Score (DS). The formulas of these criteria are based on four basis metrics which are stated in Table 3.

$$AC = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$FPR = \frac{FP}{FP + TN} \tag{11}$$

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

$$Recall = Sensivisity = \frac{TP}{TP + FN} = 1 - FNR \tag{13}$$

$$Specificity = \frac{TN}{TN + FP} = 1 - FPR \tag{14}$$

$$F - Measure = 2\frac{Precision^{*}Recall}{Precision + Recall} \tag{15}$$

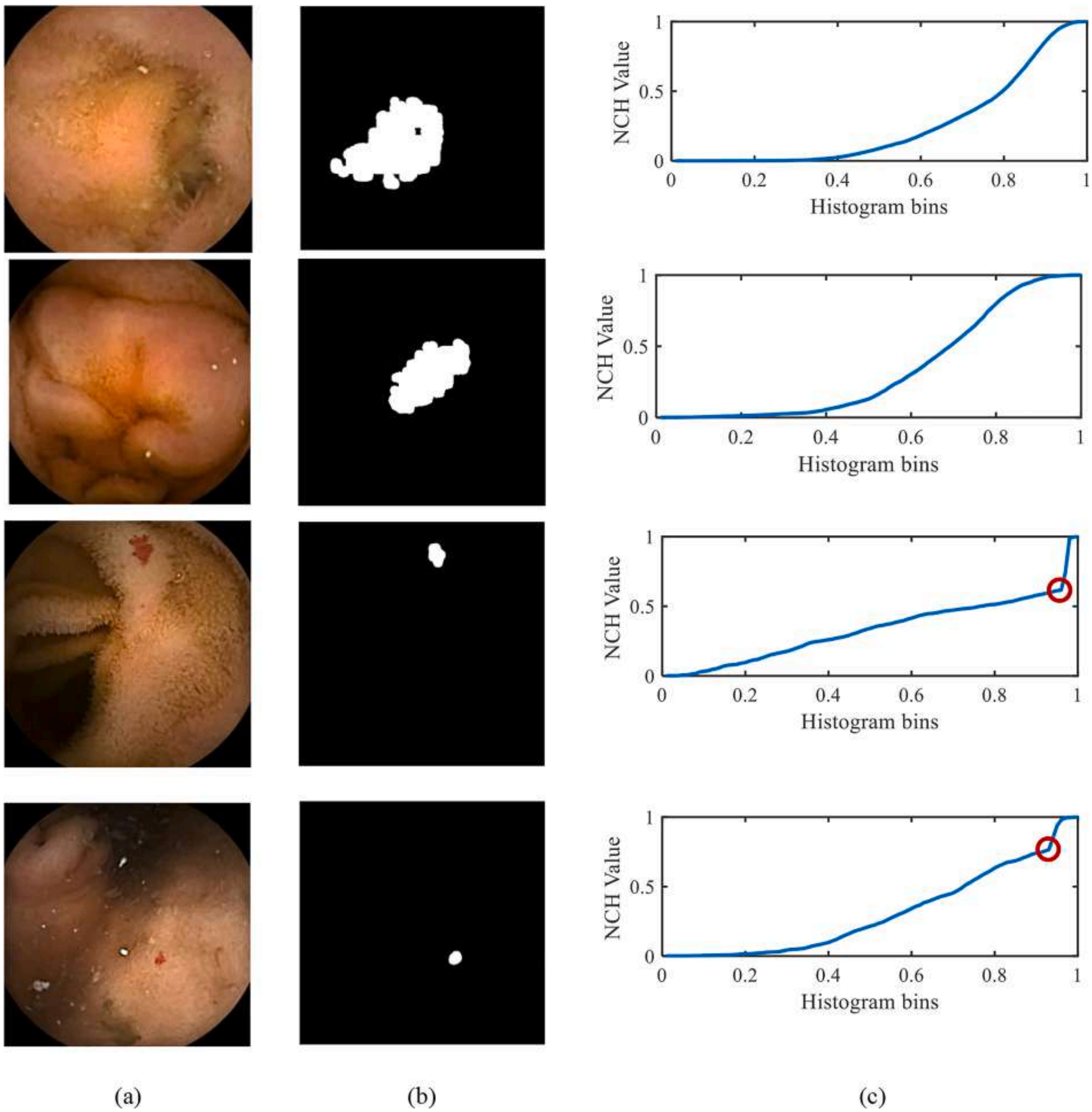


Fig. 11. Comparing sudden change in NCH for normal (first two rows) and abnormal frames (last two rows). The columns are: a) the main image, b) ROI, c) NCH curve which the sudden changes were indicated by a red circle.

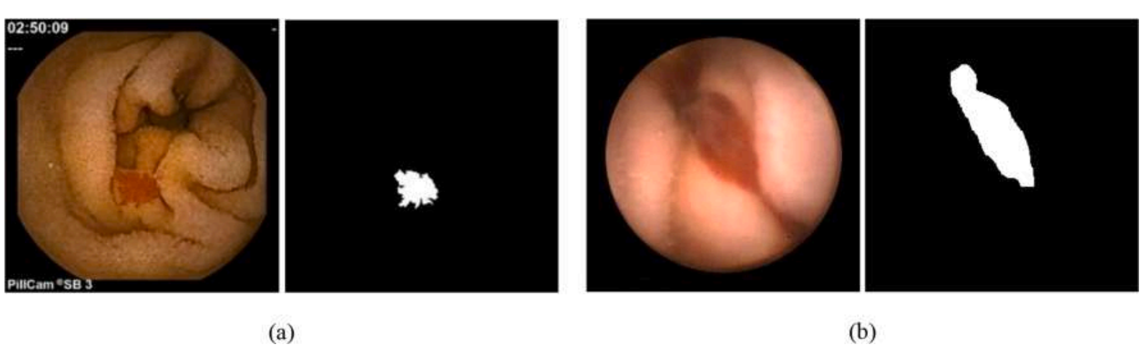


Fig. 12. Two image samples from datasets: a) AD lesion and its corresponding binary ground truth masks from the first database, b) bleeding lesion with its binary ground truth masks from the second database.

Table 2
Image sets in Dataset 2.

	Set 1	Set 2	Total
With lesion	1131	439	1570
Normal	2164	161	2325
Total	3295	600	3895

Table 3
Description of several parameters of evaluation criteria.

	Abbr.	Description
True Positive	TP	The number of correctly classified images containing abnormality
False Positive	FP	The number of incorrectly classified images containing abnormality
True Negative	TN	The number of correctly classified images without abnormality
False Negative	FN	The number of incorrectly classified images without abnormality

$$MCC = \frac{TP^*TN - FP^*FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (16)$$

$$DS = \frac{2TP}{2TP + FP + FN} \quad (17)$$

Note that the classification is more accurate when accuracy, precision, recall, Specificity, F-Measure and MCC are closer to one while FPR is closer to zero. Also, when DS is closer to one, segmentation is more accurate. The values of these evaluation criteria are a non-negative number below or equal to one.

The proposed method is also tested on a computer with an Intel(R) Core(TM) i5-7400 CPU @ 3.00 GHz and 8 GB of memory. The Test Time (TT) of the proposed method to process one image is considered as the time complexity evaluation metric.

4.3. Results

In this section, at first, the proper value for k (number of different Gaussian distributions) in the ROI selection algorithm is chosen by an experiment. Then, the results of WCE image segmentation for ROI selection using k-means and EM are compared with each other. Finally, several experiments are performed to evaluate the abnormality detection method.

In ROI selection, it was shown that image pixels can be considered as

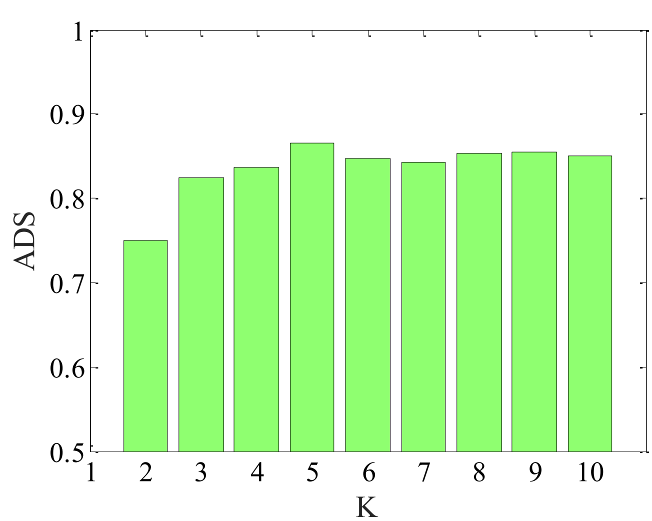


Fig. 13. Average values of ADS for different number of classes (k).

a combination of k different normal distributions. Choosing a proper value for k will greatly affect the ROI selection result. A greater value of k causes the lesion area to be divided into several segments, and as a result, only a subset of the lesion will participate in ROI. With a smaller value, some parts of the background will be present in the ROI, and extracted features will not distinguish the lesion with appropriate accuracy. For selecting the best value for k, 100 bleeding or AD images were randomly selected from the database. Then the introduced segmentation method was performed on each image for different values of k ranging from 2 to 10, and DS was calculated in each test. Finally, the Average of DS (ADS) was calculated for each k on all images. The ADS values for various k can be seen in Fig. 13. As can be seen, the best ADS value for k is five. Hence, in all subsequent experiments k is set to five.

In order to compare the k-means and EM based image segmentations in ROI selection, we randomly selected 100 bleeding or AD images and the two methods were applied on component 'a' of images in the CIELab color space with k equal to five. In k-means algorithm, at first, five pixels are selected as centers, randomly, then each pixel is assigned to the nearest center. In the next step, the mean of each pixel group is considered as a new center. The two previous steps are repeated until convergence. Finally, the pixel groups with the highest average value are considered as the lesion segment (ROI). DS (Eq. (17)) for each image and ADS over all images are calculated for the two methods. The values of ADS for EM and k-means methods are 86.67% and 45.97%, respectively. Therefore, EM based image segmentation has a better performance for ROI selection.

In order to evaluate the proposed abnormality detection method. First, the influence of each feature set on the classification task was examined. Then, different classifiers were investigated to choose the most appropriate one for this problem. Then the time complexity of the proposed method is investigated. ROI extraction is an effective step in the performance of the proposed method, hence in the next experiment, we examine the effectiveness of the ROI extraction step on the performance of the proposed method. Finally, we compared our method with other existing methods.

All tests are executed with 10-fold cross-validation. The k-fold cross-validation is a very effective method to prevent over-learning or under-learning. The number of images in our dataset is also limited. Therefore by dividing the dataset into training, test and validation sets, the number of images that can be used to learn the model will be very limited. The k-fold cross-validation is also an efficient method to learn and evaluate the models with a limited data samples. In this method, the samples are divided into k groups. In the k iterations, the model is trained and validated. In each iteration, one set is considered for validation and the rest for training. The performance of the model in each iteration is calculated on the validation set. Each set is considered as the validation set only once. Finally, the average of all evaluation criteria in k iterations are reported by k-fold cross-validation [29].

To compare the impact of different feature sets, the frames were classified using the first, the second and both feature sets. The results are shown in Table 4. As can be seen, using both features yields the best results in all metrics except FPR. Using both feature sets can considerably improve the Recall criterion which is more important than FPR in medical image diagnosis. Higher value of Recall indicates lower value of FNR which means there is an increased probability of diagnosing abnormal images.

For selecting the best classifier, all features (containing the first and the second feature sets) are fed to different classifiers. The results are reported in

Table 5 and the superiority of the MLP can be found among all the examined methods. The results also show that the proposed method is able to detect abnormality in WCE.

Time complexity of the proposed method for five different scenarios is reported in Table 6. In the first scenario, the size of image is 512×512 and in each next scenario the image is resized by half. To calculate the TT, 400 images (200 normal and 200 abnormal) are randomly selected

Table 4
Comparing the impact of each feature sets in classification.

Features	Accuracy	FPR	Precision	Sensitivity	F-Measure	MCC
The first features set (statistical features)	0.937	**0.049**	0.902	0.910	0.906	0.859
The second features set (sudden change in NCH)	0.787	0.422	0.809	0.892	0.848	0.502
Both (the first and the second feature sets)	**0.949**	0.070	**0.965**	**0.959**	**0.962**	**0.887**

Table 5
The evaluation Result of the proposed method.

Method	Accuracy	FPR	Precision	Sensitivity	Specificity	F-Measure	MCC
MLP	**0.949**	**0.070**	**0.965**	**0.959**	**0.930**	**0.962**	**0.887**
Random Tree [34]	0.900	0.158	0.921	0.929	0.842	0.925	0.774
Random Forest [35]	0.918	0.140	0.931	0.948	0.860	0.939	0.815
Decision Table [36]	0.912	0.162	0.922	0.949	0.838	0.935	0.800
Support Vector Machine [37]	0.850	0.292	0.863	0.921	0.708	0.891	0.654

Table 6
Time complexity and accuracy of proposed method in different size of input images.

Metric	512×512	256×256	128×128	64×64	32×32
ATT (second)	1.407	0.344	0.120	0.057	0.037
Accuracy	0.945	0.940	0.900	0.830	0.655

Table 8
Comparing the proposed method with the Caroppo et al. [30].

Method	Accuracy	Sensitivity	Specificity	precision	F-measure
Our proposed method	**0.988**	**0.988**	0.952	**0.975**	**0.982**
Caroppo *et al.* method [30]	0.957	0.958	**0.953**	0.970	0.964

Table 7
Effectiveness of ROI extraction.

Method	Accuracy	FPR	Precision	Sensitivity	Specificity	F-Measure	MCC
Excluding ROI extraction	0.798	0.133	0.751	0.798	0.867	0.774	0.656
Our proposed method	**0.949**	**0.070**	**0.965**	**0.959**	**0.930**	**0.962**	**0.887**

from the dataset. The average testing time of these image is reported as ATT (average Testing Time). For each experiment we also calculate the accuracy of method to find out the minimum size of the images so that the method is still able to accurately detect the lesions. As can be seen, by reducing the size of the images, the processing time is significantly reduced. In fact there is a trade-off between the accuracy of the method and the time complexity. But for acceptable performance of the proposed method, images with a minimum size of 128×128 are required.

In the next experiment, the effectiveness of the ROI extraction step is investigated. In the proposed method, ROIs are extracted by using the EM segmentation method. The EM segmentation method is an iterative and time-consuming algorithm. Therefore, it needs to show that it has a positive impact on the performance of our proposed method. To prove this claim, in an experiment we removed the ROI extraction step and the features were extracted from the whole part of the image. The results were reported in Table 7. As can be seen by omitting this step, the performance of the method in all criteria is significantly reduced.

We compared our method with the method of Caroppo *et al.* [30]. In this method, a deep transfer learning method was introduced for bleeding detection method. In their method, features were extracted and combined from three popular convolutional neural network models (VGG19, InceptionV3 and ResNet50). They extracted features and classified them using a supervised machine learning method. These methods were evaluated on all images in Dataset 2. So, we also evaluated our proposed method on the whole images in Dataset 2 and compared it with Caroppo *et al.* method in Table 8. As can be seen, our proposed method shows better performance in all metrics except specificity. It is clear that in the detection of abnormal images in medical system, missing an abnormal image (i.e. FN) is much more undesirable than recognizing normal images as the abnormal (i.e. FP). Therefore, the sensitivity is more important than the specificity in this research area. In fact, by increasing the sensitivity, the rate of the FNR decreases.

We compare the proposed method with the method of Hajabdollahi *et al.* [31]. In this bleeding detection method, at first, the informative channels in different color spaces are recognized. The detection is based on a simplified version of the convolutional neural network (CNN). Structure of the CNN was simplified by using simultaneous quantization and pruning methods. The results of this paper were reported based on the combination of two publicly available datasets [32,33] which contains 55 bleeding and 728 normal frames; five bleeding and 728 normal frames from KID dataset 1 [32] and 50 bleeding frames from Deeba's dataset [33]. We also used this combined dataset to compare our proposed method with this method. The result of comparing our proposed method with this method is shown in Table 9. The results show the superiority of our method compared to this method in all evaluation criteria.

Finally, our work was compared with Yuan's method [17] and also with our previous method in [5]. We implemented Yuan's method and evaluated these two methods using our dataset. The results reported in Table 10. It shows that our method outperforms Yuan's method and our previous work in all confiders evaluation criteria. As can be seen, Yuan's method does not achieve acceptable results on our dataset, yet our previous work yields better results than Yuan's method.

A potential improvement to our method in the future might include modifying the EM algorithm to offer more precise ROI, which would improve the overall accuracy of the proposed method.

5. Conclusions

In this paper, a novel abnormality detection method in capsule endoscopy images was proposed. In the proposed method, at first a potential ROI corresponding to lesion candidates is selected in each

Table 9
Comparing the proposed method with the method of Hajabdollahi et al. [31] on the dataset provided in [31].

Method	Accuracy	FPR	Precision	Sensitivity	Specificity	F-Measure	MCC
Proposed method	**0.983**	**0.182**	**0.984**	**1.000**	**0.818**	**0.991**	**0.890**
Hajabdollahi's *et al.* Method [31]	0.9527	0.402	0.950	1.000	0.598	0.980	0.753

Table 10
Comparing the proposed method with the Yuan's method [17].

Method	Accuracy	FPR	Precision	Sensitivity	Specificity	F-Measure	MCC
Proposed method	**0.949**	**0.070**	**0.965**	**0.959**	**0.930**	**0.962**	**0.887**
Yuan's method [17]	0.690	0.477	0.765	0.774	0.523	0.769	0.299
Our previous method [5]	0.895	0.115	0.940	0.901	0.885	0.920	0.771

frame. Then features are extracted from the ROI using a combination of color histogram analysis and statistical features. Our investigation showed that the characteristics of lesions in one channel may be more obvious than other channels in some color spaces. Hence, the first feature set was extracted from the four highest effective channels. We found that recognizing lesions in WCE which appear in various colors, merely by investigating the color specifications, yields a challenging issue, yet investigating the lesion's color as a distinctive color with respect to the background is reflected in the normalized cumulative histogram which is another feature set to characterize the lesions. Finally, the two feature sets were employed to classify the WCE frames into normal and abnormal categories by using MLP. The method was tested on a large dataset with 1800 images. The obtained results show the efficiency of the proposed method.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Zahra Amiri: Methodology, Software, Visualization, Validation, Writing – original draft. **Hamid Hassanpour:** Conceptualization, Methodology, Investigation, Writing – review & editing, Supervision. **Azeddine Beghdadi:** Data curation, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. Shahril, S. Baharun, A. Islam, Pre-processing technique for wireless capsule endoscopy image enhancement, Int. J.of Electric. Comput. Eng. 6 (2016) 1617–1626, https://doi.org/10.11591/ijece.v6i4.
[2] Z. Li, D. Carter, R. Eliakim, W. Zou, H. Wu, Z. Liao, Z. Gong, J. Wang, J.W. Chung, S.Y. Song, The Current Main Types of Capsule Endoscopy. Handbook of Capsule Endoscopy, Springer, Dordrecht, 2014.
[3] H. Chen, X. Wu, G. Tao, Q. Peng, Automatic content understanding with cascaded spatial–temporal deep framework for capsule endoscopy videos, Neurocomputing 229 (2017) 77–87, https://doi.org/10.1016/j.neucom.2016.06.077.
[4] S. Lecleire, I. Iwanicki-Caron, A. Di-Fiore, C. Elie, R. Alhameedi, S. Ramirez, M. Antonietti, Yield and impact of emergency capsule enteroscopy, severe obscure-overt gastrointestinal bleeding, Endoscopy 44 (04) (2012) 337–342, https://doi.org/10.1055/s-0031-1291614.
[5] Z. Amiri, H. Hassanpour, A. Beghdadi, A computer-aided method to detect bleeding frames in capsule endoscopy images, in: 8th European Workshop on Visual Information Processing, 2019, pp. 217–221, https://doi.org/10.1109/EUVIP47703.2019.8946168.
[6] M. Usman, G. Satrya, M. Usman, S. Shin, Detection of small colon bleeding in wireless capsule endoscopy videos, Comput. Med. Imaging Graph. 9 (5) (2016) 16–26, https://doi.org/10.1016/j.compmedimag.2016.09.005.
[7] D. Iakovidis, A. Koulaouzidis, Automatic lesion detection in capsule endoscopy based on color saliency: closer to an essential adjunct for reviewing software, Gastrointest Endoscopy 80 (5) (2014) 877–883, https://doi.org/10.1016/j.gie.2014.06.026.
[8] F. Noya, M.A. Alvarez-Gonzalez, R. Benitez, Automated angiodysplasia detection from wireless capsule endoscopy, in: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2017, pp. 3158–3161, https://doi.org/10.1109/EMBC.2017.8037527.
[9] B. Li, M.Q.H. Meng, Texture analysis for uulcer detection in capsule endoscopy images, Image Vision Comput. 27 (9) (2009) 1336–1342, https://doi.org/10.1016/j.imavis.2008.12.003.
[10] Y. Fu, W. Zhang, M. Mandal, M.-H. Meng, Computer-aided bleeding detection in WCE video, IEEE J. Biomed. Health. Inf. 18 (2) (2013) 636–642, https://doi.org/10.1109/JBHI.2013.2257819.
[11] G. Pan, G. Yan, X. Qiu, J. Cui, Bleeding detection in wireless capsule endoscopy based on probabilistic neural network, J. Medical Syst. 35 (6) (2011) 1477–1484, https://doi.org/10.1007/s10916-009-9424-0.
[12] B. Li, M. Meng, Computer–aided detection of bleeding regions for capsule endoscopy images, IEEE Trans. Biomed. Eng. 56 (4) (2009) 1032–1039, https://doi.org/10.1109/TBME.2008.2010526.
[13] F. Deeba, S.K. Mohammed, F.M. Bui, K.A. Wahid, Unsupervised abnormality detection using saliency and retinex based color enhancement, in: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Socity, 2016, https://doi.org/10.1109/EMBC.2016.7591573.
[14] G. Pan, G. Yan, X. Song, X. Qiu, BP neural network classification for bleeding detection in wireless capsule endoscopy, J. Medical Eng. Technol. 33 (7) (2009), https://doi.org/10.1080/03091900903111974.
[15] P. Szczypiński, A. Klepaczkoa, M. Pazurekb, P. Danielb, Texture and Color based Image Segmentation and Pathology Detection in Capsule Endoscopy Videos, Computer Methods and Programs in Biomedicine 113 (1) 396–411. doi:10.1016/j.cmpb.2012.09.004.
[16] G. Lv, G. Yan, Z. Wang, Bleeding detection in wireless capsule endoscopy images based on color invariants and spatial pyramids using support vectormachines, IEEE Eng. Med. Biol. Soc. (2011), https://doi.org/10.1109/iembs.2011.6091638.
[17] Y. Yuan, Li, Q. Meng, Bleeding frame and region detection in the wireless capsule endoscopy video, IEEE J. Biomed. Health Inf. 20 (2) (2015) 624–630, https://doi.org/10.1109/JBHI.2015.2399502.
[18] D. Iakovidis, A. Koulaouzidis, Software for enhanced video capsule endoscopy: challenges for essential progress, Nat. Rev. Gastroenterol. Hepatol. 12 (3) (2015) 172–186, https://doi.org/10.1038/nrgastro.2015.13.
[19] X. Jia, Q.-H.M. Meng, A study on automated segmentation of blood regionsin wireless capsule endoscopy images using fully convolutional networks, in: IEEE 14th International Symposium on Biomedical Imaging, 2017, https://doi.org/10.1109/ISBI.2017.7950496.
[20] M. Hajabdollahi, R. Esfandiarpoor, E. Sabeti, N. Karimi, S.R. Soroushmehr, S. Samavi, Multiple abnormality detection for automatic medical image diagnosis using bifurcated convolutional neural network, Biomed. Signal Process. Control 57 (2020), 101792, https://doi.org/10.1016/j.bspc.2019.101792.
[21] D.E. Diamantis, D.K. Iakovidis, A. Koulaouzidis, Look-behind fully convolutional neural network for computer-aided endoscopy, Biomed. Signal Process. Control 49 (2019) 192–201, https://doi.org/10.1016/j.bspc.2018.12.005.
[22] P.M. Vieira, B. Gonçalves, C.R. Gonçalves, C.S. Lima, Segmentation of Angiodysplasia Lesions in WCE Images using a MAP Approach with Markov Random Fields, in: 38th Annual International Conference of the Engineering in Medicine and Biology Society Orlando, FL, USA, 2016, https://doi.org/10.1109/embc.2016.7590916.
[23] M. Salehi, J. Razmara, S. Lotfi, Development of an ensemble multi-stage machine for prediction of breast cancer survivability, J. AI Data Mining 8 (3) (2020) 371–378, https://doi.org/10.22044/jadm.2020.8406.1978.
[24] F. Deeba, M. Islam, F.M. Bui, K.A. Wahid, Performance assessment of a bleeding detection algorithm for endoscopic video based on classifier fusion method and exhaustive feature selection, Biomed. Signal Process. Control 40 (2018) 415–424, https://doi.org/10.1016/j.bspc.2017.10.011.

[25] F. Rahman, B. Zarpak, Image segmentation using Gaussian mixture model, IUST Int. J. Eng. Sci. 19 (2008) 29–32.
[26] F.R.D. Velasco, Thresholding using the ISOData Clustring Algorithm, IEEE Trans. Syst., Man, Cybernet. 10 (11) (1980) 771–774, https://doi.org/10.1109/TSMC.1980.4308400.
[27] Gastrointestinal image analysis. https://endovissub2017-giana.grand-challenge.org/ (Accessed 10 July 2021).
[28] P. Coelho, A. Pereira, M. Salgado, A. Cunha, Deep learning approach for red lesions detection in video capsule endoscopies, in: International Conference Image Analysis and Recognition, 2018, pp. 553–561, https://doi.org/10.1007/978-3-319-93000-8_63.
[29] Ran Li, Shihui Xu, Sibo Li, Yongqin Zhou, Kai Zhou, Xianzhong Liu, Jie Yao, State of charge prediction algorithm of lithium-ion battery based on PSO-SVR cross validation, IEEE Access 8 (2020) 10234–10242, https://doi.org/10.1109/ACCESS.2020.2964852.
[30] A. Caroppo, A. Leone, P. Siciliano, Deep transfer learning approaches for bleeding detection in endoscopy images, Comput. Med. Imaging Graph. 88 (2021), 101852, https://doi.org/10.1016/j.compmedimag.2020.101852.
[31] M. Hajabdollahi, R. Esfandiarpoor, P. Khadivi, S.R. Soroushmehr, N. Karimi, K. Najarian, S. Samavi, Segmentation of bleeding regions in wireless capsule endoscopy for detection of informative frames, Biomed. Signal Process. Control 53 (2019), 101565, https://doi.org/10.1016/j.bspc.2019.101565.
[32] A. Koulaouzidis, D.K. Iakovidis, D.E. Yung, E. Rondonotti, U. Kopylov, J.N. Plevris, G.E. Tontini, KID Project: an internet-based digital video atlas of capsule endoscopy for research purposes, Endoscopy Int. Open 5 (06) (2017) 477–483, https://doi.org/10.1055/s-0043-105488.
[33] F. Deeba, Bleeding images and corresponding ground truth of CE images, https://sites.google.com/site/farahdeeba073/Research/resources (Accessed 11 September 2021).
[34] D. Aldous, the Continuum Random Tree. II, the Annals of Probability (1991) 1-28.
[35] A. Liaw, M. Wiener, Classification and regression by random Forest, R News 2 (3) (2002) 18–22.
[36] B.J. Cragun, H.J. Steudel, A decision-table-based processor for checking completeness and consistency in rule-based expert systems, Int. J. Man Mach. Stud. 26 (5) (1987) 633–648, https://doi.org/10.1016/S0020-7373(87)80076-7.
[37] L. Wang, Support vector machines: theory and applications, Stud. Fuzziness Soft Comput. Springer Sci. Business Media 177 (2005).