Computer Methods and Programs in Biomedicine 157 (2018) 39–47

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

Computer assisted gastric abnormalities detection using hybrid texture descriptors for chromoendoscopy images

Hussam Ali[a,*], Mussarat Yasmin[a], Muhammad Sharif[a], Mubashir Husain Rehmani[b]

[a] COMSATS Institute of Information Technology Wah, Pakistan
[b] Telecommunications Software and Systems Group (TSSG) Waterford Institute of Technology (WIT), Ireland

ARTICLE INFO

Article history:
Received 17 June 2017
Revised 25 November 2017
Accepted 10 January 2018

Keywords:
Chromoendoscopy
Co-occurrence matrix
Filter bank
Texture analysis
Gabor filter
Stomach cancer

ABSTRACT

Background and Objective: The early diagnosis of stomach cancer can be performed by using a proper screening procedure. Chromoendoscopy (CH) is an image-enhanced video endoscopy technique, which is used for inspection of the gastrointestinal-tract by spraying dyes to highlight the gastric mucosal structures. An endoscopy session can end up with generating a large number of video frames. Therefore, inspection of every individual endoscopic-frame is an exhaustive task for the medical experts. In contrast with manual inspection, the automated analysis of gastroenterology images using computer vision based techniques can provide assistance to endoscopist, by finding out abnormal frames from the whole endoscopic sequence.

Methods: In this paper, we have presented a new feature extraction method named as Gabor-based gray-level co-occurrence matrix (G2LCM) for computer-aided detection of CH abnormal frames. It is a hybrid texture extraction approach which extracts a combination both local and global texture descriptors. Moreover, texture information of a CH image is represented by computing the gray level co-occurrence matrix of Gabor filters responses. Furthermore, the second-order statistics of these co-occurrence matrices are computed to represent images' texture.

Results: The obtained results show the possibility to correctly classifying abnormal from normal frames, with sensitivity, specificity, accuracy, and area under the curve as 91%, 82%, 87% and 0.91 respectively, by using a support vector machine classifier and G2LCM texture features.

Conclusion: It is apparent from results that the proposed system can be used for providing aid to the gastroenterologist in the screening of the gastric tract. Ultimately, the time taken by an endoscopic procedure will be sufficiently reduced.

1. Introduction

Gastric cancer is a primary cause of cancer-related deaths around the world [1]. A normal screening strategy for detecting gastric malignancies involves visual inspection of tissues. The early detection of gastric malignancies can help in treating cancer timely. Several studies confirm the benefits of diagnosing cancer at its early stages [2–5].

In normal practices, gastroenterologist uses an endoscope to inspect a human's inner cavity. An endoscope is equipped with a camera and light source mounted on its distal tip [6]. There are several advancements have been made in the endoscopy technology to assist the gastroenterologist in the detection of abnormal regions [7].

Chromoendoscopy (CH) is an image-enhanced endoscopy technique; Traditionally, in chromoendoscopy, surface of mucosal wall is enhanced by spraying methylene blue. Currently, band-pass filters and image processing algorithms are used to render the effects of a dye-based CH [8]. CH makes the gastric mucosal surface of a patient more prominent thus, provides help to medical analyst in visualization of irregular patterns.

Consequently, the inspection of the gastric tract for detection of abnormalities is a critical and time-consuming task for an endoscopist. Specifically, in a case when there are too many gastric patients for screening and false detection rate increased when the endoscopist became tired [9].

* Corresponding author.
E-mail address: hussam@ciitwah.edu.pk (H. Ali).

https://doi.org/10.1016/j.cmpb.2018.01.013
0169-2607/

Sometimes cancer in its early stages is hard to detect because there are a few abnormal frames in the whole video sequence which can be easily unnoticed by an endoscopist [10]. Computer-aided analysis of endoscopy videos is still an emerging field in the medical imaging and in its early stages of development. The automated detection of abnormal frames from an endoscopic session could offers the possibility to deal with this dilemma (time taken by an endoscopic procedure). Then gastroenterologist has to analyze a few frames after abnormal frames are classified by automated system. Moreover, it will also able to provide a second opinion to medical experts about the diagnosis. The gastric environment poses many challenges to the automatic detection of tumors from gastric frames.

The dynamics of image acquisition such as uncontrolled camera movements, poor focus of the camera, illumination variations, body positions, and poor cleansing of gastrointestinal are the main reasons for miss-classification. However, even in the ideal imaging scenarios, it is hard to detect tumor regions due to lack of color space and a specific geometric structure of lesions. Therefore, there is a need for such computer aided diagnosis (CAD) systems which can cope with these issues [11].

Previously, several methods have been developed to detect gastric cancer by employing the vision techniques. These methods include various descriptors to represent characteristics of endoscopic images further learning models are trained using these features. Different types of features can be used for classification and extraction methods have been developed for gastric images. Colors are basic descriptors to represent the endoscopic frames, within different color spaces images can be represented with different color characteristics. RGB is a basic color space [9] and colors are good descriptors for detection of bleeding and ulcer. Other color features were described in color space HSV [12] and CIELAB [13].

However, for detection of cancerous regions using only color features is not sufficient due to certain limitations (e.g., illumination variation). Therefore, the texture is a good choice for representing the mucosal structures in endoscopic frames. Numerous methods of texture analysis have been reported in literature [14–16]. The texture of an image delivers the information about the spatial arrangement of frequencies. Local binary patterns (LBP) [17], gray level co-occurrence matrix (GLCM) [18], and wavelet-based features are employed in various systems which are targeted to the analysis of gastric abnormalities.

Homogeneous texture features are extracted using an array of Gabor filters which explicitly tuned to different scales, frequencies, and orientations. Further the mean and standard deviation of responses of these filters are used to represent the texture of images [19]. Similarly, the gray level co-occurrence matrix of an image is calculated by computing the occurrence of same intensities distributed on a specified distance. Moreover, the second order statistics of these matrices are calculated to represent the textual information of images [20]. In this paper, we have proposed an automated method for classifying the normal and abnormal CH images. We have used a hybrid feature extraction method to extract texture information from CH frames. Moreover, the suggested method is based on an incorporation of Gabor filters and gray level co-occurrence matrices (GLCM). A Gabor filter bank with various filters as it is represented in Fig. 1 was created. Then, GLCM was computed from the Gabor filters' responses, instead of extracting the micro-statistics (mean and standard deviation) of these responses. Furthermore, the texture of an image has represented a second order statistics calculated from GLCM of Gabor filters' responses. The performance of the proposed method is tested against the traditional feature extraction methods on the classification of CH images. Moreover, several classification models are trained and tested on proposed features.

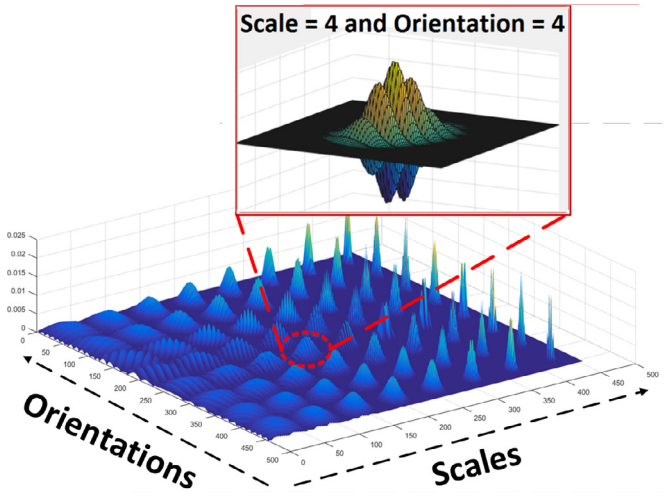

Fig. 1. A filter bank obtained by selecting different orientation (k) and scale (s) of Gabor filter to get multi-resolution responses of CH images.

1.1. Contributions

As mentioned earlier, the texture is important visual cues for analysis. Therefore, texture features are our main focus in this article. The main contributions of this paper are summarized as follows:

- A new method G2LCM for texture feature extraction has proposed for classification of CH images. Furthermore, the GLCM and HT features are also used for the classification of CH images.
- A performance comparison of existing feature extraction methods with G2LCM has given. Moreover, the performance of GLCM, HT, and G2LCM feature extraction methods are validated by training several classifiers on extracted features stated and compared.

The rest of paper is organized as follows: a brief overview of existing methods is described in Section 2. The proposed methodology is described in Section 3. Section 4 gives a summary of experimentations, and Section 5 describes results which are presented with a great detail. The average results of the proposed method are compared with results of existing features extraction methods in Section 6. The paper is finally concluded in Section 7 with some future directions.

2. Related work

Many researchers have developed several methods for automatic detection of abnormalities in gastric tract via analyzing endoscopic frames. We have grouped these methods into three categories according to their description models.

2.1. Methods use texture features to detection abnormalities in gastric-tract

A normalize gray level co-occurrence matrix (NGLCM) was computed by performing the discrete wavelet transform to represent the texture of the images further the extracted features were used to distinguish wireless capsule endoscopy (WCE) frames containing any bleeding signs. Furthermore, the classification was performed by using the support vector machine (SVM) classifier [21]. LBP and vector quantization based approach for extracting texture features was presented in [22], and magnified stomach images were labeled using this approach with a great accuracy. Likewise,

in [23] LBP features were used to classify abnormal WCE frames from the small-intestine. An auto-correlated homogeneous texture (AHT) was proposed using Gabor filters in [24]. Moreover, rotation, scale, and illumination invariant properties of Gabor filters are stated and NBI and CH images were classified using naïve Bayes and SVM classifier. Also, in our previous work, multiple classifiers were trained for classification of CH images using modified Gabor-based texture features [25].

2.2. Methods use color models to detect irregularities in gastric tract

Color histogram descriptors by specifically using YCbCr color space and k-mean clustering were used to segment the image' pixels for detection of salient bleeding areas in WCE frames. Moreover, the classification was performed by extracting color-based features from the region of interest (ROI) [26]. Similarly, the saliency-based method was used for ROI selection, further, the codebook (also known as a bag of features) approach was used to select features from three distinct classes of descriptors. Point-based, shape-based, and texture-based features are extracted from endoscopy images [27]. In the same way, first-order statistics were applied to each salient super-pixel of an endoscopic frame to represent the image color description. Furthermore, bleeding WCE frames were detected through training the SVM classifier with a radial basis function on extracted color features [28].

2.3. Hybrid models for representing image texture

WCE frames having ulcer were classified using multi-scale LBP descriptors. Weak as well as ensemble classifiers were tested with various ensemble strategies for a performance comparison [29]. Frames with ulcer colitis were detected by combining various extracted descriptors (e.g., Gabor, LBP, and MPEG-7) from endoscopic frames as described in [30], then frames were classified into normal and abnormal classes, both block and image level classifications were performed. Another method for ulcer detection from WCE frames was proposed. Log of Gabor (LoG) filters is used to extract texture of images. First, WCE frames were transformed via the contourlet transform. Then a LoG filter bank was used to extract texture features; Furthermore, the images are classified using the SVM classifier [31]. Another hybrid approach was proposed for detection of inflammation from endoscopic frames. The detection was done by analyzing the pit-patterns of the mucosal surface using hybrid texture features. Moreover, textures features were extracted from block filters in [32]. Color Hu moment and Fourier-based color features were extracted from images. Then, the random forest (RF) and KNN classifiers were used for classifying esophageal malignancies in [33]. Similarly, the RF classifier with an ensemble method was used to classify cancerous regions in high-definition endoscopy images [34].

As was previously stated, the dynamics of imaging environment of gastrointestinal tract poses numerous challenges to automated detect the abnormal regions in endoscopy frames. Consequently, existing methods have certain limitations, as color-based models are not adaptive to light variations and lack of sufficient color space makes them susceptible to image variations and causes a higher false detection-rate. In contrast, texture features are well suited for classification of gastric mucosal structures. However, existing methods are designed for general texture analysis applications. A hybrid approach would be a suitable option for efficiently representing gastric images in terms of their descriptors. With this intention, a new hybrid feature extraction technique G2LCM has been proposed by fusion of two existing traditional feature extraction methods. GLCM and Gabor filters are combined advantageously to give a better description of a texture of gastroenterology images.

The reason behind using Gabor filter is its direction and frequency selectivity properties. A bank of Gabor filters (an array of Gabor filters, each Gabor filter is tuned to a specific center frequency, orientation, and scale) mimics the mammalian visual system [24]. Thus, mean and variance of these Gabor filters give us a representation of homogeneous texture for CH images. On the other hand, motivation behind using GLCM was to provide local texture information. GLCM with an offset and direction selection capability delivers rotation and spatial invariance. Therefore, a combined features extracting method will provide features which are more robust to invariant imaging conditions.

3. Methods

As described earlier, the texture is a basic descriptor which hints about the spatial distribution of repetitive intensities of an image. Then the texture features are represented by computing the different order of statistics of these intensities The statistical methods of texture extraction involve calculation of different order of statistics to represent the textures of images. Furthermore, these features are used for training classifiers. In the suggested method, two texture feature extraction techniques are combined to extract the CH images' textures. The overview of the proposed method has manifested in Fig. 2.

After image acquisition, Gabor filters were applied to capture the intensity responses of all CH images and second-order statistics are calculated from GLCM of these responses. A Gabor filter is defined as follows:

3.1. 2D Gabor filter

A Gabor filter can be described by a Gaussian function in the spatial domain as $\dot{G}(x, y)$ with x and y spatial coordinates. The Fourier transform of this Gaussian function $\dot{G}(x, y)$ can be given as $G^f(u, v)$ as a function of frequency components u and v as described in Eq. (1).

$$\dot{G}(x, y, \sigma, \Omega) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) exp\left[-\frac{1}{2}\left\{\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right\} + 2\pi j\Omega x\right] \quad (1)$$

The Gabor filter is a result of the convolution of a Gaussian function and complex sinusoidal in Fourier domain. Moreover, the spread of this Gaussian window is controlled by the standard deviation σ_x and σ_y along the respective spatial axis. The center frequencies of Gabor filters are represented by Ω. On center frequency, the filter results in a higher response. A Gabor filter has shown in Fig. 1 which is tuned to a specific scale and orientation. Using a Gabor filter one can perform a multi-resolution spatial-frequency analysis of images due to its basic characteristics (band-pass nature and direction selectivity). The human visual perception system can be modeled using these phase and orientation selection properties of Gabor filters [35]. The Fourier transform of a Gabor filter as described with its frequency components in Eq. (2).

$$\dot{G}^f(u, v, \sigma, \Omega) = exp\left\{-\frac{1}{2}\left(\frac{(u-\Omega)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right\} \quad (2)$$

where, $\sigma_u = \frac{1}{2\pi\sigma_x}, \sigma_v = \frac{1}{2\pi\sigma_y}$

3.2. Filters dictionary design

Variations of these filters can be obtained by appropriately changing scale and orientation of a Gabor filter and can be defined as given in Eq. (3).

A Gabor filter' s bank has created with 64 filters, every filter with a different orientation, scale, and frequency. Moreover, in Gabor filter' s bank 8 scales, 8 orientations were used and each Gabor

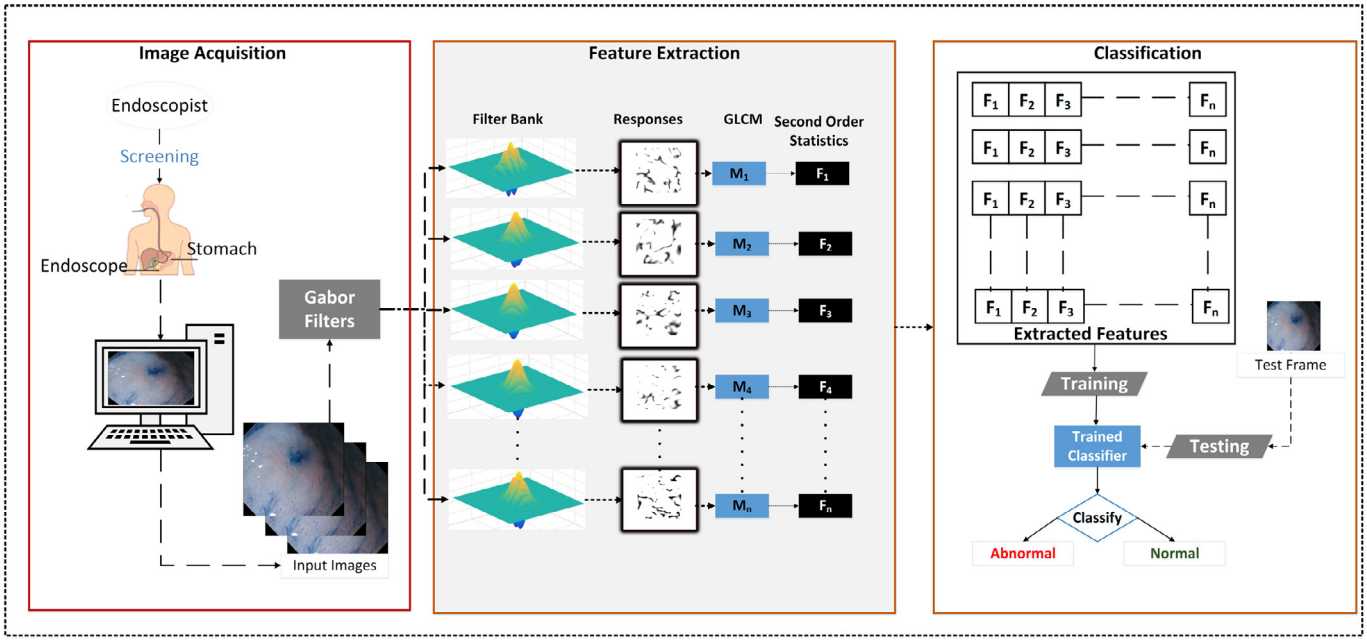

Fig. 2. Shows the flow diagram of proposed method, features from CH images were extracted to classify these images into abnormal and normal case.

filter in filter's bank is responsive to a different frequency present in the images.

Moreover, the properties of direction selectivity and band-pass nature of Gabor filters has been used in a multi-resolution analysis of frames. it also resembles the human visual cortex where every filter represents a visual neuron which is sensitive to a specific frequency. Therefore, features obtained from these filter responses will be rotation, scale, and illumination invariant.

$$\begin{aligned} g_{m,n}(x,y) &= a^{-m} g(x', y' a > 1 | m, n \in Z, \\ x' &= a^{-m}(x\cos(\theta) + y\sin(\theta)) \\ y' &= a^{-m}(y\cos(\theta) - x\sin(\theta)) \end{aligned} \tag{3}$$

The size and rotation of a Gabor filter are controlled by the a^{-m} and θ respectively. A dictionary of Gabor filters is designed to get responses of distinct frequencies from images. K represents the number of orientations in the filter bank and it can be calculated as $\theta = \frac{n\pi}{K}$. Different scales of filter bank denoted by S. The x' and y' are the invariant coordinates of a Gabor filter which are rotated and scaled to some extent. A Gabor filter from the filter bank can be referred by using its scale m and rotation n as shown in Fig. 1.

$$a = \left(\frac{\mathbf{U_h}}{\mathbf{U_l}}\right)^{\frac{1}{S-1}}, \sigma_u = \frac{(a-1)\mathbf{U_h}}{(a+1)\sqrt{(2\ln 2)}},$$

$$\sigma_v = \tan\left(\frac{\pi}{2k}\right)\left[\mathbf{U_h} - 2\ln\left[\frac{2\sigma_u^2}{\mathbf{U_h}}\right]\right]\left[2\ln 2 - \frac{(2\ln 2)^2\sigma_u^2}{\mathbf{U_h^2}}\right]^{-1/2} \tag{4}$$

3.3. Representing texture features

The responses of multiple Gabor filters were obtained by filtering every CH input image. A number of Gabor filters are created by changing the orientation and scale as described earlier. A filter bank containing multiple Gabor filters is used to filter all CH frames. Furthermore, these filters output multi-resolution responses of images. Eq. (5) describes these responses as:

$$B_{m,n} = \int I_{input}(x,y) g_{m,n} * (x - x_1, y - y_1) dx_1 dy_1 \tag{5}$$

A dictionary of Gabor filters is denoted by B and a Gabor filter with a specific scale m and rotation n is represented as $B_{m,n}$. B contains the Gabor feature space and responses can be conceptualized as a 2D matrix with rows and columns. Therefore, the $B_{m,n}$ represents the filter response with having n^th rotation and m^{th} scale components. Moreover, the homogeneous texture features are extracted by computing the higher order statistics from the filter responses by assuming the frequency content of response perceptually homogeneous. The average response of a filter $B_{m,n}$ with m^{th} scale and n^{th} rotation has shown in Eq. (6).

$$\mu_{m,n} = \left| \frac{1}{R \times C} \sum_{i=1}^{R} \sum_{j=1}^{C} B_{m,n}(i,j) \right| \tag{6}$$

The standard deviation often denoted as $\sigma_{m,n}$. It is used to capture the variation in responses from the standard given in Eq. (7).

$$\sigma_{m,n} = \sqrt{\frac{1}{R \times C} \sum_{i=1}^{R} \sum_{i=j}^{C} (|B_{m,n}(i,j)| - \mu_{m,n})^2} \tag{7}$$

A homogeneous texture features set is formed by combining mean and standard deviations, those were extracted from filters responses $\mu_{m,n}$ followed by $\sigma_{m,n}$ can be represented as:

$$F' = [\mu_0, \sigma_0, \mu_1, \sigma_1, \mu_2, \sigma_2, \ldots\ldots, \mu_{m,n}, \sigma_{m,n}] \tag{8}$$

There is no single definition of texture found in literature. Generally, the texture gives important information about nature of surface in an image. Texture hints about an image, whether it contains very smooth, or coarse regions. If regions are smooth the standard deviation of surface intensities will be close to zero and vice versa. Eq. (8) represents homogeneous texture features. Moreover, these features have been widely used for texture analysis and classification of gastric images [24]. In this paper, HT features were also extracted from CH images to compare its performance with proposed feature extraction method.

3.4. Gray level co-occurrence matrix

GLCM is used to analyze local intensity variations in an image further it helps in computing image's texture. The GLCM of an im-

age is calculated by taking account of frequencies of pixels that are appearing on a specific distance repeatedly. Therefore, the GLCM is specially designed for texture representation of images. A matrix $C_{L\times L}$ is a GLCM if each element of this matrix contains an occurrence of a pair of pixels separated by a specified distance D, and L represents the number of gray levels.

$$C_{i,j}(D,R)_{L\times L} = \frac{1}{N}\sum_{l=0}^{L-1}\left(\sum_{x=1}^{X}\sum_{y=1}^{Y}\begin{cases}1, & \text{if } I(x,y)=i\\ \text{and } I(x',y')=j\\ 0, & otherwise\end{cases}\right) \tag{9}$$

where, $D=(\Delta x,\Delta y)$, $x'=x\pm\Delta x$, $y'=y\pm\Delta y$ I denotes the input image and $\pm$ sign indicates the offset can be in both directions. Here, C represents the occurrence on a defined offset D and R denotes rotation parameters. The pixels $I_{x,y}$ and $I_{x',y'}$ are situated at a specific distance and angle. The GLCM contains joint probability of a pixel intensity occurrences. However, the matrix has only frequency information like a histogram. Further, second-order moments are computed from co-occurrence matrix to represent the image's texture. The statistics e.g., mean, variance, covariance, etc., are extracted from co-occurrence matrix. These statistics express various information globally about the surface of the images. In this paper, we have extracted contrast, correlation, energy, and homogeneity to represent the texture of CH images. A sum of differences between co-occurrence matrix elements are used to represent the contrast of the image and it also records the local variation in co-occurrence matrix.

$$Contrast = f^1 = \sum_{i=0}^{M}\sum_{j=0}^{N} C_{i,j}(i-j)^2 \tag{10}$$

In above equation, GLCM is denoted by $C_{i,j}$ and co-occurrence elements are represented by i and j The correlation is a measure of occurrences of the joint probability. This correlation is computed for specified pairs of gray values as given in Eq. (11).

$$Correlation = f^2 = \sum_{i=0}^{M}\sum_{j=0}^{N} C_{i,j}\left(\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{\sigma_i^2\times\sigma_j^2}}\right) \tag{11}$$

μ_i and μ_j are means of co-occurrence elements. Similarly, σ_i and σ_j are standard deviations along respective axes. Moreover, energy delivers the uniformity (angular second moment) of an image' surface by computing the sum of squared elements as shown in Eq. (12).

$$Energy = f^3 = \sqrt{\sum_{i=0}^{M}\sum_{j=0}^{N} C_{i,j}^2} \tag{12}$$

A measure of similarity of distribution of elements of co-occurrence matrix to its diagonals is defined by 13. This measure provides information about the repetition of certain patterns in CH images.

$$Homogeneity = f^4 = \sum_{i=0}^{M}\sum_{j=0}^{N}\frac{C_{i,j}}{1+(i-j)^2} \tag{13}$$

The texture features are normally extracted from GLCM by computing these statistics from co-occurrence matrices. In this paper, we also have extracted these features from GLCM of CH images to see the performance of these features and compare it with the performance of proposed hybrid features.

3.5. Gabor-based gray level co-occurrence matrices (G2LCM) texture features

As mentioned earlier, a Gabor filter defined in Eq. (1) to (3), respond to a specific frequency with also a distinct scale and orientation. However, the images contain frequencies related to various orientations and scales. Therefore, a filter bank has designed with multiple filters to do a multi-resolution analysis of CH images. Every filter in the filter bank responds to a frequency for which it is tuned for and with selective direction and scale. In existing Gabor-based texture extraction method, features are extracted from these filter responses as given in (5). However, we have computed the GLCM from the filter response of every Gabor filter. Then, from Eq. (5) and (9) the GLCM of a filter response can be given as:

$$M_{i,j}(D,R)_{L\times L} = \frac{1}{N}\sum_{l=0}^{L-1}\left(\sum_{x=1}^{X}\sum_{y=1}^{Y}\begin{cases}1, & \text{if } (B_{m,n}(x,y)=i\\ \text{and } B_{m,n}(x',y')=j)\\ 0, & otherwise\end{cases}\right) \tag{14}$$

where, $B_{m,n}$ is a response to a specific Gabor filter. A GLCM matrix is calculated from this filter response. Furthermore, the features set is formed by computing the second order statistics from GLCM as described in Eq. (10) to (13).

$$\Psi_{m,n} = \begin{bmatrix} M_{1,1} & M_{1,2} & M_{1,3} & \cdots & \cdots & M_{1,n}\\ M_{2,1} & M_{2,2} & & & & \\ M_{3,1} & & M_{3,3} & & & \\ \vdots & & & \ddots & & \\ \vdots & & & & \ddots & \\ M_{m,1} & & & & & M_{m,n}\end{bmatrix} \tag{15}$$

Co-occurrence matrices of every filter response are given as $\Psi_{m,n}$. The m and n point to a filter respective to a rotation and a scale. $M_{m,n}$ represents a co-occurrence matrix of a specific filter response. Furthermore, from every co-occurrence matrix, second-order statistics are calculated to form a feature set. G2LCM features extracted from a co-occurrence matrix of a Gabor filters are arranged as presented in the Eq. (16).

$$F_i = \begin{bmatrix} f^1 & f^2 & f^3 & f^4\end{bmatrix} \tag{16}$$

Statistical measures from every matrix are computed and each statistical feature's value has been concatenated with other, as a subset of features given in Eq. (16). If we convert the Ψ into one dimension, F_i contains the texture information of i^{th} GLCM. However, features set is formed by combining all statistical values into a single row vector represented in Eq. (17) to describe the texture characteristics of single CH image.

$$F = F_1, F_2, F_3 \ldots\ldots\ldots F_n \tag{17}$$

3.6. Training of classification models

We have compared the performance of multiple classifiers by training these classifiers on features which were extracted by existing texture extraction methods (Gabor-based homogeneous textures and GLCM texture features), further we have also trained these classifiers on G2LCM texture features. The motivation behind training so many classifiers is, this comparison will give us an insight how the certain classifiers respond to different texture features. Therefore, SVM, Naïve Bayes(NB), k-NN, linear discriminant (LDA), decision tree (DT), and ensemble classifiers were used for classification of CH images using extracted features.

The SVM classifies by maximizing the gap between the decision boundaries of features belong to different classes. The SVM classifier is used with a Gaussian kernel function for training and testing. The reason for selecting Gaussian kernel function is because there is a diverse range of texture features and they are non-linearly separable.

Furthermore, NB classifier is trained for classification. The NB assumes that the features are a multivalued Gaussian distribution. Moreover, NB classifier is based on Bayes theorem of finding a posterior probability of class label while the prior probabilities of features related to specific classes are already known. NB is based on

an assumption of class independence, each attribute is independent of other attribute and class only depends on the product of prior of a feature vector.

On the contrary, the k-NN classifier assigns a class label to an image by considering the features relation by calculating the distance of features in feature's space. In k-NN, An assignment of a class label to row vector based on similarity of features to a learned feature model. In this study, a k-NN is used with 5 neighboring elements and trained on texture features and a correlation function is used as a distance of measure.

Discriminant analysis is normally used for dimension reduction applications. However, it also can be used for classification purpose. We have used a linear type discriminant for a regularized linear discriminant analysis. Moreover, the LDA considers that the features are independent and normally distributed. Therefore, we have used LDA classifier to compare its performance with other classifiers. In particular, DT classifier is also trained to verify the performance of a diverse range of classifiers. The DT classifier employs the features' statistics to calculate the inter-dependence of attributes (contrary to NB where attributes are conditionally independent). The DT can be visualized as a conjunction of disjunctions of attributed values and can be represented by a tree. The DT have class labels on its leaf nodes and attribute values on its edges. The specific parameters for growing tree are N-1 number of maximum splits where N denotes the total number of examples and with minimum leaf size is 1, minimum parents are chosen to be 10.

Moreover, ensemble method is also used for classification. Multiple weak classifiers' results can be combined to support a single decision in ensemble method. Therefore, instead of relying on a single weak DT classifier an ensemble classification is used. In an ensemble classifier 200 weak classifiers take part in training and testing. Then, results of these weak classifiers are combined by a technique called adaptable boosting (Ada-Boost). Moreover, the performance results of these classifiers are presented in results and discussion sections.

4. Materials and experimentation

As mentioned before, CH is an advanced imaging modality for visualizing gastrointestinal tract. It includes staining tissues with the methylene blue, which interacts with the mucosal walls of the gastric tract and increase the visibility of gastric lesions for doctors. In this study, we used CH frames for detection of abnormalities.

We used a publicly available images data-set having images of CH of multiple patients with normal and abnormal tumorous conditions. CH data-set is composed of 176 CH frames with expert annotations and it came with two sets, one set is for training and other for testing the proposed methods. Moreover, these images were analyzed according to the taxonomy given by [36].

This images data-set is part of the analysis of images to detect abnormalities in endoscopy (AIDA-E) which is a biomedical challenge.

Data-set contains three groups, images with normal mucosal regions are in group I (Normal class) and the images with abnormal regions belong to group II and group III. Images with two condition metaplasia and dysplasia are referred to positive (Abnormal class) as shown in Fig. 3. In data-set 56 (31.8%) CH images and annotation files from Group and 120 images belongs to group II and III, further group II contains 96 images (54.5%) and 24 images (13.6%) belong to Group III(for data distribution see Table 1). The goal of this biomedical challenge is to classify the manually segmented area of each image as either abnormal or normal. However, we have used whole images for extraction of features so that the burden of medical expert should be minimized. The images obtained by utilizing an Olympus GIF-H180 endoscope at the Portuguese Institute of Oncology (IPO) Porto, Portugal while a regular clinical checkup. Visual settings of this endoscope include 140° field of view (FOV) and four-way angulation (210° up, 90° down and 100° for right and left). Furthermore, the annotations were conducted independently by two medical experts (Dr. Dinis-Ribeiro, Dr. Miguel Areia) leading to a gold-standard final annotation which has used here[1].

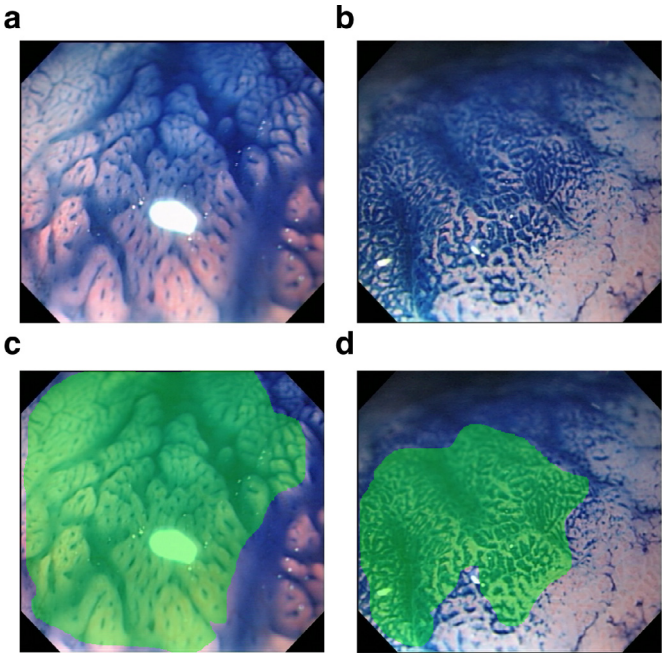

Fig. 3. (a) Shows a normal CH image, (b) an abnormal CH image, (c) area of interest in a normal frame, and in (d) the abnormal area annotated with green overlay. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Distribution of chromoendoscopy images dataset.

Class	Normal	Abnormal		Total
Group	Group I	Group II	Group III	176
No. of Images	56	96	24	
Percentage	31.80%	54.50%	13.60%	

4.1. Performance measures

The image classification's performance of the proposed features and other traditional feature extraction approaches are compared on the basis of following accuracy measures: sensitivity (Sen), specificity (Spec), accuracy (ACC) and area under the curve (AUC). The sensitivity also referred to true positive rate (TPR) and specificity is also known as a true negative rate (TNR)) are described in Eq. (18) and (20) respectively. However, the false positive rate (FPR) of a system can be derived from TNR as in Eq. (19). Moreover, the accuracy of a system will be computed by first calculating the false positive (FP), false negative (FN), true positives (TP), and true negative (TN) as mentioned in Eq. (21). The area under the curve (AUC) is calculated as shown in Eq. (22).

$$TPR = Sensitivity = \frac{TP}{TP + FN} \tag{18}$$

$$FPR = 1 - TNR \tag{19}$$

[1] https://www.aidasub-chromogastro.grand-challenge.org/description/.

Table 2
Classification performance of multiple classifiers by using homogeneous texture descriptors.

Classifiers	TN	FN	FP	TP	Sen	Spec	ACC	AUC
SVM	40	16	17	103	87%	70%	81%	0.86
NB	49	7	35	85	92%	58%	76%	0.80
KNN	34	22	20	100	82%	63%	76%	0.83
LDA	43	13	36	84	87%	54%	72%	0.77
DT	27	29	23	97	77%	54%	70%	0.79
Ensemmble	37	19	11	109	85%	77%	83%	0.88

$$TNR = Specificity = \frac{TN}{FP + TN} \tag{20}$$

$$Accuracy = \frac{TP + TN}{P + N} \tag{21}$$

$$AUC = \int_{-\infty}^{\infty} TPR(T)FPR'(T)dT \tag{22}$$

4.2. Experimentation

As we have discussed earlier, the publicly available CH images data-set came with an explicit partitioning (provided by the authors of data-set) one set is for training and other is for testing.

However, this is not a standard way of training and testing classifiers. The reason is, trained model will not well generalized with a fixed training and testing set and will not able to well perform on future data.

Therefore, in our experimentation, we used k-cross- validation method for selection of training and testing data. After the feature extraction, multiple classifiers are trained using 10-cross-validation with many iterations to achieve more concrete results. After the feature extraction phase, features from all images are extracted and then every image is represented by a row vector. Furthermore, the last element of a row vector is the class label.

In k-cross-validation, the features vectors corresponding to images are randomly divided into k subsets. Then, from the subsets, a single subset is used for training and remaining are subsets used for testing a classifier.

It is worth mentioning here when a random data is selected in k-cross-validation for training and testing, accuracies of classifiers will also change every time. Therefore, the procedure is repeated over 100 iterations to get average results. Moreover, the performance of multiple classifiers has also shown in results section which is obtained from an only a single run of 10-cross-validation. All experiments are conducted on Intel Core i5 with 2.40 GHz using the MATLAB.

5. Results

After features extraction, CH images are classified into normal and abnormal classes We have extracted Gabor-based homogeneous texture features, further we have also extracted GLCM based texture features for comparison of performance with proposed G2LCM texture features. These texture features were extracted from all CH images and multiple classifiers are trained and tested.

Furthermore, results are shown in Table 2 refers to one iteration of training and testing of classification models using 10-cross-validation. These classification results are obtained by classifying CH frames using HT descriptors. It is clear from the results that the SVM classifier best performs and have 0.86 AUC with 81% ACC and the LDA have the least performance on HT descriptors in terms of accuracy.

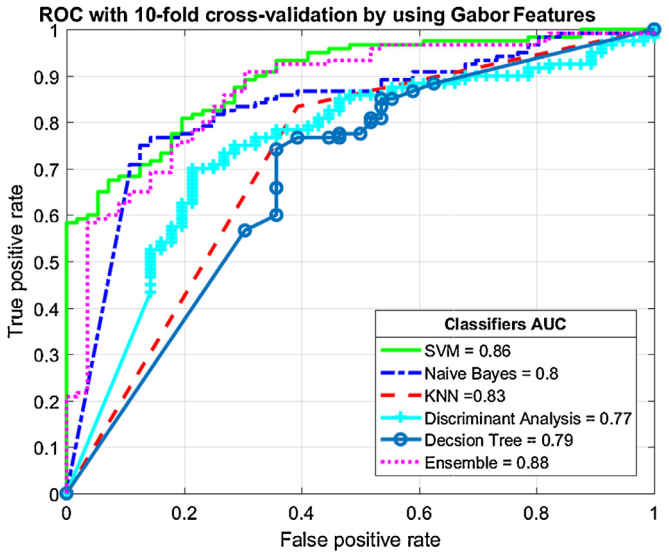

Fig. 4. ROC and AUC analysis of multiple classifiers by training and testing on HT features, where the 10-cross-validation is used.

Table 3
Classification performance of multiple classifiers by using GLCM texture descriptors.

Classifiers	TN	FN	FP	TP	Sen	Spec	ACC	AUC
SVM	38	18	17	103	85%	69%	80%	0.85
NB	50	6	29	91	94%	63%	80%	0.84
KNN	32	24	19	101	81%	63%	76%	0.82
LDA	40	16	21	99	86%	66%	79%	0.84
DT	39	17	23	97	85%	63%	77%	0.83
Ensemble	37	19	18	102	84%	67%	79%	0.85

ROC analysis of the classifiers on HT features has shown in Fig. 4, where the performance curve of SVM classifier manifest good AUC and sensitivity.

In experiments, every time a new subset of data is selected for training and testing, therefore each time the accuracy of classifiers will also change each time. Therefore, these experiments are repeated numerous time to get more accurate results. Furthermore, the average results are calculated with a margin of error and with a 96% confidence interval. As the result from multiple iterations, the SVM, NB, k-NN, LDA, DT, and ensemble classifier have achieved 81.9% ± 0.1, 76.9% ± 0.07, 76.3% ± 0.2, 76.3% ± 0.5, 67.8% ± 0.4, 72.2% ± 0.3, and 81.1% ± 0.3 respectively average accuracies with a margin of error. In same order, the average AUCs of classifiers are 0.86 ± 0.001, 0.80 ± 0.000, 0.82 ± 0.001, 0.73 ± 0.005, 0.79 ± .003, and 0.86 ± 0.002. It is clear from the results, that the SVM and ensemble methods have similar average classification performance. However, the ensemble method required training multiple classifiers, which is computationally exhaustive.

After testing the performance of Gabor-based HT features, for further verification, the classifiers are trained and tested on GLCM texture features for classification CH images. The Fig. 5 shows AUCs of classifiers on a single iteration. Similarly, the results have shown in Table 3 are from a single iteration of training and testing phase. Where all classifiers have a similar performance. However, the NB classifier has a higher AUC than other classifiers. The average accuracies of classifiers with a margin of error by using GLCM texture features with SVM 79.9% ± 0.3 ACC and 0.85 ± 0.001 AUC. The NB has 80.8% ± 0.1 average ACC with 0.85 ± 0.001 AUC. The LDA classifier 79.7% ± 0.2 average ACC with 0.85 ± 0.001 AUC. Moreover, the k-NN, DT, and ensemble classifiers have 74.7% ± 0.1, 76.8% ± 0.1, and

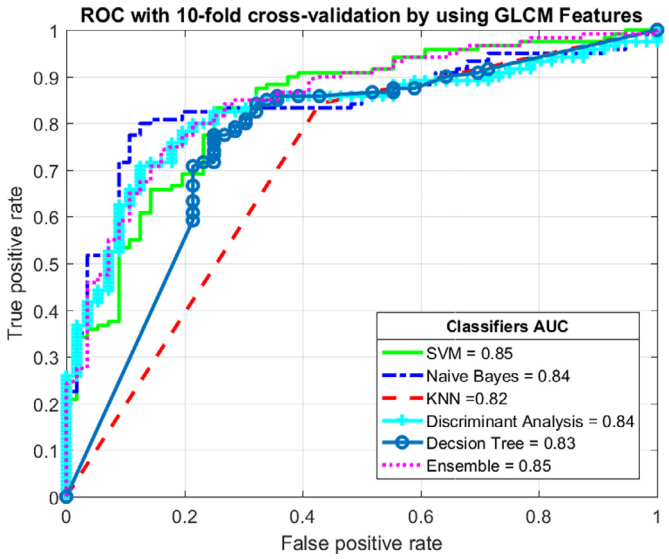

Fig. 5. ROC and AUC analysis of multiple classifiers trained and tested over GLCM texture features by using 10-cross-validation method.

Table 4
Classification performance of multiple classifiers on G2LCM descriptors.

Classifiers	TN	FN	FP	TP	Sen	Spec	ACC	AUC
SVM	45	11	10	110	91%	82%	88%	0.91
NB	47	9	34	86	91%	58%	76%	0.80
KNN	42	14	17	103	88%	71%	82%	0.87
LDA	36	20	28	92	82%	56%	73%	0.79
DT	37	19	19	101	84%	66%	78%	0.84
Ensemble	43	13	14	106	89%	75%	85%	0.89

77.6% ± 0.4 average accuracies and have 0.82 ± 0.001, 0.83 ± 0.003, and 0.84 ± 0.002 AUCs respectively.

In the same way, the results of proposed texture features G2LCM have shown in Table 4 from a single iteration. These results show the hybrid texture features have a performance higher than existing texture extraction methods for classification of CH images. Moreover, it can be seen from the results that there is a rise in the accuracy of every classifier when they are trained and tested against the proposed G2LCM texture features. Thus, this will also indicates that the proposed G2LCM has a higher discriminative power than existing texture features for classification of CH frames.

The SVM classifier has a highest 88% ACC and AUC 0.91. The AUCs of all classifiers on G2LCM texture features are depicted in Fig. 6. Therefore, we can conclude here that the SVM classifier performs best among all other classifiers specifically when used along texture descriptors. Moreover, the average performance of SVM classifier is 87% ± 0.2 ACC and 0.91 ± 0.001 AUC.

The NB, k-NN, LDA, DT, and ensemble classifiers have 77% ± 0.1, 81% ± 0.3, 73% ± 0.4, 82% ± 0.4, and 83% ± 0.2 ACCs accordingly. The corresponding AUCs of classifiers are 0.81 ± 0.001, 0.86 ± 0.002, 0.79 ± 0.004, 0.86 ± 0.003, and 0.88 ± 0.002. However, the ensemble classifier performs well among other classifiers, but has less accuracy and AUC than SVM. Thus the SVM classifier is the best choice for classification of CH images using hybrid G2LCM texture features.

6. Discussion

The average accuracy of the proposed method is compared with accuracies of those methods which were developed for classification of CH images in Table 5. An auto-correlation homogeneous texture (AHT) texture extraction method was proposed by

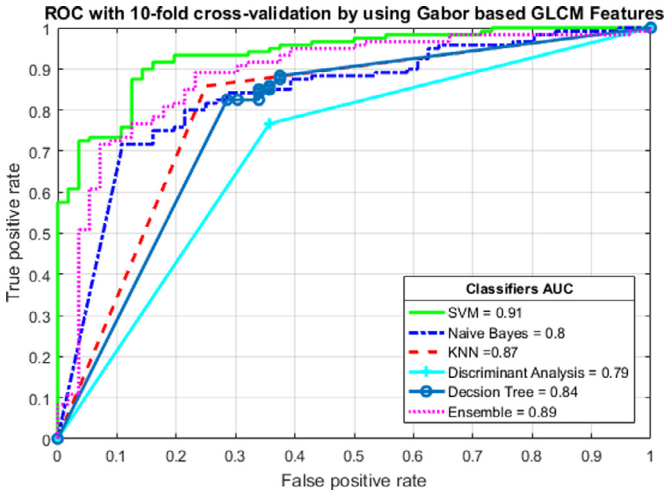

Fig. 6. ROC and AUC of multiple classifiers by using proposed Gabor co-occurrence hybrid texture features using 10-cross-validation.

Table 5
Classifiers and features extraction methods with highest average ACCs and AUCs.

Ref.	Classifier	Descriptors	ACC	AUC
[24]	SVM	AHT	83.1%	0.85
[25]	SVM	GHT	86.1%	0.91
-	SVM	HT	81.0%	0.86
-	NB	GLCM	80.8%	0.85
Proposed	SVM	G2LCM	87.2%	0.91

[24] to deal with rotation, scale, and illumination invariance. Furthermore, these features were used for classification of CH images and achieved maximum 83.1% ACC with 0.95 AUC. In our previous work, we have proposed geometric homogeneous texture (GHT) features by modifying the statistics of micro-features of Gabor filters in [25] and achieved 86.1% ACC and 0.91 AUC without any optimization.

Moreover, the performance of our proposed G2LCM texture features has compared with traditional texture extraction methods HT, GLCM features. The maximum average performance by using the HT features, SVM classifier has 81% ACC and 0.86 AUC. NB classifier has the highest performance when it was trained with GLCM texture features as shown in the previous section. However, the G2LCM method uses the combined power of GLCM and Gabor filters. GLCM extracts texture in information in the spatial domain, and on the other hand, in Gabor filters images are processed in the frequency domain. Therefore, this hybrid approach provides a good texture description of lesions in CH images, which is the main reason for higher average performance for classification CH images images.

7. Conclusion and future directions

This paper present a feature extraction method named as Gabor-based gray-level co-occurrence matrix (G2LCM) for texture representation of endoscopic images. Endoscopic images were obtained from an endoscopic procedure using chromoendoscopy (CH) and analyzed for diagnosis of gastric abnormalities by using G2LCM. Furthermore, these CH images are classified into normal and abnormal classes. The homogeneous texture (HT) and gray-level co-occurrence matrices (GLCM) features are extracted from CH images and these images are classified via multiple classifiers. Moreover, the performance of proposed texture features is compared with existing texture extraction methods. It is confirmed

from the results that the SVM classifier outperforms among all of the other state-of-the-art classifiers when it is used with G2LCM texture features. In this paper, we have only used texture features for classification of CH images. However, colors are also important visual cues for representation of CH images. In future, it would be fascinating to encompass color information along with textures to represent characteristics gastric lesions more accurately. Moreover, we have used only CH images, however, this method can also be generalized for other endoscopy modalities.

Conflict of interest

We declare that there is no conflict of interest related to this paper.

Acknowledgment

The authors would like to thank Dr. Farhan Riaz, Assistant Professor, CE&ME, National University of Sciences and Technology (NUST) Pakistan, who provided insight and expertise that greatly assisted this research.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2015, CA Cancer J. Clin. 65 (1) (2015) 5–29.
[2] R. Miyaki, S. Yoshida, S. Tanaka, Y. Kominami, Y. Sanomura, T. Matsuo, S. Oka, B. Raytchev, T. Tamaki, T. Koide, K. Kaneda, M. Yoshihara, K. Chayama, A computer system to be used with laser-based endoscopy for quantitative diagnosis of early gastric cancer., J. Clin. Gastroenterol. 49 (2) (2015) 108–115.
[3] J.H. Nam, I.J. Choi, S.-J. Cho, C.G. Kim, J.K. Jun, K.S. Choi, B.-H. Nam, J.H. Lee, K.W. Ryu, Y.-W. Kim, Association of the interval between endoscopies with gastric cancer stage at diagnosis in a region of high prevalence., Cancer 118 (20) (2012) 4953–4960.
[4] M. Lopez-Ceron, F.J.C. Van Den Broek, E.M. Mathus-Vliegen, K.S. Boparai, S. Van Eeden, P. Fockens, E. Dekker, The role of high-resolution endoscopy and narrow-band imaging in the evaluation of upper GI neoplasia in familial adenomatous polyposis, Gastrointest. Endosc. 77 (4) (2013) 542–550.
[5] S.M. Sweeney, AACR Cancer Progress Report 2014, American Association for Cancer Research, 2014, p. 126.
[6] K.C. Lau, E.Y.Y. Leung, P.W.Y. Chiu, Y. Yam, J.Y.W. Lau, C.C.Y. Poon, A flexible surgical robotic system for removal of early-stage gastrointestinal cancers by endoscopic submucosal dissection, IEEE Trans. Ind. Inf. 12 (6) (2016) 2365–2374.
[7] W.Y. Cho, J.Y. Jang, D.H. Lee, Recent advances in image-enhanced endoscopy, Clin. Endosc. 44 (2) (2011) 65–75.
[8] C.L. Leggett, P.G. Iyer, Mucosal imaging advanced technologies in the gastrointestinal tract, Tech. Gastrointest. Endosc. 17 (4) (2015) 161–170.
[9] Y. Yuan, M.-H. Meng, Automatic bleeding frame detection in the wireless capsule endoscopy images, in: IEEE International Conference on Robotics and Automation, ICRA, 2015, pp. 1310–1315.
[10] M. Liedlgruber, A. Uhl, Computer-aided decision support systems for endoscopy in the gastrointestinal tract: a review, IEEE Rev. Biomed. Eng. 4 (2011) 73–88.
[11] F. Riaz, F.B. Silva, M.D. Ribeiro, M.T. Coimbra, Impact of visual features on the segmentation of gastroenterology images using normalized cuts, IEEE Trans. Biomed. Eng. 60 (5) (2013) 1191–1201.
[12] M.M. Martins, D.J. Barbosa, J. Ramos, C.S. Lima, Small bowel tumors detection in capsule endoscopy by Gaussian modeling of color curvelet covariance coefficients, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, 2010, pp. 5557–5560.
[13] D.Y. Liu, T. Gan, N.N. Rao, G.G. Xu, B. Zeng, H.L. Li, Automatic detection of early gastrointestinal cancer lesions based on optimal feature extraction from gastroscopic images, J. Med. Imaging Health Inform, 5 (2015) 296–302.
[14] F. Riaz, A. Hassan, R. Nisar, M. Dinis-Ribeiro, M. Coimbra, Content-adaptive region-based color texture descriptors for medical images, IEEE J. Biomed. Health Inform. (2015).
[15] R. Nawarathna, J. Oh, J. Muthukudage, W. Tavanapong, J. Wong, P.C. de Groen, S.J. Tang, Abnormal image detection in endoscopy videos using a filter bank and local binary patterns, Neurocomputing 144 (2014) 70–91.
[16] M.T. Coimbra, J.S. Cunha, MPEG-7 visual descriptors contributions for automated feature extraction in capsule endoscopy, IEEE Trans. Circuits Syst. Video Technol. 16 (5) (2006) 628–637.
[17] A. Karargyris, N. Bourbakis, Identification of ulcers in wireless capsule endoscopy videos, in: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI, 2009, pp. 554–557.
[18] B. André, T. Vercauteren, A. Perchant, A.M. Buchner, M.B. Wallace, N. Ayache, Introducing space and time in local feature-based endomicroscopic image retrieval, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5853 LNC, 2010, pp. 18–30.
[19] B.S. Manjunath, W.-Y. Ma, Texture features for browsing and retrieval of image data, IEEE Trans. Pattern Anal. Mach. Intell. 18 (8) (1996) 837–842.
[20] S. Ameling, S. Wirth, D. Paulus, G. Lacey, F. Vilarino, Texture-based polyp detection in colonoscopy, Bildverarbeitung für die Medizin (2009) 346–350.
[21] A.R. Hassan, M.A. Haque, Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos, Comput. Methods Programs Biomed. 122 (3) (2015) 341–353.
[22] T.-C. Lee, Y.-H. Lin, N. Uedo, H.-P. Wang, H.-T. Chang, C.-W. Hung, Computer-aided diagnosis in endoscopy: a novel application toward automatic detection of abnormal lesions on magnifying narrow-band imaging endoscopy in the stomach, in: 35th Annual International Conference of the IEEE in Engineering in Medicine and Biology Society, EMBC, 2013, pp. 4430–4433.
[23] A.F. Constantinescu, M. Ionescu, I. Rogoveanu, M.E. Ciurea, C.T. Streba, V.F. Iovanescu, S.A. Artene, C.C. Vere, Analysis of wireless capsule endoscopy images using local binary patterns, Appl. Med. Inform. 36 (2) (2015) 31.
[24] F. Riaz, F.B. Silva, M.D. Ribeiro, M.T. Coimbra, Invariant Gabor texture descriptors for classification of gastroenterology images, IEEE Trans. Biomed. Eng. 59 (10) (2012) 2893–2904.
[25] H. Ali, M. Sharif, M. Yasmin, M.H. Rehmani, Computer-based classification of chromoendoscopy images using homogeneous texture descriptors, Comput. Biol. Med. 88 (2017) 84–92.
[26] Y. Yuan, B. Li, Q. Meng, Bleeding frame and region detection in the wireless capsule endoscopy video, IEEE J. Biomed. Health Inform. 2194 (2015). 1–1
[27] Y. Yuan, J. Wang, B. Li, M.Q.H. Meng, Saliency based ulcer detection for wireless capsule endoscopy diagnosis, IEEE Trans. Med. Imaging 34 (10) (2015) 2046–2057.
[28] D.K. Iakovidis, D. Chatzis, P. Chrysanthopoulos, A. Koulaouzidis, Blood detection in wireless capsule endoscope images based on salient superpixels, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), 2015, pp. 731–734.
[29] B. Li, M.Q.-H. Meng, J.Y. Lau, Computer-aided small bowel tumor detection for capsule endoscopy, Artif. Intell. Med. 52 (1) (2011) 11–16.
[30] A. Dahal, J. Oh, W. Tavanapong, J. Wong, P.C. De Groen, Detection of ulcerative colitis severity in colonoscopy video frames, International Workshop on Content-Based Multimedia Indexing, 2015.
[31] N.E. Koshy, V.P. Gopi, A new method for ulcer detection in endoscopic images, in: 2nd International Conference on Electronics and Communication Systems (ICECS), IEEE, 2015, pp. 1725–1729.
[32] I. Ševo, A. Avramović, I. Balasingham, O.J. Elle, J. Bergsland, L. Aabakken, Edge density based automatic detection of inflammation in colonoscopy videos, Comput. Biol. Med. 72 (2016) 138–150.
[33] L. Serpa-Andrade, V. Robles-Bykbaev, L. González-Delgado, J.L. Moreno, An approach based on Fourier descriptors and decision trees to perform presumptive diagnosis of esophagitis for educational purposes, in: International Autumn Meeting on Power Electronics and Computing (ROPEC), IEEE, 2015, pp. 1–5.
[34] M.H.A. Janse, F. van der Sommen, S. Zinger, E.J. Schoon, P.H.N. de With, Early esophageal cancer detection using RF classifiers, Proc. SPIE 9785 (2016). 97851D–1 to 8
[35] S. Marĉelja, Mathematical description of the responses of simple cortical cells, J. Opt. Soc. Am. A 70 (11) (1980) 1297–1300.
[36] A. Sousa, M. Dinis-Ribeiro, M. Areia, M. Coimbra, Identifying cancer regions in vital-stained magnification endoscopy images using adapted color histograms, in: 16th IEEE International Conference on Image Processing (ICIP), 2009, pp. 681–684.