Accepted Manuscript

Computer-based classification of chromoendoscopy images using homogeneous texture descriptors

Hussam Ali, Muhammad Sharif, Mussarat Yasmin, Mubashir Husain Rehmani

PII: S0010-4825(17)30226-3

DOI: 10.1016/j.compbiomed.2017.07.002

Reference: CBM 2713

To appear in: *Computers in Biology and Medicine*

Received Date: 17 February 2017

Revised Date: 29 June 2017

Accepted Date: 2 July 2017

Please cite this article as: H. Ali, M. Sharif, M. Yasmin, M.H. Rehmani, Computer-based classification of chromoendoscopy images using homogeneous texture descriptors, *Computers in Biology and Medicine* (2017), doi: 10.1016/j.compbiomed.2017.07.002.

Computer-based Classification of Chromoendoscopy Images using Homogeneous Texture Descriptors

Hussam Ali*, Muhammad Sharif†, Mussarat Yasmin‡ and Mubashir Husain Rehmani §
*†‡Department of Computer Science
§ Department of Electrical Engineering
COMSATS Institute of Information Technology
Wah, Pakistan
Email: *hussamalics@gmail.com, †muhammadsharifmalik@yahoo.com, ‡mussaratabdullah@gmail.com,
§mshrehmani@gmail.com

***Abstract*—**
Computer-aided analysis of clinical pathologies is a challenging task in the field of medical imaging. Specifically, the detection of abnormal regions in the frames collected during an endoscopic session is difficult. The variations in the conditions of image acquisition, such as field of view or illumination modification, make it more demanding. Therefore, the design of a computer-assisted diagnostic system for the recognition of gastric abnormalities requires features that are robust to scale, rotation, and illumination variations of the images. Therefore, this study focuses on designing a set of texture descriptors based on the Gabor wavelets that will cope with certain image dynamics. The proposed features are extracted from the images and utilized for the classification of the chromoendoscopy (CH) frames into normal and abnormal categories. Moreover, to attain a higher accuracy, an optimized subset of descriptors is selected through a genetic algorithm. The results obtained using the proposed features are compared with other existing texture descriptors (e.g., local binary pattern and homogeneous texture descriptors). Furthermore, the selected features are used to train the support vector machine (SVM), naive Bayes (NB) algorithm, k-nearest neighbor algorithm, linear discriminant analysis, and ensemble tree classifier. The performance of these state-of-the-art classifiers for different texture descriptors is compared based on the accuracy, sensitivity, specificity, and area under the curve (AUC) derived by using the CH images. The classification results reveal that the SVM classifier achieves 90.0% average accuracy and 0.93 AUC when it is employed with an optimized set of features obtained by using a genetic algorithm.

***Index Terms*—Classification, Endoscopy, Feature Extraction, Gastrointestinal, Gastric Cancer, Local Binary Patterns, Homogeneous Texture, Chromoendoscopy, Genetic Algorithm.**

I. Introduction

In recent years, there has been an increase in the concern for gastric cancer globally [1]. An inappropriate diet is one of the main causes of complications (e.g., ulcer and inflammation) in the gastrointestinal (GI) tract [2]. In addition, these abnormalities may contribute to the development of gastric cancer [3].

An early diagnosis of tumors is useful for decreasing the mortality rate [4]. For the well-timed detection of tumors, a normal clinical practice is intestinal biopsy (in which tissue samples of the mucosa are collected and analyzed) conducted by an expert to identify if there are any cancerous or abnormal cells present in the tissue samples [5]. In contrast, endoscopy is a less invasive method for screening the GI tract. An endoscope is composed of a flexible tube with a mounted camera, light source, and surgical apparatus [6]. Therefore, an endoscope is also sometimes used for performing GI biopsies [7]. Inspection of the GI tract via an endoscope is an indispensable task for the timely identification of irregularities (e.g., cancer, ulcer, and polyps) in gastric patients. Various improvements have been made in the video endoscopy technology [8]. Chromoendoscopy (CH) is an advancement of video endoscopy [9]. CH facilitates the investigation of mucosal vascular structures by spraying dyes over the mucosal surface. The dyes make the cancerous regions more prominent visually, and several clinical studies have also utilized their benefits. Digital (virtual) CH employs image processing algorithms and uses band-pass filters to render the effect of dye-based (traditional) CH. The advantage of virtual CH over traditional CH is that there is no requirement of spraying colorants. Therefore, there is no necessity for extra cleansing that is otherwise essential before performing further endoscopic procedures [10]. An endoscopic procedure performed for a single patient can consume 45 min to 8 h of time, producing more than 80,000 frames. However, majority of the frames are discarded because of degradation or high-similarity between them. Thus, for a physician it is a rigorous and lengthy process to inspect each frame individually. Moreover, the time required by an endoscopic session also depends on the GI tract target area and skills of the endoscopist [11].

Computer-aided design (CAD) support systems can play an important role in providing a better diagnosis by enhancing the details in the endoscopic videos. CAD can also provide a second opinion to a gastroenterologist on a decision taken based on a manual analysis of the endoscopic frames [12]. Textures are basic visual features for recognizing an object. They are also useful in the examination of the abnormal regions in endoscopic frames. Textures are perceptually homogeneous and have an almost repetitive structure that is a local or global pattern distributed over an entire image [13]. Texture analysis has been an interesting area for researchers owing to its various applications in medical imaging and remote sensing. Abnormalities in the GI tract can be detected through the examination of random pit patterns and vascular structures via texture analysis of the endoscopic frames [14]. Several

methods have been developed for texture representation and classification of gastric images such as local binary patterns (LBP) [15] and statistical analysis [16] of the images in the spatial domain. Some techniques are based on the curvelet analysis of the images in the frequency domain. Filters are applied to the images to obtain the texture information [17]. Gabor filters are also used to extract texture features with the least joint two-dimensional-spatial and frequency uncertainty [18]. A filter bank is designed by a combination of filters of varying sizes and orientations. The coefficients of the responses of the filters were used to obtain the texture information in [19]. Texture features are widely employed in image retrieval applications, image segmentation, image recognition, image registration, and motion tracking applications [20]. The systems presented in [21] and [14] used the arithmetic mean and standard deviation of the responses of the filters to represent the texture of an endoscopic frame.

In this study, we have proposed a feature extraction method by computing the geometric mean and geometric standard deviation of the responses of the filters to represent the textures of the endoscopic frames. The extracted texture features are then used for the classification of CH images as normal and abnormal . Moreover, feature selection is performed via a genetic algorithm (GA). Subsequently, state-of-the-art classifiers are trained by conducting a 10-cross-validation of the extracted texture features, and the results are also compared.

A. Contributions

In this study, our primary focus is to design features that are robust to various image acquisition settings and can better distinguish between the disease and normal regions in the endoscopic frames. Our main contributions are as follows:

- A new method of designing geometric homogeneous textures (GHTs) is proposed for
  texture feature extraction. This is accomplished by altering the method of capturing the responses of the filters. Basically, second-order statics are modified to generate more robust micro-features.
- Moreover, the classification results of the proposed features are compared with homogeneous texture descriptors (features obtained by the arithmetic mean and standard deviation of the responses of the filters as described in [20]) and LBP features.

The rest of the paper is organized as follows: diverse existing CADs and their comparison are given in Section II, and the proposed technique is described in Section III. Moreover, along with the experimentation details, the results are discussed in Section IV. Finally, the study is concluded in Section V.

II. Related Work

In the preceding one and a half decades, several CAD systems have been developed for the automatic examination of GI diseases. In this section, some of these existing CAD systems are described, and in particular those using texture descriptors for classification. We organized the CAD methods described in Table I into three classes on the basis of their reported corresponding domains as follows:

A. Texture Feature Extraction Methods in the Spatial Domain

Given the importance of gastric disease detection, numerous researchers have developed methods for the extraction of features in the spatial domain. Thus, uniform LBP features were extracted to classify the abnormal wireless capsule endoscopy (WCE) images of the small bowels of the patients suffering from Crohn?s disease [22]. The LBP features deliver the texture information specifically, in the spatial domain. Further, the SVM classifier is for training and testing using these descriptors. Similarly, the region of interest is selected in [23] using the k-mean clustering. Moreover, LBP features are extracted from the segmented area of the narrow band imaging (NBI) images for the detection of irregular regions in the stomach. The co-occurrence matrices are calculated by dividing the colonoscopy images into several sub-images as in [16]. The co-occurrence matrices contain the information about the occurrence of the pixels, that are located within a specified distance in a sub-image. Then, the features are calculated by the second order statistics of a co-occurrence matrix of an image (e.g., correlation, inverse difference moment, and energy). Afterward, these features are used in training and testing of an artificial neural network (ANN) for the classification of abnormal frames. The power of an ANN is exploited for an automatic learning of the features by using the deep learning method. A multilayer convolution neural network (CNN) is used to capture the important information from the images, which is later used for the classification of gastric images by using the SVM and ELM classifiers [11].

B. Texture Feature Extraction Methods in the Frequency Domain

Curvelet based methods have been widely used for the analysis of images. Therefore, the LBP features are used in combination with curvelet transformation to deliver texture information for the classification of images. Multilayer perceptron (MLP) ANN and SVM are used to analyze the wireless capsule endoscopy (WCE) images in [15]. Moreover, curvelet based LBP provides a robust and better discrimination of the ulcer regions in a dynamic imaging setting. Similar, to their previous work, multi-scale LBP with discrete wavelet transforms was used to obtain the texture features in [27]. The Small intestine tumors are analyzed using these descriptors and various classifiers such as (KNN), MLP ANN, and SVM.

In the frequency domain, multiple ring-shaped filters were used for the feature extraction in [25]. Further, a subset of important features were selected using the GA. The KNN SVM and naive Bayes (NB) classifiers were trained to classify the images with a mucosal damage, which was mostly because of a Celiac disease. A set of zoom-endoscopy images was categorized through the texture analysis of the pit patterns. The images were later classified into malignant and benign classes. Furthermore, feature selection was performed to reduce the dimensions of the feature set. Then, based on these features the colon images are classified further into six classes, corresponding to their lesion type in [17]. Similarly, the texture features were extracted via the wavelet analysis of the colonoscopy frames of the colorectal lesions detector (CoLD) system. The

TABLE I
SUMMARY OF TECHNIQUES USED IN EXISTING CADS DEVELOPED FOR GASTRIC DISEASES

Ref.	Features	GI Area	Abnormalities	Classification/Segmentation	Imaging Technology
[11]	CNN	Stomach	Cancer	ELM and SVM	WCE
[14]	Gabor Texture	Cancer	Stomach	SVM	NBI and CH
[15]	LBP	Ulcer	GI tract	SVM and ANN	WCE
[16]	NGLCM	Colon	Cancer	ANN	Colonoscopy
[17]	Fourier Features	Colon	Cancer	RDA Classifier	Zoom-Endoscopy
[21]	Log of Gabor	Ulcer	GI tract	SVM	WCE
[22]	LBP	Small bowel	Crohns disease	SVM	WCE
[23]	LBP	Stomach	Cancer	K-mean	NBI
[24]	Gabor Texture	Colon	Polyps	ANN	Colonoscopy
[25]	Fourier Features	Duodenal	Celiac disease	KNN, SVM and NB	Video Endoscopy
[26]	NGLCM	Bleeding	GI tract	SVM	WCE
[27]	Multi-Scale LBP	Small bowel	Cancer	KNN, MLP ANN, and SVM	WCE

ANN was trained and tested for classification of diminutive polyps in the abnormal frames [24]. A CoLD system is presented in [24]. The CoLD performs further wavelet analysis of the colonoscopy images , and second order statistics was conducted for the extraction of texture information. Subsequently, the ANN was used for the classification of normal and abnormal images for diminutive polyps.

C. Texture Feature Extraction Methods in the Spatial–Frequency Domain

Rotation and scale invariant texture features were proposed in [14] in which the features were extracted by using a Gabor filter? bank. In addition, using a bag-of-features approach, a texton framework was designed in their work. This texton framework was then used to classify the GI images of two complementary imaging modalities, namely, CH and NBI, into three groups using several classifiers. Furthermore, these features were used for the segmentation of the abnormal regions in gastric images by utilizing the normalized cut algorithm [28]. The contourlet transformation was performed on the images in different color spaces, and then the log-of-Gabor filter was applied to these images. The responses of the log-of-Gabor filter were used to form a feature vector. Moreover, the extracted features were used to train the SVM classifier to detect the ulcer condition in WCE images [21]. Local and global features have been extracted from the endoscopic images by using Gabor filters. By calculating the DFTs of the images, the co-occurrence matrices were derived from the transformed images. Furthermore, the SVM classifier was trained with these features to categorize the WCE images as bleeding or normal images [26]. The proposed method is also based on the Gabor wavelet analysis of images. However, the texture information is extracted from the images by calculating the geometric mean and geometric standard deviation of the responses of the filter bank instead of computing the arithmetic mean and arithmetic standard deviation. The basic concept of the Gabor method is to compute the features for an explicit scale and orientation, where the energies of different textures are concentrated in different scales and orientations [19]. Another significance of using Gabor filters instead of other filters is that they are symmetric to the human visual cortex cells. Moreover, the Gabor wavelet is also the best among the other wavelet-based features as described in [29].

III. TEXTURE FEATURE EXTRACTION FOR THE CLASSIFICATION OF GASTRIC IMAGES

A. 2D Gabor Filter

A 2D Gabor function $g(x, y)$ with spatial coordinates x and y is defined in equation 1. Let $G(u, v)$ be the Fourier transform of $g(x, y)$ as a function of frequency components u and v that are described in equation 2.

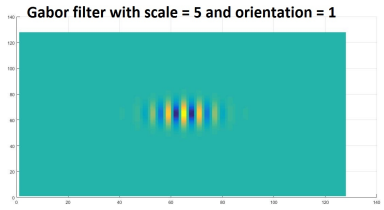
$$g(x, y, \sigma, \Omega) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) exp\left[-\frac{1}{2}\left\{\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right\} + 2\pi j\Omega x\right] \quad (1)$$

In the above equation, σ_x and σ_y are the standard deviations along the respective spatial axis. A Gabor filter is generated by the convolution of a sinusoidal with a Gaussian window where σ controls the spread of the Gaussian window. Ω represents the center frequency (the frequency of the filter resulting in a higher response).

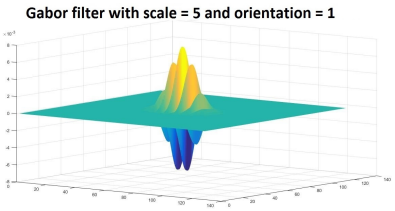
An important property of the Gabor wavelet is that it minimizes the product of its standard deviations in the time and frequency domains and reduces the uncertainty in information. Gabor filters are specifically used for the texture representation of an image. A Gabor filter, as shown in Fig. 1, achieves multiresolution spatial–frequency analysis. Moreover, because of its frequency and phase selection capability, a Gabor filter has properties similar to the human visual perception system [30]. Our aim of using the Gabor filters is to acquire texture information from the images and use it for classification. Gabor filters are related to Gabor wavelets, and filters with different scales and rotations contribute to the formation of a filter bank. These filters provide texture information images that are rotation and scale invariant. Moreover, a Gabor filter also has two important characteristics, namely, band-pass nature and direction selectivity, causing them respond only to specific spatial frequencies and orientations. Another important consideration for using Gabor filters is their capability to achieve optimum spatial–frequency localization of a signal [14]. Gabor filters have been used in various studies for texture representation of images as discussed in [31, 13].

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

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



(a)



(b)

Fig. 1. (a) 2D view of spatial kernel of Gabor filter with orientation = 0 and scale = 5 and (b) 3D view of spatial kernel of Gabor filter with orientation = 0 and scale = 5.

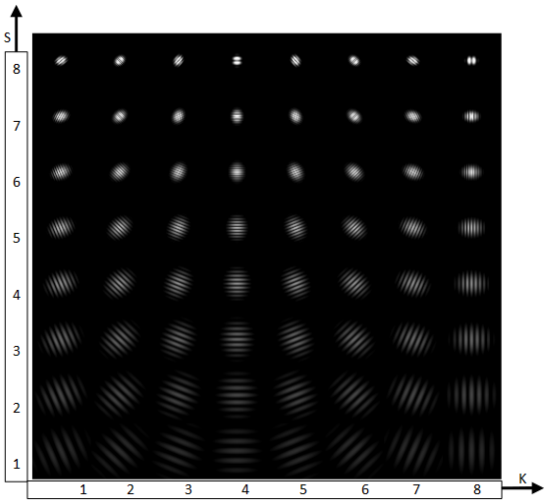


Fig. 2. Gabor filter bankthat is obtained by selecting different orientations (k) and scales (s) of Gabor filters to obtain the multiresolution responses of the CH images.

B. Designing the Gabor Filter Bank

A variation in this filter can be achieved by appropriately modifying the scale and orientation as given below.

$$\begin{aligned} \dot{\mathrm{g}}_{m,n}(x,y) &= a^{-m} g(x^{'}, y^{'} a > 1 | m, n \in Z, \\ x^{'} &= a^{-m}(xcos(\theta) + ysin(\theta)) \\ y^{'} &= a^{-m}(ycos(\theta) - xsin(\theta)) \end{aligned} \tag{3}$$

a^{-m} is the scale factor and θ represents the orientation of the filter. Subsequently, a dictionary of Gabor filters is designed to capture the responses of distinct frequencies by varying the scale and orientation. K is the number of phases in $\theta = \frac{n\pi}{K}$ and S represents the number of scales. $x^{'}$ and $y^{'}$ are the modified spatial coordinates of a Gabor filter. The new components are scaled by m, and the orientation is given by n. It is worth mentioning here that a significant amount of redundant information is present in the images. The importance of designing a filter bank is to capture different responses with diverse frequencies.

$$\begin{aligned} a = \left(\frac{\mathbf{U_h}}{\mathbf{U_l}}\right)^{\frac{1}{S-1}}, \sigma_u = \frac{(a-1)\mathbf{U_h}}{(a+1)\sqrt{(2\ln 2)}}, \\ \sigma_v = \tan(\frac{\pi}{2k}) \left[\mathbf{U_h} - 2\ln\left[\frac{2\sigma_u^2}{\mathbf{U_h}}\right]\right] \left[2\ln 2 - \frac{(2\ln 2)^2\sigma_u^2}{\mathbf{U_h^2}}\right]^{-1/2} \end{aligned} \tag{4}$$

Here, $\mathbf{U_h}$ and $\mathbf{U_l}$ are the maximum and minimum center frequencies as in Ω. In this study, values of the parameters are $\mathbf{U_h} = \mathbf{0.4}, \mathbf{U_l} = 0.05$, K = 8, and S = 8, as shown in Fig 2.

C. Texture Features Set Representation

The responses of multiple Gabor filters are obtained by filtering the CH images. Every image is filtered through numerous Gabor filters that are generated by varying the orientation and scale as mentioned earlier. The responses of these filters can be described as

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

In the above equations, B denotes the responses of a Gabor filter bank and m and n signify the scale and orientation, respectively. $B_{m,n}$ responses can be visualized as a matrix of responses of filters of

n^{th} orientation and m^{th} scale as given in equation 5. An assumption is that the responses are visually homogenous in nature. Thus, the higher order statistics of these responses are used to represents the image texture.

The arithmetic mean of a filter response in a filter bank is denoted by $\mu_{m,n}$. It can be described as

$$\mu_{m,n} = \int\int |B_{m,n}| dxdy \tag{6}$$

The arithmetic standard deviation is often denoted as $\sigma_{m,n}$. It can be defined as

$$\sigma_{m,n} = \sqrt{\int\int (|B_{m,n}(x,y)| - \mu_{m,n})^2 dxdy} \tag{7}$$

The feature set is constructed through by having all $\mu_{m,n}$ followed by $\sigma_{m,n}$ in a row vector that can be represented as

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

D. Proposed Feature Extraction

The features described in equation 8 represents a simple Gabor feature space. However, it is suitable for image-matching applications such as content-based image retrieval (CBIR) requiring only the determination of the minimum difference by comparing the distance of the features of a query image with all the features of images [14]. In contrast, the problem of classification that is not concerned with comparing features, limits the use of traditional texture features for classification of CH images.

$$\mu_{m,n}^{g} = \exp\left[\int\int \{\ln(B(x,y)_{m,n}) dxdy\}\right] \tag{9}$$

$$\sigma^g_{m,n} = \exp\left[\sqrt{\frac{1}{n}\int\int(\ln B(x,y)_{m,n} - \mu^g_{m,n})^2 dxdy}\right] \quad (10)$$

where $\mu^g_{m,n}$ denotes the geometric mean and $\sigma^g_{m,n}$ represents the geometric standard deviation of the filter responses $B(x,y)_{m,n}$. Therefore, a feature set can be formed as given in equation below

$$F^{'}_g = \left[\mu^g_0, \sigma^g_0, \mu^g_1, \sigma^g_1, \mu^g_2, \sigma^g_2, ..., \mu^g_{m,n}, \sigma^g_{m,n}\right] \quad (11)$$

The mean and standard deviation values of a single response may not differ significantly for some events (events in reference to the occurrence of abnormal regions or frames in a video). Subsequently, there exists much redundant information, and thus, some of these values can be discarded [13]. The proposed features are calculated from filter responses using the geometric mean and geometric standard deviation by assuming that the responses of the filters are a log of a normal distribution. The mean of a response is expected to be a good predictor to capture the centralizing tendency of a response specific to a class. However, the standard deviation is more informative for predicting the variations in a response from its standard value. The responses have a probability of being represented by a well-known bell-shaped normal distribution, whose sample data are, of course, symmetric around the mean. Another possibility is of a lognormal distribution, in which the logarithms of the data, and not the data, are normally distributed on a logarithmic scale, and are thus, symmetric on that scale.

1) Extraction of the LBP Features : The LBP features are extracted from the CH images for further verification and comparison of the performance achieved with the proposed features. Therefore, LBP features are used for the classification of CH images. The LBP features are defined as follows:

$$LBP = \sum_{i=0}^{M}\sum_{j=0}^{N}\left(\sum_{w=0}^{W} d(p_n - p_c) \times 2^w\right) \quad (12)$$

where W represents the neighborhood or window size, w represents the binary weight of a neighbor, p_c is the central pixel, and p_n is a neighboring pixel.

$$d_n(p) = \begin{cases} 1, & if p \geq t \\ 0, & otherwise. \end{cases} \quad (13)$$

t is a decimal value of the central pixel that is compared with the neighboring pixels, $d_n(p)$ i. The decimal version of these binaries provides the texture information locally.

E. Feature selection using a Genetic Algorithm (GA)

It is worth mentioning here that all the texture features do not advantageously contribute to the classification. Some attributes in a feature set may have similar magnitudes in some classes; consequently, these non-important features can be discarded from the feature set. The irrelevant components can only be removed by selecting salient features using an optimization method. Therefore, in the proposed method, a set of optimized features are selected through a genetic algorithm (GA). A GA is metaheuristic-based technique and an evolutionary algorithm inspired by the natural selection phenomena. In a GA, an individual has to compete to have offspring(s) and continue through to the next generations by the survival of the fittest. Initially, each chromosome represents a set of random descriptors. The corresponding accuracy of every randomly generated set is computed. Those individuals having a higher accuracy participate in the reproduction of new offspring. Therefore, a single point crossover operator is employed for generating the new descendants. The new population is formed by replacing the old low-fitness chromosomes with new offspring having a higher accuracy. Thus, new children can participate in producing better solutions in the future generations. In this study, the GA iterates through 500 generations by selecting a population of 100 chromosomes.

Algorithm 1 Genetic Algorithm for Texture Feature Selection

1) Initialize a Random Population $P = [C_1, C_2, ..., C_m]$. ▷ *randomly generating chromosomes of the size same as the feature dimension.* $C_i = [A^1, A^2, ..., A^n]$ *initializes the values of alleles* a^j *with 1 or 0 randomly for selecting and not selecting attributes, respectively.*
2) Computing the Accuracy of Selected Features. ▷ *The accuracy is computed for all the different randomly selected features in the population.*
3) Selection of the Fittest Combination from the Population. ▷ *the feature attributes having the highest accuracies are selected and single-point crossover is used in the reproduction of two offsprings.*
4) Formation of New Population ▷ *if the new offsprings have a higher fitness, then the existing chromosomes in population the chromosomes with the lowest accuracy are replaced by chromosomes have higher fitness.*
5) Repeat steps 3 to 5 for 500 generations by fixing the population size to 100.

F. Classification of CH Images into Normal and Abnormal Frames

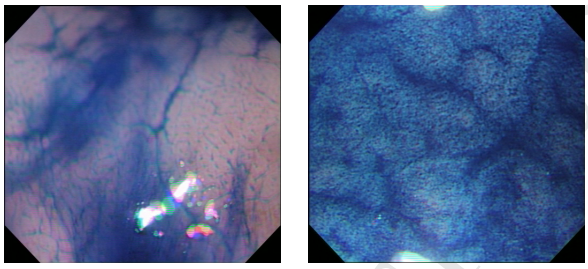
The CH images are classified into abnormal or normal classes by training various classifiers on a feature set, as given in equation 11. The reason of training multiple classifiers is to examine the performance of the proposed features and other existing texture feature sets. The SVM classifier is used for this task of binary classification with a quadratic kernel function. Subsequently, the NB classifier is trained to investigate the impact of the texture features on a probabilistic classification model. The NB classifier acquires the advantages of the class independence assumption of inter-features relationship. The NB classifier assumes that the features is not a combination of anything, but is a multivalued Gaussian distribution. Similarly, the KNN classifier with 5 neighboring elements is trained to further compare the impact of texture features on the classification accuracy. The correlation is used as a distance of measure, assuming the data are not linearly separable. In the KNN, the distance of every image feature is computed with a learned feature model to assign a label to the class having a less difference.

The extracted texture features are independent continuous variables, where the dependent variable (class) is a categorical (discrete) variable. The linear discriminant analysis (LDA) classifier assumes that the dependent variables are normally distributed; however, this is not always true. The performance of the LDA classifier is also analyzed through the training and testing on existing and proposed features.

Finally, to compare the performance of the diverse range of classifiers for the suggested features, a decision tree (DT) classifier is also trained and tested. The DT uses the sample statistics to compute the attribute interdependency (in contrast with NB where attributes are conditionally independent) based on the given data. The DT forms a tree-like structure with class labels on its leaves and attribute values on its edges. The DT classifier is also used in an ensemble method where it functions as a weak classifier. The ensemble method is employed by combining 200 weak DT classifiers. In addition, the ensemble method uses adaptable boosting (Ada-Boost) for combining the results of these weak classifiers.

G. Materials and Experimentation

1) Dataset: The proposed method was tested using the CH images of multiple patients suffering from normal and abnormal gastric conditions. The data set contained 176 CH images, and was divided into two sets, where 88 images were used for testing and the remaining images were used for training. These images can be classified according to the taxonomy given by [32]. The image dataset was part of the analysis of images to detect the abnormalities in endoscopy (AIDA-E), a biomedical challenge. The image dataset contained three groups, where the Group I included normal images and Groups II–III contained abnormal images. For simplifying the classification, we consider Group I as a negative (Normal) class, whereas images from Groups II (metaplasia) and III (dysplasia) are referred as positive (Abnormal) class, as depicted in Fig. 3.



(a) (b)

Fig. 3. (a) Shows a normal CH image and (b) is an abnormal CH image.

2) Performance measures: The performances of the proposed system and existing feature extraction methods are compared based on the following accuracy measures: sensitivity (Sen) (also known as the true positive rate (TPR)) and false positive rate (FPR) can be calculated from equation 15; specificity (Spec) (also referred to as true negative rate (TNR)) is described in equation 16; accuracy (ACC) is calculated by obtaining the true positive (TP), true negative (TN), false positive (FP), and false negative, (FN) as shown in equation 17; and area under the curve (AUC) that is computed as shown in equation 18.

$$TPR = Sensitivity = \frac{TP}{TP + FN} \tag{14}$$

$$FPR = 1 - TNR \tag{15}$$

$$TNR = Specificity = \frac{TN}{FP + TN} \tag{16}$$

$$Accuracy = \frac{TP + TN}{P + N} \tag{17}$$

$$AUC = \int_{-\infty}^{\infty} TPR(T)FPR'(T)dT \tag{18}$$

3) Experimentation: After the extraction of the features from all the images, multiple classifiers were trained on the extracted descriptors. The k-cross-validation was used for the training and testing of these classifiers, where k = 10. In a cross-validation, all the samples are divided into random k subsamples, where a single subsample is used for training and the remaining samples are used for testing. However, every time a new data is randomly selected for training and testing, the accuracies of the classifiers will fluctuate. Therefore, the average accuracies of the classifiers were calculated by repeating the experiments 100 times for all the feature sets. Moreover, the performances of the multiple classifiers discussed in the Results section are from a single iteration. All the experiments were conducted on Intel Core i3 with 2.40 GHz using MATLAB.

IV. Results and Discussions

All the above-mentioned texture features were extracted from individual CH frames. For comparing the performance of the proposed method with the existing feature extraction methods, first, every classifier was trained on the features described in equation 8. The classification results using state-of-the-art classifiers are presented in Table II.

TABLE II
CLASSIFICATION RESULTS BY USING HOMOGENOUS TEXTURE DESCRIPTORS COMPUTED USING THE MEAN AND STANDARD DEVIATION OF FILTER RESPONSES

Classifiers	TN	FN	FP	TP	Sen	Spec	ACC	AUC
SVM	32	24	11	109	82.0%	74.4%	80.1%	0.86
NB	51	5	36	84	94.4%	58.6%	76.7%	0.80
KNN	38	18	22	98	84.5%	63.3%	77.3%	0.83
LDA	38	18	44	76	80.9%	46.3%	64.8%	0.71
DT	31	25	23	97	79.5%	57.4%	72.7%	0.80
Ensemble	39	17	21	99	85.3%	65.0%	78.4%	0.84

For the further verification of the results, the experiment was repeated for 100 iterations, and the average accuracies and margin of errors were calculated with 96% confidence interval for each classifier. The SVM, NB, KNN, LDA, DT, and ensemble classifiers have average accuracies and some margin of errors of 81.0± 0.2%, 76.8± 0.1%, 76.8± 0.2%, 65.6 ± 0.5%, 72.8 ± 0.5%, and 80.3± 0.3%, respectively on the homogeneous texture (HT) features. Their corresponding AUCs are similar with values of 0.87±0.001, 0.81±0.001, 0.83±0.002, 0.72±0.005, 0.80±0.004 and 0.86±0.003, respectively. The average classification performance of these classifiers shows that the SVM classifier has a better accuracy than the other classifiers. The ensemble method also has a good average accuracy on the HT features. The textures of the CH frames may exhibit a small amount of texture difference even in frames belonging to distinct classes because two different images may have similar mucosal vascular structures. Therefore, the SVM classifier is suitable for classification

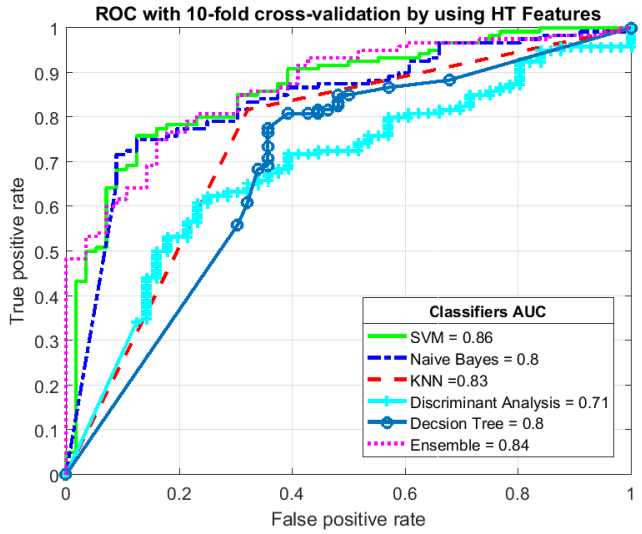


Fig. 4. ROC and AUC of multiple classifiers using homogenous texture descriptors by using the 10-cross-validation.

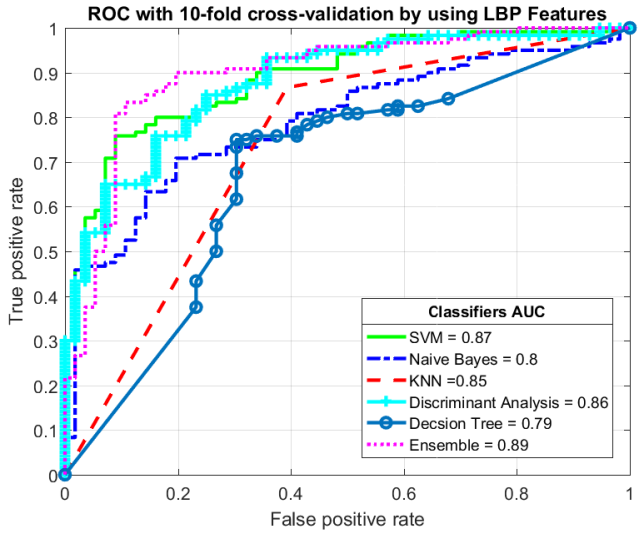


Fig. 5. ROC and AUC of multiple classifiers on the LBP features, by using a 10-cross-validation.

because the selection of the decision boundary is based on the maximum gap between the samples belonging to different classes.

The same state-of-the-art classifiers were also trained on an LBP texture feature extraction method for comparison. The SVM, NB, KNN, LDA, DT, and ensemble classifiers have average accuracies and margin of errors of 80.1± 0.2%, 71.2± 0.2%, 81.1± 0.2%, 79.0 ± 0.3%, 71.5 ± 0.5%, and 82.7± 0.3%, respectively on the LBP features; these are quite similar to the results of the HT texture features. The corresponding average AUCs of all the classifiers were also computed to have a more accurate performance. The SVM, KNN, and ensemble classifier exhibit a good performance of 0.87 ± 0.002, 0.87 ± 0.002 and 0.88± 0.002 AUC in that order, and the NB, LDA, and DT have 0.78 ± 0.002, 0.84 ± 0.003 and 0.79 ± 0.004 AUC correspondingly. It is apparent from the average results that the ensemble method will outperform other classification models in the long run. The KNN classifier has a similar accuracy and AUC as the SVM classifier. Thus, we can conclude that the SVM, KNN, and ensemble methods exhibit better classification performances on the LBP features than other classifiers. The results are shown in Table III, and Fig 5 provides us an insight of how the LBP features affect the classification of the CH images. The results are quite similar to the classification results of the HT descriptors.

TABLE III
CLASSIFICATION RESULTS OF MULTIPLE CLASSIFIERS BY USING THE LBP TEXTURE DESCRIPTORS

Classifiers	TN	FN	FP	TP	Sen	Spec	ACC	AUC
SVM	29	27	7	113	80.7%	80.6%	80.7%	0.87
NB	34	22	25	95	81.2%	57.6%	73.3%	0.80
KNN	34	22	16	104	82.5%	68.0%	78.4%	0.85
LDA	42	14	18	102	87.9%	70.0%	81.8%	0.86
DT	32	24	26	94	79.7%	55.2%	71.6%	0.79
Ensemble	40	16	12	108	87.1%	76.9%	84.1%	0.89

The results are taken from a single experiment using the LBP features, and they show that the SVM classifier achieves 80.7% accuracy and 0.87 AUC. The DT classifier exhibits the worst performance of 71.6% accuracy and 0.79 AUC, whereas the KNN classifier shows 78.1% accuracy and 0.85 AUC on the LBP features. In the endoscopic frames, there is a low inter-frame texture difference between the two classes. The features belong to two different classes having similar magnitude responses. Therefore, the results show a poor performance for similarity-based classifiers such as KNN. Another point to consider here is that the LBP features are sensitive to noise and illumination changes. The above are the main reasons for the poor performance of the LBP features [33].

By the same approach, the proposed features extracted from CH images are described in equation 11. As described in earlier sections, the geometric mean and geometric standard deviation as second order statistics of the Gabor filter? responses are used to represent the image texture. The performance of these features was evaluated by all the classifiers. The results in Table IV reveal the excellent performance of all the classifiers using the GHT features. The average accuracies of 100 iterations using 10-cross-validation and margin of error of the SVM, NB, KNN, LDA, DT, and ensemble classifiers are 86.1± 0.2%, 76.2± 0.1%, 75.1± 0.2%, 61.0 ± 0.5%, 77.5 ± 0.4%, and 83.7± 0.3%, respectively on the GHT features, whereas the corresponding average AUCs are 0.91 ± 0.001, 0.88 ± 0.001, 0.70 ± 0.003, 0.70 ± 0.006, 0.76 ± 0.005 and 0.90 ± 0.001. Here, it is apparent that the SVM and ensemble classifiers outperform the other classifiers.

According to the results shown in Table IV, on using the GHT features the SVM classifier scores 86.4% ACC and 0.90 AUC. Although, the NB classifier also exhibits a good performance with 86.4% accuracy and 0.89 AUC. The ROCs of the different classifiers by using the GHT features are presented in Fig. 6. The advantage of using the geometric mean and geometric standard deviation over the arithmetic mean and arithmetic standard deviation can be perceived from

TABLE IV
CLASSIFICATION RESULTS BY USING HOMOGENEOUS TEXTURES DESCRIPTORS COMPUTED BY THE GEOMETRIC MEAN AND GEOMETRIC STANDARD DEVIATION OF THE FILTER RESPONSES

Classifiers	TN	FN	FP	TP	Sen	Spec	ACC	AUC
SVM	47	9	15	105	92.1%	75.8%	86.4%	0.90
NB	51	5	19	101	95.3%	72.9%	86.4%	0.89
KNN	34	22	21	99	81.8%	61.8%	75.6%	0.82
LDA	33	23	43	77	77.0%	43.4%	62.5%	0.70
DT	33	23	19	101	81.5%	63.5%	76.1%	0.83
Ensemble	42	14	14	106	88.3%	75.0%	84.1%	0.88

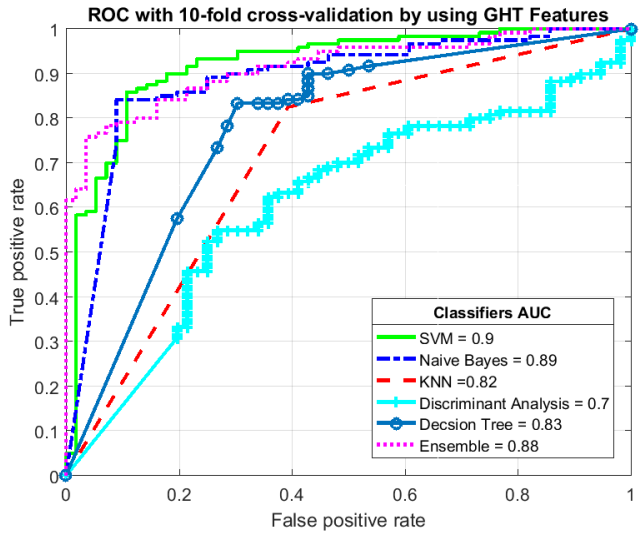


Fig. 6. ROC and AUC of multiple classifiers by using proposed geometric homogenous texture features using the 10-cross-validation. The SVM classifier has a larger AUC owing to its higher sensitivity than other classifiers.

the results. The statistical measures of the filter responses yielded high variations in The responses of the sample images for both classes. A generic feature vector of an image is obtained by calculating the mean and variance of the filter responses. The mean of the filter response provides the average location for the individual feature coefficient, and the variance yields information about its discriminative potential. If the variance of the filter responses for all the images is zero, then it is difficult to discriminate between frames related to different events. However, if the variance is high for all the frames and near zero for a single instance, then we may be able to easily distinguish images belonging to different events [13].

A genetic algorithm was used for the selection of a subset of features that can be used advantageously for classification. There are two main benefits of this type of feature selection. First is that the dimensions of the features set are reduced, and the second advantage is an increase in accuracy. Thus, only a subset of the GHT features that were important for classification were selected because some attributes may cause underfitting of the classifiers because of fewer images samples.

Consequently, a GA was used for the selection of the subset of components of the GHT vector, resulting in a good classification accuracy as illustrated by the results in Table V. The performance results confirm that the SVM classifier has an outclassed accuracy as compared with the other classifiers.

TABLE V
CLASSIFICATION RESULTS OF ALL THE CLASSIFIERS WITH AN OPTIMIZED FEATURE SELECTION USING A GENETIC ALGORITHM

Classifiers	TN	FN	FP	TP	Sen	Spec	ACC	AUC
SVM	47	9	6	114	92.7%	88.7%	91.5%	0.94
NB	50	6	24	96	94.1%	67.6%	83.0%	0.86
KNN	43	13	21	99	88.4%	67.2%	80.7%	0.85
LDA	41	15	14	106	87.6%	74.5%	83.5%	0.88
DT	44	12	19	101	89.4%	69.8%	82.4%	0.87
Ensemble	39	17	16	104	86.0%	70.9%	81.3%	0.86

The SVM classifier has 91.5% accuracy with a 0.94 AUC. The other classifiers have some improved accuracy and AUC as exhibited by the performance of all the classifiers in Fig. 7. In contrast, the mean accuracies of the SVM, NB, KNN, LDA, DT, and ensemble classifiers are 90± 0.1%, 85.3± 0.1%, 75.4 ± 0.3%, 81.2 ± 0.2%, 81.9 ± 0.4%, and 82.0± 0.3%, respectively on the GHT features selected by the GA, whereas the corresponding average AUCs are 0.93± 0.001, 0.89 ± 0.001, 0.81 ± 0.001, 0.86 ± 0.001, 0.87 ± 0.003 and 0.87 ± 0.002.

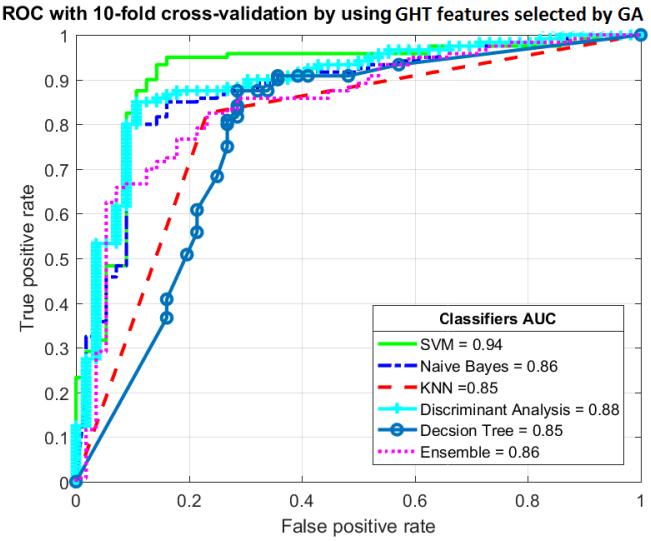


Fig. 7. ROC and AUC of multiple classifiers by using geometric texture features selected by the GA and 10-cross-validation. The performance curves show the discriminating GHT feature. The performance of SVM classifier as well as other classification models was improved by using the GA.

It is obvious from the results that the proposed features leads to a higher classification accuracy when they are used with a combination of the SVM classifier and GA optimization algorithm. The performances of other classifiers are also improved and are validated against optimally selected features. The ROC of the classifiers with the proposed features selected by the GA is shown in Fig. 7.

Table VI shows the average performance of the proposed system, where the results exhibit that the GHT features have a better discrimination potential than the traditional texture extraction methods. Furthermore, numerous attributes are less helpful and may instead have an undesirable effect on the classification of images. Therefore, the selection of important features with the GA confirms a higher performance instead

TABLE VI
COMPARISON OF THE AVERAGE ACCS AND AUCS WITH EXISTING FEATURE EXTRACTION METHODS

Classifier	Descriptors	ACC	AUC
SVM	HT	81.0%	0.86
Ensemble	LBP	82.7%	0.88
SVM	**GHT**	**86.1%**	**0.91**
SVM	**GHT with GA**	**90.0%**	**0.93**

of using the entire dimension of a feature vector.

V. CONCLUSION AND FUTURE DIRECTIONS

In the presented work, the texture features exhibited a good discrimination performance on the C) images. These features were HTs, LBPs, and proposed GHTs. The experimental results showed the significance of the GHT features with and without feature selection. Furthermore, the SVM classifier was found as the classifier with the best performance among other state-of-the-art classifiers using all the texture features. We elicited texture information from the images for the classification of cancerous images in abnormal and normal classes. Nevertheless, the texture descriptors have their own significance. However, the colors may also have some contribution in the detection of the cancerous regions specifically when using the CH images. In the future, it would be interesting to have design descriptors that include color information as well as the texture and shape. Moreover, because this method supports image level (binary) classification, it would be fascinating to include lesion level (multilevel) classes of abnormal regions (e.g. metaplasia and dysplasia). Though the proposed method is limited to CH images,

it could be generalized by using images obtained from diverse imaging modalities under several unhealthy conditions.

ACKNOWLEDGMENT

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

REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2015," *CA: a cancer journal for clinicians*, vol. 65, no. 1, pp. 5–29, 2015.

[2] R. Miyahara, Y. Niwa, T. Matsuura, O. Maeda, T. Ando, N. Ohmiya, A. Itoh, Y. Hirooka, and H. Goto, "Prevalence and prognosis of gastric cancer detected by screening in a large Japanese population: Data from a single institute over 30 years," *Journal of Gastroenterology and Hepatology (Australia)*, vol. 22, no. 9, pp. 1435–1442, 2007.

[3] R. Kiesslich and M. F. Neurath, "Endomicroscopy is born-do we still need the pathologist?," *Gastrointestinal Endoscopy*, vol. 66, no. 1, pp. 150–153, 2007.

[4] R. Swannell, "World Cancer Report 2014," *The Globe*, no. 1, pp. 6–7, 2010.

[5] X. Qi, *Computer-aided diagnosis of early cancers in the gastrointestinal tract using optical coherence tomography*. PhD thesis, Case Western Reserve University, 2008.

[6] M. Pennazio, "Capsule endoscopy: Where are we after 6 years of clinical use?," *Digestive and Liver Disease*, vol. 38, no. 12, pp. 867–878, 2006.

[7] M. B. Wallace and R. Keisslich, "Advances in Endoscopic Imaging of Colorectal Neoplasia," *Gastroenterology*, vol. 138, no. 6, pp. 2140–2150, 2010.

[8] M. Liedlgruber and A. Uhl, "Computer-aided decision support systems for endoscopy in the gastrointestinal tract: A review," *IEEE Reviews in Biomedical Engineering*, vol. 4, pp. 73–88, 2011.

[9] S. McGill, R. Soetikno, and T. Kaltenbach, "Image-enhanced endoscopy in practice," *Canadian Journal of Gastroenterology*, vol. 23, no. 11, p. 741, 2009.

[10] T. Nagahama, K. Yao, S. Maki, M. Yasaka, Y. Takaki, T. Matsui, H. Tanabe, A. Iwashita, and A. Ota, "Usefulness of magnifying endoscopy with narrow-band imaging for determining the horizontal extent of early gastric cancer when there is an unclear margin by chromoendoscopy (with video)," *Gastrointestinal endoscopy*, vol. 74, no. 6, pp. 1259–1267, 2011.

[11] J.-s. Yu, J. Chen, Z. Xiang, and Y.-X. Zou, "A hybrid convolutional neural networks with extreme learning machine for wce image classification," in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1822–1827, 2015.

[12] F. Riaz, A. Hassan, R. Nisar, M. Dinis-Ribeiro, and M. Coimbra, "Content-adaptive region-based color texture descriptors for medical images," *IEEE journal of biomedical and health informatics*, 2015.

[13] M. T. Coimbra and J. S. Cunha, "MPEG-7 visual descriptorscontributions for automated feature extraction in capsule endoscopy," *IEEE Transactions on Circuits and Systems For Video Technology*, vol. 16, no. 5, pp. 628–637, 2006.

[14] F. Riaz, F. B. Silva, M. D. Ribeiro, and M. T. Coimbra, "Invariant gabor texture descriptors for classification of gastroenterology images," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2893–2904, 2012.

[15] B. Li and M. Q.-H. Meng, "Texture analysis for ulcer detection in capsule endoscopy images," *Image and Vision Computing*, vol. 27, no. 9, pp. 1336–1342, 2009.

[16] G. D. Magoulas, V. P. Plagianakos, and M. N. Vrahatis, "Neural network-based colonoscopic diagnosis using online learning and differential evolution," *Applied Soft Computing*, vol. 4, no. 4, pp. 369–379, 2004.

[17] M. Hafner, L. Brunauer, H. Payer, R. Resch, A. Gangl, A. Uhl, F. Wrba, and A. Vécsei, "Computer-aided classification of zoom-endoscopical images using fourier filters," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 4, pp. 958–970, 2010.

[18] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *Journal of the Optical Society of America A*, vol. 2, no. 7, pp. 1160–1169, 1985.

[19] T. S. Lee, "Image representation using 2d gabor wavelets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 959–971, 1996.

[20] B. S. Manjunath and W.-Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, no. 8, pp. 837–842, 1996.

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

[22] A. F. Constantinescu, M. Ionescu, I. Rogoveanu, M. E. Ciurea, C. T. Streba, V. F. Iovanescu, S. A. Artene, and C. C. Vere, "Analysis of wireless capsule endoscopy images using local binary patterns," *Applied Medical Informatics*, vol. 36, no. 2, p. 31, 2015.

[23] T.-C. Lee, Y.-H. Lin, N. Uedo, H.-P. Wang, H.-T. Chang, and 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)*, pp. 4430–4433, 2013.

[24] D. E. Maroulis, D. K. Iakovidis, S. A. Karkanis, D. A. Karras, D. E. Maroulis, D. K. Iakovidis, S. A. Karkanis, and D. A. Karras, "Cold: a versatile detection system for colorectal lesions in endoscopy video-frames," *Computer Methods and Programs in Biomedicine*, vol. 70, no. 2, pp. 151–166, 2003.

[25] A. Vécsei, T. Fuhrmann, M. Liedlgruber, L. Brunauer, H. Payer, and A. Uhl, "Automated classification of duodenal imagery in celiac disease using evolved fourier feature vectors," *Computer Methods and Programs in Biomedicine*, vol. 95, no. 2, pp. S68–S78, 2009.

[26] A. R. Hassan and M. A. Haque, "Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos," *Computer Methods and Programs in Biomedicine*, vol. 122, no. 3, pp. 341–353, 2015.

[27] B. Li, M. Q.-H. Meng, and J. Y. Lau, "Computer-aided small bowel tumor detection for capsule endoscopy," *Artificial Intelligence in Medicine*, vol. 52, no. 1, pp. 11–16, 2011.

[28] F. Riaz, F. B. Silva, M. D. Ribeiro, and M. T. Coimbra, "Impact of visual features on the segmentation of gastroenterology images using normalized cuts," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 5, pp. 1191–1201, 2013.

[29] W.-Y. Ma and B. Manjunath, "A comparison of wavelet transform features for texture image annotation," in *International Conference on Image Processing*, vol. 2, pp. 256–259, IEEE, 1995.

[30] S. Marĉelja, "Mathematical description of the responses of simple cortical cells," *Journal of the Optical Society of America A*, vol. 70, no. 11, pp. 1297–1300, 1980.

[31] M. Häfner, R. Kwitt, A. Uhl, F. Wrba, A. Gangl, and A. Vécsei, "Computer-assisted pit-pattern classification in different wavelet domains for supporting dignity assessment of colonic polyps," *Pattern Recognition*, vol. 42, no. 6, pp. 1180–1191, 2009.

[32] A. Sousa, M. Dinis-Ribeiro, M. Areia, and M. Coimbra, "Identifying cancer regions in vital-stained magnification endoscopy images using adapted color histograms," in *16th IEEE International Conference on Image Processing (ICIP)*, pp. 681–684, 2009.

[33] B. Li, H. Jin, C. Yang, and G. Xu, "A novel color textural feature towards capsule endoscopy video summary," in *IEEE International Conference on Automation and Logistics*, pp. 766–769, 2015.

Highlights

1. A new method of designing geometric homogeneous textures (GHTs) is proposed for texture feature extraction.
2. Chromoendoscopy frames are analyzed via multiple classifiers trained on proposed feature extraction method.
3. The Genetic Algorithms is used to select an optimized set of texture features.