Received: 16 November 2019 | Revised: 28 December 2019 | Accepted: 13 January 2020
DOI: 10.1002/jemt.23447

RESEARCH ARTICLE

MICROSCOPY RESEARCH & TECHNIQUE WILEY

# Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection

Abdul Majid[1] | Muhammad Attique Khan[2] | Mussarat Yasmin[1] | Amjad Rehman[3] | Abdullah Yousafzai[2] | Usman Tariq[4]

[1]Department of Computer Science, COMSATS University Islamabad, Wah Campus, Wah Cantt, Pakistan

[2]Department of Computer Science, HITEC University Museum Road, Taxila, Rawalpindi, Pakistan

[3]AIDA Lab CCIS, Prince Sultan University Riyadh, Riyadh, Saudi Arabia

[4]College of Computer Engineering and Science, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

**Correspondence**
Muhammad Attique Khan, Department of Computer Science, HITEC University Museum Road, Taxila, Rawalpindi, Pakistan.
Email: attique@ciitwah.edu.pk

Review Editor: Peter Saggau

## Abstract

Automated detection and classification of gastric infections (i.e., ulcer, polyp, esophagitis, and bleeding) through wireless capsule endoscopy (WCE) is still a key challenge. Doctors can identify these endoscopic diseases by using the computer-aided diagnostic (CAD) systems. In this article, a new fully automated system is proposed for the recognition of gastric infections through multi-type features extraction, fusion, and robust features selection. Five key steps are performed—database creation, handcrafted and convolutional neural network (CNN) deep features extraction, a fusion of extracted features, selection of best features using a genetic algorithm (GA), and recognition. In the features extraction step, discrete cosine transform, discrete wavelet transform strong color feature, and VGG16-based CNN features are extracted. Later, these features are fused by simple array concatenation and GA is performed through which best features are selected based on K-Nearest Neighbor fitness function. In the last, best selected features are provided to Ensemble classifier for recognition of gastric diseases. A database is prepared using four datasets—Kvasir, CVC-ClinicDB, Private, and ETIS-LaribPolypDB with four types of gastric infections such as ulcer, polyp, esophagitis, and bleeding. Using this database, proposed technique performs better as compared to existing methods and achieves an accuracy of 96.5%.

KEYWORDS

CNN features, database preparation, features selection, gastric infections, handcrafted features

## 1 | INTRODUCTION

In medical image processing, gastrointestinal (GI) tract infections detection and recognition is an active research area nowadays (Khan, Rashid, Sharif, Javed, & Akram, 2019; Khan, Sharif, Akram, Yasmin, & Nayak, 2019). GI tract infections include polyps, ulcers, bleeding, and esophagitis among which ulcer and bleeding are very common (Sharif et al., 2019). In United States, since 2019, about 22% of adult population is diagnosed with gastric conditions. Estimated new cases of stomach cancer are 27,510 and deaths are 11,140 which is 1.6 and 1.8% of all cancer cases and deaths, respectively (cancer.net, 2019). Likewise, 319,160 new cases of gastric cancer are diagnosed and 160,820 deaths have occurred during 2018 (Siegel, Miller, & Jemal, 2016). The occurrence of gastric cancer worldwide exceeds 1 million in 2017 making it a third leading cause of human deaths (Lee et al., 2019). Colorectal cancer affects both sexes, men and women. This is also known as bowel cancer. In developing countries, bowel cancer causes 694,000 deaths on an average (Siegel et al., 2017). Some GI

  

diseases such as short bowel and hemorrhoids are directly linked with colorectal cancer (Ribeiro et al., 2019). The seventh common cancer disease in mature persons is esophageal cancer (Ghatwary, Ye, & Zolgharni, 2019a). The enhanced procedure to detect and treat esophageal diseases can improve the chance of survival from 19 to 80% (Menon & Trudgill, 2014). Therefore, the human mortality rate (Fu, Zhang, Mandal, & Meng, 2013) can decrease if GI tract diseases could be detected and provided treatment at an early stage.

WCE imaging is a medical imaging technique which is widely used to detect and recognize GI diseases (Iddan, Meron, Glukhovsky, & Swain, 2000). In this technique, a tiny camera captures color images of GI region. The dimension of this tiny camera is around $11 \times 30$ mm$^2$ and requires no external wires or cables. This camera is swallowed by the patient to capture the video images. These video images are transferred to the external recorder using radio telemetry (Iddan et al., 2000). To locate the capsule in GI tract, an array of eight aerials is attached to the body (Swain, 2003). However, this task still has many challenges such as time consuming and availability of an expert (Liaqat et al., 2018). Therefore, several automated CAD systems are developed by computer vision (CV) researchers to solve this issue (Iakovidis, Georgakopoulos, Vasilakakis, Koulaouzidis, & Plagianakos, 2018). The effective use of WCE technology along with CV can reduce the cost of overall healthcare systems as well as time for the detection and recognition of disease (Vasilakakis et al., 2019).

Recently, many developed CAD systems significantly help physicians in their clinics (Akram, Khan, Sharif, & Yasmin, 2018; Nasir et al., 2018; Saba, Khan, Rehman, & Marie-Sainte, 2019; Sharif, Tanvir, Munir, Khan, & Yasmin, 2018). The majority of CAD systems for GI tract diseases detection and classification are based on supervised learning algorithms (Khan, Sharif, Akram, Bukhari, & Nayak, 2019; Sharif, Li, Khan, & Saleem, 2020). These CAD systems extract features from WCE images to detect GI diseases automatically (Ghatwary, Ye, & Zolgharni, 2019b). In different CAD systems (De Souza, Afonso, Palm, & Papa, 2017; Münzer, Schoeffmann, & Böszörmenyi, 2018), handcrafted features including shape, texture, and color are taken out to detect abnormalities. Moreover, in Rajaei and Rangarajan (2011), discrete wavelet transform (DWT) and discrete cosine transform (DCT) features are extracted for the classification of these diseases. Color feature is another type of descriptor which computes color based information of gastric infections in WCE images (Li & Meng, 2009). The more recent development in the area of deep learning shows well performance for many CAD systems which previously not performed well through classical features. An important deep learning model is convolutional neural network (CNN) used for high level features extraction from raw images (Khan et al., 2018; Khan, Javed, Sharif, Saba, & Rehman, 2019).

However, due to WCE image characteristics, all computed features are not useful; therefore, the selection of best features is important. Features selection algorithms convert high-dimensional data into smaller dimensions (Pontabry, Rousseau, Studholme, Koob, & Dietemann, 2017). These algorithms help to eliminate irrelevant features and keep useful features (Kishore, 2015). An appropriate features selection method selects robust features and significantly enhances performance of the system. A lot of techniques are newly introduced to select best subgroup of features from an originally obtained vector including grass hopper, genetic algorithm (GA), entropy selection, and Particle Swarm Optimization (Khan, Lali, et al., 2019; Rashid et al., 2019). These techniques performed satisfactorily in both time computation and accuracy performance. In this study, GA is used for the selection of robust features. Generally, GA is an optimization method introduced by Woody Bledsoe (Bledsoe, 1961) and formalized mathematically by John (1992). Several researchers have shown the importance of GA in solving the dimensionality reduction and features selection problems (Ghareb, Bakar, & Hamdan, 2016; Günal, 2012; Uğuz, 2011; Uysal & Gunal, 2014). GA reduces the dimensionality of feature set and enhances the system's efficiency. As GA reduces the dimensionality, computational time is also reduced.

## 2 | RELATED WORK

In WCE imaging, a lot of work has been done to detect and classify GI diseases (Cogan, Cogan, & Tamil, 2019; Ghatwary et al., 2019b; Lee et al., 2019; Liaqat et al., 2018). Many of them are focused on supervised learning methods. In this learning, several deep learning-based pretrained models are available. Lee et al. (2019) used the most recent models named as Inception V3, VGG16, and ResNet-50 for the classification of ulcer and normal classes of GI images. From these deep networks, ResNet-50 showed the highest level of performance. A novel saliency-based segmentation method is proposed to detect and classify GI infections (Khan, Rashid, Sharif, Javed, & Akram, 2019) based on YIQ colorspace. Before saliency-based segmentation, active contour-based segmentation is implemented (Chan & Vese, 2001) based on HSI color transformation. This technique improved the segmentation and classification results. A superpixels segmentation-based CAD method (Fu et al., 2013) was presented to detect the bleeding area. This method achieved 95% classification accuracy on support vector machine (SVM) classifier.

A technique (Ghatwary et al., 2019b) is proposed based on the fusion of handcrafted Gabor features and faster region-based CNN (Ren, He, Girshick, & Sun, 2015) which detects esophageal infections in WCE images at an early stage. It is tested on Kvasir and Endo Vis MICCAI 2015 challenge dataset with an achieved recall rate as 95% and precision rate as 91%. For the first time, Fan, Xu, Fan, Wei, and Li (2018)) used AlexNet model to detect erosion and ulcer in WCE images. A novel technique (Iakovidis et al., 2018) is introduced for anomaly detection and localization using a deep CNN model. This model is used for the detection of salient region and iterative cluster unification (ICU). The deep saliency detection method detects salient regions of GI anomalies whereas to localize the anomalies, ICU algorithm is used. In this step, pointwise-cross-feature map is computed for localization. Later, color wavelet features are merged with CNN features to perform classification through SVM (Billah, Waheed, & Rahman, 2017).

Color features have an important role in agriculture and medical imaging field (Akram et al., 2018; Sharif, khan, et al., 2018). In medical imaging, color features calculate the information of infected regions. A fusion of color, HOG, and SURF features-based gastric infections

detection and classification is presented in (Liaqat et al., 2018). In this method, HSV color transformation is performed for the next segmentation process. HSV color space has resistance to illumination changes (Chen & Lee, 2012) which helps in WCE imaging for good results. Deeba, Islam, Bui, and Wahid (2018) introduced a technique based on color features and fusion of classifier to detect bleeding areas in WCE images. This method achieves an average accuracy and sensitivity of 95, and 94%, respectively. An approach is introduced based on statistical color features (Suman et al., 2017) for bleeding detection. These features are extracted from RGB images that are later classified through SVM in their relevant categories. Charfi and El Ansari (2018) discussed a method based on texture features for abnormalities detection in WCE images. Local binary pattern (LBP) and DWT features are extracted and then these features are given to SVM and multilayer perceptron for classification. LBP and DWT features were used to overcome the illumination issues.

The above studies show that CNN models increase the scale and abstract level for automatic GI tract diseases detection (Yi, Wu, Hoeppner, & Metaxas, 2017). However, the concatenation of deep CNN features with handcrafted features can enhance the performance of system (Hosseini, Lee, & Cho, 2018). The performance of Gabor features is increased after fusion with deep CNN features (Shi et al., 2018). Different studies (Chen et al., 2017; Kwolek, 2005; Luan, Chen, Zhang, Han, & Liu, 2018; Yao, Chuyi, Dan, & Weiyu, 2016) have highlighted the advantages of fusion of both these features. Therefore, in this work, CNN features are combined with color and statistical descriptors (DCT and DWT). Moreover, focus is also on best features selection to minimize the problem of computational time.

## 3 | PROBLEM STATEMENT AND CONTRIBUTION

The major challenges of GI tract diseases classification is similar patterns of infected regions. For example, low contrast ulcer images color is similar to polyp making it difficult to categorize. On the other side, low contrast ulcer regions have high similarity among healthy regions; therefore, it is not easy to classify them correctly. For correct classification, the selection of useful features is important and for that best features selection algorithm is focused in this article. Robust features produce more accurate classification results and enhance the overall efficiency of CAD system. In this work, a new CAD system is proposed based on fusion and selection of deep CNN features and well known handcrafted features. Major contributions are as follows:

- Three types of features including DCT, DWT, and strong color features (SCFs) are extracted.
- A methodology is introduced for SCFs selection which defines a threshold value to select the most discriminant features.
- CNN features are extracted through pretrained model VGG16.
- All features are fused and GA is performed for the selection of robust features.
- A new database is prepared which consists of 9,889 images of GI diseases including polyp, bleeding, esophagitis, and ulcer as well as healthy class.

## 4 | PROPOSED METHODOLOGY

The proposed GI tract diseases classification method is presented in this section with detailed mathematical expressions and visual representation. Figure 1 shows flow architecture of the proposed methodology which incorporates following steps: (a) database preparation; (b) extraction of handcrafted features, namely, DCT (Ahmed, Natarajan, & Rao, 1974), DWT (Mallat, 1989), and color; (c) VGG 16-based CNN features calculation; (d) fusion of CNN and handcrafted descriptors; (e) GA-based best features selection; and (f) classification to obtain labeled data.

### 4.1 | Database preparation

In this work, two datasets Kvasir and Private are utilized for the preparation of a new database for experimental process. The database is constructed from the combination of these two listed datasets. The prepared database includes a total of 9,889 RGB images of four stomach infections (bleeding, esophagitis, polyp, and ulcerative-colitis) along with one healthy class. Bleeding and healthy images (each class contains 2000 RGB images) are collected from private datasets (Khan, Rashid, Sharif, Javed, & Akram, 2019) whereas esophagitis and ulcerative-colitis classes are taken from publicly available dataset Kvasir (Pogorelov et al., 2017). Initially, both classes contain 1,000 images which is later increased up to 2000 by augmentation method. The polyp images are collected from Kvasir, CVC-ClinicDB (Bernal et al., 2015), and ETIS-LaribPolypDB (Silva, Histace, Romain, Dray, & Granado, 2014). The sample images of this combined database are shown in Figure 2.

### 4.2 | Handcrafted features extraction

In image processing, features are extracted to represent the information of an image and the objects. Features extraction is an important step in the area of pattern recognition (PR) and machine learning (Rashid et al., 2019). In medical field, these extracted features are fed to classifier for the classification of selected diseases (Khan et al.). In this study, three types of handcrafted features including DCT (Ahmed et al., 1974), DWT (Mallat, 1989), and SCF are extracted from the original images. The working of each feature type is described below.

#### 4.2.1 | DCT features

The most significant information of image data can be represented in just a few coefficients by employing DCT (Ahmed et al., 1974);

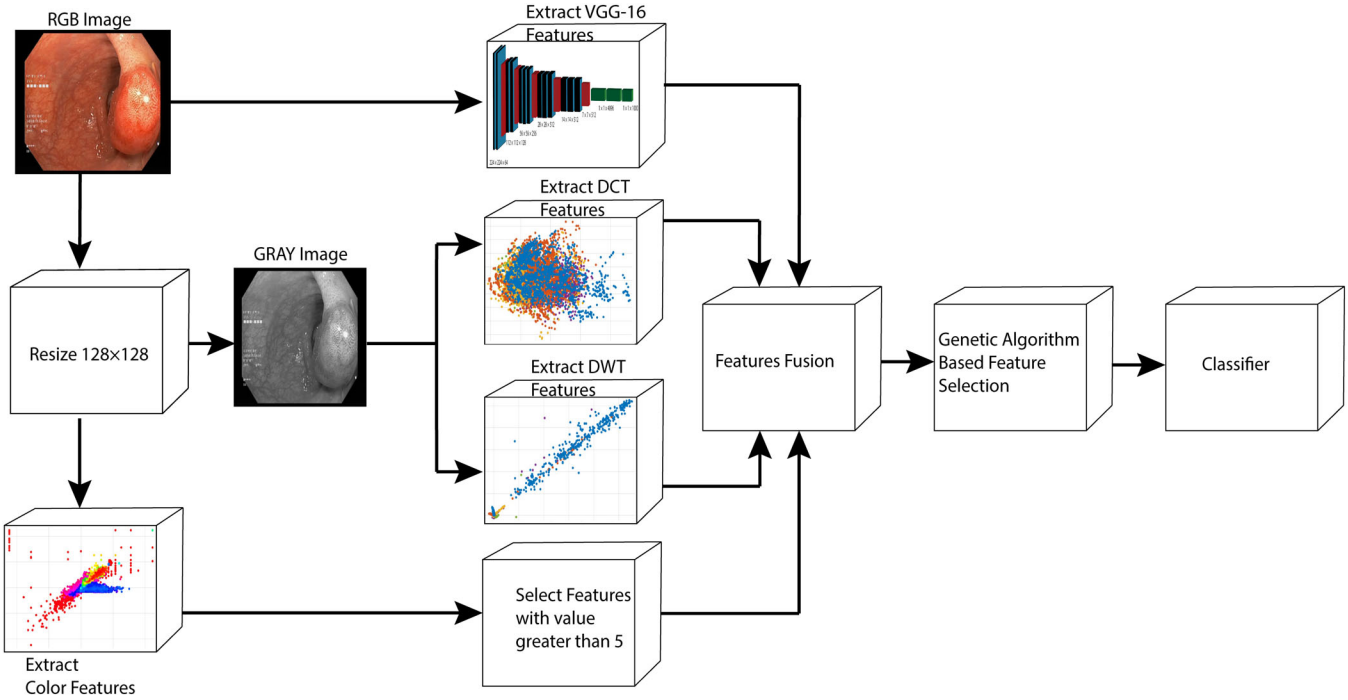


**FIGURE 1** The proposed methodology for the classification of gastrointestinal (GI) diseases [Color figure can be viewed at wileyonlinelibrary.com]

therefore, in this work, these features are computed with an objective of high accuracy results in low time complexity. Mathematically, the coefficients of an image of size $M \times N$ are expressed as follows:

$$C(u,v) = \frac{1}{\sqrt{MN}}\alpha(u)\alpha(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}A(u,v)\times\cos\left(\frac{(2x+1)u\pi}{2M}\right)\times\cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad (1)$$

where $u = 0, 1, 2, \ldots, M$, and $v = 0, 1, 2, \ldots, N$.

$$\alpha(\omega) = \begin{cases} \frac{1}{\sqrt{2}}, \omega = 0 \\ 1, \text{otherwise} \end{cases} \quad (2)$$

The matrix $A(x, y)$ shows the image intensity and $C(a, b)$ is a 2D matrix of DCT coefficients. The number of DCT features is equal to the image size and final obtained vector of DCT features is $N \times 128$ in length.

### 4.2.2 | DWT features

DWT is a powerful technique for features extraction from different directions and scales (Mallat, 1989). In this work, 2D DWT is used to extract wavelet coefficients from WCE images. The wavelet localizes frequency information which helps to produce good classification results. To calculate DWT, the input signal is passed from low and high-pass filters. For a signal $a$ and low-pass filter $l$, DWT is defined as:

$$b_{\text{low}}[m] = (a \times l)[m] = \sum_{z=-\infty}^{\infty} a[z]l[m-z] \quad (3)$$

The signal is further decomposed using a high-pass filter $h$. The output after high-pass filter is the detail of coefficients. The output is calculated as:

$$b_{\text{high}}[m] = \sum_{z=-\infty}^{\infty} a[z]h[m-z] \quad (4)$$

The above expression returns four different types of feature details such as coefficients approximation (coef-A) as well as horizontal (coef-H), vertical (coef-V), and diagonal (coef-D) coefficients. All four coefficients are fused to obtain the final $N \times 256$ DWT feature vector.

### 4.2.3 | Strong color features

The color features are most important to detect and classify GI diseases from RGB images. In this work, three types of color spaces are used for the extraction of color features. Color spaces are RGB, HSV, and LAB. These color spaces are separated into number of channels to compute five parameters like mean, variance, *SD*, skewness, and kurtosis. This process is completed for all nine channels of these three

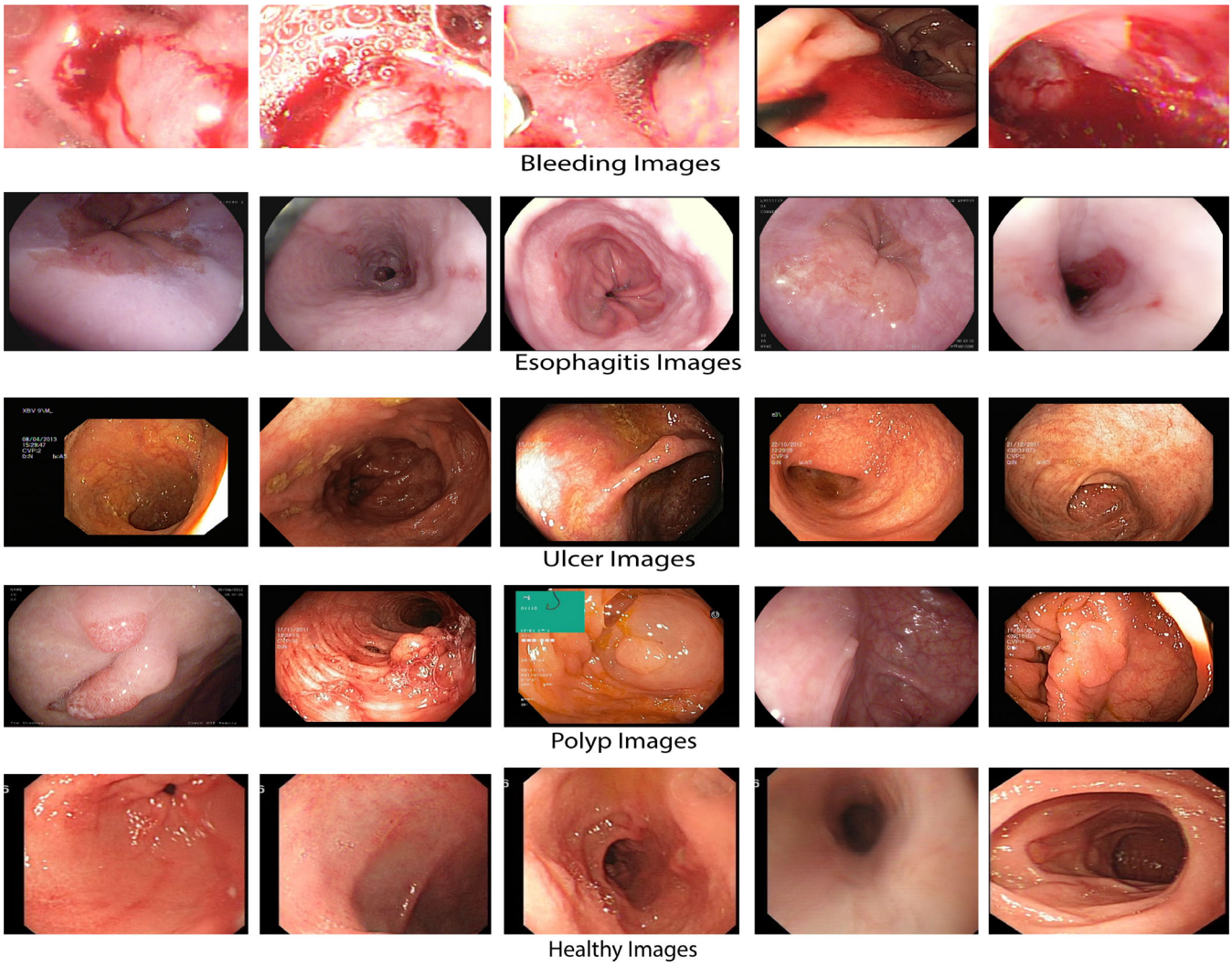


**FIGURE 2** Sample wireless capsule endoscopy (WCE) images of combined database [Color figure can be viewed at wileyonlinelibrary.com]

color spaces. After computing these parameters of each channel, they are combined using a serial-based approach to obtain a vector of dimension $N \times 6{,}500$. The combined vector contains many redundant features; therefore, it is necessary to remove these redundant features to obtain an SCF vector.

To handle this problem, a threshold function is defined based on mean value for the selection of SCFs. Color features equal or greater than the mean value are selected and remaining are eliminated from final classification. This function reduces approximately 70% irrelevant color features. Mathematically, it is defined as follows:

$$T = \begin{cases} \xi_{sc}(i) \ if & fv(i) \geq \mu \\ \text{Eliminate elsewhere} \end{cases} \tag{5}$$

where $\xi_{sc}(i)$ denotes strong selected color features of dimension $N \times 809$ (prepared dataset in this work) and $fv(i)$ represents original color vector of dimension $N \times 6{,}500$.

Finally, all handcrafted features (DWT, DCT) and SCF are combined through a simple expression given as:

$$\xi_{cd}(i) = \begin{pmatrix} \alpha(\omega)_{M\times N} \\ b_{high}[n]_{M\times N} \\ \xi_{sc}(i)_{M\times N} \end{pmatrix} \tag{6}$$

where $\xi_{cd}(i)$ denotes combined handcrafted features and $M \times N$ represents the dimension of these features. This equation shows that the features are combined in one matrix based on the size of each vector.

## 4.3 | Modified CNN features

The modified CNN features are extracted for high-level image information. For this purpose, per-trained CNN model VGG-16 is used. It

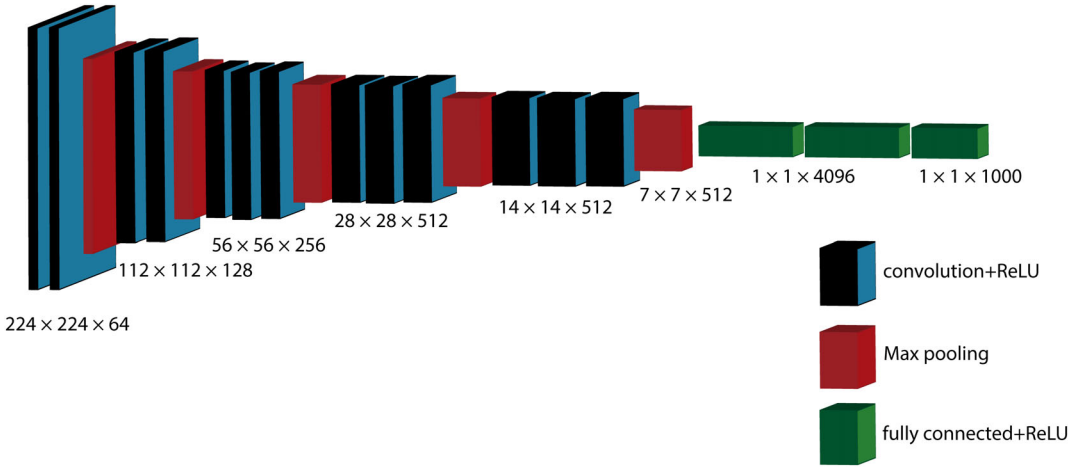


**FIGURE 3** A comprehensive architecture of VGG16 pretrained convolutional neural network (CNN) model [Color figure can be viewed at wileyonlinelibrary.com]

is a sequential model with a simple architecture. This model is originally trained on ImageNet database which consists total 1.28 million images (Deng et al., 2009). A complete architecture of VGG16 is shown in Figure 3 which describes that this model contains one input layer, 13 convolutional layers in five segments, five pooling layers and three fully connected (FC) layers. Max pooling is applied in the pooling phase. The image input layer size is 224 × 224 × 3 and it accepts RGB image. Each convolution segment is followed by max pooling layer and all hidden layers use rectified linear unit as an activation function. The window size of max pooling layer is 2 × 2. For deep features extraction, initially, this model is trained on prepared database of this work by using the concept of transfer learning (TL) which utilizes previously learned knowledge and applies this to new data hence making it easy to learn a model as compared to

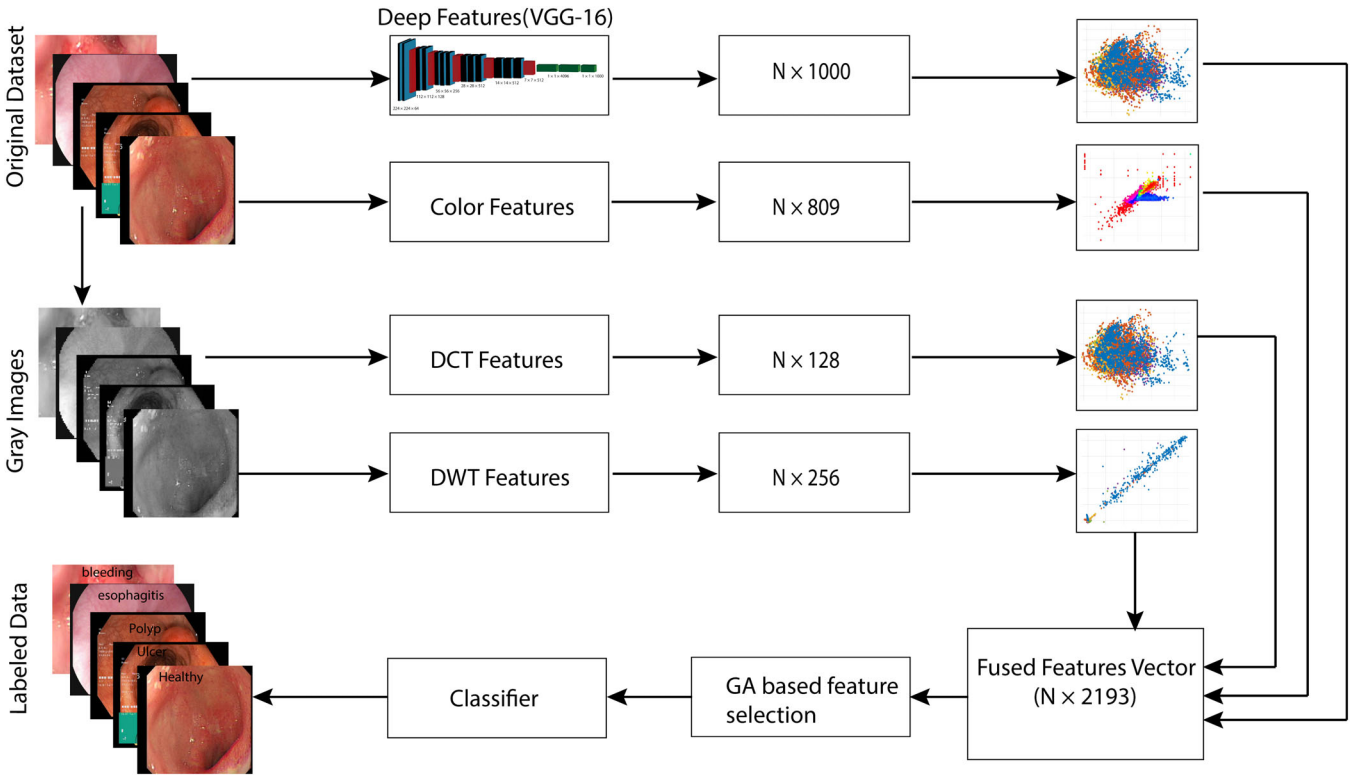


**FIGURE 4** The proposed strategy for fusion and selection of features [Color figure can be viewed at wileyonlinelibrary.com]

newly generated model from scratch. During TL, 50% of data are selected for training and remaining for testing. By using TL, features are extracted on FC layer (FC-1000) where cross entropy function is applied as an activation function which returns a vector of dimension $N \times 1{,}000$ as output, where $N$ denotes the number of training and testing images.

## 4.4 | Features fusion

Features fusion is an important step of PR field in which different features vectors are concatenated to obtain final feature vector for classification (Khan, Akram, et al., 2019; Khan, Lali, et al., 2019; Khan, Javed, Sharif, Saba, & Rehman, 2019; Khan, Rashid, Sharif, Javed, & Akram, 2019; Khan, Sharif, Akram, Bukhari, & Nayak, 2019; Khan, Sharif, Raza, Anjum, et al., 2019). Main motivation of this step is to put all the information of different descriptors in a single feature vector which may be helpful for minimum error rate. In this work, Equation (6) is used for fusion of both features types (handcrafted and CNN) to obtain a new vector of dimension $N \times 2,193$. This process is also depicted in Figure 4 which shows internal facts behind this step. In this figure, it is shown that color and deep features are extracted from original RGB images whereas DCT and DWT required gray images. Mathematically, fusion process is defined as:

$$\xi_{fd}(i) = \begin{pmatrix} \alpha(\omega)_{M\times N} \\ b_{\text{high}}[n]_{M\times N} \\ \xi_{\text{sc}}(i)_{M\times N} \\ \xi_{\text{dp}}(i)_{M\times N} \end{pmatrix} \tag{7}$$

where $\xi_{\text{fd}}(i)$ represents final fused vector of dimension $M \times N$ and $\xi_{\text{dp}}(i)_{M\times N}$ is VGG16-based deep feature vector.

## 4.5 | GA-based features selection

GA is a nature inspired approach and mostly used for features optimization. In this work, GA is used for the selection of best features from fused feature vector. The fused vector $\xi_{fd}(i)$ is given as input and this algorithm returns best features as output in the form of sets of solutions known as population. Each solution known as chromosome consists of genes that represent possible solution for the given issue. After each iteration, best solutions are generated and evaluated on the basis of their fitness function. The main architecture of GA is presented in Figure 5 and each step is described briefly below.

### 4.5.1 | Initialization

In this step, number of chromosomes is defined as 10, number of total generations is 100, mutation rate is 0.01, and crossover rate is taken as 0.8. This algorithm has run over total number of generations which shows that 100 iterations are performed in this work. Following parameters are fed to GA-fused feature vector of dimension $M \times N$, labeled data against fused vector, number of chromosomes, population size, mutation rate, and crossover rate. After passing these parameters, initial population is generated based on number of chromosomes and dimension of fused vector.

**ALGORITHM 1**

**Algorithm for GA-based Features Selection**

*Output*: $\boldsymbol{SF(i)} \leftarrow$ ***Selected Vector***
*Input*: $\boldsymbol{\xi_{fd}(i)} \leftarrow$ ***Fused Vector***
Step 1: *Parameters Initialization*

$$\textit{Population} \leftarrow N = 10$$
$$\textit{Iterations} \leftarrow T = 100$$
$$\varphi_{cr} \leftarrow 0.8$$
$$\varphi_{mr} \leftarrow 0.01$$
$$\beta \leftarrow 5$$

Start.
Step 2: *Fitness Function*

$$\textit{Type} \leftarrow \textit{FKNN}$$
$$\textit{Neighbors} \leftarrow 5$$
$$K-\textit{fold} \leftarrow 2$$
$$\textit{Distance} \leftarrow \textit{Euclidean}$$
$$\textit{Loss} \leftarrow \sum_{i=1}^{n} x_i I\{\hat{a}_i \neq a_i\}$$

Step 3: *Selection*

$$X \leftarrow \frac{x_i}{\sum(x_i)}$$
$$x_i \leftarrow \exp\left(-k_1 \times \frac{S_\beta}{X_L}\right)$$

Step 4: *Crossover*

$$\varphi_{cr} \leftarrow \textit{CrossOver}(A_1, A_2)$$

Step 5: *Mutation*

$$\textit{Type} \leftarrow \textit{Uniform}$$

Step 6: Repeat Step 2
Step 7: $\boldsymbol{SF(i)} \leftarrow$ ***Best Features***
End

### 4.5.2 | Selection

The selection of features from each iteration is key step in GA. In this work, a roulette wheel method is applied for parent selection. This method is based on probability value expressed as follows:

$$X = \frac{x_i}{\sum(x_i)} \tag{8}$$

$$x_i = \exp\left(-k_1 \times \frac{S_\beta}{X_L}\right) \tag{9}$$

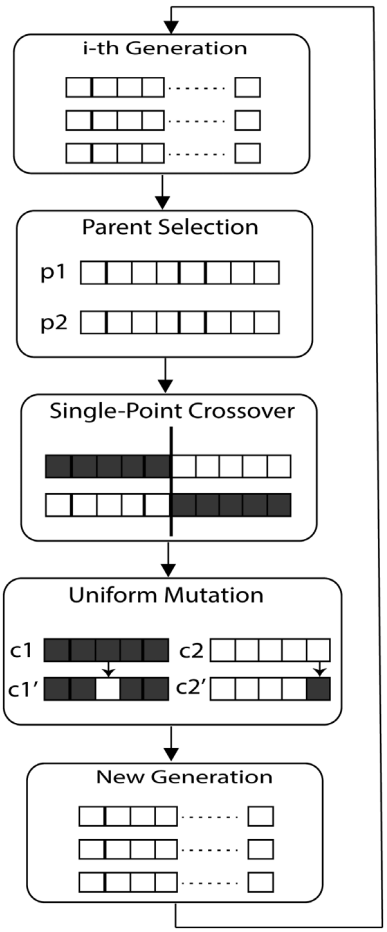


**FIGURE 5** Architecture of genetic algorithm

where $k_1$ is the pressure for selected parent, $S_\beta$ is the population in sorted way, and $X_L$ is the last population.

### 4.5.3 | Crossover rate

After selection of best features from each iteration, crossover is a next step. In this step, genes of two parents are swapped to generate a better individual. Single point crossover of value 0.8 is used in this work and presented in Figure 6. A high crossover rate may result in premature convergence of GA. Mathematically, crossover is defined as follows:

$$\varphi_{cr} = \text{Crossover}\,(A_1, A_2) \tag{10}$$

where

$$A_1 = cB_1 + (1-c) \times B_2 \tag{11}$$

$$A_2 = cB_2 + (1-c) \times B_1 \tag{12}$$

### 4.5.4 | Mutation rate and fitness function

The mutation is used to maintain genetic diversity and to avoid premature convergence (Das, Sengupta, & Bhattacharyya, 2018). The process of mutation is flipping one or more gene according to a defined mutation rate. In this work, uniform mutation of rate 0.01 is used after which fitness function "fitcknn" is defined to get more optimized results. This function describes K-nearest neighbor (KNN) classifier. In each iteration, the accuracy of selected features is analyzed for which selected features and Euclidean distance are used as input. Then, the performance of selected features is judged through error rate on cross validation with fivefold. Finally, the selected features performance is seen on the basis of loss value. Mathematically, Euclidean distance is defined as:

$$d(a,b) = \sqrt{(a_1 - b_1)^2 + \ldots + (a_n - b_n)^2} \tag{13}$$

Moreover, error for the loss function of KNN model is expressed as:

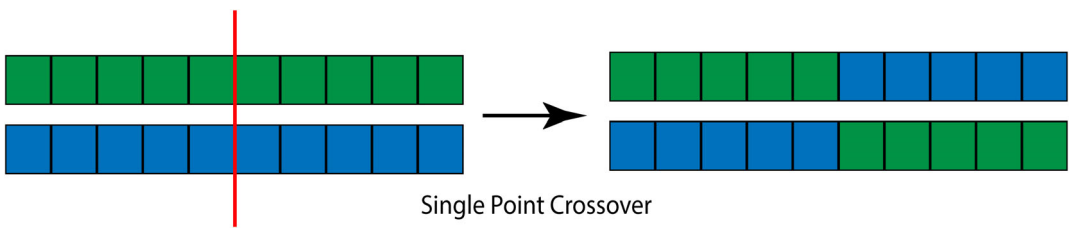


**FIGURE 6** Crossover (single point) of 0.8 rate [Color figure can be viewed at wileyonlinelibrary.com]

$$\text{Loss} = \sum_{i=1}^{n} x_i I\{\hat{a}_i \neq a_i\} \tag{14}$$

This process is continued until the all iterations are completed. After that, a new optimized vector is obtained known as best selected feature vector which is validated through Ensemble classifier.

### 4.6 | Classification

In machine learning, ensemble is a supervised learning method and combines several learning methods to produce a better performance. In this model, the features are randomly tested. Mathematically, it is represented as:

$$\hat{Z} = \exp\left(\sum_{x=1}^{X} \hat{a}_x b_x(y)\right) \tag{15}$$

where, $b_x(y) = b_1(y), b_2(y), \ldots, b_x(y)$ and $\hat{a}_x = \hat{a}_1, \hat{a}_2, \ldots, \hat{a}_x$.

## 5 | RESULTS AND ANALYSIS

In this section, a detailed description of experimental results and analysis is presented. The proposed system is tested on GI dataset prepared from different datasets, as described in Section 4.1. Five different categories are included in this dataset: (a) bleeding, (b) ulcer, (c) polyp, (d) esophagitis, and (e) healthy. To give fair analysis and comparison of the proposed system, results are collected on seven different classifiers including cubic SVM, weighted KNN (WKNN), medium Gaussian SVM (MGSVM), fine KNN (FKNN), quadratic SVM, ensemble bagged trees (EBT), and ensemble subspace KNN (ESKNN). All methods are tested on different number of features as given in Table 1. In this table, it is shown that DCT, DWT, strong 809 colors, and VGG16-based features are fused in feature set 1. In feature set 2, GA-based selection is performed on feature set 1 whereas DWT, DCT, 509 SCF, and VGG16-based features are fused in feature set 3. At the end, GA-based selection is performed on feature set 3 and best points for classification are selected.

**TABLE 1** Different scenarios for GI diseases classification

| Feature set | Description |
|---|---|
| 1 | Classification using DCT, DWT, VGG16, and 809 SCFs |
| 2 | Classification after GA-based features selection on feature set 1 |
| 3 | Classification using DCT, DWT, VGG16, and 509 strong color features |
| 4 | Classification after GA-based feature selection on feature set 2 |

Abbreviations: DCT, discrete cosine transform; DWT, discrete wavelet transform; GA, genetic algorithm; GI, gastrointestinal; SCF, strong color feature.

### 5.1 | Implementation setup and evaluation protocols

The implementation of the proposed method is performed on MATLAB2018b using desktop computer. At the start of implementation, data are partitioned into 70:30 for training and testing process. All handcrafted features are implemented through well-defined techniques, as listed in Section 4.2, whereas deep learning-based features are calculated by the pretrained model VGG16. For deep features, the cross entropy activation function is used along with few parameters such as learning rate as 0.001with minibatch size of 64. Both tests and trained sample features are obtained. The trained sample features are fed to multi SVM which are later utilized in visual testing. The test sample features are provided to the selected classifier for final classification accuracy. The performance of selected classifier is analyzed using different performance evaluation measures including accuracy, false negative rate (FNR), sensitivity, precision, F1 score, and computational time.

### 5.2 | Results

In the first scenario, DCT, DWT, SCF, and deep features are fused and fed to classifiers. The best accuracy is 95.4% achieved on ESKNN classifier. Other parameters such as FNR, sensitivity, precision and F1

**TABLE 2** Classification results using feature set 1 on prepared database

| Classifier | Accuracy (%) | FNR (%) | Sensitivity (%) | Precision (%) | F1 score (%) | Time (s) |
|---|---|---|---|---|---|---|
| Quadratic SVM | 91.5 | 8.5 | 91.4 | 91.5 | 91.45 | 211.35 |
| Cubic SVM | 92.9 | 7.1 | 93.0 | 92.78 | 92.89 | 289.49 |
| MGSVM | 90.1 | 9.9 | 89.8 | 90.0 | 89.89 | 276.29 |
| FKNN | 94.2 | 5.8 | 94.0 | 94.3 | 94.15 | 137.66 |
| WKNN | 93.0 | 7.0 | 93.2 | 93.2 | 93.2 | 136.55 |
| EBT | 92.6 | 7.4 | 92.4 | 92.43 | 92.42 | **113.37** |
| ESKNN | **95.4** | **4.6** | **95.4** | **95.48** | **95.44** | 176.21 |

Abbreviations: EBT, ensemble bagged tree; ESKNN, ensemble subspace KNN; FKNN, fine KNN; FNR, false negative rate; KNN, K-nearest neighbor; SVM, support vector machine. Bold values represent best results.

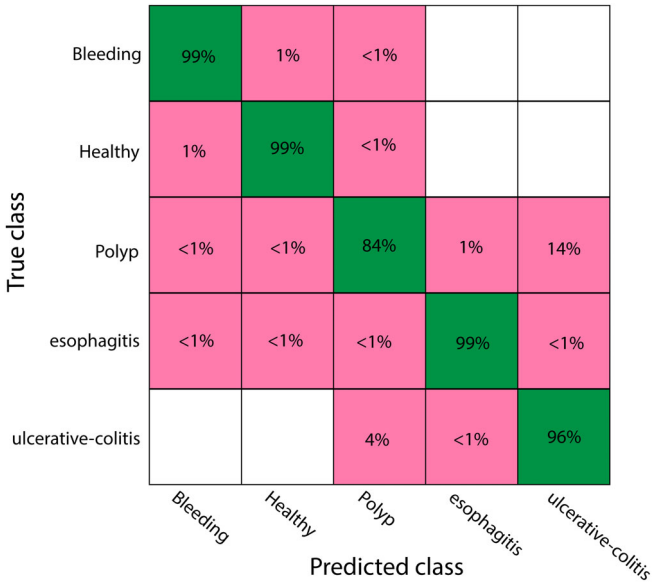


**FIGURE 7** Confusion matrix of ensemble subspace KNN (ESKNN) using fused feature set 1 [Color figure can be viewed at wileyonlinelibrary.com]

score are calculated as 4.6, 95.4, 95.48, and 95.44%, respectively, on this classifier as given in Table 2. The performance of this classifier is also validated through confusion matrix demonstrated in Figure 7. The second best achieved performance is 94.2% on FKNN using same feature set. The worst accuracy achieved on this feature set is 90.1% on MGSVM. Finally, the classification computation time of each classifier is also observed on the same feature set and plotted in Figure 8 where it is shown that EBT executed in 113.37 s which is best noted time as compared to other classifiers.

After the first experiment on all extracted features, GA is performed on fused vector to select best of them of dimension $N \times 554$, where $N$ denotes the testing set. GA-based selected vector is provided to classifiers and maximum accuracy of 96.5% is obtained on ESKNN as given in Table 3. Results on other parameters such as FNR, sensitivity, precision, and F1 score on ESKNN are 3.5, 96.5, 96.5, and 96.5%, respectively. The achieved accuracy on this classifier is also validated from confusion matrix given

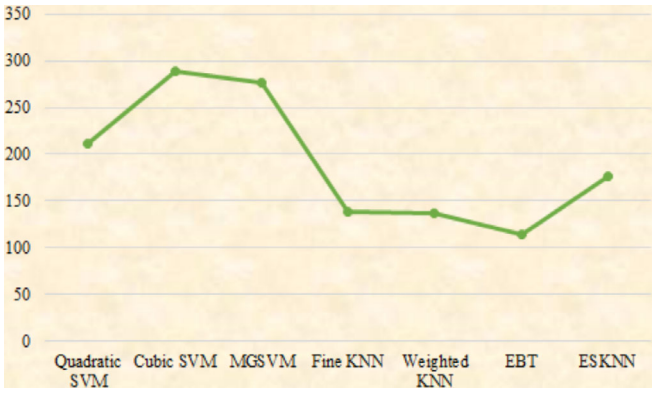


**FIGURE 8** Classification computation time for all selected classifiers on feature set 1 [Color figure can be viewed at wileyonlinelibrary.com]

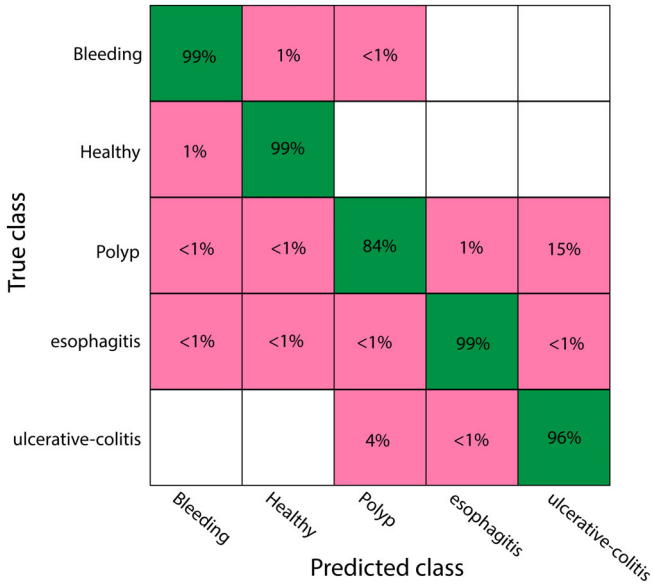


**FIGURE 9** Confusion matrix of ensemble subspace KNN (ESKNN) using genetic algorithm (GA) on fused feature set 1 [Color figure can be viewed at wileyonlinelibrary.com]

**TABLE 3** Classification results using GA on feature set 1

| Classifier | Accuracy (%) | FNR (%) | Sensitivity (%) | Precision (%) | F1 score (%) | Time (s) |
|---|---|---|---|---|---|---|
| Quadratic SVM | 91.2 | 8.8 | 91.0 | 90.98 | 90.99 | 92.044 |
| Cubic SVM | 92.5 | 7.5 | 92.6 | 92.36 | 92.48 | 103.3 |
| MGSVM | 89.8 | 10.2 | 89.6 | 89.7 | 89.65 | 116.34 |
| FKNN | 94.6 | 5.4 | 94.6 | 94.67 | 94.65 | **70.328** |
| WKNN | 93.3 | 6.7 | 93.4 | 93.48 | 93.44 | 72.463 |
| EBT | 92.4 | 7.6 | 92.2 | 92.3 | 92.25 | 88.583 |
| ESKNN | **96.5** | **3.5** | **96.5** | **96.5** | **96.5** | 108.7 |

Abbreviations: EBT, ensemble bagged tree; ESKNN, ensemble subspace KNN; FNR, false negative rate; FKNN, fine KNN; GA, genetic alogrithm; KNN, K-nearest neighbor; MGSVM, medium Gaussian SVM; SVM, support vector machine; WKNN, weighted KNN. Bold values represent best results.

in Figure 9. The second best achieved accuracy is 94.6% on FKNN with FNR as 5.4% and F1 score as 94.65%. Finally, the classification time for each selected classifier is also observed and plotted in Figure 10 which shows that the minimum noted time is 70.328 s on FKNN whereas the worst classification time is 116.34 s obtained on MGSVM.

In the third experiment, SCF, DCT, DWT, and deep features are fused in one matrix where the number of color features in this experiment is 509 in total. The fusion of these features has given $N \times 1893$ dimension vector, where $N$ denotes total number of testing set. On this feature set, maximum achieved accuracy is 95.1% on ESKNN classifier with computational time 1,560.9 s as given in Table 4. Other calculated metrics such as FNR is 4.9%, sensitivity is 95.2%, precision is 95.22%, and F1 score is 95.21%. The confusion matrix is demonstrated in Figure 11 which also verifies the performance of ESKNN. At the end, the computational time of this feature set on all classifiers is also observed and plotted in Figure 12 where it is shown that 110.56 s is best noted time whereas 246.22 s is highest.

Finally, GA is applied on fused feature set 3 and classification is performed. The maximum achieved accuracy of this feature set is 95.3% on ESKNN classifier as presented in Table 5. Figure 13 displays the performance of this classifier on feature set 4 through confusion matrix. The computational time for ESKNN is also noted which is minimized to 75.71 s. The best noted time is 57.103 s on WKNN as plotted in Figure 14. The overall evaluation of this feature set shows that the system performance is declined when color features are decreased by variance method but it consumes little lower time for execution as compared to feature set 2 which is best in this work.

## 5.3 | Analysis

In this part, the proposed classification model is analyzed and presented. Major steps of the proposed model are database preparation, extraction of features (DCT, DWT, SCF, and deep features), selection of best features using GA, and classification using Ensemble classifier. Five classes, that is, polyp, ulcer, bleeding, esophagitis, and healthy of GI tract are used. Different types of feature sets are generated for better analysis of classification performance. The description of

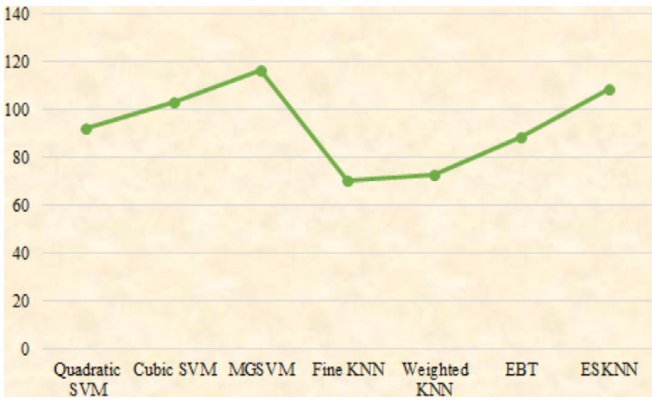


**FIGURE 10** Classification computation time for all selected classifiers using genetic algorithm (GA) based feature selection [Color figure can be viewed at wileyonlinelibrary.com]

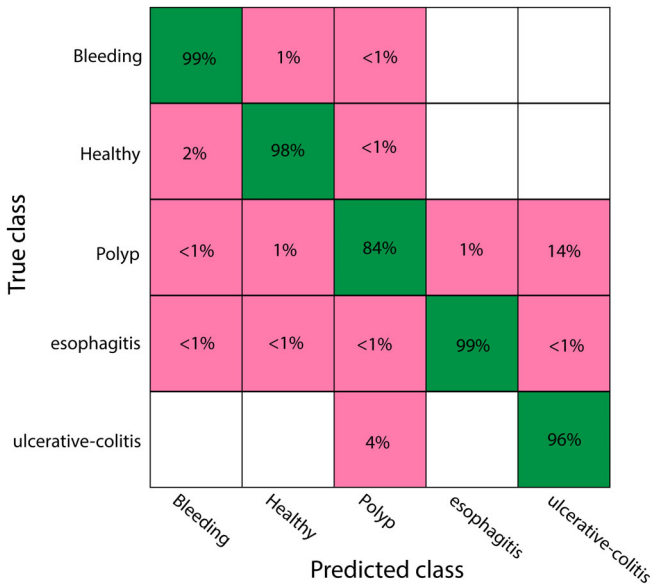


**FIGURE 11** Confusion matrix of ensemble subspace K-nearest neighbor (ESKNN) using fused feature set 3 [Color figure can be viewed at wileyonlinelibrary.com]

**TABLE 4** Classification results using feature set 3 on prepared database

| Classifier | Accuracy (%) | FNR (%) | Sensitivity (%) | Precision (%) | F1 score (%) | Time (s) |
|---|---|---|---|---|---|---|
| Quadratic SVM | 91.7 | 8.3 | 91.6 | 91.57 | 91.58 | 164.12 |
| Cubic SVM | 92.9 | 7.1 | 92.8 | 92.79 | 92.79 | 246.22 |
| MGSVM | 90.3 | 9.7 | 90.2 | 90.26 | 90.23 | 199.61 |
| FKNN | 94.7 | 5.3 | 94.8 | 94.75 | 94.77 | **110.56** |
| WKNN | 93.4 | 6.6 | 93.4 | 93.68 | 93.54 | 115.02 |
| EBT | 92.4 | 7.6 | 92.2 | 92.23 | 92.22 | 111.85 |
| ESKNN | **95.1** | **4.9** | **95.2** | **95.22** | **95.21** | 156.9 |

Abbreviations: EBT, ensemble bagged tree; ESKNN, ensemble subspace KNN; FNR, false negative rate; FKNN, fine KNN; KNN, K-nearest neighbor; MGSVM, medium Gaussian SVM; SVM, support vector machine; WKNN, weighted KNN. Bold values represent best results.

different feature sets is given in Table 1. The results of all scenarios are given in Tables 2–5 and the best results are validated through Figures 7, 9, 11, and 13. From Tables 2–5, it is clearly shown that feature set 2 gives the best accuracy but when the number of SCF is reduced, then accuracy is also reduced up to 2%. On the other end, the classification computational time is plotted in Figures 8, 10, 12, and 14 which demonstrated that feature set 4 required less time as compared to other three feature sets.

In feature set 1, the selected numbers of SCF is 809 while in feature set 3, the selected number of color features is 509. The main difference among both feature sets is number of SCF that decreases heuristically. Further, the accuracy on both features sets is checked separately and presented in Tables 2–4 which shows that the accuracy on feature set 3 is decreased only 0.23% while on the other end, the computational time of feature set 3 is improved and almost 5–10 s is decreased as compared to feature set 1. After that, GA is applied on both sets separately and it is illustrated that the accuracy of feature set 2 is improved but the execution time of feature set 4 is better. Overall, it is observed that the number of SCF provides better accuracy while on feature set 2, the accuracy decreases almost 1.13% but the computational time of fewer color features is better.

The statistical analysis is also performed by 1,000 times iterations after which results in the form of min, avg, max, *SD*, and confidence interval (CI) are given in Table 6 where it is shown that overall feature set 2 outperformed based on CI. CI is an estimation computed after performing statistical operations on the dataset. If CI is smaller, uncertainty is less about the results. As presented in Table 6, CI is high while the difference is high between minimum and maximum accuracy. The lowest difference in accuracies is for feature set 4; therefore, CI is minimum for this set which is 0.0942. Besides, the maximum difference in feature set 1 and CI is 0.2596.

A general comparison with previous models is also conducted which is given in Table 7. In previous studies, mostly results are taken using two classes of GI diseases. In this study, the proposed method is evaluated using five different classes and from the above analysis, it is clearly observed that the proposed model outperforms the previous models.

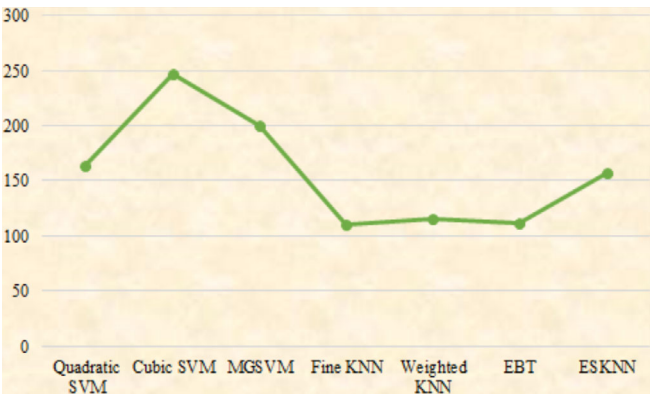


**FIGURE 12** Classification computation time for all selected classifiers on feature set 3 [Color figure can be viewed at wileyonlinelibrary.com]

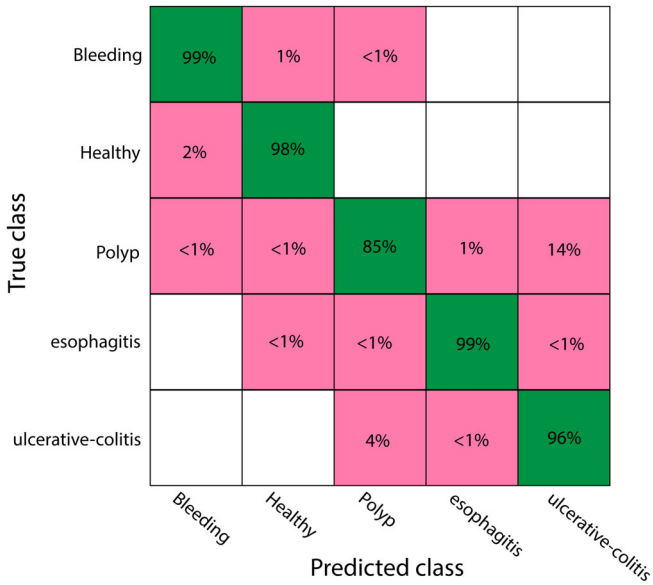


**FIGURE 13** Confusion matrix of ensemble subspace K-nearest neighbor (ESKNN) for feature set 4 [Color figure can be viewed at wileyonlinelibrary.com]

**TABLE 5** Classification results using GA on feature set 3

| Classifier | Accuracy (%) | FNR (%) | Sensitivity (%) | Precision (%) | F1 score (%) | Time (s) |
|---|---|---|---|---|---|---|
| Quadratic SVM | 91.5 | 8.5 | 91.4 | 91.36 | 91.38 | 78.278 |
| Cubic SVM | 92.9 | 7.1 | 93.0 | 92.7 | 92.85 | 89.901 |
| MGSVM | 90.3 | 9.7 | 90.2 | 90.24 | 90.22 | 97.377 |
| FKNN | 95.1 | 4.9 | 95.2 | 95.25 | 95.23 | 59.114 |
| WKNN | 93.4 | 6.6 | 93.6 | 93.58 | 93.59 | **57.103** |
| EBT | 92.0 | 8.0 | 92.0 | 91.88 | 91.94 | 73.502 |
| ESKNN | **95.3** | **4.7** | **95.4** | **95.37** | **95.39** | 75.71 |

Abbreviations: EBT, ensemble bagged tree; ESKNN, ensemble subspace KNN; FNR, false negative rate; FKNN, fine KNN; GA, genetic alogrithm; KNN, K-nearest neighbor; MGSVM, medium Gaussian SVM; SVM, support vector machine; WKNN, weighted KNN. Bold values represent best results.

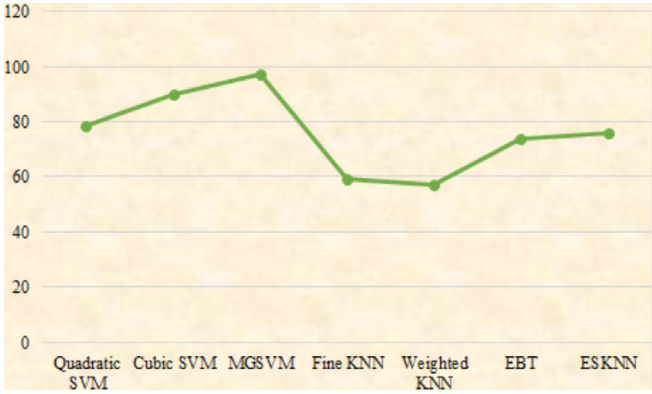


**FIGURE 14** Classification computation time for all selected classifiers on feature set 4 [Color figure can be viewed at wileyonlinelibrary.com]

**TABLE 6** Statistical analysis of the proposed method on feature set 2 using ESKNN

| Features set | Measures after 1,000 iterations | | | | |
|---|---|---|---|---|---|
| | Minimum | Average | Maximum | SD | CI |
| 1 | 94.3 | 94.8 | 95.4 | 0.4496 | 0.2596 |
| 2 | **95.8** | **96.1** | **96.5** | 0.2867 | 0.1655 |
| 3 | 94.2 | 94.6 | 95.1 | 0.3681 | 0.2125 |
| 4 | 94.9 | 95.1 | 95.3 | 0.1632 | 0.0942 |

Abbreviations: CI, confidence interval; ESKNN, ensemble subspace KNN; KNN, K-nearest neighbor. Bold values represent best results.

**TABLE 7** Comparison with existing methods

| Method | Year | Disease | Performance (%) |
|---|---|---|---|
| Lee et al. (2019) | 2019 | Normal, ulcer | Accuracy (92.62) |
| Ghatwary et al. (2019b) | 2019 | Normal, cancer | Sensitivity (95.0) |
| Fan et al. (2018) | 2018 | Ulcer, erosion | Accuracy (95.16) |
| Deeba et al. (2018) | 2018 | Normal, bleeding | Accuracy (94.5) |
| Chen et al. (2017) | 2017 | Polyps | Accuracy (90.1) |
| Proposed | **2019** | Polyp, ulcer, bleeding, esophagitis, and normal | **Accuracy (96.5)** |

## 6 | CONCLUSION

In this work, a new method is proposed to detect and classify GI diseases from endoscopic images. The proposed model comprises of (a) extraction of DCT, DWT, and SCF; (b) extraction of VGG16 features; (c) features fusion; (d) GA-based features selection; and (e) classification. The maximum classification accuracy achieved on five classes of WCE images is 96.5%. From the analysis of all the above results, the conclusion is that the proposed GA-based features selection method performs well to select robust features for classification. This features selection method also improves computational time without affecting the classification accuracy. SCF also helps to improve the performance of system. There is a slight difference between classification accuracies of minimum feature vector and high-dimensional feature vector. The accuracy can be increased by choosing more appropriate parameters of GA such as crossover and mutation rate. System performance can be affected with increase in number of classes. Moreover, sometimes, it is a possibility that during the selection process, few important features are discarded. In future, a new CAD system can be developed for the classification of GI diseases by utilizing the latest CNN model like CapsuleNet and DenseNet on more complex data. Further, the detection of these stomach deformities can be done through deep learning.

## ORCID

*Muhammad Attique Khan* https://orcid.org/0000-0001-7058-0715
*Amjad Rehman* https://orcid.org/0000-0002-3817-2655

## REFERENCES

Ahmed, N., Natarajan, T., & Rao, K. R. (1974). Discrete cosine transform. *IEEE Transactions on Computers*, *100*(1), 90–93.

Akram, T., Khan, M. A., Sharif, M., & Yasmin, M. (2018). Skin lesion segmentation and recognition using multichannel saliency estimation and M-SVM on selected serially fused features. *Journal of Ambient Intelligence and Humanized Computing*, 1–20.

Bernal, J., Sánchez, F. J., Fernández-Esparrach, G., Gil, D., Rodríguez, C., & Vilariño, F. (2015). WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics*, *43*, 99–111.

Billah, M., Waheed, S., & Rahman, M. M. (2017). An automatic gastrointestinal polyp detection system in video endoscopy using fusion of color wavelet and convolutional neural network features. *International Journal of Biomedical Imaging*, *2017*, 1–9.

Bledsoe, W. W. (1961). *The use of biological concepts in the analytical study of systems*. Paper presented at the ORSA-TIMS National Meeting.

cancer.net. (2019). Stomach Cancer: Statistics. Retrieved from https://www.cancer.net/cancer-types/stomach-cancer/statistics

Chan, T. F., & Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, *10*(2), 266–277.

Charfi, S., & El Ansari, M. (2018). Computer-aided diagnosis system for colon abnormalities detection in wireless capsule endoscopy images. *Multimedia Tools and Applications*, *77*(3), 4047–4064.

Chen, Y., & Lee, J. (2012). A review of machine-vision-based analysis of wireless capsule endoscopy video. *Diagnostic and Therapeutic Endoscopy*, *2012*, 1–9.

Chen, Y., Zhu, L., Ghamisi, P., Jia, X., Li, G., & Tang, L. (2017). Hyperspectral images classification with Gabor filtering and convolutional neural network. *IEEE Geoscience and Remote Sensing Letters*, *14*(12), 2355–2359.

Cogan, T., Cogan, M., & Tamil, L. (2019). MAPGI: Accurate identification of anatomical landmarks and diseased tissue in gastrointestinal tract using deep learning. *Computers in Biology and Medicine*, *111*, 103351.

Das, A. K., Sengupta, S., & Bhattacharyya, S. (2018). A group incremental feature selection for classification using rough set theory based genetic algorithm. *Applied Soft Computing*, *65*, 400–411.

De Souza, L. A., Afonso, L. C. S., Palm, C., & Papa, J. P. (2017). *Barrett's esophagus identification using optimum-path forest*. Paper presented at the 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI).

Deeba, F., Islam, M., Bui, F. M., & Wahid, K. A. (2018). Performance assessment of a bleeding detection algorithm for endoscopic video based on classifier fusion method and exhaustive feature selection. *Biomedical Signal Processing and Control*, *40*, 415–424.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). *Imagenet: A large-scale hierarchical image database*. Paper presented at the 2009 IEEE Conference on Computer Vision and Pattern Recognition.

Fan, S., Xu, L., Fan, Y., Wei, K., & Li, L. (2018). Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. *Physics in Medicine & Biology*, *63*(16), 165001.

Fu, Y., Zhang, W., Mandal, M., & Meng, M. Q.-H. (2013). Computer-aided bleeding detection in WCE video. *IEEE Journal of Biomedical and Health Informatics*, *18*(2), 636–642.

Ghareb, A. S., Bakar, A. A., & Hamdan, A. R. (2016). Hybrid feature selection based on enhanced genetic algorithm for text categorization. *Expert Systems with Applications: An International Journal*, *49*, 31–47.

Ghatwary, N., Ye, X., & Zolgharni, M. (2019a). Esophageal Abnormality detection using DenseNet based Faster R-CNN with Gabor features. *IEEE Access*, *99*, 1–1.

Ghatwary, N., Ye, X., & Zolgharni, M. (2019b). Esophageal abnormality detection using DenseNet based faster R-CNN with Gabor features. *IEEE Access*, *7*, 84374–84385.

Günal, S. (2012). Hybrid feature selection for text classification. *Turkish Journal of Electrical Engineering and Computer Sciences*, *20*(Suppl. 2), 1296–1311.

Hosseini, S., Lee, S. H., & Cho, N. I. (2018). Feeding hand-crafted features for enhancing the performance of convolutional neural networks. *arXiv Preprint*, 8.

Iakovidis, D. K., Georgakopoulos, S. V., Vasilakakis, M., Koulaouzidis, A., & Plagianakos, V. P. (2018). Detecting and locating gastrointestinal anomalies using deep learning and iterative cluster unification. *IEEE Transactions on Medical Imaging*, *37*(10), 2196–2210.

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

John, H. (1992). *Holland, adaptation in natural and artificial systems*. Cambridge, MA: MIT Press.

Khan, M. A., Akram, T., Sharif, M., Awais, M., Javed, K., Ali, H., & Saba, T. (2018). CCDF: Automatic system for segmentation and recognition of fruit crops diseases based on correlation coefficient and deep CNN features. *Computers and Electronics in Agriculture*, *155*, 220–236.

Khan, M. A., Akram, T., Sharif, M., Javed, K., Rashid, M., & Bukhari, S. A. C. (2019). An integrated framework of skin lesion detection and recognition through saliency method and optimal deep neural network features selection. *Neural Computing and Applications*, 1–20.

Khan, M. A., Javed, M. Y., Sharif, M., Saba, T., & Rehman, A. (2019). *Multi-Model Deep Neural Network based Features Extraction and Optimal Selection Approach for Skin Lesion Classification*. Paper presented at the 2019 International Conference on Computer and Information Sciences (ICCIS).

Khan, M. A., Lali, M. I. U., Sharif, M., Javed, K., Aurangzeb, K., Haider, S. I., … Akram, T. (2019). An optimized method for segmentation and classification of apple diseases based on strong correlation and genetic algorithm based feature selection. *IEEE Access*, *7*, 46261–46277.

Khan, M. A., Rashid, M., Sharif, M., Javed, K., & Akram, T. (2019). Classification of gastrointestinal diseases of stomach from WCE using improved saliency-based method and discriminant features selection. *Multimedia Tools and Applications*, 1–28.

Khan, M. A., Sharif, M., Akram, T., Bukhari, S. A. C., & Nayak, R. S. (2019). Developed Newton-Raphson based deep features selection framework for skin lesion recognition. *Pattern Recognition Letters*.

Khan, M. A., Sharif, M., Akram, T., Yasmin, M., & Nayak, R. S. (2019). Stomach deformities recognition using rank-based deep features selection. *Journal of Medical Systems*.

Khan, M. A., Sharif, M. I., Raza, M., Anjum, A., Saba, T., & Shad, S. A. (2019). Skin lesion segmentation and classification: A unified framework of deep neural network features fusion and selection. *Expert Systems*.

Kishore, M. (2015). An effective and efficient feature selection method for lung cancer detection. *International Journal of Computer Science & Information Technology*, *7*(4), 2074–9007.

Kwolek, B. (2005). *Face detection using convolutional neural networks and Gabor filters*. Paper presented at the International Conference on Artificial Neural Networks.

Lee, J. H., Kim, Y. J., Kim, Y. W., Park, S., Choi, Y.-i., Kim, Y. J., … Chung, J.-W. (2019). Spotting malignancies from gastric endoscopic images using deep learning. *Surgical Endoscopy*, 1–8.

Li, B., & Meng, M. Q.-H. (2009). Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments. *Computers in Biology and Medicine*, *39*(2), 141–147.

Liaqat, A., Khan, M. A., Shah, J. H., Sharif, M., Yasmin, M., & Fernandes, S. L. (2018). Automated ulcer and bleeding classification from WCE images using multiple features fusion and selection. *Journal of Mechanics in Medicine and Biology*, *18*(04), 1850038.

Luan, S., Chen, C., Zhang, B., Han, J., & Liu, J. (2018). Gabor convolutional networks. *IEEE Transactions on Image Processing*, *27*(9), 4357–4366.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *11*(7), 674–693.

Menon, S., & Trudgill, N. (2014). How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-analysis. *Endoscopy International Open*, *2*(02), E46–E50.

Münzer, B., Schoeffmann, K., & Böszörmenyi, L. (2018). Content-based processing and analysis of endoscopic images and videos: A survey. *Multimedia Tools and Applications*, *77*(1), 1323–1362.

Nasir, M., Attique Khan, M., Sharif, M., Lali, I. U., Saba, T., & Iqbal, T. (2018). An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach. *Microscopy Research and Technique*, *81*(6), 528–543.

Pogorelov, K., Randel, K. R., Griwodz, C., Eskeland, S. L., de Lange, T., Johansen, D., … Schmidt, P. T. (2017). *Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection*. Paper presented at the Proceedings of the 8th ACM on Multimedia Systems Conference.

Pontabry, J., Rousseau, F., Studholme, C., Koob, M., & Dietemann, J.-L. (2017). A discriminative feature selection approach for shape analysis: Application to fetal brain cortical folding. *Medical Image Analysis*, *35*, 313–326.

Rajaei, A., & Rangarajan, L. (2011). Wavelet features extraction for medical image classification. *International Journal of Engineering Science*, *4*, 131–141.

Rashid, M., Khan, M. A., Sharif, M., Raza, M., Sarfraz, M. M., & Afza, F. (2019). Object detection and classification: A joint selection and fusion strategy of deep convolutional neural network and SIFT point features. *Multimedia Tools and Applications*, *78*(12), 15751–15777.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster R-CNN: Towards real-time object detection with region proposal networks*. Paper presented at the Advances in Neural Information Processing Systems.

Ribeiro, M. G., Neves, L. A., do Nascimento, M. Z., Roberto, G. F., Martins, A. S., & Tosta, T. A. A. (2019). Classification of colorectal cancer based on the association of multidimensional and multiresolution features. *Expert Systems with Applications*, *120*, 262–278.

Saba, T., Khan, M. A., Rehman, A., & Marie-Sainte, S. L. (2019). Region extraction and classification of skin cancer: A heterogeneous

framework of deep CNN features fusion and reduction. *Journal of Medical Systems*, *43*(9), 289.

Sharif, M., Attique Khan, M., Rashid, M., Yasmin, M., Afza, F., & Tanik, U. J. (2019). Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. *Journal of Experimental & Theoretical Artificial Intelligence*, 1–23.

Sharif, M., Khan, M. A., Iqbal, Z., Azam, M. F., Lali, M. I. U., & Javed, M. Y. (2018). Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Computers and Electronics in Agriculture*, *150*, 220–234.

Sharif, M., Tanvir, U., Munir, E. U., Khan, M. A., & Yasmin, M. (2018). Brain tumor segmentation and classification by improved binomial thresholding and multi-features selection. *Journal of Ambient Intelligence and Humanized Computing*, 1–20.

Sharif, M. I., Li, J. P., Khan, M. A., & Saleem, M. A. (2020). Active deep neural network features selection for segmentation and recognition of brain tumors using MRI images. *Pattern Recognition Letters*, *129*, 181–189.

Shi, Q., Li, W., Zhang, F., Hu, W., Sun, X., & Gao, L. (2018). Deep CNN with multi-scale rotation invariance features for ship classification. *IEEE Access*, *6*, 38656–38668.

Siegel, R. L., Miller, K. D., Fedewa, S. A., Ahnen, D. J., Meester, R. G., Barzi, A., & Jemal, A. (2017). Colorectal cancer statistics, 2017. *CA: A Cancer Journal for Clinicians*, *67*(3), 177–193.

Siegel, R. L., Miller, K. D., & Jemal, A. (2016). Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians*, *66*(1), 7–30.

Silva, J., Histace, A., Romain, O., Dray, X., & Granado, B. (2014). Toward embedded detection of polyps in wce images for early diagnosis of colorectal cancer. *International Journal of Computer Assisted Radiology and Surgery*, *9*(2), 283–293.

Suman, S., Malik, A. S., Pogorelov, K., Riegler, M., Ho, S. H., Hilmi, I., & Goh, K. L. (2017). *Detection and classification of bleeding region in WCE images using color feature*. Paper presented at the Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing.

Swain, P. (2003). Wireless capsule endoscopy. *Gut*, *52*(Suppl 4), iv48–iv50.

Uğuz, H. (2011). A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm. *Knowledge-Based Systems*, *24*(7), 1024–1032.

Uysal, A. K., & Gunal, S. (2014). Text classification using genetic algorithm oriented latent semantic features. *Expert Systems with Applications*, *41*(13), 5938–5947.

Vasilakakis, M., Koulaouzidis, A., Yung, D. E., Plevris, J. N., Toth, E., & Iakovidis, D. K. (2019). Follow-up on: Optimizing lesion detection in small bowel capsule endoscopy and beyond: From present problems to future solutions. *Expert Review of Gastroenterology & Hepatology*, *13*(2), 129–141.

Yao, H., Chuyi, L., Dan, H., & Weiyu, Y. (2016). *Gabor feature based convolutional neural network for object recognition in natural scene*. Paper presented at the 2016 3rd International Conference on Information Science and Control Engineering (ICISCE).

Yi, J., Wu, P., Hoeppner, D. J., & Metaxas, D. (2017). *Fast neural cell detection using light-weight SSD neural network*. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.

**How to cite this article:** Majid A, Khan MA, Yasmin M, Rehman A, Yousafzai A, Tariq U. Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection. *Microsc Res Tech*. 2020;1–15. https://doi.org/10.1002/jemt.23447