Received September 29, 2020, accepted October 20, 2020, date of publication October 26, 2020, date of current version November 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3033504

Query Expansion Based on Top-Ranked Images for Content-Based Medical Image Retrieval

ALI AHMED, (Member, IEEE), AND SHARAF J. MALEBARY, (Member, IEEE)
Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh 21589, Saudi Arabia

Corresponding author: Ali Ahmed (aabdelrahim@kau.edu.sa; alikarary@gmail.com)

This work was supported by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, under Grant DF-093-830-1441.

ABSTRACT Recently the collection of varied digital image databases has increased. Many users have found that searching and retrieving required images from large collections are very difficult tasks, and successful and effective retrieval methods were developed to provide an effective and efficient search and retrieval process. In most content-based image retrieval (CBIR) methods, various visual features have been considered indirect to retrieve the images from databases. Content-based medical image retrieval (CBMIR), like any CBIR method, is a technique for retrieving medical images on the basis of automatically derived image features, such as colour and texture. Although a number of methods and approaches have been suggested, retrieval performance remains one of the most challenging problems in current CBMIR studies, this due to the well-known 'semantic gap' issue that exists between machine-captured low-level image features and human-perceived high-level semantic concepts. To bridge this gap, much research has been proposed. This study proposes two expansion methods to increase and enhance precision of the retrieval model; both of our proposed methods depend on top-ranked images. Our first expansion method is to reformulate the new expand query image based on mean values for features of top-ranked images, while the second method is based on selecting the most important features only. The expansion process was performed on eighteen colour features and twelve texture features extracted from two common medical images datasets, Kvasir and PH2. Experimental results show that our proposed methods performed better and have a precision of 95.8% and 80.7% for Kvasir and PH2 data set, respectively.

INDEX TERMS Image retrieval, feature extraction, query expansion, relevance feedback.

I. RESEARCH BACKGROUND

Imaging is a fundamental component of clinical medicine, and is commonly used for diagnosis [1], care and treatment planning [2], and patient response assessment [3]. The idea of image similarity has significant medical implications because diagnostic decision-making has historically involved the use of information from a patient (image and non-image). In recent years, the use of digital images has become increasingly popular across various sectors, including the medical, science and educational sectors. Hospitals and medical facilities generate a large number of digital images as part of their daily routines, such as X-ray, mammogram and magnetic resonance imaging (MRI). It is definitely a complex task to interpret medical images, requiring extensive knowledge. Since the CBMIR is a part of CBIR, it benefits from the advantages and developments of many methods used in the CBIR method. Kumara *et al.* [4] have developed a method for the automatic semantic annotation of medical images which leverages techniques from content-based image retrieval (CBIR); they extended CBIR methods to automatically annotate liver CT images. Also, a good combination of local texture information derived from more than one different texture descriptor proposed in [5] was widely used in CBMIR. Most content-based medical image retrieval (CBMIR) systems are focused on image similarity, whereby a user enters a query image, and the system responds by presenting the most similar image focused on a certain similarity criterion, then, the results of the related image query are shown in descending order.

The basic concept of any CBMIR method consists of two main steps or phases: extraction of the feature (offline phase) and calculation of similarity measures (online phase) [6]–[8]. Figure 1 shows the main and fundamental framework of a

The associate editor coordinating the review of this manuscript and approving it for publication was Jiju Poovvancheri.

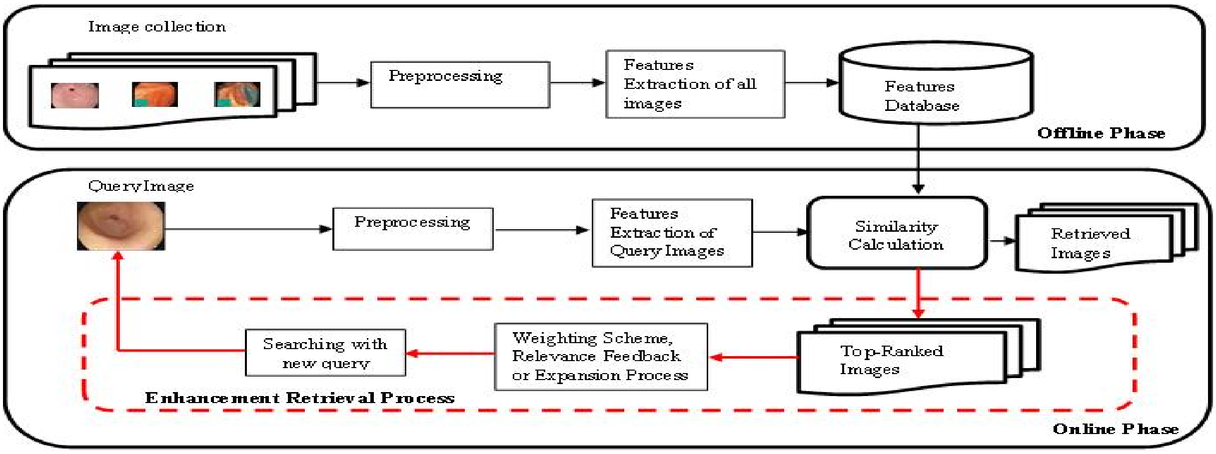


FIGURE 1. Main framework of the CBMIR method.

CBMIR system. Many enhancements were investigated and developed to enhance the retrieval performance and effectiveness of the CBMIR system; these extensions could be at the pre-processing and/or feature extraction phase [9], [10] but the most effective method might be in terms of any type of relevance feedback [11], [12] and query expansion stage as shown in red in Figure 1. Relevance feedback successfully increased the retrieval performance and improved the retrieval process in terms of recall and precision, as in our previous study [13]. Karamti [14] developed an effective retrieval method that includes a vectorisation technique combined with a pseudo-relevance model. Query expansion methods and techniques could apply to any type of multimedia object retrieval, especial for image retrieval. Carpineto and Romano [15] provide general application of query expansion in information retrieval. Chum *et al.* [16] use top-ranked images retrieved based on a single query to increase the retrieval precision by removing false positives; also the authors of the study proposed the use of the positively veri?ed results together with the original query, including the computation of a new query point and the use of multiple queries. In this study two effective enhancement methods of query expansion are proposed. The first method expands the query based on mean values for all image features of top-ranked images retrieved using a single query image. The second method filtered non-useful features of top-ranked images and used the important features only for new query image construction. The main contributions of our proposed scheme are summarised as:

(i) Extensions or enhancement method for increasing the recall and precision of CBMIR method using a simple and free user interaction expansion process.
(ii) Formulation of new query image based on expansion of colour and texture features using the mean values of top-ranked images.
(iii) Apply feature selection algorithm and removal of non-useful features of top-ranked images and formulation of new query image based on using selected important features expansion of colour and texture.

The rest of this paper is arranged as follows: Section 2 presents the previous works related to the study; Section 3 describes the basic idea of the proposed expansion methods. In Section 4, simulation results and discussions are demonstrated. Finally, Section 5 provides the conclusion of the paper.

II. RELATED STUDIES

In similar content-based searches the common approach for query expansion takes advantages of label information of top-ranked images retrieved and saved in a feedback session designed especially for this purpose. Many successful attempts of different expansion methods were found in the literature based on using local and global features, convolutional neural networks, incorporating some features with expand query and other approaches, recent review for query expansion in information retrieval found in [17]. Houle *et al.* [18] divided the expansion model into two parts or stages: offline process and online retrieval. In the offline process they use a generalised version of Laplacian score method for the computation of subjective feature spaces for individual data objects, while in the online retrieval part they rank the query according to feature scores of candidate objects from the database. Finally they replaced the original query with several top-rank initial query results. Their experiments with image sets and one voice set objects have signi?cantly outperformed their competitors. Kondylidis *et al.* [19] proposed a query expansion method on top of the formulated descriptors in their study; they treat the learned filters of the convolutional layers of a pre-trained CNN model as the detectors of the visual words. Features of more relevant images used to expand the features of the query to support final successful matching as states by Imbriaco *et al.* [20], the query expansion technique in their proposed method was particularly useful in combination with geometric verification. Chum *et al.* [21] proposed three extensions to automatic query expansion; they improved spatial verification and re-ranking by taking account of already evaluated results and proposed a method that exploits spatial context by incorporating matching features outside the initial query boundary into the query expansion. The latest study by Gordo *et al.* [22] developed a query expansion model based on a mathematical framework by treating query expansion as a discriminative learning problem, where an aggregation model is learned in a supervised manner and then they proposed (LAttQE) an aggregation model which is designed to share information between the query and top-ranked item by mean of self-attention. Top-ranked based method was widely used for the expansion process; different approaches were applied for this purpose. The user interaction method as part of a multi-modal query expansion framework developed by [23] and 'itra-expansion'' and 'inter-expansion'' for images represented by bang-of-features index based on hash image retrieval [24], were successfully implemented. This study proposed two simple free user interaction expansion methods using top-ranked images: the first method depends on the mean values of top-ranked images' features, while the second method takes advantage of the most important features of top-ranked images.

III. PROPOSED METHOD

Top-ranked images are considered the simplest and most effective method for the expansion process in the area of information retrieval (IR). In this method, query is automatically expanded to include its top-ranked neighbours, which may increase the chances of finding the most semantically related result. The main idea of our proposed expansion method is that, the top retrieved relevant images are used to provide information to expand and reformulate new query images. Here, top images retrieved by single image query are used as the basis for the expansion process in two different scenarios: first, by calculating the mean values for each feature of top images – we named this method as expansion method based on mean values – and it will implemented for both colour and texture images features, which we refer to as colour expansion based on mean retrieval (CLEBMR) method and texture expansion based on mean retrieval (TXEBMR) method, respectively. In the second scenario, instead of calculating the mean values, we apply a feature selection algorithm for the top retrieved images and we refer to this method as colour expansion based on feature selection retrieval (CLEBFSR) method when we apply it for colour features, and texture expansion based on feature selection retrieval (TXEBFSR) method when it is applied for texture features. More details with associated diagrams and pseudocodes for these two proposed methods will explained in sub-sections B and C, and in the following section we will explain the colour and texture features extraction methods that represent the most important part in any content based retrieval system.

A. FEATURE EXTRACTION

Feature values are considered to be a very important part or stage since the calculation of the similarity is based on values of colour and texture features. As in many related studies in the area of CBMIR, both colour and texture features descriptors could be used. The following two paragraphs explain colour and texture features used in this study.

1) COLOUR FEATURES

For colour feature descriptors, the original RGB images were first converted to HSV colour space [25] then six colour moments were calculated from each colour space, resulting in 18 features values. The six basic features extracted from each colour channels are: mean, variance, sknew, kurtos, smoothness and contrast. The following equations (1 to 6) describe the formula for each colour moment function; these colour moments were proposed by [26] and were successfully used in our previous study [13]. Let the value of pixel at i*th* row and j*th* column be v_{ij} and the dimension of image is (M,N) pixels, then the six colour moments functions are defined as follows:

$$\text{Mean} = \frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N} v_{ij} \tag{1}$$

$$\text{Variance} = \frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}(v_{ij}-m)^2 \tag{2}$$

$$\text{sknew} = \left(\frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(v_{ij}-m\right)^3\right)^{1/3} \tag{3}$$

$$\text{kurtos} = \left(\frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(v_{ij}-m\right)^4\right)^{1/4} \tag{4}$$

$$\text{Smoothness} = 1\frac{1}{(1+V)} \tag{5}$$

$$\text{Contrast} = \left(\frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(v_{ij}-m\right)^2\right)^{\frac{1}{2}} \tag{6}$$

2) TEXTURE FEATURES

For texture features descriptors Grey-level co-occurrence matrix (GLCM) was implemented. GLCM was introduced early by Weszka *et al.* [27]; it computed or organised by analysing the grey level of neighbouring pixels for any input image A, recent review on image texture classification method found in [28]. Let the original image A with Nr × Nc rows and columns be quantised to Ng grey level. The co-occurrence matrix will be a symmetric Ng × Ng that will describe the number of co-occurrences in a certain orientation and a certain pixel distance. In some textures the o-occurrence matrix computed for four different orientation angel $\varnothing = \{0, \pi/4, \pi/2, 3\pi/4\}$ and three distances $d = \{1, 2, 3\}$. In this study we fixed the orientation angle at $\varnothing = 0$ and $d = 1$, and there are $8 \times 8 = 64$ possible ordered combinations of values for each pixel pair and each intensity image is scaled to 8 level. Finally, and similar to the colour features extraction, each RGB image is converted to HSV colour space then four texture features (uniformity, correlation, contrast and entropy) are calculated for each HSV channels, which results in 12 feature vectors for each image. The following

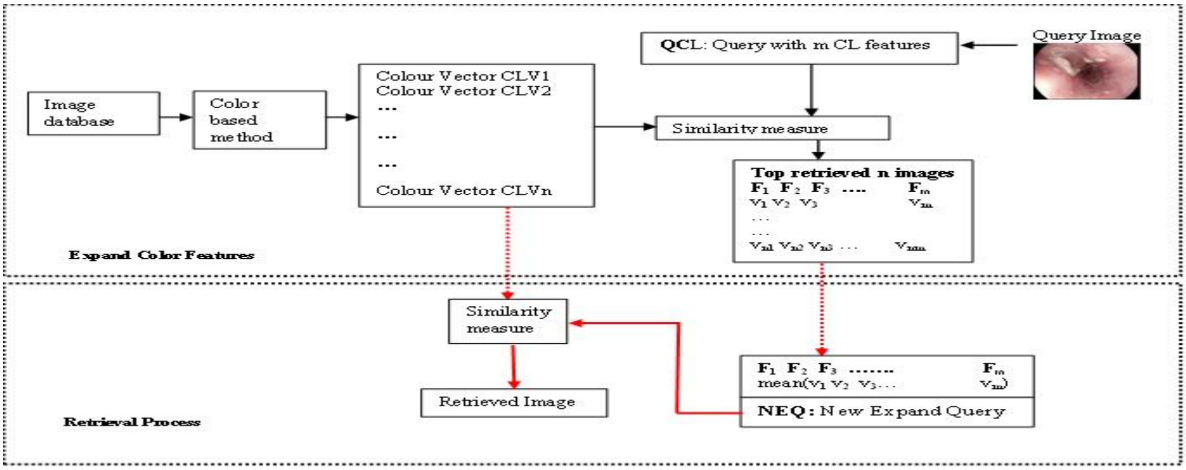


FIGURE 2. CLEBMR expansion and retrieval framework.

equations (7 to 10) describe the four texture feature functions implemented in this study. Let the p(i, j) = (p(i, j))/R is (i,j)*th* entry in normalised matrix, Ng is the number of distinct grey levels of the quantised image, μ and σ are mean and standard deviation for image intensity respectively, then the four texture functions are defined as follows, both colour moment functions are applied for each HSV channel separately.

$$\text{uniformity} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \{p(i, j)\}^2 \tag{7}$$

$$\text{correlation} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{(i - \mu)(j - \mu)\, p(i, j)}{\sigma^2} \tag{8}$$

$$\text{contrast} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)(i - j)^2 \tag{9}$$

$$\text{entropy} = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log \{p(i, j)\} \tag{10}$$

B. EXPANSION RETRIEVAL METHOD BASED ON MEAN VALUES

The first expansion method proposed in this study is implemented for colour and texture features; however, the proposed diagram in Figure 2 explains the whole idea for expansion colour features and formulating a new query to be used for the final searching process, and it could be easily considered or used for texture features as we did implicitly in this study. This method is based on the mean values of the top relevant images after a quick search using single image query (QCL) with all eighteen colour features as shown in Figure 1. The dimension of each feature vector, colour vector (CLV), for the images dataset in this case is eighteen features (m = 18). The average or mean value of each colour feature of the top ten relevant images (for each image class) is calculated and the new expand query (NEQ) is used for the final searching process and relevant images are retrieved. Simple numeric example of constructing NEQ is given in Figure 3 and the proposed algorithm for the colour expansion retrieval method is illustrated in Algorithm 1. A similar diagram framework and algorithms are implemented for the texture feature expansion.

	F1	F2	F3	...	F18
Img1	0.143	0.439	0.420	...	0.220
Img2	0.247	0.181	0.303	...	0.248
Img3	0.184	0.352	0.322	...	0.230
Img4	0.262	0.375	0.470	...	0.254
Img5	0.189	0.365	0.423	...	0.231
NEQ	**0.205**	**0.342**	**0.388**	...	**0.237**

FIGURE 3. NEQ based on mean values.

Algorithm 1 CLEBMR Expansion and Retrieval Method

Input: single image query for each class
Output: Recall and Precision
1: **Start**
2: **For** i = 1:m // number of images classes
3: Extract colour features for all images as in section (A) in the proposed method
3: Retrieve top similar (n = 10) images based on single (QCL) query image
4: Compute mean features values for top relevant images per each class
5: use mean values and formulate (NEQ) query image
6: Calculate the similarity measures
7: Calculate the avg(R)and avg(P)for each class and for whole images dataset
8: **End** //for
9: **End** // Start

C. EXPANSION RETRIEVAL METHOD BASED ON FEATURE SELECTION

Similar to the previous method, the expansion process here is based on the top retrieved image to assist in construction of a new expanded query image, but here it is based on selection of the most important features rather than on mean values. After performing a fast search using single image query (QCL), the top ten ranked (n = 10) images for each class are combined together with their class labels and used as input

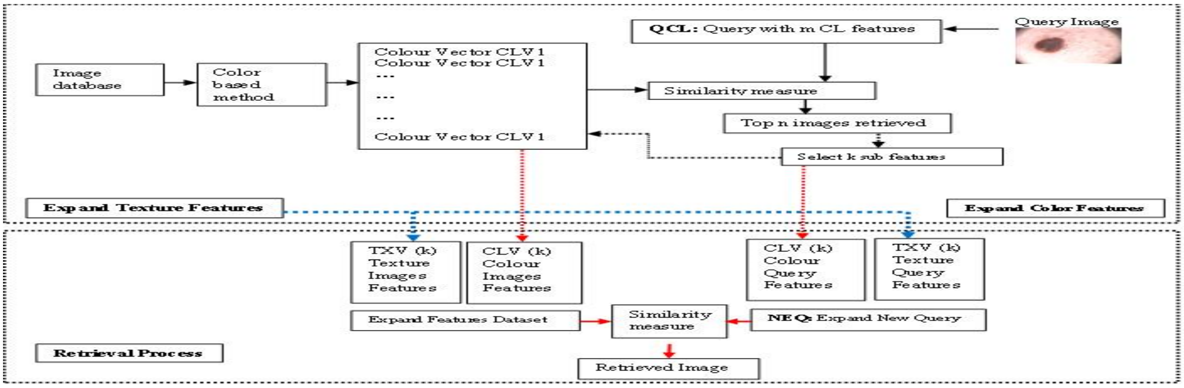


FIGURE 4. CLEBFSR expansion and retrieval framework.

for the feature selection algorithm, as in the next paragraph, which results in selecting only the sub-important features. The new dimension for most important features (k) is considered only for both the new expand query image and the whole image datasets, and the new expand query (NEQ) image is used for the final searching process. The framework diagram and algorithm are explained in Figures 4 and Algorithm 2 respectively.

Algorithm 2 CLEBFSR Expansion and Retrieval Method

Input: single image query for each class
Output: Recall and Precision
1: **Start**
2: **For** i = 1:m // number of images classes
3: Extract colour features for all images as in section (A) in the proposed method
3: Retrieve top similar (n = 10) images based on single (QCL) query image
4: Apply Feature Selection algorithm as in section (C) in the proposed method
5: Decrease dimensions of both image query and all images in dataset based on selected relevant features (k) and formulate new expand query image
6: Calculate the similarity measures
7: Calculate the avg(R) and avg(P)for each class and for whole images dataset
8: **End** //for
9: **End** // Start

1) FEATURES SELECTION

This step could be considered as a preprocess step and the main purpose is to select the most important image features from generated features. There are many choices to do that, and this study adopts the mutual information (MI) feature selection algorithm proposed by [31], and recently it has been used for classification accuracy [32]. This algorithm is a type of scoring method similar to a statistical dependency (SD) algorithm using entropy, and results in a scoring and ranking of features, according to which a chosen number of features having the highest values can be selected. The mutual information between the discretised feature values y and the class labels z is evaluated according to the following formula; the logarithm of this algorithm has been found preferable to the conventional MI measure such as SD method.

$$\mathrm{MI} = \sum\nolimits_{y \in \mathrm{Y}} \sum\nolimits_{z \in \mathrm{Z}} p\,(\mathrm{y}, \mathrm{z}) \log \left(\frac{\mathrm{p}(y, z)}{\mathrm{p}\,(y)\,\mathrm{p}\,(z)} \right) \tag{11}$$

D. SIMILARITY MEASURING

Similarity scores for both fast search and final search are calculated using Euclidean distance which has usually been used for calculating similarity. Let $\mathrm{X} = (\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_\mathrm{n})$ and $\mathrm{Y} = (\mathrm{y}_1, \mathrm{y}_2, \ldots, \mathrm{y}_\mathrm{n})$, two image vectors with n dimension, the similarity using Euclidean distance is computed as follows:

$$\text{similarity score} = \sqrt{\sum\nolimits_{\mathrm{i}=1}^{\mathrm{n}} (\mathrm{x}_\mathrm{i} - \mathrm{y}_\mathrm{i})^2} \tag{12}$$

For evaluation measure the two well-known metrics, precision and recall, were used as follows:

$$\text{Recall} = \frac{\text{number of relevant images retrieved}}{\text{total number of retrieved images}} \tag{13}$$

$$\text{Precision} = \frac{\text{number of relevant images retrieved}}{\text{total number of relevant images}} \tag{14}$$

E. IMAGES DATASETS

Experiments in this study use two of most common medical images, Kvasir and PH2 datasets. Pogorelov [33] consists of 4,000 coloured endoscopic images annotated by medical experts. The images of this dataset were grouped into eight different classes (500 images in each class) based on anatomy, pathology, or polyp removal procedures. The second dataset is a dermoscopic image database for research and benchmarking [34], and it has a set of 200 dermoscopic images divided into three classes. These datasets have been used in recent studies [35]–[37].

IV. EXPERIMENTAL RESULTS AND DISCUSSION

As we explained earlier in the main framework of the methodology there are two searching processes: fast search

TABLE 1. Kvasir dataset: colour results.

classes	CLR		CLEBMR		CLEBFSR	
	P	R	P	R	P	R
DLP	0.560	0.011	0.700	0.014	0.900	0.018
DRM	0.380	0.008	0.600	0.012	0.960	0.019
Esophagitis	0.480	0.010	0.700	0.014	0.900	0.018
Normal Caecum	0.900	0.018	0.900	0.018	1	0.020
Normal Pylorus	0.940	0.019	0.600	0.012	1	0.020
Normal Z Line	0.600	0.012	0.800	0.016	0.960	0.019
Polyps	0.580	0.012	0.700	0.014	0.980	0.020
Ulcerative Colitis	0.450	0.009	0.200	0.004	0.960	0.019
Average	**0.611**	**0.012**	**0.650**	**0.013**	**0.958**	**0.019**
	Ahmad, J., et al. [29] P@50 : **0.740**			Hu, H., et al. [30] P@10 : **0.940**		

TABLE 2. Kvasir dataset: texture results.

classes	TXR		TXEBMR		TXEBFSR	
	P	R	P	R	P	R
DLP	0.500	0.010	0.400	0.008	1	0.02
DRM	0.560	0.011	0.500	0.01	0.540	0.011
Esophagitis	0.820	0.016	1	0.020	1	0.02
Normal Caecum	0.780	0.016	0.900	0.018	0.920	0.018
Normal Pylorus	0.540	0.011	0.400	0.008	0.540	0.011
Normal Z Line	0.440	0.009	0.400	0.008	1	0.020
Polyps	0.500	0.010	0.700	0.014	1	0.020
Ulcerative Colitis	0.440	0.009	0.400	0.008	0.560	0.011
Average	**0.573**	**0.0115**	**0.588**	**0.012**	**0.820**	**0.0164**

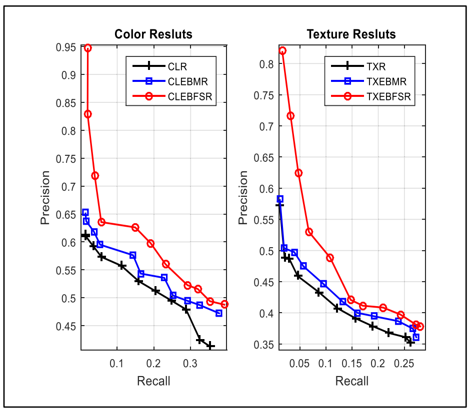


FIGURE 5. Recall and precision of colour and texture expansion for Kvasir dataset.

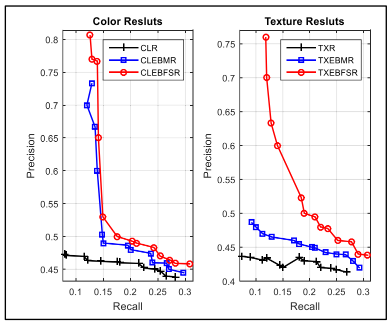


FIGURE 6. Recall and precision of colour and texture expansion for PH2 dataset.

performed using single query image selected randomly from each class of images followed by final search process based on new expanded queries.

The output retrieved images for this search are used as input for the expansion process in an automatic way without any user interaction or suggestion which represents one of the main advantages of our proposed methods. The new formulated query images will then be used for the final searching process and all evaluation metrics will calculated from the result of this search. For colour features, experiments were performed and evaluated, colour retrieval was based on colour features only (CLR) associated with our two proposed methods; colour expansion based on mean retrieval (CLEBMR) and colour expansion based on feature selection retrieval (CLEBFSR) methods. Similarly, for texture features there are three experiments also: texture retrieval based on texture feature only (TXR), texture expansion based on mean retrieval (TXEBMR) and texture expansion based on feature selection retrieval (TXEBFSR). Colour retrieval (CLR) and texture retrieval (TXR) represent the base retrieval which was enhanced and increased using two expansion methods; evaluation measures after the last searching process were calculated as top 10 retrieved images. Recall and precision values of the three retrieval methods for Kvasir dataset were given in Tables 1 and 2 respectively. In these two tables recall and precision values for each class are computed as well as the mean of recall and precision values for all eight classes. First looking for these two tables we observe that CLEBMR

TABLE 3. PH2 dataset: colour results.

classes	CLR		CLEBMR		CLEBFSR	
	P	R	P	R	P	R
DLP	0.480	0.060	0.600	0.075	0.920	0.115
DRM	0.460	0.057	0.700	0.087	0.900	0.113
Esophagitis	0.480	0.120	0.900	0.225	0.600	0.150
Average	**0.473**	**0.079**	**0.733**	**0.129**	**0.807**	**0.126**

TABLE 4. PH2 dataset: texture results.

classes	TXR		TXEBMR		TXEBFSR	
	P	R	P	R	P	R
DLP	0.380	0.048	0.260	0.025	0.710	0.120
DRM	0.420	0.053	0.400	0.050	0.760	0.060
Esophagitis	0.480	0.120	0.800	0.200	0.810	0.180
Average	**0.427**	**0.074**	**0.487**	**0.092**	**0.760**	**0.120**

TABLE 5. Ranking of retrieval methods based on Kendall W test for precision value.

Data set	Retrieval Methods	W	P	Ranking
Kvasir	Colour based retrieval	0.810	0.002	CLEBFSR > CLEBMR > CLR
	Texture based retrieval	0.508	0.020	TXEBFSR > TXEBMR > TXR
PH2	Colour based retrieval	0.770	0.090	CLEBFSR > CLEBMR > CLR
	Texture based retrieval	0.778	0.097	TXEBFSR > TXEBMR > TXR

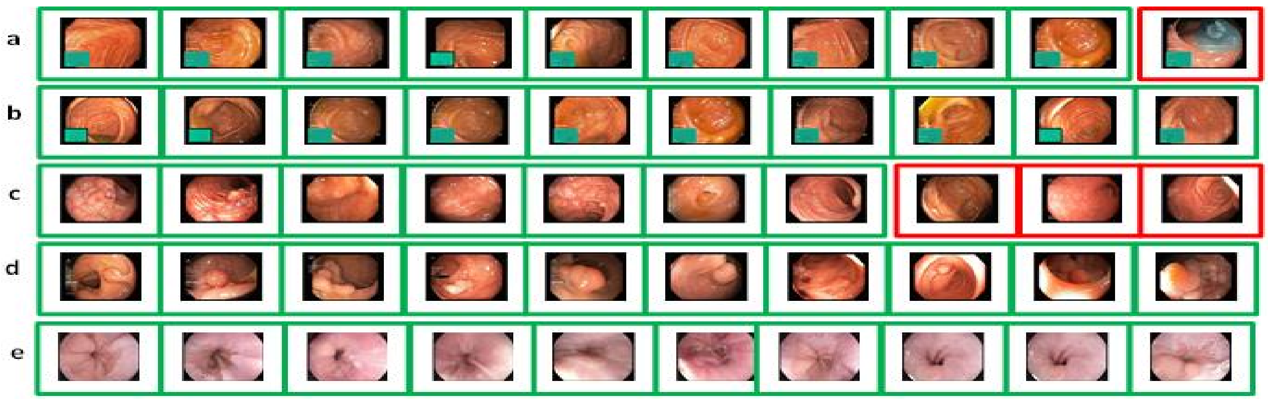


FIGURE 7. Samples image retrieved for classes of Kvasir dataset: (a) normal Caecum class (9 out of 10) using CLEBMR, (b) normal Caecum class (10 out of 10) using CLEBFSR, (c) Polyps (7 out of 10) using TXEBMR, (d) Polyps (10 out of 10) using TXEBFSR, (e) normal Z Line (10 out of 10) using TXEBFSR.

and CLEBFSR outperformed the base (CLR) method in many classes of images as well as in mean values. Recall and precision graphs for this dataset were shown in Figure 5 for both colour and texture features; red lines of the second proposed method have best behaviour and performance. Similarly, recall and values of three methods for the PH2 dataset are shown in Table 3 and Table 4. For all three classes and the mean values, our proposed method performed better, even though the measures have low values compared with the Kvasir dataset. Similarly Figure 6 shows the recall and precision graph for the PH2 dataset. Kendall W concordance test [38] provides more quantitative testing and it has been used in our previous works [39]–[42]. This test was developed to measure the level of agreement between multiple sets of rankings of the same set of objects and it was used to evaluate the effectiveness of different retrieval methods. This test was implemented in this study, where the image classes were considered judges in the present context, and the precision rates of the three different methods were considered objects. This test takes precision rates of each of the three methods as input and it has two outputs: Kendall coefficient and associated level of significance.

Kendall coefficient, ranging from 0 (no agreement between a set of ranks) to 1 (complete agreement), the second output parameter indicates whether this coefficient value could have occurred by chance. If the significant value is less than or close to 0.02 for the Kvasir dataset and 0.15 for PH2 dataset (since we used the 0.02 or 0.15 cut-off values for

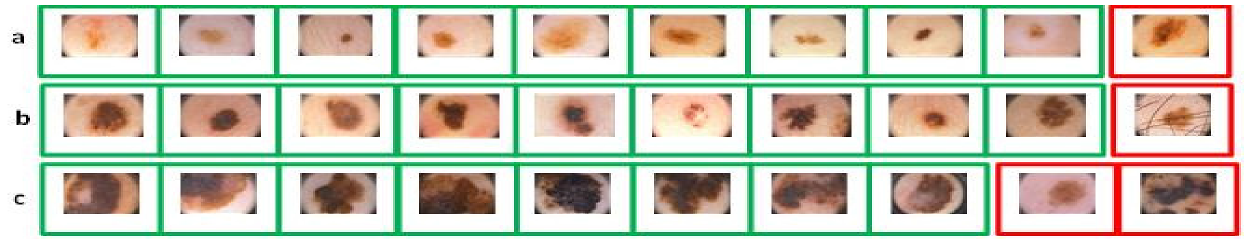
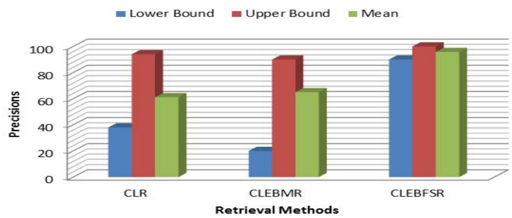
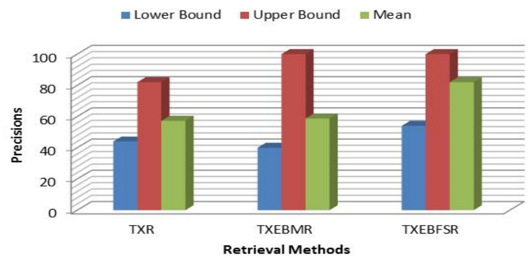


FIGURE 8. Samples image retrieved for classes of PH2 dataset: (a) DLP (9 out of 10) using CLEBFSR, (b) DRM (9 out of 10) using CLEBFSR, (c) Esophagitis (8 out of 10), TXEBMR.

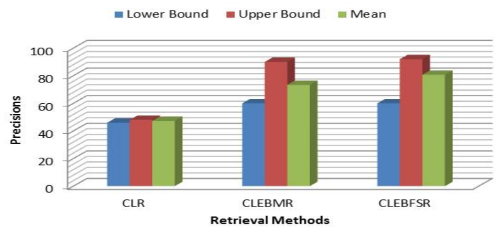


a) Confidence bounds for precision of colour retrieval methods with the Kvasir dataset

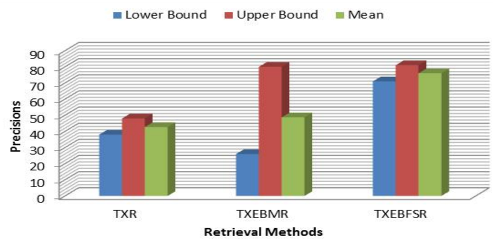


b) Confidence bounds for precision of texture retrieval methods with the Kvasir dataset

FIGURE 9. a) Confidence bounds for precision of colour retrieval methods with the Kvasir dataset. b) Confidence bounds for precision of texture retrieval methods with the Kvasir dataset.



a) Confidence bounds for precision of colour retrieval methods with the PH2 dataset



b) Confidence bounds for precision of texture retrieval methods with the PH2 dataset

FIGURE 10. a) Confidence bounds for precision of colour retrieval methods with the PH2 dataset. b) Confidence bounds for precision of texture retrieval methods with the PH2 dataset.

two datasets, respectively, which was equivalent to the top 10 retrieved images), then an overall ranking of the graded items may be given. The result of this test for all three methods with two datasets is shown in Table 5; in this table CLEBFSR has the best performance with agreement or confidence of 81% for the Kvasir dataset, while TXEBFSR performed better with 77.8% agreement for PHs dataset. Ranking of different methods with two datasets shows the good retrieving performance of the two proposed methods. Samples of top retrieved images of some classes for two datasets are shown in Figures 7 and 8. More than one class in both datasets has (10 out 10) successful retrieved images; the successful retrieved images are very important and will be helpful and assist more in the diagnosis process. Finally more statistical certainty in terms of the mean, lower and upper bounds of the confidence intervals of different retrieval methods for colour and texture features is shown in Figure 9 (a and b) and Figure 10 (a and b). These figures indicate that our proposed retrieval methods do better than using absolute colour or texture features only.

V. CONCLUSION

In this study we developed a new expansion method based on top-ranked images for content based medical image retrieval implemented on colour and texture features. The proposed expansion method is a fully automated process without any interaction or feedback from a user; the end user only needs to post a single query image. Experiments with two medical image datasets proved that these methods provide a simple way of enhancing and improving the retrieval effectiveness of related medical images. While these methods were implemented on colour and texture feature vectors, further efforts could focus on shape features or fusion combination of more than one feature spaces, but it has to be implemented with a careful and weighted scenario since the traditional fusion and combination of colour and texture results in poor retrieval performance as we have tried here. Also a hybrid or combination expansion scheme of our proposed methods in terms of formulation of a new expanded query image based on calculation of the mean of selected important features could gain good retrieval performance.

ACKNOWLEDGMENT

This research was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, under Grant No. (DF-093-830-1441). The authors, therefore, gratefully acknowledge the DSR for their technical support.

REFERENCES

[1] K. Doi, "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Computerized Med. Imag. Graph.*, vol. 31, nos. 4–5, pp. 198–211, Jun. 2007.

[2] H. Zaidi, H. Vees, and M. Wissmeyer, "Molecular PET/CT imaging-guided radiation therapy treatment planning," *Acad. Radiol.*, vol. 16, no. 9, pp. 1108–1133, Sep. 2009.

[3] C. D. Marcus, V. Ladam-Marcus, C. Cucu, O. Bouché, L. Lucas, and C. Hoeffel, "Imaging techniques to evaluate the response to treatment in oncology: Current standards and perspectives," *Crit. Rev. Oncol./Hematol.*, vol. 72, no. 3, pp. 217–238, Dec. 2009.

[4] A. Kumar, S. Dyer, J. Kim, C. Li, P. H. W. Leong, M. Fulham, and D. Feng, "Adapting content-based image retrieval techniques for the semantic annotation of medical images," *Computerized Med. Imag. Graph.*, vol. 49, pp. 37–45, Apr. 2016.

[5] N. Hor and S. Fekri-Ershad, "Image retrieval approach based on local texture information derived from predefined patterns and spatial domain information," 2019, *arXiv:1912.12978*. [Online]. Available: http://arxiv.org/abs/1912.12978

[6] M. Owais, M. Arsalan, J. Choi, and K. R. Park, "Effective diagnosis and treatment through content-based medical image retrieval (CBMIR) by using artificial intelligence," *J. Clin. Med.*, vol. 8, no. 4, p. 462, Apr. 2019.

[7] P. Tschandl, G. Argenziano, M. Razmara, and J. Yap, "Diagnostic accuracy of content-based dermatoscopic image retrieval with deep classification features," *Brit. J. Dermatol.*, vol. 181, no. 1, pp. 155–165, Jul. 2019.

[8] M. Sadeghi, P. Chilana, J. Yap, P. Tschandl, and M. S. Atkins, "Using content-based image retrieval of dermoscopic images for interpretation and education: A pilot study," *Skin Res. Technol.*, vol. 26, no. 4, pp. 503–512, Jul. 2020.

[9] A. Shinde, A. Rahulkar, and C. Patil, "Content based medical image retrieval based on new efficient local neighborhood wavelet feature descriptor," *Biomed. Eng. Lett.*, vol. 9, no. 3, pp. 387–394, Aug. 2019.

[10] P. Kaur and R. K. Singh, "A panoramic view of content-based medical image retrieval system," in *Proc. Int. Conf. Intell. Eng. Manage. (ICIEM)*, Jun. 2020, pp. 187–192.

[11] H. Qazanfari, H. Hassanpour, and K. Qazanfari, "A short-term learning framework based on relevance feedback for content-based image retrieval," in *Proc. 3rd Iranian Conf. Intell. Syst. Signal Process. (ICSPIS)*, Dec. 2017, pp. 136–140.

[12] I. Banerjee, C. Kurtz, A. E. Devorah, B. Do, D. L. Rubin, and C. F. Beaulieu, "Relevance feedback for enhancing content based image retrieval and automatic prediction of semantic image features: Application to bone tumor radiographs," *J. Biomed. Informat.*, vol. 84, pp. 123–135, Aug. 2018.

[13] A. Ahmed, "Implementing relevance feedback for content-based medical image retrieval," *IEEE Access*, vol. 8, pp. 79969–79976, 2020.

[14] H. Karamti, M. Tmar, M. Visani, T. Urruty, and F. Gargouri, "Vector space model adaptation and pseudo relevance feedback for content-based image retrieval," *Multimedia Tools Appl.*, vol. 77, no. 5, pp. 5475–5501, Mar. 2018.

[15] C. Carpineto and G. Romano, "A survey of automatic query expansion in information retrieval," *ACM Comput. Surveys*, vol. 44, no. 1, pp. 1–50, Jan. 2012.

[16] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman, "Total recall: Automatic query expansion with a generative feature model for object retrieval," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.

[17] H. K. Azad and A. Deepak, "Query expansion techniques for information retrieval: A survey," *Inf. Process. Manage.*, vol. 56, no. 5, pp. 1698–1735, Sep. 2019.

[18] M. E. Houle, X. Ma, V. Oria, and J. Sun, "Query expansion for content-based similarity search using local and global features," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 13, no. 3, pp. 1–23, Aug. 2017.

[19] N. Kondylidis, M. Tzelepi, and A. Tefas, "Exploiting tf-idf in deep convolutional neural networks for content based image retrieval," *Multimedia Tools Appl.*, vol. 77, no. 23, pp. 30729–30748, Dec. 2018.

[20] R. Imbriaco, C. Sebastian, E. Bondarev, and P. de With, "Aggregated deep local features for remote sensing image retrieval," *Remote Sens.*, vol. 11, no. 5, p. 493, Feb. 2019.

[21] O. Chum, A. Mikulik, M. Perdoch, and J. Matas, "Total recall II: Query expansion revisited," in *Proc. CVPR*, Jun. 2011, pp. 889–896.

[22] A. Gordo, F. Radenovic, and T. Berg, "Attention-based query expansion learning," 2020, *arXiv:2007.08019*. [Online]. Available: http://arxiv.org/abs/2007.08019

[23] B. Feng, J. Cao, Z. Chen, Y. Zhang, and S. Lin, "Multi-modal query expansion for Web video search," in *Proc. 33rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, 2010, pp. 721–722.

[24] Y.-H. Kuo, K.-T. Chen, C.-H. Chiang, and W. H. Hsu, "Query expansion for hash-based image object retrieval," in *Proc. 17th ACM Int. Conf. Multimedia (MM)*, 2009, pp. 65–74.

[25] U. Erkut, F. Bostancioglu, M. Erten, A. M. Ozbayoglu, and E. Solak, "HSV color histogram based image retrieval with background elimination," in *Proc. 1st Int. Informat. Softw. Eng. Conf. (UBMYK)*, Nov. 2019, pp. 1–5.

[26] P. Maheshwary and N. Srivastava, "Prototype system for retrieval of remote sensing images based on color moment and gray level co-occurrence matrix," *Int. J. Comput. Sci.*, vol. 3, pp. 20–23, Aug. 2009.

[27] J. S. Weszka and C. R. A. Dyer Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 4, pp. 269–285, Apr. 1976.

[28] L. Armi and S. Fekri-Ershad, "Texture image analysis and texture classification methods—A review," 2019, *arXiv:1904.06554*. [Online]. Available: http://arxiv.org/abs/1904.06554

[29] J. Ahmad, K. Muhammad, M. Y. Lee, and S. W. Baik, "Endoscopic image classification and retrieval using clustered convolutional features," *J. Med. Syst.*, vol. 41, no. 12, p. 196, Dec. 2017.

[30] H. Hu, W. Zheng, X. Zhang, X. Zhang, J. Liu, W. Hu, H. Duan, and J. Si, "Content-based gastric image retrieval using convolutional neural networks," *Int. J. Imag. Syst. Technol.*, Aug. 2020, doi: 10.1002/ima.22470.

[31] J. Pohjalainen, O. Räsänen, and S. Kadioglu, "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits," *Comput. Speech Lang.*, vol. 29, no. 1, pp. 145–171, Jan. 2015.

[32] A. Ahmed and S. Malebary, "Feature selection and the fusion-based method for enhancing the classification accuracy of SVM for breast cancer detection," *Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 11, p. 55, 2019.

[33] K. Pogorelov, K. R. Randel, C. Griwodz, S. L. Eskeland, T. de Lange, D. Johansen, C. Spampinato, D. T. Dang-Nguyen, M. Lux, P. T. Schmidt, and M. Riegler, "Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection," in *Proc. 8th ACM Multimedia Syst. Conf.*, 2017, pp. 164–169.

[34] T. Mendoncÿa, P. M. Ferreira, J. Marques, A. R. S. Marcÿal, and J. Rozeira, "A dermoscopic image database for research and benchmarking," presented at the Proc. PH2 IEEE EMBC, 2013.

[35] H. Kasban and D. H. Salama, "A robust medical image retrieval system based on wavelet optimization and adaptive block truncation coding," *Multimedia Tools Appl.*, vol. 78, no. 24, pp. 35211–35236, Dec. 2019.

[36] J. Ahmad, K. Muhammad, and S. W. Baik, "Medical image retrieval with compact binary codes generated in frequency domain using highly reactive convolutional features," *J. Med. Syst.*, vol. 42, no. 2, p. 24, Feb. 2018.

[37] M. E. Celebi, N. Codella, and A. Halpern, "Dermoscopy image analysis: Overview and future directions," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 2, pp. 474–478, Mar. 2019.

[38] S. Sidney, "Nonparametric statistics for the behavioral sciences," *J. Nervous Mental Disease*, vol. 125, no. 3, p. 497, 1957.

[39] A. Ahmed, F. Saeed, N. Salim, and A. Abdo, "Condorcet and borda count fusion method for ligand-based virtual screening," *J. Cheminformatics*, vol. 6, no. 1, p. 19, Dec. 2014.

[40] A. Ahmed, A. Abdo, and N. Salim, "Ligand-based virtual screening using Bayesian inference network and reweighted fragments," *Sci. World J.*, vol. 2012, pp. 1–7, May 2012.

[41] A. Abdo, F. Saeed, H. Hamza, A. Ahmed, and N. Salim, ''Ligand expansion in ligand-based virtual screening using relevance feedback,'' *J. Comput.-Aided Mol. Design*, vol. 26, no. 3, pp. 279–287, Mar. 2012.

[42] A. Abdo, N. Salim, and A. Ahmed, ''Implementing relevance feedback in ligand-based virtual screening using Bayesian inference network,'' *J. Biomolecular Screening*, vol. 16, no. 9, pp. 1081–1088, Oct. 2011.

ALI AHMED (Member, IEEE) received the B.Sc. degree in computer engineering from Karary University, Sudan, the M.Sc. degree in computer science from Khartoum University, Sudan, and the Ph.D. degree in computer science from UTM, Malaysia.

He held a postdoctoral position at the UTM. He has been worked as an assistant professor for more than four years (teaching and supervising master's degree students). He is currently an Associate Professor with the Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, Saudi Arabia. His research interests include machine learning, data mining, information retrieval, and image processing and retrieval. He is a member of the ACM and SEA, Sudan.

SHARAF J. MALEBARY (Member, IEEE) received the Ph.D. degree in computer science and engineering from the University of South Carolina, USA.

He is currently an Assistant Professor with King Abdulaziz University, Rabigh. He is a very enthusiastic individual, and has shown and proven strong leadership potentials. As a result, he serves as the Head of Department of Information Technology. Moreover, he was nominated to serve as the Vice-Dean for Graduate Studies and Scientific Research at the Faculty of Computing and Information Technology, King Abdulaziz University. His research interests include autonomous systems, wireless communications, artificial intelligence, and machine learning.