An Efficient Two-layer based Technique for Content-based Image Retrieval

Fawzi Abdul Azeez Salih[1], Alan Anwer Abdulla[2,3]*

[1]Department of Computer Science, College of Science, University of Sulaimani, Sulaimani,Iraq, [2]Department of Information Technology, College of Commerce, University of Sulaimani, Sulaimani, Iraq, [3]Department of Information Technology, University College of Goizha, Sulaimani, Iraq

ABSTRACT

The rapid advancement and exponential evolution in the multimedia applications raised the attentional research on content-based image retrieval (CBIR). The technique has a significant role for searching and finding similar images to the query image through extracting the visual features. In this paper, an approach of two layers of search has been developed which is known as two-layer based CBIR. The first layer is concerned with comparing the query image to all images in the dataset depending on extracting the local feature using bag of features (BoF) mechanism which leads to retrieve certain most similar images to the query image. In other words, first step aims to eliminate the most dissimilar images to the query image to reduce the range of search in the dataset of images. In the second layer, the query image is compared to the images obtained in the first layer based on extracting the (texture and color)-based features. The Discrete Wavelet Transform (DWT) and Local Binary Pattern (LBP) were used as texture features. However, for the color features, three different color spaces were used, namely RGB, HSV, and YCbCr. The color spaces are utilized by calculating the mean and entropy for each channel separately. Corel-1K was used for evaluating the proposed approach. The experimental results prove the superior performance of the proposed concept of two-layer over the current state-of-the-art techniques in terms of precision rate in which achieved 82.15% and 77.27% for the top-10 and top-20, respectively.

Index Terms: CBIR, Feature Extraction, Color Descriptor, DWT, LBP

1. INTRODUCTION

Beside content-based image retrieval (CBIR), digital image processing plays a vital role in numerous areas such as processing and analyzing medical image [1], image inpainting [2], pattern recognition [3], biometrics [4], multimedia security [5], and information hiding [6]. In the area of image processing and computer vision, CBIR has grown increasingly as an advanced research topic. CBIR refers to the system which retrieves similar

images of a query image from the dataset of images without any help of caption and/or description of the images [7]. There are two main mechanisms for image retrieval which are text-based image retrieval (TBIR) and CBIR [8]. TBIR was first introduced in 1970 as search and retrieve images from image dataset [9]. In such a kind of image retrieval mechanism, the images are denoted by text and then the text is used to retrieve or search for the images. The TBIR method depends on the manual text search or keyword matching of the existing image keywords and the result has been relied on the human labeling of the images. TBIR approach requires information such as image keyword, image location, image tags, image name, and other information related to the image. Human involvement is needed in the challenging process of entering information of the images in the dataset. The drawbacks of TBIR are as follows: 1- It leads to inaccurate results if human has been

Corresponding author's e-mail: Alan Anwer Abdulla, Department of Information Technology, College of Commerce, University of Sulaimani, Sulaimani, Iraq; Department of Information Technology, University College of Goizha, Sulaimani, Iraq. E-mail: alan.abdulla@univsul.edu.iq

doing datasets annotation process incorrectly, 2- single keyword of image information is not effective to transfer the overall image description, and 3- it is based on manual annotation of the images, which is time consuming [10]. Researchers introduced CBIR as a new mechanism for image retrieval to overcome the above-mentioned limitations of TBIR. It is considered as a popular technique to retrieve, search, and browse images of query information from a broad dataset of images. In CBIR, the image information, visual features such as low-level features (color, texture, and/or shape), or bag of features (BoF) have been extracted from the images to find the most similar images in the dataset [11]. Fig. 1 illustrates the general block diagram of the CBIR mechanism [12].

Fig. 1 shows the block diagram of basic CBIR system that involves two phases: Feature extraction and feature matching. The first phase involves extracting the image features while the second phase involves matching these features [13]. Feature extraction is the process of extracting features from the dataset of images and stored in feature vector and also extracting features from the query image. On the other hand, feature matching is the process of comparing the extracted features from the query image to the extracted features from images in the dataset using similarity distance measurement. The corresponding image in the dataset is considered as a match/similar image to the query image if the distance between feature

vector of the query image and the image in the dataset is small enough. Thus, the matched images are then ranked based on the similarity index from the smallest distance value to the largest one. Eventually, the retrieved images are selected according to the lowest distance value. The essential objective of the CBIR systems is improving the efficiency of the system by increasing the performance using the combination of features [9]. Image features can be categorized into two types: Global features and local features. Global features extract information from the entire image while local features work locally which are focused on the key points in images [14]. For the large image dataset, image relevant to the query image are very few. Therefore, the elimination of irrelevant images is important. The main contribution of this research is first eliminating the irrelevant images in the dataset and then finds the most similar/matches images from the rest of the remained images. The reminder of the paper is organized as follows: Section 2 discusses the related work. Section 3 introduces a background about the techniques used in the proposed approach. Section 4 presents the proposed CBIR approach. Section 5 shows the experimental results. Finally, section 6 gives the conclusions.

2. RELATED WORK

Studies related to the developed techniques of CBIR have been researched a lot and they mainly focused on analyzing and investigating the interest points/areas such as corners, edges, contours, maxima shapes, ridges, and global features [15]. Some of those developed approaches are concerned on combining/fusing certain types of the extracted features, since such kind of strategy has an impact on describing the image content efficiently [13], [16]. This section reviews the most important and relevant existing works on CBIR. The main competition in this research area is increasing the precision rate that refers to the efficiency of retrieving the most similar images correctly. Kato *et al.* were first to investigate this field of study in 1992, who developed a technique for sketch retrieval, similarity retrieval, and sense retrieval to support visual interaction [17]. Sketch retrieval accepts the image data of sketches, similarity retrieval evaluates the similarity based on the personal view of each user, and sense retrieval evaluates based on the text data and the image data at content level based on the personal view. Yu *et al.*, in 2013, proposed an effective image retrieval system based on the (BoF) model, depending on two ways of integrating [18]. Scale-invariant feature transform (SIFT) and local binary pattern (LBP) descriptors were integrated in one hand, and histogram of oriented gradients (HOG) and LBP descriptors were integrated on the other hand. The first proposed integration, namely SIFT-LBP, provides better precision rate in which reached 65% for top-20 using Jaccard similarity measurement.
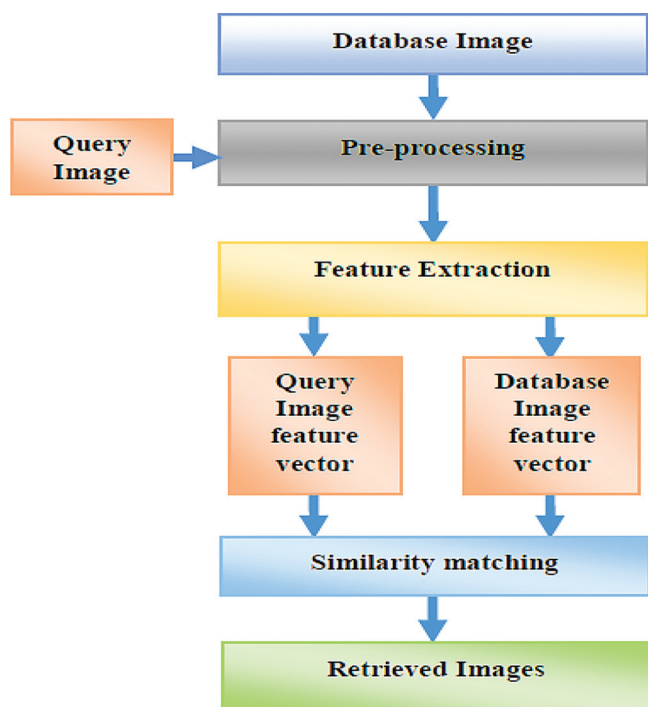

Fig. 1. General block diagram of CBIR mechanism.

Shrivastava *et al.*, in 2014, introduced a new scheme for CBIR based on region of interest ROI codes and an effective feature set consisting of a dominant color and LBP were extracted from both query image and dataset of images. This technique achieved, using Euclidean distance measurement, a precision rate of 76.9% for top-20 [19]. DWT as a global feature, and gray level co-occurrence matrix (GLCM), as local feature, was extracted and fused in the algorithm introduced by Gupta *et al.*, in 2015, and as a result, a precision rate of 72.1 % was obtained for top-20 using Euclidean distance [20]. Another technique was introduced by Navabi *et al.*, in 2017, for CBIR which based on extracting color and texture features. The technique used color histogram and color moment as color feature. The principal component analysis (PCA) statistical method was applied for the dimension's reduction. Finally, Minkowski distance measurement was used to find most similar images. As reported, this technique achieved the precision rate of 62.4% for top-20 [21]. Nazir *et al.*, in 2018, proposed a new CBIR technique by fusing the extracted color and texture features [22]. Color Histogram (CH) was used to extract a color information, and DWT as well as edge histogram descriptor (EDH) were used to extract texture features. As authors claimed, this technique achieved a precision rate of 73.5% for top-20 using Manhattan distance measurement. Pradhan *et al.*, in 2019, developed a new CBIR scheme based on multi-level colored directional motif histogram (MLCDMH) [23]. This scheme extracts local structural features at three different levels. The image retrieval performance of this proposed scheme has been evaluated using different Corel/natural, object, texture, and heterogeneous image datasets. For the Corel-1k, the precision rate of 64% and 59% was obtained for top-10 and top-20, respectively. Recently, Sadique *et al.*, in 2019, developed a new CBIR technique by extracting global and local features [7]. A combination of speeded up robust features (SURF) descriptor with color moments, as local feature, and modified GLCM, as global feature, leads this technique to obtain 70.48% of the precision rate for top-20 using Manhattan similarity measurement. Continuously, in 2019, Khawaja *et al.* proposed another technique for CBIR using object and color features [24]. Authors claimed that this technique outperformed in certain categories of the benchmark datasets Caltech-101 and Corel-1000, and it gained 76.5% of the precision rate for top-20 using Euclidean distance. Different from the previous techniques discussed above, Qazanfari *et al.*, in 2019, investigated HSV color space for developing CBIR technique [25]. As reported in this work, the human visual system is very sensitive to the color as well as edge orientation, and also color histogram and color difference histogram (CDH) are two kinds of low-level feature extraction which are meaningful representatives of the image color and edge orientation information. This proposed technique used Canberra distance measurement to measure the similarity between the extracted feature of the both query image and images in the dataset. This technique achieved 74.77% of the precision rate for the top-20 using Euclidean distance similarity measurement. Rashno *et al.*, in 2019, developed an algorithm in which HSV, RGB, and norm of low frequency components were used to extract color features, and DWT was used to extract texture features [26]. Accordingly, ant colony optimization (ACO) feature selection technique was used to select the most relevant features. Eventually, Euclidian distance measurement was used to measure the similarity between query and images in the dataset. The results reported in this work showed that this approach reached the precision rate of 60.79% using Euclidean distance for the top-20. Finally, Aiswarya *et al.*, in 2020, proposed a CBIR technique which uses a multi-level stacked Autoencoders for feature selection and dimensionality reduction [27]. A query image space is created first before the actual retrieval process by combining the query image as well as similar images from the local image dataset (images in device gallery) to maintain the image saliency in the visual contents. The features corresponding to the query image space elements are searched against the characteristics of images in a global dataset. This technique achieved the precision rate of 67% for top-10.

3. BACKGROUND

This section aims to provide detailed background information about important techniques, used in the proposed approach presented in this paper, such as SURF feature descriptor, color-based features, texture-based features, and feature matching techniques.

3.1. SURF Feature Descriptor

There are many feature descriptors available and SURF is one of the most common and significant feature descriptors in which can be considered as a local feature. In comparison with global features such as color, texture, and shape; local features can provide more detailed characters in an image. The rotation and scale invariant descriptor can perform better in terms of distinctiveness, repeatability, and robustness [12]. SURF is used in many applications such as bag of feature (BoF) which is used and success in image analysis and classification [28]. In the BoF technique, the SURF descriptor is sometimes used first to extract local features. Then K-means clustering is used to initialize M center point to create M visual words. The K-means clustering algorithm takes the feature space as input and reduces it to the M cluster as output. Then, the image is represented as a code word histogram by mapping the local features into
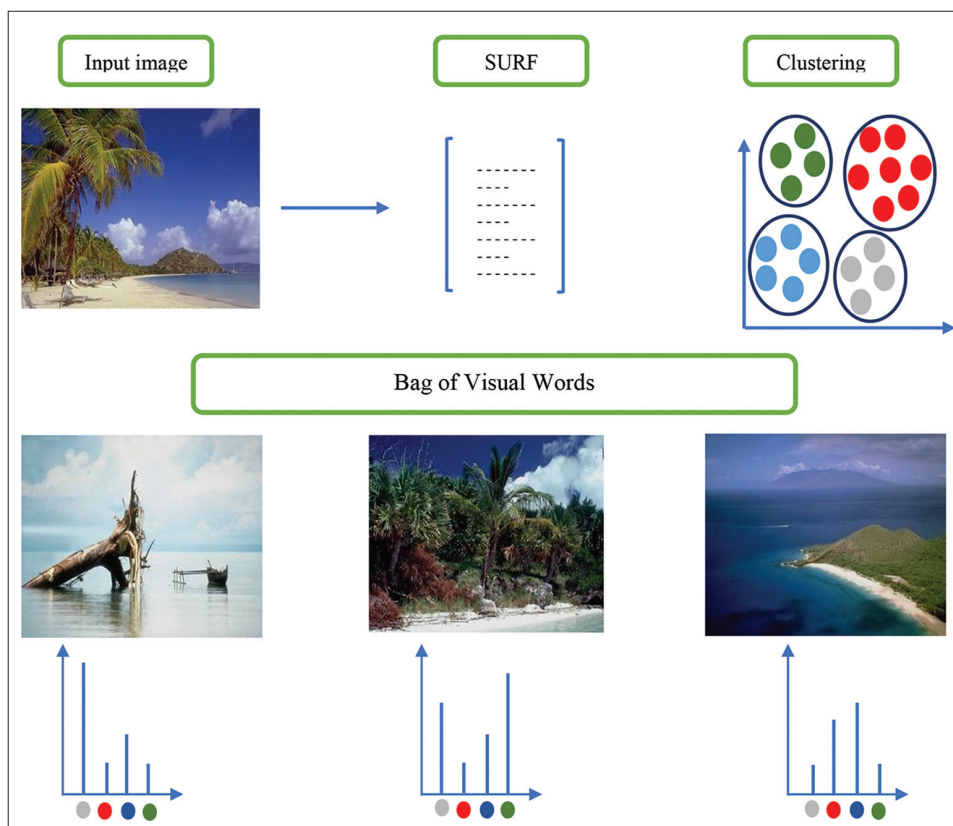
Fig. 2. Methodology of the BoF technique for representing Image in CBIR.

a vocabulary [28]. Fig. 2 illustrates the methodology of the image representation based on the BoF model.

SURF features are extracted from database images, then the k-means clustering algorithm takes feature space as input and reduces it into clusters as output. The center of each cluster is called a visual word and the combination of visual words formulates the dictionary, which is also known as codebook or vocabulary. Finally, using these visual words of the dictionary, the histogram is constructed using visual words of each image. The histogram of v visual words is formed from each image. After that, resultant information in the form of histograms is added to the inverted index of the BoF model [25].

3.2. Texture-based Features Extraction

Texture-based features can be considered as a powerful low-level feature for image search and retrieval applications. There are many works have been developed on texture analysis, classification, and segmentation for the last four decades. Yet, there is no unique definition for the texture-based features. Texture is an attribute representing the spatial arrangement of the gray levels of the pixels in a region or image. In other words, texture-based features can be used to separate and extract prominent regions of interest in an image and apply

to the visual patterns that have properties of homogeneity independent of a single color or intensity [9]. Texture analysis methods can be categorized into statistical, structural, and spectral [15]. DWT and LBP are the two methods of texture feature extraction used in this work.

3.2.1. Discrete wavelet transform (DWT)

The DWT is considered to be an efficient multiresolution technique and it is easy to compute [29]. The signal for each level decomposed into four frequency sub-bands which are: Low of low (LL), low of high (LH), high of low (HL), and high of high (HH) [30]. DWT is used to change an image from the spatial domain into the frequency domain, the structure of the DWT is illustrated in Fig. 3 [22], [26].

Wavelet transform could be applied to images as 2-dimensional signals. To refract an image into k level, first the transform is applied on all rows up to k level while columns of the image are kept unchanged. Then, this task is applied on columns while keeping rows unchanged. In this manner, frequency components of the image are obtained up to k level. These frequency components in various levels let us to better analyze original image or signal [26]. For more details about DWT, you can see [31].

3.2.2. Local binary pattern (LBP)

The concept of LBP was originally proposed by Ojala *et al.* in 1996 [29], [32]. LBP can be considered as a texture analysis approach unifying structural and statistical models. The characteristic of LBP is that LBP operator is invariant to monotonic gray-level changes [33]. In the process of LBP calculation, firstly a 3×3 grid of image is selected, and then the intensity value of the center pixel can be computed using the intensity values of its neighboring pixels based on the following equations [34]:

$$Lbp_{n\ \{1,2,3,\dots,9\}} = \sum_{k=0}^{n-1} 2^k \times fn\left(I_k - I_c\right) \qquad (1)$$

$$fn\left(a-b\right) = \begin{cases} 1, & \text{if } \left(a-b\right) \geq 0 \\ 0, & \text{Otherwise} \end{cases} \qquad (2)$$

where, n is the number of neighboring pixels around the center pixel. I_k is the intensity value of the kth neighboring pixel, and I_c is the intensity value of the center pixel. An example of LBP is presented in Fig. 4 [34].

Fig. 4 shows the LBP spectrum of the Lena image with different circular domain radius and sampling points. Correspondingly, fineness of the texture information in the obtained LBP spectrum is different. Taking the Lena image as an example, with the increase of sampling radius, the gray scale statistical value of the LBP map is sparser [34].

In the proposed approach presented in this paper, after LBP is applied on the LH and HL sub-bands of the DWT, 512 features are extracted to represent the image.

3.3. Color-based Features Extraction

Color is considered as a basic feature observation in viewing an image to reveal a variety of information [12]. Color is extremely used feature for image retrieval techniques [35], [36].
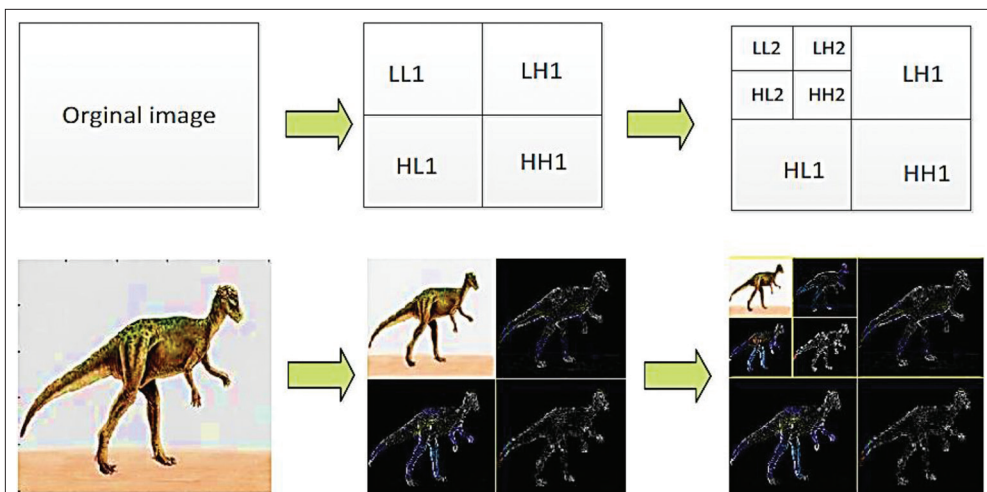

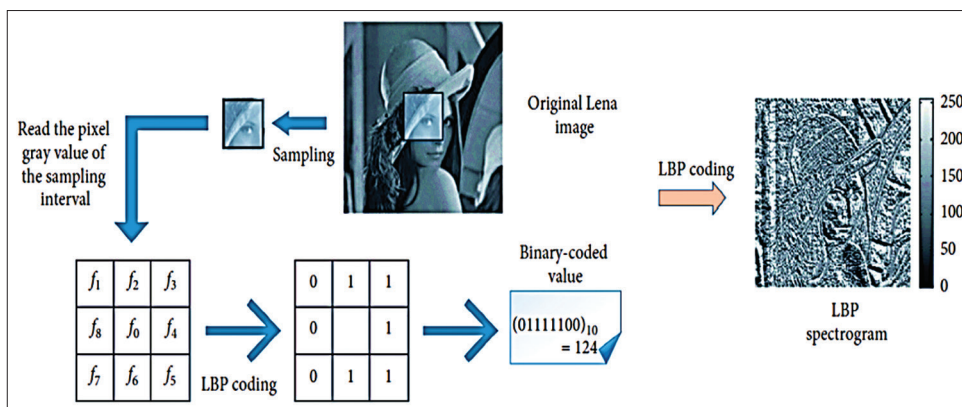**Fig. 3.** DWT sub-bands.

Fig. 4. An example of LBP operator.

Color points create color space and various color spaces based on the perceptual concepts are used for color illustration [23]. Among all color spaces, YCrCb and HSV have the mentioned perceptual characteristic. In YCbCr, the Y represents the luminance while the color is represented by Cb and Cr [37]. In our proposed approach, the mean and entropy of each component of the color spaces RGB, HSV, and YCbCr have been calculated as color features. Meanwhile, totally 18 color-based features are extracted.

3.4. Feature Matching

There are variety of similarity measurements used to determine the similarity between the query image and the images in the dataset [9]. Manhattan distance is used as a similarity measurement for both layers of the proposed approach in this work, equation (3) [36]:

$$\text{Manhattan distance (MD)} = \sum_{i=1}^{k} |x_i - y_i| \qquad (3)$$

Where x is the feature vector of query image, y is the feature vector of the dataset of images, and k is the dimension of image feature. Manhattan distance is also known as City block distance. In general, the Manhattan distance is non-negative where zero defines an identical point, and other means little similarity [9].

3.4.1. Proposed approach

This section describes the details of the proposed two-layer approach in the following steps:

1. Let the query image is denoted by Q, and $I = \{I_1, I_2, \ldots, I_n\}$ refers to the dataset which consists of n images.
2. First layer of the proposed approach involves the following steps:
   a. Q_{BoF} and I_{BoF} represent the feature vector of Q and I, respectively, after BoF technique is implemented on.
   b. To find the similarity between Q_{BoF} and I_{BoF}, Manhattan similarity measurement is used, and as a result, M most similar images to the query image are retrieved.
3. Second layer of the proposed approach, which includes the following steps, implements on the query image Q as well as the M most similar images that gained in the first layer.
   a. Extracting the following features from Q and M_i:
      • Let $L = \{l_1, l_{2\ldots}, l_{512}\}$ be the vector of 512 extracted texture-based features after LBP is applied on the LH and HL sub-bands of the DWT, 256 features extract from each sub-band.
      • Let $C = \{c_1, c_2, \ldots, c_{18}\}$ be the extracted 18 color-based features that represent the mean and entropy of the three components of RGB, HSV,

and YCbCr color spaces. Meanwhile, 6 features are extracted from each of the mentioned color spaces.
   • Let $F = L + C$ represents the feature vector of the fused of all the 530 extracted features from the previous steps.
   • Finally, Q_F and M_{Fi} represent the fused feature vector of Q and M_i, respectively.
   b. To find the similarity between Q_F and M_{Fi}, Manhattan similarity measurement is used, to retrieve the most similar images to the query image.

The block diagram of the proposed two-layer approach is illustrated in Fig. 5.

4. EXPERIMENTAL RESULTS

Experiments are conducted comprehensively in this section to evaluate the performance of the proposed approach in terms of precision rate, the most common confusion matrix measurement used in the research area of CBIR. In addition, the proposed approach is compared to the current existing works.

4.1. Dataset

Corel-1K dataset of images has been used, which is a public and well-known dataset, that contains 1000 images in the form of 10 categories and each category consists of 100 images with resolution sizes of (256 × 384) or (384 × 256) [37], [38]. The categories are arranged as follows: African, people, beaches, buildings, buses, dinosaurs, elephants, flowers, horses, mountains, and foods [37].

4.2. Evaluation Measurements

To evaluate the performance of the proposed approach, precision confusion matrix measurement has been used which determines the number of correctly retrieved images to the total number of the retrieved images from the tested dataset of images. Meanwhile, it measures the specificity of image retrieval system based on the following equation [38], [39]:

$$\text{Precision} = \frac{R_c}{R_t} \qquad (4)$$

where R_c represents the total number of correctly retrieved images and R_t represents the total number of retrieved images. In this study, top-10 and top-20 have been tested. Top-10 indicates the total number of retrieved images is 10 images, and top-20 indicates the total number of retrieved images is 20 images.
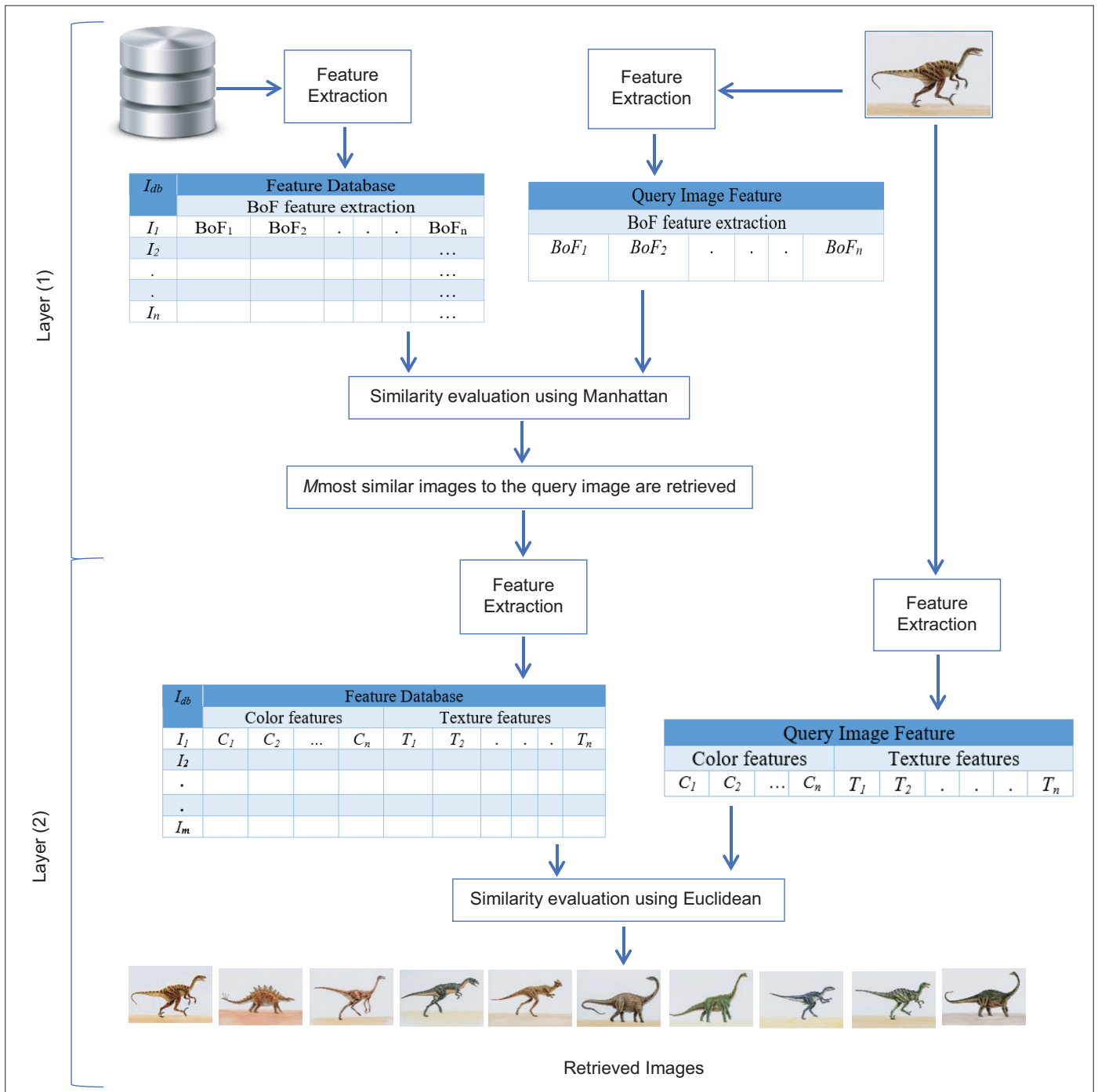
Fig. 5. Block diagram of the proposed two-layer CBIR approach.

4.3 Results

The experiments carried out in this work include two parts: (a) Single layer CBIR model and (b) Two-layer CBIR model. First part evaluates the single layer model (i.e., BoF technique) alone, and on the other hand, CBIR technique based on extracting texture and color features is evaluated.

In the second part, the proposed two-layer model has been assessed. The experiments are detailed in the following steps:
1. BoF-based CBIR technique is tested using different number of clusters, as BoF technique relies on the K-means clustering algorithm to create clusters, which is commonly called visual words. The number of clusters

cannot be selected automatically; manual selection is needed. To select the proper number of clusters, (i.e., value of k-means), the different number of clusters have been tested to obtain the best precision result of BoF technique. The precision results of different number of clusters are illustrated in the following tables.

From Tables 1 and 2, it is quite obvious that the best result is achieved when k = 500 for both top-10 and top-20.

2. The DWT sub-bands, and the concatenation of the sub-bands, have been tested as a texture feature as presented in the following tables.

From Tables 3 and 4, one can observe that the best result is obtained when the LH and HL sub-bands are concatenated for both top-10 and top-20.

3. In this step, LBP is implemented on DWT sub-bands, implementation of LBP on LH and HL sub-bands.

In the proposed method, LBP is extracted from DWT sub-bands to form a sub-novel local feature descriptor. To achieve this, we performed DWT decomposition and consider the

high frequency sub-bands HL, and LH. However, the sub-bands HL and LH also contain edge and contour details of image's significant in extracting pose and expression relevant features with the aid of LBP. We ignored the low-frequency LL and the high-frequency HH sub-band as it mostly contains the noise with negligible feature details. To preserve the spatial characteristics and to form a robust local feature descriptor, multi-region LBP pattern-based features [4] are obtained from non-overlapping regions of DWT sub-bands {HL, LH}, are statistically significant and offer reduced dimensionality with increased robustness to noise. Each of the sub-band {HL, LH} is equally divided into m non-overlapping rectangle regions R0; R1;… ; Rm, each of size (x,y) pixels. From each of these m regions, we extract local features LBP each with 256 labels separately. Local features from successive regions are concatenated to form a combining the two results in one vector with 512 features, the results in Tables 5 and 6.

From Tables 5 and 6, one can observe that the best result is obtained when the implementation of LBP on LH sub-band as well as HL sub-band is concatenated.

TABLE 1: Precision rate of BoF technique for different number of clusters for top-10

Different number of clusters	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
K=100	61.1	47.2	50.1	88.9	100	68.9	87.3	89.6	46.4	53.8	71.056
K=200	65.7	49.3	55.2	88.9	100	70.9	87	91.8	49.1	53.3	73.1
K=300	65.2	49.9	57.1	89.2	100	72.3	87.4	93.1	48.4	54.4	73.622
K=400	65.5	47.9	56.7	90.9	100	72.1	87.6	92.4	48.9	53.4	73.556
K=500	67.1	49.3	60.4	89.2	100	70.1	87.7	93.6	51.5	53.9	74.322
K=600	64.9	47.9	59.1	89.4	100	70.6	88.6	93.6	51.7	53.4	73.978
K=700	67.1	49.3	60.4	88.2	100	70.1	86.7	93.6	51.5	53.9	74.1
K=800	64.6	48	59.4	89.2	100	68.2	88.3	94.2	53.5	51	73.933
K=900	64.7	46.2	62	88.1	100	68	87.5	93.4	55.6	52.7	73.944
K=1000	64.5	47	61.5	87.6	100	66.9	88.3	94.2	53.4	50.9	73.711

TABLE 2: Precision rate of BoF technique for different number of clusters for top-20

Different number of clusters	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
K=100	54.75	46.5	42.4	84.5	100	58.55	84.9	82	39.9	40.15	63.365
K=200	56.95	48.65	42.5	86.35	100	59.85	85.5	87.05	41	42.3	65.015
K=300	58.55	48.55	45	85.7	100	61.05	85.5	88.4	42.45	41.95	65.715
K=400	58.45	48.25	47.3	87.35	100	59.2	85.65	87.65	44.65	41.35	65.985
K=500	60.5	48.75	50.3	85.6	99.95	59.05	85	87.9	47.15	40.55	66.475
K=600	60.2	48.5	50	84.8	99.95	58.85	84.7	87.7	47.1	40.05	66.185
K=700	57.95	48.5	51.6	84.85	99.95	57.5	84.75	88.45	44.75	39.65	65.795
K=800	58.3	48.7	50.55	84.3	99.95	58.95	85.95	89.3	47.35	39.15	66.25
K=900	57.65	47.95	51.7	84.35	99.95	57.6	85.05	88.75	47.7	39.3	66
K=1000	56.8	48.2	52.95	82.6	99.85	56.95	85.15	88.4	46.15	37.55	65.46

TABLE 3: Precision rate for the DWT sub-bands for top-10

DWT sub-bands	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
LL	17.2	48.4	44.6	24.85	95.1	29.4	27.25	16.5	26.15	20.25	34.97
LH	25.05	36.05	31.8	24.3	97.35	29.85	29.05	22	27.95	27.95	35.14
HL	11.55	21.55	22.55	23.15	93.3	23.2	24.25	21.25	21.65	19.7	28.22
HH	21.2	12.55	17.85	22.8	92.95	22.85	13.9	20.9	11.3	15.35	25.17
LL_LH	20.2	33.15	29.8	26.9	98	34.5	34.2	20.55	30.6	23.95	35.19
LL_HL	30.2	32.9	30.1	26.8	98.55	33.35	34.45	21.15	31.1	24.85	36.35
LL_HH	29	31.25	28.2	25.05	96.8	31.9	32.6	19.85	29.1	23.2	34.7
LH_HL	26.85	34.05	34.05	33.7	98.55	27.8	39.05	25.2	26.15	23.35	36.88
LH_HH	15.61	22.86	22.46	15.86	98.16	21.51	19.26	15.71	21.56	22.76	27.58
HL_HH	11.85	27.8	22.5	8.1	92.4	23.75	21.5	17.05	15.8	19	25.98
LL_LH_HH	16.05	32	26.7	12.3	96.6	27.95	25.7	21.25	20	23.2	30.18
LL_HL_HH	19.85	33.05	30.3	25.9	99.25	33.1	34.2	19.9	31.25	24.25	35.11
LL_LH_HH	24.66	33.86	32.11	30.71	96.2	31.91	31.01	20.71	30.06	22.06	35.329
LL_LH_HL_HH	17.6	29.2	29.5	24.3	99.7	31.8	30.6	15	30.3	23.4	33.14

TABLE 4: Precision rate for the DWT sub-bands for top-20

DWT sub-bands	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
LL	15.2	29.7	26.75	22.85	94.1	29.4	27.25	16.5	26.15	20.25	30.82
LH	20.05	23.5	22.9	15.3	97.35	21.85	19.05	15	23.95	22.95	28.19
HL	11.55	12.9	18.2	13.15	93.3	13.2	14.25	11.25	11.65	15.7	21.52
HH	17.55	8.9	14.2	19.15	89.3	19.2	10.25	17.25	7.65	11.7	21.52
LL_LH	16.55	29.5	26.15	23.25	94.35	30.85	30.55	16.9	26.95	20.3	31.54
LL_HL	25.95	28.65	25.85	22.55	95.3	29.1	30.2	16.9	26.85	20.6	32.2
LL_HH	26.45	28.7	25.65	22.5	94.25	29.35	30.05	17.3	26.55	20.65	32.15
LH_HL	22.3	29.5	29.9	29.55	94.4	23.65	34.9	21.2	22	19.2	32.66
LH_HH	11.85	19.1	18.7	12.1	94.4	17.75	15.5	11.95	17.8	19	23.82
HL_HH	7.85	15.1	14.7	8.1	92.4	13.75	11.5	7.95	13.8	15	20.02
LL_LH_HH	11.85	19.1	18.7	12.1	96.4	17.75	15.5	11.95	17.8	19	24.02
LL_HL_HH	15.95	29.15	26.4	22	95.35	29.2	30.3	16	27.35	20.35	31.21
LL_LH_HH	17.35	30.55	27.8	23.4	95.75	30.6	31.7	17.4	28.75	21.75	32.51
LL_LH_HL_HH	11.95	25.4	23.3	18.55	95.2	27.05	27.2	11.65	23.95	15.85	28.01

TABLE 5: Precision rate for LBP for top-10

DWT sub-bands	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
LL	37.35	50.55	47.8	43.4	115.75	50.6	51.7	37.4	48.75	41.75	52.51
LH	56.8	49.5	54.1	85.2	99.5	45.8	91.5	70.2	39.8	47.8	64.02
HL	74.1	54.2	58	92.9	99.9	56.3	89.1	78.1	46.5	63.8	71.29
HH	53.5	40.9	54.9	90.1	98.6	39.8	92.3	57.1	37.8	46.8	61.18
LL_LH	63.4	50.4	61.8	92.3	98.7	44.6	92.9	74.6	39.1	51.4	66.92
LL_HL	62.64	49.54	61.02	91.74	98.18	43.7	92.34	73.91	38.17	50.55	66.18
LL_HH	61.19	48.19	59.59	90.09	96.49	42.39	90.69	72.39	36.89	49.19	64.71
LH_HL	70.7	61.4	69.7	95.5	99.2	55.1	93.4	83	46.4	65	73.94
LH_HH	64.45	51.45	62.85	93.35	99.75	45.65	93.95	75.65	40.15	52.45	67.97
HL_HH	63.45	50.45	61.85	92.35	98.75	44.65	92.95	74.65	39.15	51.45	66.97
LL_LH_HH	64.55	51.75	63.05	92.95	99.75	45.95	94.05	75.45	40.15	52.45	68.01
LL_HL_HH	62.75	49.95	61.25	91.15	97.95	44.15	92.25	73.65	38.35	50.65	66.21
LL_LH_HH	74.35	57.95	67.75	95.75	99	54.05	94.15	81.85	47.55	65.05	73.75
LL_LH_HL_HH	72	55.8	65.5	93.6	98.3	52	91.9	79.7	45.5	62.9	71.72

4. Before selecting an appropriate color description, selection of color space is important and needs to choose a color model for color feature extraction process [35]. This step evaluates the impact of extracting the color

feature by testing different color space components such as: RGB, YCbCr, and HSV. In other words, mean and entropy for each color components have been calculated and the results are presented in the following tables.

The results presented in Tables 7 and 8 demonstrates that combining the extracted features of all the tested color spaces provides best precision rate.

5. Finally, all the extracted features are fused. Meanwhile, by concatenating the extracted LBP in step 3 and the extracted color feature in step 4, Table 9.

6. Eventually, the proposed two-layer approach has been tested. It includes two layers: The first layer implements BoF technique (for K=500) and M most similar images are retrieved, M is user defined. In the second layer, color and texture features are extracted from the query image and the M remained images, as a result, N most similar images are retrieved. The following tables investigate the best value of M. In other words, Tables 10 and 11 show investigating different number of M for top-10 and top-20, respectively.

Results in Tables 10 and 11 demonstrate that the best precision results are obtained for M = 100 and M=200. For

TABLE 6: Precision rate for LBP for top-20

DWT sub-bands	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
LL	35.15	48.35	45.6	41.2	113.55	48.4	49.5	35.2	46.55	39.55	50.31
LH	48.95	40.7	45.25	79.65	96.6	36.85	87.8	64.75	32.85	40.75	57.42
HL	67.55	48.7	49.2	89.85	98.4	44.7	86.05	68.9	40.7	57.35	65.14
HH	44.8	36	46.25	84.4	98.1	32.1	89.9	49.9	30.75	38.95	55.12
LL_LH	53	43.3	51	87.95	98.1	35.95	90.5	62.85	34.3	42.15	59.91
LL_HL	52.16	42.39	50.15	87.36	97.58	34.99	89.93	62.08	33.33	41.24	59.12
LL_HH	50.79	41.09	48.79	85.74	95.89	33.74	88.29	60.64	32.09	39.94	57.7
LH_HL	63.8	54.45	59.85	93.2	99.15	43.75	90.65	72.75	41	58.2	67.68
LH_HH	54.05	44.35	52.05	89	99.15	37	91.55	63.9	35.35	43.2	60.96
HL_HH	53.05	43.35	51.05	88	98.15	36	90.55	62.9	34.35	42.2	59.96
LL_LH_HH	54.1	44.65	52.15	88.65	99.15	36.9	91.55	63.8	35.4	43.35	60.97
LL_HL_HH	52.3	42.85	50.35	86.85	97.35	35.1	89.75	62	33.6	41.55	59.17
LL_LH_HH	64.5	50.6	57.8	93.3	98.9	43.35	92.45	69.75	40.7	56	66.74
LL_LH_HL_HH	62.25	48.4	55.9	90.9	97.65	41.2	90.2	67.6	38.45	53.7	64.625

TABLE 7: Precision rate for color-based features for top-10

Color spaces	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
RGB	45.8	35.5	33.2	41.6	99.4	49.6	62.6	83.7	39.4	56.4	54.72
HSV	21.1	22.65	32.8	26.7	39.5	23.9	19.9	29.45	20	22.1	25.81
YCbCr	47.6	35.9	40.1	46.9	100	53.4	65.7	56.6	39.3	51	53.65
RGB_HSV	48	37.3	41.7	48.8	99.3	52.3	60.3	84	38.5	56.7	56.69
RGB_YCbCr	50.7	37.4	38.5	50.8	99.7	55.2	68	86.4	38.8	55.6	58.11
HSV_YCbCr	41.5	33.5	41.1	46.8	99.6	48.65	59.65	56.1	35.2	44.2	50.63
RGB_HSV_YCbCr	47.8	38.9	47.6	60.3	99.7	52.4	63.9	82.2	44.6	56.8	59.42

TABLE 8: Precision rate for color-based features for top-20

Color spaces	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
RGB	40	28.8	27.55	34.5	99.5	45.45	54.6	77.7	34.5	48.15	49.075
HSV	27.1	29.3	37.7	31.2	50.4	28.1	26.3	37.1	25.2	27.5	31.99
YCbCr	39.2	30.55	34.5	39.65	99.85	47.8	60.25	49.55	35.1	44	48.045
RGB_HSV	41.05	30.85	34.55	41.05	99.15	46.15	51.8	77.25	36.25	48.2	50.63
RGB_YCbCr	42.55	30.3	31	42.25	99.55	49	58.9	78.45	33.85	48.05	51.39
HSV_YCbCr	44.4	33.4	45.5	48.9	95.2	49.8	60.5	61.4	35.3	47.2	52.16
RGB_HSV_YCbCr	41.9	33.3	37.15	55.35	99.5	47.25	56.05	73.85	38.55	48.25	53.115

this reason, different numbers of M in the range of M=100 to M=200 have also been investigated to gain better precision result, Tables 12 and 13.

From Tables 12 and 13, it is quite clear that the best result is obtained when M =110 for both top-10 and top-20. More

experiments have been done to compare the proposed approach with the state-of-the-art techniques, Table 14.

According to the results presented in Table 14, the best performance (precision rate) is achieved by the proposed approach for both top-10 and top-20. All the tested

TABLE 9: Concatenation of texture-based feature and color-based feature

Top image retrieved	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
Top-20	67.45	49.1	60.65	88.8	99.45	61.7	90.5	81.8	49	61.55	71
Top-10	67.45	56.25	69.5	88.8	99.45	61.7	90.5	81.8	49	61.55	72.6

TABLE 10: Precision rate for different number of M for top-10

Different Number of M	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
M=100	81.65	63	71.9	95.95	99.85	77.6	96.35	92.75	67.55	73.45	82.005
M=200	78.85	62.2	73.65	95.7	99.85	76.6	95.95	91.6	66.65	75.1	81.615
M=300	77.05	61.1	74.65	94.8	99.85	75.4	96.05	90.65	65.4	75.35	81.03
M=400	76.6	60.1	74.7	94.65	99.65	74.8	95.95	90.3	64.45	75.15	80.635
M=500	76.5	60.2	74.75	94.35	99.65	74.2	95.85	90.25	63.45	74.85	80.405
M=600	76.2	60.1	74.3	94	99.65	74	95.65	90	63.3	74.75	80.195
M=700	76.15	60	74.1	93.85	99.65	73.85	95.45	90	63.15	74.55	80.075
M=800	76	59.4	74	93.6	99.2	73.25	95.15	89.4	63	74.2	79.72
M=900	76.35	59.9	75.2	94.15	99.65	73.4	95.85	90	62.1	75.05	80.165
M=1000	76.3	59.3	75.2	94.1	99.25	73.2	95.55	89.7	62	74.85	79.945

TABLE 11: Precision results for different number of M for top-20

Different number of M	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
M=100	76.35	58.8	61.85	94.95	99.9	70.35	94.1	90.3	61.45	64.15	77.22
M=200	75.25	57.45	63.85	94.35	99.6	70.1	93.95	87.45	60.6	66.25	76.885
M=300	73.1	55.5	64.4	92.9	99.55	68.9	93.8	86.35	58.85	66.75	76.01
M=400	72.1	54.47	65.4	92.4	99.55	68.15	93.6	85.95	57.75	66.5	75.587
M=500	72	54.45	65.4	92	99.55	67.85	93.55	85.75	57.25	66.3	75.41
M=600	71.7	54	65.1	91.6	99.55	67.75	93.3	85.65	57.15	66.2	75.2
M=700	71.6	53.8	65	91.4	99.45	67.6	93	85.4	57.05	66.2	75.05
M=800	71.5	53.5	64.8	91.2	99	67.1	93	85.1	56.8	66	74.8
M=900	71.7	54.05	65.5	91.45	99.55	66.75	93.45	85.5	54.05	66.45	74.845
M=1000	71.2	54.05	65.2	91.15	99.05	66.65	93.15	85.4	53.85	66.35	74.605

TABLE 12: Precision rate of the proposed approach for different number of M for top-10

Different Number of M	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
M=30	82.89	64.64	70.39	95.44	99.74	76.79	96.24	94.19	68.94	69.94	81.92
M=50	83.15	64.9	70.65	95.7	100	77.05	96.5	94.45	69.2	70.2	82.18
M=70	81.9	63.95	70.85	95.6	100	77.2	96.65	93.45	68.5	72.65	82.08
M=90	81.55	62.65	71.75	95.9	99.95	77.85	96.6	93	68.45	72.65	82.04
M=110	80.95	64.25	72.6	96	99.95	77.75	96.35	92.55	68	73.05	82.15
M=130	81.85	63.8	71.05	95.8	99.92	77.65	96.4	92.4	68.2	72	81.91
M=150	81.49	63.38	71.86	95.8	99.9	77.5	96.4	92.55	68	72.9	81.98
M=170	81.1	62.25	71.29	95.33	99.58	77.3	96	92.2	67.7	72	81.48
M=190	80	62.6	70.9	94.6	99.2	76.45	95.3	91.8	67	71.7	80.96

TABLE 13: Precision rate of the proposed approach for different number of M for top-20

Different number of M	Categories										
	Africa	Beaches	Buildings	Buses	Dinosaur	Elephant	Roses	Horses	Mountains	Food	Average
M=30	75.14	57.84	59.19	94.09	99.74	69.29	93.29	90.64	58.54	59.24	75.70
M=50	75.40	58.10	59.45	94.35	100.00	69.55	93.55	90.90	58.80	59.50	75.96
M=70	76.80	59.00	60.65	94.45	99.95	69.80	94.65	90.65	59.65	62.10	76.77
M=90	76.95	58.70	61.35	94.85	99.95	69.65	94.15	90.45	61.00	63.70	77.08
M=110	76.00	59.10	62.35	95.00	99.90	70.35	93.85	90.20	61.55	64.40	77.27
M=130	76.35	58.6	60.7	94.74	99.2	69.45	94.1	90.4	61	62.2	76.674
M=150	76.4	58.9	61.5	94.7	99.35	70	94.3	90.1	61.1	61.58	76.793
M=170	76.38	58	61	94.4	99	69.4	93.2	89.9	61.69	61.99	76.496
M=190	75.25	57.6	60.5	93.7	98.93	69.33	93.1	89	60.87	61.93	76.021

TABLE 14: Precision rate of the tested CBIR techniques

Approaches	Top-10	Top-20
Proposed approach	82.15	77.27
Atif Nazir, 2018	-	73.5
Pradhan Jitesh, 2019	64.00	59.60
Khawaja and Ahmed , 2019	-	76.5
Hamed Qazanfari, 2019	-	74.77
Aiswarya, 2020	67	

state-of-the-art techniques, except technique in Ahmed and Naqvi [24], they tested their approach either for top-10 or for top-20, and this is why in Table 14 some cells are not contained the precision rate.

5. CONCLUSION

The two-layer based CBIR approach for filtering and minimizing the dissimilar images in the dataset of images to the query image has been developed in this study. In the first layer, the BoF has been used and as a result, M most similar images are remained for the next layer. Meanwhile, the most dissimilar images are eliminated and, hence, the range of search is narrowed for the next step. The second layer concentrated on concatenating the extracted both (texture and color)-based features. The results obtained by the proposed approach devoted the impact of exploring the concept of two-layer in improving the precision rate compared to the existing works. The proposed approach has been evaluated using Corel-1K dataset and the precision rate of 82.15% and 77.27% for top-10 and top-20 is achieved, respectively. In the future, certain feature extractors need to be investigated as well as feature selection techniques need to be added to select the most important feature which reflects increasing precision rate.

REFERENCES

[1] Z. F. Mohammed and A. A. Abdulla. "Thresholding-based white blood cells segmentation from microscopic blood images". *UHD Journal of Science and Technology*, vol. 4, no. 1, p. 9, 2020.

[2] M. W. Ahmed and A. A. Abdulla. "Quality improvement for exemplar-based image inpainting using a modified searching mechanism". *UHD Journal of Science and Technology*, vol. 4, no. 1, p. 1, 2020.

[3] H. Liu, J. Yin, X. Luo and S. Zhang. "*Foreword to the Special Issue on Recent Advances on Pattern Recognition and Artificial Intelligence*". Springer, Berlin, p. 1, 2018.

[4] A. Wojciechowska, M. Choraś and R. Kozik. "*Evaluation of the Pre-processing Methods in Image-Based Palmprint Biometrics*". Springer, International Conference on Image Processing and Communications, p. 1, 2017.

[5] A. A. Abdulla, S. A. Jassim and H. Sellahewa. "*Secure Steganography Technique Based on Bitplane Indexes*". 2013 IEEE International Symposium on Multimedia, 2013.

[6] A. A. Abdulla. "*Exploiting Similarities between Secret and Cover Images for Improved Embedding Efficiency and Security in Digital Steganography*". Department of Applied Computing, The University of Buckingham, United Kingdom, pp. 1-235, 2015.

[7] S. Farhan, B. K. Biswas and R. Haque. "*Unsupervised Content-Based Image Retrieval Technique Using Global and Local Features*". International Conference on Advances in Science, Engineering and Robotics Technology, p. 2, 2019.

[8] R. S. Patil, A. J. Agrawal. "Content-based image retrieval systems: A survey". *Advances in Computational Sciences and Technology*, vol. 10, 9, pp. 2773-2788, 2017.

[9] H. Shahadat and R. Islam. "A new approach of content based image retrieval using color and texture features". *Current Journal of Applied Science and Technology*, vol. 21, no. 1, pp. 1-16, 2017.

[10] A. Sarwar, Z. Mehmood, T. Saba, K. A. Qazi, A. Adnan and H. Jamal. "A novel method for content-based image retrieval to improve the effectiveness of the bag-of-words model using a support vector machine". *Journal of Information*, vol. 45, pp. 117-135, 2019.

[11] L. K. Paovthra and S. T. Sharmila. "Optimized feature integration and minimized search space in content based image retrieval". *Procedia Computer Science*, vol. 165, pp. 691-700, 2019.

[12] A. Masood, M. A. Shahid and M. Sharif. "Content-based image retrieval features: A survey". *The International Journal of Advanced Networking and Applications*, vol. 10, no. 1, pp. 3741- 3757, 2018.

[13] S. Singh and S. Batra. "*An Efficient Bi-layer Content Based Image*

Retrieval System". Springer, Berlin, p. 3, 2020.

[14] Y. D. Mistry. "Textural and color descriptor fusion for efficient content-based image". *Iran Journal of Computer Science*, vol. 3, pp. 1-15, 2020.

[15] K. T. Ahmed, A. Irtaza, M. A. Iqbal. "*Fusion of Local and Global Features for Effective Image Extraction*". Elsevier, Amsterdam, Netherlands, vol. 51, pp. 76-99, 2019.

[16] M. O. Divya and E. R. Vimina. "*Maximal Multi-channel Local Binary Pattern with Colour Information for CBIR*". Springer, Berlin, p. 2, 2020.

[17] T. Kato. "*Database Architecture for Content-based Image Retrieval*". International Society for Optics and Photonics, vol. 1662. pp. 112-123, 1992.

[18] J. Yu, Z. Qin, T. Wan and X. Zhang. "Feature integration analysis of bag-of-features model for image retrieval". *Neurocomputing*, vol. 120, pp. 355-364, 2013.

[19] N. Shrivastava. "*Content-based Image Retrieval Based on Relative Locations of Multiple Regions of Interest Using Selective Regions Matching*". Elsevier, Amsterdam, Netherlands, vol. 259, pp. 212-224, 2014.

[20] E. Gupta and R. S. Kushwah. "*Combination of Global and Local Features Using DWT with SVM for CBIR in Reliability*". Infocom Technologies and Optimization (ICRITO) Trends and Future Directions, 2015.

[21] E. Gupta and R. S. Kushwah. "Content-based image retrieval through combined data of color moment and texture". *International Journal of Computer Science and Network Security*, vol. 17, pp. 94-97, 2017.

[22] A. Nazir, R. Ashraf, T. Hamdani and N. Ali. "*Content Based Image Retrieval System by using HSV Color Histogram, Discrete Wavelet Transform and Edge Histogram Descriptor*". 2018 International Conference on Computing, Mathematics and Engineering Technologies, p. 4, 2018.

[23] P. Jitesh, A. Ashok, P. A. Kumarand and B. Haider. "Multi-level colored directional motif histograms for content-based". *The Visual Computer*, vol. 36, pp. 1847-1868, 2020.

[24] K. T. Ahmed and S. H. Naqvi. "*Convolution, Approximation and Spatial Information Based Object and Color Signatures for Content Based Image Retrieval*". 2019 International Conference on Computer and Information Sciences, 2019.

[25] H. Qazanfari, H. Hassanpour and K. Qazanfari. "Content-based image retrieval using HSV color space features". *International Journal of Computer and Information Engineering*, vol. 13, no. 10, pp. 537-545, 2019.

[26] E. Rashno. "Content-based image retrieval system with most relevant features among wavelet and color features". *Iran University of Science and Technology*, vol. pp. 1-18, 2019.

[27] K. S. Aiswarya, N. Santhi and K. Ramar. "Content-based image retrieval for mobile devices using multi-stage autoencoders". *Journal of Critical Reviews*, vol. 7, pp. 63-69, 2020.

[28] J. Zhou, X. Liu, W. Liu and J. Gan. "Image retrieval based on effective feature extraction and diffusion process". *Multimedia Tools and Applications*, vol. 78, no. 5, pp. 6163-6190, 2019.

[29] P. Srivastava. "*Content-Based Image Retrieval Using Multiresolution Feature Descriptors*". Springer, Berlin, pp. 211-235, 2019.

[30] I. A. Saad. "An efficient classification algorithms for image retrieval based color and texture features". *Journal of AL-Qadisiyah for Computer Science and Mathematics*, vol. 10, no. 1, pp. 42-53, 2018.

[31] M. S. Haji. "Content-based image retrieval: A deep look at features prospectus". *International Journal of Computational Vision and Robotics*, vol. 9, no. 1, pp. 14-37, 2019.

[32] V. Geetha, V. Anbumani, S. Sasikala and L. Murali. "*Efficient Hybrid Multi-level Matching with Diverse Set of Features for Image Retrieval*". Springer, Berlin, pp. 12267-12288, 2020.

[33] R. Boukerma, S. Bougueroua and B. Boucheham. "*A Local Patterns Weighting Approach for Optimizing Content-Based Image Retrieval Using a Differential Evolution Algorithm*". 2019 International Conference on Theoretical and Applicative Aspects of Computer Science, 2019.

[34] Y. Cai, G. Xu, A. Li and X. Wang. "A novel improved local binary pattern and its application to the fault diagnosis of diesel engine". *Shock and Vibration*, vol. 2020, p. 9830162, 2020.

[35] G. Xie, B. Guo, Z. Huang, Y. Zheng and Y. Yan. "*Combination of Dominant Color Descriptor and Hu Moments in Consistent Zone for Content Based Image Retrieval*". IEEE Access, vol. 8, pp. 146284-146299, 2020.

[36] A. C. Nehal and M. Varma. "*Evaluation of Distance Measures in Content Based Image Retrieval*". 2019 3rd International conference on Electronics, Communication and Aerospace Technology, pp. 696-701, 2019.

[37] S, Bhardwaj, G. Pandove and P. K. Dahiya. "A futuristic hybrid image retrieval system based on an effective indexing approach for swift image retrieval". *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 12, pp. 1-13, 2020.

[38] S. P. Rana, M. Dey and P. Siarry. "Boosting content based image retrieval performance through integration of parametric and nonparametric approaches". *Journal of Visual Communication and Image Representation*, vol. 58, pp. 205-219, 2019.

[39] M. K. Alsmadi. "*Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features*". Springer, Berlin, pp. 1-14, 2020.