

An Improved Content Based Image Retrieval Technique by Exploiting Bi-layer Concept



Shalaw Faraj Salih¹, Alan Anwer Abdulla^{2,3}

¹Department of Information Technology, Technical College of Informatics, Sulaimani Polytechnic University, Sulaimani, Iraq.

²Department of Information Technology, College of Commerce, University of Sulaimani, Sulaimani, Iraq.

³Department of Information Technology, University College of Goizha, Sulaimani, Iraq

ABSTRACT

Applications for retrieving similar images from a large collection of images have increased significantly in various fields with the rapid advancement of digital communication technologies and exponential evolution in the usage of the Internet. Content-based image retrieval (CBIR) is a technique to find similar images on the basis of extracting the visual features such as color, texture, and/or shape from the images themselves. During the retrieval process, features and descriptors of the query image are compared to those of the images in the database to rank each indexed image accordingly to its distance to the query image. This paper has developed a new CBIR technique which entails two layers, called bi-layers. In the first layer, all images in the database are compared to the query image based on the bag of features (BoF) technique, and hence, the M most similar images to the query image are retrieved. In the second layer, the M images obtained from the first layer are compared to the query image based on the color, texture, and shape features to retrieve the N number of the most similar images to the query image. The proposed technique has been evaluated using a well-known dataset of images called Corel-1K. The obtained results revealed the impact of exploring the idea of bi-layers in improving the precision rate in comparison to the current state-of-the-art techniques in which achieved precision rate of 82.27% and 76.13% for top-10 and top-20, respectively.

Index Terms: BoF, CBIR, Gabor, HSV, Zernike

1. INTRODUCTION

Digital image processing plays a significant role in various areas such as medical image processing [1], image inpainting [2], pattern recognition, biometrics, content-based image retrieval (CBIR), image compression, information hiding [3], and multimedia security [4]. The retrieval of similar images from a large range of images is becoming a serious challenge with the advent of digital communication technology and the growing use of the Internet. Several

penetrating and retrieval utilities are essential for end users to retrieve the images efficiently from different domains of the image databases such as medical, education, weather forecasting, criminal investigation, advertising, social media, web, art design, and entertainment. The query information is either text format or image format.

Different techniques for image retrieval have been developed and they are classified into two approaches: Text-based image retrieval (TBIR) and CBIR [5]. TBIR was first introduced in 1970 for searching and retrieving images from image databases [6]. In TBIR, the images are denoted by text, and then the text is used to retrieve or search the images. Such a system is text-based search and is generally referred to as TBIR. The TBIR method relies on the manual text search or keyword matching of the existing image keywords and the result has been dependent on the human labeling of the

Access this article online

DOI: 10.21928/uhdjst.v5n1y2021.pp1-12 E-ISSN: 2521-4217
P-ISSN: 2521-4209

Copyright © 2021 Al-Janabi, *et al.* This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Shalaw Faraj Salih, Department of Information Technology, Technical College of Informatics, Sulaimani Polytechnic University, Sulaimani, Iraq. E-mail: shalaw.faraj.s@spu.edu.iq

Received: 01-11-2020

Accepted: 23-12-2020

Published: 05-01-2021

images. TBIR approach requires information such as image keyword, image location, image tags, image name, and other information related to the image. It needs human intervention to enter the data of images in the database and that is the difficulty of the process. TBIR has the following limitations: (1) It leads to inaccurate results when human has been doing datasets annotation process wrongly, (2) single keyword of image information is not efficient to transfer the overall image description, (3) it is based on manual annotation of the images, which is time consuming [5], [7].

To overcome those mentioned limitations of TBIR, a new approach for image retrieval has been invented by researcher which is known as CBIR. CBIR can be considered as a common tool for retrieving, searching, and browsing images of a query information from a large database of digital 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 similar images in the database [8]. Fig. 1 shows the general block diagram of CBIR approach [7].

In general, CBIR entails two main steps: The feature extraction and feature matching. In the first step, features are extracted from a dataset of images and stored in a feature vector. In the second step, the extracted features from the query image are compared with the extracted features of images in the dataset using certain distance measurement. If the distance between feature vector of the query image and the image in the database is small enough, the corresponding image in the database is considered as a match/similar image to the query image. Consequently, the matched images are then ranked accordingly to a similarity index from the smallest distance value to the largest one. Finally, the retrieved images are specified according to the highest similarity, that is, lowest distance value [9].

The main objective of CBIR techniques is to improve the efficiency of the system by increasing the performance

using the combination of features [6]. Image features can be classified into two types: Local features and global features. Local features work locally which are focused on the key point in images whereas global features extract information from the entire image [10]. When the image dataset is quite large, image relevant to the query image are very few. Therefore, it is important to eliminate those irrelevant images. The main contribution of our proposed approach is filtering the images in the dataset to eliminate/minimize the most irrelevant images, then from the remaining images find the most similar/match images. In this paper, a new CBIR approach based on two layers is developed. The first layer aims in filtering the images using (BoF) strategy on the basis of extracting local features, while the second layer aims to retrieve similar images, from the remaining images, to the query image based on extracting global features such as color, shape, and texture.

The rest of the paper is organized as follows: Section 2 presents the literature review, Section 3 gives the background, Section 4 addresses the proposed approach in detail, Section 5 illustrates the experimental results, and finally, Section 6 presents the conclusion.

2. LITERATURE REVIEW

There are several CBIR techniques proposed for image retrieval applications using various feature extraction methods. Each of these techniques competes to improve the precision rate of finding the best similar images to the query image. In general, all the CBIR techniques have two main steps; the first step is feature extraction and the second step is feature matching. This section concerns the review of the most related and important existing CBIR techniques.

The concept of CBIR was first introduced by Kato in 1992 by developing a technique for sketch retrieval, similarity retrieval, and sense retrieval to support visual interaction [11]. Sketch

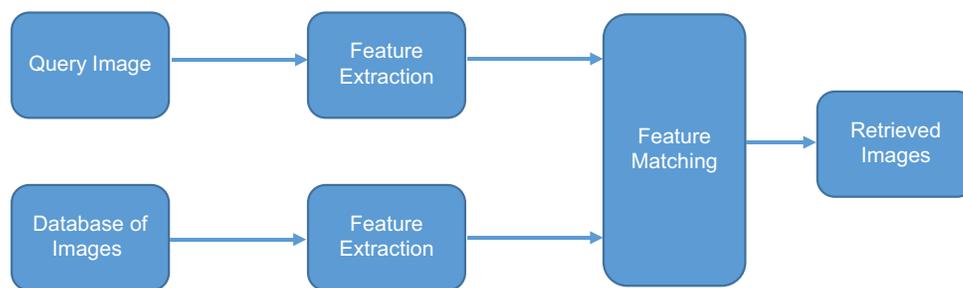


Fig. 1. General block diagram of content-based image retrieval approach.

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. In 2009, Lin *et al.* proposed a CBIR technique depending on extracting three types of image features [12]. The first feature, color co-occurrence matrix was extracted as a color feature, while for the second feature, difference between pixels of scan pattern was used for extracting texture feature, and the third feature, color histogram for K-mean was extracted which is based on color distribution. Consequently, feature selection techniques were implemented to select the optimal features not only to maximize the detection rate but also to simplify the computation of image retrieval. In addition, this proposed technique further uses sequential forward selection to select features with better discriminability for image retrieval and to overcome the problem of excessive features. Finally, Euclidean distance was used to find the similarity in the feature matching step. The results reported in this work claimed that the proposed technique reached a precision rate of 72.70% for the top-20.

Huang *et al.* proposed a new CBIR technique, in 2010, in which combined/fused the Gabor texture feature and Hue Saturation Value (HSV) color moment feature [13]. Furthermore, the normalized Euclidean distance was used to calculate the similarity between the feature vector of the query image and the feature vector of the images in the dataset. This proposed technique achieved the precision rate of 63.6% for the top-15. In 2012, Singha *et al.* proposed an algorithm for CBIR by extracting features called wavelet based color histogram image retrieval as a color and texture features [14]. The color and texture features are extracted through color histogram as well as wavelet transformation, for the combination of these features is robust to object translation and scaling in an image. This technique was used the histogram intersection distance for feature matching purposes. The results reported in this work claimed that this technique achieved a precision rate of 76.2% for the top-10. Another CBIR technique was proposed by Yu *et al.*, in 2013, that aims to investigate various combinations of mid-level features to build an effective image retrieval system based on the BoF model [15]. Specifically, this work studies two ways of integrating: 1- scale-invariant feature transform (SIFT) with local binary pattern (LBP) descriptors and, 2- histogram of oriented gradients with LBP descriptors. Based on the qualitative and quantitative evaluations on two benchmark datasets, the integrations of these features yield complementary and substantial improvement on image retrieval even with noisy background and ambiguous objects.

Consequently, two integration models are proposed, the patch-based integration and the image-based integration. Using a weighted K-means clustering algorithm, the image-based SIFT-LBP integration achieved a precision rate of 65% for the top-20. A new CBIR technique was proposed by Somnugpong *et al.*, in 2016, by combining color correlograms and edge direction histogram (EDH) features to give precedence for spatial information in an image [16]. Color correlogram treats information about spatial color correlation, while EDH provides the geometry information in the case of the same image but different color. Evaluation is performed by simple calculation like Euclidean distance between the query image and the images in the database. Researchers claimed that their proposed technique achieved 65% of precision rate for the top-15. In 2018, Al-Jubouri *et al.* proposed a new CBIR technique that addresses the semantic gap issue by exploiting cluster shapes [17]. The technique first extracts local color using YCbCr color space and texture feature using Discrete Cosine Transform coefficients. The Expectation-Maximization Gaussian Mixture Model clustering algorithm is then applied to the local feature vectors to obtain clusters of different shapes. To compare dissimilarity between two images, the technique uses a dissimilarity measure based on the principle of Kullback-Leibler divergence to compare pair-wise dissimilarity of cluster shapes. This work further investigates two respective scenarios when the number of clusters is fixed and adaptively determined according to cluster quality. The results reported in this work illustrate that the proposed retrieval mechanism based on cluster shapes increases the image discrimination, and when the number of cluster is fixed to a large number, the precision of image retrieval is better than that when the relatively small number of clusters is adaptively determined. Authors claimed that their technique achieved a precision rate of 75% for the top-10.

In 2018, Nazir *et al.* proposed a new CBIR technique in which used color and texture features [18]. The edge histogram descriptor is extracted as a local feature and discrete wavelet transform as well as color histogram features are extracted as global features. Consequently, Manhattan distance measurement was used to measure the similarity between the feature vector of the query image and the feature vector of the images in the dataset. The reported results of the work revealed that this proposed technique achieved a precision rate of 73.5% for the top-20. Pradhan *et al.*, in 2019, developed a new CBIR scheme based on multi-level colored directional motif histogram [7]. The proposed scheme extracts local structural features at three different levels. The performance of this proposed scheme has been

evaluated using different Corel/natural, object, texture, and heterogeneous image datasets. Regarding to the Corel-1k, the precision rate of 64% and 59.6% was obtained for top-10 and top-20, respectively. Qazanfari *et al.*, in 2019, proposed a CBIR technique based on HSV color space [19]. The human visual system is very sensitive to the color and edge orientation, also color histogram and color difference histogram are two kinds of low-level feature extraction which are meaningful representatives of the image color and edge orientation information. This proposed technique was used Canberra distance measurement and this work has been evaluated using three standard databases Corel 5k, Corel 10 and UKBench and achieved 61.82%, 50.67%, and 74.77% of precision rate for the top-12, respectively. In 2019, Rashno *et al.*, proposed a new technique for CBIR in which color and texture features were used. HSV, Red, green, and blue (RGB) and norm of low frequency components were used as color features, while wavelet transformation was used to extract texture features [20]. Consequently, ant colony optimization-based feature selection was used to select the most relevant features, to minimize the number of features, and to maximize F-measure in the proposed CBIR system. Furthermore, Euclidean distance measurement was used to find the similarity between query and database images. The results reported in this work demonstrate that this approach reached the precision rate of 60.79% for the top-20. In 2019, Rana *et al.* proposed a CBIR technique by fusing parametric color and shape features with nonparametric texture feature [21]. The color moments and moment invariants which are parametric feature are extracted to describe color distribution and shapes of an image. The non-parametric ranklet transformation is performed to narrate the texture features. These parametric and non-parametric features were integrated to propose a robust and effective CBIR algorithm. In this proposed work, four similarity measurements are investigated during the experiment, namely, Chi-squared, Manhattan distance or City block distance, Euclidean distance, and Canberra distance. The experimental results demonstrate that Euclidean distance metric yields better precision and recall than other distance measuring criteria. Authors claimed that their technique achieved a precision rate of 67.6% for the top-15 using Euclidean distance. Finally, Sadique *et al.*, in 2019, proposed a CBIR technique in which investigates various global and local feature extraction methods for image retrieval [22]. The proposed work uses a combination of speeded up robust features (SURF) detector and descriptor with color moments as local features, and modified grey level cooccurrence matrices as global features. Both global and local features are used as the only local features are not suitable when the variety of images is large.

Finally, fast approximate nearest neighbor search was used for matching the extracted features. Authors claimed that their proposed technique achieved a precision rate of 70.48% for the top-20.

3. BACKGROUND

This section aims to present a reasonable amount of background information about useful techniques such as (SURF) feature descriptor, color-based features, texture-based features, shape-based features, and feature matching techniques.

3.1. SURF Feature Descriptor

The most popular feature descriptor is SURF, which is also the most important one. However, there are other available feature descriptors. SURF can be considered as a local feature. Local features can provide more detailed characters in an image in comparison with global features such as color, texture, and shape. It is a rotation and scale invariant descriptor that performs better with respect to distinctiveness, repeatability, and robustness. It is also photometric deformations, detection errors, geometric, and robust to noise [23]. SURF is used in many applications such as BoF which is used and successful in image analysis and classification [24]. In BoF technique, SURF descriptor is often used to extract local feature first. In the next stage, a quantization algorithm such as K-means is separately applied to the extracted SURF features to reduce high dimensional feature vectors to clusters, which are also known as visual words. Then, K-means clustering is used to initialize the M center point to build M visual words. The K-means clustering algorithm takes feature space as input and reduces it into M cluster as output. Then, the image is represented as a histogram of code word occurrences by mapping the local features to a vocabulary [24]. The methodology of the image representation based on the BoF model is illustrated in Fig. 2.

3.2. Color-based Features Extraction

The color-based features have commonly been used in CBIR systems because of its easy and fast computation [14]. Color-based features can be extracted using a histogram of quantized values of color in Hue (H), Saturation (S), and Value (V) of the HSV color space. HSV color space is more robust to human perception as compared to the RGB color space. Due to the robustness of the HSV color space, first RGB images are converted to HSV color space and then uniform quantization is applied. Feature vectors are generated by considering the values of H=9, S=3, and

$V=3$ to form the feature vector of size 81 ($9 \times 3 \times 3$) bins. Representation of HSV color feature vector of an image is presented in Fig. 3.

3.3. Texture-based Features Extraction

Like color-based features, the texture-based features can be considered as powerful low-level features for image search and retrieval applications. There are certain works that have been done on texture analysis, classification, and segmentation for the last four decades. So far, there is no unique definition for the texture-based features. Texture is an attribute representing the spatial arrangement of the

grey levels of the pixels in a region or image. Gabor filter is one of the widely used filters for texture-based feature extraction. It is a Gaussian function modulated by complex sinusoidal of frequencies and orientations. In our proposed approach, texture features of an image are extracted using a Gabor filter for five scales (s) and six orientations (θ). The usage of multiple s and θ makes the features rotation and scaling invariant on texture feature space. Five scales and six orientations produce thirty magnitudes. Consequently, mean and standard deviation need to be calculated for each magnitude and this leads to producing sixty features as a texture descriptor [14].

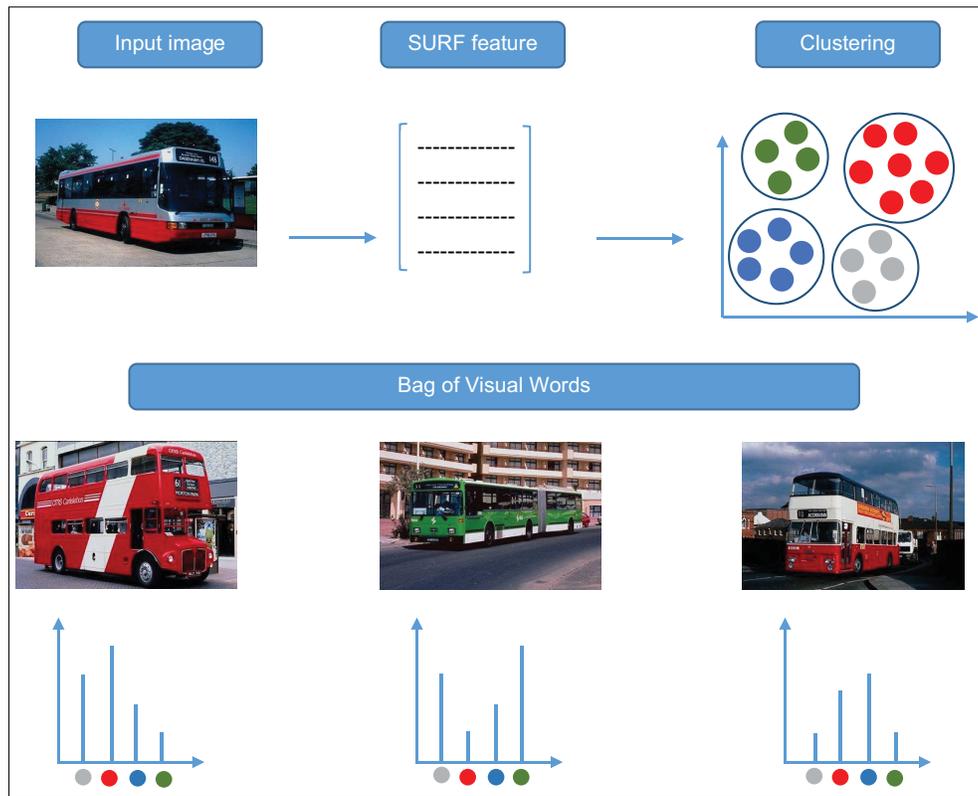


Fig. 2. Methodology of the bag of features based image representation for content-based image retrieval.

$$H = \begin{cases} 0 & h \in [1,40] \\ 1 & h \in [41,80] \\ 2 & h \in [81,120] \\ 3 & h \in [121,160] \\ 4 & h \in [161,200] \\ 5 & h \in [201,240] \\ 6 & h \in [241,280] \\ 7 & h \in [281,320] \\ 8 & h \in [321,360] \end{cases} \quad S = \begin{cases} 0 & s \in [0.00, 0.30] \\ 1 & s \in [0.31, 0.70] \\ 2 & s \in [0.71, 1.00] \end{cases} \quad V = \begin{cases} 0 & v \in [0.00, 0.30] \\ 1 & v \in [0.31, 0.70] \\ 2 & v \in [0.71, 1.00] \end{cases}$$

Fig. 3. Hue saturation value feature vector.

3.4. Shape-based Features Extraction

Shape-based features are also useful to obtain more detailed characters of the images. Shape-based features include turn angle, central angle, distance between two feature points, distance between center of mass and feature point. Zernike Moments (ZMs) are used as a shape-based feature extractor in the proposed approach. ZMs are invariant to rotation, translation, and scaling [25]. Furthermore, ZMs are robust to noise and minor variations in shape and use Zernike polynomials to form feature vectors to represent an image based on shape features [26]. The proposed approach used 21 initial ZMs to represent the images.

3.5. Feature Matching

There are certain similarity measurements that used to compute the similarity between query image and images in the database, in our proposed approach, Manhattan distance is used for the BoF, see equation (1) [27], and Euclidean distance is used for color, texture and, shape features, equation (2) [13].

$$\text{Manhattan Distance (MD)} = \sum_{i=1}^f |Q^f - D^f| \quad (1)$$

$$\text{Euclidean Distance (ED)} = \sum_{i=1}^f \sqrt{(Q^f - D^f)^2} \quad (2)$$

Where $Q^f = (Q^f_1, Q^f_2, \dots, Q^f_{L-1})$ is the feature vector of query image, $D^f = (D^f_1, D^f_2, \dots, D^f_{L-1})$ is the feature vector of the database of images, and L is the dimension of image feature.

Next section will present the proposed approach in detail.

4. PROPOSED APPROACH

This section presents the detailed steps of the proposed bi-layer approach as follows:

- 1- Let Q be the query image, and $I_{db} = \{I_1, I_2, \dots, I_n\}$ be the database of n images.
- 2- First layer entails the following steps:
 - a) Q_{BoF} and I_{BoF} represent feature vector of Q and I_{db} , respectively, after BoF technique is applied.
 - b) Manhattan similarity measurement is used to find the similarity between Q_{BoF} and I_{BoF} and as a result, M most similar images to the query image are retrieved.

- 3- Second layer will implement on the query image Q and the M most similar images M_i that were retrieved/obtained from the first layer. It includes the following steps:
 - a) Extract the following features from Q and M_i :
 - Let $C = \{c_1, c_2, \dots, c_{81}\}$ be the extracted 81 color-based features that represent the 81 bins of the quantized HSV color space.
 - Let $T = \{t_1, t_2, \dots, t_{60}\}$ be the extracted 60 texture-based features using Gabor filter.
 - Let $S = \{s_1, s_2, \dots, s_{21}\}$ be the extracted 21 shape-based features using ZMs.
 - Let $F = C + T + S$ be the feature vector of the fused/combination of all the extracted features above.
 - Finally, Q_F and M_{Fi} represent the fused feature vector of Q and M_i .
 - b) Euclidean similarity measurement is used to find the similarity between Q_F and M_{Fi} to retrieve the N most similar match images to the query image.

The block diagram of the proposed bi-layer approach is illustrated in Fig. 4.

5. EXPERIMENTAL RESULTS

In this section, experiments are performed comprehensively to assess the performance of the proposed approach in terms of precision rate, the most common confusion matrix measurement used in the CBIR research area. Furthermore, the proposed approach is compared to the most recent existing works.

5.1. Dataset

The experiments are conducted on the public and well-known dataset called Corel-1K that contains 1000 images in the form of ten categories and each category consists of 100 images with resolution sizes of (256×384) or (384×256) [28]. The categories are organized as follows: African, people, beaches, buildings, buses, dinosaurs, elephants, flowers, horses, mountains, and foods.

5.2. Evaluation Measurements

Precision confusion matrix measurement is used to assess the performance of the proposed approach. The precision determines the number of correctly retrieved images over the total number of the retrieved images from the tested database of images and it measures the specificity of image retrieval system, as presents in the following equation [21]:

$$\text{Precision} = \frac{R_c}{R_r} \quad (3)$$

where R_c represents the total number of correctly retrieved images and R_r represents the total number of retrieved

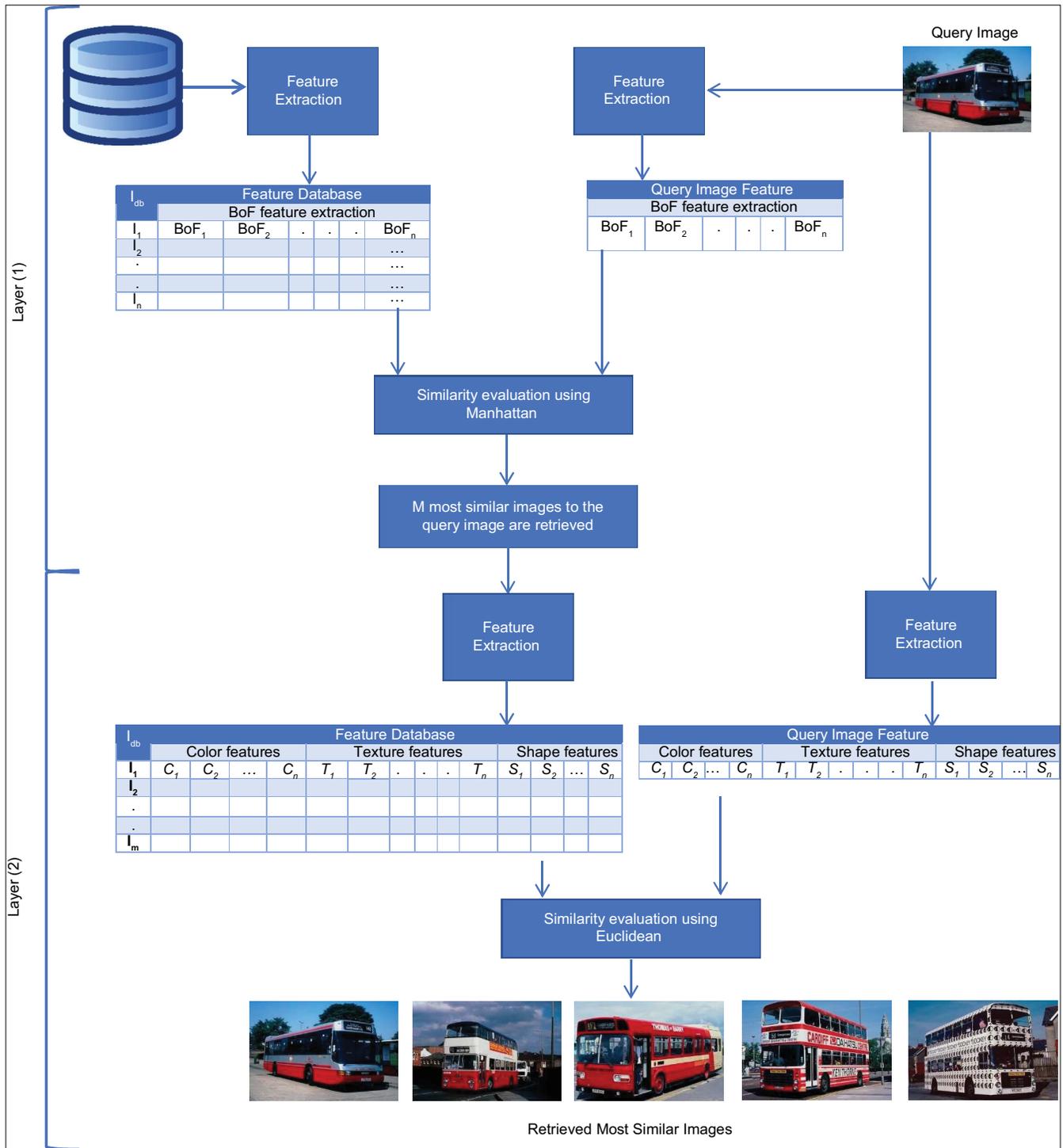


Fig. 4. Block diagram of the proposed bi-layer content-based image retrieval system for top-5.

images. Precision can also be expressed in the following equation.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

Where TP represents true positive and FP represents false positive. In this work, top-10 and top-20 have been tested. Top-10 means the total number of retrieved images is 10 images, and top-20 means the total number of retrieved

images is 20 images. Figs. 5 and 6 present examples for the query image based on top-10 and top-20.

5.3. Results

The experiments conducted in this work involve two phases: (a) Single layer CBIR model and (b) Bi-layer CBIR model. In the first phase, the single layer model (i.e., BoF technique) is evaluated alone, and on the other hand, CBIR technique based on extracting shape, texture, and color features is evaluated. In the second phase, the proposed bi-layer model is evaluated. The experiments are detailed in the following steps:

1. BoF-based CBIR technique is tested with different number of clusters, as BoF technique depends on the K-means clustering algorithm to create clusters, which is commonly called visual words. The number of clusters cannot be selected automatically; it needs manual selection. 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 numbers of clusters are illustrated in the following tables.

From Tables 1 and 2, one can observe that the best result is obtained when $k = 500$ for both top-10 and top-20.

2. The proposed CBIR technique that relies on extracting shape, color, and texture features has been tested, and the results are presented in Table 3.
3. The proposed bi-layer approach has been tested. It includes two layers: First layer implements BoF technique (for $K=500$) and M most similar images are retrieved, M is user defined. In the second layer, shape, color, and texture features are extracted from the query image and the M images, as a result, N most similar images are retrieved. The following tables investigate the best value of M . In other words, Tables 4 and 5 show testing different number of M for top-20 and top-10, respectively.

Results in Tables 4 and 5 demonstrate that the best precision results are obtained when $M = 200$. For this reason, different small numbers of M , in the range of $M = 100$ to $M = 300$, are investigated to gain better precision results, Tables 6 and 7.

From Tables 6 and 7, it is quite clear that the best result is obtained when $M = 225$ for both top-20 and top-10.

TABLE 1: Precision rate of BoF technique for different number of clusters for top-20

| Categories | Different number of clusters | | | | | | | | | |
|------------|------------------------------|-------|--------|--------|---------------|-------|--------|-------|-------|--------|
| | K=100 | K=200 | K=300 | K=400 | K=500 | K=600 | K=700 | K=800 | K=900 | K=1000 |
| Africa | 52.05 | 55.85 | 55.25 | 56.25 | 55.85 | 56.35 | 55.8 | 55.5 | 54.15 | 55.5 |
| Beaches | 44.35 | 45.7 | 45.85 | 46.35 | 47.2 | 45.4 | 47.15 | 45.35 | 46.65 | 48.15 |
| Buildings | 41.3 | 44.75 | 46.5 | 47.55 | 49.05 | 50.4 | 51.25 | 52.25 | 52.55 | 52.15 |
| Buses | 83.5 | 86.15 | 85.15 | 86.3 | 86.75 | 85.4 | 84.75 | 84.4 | 84.65 | 83.5 |
| Dinosaur | 100 | 100 | 100 | 100 | 100 | 99.95 | 99.95 | 100 | 99.95 | 99.95 |
| Elephant | 55.85 | 59.95 | 62 | 61.45 | 60.5 | 59.65 | 58.55 | 57.85 | 57.15 | 58.1 |
| Roses | 84.35 | 84.4 | 84.95 | 85.4 | 85.45 | 84.55 | 85.75 | 86.15 | 84.1 | 86.1 |
| Horses | 85.9 | 86.4 | 87.9 | 88.6 | 89.35 | 87.85 | 88.9 | 88.15 | 88.55 | 88.7 |
| Mountains | 39.5 | 40.7 | 41.9 | 42.5 | 43.1 | 45.95 | 45 | 46.95 | 45.6 | 45.55 |
| Food | 39.45 | 42 | 41.05 | 40.35 | 41 | 39.3 | 39.35 | 38.2 | 38.45 | 37.25 |
| Averages | 62.625 | 64.59 | 65.055 | 65.475 | 65.825 | 65.48 | 65.645 | 65.48 | 65.18 | 65.495 |

TABLE 2: Precision rate of bof technique for different number of clusters for top-10

| Categories | Different number of clusters | | | | | | | | | |
|------------|------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| | K=100 | K=200 | K=300 | K=400 | K=500 | K=600 | K=700 | K=800 | K=900 | K=1000 |
| Africa | 58.9 | 60.4 | 61 | 62.1 | 62.6 | 60.8 | 62.1 | 61.4 | 61.9 | 59 |
| Beaches | 50.8 | 52 | 51.6 | 51.3 | 52.5 | 50.6 | 51.4 | 51.7 | 50.5 | 51.1 |
| Buildings | 50.1 | 54.4 | 56.1 | 58.1 | 58.4 | 59.7 | 59.5 | 58.3 | 60.6 | 61 |
| Buses | 89.2 | 89.5 | 89.7 | 89.6 | 89.2 | 88.4 | 88.4 | 88.6 | 87.9 | 87.4 |
| Dinosaur | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Elephant | 69.1 | 71.3 | 73.2 | 69.8 | 71.4 | 69.6 | 70.5 | 69.7 | 67.3 | 66.6 |
| Roses | 86.6 | 88.3 | 88.5 | 88.4 | 87.8 | 88.9 | 88 | 87.6 | 88.4 | 89.3 |
| Horses | 88.8 | 91.2 | 93.3 | 93.2 | 93.7 | 93.9 | 93.7 | 93.1 | 94.4 | 94 |
| Mountains | 46 | 49.1 | 50.6 | 50.2 | 50.4 | 52.4 | 50.6 | 53 | 52.2 | 52.5 |
| Food | 46.7 | 49 | 50.5 | 49.4 | 48.6 | 47.8 | 48.4 | 46.6 | 44.5 | 45.8 |
| Averages | 68.62 | 70.52 | 71.45 | 71.21 | 71.46 | 71.21 | 71.26 | 71 | 70.77 | 70.67 |



Fig. 5. Query result for top-10.



Fig. 6. Query result for top-20.

TABLE 3: Precision rate for the tested feature extractors for top-20 and top-10

| Categories | Top-20 | Top-10 |
|------------|--------|--------|
| Africa | 70.55 | 75.4 |
| Beaches | 36.55 | 43.6 |
| Buildings | 42.4 | 50.9 |
| Buses | 72.85 | 79.8 |
| Dinosaur | 92.45 | 95.9 |
| Elephant | 40.5 | 54.3 |
| Roses | 58.9 | 72.2 |
| Horses | 84.9 | 89.5 |
| Mountains | 34.9 | 40.8 |
| Food | 65.7 | 70.5 |
| Averages | 59.97 | 67.29 |

The ratio of correctly retrieved images over the total number of images of the semantic class in the image database is known as recall and it measures the sensitivity of the image retrieval system, equation (5):

$$Recall = \frac{R_c}{T_s} \quad (5)$$

Where R_c is the total number of retrieved images and T_s is the total number of images in the semantic class in the database. More experiments have been done to compare the proposed approach with the state-of-the-art

TABLE 4: Precision rate for different number of M for top-20

| Categories | Different number of M | | | | | | | | | |
|------------|-----------------------|--------|-------|-------|--------|-------|-------|-------|--------|--------|
| | M=100 | M=200 | M=300 | M=400 | M=500 | M=600 | M=700 | M=800 | M=900 | M=1000 |
| Africa | 78.65 | 77.75 | 76.75 | 76.6 | 76.2 | 76.5 | 76.65 | 76.65 | 76.3 | 76.25 |
| Beaches | 55.7 | 56.65 | 56.55 | 56.5 | 55.5 | 54.6 | 54 | 53.85 | 53.7 | 53.55 |
| Buildings | 56.1 | 56.05 | 56.35 | 56.35 | 56.15 | 56.65 | 57 | 56.6 | 56.5 | 56.55 |
| Buses | 94.45 | 94.25 | 93.8 | 93.45 | 93.1 | 92.55 | 92.25 | 92.15 | 91.85 | 91.8 |
| Dinosaur | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Elephant | 64.05 | 64.25 | 64.5 | 64.8 | 65.15 | 65.15 | 65.05 | 64.65 | 64.65 | 64.45 |
| Roses | 91.35 | 91.15 | 90.7 | 89.75 | 89.3 | 88.75 | 88.6 | 88.3 | 88.05 | 87.9 |
| Horses | 94.65 | 94.95 | 94.9 | 94.9 | 94.8 | 94.7 | 94.6 | 94.6 | 94.6 | 94.6 |
| Mountains | 52 | 52.65 | 52.15 | 52 | 50.9 | 50.65 | 50 | 49.55 | 49.25 | 49.3 |
| Food | 68.55 | 72.15 | 72.6 | 72.75 | 72.35 | 71.55 | 70.85 | 70.85 | 70.55 | 70.4 |
| Averages | 75.55 | 75.985 | 75.83 | 75.71 | 75.345 | 75.11 | 74.9 | 74.72 | 74.545 | 74.48 |

TABLE 5: Precision results for different number of M for top-10

| Categories | Different number of M | | | | | | | | | |
|------------|-----------------------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| | M=100 | M=200 | M=300 | M=400 | M=500 | M=600 | M=700 | M=800 | M=900 | M=1000 |
| Africa | 83.7 | 82.4 | 82.8 | 82.5 | 82 | 82 | 82 | 81.7 | 81.7 | 81.6 |
| Beaches | 63.4 | 64 | 63 | 62.6 | 61.6 | 61 | 60.4 | 60 | 59.8 | 59.8 |
| Buildings | 68.4 | 67.5 | 66.6 | 66.6 | 66.3 | 66.9 | 67 | 66.6 | 66.8 | 66.8 |
| Buses | 97.5 | 96.8 | 96.3 | 96 | 95.6 | 95.1 | 94.6 | 94.6 | 94.6 | 94.6 |
| Dinosaur | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Elephant | 77.8 | 78.4 | 79.4 | 78.9 | 79.1 | 79 | 78.7 | 78.5 | 78.5 | 78.3 |
| Roses | 94.4 | 94.2 | 94 | 93.6 | 93.3 | 93 | 92.9 | 92.6 | 92.5 | 92.5 |
| Horses | 97.6 | 97.4 | 97.4 | 97.4 | 97.3 | 97.3 | 97.2 | 97.1 | 97.1 | 97.1 |
| Mountains | 60.4 | 62.4 | 61.3 | 60.3 | 59 | 58.6 | 58.3 | 57.6 | 56.9 | 57 |
| Food | 77.8 | 78.6 | 79 | 78.9 | 78.6 | 78.3 | 77.4 | 77.5 | 77.4 | 77.4 |
| Averages | 82.1 | 82.17 | 81.98 | 81.68 | 81.28 | 81.12 | 80.85 | 80.62 | 80.53 | 80.51 |

TABLE 6: Precision results for different number of M in the range 100–300 for top-20

| Categories | Different number of M | | | | | | | | | |
|------------|-----------------------|-------|--------|-------|--------|-------|--------|-------|-------|--|
| | M=100 | M=125 | M=150 | M=175 | M=200 | M=225 | M=250 | M=275 | M=300 | |
| Africa | 78.65 | 78.4 | 78.15 | 77.7 | 77.75 | 78 | 77.4 | 77.2 | 76.75 | |
| Beaches | 55.7 | 56.35 | 56.75 | 57 | 56.65 | 57.1 | 56.3 | 56.3 | 56.55 | |
| Buildings | 56.1 | 56.45 | 56.55 | 56.5 | 56.05 | 56.4 | 56.45 | 56.05 | 56.35 | |
| Buses | 94.45 | 94.8 | 95.05 | 94.6 | 94.25 | 94.55 | 94.2 | 94.1 | 93.8 | |
| Dinosaur | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | |
| Elephant | 64.05 | 64.5 | 63.9 | 64.1 | 64.25 | 64.3 | 64.2 | 64.4 | 64.5 | |
| Roses | 91.35 | 91.4 | 91.3 | 91.5 | 91.15 | 90.95 | 90.9 | 90.6 | 90.7 | |
| Horses | 94.65 | 94.9 | 95.05 | 94.95 | 94.95 | 94.95 | 94.85 | 94.95 | 94.9 | |
| Mountains | 52 | 52.8 | 52.6 | 52.3 | 52.65 | 52.65 | 52.55 | 52.4 | 52.15 | |
| Food | 68.55 | 69.4 | 71.3 | 71.45 | 72.15 | 72.4 | 72.6 | 72.7 | 72.6 | |
| Averages | 75.55 | 75.9 | 76.065 | 76.01 | 75.985 | 76.13 | 75.945 | 75.87 | 75.83 | |

techniques, Table 8. In all experiments, each image in the Corel-1K database is used as a query image. The retrieval performance of tested techniques is measured in terms of average retrieval precision (ARP) and average retrieval recall (ARR). The higher ARP and ARR values mean the better performance.

According to the results in Table 8, the best result is achieved by the proposed approach for both top-10 and top-20. All the tested state-of-the-art techniques, except technique in Al-Jubouri and Du [7], they tested their approach either for top-10 or for top-20, and this is why in Table 8 some cells do not contain the ARP and ARR results.

TABLE 7: Precision results for different number of M in the range 100–300 for top-10

| Categories | Different number of M | | | | | | | | |
|------------|-----------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| | M=100 | M=125 | M=150 | M=175 | M=200 | M=225 | M=250 | M=275 | M=300 |
| Africa | 83.7 | 83.2 | 82.7 | 82.4 | 82.4 | 82.1 | 82.3 | 82.7 | 82.8 |
| Beaches | 63.4 | 63.6 | 63.6 | 63.9 | 64 | 64.1 | 63.1 | 63.1 | 63 |
| Buildings | 68.4 | 68.3 | 68.3 | 67.7 | 67.5 | 67.4 | 66.7 | 66.8 | 66.6 |
| Buses | 97.5 | 97.3 | 97.2 | 97.2 | 96.8 | 96.9 | 96.6 | 96.6 | 96.3 |
| Dinosaur | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Elephant | 77.8 | 78.6 | 78.4 | 78.3 | 78.4 | 79.1 | 79 | 79.4 | 79.4 |
| Roses | 94.4 | 94.4 | 94.6 | 94.5 | 94.2 | 94.1 | 93.9 | 93.9 | 94 |
| Horses | 97.6 | 97.4 | 97.4 | 97.3 | 97.4 | 97.3 | 97.4 | 97.4 | 97.4 |
| Mountains | 60.4 | 61 | 61.5 | 61.7 | 62.4 | 62.7 | 62.5 | 61.8 | 61.3 |
| Food | 77.8 | 77.8 | 78.5 | 77.8 | 78.6 | 79 | 79 | 78.7 | 79 |
| Averages | 82.1 | 82.16 | 82.22 | 82.08 | 82.17 | 82.27 | 82.05 | 82.04 | 81.98 |

TABLE 8: ARP results of tested CBIR techniques

| Approaches | Top-10 | | Top-20 | |
|-------------------|--------|------|--------|-------|
| | ARP | ARR | ARP | ARR |
| Proposed approach | 82.27 | 8.22 | 76.13 | 15.22 |
| [7] | 64.00 | 6.40 | 59.60 | 11.92 |
| [18] | - | - | 73.5 | 14.7 |
| [21] | 67.60 | 6.76 | - | - |
| [22] | - | - | 70.48 | 14.09 |

6. CONCLUSION

This paper has developed an effective CBIR technique for retrieving images from a wide range of images in the dataset. The proposed approach involves two layers; in the first layer, all images in the database are compared to the query image based on the BoF technique, and as a result, 225 most similar images to the query image are selected. Color, texture, and shape features are used in the second layer to extract significant features from the selected 225 images to retrieve the most similar images to the query image. The obtained results depicted that the proposed approach has reached an optimal average precision of 82.27% and 76.13% for top 10 and top 20, respectively.

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, pp. 9-17, 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, pp. 1-8, 2020.
- [3] A. A. Abdulla, H. Sellahewa and S. A. Jassim. "Secure Steganography Technique Based on Bitplane Indexes". 2013 IEEE International Symposium on Multimedia, United States, pp. 287-291, 2013.
- [4] A. A. Abdulla. "Exploiting Similarities Between Secret and Cover Images for Improved Embedding Efficiency and Security in Digital Steganography, PhD Thesis". 2015. Available from: <http://www.bear.buckingham.ac.uk/149>. [Last accessed on 2020 Dec 15].
- [5] 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 Science*, vol. 45, no. 1, pp. 117-135, 2019.
- [6] S. Hossain and R. Islam. "A new approach of content based image retrieval using color and texture features". *Current Journal of Applied Science and Technology*, vol. 51, no. 3, pp. 1-16, 2017.
- [7] J. Pradhan, A. Ajad, A. K. Pal and H. Banka. "Multi-level colored directional motif histograms for content-based". *The Visual Computer*, vol. 36, pp. 1-22, 2019.
- [8] L. K. Pavithra and T. S. Sharmila. "Optimized feature integration and minimized search space in content based image retrieval". *Procedia Computer Science*, vol. 165, pp. 691-700, 2019.
- [9] H. F. Atlam, G. Attiya and N. El-Fishawy. "Comparative study on CBIR based on color feature". *International Journal of Computer Applications*, vol. 78, no. 16, pp. 9-15, 2013.
- [10] Y. D. Mistry. "Textural and color descriptor fusion for efficient content-based image". *Iran Journal of Computer Science*, vol. 3, pp. 1-15, 2020.
- [11] T. Kato. "Database architecture for content-based image retrieval". *International Society for Optics and Photonics*, vol. 1662, pp. 112-123, 1992.
- [12] C. H. Lin, R. T. Chen and Y. K. Chan. "A smart content-based image retrieval system based on color and texture feature". *Image and Vision Computing*, vol. 27, no. 6, pp. 658-665, 2009.
- [13] Z. C. Huang, P. P. Chan, W. W. Ng and D. S. Yeung. "Content-based image retrieval using color moment and Gabor texture feature". *2010 International Conference on Machine Learning and Cybernetics*, vol. 2, pp. 719-724, 2010.
- [14] M. Singha and K. Hemachandran. "Content based image retrieval using color and texture". *Signal and Image Processing*, vol. 3, no. 1, p. 39, 2012.
- [15] 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.
- [16] S. Somnugpong and K. Khiewwan. "Content-based Image

- Retrieval Using a Combination of Color Correlograms and Edge Direction Histogram*". 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), Thailand, pp. 1-5, 2016.
- [17] H. Al-Jubouri and H. Du. "A content-based image retrieval method by exploiting cluster shapes". *Iraqi Journal for Electrical And Electronic Engineering*, vol. 14, no. 2, pp. 90-102, 2018.
- [18] 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 (iCoMET), Pune, pp. 1-6, 2018.
- [19] 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.
- [20] A. Rashno and E. Rashno. "Content-based image retrieval system with most relevant features among wavelet and color features". *arXiv preprint arXiv*, vol. 2019, pp. 1-18.
- [21] 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.
- [22] F. Sadique, B. K. Biswas and S. M. Haque. "Unsupervised Content-based Image Retrieval Technique Using Global and Local Features". 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Bangladesh, pp. 1-6, 2019.
- [23] S. Jabeen, Z. Mehmood, T. Mahmood, T. Saba, A. Rehman and M. T. Mahmood. "An effective content-based image retrieval technique for image visuals representation based on the bag-of-visual-words model". *PloS One*, vol. 13, no. 4, pp. 1-24, 2018.
- [24] 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.
- [25] F. Rajam and S. Valli. "A survey on content based image retrieval". *Life Science Journal*, vol. 10, no. 2, pp. 2475-2487, 2013.
- [26] J. Olaleke, A. Adetunmbi, B. Ojokoh and I. Olaronke. "An appraisal of content-based image retrieval (CBIR) methods". *Asian Journal of Research in Computer Science*, vol. 3, pp. 1-15, 2019.
- [27] M. Sharma and A. Batra. "Analysis of distance measures in content based image retrieval". *Global Journal of Computer Science and Technology*, vol. 14, no. 2, p. 7, 2014.
- [28] J. Li and J. Z. Wang. "Automatic linguistic indexing of pictures by a statistical modeling approach". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1075-1088, 2003.