International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 9, No.4 (July-2020)

http://dx.doi.org/10.12785/ijcds/090412

An Employment of Neural Network Classifiers to evaluate the Performance of Color Feature Descriptors in an Image Retrieval System: An Experimental Survey

Shikha Bhardwaj[1], Gitanjali Pandove[2] and Pawan Kumar Dahiya[3]

[1] Department, of ECE, DCRUST, Murthal, Sonepat and UIET, KUK, Haryana, India
[2,3]Department of ECE, DCRUST, Murthal, Sonepat, India

Received 12 Aug. 2019, Revised 11 May 2020, Accepted 20 Jun. 2020, Published 1 Jul. 2020

Abstract: Content-based Image Retrieval (CBIR) is a system utilized to index and redeem images from a massive image repository, according to the user's preference. This paper focusses on an experimental analysis on varied color descriptors namely Color Moment, Color auto-Correlogram, Color Histogram, Color Coherence Vector and Dominant Color Descriptor utilized for color feature extraction in the two phases. In the first phase, Average Precision is computed for each technique one by one and in the second phase, a Cascade forward back propagation neural network (CFBPNN) is used in combination with each color descriptor to again calculate the performance of each technique. The classification accuracy of each color descriptor is computed by using both CFBPNN and Patternnet neural network and the results of these neural network classifiers are compared. The results of these analytical experimentations depict that a framework of Color auto-correlogram and CFBPNN outperforms the other color descriptors by obtaining average precision of 97%, 90.5% and 89.5% on Corel-1K, Corel-5K and Corel-10K benchmark datasets respectively.

Keywords: Color moment, Color auto-correlogram, Color histogram, Color coherence vector, Dominant color descriptor, Cascade forward back propagation neural network, Patternnet neural network, Similarity matching

1. INTRODUCTION

Digital imaging has now a days become a predominant part in many aspects of image analysis and processing. Digital images are rather used in a variety of applications like face detection, medical diagnostics, satellite imaging, art, crime detection, remote sensing etc [1]. Due to the proliferation of digital images on the online network because of camera mobiles, laptops, hand held electronic devices and other sophisticated imaging techniques, it is very cumbersome to retrieve a particular drawbacks because of various delusions like misspellings, usage of synonyms, homonyms etc.

In comparison to Text based image retrieval (TBIR), Content-based Image Retrieval is a methodology in which images are indexed, searched and retrieved on the basis of basic image features like color, texture, shape, edge and spatial information etc.[2].The features can be procured both locally and globally [3]. Global feature extraction is focused on features like color, texture etc. of the whole image but local feature extraction is concentrated on a special region or an object.

Color is the utmost fundamental characteristic that can be withdrawn from an image and is indispensable to human perception. Categories of colors and different objects associated with color has a specific wavelength associated with it, which is reflected and forms a specific color in human's retina. Color Histogram (CH) [4] is the cardinal color descriptor which is used for image from a large database. Therefore, the answer to this peculiar problem is Content-based Image Retrieval (CBIR) system. Images can be retrieved in two basic ways. The conventional method of image retrieval is called as TBIR, in which images are recovered on the basis of text and keywords comment or annotation [5]. But TBIR suffers from many color feature extraction. It is rotation and scale invariant but it has few pitfalls too.

Texture is another feature of an image, which defines the inborn appearance patterns of an image. A large quantity of techniques like Discrete Wavelet Transform (DWT), Curvelet Transform [6], Gabor Transform, Gray level Co-occurance Matrix (GLCM) [7], Color Co-occurance Matrix (CCM), Tamura features etc. are being

E-mail:shikpank@yahoo.com, pandove_geet@yahoo.co.in, pawan.dahiya@gmail.com*

used for texture extraction. Shape is a feature which bears semantic information and can be grouped into region based and boundary based. B-splines, Curvature scale space (CSS), Hough transform, Fourier Transform, Zernike Moments etc. are handful of shape descriptors used in CBIR system [1]. Edge detection is an approach by which identification and location of keen disruption in image is acquired. Robert edge detector, Sobel detector, Canny edge detectors are a few in the count. A large number of low feature extraction methods are deficient of spatial information. Region of interest (ROIs), graph/tree based etc. representation are few of spatial information representing methods.

But all these low level feature extraction techniques lack the lofty intelligence of human perception. Therefore, a wide gap called as Semantic gap exists between them. This gap can be narrowed by using different Intelligent or Machine learning techniques based on relevance feedback, support vector machines (SVM), deep learning, etc.

Neural network (NN) is considered as an exceptional modelling system which can be utilized for the purpose of analysis and experimentation. Feed forward neural network (FFNN) and Cascade forward neural network (CFNN) are the two main classes of neural network. The main characteristic of the CFNN is that every input layer neuron is connected to every neuron of the hidden layer as well as to each neuron of the output layer [8]. The foremost advantage of this network is that gratifies the non-linear interrelation between the input and the output layer while maintaining the linear association between the two.

The main contributions of this paper are as follows:

- To analyze the various color feature descriptors specifically, Color Moment, Color auto-Correlogram, Color Histogram, Color Coherence Vector and Dominant Color Descriptor utilized in CBIR system.
- To calculate accuracy and average precision of all these feature extraction techniques individually.
- To enhance the retrieval accuracy of the utilized descriptors, a Cascade forward back propagation neural network (CFBPNN) is being added to all these techniques.
- The retrieval accuracy of each descriptor is evaluated by using both CFBPNN and Patternnet neural network.
- Finally, three benchmark datasets have been tested by using this comparative analysis.

The remaining section of this paper is organized as follows: Related state-of-the-art work of different color feature extraction techniques is given in Section 2. Preliminaries in the form of different color descriptors, utilized neural network model is given in section 3. Section 4 gives detailed description about methodology and methods. Conclusion and future trends are given in Section 5.

2. Related State of the Art Work

This section dispenses the state-of-the-art procedures for color feature retrieval in CBIR systems. A method utilized for color extraction is Color Difference Histograms (CDH) [5] which is typically different from traditional method of Color Histogram [9]. This method is used to measure the color difference between two points which is perceptually uniform over the HSV color space. Again, Color coding (CC) with Bi cubic interpolation (BCI) is being used as color feature extraction tool where BCI is used for query image and database image rescaling and CC for color feature extraction [10] with Gray level Difference Method (GLDM), 4 levels of Discrete Wavelet Transform (DWT) and Hu moments.

Chandan Singh et al. [11] described about a CBIR system where color histogram is utilized as a color descriptor. Color histogram is a graphical characterization of pixels in an image. In addition to color histogram, block variation of local correlation coefficients (BVLC) and block difference of Inverse Probabilities (BDIP) are adopted for texture extraction. The preciseness of Color Histogram based technique can be further enhanced with the addition of color coherence vector (CCV) [12].

Efficiency and reliability can be extended to a great level by an amalgamation of various color descriptors in an individual CBIR system. Color moments, color histogram and color coherence vector (CCV) are used in a fusion system mode, which has an extravagant results as compared to a single color extraction CBIR system [13]. Color auto correlogram is also a phenomenal technique for the extraction of color features where spatial relation between two identical color pixels is calculated. Color auto correlogram with color moments is being deployed with support vector machines (SVM) to constitute a hybrid CBIR system with texture feature extraction methods [14].

Wei-Ta chen et al. [15] proposed a robust color feature extraction methodology. For thresholding, binary quaternion-moment-preserving (BQMP) is deployed which conjointly with extraction methods are used to retrieve color features up to third moment. The methods used for color extraction are variable cardinality (VC) and fixed cardinality (FC). To improve the retrieval accuracy of a typical CBIR system, an approach based on weighted adjacent structure (WAS) merged with hypergraph is also being used. Jointly color difference histogram and micro-structure descriptor have been deployed for similarity calculation [16]. Dominant color descriptor (DCD) is among the vital color descriptors used in CBIR where color space is divided into course partitions. Each partition has a partition center and its percentage. Integration of DCD, gray level co-occurance matrix (GLCM) and fourier descriptors is deployed for the constitution of Hybrid CBIR system [17].

Color is considered as a principal trait in an image analysis process. Different color spaces can be used for image representation. RGB color space is the most basic of these color spaces because of the image availability in the same format. Color information can be represented by using three separate 1D histograms or one individual 3D histogram [18]. Histogram normalization can also applied to get accurate results.

Deep learning is currently a popular technique based on numerous neural network algorithms which can also being used for image retrieval systems. Techniques based on artificial neural networks (ANN), convolutional neural networks (CNN), boltzman machines, etc. are being used. Rani Saritha et al. [19] proposed a CBIR system based on deep belief networks (DBN), a form of Deep learning algorithm, which is being used for the extraction of both spatial and frequency features of the datasets. High level of abstraction is being achieved by using a multi-layer neural units together.

Based on the popularity and the supremacy of deep learning techniques, deep learning is also successfully being used for biomedical images. A CBIR system can be developed on the basis of varied types of deep neural networks. Restricted boltzmann machines, deep belief networks, auto encoders are some of its types [20]. Another technique based on Deep CNN is proposed to extract features of an image using convolution layers which subsequently employs max pooling layers for its execution. Also, once a model is trained using CNN, again retraining is being employed, based on three different techniques, i.e, completely unsupervised retraining, retraining based on the presence of relevance information and retraining on the feedback obtained during relevance calculation.[21].

In this paper, different color descriptors are analyzed and compared on three standard benchmark databases, to detect the best amongst them in terms of evaluation metrics like precision and recall.

3. PRELIMINARIES

In CBIR systems, color is the utmost cue that approximates for feature extraction. Various types of color descriptors are being used in CBIR systems but here five types of descriptors will be compared and analyzed. The type of convolutional neural network (CNN) which is used as a classifier is also described here.

A. Color histogram

Color histogram (CH) [22] is the primary implemented color descriptor in CBIR systems. In this technique, colors are represented in a particular image which defines the total pixels in the same image. It is a type of bar graph which counts the number of pixels in fixed intervals called as bins. Its two axis namely X axis represents bar graphs of tonal colors and Y axis signifies the number of pixels in that particular bin. Histogram normalization, equalization and quantization are among the few post processing steps which can be applied after histogram calculation. These processes solves the issue of complexity. Histogram can be computed in two ways: (i) Global color histogram (GCH) (ii) Local color histogram (LCH). In GCH, histogram of an entire image is calculated least concerning about color distribution in a particular part of an image where as an image is split into different blocks and histogram calculation of all these blocks is done. It gives more statistics of an image but it is time consuming and costly. A general form of color [23] histogram can be denoted as follows:

$$h(p,q,r) = N.Prob\ (P = p, Q = q, R = r) \qquad (1)$$

where the three color channels of an image are given by

Q and R and N signifies the number of pixels of an image Color histograms are invariant to scale, translation and rotation but suffers from the major limitation that these do not include spatial information from images. To overcome this limitation of color histogram, many other descriptors like color moment, color signatures etc. can be used.

B. Color moment

Color moments (CM) are metrics which signify the color distributions in an image. According to the probability theory, its distribution can be effectively characterized by its moments. Different moments specify diverse analytical and statistical measures. This color descriptor is also scale and rotation invariant but it includes the spatial information from images [24].

If *Iij* specifies the *ith* color channel and *jth* image pixel, number of pixels in an image are *N*, than index entries associated with the particular color channel and region *r* is given by first color moment, which signifies average color in an image denoted as:

$$Mean(E_{r,i}) = \frac{1}{N} \qquad (2)$$

After mean, successive color moment is given by Standard deviation and is calculated by taking square root of variance of color distribution in an image. It is given as:

$$Standard\ deviation(\sigma_{r,i}) = \left(\frac{1}{N}\sum_{j=1}^{N}(I_{ij} - E_{r,i})^2\right)^{\frac{1}{2}} \quad (3)$$

Third moment is called as skewness and signifies that how asymmetric is the color distribution. It is given as:

$$Skewness\ (S_{r,i}) = \left(\frac{1}{N}\sum_{j=1}^{N}(I_{ij} - E_{r,i})^3\right)^{\frac{1}{3}} \quad (4)$$

Fourth moment is denoted by Kurtosis and it signifies the color distribution shape, emphasizing particularly on tall shape or flatness of the color distribution.

C. Color auto-correlogram

A color correlogram (CAC) is a color descriptor which gives information specifying contiguous relationship within color pixels of an image and also with the change in distance. Correlogram can also be defined as a stored table, indexed by color pair *(i,j)*, which approximates the likelihood of detecting a pixel *j* from pixel *i* at a span of distance denoted by *d*. whereas in auto-correlogram, it specifies the possibility of detecting a pixel *i* from the same pixel at a distance *d* [25]. Therefore, it can be said that auto-correlogram computes the spatial relationship between identical colors or levels.

Let *d* be the distance between two image pixels, then correlogram of image *I*, denoted for a color pair c_i, c_j at a distance of *d* is given by formula:

$$\gamma_{c_i,c_j}^{d}(I) \triangleq Pr_{p_1 \in Ic_i, p_2 \in I}\left[p_2 \in Ic_j \,\|p_1 - p_2 = d|\right] \quad (5)$$

Auto-correlogram computes the contiguous correlation between indistinguishable levels only and size of auto-correlogram is given by *O (md)* which is typically less than the requirement for correlogram. The auto-correlogram for a color image is given as:

$$\alpha_c^d(I) = \gamma_{c,c}^d(I) \quad (6)$$

In the above formula, for identical color given by c and a distance of *d*, the probability of pixels p_1 and p_2 can be found [26]. Thus, it can be concluded that color correlogram takes into account the spatial relationship among color pixels. It further can be employed to describe the overall dispensation of native spatial correlation of colors, depending upon the value of distance d to be chosen.

D. Color coherence vector

Color coherence vector (CCV) is another important color descriptor used in CBIR systems. According to coherency, pixels are categorized as either coherent or incoherent [27].

For the computation of CCV, the initial steps are similar to those of color histogram. First step is to blur the given image slightly with the average values in place of pixel values, specified in a confined proximity containing eight adjoining pixels. Then, in next step color space is divided into diverse parts, such that only *n* different colors are present in an image. Next, pixels belonging to a particular color region are categorized as either coherent or non-coherent and connected components are computed for finding pixel groups. Finally, single pixel will correspond to single connected component. Now, depending upon the size of connected component, pixels are classified as coherent and incoherent. A fixed value is chosen as Γ and pixels whose connected component size overreach Γ are denoted as coherent pixels and otherwise the pixels are incoherent. Suppose out of j_{th} discrete color pixels, number of coherent pixels are denoted by α_j and incoherent pixels by β_j. Finally, the total quantity of pixels with that particular color is given by $(\alpha_j+\beta_j)$, so according to color histogram technique, results obtained will be given by:

$$(\alpha_1+\beta_1\ldots..+\alpha_n+\beta_n) \quad (7)$$

Now, in CCV instead of each color, color pair (α_j, β_j) is computed, which is the coherency pair for the j_{th} color. The color coherence vector pairs for an image is given as:

$$[(\alpha_1,\beta_1)\ldots.(\alpha_n,\beta_n)] \quad (8)$$

"(8)" describes the coherence vector pair, one for each separate color.

E. Dominant color descriptor

A Dominant color descriptor (DCD) is a color feature extraction technique which designates the dominant color of an image. In this method, based on individual or several dominant values, the large database can be browsed by the user for an effective image retrieval. DCD and MPEG-7 are among the various multimedia content representation techniques and the related distance metrics can be used for an efficient image retrieval. DCD is an effective color descriptor as compared to other traditional color descriptors used in CBIR systems. DCD is given as under:

$$F=\{(c_i,p_i,v_i),s\} \quad i=1.....N \tag{9}$$

Where N is the number of colors and $\sum_{i=1}^{N} pi = 1$, c_i is color index, p_i is percentage, v_i is color variance, s is spatial coherency. Here last two parameters are optional [24].

F. Cascade forward back propagation neural network

A Convolutional neural networks (CNN) are a type of Artificial neural networks (ANN) [28] which uses multiple cascaded layers for its functioning**.** The modelling of a neural network can be in the form of: (1) Feed Forward Back Propagation Neural Network (FFBPNN) (2) Cascade Forward Back Propagation Neural Network (CFBPNN).

In FFBPNN, a signal is transferred from an input layer to a hidden layer based on various sigmoid neurons in the form of weighted signals and then succeeded by linear neurons of an output layer. In CFBPNN, each neuron is related to all its preceding neurons. In this network, only one neuron is initially used for the process of learning and there is an automatic addition of neurons during the training process. As the training error decreases, there is a cumulative increase in the number of hidden layers.

CFBPNN are based on the working aspect of MLP. In this network, a weighted signal travels from an input layer to a hidden layer and the signal is being distributed to all the neurons present in the hidden layer. Then, this signal is being processed by a non-linear activation function here. Similarly, the output from the hidden layer is being sent to an ouput layer in the form of weighted sum and gets processed by the output's layer activation function. Generally, there is an identity mapping between the activation function of the input and output layer, so the signal obtained is same. A model with m inputs denoted by $\boldsymbol{x_i}$, where i = 1….m, one output neuron z and k neurons in the hidden layer is constituted. The general architecture of CFBPNN is shown in Fig. 1 which depicts an input layer containing three neurons, hidden layer of two neurons and an output layer of single neuron.

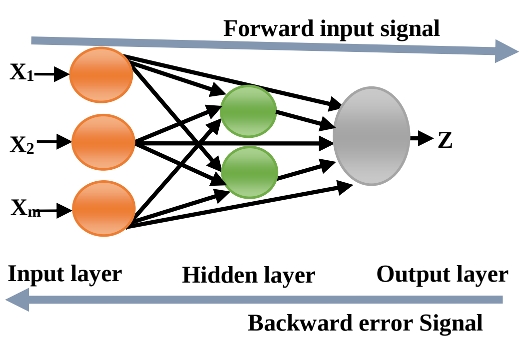

Figure 1. A basic Cascade forward back propagation neural network

The output of the neurons in the hidden layer is expressed by *yj*, where j=1, ..., k. The inputs x_1, ..., x_m are distributed to neurons in the hidden layer. $\boldsymbol{g^0}$ is the activation function of the neuron in the output layer and $\boldsymbol{g^h}$ is the activation function of the hidden layer. The mathematical equation for the above described model can be written as:

$$\boldsymbol{z = g^0\left(\sum_{j=1}^{k} w_j^0\, g_j^h\left(\sum_{i=1}^{m} w_{ji}^h x_i\right)\right)} \tag{10}$$

In the above equation "(10)" $\boldsymbol{g^0}$ is an activation function of output layer while $\boldsymbol{g_j^h}$ is the activation function of the hidden layer. If a bias is added to the above "(10)" then, it will change to:

$$\boldsymbol{z = g^0\left(w^b + \sum_{j=1}^{k} w_j^0\, g^h\left(w_j^b + \sum_{i=1}^{m} w_{ji}^h x_i\right)\right)} \tag{11}$$

Here, $\boldsymbol{w^b}$ is the weight from bias to output and $\boldsymbol{w_j^b}$ is the weight from bias to the hidden layer. Cascade forward neural network (CFNN) is a type of deep neural network which uses back propagation algorithm for its operation. This network has been used in an intrusion detection system. These models can also be used to increase the image quality and obtain the desired compression in any image related system [29].

In Cascade forward neural network, each neuron of an input layer is connected to the progressive neurons of the hidden layer and also to each neuron of an output layer. Thus, all the input neurons are related to each and every neuron of the neural network. The generalized equation of the CFBPNN model can be deduced from the above given "(10)" and "(11)" and is given as:

$$z = \sum_{i=1}^{m} g^i\, w_i^i x_i + g^0\left(\sum_{j=1}^{k} w_j^0\, g_j^h\left(\sum_{i=1}^{m} w_{ji}^h x_i\right)\right) \tag{12}$$

In "(12)" g^i denotes the activation function from an input layer towards an output layer and w_i^i depicts the weights used in this function. If the activation function of each neuron present in the hidden layer becomes g^h on the application of bias information. Then, the above equation changes to:

$$z = \sum_{i=1}^{m} g^i\, w_i^i x_i + \left(w^b \sum_{j=1}^{k} w_j^0\, g^h\left(w_j^b \sum_{i=1}^{m} [\![w_{ji}^h x_i)\right)\right)]\!] \tag{13}$$

Here, CFBPNN model is used as a classifier on different benchmark datasets.

4. Methodology and Methods

A. Methodology

The proposed methodology consists of an experimentation which is conducted in two phases i.e, without CFBPNN and with CFBPNN. An initial analysis is done independently on each technique and secondly by adding neural network classifier (CFBPNN) to each technique and average precision has been calculated in both the cases. Thus, the two phases of the experimentation are as follows:

1) 1st phase of experimentation: In this phase of an experimentation, color attributes are extracted from a query image and from the complete database using CM, CAC, CH, CCV and DCD one by one. Then, similarity is calculated using a specific distance metric between feature vectors of a query and database images and average precision is calculated for each technique. Finally, top 10 images are retrieved. The architectural framework for this phase is given in Fig. 2

2) 2nd phase of experimentation: This phase consist of feature extraction using each color descriptor in an amalgamation with a CFBPNN classifier. It is further divided into two stages: Training and Testing stage. Its architectural framework is given in Fig. 3.

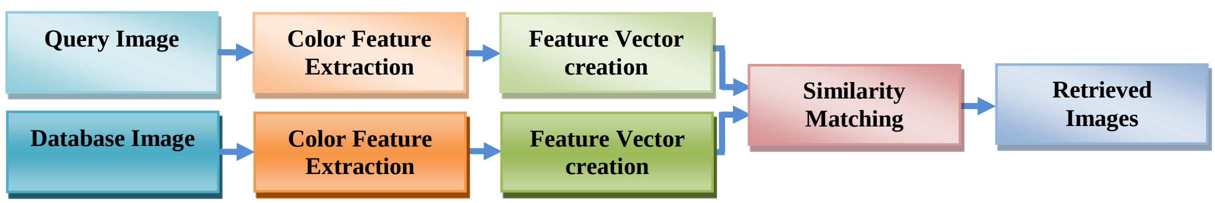

Figure 2. Architectural framework of 1st phase of implemetatuion

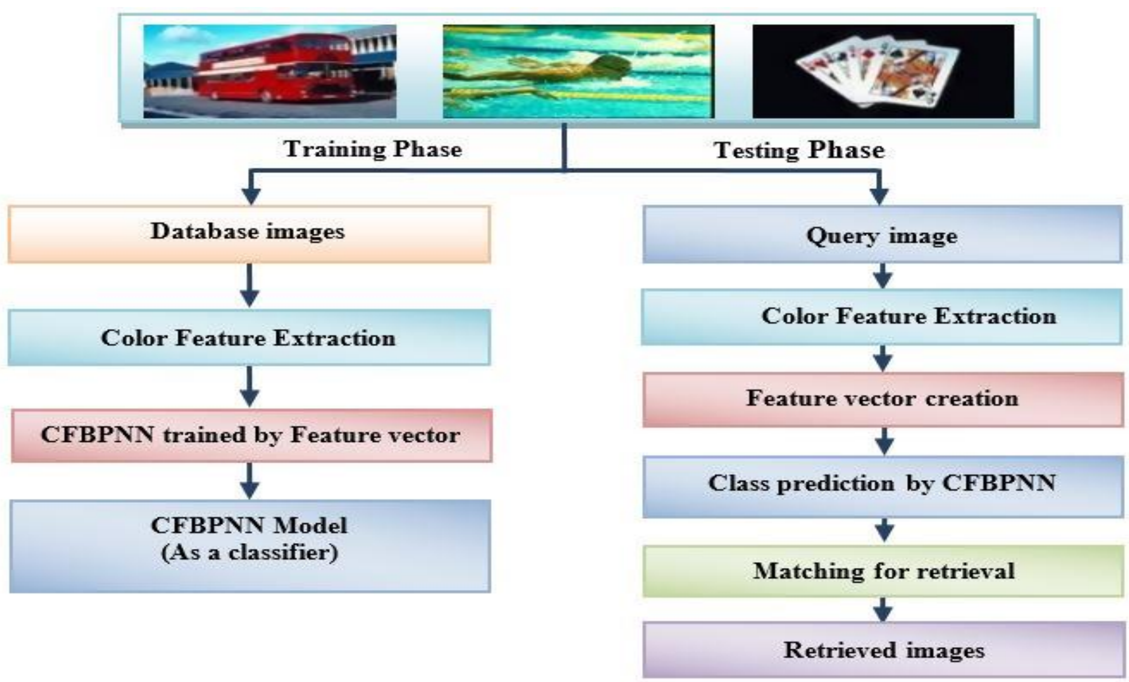

Figure 3. Architectural framework of 2nd phase of implemetatuion

a) Training stage: This stage is concerned with the training of a CFBPNN model. Every neural network is required to be trained through the total images in the database. In this phase, color features are withdrawn from the whole dataset and feature vector is formed for each technique which is applied as an input to a Cascade Forward back propagation neural network.

This network is composed of layers employing DOTPROD weight function, net input function as NETSUM and the prescribed transfer function. The weights of an initial layer are coming from the applied inputs and every succeeding layer has a weight coming from the inputs of all preceding layer. The last layer is denoted as an an output layer. Each layer of the network has a bias associated with it. INITNW function is used to initialize the bias and weights of the neural network.

TRAINS function is utilized for an adaptation and weight updation.

b) Testing stage: The main aim of this phase is concerned with the testing of the experimental system by using a pre-trained CFBPNN model with a specific query image. This testing phase is further classified into three main stages:

- *Feature extraction stage:* In this stage, color features are withdrawn from a query image using the utilized color descriptor independently. Then, a matrix of extracted features is constructed. Every image of a specific dataset is considered as a query image.
- *Classification stage:* The second stage of this testing process is based on classification. Here, the obtained feature vector from the query image is applied in the form of an input to a pre-trained CFBPNN model. As it has initially been trained with the complete dataset, where accurate classification of the whole dataset has already been done with the formation of different categories on the basis of similar semantic features. The obtained feature vector from the query image is matched with various categories of the dataset and an accurate class or category prediction is obtained by the CFBPNN model.
- *Similarity matching stage:* As, the exact category of the given query image has been predicted by CFBPNN model. Now, in similarity matching stage, the query image is matched with all the images of the predicted category by using a specific distance metric technique. Then, based on the retrieval requirements of the user, top ranked N images are retrieved. These top N images are ranked based on the usage of distance metric technique. The two phases of the implementation can be understood by means of an algorithm as under:

Phase I
Step 1: *Feature extraction by using five color descriptors independently and formation of feature vector for each of the utilized technique.*
Step2: *Submission of query image by the user and again feature extraction by using all color descriptors, one by one.*
Step 3: *Matching of the feature vector of the given query image with the obtained feature vectors by using a specific distance metric. Repetition of this step for each technique.*
Step 4: *Ranking and retrieval of top N images by the user.*

Phase II
Step 1: *Training of a neural network model, CFBPNN with the whole dataset of images.*
Step 2: *Repetition of Step 2, from the previous stage.*
Step 3: *Class prediction of the query image by the CFBPNN model and similarity matching is done.*
Step 4: *Ranking and retrieval of desired top N images.*

Some of the prominent distance metrics [30] which are used in similarity calculation are given under:

$$Distance_{Euclidean} = \sqrt{\sum_{j=1}^{n}(|I_j - D_j)^2|} \quad (14)$$

$$Distance_{Manhattan} = \sum_{j=1}^{n}|I_j - D_j| \quad (15)$$

$$Distance_{Minkowski} = [\sum_{j=1}^{n}(|I_j - D_j)^{1/P}|] \quad (16)$$

Here, *Ij* denotes the input query image and *Dj* depicts all database images related to Euclidean, Manhattan and Minkowski distances.

B. Methods

In order to carry the experimentation, the set-up is as follows:

1) Experimental set-up

For the impetus of an analysis and experimentation, the experiments are conducted over three benchmark databases namely Corel-1K, Corel-5K and Corel-10K. The description of these datasets is as follows: Sample images from each dataset is given in Fig. 4(a-c).

Figure 4(a). Sample images from Corel-1K dataset

Figure 4(b). Sample images from Corel-5K dataset

Figure 4(c). Sample images from Corel-10K dataset

1st Dataset: Corel-1K: The initial database analysed for experimentation is Corel-1K which is composed of 1000 different images. It has 10 categories and 100 images are present in each category with each image bearing the size of either 256 × 384 or 384 × 256 [31].

2nd Dataset: Corel-5K: The next dataset is Corel-5K and consists of 5000 images. There are 50 categories in this dataset and again every group has 100 images. The size of each image is either 256 × 384 or 384 × 256 [32].

3rd Dataset: Corel-10K: The third dataset is Corel-10K and there are 10,000 images in this dataset. This database has different groups of 100 images and every set has further 100 images in it. Every image bears the size of either 256 × 384 or 384 × 256 [32].

All the experiments are conducted over MATLAB R2017a, core i3 processor, 4 GB memory, 64 bit windows. For the formation of a query image, each and every single image of all the databases is utilized.

2) Retrieval results

The results based on the retrieval of desired images are obtained by taking each and every image from all the datasets i.e Corel-1K, Corel-5K and Corel-10K as a query image. Then, CM, CAC, CH, CCV and DCD techniques are used for the withdrawal of color attributes one by one. Lastly, top 10 images are retrieved in two ways: (1) Without applying any deep learning classification model (2) With the application of pre-trained CFNN back propagation model as a classifier. Precision and Recall [33-34] are the most well-known evaluation metrics. These are defined using the given equations:

$$Precision = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved} \quad (17)$$

$$Recall = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ relevant\ images\ in\ the\ database} \quad (18)$$

In order to carry out the experimentation, Average precision has been calculated on all the employed datasets using all five color descriptors and the results are given in Table 1, 2 and 3. These results are obtained firstly without using CFBPNN model and again with CFBPNN model to find out the best color descriptor in all the conditions.

TABLE 1. AVERAGE PRECISION (%) ON COREL-1K DATASET

Color Descriptor	Without CFBPNN	With CFBPNN
Color moment	83.5	94.5
Color auto-correlogram	89.95	97
Color histogram	79	90.5
Color coherence vector	68.7	84.6
Dominant color descriptor	77	89

TABLE 2. AVERAGE PRECISION (%) ON COREL-5K DATASET

Color Descriptor	Without CFBPNN	With CFBPNN
Color moment	68.9	86.2
Color auto-correlogram	80	90.5
Color histogram	66.7	84.6
Color coherence vector	60.2	80.1
Dominant color descriptor	63.4	82.7

TABLE 3. AVERAGE PRECISION (%) ON COREL-10K DATASET

Color Descriptor	Without CFBPNN	With CFBPNN
Color moment	65.2	84
Color auto-correlogram	77.2	89.5
Color histogram	63.5	81.9
Color coherence vector	57.6	78.9
Dominant color descriptor	60.2	80.1

From Table 1, 2 and 3, it can be concluded that the average precision is enhanced to a great extent by incorporating this neural network model as a classifier to a basic color descriptor. It can also be concluded that out of five color descriptors, Color auto-correlogram has the highest average precision on all the three datasets in both phases of experimentation.

Five prominent distance metrics are utilized for the purpose of testing. Average precision obtained on the three datasets utilizing the three distance metrics after the implementation of CFBPNN is shown in Table 4.

TABLE 4. AVERAGE PRECISION (%) VS DISTANCE METRICS

Datasets	Euclidean	Manhattan	Minkowski	Cosine
Corel-1K	97	96.2	89.9	89.6
Corel-5K	90.5	85.6	81.8	62.1
Corel-10K	89.5	82.7	77.5	61.7

From Table 4, it can be seen that the average precision obtained by using Euclidean distance metric outperforms the other distance metrics. Since Manhattan distance is a distinctive case of Minkowski distance. It produces innumerable false negatives and does not yield accurate results. Cosine distance also takes long computational time. Euclidean distance metric is based on weighted and normalized attributes and has speedy computational performance. Therefore, the Euclidean distance metric gives precise results and is being used here.

Recall is also an eminent parameter used to measure the effectiveness of any feature extraction technique. Precision vs Recall curves [33] obtained on the three datasets are shown in Fig. 5. From the above experimentation, it can be clearly visualized that CAC outperforms other color descriptors in terms of average precision both without using CFBPNN and with CFBPNN. Finally, the combination of CAC and CFBPNN is selected to plot the precision vs recall curves by varying the number of retrieved images from 10 to 50 on all the three datasets.

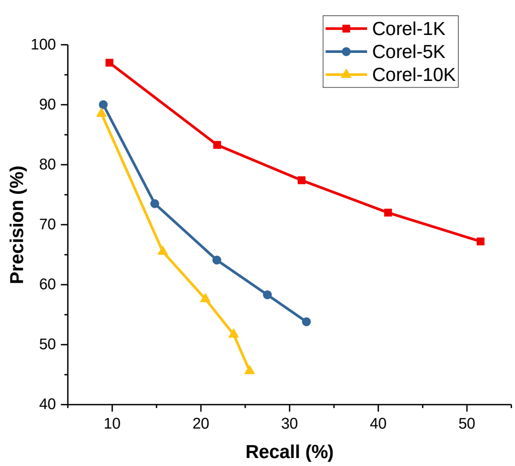

Figure 5. Precision vs Recall curves by using CAC and CFBPNN

In order to retrieve the desired images, separate Graphical user interfaces (GUI's) have been designed for three datasets. In order to show the retrieval accuracy of CAC with CFBPNN, a GUI image depicting the top 10 results for Corel-1K dataset is given in Fig. 6(a-b).

In order to check the classification accuracy of each color feature extraction technique, two types of neural network classifiers have been tested. There are two ways to model the neural network. One is Cascade forward back propagation neural network where input from a preceding forward back propagation neural network (FFBPNN) which can be trained to classify the given inputs according to the classes of targets. In case of Patternnet classifier, all zero vectors are present in target data except the desired class which has to be represented by the same neural network classifier. This desired class has a value of 1 presented in it.

Figure 6(a). Retrieval results on Corel-1K by using CAC only

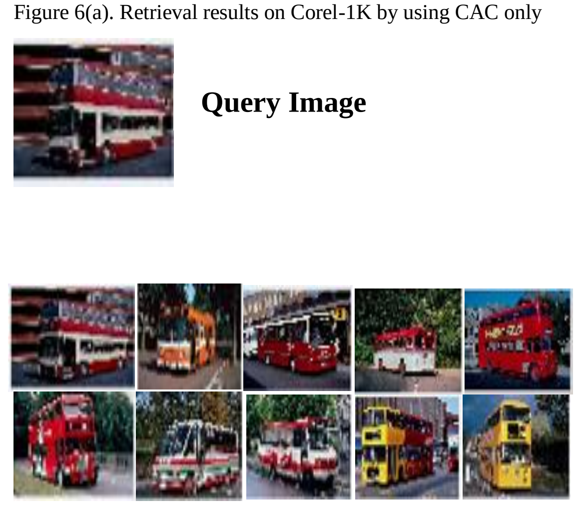

Figure 6(b). Retrieval results on Corel-1K by using CAC with CFBPNN

From Fig. 6(a-b), it can be seen that all the 10 images belonging to the same category of the query image were retrieved by using a combination of CAC and CFBPNN in comparison to the images retrieved by using only Color auto correlogram (CAC). The images which were retrieved without using any neural network classifier have less accuracy as the all top 10 images did not belong to the same native category. Thus, it can be concluded that the combination of color descriptor and neural model based classifier increases the accuracy of the system.

The plots of the classification accuracy obtained on the three benchmark datasets by utilizing the two Neural network classifiers is shown in Fig. 7(a-c). This accuracy has been calculated on all the five color feature descriptors by increasing the hidden layer of neurons from 10 to 15 in each neural network classifier.

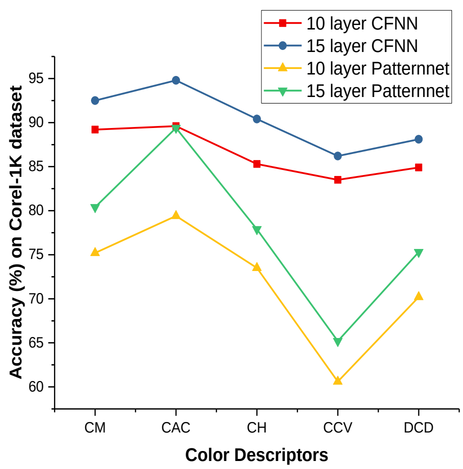

Figure 7(a). Accuracy on Corel-1K dataset

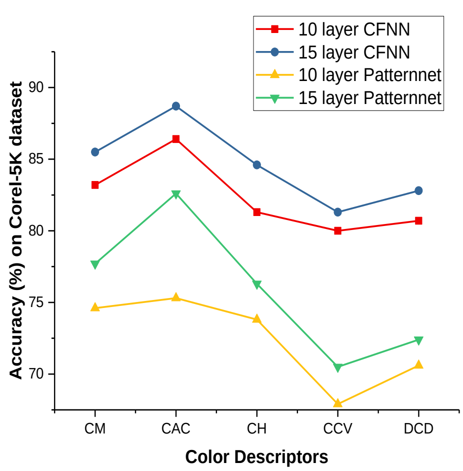

Figure 7(b). Accuracy on Corel-5K dataset

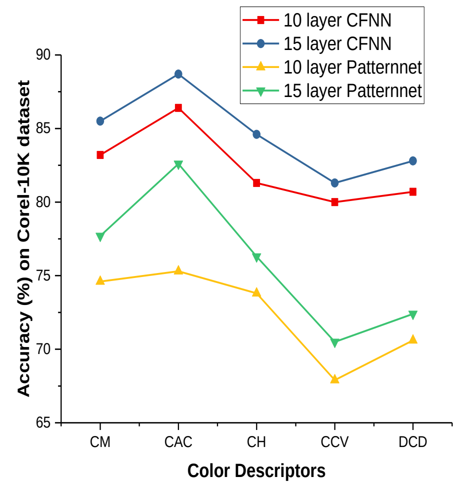

Figure 7(c). Accuracy on Corel-10K dataset

From this Fig. 7, it is clear that the classification accuracy of CFNN classifier is significantly larger than the accuracy obtained through patternnet classifier.

Also, it can be concluded that as the number of neurons present in the hidden layer are increased from 10 to 15, there is a cumulative increase in the classification accuracy of each utilized color feature descriptor. The Levenberg-marquardt algorithm is utilized here, for the training purpose of a CFBPNN classifier. Various parameters utilized for experimentation are given in Table 5.

TABLE 5. VARIOUS WORKING PARAMETERS

Size of each descriptor/No. of features in each	CM (9), CAC (256), CH (32), CCV(6) and DCD(15)
Algorithm used	Cascade-forward back propagation neural network
Training technique	Levenberg-Marquardt

Retrieval time of the CBIR system is an important parameter to be considered. Therefore, to validate this issue, the retrieval time for all the utilized color descriptors in combination with CFBPNN, on Corel-1K dataset is given in Table 6. To compare the precision accuracy of CAC descriptor with CFBPNN on various training and testing ratios of the dataset, an experimentation has also been done and the results are given in Table 7.

TABLE 6. RETRIEVAL TIME (IN SECONDS) ON COREL-1K

Technique	Retrieval time based on Corel-1K dataset
Color moment+ CFBPNN	0.24
Color auto-correlogram + CFBPNN	0.188
Color histogram + CFBPNN	0.32
Color coherence vector+ CFBPNN	1.24
Dominant color descriptor + CFBPNN	0.93

TABLE 7. AVERAGE PRECISION (%) OF CAC-CFNN COMBINATION

Training Vs Testing Ratio	Corel-1K	Corel-5K	Corel-10K
50: 50	88.5	81.36	78.4
60: 40	90.4	83.7	80.1
70: 30	95.2	87.2	85.3

From the above discussion, it is already clear that the highest average accuracy is obtained by using the color auto-correlogram, color feature extraction technique. Therefore, the observations obtained by using color auto-correlogram with CFNN classifier on Corel-1K dataset, consisting of 1000 images are shown in Fig 8(a-b).

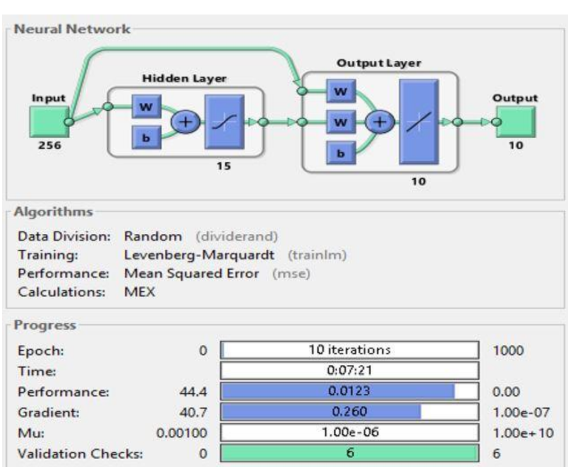

Figure 8(a). Simulation of CFBPNN

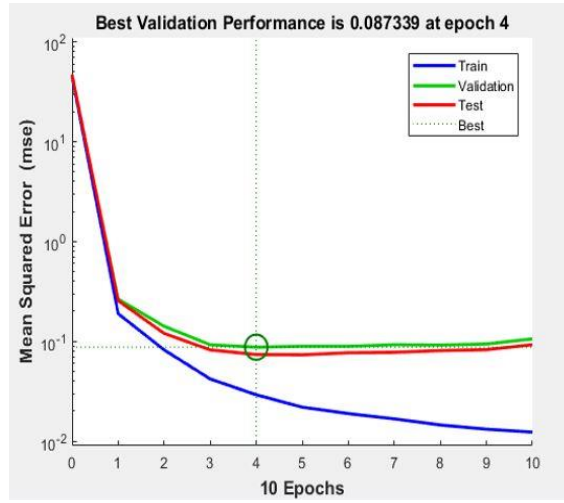

Figure 8(b). Validation graph

The best validation performance plot between Mean Square error and epochs depict that the best performance of the neural network is obtained at an epoch 4. After this epoch 4, the validation remains almost constant.

Thus, this work is an effective framework of CAC and CFBPNN. In real life or practical applications, this work can be utilized for the retrieval of desired images from large art collections gallery, photograph archives, textile industry, medical diagnosis, etc.

5. CONCLUSION WITH FUTURE WORK

This paper depict an experimental and analytical analysis on chosen five color feature descriptors which are Color moment (CM), Color auto-correlogram (CAC), Color histogram (CH), Color coherence vector (CCV) and Dominant color descriptor (DCD). Initially, these color extraction techniques are used and tested one by one to extort the features of an image and average precision is calculated for each of them. In the second phase of an experimentation, a Cascade forward back propagation neural network (CFBPNN) has been used as a classifier in combination with each of the used color feature descriptor and the performance of each of these techniques is again evaluated in terms of average precision. It can be easily evaluated from these two phases of the experimentation that an amalgamation of a neural network with the feature extraction technique leads to a remarkable enhancement in the performance of each technique. It can be concluded that a framework of Color auto-correlogram and CFBPNN outperforms the other color descriptors by obtaining average precision of 97%, 90.5% and 89.5% on Corel-1K, Corel-5K and Corel-10K benchmark datasets respectively. In future, the same technical and comparative analysis will be adopted for texture feature descriptors to determine the optimum texture feature extraction technique. Finally, the proposed work will be extended for the designing of a hybrid CBIR system by combining color, texture and shape feature extraction techniques. Also, in future, different Deep learning classifiers can be tested and compared to analyze the performance of the hybrid CBIR system on more benchmark datasets like COIL, Oxford flower, ZUBUD, MIT-Vistex, INRIA-Holidays, etc.

References

[1] Alzu'bi A., Amira A., and Ramzan N. (2015). Semantic content-based image retrieval: A comprehensive study. *Journal of Visual Communication and Image Representation* 32: 20-54.

[2] Pavithra L.K., and Sharmila T.S. (2017). An efficient framework for image retrieval using color, texture and edge features. *Computer and Electrical Engineering,* 0: 1–14.

[3] Naveena A.K, and Narayanan N.K. (2016). Image Retrieval using combination of Color, Texture and Shape Descriptor. Proceedings in *IEEE International Conference on Next Generation Intelligent Systems (ICNGIS)* 1-5.

[4] Bayan B., and Lamiaa A. E. (2018). Arabic Manuscript Content Based Image Retrieval: A Comparison between SURF and BRISK Local Features. *International Journal of Computing and Digital Systems,* 7(6): 355-364.

[5] Guang-Hai Liu, and Jing-Yu Yang. (2013). Content-based image retrieval using color difference histogram. *Pattern Recognition* 46(1): 188-198.

[6] Fadaei S., Amirfattahi R., and Ahmadzadeh M. R. (2017). New content-based image retrieval system based on optimised integration of DCD, wavelet and curvelet features. *IET Image Procesing*, 11(2): 89–98.

[7] Bhardwaj S., Pandove G., and Dahiya P. K. (2019). An Intelligent Multi-resolution and Co-occuring local pattern generator for Image Retrieval. *EAI Endorsed Transactions on Scalable Information Systems*, 6(22): e1, 1–12.

[8] Waristo B., Santoso R., Suparti, and Yasin H. (2018). Cascade forward neural network for time series prediction. *Journal of Physics Conference Series*, 1025(012097): 1-9.

[9] Kumar A.R, and Saravanan D. (2013). Content-based Image Retrieval using Color Histogram. *International Journal of Computer Science and Information Technolgy* 4(2): 242–245.

[10] Pandey D. (2017). An Efficient Low Level Features of CBIR using Wavelet, GLDM and SMOSVM Method. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 10(3): 13–22.

[11] Singh C., and Preet Kaur K. (2016). A fast and efficient image retrieval system based on color and texture features. *Journal of Visual Communication and Image Representation,* 41: October, 225–238.

[12] Chaudhari R, and Patil A.M. (2012). Content-based Image Retrieval Using Color and Shape Features. *International Journal of Advance Research in Electrical, Electronics and Instrumentation Engineering* 1(5): 386-392.

[13] Shahbahrami A. (2008). Comparison Between Color and Texture Features for Image Retrieval. *Journal of Iran*, 27648: 1-9.

[14] Tripathi N. (2016). A New Technique For CBIR with Contrast Enhancement using Multi-Feature and Multi Class SVM Classification. Proceedings in *International conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, 2031–2036.

[15] Chen Wei-Ta, Liu Wei-Chuan, and Chen Ming-Syan. (2016). Adaptive color feature extraction based on image color distributions. *IEEE Transactions on Image Processing*, 19(8): 2005–2016.

[16] Suliang Y, Dongmei N, Liangliang Z, et al. (2018). Colour image retrieval based on the hypergraph combined with a weighted adjacent structure. *IET Computer Vision* 12(5): 563-569.

[17] Lande M.V, Bhanodiya P, and Jain P. (2014). An Effective Content-Based Image Retrieval Using Color, Texture and Shape Feature. Proceedings in *Intelligent Computing, Networking, and Informatics, Advances in Intelligent Systems and Computing, Springer* 1163–1170.

[18] Jain A.K, and Vailaya A. (1996). Image retrieval using color and shape." *Pattern Recognition*, 29(8): 1233-1244.

[19] Saritha R, Varghese P, and Ganesh K. (2018). Content-based image retrieval using deep learning process. *Cluster Computing* /doi.org/10.1007/s10586-018-1731-0, 1-14.

[20] Weibo L, Zidong W, Xiaohui L, et al. (2017). A survey of deep neural network architectures and their applications." *Neurocomputing,* 234: 11-26.

[21] Maria T, and Anastasios T. (2018). Deep convolutional learning for Content-based Image Retrieval." *Neurocomputing,* 275: 2467-2478.

[22] Tanya, S. J., and Ali, J. D. (2018). Computer-Aided Diagnosis Psoriasis Lesion Using Skin Color and Texture Features. *International Journal of Computing and Digital Systems,* 7(3): 145-154.

[23] Jain N, and Salankar S.S. (2014). Color & Texture Feature Extraction for Content-based Image Retrieval. *IOSR Journal of Electrical and Electronics Engineering,* 53–58.

[24] Patel J.M. (2016). A Review on Feature Extraction Techniques in Content-based Image Retrieval. Proceedings in *IEEE International Conference*, 2259–2263.

[25] Singh S.M, and Hemachandran K. (2012). Content-Based Image Retrieval using Color Moment and Gabor Texture Feature. *International Journal of Compter Science Issues,* 9(5): 99–309.

[26] Vinayak V. (2017). CBIR System using Color Moment and Color Auto-Correlogram with Block Truncation Coding. *International Journal of Computer Applications,* 161(9): 1–7.

[27] Pass G, Zabih R, and Miller J. (1998). Comparing images using color coherence vectors. Procedings in *fourth ACM International Conference Multimedia (MULTIMEDIA '96)* 1–14.

[28] Huthaifa, M. Saher and Hazem, H. (2018). A Flower Recognition System Based On Image Processing And Neural Networks, *Int. j.of scientific & technology research* , 7: 1-9.

[29] Omaima N.A, Tamimi A.A, and Alia M.A. (2013). Face Recognition System Based on Different Artificial Neural Networks Models and Training Algorithms. *International Journal of Advance Computer Sciene and Applications,* 4: 40-47.

[30] Mistry, Y., Ingole, D.T., and Ingole, M.D. (2017). Content based image retrieval using hybrid features and various distance metric, *J. Electr. Syst. Inf. Technol.*, 2016, 1–15.

[31] http://Corel-1K.ist.psu.edu/ docs/related/

[32] http://www.ci. gxnu.edu.cn/cbir/

[33] Bhardwaj S., Pandove G., and Dahiya P. K. (2020). 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*. 12: 1-13.

[34] Naghashi, V. (2018). Co-occurrence of adjacent sparse local ternary patterns : A feature descriptor for texture and face image retrieval,*Opt. - Int. J. Light Electron Opt.*, 157: 877–889.

[35] Elnemr, H.A. (2016). Combining SURF and MSER along with Color Features for Image Retrieval System Based on Bag of Visual Words. *J. of Comp. sciences*, 213-222.

Shikha Bhardwaj received her B.Tech degree in ECE from Kurukshetra University, Haryana, India in 2006, Master's degree in ECE from M.M University, Ambala, Haryana, India in 2009. Now, pursuing Ph.D in Electronics and Communication Engineering Department from DCRUST University, Murthal, Haryana, India. Since 2011, she is currently working as an Assistant Professor in University Institute of Engineering & Technology, Kurukshetra University, Haryana, India. Her major research areas are Image processing, Artificial intelligence, Machine learning etc.

Gitanjali Pandove received her Ph.D degree in ECE. She is working as Professor in ECE Deptt., in DCRUST, Murthal, Haryana, India. Her research areas are optical communication, digital image processing, power systems, etc. She has more than 70 publications in various international reputed journals.

Pawan Kumar Dahiya received his Ph.D degree in ECE. He is working as a Professor in ECE Deptt., in DCRUST, Murthal, Haryana, India. His research areas are Digital and Embedded System Design, ANPR, VANETs, Soft Computing Techniques, Image Processing, Internet of Things (IoT), etc. He has more than 75 publications in various national and international reputed journals.