EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

# A Competent and Novel Approach of Designing an Intelligent Image Retrieval System

Shefali Dhingra[1,*], Poonam Bansal[2]

[1]Ph.D Research Scholar, Guru Gobind Singh Indraprastha University, New Delhi, India
Assistant Professor, University Institute of Engineering & Technology, Kurukshetra University, Kurukshetra, India
[2]Professor, Maharaja Surajmal Institute of Technology, New Delhi, India

## Abstract

A prominent technique used to search imperceptible images from any vast repository is denoted by Content based image retrieval (CBIR) system. The most eminent features of CBIR system are Texture, Color and Shape and in this paper these features are extracted using Local binary pattern, Color moment and Automatic segmentation process respectively. The features of these descriptors are combined for the formation of a hybrid feature vector by utilizing the process of normalization. Then, for enhancing the retrieval accuracy of the proposed system, Extreme learning machine (ELM) has been utilized as a classifier. The proposed hybrid CBIR system with ELM leads to an evolution of an intelligent system. Various evaluation metrics like Precision, Recall, Accuracy, etc. have been calculated on the proposed system on standard datasets. Average precision of 90%, 78%, 72% and 88.70% has been obtained respectively on Corel-1K, Corel-5K, Corel-10K and GHIM datasets which is significantly higher than the related techniques.

**Keywords:** Content based Image retrieval (CBIR), Local Binary Pattern (LBP), Precision, Deep Learning Algorithm, Extreme learning Machine, Neural Networks

Received on 24 February 2019, accepted on 24 September 2019, published on 30 September 2019



doi: 10.4108/eai.13-7-2018.160538

*Shefali Dhingra. Email: munjalshefali@gmail.com

## 1. Introduction

Due to breakneck increase in digital images in different fields such as remote sensing, medical imaging, crime prevention etc., large number of digital image libraries exists with huge amount of images. So for proper indexing and retrieving these images from the large databases an adequate and convincing retrieval system is required. These image retrieval systems are chiefly of two types which are text based and image based. Conventionally text based system was used by which images were retrieved through specific keywords which are entered manually. But this system proved to be not so accurate in case of large image databases. In this human annotation is required which is a very burdensome task. So to overcome this stiff task CBIR system comes into action in which images are retrieved automatically by the help of low level features of the image such as color, shape and texture [1]. The CBIR retrieves the relevant images by inputting the query image. The similar features are extracted from the query and database images and then these features are utilized to build a feature vector.

After the formation of feature vector the similarity is calculated between the images by various means to find out the relevant images from the database. The traditional image retrieval approaches depends on the color feature. Generally color is an observable feature of an image or any object. For particular geometrical transformations, it is invariant. This color feature is used in terms of color correlogram[2], color histogram, dominant color descriptor (DCD) [3], color moments and many more. Though, the feature of color clearly shows the perceptible property which corresponds to the object but it does not differentiate among textures and shapes of the objects. Sometimes two images can have similar histogram but at the same time it cannot be expressing the same objects or images. Then texture is one of the features that is being used from the long period. Generally texture feature can be extracted by Discrete wavelet transform (DWT), Gabor filter [4], Fourier transform[5] etc. The modern developments in this field is using local patterns for example Local Ternary Pattern [6] and Local Binary Pattern. After color and texture, shape is another main feature. Generally, shape feature can't be exploited as a single feature; it is generally used when integrated with color or texture for obtaining the desired results. When the image is a complex one, only the use of primary feature or a single feature will not be sufficient. Because a single feature is not being able to capture the variable details present in the images. To overwhelm this problem combination of features are employed. The basic block diagram of CBIR system is shown in Figure 1. Along with feature selection, the optimization of features is also very important. An Antlion optimization algorithm is proposed based on the chaos principles is used for the feature selection. This algorithm generates the fitness function which enhances the classification and also helps in reducing the number of the features [7].

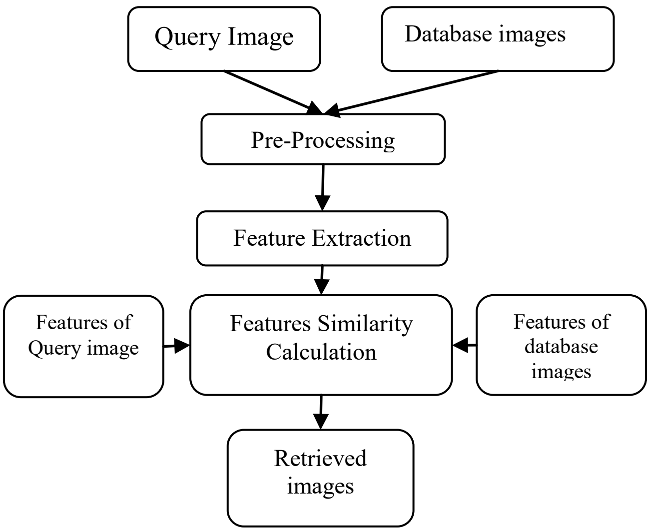


**Figure 1.** Basic block diagram of CBIR system

This work proposes hybrid CBIR system. The features such as shape, texture and color feature are extracted. The color features are extracted by deploying color moment. The texture features are extracted by using LBP. The shape features are extracted by using various parameters such as centroid, area, perimeter and solidity of the image which provides the approximate idea of the shape of an image. After extracting all these three features from the images these are fused together through normalization and a fused feature vector is formed of all the images which is undoubtedly more competent than single feature vector

The semantic gap is the major limitation of these systems that is the difference in the low level features of the image captured by the machine and human perception. To blown away this gap deep learning algorithms are now employed. Deep learning algorithms are the subset of machine learning algorithms. These algorithms have a very deep architecture with multiple layers of transformations. They work like the human brain as they train the computers according to human understandings.

Deep learning networks uses the neural networks and the deep term means the number of hidden layers used in the network. The training of these models is done by using the labelled datasets without any manual intervention. There are various applications in which these networks are used such as automated driving, image classification in large datasets, medical research and many more. In the recent literature, these algorithms have been used in the CBIR system for the feature extraction but very less consideration is given on the issue of semantic gap which decreases the overall accuracy of the system.

The proposed framework here is different from others as in this system firstly the low level features are extracted from different techniques and a hybrid CBIR system is obtained. After that to increase the accuracy of the system and to overcome the semantic problem machine learning algorithm is applied in large datasets which works as a classifier. While in the literature these algorithms are mostly used for the extraction of features. This is the main novelty of this paper,

To improve the accuracy of the overall hybrid system an intelligent deep learning algorithm is deployed i.e. Extreme Learning machine. The ELM is a feed forward neural network with a single hidden layer. With least human interference it provides better performance along with higher speed as compared with other feed forward neural networks [8]. This machine learning algorithm can be used for performing various tasks such as classification, regression, pattern recognition etc. In this paper ELM is modelled as a classifier for the classification of images into various classes so that the searching of images becomes an easy task and the system becomes more accurate. At last, the images are retrieved by calculating the similarity using Euclidean distance measure.. The main contribution of this work is listed below:

- To extract features (color, texture and shape) from images and a combined feature vector is obtained.
- To classify the optimized features, Extreme learning Machine is proposed.
- Finally, the retrieval process is carried out by the similarity calculation between query and database images.

This paper is structured in the following way: The Section 1 provide the introduction about Image Processing, Image

Retrieval, CBIR technique, and simple overview of the proposed approach. Section 2 reveals about the literature which is interrelated to the proposed work of Hybrid and intelligent CBIR system. Section 3 gives the details about feature extraction techniques and classifier used. In the next section the main proposed work by an overall flow diagram, and algorithm is explained. Section 5 is the result and discussion section. In this section the outcomes of an intelligent CBIR system and the performance comparison with the existing is given with graphical representation. Section 6 provides the concluding summary the work.

## 2. Related Works

This section gives the review on various previous works mainly on hybrid CBIR systems. Different features extraction techniques that are used in these systems have their own pros and cons and when these features are properly selected and combined together the system become more compelling. The two features were combined i.e. discrete wavelet transforms and edge histogram descriptor for texture and shapes respectively [9]. The edge histogram was applied on the selected wavelet coefficients which come after the decomposition of the input image. This combination had increased the performance of the system. Sometimes only the single texture feature provides the good results such Local tri-directional pattern was proposed in [10] which was the texture feature and uses the pixel intensity in three directions. It was then compared with other existing systems based on image retrieval applications

Hassan Farsi et al. [11] described a new method which combines the DWT and Handamard matrix which results in increasing the accuracy and speed of the CBIR system. For measuring the output performance of the above system combination of recall, precision and normalized rank are evaluated. The results obtained by HDWT provide better performance as compared with color layout descriptor, Haar wavelet transform, histogram intersection and dominant color descriptor. Sadegh Fadaei et al. proposed the CBIR scheme on the basis of optimized combination of texture and color features to improve the precision value of image retrieval. Dominant Color Descriptor (DCD) features were extracted from HSV color space and to extract texture features wavelet and curvelet were applied and finally these two features are combined optimally by optimization algorithm which is particle swarm optimization algorithm [12].

Feature selection and the optimization of the features is the most crucial task for designing of the systems. Various researches have been done in this area also. A chaotic crow search algorithm is introduced in [13] to overcome the limitations of simple crow search algorithm. The entrapment in local optima and low convergence rate are the basic problems with simple CSA. In Mutasem K. Alsmadi proposed an efficient CBIR system using memetic algorithm for retrieval of images from the large database. The important features like color signature, shape and texture were extracted and memetic algorithm was applied for the similarity measurement between query image and images in the database [14].

In [15] the CBIR was presented by developing the Ordered-Dither Block Truncation Coding (ODBTC). The image features were extracted by two techniques which are Bit Pattern Features (BPF) and Color Co-occurrence Feature (CCF). These features were developed from the encoded data streams of ODBTC in a direct manner without the execution of decoding process. The results from the experiments showed that the proposed approach was superior to the existing approaches such as Block Truncation Coding (BTC) image retrieval systems. Srivastava et al. [16] proposed a multiscale LBP technique for the retrieval of images from large databases. Final feature vector is obtained from GLCM technique. The results are evaluated from many benchmark datasets and proved to be much effective than single scale LBP technique used for extracting texture of the images. In [17] hybrid system is designed by using color and texture extraction algorithms. The color features are represented by color histogram at the same time to describe the entire image texture feature extraction was carried out by the Gabor filter which has already proven its effectiveness in defining the visual content through the analysis of multi-resolution. Guo et al. proposed `the new technique for indexing of the color images by using the same ODBTC technique which obtains the features of the images. By the use of two ODBTC quantizers the descriptors of image contents are generated. PSO algorithm is deployed here for determining the proper similarity constants which in turn improves the accuracy of image retrieval [18]

For large datasets image classification is an important task in CBIR systems. In this work various classification methods along with their pros and cons are discussed such as SVM, SLFN and experimented [19].

In [20] an overview of machine learning algorithms is presented along with the various applications in the field of image processing. Various advanced deep learning algorithms along with their challenges and designing issues are discussed. [21] proposed a feed-forward back propagation neural network for global image properties which was based on CBIR. Initially, the neural network has been trained on image features. The considered image features are the color histogram that acted as a color descriptor. The GLCM acted as a texture descriptor. The edge histogram acted as edge descriptor. The results have been shown significant developments in recall and precision of image retrieval on simple databases.

For the purpose of classification of the optimal feature a new algorithm is proposed by Emary eta al. [22] This work is based on the binary version of Gray wolf optimization technique. This binary GWO is the combination of two approaches were used here for determining the minimum features from the complete feature set and for increasing the accuracy of the system.

In another paper two innovative approaches for image descriptors were introduced. This was based on scale invariant feature transform (SIFT) algorithm. By using SIFT; the image features have been extracted. After this, the k-means clustering has been applied for extracting the feature matrix. This proposed approach has the advantage of highly optimizing the memory storage in addition to advantage of SIFT features [23].

Mohamed et al. presented an innovative approach CBIR-PE. It has been developed to overcome the issues such as data security, bandwidth requirement and storage space. This approach implemented a combination of various clustering approaches WBCT-FCM that utilized a wavelet based contourlet transform (WBCT) as well as and fuzzy clustering-means (FCM) algorithm. These methods are experimented on COREL- 1K database and the outcomes from these experiments showed that this proposed approach attained improved results I terms of accuracy [24].

In [25] the features were not compared for similarity measurement but the convex hull geometry of the images are compared for the retrieval of relevant images. The convex hull area ratio was used as the relative metric for the comparison of the images. In this method the convex hull values of the images are extracted and CHAR is calculated. Images are sorted in terms of decreasing values of CHAR means higher the value of CHAR, maximum is the similarity

With the invention of intelligent systems the soft computing (SC) techniques are now explored very much in various applications. They help in providing the solution of complex problems by dealing with the approximate model. The basic components of soft computing techniques are machine learning, fuzzy logic and evolutionary computation [26]. Soft computing are used in various areas such as designing of smart systems, cloud computing, big data etc.

There are basically two main assets of soft computing techniques. First is it provides the acceptable solution of the complex problem according to the human perception and second these techniques are able to solve the non-linear problems. The comparison between all the three constituents of SC techniques is also done in this paper [27].

In [28] a CBIR system is designed based of fuzzy component of soft computing technique. For the selection of best features from the images a new extraction technique is proposed known as incremental filtering algorithm. As the fuzzy logic results matches with the human understanding so this algorithm reduces the problem of semantic gap. Manoharan Subramanian et al. compare all the three components of SC techniques like fuzzy logic, machine learning and evolutionary computation by incorporating them into CBIR systems [29].

## 3. Feature Extraction and Classification

The feature extraction techniques that are recommended in this work are color moment for color, Local binary pattern (LBP) for texture and auto segmentation is used for the shape extraction. These techniques are explained in detail in the below headings.

### 3.1 Color moment for color feature extraction

Color moment technique is preferentially selected for capturing the details of the images due to its lowest complexity and quicker response than other methods such as histogram based or dominant color descriptor. Moreover, it also increases the effectiveness of the system [30].

It provides the statistical measures which are able to express all the important details present in the image. Only in two compressed forms it gives the information regarding the pixel distribution in the images [31].

First and second order moments i.e. mean and standard deviation are globally calculated from RGB color space which are shown in Equations (1) and (2). Mean provides average information of the color in the image and standard deviation is the number of pixels which varies from the mean.

$$\text{Mean (Ir)} = \frac{1}{XxY}\sum_{i=1}^{X}\sum_{j=1}^{Y} Pcij \quad , \text{r=[R, G, B]} \tag{1}$$

Ir = color channel information

X, Y= row and column size of image

$Pcij$=image pixel value in ith row and jth column

$$\text{Std } (I_r) = \left(\frac{1}{XxY}\sum_{i=1}^{X}\sum_{j=1}^{Y}(P_{cij} - Mean(I_r))^2\right)^{\frac{1}{2}}, \text{ r =[R, G, B]} \tag{2}$$

### 3.2 Texture feature using Local Binary Pattern

Texture is also the dominant feature in CBIR systems for retrieval of relevant images. LBP feature extraction technique for the texture is widely used in various applications of image processing due to its simplicity, performance and implementation. So the designed hybrid CBIR system uses LBP for the extraction of texture from the images [32].

The LBP texture descriptor is extensively used due to its rotational and illumination invariant properties. Firstly, pre-processing step is performed in which the RGB image is transformed to grey scale. After this, the image is divided into smaller sub matrices of size 3×3 from which the feature extraction take place. All the features obtained by these smaller sub matrices are integrated to form the one feature histogram which represents the whole image [33].

Its working is based on the difference between the centre value of the pixel and the neighbour pixel. The calculation of LBP is shown in Figure 2. The binary code for every pixel is produced by the step of thresholding of neighbour pixel with the centre pixel as given in Equation (3) and Equation (4).

$$LBP_N = \sum_{i=0}^{N-1} f(P_i - CP)\, 2^i \tag{3}$$

$$\text{f (p)} = \begin{cases} 1; P \geq 0 \\ 0; P < 0 \end{cases}, \text{ N is total neighbouring pixels.} \tag{4}$$

$$Hist_k = \sum_{x=1}^{M} \sum_{y=1}^{M} f_2 \left(LBP(x,y), k\right) \quad (5)$$

In this figure first 3×3 sub block is taken from the size 5×5. The thresholding value i.e. the centre pixel value is 34 for other 8 pixels and the code produced for this is 92. Local binary patterns are calculated for the complete image and after that the histogram is framed which describes the complete texture of the image. Histogram of every bin is computed by summation of the image pixel numbers as shown in Equation (5) [34].

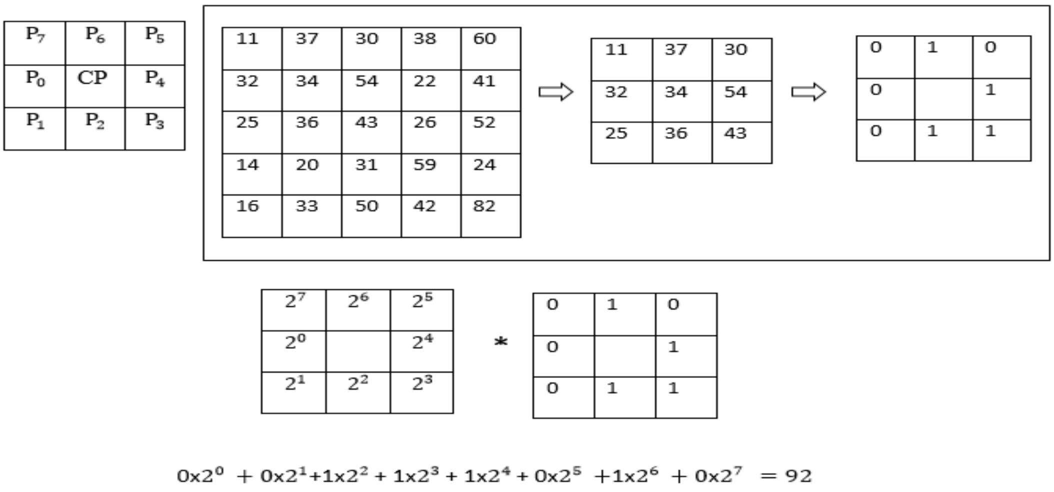


**Figure 2.** Calculation of Local Binary Pattern for every pixel

## 3.3 Shape feature Extraction

As color and texture, shape is also the considerable feature for these systems when integrated with color or texture. However, its extraction from the images is a tough task as for doing this job one dimensional data is lost when a real 3-D object gets projected over the 2-D image plane [35]. The shape features are mainly extracted by using the region areas and contours. The models used for contours and areas are basically Fourier descriptors, spline fitting curves, gaussian curves etc. But the limitations with these models are that some boundary of the shape information is missed out and the performance of the system goes down [36, 37].

In this paper the shape is extracted by using different aspects with the calculation of various shape parameters such as mass, centroid, dispersion, solidity, variance, mean etc [38]. These obtained parameters provide an approximate idea of the shape of the image. In this technique the image is segmented into classes. The most important attributes of the shape i.e. mass, centroid and dispersion are evaluated and stored as a shape feature vector.

These are described in Equations (6), (7) and (8).

Mass is defined as the total number of pixels present in a single class.

$$\text{Mass} = \sum_{xy} m(x,y) \quad (6)$$

Where,

$$m = \begin{cases} 1, \ if\ i(x,y) \in c \\ 0, if\ i(x,y) \notin c \end{cases} \quad (7)$$

Centroid is defined as the mask's centre where m is the mask of the cluster which is given as c, over the image i(x, y). The $x_c$ and $y_c$ are the co-ordinates which are given as

$$x_c = \frac{\sum_{xy} x * m(x,y)}{mass} \quad (8)$$

$$y_c = \frac{\sum_{xy} y * m(x,y)}{mass} \quad (9)$$

Dispersion is defined as the summation of all the regions of the class from the calculated centroid. Euclidean measure is used for this distance measurement.

$$D = \sum_i dist\left(d_c, d_{i,c}\right) \quad (10)$$

Where, dist $(d_c, d_{i,c})$ is the Euclidean distance

$d_c$ Represents the centroid having class c

$d_{i,c}$ Represents the centroid of region of class c

### 3.4 Extreme Learning Machine as a Classifier

In order to overcome the constraints of single layer feed forward networks or back propagation networks which have to face complications during tuning the control parameters manually, the Extreme learning machine is deployed which is implemented automatically without any human intervention.

In ELM the learning parameters such as weights, input nodes and biases are assigned inconstantly in which no tuning is required. The output is determined empirically by the inverse operation and only the number of hidden nodes needed should be defined. So in comparison with other learning algorithms, the ELM has a quicker learning speed, much better performance and that too with lowest manual interference [8].

ELM algorithm has been employed successfully in many areas such as classification, regression, pattern recognition, image classification and many more. In this work ELM has been used as a classifier for the classification of images in large databases. Here Radial basis function (RBF) is used as an activation function due to its universal approximation ability [39].

The database images are trained by ELM classifier and the testing images (query image) are classified into the particular class after which the similarity between the query and training images is carried out and relevant images are retrieved.

The ELM is described below in detail.

If there are X different samples of training ($a_i$, bi ) $\in R^n \times R^m$ ( i= 1, 2, ……X), then the output of Single Layer Feed forward Network with the hidden nodes $\tilde{X}$ (RBF) is given by Equation (11).

$$O = \sum_{i=1}^{\tilde{X}} \beta_i f_i(a_j) = \sum_{i=1}^{\tilde{X}} \beta_i f_i\left(a_j\,;\, c_i, d_i\right) \qquad (11)$$

j = 1, 2……X

$c_i = [c_{i1}, c_{i2} \dots\dots c_{in}]^T$ and $d_i$are the randomly generated parameters through the jth hidden node.

$\beta_i = [\beta_{i1}, \beta_{i2} \dots\dots \beta_{im}]^T$ are the connection links between the hidden and output nodes.

Activation function of ELM is given by $f\left(a_j;\, c_i, d_i\right)$

Now, the Equation (12) is the result when $c_i \cdot a_j$ is the inner product of $c_i$ and $a_j$

$$\mathrm{H}\beta = O \qquad (12)$$

Where

$$\mathrm{H}=\begin{bmatrix} f(c_1 \cdot a_1 + d_1) & \cdots & f(c_{\tilde{X}} \cdot a_1 + d_{\tilde{X}}) \\ \vdots & & \vdots \\ f(c_1 \cdot a_X + d_1) & \cdots & f(c_{\tilde{X}} \cdot a_X + d_{\tilde{X}}) \end{bmatrix}_{X x \tilde{X}}$$

The output matrix of hidden layer is H

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{X}}^T \end{bmatrix}_{\tilde{X} xm} \quad O = \begin{bmatrix} O_1^T \\ \vdots \\ O_X^T \end{bmatrix}_{Xxm}$$

The learning parameters $C_i$ and $d_i$ are randomly assigned without taking consideration of the input.

By finding the least square solution the output weights are calculated in Equation 13.

$$\hat{\beta} = H^{\dagger} T \qquad (13)$$

Where $H^{\dagger}$ is defined as the generalized function of H given by Moore Penrose. Hence the output weights are simply calculated by mathematical transformation avoiding any complicated training phase. The ELM model is shown in Figure 3.

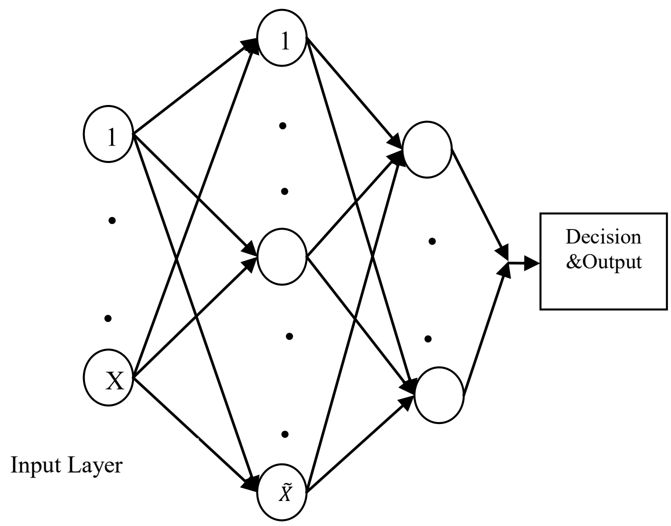


**Figure 3.** Basic ELM model

## 4. Proposed Work

To reduce the issue of semantic gap and to increase the accuracy a smart hybrid CBIR system is proposed in this paper which is based on the hybridization of low level features and ELM classifier. The leading feature extraction techniques of color, shape and texture are used to extract the visual features from all the images and are merged together to form a single efficient feature vector. The database images are trained by using ELM classifier as it has an admirable property of multi-class classification due to its superior scalability and least computational complexity. After the classification process the similarity between the query and database images is calculated by Euclidean distance and top ten images are retrieved. The block diagram of the complete proposed work including training and

testing is shown in Figure 4a and 4b. In the training phase the ELM is trained by database images and a classifier model is developed.
In the testing phase the query image is searched out from the most suitable class out of all the classes formed by ELM model and images are retrieved.

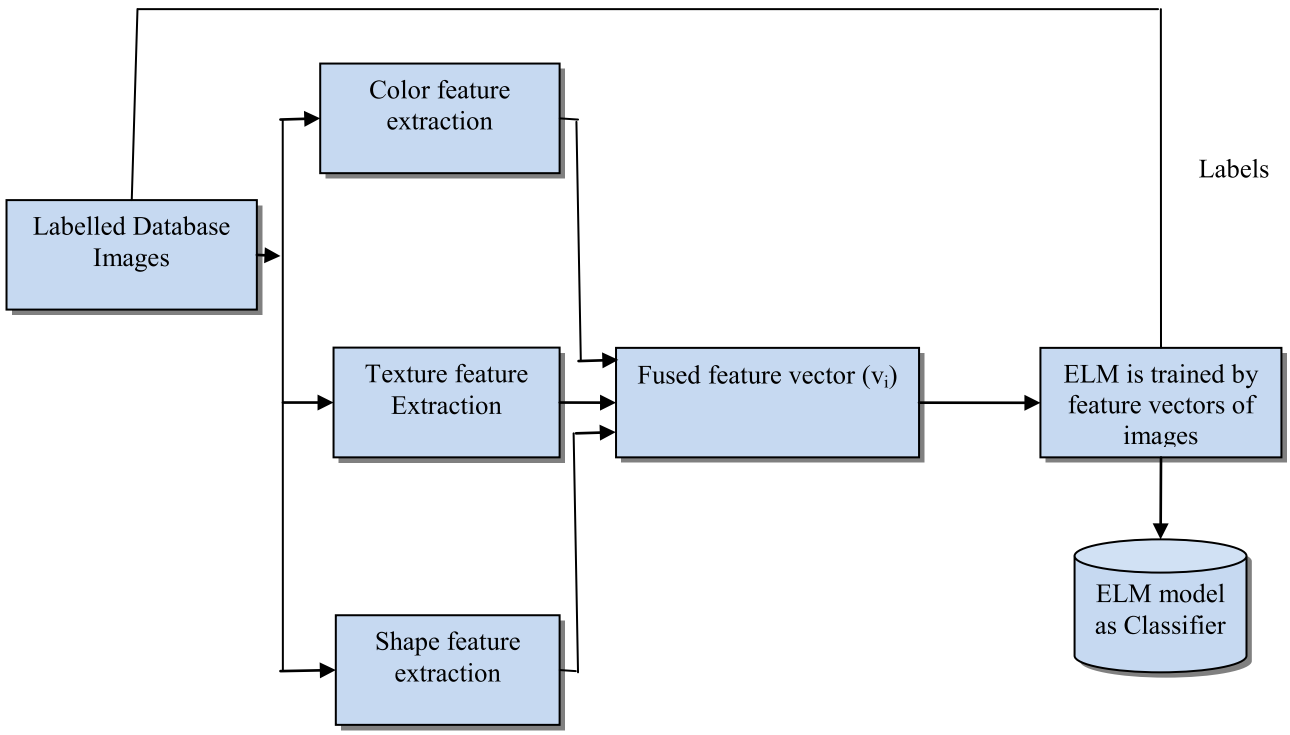


**Figure 4 a.** Training phase of the proposed system

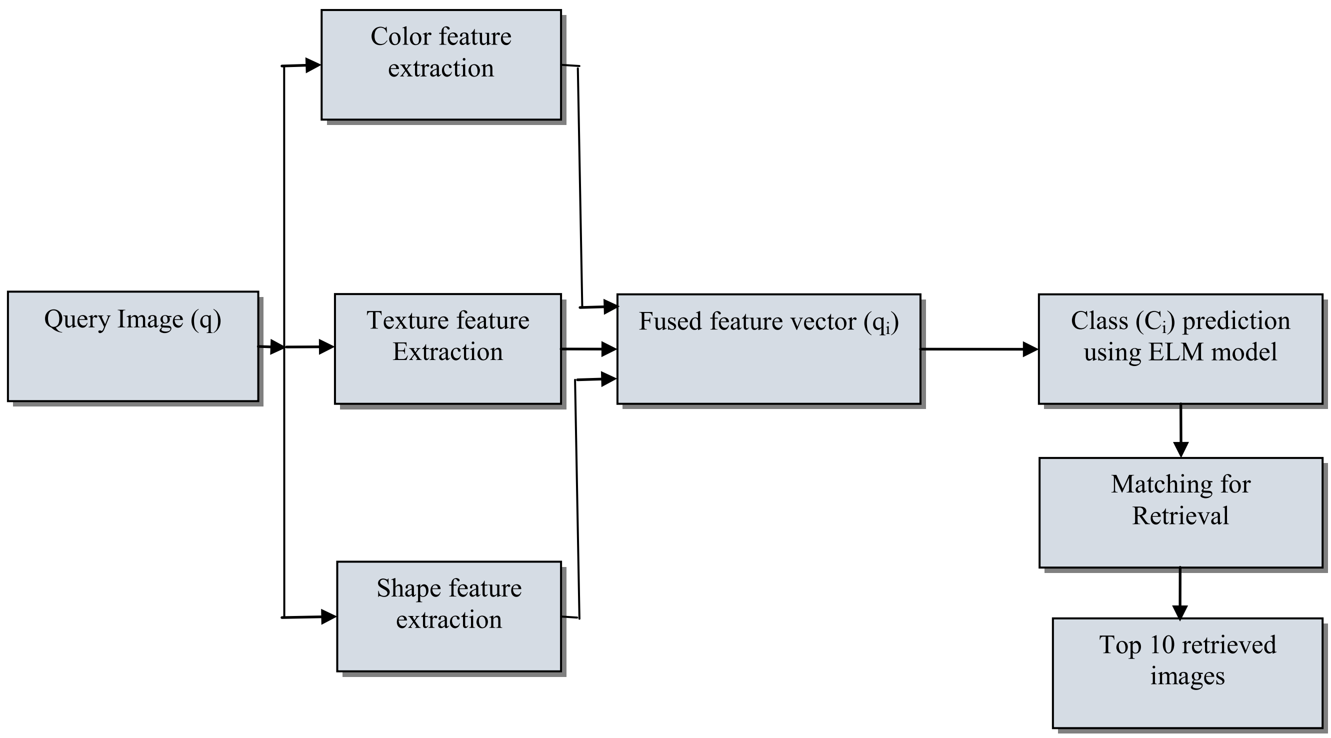


**Figure 4 b**. Testing phase of the proposed system

The algorithm of the designed framework is given below.

**Algorithm 1**
Input: Databases with labelled images.
Step 1: Evaluate the visual features $v_i$ of database images by the combination of color, shape and texture where (i=1, 2, 3) and stored as a feature vector.
Step 2: Select the query image q.
Step 3: Compute the similar visual feature vector $q_i$.
Step 4: Train the ELM with features of the images by one to one method and multiple classes are formed by ELM classifier.
Step 5: Classification of the query image is done into the most pertinent class say $C_i$ of the database.
Step 6: Similarity between the query image q and images in class $C_i$ is computed by using Euclidean distance.
Step 7: Sorting of images in class is done depending on the distance and top 10 images are retrieved.

## 5. Experimental Set up & Results

### 5.1 Experimental Set up

Experiments for the proposed system have been conducted on four benchmark datasets which are Corel-1K, 5K, 10K and GHIM-10K. The details of these datasets are given in Table 1. The implementations are done on MATLAB-15A with 64 bit windows and 2GB memory.

Table 1. Details of datasets

| Dataset | Categories | Images per category | Total Images |
|---|---|---|---|
| GHIM-10K | 20 | 500 | 10,000 |
| Corel-5K | 50 | 100 | 5000 |
| Corel-10K | 100 | 100 | 10,000 |
| Corel-1K | 10 | 100 | 1000 |

Description of Corel-1Kdataset:

The Corel-1K dataset contains of 1000 images. The images are categorized into various groups such as buses, dinosaurs, flowers, beaches etc. The image size is 256 × 384 otherwise 384 × 256. The sample images per category are shown in Figure 5. (http://wang.ist.psu.edu/docs/related/).

**Figure 5.** Sample images of Corel-1k database

Description of Corel-5K dataset:
The Corel-5K datasets contains around 5000 images. These images are categorized into fifty groups such as animals, cars, ships, butterflies etc.The image size is 128 × 187 or 187 × 128 and is presented in Figure 6. (http://www.ci.gxnu.edu.cn/cbir/).

**Figure 6.** Sample Image of Corel-5k database

Description of Corel-10K dataset:
The Corel-10K datasets contains 10, 000 images. The groups are fruits, birds, globes, coins and many more some of which are displayed in Figure 7. The image size is 128 × 187 or 187 × 128(http://www.ci.gxnu.edu.cn/cbir/).

**Figure 7.** Sample Images of Corel-10k database

Description of GHIM-10K dataset:
The GHIM-10k dataset contains of 10,000 images. The images such as airplanes, insects bikes, cars etc are categorized into twenty groups. The image size is 300× 400 or 400 × 300 which are in JPEG format. Sample images are shown in Figure 8. (http://www.ci.gxnu.edu.cn/cbir/).

**Figure 8.** Sample Image of GHIM-10k database

## 5.2 Experimental Results

In order to check out the competence of the proposed framework the most important parameters of the CBIR systems which are precision, recall and accuracy are evaluated for different benchmark datasets which are described in the above section.

$$\text{Precision} = \frac{\text{No. of relevant images Retrieved}}{\text{Total No. of images Retrieved}}$$

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{No. of relevant images in database}}$$

Firstly the experiment is conducted on Corel-1K database which consists of 10 different categories for evaluating the performance parameters of the system. Random twenty images are selected from each category as query images and the average value of precision and recall is calculated. The average value of precision of each category on the basis of retrieved images is displayed in Figure 9. For the similarity calculation between the query and the classified images euclidean distance is applied which is given in Equation 14.

$$D_E = \sqrt{\sum_{i=1}^{n} (|q_i - C_i|)^2} \qquad (14)$$

Where $q_i$ and Ci are the feature vectors of query and classified images respectively.

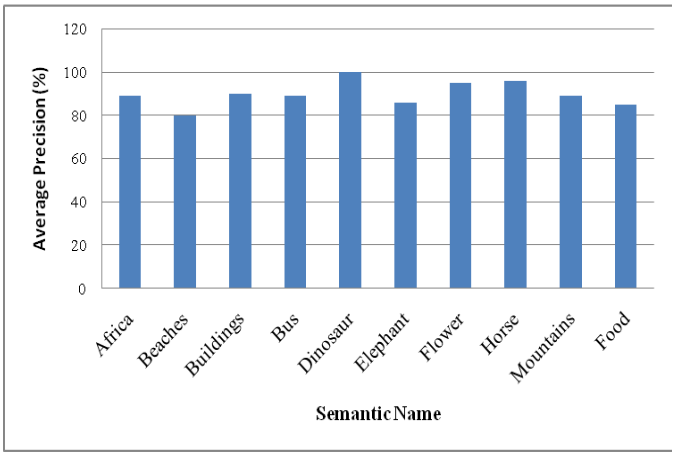


**Figure 9.** Average value of precision of 10 categories of Corel-1K database

Similarly the average value of recall is evaluated of all the categories by the above given equation. If the obtained resuls are relevant and appropriate then the recall values increases with increase in precision which is shown in Figure 10.

The average value of precion along with the average value of recall is used for the assesment of every CBIR system. The precision versus recall graph of Corel-1K database is shown in Figure 10.

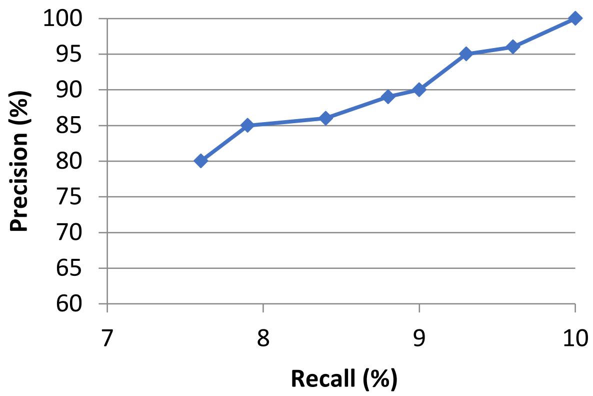


**Figure 10.** Precision versus Recall graph of Corel-1K database

Similarly, the experiments with the same procedure are carried out on other large datasets such as Corel-5K, Corel-10K and GHIM-10K and the average value of precision is estimated by taking the random 20 images from each category with the retrieval of top 10 images.

When the three low level features are combined and ELM is used as a classifier the results obtained are much superior in terms of accuracy and precision. The average precision of Corel-1K, is 90% and for GHIM-10k, Corel-10k and Corel-5K are 88.7%, 74% and 79% respectively.

### 5.2.1 Performance Comparison of the proposed method and major gaps reported in the literature

The most relevant parameter to measure the efficiency of CBIR systems is precision. It provides the relevant image ratio for the retrieved resultant images. In order to authenticate the novelty in terms of precision of our proposed work, its comparison with other latest state-of-art methods is tabulated in Table 2 for Corel datasets and in Table 3 of GHIM-10K dataset . The average value of precision of all the datasets with other authors are graphically correlated in Figure11.

The most critical issue in the literature is the proper selection of feature extraction techniques, semantic gap and the complete response time of CBIR system. The proposed CBIR system swamped two of these obstacles. The designed framework has been compared with other latest methods.
The first method which is compared is the Guo et al. [16]. It uses the Ordered Dither Block Truncation Coding to

generate the features of the images. For capturing the contrast and color distribution of the images color co-occurrence feature is used and for detecting the image edges bit pattern feature is used. PSO is used to calculate the similarity constants between the images. The experiments are conducted on Corel datasets. Although, the system is less complex and average precision values are acceptable for small datasets but for large datasets the precision values are very less. Moreover, the retrieval time is also high. The next method that is used for the comparison is Pavithra et al. [25]. In this framework, firstly the color feature is extracted from the whole dataset and comparable images are selected as per the query image. After this step texture and edge features are extracted using LBP and canny edge detector respectively. This system increases the accuracy of the hybrid CBIR systems and the retrieval time is also reduced as the searching space of the images is reduced but nothing is done here to reduce the problem of semantic gap. The performance can be increased and can be applied on different types of databases when the system is integrated with some machine learning algorithm.

The third paper that is correlated here is of Fadaei et al. who proposed the CBIR scheme on the basis of optimized combination of texture and color features to improve the accuracy and time of the image retrieval. Dominant Color Descriptor (DCD) features were extracted from HSV color space and to extract texture features wavelet and curvelet were applied and finally these two features are combined optimally by optimization algorithm which is particle swarm optimization algorithm [11]. Similarly as the above case the precision is better for smaller datasets only but for large datasets such as Corel-5K and Corel-10K the average precision is approximately only 59 and 50% only.

The last paper which is compared with the proposed system is of Walia et al. [33]. In this method color difference histogram and angular radial transform techniques are used for extracting the low level features. In spite of the fact it provides very less retrieval time as compared with other methods but at the expense of increased complexity. Further in this system, first 30 images are shortlisted for texture and shape detection and hence it is not supported for larger datasets.

To overcome all these limitations the proposed system here combines all the three low level features through normalization by using suitable techniques. After this the machine learning algorithm is applied on this hybrid CBIR system which reduces the semantic gap and increases the precision of the system even for larger datasets also. Such as in case of Corel-1K dataset precision is 90% and accuracy is 96% which is higher than other proposed techniques.

The graphical comparison of average precision of authors Guo et al. , Pavithra et al., Fadaei et al., and Walia et al., with the proposed framework is shown in Figure 11 for Corel datasets.

Table 2. Comparison of average precision of proposed method with other methods for Corel datasets

| Corel-1K | Guo [18] | Pavithra [32] | Fadaei [12] | Walia [40] | **Proposed** |
|---|---|---|---|---|---|
| Africa | 82.5 | 81 | 72.4 | 73 | **89** |
| Beaches | 46.6 | 66 | 51.15 | 39.25 | **80** |
| Buildings | 64.8 | 78.75 | 59.55 | 46.25 | **90** |
| Bus | 89.4 | 96.25 | 92.35 | 82.5 | **89** |
| Dinosaur | 99.3 | 100 | 99.9 | 98 | **100** |
| Elephant | 72.2 | 70.75 | 72.7 | 59.25 | **86** |
| Flower | 94.9 | 95.75 | 92.25 | 86 | **95** |
| Horse | 89.2 | 98.75 | 96.6 | 89.75 | **96** |
| Mountains | 42.7 | 67.75 | 55.75 | 41.75 | **89** |
| Food | 76.7 | 77.25 | 72.35 | 53.45 | **85** |
| Average | 76.5 | 83.225 | 76.5 | 66.92 | **90** |
| Corel-5K(AP) | 64.5 | 68.6 | 58.97 | 56.72 | **78** |
| Corel-10K(AP) | 52 | 59.98 | 49.5 | 40.78 | **72** |

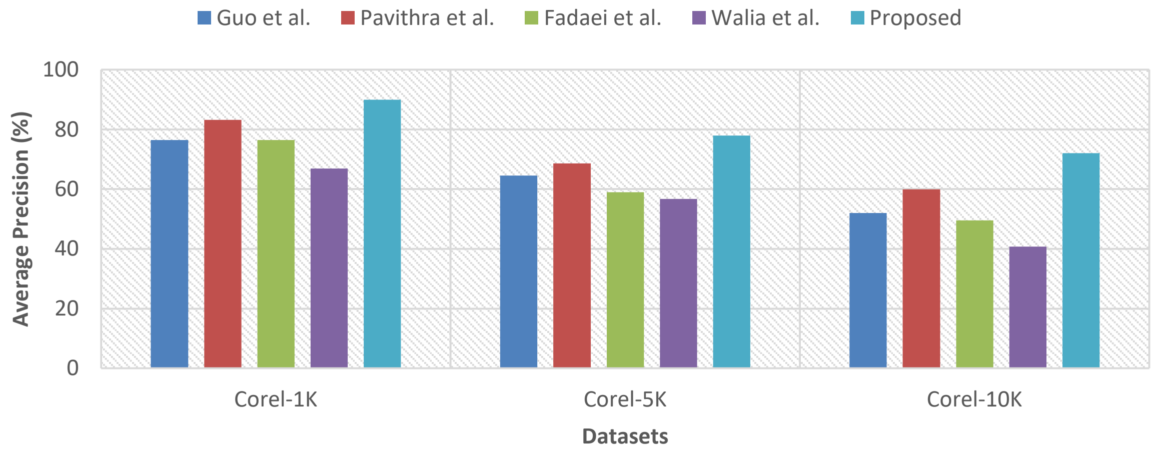


**Figure 11**. Performance Comparison of the designed system on different datasets with other techniques

Table 3. Comparison of Average precision of proposed method with other methods for GHIM-10 K database

| Dataset | Zhu et al. [41] | Srivastava et al.[16] | Liu Y et al.[6] | **Proposed** |
|---|---|---|---|---|
| GHIM-10K | 43.89 | 76.99 | 52.02 | **88.70** |

### 5.2.2 Accuracy of datasets

The confusion matrix is generated of each database by using ELM classifier and from this the accuracy of the systems is calculated. The confusion matrix which is also known as an error matrix is used in the machine learning algorithms to measure the performance of the algorithms as a classifiers. In this matrix every row is used to represent the instances of the predicted class and every column represents the instances of the actual class. Many performance parameters can be evauated using the confusion matrix such as error rate, specificity, sensitivity and especially the accuracy of the system. It is calculated as the ratio of number of correct predictions to the number in total dataset. As for example the confusion matrix of Corel-1K dataset is shown in Figure 12. Similarly,the accuracy of all the databases Corel-5K, Corel-10K and GHIM-10K is presented in Figure 13 which is much better than other hybrid systems.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 92 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 0 | 96 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 1 |
| 1 | 1 | 95 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 98 | 0 | 1 | 1 | 0 | 0 | 0 |
| 0 | 1 | 1 | 0 | 96 | 0 | 0 | 1 | 1 | 0 |
| 0 | 0 | 0 | 1 | 0 | 96 | 2 | 0 | 0 | 1 |
| 1 | 0 | 0 | 0 | 0 | 1 | 98 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 96 | 1 | 2 |
| 1 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 95 | 0 |
| 0 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 96 |

**Figure 12.** Confusion Matrix obtained for Corel-1K dataset by ELM

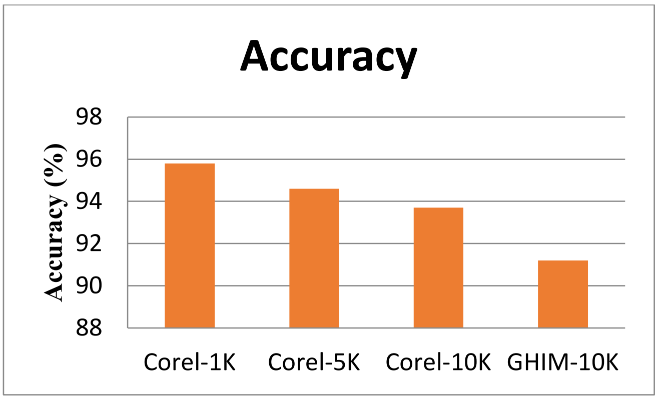


**Figure 13.**.Accuracy of all datasets

The above graph illustrates the comparison of average accuracy value in percentage among experimented datasets. The results of accuracy comparison are having smaller variations between the datasets. The superior value belongs to the Corel- 1K database which is 95.8%. The GHIM-10K has the accuracy of 91.2% and Corel-10K, 5K have 93.7% and 94.6% respectively.

### 5.2.3 Image retrieval results

The retrieval results of the designed framework for different datasets are shown in Figures by taking the random image from any category. In Figure 14 top 10 images are retrieved by taking the query image from flower category from Corel-1K dataset. In Figure 15, the query image is taken from the car category through GHIM-10k dataset and the top 10 images are retrieved by the designed framework. Similarly is the case of Figure 16 the relevant images are retrieved by taking the query image from dog category.

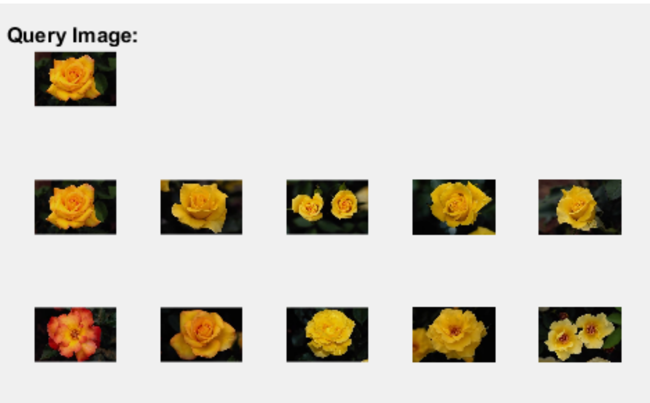


**Figure 14.** Retrieved top 10 images by proposed method from Corel-1K dataset

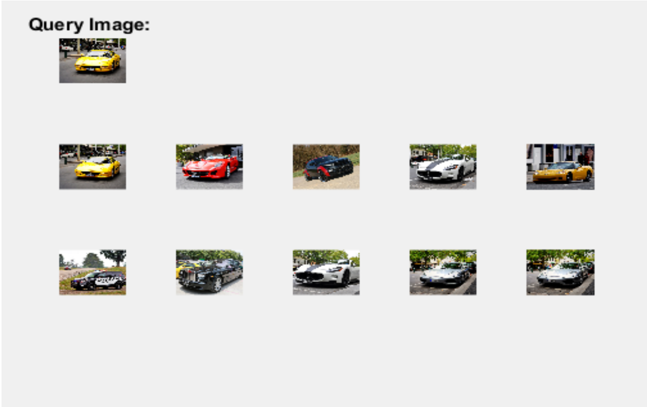


**Figure 15.** Top 10 images retrieved by the proposed method from GHIM-10K dataset

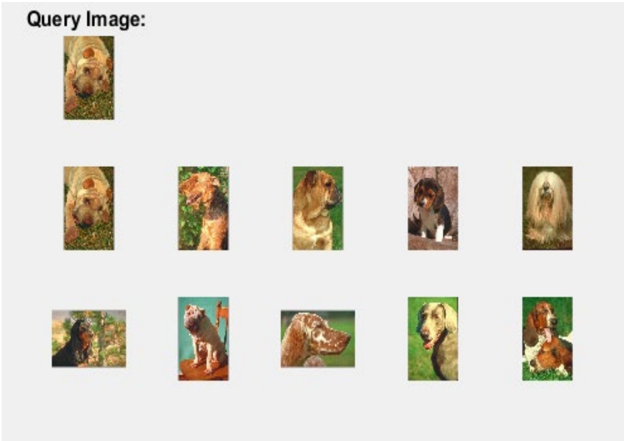


**Figure 16.** Top 10 images retrieved by the proposed method from Corel 5K dataset

## 6. Conclusion

An innovative hybrid framework is proposed here for the retrieval of comparable images from big databases by using shape, texture and color features. To run-over the problem of semantic gap and to increase the accuracy of the system, machine learning algorithm is applied which functions as a classifier. The feature extraction techniques and ELM plays a leading role in this designed system. Color moment, LBP and parameters of the images are used for extracting the color, texture and shape respectively and after this these features are combined for designing a hybrid image retrieval system. The ELM classifier reduces the search space by making the different classes and helps in the suitable selection of the images even from the big databases also. The deigned system has been implemented on Corel-1K, 5K, 10K and GHIM-10K datasets and their precision rate along with accuracy is compared with other state of art methods which spectacles that it is much more competent and effective than other systems.

## References


[1] Hemachandran K, Paul A, Singha M. (2012) Content-based image retrieval using the combination of the fast wavelet transformation and the colour histogram. *IET Image Processing* 6(9):1221–1226.

[2] Gandhani S, Singhal N.(2015) Content Based Image Retrieval : Survey and Comparison of CBIR System based on Combined Features. *International Journal of Signal Processing, Image Processing & Pattern Recognition* 8(11):417–422.

[3] Lande M V, Bhanodiya P, Jain P.(2014) An Effective Content Based Image Retrieval using Color, texture and Shape. *Journal of Intelligent Computing, Networking, and Informatics* 243:1163–1170.

[4] Yadav AK, Roy R, Kumar AP.(2015) Survey on Content-based Image Retrieval and Texture Analysis with Applications. *International Journal of Signal Processing, Image Processing & Pattern Recognition* 7(6);41-50

[5] Ke CJS. (2017) Content-Based Image Retrieval Based on Shape Similarity Calculation. *3D Resolution.* DOI 10.1007/s13319-017-0132-0; 1-19

[6] Yu L, Feng L, Wang H, Li L, Liu Y, Liu S. (2017)Multi-trend binary code descriptor: a novel local texture feature descriptor for image retrieval. *International Journal of Signal, Image Video Processing.* DOI 10.1007/s11760-017-1152-1;1–8.

[7] Zawbaa, Hossam M., E. Emary, and Crina Grosan. (2016) Feature selection via chaotic antlion optimization. *PloS* one 11.3 e0150652.

[8] Ding S, Xu X. (2014) Extreme learning machine: algorithm, theory and applications. *Neural Computing & Applications;* 1-11

[9] Verma M, Raman B. (2016) Local tri-directional patterns: A new texture feature descriptor for image retrieval. *Digital Signal Process A Rev Journal* 51:62–72

[10] Maheshwari SMRP, Balasubramanian R. (2012) Directional local extrema patterns : a new descriptor for content based image retrieval. *International Journal of Multimedia Information Retrieval*;191–203.

[11] Mohamadzadeh S, Farsi H. (2013) Colour and texture feature-based image retrieval by using Hadamard matrix in discrete wavelet transform. *IET Image Processing* 7(3);212–218.

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

[13] Sayed, Gehad Ismail, Aboul Ella Hassanien, Ahmad Taher Azar.(2019) Feature selection via a novel chaotic crow search algorithm. *Neural Computing and Applications* 31(1); 171-188.

[14] Alsmadi MK. (2017) Query-sensitive similarity measure for content-based image retrieval using meta-heuristic algorithm. *Journal of King Saud University - Computer & Information Sciences*; 1-9

[15] Guo J, Member S, Prasetyo H. (2014) Content - Based Image Retrieval Using Features Extracted From Halftoning - Based Block Truncation Coding. *IEEE Transactions on Image Processing;*7149(c).

[16] Srivastava P, Khare A. (2017) Utilizing multiscale local binary pattern for content-based image retrieval. *International Journal of Multimedia Tools & Applications* DOI 10.1007/s11042-017-4894-4; 1-27

[17] Jain N, Salankar SS. (2014) Color & Texture Feature Extraction for Content Based Image Retrieval.*IOSR Journal of Electrical & Electronic Engineering* ;53–58.

[18] R JVCI, Guo J, Prasetyo H, Su H.(2013) Image indexing using the color and bit pattern feature fusion. *Journal of Visual Communication Image Represent*ation 24(8) ;1360–79

[19] Lu B, Duan X, Wang C. (2014)A Novel Approach for Image Classification Based on Extreme Learning Machine *in proceedings of IEEE Conference* ; 381-384

[20] Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS. (2016) Deep learning for visual understanding : A review. *Neurocomputing* 187;27–48.

[21] Nagathan A. (2014) Content-Based Image Retrieval System using Feed-Forward Backpropagation Neural Network.*International Journal of Computer Science & Network Security* 14(6);70–77.

[22] Emary, Eid, Hossam M. Zawbaa Aboul Ella Hassanien.(2016) Binary grey wolf optimization approaches for feature selection. *Neurocomputing* 172 ; 371-38*1*

[23] Ali G, Giveki D. (2015) Content based image retrieval system using clustered scale invariant feature transforms. *Optik- International Journal of Light Electron* 126(18);1695–1709

[24] Vogel J, Schiele B. (2006) Performance evaluation and optimization for content-based image retrieval. *Pattern Recognition* 39(5); 897–909.

[25] Mathew SP, Balas VE, Zachariah KP. (2015) A Content-based Image Retrieval System Based On Convex Hull Geometry Background of the Research Work. *Acta Polytechnica Hungarica*12(1);103–116.

[26] Ibrahim Dogan (2016) An Overview of Softcomputing. 12th *International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS,Procedia Computer Science;* 34-38

[27] Philip O.Omolaye, Joseph M. Mom, Gabriel A. Igwue (2017) A Holistic Review of Soft Computing Techniques. *Journal of Applied and Computational Mathematics*6(2); 93-110

[28] V. Yadaiah , R .Vivekanandam, .Rajaram Jatothu (2018) A Fuzzy Logic Based Soft Computing Approach in CBIR System Using Incremental Filtering Feature Selection to Identify Patterns. *International Journal of Applied Engineering Research* 13(5)2432-2442.

[29] Manoharan Subramanian (2012) A Comparison and Analysis of Soft Computing Techniques for Content based Image Retrieval System *International Journal of Computer Applications.*59(13); 13-18

[30] Srivastava D, Wadhvani R, Gyanchandani M. (2015 ) A Review : Color Feature Extraction Methods for Content Based Image Retrieval.*International Journal of Computational Engineering & Management* 18(3);9–13.

[31] Shahbahrami A. (2008 ) Comparison Between Color and Texture Features for Image Retrieval. *6th International Symposium on Applied Machine Intelligence and Informatics* 27648. ; 221-224

[32] Pavithra LK, Sharmila TS. (2017) An efficient framework for image retrieval using color, textureand

edge features. *Journal of Computers & Electrical Engineering* 0;1–14
[33] Liao S, Law MWK, Chung ACS.(2009) Dominant local binary patterns for texture classification. *IEEE Transactions on Image Processing* 8(5);1107–1118.
[34] Huang D, Member S, Shan C, Ardabilian M.(2014) Local Binary Patterns and Its Application to Facial Image Analysis : A Survey. *IEEE Transactions on System Man & Cybernatics* 41(6); 765-780
[35] Chaudhari R, Patil AM.(2012) Content Based Image Retrieval Using Color and Shape Features. *International Journal of Advances in Research Electrical Electronics Instrumentation Engineering* 1(5); 386–92.
[36] Shirazi SH, Razzak MI. (2016) Content-Based Image Retrieval Using Texture Color Shape and Region .*International Journal od Advanced Computer Science and Applications*. 7(1); 418–26.
[37] Puviarasan N, Bhavani R, Vasanthi A.(2016) Image Retrieval Using Combination of Texture and Shape Features. *in proceedings of IEE International Conference on Next Generation Intelligent Systems* 3(3); 5873–5877
[38] Annrose J, Christopher CS. (2018) An efficient image retrieval system with structured query based feature selection and filtering initial level relevant images using range query; *Optik;*1053–64.
[39] Huang G, Member S, Zhou H, Ding X, Zhang R.(2012) Extreme Learning Machine for Regression and Multiclass Classification. *IEEE Transactions on Systems, man & Cybernatics* 42(2);513–529.
[40] Walia E, Vesal S, Pal A. (2014) An effective and fast hybrid framework for color image retrieval. *Journal of sensing and Imaging* 15(93); 1–23
[41] Zhu C, Bichot C, Chen L (2010) Multi-scale Color local binary pattern for visual object class recognition. *In proceedings of International Conference of Pattern Recognition;* 3065-3068.