

A Genesis of a Meticulous Fusion based Color Descriptor to Analyze the Supremacy between Machine Learning and Deep learning

Shikha Bhardwaj*

Department of Electronics & Communication Engineering, DCRUST, Murthal, Sonapat, India & Department of Electronics & Communication Engineering, UIET, Kurukshetra University, Kurukshetra, India
E-mail: shikpank@yahoo.com

Gitanjali Pandove

Department of Electronics & Communication Engineering, DCRUST, Murthal, Sonapat, India
E-mail: pandove_geet@yahoo.co.in

Pawan Kumar Dahiya

Department of Electronics & Communication Engineering, DCRUST, Murthal, Sonapat, India
E-mail: pawan.dahiya@gmail.com

Received: 17 June 2019; Revised: 05 August 2019; Accepted: 07 September 2019; Published: 08 April 2020

Abstract—The tremendous advancements in digital technology pertaining to diverse application areas like medical diagnostics, crime detection, defense etc., has led to an exceptional increase in the multimedia image content. This bears an acute requirement of an effectual retrieval system to cope up with the human demands. Therefore, Content-based image retrieval (CBIR) is among the renowned retrieval systems which uses color, texture, shape, edge and other spatial information to extract the basic image features. This paper proposes an efficient and unexcelled hybrid color descriptor which is an amalgamation of color histogram, color moment and color auto-correlogram. In order to determine the predominance between machine learning and deep learning, two machine learning models, Support vector machine (SVM) and Extreme learning machine (ELM) have been tested. Whereas from deep learning category, Cascade forward back propagation neural network (CFBPNN) and Patternnet have been utilized. Finally, from these divergent tested algorithms, CFBPNN attains the highest accuracy and has been selected to enhance the retrieval accuracy of the proposed system. Numerous standard benchmark datasets namely Corel-1K, Corel-5K, Corel-10K, Oxford flower, Coil-100 and Zurich buildings have been tested here and average precision of 97.1%, 90.3%, 87.9%, 98.4%, 98.9% and 82.7% is obtained respectively which is significantly higher than many state-of-the-art related techniques.

Index Terms—Color moment, Color histogram, Color correlogram, Support vector machine, Extreme learning machine, Cascade forward back propagation neural network, Patternnet neural network.

I. INTRODUCTION

Digital imaging has now a days evolved into a paramount section in assorted areas of image evaluation and processing [1]. Due to this ever increasing digital image content, the retrieval or recovery of desired images becomes too strenuous for users. For the ease of users, there is a cardinal necessity of a technology for the effective retrieval of images. And the solution to this dilemma is known by a technique known as Content Based Image Retrieval (CBIR) system. It extracts basic visual attributes like texture, shape, color etc., both from the query image and total database images.

Images can be recovered in two basic ways: (1) Text-Based Image Retrieval (TBIR) (2) Content-Based Image Retrieval (CBIR). TBIR is the traditional method of retrieving images and is based on text or keyword manual annotation of images. But, it suffers from vast errors due to use of synonyms, homonyms, misspellings etc. [2]. In contrast to TBIR, Content-Based is an efficient technique of retrieving images from a huge repository of collected images. This system eliminates majority of the pitfalls present in TBIR system. In this system, an image recovery is based on the presence of various basic image features.

A basic CBIR system consists of a query image given by the user and based on this query image, different image features are extracted and a feature vector is created. The same procedure is applicable to the utilized image dataset and again its feature vectors are formed. These calculated feature vectors relating to both the query image and the complete dataset are matched in the

similarity matching stage of the system by using a specific distance metric, resulting in visually similar images. Both local and global type of features can be withdrawn from an image. Global feature extraction is concentrated on the features of the complete image while a local extraction is focused on a specific part of an image.

Among all the basic image features, color is the utmost indispensable attribute that can be easily perceived by human eye. Each color and every object associated with color has a specific wavelength related to it, which gets reflected and forms an image in human retina. For color feature extraction, there are numerous techniques [3] which can be applied to CBIR system. But, the challenge here is to find the optimal technique for feature recovery and also the best amalgamation of retrieval methods. Numerous techniques can be used for color feature extraction like Dominant color descriptor (DCD) [4], Color autocorrelogram, Color coherence vector, Distance coherence vector, color coding, etc. Texture is another important attribute of an image which defines the inborn appearance patterns of an image. A large quantity of techniques like Local binary pattern (LBP) [5], Discrete Wavelet Transform (DWT), Curvelet Transform, Gabor Transform, Gray level Co-occurrence Matrix (GLCM), Tamura features etc. have been used for texture extraction. The third prominent feature of an image is shape which depicts the semantic information and further it can be disintegrated into two parts: Region based and Boundary based. Curvature scale space (CSS), Hough transform, Fourier Transform, B-splines, Zernike Moments etc. are few of the shape descriptors used in CBIR system [1]. There are many edge detection methods like Robert, Sobel, Canny etc. which can be used for the extraction of edge information from an image. In order to acquire the spatial information of an image, different techniques like Regions of interest (ROIs), graph/tree based methods etc. can be deployed.

But, these low level feature extraction techniques lack the high-level human intelligence and therefore a wide gap called as Semantic gap exists between the low level feature extraction techniques and lofty human perception. To reduce this gap, prominent machine learning algorithms like SVM, Naïve Bayes, Random forest, K-means, K-NN etc. have been used. But, the accuracy obtained by using these models is still inadequate. To further enhance the performance of retrieval system, the interest of the researcher community has been shifted from machine learning to deep learning. Although, deep learning is a subset of Machine learning, but it has superior results as compared to many machine learning models. Different Deep learning techniques like Deep Boltzman machines, Convolutional neural networks (CNN), Bag of Visual words (BoVW) etc. have been used for an accurate retrieval. Therefore, an experimental analysis is performed here on some of the machine learning and deep learning models to find the best algorithm, which could be successfully used to reduce the semantic gap present in an image retrieval system.

The main contributions of this paper are as follows:

- To form an efficient hybrid color descriptor which extracts the maximum color information from an image.
- An experimental based comparison between various machine learning and deep learning models has been done in order to find the best model.
- Lastly, this best model has been selected to be used with the proposed hybrid color descriptor in order to enhance the retrieval accuracy of the system.
- Finally, the proposed system has been tested on six benchmark CBIR datasets with varied similarity measures.

The remaining section of this paper is organized as follows: Related state-of-the-art work of different hybrid CBIR systems based on various machine and deep learning techniques is given in Section II. Preliminaries in the form of utilized color descriptors with other techniques are given in section III. Section IV gives description of the proposed methodology and experimental setup and results are given in section V. Conclusion with Future trends are given in Section VI.

II. STATE -OF-THE-ART RELATED WORK

This section dispenses the state-of-the-art procedures for image retrieval in CBIR systems. A hybrid CBIR system which utilizes Dominant color descriptor (DCD) for color feature retrieval and combination of curvelet and wavelet features for texture extraction has been proposed by Fadaei et al. [4]. Color and edge directivity descriptor (CEDD) [6] is another technique which is utilized for the extraction of both color and texture features where as Discrete wavelet transform (DWT) is employed for the analysis of shape features in another image retrieval system.

Various spatial domain and frequency domain features have been also used for the formation of a hybrid retrieval system [7]. Pavitra et al. [8] develops a fusion based system, where Color moment is used as a filter for the selection of limited images based on the value of a moment. Later, Local binary pattern (LBP) and Canny edge detector has been used to extract texture and edge information from an image.

Tetrolet transform [9] has been used by Pradhan et al. for the extraction of textual content of an image Edge joint histogram and color channel correlation histogram have been used respectively to extract the shape and color features of an image. This system is realized in the form of a three-level hierarchical system. Sometimes, an image is needed to be pre-processed before the extraction of its features. Segmentation and the removal of noise are the basic pre-processing steps in any image retrieval system [10]. After these steps, color and texture features have been extracted from the given dataset of images. Another hybrid system where both image enhancement and

segmentation has been done initially on the image dataset has been proposed by Huthaifa et al. [11]. Afterwards, color, texture and shape features have been extracted while Artificial neural networks (ANN) have been used for classification purpose.

Image segmentation is a process which is based on obtaining information from a specific part of an image and therefore be called as Region based image retrieval (RBIR). Edge-integrated minimum spanning tree (EIMST) is an algorithm which can be used for RBIR and has been developed by Yang liu [12]. Various global descriptors like color layout descriptor (CLD), Edge histogram descriptor (EHD), Color and edge directivity descriptor (CEDD), etc have been analyzed by lakovidou et al. [13]. These descriptors have been tested to analyze their performance in CBIR systems.

Clustering is another approach which can be used to separate the given dataset into specified parts and the efficiency of the CBIR system can be enhanced. Agglomerative hierarchical clustering partitions the database into clusters which can be visually separated [14]. Many more clustering approaches like K-means, Density based, etc. have also been used in image retrieval tasks. Indexing or Dimensionality reduction has also been considered as an important technique in CBIR systems which aims to maintain proper index of every image present in the database and hence reduces the computation time in similarity measurement. Such an indexing based hybrid system has been proposed by Arousi et. al [15] where 2D histogram with statistical moments has been used for color feature extraction and Gray level co-occurrence matrix (GLCM) for texture analysis.

Speeded up robust features (SURF), Maximally stable external regions (MSER), color correlograms and Improved color coherence vector (ICCV) have been used to create a multidimensional feature matrix [16]. In order to quantize the obtained feature vector, Bag of visual words (BoVW) has also been used. To capture the human visual perception, Perceptual uniform descriptor (PUD) has been proposed by Shenglan Liu et al. [17]. This descriptor is based on a combination of gradient direction features with color attributes of an image.

In order to increase the performance of CBIR systems, many machine learning algorithms have been developed. A combination of neural network (NN), K Nearest neighbors (KNN) and Support vector machine (SVM) have been used to develop an image retrieval system [18]. The performance of many classifiers like Naïve Bayes, Decision tree, Random forest, Logistic regression, SVM have been compared for an analysis based on text reviews [19]. Now, in this era, the focus of the researcher section has been shifted from machine learning to deep learning. Although, this deep learning is itself a subset of machine learning but it is based on the conception of human neurological nervous system. Therefore, more human perception and intelligence can be added to a system.

Hybrid deep learning architecture (HDLA), [20], is a technique based on deep learning which can be used for an effective image retrieval. This model uses Boltzmann

machines in the upper layers and Softmax model in the lower levels. Another technique based on Deep Convolutional Neural Network (DCNN) [21] has been developed which extracts the 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.

Different types of deep learning techniques like Deep neural networks, Restricted Boltzmann machines, Deep Belief networks (DBN), Auto encoders etc. can be used for the formation of a CBIR system [22]. A description regarding the prominent machine learning techniques is given in Table 1.

These machine learning models can be categorized under Supervised, Unsupervised and Reinforcement learning. In case of deep learning, the famous network models are given in Table 2.

These discussed hybrid systems in literature suffers from some drawbacks and there is no provision to compare the working performance of machine learning and deep learning algorithms. The proposed system is a hybrid combination of three color feature descriptors which extracts the maximal color information from an image. Moreover, a testing based comparison is also being done between some of the techniques of machine learning and deep learning to find the dominant out of the two.

Table 1. Some of the prominent machine learning techniques

Technique used	Brief description
Decision Trees	Uses a tree-like graph for various decisions and based on varied - consequences
Linear regression	To plot a relationship between a dependent and an independent variable
Support vector machine (SVM)	Find the optimal hyper-plane which divides the given points into two parts with maximal distance
Naïve Bayes	The presence of any specific attribute is unrelated to any other feature is assumed.
K-Nearest neighbor (KNN)	It works by taking the majority opinion of its K neighbors.
Ensemble methods	Works by the formation of many classifiers and then classify the data according to the obtained predictions.
K-means	Data is partitioned into different clusters with homogenous data in an individual cluster
Random forest	Each tree is classified and then each tree votes for a specific class in the classification procedure.
Principal component analysis (PCA)	Uses orthogonal transformation to convert observations into linearly uncorrelated variables.
Singular value decomposition (SVD)	It is a factorization of real complex matrix
Independent component analysis (ICA)	It is a generative model for the given multivariate data

Table 2. Some of the prominent methods of Deep learning

Technique used	Brief description
Convolutional neural network (CNN)	Consists of convolution, pooling and various other layers for performing tasks
Recurrent neural network (RNN)	Uses the information from previous layers to predict its current outcome
Deep Reinforcement learning	It is based on the concept of an agent communicating with an environment
Fully connected neural networks	Here, every neuron is connected to the other neuron of the preceding and succeeding layer
Boltzman machine	These structures are based on the working procedure of RNN
Deep belief networks (DBN)	These have connections among layers but not between different units present in those layers
Autoencoders	These are specifically designed to learn the encoding of given data with the main aim of reduction in dimension of data
Bag of visual words (BoVW)	It is vector of occurrence counts of words and can be used for different applications

III. PRELIMINARIES

In CBIR systems, color is considered as the prominent feature of an image which can be effectively used for the retrieval of images. It is the most basic feature of visual content. Therefore, in this paper three eminent color descriptors namely color histogram, color moment and color auto-correlogram have been fused together for the formation of a hybrid color based descriptor. Moreover, to justify the supremacy between machine learning and deep learning, varied algorithms related to both are tested and analyzed and the best amongst them has been chosen to enhance the efficiency of the proposed technique. Three color descriptors and a comparison between machine learning and deep learning techniques have been given in this section.

A. Color Histogram

Color histogram (CH) is the basic technique utilized for extracting color information of an image. It is generally based on depicting total pixels in an image. It is a type of graphical representation which counts the number of pixels in fixed intervals called as bins. Bar graph has two axes where X axis represents the tonal colors and Y axis signifies the number of pixels in that particular bin. Some post-processing steps like Histogram normalization, equalization and quantization can also be applied to it. There are two techniques of calculating a Histogram (i) Global color Histogram (GCH) (ii) Local color Histogram (LCH) [23]. When a histogram of an entire image is calculated least concerning about the color distribution in a particular part of an image, it is specified as GCH whereas when an image is divided into many blocks and histogram of all these blocks is calculated, it is called as LCH. A general form of color histogram can be denoted as follows:

$$h(p, q, r) = N \cdot \text{Prob}(P=p, Q=q, R=r) \quad (1)$$

where P, Q and R denotes the three color channels of an image and N signifies the number of pixels of an image. Color histogram can also be used for the calculation of mean and standard deviation of a particular category of images or large datasets when images are difficult to count. The total number of pixels in an image is equal to the aggregate of gross elements and given by:

$$\text{Number of Pixels} = \sum_{i=0}^{255} H[i] \quad (2)$$

In “Equation (2)” H signifies the histogram of an image. Now, mean and standard deviation can be computed by making use of the following formulas:

$$\text{Mean}(\bar{x}) = \frac{\sum_{i=0}^{255} i * H[i]}{\text{Number of pixels}} \quad (3)$$

$$\text{Standard deviation}(\delta) = \left(\frac{1}{\text{Number of pixels}} \sum_{i=0}^{255} H[i] * (i - \bar{x})^2 \right) \quad (4)$$

Here, Color histogram has been chosen as an initial technique for color extraction because identification of different forms of data, frequency of occurring of data becomes easier by using this technique. It also gives precise information about the skewness of the data. Moreover, the computation of the histogram features is very fast and even for the large datasets, it can be efficiently calculated [24]. Color histograms are also invariant to scale, translation and rotation but they do not provide any spatial information regarding an image. Therefore, to overcome this limitation of color histogram, other descriptors like color moment and color auto-correlogram has been used.

B. Color Moment

Color moments (CM) are functions which denotes color distributions in an image. Divergent moments indicate diverse analytical and statistical measures. This color descriptor is also scale and rotation invariant but unlikely color histogram, it includes the spatial information from images [25].

The *i*th color channel and *j*th image pixel is represented by *I_{ij}* and *N* denotes the number of pixels in an image then 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:

$$\text{Mean}(E_{r,i}) = \frac{1}{N} \sum_{j=1}^N I_{ij} \quad (5)$$

After mean, the next successive moment is given by Standard deviation and is equal to the square root of variance of color distribution. It is given as:

$$\text{Standard deviation}(\sigma_{r,i}) = \left(\frac{1}{N} \sum_{j=1}^N (I_{ij} - E_{ri})^2 \right)^{\frac{1}{2}}. \quad (6)$$

Third moment is called as skewness and gives information regarding an image about its asymmetric characteristic. It is given as:

$$\text{Skewness}(S_{r,i}) = \left(\frac{1}{N} \sum_{j=1}^N (I_{ij} - E_{ri})^3 \right)^{\frac{1}{3}}. \quad (7)$$

Fourth moment is denoted by Kurtosis and it signifies about the shape of the color distribution like whether the obtained shape is tall or flat.

These color moments captures the spatial information of the given image and gives more accurate information regarding the color content of an image. This technique is very fast, efficient, robust and scalable in its performance.

C. Color Auto-correlogram

A color auto-correlogram (CAC) is a technique which gives information about contiguous relationship within color pixels of an image and also with change in distance.

Correlogram can also be defined by means of a stored table which is indexed by color pair (i,j) , where the chances of detecting a pixel j from pixel i at a span of distance is denoted by d . whereas in auto-correlogram, it specifies the possibility of detecting a pixel i from the same pixel at a distance d [26]. Therefore, it can be said that auto-correlogram captures and computes the spatial relationship between identical colors or levels.

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

$$\gamma_{c_i, c_j}^d(I) \triangleq \Pr_{p_1 \in I_{c_i}, p_2 \in I} [P2 \in I_{c_j} \mid |P1 - P2| = d] \quad (8)$$

The above equation gives the information regarding probability of pixel p_1 of level c_i is at a distance d from another pixel p_2 of level c_j . Auto-correlogram computes the contiguous correlation between indistinguishable levels only. The auto-correlogram for a color image is given as:

$$\alpha_c^d(I) = \gamma_{c_i, c_j}^d(I). \quad (9)$$

In the above equation, with the same level c_i and a distance of d , the probability of pixels p_1 and p_2 are given [27]. Thus, it can be concluded that color correlogram takes into account, spatial relationship among color pixels.

Similar like color moments, color auto-correlogram includes the spatial correlation of different colors and moreover, it can be used to describe the local spatial correlation in the global form. Also, it has very fast execution with simple computation. Based on the characteristics and advantages of the three color descriptors, these have been selected to form a fusion based color descriptor.

D. Machine learning (SVM vs ELM)

Machine learning is a trending topic now-a-days. It is generally an application of artificial intelligence which provides the capability of automatic learning to the computer systems. Generally, machine learning models are classified into two types: Supervised and Unsupervised. In case of supervised model, learning is based on the labelled data while unlabeled or unclassified data is utilized in the case of unsupervised learning.

To evaluate the performance of machine learning, two of its precise and popular models particularly, Support vector machine (SVM) and Extreme learning machine (ELM) has been tested and analyzed on various benchmark CBIR datasets.

Naive's Bayes is also used effectively as a classifier but for the transformation of obtained features into discrete features, a binning technique is required. Random forest, KNN are also used as classifiers but lacks in accurate classification due to some issues. Therefore, SVM and ELM are chosen to find the best technique in machine learning algorithms.

SVM is a machine learning algorithm which aims to find a hyperplane which in turn classifies the data efficiently into two parts based on the maximal distance of the hyperplane. Different types of kernel functions like Sigmoid, Radial basis function (RBF), linear, polynomial [28] etc. can be chosen for the functioning of a SVM. These are widely used due to their ease of working. But, they consume more training time and no multi-class SVM is directly available. Therefore, a second machine learning model namely ELM has been tested on the proposed hybrid color descriptor.

An Extreme learning machine (ELM) is a single-layer feed-forward neural network (SLFN) which can be used for classification and regression applications. It is a supervised learning model based on labeled data. It was initially introduced by Huang at el. [29]. ELM has been successfully utilized in many research problems like pattern recognition, classification, fault identification, over-fitting, etc. It is based on a single hidden layer of neurons and the total number of nodes present in this layer is a principal parameter to be decided. ELM has many advantages in contrast to many related traditional techniques like: least human involvement, swift learning speed, complimentary universal capability, convenience to use, varied kernel functions, etc. [30]. The basic diagram of ELM is shown in Fig.1.

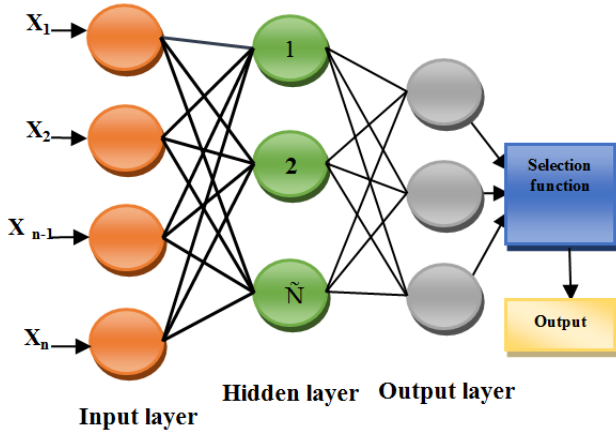


Fig.1. A basic Extreme learning machine model

This ELM is indeed based on a single hidden layer of neurons but the hidden layers cannot be increased in this case. Therefore, to increase the number of hidden neurons, for the purpose of obtaining more accurate results, the interest of the researcher community has been shifted from machine learning to deep learning.

E. Deep learning (Patternnet neural network vs CFBPNN)

The achievements in machine learning have now transformed into an eminent domain, known as Deep learning. Deep learning techniques are influenced by the various patterns of processing and communication involved in human biological nervous system. Deep learning is a part of machine learning which uses a cascade of multiple layers for its operation. It can be used for feature extraction, classification and many other functions. These deep learning networks can be trained in a supervised or in an un-supervised manner. Two neural network models namely Pattern recognition neural network and Cascade forward back propagation neural network (CFBPNN) have been tested in this proposed work.

Patternnet is a type of Feed forward back propagation neural network (FFBPNN) which can be trained to classify the given inputs according to the classes of targets [31]. 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.

Cascade forward back propagation neural network (CFBPNN) uses back propagation algorithm for its operation [32]. In these networks, 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 CFNN model is given as:

$$z = \sum_{j=1}^m g^j w_j^j x_j + g^o (w^b + \sum_{i=1}^l w_i^o g^h (\sum_{j=1}^m w_{ij}^h x_j)) \quad (10)$$

In “(10)”, g^j denotes the activation function from an input layer towards an output layer and w_j^j 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:

$$z = \sum_{j=1}^m g^j w_j^j x_j + g^o (w^b + \sum_{i=1}^l w_i^o g^h (w_i^b \sum_{j=1}^m w_{ij}^h x_j)) \quad (11)$$

The general architecture of CFBPNN is shown in Fig. 2 which depicts input layer containing three neurons, hidden layer of two neurons and an output layer.

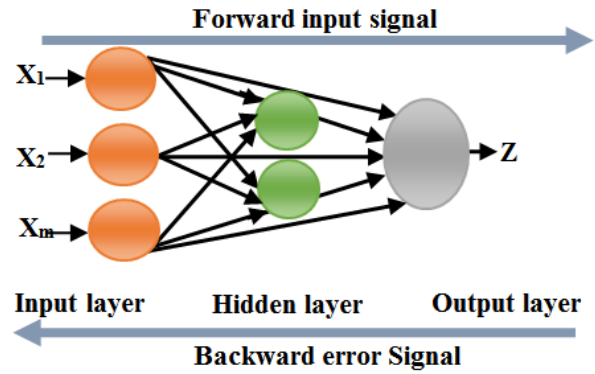


Fig.2. A basic Cascade forward back propagation neural network

IV. PROPOSED METHODOLOGY

The proposed system consists of a color based fusion retrieval system in which three color feature descriptors namely color histogram (CH), color moment (CM) and color autocorrelogram (CAC) have been combined.

Color histogram has very fast computation and is rotation and scale invariant. Autocorrelogram includes the spatial correlation between colors and can be used to depict the global distribution of local spatial content. Moreover, this technique performs well even if the provided color information is bristly in its characteristics. Similarly, color moments also include the spatial information between colors. The greatest advantage of using this technique is that the complete color distribution table is not required and color indexes are maintained. Therefore, the triad of these techniques jointly becomes incredible.

There is an interminable competition between machine learning and deep learning since many years. To find the supremacy among these learning techniques, two models each belonging to both the heads have been tested and analyzed on seven benchmark CBIR datasets.

Firstly, two machine learning models particularly SVM and ELM have been trained and then tested on all the utilized datasets. Then, in the second phase of the experimentation, two deep learning models namely

Pattern recognition neural network and CFBPNN have been trained and tested. From these two sets of the experimentation, various parameters like Precision, Recall and Accuracy has been calculated on all the four models. From these obtained results, it is well transparent and clear that deep learning has superior results as compared to machine learning and amongst the deep learning models, CFBPNN outperforms Patternnet and finally this model has been added to the designed hybrid color descriptor to enhance the retrieval performance of the system.

Lastly, after this experimentation, the proposed model finally is an amalgamation of Color Histogram, Color Moment and Color autocorrelogram. And to increase the efficiency of the proposed hybrid color descriptor, Cascade forward back propagation neural network has been used as a classifier. Broadly, the final proposed technique can be divided into two steps: Training and Testing.

A. Training stage

This initial stage is basically related to training of the neural network model i.e. CFBPNN. In this stage, the three utilized techniques namely CH, CM and CAC are utilized to extract color features and three independent feature vectors are formed. These feature vectors are finally combined to form a hybrid feature vector (HFV) by the process of normalization. Normalization brings the feature dimensions into a common range.

This HFV is applied as an input to a Cascade Forward back propagation neural network and the prescribed model gets trained with the complete dataset of images. 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 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 testing of the proposed system by using a query image from each dataset is further divided into three main sections.

(1) Feature extraction and normalization section

In this section again color features of a specific query image are extracted using the three techniques namely CH, CM and CAC and three respective feature vectors are formed. After feature extraction, the three obtained feature vectors are combined to form a Hybrid feature vector (HFV) by the process of normalization. The basic architectural implementation is given in Fig. 3.

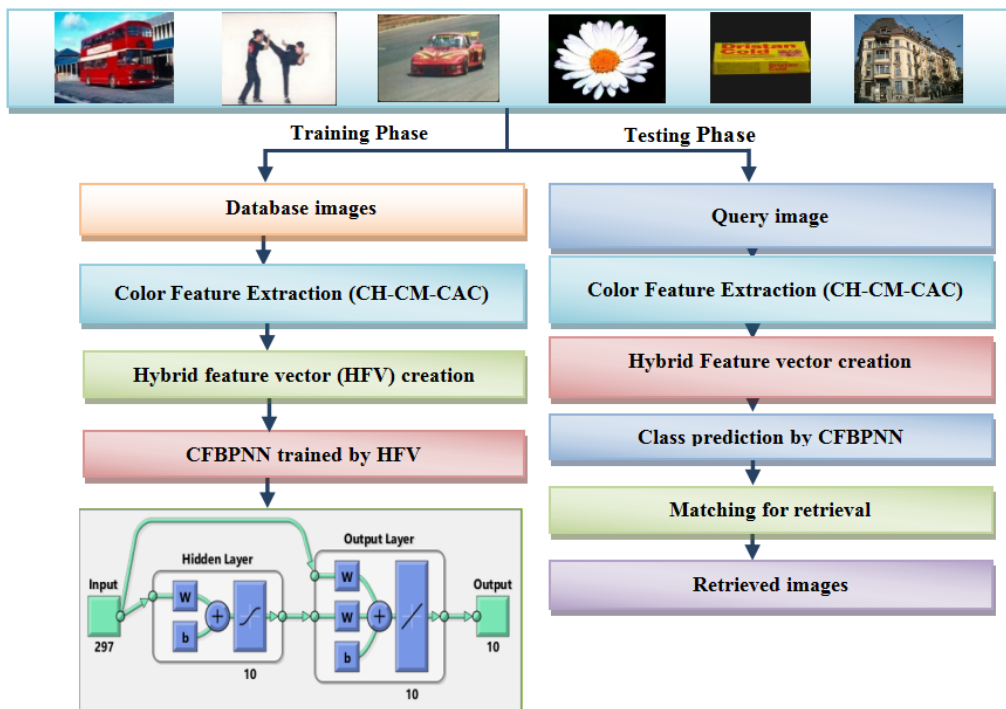


Fig.3. Architectural set-up of the proposed system

(2) Classification section

The obtained HFV is applied to the input of a pre-trained CFBPNN. This neural network model performs an efficient classification and class prediction based on the query image.

(3) Similarity matching section

In this section, the similarity between a query image and the categories formed by CFBPNN is calculated based on

a specific distance metric. Then, top N images are retrieved after this step. The prominent distance metrics [33] which are utilized in the proposed technique are given as under:

$$dis_{Euclidean} = \sqrt{\sum_{i=1}^n (|I_i - D_i|)^2} \quad (12)$$

$$dis_{Manhattan} = \sum_{i=1}^n |I_i - D_i| \quad (13)$$

$$dis_{\cosine} = 1 - \cos \theta = \frac{X \cdot Y}{\|X\| \|Y\|} \quad (14)$$

$$dis_{Minkowski} = \left[\sum_{i=1}^n (|I_i - D_i|)^p \right]^{\frac{1}{p}} \quad (15)$$

Here, I_i denotes the input query image and D_i depicts all database images related to Euclidean, Manhattan and Minkowski distances.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In order to analyze the retrieval efficiency of the proposed technique, various benchmark datasets has been tested and analyzed. All the experiments are performed in MATLAB R2017a, core i3 processor, 4 GB memory, 64-bit windows. A brief description of each dataset is given below and its sample images in Fig. 4.

Dataset 1: Corel-1K: This database consists of 1000 images with 10 categories in it and each category has 100 images in it. The size of each image is 256×384 or 384×256 . (<http://wang.ist.psu.edu/docs/related/>).

Dataset 2: Corel-5K: It is the second utilized dataset which has 5000 images and 50 categories with each category consisting of 100 images. Each image bears the size of 256×384 or 384×256 . (<http://www.ci.gxnu.edu.cn/cbir/>).

Dataset 3: Corel-10K: This database has 10000 images and 100 categories in it. Each category has further 100 images in it. The size of an image is 256×384 or 384×256 . (<http://www.ci.gxnu.edu.cn/cbir/>).

Dataset 4: Oxford flower: It consists of 1360 images with 17 categories. Each category has 80 images in it. . <http://www.robots.ox.ac.uk/~vgg/data/flowers/>

Dataset 5: Coil-100: This dataset consists of 7200 images with 100 categories and each category has 72 images in it. <https://www.kaggle.com/jessicali9530/coil100>

Dataset 6: Zubud: It is dataset of Zurich buildings with 1005 images and 5 images in each category. There are

201 categories in this dataset. <http://www.vision.ee.ethz.ch/showroom/zubud/>



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



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



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



Fig.4(d). Sample images from Oxford flower dataset

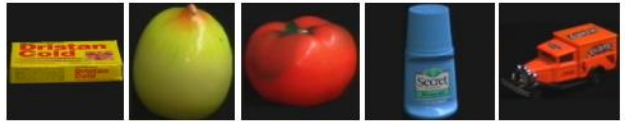


Fig.4(e). Sample images from Coii-100 dataset



Fig.4(f). Sample images from Zubud dataset

A. Evaluation metrics

There are many evaluation metrics which can be used to judge the performance of any CBIR system. [34]. But, the prominent evaluation metrics are:

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \quad (16)$$

$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Number of relevant images in the database}} \quad (17)$$

$$\text{F-Measure} = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (18)$$

B. Retrieval results

The retrieval results of the desired images are obtained by considering each and every image of all the datasets as the query images.

In order to check the efficiency of the proposed color hybrid descriptor with the two machine learning techniques, an experimental analysis has been done and the results are given in Table 3.

Table 3. Average Precision (%) on two machine learning models

Dataset	Used Technique		
	Hybrid Only	Hybrid with SVM	Hybrid with ELM
Corel-1K	92.7	89	94.6
Corel-5K	70.5	65.5	87.2
Corel-10K	68.2	64.4	85.1
Oxford flower	93.52	70.8	95.2
Coil-100	94.7	79.6	95.8
Zubud	79.7	70.5	80.13

From Table 3, it can be concluded that among the two machine learning models namely SVM and ELM, the ELM has superior results as compared to the previous one. This is so because, it is based on a feed forward neural network and contains a single hidden layer of neurons. Moreover, SVM takes much time in training of images and the results obtained are not so precise.

Similarly, to evaluate the proposed hybrid descriptor with the two deep learning models, these two have been experimentally tested on the proposed model and the results are given in Table 4.

Table 4. Average Precision (%) on two deep learning models

Dataset	Used Technique		
	Hybrid Only	Hybrid with Patternnet	Hybrid with CFBPNN
Corel-1K	92.7	95.2	97.1
Corel-5K	70.5	88.5	90.3
Corel-10K	68.2	86.4	87.9
Oxford flower	93.52	97.05	98.4
Coil-100	94.7	96.2	98.9
Zubud	79.7	80.5	82.7

From Table 4, it can be seen that, among the two deep learning models viz: Patternnet and CFBPNN, the latter one has enhanced results and therefore, this deep learning model has been finally chosen to be used with the proposed hybrid color descriptor as a classifier.

Finally, it can be seen that between machine learning and deep learning models, the superior results have been obtained by using deep learning techniques and among the two deep learning models, CFBPNN has higher precision than Pattern recognition neural network. The average precision of various techniques in different levels of the proposed system is given in Fig. 5. Therefore, CFBPNN Various working parameters of the proposed system are given in Table 6.

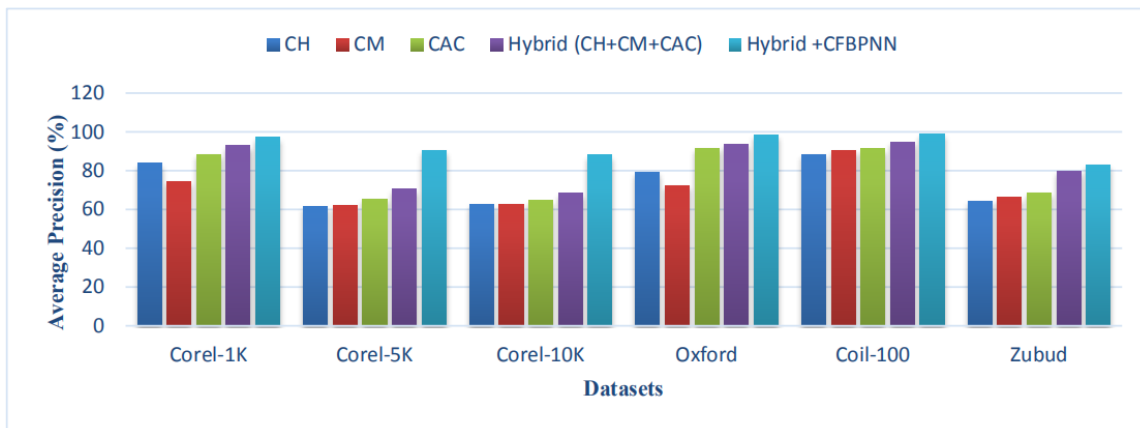


Fig.5. Performance of the proposed color descriptor (Average Precision (%) during various levels)

has been chosen to work as a classifier with the proposed color fusion descriptor to enhance the retrieval accuracy of the system.

From Fig. 5, it is evident that that efficiency of the designed system increases as more and more preciseness is being added to it. The results of the hybrid descriptor are enhanced than the independent color techniques but the results of the proposed technique i.e. combined hybrid and CFBPNN are much more augmented than the hybrid descriptor on all the utilized datasets.

Various types of distance metrics can be utilized for calculating the similarity between images. Here, in this proposed work, four prominent distance metrics have been evaluated on the proposed technique and the results in terms of Average precision are given in Table 5.

Table 5. Average Precision (%) vs distance metrics

Dataset	Distance metrics			
	Euclidean	Manhattan	Cosine	Minkowski
Corel-1K	91.3	92.7	89.6	90.7
Corel-5K	67.5	70.5	62.1	66.7
Corel-10K	65.4	68.2	61.7	64.6
Oxford flower	92.94	93.52	91.76	92.0
Coil-100	93.2	94.7	90.1	91.5
Zubud	76.5	79.7	70.3	75.2

Here, the results shown in Table 5 are obtained by using color fusion descriptor with various distance metrics. Here, Manhattan distance metric has superior

results as compared to other metrics because it is based on absolute value results in comparison to squared value results. Moreover, it is fast in computation.

Precision and Recall are both considered as important measures to evaluate the performance of any CBIR system. Therefore, Precision vs recall curves by varying the number of retrieved images from 10 to 50 on five benchmark datasets except Zubud are given in Fig. 6(a). and 6(b), as it has only 5 images per category.

Table 6. Working parameters of the proposed system

Fused features	CM (9) + CH (32) + CAC (256) = 297
Number of Output (Classes)	10 (Corel-1K), 50 (Corel-5K), 100 (Corel-10K), 17 (Oxford), 100 (Coil-100), 201 (Zubud)
ELM Kernel Type	Radial basis function (RBF)
SVM Type	Multi-class
Patternnet Training	Scaled Conjugate gradient
CFBPNN Training	Levenberg-Marquardt

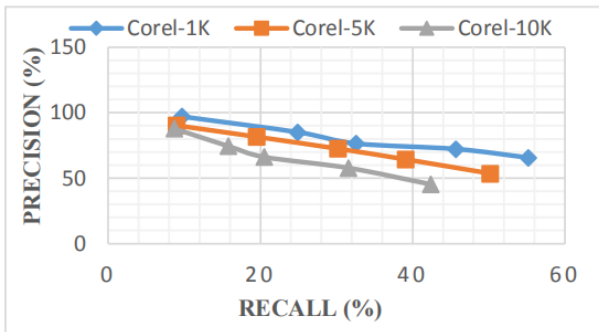


Fig.6(a). Precision vs Recall curves on Corel-1K, Corel-5K and Corel-10K

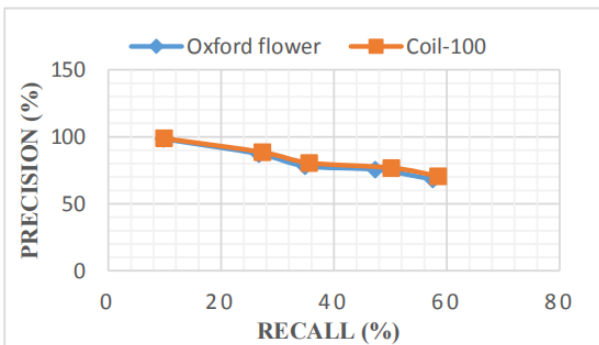


Fig.6(b). Precision vs Recall curves on Oxford and Coil-100

From Fig. 6(a). and 6(b), it can be seen that as the number of retrieved images increases, the value of Precision comes down and Recall increases. Separate Graphical user interface (GUI's) have been designed for each of the six datasets. Top ten images are retrieved for each utilized dataset except zubud dataset, where top 5 images are retrieved because it has 5 images in each category. To show the retrieval accuracy, the GUI's for three utilized dataset have been given in Fig. 7 (a-c).

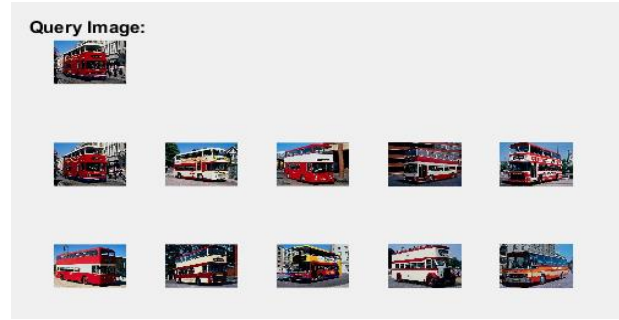


Fig.7(a). Retrieval results for Corel-1K dataset



Fig.7(b). Retrieval results for Oxford flower dataset

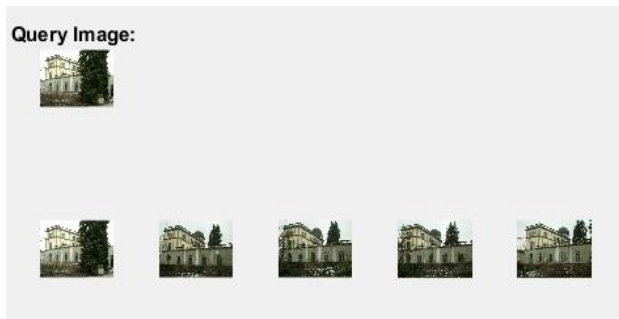


Fig.7(c). Retrieval results for ZUBUD dataset

From Fig 7(a-c), it can be concluded that the top N images are retrieved from the desired category of the dataset which in-turn belongs to the native category of the query image. Thus, the proposed system is efficient in retrieving the desired images from the required dataset.

In order to validate the performance of the proposed hybrid color descriptor, the proposed system has been compared with many state-of-the-art techniques. The main considerations of these comparisons are:

Majority of the hybrid CBIR systems lack the presence of any Deep learning technique but the proposed system encapsulates CFBPNN, which plays a major role in increasing the retrieval accuracy of the system. The comparative analysis of the proposed technique based on three datasets namely Oxford flower, COIL-100 and ZUBUD is shown in Table 7, 8 and 9.

Table 7. Comparative analysis with the related methods

Dataset	Average Precision (%)				
	Hybrid CH-CM-CAC	Ref. [21]	Ref. [10]	Ref. [11]	Proposed
Oxford flower	93.52	89.02	65	81.6	98.4

Table 8. Comparative analysis with the related methods

Dataset	Average Precision (%)				
	Hybrid CH-CM-CAC	Ref. [15]	Ref. [16]	Ref. [17]	Proposed
Coil-100	94.7	95	82.5	96.5	98.9

Table 9. Comparative analysis with the related methods

Dataset	Average Precision (%)				
	Hybrid CH-CM-CAC	Ref. [13]	Ref. [14]	Ref. [12]	Proposed
ZUBUD	79.7	75.85	70	79.12	82.7

In order to show the effectiveness of the proposed system, the retrieval time of the developed system based on different utilized datasets has been compared to many related techniques as shown in Table 10.

Thus, the proposed system has faster retrieval rate as compared to some related techniques.

The comparison based on the three datasets namely Corel-1K, Corel-5K and Corel-10K with the state-of-the-art techniques is given in Table 11. The compared techniques suffer from many disadvantages like: The compared hybrid descriptors lacks an amalgamation of any intelligent technique which in turn helps in retrieving the higher level semantic information, thereby reducing the “semantic gap”. Moreover, the proposed hybrid color descriptor has been tested and evaluated on both well-known machine learning and deep learning models on many datasets. This, comparative analysis between machine learning and deep learning is absent in most of the compared systems.

Table 10. Retrieval time analysis

Technique used	Retrieval time (in seconds)	
[4]	1.74	
[9]	1.848	
[15]	1.545	
Proposed system on utilized datasets	Corel-1K	0.52
	Corel-5K	1.26
	Corel-10K	1.54
	ZUBUD	0.378
	Oxford	0.249
	COIL-100	0.752

Table 11. Comparative analysis with the related methods

Dataset	Semantic name	Average Precision (%)						
		Hybrid CH-CM-CAC	Ref. [7]	Ref. [9]	Ref. [4]	Ref. [8]	Ref. [6]	Proposed
Corel-1K	Africa	86	75	95	72.4	81	74.5	96
	Beach	95	83	60	51.5	66	69.3	98
	Building	85	71	55	59.55	78.75	85.1	90
	Bus	99	84	100	92.35	96.25	95.4	95
	Dinosaur	100	95	100	99	100	100	100
	Elephant	88	71	90	72.7	70.75	83.3	100
	Flower	100	93	100	92.25	95.75	98	98
	Food	97	72	100	72.35	77.25	92.6	100
	Horse	87	75	100	96.6	98.75	94.2	98
	Mountain	90	81	75	55.75	67.75	75.9	96
	Average	92.7	80.0	87.5	76.5	83.22	77.8	97.1
Corel-5K	Average	70.5	65.2	72.56	55.25	68.6	59.5	90.3
Corel-10K	Average	68.2	60.54	65.45	49.58	59.98	52.45	87.9

Thus, from these comparisons with the related techniques, it can be concluded that the proposed technique has superior results and can perform precisely on both small and large benchmark datasets.

VI. CONCLUSION AND FUTURE TRENDS

This paper carries out a comparable study between some machine learning and deep learning models. In order to check the supremacy between the two, a hybrid color descriptor is proposed here which is based on a

combination of color histogram, color moment and color auto-correlogram. Color histogram is invariant to scale and angle rotation while both color moment and color auto-correlogram precisely captures spatial information of an image. This triad of feature extraction techniques together produces an efficient hybrid descriptor. To evaluate machine learning techniques, Support vector machine and Extreme learning machine have been tested while from deep learning category, Pattern recognition neural network and Cascade forward back propagation neural network have been tested. Among, these CFBPNN has superior results as compared to other models. To

validate these facts, simulation results are provided which are finally compared with many latest hybrid CBIR systems. Our future work will be focused on developing a hybrid CBIR system by extracting color, texture and shape features of an image. Last but not the least, the concept of Internet of things (IoT) can be used for the online transfer of desired images.

REFERENCES

- [1] A. Alzu'bi, A. Amira, and N. Ramzan, "Semantic content-based image retrieval: A comprehensive study", *J. of Vis. Commun. and Image Representation*, vol. 32, pp. 20-54, 2015
- [2] A. K. Naveena, and N. K. Narayanan, "Image Retrieval using combination of Color, Texture and Shape Descriptor", Proceedings in *Next Generation Intelligent Systems (ICNGIS)*, *IEEE*, 2016, pp. 1-5.
- [3] A. R. Kumar, and D. Saravanan, "Content Based Image Retrieval using Color Histogram", *Int. J. of Comp. Science and Information Technology*, vol.4, no.2, pp. 242-245, 2013.
- [4] S.Fadaei, R. Amirfattahi, and M. R. Ahmadzadeh, "New content-based image retrieval system based on optimised integration of DCD, wavelet and curvelet features", *IET Image Processing*, vol. 11, no. 2, pp. 89-98, 2017.
- [5] V. Naghashi, "Co-occurrence of adjacent sparse local ternary patterns : A feature descriptor for texture and face image retrieval", *Opt. – Int. J. of Light Electron Optics*, vol. 157, pp. 877-889, 2018.
- [6] M. A. Ansari, D. Dixit, Kurchaniya, and P. K. Johari, "An Effective Approach to an Image Retrieval using SVM Classifier", *Int. J. of Computer Sciences and Engg.*, vol. 5, pp. 64-72, 2018.
- [7] Y. Mistry, D.T. Ingole, and M.D. Ingole, "Content based image retrieval using hybrid features and various distance metric", *J. of Electrical System of Information Technology*, vol. 2016, pp. 1-15, 2017.
- [8] L. K. Pavithra, and T.S. Sharmila, "An efficient framework for image retrieval using color, texture and edge features", *Comp. and Electrical Engg.*, vol. 0, pp.1-14, 2017.
- [9] J. Pradhan, S. Kumar S, and H. Banka, "A hierarchical CBIR framework using adaptive tetrolet transform and novel histograms from color and shape features", *Digit. Signal Process. A Revolution Journal*, vol. 82, pp. 258-281, 2018.
- [10] H. Riddhi, Shaparia, M. Narendra, P. Zankhana, and H. Shah, "Flower Classification using Texture and Color Features", *Kalpa Publications in Computing*, vol. 2, pp. 113-118, 2017.
- [11] A. Huthaifa, M. Saher, and H. Hazem, "A Flower Recognition System Based On Image Processing And Neural Networks", *Int. J. of scientific & technology research*, vol. 7, pp. 1-9, 2018.
- [12] L. Yang, H. Lei, W. Siqi, L. Xianglong, and L. Bo, "Efficient Segmentation for Region-based Image Retrieval Using Edge Integrated Minimum Spanning Tree", Proceedings in *23rd International Conference on Pattern Recognition (ICPR)*, 2016, 4-8 Dec. ancun, Mexico: pp. 1-6.
- [13] C. Iakovidou, N. Anagnostopoulos, A. Kapoutsis, et al., "Localizing global descriptors for content based image retrieval", *EURASIP Journal on Advances in Signal Processing*, 10.1186/s13634-015-0262-6: pp.1-20, 2015.
- [14] P. Shreelekha, K. Pritee, and Y. Haruo, "Clustering of hierarchical image database to reduce inter-and intra-semantic gaps in visual space for finding specific image semantics", *J. of Visual Communication and Image Representation*, vol. 38, pp. 704-720, 2016.
- [15] E. Mehdi, E. Aroussi, N. El. Houssif, "Content-Based Image Retrieval Approach Using Color and Texture Applied to Two Databases (Coil-100 and Wang)", *Springer International Publishing*, https://doi.org/10.1007/978-3-319-76357-6_5, pp. 49-59, 2018.
- [16] A. Heba, "Combining SURF and MSER along with Color Features for Image Retrieval System Based on Bag of Visual Words", *J. of Computer Sciences*, pp.1-10, 2016.
- [17] L. Shenglan, W. Jun, F. Lin, et al., "Perceptual uniform descriptor and Ranking on manifold: A bridge between image representation and ranking for image retrieval", *J. of Latex*, arXiv:1609.07615v1 [cs.CV] 24 Sep 2016, pp. 1-14, 2016.
- [18] S. Sandeep, and R. Rachna, "Content Based Image Retrieval using SVM, NN and KNN Classification", *Int. J. of Advanced Research in Comp. and Comm. Eng.*, vol. 4, no. 6, pp. 549-552, 2015.
- [19] P. Tomas, and M. Virginijus, "Comparison of Naïve Bayes, Random Forest, Decision Tree, Support Vector Machines, and Logistic Regression Classifiers for Text Reviews Classification", *Baltic J. Modern Computing*, vol. 5, no.2, pp. 221-232, 2017.
- [20] K.S. Arun, and V. K. Govindan, "A Hybrid Deep Learning Architecture for Latent Topic-based Image Retrieval", *Data Science Eng.*, vol. 3, no.2, pp.166-195, 2018.
- [21] K. Lin, H.F. Yang, and C.S. Chen, "Flower Classification with Few Training Examples via Recalling Visual Patterns from Deep CNN", Proc. In *International Conference on Computer vision, graphics and image process.* (CVGIP), 2015, pp. 1-9.
- [22] L. Weibo, W. Zidong, L. Xiaohui, Z. Nianyin, L. Yurong, and E.A. Fuad, "A survey of deep neural network architectures and their applications", *Neurocomputing*, vol. 234, pp. 11-26, 2017.
- [23] S. Kodituwakku, and S. Selvarajah, "Comparison of color features for image retrieval", *Indian J. of Comp. Science*, vol. 1, no. 3, pp. 207-211, 2004.
- [24] N.M. Sanmukh, and M.L. Tejaswini, "Color Histogram Features for Image Retrieval Systems", *Int. J. of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 4, pp.10941-10946, 2014.
- [25] S.M. Singh, and K. Hemachandran, "Content -Based Image Retrieval using Color Moment and Gabor Texture Feature", *Int. J. of Computer Science Issues*, vol. 9, no. 2, pp. 99-309, 2012.
- [26] V. Vinayak, "CBIR System using Color Moment and Color Auto-Correlogram with Block Truncation Coding", *Int. J. of Computer Applications*, vol. 161, no. 9, pp.1-7, 2017.
- [27] J. Huang, S. K. Kumar, M. Mitra M, et al., "Image indexing using color correlograms", Proceedings in *IEEE Computer Soc. Conf. Computer Vision Pattern Recognition*, 1994, 191(3-4): 7, pp.62-768.
- [28] Fu. Ruigang, Li. Biao, G. Yinghui, and P. Wang, "Content-Based Image Retrieval Based on CNN and SVM", Proceedings in *2nd IEEE International Conference on Computer and Communications*, 2016, pp. 1-5.
- [29] M. Mansourvar, S. Shamshirband, R.G. Raj, R. Gunalan, and I. Mazinani, "An Automated System for Skeletal Maturity Assessment by Extreme Learning Machines", *PLOS One*, doi:10.1371/journal.pone.0138493, pp. 1-14, 2015.

- [30] S. Liu, H. Wang, J. Wu, and L. Feng, "Incorporate Extreme Learning Machine to content-based image retrieval with relevance feedback", In Proceedings of *11th World Congress on Intelligent Control and Automation* Shenyang, China, 29 June-4 July (IEEE), 2015, pp. 1010-1013.
- [31] Z. Weixun, N. Shawn, Li. Congmin, S. Zhenfeng, "PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval", *ISPRS Journal of Photogrammetry and Remote Sensing*, pp.1-13, 2018.
- [32] N.A. Omaina, A.A.Tamimi, and M.A. Alia, "Face Recognition System Based on Different Artificial Neural Networks Models and Training Algorithms", *Int. J. of Advances in Computer Science and Applications*, vol. 4, pp. 40-47, 2013.
- [33] N. Arora, A. Ashok, and S. Tiwari, "Efficient Image Retrieval through Hybrid Feature set and Neural Network", *I.J. Image, Graphics and Signal Processing*, vol. 1, pp. 44-53, 2019.
- [34] S. P. Rana, M. Dey, and P. Siarry, "Boosting content based image retrieval performance through integration of parametric & nonparametric approaches", *J. Vis. Commun. Image R.*, vol. 58, pp. 205-219, 2019.
- Machine Learning and Deep learning", *International Journal of Intelligent Systems and Applications(IJISA)*, Vol.12, No.2, pp.21-33, 2020. DOI: 10.5815/ijisa.2020.02.03

Authors' Profiles



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.



Dr. 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.



Dr. Pawan Kumar Dahiya received his Ph.D degree in ECE. He is working as an Assistant 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 50 publications in various national and international reputed journals.

How to cite this paper: Shikha Bhardwaj, Gitanjali Pandove, Pawan Kumar Dahiya, "A Genesis of a Meticulous Fusion based Color Descriptor to Analyze the Supremacy between