A New Approach for CBIR based on Color Coding-MTSD with Pattern Recognition Neural Network

Nisha Rajput[1] Bhupendra Verma[2]
[1,2]Department of Digital Communication
[1,2]NITM Datiya, India

Abstract— Content Based Image Retrieval (CBIR) is the method which uses visual data to find the images from a huge database on the foundation of user's input in conditions of query image and usually some effective features for color string, HSV histogram, edge orientation and intensity map are extracted. This study proposes an picture characteristic illustration manner centered on Color Coding-Multi-Trend Structure Descriptor (CMTSD) and Pattern Recognition Neural Network (PRNN). We overview the outcome on Corel dataset which includes a thousand and 5000 pictures. And the experimental results based on precision, accuracy and feature extraction (FE) time. The proposed algorithm is in similarity with MTSD procedure and it is better than MTSD approach. The distance between two matrixes is calculated using different similarity measures, namely, L1, Euclidean distance (ED), ChebyChev, Hamming and Jaccard distance. Neural network (NN) is used for categorization of query picture as per training database. At first NN is trained about the image features in the database. The training is done by using pattern network algorithm. This trained database is used for classification of the query image. According to retrieved image class further CMTSD based CBIR is used for retrieving similar images.

Key words: MTSD, CBIR, Neural Network, Color Coding, Precision, Accuracy

I. Introduction

With growing development of image acquisition devices, the volume of image database is expanding rapidly. An efficient picture retrieving instrument is required by the customers of more than a few domains comparable to far flung sensing, crime prevention, trend, medical diagnosis and many others. Basically, image retrieving tools are defined using the following terms: query, result presentation, features, and matching. For image retrieval purpose, the various image retrieving tool has been developed, see. For picture retrieving motive, 2 frameworks had been developed, text-based image retrieval (TBIR) and CBIR [1].

CBIR remove the picture from database by analyzing its contents. The content might refers to texture, color, shape or the spatial layout of an picture. Fig. 1 presents the block diagram of CBIR. It performs 2 most important tasks:-

1) Feature Extraction (FE)
2) Similarity Measurement (SM)

The effectiveness of CBIR techniques will depend on the performance of the algorithm for extracting the elements of picture [1]. In this, the visual characteristics of the pictures are represented by feature vectors and hence form a feature database.

A. Multi-Trend Structure Descriptor (MTSD)

Juleszs textons thought suggests that pictures encompass particular neighborhood structures which most often show a distinctive quantity of comparison in analogous pictures. For this reason, regional structures can also be considered as a fundamental picture property. By means of local structures, low level features blend well in pictures. Thus, it is suitable to use local structures to mine the internal image correlations. In this paper, the proposed algorithm MTSD integrates color, facet orientation and intensity info as an entire for picture illustration.

1) Visual Feature Extraction and Quantization

Lots of visible knowledge is hidden in pictures. In our proposed procedure, color, Edge orientation and depth understanding are selected for function illustration.

Color understanding is extracted in the HSV color area, on account that it's regarded to be essentially the most compatible color space to imitate a human's visual system. present are 3 add-ons in HSV color space:

Hue (H), Saturation (S) and Value (V). Based on the theory that human eyes cannot distinguish large numbers of colors at the same time, color quantization is used in our method. The purpose of color quantization is to assign a distinctive set of colors to represent a picture maximum valuable info. We divide H, S and V color channels into 12, three and 3 boxes, respectively, therefore resulting in 12* 3*3=108 combos got in complete.

2) Local Structure and Multi-Trend Definition

The foremost concern of MTSD is ways to outline multi-trend structure. Many nearby constructions had been designed for picture content evaluation and FE. In the proposed algorithm, we prefer 3* three blocks as the elemental neighborhood structure. Taking orientations into account, there are more than a few positions between pixels within the nearby constitution; as a consequence, we define four neighborhood structure patterns in accordance to four angle directions in 3 * 3 blocks: 0°, 45°, 90° and 135°.

3) Image Feature Representation

It can be considered as the most critical problem for picture retrieval, and good feature representations have Excessive discrimination energy, and toughen picture retrieval efficiency extensively. As mentioned before, taking advantage of local spatial structure information can contribute to performance improvement, therefore, how to integrate visual features in local structure to explore local spatial structure information for picture feature representation is a challenging issue. Based on the defined local structure and multi-trend together with color, edge orientation and intensity information.[2]

B. Neural Networks (NN)

It is a network of "neuron like" units called nodes. This neural computing manner is utilized in fields of classification, optimize, and manage idea and for solving regression issues. NN are very robust in the case of classification problems where detection and consciousness of goal is required NN is preferred over other techniques due to its dynamic nature. Dynamic nature is carried out by means of adjusting the weights in step with ultimate output and applied enter information. This adjustment of weights takes the position iteratively unless desired output is received. And this weight adjustment of the network is often called "studying" of NN. [3].

C. HSV Color histogram

Color function is among the primary matters to entry the picture. The color of a picture is represented from the noted color spaces like RGB, XYZ, YIQ, L*a*b, U*V*W, YUV and HSV [1]. HSV color space gives the best CH feature, among the different color spaces [1]. HSV color gap is represented by three accessories corresponding to Hue (H), Saturation(S), and value (V).[4].

$$H = cos^{-1}\left\{\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)+(R-B)(G-B)}}\right\}$$

$$S = 1 - \frac{3}{R+G+B}[\min(R, G, B,)]$$

$$V = \frac{1}{3}[R+G+B]$$

II. LITERATURE SURVEY

This paper proposes a picture feature representation procedure, particularly MTSD, which is constructed, founded in the neighborhood and multi-development constructions. The scene competencies, such as color, area orientation and depth map are viewed and quantized, and with the regional constitution as a bridge, we use multi-trend to detect color, aspect orientation and intensity map respectively for FE. MTSD can characterize not only the low-level features, such as color, shape and texture, but also the local spatial structure information.

This system makes use of the contents of the picture information for segmenting, indexing, retrieval and browsing of imperative images from picture repository. This paper most commonly concentrates on the indexing section of the image retrieval approach for development of an effective indexing algorithm of CBIR systems.

This paper describes a new approach in which each part and color functions of the pictures are taken into consideration for generation of characteristic vectors. DWT is used to preserve the detailed contents of the images along with the reduction of the size of the feature vector. A comparison study on the effect of the proposed method on YCbCr and HSI color spaces is also presented in this paper.

On this paper, HSV founded texton histogram (HSV-TH) is proposed for CBIR. The HSV-MT is proposed not just like the RGB established text on histogram (RGB-TH). The proposed HSV-TH approach is founded on Julesz's textons conception, and it works directly with nature pictures as shape descriptor and a color texture descriptor. HSV-TH integrates the benefits of co- incidence matrix and the histogram by way of the use of representing the characteristic of co-prevalence matrix utilizing histogram.

Color Histogram (CH) founded on HSV and color Moments (CM) are generally utilized in picture retrieval. In this paper, we focus on the study concerning the picture retrieval and suggest a brand new color function, referred to as Cascade CMs. By dividing the image into blocks, we add spatial information of the picture into the color features. The Cascade CMs feature is formed by cascading the CMs of each block. The experimental results exhibit that Cascade CMs is better than HSV CH and CM when taking the retrieval precision into consideration.

Color is a vital feature of color pictures. Various color models are there like, RGB, CMYK, CIE Lab, CIE XYZ, HSV, HSL, etc. Madhura C proposed that, converting RGB images into some other color space like Gray, Lab, YCbCr, CMY, HSV, HIS and then processing them gives good results.

III. PROPOSE WORK

In this algorithm, now we have proposed a mixture of color, shape and texture features. In this approach, the previous work is enhanced to achieve better precision and reduce feature extraction time. This approach consists of the next steps: (i) Pre Processing (ii) feature Extraction (iii) Classification

A. Pre Processing

In this process, take a query image which is resized with 192*128 sizes.

B. Feature Extraction

1st, resize all frames to curb the consequences of variation in dimension. On the grounds that the frames will have distinctive sizes, all frames are normalized to an ordinary measurement (i.e. 20 × 20 pixels) on this step. Herein, all frames is resized by way of the bi cubic interpolation process. After the transfer, each frame will become a 2D number, array, and then we will convert the 2D number array to a number as below: 66666666666646566666…11111111111111111111 (20×20 = 400 numbers) we can perceive the power of discrimination between different frames because 400 numbers present 6400 permutations. Color information is extracted in HSV color space, because it is more suitable for human eye perception. The human eye cannot distinguish large quantity of colors even as; color quantization is used on this process. Color quantization is to assign a certain group of colors to indicate the image with maximum useful information. Split the image into three components, H, S and V and assign 12, 3, 3 bins to each component, thus resulting in 12*3*3=108. Edge orientation is used for shape features and it identifies the object edges and image information. In this algorithm, Sobel operator is used to produce the edge orientation map. Intensity information is given by the V color channel, which represents the brightness of color in HSV color space. By means of intensity, uniform quantization, the intensity map M_I is obtained. Finally merged all features. This process is also executed for database images. Store all features into matrix file.

C. Classification

After the feature extraction process, read the stored database of features and query image features. The database is divided into 10 classes in The Corel-1k dataset and 50 classes in Corel-5k dataset: Africa, Beach, Monuments, Elephant, Horses, Building, Food, Flower, Mountain, and

Dinosaur etc. Then randomly select the training dataset and test dataset for classification. After database classification we will get the class labels (each class contains related images). The educational process includes production, configuring a 3-layered NN and making it be taught about the extracted elements of training picture. Training set includes all the images from an image database. The learning process is carried out using pattern network algorithms, which include computing error, updating weights in order to minimize the error. Given enter question picture contains the caption. Then examine enter question picture caption with the every type label within the gigantic database. In order that we can without difficulty identify that the query enter picture belongs to certain category label or now not. If the input query picture belongs to any person of the class label within the database, now we will choose handiest that matched elegance pictures inside the database and carry out retrieval operations on that selected particular part of the database based on the similarity matching.

1) Proposed Algorithm

1) Consider RGB image as a query image $Q(x,y)$ with the size of M*N.
2) Resize the image with 20*20 size and apply cubic interpolation on the image.
3) We have know-how that the values of r, g, and b are completely wonderful with the altered illumination conditions. Nonetheless, the relative values between r (i), g (i), and b (i) are very identical. For that reason, make use of 6 rules to transfer each and every frame to a color string as follows:
   1) if a pixel R > G > B, next allots the pixel as 1;
   2) if a pixel R > B > G, after that allocates the pixel as K(i,j,1)= 2;
   3) if a pixel G > R > B , next allots the pixel as K(i,j,2)= 3; if a pixel G >= B >= R , next gives the pixel as K(i,j,2)= 4;
   4) if a pixel B >= R >= G , then allocates the pixel as K(i,j,3)= 5;
   5) if a pixel B >= G >= R , next gives the pixel as K(i,j,3)= 6;
4) We split H, S and V color channels into 12, three and 3 containers, respectively, thus leading to 12*three*3=108, combinations bought in L_H.
5) Calculate the histogram for three components using Edge Histogram. Initiate the filters for the 5 types of edges:
   – f1(:,:,1) = [1 2 1;0 0 0;-1 -2 -1]; %vertical
   – f1(:,:,2) = [-1 0 1;-2 0 2;-1 0 1]; %horizontal
   – f1(:,:,3) = [2 2 -1;2 -1 -1; -1 -1 -1];% 45 diagonal
   – f1(:,:,4) = [-1 2 2; -1 -1 2; -1 -1 -1];%135 diagonal
   – f1(:,:,5) = [-1 0 1;0 0 0;1 0 -1]; % non directional
6) Calculate the max sobel gradient and index of the orientation
   $$[m,i] = \max(g_{im},3)$$
   Where m is max sobel gradient, i is index of the orientation, g_{im} is query image
7) Find the Edge utilizing canny procedure
8) Multiply edge image with the types of orientations detected by the Sobel masks:
   $$im = (i \times edim)$$
   Where im is detected image, edim is canny image
   $$L_E = hist(im(:,5)$$
   Where L_E is final histogram of image on five directions
9) We define $M_I(x,y)$ as the intensity quantized value in the position of (x,y) , where $0 \le x \le M-1, 0 \le y \le N-1$, and $M_I(x,y) = \beta$, $\beta \in (0,1 \ldots \ldots N_I - 1)$,where $N_I = 20$ in this paper.
10) Repeat Step 1 to Step 9 until all images in the dataset.
11) Determine the similarity matrix of query image and image dataset using L1 distance, Euclidean, Chebychev, Jaccard and Hamming distance.
12) Then, by splicing $[L_C\, L_H\, L_E\, L_I]$ together, the histogram of an image can be represented by $L_T = [L_C\, L_H\, L_E\, L_I]$
13) In this manner, L_T will be 189 dimensions instead of 974 dimensions, which can save the computation time and memory space to a great extent.
14) Compute the distinction or comparison between two vectors.

$$d^{JAS}(i,j) = \frac{j11}{j01+j10+j11}$$

In the equation d^{JAD} is the Jaccard gap between the objects *i* and *j*. For 2 info records with *n* binary variables *y* the variable index *k* ranges from 0 to *n*-1. Four different combinations between $y_{i,k}$ and $y_{j,k}$ can be distinguished when comparing binary variables. These combinations are (0/0), (0/1), (1/0) and (1/1). The sums of these combinations can be grouped by:

– J_{01}: the total number of variables being 0 in y_i and 1 in y_j.
– J_{10}: the total number of variables being 1 in y_i and 0 in y_j.
– J_{11}: the total number of variables being 1 in both y_i and y_j.
– J_{00}: the total number of variables being 0 in both y_i and y_j.

As each and every paired variable belongs to one of these groups it may be effortlessly visible that:

$$J_{00} + J_{01} + J_{10} + J_{11} = n$$

because the Jaccard similarity is founded on joint presence, J_{00} is discarded.

15) Classify the images using pattern recognition NN classifier.
16) Calculate correctness, FE time, precision, and keep in mind of retrieved pictures. In the subject of pictures retrieval, two foremost indices are probably used to assess the efficiency of methods: precision and consider. Precision P is outlined because the ratio between the number of retrieved primary pictures to the number of the retrieved photos awarded to users, while do not forget R is the ratio between the quantity of the retrieved founded pictures and the total number of the crucial images in the dataset.

$$Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)}$$

$$P = \frac{No.\ of relevant image retrieved}{Total number of image retrieved}$$

$$R = \frac{No.\ of relevant image retrieved}{number of image in the datebese}$$

Where Confusion matrix consists of 4 variables particularly:

1) True positive (TP): The percentage of total number of genuine images taken as genuine, closer to 1 is better.
2) True negative (TN): The percentage of total number of genuine images taken as forged, closer to 1 is better.

3) False positive (FP): The percentage of forged images taken as genuine, closer to 0 is better.
4) False negative (FN): the percentage of forged images taken as forged, closer to 0 is better.

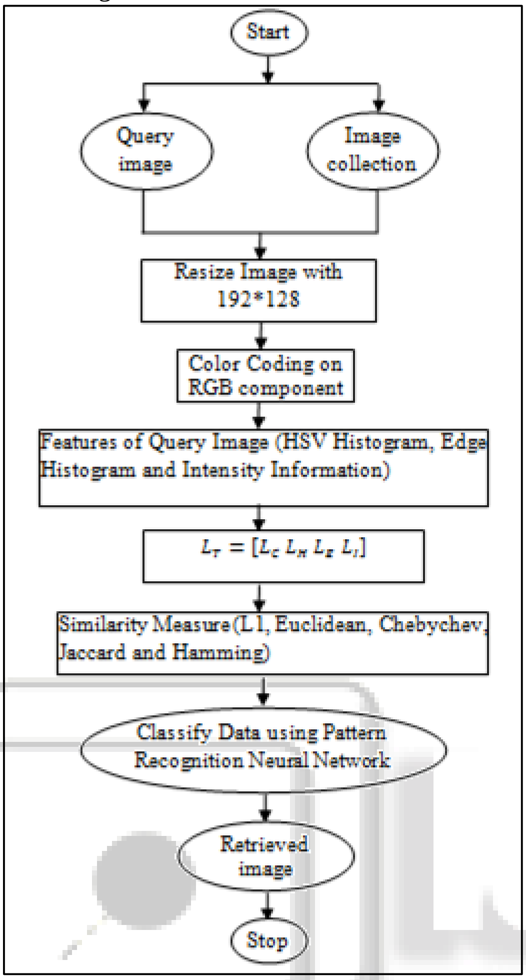

Fig. 1: Flow Process of Proposed System

IV. RESULS ANALYSIS

The experimental results performed on following datasets, Corel-1k and Corel5k are used. On Corel-1k dataset, each and every category consist one hundred pictures of dimension 192×128 in JPG structure and equal as Corel-5k dataset.

This algorithm is performed on popular similarity metrics, such as L1 distance, Euclidean distance, Chebyshev distance, Hamming and Jaccard. Table 3 suggests the normal retrieval precisions and accuracy of MTSD with specific similarity metrics on 2 datasets. As can be visible from table 3, our similarity metric outperforms the other 3 metrics. Euclidean distance is widely used for similarity measure, but it is time consuming for the square operation calculation. Besides, Euclidean distance greatly increases the difference between two dissimilar feature vectors. Chebyshev distance performs the worst and is not suitable for the proposed algorithm. L1 gap is extra efficient for calculation, nonetheless, it has the issue of dropping the rotation invariant property, and its performance is lower than our metric obviously. Our similarity metric may also be viewed as a Jaccard metric, which takes into consideration both the computational complexity and the experimental performance, hence,it is more appropriate than the other four metrics.

Table 1: Corel-1k Dataset

Table 2: Corel-5k Dataset

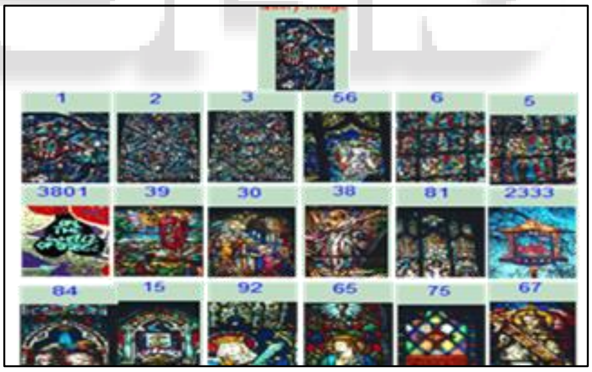

Fig. 2: Retrieval of Images for Churches by Proposed Algorithm on the Corel-5k dataset

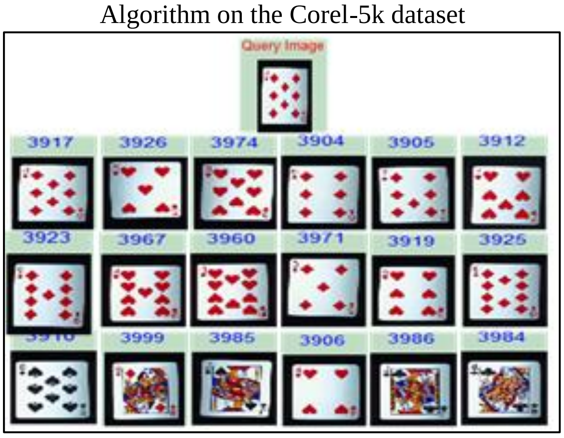

Fig. 3: Retrieval of Images for Playing Cards by Proposed Algorithm on the Corel-5k dataset.

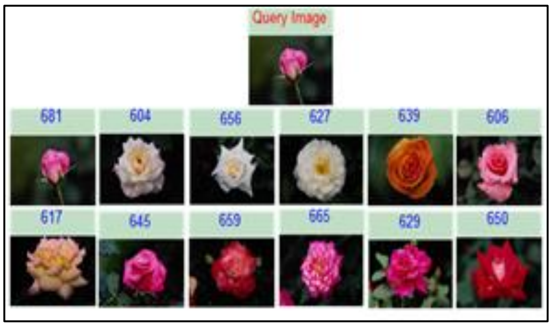

Fig. 4: Retrieval of Images for Flower by Proposed Algorithm on the Corel-1k dataset.

Fig. 5: Retrieval of Images for Racing Car by Proposed Algorithm on the Corel-5k dataset.

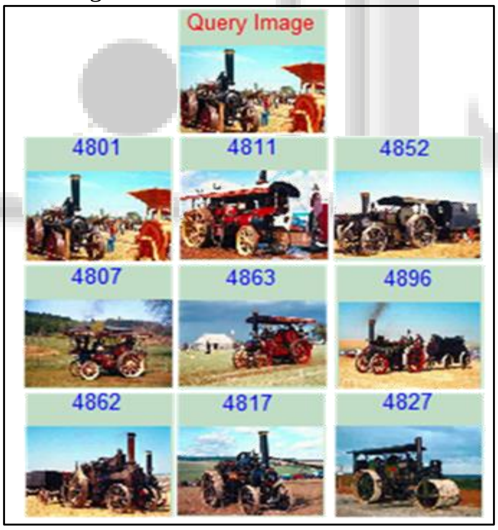

Fig. 6: Retrieval of Images for Agricultural Vehicle by Proposed Algorithm on the Corel-5k dataset.

Fig. 7: Retrieval of 20 Images for Statues by Proposed Algorithm on the Corel-5k dataset.

Figs. 2 and 7 show two examples of image retrieval on Corel-5k dataset by MTSD. In Fig. 6, the query image is agricultural vehicles and the entire top 16 retrieval pictures exhibit excellent fit of color and shape to the questionable picture. In Fig. 7, the query snapshot is swimmers and the entire high sixteen retrieval pictures showing just right match of texture and color to the question picture.

Fig2 to Fig7 show six examples of CBIR on two datasets by CMTSD. In Fig7, the query imageis statues and all the top 20 matched images whcih shows good match of color, texture and shape to the query image. In Fig 6, the query image is agricultural vehicle and all the top 9 matched images. In Fig2, the query image is churches and all top 18 retrieval images but some images are false match.

Dataset	MTSD Accuracy (%)	Proposed Accuracy (%)	MTSD Precision (%)	Proposed Precision (%)
Corel-1k	68.50	81.33	55.17	84.61
Corel-5k	52.42	63.2	53.44	67.18

Table 3: The average retrieval precision and accuracy by various methods on two datasets (the best values are in boldface)

Methods	Dimension	Feature Extraction Time (sec.)
MTSD	179	9.1615
Proposed	189	2.1698

Table 4: The dimension of the feature vector and feature extraction time.

The quantization level of edge orientation is 9	The quantization level of color is 108			
	MTSD Precision (%)	Proposed Precision (%)	MTSD Accuracy (%)	Proposed Accuracy (%)
Corel-1k	55	77.61	64.86	82.66
Corel-5k	50.98	67.59	52.97	71.33

Table 5: The average retrieval precision and accuracy of MTSD and Proposed under quantization levels of color and edge orientation when the quantization level of intensity is fixed to 20 (the best values are in boldface).

Category	Proposed Precision of Similarity Metrics (%) on Corel-5k dataset				
	L1	Euclidean	ChebyChev	Hamming	Jaccard
Church	72.61	71.98	68.62	67.45	70.55
Playing Cards	66.12	70.12	62.96	69.27	63.84
Flag	69.6	69.35	64.61	65.20	62.80
Agricultural Vehicle	61.44	70.22	63.82	67.2	67.07
Trees	64.70	70.57	66.48	70.53	68.03

Table 6: The average retrieval precision of Proposed Method on different similarity metrics

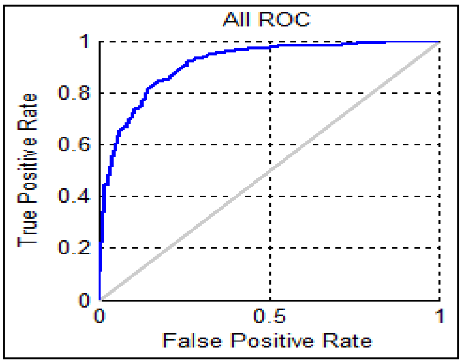

Fig. 8: ROC by Proposed Algorithm on the Corel-1k dataset

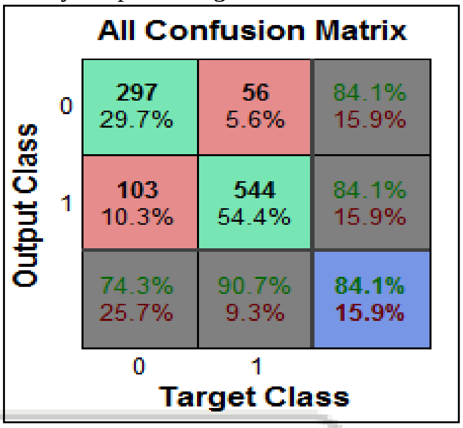

Fig. 9: Confusion Matrix by Proposed Algorithm on the Corel-1k dataset

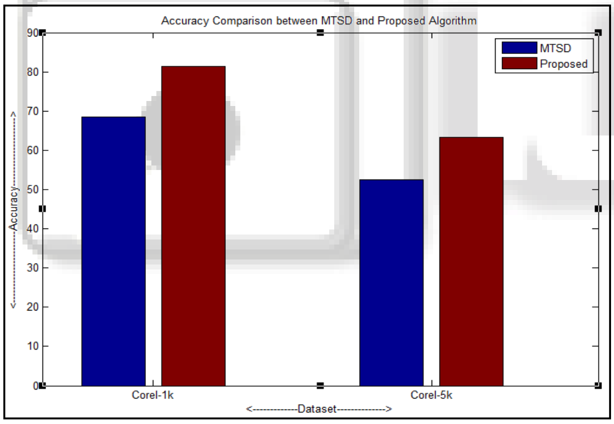

Fig. 10: Accuracy is comapred By MTSD and Proposed Method

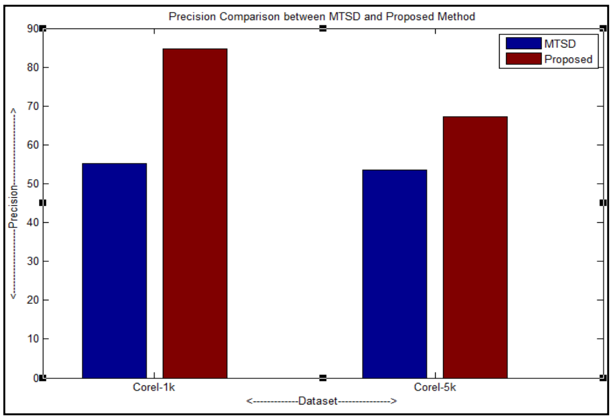

Fig. 11: Precision is comapred By MTSD and Proposed Method

V. CONCLUSION

After comparing the proposed method with previous method, it is found that Pattern Recognition NN gives an optimal result as compared to the first method. In proposed system, the image classes used by us give very good images retrieval accuracy. Our system performance is good in terms of precision value, time and accuracy value and it shown in graph. At the same time the accuracy is uniform for all the classes making PRNN a better choice. Here we use different types of distances like L1, Euclidean, Chebychev, Hamming and Jaccard to calculate similarity between two images. In this system, the overall precision and accuracy has reached up to **77.61**% as well as **82.66%** for Corel-1k dataset and **67.59%** as well as **71.33%** for Corel-5k dataset. Previously, did not use any model training, that performance was not better on the basis of distance measure only. The experimental results on two datasets demonstrate that CMTSD This technique focuses simplest on retrieval of picture files, but in destiny, this work may be more desirable to retrieve the audio and video record with the aid of the usage of identical capabilities or editing them.

REFERENCES

[1] Komal Juneja, Akhilesh Verma, Savita Goel, Swati Goel, "A Survey on Recent Image Indexing and Retrieval Techniques for Low-level Feature Extraction in CBIR systems". 978-1-4799-6023-1/15 $31.00 © 2015 IEEE

[2] Meng Zhao, Huaxiang Zhang, Jiande Sun, "A novel image retrieval method based on multi-trend structure descriptor". 1047-3203/_ 2016 Elsevier Inc. All rights reserved.

[3] Arvind Nagathan, Manimozhi, Jitendranath Mungara, "Content-Based Image Retrieval System using Feed-Forward Backpropagation Neural Network". IJCSNS. 2014.

[4] R. Usha , K. Perumal , "Content Based Image Retrieval using Combined Features of Color and Texture Features with SVM Classification". International Journal of Computer Science & Communication Networks.2014

[5] Jiande Sun, Huaxiang Zhang , Meng Zhao, "A novel image retrieval method based on multi-trend structure descriptor". 2016 Elsevier

[6] Md. Khalid Imam Rahmani, M. A. Ansari, Amit Kumar Goel, "An Efficient Indexing Algorithm for CBIR". 978-1-4799-6023-1/15 $31.00 © 2015 IEEE

[7] Nooby Mariam, Rejiram R, "A Modified Approach in CBIR Based on Combined Edge Detection, Color and Discrete Wavelet Transform". 978-1-4799-8792-4/15/$31.00_c 2015 IEEE

[8] Qin-Jun Qiu, Yong Liu, Da-Wei Cai, Jia-Zheng Tan, "Modified Color Texton Histogram for Image Retrieval". 978-0-7695-5125-8/13 $26.00 © 2013 IEEE

[9] Tianfa Weng, Yule Yuan, Ling Shen, Yong Zhao, "Clothing Image Retrieval Using Color Moment". 978-1-4799-0561-4/13/$31.00 ©2013 IEEE

[10] Madhura C., Dheeraj D., "Feature Extraction for image retrieval using colour spaces and GLCM", IJITEE.2013.