Semiconductor Optoelectronics, *Vol. 42 No. 2 (2023), 1204-1215*
https://bdtgd.cn/

Semiconductor Optoelectronics
ISSN: 1001-5868

DEVELOPMENT OF A HYBRID CONTENT-BASED IMAGE RETRIEVAL SYSTEM USING A TWO LEVEL HIERARCHICAL CLASSIFICATION MECHANISM

Shikha Bhardwaj
1UIET, Kurukshetra University, India

Shefali Dhingra
UIET, Kurukshetra University, India

Abstract: In the today's era, as the multimedia content is increasing swiftly on the internet, therefore, the requirement of an effective image retrieval system is urgently needed to retrieve or sort the desired images from a large repository of images. And this system is called as Content-based image retrieval (CBIR) system. In this paper, a hybrid CBIR system is formed by an amalgamation of two color-base retrieval techniques with a texture based technique, namely color moment (CM) and color auto-correlogram (CAC) with Gray level co-occurrence matrix (GLCM) respectively. Then, to add intelligence to the system, a two level hierarchical classification mechanism is added to the developed hybrid system. In the first level of hierarchy, two classifiers are used, viz., Extreme learning machine (ELM) and Support vector machine (SVM). In the second level of hierarchy, several ELM's are used together to remove remaining ambiguous images. Each classifier is trained with the images of the dataset. Two datasets are utilized for the experimentation purpose, namely Corel-10K and Oxford flower. To measure the efficiency of the proposed system, various parameters like Precision, Recall and Accuracy are evaluated. Average precision of 97.8% and 98.22% respectively on Corel-5K and Oxford flower benchmark datasets. The experimental analysis portrays that the implemented technique outmatches many state-of-the-art related approaches depicting varied hybrid CBIR system.
Keywords: Image retrieval, Color moment, Color correlogram, Gray level co-occurrence matrix, Extreme learning machine, Support vector machine, Two-level hierarchical classification.

Introduction

The steady growth of the internet, the falling prices of storage devices and the increasing pool of available computing power, etc. contributes to the large repository of digital information which is available on the internet efficiently. Evidently, manually interpreting of images from big image databases is a cumbersome and costly task [1]. Even if, the image is searched from the world's biggest databases, it is not possible to match user's expectation. The majority of image matches mostly rely on annotated text linked to that image. Due to the presence of thousands of irrelevant images to the query image, it is a complicated task to discover an image from the large database [2]. Therefore, the solution to this problem is to automatically extract the content of the images and the system based on this mechanism is called as Content based image retrieval (CBIR) system [3]. A basic CBIR system is given in Fig. 1.

Color, shape, texture, and any other information that can be received from an image itself is referred to as content in CBIR (Content-Based Image Retrieval).
There are two important stages in CBIR: one is the indexing stage and another is the retrieval stage. But, if a system can filter images based on their content, it may provide more precise outcomes. To develop a hybrid CBIR system, many types of features can be extorted depending on the requirements of the user. Color is the basic attribute of an image and techniques like color histogram [4], color coherence vector, dominant color descriptor, color correlogram, etc, can be used for its extraction. Texture defines different patterns of an image. Different texture descriptors like discrete wavelet transform (DWT), curvelet transform [5], gabor transform, color co-occurance matrix (CCM), Tamura features etc. have been used for texture extraction. Different edge detection techniques like Robert, Prewitt, Sobel, etc. can also be utilized for identification and location of keen disruption Shape is also a prominent feature of an image which bears the semantic information and has been categorized into region based and boundary based. B-splines, curvature scale space (CSS), hough transform, fourier transform, zernike moments etc. are handful of shape descriptors used in CBIR system [6]. But, when two or more attributes are combined together for the development of a hybrid system, more accurate results are obtained and moreover by adding a two level hierarchical classification mechanism adds human like intelligence to the system.
The main contributions of this paper are:

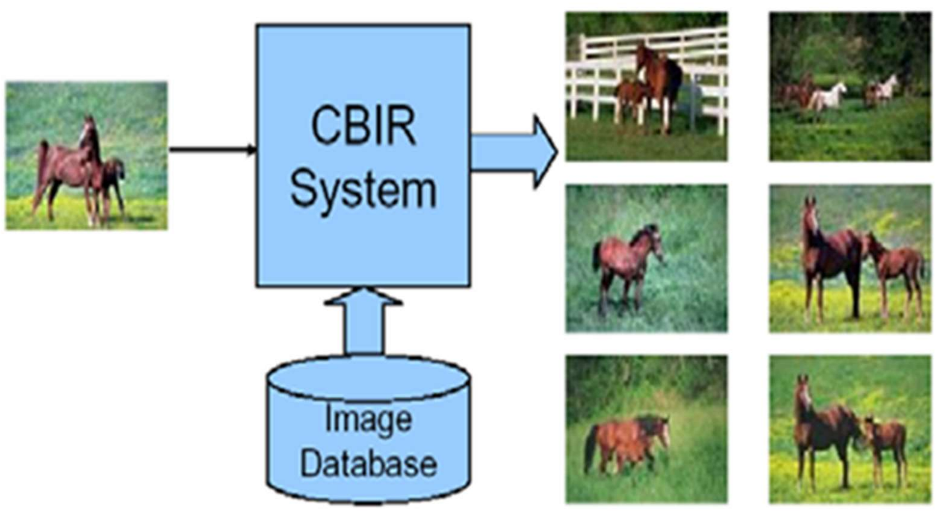

Figure. 1 Basic block diagram of CBIR System

- Extraction of two color based attributes, namely color moment and color auto correlogram for the formation of a color-based hybrid system.
- Addition of a texture based technique, Gray level co-occurrence matrix to the already developed color based system.
- Finally, a two level hierarchical classification mechanism is added to the final system to

remove the ambiguous images.

The organization of the paper is as follows: Related work is presented in section 2. Methodology Used and is given in section 3. In section 4, Performance evaluation and results are given. Finally, the paper ends with conclusion and future trends, which is presented in section 5.

State-of-the-art related work

The proposed work has been compared with many latest and related state-of-the-art techniques based on an image retrieval system. The detailed description is given as under:

S. Singh et al. 2020 [7] discussed about a layered based picture recovery framework. A proficient Bi-CBIR (bi-layer content-based picture recovery) framework is introduced that utilizes the initial layer for picture shifting and another layer for retrieval of images. When the surface and shape attributes from the initial layer and the shape and shading attributes of the another layer are deliberated merged, a retrieval system is formed. But, the shape suffers from many ringing and boundary defects.

L. K. Pavithra et al. 2017 [8] proposed another structure for effective recovery of comparative pictures utilizing texture, surface and edge features. From the chosen images, the texture and edge are extracted utilizing LBP (Local Binary Pattern) and Canny edge detector. The productivity of the hybrid structure moderately gives 83.22%, 68.60% and 59.98% retrieval rate on Wang's, Corel-5K and Corel-10K dataset. But, canny edge is itself a cumbersome and time consuming technique.

D. Kishore 2020 et al. [9] proposed a hybrid framework for CBIR machine which utilizes CS-SCHT (Conjugate Symmetric-Sequence Complex Hadamard Transform) dc coefficients, from attributes and color capabilities. HVS (Human Visual System) color quantization, ELM (Extreme Learning Machine), and statistical parameters are also used. The average precision and average recall acquired are 0.89 and 0.72 respectively. But, usage of a single classification technique, does not provide high end precision.

L. Putzu et al. 2020 [10] presented two methodologies to make a new system, one is based on the usage of CNN (convolutional neural network) as an extractor and the other is based on the utilization of CNN as a classifier. But, deploying of CNN is time consuming and complex in nature.

S. Bhardwaj et al. 2020 [11] proposed an effective color descriptor that is an integration of color auto-correlogram, CH (color histogram), and CM (color moment). SVM (support vector machine) and ELM (extreme learning machine) are two models that have been tried in order to decide the prevalence among machine learning. PatternNet and CFBPNN (Cascade Forward Back Propagation Neural Network) have been used. According to the analysis, Cascade Forward Back Propagation Neural Network (CFBPNN) has better outcomes compared to different models. But, usage of classifier in single layer, does not provide high recall and precision outcomes.

S. Fadaei et al. 2020 [12] proposed a system based on Zernike and Wavelet functions. This technique was used to eliminate all insignificant images from database. Then, Particle swarm optimization (PSO) is used. The PSO itself does not provide correct and accurate solutions to the problem.

S. Dhingra et al. 2020 [13] describes three different texture based techniques for the extraction on three distinctive datasets that is Corel-10 K, Wang, and Corel-5 K. The GLCM (gray-level co-occurrence matrix), Gabor transform, curvelet, DWT (discrete wavelet transform), and LBP (local binary pattern) are utilized here. But, texture techniques cannot alone provide precise results.

In this paper, a dual-stage model is used to extract and represent an image in a way that matches human perception, improving and expanding the CDH (color difference histogram) method. Using the properties of anti-color and Hue saturation value (HSV) color spaces, a new visual descriptor based on fused perceptual color information had been proposed by Z. Wei et al. 2020 [14]. Its goal is to create contents of the picture utilizing rapidity, edge, and hue orientation features in rival color & HSV color spaces, allowing it the strength to delineate color, shape, surface, spatial characteristics. No classification technique is utilized in this paper.

Methodology Used

For the development of the proposed system, initially two color descriptors namely, Color moment (CM) and Color auto-correlogram (CAC) are used for extracting color attributes of an image. Texture is also considered as an important element in extracting the features of an image. Therefore, to extract texture feature, Gray level co-occurrence matrix (GLCM) has been used. Now, a hybrid CBIR system is developed which is a combination of two color descriptors and a texture descriptor (CM+CAC+GLCM). Thus, a hybrid CBIR system is formed.

Now, a two level hierarchical classification mechanism is added to developed hybrid system. The first level consists of two classifiers, namely ELM and SVM, arranged in parallel and in the second level three classifiers, particularly ELM's are arranged in a sequential manner to make the final decision. The pair of classifiers in the first level receive input as a query image, as submitted by the user and their response is compared. If the compared result matches to each other, the particular image is chosen and if the comparison result is divergent, there must be some ambiguity in the query image submitted by the user. Now, this ambiguity is being resolved in the second level, where three ELM,s are arranged in a sequential manner and by passing through each one of them, a final unambiguous image is obtained, after matching with the query image as shown in Fig. 2. A brief description of all the applied techniques is given herewith:

3.1 Color Moment

Color moments (CM) are metrics that signify color distributions in an image. According to probability theory, its distribution can be effectively characterized by its moments. Color moment can be computed for any color space model. Different moments specify diverse analytical and statistical measures. This color descriptor is also scale and rotation invariant but it includes the spatial information from images [15-16]. Mean, standard deviation, skewness and kurtosis can be calculated by using color moment.

3.2 Color-Auto Correlogram

A color correlogram (CAC) is a color descriptor which gives information specifying contiguous relationship within color pixels of an image and also with the change in distance. Correlogram can also be defined as a stored table, indexed by color pair (i,j), which approximates the likelihood of detecting a pixel j from pixel i at a span of distance denoted by d, whereas in auto-correlogram, it specifies the possibility of detecting a pixel i from the same pixel at a distance d [17]. Therefore, it can be said that auto-correlogram computes the spatial relationship between identical colors or levels.

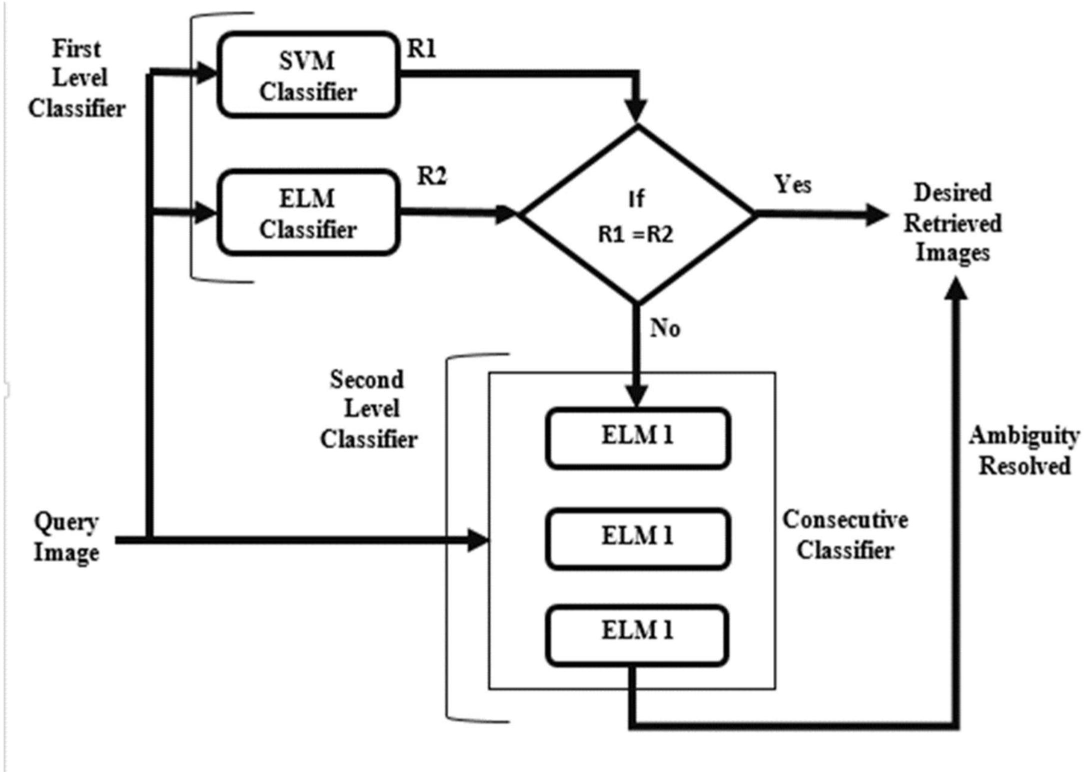

Figure. 2 Basic block diagram of applied Methodology

3.3 Gray level Co-occurrence Matrix

It is one of the precise texture extraction methods and computes the statistical properties within an image. It has fast computation and lesser complexity. A gray level co-occurrence framework for image disintegration has been proposed by Haralick et al. in the year 1973. Furthermore, GLCM is used to separate the features based on two phases: the initial phase is used to calculate GLCM and the subsequent phases to find different orientations of GLCM. It is referred to as co-occurrence distribution [18]

3.4 Support Vector Machine

In SVM, it designs each information object as a point in n-dimensional space (where n is a number of the ones you highlighted) with the value of every section being the value of a distinct organization. Then, on that point, it characterizes by tracking the hyperplane that isolates the two classes fairly well. The co-ordinates of individual perception are just supported vectors [19]. The SVM (support vector machine) classifier isolates the two classes (hyper-plane/line). The figure 3 shows the Support Vector Machine (SVM) algorithm.

3.5 Extreme Learning Machine

Many types of classifiers have been used based on image retrieval to acquire a specific accuracy during the classification of the given database. Extreme Learning Machine (ELM) is the latest learning algorithm which is single hidden layer feed-forward neural network. ELM depends on minimization hypothesis and its learning cycle needs just a solitary prominence [20]. The algorithm avoids multiple iterations and local minimization. ELM designates learning parameters such as weights, input nodes, and biases in an adequate way, containing no alteration [20]. Only the number of hidden nodes involved should be indicated, as the output is estimated factually by the independent variable. In comparison to several variant learning algorithms, ELM has a rapid learning rate, considerably superior efficiency, or reduced management interference, as shown in fig.4

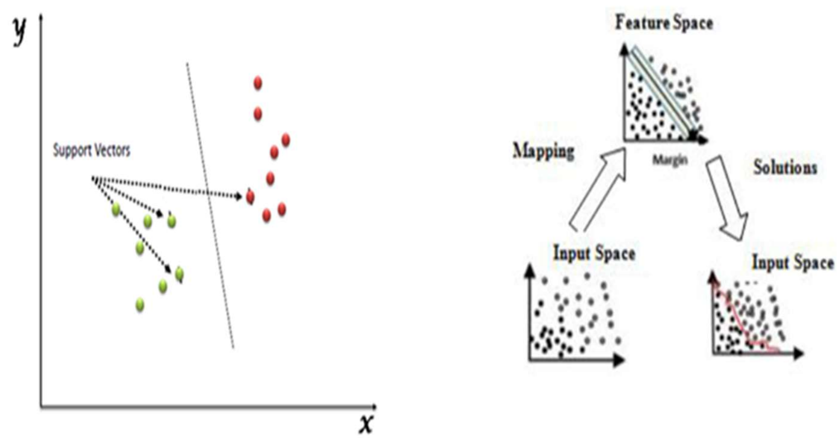

Figure. 3 Fundamentals of SVM

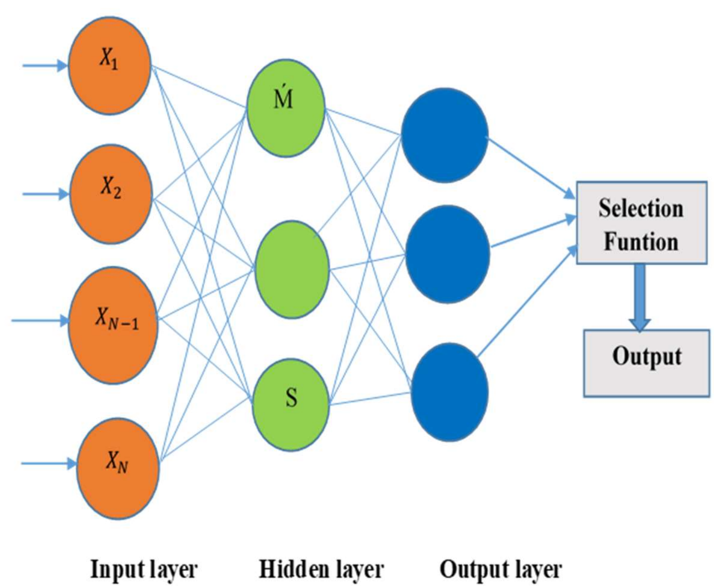

Figure. 4 Fundamentals of ELM

Performance Evaluation and Results

Various performance metrics and results are discussed in this section.

4.1 Performance Metrics

There are various evaluation metrics that can be used to evaluate the performance of a CBIR system.

But, the most prominent performance metrics are:

- **Precision:-** The Precision is defined as the fraction of the retrieved images that are indeed relevant for the query [21]:

$$Precision\ (P) = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved} \quad (1)$$

- **Recall:-** The Recall is the fraction of relevant images that is returned by the query [21]:

$$Recall\ (R) = \frac{Number\ of\ relevant\ images\ retrieved}{Number\ of\ relevant\ images\ in\ the\ database} \quad (2)$$

- **F-Measure:-** Precision and Recall are combined to find the F-score / F-measure to calculate the performance of the system [18].

$$F - Measure = 2\ \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

4.2 Results and Discussion

For the purpose of analysis and experimentation, two benchmark datasets have been utilized. All the experiments are performed in MATLAB R2018a, core i3 processor, 4 GB memory, 64 bit windows.

1st Dataset: Corel-5K: The first dataset is Corel-5K and consists of 5000 images. There are 50 categories in this dataset and again every group has 100 images. The size of each image is either 256 × 384 or 384 × 256. (http://www.ci.gxnu.edu.cn/cbir/)

2nd Dataset: Oxford flower dataset: The second dataset is Oxford flower dataset and there are 1360 images in this dataset and it has 17 image categories with each category consisting of 80 images. http://www .robot. ox.ac.uk/~vgg/ data/flowers/

Precision is the most important evaluation metric of a CBIR system. Therefore, it is calculated at each stage. The precision of the proposed system during the different stages of the implementation on the two datasets is given in Table 1.

From Table 1, it can be seen that hybrid CBIR system has enhanced precision as compared to independent techniques of color and texture. But, as two level classification is added to the system, the results are further improved. As seen 1st level classification gives precision of 95.4% and 96.3% on Corel-5K and oxford flower dataset. While when the system passes from 2nd level of classification, precision of 97.8% and 98.22% on Corel-5K and Oxford flower dataset respectively.

To match a given query image with the database image, a distinct distance metric is required. This distance metric is used for similarity matching. For this purpose, three prominent distance metrics are utilized for the purpose of calculating the similarity between a query and database images. The results are given in Table 2

Table 1. Average Precision (%) on the proposed system during various phases of its implementation

DATASET	CM	CAC	GLCM	CM+CAC+GLCM (Hybrid CBIR)	1st classifier level	2nd classifier level	Final System
Corel-5K	82.1	87.5	86.3	90.56	94.4	96.1	97.8

Oxford flower	77.5	81.23	82.5	92.33	95.3	97.22	98.22

Table 2. Average Precision (%) on the proposed system by using three prominent distance metrics

Datasets	Euclidean	Manhattan	Minkowski
Corel-1K	97.8	95.2	94.33
Oxford Flower	98.22	96.6	95.5

From Table 2, it can be seen that Manhattan is a distinctive case of Minkowski distance and it leads to the production of innumerable false negatives which do not yield accurate results. Euclidean distance metric is based on weighted and normalized attributes and has speedy computational performance. Average precision obtained by Euclidean distance metric outperforms others, so it is selected for similarity matching.

By obtaining a prominent and an effective value of accuracy, it can be concluded that the machine learning model has done an accurate classification. This performance of classification can be studied by a confusion matrix. A table which describe the performance of a classification model on a set of test data for which the true values are known. Therefore, the obtained confusion matrices for oxford flower dataset is given in Table 3.

Table 3. Confusion Matrix on Oxford flower dataset

Category Confusion Matrix (Oxford Flower Dataset)																	
1	78	1	1	0	0	0	1	1	0	0	0	0	0	1	2	1	0
2	0	76	0	1	0	2	0	0	0	0	0	0	1	0	0	0	1
3	0	0	79	0	0	0	0	0	1	0	0	0	0	1	0	1	0
4	1	0	0	74	0	0	0	1	0	1	0	0	1	0	1	0	1
5	0	0	0	0	76	1	1	0	0	0	0	0	2	1	0	0	1
6	0	0	0	0	0	80	0	0	0	0	0	0	0	0	1	0	1
7	0	0	1	0	0	0	80	0	1	0	0	0	1	0	1	0	0
8	0	0	0	0	0	0	0	79	0	1	0	0	1	1	0	1	0
9	0	0	1	0	0	0	0	1	78	0	0	0	0	0	0	1	0
10	1	0	0	0	0	0	2	0	1	80	0	0	1	1	1	1	2
11	0	1	0	0	0	0	0	0	0	0	79	0	0	0	0	0	1
12	0	0	1	0	0	0	0	0	1	0	0	80	0	1	1	0	0
13	0	1	0	0	0	0	1	0	0	1	1	1	78	1	0	1	1
14	0	1	2	2	0	2	0	1	1	0	1	0	1	80	0	0	1
15	1	0	1	1	1	0	0	1	0	1	0	1	0	1	79	1	0
16	0	0	0	0	0	1	1	0	1	2	1	0	1	0	1	76	0
17	0	2	0	2	1	0	0	0	0	0	0	1	0	1	0	1	80

Accuracy obtained by using a two-level hierarchical classification system is shown with the help of a graph as shown in fig. 5. From this graph, it can be seen that as a query image passes through the second level of classification, more accurate results are obtained. And among the two datasets, oxford flower dataset provides more accuracy on the applied technique.

Figure. 5 Accuracy Vs various levels of Classifiers

Separate Graphical user interfaces (GUI's) have been designed for the two datasets by retrieving top 10 images in each case as shown in Fig. (6a-b).

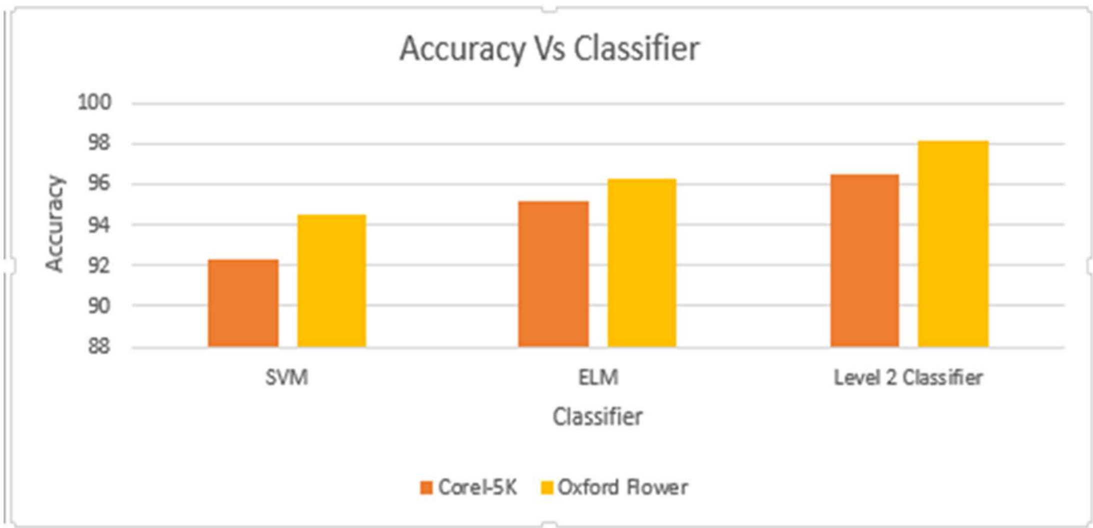

Figure. 6(a) Retrieval of top 10 images from Corel-10K dataset

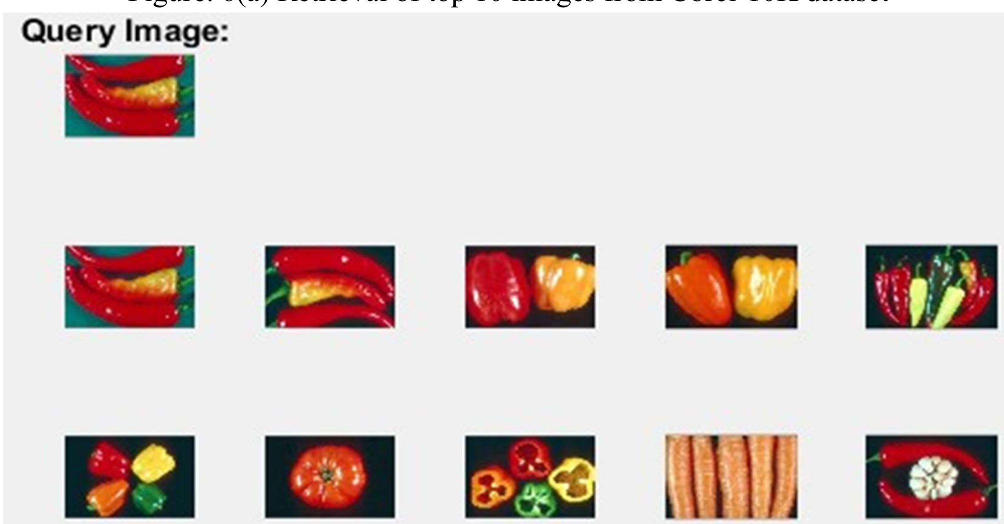

Figure. 6(b) Retrieval of top 10 images from Oxford flower dataset

Therefore, from these retrieval results, it can be concluded that the images from the same native category of the query image are recovered.

Conclusion and Future Trends

This paper depict an innovative technique initially for the formation of a hybrid CBIR system which is formed by an amalgamation of two color descriptors with a texture extraction technique namely Color moment (CM) and Color auto-correlogram (CAC) with Gray level co-occurence matrix (GLCM) respectively. Then, the results from the hybrid system are applied to the first level of classification, where two classifiers namely Extreme learning machine (ELM) and Support vector machine (SVM) are placed in parallel. If the results obtained from both the classifiers matches to each other, then a valid output is obtained. If there is an ambiguity, then, the result is passed to second level of classifier which consists of three ELM's arranged in a sequential manner. After passing through the second level, an unambiguous image is obtained which is matching image to the query image. Our future work will be focused on developing a hybrid CBIR system by the usage of deep learning techniques like auto-encoders, deep belief networks, CNN, etc. There are many real-life multimedia applications like medical diagnosis, crime detection for the extraction of fingerprints, face recognition, pattern recognition etc, where this system can be utilized. Last but not the least, the concept of Internet of things (IoT) can be used for the online transfer of desired images and for capturing real life data.

References

[1] N. Varma, A. Mathur, 2020, "A Survey On Evaluation of Similarity Measures forContent-based Image Retrieval Using Hybrid Features", IEEE International Conference on Smart Electronics and Communication, pp. 978-1-7281-5461-9.

[2] Y. Mistry, D. T. Ingole, M. D. Ingole, 2018, "Content-based image retrieval using hybrid features and various distance metric", Journal of Electrical Systems and Information Technology, vol. 5, pp. 874–888.

[3] K. Bharathi, Dr. M. C. Mohan, 2017, “Content-Based Image Retrieval: An Overview of Architecture, Challenges and Issues”, International Journal of Engineering Research in Computer Science and Engineering (IJERCSE), vol. 4, iss. 12, pp. 2394-2320.
[4] S. Shinde, M. Desai, M. Gupte, D. Racha, 2018, “Content-Based Image Retrieval: Using Multi-Sequential Search”, IEEE 2nd International Conference on Inventive Communication and Computational Technologies, pp. 978-1-5386-1974-2.
[5] N. Shrivastava, V. Tyagi, 2014, “An efficient technique for retrieval of color images in large databases”, Computers and Electrical Engineering, pp.1-9.
[6] J. X. Zhou, X. D. Liu, T. W. Xu, J. H. Gan, W. Q. Liu, 2016, “A new fusion approach for content-based image retrieval with color histogram and local directional pattern”, Int. J. Mach. Learn. & Cyber, pp.1-12.
[7] S. Singh, S. Batra, 2020, “An efficient bi-layer content-based image retrieval System”, Multimedia Tools and Applications, pp. 1-13.
[8] L. K. Pavithra , T. S. Sharmila, 2017, “An efficient framework for image retrieval using color, texture, and edge features”, Computers and Electrical Engineering, pp. 1-14.
[9] D. Kishore, C. S. Rao, 2020, “A Multi-class SVM Based Content-Based Image Retrieval System Using Hybrid Optimization Techniques”, International Information and Engineering Technology Association, vol. 37, pp. 217-226.
[10] L. Putzu, L. Piras, G. Giacinto, 2020, “Convolutional neural networks for relevance feedback in content-based image retrieval”, Multimedia Tools and Applications, vol. 79, pp. 26995–27021.
[11] S. Bhardwaj, G. Pandove, P. K. Dahiya, 2020, "A Genesis of a Meticulous Fusion based Color Descriptor to Analyze the Supremacy between Machine Learning and Deep Learning", International Journal of Intelligent Systems and Applications (IJISA), vol. 12, no. 2, pp. 21-33.
[12] S. Fadaei, A. Rashno, 2020, “Content-Based Image Retrieval Speedup Based on Optimized Combination of Wavelet and Zernike Features Using Particle Swarm Optimization Algorithm”, International Journal of Engineering, vol. 33, iss. 5, pp. 1000-1009.
[13] S. Dhingra P. Bansal, 2020,“Experimental analogy of different texture feature extraction techniques in image retrieval systems”, Multimedia Tools and Applications, pp. 79:27391–27406.
[14] G. H. Liu, Z. Wei, 2020, “Image Retrieval Using the Fused Perceptual Color Histogram”, Computational Intelligence and Neuroscience, pp.1-7.
[15] A. Khokher, R. Talwar, 2016, “A fast and effective image retrieval scheme using color, texture, and shape-based histograms”, Multimed Tools Appl, pp. 1-16.
[16] P. Saxena, 2021, “An Overview of Feature Extraction Techniques in Content-Based Image Retrieval Systems”, International Journal of Current Engineering and Technology, vol. 11, no. 1, pp. 2347 – 5161.
[17] R. Rajkumar, M. V. Sudhamani, 2019, “Retrieval of Images using Combination of Features as Color, Color Moments and Hu Moments”, Advances in Image and Video Processing, vol. 7, no. 5, pp. 9-21.
[18] S. Bhardwaj, G. Pandove, P. K. Dahiya, 2020, “An Extreme Learning Machine-Relevance Feedback Framework for Enhancing the Accuracy of a Hybrid Image Retrieval System”, International Journal of Interactive Multimedia and Artificial Intelligence, vol. 6, no. 2, pp. 1-12.

[19]L. K. Pavithra, T. S. Sharmila, 2019, “An Improved Seed Point Selection-Based Unsupervised Color Clustering for Content-Based Image Retrieval Application”, Computational Intelligence, Machine Learning and Data Analytics, The Computer Journal, pp. 1-14.

[20] S. Dhingra, P. Bansal, 2019, “An Intelligent Multi-Resolutional and Rotational Invariant Texture Descriptor for Image Retrieval Systems”, ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, vol. 8, no. 2, pp. 33-49.

[21] D. Soni, K. J. Mathai, 2015, “An Efficient Content-Based Image Retrieval System based on Color Space Approach Using Color Histogram and Color Correlogram”, IEEE 2015 Fifth International Conference on Communication Systems and Network Technologies, pp. 1-14.