IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 1, 2020, accepted June 17, 2020, date of publication June 22, 2020, date of current version July 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3003911

Image Retrieval Scheme Using Quantized Bins of Color Image Components and Adaptive Tetrolet Transform

NAUSHAD VARISH[1], ARUP KUMAR PAL[2], (Member, IEEE), ROSILAH HASSAN[3], (Senior Member, IEEE), MOHAMMAD KAMRUL HASAN[3], (Senior Member, IEEE), ASIF KHAN[4], NIKHAT PARVEEN[1], DEBRUP BANERJEE[1], VIDYULLATHA PELLAKURI[1], AMIN UL HAQIS[4], AND IMRAN MEMON[5]

[1]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (KLEF), Guntur 522502, India
[2]Department of Computer Science and Engineering, Indian Institute of Technology (ISM) Dhanbad, Dhanbad 826004, India
[3]Center for Cyber Security, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), Bangi 43600, Malaysia
[4]School of Computer Science and Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China
[5]Department of Computer Science, Bahria University—Karachi, Karachi 75260, Pakistan

Corresponding authors: Naushad Varish (naushad.cs88@gmail.com) and Asif Khan (asifkhan@uestc.edu.cn)

This paper supported was under the Dana Impak Perdana (DIP) Grant Scheme DIP-2018-040, University Kebangsaan Malaysia (UKM).

ABSTRACT In this paper, a three stage hierarchical image retrieval scheme using a color, texture and shape visual contents (or descriptors) is proposed, since single visual content is not produce an adequate retrieval results effectively. This scheme has reduced the searching space during the image retrieval process at a certain extent due to the hierarchical mode. In initial stage, the shape feature descriptor has been computed by simple fusion of histograms of gradients and invariant moments of segmented image planes. The shape based retrieval process has reduced the search space by discarding the non-relevant images from the universal dataset (or original dataset) effectively and kept the retrieved images into the intermediate dataset. In the second stage, the texture feature descriptors have been computed from the intermediate sub-image dataset by applying the adaptive tetrolet transform on image plane of preprocessed HSV color image. This transform provides the multi-resolution images with finer details by employing the tetrominoes and the proper arrangement of tetrominoes contributes the effective local geometry of image plane. The gray level co-occurrence matrix based texture feature descriptor is obtained by computing second order statistical parameters from each decomposed sub-image. At this stage, the most of the irrelevant images are discarded by retrieving the images from intermediate dataset but still some undesired images are left, those will be handled at the last stage. At this stage, fused color information is captured by applying the color autocorrelogram on both the non-uniform quantized color components of the preprocessed HSV color image. Finally, the color feature descriptor produces the desired retrieval results by discarding the irrelevant images from the texture based sub-image dataset. The proposed scheme has also low computational overhead due to the use of three descriptors at different stages separately. The retrieved results show the better accuracy as compared to the other related visual contents based image retrieval schemes.

INDEX TERMS Image retrieval, adaptive tetrolet transform, gray level co-occurrence matrix, statistical moments.

I. INTRODUCTION

Content-based Image Retrieval (CBIR) [1], [2] has become proficient research area from the last two decades due to the rapid advancement of multimedia data (i.e image, video and sound) through different sources like social networking sites, high speed Internet, smart phones and image capturing devices. The multimedia data like images are commonly used, shared and accessed among several users which leads to increase the size of the digital image repositories or databases

The associate editor coordinating the review of this manuscript and approving it for publication was Maurizio Tucci.

continuously. Therefore, searching and retrieving of such kinds of images from large digital repositories is one of the tedious task. Text based image retrieval scheme is unable to produce an efficient results to the user because images are annotated or saved based on their texts, keywords or tag numbers. This annotation process is clumsy and time consuming for large scale image databases due to the human involvement [3]. Hence, to overcomes the flaws and limitations, CBIR is an alternative solution, where searching and retrieving of the desired images from the entire digital repository is done based on visual descriptors like color, texture and/or shape of the given query image. The most of the earlier CBIR schemes have been developed based on either color, texture or shape contents/descriptors [4]. However, the single feature descriptor is not always sufficient to provide best retrieval results. Therefore, some researchers have clubbed more than one image features in some proper order or simultaneously for improving the retrieval accuracy. The selection of appropriate image features are also important in CBIR scheme because it affects the image retrieval performance.

Since in traditional CBIR schemes, the similarity distances or measures have been computed between the color, texture and shape visual feature descriptors of images individually and a single similarity measure using fusion or normalization approaches have been determined by assigning appropriate weights to the particular distance of specific feature descriptor [5]. Assignment of the proper proportionate of weights to the distances for these three visual contents descriptors is tedious work and it also requires high computational overhead to decide a suitable distance weight during image retrieval matching. Moreover, the traditional CBIR scheme searches the most desired images in the whole image database using the visual feature descriptor of the query image. Therefore, in this work, authors have designed and developed a hierarchical CBIR schemes using visual feature contens i.e., color, texture and shape in three stages and it reduces the search space of database at certain level. The main contribution of the proposed image retrieval scheme is highlighted as:

1. A three stage hierarchical image retrieval scheme is proposed using three different kinds of visual feature descriptors i.e., color, texture and shape. It reduces the computational overhead and search space of an image dataset due to the adaptation of three feature descriptors in hierarchical mode.
2. In the first stage, shape visual feature descriptor is computed by simple fusion of two kinds of shape representatives i.e., hue invariant moments and histograms of gradients (HOGs) contents of the preprocessed image. At this stage, the top desired retrieved images from universal dataset are stored in new dataset based on the shape visual feature descriptors of the query image and universal dataset images using suitable distance values. This new intermediate dataset would be considered as a input for the next stage.
3. In the second stage, the textural visual feature descriptors have been computed using tetrolet transform and GLCM contents of the preprocessed images of the new intermediate dataset. Subsequently, the top most images have been retrieved from the intermediate dataset by discarding the some available non-relevant images and kept these images in another dataset. But, still this dataset has few non-relevant image, those will be handled in the final stage.
4. In the final stage, color visual feature descriptor is obtained by simple fusion of autocorrelogram based two color visual contents of the preprocessed color image. Based color descriptors, the most desired images have been extracted from the second stage image dataset by discarding the non-relevant images. These retrieved images are the final outcomes of the proposed CBIR system.
5. In each stage, the retrieval of images is performed using minimum computed Euclidean distances between the visual feature descriptors of target images and query image. The experimental results are tested for various possible orders of stages i.e., color(C), texture(T) and shape(S) visual feature descriptors), where the satisfactory retrieval results have been achieved in most of the cases.

The rest parts of paper is arranged as follows: Section 2 presents the related works, Section 3 describes the fundamental behind CBIR schemes. Section 4 elaborates the proposed CBIR methodology. In Section 5, simulation results and discussions are demonstrated. Finally, Section 5 provides the conclusion of the paper.

II. RELATED WORKS

In this Section, we will present the literature survey of related CBIR schemes based on visual feature descriptors one after another.

A. SHAPE VISUAL FEATURE DESCRIPTORS

Shape is one of important features used in CBIR applications since it is significantly finds the object region in the image. The image retrieval scheme is basically carried out using the semantic visual contents which can be easily recognizable by shape features. It also provides much accurate results if the images of the databases consist of object, various shapes, different structures, more edges in many directions etc. El-ghazal *et al.* [6] have developed shape based image retrieval scheme, where shape signatures have been extracted based on the Fourier descriptor (FD) and farthest point distance (FPD) technique. In this scheme, the shape signatures have been computed at each point on a shape contour and it also achieved the scale, translation invariant properties of image which further improves the retrieval accuracy. But during acquiring the invariant properties, some valuable information have been lost. Hence another shape based image retrieval scheme has been suggested by Sokic *et al.* [7], [8] which overcomes the problem of existing scheme [6] and

it also preserves the invariance properties of image. They have adopted only the phase of Fourier coefficients and it has been used for the specific points (or pseudomirror points) as a shape orientation reference. The shape signature is also invariant under translation, scaling and rotation due to the phase-preserving Fourier descriptors. Wu *et al.* [9] have computed Tchebichef moments based on Tchebichef polynomials and power series. The obtained Tchebichef moments have translation and scale invariant properties of image. The selected invariant moments have constructed the shape feature vector for the shape based image retrieval application. Srivastava *et al.* [10], [11] have suggested an image retrieval scheme, where the shape moments have been extracted from the details sub-band images of discrete wavelet transform (DWT). In this scheme, several local binary pattern (LPB) images at each resolution of DWT are computed and subsequently moments of each LBP sub-band images for devising image retrieval scheme. Pradhan *et al.* [12], [13] have employed graph based visual saliency map to identify the object region and subsequently the feature vector is constructed from object region. In their scheme, it gives adequate retrieval results for object based images but it fails to provide satisfactory results for other category images. So in general, CBIR schemes based on shape features/contents are simple and faster but alone shape features have not enough capacity to identify image contents with different structures and various objects.

B. TEXTURE VISUAL FEATURE DESCRIPTORS

Texture features are another significant visual contents which has vital role play for developing CBIR scheme because it has the powerful discriminating property to distinguish the image contents. Textures are referred as pattern of pixels for representing the objects and region of interest (ROI) of images. and it provides some important properties of image such as smoothness, regularity, homogeneity, coarseness, etc. Wang *et al.* [14], [15] have developed CBIR scheme using texture visual contents, where the texture visual contents have been extracted from each region with 8-connectivity using co-occurrence matrix. Manthalkar *et al.* [16] have suggested an image retrieval scheme using wavelet packet transform, where the texture feature descriptor is computed by collecting the means and standard deviations of vertical and horizontal decomposed multi-resolution images at each level. In this work, the rotation and scale invariant properties are also validated for 12 directions and 5 scales in standard texture image database. Rakvongthai *et al.* [17] have computed texture feature descriptor in the noisy environment based on the complex wavelets using the statistical values where each sub-band coefficients are modeled by the standard statistical distribution. To improve the retrieval accuracy, they have combined the magnitude and phase information of complex sub-band coefficients. Krishnamoorthi *et al.* [18] have considered orthogonal polynomials model, where the coefficients are re-arranged into subband images and some statistical values have been determined from all decomposed sub-band images for the formation the texture feature descriptor. The above discussed texture based image retrieval schemes have been developed using the traditional wavelet transforms. The 2D-DWT is a separable and it is applied on each row followed by columns independently. Therefore, DWT achieves the horizontal, vertical and diagonal information of the image whereas it fails to extract the optimal information of image after certain decomposition levels. The DWT is also unable to provide the significant image features when the images have more local structures and geometric shapes in different orientations. Some image feature extraction techniques based on the wavelets [19], [20] have been developed for optimal image representation with more directional sensitivity, but they have some drawbacks. To overcome the drawbacks, in presented work, authors have considered the ATT [21] which represents the image in a very optimum way, where the tetrolets are derived based on the Haar wavelets. Tetrolets represent tetrominoes which is nothing but the local shapes of image. Raghuwanshi *et al.* [21], [22] have suggested tetrolet transform based image retrieval scheme, where they have constructed the texture feature descriptor by computing the means and standard deviations from detailed sub-band images upto 4 levels, but the direct computation of statistical values lacks the spatial information among tetrominoes of the sub-images. Hence in our presented work, we have considered co-occurrence matrix which provides the spatial relationship between the pixel values of tetrominoes at multiple decomposision levels and low dimensional texture feature descriptor have been constructed by computing the second order statistical values from each co-occurrence matrix of tetrolet based sub-images. The shape and texture visual descriptors together provide the better retrieval results for shape texture image retrieval databases. But, it does not provide always good retrieval results, Hence, third low level visual content i.e., color have been incorporated to enhance the performance of the proposed image retrieval scheme [23], [23].

C. COLOR VISUAL FEATURE DESCRIPTORS

In many CBIR applications, color visual contents are generally used to browse the most relevant images from the digital repository. In color visual features/contents are invariant with image scaling, orientations of objects, regions of interest and other pixel transformations. In the literature review, some CBIR schemes [24], [25] have been developed using color information based on histogram, dominant color descriptor, correlogram and color coherent vector techniques. Wang *et al.* [26] have initially converted an RGB color space into YCbCr color image and have found the image feature points using Harris-Laplace detector. These points have represented the local feature regions (LFRs) of an image and subsequently, each LFR is uniformly quantized. Finally, quantized histogram of LFRs has been used as a color feature descriptor to develop a CBIR scheme. Liu *et al.* [27] have extracted a color information using

color difference histogram (CDH) technique, where, color components of L*a*b color space are quantized into uniform bins and an edge orientations are also quantized into uniform bins. Thereafter, simultaneously CDH features have been extracted from obtained edge orientation image and quantized L*a*b image components. Some other color histogram based CBIR schemes have been found in [28], [29] but the histogram based color information is unable to exploit the spatial relationship among the colors (or pixel values). To find the spatial relationship among the pixel values, color autocorrelogram [1] is an alternative solution which overcomes the limitations of the conventional color based approaches. So, in the presented paper, the authors have adapted color autocorrelogram technique for extracting some significant color information from the HSV color image.

D. HYBRID VISUAL FEATURE DESCRIPTORS

The most of the previously discussed image retrieval schemes have been developed using three visual feature descriptors i.e., color, texture and shape alone but, only single descriptor does not carry much information of image, which indicates that it is not good enough to produce the significant retrieval results in large and complex datasets. Therefore, to enhance the retrieval accuracy, several image retrieval schemes have been developed using combinations of two or three visual feature descriptors together are found in literature survey. Seetharaman *et al.* [1] have computed the color and texture visual contents from the optimal decomposed sub-images of DWT. They have extracted color information using color autocorrelogram technique from two quantized color components of HSV color image and the texture visual contents using co-occurrence matrix are extracted by computing the statistical values from Value image plane of HSV color image. Rahimi *et al.* [30] have designed image retrieval scheme using color and texture information, where they determined spatial co-relation with the pixel values of red, green and blue color components using Color Ton (DCTon) technique whereas the texture information using DT-CWT and SVD tools is obtained by dividing image plane into sub-images manually. This scheme is not efficient because manual image segmentation is feasible. A number of image retrieval schemes using concatenation of all three visual feature descriptors i.e., color, texture and shape together are also found in literature review. Wang *et al.* [31] have also integrated of color, texture and shape visual information simultaneously for developing a CBIR scheme, where color information has been extracted from 8 coarse partitions/regions of RGB color image. The texture visual feature descriptor has been computed by taking the energies and standard deviations from the steerable filter decomposed sub-images at 4 directions. Lastly, shape visual feature descriptor is computed by collecting the scale and translation invariant Pseudo-Zernike moments. Similarly, Huang *et al.* [32] have developed image retrieval scheme using color, texture and shape visual feature descriptors, where color visual contents have been extracted from an RGB and HSV color images. The shape visual contents have been obtained by combining two types of zernike moments of the image plane. Lastly, texture visual contents are computed using second order statistical parameters of the co-occurrence matrix of image plane. Shrivastava *et al.* [33] have used all three visual feature descriptors, where color and texture visual feature descriptors have been extracted using the modified LBP histogram of quantized color image and gray scale image respectively. The shape visual feature descriptors have been extracted based on the binary edge map of each block using Sobel edge detector operator and the combined three visual feature descriptors represent a single feature descriptor. The combined visual feature descriptors have been used in retrieval of desired images from digital repository.

III. BASIC CONCEPTS

This Section presents the basic concepts like color autocorrelogram, Adaptive Tetrolet Transform (ATT), co-occurrence matrix and the histogram of gradients. In the proposed CBIR scheme, these concepts play important role for extracting significant visual feature descriptors.

A. COLOR AUTOCORRELOGRAM

A color autocorreloram [1] is a special case of color correlogram technique [34] which provides the spatial co-relation between the similar color values at a particular distance while the color correlogram provides the spatial correlation among all combination of distinct pixel pair values. So color autocorreloram based visual feature descriptor performs well and has low dimension as compare to the correlogram descriptor. Let us consider I be the color image of size $n \times n$ with pixel position $p(x, y)$ and it consist of color set C having M number of colors. The color set C is quantized into different number of colors $C_i, i = 1, 2 \ldots m, m < M$. The fixed distance k between two colors C_1 and C_2 is computed by $|p_1 - p_2| = \max\{|x_1 - x_2|, |y_1 - y_2|\}$, where $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ are the pixel positions of the corresponding colors in the image. For extraction of color visual information, color correlogram is calculated as

$$\gamma^{(k)}_{C_i, C_j}(I) \triangleq \Pr_{p_1, p_2 \in I} \left\{ (p_1, p_2) \, | p_1 \in I_{C_i}, p_2 \in I_{C_j} \, | p_1 - p_2 | = k \right\} \tag{1}$$

where $(i, j) \in \{1, 2, \ldots, m\}$, element $\gamma^{(k)}_{C_i, C_j}(I)$ represents the probability of pixels with two different colors C_i and C_j at a fixed distance k. The length of color correlogram feature descriptor is $m^2 \times K$ where K denotes the total number distinct distances. Then color autocorrelogram of I image is computed as

$$\alpha^{(k)}_{C_i}(I) \triangleq \gamma^{(k)}_{C_i, C_i}(I) \tag{2}$$

where $\alpha^{(k)}_{C_i}(I)$ is probabilities of pixels between two similar colors C_i in image and it represents low dimensional color visual feature descriptor.

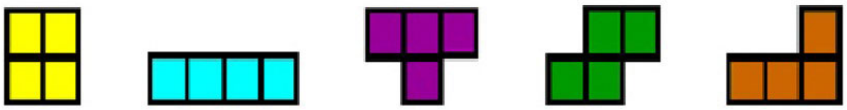


FIGURE 1. Five free tetromeno patterns.

B. TETROLET TRANSFORM

The traditional 2D wavelet transforms are separable and they are employed on each row followed by each column independently by using pair of low and high pass filters. Therefore, wavelet transform such as DWT which achieves the horizontal,vertical and diagonal information of the image but it fails to extract the optimal/local information of image after certain decomposition levels when images have more local structures and geometric shapes in different orientations. Some image classification methods based on several wavelets [19], [35] have been developed for optimal image representation with more directional sensitivity. The Adaptive Tetrolet Transform(ATT) [36] overcomes the problems of the above said transformation tools. Therefore, in this paper, authors have considered the ATT which represent an image an optimally where the tetrolets are derived based on the Haar wavelets. Tetrolets represent tetrominoes which is nothing but the geometric shapes of the images. The concept of tetrolets has been introduced by the Krommweh [36] while Golomb *et al.* [37] suggested the idea of tetrominoes. Tetrominoes are geometric shapes those are formed by taking the union of four unit squares where each unit represent the pixel values and edges of each are connected with each other. The five various shapes are depicted in Fig 1 with respect to their rotations and reflections characteristics. For more detail, let input image is a^{r-1} and it is divided into $Q_{i,j}$, blocks of 4×4, $i, j = 0, 1, 2, 3$. For each block, Qi, j considering all 117 solutions and for each tiling c, 4 low pass coefficients and 12 high pass coefficients have been computed. Thereafter, high pass and low pass coefficients are re-arranged into an 2×2 blocks and stored both coefficients to form the subband components of image a^{r-1}. Apply tetrolet transform on approximate image(low subband component) recursively without changing high pass components/detailed image. The approximate image decomposed at each level is computed as

$$\begin{aligned} A^{r,(c)} &= (a^{r,(c)}[s])_{S=0}^{3} \\ a^{r,(c)}[s] &= \sum_{(m,n)\in I_s^{(c)}} \in [0,\ L(m,n)]a^{r-1}[m,n] \end{aligned} \tag{3}$$

For next level decomposition, only low pass sub image is adopted while the high pass sub images have kept as it is, those are considered for texture analysis. The three detailed sub-images at each decomposition level is computed as

$$\begin{aligned} W_l^{r,(c)} &= (w^{r,(c)}[s])_{s=0}^{3} \\ w^{r,(c)} &= \sum_{(m,n)\in I_s^{(c)}} \in [l,\ L(m,n)]a^{r-1}[m,n] \end{aligned} \tag{4}$$

Haar wavelet transform matrix W has four fixed 2×2 squares with 117 solution for disjoint covering of a 4×4 board.

$$W = (\in [m,n])_{m,n=0}^{3} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \tag{5}$$

where L is the bijective function which is applied for mapping the four index pairs (m, n) of $I_s^{(c)}$ with the values 0, 1, 2 and 3 in unique fashion i.e. decreasing order. At each, decomposition level, the function L assigns values from 0 to 3 to high and low pass sub images. The four tetrominoes I_0, I_2, I_2, I_3, subset of $I_s^{(c)}$ are mapped by applying the mapping function L into a unique order (0, 1,2,3). The wavelets based on haar function decomposed image into several number of fixed sized blocks which lacks to describe the local geometric pattern/shapes of the image. The tetrolet transform covers the much directional sensitivity where the local geometric shapes have been extracted significantly. As for as shapes of image are concern, tetrolet transform automatically captures the local geometric patterns and positions of the patterns which is not possible in most of the available transformation tools. The haar wavelet coefficients of a 4×4 block is shown in Fig. 2 **(a)** while the 117 solutions of tiling with any four out of five free geometric patterns/shapes with their local structure are depicted in Figs. 2**(b)-(c)** respectively.

C. GRAY LEVEL SPATIAL DEPENDENCE MATRIX

The Gray Level Spatial Dependence Matrix is also known as Gray Level co-occurrence matrix(GLCM) and it is introduced by Haralick *et al.* [38] in 1973. The GLCM is computed by transforming a image into a small matrix based on the spatial relationship among the pair of pixel values of the original image. The mutual occurrence of the pair of pixel values at particular orientations (specially horizontal vertical and diagonal) have been computed for forming the dependence matrix. Thereafter several statistical values from the each GLCM are computed for texture classification [19] of the images. Let us consider an image X of size $N_x \times N_y$ and their corresponding GLCM matrix is defined as $G_D^{\theta} = \left[g_D^{\theta}(i,\ j)\right]_{Q\times Q}$ where D and θ (in radian) are the distances and orientations/directions respectively, Q represents the quantization level of the original image. The values of the co-occurrence matrix with four directions and fixed distance D are computed as follows:

$$\begin{aligned} &g_D^0(i,j) \\ &= \#\left\{((a,b)((c,\ d)) \,\middle|\, \begin{matrix} a-c=0,\ |b-d|=D, \\ X(a,\ b)=i,\ X(c,\ d)=j \end{matrix}\right\} \end{aligned}$$

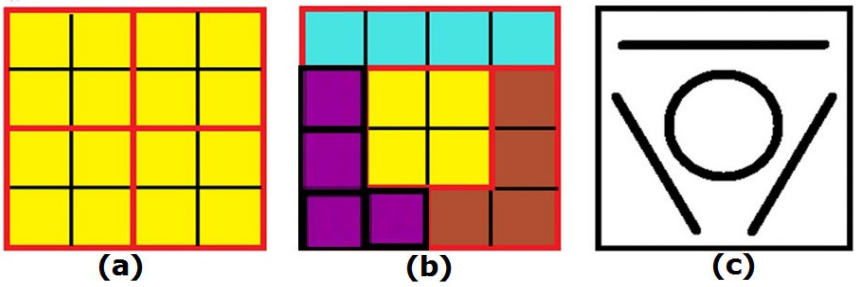


FIGURE 2. (a) The coefficients of the 2D Haar wavelets in fixed squares; (b) One of the disjoint covering from 117 kinds of tilings. (c) The local structure of patter.

$$(a)\quad \begin{pmatrix} 0 & 1 & 3 & 2 \\ 1 & 3 & 1 & 0 \\ 3 & 0 & 1 & 2 \\ 0 & 2 & 0 & 3 \end{pmatrix} \qquad \begin{pmatrix} \#(0,0) & \#(0,1) & \#(0,2) & \#(0,3) \\ \#(1,0) & \#(1,1) & \#(1,2) & \#(1,3) \\ \#(2,0) & \#(2,1) & \#(2,2) & \#(2,3) \\ \#(3,0) & \#(3,1) & \#(3,2) & \#(3,3) \end{pmatrix}$$

$$(b)\quad \begin{pmatrix} 0 & 3 & 2 & 2 \\ 3 & 0 & 1 & 3 \\ 2 & 1 & 0 & 1 \\ 2 & 3 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0.1250 & 0.0833 & 0.0833 \\ 0.1250 & 0 & 0.0417 & 0.1250 \\ 0.0833 & 0.0417 & 0 & 0.0417 \\ 0.0833 & 0.1250 & 0.0417 & 0 \end{pmatrix}$$

$$(c)\quad \begin{pmatrix} 0 & 2 & 3 & 2 \\ 2 & 2 & 0 & 3 \\ 3 & 0 & 0 & 1 \\ 2 & 3 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0.0833 & 0.1250 & 0.0833 \\ 0.0833 & 0.0833 & 0 & 0.1250 \\ 0.1250 & 0 & 0 & 0.0417 \\ 0.0833 & 0.1250 & 0.0417 & 0 \end{pmatrix}$$

$$(d)\quad \begin{pmatrix} 2 & 2 & 1 & 0 \\ 2 & 2 & 2 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \qquad \begin{pmatrix} 0.1111 & 0.1111 & 0.0556 & 0 \\ 0.1111 & 0.1111 & 0.1111 & 0 \\ 0.0556 & 0.1111 & 0 & 0 \\ 0 & 0 & 0 & 0.2222 \end{pmatrix}$$

$$(e)\quad \begin{pmatrix} 2 & 1 & 0 & 2 \\ 1 & 2 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 2 & 2 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0.1111 & 0.0556 & 0 & 0.1111 \\ 0.0556 & 0 & 0.0556 & 0.1111 \\ 0 & 0.0556 & 0 & 0.0556 \\ 0.1111 & 0.1111 & 0.0556 & 0 \end{pmatrix}$$

FIGURE 3. Computation of GLCM and their corresponding Normalized GLCM with distance D = 1 (a) Grayscale matrix of four values ranging from 0-3 and its corresponding GLCM; Different GLCM and their normalized GLCM with (b) Horizontal $\theta = 0$ (c) Vertical $\theta = \pi/2$ (d) Right diagonal $\theta = \pi/4$ (e) Left diagonal $\theta = 3\pi/4$ directions respectively.

$$\begin{aligned} &g_D^{\pi/4}(i,j) \\ &= \#\left\{((a,b)((c,\ d))\ \middle|\ \begin{array}{l} a-c=D,\ |b-d|=D \\ or\ |a-c|=-D,\ |b-d|=-D, \\ X(a,\ b)=i,\ X(c,\ d)=j \end{array}\right\} \\ &g_D^{\pi/2}(i,j) \\ &= \#\left\{((a,b)((c,\ d))\ \middle|\ \begin{array}{l} |a-c|=D,\ b-d=D, \\ X(a,\ b)=i,\ X(c,\ d)=j \end{array}\right\} \\ &g_D^{3\pi/4}(i,j) \\ &= \#\left\{((a,b)((c,\ d))\ \middle|\ \begin{array}{l} a-c=D,\ b-d=-D \\ or\ a-c=-D,\ b-d=D, \\ X(a,\ b)=i,\ X(c,\ d)=j \end{array}\right\} \end{aligned} \tag{6}$$

where $X(a,b)$ and $X(c,d)$ are pixel values of the image at positions (a, b) and (c, d) for all $(a,b)(c,\ d) \in N_x \times N_y$. The computation of different GLCMs and their normalized GLCMs of 4×4 grayscale matrix with distance $D = 1$ are depicted in Fig. 3, where the grayscale values are ranging from 0-3. The size of the GLCM depends on the quantized level of the image. In this research work, the authors have quantized the image into 8 gray levels for the simplicity of computation because the images in nature are huge in size due to the availability of modern image capturing devices and fast Internet technology. Hence the size of GLCM is 8 which is too small as compared to original size of the image. Then the conditional/co-occurrence probability can be computed as

$$P_r(x) = \left\{C_{ij}\ |(D,\ \theta)\right\}$$

where *Cij* (NGLCM), between gray levels i and j is defined as:

$$C_{ij} = \frac{g_{ij}}{\sum_{i,j=1}^{G} g_{ij}}$$

The variable g_{ij} represent the elements of the GLCM and $\sum_{i,j=1}^{G} g_{ij}$ represents the sum of elements. The different GLCMs are computed for various values of (D, θ). The distance (D) between the pixels are normally chosen as less than or equal to 10 and four orientations (i.e. $\theta = 0, \pi/4, \pi/2, 3\pi/4$ radians) for the easy computation. Some statistical parameters like contrast, correlation, energy, homogeneity are computed for different values of angles and distances for discriminating the texture features where the scheme [38] have validated that the average of these statistics in various orientations along with distances are capable to classify the image texture properties.

D. ACTIVE CONTOUR AND HISTOGRAM OF GRADIENTS

Chan and Vese [39] have suggested a technique based on restricted Mumford-Shah model [40] for segmentation which divides image into foreground(object) and background parts. It is a region based segmentation model where image intensities lies inside and outside of closed curve. Let X and C are input image and the closed curve respectively, then the energy function is defined as

$$E^{CV}(C, c_1, c_2) = \lambda_1 \int_{out(C)} |X(x) - c_1|^2 dx + \lambda_2 \int_{in(C)} |X(x) - c_2|^2 dx + u.length(C) \quad (7)$$

where $X(x)$ represents the intensity value of image at a point x, The terms λ_1, λ_2 and u are positive constants. The constants c_1 and c_2 are used for approximating the image intensity values inside and outside of the closed curve C respectively. In equation (7), the first and second terms have been computed the curve (contour) C to find the desired boundaries of the objects(ROIs) which means that these two data (internal energies) controls the smoothness property of the contour C while the third term(external energy) represents the euclidean length which is used to compute the sharp contour. The energy E^{CV} must be minimum if the contour C is situated at the boundaries of ROI. The image will be wrongly segmented if the pixel values of the inside and outside contour are not similar since the constants c_1 and c_2 represent the global characteristic of pixel values of the image. In the presented work, the grayscale image is segmented based on above discussed model and it is separated into foreground and background images individually. Moreover, some images have the prominent foreground and low background, where foreground part carries most significant information and background part carries less significant information and vice-versa. Hence, it is difficult to describe background and foreground parts of the image significantly, specially in case of large and complex image databases. Therefore, we have also considered background and foreground parts/images simultaneously for the formation of the feature vector.

For significant shape feature extraction, the histogram of oriented gradients (HOGs) [41] of foreground image have been computed. The HOGs are significant image descriptors which are rotation invariant and have been used in several image recognition applications such as face detection [42], pedestrian detection [43], image retrieval scheme [44], etc.. The previous HOG features based methods [41], [45] have counts the occurrences of the gradient directions or edge orientations of a fixed size small spatial regions (rectangular blocks or cells) of the image window. In this paper, the histogram of gradients of 50.00 % overlapping rectangular blocks(cells) of the foreground image have been computed because overlapping blocks provides the more significant details of image than non-overlapping blocks. Here, overlapping blocks means that an each of them performs more than once than for the construction of the HOG feature descriptor/vector. Afterwards the background image is adopted for further processing because it also carries some prominent image information. Here, another shape based features from background image using hue invariant moments [46] have been computed. The hue moments is one of the most successful approach for computing an image visual features for object recognition [47] based applications like face expression [48], hand gesture [49], speech emotion recognition [50]. Since background image contains number of small objects and they are not always visualize with respect to their location, size and directions. Therefore, hue invariant moments [49] are adopted for shape features extraction process. The single shape feature vector is obtained by proficient combination of two features effectively in presented paper.

IV. PROPOSED CBIR SCHEME

In this section, low level visual features such as color, texture and shape of the image are computed one after another. Since images in nature are not accurate and vivid due to the light condition, poor illumination and/or high temperature, camera motion, analog-to-digital converter errors etc. So, they have some bad quality and blurred pixels. Therefore, in this paper, authors have performed some preprocessing techniques on images before direct extracting the visual features. For shape features extraction, initially RGB color image is converted into grayscale image and then it is segmented based on the active contour model which has been discussed in Section III-D. Further, the segmented grayscale image is separated into foreground and background images/parts. For the formation of the shape feature vector, HOGs features of the foreground part and hue moments of background part are computed. For the color and texture features extraction, an RGB color is preprocessed using Laplacian operator and converted into HSV color image. The conversion and sharpening process of HSV color image is depicted in Figure 4. For the texture feature extraction, the V component of HSV color image is considered, where tetrolet transform has been employed to obtain the sub-band images and applied the corresponding GLCM tools to the each tetrolet decomposed sub-images. For forming the texture feature vector, the second

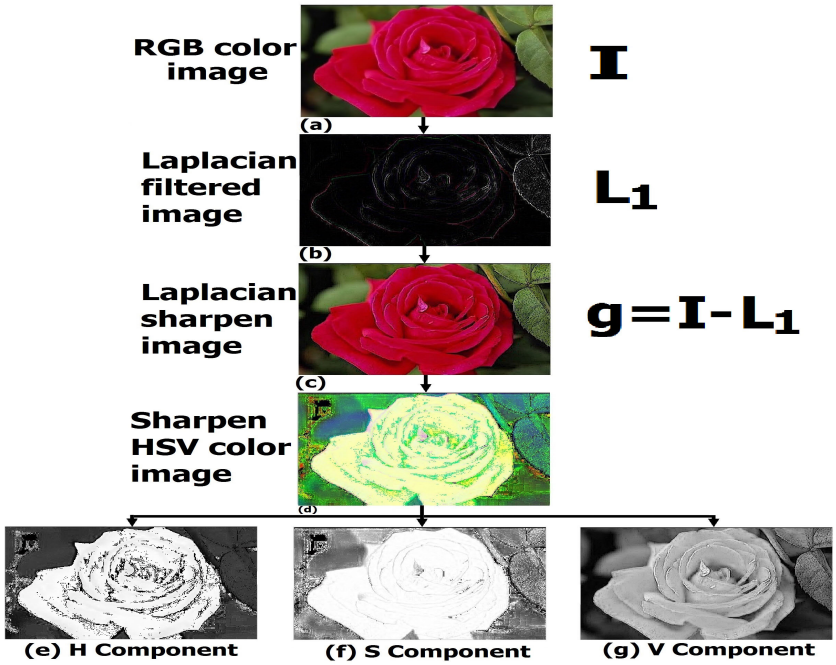


FIGURE 4. Preprocessed components of HSV color image.

order statistical parameters like contrast, correlation, angular moment and diagonal distribution of each GLCM are computed. After that, the color features are directly extracted from the non-uniform quantized H and S components of sharpened HSV color image where the quantization of each color component is done by Lloyd algorithm [51]. For forming the color feature vector, the color autocorrelogram of the preprocessed H and S components has been computed.

Let D be the universal image dataset having N number of images and dataset R_S is a set of K number of images those are retrieved from an universal database based on the shape features. This R_S dataset has most relevant images and very few non-relevant images and sometimes it has only the relevant images. Then a set R_S can be defined by an equation

$$R_S = \{x : x \in D\}, n(R_S) = K, \ where\ K < N, hence\ R_S \subset D \tag{8}$$

The size of the R_S dataset is much less than that of original database. Similarly, texture based approach will subsequently retrieve the L number of images from a R_S dataset. Let R_T be a set of top M number of retrieved images using the texture features. Then a R_T can be defined as

$$R_T = \{x : x \in R_S\}, n(R_T) = M, \ where\ M < K, hence R_T \subset R_S \tag{9}$$

Similarly, let R_C be a set of L number of retrieved images from a dataset R_T using color based approach. Then R_C can be defined as

$$R_C = \{x : x \in R_T\}, n(R_C) = L, \ where\ L < M, hence\ R_C \subset R_T \tag{10}$$

From equations (8), (9) and (10), the relation among D, R_S, R_T and R_C can be given as

$$R_C \subset R_T \subset R_S \subset D \tag{11}$$

the dataset R_C is the final output of the proposed CBIR system. The presented research work requires low computational overhead for retrieving the top most images from the dataset because each stage reduces the size of the dataset. The schematic block diagram of proposed image retrieval system is shown in Fig. 5. The three approaches/stages of proposed scheme based on the shape, texture and color features are discussed in the following subsections in detail.

A. SHAPE BASED APPROACH

In this approach, initially, the query image and images in the dataset D are segmented based on Chan and Vese [39] method which divides images into foreground(object) and background parts/images. In this work, authors have concatenated the feature vectors of foreground(object) and background parts into single feature vector which represent the contents of whole image. The creation of shape feature vector is shown in Figure 6. Now, histogram of gradients(HOGs) and invariant moments based shape features have been computed from two separated parts of the grayscale image. The foreground image is considered for construction of the HOG feature vector using 50% overlapping rectangular blocks (or cells) because overlapping blocks provides more significant image information than the non-overlapping blocks. The 9 bin histogram of signed gradient from each overlapping rectangular block is calculated by quantizing the angles 180^0, to -180^0 or 0^0, to,360^0 over a range of 40 degree per bin, where each bin value consists of a magnitude of gradient.

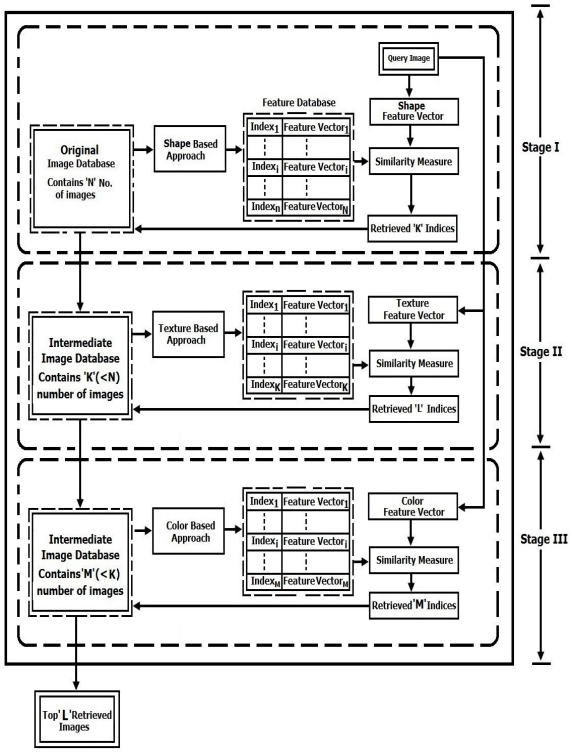


FIGURE 5. Schematic block diagram of proposed CBIR system.

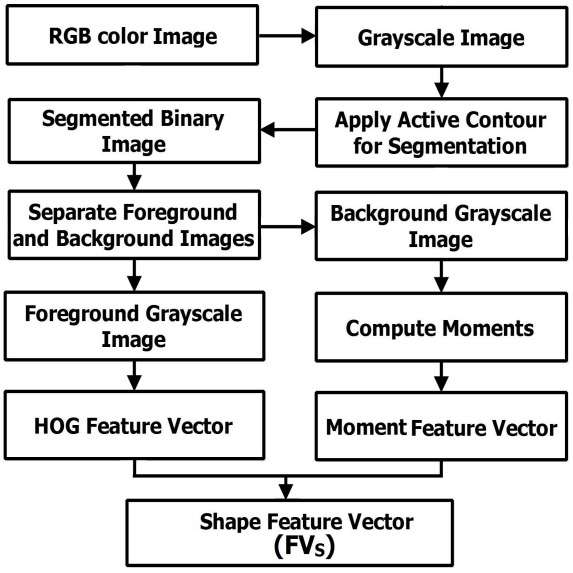


FIGURE 6. Block diagram of shape feature extraction.

The HOG descriptors are computed based on the signed gradients values, where each cell participated more than once for construction of the single HOG feature vector and in order to avoid illumination and contrast changes, the feature vector of each cell has been normalized. For the computation of HOG features, the foreground image has been divided into 9 non-overlapping rectangular cells and constructed the gradient histograms of 9 bin per cell. The 81-dimensional hog feature vector FV_{hog} is constructed by combining 9 gradient histograms with 9 bins. The HOG descriptors have several other merits like it provides an edge or image gradient structure that represents the local shape features of image and it also controls the invariance property of local geometric shape and other transformations such as translation and rotations. Next, we have computed seven moments based shape features from the background part of the segmented image. These moments are independent to location, orientations and scale. Several CBIR systems [52] are available in literature review that perform the manual preprocessing for the setting of image size and alignment of image but moment based features can obtain this automatically. Therefore seven moments which are rotation, location and size invariant are described as:

The moments of order p and q of discrete image $f(x, y)$ is defined as

$$m_{pq} = \sum_{x} \sum_{y} x^p y^q f(x, y), \quad \forall\ p,\ q = 0,\ 1,\ 2 \tag{12}$$

where x and y are spatial coordinates of the image. The central moments are defined as

$$\mu_{pq} = \sum_{x} \sum_{y} (x - \bar{x})\ (y - \bar{y}) f(x, y) \tag{13}$$

where $\bar{x} = m_{10}/m_{00}$, $\bar{y} = m_{01}/m_{00}$, are known as the center of region. Hence center moments of order three can be calculated as:

$$\begin{aligned}
\mu_{00} &= m_{00} \\
\mu_{10} &= 0 \\
\mu_{01} &= 0 \\
\mu_{11} &= m_{11} - \bar{y} m_{10} \\
\mu_{20} &= m_{20} - \bar{x}^2 m_{10} \\
\mu_{02} &= m_{02} - \bar{y} m_{01} \\
\mu_{30} &= m_{30} - 3\bar{x} m_{20} + 2m_{10}\bar{x}^2 \\
\mu_{21} &= m_{21} - 2\bar{x} m_{11} - \bar{y} m_{20} + 2\bar{x}^2 m_{01} \\
\mu_{12} &= m_{12} - 2\bar{y} m_{11} - \bar{x} m_{02} + 2\bar{y}^2 m_{10} \\
\mu_{03} &= m_{03} - 3\bar{y} m_{02} + 2\bar{y}^2 m_{01}
\end{aligned} \tag{14}$$

The central moments of order p and q are normalized as

$$\mu_{pq} = \mu_{pq}/\mu^{\gamma}{}_{00}, \quad \forall p, q = 0,\ 1,\ 2, \ldots \tag{15}$$

where $\gamma = (p + q)/2 + 1$.

The set of seven moments $(\phi_1 - \phi_7)$ for $(p + q) = 2, 3, \ldots$ can be calculated as follows:

$$\begin{aligned}
\phi_1 &= \mu_{20} + \mu_{02} \\
\phi_2 &= (\mu_{20} + \mu_{02})^2 + (4\mu_{11})^2 \\
\phi_3 &= (\mu_{30} + 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \\
\phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} - \mu_{03})^2 \\
\phi_5 &= (\mu_{30} + 3\mu_{12}) + (\mu_{30} + \mu_{12}) \\
&\quad \times \left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right]
\end{aligned}$$

$$
\begin{aligned}
&+ (3\mu_{21} + \mu_{03})(\mu_{21} + \mu_{03}) \\
&\times \left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] \\
\phi_6 &= (\mu_{20} - \mu_{02}) \\
&\times \left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right] \\
&+ 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \\
\phi_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} - \mu_{12}) \\
&\times \left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2\right] \\
&- (\mu_{30} - 3\mu_{03})(\mu_{21} + \mu_{03}) \\
&\times \left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2\right] \qquad (16)
\end{aligned}
$$

where six moments $\phi_1 - \phi_6$ are the invariant with size, orientations and translations and ϕ_7 is invariant to skew which is used to distinguish the mirror images. The set of seven normalized central moments form a feature vector which is independent to translation, orientations and size changes in image. It is obtained as:

$$FV_{mu} = [\phi_1,\ \phi_2, \phi_3, \phi_4, \phi_5,\ \phi_6,\ \phi_7] \qquad (17)$$

Finally, the single shape feature vector is computed by concatenating the hog feature vector FV_{hog} and moment vector FV_{mu} which is used for retrieving the top most images from an universal image dataset D. The retrieved images are stored in the intermediate dataset R_S and it will be input for the next stage where the undesirable images are filtered out. The entire process for the proposed shape based approach is given in the form of algorithmic steps as follows: The algorithmic steps for the texture based image retrieval scheme is gives as:

The images in the dataset R_S will be input for the texture based approach. As we know that the number of images in dataset R_S are less than the images in the original dataset D. Therefore, the computational overhead for the construction of the texture feature vectors will be low due to the reduced size dataset R_S. The texture based approach is employed on the images of R_S dataset and retrieved the top most relevant images by discarding some available the non-relevant images.

B. TEXTURE BASED APPROACH

In this approach, the query image and intermediate databases images are decomposed using tetrolet transform to overcome the limitations of the basic transforms like DWT [53], DT-CWT [19], shearlets [35] and contourlets [54] for selecting more directional local geometric features. The texture features of these local geometrical shape have been extracted in [21] by computing mean and standard deviations from each tetrolet decomposed sub-images where the correlation or spatial relationship among the pixel values was lost. Therefore, in order to recover the correlation between them, in this paper, we have computed the GLCMs of an each decomposed sub-image of tetrolet transform. In this work, the GLCMs of each sub-image have been considered in four various directions(or orientations) with unit distance between pixel values, since texture discrimination property depends on the selection

Algorithm 1 Shape Features Based Image Retrieval

Begin

1. Take RGB image I as an input and convert it into grayscale image.
2. Apply active contour method for segmenting the grayscale image and separate foreground and background parts for extracting the shape features.
3. Compute the HOG features and invariant moments from foreground and background parts respectively.
4. The single shape feature vector is obtain by concatenating hog feature vector FV_{hog} and moment feature vector FV_{mu} as
$$FV_s = [FV_{hog}, FV_{mu}]$$
5. Construct the shape feature vectors of all images in the dataset D and query image shape feature vector by using above steps.
6. The Euclidean distance between the query image shape feature vector $FV_s(Q)$ and target image shape feature vector $FV_s(T)$ in the dataset D is computed as:
$$\Delta D_s = \sqrt{\sum_{z=1}^{ds} (FV_s^z(Q) - FV_s^z(T))^2},\ \forall z = 1, 2, \ldots, ds$$
where ds represents the dimension of the shape feature vector.
7. Sort the computed distances in an non-decreasing order and retrieved the top $K(< N)$ number of images from dataset D based on their sorted distances. The retrieved images are stored in a dataset R_S.

End

of different orientations and distances. See Section III-C for more detail. Since the size of the GLCM depends on Q (Q is the quantization level) which means that its size is $Q \times Q$, we will not compute directly a feature vector of dimension $Q \times Q$. Therefore, for the formation of low dimensional texture feature vector, second order statistical parameters like contrast, correlation, angular second moment(or energy) and diagonal distribution(or homogeneity) have been computed from all GLCMs of tetrolet decomposed sub-images. The second order statistical parameters are widely used for the classification of images by characterizing their basic properties. These four statistical values from the GLCM of each kk^{th} level decomposed sub-image are computed as follows:

$$\textit{Contrast} \qquad f_1^{kk} = \sum_{i,j=1} C_{ij}(i-j)^2$$

$$\textit{Correlation} \qquad f_2^{kk} = \sum_{i,j=1} \frac{(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j} C_{ij}$$

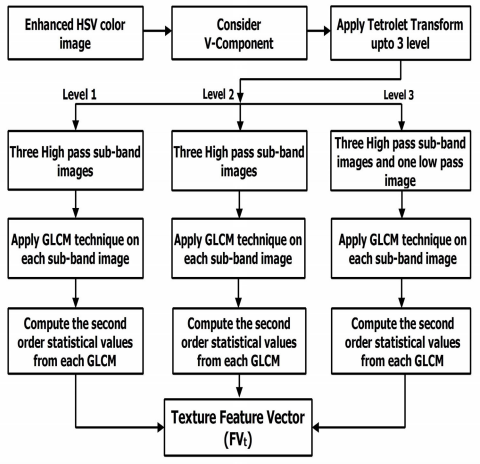


FIGURE 7. Block diagram of texture feature extraction.

$$\begin{aligned} &\textit{Energy} && f_3^{kk} = \sum_{i,j=1} C_{ij}^2 \\ &\textit{Homogeneity} && f_4^{kk} = \sum_{i,j=1} \frac{C_{ij}}{1+|i-j|} \end{aligned} \tag{18}$$

where μ_i and μ_j denote the mean values; σ_i and σ_j represent the standard deviations. Matrix Cij is NGLCM of the kk^{th} decomposed sub-image. The contrast feature f_1^{kk} measures the intensity variations between a gray tone value and its neighbor gray tone value over the kk^{th} tetrolet decomposed sub-image and it represents the contrast between of gray tone values for the texture features; Correlation feature f_2^{kk} measures how correlated a gray tone value to its neighboring gray tone values of the sub-band image. Energy feature f_3^{kk} is computed by the summing up the squared GLCM elements. The another names of Energy are angular second moment, uniformity or uniformity of energy. It becomes one if images are reported as constant or uniform. The forth feature f_4^{kk} measures the distribution of the diagonal elements in the GLCM and find how much the elements are closed with each other. It will be 1 if GLCM has only diagonal elements. Finally, texture feature vector FV_t is obtained by the combining the above statistical features at kk^{th} decomposition level as

$$FV_t = [f_1^{kk}, f_2^{kk}, f_3^{kk}, f_4^{kk}] \tag{19}$$

In the proposed work, the tetrolet transform has been applied upto three levels and the four statistical features (discussed in equation (18)) are computed from the high pass sub-band image at each level of decomposition while these are computed from the low pass sub-band image at last decomposition level only. In this way, the length/dimension of the texture feature vector FV_T is 1×40 which is too small as compared to the original dimension of image. The whole process for creation of texture feature vector is depicted in Figure 7.

The algorithmic steps for the texture based image retrieval scheme is gives as:

Algorithm 2 Texture Features Based Image Retrieval

Begin

1. Get an enhanced HSV color image and it is decomposed into its three H, S and V color components. Consider V-component for extracting significant texture features.
2. Obtain the sub-band images of V-component by applying tetrolet transform upto kk^{th} level with all 117 possible combinations where the high pass/detailed and low pass/approximate coefficients are arranged by using equations (3) and (4). The bijective method is used for mapping of index pairs.
3. An arrangement of coefficients have constructed one low pass sub-band image which is further decomposed in next level and three high pass sub-band images those are kept entirely at each level without changing their pixel values.
4. Apply the GLCM approach to the high pass sub-band images at each level whereas it is employed only on the low pass sub-band image at last level.
5. The texture feature vector (FV_t) is obtained by computing the four second order statistical parameters from the kk^{th} level of decomposed sub-band images by using equation (19) as:
$$FV_t = [f_1^{kk}, f_2^{kk}, f_3^{kk}, f_4^{kk}]$$
6. Construct the texture feature vectors of all images in the dataset R_S and query image texture feature vector by using above steps.
7. The Euclidean distance between the query image texture feature vector $FV_t(Q)$ and target image texture feature vector $FV_t(T)$ in the dataset R_S is computed as:
$$\Delta D_t = \sqrt{\sum_{z=1}^{dt} (FV_t^z(Q) - FV_t^z(T))^2}, \quad \forall z = 1, 2, \ldots, dt$$
where dt represents the dimension of the texture feature vector.
8. Sort the computed distances in non-decreasing order and retrieved the top $L(< K)$ number of images from dataset R_S based on their sorted distances. The retrieved images are stored in a dataset R_T.

End

The $L(< K)$ number of images based on texture features are retrieved and stored in the dataset R_T. The dataset R_T has most of relevant images and few non-relevant images. To filter out the undesired images, the dataset R_T will be

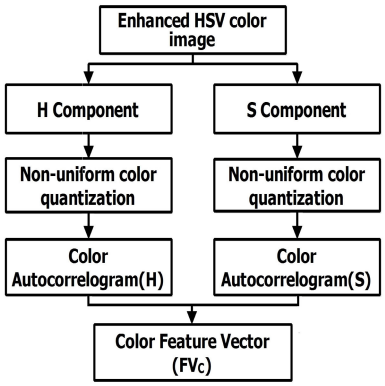


FIGURE 8. Block diagram of color feature extraction.

further processed by the color features based approach. Since dataset $R_C \subset R_T$, therefore, it requires low computational cost for constructing the color feature vector.

C. COLOR BASED APPROACH

Color is the most prominent feature and plays an significant role in image retrieval system. In this work, an HSV color space is adopted for extracting the color features because it is very close to the human visual perception. In HSV color image [53], color information is directly based on the two color components: H and S where each color component is represented by 255 different colors/bins while the V component is considered for texture feature. As a result, it is computationally high and time consuming process and there is no need to represent all different colors directly. Hence to reduce the computational overhead, the H and S components of HSV color space are quantized into non-uniform bins or levels by using Lloyd algorithm [51] where both H and S components are represented by only 8 bins. Since histogram based bin to bin matching is not feasible in CBIR because two different images may have same histograms and it is also lost the correlation among the neighboring pixel values of the image. In order to overcome the limited, the color autocorrelogram features are computed which provides the spatial correlation between identical colors of the image at a particular distance and it is also outperforms over the color correlogram features. Figure 8 shows the whole process for the formation of the color feature vector.

The color autocorrelogram $\alpha^k(l)$ is defined as the probability of calculating a pixel p' of similar color at specific distance k form an other given pixel p of the l^{th} color. It is computed as follows

$$\alpha^k(l) = \Pr\left[p' \in I \,\middle|\, |p - p'| = k \text{ and } p' \in I(l) \text{ for } p \in I(l) \subset I\right], \quad l \in \{0, 1, \ldots, L-1\} \qquad (20)$$

where $Pr[.]$ represent the probability of occurrence of the same colors, Image I is the set of pixels of original image and $I(l)$ is set of pixels of quantized image where it is re-quantized with L levels/bins. The algorithm steps of the proposed image retrieval scheme based on color features are presented as follows:

Algorithm 3 Color Features Based Image Retrieval

Begin

1. Get an enhanced HSV color image and decomposed into its three hue(H), Saturation(S) and value(V) components.
2. Quantize H and S components into 8 number of levels/bins using non-uniform quantization process.
3. Find the color feature vectors of the quantize H and S components by using equation (20) as follows:
   $FV_c^H = \alpha^k(H, l), \quad FV_c^S = \alpha^k(S, l)$
   where $\alpha^k(H, l)$ and $\alpha^k(S, l)$ are the color autocorrelograms of the H and S components while variable k represent the distance between two identical colors, here l is the identical color)
4. The single color feature vector (FV_c) is obtained by concatenating the features vectors of H and S components as:
   $$FV_c = \left[FV_c^H,\ FV_c^S\right]$$
5. Construct the color feature vectors of database images and query image by using above steps.
6. The Euclidean distance between the query image and target images in the dataset based on their color feature vectors is computed as:
   $$\Delta D_c = \sqrt{\sum_{z=1}^{dc} (FV_c^z(Q) - FV_c^z(T))^2}, \quad \forall z = 1, 2, \ldots, dc$$
   where $FV_c(Q)$ and $FV_c(T)$ are the feature vectors of the query and target images respectively and dc represents the dimension of the feature vector.
7. Sort the computed distances in an ascending order and retrieved the top $K(< N)$ number of images from dataset R_T based on their sorted minimum distances. The retrieved images are stored in a dataset R_C.

End

The retrieved images in the dataset R_C is the final output of the proposed image retrieval system.

V. EXPERIMENTATION

The experiments are validated on MatLab2011b with 32-bit Microsoft Windows 7 OS, platform Intel core i5-2365 with 1.65 processor and 6 GB random access memory. The benchmark databases, retrieval performance and results discussions are described in the following subsections.

A. DATASETS DESCRIPTION

The retrieval performance of any CBIR system is directly depends on image characteristic like color, texture, shape, size, object location, quality, overlapping of ROIs and

FIGURE 9. Sample images of the Corel-1K dataset.

FIGURE 10. Sample images of the GHIM-10K dataset.

cluttering [55]. In the proposed CBIR system, two widely used benchmark databases with their image categories are considered where the images in the categories have the various properties like diverse contents with different structures, several kinds of texture patterns, small objects in foreground and background, spatial image features, shapes and colors. The images of these datasets have taken from different areas to validate the accuracy of presented CBIR systems. The two datasets Corel-1K [56] and GHIM-10K [57] are widely used in standard existing CBIR systems for evaluation of their performance. The Corel-1K dataset has 1000 images with 10 categories/groups where each category consists of 100 similar types of images. The semantic names of the images of each category are people, flowers, elephants, dinosaurs, beaches, buildings, mountains, foods, horses and buses with the resolution 256×384 or 384×256 pixels. In this dataset, the images of some categories are mixed with the other categories images and also they have various content with different structures. Hence it is not easy to retrieve the most appropriate images based on the user's need.

The second dataset GHIM-10K is more challenging and 10 times larger than the Corel-1K dataset. This dataset consists 10000 images with 20 categories/groups where each category has 500 similar types of images with diverse contents and various objects. The semantic names of images in each group are sunsets, bikes, forts, ships, flies, cars etc. with resolution of 300×400 pixels or 400×300 pixels. The Figures 9 and 10 are depicted the sample images of Corel-1K and GHIM-10K dataset where single image has been taken from each category.

B. PERFORMANCE EVALUATION

Precision and recall are two standard matrices those are used to measure the performance of image retrieval system. Our proposed CBIR system has been developed in three stages i.e. color based approach(C), texture based approach(T) and

shape based approach(S) where each stage filtered out the non-relevant images keeping the top relevant images. In the final stage, the precision and recall are computed based on the number of retrieved images and relevant images available in the dataset. Let variables X and Y are the total number of relevant images and non-relevant images retrieved from the dataset and variable Z denotes the total number of relevant images present in the dataset. Then, the precision and recall matrices are defined as

$$P(q) = \frac{X}{X+Y}$$
$$R(q) = \frac{X}{X+Z}$$

where $P(q)$ denotes the ratio of total number of retrieved images relevant to query image q and total number of retrieved images from the dataset; $R(q)$ is the ratio of the total number of retrieved relevant images and total number of relevant images to query q present in the dataset. But above two metrices are not sufficient to evaluate the overall performance of CBIR system. Hence harmonic mean of precision and recall is defined which is known as F-score or F-measure. It is computed as

$$F(q) = \frac{2 \times P(q) \times R(q)}{P(q) + R(q)}$$

where F-score $F(q)$ is a single value which represent the overall performance of the CBIR system.

C. COMPUTATIONAL ANALYSIS

The proposed CBIR system can generate the good retrieval results if an user already knows about the distribution of images in the dataset. The previous awareness of the size of dataset may facilitate for putting the appropriate values of the variables K, M and L where L is final output(i.e. retrieved images) of the image retrieval system while K and M are intermediate retrieval results. The suitable values of K, M and L should provides the better retrieval performance of the presented CBIR system. Here, the final retrieval results depends on the selected value of K. In the first stage, the variable K divides the universal dataset into two parts; first part known as relevant search area which has K number of images while the other part is known as irrelevant search area having $(N - K)$ images where N is the total number of images available in the universal dataset. The better retrieval results means here is that the variable K has the most of the relevant images if it has not enough related images, it is required to increase the value of K until the good precision is attained. In the traditional image retrieval systems, images are searched in the whole dataset independently for each feature space where feature spaces are made by color, texture and shape features individually. Let N denotes the total images in the universal dataset and variable F represent the total feature spaces used for indexing of the images in the dataset. Then, the computational time for earlier image retrieval systems is approximately computed as: $T = N \times F$. Let $N = 1000$ and user wants to compute the three feature spaces for color, texture and shape features, then the estimated time T for traditional CBIR systems is equal to $1000 \times 3 = 3000$ units. It will be much time consuming process if the dataset is too large but the computational time for the proposed CBIR system rely on the values of the variable K, M and L. The proposed system is divided into three techniques i.e. color, texture and shape where all the images are searched from the universal dataset at first technique only and significant number of non-relevant images are discarded to create the reduced sized intermediate datasets at the subsequent techniques based on the values K, M and L. In the present research work, authors have reduced the dataset sizes based three color, texture and shape feature spaces by filtering out the non-relevant images during retrieval process. Hence, the estimated computational time T of the presented CBIR system is defined as:

$$T = N + K + L$$

For example the approximate computational time for Corel-1K is computed as $T = 1000 + 100 + 40 = 1140$ units where $N = 1000, K = 100$ and $M = 40$. For dataset GHIM-10K, the time $T = 10000 + 500 + 200 = 10700$ units where $N = 10000, K = 500$ and $M = 200$. The values of the variable K and M varies according to the user's requirement and the suitable values have been chosen which produces the higher precision. Hence, we have found that the computational time for the proposed CBIR system is low as compared to the conventional image retrieval systems.

D. RETRIEVAL RESULTS AND DISCUSSIONS

The proposed method retrieves the top most relevant images based on the color(C), texture(T) and shape(S) visual features in the hierarchical mode using three stages where the orders of the visual features are taken as:

$$O_1 = (C{-}T - S)$$
$$O_2 = (C{-}S - T)$$
$$O_3 = (T{-}C - S)$$
$$O_4 = (T{-}S - C)$$
$$O_5 = (S{-}C - T)$$
$$O_6 = (S{-}T - C)$$

The retrieval results in terms of precision, recall and F-score based on above six order ($i.e. O_1, O_2, O_3, O_4, O_5$ and O_6) are shown in Figs 11, 12 and 13 for Corel-1K dataset where image category ID represent the categories of the images of the dataset and dinosaur images has the best retrieval results among all other categories while the mountain images has the worst results but order O_6 provides the best average retrieval accuracy among other five orders of the visual features. Similarly, the experimental results are validated on GHIM-10K dataset where order O_6 also provides the best retrieval rate than other orders. Therefore, an order $O_6 = (S{-}T - C)$ will be adopted for further discussions to presents the final retrieval rates. Table 1 shows the average metrics(i.e. precision, recall and F-score) for Corel-1K and GHIM-10K

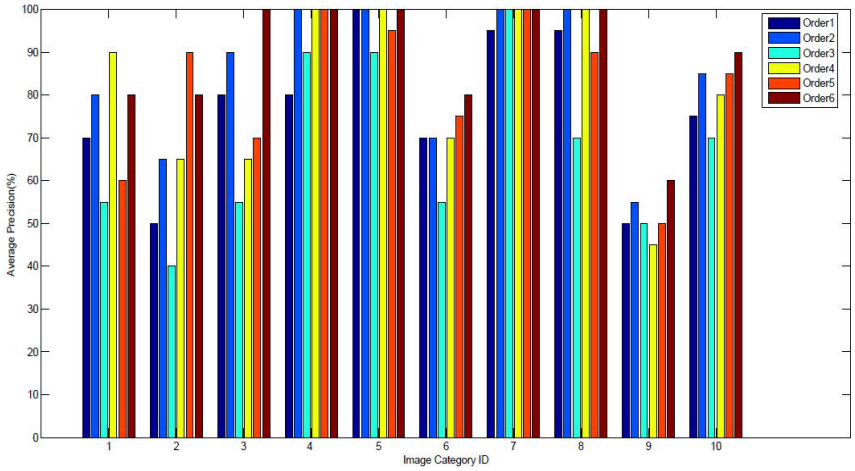


FIGURE 11. Average precision based on orders of the visual features for Corel-1K dataset.

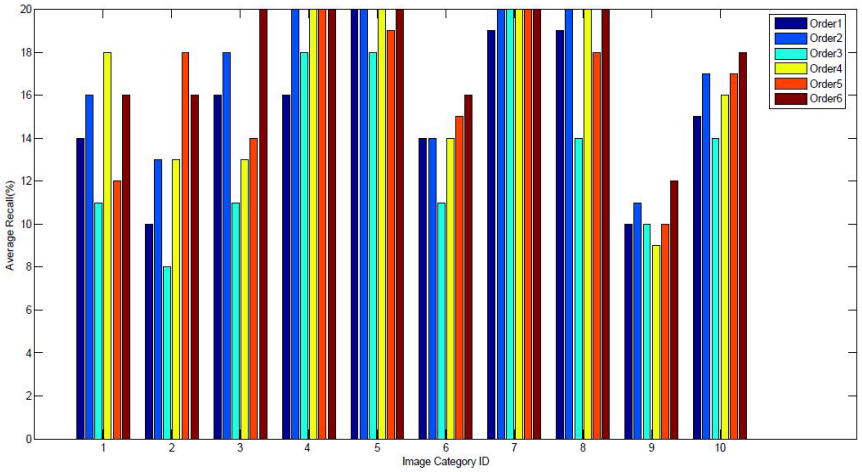


FIGURE 12. Average recall based on orders of the visual features for Corel-1K dataset.

TABLE 1. Average precision, recall and F-score for two standard datasets.

Order of features	Corel-1K			GHIM-10K		
	Precision	Recall	F-score	Precision	Recall	F-score
$O_1(C-T-S)$	76.50	15.30	25.50	67.50	2.70	5.19
$O_2(C-S-T)$	84.50	16.90	28.17	84.25	3.23	6.48
$O_3(T-C-S)$	67.50	13.50	22.50	55.00	2.20	4.23
$O_4(T-S-C)$	81.50	16.30	27.17	86.75	3.47	6.67
$O_5(S-C-T)$	81.50	16.30	27.17	65.75	2.63	5.10
$O_6(S-T-C)$	89.00	17.80	29.49	90.00	3.60	6.92

datasets where average metrics for ten image categories of Corel-1K dataset are 89.00%(precision), 17.80%(recall) and 29.49%(F-score) while average(precision, recall and F-score) for twenty image categories of GHIM-10K are 90.00%, 3.60% and 6.92%. These average metrics are the best among the other five orders. The produced retrieval rate for presented CBIR scheme will be acceptable for any modern scenario of image retrieval systems.

The retrieval results based on order O_6 of the visual features are discussed as follows; The order O_6 produces the best retrieval results in most of the instances because shape features based on active contour model, HOGs and invariant moments are an efficient and proficients for large image dataset with various objects and structures. Initially, shape based approach is adopted for retrieving the top most K number of images from the universal dataset and stored these images in a new dataset R_S which contains most relevant images and few non-relevant images. Next, in second stage, texture based approach is considered to retrieve the top M number of images from the dataset R_S by discarding the non-relevant images and retrieved images are kept in R_T dataset. Here, the dataset R_T has very few non-relevant images which will be further filtered out in the last stage. Hence, in this stage, color based approach is used to retrieve the most relevant images from the dataset R_T by filtering out the available non-relevant images and retained the L number

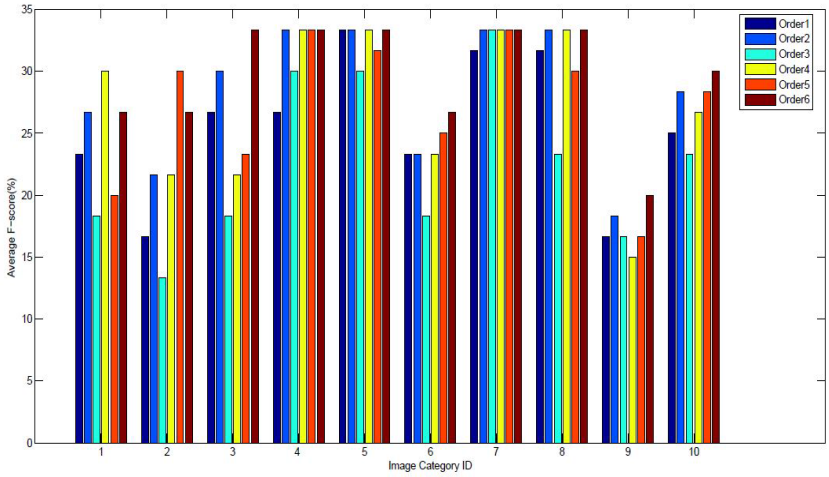


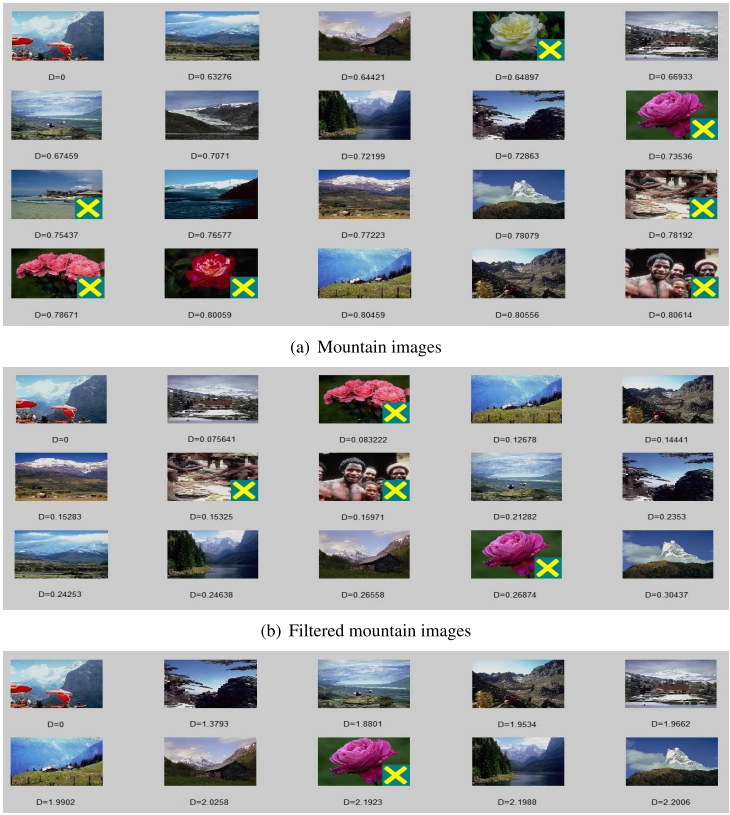
FIGURE 13. Average F-score based on orders of the visual features for Corel-1K dataset.

of retrieved images in R_C dataset. The images in R_C dataset is the final output of proposed system. The filtering process of non-relevant images from reduced sized datasets R_S and R_T are explained as where each dataset has more number of relevant images than that of non-relevant images. For example, we set $n(R_S) = K = 20$, $n(R_T) = M = 15$, and $n(R_C) = L = 10$, in more detail, initially, the top 20 images are retrieved based on the shape based approach, after that the texture based approach is employed for retrieving the top 15 images from the $n(R_S) = K = 20$ images by filtering out the non-relevant images. Lastly, still some non-relevant images are available in $n(R_T) = M = 15$ retrieved images, so top 10 images are retrieved from $n(R_T)$ dataset by filtering out the remaining non-relevant images. For corel-1K image dataset, Figs 14 and 15 present the filtering process of non-relevant images for mountains and elephants, where upper left images of window are queries. The symbol $(\times)$ denotes the irrelevant images while all remaining images are the relevant to the query image. Fig. 14(a) shows the 13 relevant retrieved images and 7 non-relevant retrieved images while Fig. 14(b) represent 11 relevant and 4 non-relevant images. Therefore, the total number of filtered out non-relevant images are 7-4 = 3 from Figs. 14(a) to 14(b) while Fig 14(c) has only 1 non-relevant image. Here, finally, total number of filtered irrelevant images are 7-1 = 6. Similarly, Fig. 15 depicted the filtering process for elephant images where only two non-relevant images are left in Fig. 15(c). Here, total number of non-relevant images filtered out are 8-2=6. Hence, we conclude that the performance of the proposed system is increased because most of the relevant images for mountain and elephant images are kept in the final retrieval stage. Similarly, the filtering process for insects and bikes of GHIM-10K dataset is shown in Figs. 16 and 17 respectively, where insect images has 3 non-relevant images while bikes has no irrelevant image at final stage. Hence, the proposed CBIR system has been produced the good retrieval rate for large and complex dataset and it also requires low computational overhead because proposed method is divided into stages where each stage kept most relevant images by filtering out the non-relevant images.

TABLE 2. Precision, recall and F-score for top L retrieved images from Corel-1K dataset.

Category	L=10			L=20		
	Precision	Recall	F-score	Precision	Recall	F-score
People	100.00	10.00	18.18	80.00	16.00	26.67
Beaches	90.00	9.00	16.36	80.00	16.00	26.67
Buildings	100.00	10.00	18.18	100.00	20.00	33.33
Buses	100.00	10.00	18.18	100.00	20.00	33.33
Dinosaurs	100.00	10.00	18.18	100.00	20.00	33.33
Elephants	90.00	9.00	16.36	80.00	16.00	26.67
Flowers	100.00	10.00	18.18	100.00	20.00	33.33
Horses	100.00	10.00	18.18	100.00	20.00	33.33
Mountains	70.00	7.00	12.72	55.00	11.00	18.33
Foods	100.00	10.00	18.18	90.00	18.00	30.00

Here, final retrieval results of the presented CBIR system are presented. Table 2 shows the retrieval performance in terms of precision, recall and F-score for top L = 10/20 retrieved images from Corel-1K dataset where top L = 10 images produces the satisfactory average metrics i.e 95.00%(precision) 9.50%(recall) and 17.27%(F-score) while in case of L = 20 images, average metrics are 88.50%, 17.70% and 29.49%. In this table, the mountain images has the worst results i.e. 70.00% precision for top 10 images and 60.00% precision for top 20 retrieved images while buildings, buses, dinosaurs, flowers and horses has the best retrieval results. The mountain images has more complex structure with diverse contents and it is mixed with other image categories like beaches, elephants and buildings. So the proposed method is unable to classify the mountain image category appropriately to retrieved the desire images. For GHIM-10K dataset, the average retrieval rate (precision, recall and F-score) for top L = 10/20 retrieved images are shown in Table 3, where we have found that the retrieval rate is high for L = 10 images than L = 20. Hence it is clear that if the



(a) Mountain images

(b) Filtered mountain images

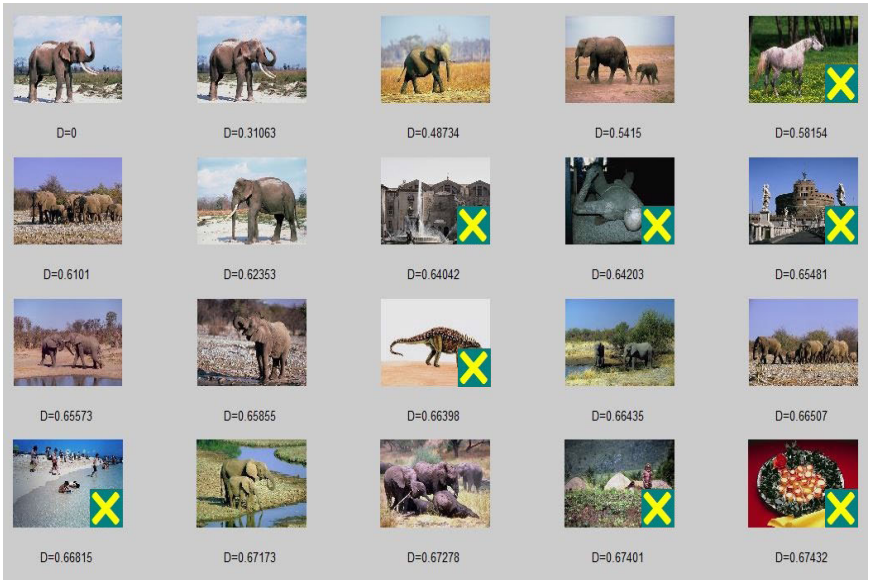
(c) Final filtered mountain images

FIGURE 14. Filtering process of mountain images: (a) Retrieved images from universal dataset using shape based approach. (b) Retrieved filtered images from the set of (a) images using texture based approach (c) Retrieved final images from the set of (b) images using Color based approach.

TABLE 3. Precision and recall and F-score for top L images retrieved from GHIM-10K dataset.

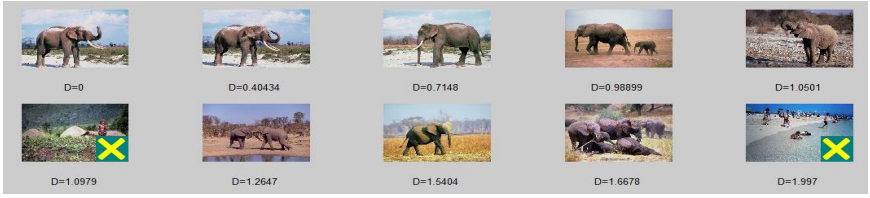
Category	Top L=10			Top L=20		
	Precision	Recall	F-score	Precision	Recall	F-score
Fireworks	100.00	2.00	3.92	100.00	4.00	7.69
Buildings	100.00	2.00	3.92	95.00	3.80	7.31
Walls	90.00	1.80	3.52	80.00	3.20	6.15
Cars	100.00	2.00	3.92	100.00	4.00	7.69
Flies	90.00	1.80	3.52	80.00	3.20	6.15
Mountains	80.00	1.60	3.13	65.00	2.60	5.00
Flowers	100.00	2.00	3.92	100.00	4.00	7.69
Trees	100.00	2.00	3.92	100.00	4.00	7.69
Green Grounds	100.00	2.00	3.92	100.00	4.00	7.69
Beaches	100.00	2.00	3.92	85.00	3.40	6.54
Airoplanes	90.00	1.80	3.52	85.00	3.40	6.54
Butterflies	100.00	2.00	3.92	90.00	3.60	6.92
Forts	90.00	1.80	3.52	85.00	3.40	6.54
Sunsets	100.00	2.00	3.92	100.00	4.00	7.69
Bikes	100.00	2.00	3.92	100.00	4.00	7.69
Boats	100.00	2.00	3.92	100.00	4.00	7.69
Ships	90.00	1.80	3.52	85.00	3.40	6.54
Chickens	100.00	2.00	3.92	95.00	3.80	7.31
Insects	80.00	1.60	3.13	70.00	2.80	5.38
Horses	90.00	1.80	3.52	85.00	3.40	6.54

number of retrieved images are increases then retrieval rate would be decreases. In this, table, fireworks, cars, flowers, trees, green grounds, sunsets, bikes, boats have generated the 100.00% precisions for top 10 and 20 images while the mountains and insects have the lowest precisions. The average metrics of 20 image categories of GHIM-10K dataset are 95.00%



(a) Elephants images

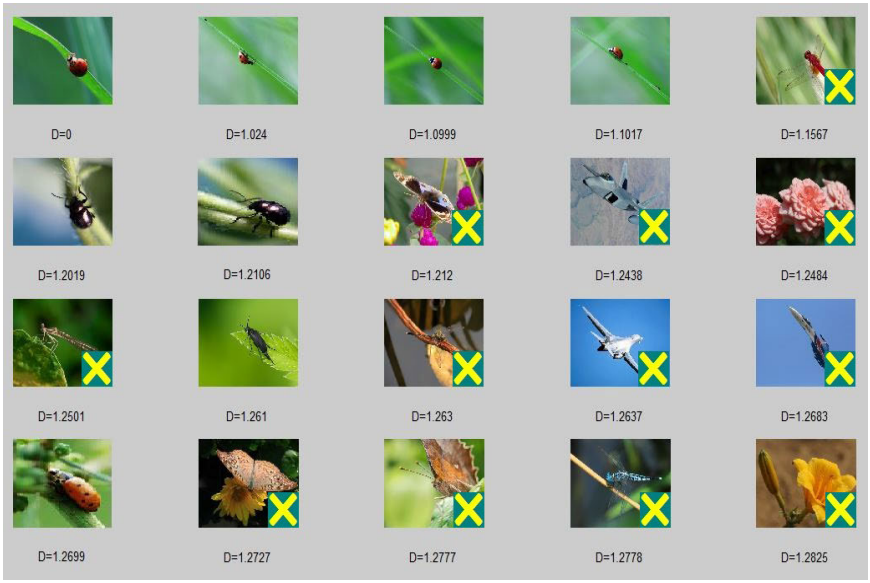
(b) Filtered elephants images



(c) Final filtered elephants images

FIGURE 15. Filtering process of elephant images: (a) Retrieved images from universal dataset using shape based approach. (b) Retrieved filtered images from the set of (a) images using texture based approach (c) Retrieved final images from the set of (b) images using Color based approach.

precision, 1.90% recall and 3.72% F-score and these averages are 90.00% precision, 3.60% recall and 6.92% F-score for top 10 images. To check the relative performance of the proposed CBIR system, the experimental results have been compared with some available state-of-the-art image retrieval systems in terms of precision, recall and F-score. These comparative methods have been developed by Walia *et al.* [58], Irtaza *et al.* [59], Ashraf *et al.* [60] Guo *et al.* [61], Zeng *et al.* [62], Fadaei *et al.* [63], Zhou *et al.* [64], Mistry *et al.* [65] and Ahmed *et al.* [66]. The discussion and details of existing CBIR systems along with the proposed system are described as follows; Walia *et al.* [58] have extracted the color, texture and shape visual features from image and efficient combination of these features have been used for retrieving the most relevant images from the dataset. In their work, color and texture visual features have been extracted based on modified color difference histogram (MCDH) where MCDH is computed using color quantization and partial derivatives. The Angular Radical Transformation (ART) has used for extracting local and global shape features.



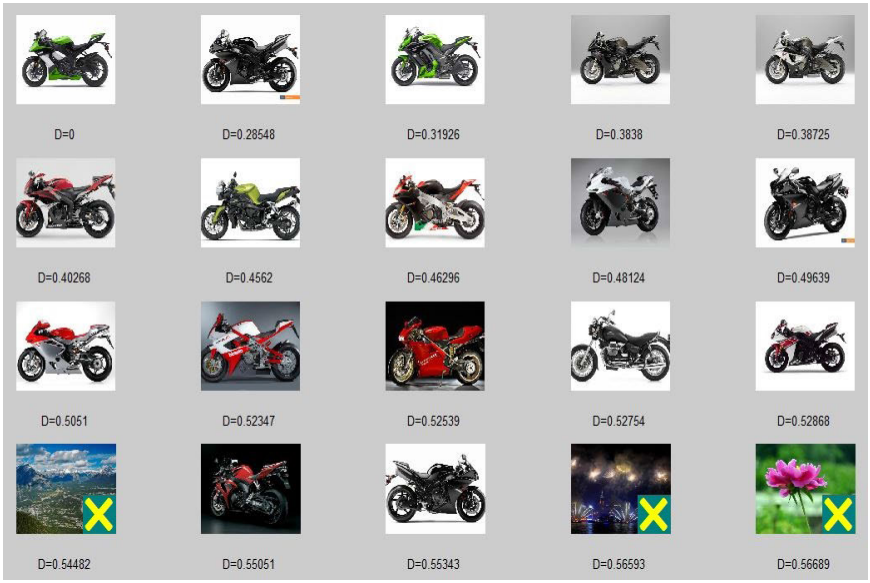
(a) Insect images

(b) Filtered insect images

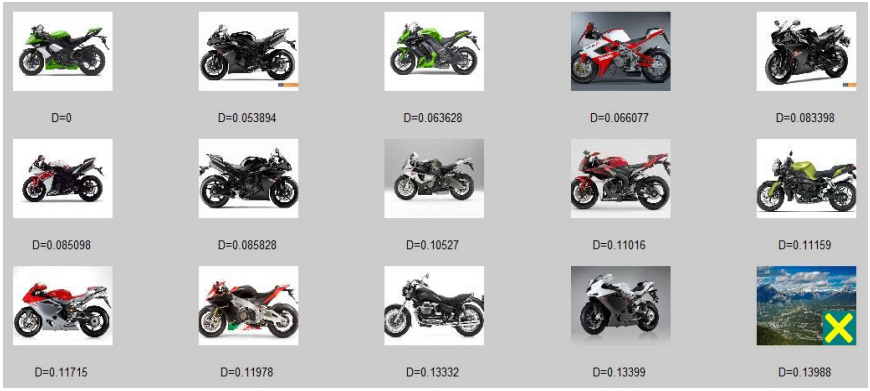
(c) Final filtered insect images

FIGURE 16. **Filtering process of insect images: (a) Retrieved images from universal dataset using shape based approach. (b) Retrieved filtered images from the set of (a) images using texture based approach (c) Retrieved final images from the set of (b) images using Color based approach.**

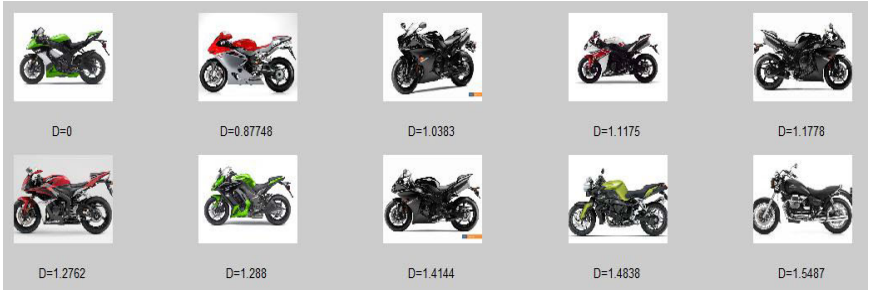
Once the features are extracted, they have used three normalization approaches on distances and single distance is obtained by assigning weights to the individual distances of MCDH and ART features for retrieving the final results. The distances based on weights are not efficient and it is computationally high. This method have achieved the best precision i.e. 100% for some image categories while very low precision i.e. 38.00% for food images have been obtained. Hence, it returns undesirable results in most of cases such as people, food and building images. Hence, our proposed CBIR system reduces the limitations of weighting distances and produces the best retrieval results for most of the category images. Irtaza *et al.* [59] proposed neural network based CBIR system where they have extracted efficient texture



(a) Bikes images



(b) Filtered Bikes images



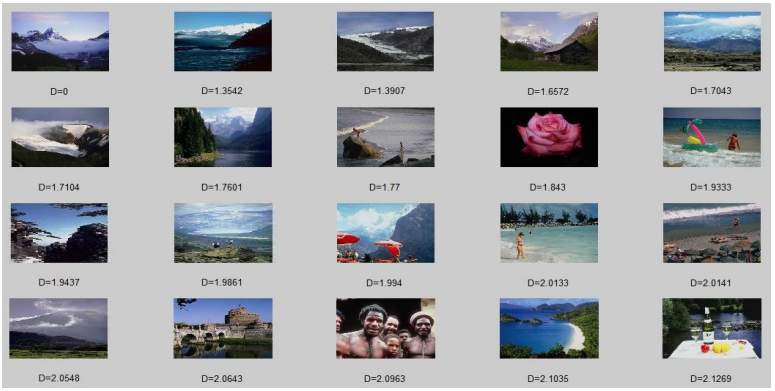
(c) Final filtered Bikes images

FIGURE 17. Filtering process of bikes images: (a) Retrieved images from universal dataset using shape based approach. (b) Retrieved filtered images from the set of (a) images using texture based approach (c) Retrieved final images from the set of (b) images using Color based approach.

image features using the concept of in-depth texture analysis, wavelet packets and Gabor filters. Further, for better retrieval, K-nearest neighbors based partial supervised learning algorithm is adopted for semantically correct image classification. In their scheme, flower images have best retrieval results i.e. 94.00% precision, 18.80% recall and 31.33% F-score and bad retrieval accuracy (i.e. 60.00% precision, 12.00% recall and 20.00% F-score) for beach images. Ashraf *et al.* [60] have extracted the color and texture features from RGB and YCbCr color spaces. The color features have been computed based on color histograms and wavelet coefficients while the texture visual features have been calculated using Bandelet transform, Gabor filters and the artificial neural network (ANN). The Bandelet transform have calculated the major

(a) Buses

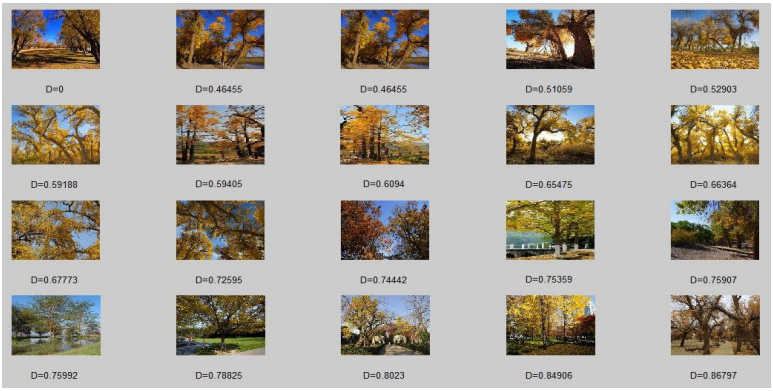
(b) Foods



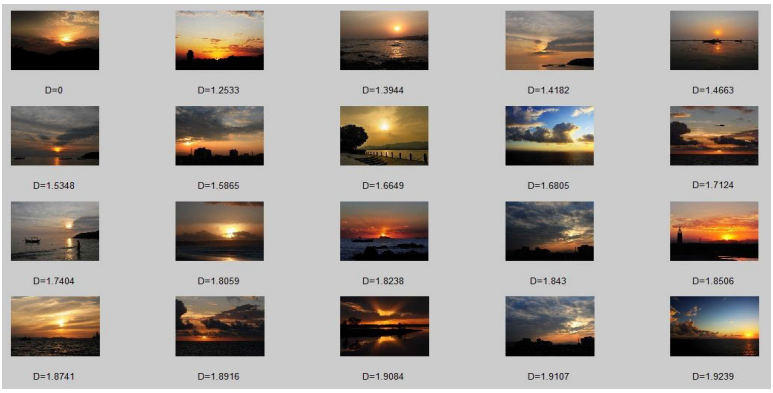
(c) Mountains

FIGURE 18. **Retrieved images for three categories of corel-1K dataset from (a)-(c) where upper top left images are the queries.**

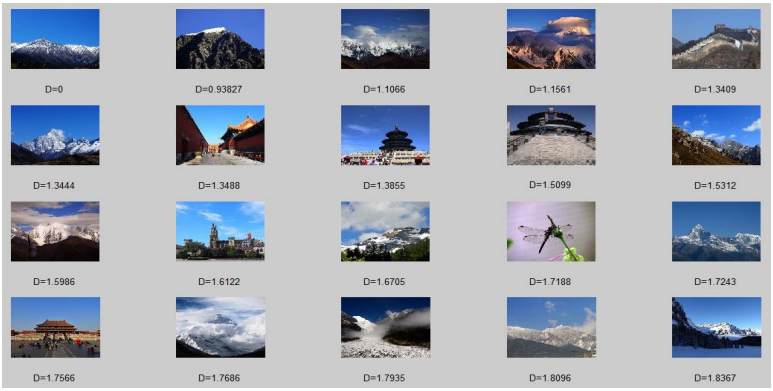
objects of the image which represent the image effectively in a compact form. For efficient retrieval results, the color, texture features together with ANN technique have been combined. The average satisfactory retrieval rates are 82.00% precision, 16.40% recall and 21.67% F-score, but still it needs some improvements over few other category images. The main drawback of this discussed method is double use of ANN technique and usage of complex transform during whole process of CBIR. Guo *et al.* [61] have developed the efficient image retrieval system based on color co-occurrence and local binary pattern image features where these features have been extracted by using ordered dither bock truncation coding quantizers and bitmap. They have also used the weighting constants between two images during similarity measures for efficient image retrieval. The main demerit of their scheme is to perform several experiments of various image block size and setting the different values of weighting constants to check the final retrieval results but still they have reported the good average precision. Zeng *et al.* [62] have computed image features based on quantized color histogram and spatiogram approaches where the color image is quantized into the various color bins by using Gaussian Mixture Models (GMM). The Expectation-Maximization (EM) algorithm has been adopted to estimate the components of the GMMs based



(a) Trees



(b) Sunsets



(c) Mountains

FIGURE 19. Retrieved images for three categories of GHIM-10K dataset from (a)-(c) where upper top left images are the queries.

on maximum likelihood estimator. Further, the EM-Bayesian Information criteria (EM-BIC) has used to find the suitable number of quantized color bins automatically. The histogram based spatiogram approach represents the distributions of colors those are spatially weighted by the location of pixels. In this scheme, dinosaurs, flowers and horses category images have 100.00%, 94.80% and 91.80% precisions, while the beaches, elephants and buildings images have achieved the worst precisions i.e. 65.20%, 70.50%, and 70.60% because beaches, elephants and buildings images have various structures with diverse contents. Fadaei *et al.* [63] have suggested a color and texture visual features based image retrieval system in which the color information based on the DCD of the quantized HSV color space has been extracted and two texture features are computed by using curvelet and wavelet transform. A single distance has been obtained by assigning the weights to the individual distances of three visual features where an appropriate weights are selected by using particle swarm optimization algorithm. The CBIR system based on the three feature spaces and weighting distances requires high overhead. To overcome the problem, hierarchical image retrieval system has been

TABLE 4. Comparison of proposed system with some other standard existing CBIR system in terms of precision, recall and F-score for top 20 retrieved results.

Category	Metrics (%)	Walia (2014)	Irtaza (2014)	Ashraf (2015)	Guo (2015)	Zeng (2016)	Fadaei (2016)	Zhou (2016)	Mistry (2017)	Ahmed (2017)	Proposed Method
Peoples	Precision	51.00	65.00	65.00	84.70	72.50	72.40	85.50	81.00	90.00	80.00
	Recall	10.20	13.00	13.00	16.94	14.50	14.47	17.00	16.20	18.00	16.00
	F-score	17.00	21.67	21.67	26.88	24.17	23.44	28.33	26.72	30.00	26.67
Beaches	Precision	90.00	60.00	70.00	46.60	65.20	51.15	53.40	92.00	92.00	80.00
	Recall	18.00	12.00	14.00	9.32	13.04	10.23	11.00	18.40	18.00	16.00
	F-score	30.00	20.00	23.33	15.06	21.73	16.72	18.22	30.11	30.11	26.67
Buildings	Precision	58.00	62.00	75.00	68.20	70.60	59.55	72.80	79.00	88.00	100.00
	Recall	11.60	12.40	15.00	13.64	14.12	11.91	14.00	15.80	18.00	20.00
	F-score	19.33	20.67	25.00	21.83	23.53	23.53	23.44	25.21	29.89	33.33
Buses	Precision	78.00	85.00	95.00	88.50	89.20	92.35	85.10	93.00	98.00	100.00
	Recall	15.60	17.00	19.00	17.70	17.84	18.47	17.00	18.60	20.00	20.00
	F-score	26.00	28.33	31.67	28.49	29.73	30.11	28.33	30.16	33.22	33.33
Dinosaurs	Precision	100.00	93.00	100.00	99.20	100.00	99.90	100.00	99.00	97.00	100.00
	Recall	20.00	18.60	20.00	19.84	20.00	19.87	20.00	19.80	19.00	20.00
	F-score	33.33	31.00	33.33	31.88	33.33	31.88	33.33	31.88	31.78	33.33
Elephants	Precision	84.00	65.00	80.00	73.30	70.50	72.70	68.70	79.00	85.00	80.00
	Recall	16.80	13.00	16.00	14.66	14.10	14.54	14.00	15.80	17.00	16.00
	F-score	28.00	21.67	21.67	23.49	23.50	23.44	23.22	25.21	28.33	26.67
Flowers	Precision	100.00	94.00	95.00	96.40	94.80	92.25	94.20	99.00	93.00	100.00
	Recall	20.00	18.80	19.00	19.28	18.96	18.45	19.00	19.80	19.00	20.00
	F-score	33.33	31.33	31.67	31.72	31.60	30.11	31.61	31.88	31.55	33.33
Horses	Precision	100.00	77.00	90.00	93.90	91.80	96.60	99.60	85.00	86.00	100.00
	Recall	20.00	15.40	18.00	18.78	18.36	19.32	20.00	17.00	17.00	20.00
	F-score	33.33	25.67	30.00	30.16	30.60	31.72	33.28	28.33	28.39	33.33
Mountains	Precision	84.00	73.00	75.00	47.40	72.25	55.75	55.70	88.00	84.00	60.00
	Recall	16.80	14.60	15.00	9.48	14.45	11.15	11.00	17.60	17.00	12.00
	F-score	28.00	24.33	25.00	15.11	24.08	18.33	18.33	28.49	28.28	20.00
Foods	Precision	38.00	81.00	75.00	80.60	78.80	72.35	86.50	80.00	92.00	90.00
	Recall	7.60	16.20	15.00	16.12	15.76	14.47	17.00	16.00	18.00	18.00
	F-score	12.67	27.00	25.00	26.67	26.27	23.44	28.39	26.67	30.11	30.00
Average	Precision	78.30	75.00	82.00	77.90	80.57	76.50	80.15	87.50	90.40	89.00
	Recall	15.66	15.10	16.40	15.58	16.11	15.29	15.90	17.50	18.10	17.80
	F-score	26.10	25.17	27.33	25.13	26.85	24.77	26.65	28.47	30.17	29.67

proposed in presented paper where we have found the huge improvement over the most of the category images. Zhou *et al.* [64] have extracted the visual color and texture features by employing discrete cubic partitioning approach to the HSV color image. Firstly, they have used hierarchical mapping on image data based on the set of the hierarchical operators instead of direct image feature extraction. Finally, the extracted visual features have been integrated in 2D CS model which retains most significant features. At the last ranking scheme has been adopted in final image retrieval process. In this scheme, the overall average precision is 80.15%. Mistry *et al.* [65] have presented an image in a compact form by extracting the hybrid features where they have extracted color autocorrelogram, color moments and HSV color histogram image features in spatial domain and frequency domain based image features are extracted by using the stationary and Gabor wavelets decomposed images. Further, to improve the retrieval performance, they have also incorporated the binarized statistical image features and HSV color model based color and edge directivity descriptor (CEDD) are extracted which represents the color and texture visual features of image efficiently. For similarity matching, they have performed various extracted features on several distances like Euclidean, City block, Minkowski and Mahalanobis distances and found that the concatenation of all features (hybrid features) based on Euclidean distance has provided the good retrieval accuracy where dinosaurs and flowers have got the best retrieval rates i.e. 99.00% precisions while the worst retrieval rate (i.e. 80.00% precision) have been given by th food images but the overall average precision is 87.50 %. Ahmed *et al.* [66] have extracted the images features based on the location of the interest key points where the key points represents the various level of invariance property of the image feature descriptors. They have also computed the global features using optimizing sliding window technique and texture features using uniform local binary pattern approach. The efficient image features from above extracted features have been retained based on principle component analysis and support vector machine those will be finally used in the retrieval purpose. This method has been generated better retrieval accuracy as compared to the other state of art methods on which bus images has the best results i.e. 98.00%(precision) while the mountain images has worst results i.e. 84.00%(precision).

The overall average performance of this CBIR system is high (i.e. 90.40% precision) among all other state of art methods including the our presented scheme where the precision of the our proposed method is approximately same with minor difference. The retrieval results in terms of precision, recall and F-score with some standard image retrieval system are shown in Table 4 where buildings, buses, dinosaurs flowers and horses image categories have obtained remarkable and highest precisions, recalls and F-scores than other existing schemes while mountain images has lowest metrics but overall the proposed system has given huge increment in the mean precision, recall and F-score and outperformed over the developed state-of-the-art CBIR systems. Finally, the simulation results for top 20 images retrieved from Corel-1K and GHIM-10K datasets are presented for two best and one worst category images. Figure 18 shows the retrieved results for Corel-1K dataset in which the two best i.e. buses and foods images have 100.00% and 90.00% precisions while the mountains images has worst precision i.e. 60.00%. In case of GHIM-10K dataset, the retrieved results of three image categories are depicted in Fig 19, where trees and sunsets have got the best precisions while the mountain images has reported 65.00% lowest precision.

VI. CONCLUSION

In this paper, a novel CBIR system has been proposed in a hierarchical mode based on three visual features like color, texture and shape for retrieving the most relevant images from the large scale of image dataset where in each stage the retrieval process discards the irrelevant images by filtering process and as a results the search space is reduced in subsequent stages. The proposed image retrieval scheme is not only considering all categories image features like shape, texture and color in hierarchical mode but also using some efficient feature extraction mechanisms on the preprocessed image datasets for improving the retrieval rates. The proposed method has given the facility to reduce the size of intermediate datasets based on the user's interest and returned the desired number of images to the user by the setting the suitable values of parameters K, M and L. So the presented image retrieval system is useful for large size of image datasets because it reduces the search space. The final output of the retrieval system is independent of the weights which are assigned to the individual similarity distances of color, texture and shape features. The presented CBIR system has validated the results on two benchmark datasets and provides satisfactory outcomes. In addition, the experimental results are also compared with some other state-of-the-art method and show that our system outperforms in most of the category images.

CONFLICTS OF INTEREST

This paper supported is under the Dana Impak Perdana (DIP) Grant Scheme DIP-2018-040, University Kebangsaan Malaysia (UKM).

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

The authors confirm that the data supporting the findings of this study are available within the article.

REFERENCES

[1] K. Seetharaman and M. Kamarasan, "Statistical framework for image retrieval based on multiresolution features and similarity method," *Multimedia Tools Appl.*, vol. 73, no. 3, pp. 1943–1962, Dec. 2014.
[2] M. K. Alsmadi, "Content-based image retrieval using color, shape and texture descriptors and features," *Arabian J. Sci. Eng.*, pp. 1–14, Feb. 2020.
[3] S. Gupta, P. P. Roy, D. P. Dogra, and B.-G. Kim, "Retrieval of colour and texture images using local directional peak valley binary pattern," *Pattern Anal. Appl.*, pp. 1–17, Apr. 2020.
[4] Y. D. Mistry, "Textural and color descriptor fusion for efficient content-based image retrieval algorithm," *Iran J. Comput. Sci.*, pp. 1–15, Apr. 2020.
[5] A. K. Bhunia, A. Bhattacharyya, P. Banerjee, P. P. Roy, and S. Murala, "A novel feature descriptor for image retrieval by combining modified color histogram and diagonally symmetric co-occurrence texture pattern," *Pattern Anal. Appl.*, pp. 1–21, Jun. 2019.
[6] A. El-ghazal, O. Basir, and S. Belkasim, "Farthest point distance: A new shape signature for Fourier descriptors," *Signal Process., Image Commun.*, vol. 24, no. 7, pp. 572–586, Aug. 2009.
[7] E. Sokic and S. Konjicija, "Phase preserving Fourier descriptor for shape-based image retrieval," *Signal Process., Image Commun.*, vol. 40, pp. 82–96, Jan. 2016.
[8] F. A. Jalin, N. E. Othman, R. Hassan, A. H. A. Rahman, D. P. D. Sikumbang, and K. A. A. Bakar, "An implementation study of DMM PMIPV6 protocol on dual-stack network environment," *Asia–Pacific J. Inf. Technol. Multimedia*, vol. 7, no. 1, pp. 29–44, 2018.
[9] H. Wu and S. Yan, "Computing invariants of tchebichef moments for shape based image retrieval," *Neurocomputing*, vol. 215, pp. 110–117, Nov. 2016.
[10] P. Srivastava and A. Khare, "Integration of wavelet transform, local binary patterns and moments for content-based image retrieval," *J. Vis. Commun. Image Represent.*, vol. 42, pp. 78–103, Jan. 2017.
[11] A. Khan, J. P. Li, N. Ahmad, S. Sethi, A. U. Haq, S. H. Patel, and S. Rahim, "Predicting emerging trends on social media by modeling it as temporal bipartite networks," *IEEE Access*, vol. 8, pp. 39635–39646, 2020.
[12] J. Pradhan, A. K. Pal, and H. Banka, "A prominent object region detection based approach for CBIR application," in *Proc. 4th Int. Conf. Parallel, Distrib. Grid Comput. (PDGC)*, 2016, pp. 447–452.
[13] M. Abdelhaq, R. Hassan, and R. Alsaqour, "Using dendritic cell algorithm to detect the resource consumption attack over MANET," in *Proc. Int. Conf. Softw. Eng. Comput. Syst.* Springer, 2011, pp. 429–442.
[14] X.-Y. Wang, Z.-F. Chen, and J.-J. Yun, "An effective method for color image retrieval based on texture," *Comput. Standards Interfaces*, vol. 34, no. 1, pp. 31–35, Jan. 2012.
[15] S. M. Adam and R. Hassan, "Delay aware reactive routing protocols for QoS in MANETs: A review," *J. Appl. Res. Technol.*, vol. 11, no. 6, pp. 844–850, Dec. 2013.
[16] R. Manthalkar, P. K. Biswas, and B. N. Chatterji, "Rotation and scale invariant texture features using discrete wavelet packet transform," *Pattern Recognit. Lett.*, vol. 24, no. 14, pp. 2455–2462, Oct. 2003.
[17] Y. Rakvongthai and S. Oraintara, "Statistical texture retrieval in noise using complex wavelets," *Signal Process., Image Commun.*, vol. 28, no. 10, pp. 1494–1505, Nov. 2013.
[18] R. Krishnamoorthi and S. Sathiya Devi, "A multiresolution approach for rotation invariant texture image retrieval with orthogonal polynomials model," *J. Vis. Commun. Image Represent.*, vol. 23, no. 1, pp. 18–30, Jan. 2012.
[19] P. Yang and G. Yang, "Feature extraction using dual-tree complex wavelet transform and gray level co-occurrence matrix," *Neurocomputing*, vol. 197, pp. 212–220, Jul. 2016.
[20] M. K. Hasan, M. M. Ahmed, A. H. A. Hashim, A. Razzaque, S. Islam, and B. Pandey, "A novel artificial intelligence based timing synchronization scheme for smart grid applications," *Wireless Pers. Commun.*, pp. 1–18, Apr. 2020.

[21] G. Raghuwanshi and V. Tyagi, "Texture image retrieval using adaptive tetrolet transforms," *Digit. Signal Process.*, vol. 48, pp. 50–57, Jan. 2016.

[22] S. Islam, O. O. Khalifa, A.-H.-A. Hashim, M. K. Hasan, M. A. Razzaque, and B. Pandey, "Design and evaluation of a multihoming-based mobility management scheme to support inter technology handoff in PNEMO," *Wireless Pers. Commun.*, May 2020.

[23] A. Khan, J. P. Li, S. Nazir, N. Ahmad, N. Varish, A. Malik, and S. H. Patel, "Partial observer decision process model for crane-robot action," *Sci. Program.*, vol. 2020, Feb. 2020.

[24] R.-H. Miao, J.-L. Tang, and X.-Q. Chen, "Classification of farmland images based on color features," *J. Vis. Commun. Image Represent.*, vol. 29, pp. 138–146, May 2015.

[25] S. G. Shaila and A. Vadivel, "Indexing and encoding based image feature representation with bin overlapped similarity measure for CBIR applications," *J. Vis. Commun. Image Represent.*, vol. 36, pp. 40–55, Apr. 2016.

[26] X.-Y. Wang, J.-F. Wu, and H.-Y. Yang, "Robust image retrieval based on color histogram of local feature regions," *Multimedia Tools Appl.*, vol. 49, no. 2, pp. 323–345, Aug. 2010.

[27] G.-H. Liu and J.-Y. Yang, "Content-based image retrieval using color difference histogram," *Pattern Recognit.*, vol. 46, no. 1, pp. 188–198, Jan. 2013.

[28] A. K. Tiwari, V. Kanhangad, and R. B. Pachori, "Histogram refinement for texture descriptor based image retrieval," *Signal Process., Image Commun.*, vol. 53, pp. 73–85, Apr. 2017.

[29] C. Reta, I. Solis-Moreno, J. A. Cantoral-Ceballos, R. Alvarez-Vargas, and P. Townend, "Improving content-based image retrieval for heterogeneous datasets using histogram-based descriptors," *Multimedia Tools Appl.*, pp. 1–31, Apr. 2018.

[30] M. Rahimi and M. Ebrahimi Moghaddam, "A content-based image retrieval system based on color ton distribution descriptors," *Signal, Image Video Process.*, vol. 9, no. 3, pp. 691–704, Mar. 2015.

[31] X.-Y. Wang, Y.-J. Yu, and H.-Y. Yang, "An effective image retrieval scheme using color, texture and shape features," *Comput. Standards Interfaces*, vol. 33, no. 1, pp. 59–68, Jan. 2011.

[32] M. Huang, H. Shu, Y. Ma, and Q. Gong, "Content-based image retrieval technology using multi-feature fusion," *Optik*, vol. 126, no. 19, pp. 2144–2148, 2015.

[33] N. Shrivastava and V. Tyagi, "Region based image retrieval using integrated color, texture and shape features," in *Information Systems Design and Intelligent Applications*. Springer, 2015, pp. 309–316.

[34] J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Spatial color indexing and applications," *Int. J. Comput. Vis.*, vol. 35, no. 3, pp. 245–268, 1999.

[35] K. Guo and D. Labate, "Optimally sparse multidimensional representation using shearlets," *SIAM J. Math. Anal.*, vol. 39, no. 1, pp. 298–318, Jan. 2007.

[36] J. Krommweh, "Tetrolet transform: A new adaptive Haar wavelet algorithm for sparse image representation," *J. Vis. Commun. Image Represent.*, vol. 21, no. 4, pp. 364–374, May 2010.

[37] S. Golomb, "Polyominoes, Charles Scribner's sons, New York, 1965," Tech. Rep., 1994.

[38] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[39] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[40] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, Jul. 1989.

[41] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2005, pp. 886–893.

[42] O. Déniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," *Pattern Recognit. Lett.*, vol. 32, no. 12, pp. 1598–1603, Sep. 2011.

[43] O. L. Junior, D. Delgado, V. Goncalves, and U. Nunes, "Trainable classifier-fusion schemes: An application to pedestrian detection," in *Proc. 12th Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2009, pp. 1–6.

[44] N. Chifa, A. Badri, Y. Ruichek, A. Sahel, and K. Safi, "Efficient combination of color, texture and shape descriptor, using slic segmentation for image retrieval," in *Artificial Intelligence and Computer Vision*. Springer, 2017, pp. 69–80.

[45] T. Kobayashi, A. Hidaka, and T. Kurita, "Selection of histograms of oriented gradients features for pedestrian detection," in *Proc. Int. Conf. Neural Inf. Process.* Springer, 2007, pp. 598–607.

[46] K. Iqbal, M. O. Odetayo, and A. James, "Content-based image retrieval approach for biometric security using colour, texture and shape features controlled by fuzzy heuristics," *J. Comput. Syst. Sci.*, vol. 78, no. 4, pp. 1258–1277, 2012.

[47] M. Rizon, Y. Haniza, S. Puteh, A. Yeon, M. Shakaff, S. A. Rahman, M. M. Rozailan, Y. Sazali, D. Hazri, and M. Karthigayan, "Object detection using geometric invariant moment," Tech. Rep., 2006.

[48] Y. Zhu, L. C. De Silva, and C. C. Ko, "Using moment invariants and HMM in facial expression recognition," *Pattern Recognit. Lett.*, vol. 23, nos. 1–3, pp. 83–91, Jan. 2002.

[49] M. Wang, W.-Y. Chen, and X. D. Li, "Hand gesture recognition using valley circle feature and Hu's moments technique for robot movement control," *Measurement*, vol. 94, pp. 734–744, Dec. 2016.

[50] Y. Sun, G. Wen, and J. Wang, "Weighted spectral features based on local hu moments for speech emotion recognition," *Biomed. Signal Process. Control*, vol. 18, pp. 80–90, Apr. 2015.

[51] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA, USA: Kluwer, 1992, pp. 309–343.

[52] W. Khan, S. Kumar, N. Gupta, and N. Khan, "A proposed method for image retrieval using histogram values and texture descriptor analysis," *Int. J. Soft Comput. Eng.*, pp. 2307–2317, May 2011.

[53] Y. D. Chun, N. C. Kim, and I. H. Jang, "Content-based image retrieval using multiresolution color and texture features," *IEEE Trans. Multimedia*, vol. 10, no. 6, pp. 1073–1084, Oct. 2008.

[54] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

[55] C. Oertel, B. Colder, J. Colombe, J. High, M. Ingram, and P. Sallee, "Current challenges in automating visual perception," in *Proc. 37th IEEE Appl. Imag. Pattern Recognit. Workshop*, 2008, pp. 1–8.

[56] J. Li and J. Z. Wang, "Real-time computerized annotation of pictures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 6, pp. 985–1002, Jun. 2008. [Online]. Available: http://wang.ist.psu.edu/docs/home.shtml

[57] G.-H. Liu, J.-Y. Yang, and Z. Li, "Content-based image retrieval using computational visual attention model," *Pattern Recognit.*, vol. 48, no. 8, pp. 2554–2566, Aug. 2015. [Online]. Available: http://www.ci.gxnu.edu.cn/cbir/Dataset.aspx

[58] E. Walia and A. Pal, "Fusion framework for effective color image retrieval," *J. Vis. Commun. Image Represent.*, vol. 25, no. 6, pp. 1335–1348, Aug. 2014.

[59] A. Irtaza, M. A. Jaffar, E. Aleisa, and T.-S. Choi, "Embedding neural networks for semantic association in content based image retrieval," *Multimedia Tools Appl.*, vol. 72, no. 2, pp. 1911–1931, Sep. 2014.

[60] R. Ashraf, K. Bashir, A. Irtaza, and M. Mahmood, "Content based image retrieval using embedded neural networks with bandletized regions," *Entropy*, vol. 17, no. 6, pp. 3552–3580, May 2015.

[61] J.-M. Guo and H. Prasetyo, "Content-based image retrieval using features extracted from halftoning-based block truncation coding," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1010–1024, Mar. 2015.

[62] S. Zeng, R. Huang, H. Wang, and Z. Kang, "Image retrieval using spatiograms of colors quantized by Gaussian mixture models," *Neurocomputing*, vol. 171, pp. 673–684, Jan. 2016.

[63] 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 Process.*, vol. 11, no. 2, pp. 89–98, Feb. 2017.

[64] Y. Zhou, F.-Z. Zeng, H.-M. Zhao, P. Murray, and J. Ren, "Hierarchical visual perception and two-dimensional compressive sensing for effective content-based color image retrieval," *Cognit. Comput.*, vol. 8, no. 5, pp. 877–889, Oct. 2016.

[65] Y. Mistry, D. T. Ingole, and M. D. Ingole, "Content based image retrieval using hybrid features and various distance metric," *J. Electr. Syst. Inf. Technol.*, vol. 5, no. 3, pp. 874–888, Dec. 2018.

[66] K. T. Ahmed, A. Irtaza, and M. A. Iqbal, "Fusion of local and global features for effective image extraction," *Appl. Intell.*, vol. 47, no. 2, pp. 526–543, 2017.

NAUSHAD VARISH received the B.Sc. (Hons.) and M.C.A degrees from Aligarh Muslim University, Aligarh, India, in 2008 and 2012, respectively, and the Ph.D. degree from the Department of Computer Science and Engineering, Indian Institute of Technology (Indian School of Mines) Dhanbad, India, in 2018. He is currently working as an Associate Professor with the Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (KLEF), India. His research interests include content-based image retrieval (CBIR), wavelets, image processing, and its applications.

ARUP KUMAR PAL (Member, IEEE) received the Ph.D. degree in computer science and engineering from the Indian School of Mines, Dhanbad, in 2011. He is currently working as an Assistant Professor with the Department of Computer Science and Engineering, Indian Institute of Technology (Indian School of Mines). He has around seven years of teaching and research experiences, and contributed a number of research articles in several journals and conference proceedings of National and International reputes. His main research interests include vector quantization, image compression, image cryptosystem, steganography, watermarking, and CBIR.

ROSILAH HASSAN (Senior Member, IEEE) received the B.Sc. degree in electronic engineering from Hanyang University, Seoul, South Korea, the Master of Electrical (M.E.E.) Engineering degree in computer and communication from Universiti Kebangsaan Malaysia (UKM), Malaysia, in 1999, and the Ph.D. degree in mobile communication from the University of Strathclyde, U.K., in May 2008. She is currently an Associate Professor with the Faculty of Information Science and Technology, UKM. She worked as an Engineer with Samsung Electronic Malaysia in Seremban, Malaysia, before joining UKM, in 1997. Besides being a Senior Lecturer, she got experience in management post for the university as the Deputy Director of Academic Entrepreneurship for over seven years. She also is the Head of the Network and Communication Technology (NCT) Lab in her Faculty. Her research interests are in wireless communications, networking, the IoT, and big data. She has had experience as an External Examiner for Master and Ph.D. for both national and international level. She is also an Active Member of IEEE, Malaysia Society for Engineering (MySET), Malaysian Board of Technologists (MBoT), and IET.

MOHAMMAD KAMRUL HASAN (Senior Member, IEEE) received the B.Sc. degree in 2003, the M.Sc. degree in communication engineering from the Department of Electrical and Computer Engineering, International Islamic University, Malaysia, in 2012, and the Doctor of Philosophy (Ph.D.) degree in electrical and communication engineering from the Faculty of Engineering, International Islamic University, Malaysia, in 2016. He is currently working as a Senior Lecturer with the Network and Communication Technology research cluster, Center for Cyber Security, Universiti Kebangsaan Malaysia (UKM). He has been experienced working in electrical and communication industries, after completing his B.Sc. degree. His undertakings included supervision, installation, and configuration of the corporate information-centric network, servers; WAN, WLAN, LAN, fault handling, and programming of the ladder logic design for the industrial LAN connected PLC devices. He has demonstrated excellent research outcomes and received several awards including Gold medal, scholarships, and best paper. He is well versed specialized with elements pertaining to cutting-edge information-centric network; Wireless Communication and Networking, microcontroller-based electronic devices, smart grid WAM systems, microprocessor systems, and the IoT. He has been a Senior Member of the Institute of Electrical and Electronics Engineers, since 2013, and a member of Institution of Engineering and Technology, and the Internet Society. He has served the IEEE IIUM student branch as the Chair, from 2014 to 2016. He has actively participated in many events/workshops/training for the IEEE and IEEE humanity programs in Malaysia. He reviews SCIE journals, including IEEE Access, *Telecommunication Systems*, Hindawi, SAGE, *ETRI Journal*, and *KSII Transactions on Internet and Information Systems*.

ASIF KHAN received the B.Sc. (Hons.) and M.C.A. (Master of Computer Science and Application) degrees from Aligarh Muslim University, India, and the Ph.D. degree (Hons.) in computer science and technology from the University of Electronic Science and Technology of China (UESTC), China, in 2016. He was an Adjunct Faculty with the University of Bridgeport, USA, for China Program in summer 2016. Previously, he was a Visiting Scholar for Big data Mining and Application at the Chongqing Institute of Green and Intelligent Technology (CIGIT), Chinese Academy of Sciences, Chongqing, China. He is currently working as a Postdoctoral Scientific Research Fellow with UESTC. He is also holding a position of Assistant Professor with BSA Crescent University, India. He is a Contributor to many international journals with robotics and vision analyses about the contemporary world in his articles. His interests include machine learning, robotics vision, and new ideas regarding vision based information critical theoretical research. He awarded by UESTC Academic Achievement Award and Excellent Performance Award, from 2015 to 2016.

NIKHAT PARVEEN is currently an Associate Professor with the Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India. She has more than ten years of teaching experience and six years of research experience. Her areas of interest are security software, security testing software engineering, and requirement engineering. She is also working in the area of soft computing and big data security optimization. She has also published and presented articles in refereed journals and conferences. She is a member of ACM, CSI, and IAENG.

DEBRUP BANERJEE received the B.E. degree from Shivaji University, India, the M.S. degree in computer science from Hampton University, USA, and the Ph.D. degree in electrical and computer engineering from Old Dominion University, USA, in 2017. He is currently an Assistant Professor with the Koneru Lakshmaiah Education Foundation, Guntur, India. He is involved in teaching soft computing, machine learning, and computer vision. His current research interests include applied machine learning and computer vision. He has also worked for a renowned multinational company in the past as a Software Engineer.

VIDYULLATHA PELLAKURI received the degree from S. V. University, Tirupathi, the Master of Technology degree in computer science from JNTU-Hyderabad, and the Ph.D. degree from the Department of Computer Science and Engineering, Koneru Lakshmaiah Education and Foundation, Guntur. She is currently an Associate Professor with the Department of Computer Science and Engineering, Koneru Lakshmaiah Education and Foundation. She is having national patents on IoT. Her research and publication interests include artificial intelligence, machine learning, and soft computing techniques. Her present research interests include image analysis, big data analytics, and the IoT. Her research has been chronicled in over 30 journal publications and international conferences. She is a Life Time Member of IAENG, IACSIT, and SDIWC.

AMIN UL HAQIS received the M.S. degree in computer science. He is currently pursuing the Ph.D. degree from the School of Computer Science and Engineering, UESTC, China. He has a vast academic, technical, and professional experience in Pakistan. He is also a Lecturer with Agricultural University Peshawar, Pakistan. His research interests include machine learning, medical big data, the IoT, e-health and telemedicine, and algorithms. He is the author of some research articles. He is associated with Wavelets Active Media Technology and Big Data.

IMRAN MEMON received the M.E. degree in information security engineering from the University of Electronic Science and Technology of China, in 2013. His research interests include location-based services, network security, and location-aware computing.