Journal of King Saud University – Computer and Information Sciences xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

A hybrid CBIR system using novel local tetra angle patterns and color moment features

Umer Ali Khan, Ali Javed *

Department of Software Engineering, University of Engineering and Technology, Taxila 47050, Pakistan

ARTICLE INFO

Article history:
Received 2 December 2021
Revised 6 July 2022
Accepted 6 July 2022
Available online xxxx

Keywords:
Local tetra angle patterns (LTAPs)
Texture features
Color features
Hybrid features
Genetic algorithm
Support vector machines

ABSTRACT

The exponential growth of communal media platforms including Facebook and Twitter, and the accessibility of low-cost digital capturing devices have generated an enormous number of multimedia content including images. Effective handling of such massive image collection has boosted the development of content-based image retrieval (CBIR) approaches. Researchers have suggested both machine learning and non-learning-based techniques for CBIR. However, machine learning-based methods outperform the non-learning-based methods in the CBIR domain. The CBIR demands the development of reliable descriptors to attain the most appropriate images from the depository and better address the semantic gap problem. To counter these problems, we suggest a novel second-order Local Tetra Angle Patterns (LTAP) to better capture the texture features from the image. LTAPs are computed from adjacent pixels of 0°, 45°, 90°, and 135° using the second-order directional derivatives. Further, we propose a hybrid feature vector by concatenating LTAPs and RGB color features and using the genetic algorithm (GA) to select the finest appropriate features that enhance the image retrieval performance of our system. We employed our hybrid descriptor with the GA to optimize the support vector machine (SVM) for the image classification task and used the Chi-square quadratic distance measure to determine the resemblance between the query image and the images in the repository. Experimental results on three standard datasets including the Corel 1 k, Oxford flower, and CIFAR-10 indicate the effectiveness of the presented system over the contemporary CBIR methods.

1. Introduction

The evolution of social media platforms and economical digital capturing devices have flooded cyberspace with multimedia content such as images, videos, audios, etc., in the last decade. Management and analysis of such enormous data to retrieve the most relevant content have motivated the researchers to develop effective content-based image retrieval methods (Shrivastava and Tyagi, 2014). Several techniques have been developed related to image retrieval based on the image contents (Mansoori et al., 2013; Liuand and Yang, 2013). Image retrieval techniques can be mainly divided into two groups. The first category is annotation-based image retrieval (ABIR), and the second category is content-based image retrieval (CBIR). ABIR techniques have some drawbacks e.g., human annotation is subjective, the laborious activity of manual image annotation, difficult to annotate the complex images. On the other hand, CBIR avoids the above issues that are related to image retrieval using annotation-based keywords. Because CBIR retrieves the images by means of their visual contents. (Lin et al., 2011; Hsiao et al., 2010).

A large number of CBIR applications have been deployed successfully in numerous domains for example surveillance systems (Robles-Serrano et al., 2021; Hafiane et al., 2006), topographical information application (Kareem Jebur, 2021; Rahman, 2006), remote detecting and sensing (Ye et al., 2019), architectural design (Bharadi and Meena, 2015), object recognition systems (Zhang et al., 2016), medical image retrieval systems (Rao and Prasad, 2021; Ashraf et al., 2020; K., a., k., n., & d., r. d., 2021; Naik, et al., 2009; Chang et al., 2013), etc. Content-based image retrieval methods use various features for example texture, shape, color, or any other spatial data that can be extracted from an image (Ashraf et al., 2020; Shukran et al., 2021; Machhour and Nasri, 2020; Elalami, 2014; Jiang and Kim, 2021; Al-Jubouri and H., 2020; Jiji

* Corresponding author.
E-mail address: ali.javed@uettaxila.edu.pk (A. Javed).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2022.07.005
1319-1578/

Please cite this article as: U.A. Khan and A. Javed, A hybrid CBIR system using novel local tetra angle patterns and color moment features, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2022.07.005

and Durairaj, 2015; Wang et al., 2011; Celebi and Aslandogan, 2004; Rahman, 2006). CBIR methods are further separated into two broad categories that use local and global image features to retrieve the images (Shrivastava and Tyagi, 2014; Shen and Wu, 2013; Talib et al., 2013). However, CBIR methods are still facing different challenging tasks e.g., the semantic gap between the visual contents of high-level and low-level features of an image, classification of images into relevant semantic class, class imbalance issues, and visual diversity of images that exists within a semantic class as well as between different semantic classes (Liuand and Yang, 2013; Wang et al., 2014; Zou et al., 2016; Bian et al., 2017).

Researchers in the field of CBIR are practicing extensively both the machine learning and non-learning methods. These methods work in both the online and offline CBIR architectures. In online CBIR architecture, any new image can be provided to find similar images from the images dataset, and hence computation of images to the user query greatly depends upon the relevancy of the query image to the dataset images. Similarly, new images can be added dynamically to the dataset which increases the images collection of the repository continuously and results in increased computational cost and image retrieval time. Whereas, in off-line CBIR architecture, the query image is taken from the available collection of images in the repository without including any new images in real-time. Hence, the offline model overcomes the limitations of the on-line based approach. Therefore, we also adopted the offline CBIR architecture to evaluate and report the retrieval results of our system. (Hameed and Abdulhussain, 2021; Bhardwaj* et al., 2020; Kaur and Devendran, 2020; Zagoris et al., 2021; Saoudi and Jai-Andaloussi, 2021; Lonescu and Ralescu, 2004; Naik, et al., 2009; Ballerini, 2009) implements CBIR systems using the non-learning-based methods by extracting different low-level features for example texture, shape, color, spatial features, etc. These methods compute the similarity by considering the distance between the query image and dataset images. However, the retrieval performance of these systems degrades due to the increasing semantic gap between the low- and high-level features and the increase of semantic classes in the dataset. Moreover, the computational cost of non-learning-based methods also increases in the presence of largescale datasets. To overcome these issues related to non-learning-based methods, investigators in the field of CBIR employed different machine learning techniques that used both the supervised (Fadaei et al., 2016; Pourreza and Kourosh, 2016; Montazer and Giveki, 2015; Ashraf et al., 2016; Fakheri et al., 2013; Sankar et al., 2017) and unsupervised (Zhu et al., 2017; Chaudhuri et al., 2016) methods with the assistance of different low-level feature descriptors for images classification and retrieval tasks. Though these machine learning-based methods outperform the non-learning methods. However, these machine learning-based methods are computationally more complicated as compared to non-learning-based CBIR methods.

The MPEG-7 Visual group divides the features into different categories that are texture, color, and shape to represent a digital image. In comparison with other image features such as texture and shape, color features are very stable and robust. It is not sensitive to rotation, translation, and scale changes. Moreover, the color feature calculation is relatively simple. Texture provides surface characteristics for the analysis of many types of images including natural scenes, remotely sensed data, and biomedical modalities. It plays an important role in the human visual system for recognition and interpretation. The texture is imperative and provides a prominent visual representation of an image by giving information about the spatial arrangement of intensity values in an image. The real-world images are composed of different kinds of objects having different surface patterns within the image. The surface pattern of an object in an image or the entire image is known as the texture (Yue et al., 2011; Chang et al., 2001; Lindsay et al., 2008).

In this research work, a supervised machine learning method through the fusion of color and texture features is presented that uses the genetic algorithm to enhance support vector machine performance for image categorization task and uses Chi-square quadratic distance measure to measure the resemblance between the query and images of dataset. In the current article, we present novel second-order Local tetra angle patterns (LTAPs) features to attain the texture features from the input images. We employ the second-order derivates over the first-order derivatives for LTAP extraction as higher-order derivatives are more sensitive to noise (Baochang et al., 2010). The image retrieval performance of first-order derivative features is lower than that of second-order derivative features. This performance improvement exists because second-order derivative features extract more information as compared to first-order derivative features. However, higher-order derivatives beyond the second order are unable to provide more effective image representation and thus the performance of the image retrieval system decreases. LTAPs are computed using 0°, 45°, 90°, and 135° neighboring pixels using the 2nd order directional derivatives. Our proposed method includes the diagonal pixels to compute the novel LTAP directional derivative, magnitude, and the direction of the central pixel with respect to its neighboring pixels. Further, we propose a fusion feature vector by concatenating the LTAP and three-color moment features for image description. We employed the genetic algorithm to choose the most appropriate features that enhance the image retrieval performance of the suggested system. Our main contributions to achieve an effective CBIR system are the following:

- We propose a novel local tetra angle patterns descriptor using the 2nd order directional derivatives to extract more distinctive low-level texture features from the input images.
- We employ a Genetic algorithm-based method to pick the most relevant features.
- We optimize the multiclass support vector machine through a genetic algorithm (GA).
- We propose a fusion feature vector of low-level texture and color image features to lessen the semantic gap between the low-level and high-level image features.
- Extensive experimentation was performed on three standard datasets i.e., Corel 1 k, oxford flower, and CIFAR-10 to show the efficacy of the proposed system over contemporary CBIR techniques.

2. Literature review

This section presents a thorough exploration of prevailing state-of-the-art CBIR methods. Prevailing CBIR methodologies have employed numerous local and global features for CBIR. Local descriptors have proved to be more robust than global descriptors under different conditions i.e., rotation, viewpoint changing, scaling, etc. On the contrary, global descriptors have less computational cost than local descriptors. We categorized our literature review section as computer vision-based methods and AI-based methods.

2.1. Computer Vision-Based methods

Existing methods have used different low-level features i.e., color, shape, texture, spatial layout, interest point features i.e., SIFT, SURF, HOG, FAST, Harris edge detector, and binary features i.e., BRISK and FREAK (Latif et al., 2019; Iakovidou et al., 2019). Liu et al. (Liu et al., 2011) used color and edge orientation features to

compute the microstructure descriptor (MSD) but this approach lacks to capture the relationship between dissimilar objects. Xia et al. (Xia et al., 2016) suggested a cloud-based model to secure the user's image privacy using a watermarking technique. The employed watermarking technique was unable to handle the distortions in geometric features. Methods used in (Ali et al., 2016; ElAlami, 2011; Lin et al., 2009) extracted the color features using YCbCr and HSV color space, salient point features using SIFT, spatial layout using the orientation of edge pixels, and employed the artificial neural network (ANN) and support vector machine (SVM) for classification and image retrieval task.

Existing CBIR approaches (Irtaza et al., 2018; Irtaza et al., 2014; Latha and Raj, 2019; Zang et al., 2018; Scott et al., 2017; Wei and Wang, 2019; Youssef, 2012; Rao et al., 2011; Wang et al., 2001; Chen et al., 2005) have employed various hybrid feature vectors to retrieve the images. Feature fusion of color and texture was used with the GA to optimize the class imbalance problem (Irtaza et al., 2018; Irtaza et al., 2014; Latha and Raj, 2019). High-level image depiction was employed using the object model (OM), but this methodology suffers from the dimensionality curse issue (Zang et al., 2018). Youssef et al. (Youssef, 2012) extracted the color and texture features through the fusion of region-based vector codebook sub-band clustering (RBSC) and curvelet transform. Rao et al. (Rao et al., 2011) used a fusion feature descriptor comprising of gray level co-occurrence matrix (GLCM) to extract the texture features, an edge gradient vector to extract the shape features, and SIFT and SURF descriptors to extract the visual features. SIFT and SURF descriptors caused the overfitting on behalf of image retrieval once employed with a support vector machine. Wang et al. (Wang et al., 2001) employed semantic classification and region-based matching through wavelet transform to match the images but this method also suffered from the imprecise texture classification. Chen et al. (Chen et al., 2005) proposed a clustering-based unsupervised method to handle the problem of uncertain texture modeling and achieved better results as compared to (Wang et al., 2001) however, unable to obtain accurate segmentation of the clusters.

Existing CBIR approaches (Song et al., 2014; Ashraf et al., 2014; Huang and Dai, 2003; Mehmood et al., 2018; Jhanwar et al., 2004; Tian et al., 2014; Zeng et al., 2016) have widely used a bag of visual words model (BoVWM) for image retrieval, object recognition, and image annotation tasks. BoVWM uses image features to construct the visual words and develop a visual words library. Histograms of this visual words library or bag of visual words are computed to find the images similar to the user query image. However, bag of visual words models also faces different issues such as these models ignore spatial information and lacking semantic meaning (Song et al., 2014; Ashraf et al., 2014). In (Lin et al., 2011; Ashraf et al., 2020; Elalami, 2014), edge orientation histogram and color SIFT features were used to develop a codebook to retrieve and classify the images. In (Khan et al., 2021), DWT and color features were used with the SVM, artificial neural network, and Gabor filters to classify and retrieve the images. Adaptive searching, region matching, texture features, and motif co-occurrence were used with the SVM and bag of visual words model to categorize and retrieve the images in (Huang and Dai, 2003; Mehmood et al., 2018; Jhanwar et al., 2004). SIFT, local binary features, and curvelet transform were used to extract the features while most similar principles of highest priority and k-nearest neighbors were used for image classification and retrieval (Tian et al., 2014; Zeng et al., 2016).

2.2. AI-Based methods

Existing CBIR approaches (Scott et al., 2017; Wei and Wang, 2019; Shi et al., 2018; Alzubi et al., 2017; Guo et al., 2010) have widely employed various deep learning-based methods using the convolutional neural networks (CNN) for image retrieval. A hashing algorithm was used for features extraction and a pairwise matrix model was presented that used a deep learning framework to formulate an objective function to learn the binary image representation for the classification and retrieval of images (Shi et al., 2018). This hashing algorithm-based method is further enhanced by the addition of online and offline learning methods. In the offline learning stage, visual and texture features were extracted to represent the images. The offline learning stage used a static dataset while the online learning stage used the updated dataset continuously (Alzubi et al., 2017). Two parallel unsupervised CNN-based models were used to extract the features. This model boosts the image retrieval performance and reduces the image retrieval time as this method used different quantized levels to reduce the image features representation (Guo et al., 2010). These deep learning-based CBIR methods have shown remarkable performance for CBIR but suffer due to the complex training process and a massive amount of training time (Scott et al., 2017; Wei and Wang, 2019).

Existing CBIR methods (Tao et al., 2006; Lai and Chen, 2011; Saadatmand-Tarzjan and Moghaddam, 2007; Irtaza et al., 2015) have employed various relevance feedback (RF) methods with a genetic algorithm, evolutionary computation, and fusion of texture and color features for image retrieval and classification tasks. However, RF-based methods are unable to get user support due to the involvement of users in the image retrieval process. Moreover, relevance feedback-based methods are computationally more time-consuming because these methods adopt the iterative methodology and need multiple iterations to improve the results. As iterations increase, the user feedback also increases to obtain the required images similar to the query image. Therefore, content-based image retrieval methods that work autonomously on the basis of user query images are gaining more popularity among researchers. Prevailing CBIR methods have certain limitations such as intra-class similarity, the gap between high-level and low-level features, and lack of feature wdb "nothing":ttagdel "_run.sc": wdbmethodologyIn this section, we discuss the proposed method for CBIR. We extract the texture features using our novel LTAP and RGB color features using the first three lower-order color moments. We then obtain a fusion feature vector by concatenating the texture and color features. We optimize this fused feature vector using the GA to choose the most suitable features and use them to train the SVM for image classification. We compute the similarity between images of the dataset and query image using the Chi-square quadratic distance measure. Fig. 1 shows the architecture of our proposed method. In our proposed methodology, we employed the off-line CBIR architecture and use all images from the pre-stored standardized images dataset that are freed from any noise. Therefore, we don't need to apply any filter or pre-treatment of images to extract the features. However, we applied the min–max normalization scheme to normalize our feature descriptor as we concatenated two different types of features (texture and color) to form the hybrid feature vector.*3.1. Features extraction*

To develop a robust CBIR method that can provide better retrieval performance on multiple image repositories, we need to develop an effective feature descriptor. In this work, we present a fusion feature descriptor comprising our novel LTAP features and three-color moments capable of effective representation of input images for CBIR. The detail of the features extraction process is given in subsequent sections.

3.1.1. Local tetra Angle patterns features. In this work, we suggest a novel descriptor LTAPs to extract the texture features from the input image. LTAPs present the texture spatial structure using

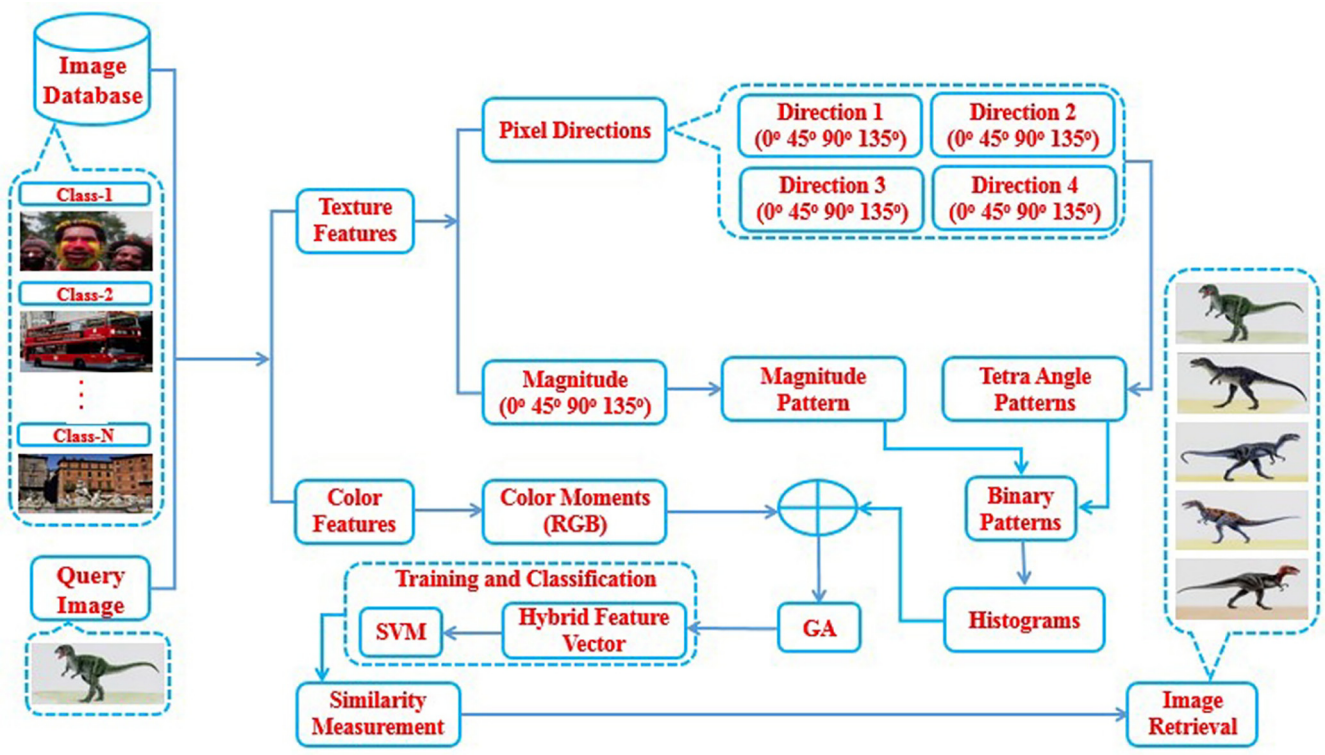

Fig. 1. Architectural drawing of proposed content-based image retrieval methodology.

the direction of central gray pixel Z_o. In order to compute the LTAPs, we compute 2nd order derivatives at 0°, 45°, 90°, and 135° neighboring pixels around the central pixel Z_o. We use the 2nd order derivatives for features extraction as the 2nd order derivatives compute more detailed discriminative features from an image than the first-order derivatives. Moreover, higher-order derivatives beyond the second-order are more sensitive to noise (Baochang et al., 2010) and degrade the performance of the image retrieval system.

To compute the 2nd order derivatives at 0°, 45°, 90°, and 135° directions, first we need to compute the first-order derivatives at 0°, 45°, 90°, and 135° directions around the central pixel Z_o. First-order derivatives at 0°, 45°, 90°, and 135° directions for an image $I(Z)$ around the central pixel Z_o can be denoted as $I'_\theta(Z)$ where θ = 0°, 45°, 90°, and 135°. Let Z_o be the central pixel in the image $I(Z)$. Where Z_i shows eight neighboring pixels around the central pixel Z_o and i = 1, 2, 3, …, 8 denotes the index of eight immediate adjacent pixels as shown in Fig. 2.

Four first order derivatives at 0°, 45°, 90° and 135° around the central pixel are computed as follows:

$$I'_{0^\circ}(Z_o) = I(Z_4) \; - -I(Z_o) \tag{1}$$

$$I'_{45^\circ}(Z_o) = I(Z_3) \; - -I(Z_o) \tag{2}$$

$$I'_{90^\circ}(Z_o) = I(Z_2) \; - -I(Z_o) \tag{3}$$

$$I'_{135^\circ}(Z_o) = I(Z_1) \; - -I(Z_o) \tag{4}$$

Moreover, the directions of central pixel Z_0 with respect to its neighboring pixels are computed as follows:

$$\begin{aligned} &1, \sum \left(I'_{0^\circ}(Z_i), I'_{45^\circ}(Z_i)\right) \geqslant 0 \; and \sum \left(I'_{90^\circ}(Z_i), I'_{135^\circ}(Z_i)\right) \geqslant 0. \\ &I'_{Dir}(Z_i) = 2, \sum \left(I'_{0^\circ}(Z_i), I'_{45^\circ}(Z_i)\right) < 0 \; and \sum \left(I'_{90^\circ}(Z_i),\; I'_{135^\circ}(Z_i)\right) \geqslant 0. \\ &3, \sum \left(I'_{0^\circ}(Z_i), I'_{45^\circ}(Z_i)\right) < 0 \; and \sum \left(I'_{90^\circ}(Z_i), I'_{135^\circ}(Z_i)\right) < 0. \\ &4, \; \sum \left(I'_{0^\circ}(Z_i), I'_{45^\circ}(Z_i)\right) \geqslant 0 \; and \sum \left(I'_{90^\circ}(Z_i), I'_{135^\circ}(Z_i)\right) < 0. \end{aligned} \tag{5}$$

Eq. (5) shows that possible directions for the central pixel are 1, 2, 3, or 4 and eventually image converts into four values or four directions. While 2nd order directional derivatives LTAPs″ (Z_o) are computed as follows:

$$LTAPs''(Z_o) = \{f(I'_{Dir}(Z_o), I'_{Dir}(Z_1)), f(I'_{Dir}(Z_o), I'_{Dir}(Z_2)) \ldots f(I'_{Dir}(Z_o), I'_{Dir}(Z_i))\} \tag{6}$$

Where i = 1, 2, 3 … 8.

$$f(I'_{Dir}(Z_o), I'_{Dir}(Z_i)) = \{0, \quad if \quad I'_{Dir}(Z_o) = I'_{Dir}(Z_i) I'_{Dir}(Z_i), Else\} \tag{7}$$

From the Eqs. (6) and (7), we obtain the 8-bit tetra pattern code for each central pixel. In this way, we get four different patterns, and then each pattern is converted into three 8-bit binary patterns. Thus, we get a total of 12 (4 × 3) binary patterns. Suppose central pixel direction $I'_{Dir}(Z_o)$ computed from Eq. (7) is "1" and LTAPs″ can be computed in three binary patterns as follows:

$$LTAPs''|_{Direction=2,3,4} = \sum_{i=1}^{8} I(Z_o) \times f(LTAPs''(Z_i))| \underset{Direction=2,3,4}{} \tag{8}$$

$$LTAPs''|_{Direction=\Phi} = \left\{ \begin{matrix} 1, & if\ LTAPs''(Z_i) = \Phi \\ 0, & Else \end{matrix} \right\} \tag{9}$$

Fig. 2. Arrangement of eight neighbors around central pixel Z_o.

where Φ = 2,3,4.

In the same way as in Eqs. (8) and (9), binary patterns for directions 2, 3, and 4 are computed and a gray value is assigned to the corresponding pixel. Guo et al. (Corel 1K dataset, available on http://wang.ist.psu.edu/docs/related/, accessed on June 05, 2021) in their work used the magnitude component to propose a magnitude local binary pattern (LBP) and proved that the magnitude component computes more useful information. Therefore, we also computed the 13th binary pattern as a magnitude pattern (MP) of first-order derivatives at 0°, 45°, 90°, and 135° around the central pixel as:

$$M_{I'(Z_i)} = \sqrt{(I'_{0^\circ}(Z_i))^2 + (I'_{45^\nu}(Z_i))^2 + (I'_{90^\circ}(Z_i))^2 + (I'_{135^\circ}(Z_i))^2} \tag{10}$$

$$MP = \sum_{i=1}^{8} I(Z_o) \times f\left(M_{I'(Z_i)} - M_{I'(Z_o)}\right) \tag{11}$$

$$M_{I'(Z_i)} - M_{I'(Z_o)} = \begin{cases} 1, & if \ M_{I'(Z_i)} > M_{I'(Z_o)} \\ 0, & Else \end{cases} \tag{12}$$

Using Eq. (5), the direction of the central pixel is computed and further used to calculate the tetra pattern bits. Fig. 3 demonstrates the computation of tetra pattern bits.

After identifying Binary Patterns (BPs) that are 12 texture binary patterns and 13th texture magnitude pattern, we compute the histogram of texture binary patterns (TBPs) and texture magnitude pattern (TMP) using Eq. (14). We obtain texture feature vector [TFV] through the concatenation of TBPs and TMP histograms as shown in Eq. (17).

$$H_S(l) = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} f\left(BP_S(i, j), l\right) \tag{13}$$

Where $l \in [0, s(s-1)+2]$ is the maximum LBPs pattern value, $M \times N$ represents the size of the input image $x = LBP_S(i, j)$, $y = l$ and function f is defined as:

$$f(x, y) = \begin{cases} 1, & x = y \\ 0, & otherwise \end{cases} \tag{14}$$

Fig. 3 shows that LTAPs bit is coded to "0" if the central and neighboring pixel has the same direction. If the direction of the central and neighboring pixel is different then the bit is coded as the direction of the neighboring pixel. Using the same methodology, LTAP bits are computed for the central pixel having directions 2, 3, and 4 as shown in Fig. 3.

An example of 2nd order LTAPs computation is illustrated in Fig. 4 where the central pixel is marked with red color having a

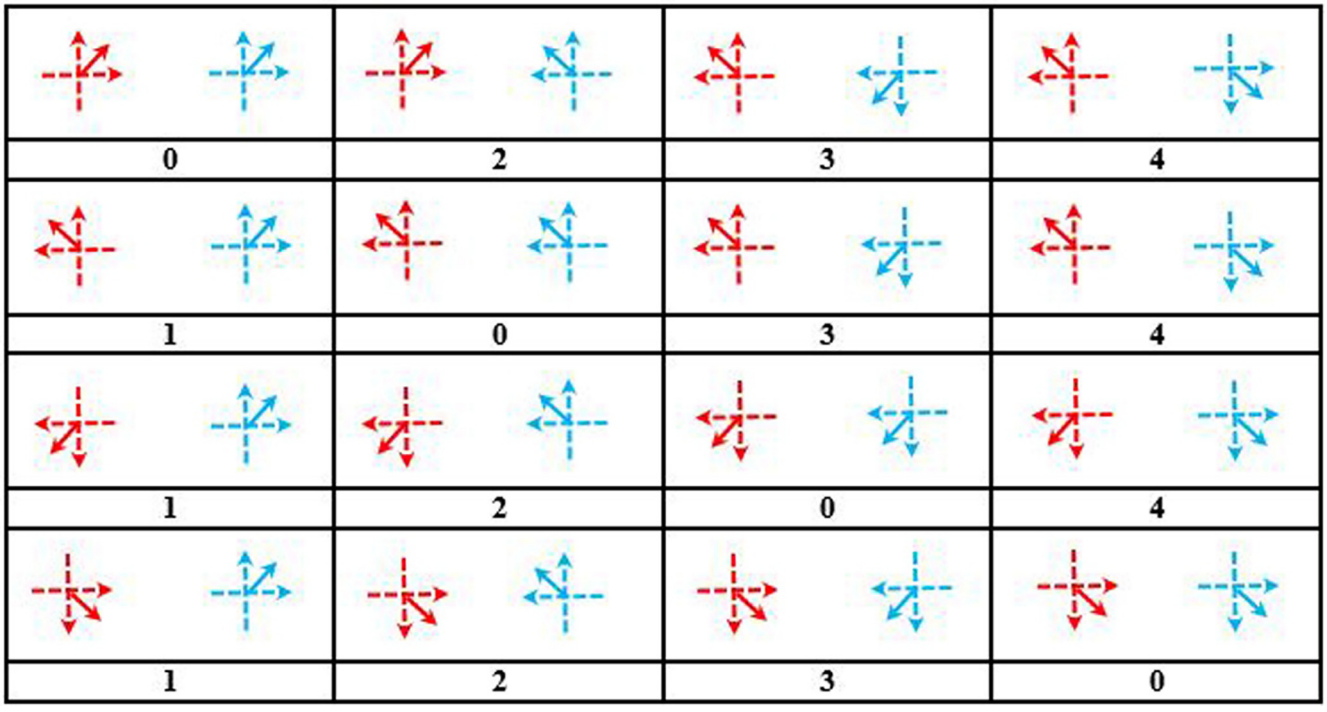

Fig. 3. Computation of Local Tetra Angle Patterns (LTAPs) bits for a central pixel having directions "1, 2, 3, and 4". Red arrow shows the direction of the central pixel and the blue arrow shows the direction of neighboring pixels.

direction "1". We compute the direction of all eight neighboring pixels marked with blue color using the Eq. (5). Then we compute the 2nd order LTAPs using the Eq. (6). We can see in Fig. 4 that when the direction of the central pixel is different from its neighboring pixel, LTAPs are coded with the direction of the neighboring pixel and when the direction of the central pixel is the same as its neighboring pixel, LTAPs are coded with "0". After computing the LTAPs, each pattern is coded in three binary patterns in all directions. To compute the binary pattern in direction "2", it is coded with "1" whereas all other directions are coded with "0". Similarly, binary patterns are computed for other directions i.e., "3" and "4". In order to compute the magnitude pattern, the magnitude of the central pixel and its neighboring pixels are compared using the Eq. (12). The magnitude pattern is coded with bit "1" when the magnitude of the central pixel is less than the magnitude of its neighboring pixel and coded with "0" when the magnitude of the central pixel is greater than the magnitude of its neighboring pixel. The process of texture features extraction using the LTAPs is presented in Algorithm 1.

Our proposed features extraction method is freed from any geometrical transformation of pixels since we always compute the derivatives in four fixed directions from the central pixel. Moreover, considering a fact that uniform patterns are less sensitive to noise, we compute the LTAP by defining a single bin for each uniform pattern and one bin for all non-uniform patterns. This makes our LTAP features robust to noise and geometrical transformations while representing the images for retrieval purposes. Thus, we argue that the proposed LTAP features can reliably be used to capture the distinctive traits from the input images even in the presence of noise and geometrical transformations.

Algorithm 1 Extraction process of texture features from gray scale image using LTAPs.

3.1.2. Color moments features. We also extracted the color moments features to further enhance the performance of LTAPs for CBIR. For this purpose, we employed the mean, standard deviation, and skewness (first 3 lower-order color moments) to attain the color features from an RGB image. Overall, 9 color moments are computed for the red, green, and blue channels for each RGB image. The extraction process of color moment features from the input RGB image is provided in Algorithm 2.

3.1.3. Hybrid feature vector. After extracting the LTAPs and color moments features, we normalized our feature descriptor through a min–max normalization scheme to scale both the feature descriptors' values in the range of 0 to 1. Then we obtained the hybrid feature (HFV) vector by the concatenation of texture feature vector (TFV) and color feature vector (CFV). The dimension of HFV

Begin
Input: *f(x, y)*: Gray Scale Image
Output: TFV (Texture Feature Vector)
For each image **from** image dataset
1. First order derivatives are computed: $I'^{o}_{\theta}(Z_o) = I(Z_i) - I(Z_o)$
Where i = 1, 2, 3 . . . 8 and θ = 0°, 45°, 90° and 135°
2. Directions of central pixel Z_i where i = 0 are computed:

$$I'_{Dir}(Z_i) = \left\{ \begin{array}{l} 1,\ \sum (I'_{0^\circ}(Z_i), I'_{45^\circ}(Z_i)) \geqslant 0\ and\ \sum (I'_{90^\circ}(Z_i), I'_{135^\circ}(Z_i)) \geqslant 0. \\ 2,\ \sum (I'_{0^\circ}(Z_i), I'_{45^\circ}(Z_i)) < 0\ and\ \sum (I'_{90^\circ}(Z_i), I'_{135^\circ}(Z_i)) \geqslant 0. \\ 3,\ \sum (I'_{0^\circ}(Z_i), I'_{45^\circ}(Z_i)) < 0\ and\ \sum (I'_{90^\circ}(Z_i), I'_{135^\circ}(Z_i)) < 0. \\ 4,\ \sum (I'_{0^\circ}(Z_i), I'_{45^\circ}(Z_i)) \geqslant 0\ and\ \sum (I'_{90^\circ}(Z_i), I'_{135^\circ}(Z_i)) < 0. \end{array} \right\}$$

3. 2nd order directional derivative patterns (LTAPs) are computed:

$$LTAPs''(Z_o) = \{f(I'_{Dir}(Z_o), I'_{Dir}(Z_1)), f(I'_{Dir}(Z_o), I'_{Dir}(Z_2)) \ldots f(I'_{Dir}(Z_o), I'_{Dir}(Z_i))\}$$

$$f(I'_{Dir}(Z_o), I'_{Dir}(Z_i)) = \left\{ \begin{array}{ll} 0, & if\ I'_{Dir}(Z_o) = I'_{Dir}(Z_i) \\ I'_{Dir}(Z_i), & Else. \end{array} \right\}$$

4. *LTAPs* are converted into binary patterns where direction of central pixel $I'_{Dir}(Z_o) = 1$:

$$LTAPs''|_{\text{Direction}=2,\ 3,\ 4} = \sum_{i=1}^{8} I(Z_o) \times f(LTAPs''(Z_i))|_{\text{Direction}=2,\ 3,\ 4}$$

$$LTAPs''|_{Direction=\Phi} = \left\{ \begin{array}{ll} 1, & if\ \ LTAPs''(Z_i) = \Phi \\ 0, & Else \end{array} \right.$$

Where $\Phi = 2, 3, 4$
5. Magnitude Pattern (MP) is computed:

$$M_{I'(Z_i)} = \sqrt{(I'_{0^\circ}(Z_i))^2 + (I'_{45^\circ}(Z_i))^2 + (I'_{90^\circ}(Z_i))^2 + (I'_{135^\circ}(Z_i))^2}$$

$$MP = \sum_{i=1}^{8} I(Z_o) \times f\left(M_{I'(Z_i)} - M_{I'(Z_o)}\right)$$

$$M_{I'(Z_i)} - M_{I'(Z_o)} = \left\{ \begin{array}{ll} 1, & if\ \ M_{I'(Z_i)} > M_{I'(Z_o)} \\ 0, & Else \end{array} \right.$$

6. TFV is computed through histogram as following:

$$H_S(l) = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} f(LBP_S(i, j), l)$$

where $l \epsilon [0, s(s-1)+2]$ is the maximum LBPs pattern value, $M \times N$ represents the size of input image, $x = LBP_S(i,j)$, $y = l$ and function f is defined as:

$$f(x,y) = \left\{ \begin{array}{l} 1, x = y \\ 0, otherwise \end{array} \right.$$

$$TFV = [\, Hist_{TBPs} || Hist_{MBP}]$$

End For
End

is [1 × 1288] for each semantic class comprising 100 different images. While the dimension of TFV for each image of the semantic class is [1 × 1024] and the dimension of CFV for each image of the semantic class is [1 × 264] for the Corel image dataset. The hybrid feature vector representation is as under where 'n' represents the total number of images in a dataset.

$$\begin{bmatrix} a_{(1,1)}a_{(1,2)}a_{(1,3)}a_{(1,4)}\ldots a_{(1,1024)} \\ a_{(2,1)}a_{(2,2)}a_{(2,3)}a_{(2,4)}\ldots a_{(2,1024)} \\ a_{(3,1)}a_{(3,2)}a_{(3,3)}a_{(3,4)}\ldots a_{(3,1024)} \\ \vdots \\ a_{(n,1)}a_{(n,2)}a_{(n,3)}a_{(n,4)}\ldots a_{(n,1024)} \\ [TFV] \end{bmatrix} \Bigg\| \begin{bmatrix} b_{(1,1)}b_{(1,2)}b_{(1,3)}b_{(1,4)}\ldots b_{(1,264)} \\ b_{(2,1)}b_{(2,2)}b_{(2,3)}b_{(2,4)}\ldots b_{(2,264)} \\ b_{(3,1)}b_{(3,2)}b_{(3,3)}b_{(3,4)}\ldots b_{(3,264)} \\ \vdots \\ b_{(n,1)}b_{(n,2)}b_{(n,3)}b_{(n,4)}\ldots b_{(n,264)} \\ [CFV] \end{bmatrix}$$

$$= \begin{bmatrix} c_{(1,1)}c_{(1,2)}c_{(1,3)}c_{(1,4)}\ldots c_{(1,1288)} \\ c_{(2,1)}c_{(2,2)}c_{(2,3)}c_{(2,4)}\ldots c_{(2,1288)} \\ c_{(3,1)}c_{(3,2)}c_{(3,3)}c_{(3,4)}\ldots c_{(3,1288)} \\ \vdots \\ c_{(n,1)}c_{(n,2)}c_{(n,3)}c_{(n,4)}\ldots c_{(n,1288)} \\ [HFV] \end{bmatrix}$$

Algorithm 2 Extraction process of color features from RGB image using **color** moments.

Begin.
Input: RGB Image,p_{ij} : R, G, B Component of Input Image.
Output: CFV (Color Feature Vector).
1. E_i : Mean value,σ_i : Standard deviation, s_i : Skewness.
For each image **from** image dataset.
While.
2. $E_i = \frac{1}{N}\sum_{N}^{j=1} p_{ij}$
E_i is computed for all three color-channels of RGB image.
$\sigma_i = \sqrt{\frac{1}{N}\sum_{N}^{j=1}(p_{ij} - E_i)^2}$
2. σ_i is computed for all three color-channels of RGB image.
3. $s_i = \sqrt[3]{\frac{1}{N}\sum_{N}^{j=1}(p_{ij} - E_i)^3}$
s_i is computed for all three color-channels of RGB image.
End While.
4. $CFV = E_i \| \sigma_i \| s_i$
End For.
End.

3.2. Optimization of SVM through GA

The fact that the existing learning-based methods (Ashraf et al., 2020; Fadaei et al., 2016; Pourreza and Kourosh, 2016; Montazer and Giveki, 2015; Ashraf et al., 2016; Fakheri et al., 2013; Sankar et al., 2017; Zhu et al., 2017; Chaudhuri et al., 2016; Khan et al., 2021; Irtaza et al., 2018; Irtaza et al., 2014) have proved to be more effective over similarity matching based methods for CBIR motivated us to develop a learning-based system. For this purpose, we employed the genetically optimized SVM to classify and retrieve the images that closely resembles to the query image. Genetic algorithm provides best combination of features to the SVM for training. Genetic algorithm uses a fitness function given in Eq. (13) to analyze the distance and standard deviation between the parent and children features and provides the features with minimum distance and standard deviation to the SVM for training purpose. Selection process of best features through the GA is given in Algorithm 3.

$$F(x) = Minimum\_Standard\ Deviation\left(Minimum\_Dist\left|\sqrt{\sum_{i=1}^{2}(p_i - c_i)^2\sum_{j=2}^{1}(p_i - c_j)^2}\right|\right) \quad (16)$$

Algorithm 3 Selection of best children as input to Support Vector Machine using GA.

Begin.
Input IP: Initial Population, ***Pc*:** Population Chromosomes.
CFV: Color Feature Vector, **TFV:** Texture Feature Vector.
N: Number of Genes, **M:** Population Size, **σ:** Standard Deviation.
Output GP: Generated Population.
IP = [CFV^j + TFV^k], Where: $j \in \{1, 2 \ldots N\}$ and $k \in \{1, 2 \ldots N\}$.
$Pc = [\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3} \ldots \alpha_{i,N}]$, Where: $i \in \{1, 2 \ldots M\}$.
$P^* = \{\Phi\}$.
While (Sizeof (GP) ≠ Sizeof (IP)).
Select: p_1, p_2 from IP.
Crossover: p_1, p_2 to obtain h_1, h_2.
Mutate: if (h_1==h_2) to obtain c_1, c_2.
else $c_1 = h_1$ & $c_2 = h_2$.
if (Chi_dis(p_1, c_1) + Chi_dis(p_2, c_2) ≤ Chi_dis(p_1, c_2) + Chi_dis (p_2, c_1)).
{.
if ($\sigma(c_1) < \sigma(p_1)$) then $GP = GP\ U\ \{c_1\}$ **else** $GP = GP\ U\ \{p_1\}$.
if ($\sigma(c_2) < \sigma(p_2)$) then $GP = GP\ U\ \{c_2\}$ **else** $GP = GP\ U\ \{p_2\}$.
}.
else.
{.
if ($\sigma(c_1) < \sigma(p_2)$) then $GP = GP\ U\ \{c_1\}$ **else** $GP = GP\ U\ \{p_2\}$.
if ($\sigma(c_2) < \sigma(p_1)$) then $GP = GP\ U\ \{c_2\}$ **else** $GP = GP\ U\ \{p_1\}$.
}.
End While.
IP = GP.
SVM = IP.
End.

Feature vector obtained from the GA are then fed to the SVM. We designed a multiclass SVM classifier using one against all approach (OAA) to classify the image into relevant semantic class of the repository. We used the SVM tuned on the quadratic kernel function for image classification. For OAA representation, consider the two classes where one class belongs to the query image class and second includes the images of all remaining classes (Stejić et al., 2003; Yildizer et al., 2012). We can represent the binary classification scenario as follows:

$$\{(x_i\ y_i)\}_{i=1}^{N}\ y_i = \{+1, -1\} \quad (17)$$

where x_i denote input features and y_i denote the corresponding labels while weight vector 'w' and bias 'b' are used to determine the hyperplane as follows:

$$W^T . x + b = 0 \quad (18)$$

SVM divides the input set in two classes as:

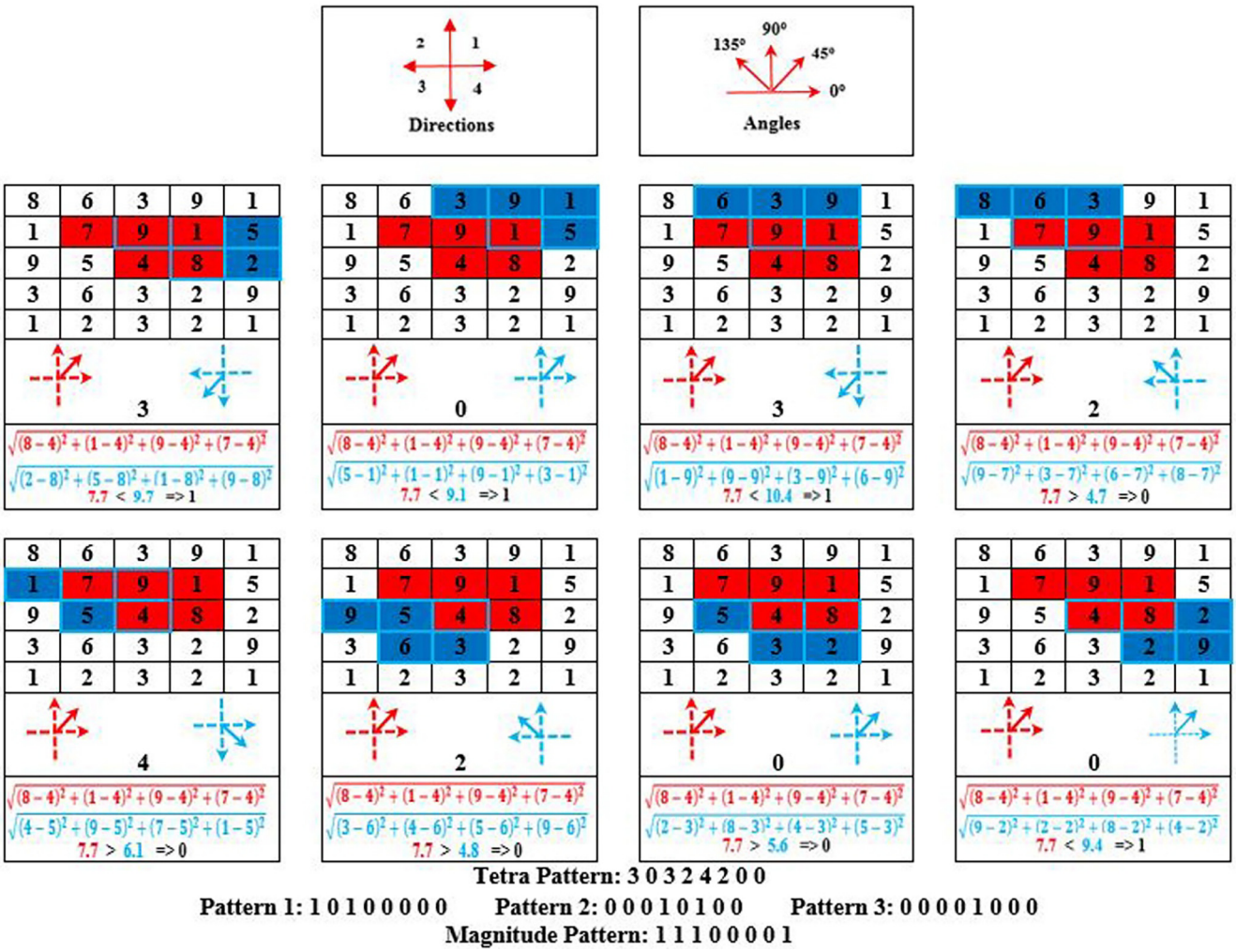

Fig. 4. Example to compute local tetra pattern at θ = 0°, 45°, 90°, and 135° and magnitude pattern.

$$W^T \cdot x + b \geqslant +1 \tag{19}$$

$$W^T \cdot x + b \leqslant -1 \tag{20}$$

SVM used the kernel version of Wolfe dual problem with Lagrangian multiplier α_i for binary classification as:

$$W(\alpha) = \sum_{i=1}^{m} \alpha - \frac{1}{2} \sum_{i,\ j=1}^{m} \alpha_i \alpha_j\, y_i\, y_j\, K(x_i\ ,\ x_j) \tag{21}$$

where $\alpha_i \geq 0$ and.

$$\sum_{i,j=1}^{m} a_i y_i = 0 \tag{22}$$

SVM hyperplane decision function $g(x)$ is represented as:

$$F(x) = Sgn[\ g(x)\ ] \tag{23}$$

$$g(x) = \sum_{i=1}^{m} \alpha_i y_i\ K(x\ ,\ x_i) + b \tag{24}$$

A high value of $g(x)$ denotes a more accurate prediction and vice versa. In the same way, support vector machines are trained for all the semantic classes to form a multiclass SVM classifier.

3.3. Image retrieval

The quadratic distance measure is a statistical measure to test the goodness of fit. Pearson's chi-square quadratic distance measure (Chen and Zhang, 2006) is a statistical method that is often used to measure the similarity between two feature matrices. In our proposed method, we also employed the chi-square method to find similarities between images of the dataset and query images. Images that resemble the user query image are retrieved. Chi-square is computed as follows:

$$X^2 = \frac{1}{2} \sum_{i=1}^{n} \frac{(x_i\ -\ y_i)^2}{(x_i + y_i)^2} \tag{25}$$

where '*x' and 'y'* denote the hybrid feature vector of user query image and repository images respectively, '*d*' is used to form the distance matrix between dataset images and the query image. A smaller value of distance matrix *'d'* means more similarity between the repository images and query images. Fig. 5 illustrates the computation of feature descriptors and the image retrieval process.

4. Experiments and results

In this section, we provide details of the various experimentations designed to measure the performance of our system. We assessed our suggested method on three standard datasets, namely

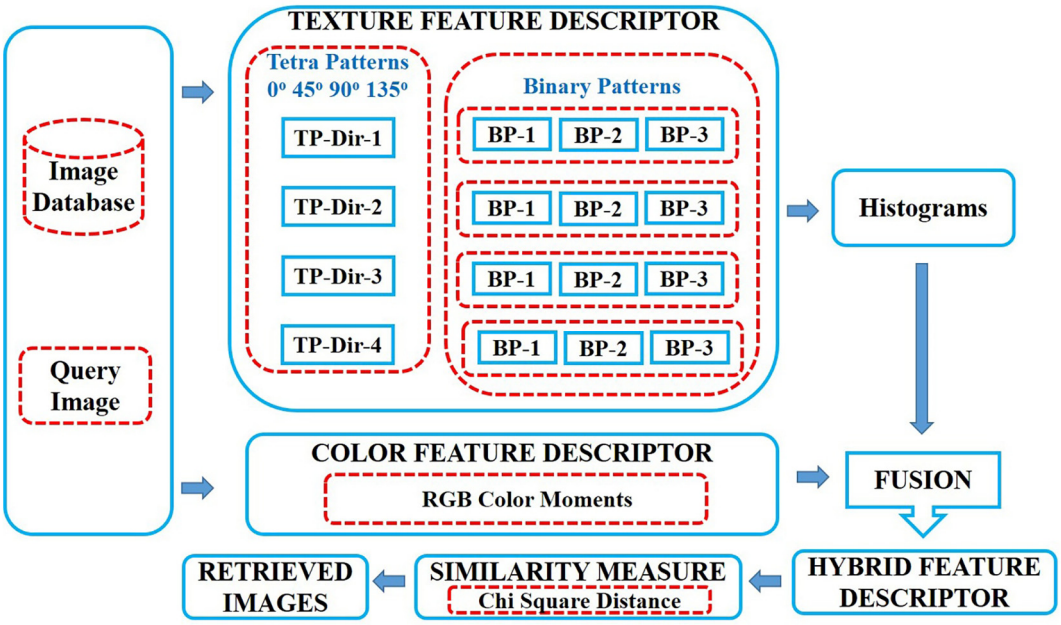

Fig. 5. Computation of feature descriptors and image retrieval process illustration.

Corel 1 k (Flower, 2021), Oxford flower (CIFAR-10 dataset, available on https://www.cs.toronto.edu/~kriz/cifar.html, accessed on June 29, 2021), and CIFAR-10 (Cord et al., 2008). Moreover, the performance of the suggested methodology is investigated in terms of mean precision and recall ensuring the effectiveness of our proposed methodology over the existing contemporary methods. Details of the data sets are also provided in the next section.

4.1. Dataset

We selected three standard and diverse datasets to investigate the performance of our method. Example images from each dataset are shown in Figs. 6, 7, and 8. Corel 1 k data set (Flower, 2021) contains 1000 color images and is divided into 10 different semantic classes such as African People, Beach, Monuments, Buses, Food, Dinosaurs, Elephants, Flowers, Horses, and Mountains. Each class contains 100 images having a resolution of either 256 × 384 or 384 × 256. Corel 1 K dataset consists of a heterogeneous collection of images ranging from natural scenes to outdoor sports to different animals and thus this dataset is suitable for the evaluation of an image retrieval system. Oxford flower dataset (CIFAR-10 dataset, available on https://www.cs.toronto.edu/~kriz/cifar.html, accessed on June 29, 2021) contains 1360 color images and is

Fig. 6. Sample images from 10 semantic classes of Corel 1 k dataset.

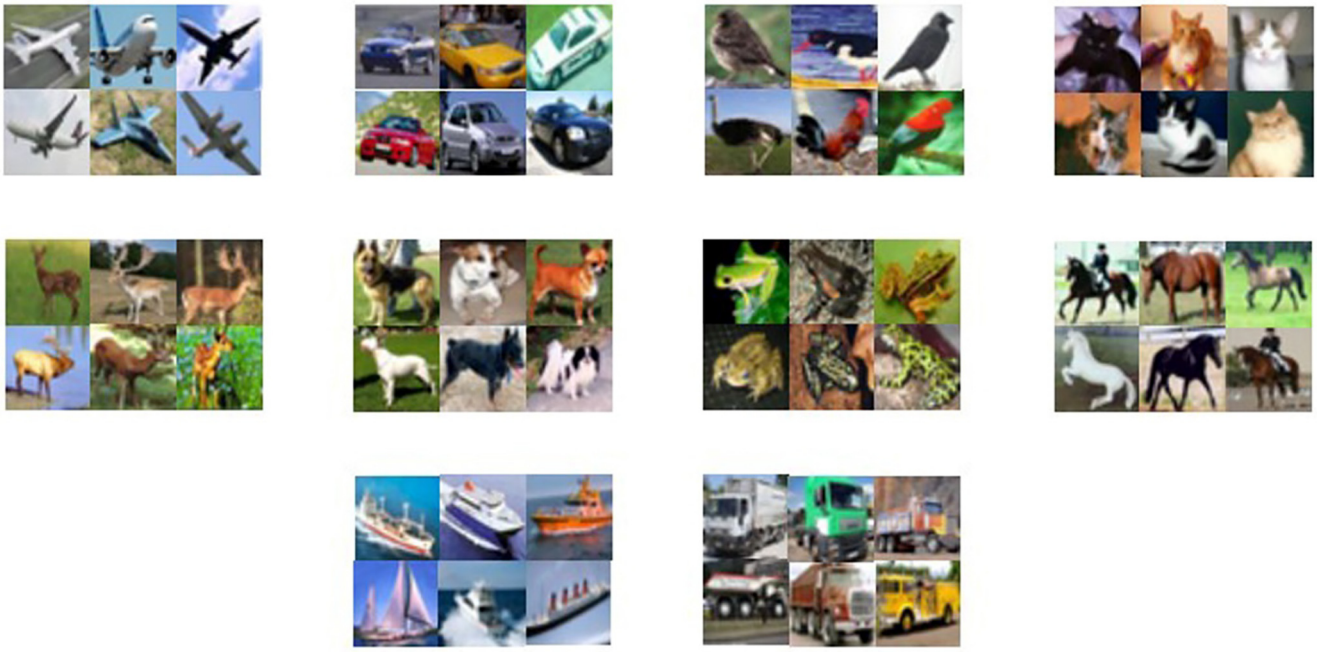

Fig. 7. Sample images from 10 semantic classes of CIFAR-10 dataset.

Fig. 8. Sample images from 17 semantic classes of the Oxford flower dataset.

divided into 17 different semantic classes such as Buttercup, Clot's Food, Daffodil, Daisy, Dandilion, Firtillary, Iris, Pansy, Sunflower, Windowflower, Snowdrop, Lillyvaley, Bluebell, Crocus, Tigerlilly, Tulip, and Cowslip. Each class contains 80 images of different sizes. This dataset is divided into three predefined datasets such as a training dataset comprising 40 images per class, a validation dataset consisting of 20 images per class, and a testing dataset comprising 20 images per class. Oxford flower dataset contains large variations with different poses and light conditions and therefore this dataset is diverse enough to be used for the evaluation of any CBIR system. While CIFAR-10 data set (Cord et al., 2008) consists of 60,000 small-size images and 10 semantic classes such as Airplane, Automobile, Bird, Cat, Deer, Dog, Frog, Horse, Ship, and Truck. Each class contains 6000 images having a resolution of 32×32. This dataset is further divided into a training dataset comprising 5000 images per class and a testing dataset comprising 1000 images per class. This dataset also contains images with high similarity and different variations and can effectively be used to evaluate the image retrieval system.

4.2. Performance evaluation parameters

In the CBIR system, the determination of the retrieval performance of the system is a critical problem. A number of different approaches are used by researchers for the evaluation of their sys-

tems. We used here the most common evaluation approaches for the evaluation of our system, i.e., precision, recall, and precision-recall graph (PR Graph). Precision measures the performance of the system against the models returned by the system while recall measures the performance of the system against all models in the dataset. We computed the precision and recall as follows:

$$Precision = \frac{TP}{TP + FP} \tag{26}$$

$$Recall = \frac{TP}{TP + FN} \tag{27}$$

TP means true positive and represents those images that are returned by the system and are similar to the query image. FP means false positive and represents those images returned by the system but are not similar to the query image. FN means false negative and represents those images that are not returned by the system but are similar to the user image of a query. Precision or recall alone is insufficient for the performance evaluation of any system. So, both measures are used to evaluate the performance of the proposed system.

P-R Graph is another most widely used measure, it is easy to interpret and contains the information about both precision and recall of the system together. A PR graph is also used to compare the performance of the proposed system with other systems. We prefer these measures because they explicitly show the number of relevant and irrelevant images against the user query image and are also widely adopted by the contemporary CBIR methods. Therefore, these parameters are supportive of fairly evaluating the performance of the CBIR system and also helpful for a reliable performance comparison of different contemporary content-based image retrieval methods.

4.3. Experimental outcomes and discussions of proposed method

The purpose of this experiment is to assess the performance of our method on a variety of diverse CBIR datasets. We planned a three-stage experiment to evaluate our method in different datasets. In the first phase of this experimentation, we assessed the performance of our system on the Corel 1 K dataset using the evaluation parameters mean accuracy and mean recall. The results of our experimentation are described in Table 1. The mean precision and recall evaluation parameters are calculated when the top 20 and 80 images are obtained from each semantic category. We achieved the mean precision of 0.945 and 0.65 at retrieval of 20 and 80 images, respectively. Similarly, we attained the mean recall of 0.189 and 0.52 at retrieval of 20 and 80 images, respectively. It can be seen from Table 1 that we obtained the best results for the semantic classes of elephants and dinosaurs because these two semantic classes have high inter- and intra-class dissimilarity and highly discriminatory characteristics. Whereas we score poorly in the semantic classes of Beach and African Peoples because these two semantic classes have high intra- and intra-class similarities and have less distinct features. As a result, false positives increase, and the image retrieval performance of the suggested system decreases.

In 2nd phase of this experimentation, we assessed the performance of the proposed methodology on the Oxford flower dataset that contains 17 semantic classes with 80 images in each class. In this stage, we tested the execution of the suggested system on a different number of images returned against the query image, and the results are reported in Table 2. It can be observed from Table 2 that as the number of images attained by the suggested system increases, the performance of the system decreases and vice versa. It happens because when the number of images retrieved against the query image increases then features similarity along with inter- and intra-class similarity also increases. As a result, false positive images attained by the system increase, and the performance of our system declines. On the contrary, as the number of images returned by the proposed system decreases, the performance of the proposed system increases. This happens because, as the number of images returned by the proposed system decreases, feature similarity, inter- and intra-class dissimilarity decrease.

In the 3rd phase of this experimentation, we assessed the performance of the proposed system on the CIFAR-10 dataset, and the outcomes are described in Table 3 for retrieval of the top 20 images. This dataset comprises 10 semantic classes with 6000 images in each class. It can be realized that the retrieved outcomes in terms of mean precision and mean recall values are different due to the similarity of features and inter- as well as intra-class similarity as described earlier. It can be seen from Table 3 that we attained the mean precision value of 0.931 and the mean recall value of 0.186. It can also be observed from Table 3 that we got the best results for automobile, bird, cat, and fish classes whereas, got the worst results for frog, dog, and truck classes. As similarity increases, the number of false positives also increases, and hence performance of the proposed system decreases. This performance difference in different semantic classes occurs because as the similarity of features increases, inter- and intra-class similarity increases, false positive images returned from the proposed system also increase, and ultimately retrieval performance of our system decreases.

4.4. Assessment of proposed method using Corel 1 K data set

This experimentation is designed to match the performance of our system with other state-of-the-art image retrieval methodologies on the Corel 1 K dataset. Results in terms of precision and recall of this experimentation are stated in Tables 4 and 5 according to the semantic class, respectively. Comparative analysis of the

Table 1
Proposed method outcomes of evaluation parameters on Corel 1 K dataset on retrieval of 20 and 80 images.

No of Returned Images	Precision		Recall	
	20 Images	80 Images	20 Images	80 Images
African People	0.90	0.64	0.18	0.51
Beach	0.90	0.63	0.18	0.50
Buildings	0.90	0.65	0.18	0.52
Buses	01	0.66	0.20	0.53
Dinosaurs	01	0.69	0.20	0.55
Elephants	01	0.66	0.20	0.53
Flowers	0.95	0.68	0.19	0.54
Horses	0.95	0.64	0.19	0.51
Mountains	0.90	0.65	0.18	0.52
Food	0.95	0.64	0.19	0.51
Mean	**0.945**	**0.65**	**0.189**	**0.52**

Table 2
Proposed method results of evaluation parameters on Oxford Flower dataset on retrieval of different number of images.

No of Returned Images	Precision							Recall						
	20	30	40	50	60	70	80	20	30	40	50	60	70	80
Butter cup	0.90	0.83	0.85	0.85	0.75	0.66	0.65	0.22	0.31	0.45	0.45	0.56	0.60	0.65
Colts foot	0.90	0.93	0.70	0.70	0.75	0.80	0.65	0.22	0.35	0.40	0.45	0.56	0.65	0.65
Daffodil	0.95	0.83	0.85	0.85	0.65	0.66	0.80	0.23	0.31	0.45	0.45	0.49	0.60	0.80
Daisy	0.95	0.90	0.90	0.80	0.80	0.75	0.75	0.23	0.33	0.45	0.50	0.60	0.55	0.75
Dandelion	0.80	0.93	0.85	0.80	0.75	0.60	0.55	0.22	0.35	0.45	0.50	0.56	0.52	0.55
Fritillary	0.90	0.90	0.80	0.76	0.70	0.70	0.65	0.22	0.33	0.45	0.45	0.52	0.50	0.65
Iris	0.95	0.93	0.90	0.85	0.80	0.75	0.70	0.23	0.35	0.45	0.45	0.60	0.55	0.70
Pansy	0.95	0.90	0.90	0.80	0.75	0.70	0.65	0.23	0.33	0.45	0.50	0.56	0.60	0.65
Sunflower	0.90	0.80	0.75	0.70	0.75	0.66	0.80	0.22	0.30	0.42	0.45	0.56	0.60	0.80
Windflower	0.85	0.83	0.75	0.70	0.70	0.70	0.65	0.23	0.31	0.42	0.45	0.52	0.45	0.65
Snowdrop	0.90	0.80	0.75	0.70	0.65	0.80	0.65	0.22	0.30	0.42	0.45	0.49	0.65	0.65
Lily Valley	0.90	0.76	0.85	0.80	0.60	0.80	0.65	0.22	0.30	0.45	0.50	0.45	0.65	0.65
Blue Bell	0.85	0.83	0.75	0.70	0.65	0.66	0.80	0.23	0.31	0.40	0.45	0.49	0.60	0.80
Crocus	0.95	0.90	0.90	0.80	0.80	0.80	0.75	0.23	0.33	0.45	0.50	0.55	0.65	0.75
Tiger Lily	0.90	0.76	0.85	0.75	0.75	0.60	0.65	0.22	0.30	0.45	0.50	0.56	0.52	0.65
Tulip	0.90	0.83	0.80	0.75	0.75	0.70	0.70	0.22	0.31	0.45	0.50	0.56	0.60	0.70
Cowslip	0.95	0.90	0.85	0.80	0.75	0.75	0.70	0.23	0.33	0.42	0.45	0.56	0.55	0.70
Mean	**0.906**	**0.856**	**0.824**	**0.771**	**0.726**	**0.711**	**0.691**	**0.225**	**0.321**	**0.437**	**0.471**	**0.541**	**0.579**	**0.691**

Table 3
Proposed method results of evaluation parameters on CIFAR-10 dataset.

Semantic Class	**Precision**	**Recall**
Airplane	0.937	0.187
Automobile	0.979	0.196
Bird	0.979	0.196
Cat	0.979	0.196
Deer	0.937	0.187
Dog	0.875	0.175
Frog	0.833	0.166
Horse	0.937	0.187
Ship	0.979	0.196
Truck	0.875	0.175
Mean	**0.931**	**0.186**

Table 4
Comparative analysis of proposed method precision value on retrieval of top 20 images.

Semantic Class	Africa	Beach	Buildings	Buses	Dinosaurs	Elephants	Flowers	Horses	Mountains	Food	Mean
Proposed Method	0.900	0.900	0.900	1.00	1.00	1.00	0.950	0.950	0.900	0.950	**0.945**
Khan et al. (2021)	0.850	0.850	0.850	0.950	1.00	1.00	0.950	0.950	0.850	0.850	**0.910**
Latha and Raj (2019)	0.700	0.910	0.870	0.920	1.00	0.830	0.810	0.800	0.760	0.920	**0.852**
Irtaza et al. (2018)	0.830	0.720	0.860	1.00	0.970	0.820	0.860	0.820	0.690	0.900	**0.847**
Mehmood et al. (2018)	0.730	0.740	0.770	0.940	0.970	0.900	0.910	0.930	0.810	0.810	**0.851**
Zeng et al. (2016)	0.720	0.650	0.700	0.890	1.00	0.700	0.940	0.910	0.720	0.780	**0.801**
Elalami (2014)	0.726	0.593	0.587	0.891	0.772	0.993	0.702	0.928	0.856	0.562	**0.761**
Tian et al. (2014)	0.750	0.380	0.540	0.970	0.990	0.660	0.920	0.870	0.590	0.620	**0.729**
Irtaza et al. (2014)	0.650	0.600	0.620	0.850	0.930	0.650	0.940	0.770	0.730	0.810	**0.755**
Ali et al. (2016)	0.690	0.540	0.630	0.890	0.980	0.480	0.920	0.890	0.470	0.700	**0.719**

Table 5
Comparative analysis of proposed method recall value on retrieval of top 20 images.

Semantic Class	Africa	Beach	Buildings	Buses	Dinosaurs	Elephants	Flowers	Horses	Mountains	Food	Mean
Proposed Method	0.180	0.180	0.180	0.200	0.200	0.200	0.190	0.190	0.180	0.190	**0.189**
Khan et al. (2021)	0.170	0.170	0.170	0.190	0.200	0.200	0.190	0.190	0.170	0.170	**0.182**
Latha and Raj (2019)	0.070	0.091	0.087	0.092	0.100	0.083	0.081	0.080	0.076	0.092	**0.085**
Irtaza et al. (2018)	0.170	0.140	0.170	0.200	0.190	0.160	0.170	0.160	0.140	0.180	**0.169**
Mehmood et al. (2018)	0.147	0.148	0.154	0.189	0.194	0.180	0.182	0.187	0.162	0.162	**0.170**
Zeng et al. (2016)	0.145	0.130	0.121	0.178	0.200	0.141	0.189	0.183	0.144	0.157	**0.158**
Elalami (2014)	0.161	0.103	0.141	0.126	0.148	0.109	0.163	0.129	0.144	0.136	**0.161**
Tian et al. (2014)	0.150	0.080	0.110	0.190	0.130	0.130	0.180	0.170	0.120	0.130	**0.146**
Irtaza et al. (2014)	0.130	0.120	0.124	0.170	0.186	0.130	0.188	0.154	0.146	0.162	**0.151**
Ali et al. (2016)	0.139	0.108	0.127	0.179	0.197	0.176	0.184	0.178	0.146	0.141	**0.157**

proposed method and other state-of-the-art methods in terms of precision and recall is given in Figs. 9 and 10. Tables 4 and 5 clearly describe that the proposed system achieved improved results in terms of precision and recall. More specifically, the proposed system achieved the best results. We obtained the mean precision of 0.945 and the mean recall of 0.189. Our prior work (Khan et al., 2021) performed second best with mean precision and mean recall of 0.910 and 0.182, respectively. Whereas (Ali et al., 2016) achieved the lowest mean precision of 0.719 and (Latha and Raj, 2019) achieved the lowest mean recall of 0.085. Tables 4 and 5 also show that the proposed system obtained different results in terms of precision and recall for different semantic classes. This difference occurs in the results of different semantic classes depending on the features of respective classes. The semantic class having highly distinct and dissimilar features provides better results as compared to the semantic class having high similarity within the features.

We also generated the PR curve to show the efficacy of the presented method over other contemporary CBIR methods as shown in Fig. 11 to show the efficacy of our technique over other state-of-the-art CBIR methodologies. From the PR curve shown in Fig. 10, we can see that the presented method obtained the best performance for CBIR, while our prior method (Khan et al., 2021) obtained the second-best results with a small margin. Whereas,

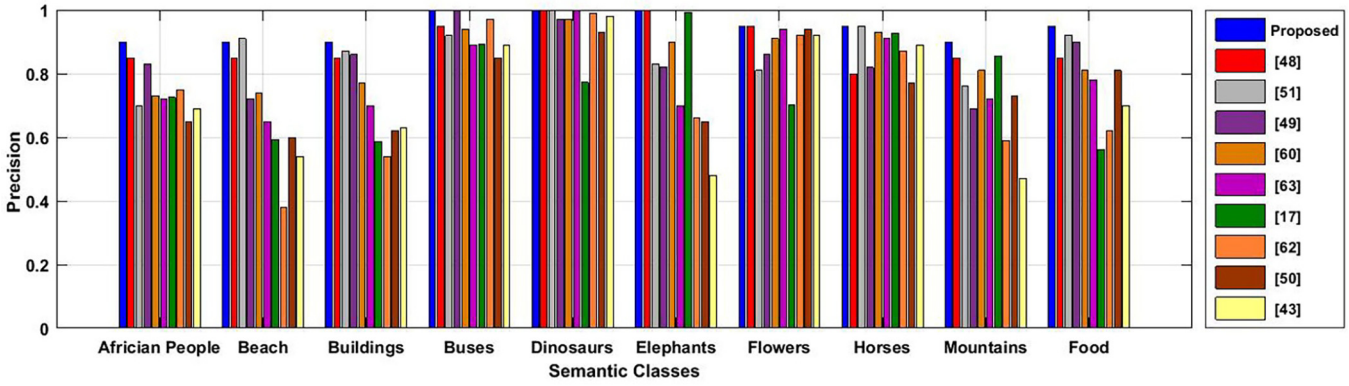

Fig. 9. Classwise comparative analysis of precision outcomes of suggested method with other CBIR methods.

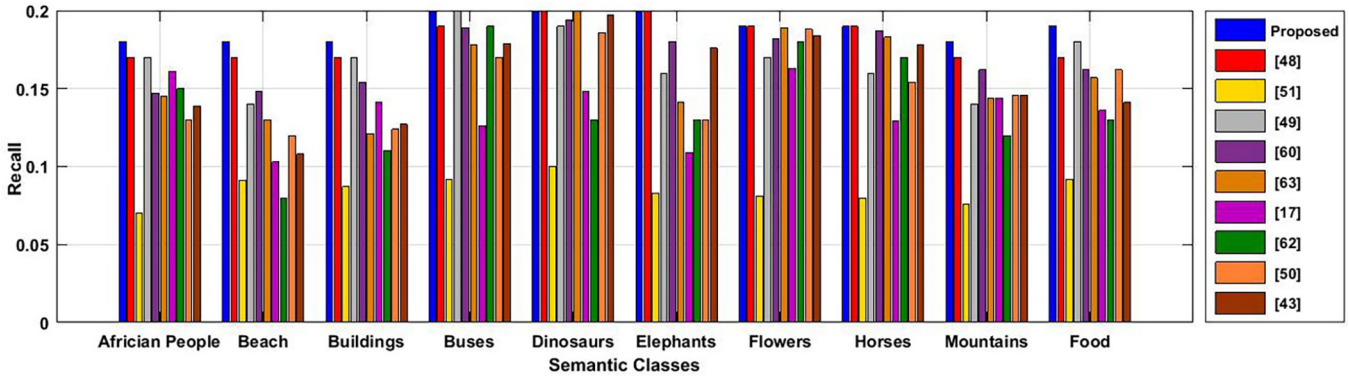

Fig. 10. Classwise comparative analysis of recall outcomes of the suggested method with other CBIR methods.

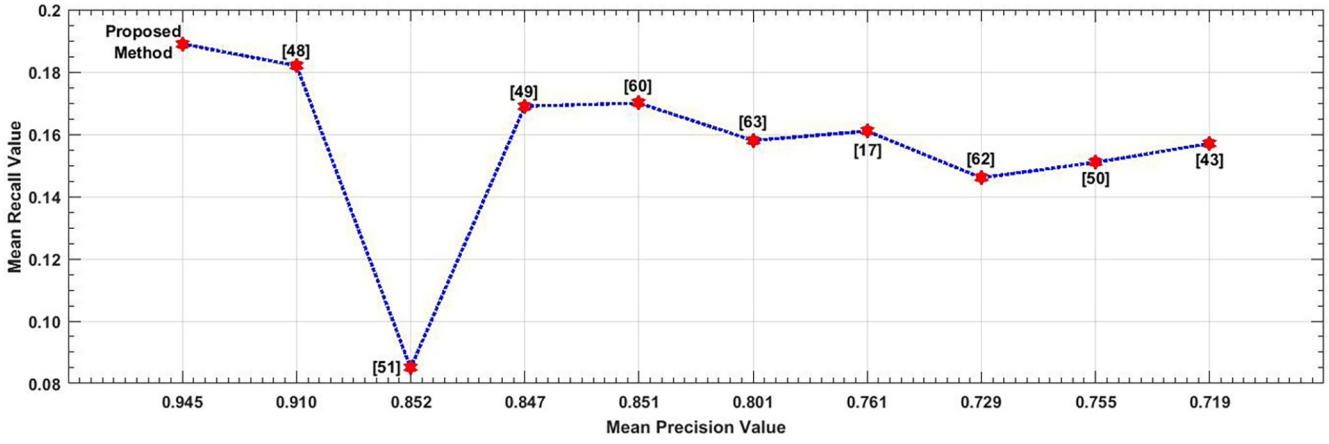

Fig. 11. Comparative analysis of mean precision and mean recall outcomes of suggested method with other CBIR methods.

(Latha and Raj, 2019) was the worst performer method for CBIR on Corel 1 k image dataset. We can see from the PR curve that our proposed method outperforms all the comparative state-of-the-art CBIR methods in terms of precision and recall. Our proposed method improves the precision by 3.7 %, 9.8 %, 10.3 %, 9.9 %, 15.2 %, 19.4 %, 22.8 %, 20.1 % and 23.9 %, respectively to the work of (Khan et al., 2021; Irtaza et al., 2014; Latha and Raj 2019; Irtaza et al., 2018; Mehmood et al., 2018; Zeng et al., 2016; Elalami, 2014; Tian et al., 2014; Ali et al., 2016). Similarly, our proposed method improves the recall by 3.7 %, 55 %, 10.5 %, 10 %, 16.4 %, 14.8 %, 22.7 %, 20.1 % and 16.9 %, respectively to the work of (Khan et al., 2021; Irtaza et al., 2014; Latha and Raj 2019; Irtaza et al., 2018; Mehmood et al., 2018; Zeng et al., 2016; Elalami, 2014; Tian et al., 2014; Ali et al., 2016).

Figs. 12 and 13 show the retrieved images from Corel 1 K images dataset that are obtained through the proposed system with low precision values and as well as high precision values. While the first image in both figures is a query image and the remaining images are retrieved images that are similar to the query image. Fig. 12 shows the image retrieval results of Africa semantic class with low precision values. We can see that out of the 20 retrieved images, 18 images are similar to the query image while image 4 in the first row and image 20 in the last row are irrelevant to the query image. Similarly, we can see in Fig. 13 with high precision values that all of the 20 images from the bus semantic class are similar to the user query image.

4.5. Assessment of proposed method using Oxford flower data set

This experimentation is designed to compare the performance of the presented method with other contemporary CBIR approaches on the oxford flower dataset. In this relative analysis, the performance of the suggested and comparative methods is assessed on a different number of retrieved images. The outcomes in terms of precision and recall are reported in Table 6. From the results presented in Table 6, we can see that our method obtained the maximum precision and recall values over the contemporary methods irrespective of the number of retrieved images. Moreover, we also observed that our system performed the best when the number of retrieved images was set to 20. More specifically, the performance of the suggested system declines as the number of images attained from the system increases. Because intra-class feature similarity grows as the number of attained images rises that results in an increase of false positive retrieved images and degrade the performance of the image retrieval system.

Fig. 12. Image retrieval results of Africa semantic class with low precision and recall values.

Fig. 13. Image retrieval results of bus semantic class with high precision and recall values.

Table 6
Comparative analysis of suggested method mean precision and mean recall value on retrieval of different no images.

#Returned Images	Proposed Method		(Khan et al., 2021)		(Elalami, 2014)		(ElAlami, 2011)		(Lin et al., 2009)		(Jhanwar et al., 2004)		(Huang and Dai, 2003)	
	P	R	P	R	P	R	P	R	P	R	P	R	P	R
20	0.906	0.225	0.874	0.203	0.821	0.173	0.793	0.161	0.764	0.153	0.603	0.124	0.621	0.129
30	0.856	0.321	0.826	0.306	0.804	0.261	0.765	0.241	0.701	0.228	0.491	0.159	0.502	0.163
40	0.824	0.437	0.776	0.424	0.767	0.324	0.742	0.304	0.682	0.298	0.484	0.193	0.491	0.204
50	0.771	0.471	0.726	0.450	0.706	0.413	0.687	0.386	0.668	0.367	0.479	0.235	0.488	0.246
60	0.726	0.541	0.694	0.503	0.687	0.452	0.656	0.418	0.643	0.399	0.458	0.246	0.462	0.289
70	0.711	0.579	0.674	0.541	0.665	0.489	0.632	0.435	0.616	0.422	0.436	0.263	0.453	0.328
80	0.691	0.641	0.653	0.585	0.642	0.542	0.605	0.476	0.597	0.456	0.427	0.303	0.442	0.345

4.6. Assessment of proposed method using CIFAR-10 data set

We also performed an experiment to assess the efficacy of the presented method over the comparative methods on the CIFAR-10 images dataset. This dataset contains 10 semantic classes with 6000 images per class. The class-wise and mean outcomes in terms of precision and recall are computed on the CIFAR-10 dataset for the proposed and relative methods and outcomes are reported in Table 7. It can be seen that our method obtained the best results, while our prior method (Khan et al., 2021) achieved the second-best performance. Whereas, (Jhanwar et al., 2004) obtained the lowest results on the Cifar-10 images dataset. These results noticeably exhibit the efficacy and usefulness of the proposed method over other state-of-the-art CBIR approaches. Fig. 14 shows the precision-recall curve of the presented and contemporary CBIR methods that also reveal the supremacy of our method over other modern and advanced CBIR approaches.

Table 7
Comparative analysis of suggested method mean precision and mean recall value.

semantic class	Proposed Method		(Khan et al., 2021)		(ElAlami, 2011)		(Lin et al., 2009)		(Jhanwar et al., 2004)		(Huang and Dai, 2003)	
	P	R	P	R	P	R	P	R	P	R	P	R
Airplane	0.937	0.187	0.916	0.183	0.762	0.178	0.732	0.167	0.521	0.129	0.568	0.146
Automobile	0.979	0.196	0.979	0.196	0.705	0.235	0.681	0.224	0.432	0.138	0.395	0.127
Bird	0.979	0.196	0.979	0.196	0.816	0.210	0.789	0.196	0.365	0.153	0.464	0.158
Cat	0.979	0.196	0.958	0.192	0.921	0.137	0.910	0.136	0.861	0.115	0.826	0.118
Deer	0.937	0.187	0.937	0.187	0.918	0.115	0.872	0.110	0.821	0.084	0.782	0.112
Dog	0.875	0.175	0.875	0.175	0.854	0.169	0.756	0.163	0.435	0.151	0.482	0.136
Frog	0.833	0.166	0.791	0.158	0.768	0.134	0.743	0.128	0.692	0.107	0.719	0.121
Horse	0.937	0.187	0.916	0.183	0.893	0.153	0.842	0.148	0.692	0.128	0.736	0.125
Ship	0.979	0.196	0.979	0.196	0.754	0.258	0.682	0.231	0.425	0.162	0.568	0.182
Truck	0.875	0.175	0.833	0.166	0.826	0.162	0.764	0.153	0.642	0.141	0.689	0.147
Mean	**0.931**	**0.186**	**0.916**	**0.183**	**0.821**	**0.175**	**0.777**	**0.165**	**0.589**	**0.130**	**0.622**	**0.137**

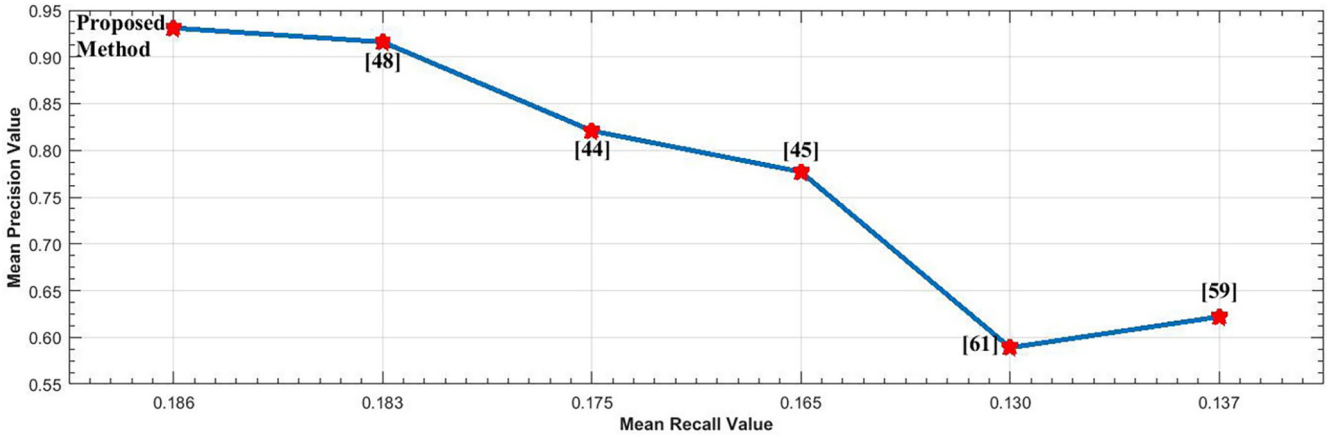

Fig. 14. Comparative analysis of the suggested method with other CBIR approaches.

4.7. Discussion

The purpose of this section is to provide a more in-depth analysis of experimental findings. To evaluate the performance of our proposed methodology we performed our experimentation on three different image datasets. In our first experiment, we used Corel 1 K images dataset. Corel 1 K dataset consists of a heterogeneous collection of images ranging from natural scenes to outdoor sports to different animals and thus this dataset is suitable for the evaluation of an image retrieval system. In this experiment, we evaluated the performance of our proposed system in retrieving the top 20 images against a user query image. Our findings show that we achieved the best results in terms of mean precision and mean recall in Dinosaurs and Elephants semantic classes, while achieved poor results in Beach and African People semantic classes. This happens because the classes where we achieve lower results have similar color and texture features. While the classes where we achieved better results have a distinct color and texture features.

In our 2nd experiment, we employed the Oxford Flower images dataset. In this experiment, we evaluated the performance of our system in retrieving a different number of images against a user query image. In this experiment, our findings show that as the number of retrieved images increases against the user query image, the performance of the proposed system declines. We achieved the best results when we retrieve 20 images against the user query image and the worst results when we retrieve 80 images. This happens due to the inter- and intra-class similarity of image features. As the feature's similarity increases, the performance of the proposed system decreases in terms of mean precision and mean recall. Our findings show that the inter- and intra-class feature similarity increases as the number of retrieved images increases. Therefore, our proposed system performed best when we retrieve 20 images and performed poorly when we retrieve 80 images against the user query image.

In our 3rd experiment, we used the CIFAR-10 images dataset. This dataset consists of low-resolution tiny images of size 32×32 with high similarity and different variations. In this experiment, we also retrieve the top 20 images against a user query image. Our findings show that the classes which contain high inter- and intra-class features similarity show poor results in terms of mean precision and mean recall and the classes which contain highly distinct features provide the best results. This happens because as the similarity between image features increases, the false positive retrieval rate also increases, and the image retrieval performance of the proposed system declines. In our experiments, we compared our system with modern state-of-the-art CBIR systems to evaluate the supremacy of the proposed system. The detailed results show the reliability of our proposed system for image retrieval applications.

Apart from the effectiveness of the proposed CBIR method, we also investigated the computational efficiency aspect of our method. The average time complexity to extract the texture and color features from a single image in our proposed method is 29.63 s. The specifications of the computing machine used to conduct the experiments are as: Windows 10 OS, 8 GB RAM, Intel Core i3 processor @ 2.66 GHz. Considering the fact that we have used a

CPU-based computing machine, this time complexity is fairly reasonable and computational efficiency can be significantly improved by using a GPU-based computing machine.

5. Conclusion

This research article has proposed an effective CBIR method able to recover semantically similar images against a user query image. To accomplish this, we suggested a fused feature vector encompassing texture and color features. We proposed a novel features extraction scheme LTAP for texture features representation while the first three-color moments i.e., mean, standard deviation, and skewness are used to extract the color features. We also employed a genetic algorithm through a support vector machine for the grouping of semantically related images. We applied the Chi-square quadratic distance measure to compute the similarity of images against the user query image. Performance evaluation on three standard and diverse datasets i.e., Corel 1 k dataset, Oxford flower dataset, and CFIAR-10 signify the effectiveness of our system over the contemporary CBIR systems. Experimental results illustrate that the presented system can reliably be used to attain semantically similar images in the presence of diverse and different image classes. From our perspective, the proposed research work is devised using a supervised learning method. An image retrieval system cannot perform well if it is not explicitly trained in an appropriate manner using an effective feature descriptor. But there still exists scope for future work to retrieve the images using unsupervised learning methods so that a system can perform using implicit training and implicit features extraction in an autonomous way without any human intervention. In the future, we plan to include user feedback regarding the guidance and acceptance of retrieved images from the CBIR system to further improve retrieval performance.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Multimedia Signal Processing Research Lab at the University of Engineering and Technology Taxila, Pakistan.

References

Ali, N., Bajwa, K.B., Sablatnig, R., Chatzichristofis, S.A., Iqbal, Z., Rashid, M., Habib, H. A., Rubin, D.L., 2016. A novel image retrieval based on visual words integration of sift and surf. PLoS ONE 11 (6).

Al-Jubouri, A., H., 2020. Integration Colour and texture features for content-based image retrieval. Int. J. Modern Educ. Comput. Sci. 12 (2), 10–18. https://doi.org/10.5815/ijmecs.2020.02.02.

Alzubi, A., Amira, A., Ramzan, N., 2017. Content-based image retrieval with compact deep convolutional features. Neuro Comput. 249, 95–105.

Ashraf, R., Mahmood, T., Irtaza, A., Bajwa.,, K.,, 2014. A novel approach for the gender classification through trained neural networks. J. Basic Appl. Sci. Res. 4, 136–144.

Ashraf, R., Bajwa, K.B., Mahmood, T., 2016. Content-based image retrieval by exploring bandletized regions through support vector machines. J. Inf. Sci. Eng. 32 (2), 245–269.

Ashraf, R., Ahmed, M., Ahmad, U., Habib, M.A., Jabbar, S., Naseer, K., 2020. Mdcbir - mf: multimedia data for content-based image retrieval by using multiple features. Multimedia Tools Appl. 79 (13), 8553–8579.

Ashraf, R., Habib, M.A., Akram, M., Latif, M.A., Malik, M.S.A., Awais, M., Dar, S.H., Mahmood, T., Yasir, M., Abbas, Z., 2020. Deep convolution neural network for big data medical image classification. IEEE Access 8, 105659–105670.

L. Ballerini et al. A query-by-example content-based image retrieval system of non-melanoma skin lesions MCBR-CDS 2009 London UK 31 38.

Baochang, Z., Yongsheng, G., Sanqiang, Z., Jianzhuang, L., 2010. Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. In IEEE Trans.Image Process. 19 (2), 533–544. https://doi.org/10.1109/TIP.2009.2035882.

Bharadi, V.A., Meena, M., 2015. Novel architecture for CBIR SAAS on azure cloud. In 2015 International Conference on Information ProcessIng (ICIP).

Bhardwaj*, S., Pandove, G., Dahiya, P.K., 2020. A journey from basic image features to lofty human intelligence in content-based image retrieval: motivation, applications and future trends. Int. J. Recent Technol. Eng. (IJRTE) 9 (2), 990–998.

Bian, X., Chen, C., Tian, L., Du, Q., 2017. Local and global features for high-resolution scene classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.. https://doi.org/10.1109/JSTARS.2017.2683799.

Celebi, M.E., Aslandogan, Y.A., 2004. Content-based image retrieval incorporating models of human perception. ITCC, vol 2. Las Vegas, Nevada, USA, pp. 241–245.

Chang, S.-F., Sikora, T., Purl, A., 2001. Overview of the MPEG-7 standard. IEEE Trans. Circuits Syst. Video Technol. 11 (6), 688–695. https://doi.org/10.1109/76.927421.

Chang, B.-M., Tsai, H.-H., Chou, W.-L., 2013. Using visual features to design a content-based image retrieval method optimized by particle swarm optimization algorithm. Eng. Appl. Artif. Intell. 26 (10), 2372–2382.

Chaudhuri, B., Demir, B., Bruzzone, L., Chaudhuri, S., 2016. Region-Based Retrieval of Remote Sensing Images Using an Unsupervised Graph Theoretic Approach. IEEE Geosci. Remote Sens. Lett. 13 (7), 987–991.

Chen, X., Zhang, C., 2006. An interactive semantic video mining and retrieval platform application in Transportation surveillance video for incident detection. In Sixth International Conference on Data Mining (ICDM'06). IEEE. 129-138.

Chen, Y., Wang, J.Z., Krovetz, R., 2005. CLUE: cluster based retrieval of images by unsupervised learning. IEEE Trans. Image Process. 14, 1187–1201.

CIFAR-10 dataset, available on https://www.cs.toronto.edu/~kriz/cifar.html, accessed on June 29, 2021.

Cord, M., Gosselin, P.-H., 2008. In: Cognitive TechnologiesMachine Learning Techniques for Multimedia. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 115–138.

Corel 1K dataset, available on http://wang.ist.psu.edu/docs/related/, accessed on June 05, 2021.

ElAlami, M.E., 2011. A novel image retrieval model based on the most relevant features. Knowledge Based Systems 24 (1), 23–32.

Elalami, M.E., 2014. A new matching strategy for content-based image retrieval system. Appl. Soft Comput. 14, 407–418.

Fadaei, S., Rassoul, A., Mohammad, R.A., 2016. New content-based image retrieval system based on optimized integration of DCD wavelet and curvelet features. IET Image Proc..

Fakheri, M., Sedghi, T., Shayesteh, M.G., Amirani, M.C., 2013. Framework for image retrieval using machine learning and statistical similarity matching techniques. IET Image Proc 7 (1), 1–11.

Oxford Flower dataset, available on https://www.robots.ox.ac.uk/~vgg/data/flowers/17/index.html, accessed on June 12, 2021.

Guo, Z., Zhang, L., Zhang, D., 2010. A completed modeling of local binary pattern operator for texture classification. IEEE Trans. Image Process. 19 (6), 1657–1663.

Hafiane, A., Chaudhuri, S., Seetharaman, G., Zavidovique, B., 2006. Region based CBIR in GIS with local space filling curves to spatial representation. Pattern Recogn. Lett. 27 (4), 259–267.

Hameed, I.M., Abdulhussain, S.H., 2021. An efficient multistage CBIR based on squared krawtchouk-tchebichef polynomials. IOP Conf. Ser. Mater. Sci. Eng. 1090 (1), 012100–012112. https://doi.org/10.1088/1757-899x/1090/1/012100.

Hsiao, M.J., Huang, Y.P., Tsai, T., Chiang, T.W., 2010. An efficient and flexible matching strategy for content-based image retrieval. Life Sci. J. 7. https://doi.org/10.1016/j.eswa.2011.03.014.

Huang, P.W., Dai, S., 2003. Image retrieval by texture similarity. Pattern Recog 36 (3), 665–679.

Iakovidou, C., Anagnostopoulos, N., Lux, M., Christodoulou, K., Boutalis, Y., Chatzichristofis, S.A., 2019. Composite description based on salient contours and color information for CBIR tasks. In IEEE Trans. Image Process. 28 (6), 3115–3129. https://doi.org/10.1109/TIP.2019.2894281.

A. Irtaza S. Adnan K. Ahmed A. Jaffar A. Khan A. Javed M. Mahmood An Ensemble Based Evolutionary Approach to the Class Imbalance Problem with Applications in CBIR Applied Sciences 8 4 2018 495 Available: 10.3390/app8040495.

Irtaza, A., Jaffar, M.A., Aleisa, E., Choi, T.S., 2014. Embedding neural networks for semantic association in content based image retrieval. Multimed Tools Appl. 72 (2), 1911–1931.

Irtaza, A., Jaffar, M.A., Muhammad, M.S., 2015. Content based image retrieval in a web 3.0 environment. Multimed Tools Appl. 74 (14), 5055–5072.

Jhanwar, N., Chaudhuri, S., Seetharaman, G., Zavidovique, B., 2004. Content based image retrieval using motif co-occurrence matrix. Image Vis Comput. 22 (14), 1211–1220.

Jiang, D., Kim, J., 2021. Image retrieval method based on image feature fusion and discrete cosine transform. Appl. Sci. 11 (12), 5701–5728. https://doi.org/10.3390/app11125701.

Jiji, G.W., Durairaj, P.J., 2015. Content based image retrieval techniques for the analysis of dermatological lesions using particle swarm optimization technique. Appl. Soft Comput. 30, 650–662.

K., A., K., N., & D., R. D., 2021. A Content-Based Approach to Medical Image Retrieval. AI Innovation in Medical Imaging Diagnostics, 114–136. https://doi.org/10.4018/978-1-7998-3092-4.ch007.

Kareem Jebur, A., 2021. Uses and applications of geographic information systems. Saudi J. Civ. Eng. 5 (2), 18–25. https://doi.org/10.36348/sjce.2021.v05i02.001.
Kaur, R., Devendran, D.V., 2020. Content based image retrieval: a review. Int. J. Innovat. Technol. Explor. Eng. 9 (10), 222–228. https://doi.org/10.35940/ijitee.j7453.0891020.
Khan, U.A., Javed, A., Ashraf, R., 2021. An effective hybrid framework for content based image retrieval (CBIR). Multimed Tools Appl. 80 (17), 26911–26937.
Lai, C.-C., Chen, Y.-C., 2011. A user-oriented image retrieval system based on interactive genetic algorithm. IEEE Trans. Instrum. Meas. 60 (10), 3318–3325.
Latha, D., Raj, Y.J., 2019. Hybrid cbir method using statistical dwt-entropy and popmv-based feature sets. IET Image Process 13 (12), 2031–2044.
Latif, A., Rasheed, A., Sajid, U., Ahmed, J., Ali, N., Ratyal, N. I, Zafar, B., Daar, S. H., Sajid, M., & Khalil, T., 2019. Content Based Image Retrieval and Feature Extraction: A Comprehensive Review, Mathematical Problems in Engineering, 21 pages, Article ID 9658350, https://doi.org/10.1155/2019/9658350.
Lin, C.H., Chen, R.T., Chan, Y.K., 2009. A smart content-based image retrieval system based on color and texture feature. Image Vis Comput 27 (6), 658–665.
Lin, C.H., Huang, D.C., Chan, Y.K., Chen, K.H., Chang, Y.J., 2011. Fast color spatial feature-based image retrieval methods. Expert Syst. Appl. 38 (09), 11412–11420.
Lindsay, B.G., Markatou, M., Ray, S., Yang, K., Chen, S.-C., 2008. Quadratic distances on probabilities: a unified foundation. Ann. Statist 36 (2), 983–1006. https://doi.org/10.1214/009053607000000956.
Liu, G.H., Li, Z.Y., Zhang, L., Xu, Y., 2011. Image retrieval based on micro-structure descriptor. Pattern Recogn. 44 (9), 2123–2133.
Liuand, G.H., Yang, J.Y., 2013. Content based image retrieval using color difference histogram. Pattern Recogn. 46 (1), 188–198.
Lonescu, M., Ralescu, A., 2004. Fuzzy hamming distance in content based image retrieval system. IEEE Int. Conf. Fuzzy Syst. Hungary. https://doi.org/10.1109/FUZZY.2004.1375443.
Machhour, N., Nasri, M., 2020. Efficient Image Retrieval Based on Support Vector Machine and Genetic Algorithm Using Color, Texture and Shape Features. In: 2020 6th IEEE Congress on Information Science and Technology (CiSt), pp. 284–289. https://doi.org/10.1109/cist49399.2021.9357167.
Mansoori, N.S., Nejati, M., Razzaghi, P., Samavi, S., 2013. Bag of visual words approach for image retrieval using color information. In: In ProceedIngs of the 2013 21st Iranian Conference on Electrical EngIneerIng, pp. 1–6. https://doi.org/10.1109/IranianCEE.2013.6599562.
Mehmood, Z., Anwar, S.M., Altaf, M., 2018. A novel image retrieval based on rectangular spatial histograms of visual words. Kuwait J. Sci. 1, 45.
Montazer, G.A., Giveki, D., 2015. Content based image retrieval system using clustered scale invariant feature transforms. Optik 126 (18), 1695–1699.
Naik, J. et al., 2009. A boosted distance metric: application to content based image retrieval and classification of digitized histopathology. Medical Imaging, Computer-Aided Diagnosis, vol.7260, no.1. Orlando, FL, USA.
Naik, J. et al., 2009. A boosted distance metric: application to content-based image retrieval and classification of digitized histopathology, Proc. SPIE 7260, Medical Imaging 2009: Computer-Aided Diagnosis, 72603F, https://doi.org/10.1117/12.813931.
Pourreza, A., Kourosh, K., 2016. A partial-duplicate image retrieval method using color-based SIFT. Electrical Engineering (ICEE), 24th Iranian Conference on IEEE.
Rahman, M. et al., 2006. Image retrieval-based decision support system for dermatoscopic images. In: CBMS. IEEE Computer Society. Salt Lake City, Utah, USA, pp. 285–290.
Rahman, M.M. et al., 2006. Image retrieval-based decision support system for derma to scopic images, in 19th IEEE Symposium on Computer Based Medical Systems (CBMS-06). IEEE, 285-290.
Rao, R.V., Prasad, T.J.C., 2021. Content-based medical image retrieval using a novel hybrid scattering coefficients – Bag of visual words - DWT relevance fusion. Multimedia Tools Appl. 80 (8), 11815–11841. https://doi.org/10.1007/s11042-020-10415-5.
Rao, M.B., Rao, B.P., Govardhan, A., 2011. Content based image retrieval system based on dominant color, texture and shape. Int. J. Eng. Sci. Technol. (IJEST) 4, 2887–2896.
Robles-Serrano, S., Sanchez-Torres, G., Branch-Bedoya, J., 2021. Automatic detection of traffic accidents from video using deep learning techniques. Computers 10 (11), 148. https://doi.org/10.3390/computers10110148.
Saadatmand-Tarzjan, M., Moghaddam, H.A., 2007. An ovelevolutionary approach for optimizing content based image indexing algorithms. IEEE Trans. Syst. Man Cybern. 37 (1), 139–153.
Sankar, S.P., Vishwanath, N., Lang, H.J., Karthick, S., 2017. An effective content based medical image retrieval by using abc based artificial neural network (ANN). Curr. Med. Image Rev 13 (3), 223–230.
Saoudi, E.M., Jai-Andaloussi, S., 2021. A distributed content-based video retrieval system for large datasets. J. Big Data 8 (1). https://doi.org/10.1186/s40537-021-00479-x.
Scott, G.J., England, M.R., Starms, W.A., Marcum, R.A., Davis, C.H., 2017. Training deep convolutional neural net- works for land±cover classification of high-resolution imagery. IEEE Geosci. Remote Sens. Lett. 14 (4), 549–553. https://doi.org/10.1109/LGRS.2017.2657778.
Shen, G.L., Wu, X.J., 2013. Content based image retrieval by combining color texture and CENTRIST. In Proceedings of the Constantinides International Workshop on Signal Processing (CIWSP'13).1-4.
Shi, X., Sapkota, M., Xing, F., Liu, F., Cui, L., Yang, L., 2018. Pairwise based deep ranking hashing for histopathology image classification and retrieval. Pattern Recogn. 81, 14–22.
Shrivastava, N., Tyagi, V., 2014. Content based image retrieval based on relative locations of multiple regions of interest using selective regions matching. Inf. Sci. 259, 212–224. https://doi.org/10.1016/j.ins.2013.08.043.
Shukran, M.A.M., Abdullah, M.N., Yunus, M.S.F.M., 2021. New approach on the techniques of content-based image retrieval (cbir) using color, texture and shape features. J. Mater. Sci. Chem. Eng. 09 (01), 51–57. https://doi.org/10.4236/msce.2021.91005.
Y. Song I.V. McLoughlin L.-R. Dai W.-B. Du Local coding based matching kernel method for image classification Plos one 9 8 2014 10.1371/journal.pone.0103575 e103575 e103575.
Stejić, Z., Takama, Y., Hirota, K., 2003. Genetic algorithm-based relevance feedback for image retrieval using local similarity patterns. Inf. Process. Manage. 39 (1), 1–23.
Talib, A., Mahmuddin, M., Husni, H., George, L.E., 2013. A weighted dominant color descriptor for content-based image retrieval. J. Vis. Commun. Image Represent. 24 (3), 345–360.
Tao, D., Tang, X., Li, X., 2006. A symmetric bagging and random sub space for support vector machines-based relevance feedback in image retrieval. IEEE Trans. Pattern Anal. Mach. Intell. 28, 1088–1099.
Tian, X., Jiao, L., Liu, X., Zhang, X., 2014. Feature integration of eodh and color-sift: application to image retrieval based on code book. Sig Process Image Commun 29 (4), 530–545.
Wang, J.Z., Li, J., Wiederhold, G., 2001. SIMPLIcity: semantics-sensitive integrated matching for picture libraries. IEEE Trans. Pattern Anal. Mach. Intell. 09, 947–963.
Wang, X.Y., Li, Y.W., Yang, H.Y., Chen, J.W., 2014. An image retrieval scheme with relevance feedback using feature construction and SVM re-classification. Neuro Comput. 127, 214–230.
Wang, X.-Y., Yu, Y.-J., Yang, H.-Y., 2011. An effective image retrieval scheme using color, texture and shape features. Comput. Standards & Interfaces 33 (1), 59–68.
Wei, W., Wang, Y., 2019. Color image retrieval based on quaternion and deep features. In IEEE Access 7, 126430–126438. https://doi.org/10.1109/ACCESS.2019.2938000.
Xia, Z., Wang, X., Zhang, L., Qin, Z., Sun, X., Ren, K., 2016. A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing. IEEE Trans. Inf. Forensics Secur. 11 (11), 2594–2608.
Ye, F., Dong, M., Luo, W., Chen, X., Min, W., 2019. A new re-ranking method based on convolutional neural network and two images to class distances for remote sensing image retrieval. IEEE Access 7, 141498–141507. https://doi.org/10.1109/ACCESS. 2019.2944253.
Yildizer, E., Balci, A.M., Hassan, M., Alhajj, R., 2012. Efficient content-based image retrieval using multiple support vector machines ensemble. Exp. Syst. Appl. 39 (3), 2385–2396.
Youssef, S.M., 2012. ICTEDCT- CBIR: integrating curvelet transform with enhanced dominant colors extraction and texture analysis for efficient content based image retrieval. Comput. Electr. Eng.. https://doi.org/10.1016/j. compeleceng. 2012. 05. 010.
Yue, J., Li, Z., Liu, L., Fu, Z., 2011. Content-based image retrieval using color and texture fused features. Math. Comput. Modell. 54 (3–4), 1121–1127.
Zagoris, K., Amanatiadis, A., Pratikakis, I., 2021. Word spotting as a service: an unsupervised and segmentation-free framework for handwritten documents. J. Imag. 7 (12), 278. https://doi.org/10.3390/jimaging7120278.
Zang, M., Wen, D., Liu, T., Zou, H., Liu, C., 2018. A pooled Object Bank descriptor for image scene classification. Expert Syst. Appl. 94, 250–264. https://doi.org/10.1016/j.eswa.2017.10.057.
Zeng, S., Huang, R., Wang, H., Kang, Z., 2016. Image retrieval using spatiograms of colors quantized by Gaussian mixture models. Neuro computing 171, 673–684.
Zhang, C., Wen, G., Lin, Z., Yao, N., Shang, Z., Zhong, C., 2016. An effective bag of visual word scheme for object recognition. In 9th International Congress on Image and Signal ProcessIng, Bio Medical EngIneerIng and Informatics (CISP.
Zhu, L., Shen, J., Xie, L., Cheng, Z., 2017. Unsupervised visual hashing with semantic assistant for content-based image retrieval. IEEE Trans. Knowl. Data Eng. 29 (2), 472–486.
Zou, J., Li, W., Chen, C., Du, Q., 2016. Scene classification using local and global features with collaborative representation fusion. Inf. Sci. 209–226. https://doi.org/10.1016/j.ins.2016.02.021X.