

Research Article

Noise Resilient Local Gradient Orientation for Content-Based Image Retrieval

Samina Bilquees,¹ Hassan Dawood¹,¹ Hussain Dawood²,² Nadeem Majeed³,³ Ali Javed⁴,⁴ and Muhammad Tariq Mahmood⁵

¹Department of Software Engineering, University of Engineering and Technology, Taxila 47080, Pakistan

²Department of Computer and Network Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah 21589, Saudi Arabia

³Punjab University College of Information Technology (PUCIT), University of the Punjab, Lahore, Pakistan

⁴Department of Computer Science, University of Engineering and Technology, Taxila 47050, Pakistan

⁵Future Convergence Engineering, School of Computer Science and Engineering, Korea University of Technology and Education, Cheonan, Republic of Korea

Correspondence should be addressed to Muhammad Tariq Mahmood; tariq@koreatech.ac.kr

Received 29 April 2021; Revised 25 June 2021; Accepted 5 July 2021; Published 14 July 2021

Academic Editor: E. Bernabeu

Copyright © 2021 Samina Bilquees et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In a world of multimedia information, where users seek accurate results against search query and demand relevant multimedia content retrieval, developing an accurate content-based image retrieval (CBIR) system is difficult due to the presence of noise in the image. The performance of the CBIR system is impaired by this noise. To estimate the distance between the query and database images, CBIR systems use image feature representation. The noise or artifacts present within the visual data might confuse the CBIR when retrieving relevant results. Therefore, we propose Noise Resilient Local Gradient Orientation (NRLGO) feature representation that overcomes the noise factor within the visual information and strengthens the CBIR to retrieve accurate and relevant results. The proposed NRLGO consists of three steps: estimation and removal of noise to protect the local visual structure; extraction of color, texture, and local contrast features; and, at the end, generation of microstructure for visual representation. The Manhattan distance between the query image and the database image is used to measure their similarity. The proposed technique was tested using the Corel dataset, which contains 10000 images from 100 different categories. The outcomes of the experiment signify that the proposed NRLGO has higher retrieval performance in comparison with state-of-the-art techniques.

1. Introduction

The advent of multimedia tools has made it easier to access a wider range of images with a variety of information. CBIR (content-based image retrieval) uses low-level image properties to search a huge database for images that suit the user's demands. Color, edges, and orientation are considered low-level image attributes [1]. CBIR is widely used in different applications like medical imaging [2], E-commerce [3], and digital libraries [4].

Text-based image retrieval (TBIR) and sketch-based image retrieval are two image retrieval methodologies. TBIR search images on the basis of labels and keywords. It is

mostly used in Google Images. Therefore, it becomes difficult to describe whole image content in words and it may show irrelevant content. Manual annotation becomes difficult when there is large database [5]. A sketch-based image retrieval system is used to find images on the basis of sketch content drawn by user [6]. Soft Histogram of Edge Local Orientation (S-HELO) uses local orientation computation for improving the performance of sketch-based image retrieval [7]. The proposed histogram of line relationship (HLR) solves the appearance problem between sketches and images. It removes the noisy edges by choosing suitable edge shape that best corresponds to object boundaries [8]. Manual annotations of text-based and user-based sketch

techniques cause incompatible image retrieval. Xie et al. proposed analogy relevance feedback method based on the user intentions that is helpful in maximizing retrieval results [9]. Case Based Long Term Learning (CB-LTL) method of relevance feedback [10] incorporates user intentions and increases retrieval performance. Mandal et al. proposed signature-based bag of visual words (S-BoVW) that employs jumble of words with image's texture and color to retrieve similar images [11]. BoVW based methods improve the retrieval performance but their computation is expensive.

In this paper, we propose Noise Resilient Local Gradient Orientation (NRLGO) that overcomes the noise factor within the visual information and strengthens the CBIR to retrieve accurate and relevant results. In our proposed NRLGO, the color, texture, and local contrast information were used for feature representation of visual data. While semantic attributes are estimated through calculating the correlation between visual features and regular structure data. NRLGO has noise resilience characteristics and saves the local structures of visual information from noise by finding an uncertain state. The major contribution of NRLGO is summarized as follows:

- (1) The NRLGO holds a noise resilient attribute by finding the uncertain bit within the visual information which is significant to increase the CBIR system's performance.
- (2) NRLGO computes pixel value and difference and then multiplies them by the same constant value. In this way, the bright changes will be canceled and the discriminative feature representation is improved.
- (3) Multiresolution gradient orientations of NRLGO improve the gradient magnitude by removing noise factor and detecting the edges.
- (4) NRLGO estimates the gradient orientation through dividing the θ value into T dominant orientations. Thus, gradient orientation strengthens the visual feature representation of image.

$$\arctan 2(u_s^{11}, u_s^{10}) = \begin{cases} \theta, & u_s^{11} > 0 \text{ and } u_s^{10} > 0, \\ \pi - \theta, & u_s^{11} > 0 \text{ and } u_s^{10} < 0, \\ \theta - \pi, & u_s^{11} < 0 \text{ and } u_s^{10} < 0, \\ -\theta, & u_s^{11} < 0 \text{ and } u_s^{10} > 0. \end{cases} \quad (1)$$

The remainder of the paper is structured as follows: The related work is given in Section 2. The suggested NRLGO is presented in Section 3 which discusses the attributes of NRLGO against noise, illumination variation, and rotation variation. Section 4 summarizes the experimental findings and our contribution against state-of-the-art techniques. In Section 5, discussion is given, and Section 6 concludes our contribution.

2. Related Work

The CBIR system plays a vital role in retrieving images that are related from large dataset that are semantically and contextually similar to search image. The CBIR systems

extract visual feature representation and estimate the similarity distance for scoring the images that are comparable from the visual database. In the CBIR systems, visual feature descriptors were used in two ways: local and global descriptors. The local descriptors extract image's local interest points while global descriptors take the complete image for feature extraction [12]. Different CBIR methods are used to achieve efficient retrieval performance. The scale invariant feature transform (SIFT) [13] was introduced as a local descriptor, and it is invariant to scale and rotation. To deal with the difficulty of image retrieval using a high-dimensional feature vector, the PCA-SIFT [14] was used. The Hessian matrix and distance ratio were used to solve the computational issue of speeded up robust features (SURF) [15]. The local binary pattern (LBP) based descriptor was proposed by [16]; however, LBP is applicable for gray-scale images only. Moreover, LBP does not perform accurate prediction in presence of noise. This limitation of LBP was overcome by multichannel decoded LBP [17]. The multichannel decoded LBP [17] performs computation on the color and gray-scale images. To further improve the performance of LBP, the fusion of color histogram with LBP was applied to achieve better retrieval performance for color images [18]. To minimize noise in regular patterns, the local ternary pattern (LTP) [19] was developed. The LTP extracts the spatial features in three directions while LBP extract textual features in two directions. The extended local ternary pattern (ELTP) [20] is robust and overcomes the noise within the images. The scale invariant local binary pattern (SILTP) [21] was used to extract local features for gray-scale images and estimate the pixel difference for complex background scenes. Edge detection Sobel filter calculates orientation of edges but has low signal to noise ratio [22]. Edges obtained from Sobel filter have large thickness which gives mismatched images. Canny filter [23] solved the issue of thickness of edges. It has good signal to noise ratio.

Color and image edges are sensitive to the human visual system. In HSV color space, CDH is utilized to improve human visual perception. Entropy is used for feature selection, and correlation is developed between the features after feature selection [24]. The use of visual feature discrimination to measure the similarity between the image passed by the user and images in the database is proposed in an approach based on weight learner [25]. Spatial pyramid matching (SPM) is proposed for spatial distribution of images that increased the retrieval accuracy, but when image alignment is not done properly, it results in rotation and translation variance. Geometric relation is established on the basis of center of image among a group of similar words [26]. Color feature is fused with texture feature to achieve highest retrieval performance as a single attribute is not more resilient for obtaining efficient retrieval performance. As a result, for color extraction, color moment is used, and Gabor descriptor is used to extract the texture feature. After that, features are represented using the Color and Edge Directivity Descriptor (CEDD) [27]. The images are first transformed to RGB color space. Texture information is obtained with the help of evaluated local binary pattern and by predefined prepattern unit. This texture information is then

fused with color channels to increase the retrieval performance [28]. Feature extraction [29] is described by fusing top-hat transforms and local binary patterns for other image processing cases to analyze the color images. Top-hat transform is used to extract the shape, and color local binary pattern is used for texture classification.

In [30], extraction of color feature is done by color difference histogram, and Gabor descriptor is used for the texture classification. Joint property of color with the texture increases the retrieval accuracy. Rotation invariant and gray-scale invariant property is achieved by using the texture feature with color feature. Another technique that uses the color and texture features is developed for image retrieval. HSV and Lab color space are used to extract the color attribute, and texture classification is done by using Haralick features in RGB and gray-scale images [31]. Plant disease [32] is identified by using color, and texture classification is done by using gray-level cooccurrence matrix (GLCM).

CBIR systems are used to represent the visual information in global context based on different deep learning techniques. Bilinear convolutional neural network was proposed for efficient retrieval performance that consists of CNN architecture with bilinear root pooling [33]. Similarly, another deep learning method for CBIR which extracts global features through two parallel CNN models was presented. In [34] SPoC global descriptor was developed through aggregation of deep learning local features for CBIR. The cross-dimensional CNN weights were aggregated to represent the global information for CBIR [35]. The triplet network [36] was used for optimizing the ReLU max aggregation convolutions (R-MAC). The R-MAC involves the pooling of regions of image to cover the point of interest. Moreover, the deep CNN model [37] was used to compress the image descriptors and activate the layers of CNN. The multilayered CNN was used for features extraction, and features were encoded with VLAD encoding scheme [38]. The CNN based CBIR systems require large amount of data for model learning and robust machines for computation of loss and hyperparameters, while local features and distance based CBIR approaches require no clustering and are computationally less expensive than deep learning methods.

The CBIR systems have been studied for decades. Still, occlusion, cluttered background, viewpoint variation, and noise make the retrieval process a challenging task. The literature reveals that all the existing methods for CBIR systems were established using clear visual information, and little effort was made to remove the noise within the visual information. However, presence of noise in visual information is responsible for degraded performance of local or global descriptors (deep learning methods). This is due to the fact that the visual information might be occluded with noise, or maliciously designed data was introduced within the visual information to deceive the CBIR models. Therefore, noise resilient local descriptor is required to overcome the noise or artifacts within the images for CBIR.

3. Noise Resilient Local Gradient Orientation

The CBIR systems compute the discriminative low-level attribute of image including color, edges, and patterns. The

distribution of pixels within the visual information is uniform and represents the discriminative attributes of visual information like color, texture, and patterns. The uniform distribution of pixels is sensitive to small uncertain changes and these small uncertain changes disturb the uniform pixel distribution and cause nonuniform patterns within the image. This uncertain noise confuses the CBIR models to retrieve relevant information. Therefore, we proposed Noise Resilient Local Gradient Orientation (NRLGO) to overcome the noise patterns and sustain the uniform pattern of the pixels. The proposed NRLGO consists of three steps: estimation of noise and protection of the local visual structure; then low-level feature extraction (color, texture, and contrast); and generation of microstructures, as illustrated in Figure 1. The nonuniform noise signal was estimated, and resilient model was established through error correction method. Then, discriminative color and texture features were estimated through quantizing HSV color space and local gradient orientation (LGO). Moreover, NRLBP extracts contrast information through quantizing V values of HSV. The Manhattan distance between the search image and the database image is used to determine their similarity. The proposed NRLGO method is completely described in Figure 2.

3.1. Noise Resilient Local Binary Pattern. The proposed NRLBP establishes a noise resilient attribute through computing the small pixel difference and estimating the uncertain bit Y and uncertain state. The uncertain bit Y and uncertain state were estimated on the basis of corrected noise-free bits within the image. The intensity difference x_p is calculated between center pixel and its corresponding neighboring pixel through the three-state fuzzy logic's B_{p-1}^N representation, and s is threshold which is chosen as 2 for NRLBP [39].

$$B_{p-1}^N = \begin{cases} 1, & \text{if } x_p \geq s, \\ U, & \text{if } |x_p| < s, \\ 0, & \text{if } x_p \leq -s. \end{cases} \quad (2)$$

The noise factor is considered as uncertain bit and occurs either from 0 to 1 or from 1 to 0. The uncertain bit $Y = (y_1, y_2, \dots, y_n)$ represents the feature vector where $y_i, y_i \in \{0, 1\}$ is formed by variable n . The uncertain state $U(Y)$ can be mathematically represented by the following equation:

$$B_{p-1}^N B_{p-2}^N, \dots, B_1^N B_0^N = U(Y). \quad (3)$$

The regular visual patterns involve edges, edge-ends, corners, etc. The regular patterns exist more frequently than irregular patterns within the images. As regular expression for normal visual patterns exists, it is possible to estimate the regular expression of uncertain state U from the regular visual patterns present within the image.

NRLBP corrects the irregular patterns to regular patterns by removing noise on the basis of error correction mechanism. For example, the image shown in Figure 1 holds an

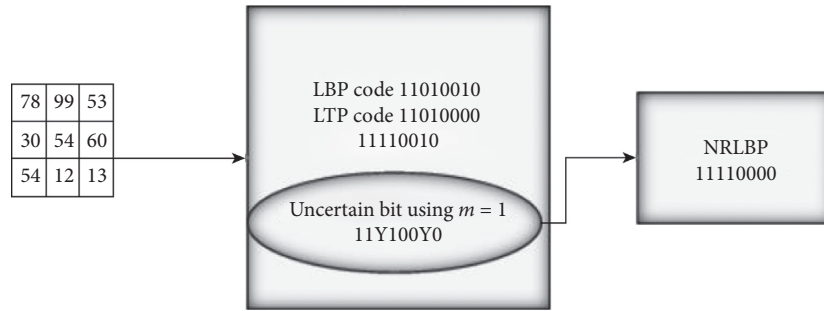


FIGURE 1: NRLBP encoding scheme in addition to comparison of LBP and LTP. NRLBP uniform code is formed by finding the uncertain bit Y at $m = 1$, and s is chosen as 2 for LTB and NRLBP.

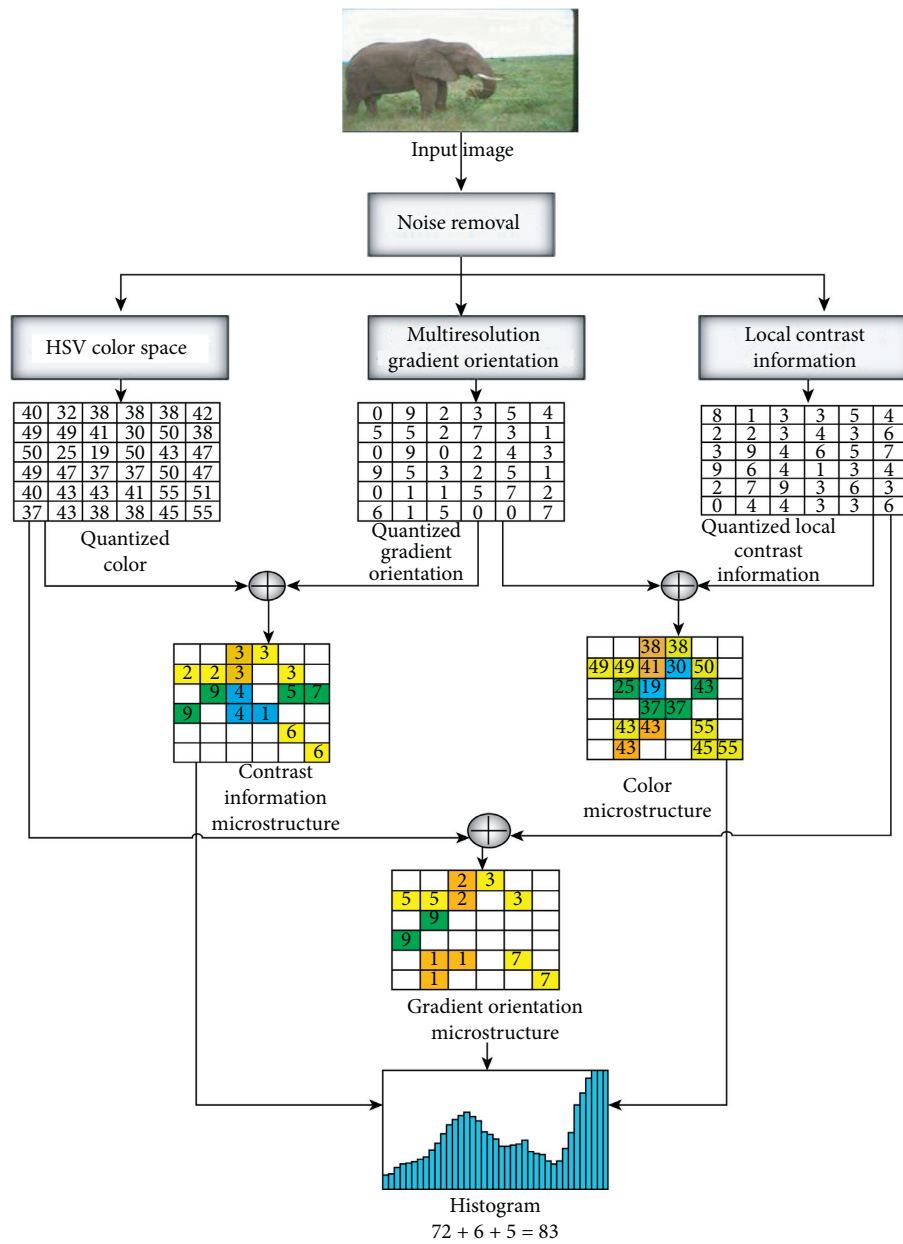


FIGURE 2: Visual representation of proposed NRLGO descriptor for CBIR system. The proposed NRLGO consists of three steps: estimation and removal of noise to protect the local visual structure; then the color and texture feature extraction; and, at the end, microstructure generation.

uncertain code 11Y100Y0. In order to remove the noise, there is a need to predict the uncertain bit and then modify the uncertain bit to form uniform patterns, for example, 11110000 and 11111000. Suppose that φ_r represents the set of all regular LBP patterns. For instance, the image consists of 58 regular patterns. On the basis of uncertain state $U(Y)$, a list of NRLBP codes is provided as follows:

$$R_{\text{NRLBP}} = \{U(Y)|y \in \{0, 1\}^t\}, \quad U(Y) \in \varphi_r. \quad (4)$$

The uniform code of 11110000 is obtained after error correction, and noise was removed. After that, the histogram of NRLBP for a local region of image was formed, and the number of elements in R_{NRLBP} is represented by n . The bin of histogram was representing each element of R_{NRLBP} which was added by $1/n$ when $t > 0$. When $t = 0$, the irregular bin was added by 1. The same process was repeated for each pixel in the region. Furthermore, NRLBP has reduced histogram bins from 118 to 59 bins. The feature vector of the local image region can be generated by summing the feature vectors collected from each pixel in the image region.

3.2. Feature Extraction. Feature extraction includes the extraction of color, texture, and local contrast information. Converting RGB to HSV color space extracts the color characteristic. HSV (high-saturation value) color space is quantized into 72 colors with, 8, 3, and 3 bins. Texture classification is done by using multiresolution gradient orientation. It improves the gradient magnitude by removing noise factor and detecting the edges. The V value of the HSV color space is used to acquire local contrast information.

3.2.1. Color Feature. The chrominance signals of visual information hold significant discriminative information to represent the images in CBIR models. To reflect the chrominance attribute of visual information, different color spaces are used like YCbCr, HSV, HIS, Lab. The RGB color model is made up of three components: red, green, and blue. In proposed method, HSV color space was selected to segregate the hue, saturation, and intensity of image chrominance.

In addition, the HSV color model represents the chrominance attribute of image, and it has cylindrical representation with hue (H), saturation (S), and value (V). The hue component of the HSV color space describes the wavelength of colors ranging from 0 to 360, beginning with red at 0, yellow at 60, green at 120, cyan at 180, blue at 240, and magenta at 300. The saturation component describes the chrominance saturation level between 0 and 1, where 0 denotes gray and 1 denotes primary color. The value V component describes the intensity values between 0 and 1, where 1 reflects black and 1 reflects white. The difference between HSV and RGB is that the HSV separates the intensity values from the color element. The RGB system is based on three color elements: red, green, and blue.

The HSV color space is color invariant because the V component of HSV represents the image's brightness and is

independent of color [24]. With 8, 3, and 3 bins, the HSV color space is quantized into 72 colors. The quantization is essential in persevering images from light and intensity and also reduces the time complexity.

$$\begin{aligned} q_h &= \left\{ h \times \left(\frac{b_h}{\max_h} \right) \right\}, \\ q_s &= \left\{ s \times \left(\frac{b_s}{\max_s} \right) \right\}, \\ q_v &= \left\{ v \times \left(\frac{b_v}{\max_v} \right) \right\}, \end{aligned} \quad (5)$$

where b_h , b_s , and b_v represent the quantized bins of hue, saturation, and value component of HSV [12].

$$cm = q_h \times (b_s \times b_v) + q_s \times b_v + q_v. \quad (6)$$

The quantized values for hue (h), saturation (s), and value (v) of HSV in color map cm are q_h , q_s , and q_v . n_c is the number of quantized colors. s and t are spatial coordinates in quantized color map q_{cm} .

$$q_{cm}(j) = \{(s, t) | (s, t) \in_{cm} = j, \quad 0 \leq j \leq n_c - 1\}. \quad (7)$$

3.2.2. Local Gradient Orientation (LGO). For gradient computation, the orientation component of Weber local descriptor (WLD) was applied to HSV color model. This is due to the fact that by using the gray-scale images majority of chrominance information is lost. Therefore, the local gradient orientations were extracted, which holds rotation invariant attribute from chrominance part and is beneficial to obtain discriminative features. The gradient orientation [40] was calculated from the angle between the reference axis and the vector in horizontal and vertical location x .

$$\theta(y_c) = \alpha_s^1 = \arctan\left(\frac{u_t^{11}}{u_t^{10}}\right), \quad (8)$$

$f_{10} = \begin{bmatrix} -1 \\ +1 \end{bmatrix}$ and $f_{11} = \begin{bmatrix} -1 \\ +1 \end{bmatrix}$. When filters f_{10} and f_{11} are applied to an input image, then outputs u_t^{11} and u_t^{10} are obtained. $u_t^{10} = y_5 - y_1$ and $u_t^{11} = y_7 - y_3$.

θ is quantized into T gradient orientations. Firstly, mapping is done before quantization $f: \theta \rightarrow \theta'$: $\theta' = \arctan 2(u_s^{11}, u_s^{10}) + \pi$ and

$$\arctan 2(u_s^{11}, u_s^{10}) = \begin{cases} \theta, & u_s^{11} > 0 \text{ and } u_s^{10} > 0, \\ \pi - \theta, & u_s^{11} > 0 \text{ and } u_s^{10} < 0, \\ \theta - \pi, & u_s^{11} < 0 \text{ and } u_s^{10} < 0, \\ -\theta, & u_s^{11} < 0 \text{ and } u_s^{10} > 0, \end{cases} \quad (9)$$

$\theta \in [-\pi/2, \pi/2]$, and $\theta' \in [0, 2\pi]$. Mapping takes value of θ under consideration.

Quantization function is given as follows.

$\varphi_s = f_r(\theta') = (2q/Q)\pi$, and $q = \text{mod}(\theta'/(2\pi/T) + t, (1/2))$. If $T=8$ then $\varphi_s = (s\pi)/4$, ($s = 0, 1, T-1$).

The multiscale gradient orientation is useful for extracting gradient orientation to describe different granular aspects of visual representation. Multiscale gradient orientation is able to improve the discriminative ability of a single resolution. It is obtained by using square neighbors of P pixels, considering the length of $(2R+1)$. As illustrated in Figure 3, P denotes the set of neighbors, and R denotes the spatial resolution.

The multiscale gradient orientation was obtained through combining the histograms of different operators at varying P and R . By varying the P and R , whole image pixel values are considered. A multiscale analysis of WLD orientation can be done using the data generated by several operators of varying scales (P, R). Despite the fact that the operator is based on a squared symmetric neighbor set of P members on a square with side length $(2R+1)$, it can also be used for a circular situation. In general, it can improve the discrimination of a single resolution (P, R).

The WLD orientation is used for texture estimation, and it also eliminates the effect of noise. The WLD orientation reduces the noise and illumination variation by computing the differences between the pixel and its neighbors. WLD orientation is tolerant of the presence of noise in an image. The influence of noise is reduced by using a WLD orientation, which is analogous to smoothing in image processing. Furthermore, the sum of its p -neighbor differences is divided by the current pixel's intensity, reducing the influence of noise in an image.

WLD orientation has also been established to mitigate the impact of changing brightness. It calculates the differences between the center pixel and its neighbors. As a result, a brightness modification that adds a constant to each image pixel has no effect on the disparities in values. WLD orientation, on the other hand, is responsible for dividing the differences. Thus, when each pixel value is multiplied by a constant, the differences are also multiplied by the same constant, canceling out the contrast change. As a result, the description is not affected by changes in brightness.

Finally, quantized orientation map incorporating multiscale granularity is represented as follows:

$$q_{os} = \frac{1}{2}(o_{mg}) \times \frac{N_o}{180}, \quad (10)$$

$$q_{os}(f) = \{(s, t) | s, t \in q_{os} = j\}, \quad 0 \leq N_o = 1,$$

where s and t are spatial coordinates of orientation structure and j is the resultant value obtained across quantized value $N_o = 6$.

3.2.3. Local Contrast Information. The pixel intensity is used to describe the local contrast information [41], as the visual information depends on intensity, and specific contrast range improves the visibility of visual information and increases the performance of CBIR systems. The V element of HSV color model was used to extract the intensity feature. This is due to the fact that V component of HSV color model

called luma separates the chrominance element of image. The local contrast information map for input image $f(s, t)$ is calculated as [42]

$$\text{lcm}(s, t) = \sum_{s=1}^U \sum_{t=1}^V \max(p(s, t), \max(H(s, t), b(s, t))). \quad (11)$$

From the mathematical expression represented in (11), the constants s and t are the coordinates of input image $f(s, t)$. The dimensions of input image are represented by m and n . Then, the local contrast information was quantized at 10 levels to retain the salient information. The final intensity map was obtained and mathematically represented by the following expression:

$$\begin{aligned} q_{\text{lcm}}(s, t) &= IM(s, t) * n_i, \\ q_{\text{lcm}}(s, t)_l &= \{(s, t) | (s, t) \in q_{\text{lcm}}(s, t)\} = l, \quad 0 \leq l \leq n_i - 1. \end{aligned} \quad (12)$$

3.3. Interlinked Microstructure Identification. Natural images contain large color, edge, and shape attribute, which are considered as low-level image attributes. More relevant images are retrieved from large datasets on the basis of color, edge, and shape attribute of image. Human eye shows sensitivity toward color and orientation features. Orientation reflects an informative content in image. Strong orientation gives a uniform pattern. As most natural standard images do not contain strong orientation, there is no uniform pattern. Natural images contain uniform and non-uniform patterns contributing spatial information, forming the microstructures. Orientation is not enough to discuss the image's spatial characteristics, so color, edge, orientation, and local contrast information are fused to describe the richer image contents. In this paper, microstructures are proposed on the basis of color, orientation, and local contrast information attributes. Color feature and microstructure which is obtained by associating texture and local contrast information are joined to obtain micro-color structure. Orientation attributes and microstructure which is derived by linking color and intensity features are joined to obtain micro-orientation structure. Micro-intensity map is derived by intensity features and micro-map which is derived by linking orientation and intensity features. As microstructures are combination of color, orientation, and intensity features, they will increase the retrieval performance. Quantized orientation and quantized intensity ranges are used to derive micro-color map. The whole orientation $O_{\text{orientation}}$ is divided into 3×3 small grids [43]. Orientation is scale and rotation invariant, so it is used to derive color map. As orientation is quantized at level 6, it can vary from 0 to 5. For horizontal and vertical location, pixel of 3lengths is defined for each grid. Suppose that $G_{ii} = G_{ii} + G_{ii+1} + G_{ii+2} + \dots + G_P$ are the grids which involve $O_{\text{orientation}}$ where $ii = 1, 2, \dots, P$ where P represents length of grid. Similarly, suppose that $G_{jj} = G_{jj} + G_{jj+1} + G_{jj+2} + \dots + G_Q$ are the grids which involve $O_{\text{intensity}}$ where $jj = 1, 2, \dots, Q$

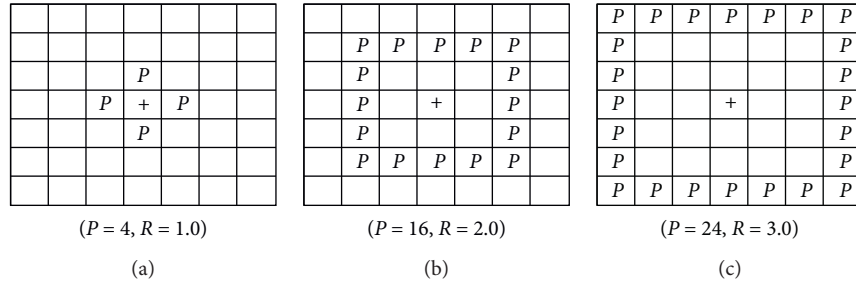


FIGURE 3: Squared symmetric neighborhood for different (P, R) . P represents neighbors and R represents the spatial resolution. Multi-resolution analysis is done by choosing P as 4, 16, and 24 and R as 1.0, 2.0, and 3.0.

where Q represents length of grid. Suppose that G_{ii} is the grid of length 3×3 moved on quantized orientation and is the grid of length 3×3 moved on quantized local contrast information. Relationship is established between the pixels in the center and the pixels in the surrounding areas as a result of resemblance measures, which are helpful in finding the regular patterns. Suppose that cp_o , cp_c are the central pixels and k_i and k_j are neighboring pixels of quantized orientation and intensity microstructures. While moving grid of 3×3 , neighbors of quantized orientation and intensity have the same value as central pixels of orientation and intensity; then, uniform or regular patterns and microstructure basic block are obtained. If no values of neighboring pixels of k_i and k_j are the same as those of the central pixels cp_o , cp_c , then irregular patterns are formed and microstructure basic block is not obtained. Five steps are followed to achieve a finally single image's microstructure.

- (1) Starting from $(0, 0)$, a grid of size 3×3 is moved on both quantized orientation and intensity from the left to the right and from the top to the bottom. The micro-map is established at $(0, 0)$ and is labeled as $M^1(s, t)$ on the basis of regular patterns in both quantized orientation and intensity where $0 \leq s \leq m-1$, $0 \leq t \leq n-1$.
- (2) At $(0, 1)$, a grid of size 3×3 is moved on both quantized orientation and intensity from the left to the right and from the top to the bottom. The micro-map is established at $(0, 1)$ and is labeled as $M^2(s, t)$ on the basis of regular patterns in both quantized orientation and intensity where $0 \leq s \leq m-1$, $1 \leq t \leq n-1$.
- (3) At $(1, 0)$, a grid of size 3×3 is moved on both quantized orientation and intensity from the left to

the right and from the top to the bottom. The micro-map is established at $(1, 0)$ and is labeled as $M^3(s, t)$ on the basis of regular patterns in both quantized orientation and intensity where $1 \leq s \leq m-1$, $0 \leq t \leq n-1$.

- (4) At $(1, 1)$, a grid of size 3×3 is moved on both quantized orientation and intensity from the left to the right and from the top to the bottom. The micro-map is established at $(1, 1)$ and is labeled as $M^4(s, t)$ on the basis of regular patterns in both quantized orientation and intensity where $1 \leq s \leq m-1$, $1 \leq t \leq n-1$.
- (5) Final micro-map of image is obtained and demonstrated as (x) by combining all four maps.

4. Feature Quantization

Extraction of discriminative set of attributes for retrieving images is a difficult task. After obtaining discriminative set of features, the second task is the image's attribute representation, that is, how to represent the extracted features. In this work, microstructures are suggested for feature representation. The main steps involved in forming correlated microstructures are shown in Figure 4.

4.1. Color Feature. Microstructure's value for the input image $f(s, t)$ is demonstrated as $f^1(s, t) = a$, where $a \in 0, 1, 2, \dots, m-1$ and m reflects the dimensionality of micro-color structure which is obtained on the basis of the relationship between microstructure of quantized orientation and intensity. The equation used to derive the microstructure features is as follows:

$$G(a) = \frac{L_1\{f(Pa_0) = a_0, f(Pa_i) = a_i | Pa_i - Pa_0 = 1\}}{8 \bar{L}_1\{f(Pa_0) = a_0\}}, \quad a_0 = a_i, i = \{1, 2, 3, \dots, 8\}. \quad (13)$$

Every block of input image is 3×3 , supposing that Pa_0 reflects the central pixel with position $Pa_0 = (s_0, t_0)$ and Pa_i reflects the neighboring pixels of central pixel with position $Pa_i = (s_i, t_i)$. The value for the central pixel is $Pa_0 = e$

neighboring pixels and denotes the number of cooccurring values for a_0 and a_i . $\bar{L}_1$ is used to represent the number of occurrences of a_0 . The micro-color structure is derived by establishing the relationship between microstructure image

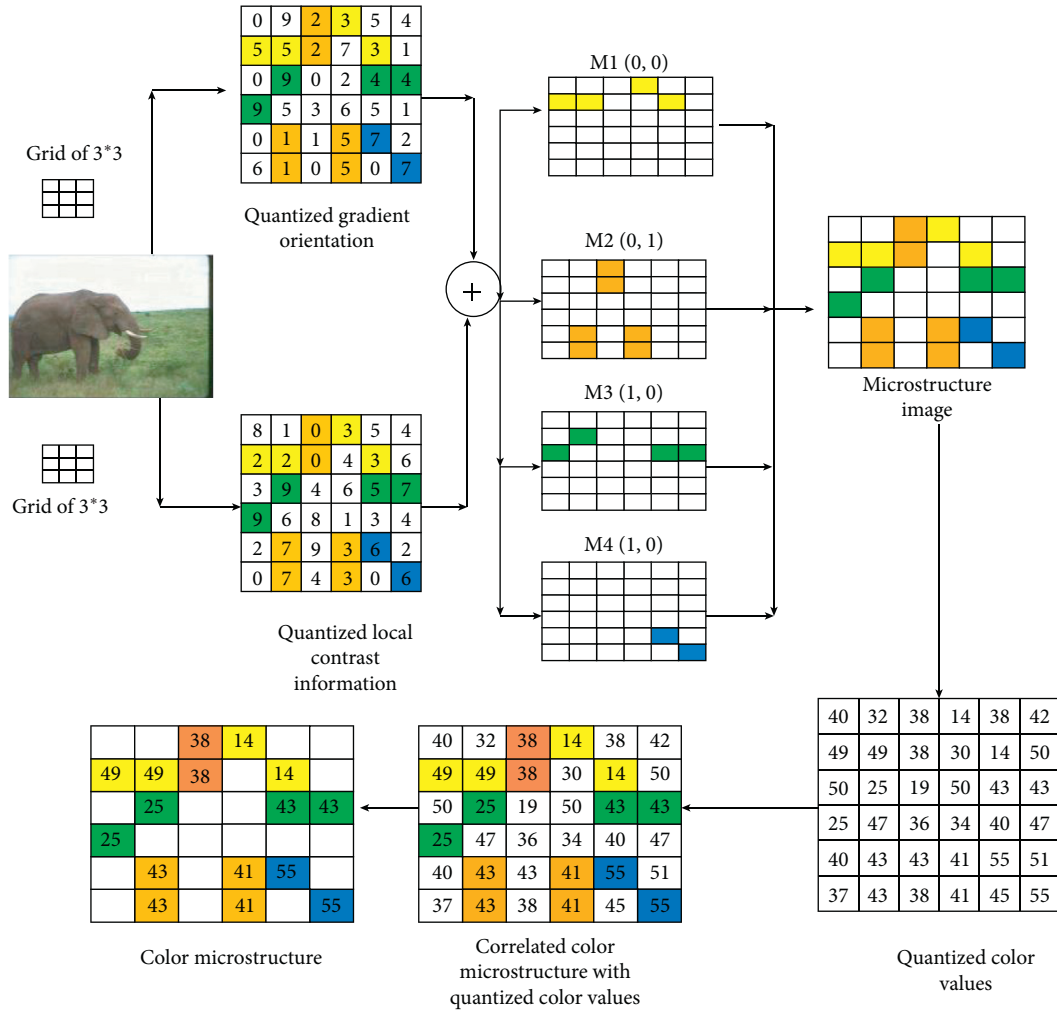


FIGURE 4: Correlated microstructure. It is formed by correlating the quantized color, orientation, and local contrast information quantized values.

$f^1(s, t)$ and quantized color structure q_{cs} . Final micro-color structure is obtained from quantized color map's pixel values, which are in regular region present in $f^1(s, t)$. The features of micro-color structure have 72 dimensions.

4.2. Orientation Feature. Microstructure's value for the input image $f(s, t)$ is demonstrated as $f^2(s, t) = b$, where

$$G(b) = \frac{L_2\{f(Pb_0) = a_0, f(Pb_i) = a_i | Pb_i - Pb_0 = 1\}}{8\overline{L_2}\{f(Pb_0) = b_0\}}, \quad b_0 = b_i, i = \{1, 2, 3, \dots, 8\}. \quad (14)$$

Every block of input image is 3×3 , supposing that Pb_0 reflects the central pixel with position $P_o = (s_o, t_o)$ and Pb_i reflects the neighboring pixels of central pixel with position $P_i = (s_i, t_i)$. The value for the central pixel is $Pb_0 = e$ neighboring pixels and denotes the number of cooccurring

$b \in 0, 1, 2, \dots, m-1$, and m reflects the dimensionality of micro-color structure which is obtained on the basis of the relationship between microstructure of quantized orientation and intensity. The equation used to derive the micro-structure features is as follows:

values for b_o and b_i . $\overline{L_2}$ is used to represent the number of occurrences of b_o . The micro-color structure is derived by establishing the relationship between microstructure image $f^2(s, t)$ and quantized orientation structure q_{os} . Final micro-color structure is obtained from quantized color map's pixel

values, which are in regular region present in $f^2(s, t)$. The features of micro-orientation structure have 6 dimensions.

4.3. Local Contrast Information. Microstructure's value for the input image $f(s, t)$ is demonstrated as $f^3(s, t) = c$, where

$$G(c) = \frac{L_3\{f(Pc_0) = c_0, f(Pc_i) = c_i | Pc_i - Pc_0 = 1\}}{8L_3\{f(Pc_0) = c_0\}}, \quad c_0 = c_i, i = \{1, 2, 3, \dots, 8\}. \quad (15)$$

Every block of input image is 3×3 , supposing that Pc_0 reflects the central pixel with position $P_o = (s_o, t_o)$ and Pc_i reflects the neighboring pixels of central pixel with position $P_i = (s_i, t_i)$. The value for the central pixel is $Pc_0 = e$ neighboring pixels and denotes the number of cooccurring values for c_0 and c_i . L_3 is used to represent the number of occurrences of c_0 . The micro-color structure is derived by establishing the relationship between microstructure image $f^3(s, t)$ and quantized local contrast structure q_{lcs} . Final micro-color structure is obtained from quantized color map's pixel values, which are in regular region present in $f^3(s, t)$. The features of micro local contrast information structure have 10 dimensions.

5. Experimental Results and Findings

In this section, experiment is performed on Corel dataset to evaluate the retrieval performance of NRLGO method. The performance of NRLGO is given as follows.

5.1. Dataset. Different databases are used for evaluating the retrieval accuracy of images. In our proposed work, Corel database is used for measuring efficiency of NRLGO. Corel database consists of Corel 1k, Corel 5k, and Corel 10k. Corel 1k contains 1000 images of different contents such as Africans, mountains, elephants, buildings, beaches, and horses. There are 10 categories and 100 images per category with dimension of 256×384 . Corel 5k is derived from Corel 10k and contains 5000 images of different contents such as waves, trees, lions, ducks, and food. There are 50 categories and 100 images per category. Corel 10k is a large dataset and contains 10000 images of different contents such as 250 cars, flowers, trains, furniture, butterflies, and tractors. There are 100 categories and 100 images per category with dimension of 192×128 .

5.2. Evaluation Parameters. Precision and recall are the most important parameters for evaluating retrieval performance. Precision of input image is described as ratio of number of images retrieved that are similar to total number of images retrieved. Recall of input image is defined as ratio of similar images retrieved to total number of similar images.

$c \in 0, 1, 2, \dots, m-1$, and m reflects the dimensionality of micro-color structure which is obtained on the basis of the relationship between microstructure of quantized orientation and intensity. The equation used to derive the micro-structure features is as follows:

$$P_I = \frac{N_s}{M}, \quad (16)$$

$$R_I = \frac{N_s}{M_s},$$

where M reflects the total number of images that have been retrieved and M_s reflects the total number of similar images that have been retrieved. Average precision and recall are also calculated to measure the image retrieval performance being 100; N_c describes the total number of images of Corel dataset in each category; and values 10, 50, and 100 describe the number of categories of Corel dataset. Average retrieval precision and recall are calculated by (18) and (20).

$$RP_I = \frac{1}{M_c} \sum_{j=1}^{M_c} P_I, \quad (17)$$

$$ARP = \frac{1}{Q_c} \sum_{k=1}^{Q_c} RP_I, \quad (18)$$

$$RR_I = \frac{1}{M_c} \sum_{j=1}^{M_c} R_I, \quad (19)$$

$$ARR = \frac{1}{Q_c} \sum_{k=1}^{Q_c} RR_I. \quad (20)$$

F-score has characteristics of both precision and recall, and it combines precision and recall into single similarity measure as shown in the following:

$$F_s = \frac{2 \times P_I \times R_I}{P_I + R_I}. \quad (21)$$

5.3. Distance Metric. Manhattan distance (L1) is used to measure the similarity between the search image and the images in dataset by using the following equation:

$$D(u, v) = \sum_{i=1}^{N_i} |u_i - v_i|, \quad (22)$$

$u_i = u_1, u_2, \dots, u_N$ is the collection of attributes extracted from database's images and $v_i = v_1, v_2, \dots, v_N$ is the collection of attributes extracted from search image, where W_i and $V_j = [1, 2, 3, \dots, N_i]$, where $N_i = 83 (72 + 6 + 5)$ is the dimension of feature vector. Manhattan distance is good to measure similarity for a large image database by reducing the computational costs.

5.4. Performance of NRLGO. Color, texture, and local contrast information are used to retrieve relevant images from large dataset. In proposed NRLGO, color, edges, and local contrast information are extracted from the color image. The effectiveness of NRLGO is obtained by using 192, 128, 108, and 72 color quantization levels; 6, 12, 18, 24, 30, and 36 orientation quantization levels; and 10, 15, and 20 local contrast information quantization levels.

Tables 1–3 demonstrate the efficient performance on Corel dataset with 5000 images by changing color and texture levels by keeping local contrast information values fixed at 5, 10, and 15. In Table 1 highest precision of 65.82% is achieved when $N_c = 72$ and $N_o = 6$, and in Table 2 highest precision of 65.32% is achieved when $N_c = 128$ and $N_o = 6$. Moreover, in Table 3 highest precision of 65.58% is achieved when $N_c = 128$ and $N_o = 24$. So, in the proposed NRLGO, we obtained the descriptor of dimensions $(72 + 6 + 5 = 83)$ by setting values for color $N_c = 72$, texture $N_t = 6$ and local contrast information $N_{ci} = 5$. It is observed that, in Tables 1–3, ARP decreases when $N_c = 128$; N_{ci} is set to 5, 10, and 15; and N_o values are selected from 6 to 36. This decrement is basically because of sensitivity to visual system toward continuously changing texture orientations. In some cases, nonuniform pattern of average retrieval precision is noticed; for example, in Table 3, when $N_c = 72$ and $N_{ci} = 15$, average retrieval precision firstly increases for N_o from 6 to 18, then decreases at $N_o = 24$, and again increases for N_o from 30 to 36. This irregular pattern occurred because of noise due to quantization that increases variability within the class and decreases the performance of NRLGO.

Table 4 shows the ARP obtained on Corel 5k by varying N_o from 6 to 36 and N_c from 16 to 32, 64, and 128 in RGB color space. Highest precision is achieved when color quantization is 128 and texture level is 6 at intensity level 5. It is shown from the table that HSV color space gave higher retrieval performance. As RGB color channels ignore the color characteristics, irregular pattern is obtained. Thus, by using HSV color space, average precision and recall have been increased. HSV color space is used in NRLGO. Not only is it effective toward user intentions, it also takes standard color characteristics into considerations. Hence, HSV color space increases average retrieval performances. Table 5 demonstrates retrieval accuracy obtained on Corel 1k, Corel 5k, and Corel 10k by using different similarity measures. In the proposed method, NRLGO, bin-by-bin similarity measures are applied. Similarity is measured between the search image and database image by using Manhattan distance. L1 increases the retrieval performance: performance of 83.5% on Corel dataset with 1000 images, 65.82% on Corel dataset with 5000 images, and 53.07% on

Corel dataset with 10000 images. L1 does not perform any square or square root calculation, so it has high performance for large datasets, while L2 has poor performance at higher distance values and has high computation. It is clear from Table 5 that average retrieval accuracy is smaller on Corel dataset with 5000 images and Corel dataset with 10000 images by using Euclidean distance.

Euclidean distance includes the square operation, so it does not perform well on larger distance values and is also computationally expensive. Furthermore, on Corel 1k, retrieval accuracy is less at square chord. Color, texture, and local contrast information attributes are selected for proposed method NRLGO. However, NRLGO performance is observed at different color, texture, and local contrast information's combinations. ARP and ARR at 7 different combinations are shown in Table 6 for performance comparisons.

It is evident that when the combination of color, texture, and local contrast information is used, better average retrieval precision (ARP) and average retrieval recall (ARR) are achieved. Besides, poor performance is achieved when using only color, texture, or local contrast information. Local contrast information is an essential feature as it distinguishes images on the basis of difference in illumination; however, it is less emphasized for efficient image retrieval methods. Standard images contain more color and intensity features. Therefore, contrast information is useful to distinguish between the images with the same contrast and the images with different contrast information. Proposed method NRLGO gave better retrieval performance in all combinations of color, texture, and local contrast information as compared to other descriptors on Corel dataset with 1000, 5000, and 10000 images. The retrieval accuracy of NRLGO descriptor is observed on Corel 1k in comparison with MTH [44], MSD [43], CMSD [2], CDH [45], SED [46], and ENN [47] descriptors. Table 7 shows the efficient retrieval accuracy of NRLGO in comparison to the state-of-the-art techniques at different categories of Corel dataset with 1000 images. Texture is penetrating feature in beach and mountain images. It is observed that beach category has variation in textural classifications. Therefore, human perception is more sensitive toward textural variations.

Table 7 shows that dinosaur's category has 100% retrieval performance. Category-wise result evaluation shows that average retrieval precision is low at beach category, i.e., 50.75%, in comparison with MTH, MSD, SED, ENN, CDH, LeNET-F6, and CMSD. In beach category, most of mountain category images are retrieved, so this results in decrement of ARP in both mountain and beach category. In Table 8, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 images are retrieved; a total of 100 images are retrieved per category; and R is the relevant images retrieved. Highest precision is achieved at top 10 images retrieved, and lowest precision is achieved at 100 images retrieved. To achieve multiresolution property that increases the retrieval accuracy, we performed experiment by varying P and R as shown in Table 9. Retrieval accuracy is increased when we select the value of P equal to 24 and R equal to 3.0. Similarly, we used different threshold values and achieved highest precision and recall when we

TABLE 1: Average retrieval precision and average retrieval recall of NRLGO at varying values of gradient orientation and color quantization levels by keeping local contrast information fixed at 5 on Corel 5k in HSV.

Color quantization	Texture classification (Corel 5k, contrast information = 5)											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
192	62.1	63.11	63.10	62.60	62.68	62.79	7.46	7.57	7.57	7.51	7.52	7.53
128	63.49	64.12	63.80	64.23	63.16	63.54	7.61	7.69	7.65	7.70	7.58	7.62
108	62.73	62.51	64.05	62.66	62.70	62.40	7.52	7.50	7.68	7.52	7.52	7.48
72	65.8	63.3	63.6	63.4	63.6	64.5	7.89	7.59	7.64	7.61	7.64	7.74

TABLE 2: Average retrieval precision and average retrieval recall of NRLGO at varying values of gradient orientation and color quantization levels by keeping local contrast information fixed at 10 on Corel 5k in HSV.

Color quantization	Texture classification (Corel 5k, contrast information = 10)											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
192	64.05	64.26	63.94	64.01	63.47	63.35	7.68	7.71	7.67	7.68	7.61	7.60
128	65.32	64.38	65.04	65.05	64.62	64.56	7.89	7.72	7.80	7.80	7.75	7.74
108	64.71	64.15	63.68	65.07	63.68	63.63	7.76	7.69	7.64	7.80	7.64	7.63
72	64.87	65.30	64.67	65.40	64.35	65.37	7.78	7.83	7.76	7.85	7.72	7.84

TABLE 3: Average retrieval precision and average retrieval recall of NRLGO at varying values of gradient orientation and color quantization levels by keeping local contrast information fixed at 15 on Corel 5k in HSV.

Color quantization	Texture classification (Corel 5k, contrast information = 15)											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
192	64.59	63.76	63.96	64.21	64.06	63.85	7.75	7.65	7.67	7.70	7.68	7.66
128	65.47	65.51	64.94	65.58	64.56	64.75	7.85	7.86	7.79	7.87	7.74	7.77
108	64.80	64.30	64.94	63.90	64.69	65.09	7.77	7.71	7.79	7.66	7.76	7.81
72	65.30	65.32	65.46	64.75	65.45	65.16	7.83	7.83	7.85	7.77	7.85	7.82

TABLE 4: Average retrieval precision and average retrieval recall of NRLGO at varying values of gradient orientation and color quantization levels by keeping local contrast information fixed at 5 on Corel 5k in RGB.

Color quantization level	Texture classification (Corel 5k, contrast information = 5)											
	Precision (%)						Recall (%)					
	6	12	18	24	30	36	6	12	18	24	30	36
128	60.1	60.5	59.9	59.5	59.8	59.7	7.07	7.10	7.11	7.06	7.13	7.09
64	58.5	57.4	57.8	55.2	56.9	58.5	6.80	6.69	6.83	6.81	6.81	6.85
32	53.5	50.2	55.0	55.6	54.7	49.4	6.20	6.23	6.31	6.30	6.27	6.25
16	44.4	48.7	46.6	47.4	47.6	48.5	5.30	5.39	5.47	5.50	5.55	5.57

TABLE 5: Average retrieval precision and average retrieval recall of NRLGO by varying combinations of gradient orientation and color quantization levels by keeping contrast information fixed at 5 on Corel 5k in RGB.

Dataset	Performance	Distance or similarity metrics					
		Chi-square	Square chord	MTH	Euclidean distance	Shepherd	Manhattan
Corel 1k	Precision (%)	73.5	76.70	83.45	97.91	83.20	83.50
	Recall (%)	8.64	9.20	10.01	11.75	9.98	10.02
Corel 5k	Precision (%)	55.42	52.89	59.64	19.91	64.00	65.8
	Recall (%)	6.83	6.34	7.15	2.39	7.68	7.89
Corel 10k	Precision (%)	47.91	40.96	49.93	10.00	52.72	53.07
	Recall (%)	5.53	4.91	5.99	12.00	6.32	6.36

TABLE 6: Performance of NRLGO at Corel dataset with 1000, 5000, and 1000 images at different combinations of color, gradient orientation, and local contrast information vector.

Dataset	Performance	Color	Texture	Contrast	Texture + contrast	Color + texture	Color + contrast	Proposed
Corel 1k	Precision (%)	76.8	54.6	52.8	59.4	79.5	78.25	83.5
	Recall (%)	9.02	6.38	6.10	6.90	9.22	9.30	9.96
Corel 5k	Precision (%)	60.2	27.5	31.5	39.5	61.6	61.8	65.8
	Recall (%)	7.15	3.06	3.89	4.50	7.30	7.25	7.89
Corel 10k	Precision (%)	48.5	20.5	25.9	28.6	50.7	49.3	53.07
	Recall (%)	5.51	2.30	3.05	3.45	5.80	5.95	6.36

TABLE 7: Category-wise performance comparison of NRLGO at Corel 1k with state-of-the-art descriptors.

Category	Performance	MTH [44]	MSD [43]	SED [46]	ENN [47]	CDH [45]	CMSD [12]	LeNet-F6 [48]	Proposed
African	Precision	69.17	83.33	82.50	85	77.50	86.66	80	88.25
	Recall	8.30	10.00	9.90	9.00	9.30	10.40	8.00	10.30
Beach	Precision	61.67	43.33	28.33	75	56.67	42.08	60	50.75
	Recall	7.40	5.20	3.40	7.00	6.80	5.05	6.5	5.15
Building	Precision	45.83	63.33	47.50	70	47.50	81.66	75	85.00
	Recall	5.50	7.60	5.70	6.00	5.70	9.80	8.00	9.95
Bus	Precision	68.33	76.67	73.33	75	71.67	81.66	80	83.70
	Recall	8.20	9.20	8.80	7.00	8.60	9.80	9.00	9.95
Dinosaur	Precision	100.00	100.00	90.00	100	100	100	100	100
	Recall	12.00	12.00	10.80	12	12.00	12.00	12.00	12.00
Elephant	Precision	70.83	65.00	55.00	75	62.50	72.08	85	100
	Recall	8.50	7.80	6.60	7.00	7.50	8.65	10.00	12.00
Flower	Precision	75.00	86.67	72.50	80	60.83	84.16	75	86.20
	Recall	9.00	10.40	8.70	8.00	7.30	10.10	8.00	10.20
Horse	Precision	100.00	97.50	62.50	85	91.67	94.16	90	95.20
	Recall	12.00	11.70	7.50	9.00	11.00	11.30	11.00	11.40
Mountain	Precision	39.17	29.17	40.00	65	44.17	50.00	60	60.05
	Recall	4.70	3.50	4.80	5.00	5.30	6.00	6.5	6.10
Food	Precision	52.50	76.67	64.17	55	45.00	92.91	75	94.9
	Recall	6.30	9.20	7.70	3.00	5.40	11.15	8.00	11.38
Average	Precision	68.25	72.16	61.58	76.5	65.75	78.54	78	83.5
	Recall	8.19	8.66	7.39	7.3	7.89	9.42	8.7	9.96

TABLE 8: Top 10 to 100 images are retrieved using NRLGO method for precision parameter.

P class	10	20	30	40	50	60	70	80	90	100
African	0.65	0.90	0.73	0.75	0.70	0.66	0.64	0.51	0.61	0.60
Beach	0.40	0.75	0.66	0.60	0.60	0.54	0.52	0.50	0.47	0.46
Building	1	0.95	0.86	0.70	0.72	0.71	0.71	0.58	0.67	0.58
Bus	1	1	1	1	0.98	0.91	0.91	0.86	0.82	0.81
Dinosaur	1	1	1	1	1	1	1	1	1	1
Elephant	1	0.50	0.76	0.60	0.60	0.51	0.48	0.45	0.43	0.33
Flower	1	0.95	0.93	0.87	0.78	0.71	0.64	0.67	0.55	0.52
Horse	1	0.60	1	1	0.88	0.96	0.93	0.98	0.84	0.82
Mountain	0.60	0.70	0.60	0.57	0.44	0.51	0.48	0.57	0.46	0.44
Food	0.70	0.60	0.60	0.57	0.62	0.60	0.58	0.57	0.56	0.53
Average (R/P)	0.835	0.795	0.741	0.76	0.732	0.71	0.68	0.669	0.64	0.609

selected the threshold value equal to 2. NRLGO extracts noise-robust features that remain invariant to variation in scale, illumination, and orientation. Experiment is performed on Corel dataset by adding Gaussian noise for texture recognition. Gaussian Noise of 5, 10, and 15 (%) is added to dinosaur color image

In Table 9, by using different values of P and R, the multiresolution property is analyzed. Highest accuracy is

observed when we choose the value of P equal to 24 and R equal to 3.

In Table 10, different values of s are used. Highest precision and recall are observed when we set threshold value equal to $s=2$.

NRLGO firstly converts the noisy color images to gray-scale images and then extracts the edges perfectly even in the presence of noise. It extracts the edges under different noise

TABLE 9: Retrieval accuracy observed at 9 different angles at Corel 1k dataset.

(P, R)	(8, 1)	(16, 2)	(24, 3)	Average
Accuracy	82.4	83.5	84.6	83.5

TABLE 10: Retrieval accuracy observed at different threshold values using Corel 1k dataset. Highest retrieval accuracy is achieved at $s = 2$.

Noise resistant	Threshold values (s)					
	1	2	3	4	6	10
Precision (%)	83.4	83.5	82.8	82.3	82.7	83.2
Recall (%)	9.91	9.96	9.94	9.88	9.75	9.94

values as shown in Figure 5. Accordingly, the accuracy of NRLGO is good in comparison to state-of-the-art descriptors on Corel dataset with 1000 images.

6. Discussion

An image retrieval approach, based on Noise Resilient Local Gradient Orientation, that reduces the noise is proposed. Experiment is performed using Corel dataset, i.e., Corel dataset with 1000, 5000, and 10000 images, and the NRLGO achieves higher retrieval accuracy in comparison to state-of-the-art image techniques based on feature extraction techniques.

6.1. Comparison with Other Techniques. Natural images consist of regular and irregular textural patterns. Different textural descriptors like GLCM [25], Gabor features [28], SQ-SpatioGram [49], CMSD [2], and GMM-mSpatioGram [49] are proposed to check the performance on these regular and irregular patterns. It is observed that these textural descriptors perform well on regular patterns; their performance is not good at irregular patterns.

Table 11 demonstrates the retrieval accuracy of proposed NRLGO descriptor with uniform codes as compared to state-of-the-art methods, namely, GLCM [50], SQ-SpatioGram [49], GMM-mSpatioGram [49], SED [25], CMSD [2], and CPV-THF [42]. Experimental result shows that increase in retrieval accuracy of 27.8%, 24.13%, 14%, 9.5%, 2.66%, 1.9%, and 5.5% is achieved as compared to GLCM, SQ-SpatioGram [49], GMM-mSpatioGram [49], SED, CMSD, and CPV-THF on Corel 5k. On Corel 10k, for GLCM, SED, CMSD, CPV-THF, and LeNET-F6 [48], retrieval accuracy increase of 21.17%, 15.43%, 5.82%, 6.67%, 2.82%, and 79% is observed. Table 12 shows retrieval performance of proposed descriptor with respect to texture, color, and shape in comparison with state-of-the-art descriptor at Corel dataset with 5000 images and Corel dataset with 10000 images. Retrieval accuracy of NRLGO descriptor as compared to other descriptors is 29.6%, 26.4%, 29.98%, 17.36%, 27.32%, 5.52%, 2.82%, and 12.06% greater on Gabor [51], EHD [8], color moment [18], LBP [52], STH [53], and MTSD [54] at Corel dataset with 5000 images. At Corel dataset with 10000

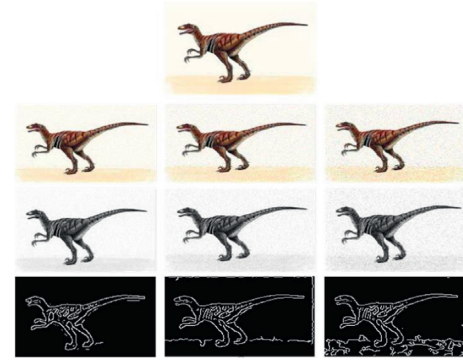


FIGURE 5: Extraction of edges in presence of Gaussian Noise. Firstly, Gaussian noise of 5, 10, and 15 (%) is inserted. Secondly, conversion of RGB to gray scale is done; then, NRLGO is applied to detect the edges.

images, the accuracy is 23.97%, 20.77%, 26.69%, 18.06%, 5.04%, and 1.12% on Gabor, EHD, color moment, LBP, STH, and MTSD.

Increase in retrieval precision and recall shows that the proposed method outperforms the shape, orientation, and color techniques. Figures 6–8 demonstrate performance of NRLGO with state-of-the-art techniques in form of curve using precision and recall parameters. x -axis is labeled as recall (%) and y -axis as precision (%). Precision-recall curves are drawn opposite to each other. For example, if average retrieval precision of descriptor is high, then obtained curve is drawn away from the point of origin. If ARP is low, then precision-recall will be short. We have also measured performance of NRLGO on the basis of F-score (%) with state-of-the-art methods like GA [55], GMM-mSpatioGram [49], ODBTC [56], Tetralet Transform [57] and BiCBIR [58] as shown in Table 13 with the increase of F-score of 0.087%, 0.062%, 0.053%, 0.03%, and 0.015%.

When there is uncertainty in the performance of a descriptor, curve with different turning point is obtained. When the performance of two descriptors is the same in image category, then the curves overlapped at some point. NRLGO outperforms on Corel 1000, Corel 5000, and Corel 10000 the other state-of-the-art descriptors as shown in Figure 6–8. These figures indicate that the ARP of proposed NRLGO is high, so precision-recall curve lies away from the point of origin. Proposed method NRLGO has a curve with few turning points and outperforms other descriptors for image categories of each dataset. MTH, SED, and MTH are textural descriptors used for efficient image retrieval including the correlation between texture and color, so local contrast information is missing. As correlation achieved by using intensity feature, the poor performance of MTH, SED, and STH is achieved. GLCM, EHD, and Gabor features are textural descriptors; they gave poor performance on natural images because using texture descriptor demonstrates limited texture classification of an image. Gabor filter is also used in textural classification due to the high relation between its cooccurrence and texture. EOAC does not consider textural characteristics of an image. MSD established correlation between color and orientation only and does not

TABLE 11: Performance of NRLGO is compared with uniform pattern-based state-of-the-art methods on Corel dataset with 5000 images and Corel dataset with 10000 images.

Dataset	Performance	CPV-THF [42]	GLCM [50]<	SQs [49]	GMMs [49]	SED [46]	CMSD [12]	Proposed
Corel 5k	Precision (%)	63.90	38.0	41.67	51.80	56.3	63.14	65.8
	Recall (%)	7.66	1.08	5.13	6.22	21.3	7.57	7.89
Corel 10k	Precision (%)	52.28	31.9	37.64	47.25	46.4	50.25	53.07
	Recall (%)	6.27	.63	4.51	5.67	18.3	6.03	6.36

TABLE 12: Performance comparison of NRLGO with color, texture, and shape based state-of-the-art techniques on Corel 5k and Corel 10k.

Database	Parameters	Gabor [51]	EHD [8]	CM [18]	LBP [52]	STH [53]	MTSD [54]	Proposed
Corel 5k	Precision (%)	36.2	39.4	35.82	48.44	60.28	62.98	65.8
	Recall (%)	4.35	4.74	4.29	28.87	7.23	9.45	7.89
Corel 10k	Precision (%)	29.1	32.3	26.38	35.01	48.03	51.95	53.07
	Recall (%)	3.50	3.88	3.16	18.59	5.76	7.79	6.36

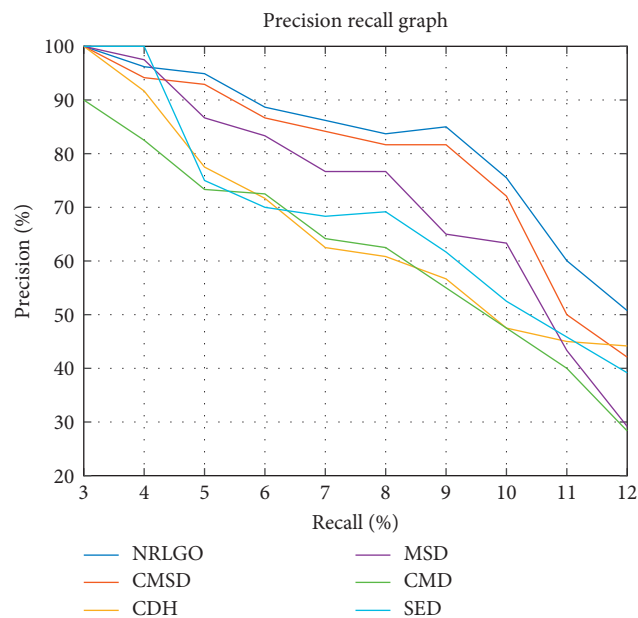


FIGURE 6: Performance of NRLGO is compared with state-of-the-art descriptors on Corel 1k using precision and recall parameter.

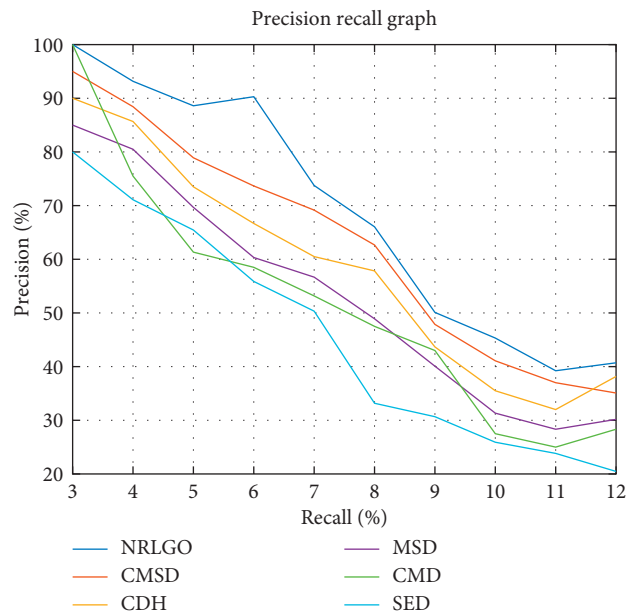


FIGURE 7: Performance of NRLGO is compared with state-of-the-art descriptors on Corel 5k using precision and recall parameters.

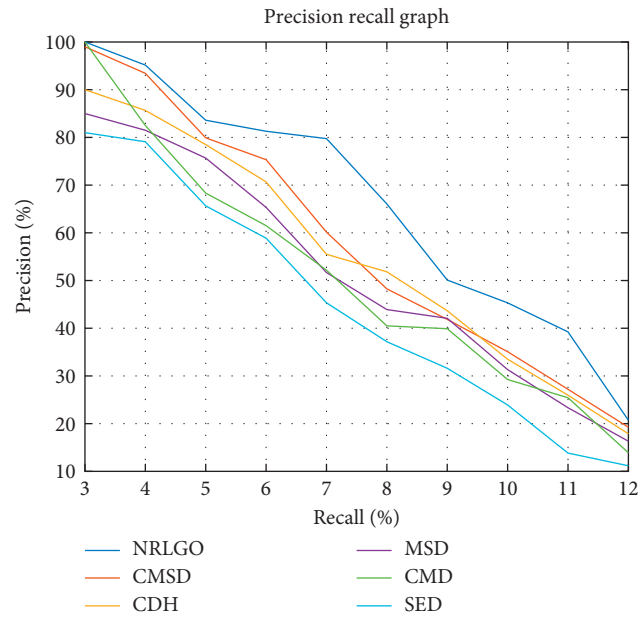


FIGURE 8: Performance of NRLGO is compared with state-of-the-art descriptors on Corel 10k using precision and recall parameters.

TABLE 13: F-score is compared with state-of-the-art descriptors at Corel 1k.

Category	GA [55]	ODBTC [56]	GMM + spatiogram [49]	Tetrolet transform [57]	BiCBIR [58]	Proposed
African	0.187	0.282	0.242	0.317	0.317	0.321
Beach	0.179	0.155	0.217	0.200	0.250	0.265
Building	0.203	0.227	0.235	0.183	0.283	0.311
Bus	0.298	0.295	0.297	0.333	0.333	0.352
Dinosaur	0.328	0.331	0.333	0.333	0.333	0.30
Elephant	0.193	0.244	0.235	0.300	0.300	0.342
Flower	0.300	0.321	0.316	0.333	0.333	0.321
Horse	0.260	0.313	0.306	0.333	0.333	0.342
Mountain	0.171	0.158	0.241	0.250	0.267	0.286
Food	0.231	0.269	0.263	0.333	0.317	0.380
Average	0.235	0.260	0.269	0.292	0.307	0.322

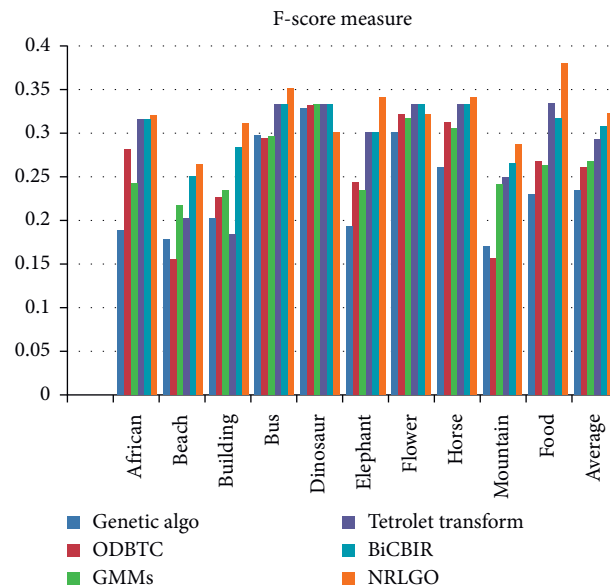


FIGURE 9: F-score is calculated for each category in Corel 1k and compared with the state-of-the-art techniques.

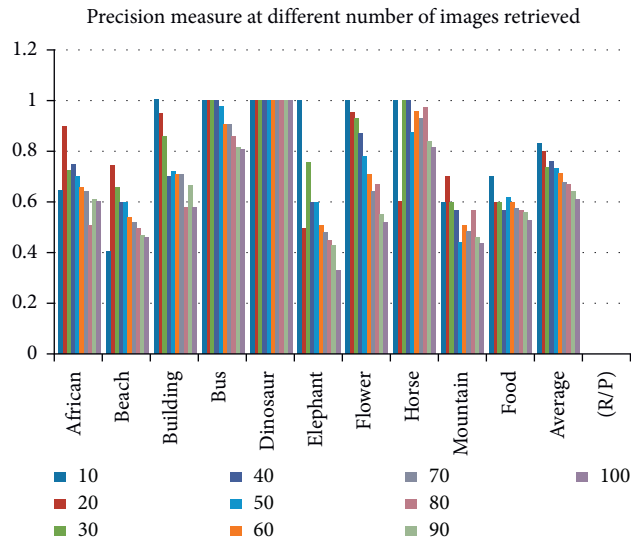


FIGURE 10: Precision is calculated for the number of images retrieved at Corel 1k.



FIGURE 11: Retrieval of top 12 images as a result of search image selected from African category.

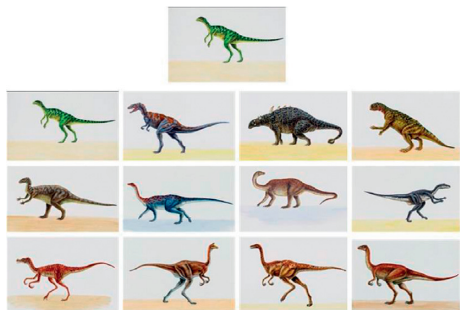


FIGURE 12: Retrieval of top 12 images as a result of search image selected from dinosaur category.



FIGURE 13: Retrieval of top 12 images as a result of search image selected from bus category.



FIGURE 14: Retrieval of top 12 images as a result of search image selected from flower category.

take local contrast information into account. The color moment (CM) approach considers only the spatial data of pixels that are near boundary of an image. LBP takes only textural characteristics of an image. A GUI application is established; it takes query image from African, dinosaur, bus, and flower category by the user and then retrieves 12 images from each category. Performance of NRLGO is observed on the basis of image's features like color, texture, and shape (uniform-nonuniform).

In Figure 9, F-score at Corel 1k is calculated and then its average is obtained. It shows that highest F-score of 0.322 is achieved by using NRLGO method in comparison with state-of-the-art descriptors. Figure 10 shows that highest precision of 0.835 is achieved at retrieval of top 10 images and lowest precision of 0.60 is achieved at retrieval of 100 images. It shows that retrieval performance decreased when the retrieval number of images increased from 10 to 100.

Retrieval of top 12 images done against the search images selected from the African, dinosaur, bus, and flower category is shown in Figures 11–14.

In the proposed NRLGO method, NRLBP solves the issue of small pixel difference. It is used to save the local structures from noise by finding an uncertain state. It finds the value of uncertain state on the basis of corrected bits of LBP code. Uniform codes describe the image local structure and nonuniform codes describe the noise patterns. Thus, for finding uncertain state, error correction method that recovers the nonuniform patterns is used. Color feature is obtained in HSV color space by quantizing input image into 72 quantization levels. Texture classification is done by using local gradient orientation. Multiscale gradient orientation is used to extract multiple granularity features. It will reduce noise and illumination variations. V component of HSV is used to extract the intensity attribute. Similarity measure and efficient indexing are done by using Manhattan distance. Color, texture, and local contrast information micro-maps are used to demonstrate extracted features. They describe the important content that is within the uniform region and hide the irrelevant content.

7. Conclusion

A feature descriptor named Noise Resilient Local Gradient Orientation is proposed in this paper for improving the retrieval performance. NRLGO relies on noise removal,

color, texture classifications, and local contrast information. Small pixel difference causes noise in color image which leads to changing code abruptly resulting in poor feature extraction. NRLBP protects the local structure from noise by finding the uncertain state first. HSV color space is quantized into 72 levels to obtain the color attribute. Texture classification is done by using LGO approach at 6 quantization levels, which quantize orientation further into supreme orientation. V component of HSV is quantized into 10 levels to obtain the local contrast information. Incorporation of multiresolution gradient orientation gave better texture detection, which increases the relationship between color, texture, and local contrast information. To achieve the resemblance between search image and image in the database, Manhattan distance is used. Correlation is established between color, texture, and local contrast information. For image characteristic representation, correlation is developed between the color, orientation, and local contrast data. Microstructures are developed on the basis of correlation information to develop more fine details in subject field. Multiresolution orientation increased the retrieval performance of proposed NRLGO. NRLGO extracts noise-robust features that remain invariant to variation in scale, illumination, and orientation. Experimental results show that NRLGO has improved retrieval performance of texture, local contrast information, and color features in comparison with state-of-the-art descriptors on Corel with 1000 images, 5000 images, and 10000 images.

Data Availability

The dataset Corel 5k is used during this study which is publicly available at corel5k.2009.1111.tar.bz2 and <https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval>.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by Education and Research Promotion Program of KOREATECH (2021).

References

- [1] R. Ashraf, "Content based image retrieval by using color descriptor and discrete wavelet transform," 2018.
- [2] A. Shinde, A. Rahulkar, and C. Patil, "Content based medical image retrieval based on new efficient local neighborhood wavelet feature descriptor," *Biomedical Engineering Letters*, vol. 9, no. 3, pp. 387–394, 2019.
- [3] H. Road, "E-commerce," 2016.
- [4] G. Wan and Z. Liu, "Content-based information retrieval and digital libraries," *Information Technology and Libraries*, vol. 27, no. 1, pp. 41–47, 2008.
- [5] J. Ren and Y. Shen, "A novel image retrieval based on representative colors," *Image Processing and Computer Vision*, vol. 16, pp. 102–107, 2003.

- [6] M. Indu and K. V. Kavitha, "Survey on sketch based image retrieval methods," in *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, pp. 1–4, Nagercoil, India, March 2016.
- [7] S. A. Orand, "Sketch based image retrieval using a soft computation of the histogram of edge local orientations (s-helo)," *Computer Vision Research Group*, vol. 32, pp. 2998–3002, 2014.
- [8] S. Wang, J. Zhang, T. X. Han, Z. Miao, and Z. Miao, "Sketch-based image retrieval through hypothesis-driven object boundary selection with HLR descriptor," *IEEE Transactions on Multimedia*, vol. 17, no. 7, pp. 1045–1057, 2015.
- [9] H. Xie, Y. Ji, and Y. Lu, "An analogy-relevance feedback cbir method using multiple features," 2016.
- [10] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi, "Long term learning in image retrieval systems using case based reasoning," *Engineering Applications of Artificial Intelligence*, vol. 35, pp. 26–37, 2014.
- [11] R. Mandal, P. P. Roy, U. Pal, and M. Blumenstein, "Bag-of-visual-words for signature-based multi-script document retrieval," *Neural Computing and Applications*, vol. 31, no. 10, pp. 6223–6247, 2019.
- [12] H. Dawood, M. H. Alkinani, A. Raza, H. Dawood, R. Mehboob, and S. Shabbir, "Correlated microstructure descriptor for image retrieval," *IEEE Access*, vol. 7, no. c, pp. 55206–55228, 2019.
- [13] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 1–28, 2004.
- [14] Y. Ke and R. Sukthankar, "PCA-SIFT: a more distinctive representation for local image descriptors," 2004.
- [15] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [16] O. A. Vatamanu, M. Frandescu, D. Lungeanu, and G. I. Mihalaş, "Content based image retrieval using local binary pattern operator and data mining techniques," *Studies in Health Technology and Informatics*, vol. 210, pp. 75–79, 2015.
- [17] S. R. Dubey, S. K. Singh, R. K. Singh, and S. Member, "Multichannel decoded local binary patterns for content-based image retrieval," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4018–4032, 2016.
- [18] S. R. Singh and S. Kohli, "Enhanced CBIR using color moments, HSV histogram, color auto correlogram, and gabor texture," *International Journal of Computer Systems*, vol. 2, no. 5, pp. 161–165, 2015.
- [19] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," 2007.
- [20] W.-H. Liao, "Region description using extended local ternary patterns," in *Proceedings of the 2010 20th International Conference on Pattern Recognition*, pp. 1007–1010, Istanbul, Turkey, August 2010.
- [21] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikäinen, and S. Z. Li, "Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes," in *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1301–1306, San Francisco, CA, USA, June 2010.
- [22] S. D. Thepade and Y. D. Shinde, "Improvisation of content based image retrieval using color edge detection with various gradient filters and slope magnitude method," in *Proceedings of the 2015 International Conference on Computing Communication Control and Automation*, pp. 625–628, Mumbai, India, February 2015.
- [23] P. Bao, L. Lei Zhang, and X. Xiaolin Wu, "Canny edge detection enhancement by scale multiplication," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 9, pp. 1485–1490, 2005.
- [24] H. Qazanfari, H. Hassanpour, and K. Qazanfari, "Content-based image retrieval using HSV color space features," *International Journal of Computer and Information Technology*, vol. 13, no. 10, pp. 537–545, 2019.
- [25] A. Al-mohamade and O. Bchir, "Multiple query content-based image retrieval using relevance feature weight learning," 2020.
- [26] B. Zafar, R. Ashraf, N. Ali et al., "A novel discriminating and relative global spatial image representation with applications in CBIR," *Applied Sciences*, vol. 8, no. 11, pp. 2242–2323, 2018.
- [27] R. Ashraf, M. Ahmed, U. Ahmad et al., "MDCBIR-MF: multimedia data for content-based image retrieval by using multiple features," *Multimedia Tools and Applications*, vol. 79, no. 13–14, pp. 8553–8579, 2020.
- [28] N. Hor and S. Fekri-Ershad, "Image retrieval approach based on local texture information derived from predefined patterns and spatial domain information," 2019, <http://arxiv.org/abs/1912.12978>.
- [29] F. Tajeripour, M. Saberi, and S. Fekri-Ershad, "Developing a novel approach for content based image retrieval using modified local binary patterns and morphological transform," *The International Arab Journal of Information Technology*, vol. 12, no. 6, pp. 574–581, 2015.
- [30] N. Tadi Bani and S. Fekri-Ershad, "Content-based image retrieval based on combination of texture and colour information extracted in spatial and frequency domains," *The Electronic Library*, vol. 37, no. 4, pp. 650–666, 2019.
- [31] M. Singha, "Content based image retrieval using color and texture," *Signal & Image Processing: An International Journal*, vol. 3, no. 1, pp. 39–57, 2012.
- [32] N. Ahmed, H. M. S. Asif, and G. Saleem, "Leaf image-based plant disease identification using color and texture features," 2021, <http://arxiv.org/abs/2102.04515>.
- [33] A. Alzu'bi, A. Amira, and N. Ramzan, "Content-based image retrieval with compact deep convolutional features," *Neurocomputing*, vol. 249, pp. 95–105, 2017.
- [34] A. B. Yandex and V. Lempitsky, "Aggregating local deep features for image retrieval," in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1269–1277, Santiago, CH, USA, August 2015.
- [35] Y. Kalantidis, C. Mellina, and S. Osindero, "Cross-dimensional weighting for aggregated deep convolutional features," *Lecture Notes in Computer Science*, vol. 9913, pp. 685–701, 2016.
- [36] B. Leibe, J. Matas, N. Sebe, and M. Welling, "Preface," *Lecture Notes in Computer Science(Including Subser. Notes Bio-informatics)*, vol. 9906, 2016.
- [37] M. Tzelepi and A. Tefas, "Deep convolutional learning for content based image retrieval," *Neurocomputing*, vol. 275, pp. 2467–2478, 2018.
- [38] J. Y.-H. Ng, F. Yang, and L. S. Davis, "Exploiting local features from deep networks for image retrieval," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 53–61, Boston, MA, USA, June 2015.
- [39] J. Ren, X. Jiang, J. Yuan, S. Member, and J. Yuan, "Noise-resistant local binary pattern with an embedded error-correction mechanism," *IEEE Transactions on Image Processing*, vol. 22, no. 10, pp. 4049–4060, 2013.

- [40] J. Shiguang Shan, W. L. D. Chu He, Z. Guoying, M. Pietikäinen, C. Xilin, and G. Wen, "WLD: a robust local image descriptor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1705–1720, 2010.
- [41] J. Qiu, H. Xu, and Z. Ye, "Color constancy by reweighting image feature maps," *IEEE Transactions on Image Processing*, vol. 29, pp. 5711–5721, 2020.
- [42] A. Raza, H. Dawood, H. Dawood, S. Shabbir, R. Mehboob, and A. Banjar, "Correlated primary visual texton histogram features for content base image retrieval," *IEEE Access*, vol. 6, pp. 46595–46616, 2018.
- [43] G.-H. Liu, Z.-Y. Li, L. Zhang, and Y. Xu, "Image retrieval based on micro-structure descriptor," *Pattern Recognition*, vol. 44, no. 9, pp. 2123–2133, 2011.
- [44] G.-H. Liu, L. Zhang, Y.-K. Hou, Z.-Y. Li, and J.-Y. Yang, "Image retrieval based on multi-texton histogram," *Pattern Recognition*, vol. 43, no. 7, pp. 2380–2389, 2010.
- [45] G.-H. Liu and J.-Y. Yang, "Content-based image retrieval using color difference histogram," *Pattern Recognition*, vol. 46, no. 1, pp. 188–198, 2013.
- [46] X. Wang and Z. Wang, "A novel method for image retrieval based on structure elements' descriptor," *Journal of Visual Communication and Image Representation*, vol. 24, no. 1, pp. 63–74, 2013.
- [47] 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, 2015.
- [48] H. Liu, B. Li, X. Lv, and Y. Huang, "Image retrieval using fused deep convolutional features," *Procedia Computer Science*, vol. 107, pp. 749–754, 2017.
- [49] 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, 2016.
- [50] E. M. Gebejes, "Master, and a samples, "texture characterization based on grey-level Co-occurrence matrix," *Informatics and Management Science*, vol. 65, pp. 375–378, 2013.
- [51] Z.-C. Huang, P. P. K. Chan, W. W. Y. Ng, and D. S. Yeung, "Content-based image retrieval using color moment and Gabor texture feature," in *Proceedings of the 2010 International Conference on Machine Learning and Cybernetics*, pp. 719–724, Qingdao, China, July 2010.
- [52] P. Srivastava and A. Khare, "Integration of wavelet transform, Local Binary Patterns and moments for content-based image retrieval," *Journal of Visual Communication and Image Representation*, vol. 42, pp. 78–103, 2017.
- [53] A. Raza, T. Nawaz, H. Dawood, and H. Dawood, "Square texton histogram features for image retrieval," *Multimedia Tools and Applications*, vol. 78, no. 3, pp. 2719–2746, 2019.
- [54] M. Zhao, H. Zhang, and J. Sun, "A novel image retrieval method based on multi-trend structure descriptor," *Journal of Visual Communication and Image Representation*, vol. 38, pp. 73–81, 2016.
- [55] M. E. Elalami, "A novel image retrieval model based on the most relevant features," *Knowledge-Based Systems*, vol. 24, no. 1, pp. 23–32, 2011.
- [56] J. M. Jing-Ming Guo and H. Prasetyo, "Content-based image retrieval using features extracted from halftoning-based block truncation coding," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 1010–1024, 2015.
- [57] J. Pradhan, S. Kumar, A. K. Pal, and H. Banka, "A hierarchical CBIR framework using adaptive tetrolet transform and novel histograms from color and shape features," *Digital Signal Processing*, vol. 82, pp. 258–281, 2018.
- [58] S. Singh and S. Batra, "An efficient bi-layer content based image retrieval system," *Multimedia Tools and Applications*, vol. 79, no. 25–26, pp. 17731–17759, 2020.