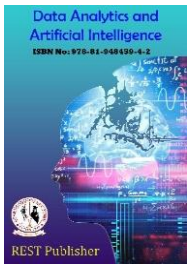
Data Analytics and Artificial Intelligence
Vol: 3(7), 2023
REST Publisher; ISBN: 978-81-948459-4-2
Website: http://restpublisher.com/book-series/daai/



Content Based Image Retrieval System by Concentrating both Local and Global features with Enhanced Deep Learning Neural Network Classifier

N. Parvin
Adhiyaman Arts and Science College for Women,Uthangarai,Tamilnadu.,India
*Corresponding Author Email: parvinfairose@gmail.com

Abstract: *Content-Based Image Retrieval (CBIR) is an efficient system for searching the desired image from a large database with the help of the various statistical features extracted from the image. The existing methods developed for image retrieval have failed to meet the user requirements, mainly due to the inability of retrieving accurate images and consumes more time for retrieving. To overcome these issues, this paper proposes an RWSM similarity calculation measure and global and local features for retrieving the images. The proposed methodology consists of preprocessing, feature extraction, feature reduction, classification, and similarity calculation phases. All the input images are enhanced by using the PDPHE algorithm. After that, two levels of features, such as local features and global features, are extracted from the images and the dimensionalities of the features are reduced by using the ESVD method. Then, the reduced features are given as input to the EDLNN classifier, which classifies the images into different classes. The same processes are done for the query image too; and in the classification phase, the query image is tested with the database images to obtain the same class images. For retrieving very similar images, the similarity is calculated in between the obtained class of database image and the query image features by using the RWSM method, and then, the higher score image is retrieved as the output. The experimental evaluation of the proposed method demonstrates that the proposed image retrieval system achieves better results than the existing methods.*

Key words: *Partitioning Double Plateau Histogram Equalization (PDPHE), Enhanced Singular Value Decomposition (ESVD), Enhanced Deep Learning Neural Network (EDLNN), Gaussian Mutated Cuttlefish optimization (GMCO), Renyi Weighted Similarity Measure (RWSM).*

1. INTRODUCTION

The explosive growth of massive digital devices has produced a huge amount of digital images in the last two decades [1]. The majority of these data are images, which are being used in fields like satellite images, medical images, video, and still images repositories [2]. Consequently, retrieving and managing the desired images from large image databases has become the main challenge in computer vision [3]. To index and search suitable images from the rapidly increasing digital image collections, an image retrieval system are used [4]. Indexing and searching are the two major tasks for retrieval [5]. Image retrieval is the task of searching similar images of a certain type from the image datasets [6]. Two types of approaches are used to search an image collection; First one is based on a search by metadata known as Text Base Image Retrieval (TBIR), another one is based on Content information in the image, also known as Content-Base Image Retrieval (CBIR) [7]. A text-based image retrieval system has two issues. Firstly, the annotation task takes a long time, and secondly, assigning keywords for image annotation is subjective. These two drawbacks led to the development of a new system, which is CBIR [8]. CBIR refers to the process of obtaining images that are relevant to a query image from a large collection based on their visual content [9, 10]. The visual contents of an image include color, texture, shape, etc. to denote the catalog of the image [11]. In the present day, many real-world image-based retrieval applications used CBIR strategies. For example, the physician may use the CBIR strategy to retrieve comparative patient issues from a database, so that the best decision can be taken by the physician in the treatment of patients [12]. The CBIR performs two important tasks; the first task is feature extraction and the other is similarity or distance measurement to retrieve the most similar images from the database [13]. Similarity measures are used to quantify the similarity between the image

descriptors [14]. From the view of feature extraction, the optimal feature extraction from the images is the most vital part of the CBIR system, and it is divided into two categories called local and global features [15]. One is based on global features extracted from the entire image, and another is based on local or semantic features extracted from image regions [16]. The success of an image retrieval system depends on the effectiveness of the features that are used to represent the image descriptors [17]. These primitives are derived to represent the images in such a way that could express the query of a human mind properly through the numerical form [18]. In the past few years, a large number of feature extraction techniques have been developed based on the visual contents of an image, such as shape, color, texture [19]. They resulted in producing irrelevant images during the image retrieval process [20], problems in the extraction of a well-suited feature for a system, and poor accuracy. In order to dumb these issues, this research methodology proposes an RWSM similarity calculation measure and global and local features-based CBIR system for retrieving the images. The structure of this paper is divided into four sections: section 2 exhibits the related work, and the proposed framework for the content-based image retrieval system is presented in section 3. The experimental results are given in section 4, and section 5 concludes the paper with future work.

2. RELATED WORK

Sachendra Singh and Shalini Batra [21] presented a Bi-layer Content-Based Image Retrieval (BiCBIR) system, which consisted of two modules: the first module extracted the features of dataset images in terms of color, texture, and shape. The second module consisted of two layers: initially, all images were compared with query image for shape and texture feature space, and indexes of M most similar images to the query image were retrieved. Next, M images retrieved from the previous layer were matched with query image for shape and color feature space, and F images similar to the query image were returned as an output. Experimental results showed that the BiCBIR system outperformed the available state-of-the-art image retrieval systems. The drawback of the method is that the extracted features were not described in the image context, which degraded the performance of the system. Munish Kumar et al. [22] represented a CBIR system that considered two local image feature extraction methods, namely, SIFT and ORB. Scale Invariant Feature Transform (SIFT) was used for detecting features and feature descriptors of an image. Oriented Fast Rotated and BRIEF (ORB) used Features from Accelerated Segment Test (FAST) keypoint detector and binary Binary Robust Independent Elementary Features (BRIEF) descriptor of an image. K-Means clustering algorithm was also used for analyzing the data. Locality Preserving Projection (LPP) was employed to reduce the length of the feature vector. For classification, two classifiers, namely, BayesNet and K-Nearest Neighbours (K-NN) were considered. Wang image datasets were used for experimentation work and accomplished the highest precision rate of 88.9%. The SIFT features are mathematically complicated and do not work well for the blur images, which led to lower performance. R. Rani Saritha et al. [23] employed a multi-feature image retrieval method by combining the features of the color histogram, edge, edge directions, edge histogram, and texture features, etc. The Deep Belief Network (DBN) method of deep learning was used to extract the features and classification. After pre-processing steps like selection removal, its above features were extracted and were stored as small signature files. Those signatures were compared with the content-based signature. In the similarity measure, the distances between the different features were measured. Appropriate weights were applied to normalize the distance coefficients. Those normalized coefficients were sorted and indexed based on the distance values and their optimized state of functioning. The method was tested through simulation in comparison and the results showed a huge positive deviation towards its performance. The method had an insufficient feature set where the DBN required a large number of data for better performance. Mutasem K. Alsmadi [24] demonstrated the extraction of vast robust and important features from the images database and the storage of these features in the repository in the form of feature vectors. The features were extracted from a specific Query Image (QI). Accordingly, an innovative similarity evaluation with a metaheuristic algorithm (genetic algorithm with simulated annealing) was attained between the QI features and those belonging to the database images. For an image entered as QI from a database, the distance metrics were used to search the related images. Furthermore, the results' precision-recall value was calculated to evaluate the system's efficiency. The CBIR system demonstrated better precision and recall values compared to other state-of-the-art CBIR systems. The system may produce inaccurate results at the time of images having more noises. Prashant Srivastava1 & Ashish Khare [25] represented a CBIR technique through multiscale Local Binary Pattern (LBP). Local Binary Pattern of different combinations of eight neighborhood pixels was computed at multiple scales. The final feature vector was constructed through Gray Level Co-occurrence Matrix (GLCM). The large-scale dominant features of some textures were captured. Similarity measurement was done to retrieve visually similar images from a large database. The performance of the proposed technique was tested on five benchmark datasets, namely, Corel1K, Olivia-2688, Corel-5K, Corel-10K, and GHIM-10K, and measured in terms of precision and recall. The experimental results demonstrated that the method outperformed other multiscale LBP techniques as well as some of the other state-of-the-art CBIR methods. The method focussed mainly on extracting LBP features that produced rather long histograms making the retrieval time longer.

3. PROPOSED CONTENT-BASED IMAGE RETRIEVAL SYSTEM

A technique used for automatic retrieval of images in a large database that perfectly matches the query image is called a Content-Based Image Retrieval (CBIR) system. Image querying refers to the predicament of finding objects that are pertinent to a user inquiry within image databases. But the existing research methodologies lack accurate retrieval and take more time to retrieve the image. So, this research methodology proposes an RWSM similarity calculation measure and global and local features. Initially, all the input images are pre-processed. The pre-processing step consists of the resizing and contrast enhancement step. Here, the contrast of the image is enhanced by using the PDPHE. After that, the two-level features are extracted: local features and global features. After feature extraction, the unnecessary dimensionality of the features is reduced by using ESVD. Thus, the features are given as input to the EDLNN, which classifies the classes of the image database. This kind of classification is helpful to reduce the time of retrieval. The above-mentioned processes are done for the query image also. In the classification phase, the query image is tested with the trained output of the database images. After testing, the same class image is predicted as a query image class. After classification, very similar images are retrieved by using the RWSM, which is calculated in between the predicted class of database image and the query image features, and then the higher score image is retrieved as the output. The block diagram for the proposed method is shown in figure 1,

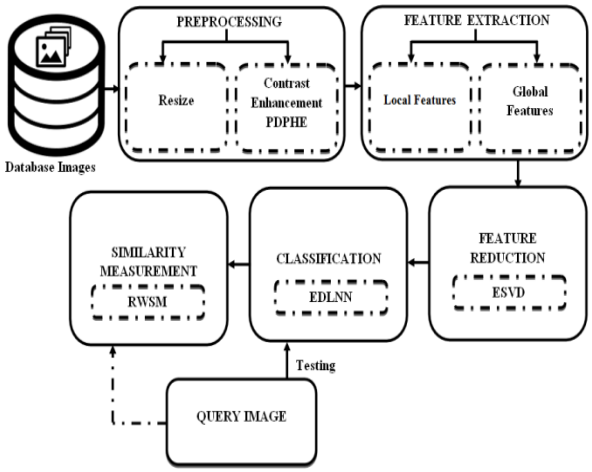


FIGURE 1. Block diagram of the proposed methodology

3.1. Preprocessing: Pre processing is the primary step needed for improving the quality of the image. In this work, preprocessing step includes resizing and contrast enhancement. Resizing is altering the size of an image without cutting anything out. In this work, resizing is handled for the reason of avoiding error during the retrieval process and the resized image is denoted as I . After resizing the contrast of the image is enhanced by PDPHE. Contrast enhancement is a process that makes the image features stand out more clearly. To improve the quality of an image, it is required to perform the operations like contrast enhancement and removal of noise. Here, the Partitioning Double Plateau Histogram Equalization (PDPHE) method is used for the purpose of contrast enhancement. In a normal double plateau histogram algorithm, the over enhancement problem does not arise, but in the upper and low-intensity level adjustment, the noise may be raised. So that this research methodology first partitioned the histogram of the image into four sub-quadrant histograms based on the median value of the original image.

The histogram of the image is,

$$H(I_{\Delta}) = n \tag{1}$$

Where, H is the histogram of input images (I) and n is the number of pixels having intensity level I_{Δ} . $H(I_{\Delta})$ is partitioned into two sub-histograms and the medians from two sub-histograms are used as the separating points which divides the partitioned histogram into two smaller sub-histograms. Thus, the four sub-histograms are obtained as,

$$H_k(I_{\Delta}) = \{H_1(I_{\Delta}), H_2(I_{\Delta}), H_3(I_{\Delta}), H_4(I_{\Delta})\} \tag{2}$$

$$\Delta = \tau\{I_{wd} \times I_{ht}\} \tag{3}$$

Where, Δ is intensity that is set to τ for the total number of pixels in the image histogram, I_{wd} is the width of the image, I_{ht} is the height of the image. The partitioned sub-quadrant histograms are normalized and applied to DPHE. The normalization is performed as,

$$H_k^{nor}(I_\Delta) = \frac{n}{I_{wd} \times I_{ht}} \tag{4}$$

After normalizing the histogram level, the image intensity is adjusted by upper and lower intensity values as follows,

$$I^H = \begin{cases} \lambda_{up} & if\ (H_k^{nor}(I_\Delta) \geq \lambda_{up} \\ H_k^{nor}(I_\Delta) & if\ (\lambda_{down} \leq H_k^{nor}(I_\Delta) \leq \lambda_{up} \\ \lambda_{down} & if\ (0 \leq H_k^{nor}(I_\Delta) \leq \lambda_{down} \\ 0 & if\ (H_k^{nor}(I_\Delta) = 0) \end{cases} \tag{5}$$

Where, λ_{up} is the upper threshold and λ_{down} is the lower threshold, I^H is the histogram after double-plateau histogram equalization.

3.2. Feature Extraction: In this section, two levels of features: the local features and global features are extracted from the preprocessed image I^H . The local features include Color features, Rotated Local Binary Pattern (RLBP), Novel Statistical Features (NSF). The Bag of Visual Word (BoVW) and Speeded Up Robust Feature (SURF) are known as global features.

3.3. Color features: In order to extract the color features, first the model of the color space is determined, and the hue, saturation, and, lightness components of the image are quantized. After quantization, the one-dimensional image $I^{H(1d)}$ is obtained as,

$$I^{H(1d)} = Q_l(h)Q_l(v)Q^h + Q_l(v)Q^v \tag{6}$$

Where, Q_l is the quantization level of h and v, Q^h is the color quantization, Q^v is the saturation quantization. According to the quantization level, the final histogram, which describes the feature vector of the image, is obtained

***3.4RLBP*:** In RLBP, first, the difference between the centre pixel and neighbouring pixels is calculated where the neighbouring is considered as circular. After that, the magnitude of the difference is calculated to determine the dominant direction. Dominant direction is known as the index in the circular neighbourhood, which has the maximum difference. The dominant direction $\vec{D}_{dir}$ is defined as follows,

$$\vec{D}_{dir} = i(|\, G_{np} - G_{cp}\, |_{\max}) \tag{7}$$

Where, i denotes the index in the circular neighbourhood, G_{np}, G_{cp} are the gray values of the centre and neighbourhood pixels of the input image I^H . While an image undergoes a rotation the dominant direction in a neighbourhood also rotates in the same angle. The $\vec{D}_{dir}$ is taken as the reference and the weights for the neighbourhood are assigned with respect to it. Then, the RLBP is represented as,

$$RLBP(r,p) = \sum_{np=0}^{p-1} q(G_{np} - G_{cp}).2^{|np - \vec{D}_{dir}, p|} \tag{8}$$

Where, r is the radius of the circular neighbourhood, p is the number of neighbours. The local binary pattern q can be obtained using,

$$q(G_{np} - G_{cp}) = \begin{cases} 1 & if\ (G_{np} \geq G_{cp)} \\ 0 & if\ (G_{np} < G_{cp)} \end{cases} \tag{9}$$

If the gray value of the neighbouring pixels is greater than or equal to the gray value of central pixel, then 1 is assigned for those pixels, otherwise, the value is assigned as zero.

***3.5NSF*:** The correlation between the adjacent pixels (i, j) is known as statistical features, which are obtained by computing, the mean (μ), standard deviation (σ), skewness (S_q), kurtosis (K_r), coefficient of variance (C_{var}), covariance (C_{cv}), correlation (C_{cor}), and entropy (E_t).

The mean (μ) and standard deviation (σ) can be calculated as,

$$\mu = \frac{\sum_{i,j=0}^{n} I_{ij}}{n} \tag{10}$$

$$\sigma^2 = \frac{\sum_{i,j=0}^{n} (I_{ij} - \mu)^2}{n-1} \tag{11}$$

The skewness (S_q), kurtosis (K_r), and coefficient of variance (C_{var}) are calculated as,

$$S_q = \frac{\sum_{i,j=0}^{n} \frac{(I_{ij} - \mu)^3}{n}}{\sigma^3} \tag{12}$$

$$K_r = \frac{\sum_{i,j=0}^{n} \frac{(I_{ij} - \mu)^4}{n}}{\sigma^4} \tag{13}$$

$$C_{\text{var}} = \frac{\mu}{\sigma} \tag{14}$$

Then, the covariance C_{cv}, correlation (C_{cor}), and entropy (E_t) are measured as follows,

$$C_{cv} = \frac{\sum (I_i - \mu)(I_j - \mu)}{n-1} \tag{15}$$

$$C_{cor} = \sum_{i,j=0}^{n} I_{ij} \frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i)^2 (\sigma_j)^2}} \tag{16}$$

$$E_t = \sum_{i,j=0}^{n} I_{ij}(-\ln(I_{ij})) \tag{17}$$

From these computations, the NSF features set f_{NSF} is obtained.

***3.6 BoVW*:** In BoVW, a large number of local descriptors (or) interest points are extracted from the image to find the words in an image. The key points are detected using SIFT descriptor. The extracted descriptors are quantized into visual words by using the k-means clustering algorithm. The distance between key point descriptors and centre point in clustering is defined as,

$$\text{ƛ}(k_p, c_p) = \sqrt{\sum_{i=1}^{n} (k_p - c_p)^2} \tag{18}$$

Where, k_p is the key point descriptor, c_p is centre point from clustering, $\hat{\lambda}$ denotes the distance between k_p , c_p . The closest visual words around extracted regions are detected from the centre point obtained from clustering. Then, the local features are compared with the visual words to generate the visual-word vocabulary describing different local patterns in images.

***3.7 SURF*:** SURF contains two stages, namely, keypoint detection and key point description. In order to detect the key point, the Hessian matrix corresponding to the pixel position of the image (x, y) and scale Φ is calculated as,

$$H(x, y, \Phi) = \begin{bmatrix} Q_{xx}(x, y, \Phi) & Q_{xy}(x, y, \Phi) \\ Q_{xy}(x, y, \Phi) & Q_{yy}(x, y, \Phi) \end{bmatrix} \tag{19}$$

Where, $Q(x, y, \Phi)$ denotes the second-order Gaussian derivatives along with the vertical (y), horizontal (x), and diagonal (xy) directions. Since the SURF is invariant to rotation, the SURF descriptor is extracted by assigning orientation where Haar wavelet responses are computed to reduce both feature extraction and matching time.
Finally, the extracted features are expressed as,

$$E_f = f_i,\ i = 1,2,....n \tag{20}$$

Where, E_f is the extracted feature set and f_i represents n number of features.

3.8 Feature Reduction: From the extracted features, E_f the unwanted features are reduced by using Enhanced Singular Value Decomposition (ESVD) algorithm. In a normal SVD algorithm, there is no strength analysis for feature dimensions which may produce unnecessary feature dimensions also. So that this research methodology uses the Jacobian matrix for analyzing the strength and gives better results.

In SVD, the data matrix D of size $p \times q$ is constructed using E_f and is decomposed into a product of three matrices as,

$$D = M_u M_d Z^T \tag{21}$$

Where, M_u is the unitary matrix, M_d is the diagonal matrix of singular values m_i , and Z^T is the matrix where each column of the matrix is the eigenvectors. Thereafter, the strongest singular values are identified by using the Jacobian matrix, which analyses the strength for each feature dimension. Hence, the SVD on Jacobian matrix is computed as,

$$J = \frac{ZM_d{}^2}{p-1} Z^T \tag{22}$$

Where, J is the Jacobian matrix, Z is the right singular vectors. In order to reduce the dimensionality of the feature from q to $a\,(a < q)$, first a columns of M_u, and $a \times a$ upper-left part of M_d are selected and their product is obtained. Then, the original data with only the largest a singular values D_a is reconstructed as,

$$D_a = M_{u(a)} M_{d(a)} Z^T_{(a)} J \tag{23}$$

Then, form the dimensionality reduced matrix D_a , the important features are obtained and the feature vector is denoted as, $\overline{E}_f$ which is subjected to further classification.

3.9Classification: The selected feature vector $\overline{E}_f$ is fed to the classification phase. In classification, the database contains more number of output images under different categories. These categories are classified into particular classes, which reduce the retrieval time of the system. For classification, this research methodology uses the

Enhanced Deep Learning Neural Network (EDLNN). A neural network consists of three types of layers of neurons, namely: the Input Layer, the Hidden Layer, and the Output Layer. DLNN, which is an enhanced version of Artificial Neural Network (ANN) has multiple hidden layers between the input and output layers to perform the learning process. In this, the weight value initialization is done by using the Gaussian Mutated Cuttlefish optimization (GMCO) algorithm instead of random weight selection where the random selection results in poor accuracy. Then, in the classifier, the Radial Basis Function (RBF) activation function is used for the reason of many layers presented in the hidden unit. The normal activation function is not well suitable for more number of layer output.

In DLNN, the feature vector $\overline{E}_f$ is inputted to the input layer $Y_{i/p}$ as,

$$Y_{i/p} = \overline{E}_f \tag{24}$$

The feature vector of the input layer is given to the hidden layer, which assigns the weight value for the inputs. Then, the output of the hidden layer is given as,

$$Y_{hid\_o/p} = \eta_m + \sum_{m=1}^{N} Y_{i/p}\psi_m \tag{25}$$

Where, η_m is the activation function, ψ_m is the optimized weight value of input and hidden layer. In DLNN, the weight values are selected randomly, which leads to poor accuracy of the classification. In order to improve the accuracy, the weight values are optimized using the GMCO method, and the step is explained below. The GMCO is inspired by the social behaviour of the Cuttlefish especially, its ability to change its color. The CO algorithm has a limited global search space, which will be enhanced by applying the Gaussian mutation. So, the proposed CO method is termed GMCO. The CO considers two main processes: reflection and visibility, which are used to produce the patterns and colours seen in cuttlefish from different layers of cells including chromatephores, leucophores, and iridophores. The population (i.e., the weight values assigned between the input and hidden layer) is generated randomly as,

$$\overline{\psi_m} = \{\psi_1, \psi_2, \psi_3, ..., \psi_N\} \tag{26}$$

The Gaussian mutation ϑ is performed for increasing the global search area as,

$$\vartheta = \xi_g . \hbar_{\overline{\psi_m}} \tag{27}$$

Where, ξ_g denotes the Gaussian distribution, $\hbar$ is the position information of the weight values in the population.

The new solution of the global search space S^{new} is generated as,

$$S_{new} = \vartheta(U + V) \tag{28}$$

Where, U, V are known as, reflection and visibility and are expressed as follows,

$$U = \alpha * \psi_{m^1}[l].pts[b] \tag{29}$$

$$V = \beta * (B.pts[b] - \psi_{m^1}[l].pts[b]) \tag{30}$$

Where, ψ_{m^1} is the group of cells, $pts[b]$ is the b^{th} point of the l^{th} cell in ψ_{m^1}, $B.pts[b]$ is the best solution points, α, β are the degree of reflection and visibility. Then, α, β are defined as,

$$\alpha = \delta() * (c_1 - c_2) + c_2 \tag{31}$$

$$\beta = \delta() * (e_1 - e_2) + e_2 \tag{32}$$

Where, c_1, c_2, e_1, e_2 are the constant values and $\delta()$ is the random function. After that, the fitness is evaluated based on the accuracy of the classifier to find the difference between the best solution and the current solution. The new solution is generated as,

$$S^{new} = \begin{cases} S^{cur} & if\ (cfn < fn) \\ S^{best} & if\ (bfn < fn) \end{cases} \tag{33}$$

Where, S^{cur}, S^{best} are the current solution and best solution, cfn, bfn are the current fitness and best fitness. The difference between the best solution and the average of the best points is estimated to generate the new search area as,

$$U_b = \alpha * B.pts[b] \tag{34}$$

$$V_b = \beta * B.pts[b] - b_{avg} \tag{35}$$

Where, b_{avg} is the average value of the best points. In the same way as finding the best solution, the optimal weight values are selected using the GMCO method. The pseudo code for the GMCO method is shown in figure 2,

Input: Initialized weight values ψ_m
Output: optimized weight values ψ_m

Begin
Initialize population $\overline{\psi_m}$, reflection U, visibility V, maximum number of iteration $i_{\max}$
Calculate fitness for the population
Set $i = 0$
While $(i \le i_{\max})$
Evaluate Gaussian Mutation
Update U, V
Update new solution by global search
Evaluate fitness of new solution S_{new}
If $(cfn < fn)$
{
Replace $S^{cur} = S^{new}$
}
Else
{
$S^{best} = P^{new}$
}
End if
Update best solution
Calculate average points of the best solution
Calculate fitness of the current best solution
Set $i = i + 1$
End while
Return optimized weight values Ψ_m
End

FIGURE 2. Pseudo code of GMCO method

Since, n number of hidden layer are used in EDLNN, the activation function η_m is not suitable for the hidden layer output. In order to solve this problem, the radial activation function is used and the output is written as,

$$\chi(r) = \sum_{m=1}^{N} \psi_m . \| \overline{E}_f - Y_{hid\_o/p} \| \tag{36}$$

Where, $\overline{E}_f$ is the input vector, $Y_{hid\_o/p}$ is the hidden layer output weighted by the coefficient ψ_m, $\chi(r)$ is the radial activation function. The output of EDLNN contains different classes of database images.

3.10 Similarity calculation:Here, the query image is taken as the input. The same process, such as preprocessing, feature extraction, feature reduction, and classification are repeated for the input query image. Then, similar classes are retrieved by testing the query image with the trained output database images. The most similar images from the similar classes are retrieved by calculating the similarity between the features of database images and the query image. For the similarity calculation, the Renyi Weighted Similarity Measure (RWSM) calculation is used. The weighted similarity matching scheme allocates the different weight values to the different image features based on their quality of information. The Renyi entropy function strongly calculates the weight value of the features and improves the robustness of the system.

In RWSM, the similarity T_{qb} between the feature vector of the query image and database image calculated as,

$$T_{qb} = \sqrt{(\sum_{i=1}^{R}\sum_{t=1}^{O} |(FV_i)_{que} - (FV_t)_{db}|)} \quad (37)$$

Where, $(FV_i)_{que}, (FV_i)_{db}$ are the feature vectors of query image and database images, R, O are the number of feature vectors for query image and database images. Then, the total feature similarity T_{tfs} is calculated as,

$$T_{tfs} = \varpi \times T_{qb} \quad (38)$$

The weight factor ϖ can be calculated using the Renyi entropy function, $\varphi(\varpi)$ which is defined as,

$$\varphi(\varpi) = (1-\theta)^{-1} \log(\sum E_k^{\theta}) \quad (39)$$

Where, θ is the order of the Renyi entropy function, E_k is the corresponding probabilities. Finally, the top x images from the database are retrieved, which have minimum T_{tfs} values.

4. RESULT AND DISCUSSION

In this section, the performance of the proposed model is analyzed. The proposed RWSM based image retrieval system is developed in the working platform of MATLAB.

4.1 Database Description: The experiments conducted in this study employed the Corel dataset (A) comprising of 1000 diverse images, and each image in the dataset is 384 * 256 or 256 * 384 size-wise. As such, the outcomes were stated with the use of 10 semantic sets and each set contains one hundred images. The Corel dataset is of the following semantic categories: Mountains, Africa, Dinosaurs, Buildings, Buses, Food, Elephants, Horses, Beach and Flowers.

4.2 Performance Analysis: This section contains the performance analysis of the proposed method with existing Genetic Algorithm (GA)-Based Classifier Comity Learning (GCCL) [26], weighted average of triangular histograms (WATH) [29], multichannel decoder based local binary pattern (MDLBP) [28], and Fused Information Feature-based Image Retrieval System (FIF-IRS) [29] methods based on precision, recall, f-measure, sensitivity, specificity, accuracy, and retrieval time.

TABLE 1. analysis the performance of the proposed method with the existing algorithms based on precision

Classes	Proposed method	GCCL[26]	WATH[27]	MDLBP[28]	FIF-IRS[29]
Buses	100	100	95.74	95.19	95.87
Mountains	90.63	69.54	85.78	45.65	63.54
Beach	80.12	72.56	79.96	55.82	60.74
Elephants	97.26	82.78	89.54	63.43	95.48
Food	92.12	90.13	84.92	70.97	71.53
Flowers	100	86.76	86.87	93.23	100
Africa	90.34	83.02	77.82	75.64	82.76
Horses	100	82.93	89.41	89.32	100
Dinosaurs	100	97.63	98.12	97.65	100
Building	90.23	86.75	80.75	67.19	67.49

Discussion: table 1 analyze the precision value of the proposed and existing methods. The precision value is obtained for ten classes namely, busses, mountains, beach, elephants, food, flowers, Africa, horses, dinosaurs, and building. For all classes, the proposed method has the values of 100%, 90.63%, 80.12%, 97.26%, 97.12%, 100%, 90.34%, 100%, and 100%. The precision values of the existing methods are lower when compared to the proposed method for all ten classes. The average precision of the proposed method is 94.07%. The existing methods have the value of 85.21% for GCCL, 86.89% for WATH, 75.41% for MDLBP, and 83.74% for FIF-IRS. In the above analysis, the proposed method has higher values for precision than the existing methods. The performance analysis based on precision is graphically represented in figure 3,

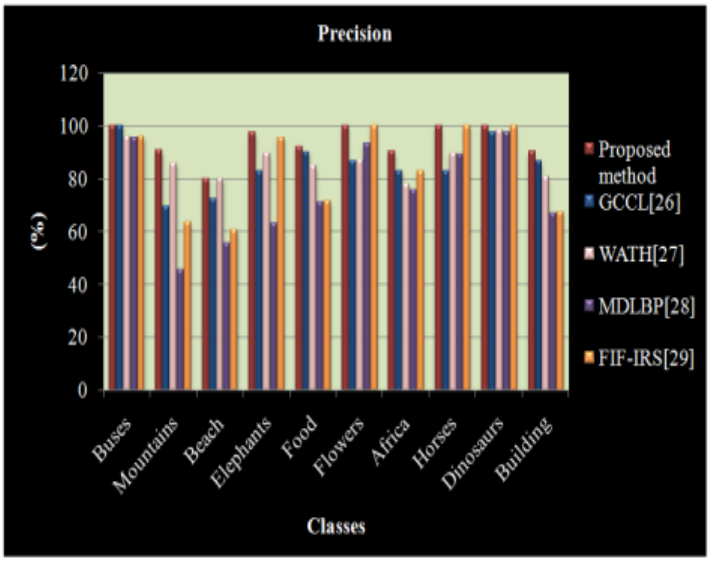


FIGURE 3. Pictorial representation for the performance analysis based on precision

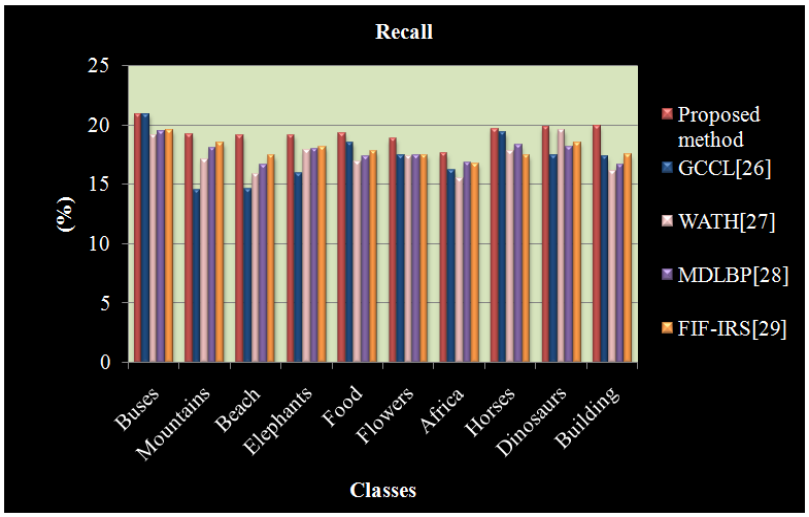


FIGURE 4. Graphical representation of the proposed and existing methods with respect to recall

Discussion*:* in figure 4, the recall values of the proposed and existing methods are analyzed. The proposed method achieves the highest recall value of 20.97% for busses where the existing methods have 20.95% for GCCL, 19.15% for WATH, and 19.53% for MDLBP, 19.61% for FIF-IRS. For the remaining classes, the proposed method has larger values and the existing GCLL method has the lowest value than other methods. The average recall value of the proposed method is 19.39% whereas, for existing GCCL, WATH, MDLBP, and FIF-IRS, the values attained are 17.28%, 17.37%, 17.54% and 17.95%. From the above analysis, it is known that the proposed method has better performance than the existing methods.

Discussion: the analysis for f-measure values of the proposed and existing methods is shown in figure 5. The f-measure value of the proposed method is higher for the classes, Africa, horses, and dinosaurs, which are 97.82%, 94.74%, 93.52%, 94.35%. The existing methods have f-measure values lesser than the proposed method. In the existing methods, the FIF-IRS methods have a medium level of f- measure, and GCCL and WATH methods have the lowest f-measure values than the existing methods. The average f-measure values of the proposed and existing methods are, 91.16%, 86.21%, 84.52%, 83.43%, and 84.42%. The analysis-based f-measure value delivers that the proposed method yields a reasonable performance than the existing methods.

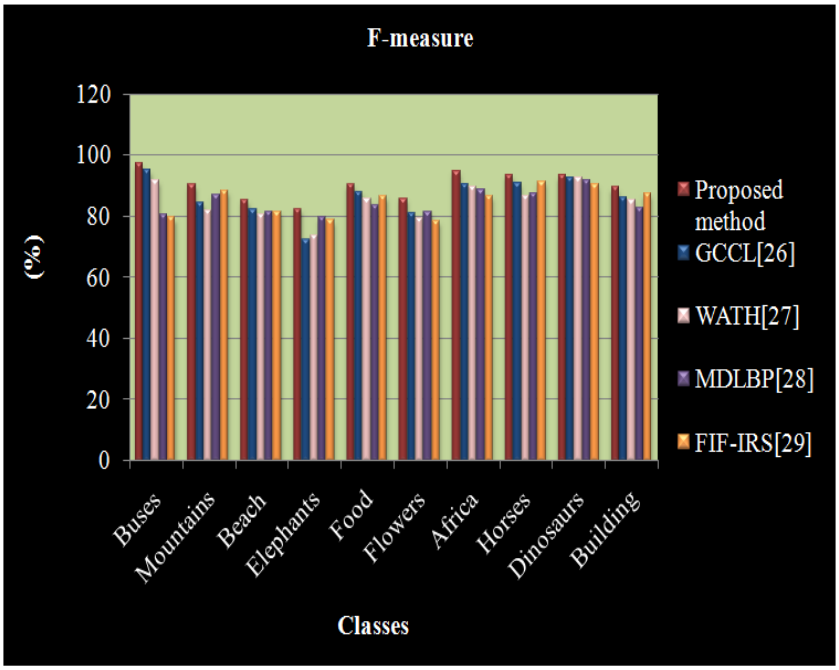


FIGURE 5. Demonstrates the performance analysis of the proposed and existing methods with reference to the F-measure

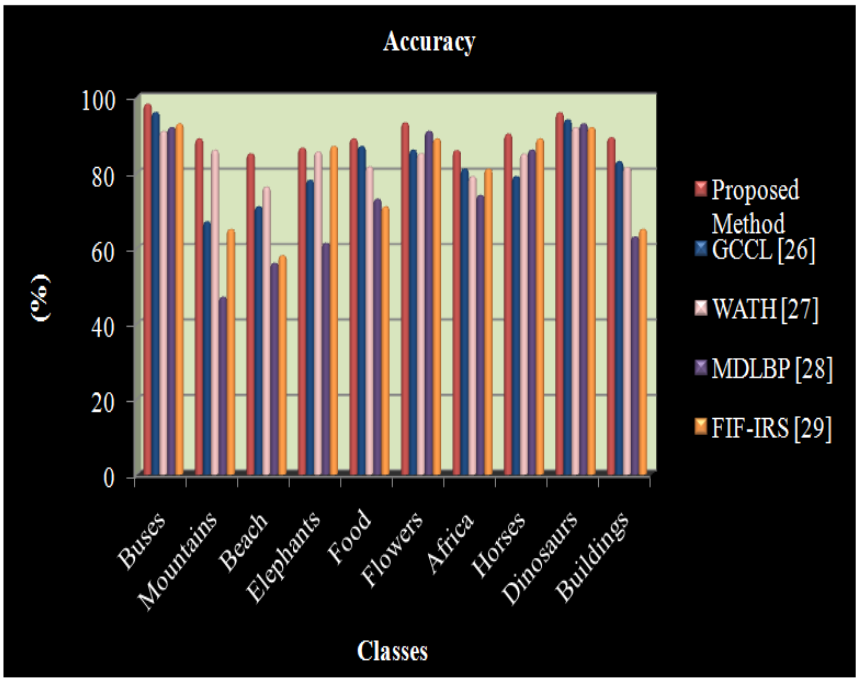


FIGURE 6. Illustrates the performance of the proposed method and existing methods in terms of accuracy

Discussion: the accuracy of the proposed and existing method is analyzed in figure 6. The proposed method has the highest accuracy of 98.23% for the class busses. For the same class, the existing methods have the accuracy of 96%, 91%, 92%, and 93% for GCCL, WATH, MDLBP, and FIF-IRS methods. Same as that the lowest accuracy is attained for the class beach, which is 85% for the proposed method, 71% for GCCL, 76.23% for WATH, 56% for MDLBP, and 58% for FIF-IRS. The average accuracy of the proposed method is 90.238, which is higher than the existing methods. The existing methods have an accuracy of 82.2% for GCCL, 84.28% for WATH, 73.652% for MDLBP, and 79% for FIF-IRS, which are smaller than the proposed method.

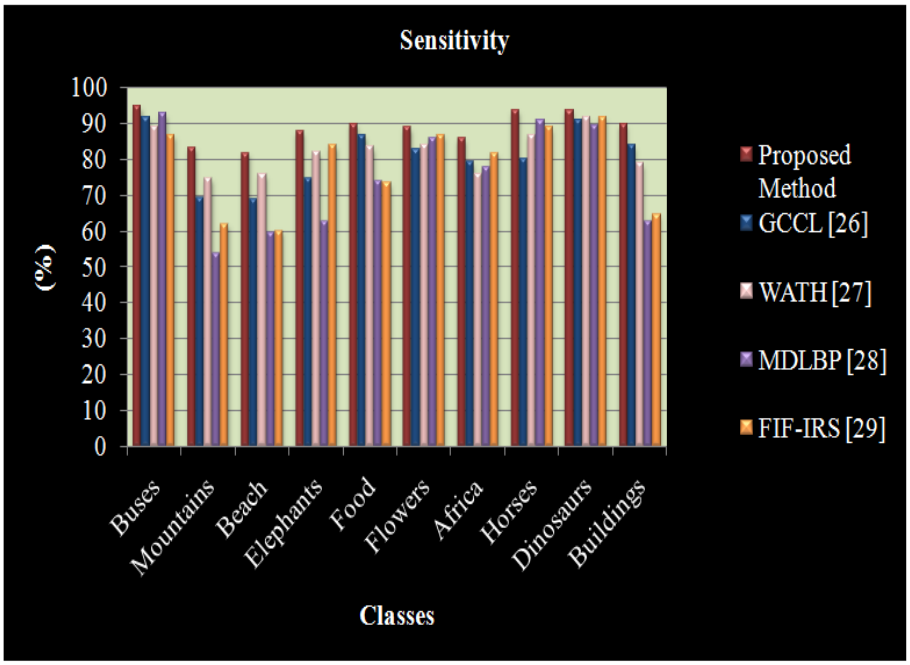


FIGURE 7. Performance analysis with respect to sensitivity

Discussion: figure 7 shows the sensitivity of the proposed and existing methods. For all ten classes, the proposed method has the values of 95%, 83.55%, 82%, 88%, 90.1%, 89%, 86%, 93.85%, 94% and 90% which are larger when compared to the existing methods. The average sensitivity of the proposed method is 89.15%. In existing methods, the GCCL and WATH have a medium level of sensitivity, such as 82.366% and 75.142%. The MDLBP and FIF-IRS methods have a lower sensitivity level of 78.167% and 78.167%. The analysis for sensitivity level shows that the proposed method has better performance than the existing methods.

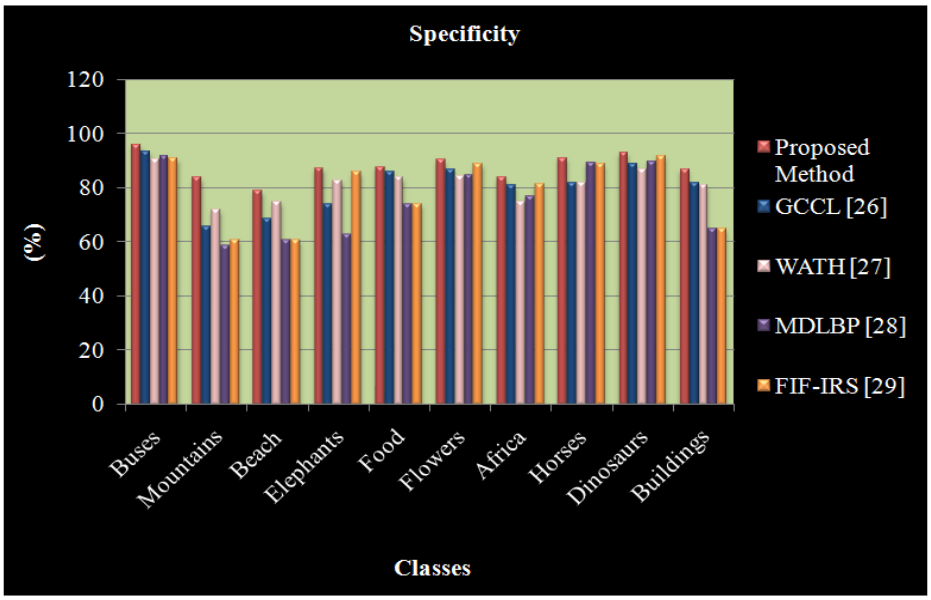


FIGURE 8. Performance analysis of the proposed and existing methods based on specificity

Discussion: the specificity of the proposed and existing methods is analyzed in figure 8. For specificity, the proposed method has the highest value for the class busses, which is 96.25%, and the lowest value of 74% is achieved for the class beach. The existing methods have busses and beach of 93.75% and 69% for GCCL, 90.5% and 75% for WATH, 92.13% and 61% for MDLBP, and 91.02% and 60.95% for FIF-IRS. In terms of specificity, the proposed method attains an average value of 88.035; whereas, the average of the existing methods is 80.995% for GCCL, 81.41% for WATH, 75.578% for MDLBP, and 78.967% for FIF-IRS. The analysis shows that the proposed method is superior to the existing methods.

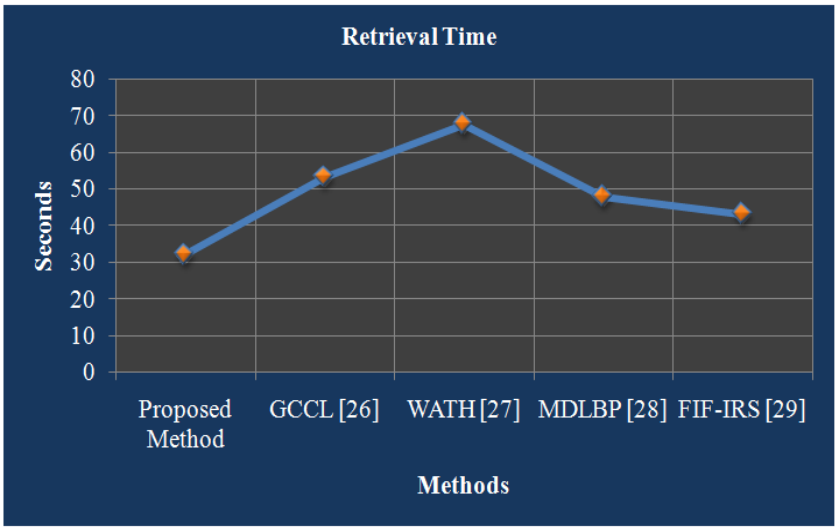


FIGURE 9. Pictorial representation for the retrieval time of proposed and existing methods

Discussion: the retrieval time of the proposed and existing methods is represented in figure 9. The time taken for the process of retrieving the image is known as retrieval time. The retrieval time of the proposed method takes 39.19 seconds for retrieval. The existing GCCL, WATH, MDLBP, and FIF-IRS methods take 53.23 seconds, 67.61 seconds, 47.93 seconds, and 43.15 seconds for retrieval. The overall analysis demonstrates that the proposed method takes lesser time than the existing methods.

5. CONCLUSION

The aim of Content-Based Image Retrieval (CBIR) is to extract similar images of a given image from huge databases by matching a given query image with the images of the database. This research methodology proposed a CBIR system based on RWSM and local and global features. The method contains five phases, namely, preprocessing, feature extraction, feature reduction, classification, and similarity measurement. For analysing the performance of the proposed method, the Corel dataset (A) is taken. The performance of the proposed method is compared with the existing methods, such as GCCL, WATH, MDLBP, and FIF-IRS methods, and the analysis is done based on the performance metrics. In this analysis, the proposed method achieves a better result than the existing methods based on all the performance metrics. The average precision and accuracy rate of the proposed method is 94.07% and 90.238%. The retrieval time of the proposed method is 39.19 seconds, which is lower than the existing methods. In the future, the robust features will be concentrated on making the system more efficient, and advanced algorithms are used for reducing the retrieval time.

REFERENCES

[1]. Salahuddin Unar, Xingyuan Wang, Chunpeng Wang and Yu Wang, “A decisive content based image retrieval approach for feature fusion in visual and textual images”, Knowledge-Based Systems, vol. 179, pp. 8-20, 2019.

[2]. Mukul Majhi, Arup Kumar Pal, SK Hafizul Islam and Muhammad Khurram Khan, “Secure content-based image retrieval using modified Euclidean distance for encrypted features”, Transactions on Emerging Telecommunications Technologies vol. 32, no. 2, pp. 1-35, 2021.

[3]. Saliha Mezzoudj, Ali Behloul, Rachid Seghir and Yassmina Saadna, “A parallel content-based image retrieval system using spark and tachyon frameworks”, Journal of King Saud University-Computer and Information Sciences, vol. 33, no. 2, pp. 141-149, 2019.

[4]. Pavithra L. K and Sree Sharmila T, “An efficient framework for image retrieval using color texture and edge features”, Computers & Electrical Engineering, vol. 70, pp. 580-593, 2018.

[5]. Mansoor Hussain D and Surendran D, “The efficient fast-response content-based image retrieval using spark and MapReduce model framework”, Journal of Ambient Intelligence and Humanized Computing, vol. 12, pp. 4049-4056, 2020.

[6]. Khawaja Tehseen Ahmed, Shahida Ummesafi and Amjad Iqbal, “Content based image retrieval using image features information fusion”, Information Fusion, vol. 51, pp. 76-99, 2019.

[7]. Rehan Ashraf, Mudassar Ahmed, Sohail Jabbar, Shehzad Khalid, Awais Ahmad, Sadia Din and Gwangil Jeon, "Content based image retrieval by using color descriptor and discrete wavelet transform", Journal of Medical Systems, vol. 42, no. 3, pp. 1-12, 2018.
[8]. Muhammad Yousuf, Zahid Mehmood, Hafiz Adnan Habib, Toqeer Mahmood, Tanzila Saba, Amjad Rehman and Muhammad Rashid, "A novel technique based on visual words fusion analysis of sparse features for effective content-based image retrieval", Mathematical Problems in Engineering, 2018, doi.org/10.1155/2018/2134395.
[9]. Maria Tzelepi and Anastasios Tefas, "Deep convolutional learning for content based image retrieval", Neurocomputing, vol. 275, pp. 2467-2478, 2018.
[10]. Samaneh Ghodratnama and Hamid Abrishami Moghaddam, "Content-based image retrieval using feature weighting and C-means clustering in a multi-label classification framework", Pattern Analysis and Applications, vol. 24, no. 1, pp. 1-10, 2021.
[11]. Mary I. Thusnavis Bella and Vasuki A, "An efficient image retrieval framework using fused information feature", Computers & Electrical Engineering, vol. 75, pp. 46-60, 2019.
[12]. Payal Chhabra, Naresh Kumar Garg and Munish Kumar, "Content-based image retrieval system using ORB and SIFT features", Neural Computing and Applications, vol. 32, no. 7, pp. 2725-2733, 2020.
[13]. Mohamed Abd El Aziz, Ahmed A. Ewees and Aboul Ella Hassanien, "Multi-objective whale optimization algorithm for content-based image retrieval", Multimedia Tools and Applications, vol. 77, no. 19, pp. 26135-26172, 2018.
[14]. Meenakshi Garg and Gaurav Dhiman, "A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants", Neural Computing and Applications, vol. 33, pp. 1311-1328, 2021.
[15]. Ruqia Bibi, Zahid Mehmood, Rehan Mehmood Yousaf, Tanzila Saba, Muhammad Sardaraz and Amjad Rehman, "Query-by-visual-search: multimodal framework for content-based image retrieval", Journal of Ambient Intelligence and Humanized Computing, vol. 11, no. 11, pp. 5629-5648, 2020.
[16]. Juxiang Zhou, Xiaodong Liu, Wanquan Liu and Jianhou Gan, "Image retrieval based on effective feature extraction and diffusion process", Multimedia Tools and Applications, vol. 78, no. 5, pp. 6163-6190, 2019.
[17]. Subash Kumar TG and Nagarajan V, "Local curve pattern for content-based image retrieval", Pattern Analysis and Applications, vol. 22, no. 3, pp. 1233-1242, 2019.
[18]. Soumya Prakash Rana, Maitreyee Dey and Patrick Siarry, "Boosting content based image retrieval performance through integration of parametric & nonparametric approaches", Journal of Visual Communication and Image Representation, vol. 58, pp. 205-219, 2019.
[19]. Megha Agarwal, Amit Singhal and Brejesh Lall, "Multi-channel local ternary pattern for content-based image retrieval", Pattern Analysis and Applications, vol. 22, no. 4, pp. 1585-1596, 2019.
[20]. Uzma Sharif, Zahid Mehmood, Toqeer Mahmood, Muhammad Arshad Javid, Amjad Rehman and Tanzila Saba, "Scene analysis and search using local features and support vector machine for effective content-based image retrieval", Artificial Intelligence Review, vol. 52, no. 2, pp. 901-925, 2019.
[21]. Sachendra Singh and Shalini Batra, "An efficient bi-layer content based image retrieval system", Multimedia Tools and Applications, vol. 79, pp. 1-29, 2020.
[22]. Munish Kumar, Payal Chhabra and Naresh Kumar Garg, "An efficient content based image retrieval system using BayesNet and K-NN", Multimedia Tools and Applications, vol. 77, no. 16, pp. 21557-21570, 2018.
[23]. Rani Saritha R, Varghese Paul and Ganesh Kumar P, "Content based image retrieval using deep learning process", Cluster Computing, vol. 22, no. 2, pp. 4187-4200, 2019.Mutasem K Alsmadi, "Content-based image retrieval using color, shape and texture descriptors and features", Arabian Journal for Science and Engineering, vol. 45, no. 4, pp. 3317-3330, 2020.
[24]. Prashant Srivastava and Ashish Khare, "Utilizing multiscale local binary pattern for content-based image retrieval", Multimedia Tools and Applications, vol. 77, no. 10, pp. 12377-12403, 2018.
[25]. Aun Irtaza, Syed M. Adnan, Khawaja Tehseen Ahmed, Arfan Jaffar, Ahmad Khan, Ali Javed and Muhammad Tariq Mahmood, "An ensemble based evolutionary approach to the class imbalance problem with applications in CBIR", Applied Sciences, vol. 8, no. 4, pp. 1-26, 2018.
[26]. Zahid Mehmood, Toqeer Mahmood and Muhammad Arshad Javid, "Content-based image retrieval and semantic automatic image annotation based on the weighted average of triangular histograms using support vector machine", Applied Intelligence, vol. 48, no. 1, pp. 166-181, 2018.
[27]. Shiv Ram Dubey, Satish Kumar Singh and Rajat Kumar Singh, "Multichannel decoded local binary patterns for content-based image retrieval", IEEE Transactions on Image Processing, vol. 25, no. 9, pp. 4018-4032, 2016.
[28]. Mary I. Thusnavis Bella and Vasuki A, "An efficient image retrieval framework using fused information feature", Computers & Electrical Engineering, vol. 75, pp. 46-60,2019.

