



A Novel CBIR-Based System using Texture Fused LBP Variants and GLCM Features

Meenakshi Garg, Manisha Malhotra, Harpal Singh

Abstract: This paper presents Multiple-features extraction and reduction for Content-Based Image Retrieval (CBIR). At first, discrete wavelet transform (DWT) is applied on the R, G and B channels to get multi-level decomposition of the image in which approximation and detailed coefficients are extracted. Approximation coefficients contain the main content of the image while detailed coefficients provide the local noise variation in the image. Further computationally efficient and rotation-invariant Dominant-Rotated Local Binary Pattern called a texture descriptor is applied on all approximation and detailed coefficients. By calculating the descriptor relative to reference in a local neighbor patch, a rotation invariance feature image is obtained. The proposed methodology contains the complete structural information extracted by the Local Binary Patterns and also extract the extra information using the information of magnitude, thus attaining extra discriminative power. Then, by getting the Dominant Rotated Local Binary Pattern image, concept of GLCM has been used to extract the statistical characteristics for the classification of texture images. GLCM directly works with the intensity of the images and also provides the spatial relationship of the pixels in the image by calculating frequency of occurring of similar patterns in different directions which makes it useful for the extraction of texture characteristics. GLCM has been improvised into a generalized co-occurrence matrix that extracts significant spatial properties from the distribution of local maxima. Further Median Robust Extended Local Binary Pattern is extracted out of the approximation and detailed coefficients and also the histogram is calculated out of them. Unlike the traditional LBP method and several variants of the LBP, Median Robust Extended Local Binary Pattern compares the local median of images instead of intensities of the raw images. It is a multiscale LBP-type descriptor that proficiently compares the image medians with novel sampling schemes, capable of capturing both microstructure and macrostructure. Further feature concatenation is applied in which GLCM features of DRLBP and histogram of MRELBP are combined for getting a large feature vector. Further, we applied a mutual information concept to sort out the most differentiable features for all categories of CORAL dataset which are further fed to particle swarm optimization-based feature selector which reduced the number of features that can be used in the classification phase. PSO uses the SVM classifier in evaluating its objective function which is average precision value for the selected features. PSO tries to increase precision value for all the categories and provides a feature vector with a large precision value when classify by SVM.

Further three classifiers are trained and tested named SVM, KNN and decision tree in which SVM gives high accuracy and precision rates of classification. Experimental results show above 94% accuracy and .80 to .90 precision values for most of the categories of CORAL dataset.

Keywords: CBIR, MLBP, DRLBP, PSO, SVM, Feature selection, classification etc.

I. INTRODUCTION

Image Retrieval based on Content (CBIR) is the most significant area of image processing and computer vision. It has been utilized in various fields of application such as medicine, health, cultural heritage, crime prevention, etc. SPIRS [1] and IRMA [2] in medication etc. CBIR is a well-defined technique for the search and retrieval of images in a large data-set based upon its visual contents. Image retrieval is characterized by the local or global characteristics based on visual contents. Global characteristics are defined by the properties of the images, like color, shape, and texture [3]. In CBIR, Color is a commonly used visual feature and mostly explored in the literature. The key objective is that human beings tend to distinguish images mainly by colored lines [4]. Texture also is an important property of surfaces in images and defined by similarity of the visual-patterns which represents the most important information related to the image surface such as bricks, tiles, clouds etc. Such descriptors are also fit for the recovery of medical images. The shape descriptors do not mean that it describes the whole image shape, but it describes the shape of particular region of an image. Forms are frequently utilized for segmentation [5] or contour detection [6]. Methods that are used for the shape descriptors are invariance for translation, rotation, and scaling [7]. Local features of an image have been used effectively for the recognition and classification of object categories and retrieved from a collection of points of interest and regions [8]. Since 1990s, the recovery of images from the data-sets by using the visual contents has been turned into an especially dynamic research topic. However, most studies do not adequately consider the semantic aspect of images, which comprises the fundamental of the semantic gap between the outcomes returned by the system and the perception of the user.

The CBIR model contains three main steps as shown in figure1. Firstly, feature extraction from an image and selection of features are performed. Secondly, the similarity measures are calculated. Finally, indexing of features and recovery are performed. The methods of extracting and selecting the characteristics rely upon the particular region of an image or upon the entire image [9]. Descriptors are often evaluated by using the spatial information texture, shape and colors etc. and global descriptors are utilized for image-retrieval.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Meenakshi Garg*, University Institute of Computing, Chandigarh University, Mohali, Punjab, India. Email: meenagarg82@gmail.com

Manisha Malhotra, University Institute of Computing, Chandigarh University, Mohali, Punjab, India. Email: mmanishamalhotra@gmail.com

Harpal Singh, Department of Electronics and Communication Engineering, CEC Landran, Punjab, India. Email: Harpal.ece@cgce.edu.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

The utilization of the local descriptors increases in the recent years as these are remain constant with similar characteristics, with the difference that the local descriptors are extracted from the image regions instead of the entire image [10].

In addition, local descriptors are extracted from the entire image in which specifically utilized for CBIR are I: Scale Scaling (SIFT) [11], Robust Accelerating Characteristics (SURF) [12], HoG Histograms (HoD) [13], Local-Binary patterns (LBP) [14]. In addition to these DFLs, the variants of LBP's have been proposed as an effective version of the traditional Local Binary Patterns [15].

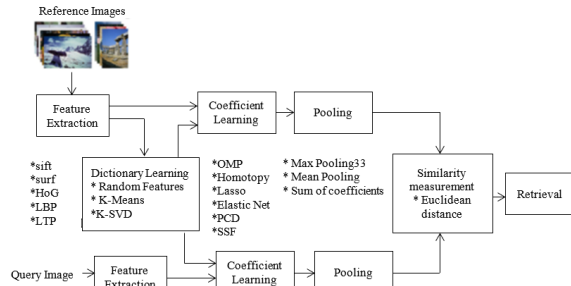


Figure 1: CBIR framework.

The second and the main step of the CBIR models is the measure of similarity. This attempts to decrease the semantic gap [10]. In the CBIR's initial years, the emphasis was placed on the different similarity measures required for each individual characteristic. The similarity measures like Mahalanobis and Euclidean distances are used to calculate the resemblance between the image descriptors. Although, learning of similarities has become popular instead of measuring the similarity recently used [10], [16]. In addition, machine learning is also becoming popular in modern research for CBIR [17] [18]. The machine performance could be improved by taking large datasets for training [19], [20]. In recent years, SAR is typically utilized in the CBIR approaches in second stage. In SAR approach, a dictionary based set of scattered coefficients are represented by the input signals with simple combination.

II. RELATED WORKS

LBP texture features are based on local characteristics for difference in the pixel value of boundary pixels in a window to a centered reference pixel and are created by the weighted combination in the various directions. Meanwhile, the local binary pattern uniform has produced inappropriate patterns for the uneven-edge shapes and textures.

Liao et al. [21], proposed a method of dominant low back pain. The Dominant local binary pattern is very robust for the image rotation and histogram equalization and is more robust against noise than the local binary pattern.

Gue et al. [22], suggested that for the preservation of the global spatial information of images, a local binary pattern variance descriptor is used. Afterwards, they built up more complete local binary pattern descriptor than conventional local binary pattern for the local texture characteristics [23]. On the basis of the combination of the local binary pattern and the Haar wavelets.

Su et al. [24], suggested various characteristics of the Haar local binary model (SLBHP). In SLBHP, the polarity of difference is considered rather than to the calculate the

amplitude difference between the neighboring and the referenced pixel, is taken into account.

Ahonen et al., suggested the histograms which are flexible for the LBP (SLBP) that are noise resistant as compare to the conventional LBP [25].

Lakovid et al. [26], proposed a diffused local binary configuration (FLBP), another form of LBP, where local texture patterns are represented by the fuzzy logic. The Oriented Gradient Histogram (HOG) is used to extract the texture features of the local information that depends upon the local intensity gradient distribution and suits in applications of object detection [27]. Also, appearance of the object and the shape is characterized by the local gradient distribution. At First, images are divided in the numerous regions on the basis of computed gradient. At last, each block's normalized histogram is determined as a final feature vector. In this method, lots of memory is consumed by the feature vector because it is very long.

Chandrasekhar et al. [28], proposed a compressed HOG method that was a low-throughput HOG type. Afterwards, several quantization schemes for compressed HOG are presented [29].

Wang et al. [30], showed that some databases, results in a considerable improvement in the recovery performance by combining the features of LBP and HOG.

Tan et al. [31], showed that the characteristics of LBP were very noise-sensitive in image regions that are uniform and introduced the Local Ternary feature based model.

Gou et al. [32], represents the local derivative model in which deep details are extracted using second-order derivatives or higher. Such LDP's describes the derived address changes for a considered pixel and its adjoined pixels and are not rotationally invariant.

Guo et al. [33] used a Local-Directional Drift Model in which pattern histogram is constructed as characteristics in order to obtain rotation invariant features. Though, the derivative order of the reference pixel and its neighbor pixel is computed without regard to the LDDP address [33]. The characteristics that are extracted from these methods are simply based on the positive and the negative directions of the edges. Further, improvements can be made by differentiating the edges in multi directions to obtain the higher-order yield.

Bakshi et al. [34] showed that by applying the Fourier transformation to local ternary patterns, the accuracy of retrieval of the medical image can be improved because of the strong redundancy in these features.

Hung et al. [35] used the local standard vector in which LVP models are generated by calculating multiple point derivatives in the different directions by combining these derivatives. Comparative Space Transformation, a new algorithm, is introduced to reduce size of feature vectors. By implementing linear-dynamic decision function, the Comparative Space Transformation extracts the functions with more information and also noise resistant.

Generally, LBP, LIP, LDP, LTrP LVP and closest to the gray level pixel and the same degree of the neighbors, they are the difference ways, rest upon the referenced pixel. These patterns are called different things to different combinations of gray pixel in the circle or square. Text can be displayed in many natural system elements intensities in direction,

they have a limited capacity to represent the texture data. Add to that the last encoding referenced pixel elements and three or four above calculations and detrimental to the image information. Therefore, the more the image data retrieved if multilevel encoding is used instead of binary coding.

III. FEATURE EXTRACTION

3.1 DWT

Wavelets are the small waves with changing frequency. Its duration is very limited [36]. These are used in multi-resolution analysis. At multiple resolution all the images are analyzed and represented. Small sized and low contrasted objects are observed at the high resolution and huge are huge and high contrasted objects are analyzed at the coarse level. The idea of multi-resolution analysis is useful when a picture comprises of small and large and also has low and high contrast objects. All images at multi-resolution are useful. Undetected features at one resolution can be identified at other resolution. DWT is simpler to implement the and gives adequate data for the analysis and for the synthesis. The signals are decomposed into the approximation coefficients and in the detailed coefficients in which detailed coefficients are computed in horizontal, diagonal direction and also separately formed the feature vector. This feature vector can be utilized for retrieving the same type of images and can be combined with other for CBIR. At different resolutions the signals with different bands are analyzed and at the next resolution, approximation coefficients are further brake which creates three detailed co-efficient. Features left in last step decomposition are extracted by the coefficients and form the feature vector for retrieval.

B. Dominant Rotated Local Binary Pattern (DRLBP)

It is an effective texture rotation invariant descriptor that is used in feature extraction phase along with median local binary patterns. This property is attained by calculating the descriptor with course to some local reference which is high value of pixel taken as dominant pixel. This Dominant direction is invariant to the monotonic gray value variation [37].

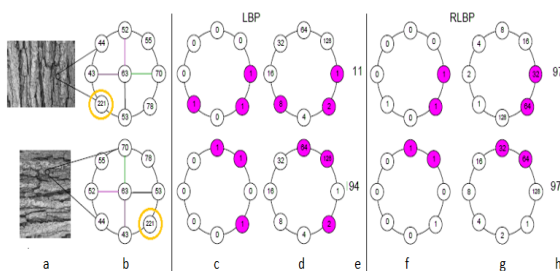


Figure 2: Effect of rotation on the LBP and the RLBP operator [37].

$$Localbinarypattern_{R,P} = \sum_{p=0}^{P-1} S(x_p - x_c) \cdot 2^p, s(x_p - x_c) = \begin{cases} 1 & \text{if } x_p \geq x_c \\ 0 & \text{if } x_p < x_c \end{cases} \quad (1)$$

$$D = \arg \max |x_p - x_c| \\ p \in (0, 1 \dots P-1) \quad (2)$$

$$RotatedLBP_{R,P} = \sum_{p=0}^{P-1} S(x_p - x_c) \cdot 2^{mod(p-D,P)} \quad (3)$$

In equations above, the weighting term depends upon the D. The used mod operator swings the weights in circular fashion by starting from the dominant pixel such that pattern of weights remained same irrespective of the rotation of the image. The change causes rotational invariance because the weights are now dependent on neighborhood and not on a preselected array. In Figure 2, effect of rotation on the characteristics of the RLBP are shown. Red color shows the pixels above the threshold value, and the yellow color shows the pixels corresponding to the dominant direction D. The bit corresponds to the index D always takes the weights that are less than 1, and the other weights are moved from circular way. Figure 2 (g), shows that the weights that are correspond to the dominant position are same for the original images and for the rotated images, while all these pixels are at the different locations. Therefore, in this case, the RLBP values that are obtained for two different rotational neighborhoods are similar, for example, the riLBP parameter produces the same function values for the rotated image and for the host image as rotated and host versions constitutes bit pattern to the same value of the function. If modified-neighborhood has been considered with the value 221 exchanged with the value -pixel 78, and it originates separate pattern and must be represented with a different characteristic value. The riLBP operator also provides the same function value because all the rotated patterns give the same values, irrespective of magnitude of pixel-difference. Therefore, edge information is ignored by the riLBP and takes into account when calculates the RLBP function.

C. Median Robust Extended Local Binary pattern (MRELBP)

The authors in [38] proposed the ELBP method, in which four descriptors similar to LBP are used based on Center-pixel Intensity, Neighborhood-pixel Intensity, Radial Differences and Angular Differences [39].

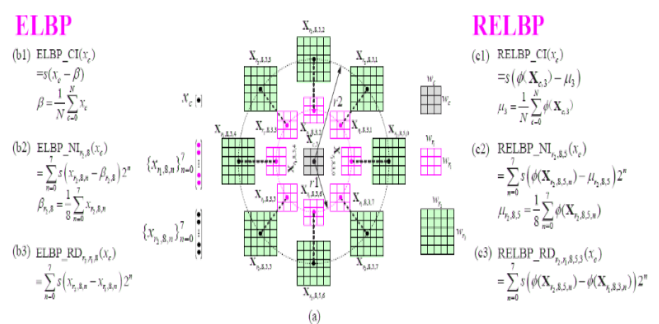


Figure 3: Represents the MRELBP descriptor with r_2 radius and x_c center pixel [38].

Main difference between ELBP method [38] and MRELBP method, In ELBP method, only the values of the single pixel are used, while in MRELBP method, windowed, scaled, and median approach is used [39].

ELBP method [38] is vulnerable to image noises, hence first approach is replacing the intensity values of individual pixel with that of representative over that region which is termed as Median Robust Extended LBP's (MRELBP). In this, raw pixel values are replaced with the median obtained from local-patches.

In modified ELBP descriptor [38], individual pixel intensities at the sampled points are replaced by the median values. Similar to existed ELBP, Images are normalized to zero-mean, unit-variance and standard encoding scheme is used. Joint histograms of MRELBP CI, $MRELBP_NI_{r,p}^{riu2}$ and $MRELBP_RD_{r,p}^{riu2}$ are used to represent the textured image as illustrated in Figure. 2 and in Figure 4.

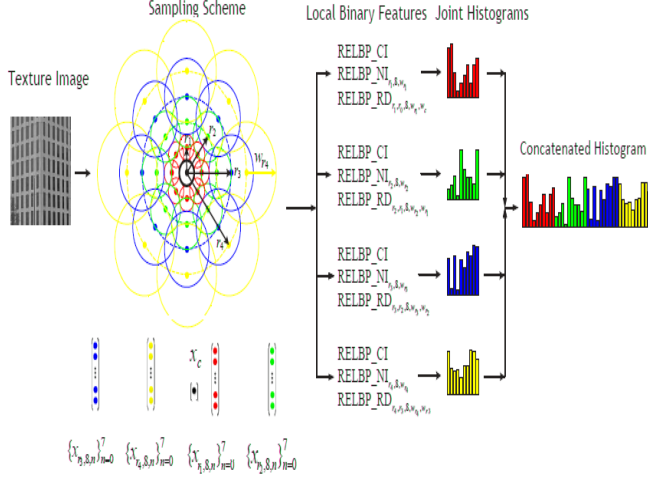


Figure 4: Evaluation procedure of multi-scale MRELBP descriptor and concatenated feature-set [38].

1. MRELBP CI

$$MRELBP_CI(x_c) = s(\phi(X_{c,w}) - \mu_w) \quad (4)$$

where $X_{c,w}$ represents local-patch centered around central pixel x_c considered for feature extraction and the function $\phi(X)$ determine median of local-patch and μ_w is the mean of the image.

2. MRELBP NI

$$MRELBP_NI_{ra,p}(x_c) = \sum_{n=0}^{p-1} s(\phi(X_{ra,p,w_{ra},n}) - \mu_{ra,p,w_{ra}}) 2^n$$

$$\mu_{ra,p,w_{ra}} = \frac{1}{p} \sum_{n=0}^{p-1} \phi(X_{ra,p,w_{ra},n}) \quad (5)$$

where $X_{ra,p,w_{ra},n}$ denotes the patch of size $w_{ra} \times w_{ra}$ centered at pixel $X_{ra,p,n}$

3. MRELBP RD

$$MRELBP_RD_{ra,ra-1,p,w_r,w_r-1} = \sum_{n=0}^{p-1} s(\phi(X_{ra,p,w_r,n}) - \phi(X_{ra-1,p,w_r-1,n}) 2^n) \quad (6)$$

where $X_{ra,p,w_{ra},n}$ denotes the patches around neighbor pixels $X_{ra,p,n}$ arranged on circle around centered pixel x_c having radius ra .

D. Gray level co-occurrence matrix

In this, host image is converted into a matrix which responds to the relation of pixel values on a certain distance. It evaluated the mutual-occurrence of two pixel pair values with particular distances and directions in which horizontal, vertical and diagonal directions are considered [40].

$$G_{dist}^{theta}(i, j) = \{((x, y), (m, n) | I(x, y) = i, I(m, n) = j) \mid (x, y), (m, n) \in N_x \times N_y\} \quad (7)$$

$$(m, n) = (x + dist \times theta_1, y + dist \times theta_2) \quad (8)$$

where G_{dist}^{theta} is the gray-level co-occurrence matrix of distance $dist$ and angle $theta$. $I(x, y)$ and $I(m, n)$ are the pixel intensity at position (m, n) . Values of $theta_1$ and $theta_2$ depend on direction and evaluated using Table 1.

Figure 5 presents an example to calculate GLCM in which first matrix is matrix transformation and second matrix is host image. Consider a pixel pair (2,2) with distance one and theta zero degree marked in red in matrix one finds three times in the host image which sets the GLCM value for this pair as three. Other pairs are also computed in similar fashion

Table 1: Theta one and two correspond to theta in gray-level co-occurrence matrix [40]

Theta (deg)	Theta1	Theta 2
zero	0	1
Forty-five	-1	1
Ninety	-1	0
One-thirty- five	-1	-1

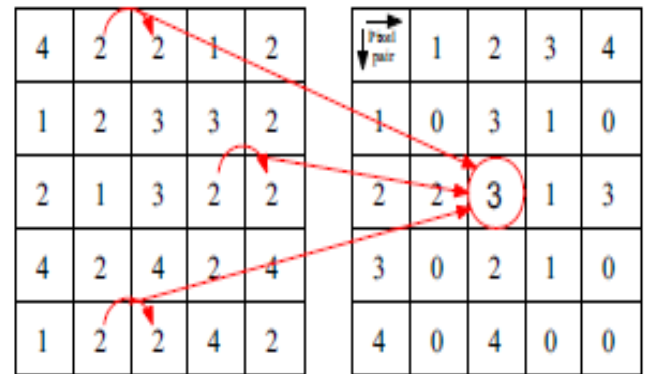


Figure 5: An example GLCM computation [40]

In existed works, usually local binary patterns are considered as histograms for feature extraction and does not includes the patterns with their frequency. By taking GLCM along different directions and distances, an improved feature extraction phase is introduced in which Dominant Rotated Local Binary Pattern (DRLBP) is used. First DRLBP is calculated from multi-scale decomposed detailed and approximation coefficients obtained from red, green and blue components of original image using discrete wavelet transform. Then GLCM pattern map is obtained by considering four direction and two distances. Then different combinations obtained by correlating different distances and directions are explored in order to obtain better feature extraction. The combinations observed are displayed in Figure 6.

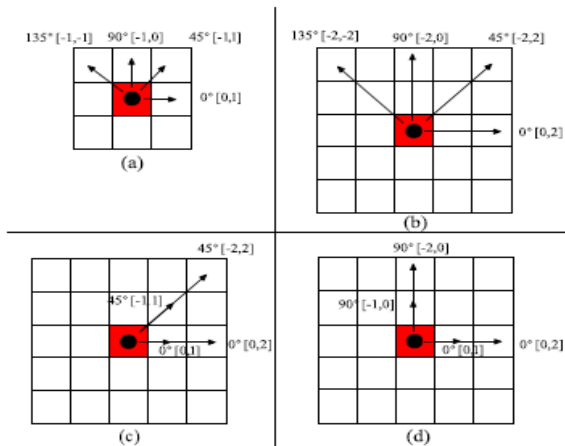


Figure 6: Different combinations obtained using different distance and directional characteristics [40].

1. In Fig. 6(a), GLCM is obtained by taking distance value one and four directions as [0 45 90 135] degree
 2. In Fig. 6(b), GLCM is obtained by taking distance value two and four directions as [0 45 90 135] degree
 3. In Fig. 6(c), GLCM is obtained by taking both distance value one in directions zero and forty-five degree.
 4. In Fig. 6(d), GLCM is obtained by taking both distance value one but in directions zero and ninety degree
- All four matrices have been concatenated to get single feature-vector.

$$Im g = DRLBP(dwtcoefficients)$$

$$GLCM_{dist}^{\theta}(Im g) = G_{dist}^{\theta}(i, j) \forall (i, j) \in Im g \quad (9)$$

where $Im g$ is the DRLBP image-map. Four combinations of GLCM as explained above, have been used to create four different feature vectors.

$$FV_1(I) = [GLCM_1^{0^\circ} GLCM_1^{45^\circ} GLCM_1^{90^\circ} GLCM_1^{135^\circ}] \quad \dots (10)$$

$$FV_2(I) = [GLCM_2^{0^\circ} GLCM_2^{45^\circ} GLCM_2^{90^\circ} GLCM_2^{135^\circ}] \quad (11)$$

$$FV_3(I) = [GLCM_1^{0^\circ} GLCM_1^{45^\circ} GLCM_2^{0^\circ} GLCM_2^{45^\circ}] \quad (12)$$

$$FV_4(I) = [GLCM_1^{0^\circ} GLCM_1^{90^\circ} GLCM_2^{0^\circ} GLCM_2^{90^\circ}] \quad (13)$$

By obtaining GLCM in different directions and distances, more spatial information is obtained present in neighbor pixels which extracts frequency of similar patterns in different directions.

E. Steps in Feature extraction phase

The steps involved in feature extraction phase have been given in this section along the flowchart containing sequence of steps, Figure 7.

Input: Image. Output: Feature vector.

1. Separate Red, Green and Blue components and apply Discrete wavelet transform based multi-scale decomposition.
2. Apply MELBP and DRLBP map of approximation and detailed coefficients at level one.
3. BY taking DRLBP image for all type of coefficients, obtain the GLCM features as below for each R, G and B component
- Evaluate gray level co-occurrence matrixes with distance one and two and along four directions 0, 45, 90 and 135.
- Convert all the matrices obtained above into one dimensional vector.
- Concatenate all vectors into a single feature vector.
4. Concatenate GLCM features for all type of coefficients
5. Obtain Histogram features of DRLBP and MELBP images and concatenate with GLCM concatenated vector

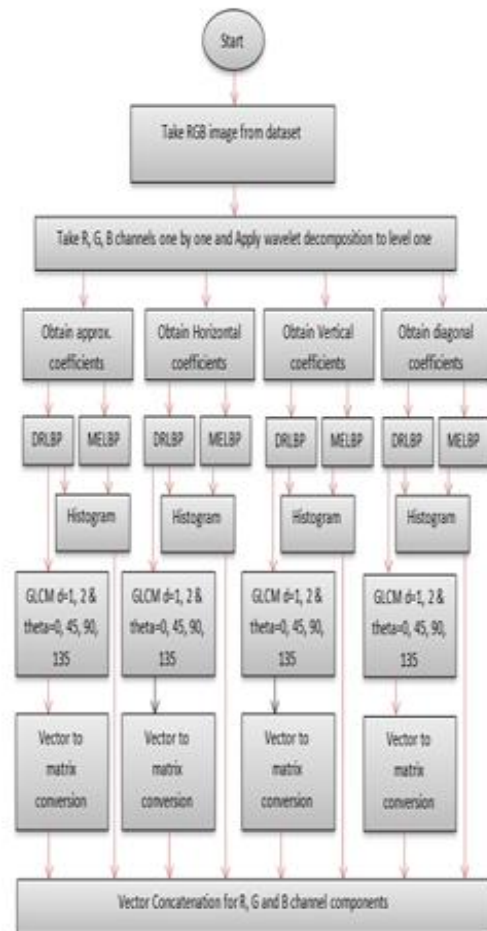


Figure 7: Flowchart of feature extraction phase

IV. FEATURE SELECTION USING PSO (PARTICLE SWARM OPTIMIZATION)

A. Proposed Multi-class Particle Swarm Optimization for feature selection

It is a random optimization process that selects the best features among the un-important features after processing in which desired number of required features are set initially. It randomly selects different combinations of features and an objective function is evaluated from those features which works as a fitness function used to carry out decision making in best feature selection. Initial PSO used binary class features but in this multi-category PSO has been adopted which can work on multi-classes. PSO is a meta-heuristic algorithm, developed by Kennedy and Eberhart in 2001 [41] widely used for distinct, continuous or combinatorial optimization applications. Inspire from flying pattern of a flock of birds, a particle in it denotes a single solution and collection of all is named as swarm. Each particle has its own current velocity and own best configuration p -Best achieved till last iteration and that particle which provides current global best in whole swarm is noted by G -best. In new iteration, each particle sets its own velocity in order to come more close to its p -best as well as G -best of the swarm [42]. The velocity of each particle, v , is updated according to the following equation:

$$velocity_{i,j}(t+1) = w \times velocity_{i,j}(t) + c_p \times r_p \times (p_Best_{i,j} - x_{i,j}(t)) + c_g \times r_g \times (g_Best_j - x_{i,j}(t)) \quad (14)$$

where x the current position of the particle, $velocity_{i,j}$ is velocity of the i^{th} particle dimension j , w is a constant named as momentum that adjusts velocity of current iteration with respect to the previous one, c_p , and c_g are fix-constants initiated in the start. In addition, the ability of the algorithm for exploration /exploitation can adjust by modification of c_p and c_g in new iteration. Updated position of i^{th} particle in dimension j is taken as in equation below:

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1) \quad (15)$$

The flow chart of the PSO procedure is represented in Figure 8.

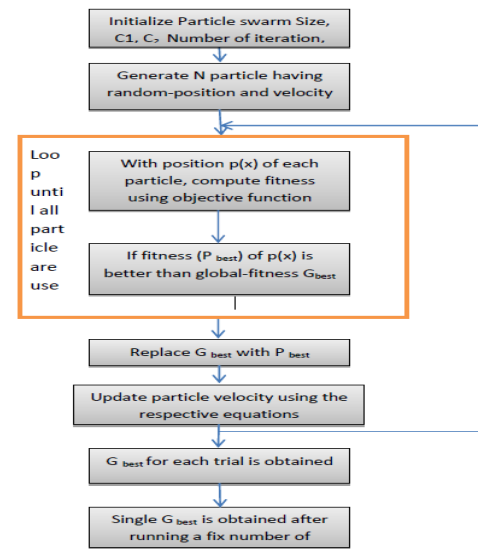


Figure 8:Flowchart of proposed method

In order to select the best particle in each iteration, objective function is required based on which a fitness value is calculated for each particle. In classification-based optimization, mainly area under curve (AUC) is used as objective function which is evaluated for the selected features by PSO depending upon the number of classes or categories in the dataset that need to be classified. Sometimes precision or sensitivity measures are also used and optimum best is figured out from the particles. The objective function applied is assumed in form of eq. as below

$$Objectivefunction = \max i mize(mean(\sum_{i=1}^N AUC))$$

...(16)

Where, N is the Image categories number.

Setting parameters

Weight parameter is not kept constant and updates in new iteration by using previous iterations weight [42]:

$$weight = w_{max} \times (\frac{w_{max} - w_{min}}{It_{max}} it) \quad (17)$$

Where, w_{max} is set as initial value of weighting-coefficient, w_{min} is final value of the weighting-coefficient, it_{max} is maximum no. of iterations and 'it' is current iteration.

V. FEATURE TRAINING STEP FOR CBIR

The final stage of text categorization process involves the training and testing the dataset using the selected feature set given by the feature selection phase using Adaptive particle swarm optimization. Three machine learning classifiers named as DT, SVM, and k-NN are used for classification. As there is need in fitness function evaluation in PSO, decision tree is used to get the area under curve in which the chosen class was marked as true class whereas rest are marked as false. A Similar procedure is adopted for rest of the classes, after which AUC average is used as the objective function and the highest value is selected for the local best.

A. Support vector machine (SVM)

It is a machine learning method based on the learning theory, (Vapnik-Chervonenkis dimension theory) and the principle of structural-risk minimization. It can effectively deal with the regression (time series analysis), pattern recognition (classification problem, discriminant analysis) and many other problems. SVM is popular in the field of prediction and comprehensive evaluation. It works well to solve the small sampling problems, non-linear problems, and the high-dimensional pattern recognition problems. It also has many unique advantages, overcoming by the “curse of dimensionality”, “over-learning” and other problems.

B. K-nearest neighbor (KNN)

KNN is one of the common examples of learning-based techniques. In this technique, K is the number of measured neighbors that is often odd, and the distance to such neighbors is known based on the Euclidean distance values. The fundamental supposition in this technique is that, each sample are real points in the space of n dimensions.

In general, this procedure is applied for two reasons: to approximate the distribution density of the training data set used and to categorize the tested data set depending upon the training of the dataset [43].

C. Decision tree (DT)

The decision tree (Quinlan 1986) [44] is one of the best known machine learning systems and its main objective is to estimate the objective tasks with discrete values. This tree is called a decision tree, since it represents the decision-making method for regulate an input sample group. The decision tree could be a good selection method for opinion mining. It has a very decent performance in contradiction with high-volume data.

VI. RESULTS AND DISCUSSION

Proposed method has been evaluated using a general-purpose WANG data-set which contains 1000 Corel-images having ten different categories of subjects. It is available in size 384×256 or 256×384 in JPEG format. In this dataset, Table 2, 100 images in each of 10 different categories like Africa, Beaches Buses, Horses, Buildings, Dinosaurs, Flowers, Elephants, Food and Mountains are used. Mostly this data-set is used for testing many CBIR systems. It is used commonly because of the good size and the availability of class information used in the data-set.

Table 2: Table showing percentage of samples classified accurately to class

Class	Number of images
Africa	100
Beaches	100
Building	100
Bus	100
Dinosaur	100
Elephants	100

Flowers	100
Food	100
Horses	100
Mountains	100

The classification of view could be assessed utilizing four indexes considered on the base of the subsequent calculations: precision, accuracy, f-measure and recall [45]. Many documents noticed and the most of the categories provides the effectiveness of the presented method based on sensitivity of particular class and specificity of rest of the documents. Hence Detection Accuracy (DA) or recall value which defines sensitivity parameter can effectively represent the accuracy of text classifier. Along with recall, precision, F-measure and Accuracy parameters are also used. The formulas

$$\text{Precision} = \frac{\text{True_Positive}}{(\text{True_Positive} + \text{False_positive})} \quad (18)$$

$$\text{Recall} = \frac{\text{True_positive}}{(\text{True_positive} + \text{False_negative})} \quad (19)$$

$$\text{F_Measurement} = \frac{2 * \text{Precision} * \text{Recall}}{(\text{Precision} + \text{Recall})} \quad (20)$$

$$\text{Accuracy} = \frac{\text{True_Positive} + \text{True_Negative}}{(\text{True_Positive} + \text{True_Negative} + \text{False_positive} + \text{True_Negative})} \quad (21)$$

Where true_positive is suitably categorized subjects in a category, False_positive as appropriately classified subjects in a category and false_negative as correctly classified documents of rest categories; Results for CBIR on Corel Dataset are shown in tables and figures below.

Table 3: Performance evaluation of CBIR using Decision Tree classifier

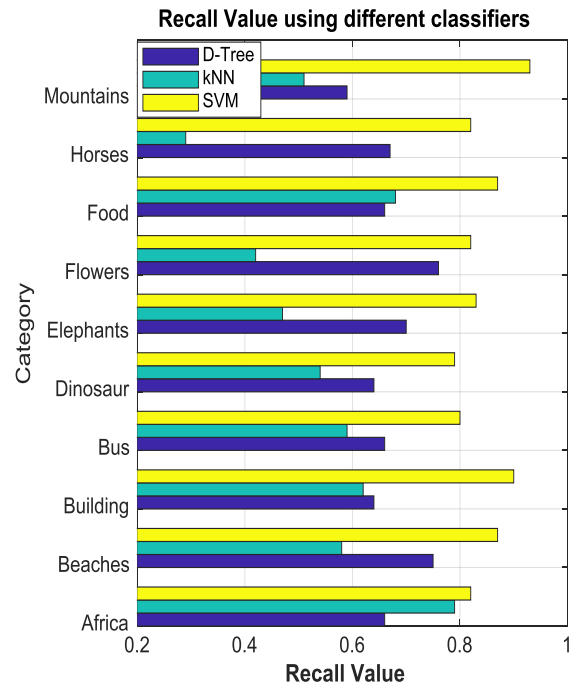
Parameters	TP	TN	FP	FN	Precision	Recall	F-Measure	Accuracy
Class	Classification Results Using Decision Tree							
Africa	66	867	33	34	0.666	0.66	0.663	0.933
Beaches	75	841	59	25	0.559	0.75	0.641	0.916
Building	64	863	37	36	0.633	0.64	0.636	0.927
Bus	66	884	16	34	0.804	0.66	0.725	0.950
Dinosaur	64	878	22	36	0.744	0.64	0.688	0.942
Elephants	70	867	33	30	0.679	0.70	0.689	0.937
Flowers	76	852	48	24	0.612	0.760	0.678	0.928
Food	66	887	13	34	0.835	0.660	0.737	0.953
Horses	67	852	48	33	0.582	0.670	0.623	0.919
Mountains	59	882	18	41	0.766	0.590	0.666	0.941

Table 4: Performance evaluation of CBIR using k-Nearest Neighbor

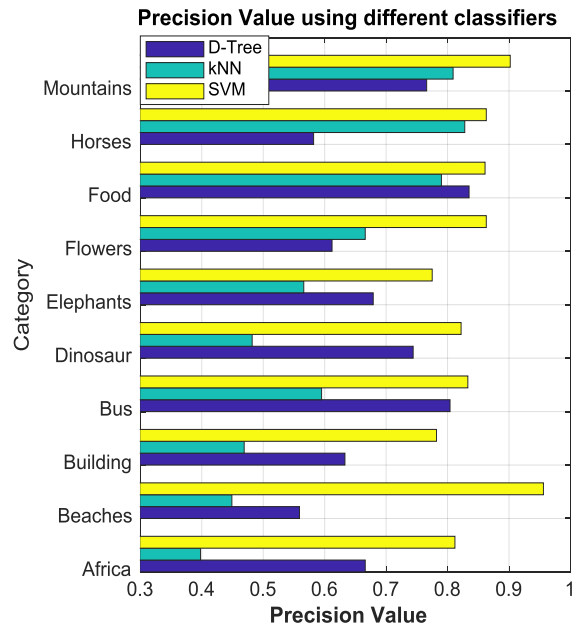
Parameters	TP	TN	FP	FN	Precision	Recall	F-Measure	Accuracy
Class	Classification Results Using k-Nearest Neighbor							
Africa	79	781	119	21	0.398	0.79	0.530	0.86
Beaches	58	829	71	42	0.449	0.58	0.506	0.887
Building	62	830	70	38	0.469	0.62	0.534	0.892
Bus	59	860	40	41	0.595	0.59	0.592	0.919
Dinosaur	54	842	58	46	0.482	0.54	0.509	0.896
Elephants	47	864	36	53	0.566	0.47	0.513	0.911
Flowers	42	879	21	58	0.666	0.42	0.515	0.921
Food	68	882	18	32	0.790	0.68	0.731	0.95
Horses	29	894	6	71	0.828	0.29	0.429	0.923
Mountains	51	888	12	49	0.809	0.51	0.625	0.939

Table 5: Performance evaluation of CBIR using Support Vector Machine

Parameters	TP	TN	FP	FN	Precision	Recall	F-Measure	Accuracy
Class	Classification Results Using SVM							
Africa	82	881	19	18	0.812	0.82	0.815	0.963
Beaches	87	896	4	13	0.956	0.87	0.910	0.983
Building	90	875	25	10	0.782	0.90	0.837	0.965
Bus	80	884	16	20	0.833	0.80	0.816	0.964
Dinosaur	79	883	17	21	0.822	0.79	0.806	0.962
Elephants	83	876	24	17	0.775	0.83	0.801	0.959
Flowers	82	887	13	18	0.863	0.82	0.841	0.969
Food	87	886	14	13	0.861	0.87	0.865	0.973
Horses	82	887	13	18	0.863	0.82	0.841	0.969
Mountains	93	890	10	7	0.902	0.93	0.916	0.983


Figure 9: Bar graphs for Recall Value for CBIR on Core Dataset

Proportion of appropriately expected positive observations to all observations in definite class true is identified as recall. SVM gives high recall ratio for all ten categories in coral dataset which falls in approx. 0.8 and 0.93 in which Beaches, Building, Food and mountain category has 0.87, 0.90, .87 and .93 recall value. Table 3, Figure 9, Rest of the categories has approximately .80 recall value when fitness function in PSO is based on SVM classification. Decision tree and kNN classifiers has significantly low value of recall parameter as compared to SVM classifier.


Figure 10: Bar graphs for Precision Value for CBIR on Core Dataset

Proportion of expected positive observations to the whole positive observation is identified as precision.

Precision value falls in b/w 0.7898 and .96 for SVM, 0.40 to 0.81 for kNN and 0.559 to 0.835 for decision tree, Table 4, Figure 10, Overall SVM gives efficient precision value for almost all categories in which Beaches, flowers, horses and mountain has high measured values of precision noted as 0.956, .863, .863 and 0.902 respectively.

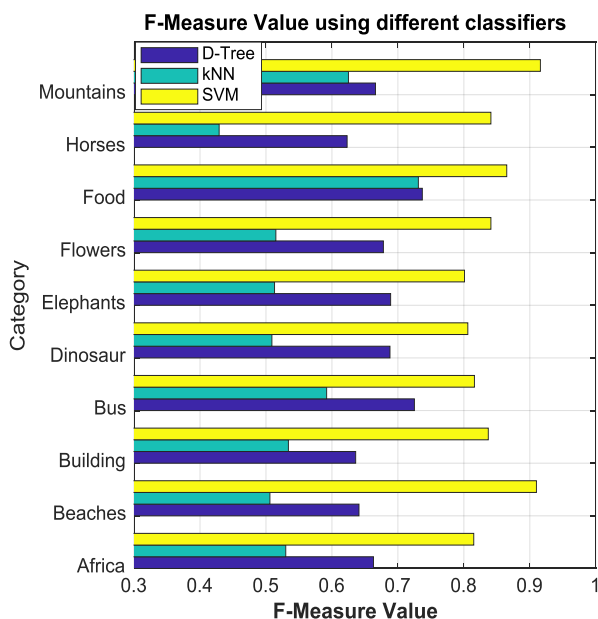


Figure 11: Bar graphs for F-Measure Value for CBIR on Corel Dataset

Weighted average of precision and recall is called f-score. It is additional significant limit than accuracy once having an uneven class distribution in data. SVM classifier has high value of F-score for all tested categories in 0.801 and 0.906. Bus, Dinosaur and Elephants category have F-score value close to 0.80 whereas beaches and mountain classes shows 0.916 F-score value. Table 5, Figure 11. Decision tree shows moderate results for F-score which falls from 0.623 and 0.737 whereas kNN shows poor classification results when F-score parameter is considered.

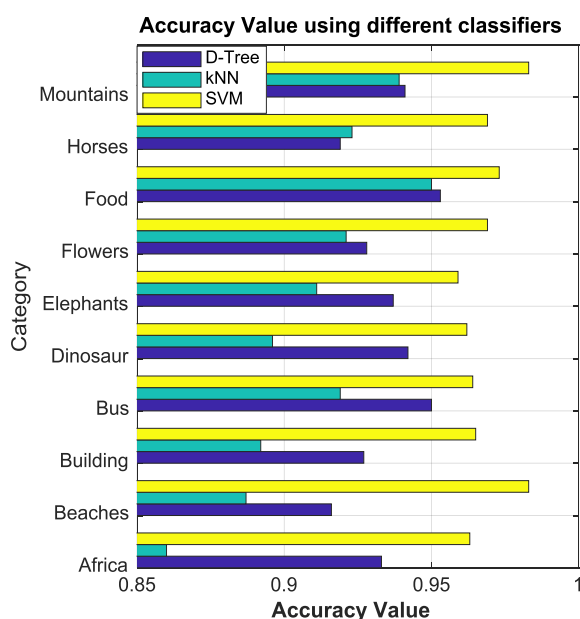


Figure 12: Bar graphs for Accuracy value for CBIR on Corel Dataset

This is the proportion of true negative plus true positives to the true negatives plus true positives plus false negative plus false positive. It computes how much measurement of cases is appropriately classified. SVM gives highest accuracy among all three tested classifiers which is 0.963, 0.983, 0.965, 0.964, 0.962, 0.959, 0.969, 0.973, 0.969 and 0.983 respectively for the images from the data-set are Africa, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers, Food, Horses and Mountains category. Figure 12, It has been found that SVM gives best performance among all three classifiers when fitness function in PSO based feature reduction is chosen in maximizing average of AUC for all categories using the SVM classification method.

VII. CONCLUSION

In this work, multi-feature extraction has been carried out from which most differentiating features are extracted using PSO optimizer. Classification results are carried out on Corel dataset which consists of ten categories and are shown using four presentation parameters i.e. precision, accuracy, f-score and recall in which three classification methods Support vector machine, Decision Tree and k-nearest neighbor are used for classification. Proposed method constitutes four steps i) Decomposition in which multi scale decomposition is carried out using discrete wavelet transform for R, G and B channels separately ii) feature extraction in which DRLBP texture image is generated and GLCM is obtained along with MLBP texture features of detailed and approximation coefficients. Concatenation is carried out from this feature set for all three R, and B channels iii) feature reduction in which PSO algorithm is used to extract the most differentiating features iv) classification in which three classifiers are used for evaluating the category of tested images. PSO uses SVM classifier in evaluating its objective function which is maximizing the average AUC for the selected features. PSO tries to increase AUC value for all the categories and provides feature vector with large AUC value when classify by SVM. Experimental results show that SVM is the best which shows high values of recall, F-score and accuracy parameter.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. Hsu W, Rodney Long L, Antani SK (2007) SPIRS: A framework for content-based image retrieval from large biomedical databases. (MEDINFO)- Proceedings of the 12th World Congress on Health (Medical) Informatics - Building Sustainable Health Systems, pp 188–192
2. Deserno TM, Guld MO, Plodowski B, Spitzer K, Wein BB, Schubert H, Ney H, Seidl T (2008) Extended query refinement for medical image retrieval. Journal Digital Imaging pp 280–289
3. Schettini R, Ciocca G, Gagliardi I (2009) Feature extraction for content-based image retrieval. Encyclopedia of Database Systems, pp 1115–1119
4. Yue J, Li Z, Lu L, Fu Z (2011) Content-based image retrieval using color and texture fused features. Mathematical and Computer Modelling, pp 1121–1127
5. Wang Y, Gong M, Wang T, Cohen-Or D, Zhang H, Chen B (2013) Projective analysis for 3D shape segmentation. ACM Transactions, pp 192–192

6. Mandal S, Sudarshan VP, Nagaraj Y, Dean-Ben XL, Razansky D (2015) Multiscale edge detection and parametric shape modeling for boundary delineation in optoacoustic images. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp 707–710
7. Zhang D, Lu G (2001) Content-Based Shape Retrieval Using Different Shape Descriptors: A Comparative Study. Proceedings of the IEEE International Conference on Multimedia and Expo
8. Anh ND, Bao PT, Nam BN, Hoang NH (2010) A New CBIR System Using SIFT Combined with Neural Network and Graph-Based Segmentation. Intelligent Information and Database Systems, Second International Conference, ACIIDS, pp 294–301
9. T. Dharani, I.L. Aroquiaraj, A survey on content-based image retrieval, in: Pattern Recognition, Informatics and Mobile Engineering (PRIME), 2013 International Conference on, 2013, pp. 4 85–4 90
10. W. Huang, Y. Gao, K. Chan, A review of region-based image retrieval, J. Signal Process. Syst. 59 (2) (2010) 143–161
11. D. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110
12. H. Bay, A. Ess, T. Tuytelaars, L.V. Gool, Speeded-up robust features (SURF), Computer Vision Image Understanding 110 (3) (2008) 346–359
13. N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, 2005, pp. 886–893
14. [14] T. Ojala, M. Pietikainen, D. Harwood, Performance evaluation of texture measures with classification based on kullback discrimination of distributions, in: Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision and Image Processing., Proceedings of the 12th IAPR International Conference on, volume 1, 1994, pp. 582–585, doi: 10.1109/ICPR.1994.576366.
15. X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, Image Process., IEEE Trans. 19 (6) (2010) 1635–1650
16. H. Chang, D.-Y. Yeung, Locally linear metric adaptation with application to semi-supervised clustering and image retrieval, Pattern Recognit. 39 (7) (2006) 1253–1264.
17. D. Zhang, M.M. Islam, G. Lu, A review on automatic image annotation techniques, Pattern Recognit. 45 (1) (2012) 346–362.
18. Y. Gao, K.L. Chan, W.-Y. Yau, Learning in content based image retrieval - a brief review, in: Information, Communications Signal Processing, 2007 6th International Conference on, 2007, pp. 1–5
19. S.R. Buluş, M. Rabbi, M. Pelillo, Content-based image retrieval with relevance feedback using random walks, Pattern Recognit. 44 (9) (2011) 2109–2122. <http://dx.doi.org/10.1016/j.patcog.2011.03.016> . Computer Analysis of Images and Patterns.
20. M.K. Kundu, M. Chowdhury, S.R. Buluş, A graph-based relevance feedback mechanism in content-based image retrieval, Knowl. Based Syst. 73 (2015) 254–264. <http://dx.doi.org/10.1016/j.knsys.2014.10.009> .
21. S. Liao, M.W. Law, A.C. Chung, Dominant local binary patterns for texture classification, IEEE Trans. Image Process. 18 (5) (2009) 1107–1118.
22. Z. Guo, L. Zhang, D. Zhang, Rotation invariant texture classification using LBP variance (LBPV) with global matching, Pattern Recognit. 43 (3) (2010a) 706–719
23. Z. Guo, L. Zhang, D. Zhang, A completed modeling of local binary pattern operator for texture classification, IEEE Trans. Image Process. 19 (6) (2010b) 1657–1663.
24. S.-Z. Su, S.-Y. Chen, S.-Z. Li, S.-A. Li, D.-J. Duh, Structured local binary haar pattern for pixel-based graphics retrieval, IET Electron. Lett. 46 (14) (2010) 996–998
25. T. Ahonen, M. Pietikinen, Soft histograms for local binary patterns, in: Proceedings of the Finnish signal processing symposium, (FINSIG 2007), 5, 2007, pp. 1–4.
26. D.K. Iakovidis, E.G. Keramidas, D. Maroulis, Fuzzy Local Binary Patterns for Ultrasound Texture Characterization, in: International Conference Image Analysis and Recognition, 2008, pp. 750–759. Berlin
27. N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, 2005 IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit. 1 (2005) 886–893.
28. V. Chandrasekhar, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, B. Girod, CHog: Compressed histogram of gradients a low bit-rate feature descriptor, in: IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 2504–2511.
29. V. Chandrasekhar, Y. Reznik, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, B. Girod, Quantization Schemes for Low Bitrate Compressed Histogram of Gradients Descriptors, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, 2010, pp. 33–40.
30. X. Wang, T.X. Han, S. Yan, An HOG-LBP Human Detector with Partial Occlusion Handling, in: 2009 IEEE 12th International Conference on Computer Vision, 2009, pp. 32–39.
31. X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process. 19 (6) (2010) 1635–1650.
32. B. Zhang, Y. Gao, S. Zhao, J. Liu, Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor, IEEE Trans. Image Process. 19 (2) (2010) 533–544.
33. Z. Guo, Q. Li, J. You, D. Zhang, W. Liu, Local directional derivative pattern for rotation invariant texture classification, Neural Comput. Appl. 21 (8) (2012) 1893–1904.
34. A. Oberoi, V. Bakshi, R. Sharma, M. Singh, A framework for medical image retrieval using local tetra patterns, Int. J. Eng. Technol. 5 (1) (2013) 27–36.
35. K.-C. Fan, T.-Y. Hung, A novel local pattern descriptor-local vector pattern in high-order derivative space for face recognition, IEEE Trans. Image Process. 23 (7) (2014) 2877–2891.
36. Prashant Srivastava, Ashish Khare, "Integration of wavelet transform, Local Binary Patterns and moments for content-based image retrieval" Journal of Visual Communication and Image Representation, Volume 42, 2017, Pages 78–103
37. Rakesh Mehta, Karen Egiazarian, "Dominant Rotated Local Binary Patterns (DRLBP) for texture classification" Pattern Recognition Letters, Volume 71, 2016, Pages 16–22
38. L. Liu, L. Zhao, Y. Long, G. Kuang, and P. Fieguth, "Extended local binary patterns for texture classification," Image and Vision Computing, vol. 30, pp. 86–99, 2012.
39. L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang and M. Pietikainen, "Median Robust Extended Local Binary Pattern for Texture Classification," in IEEE Transactions on Image Processing, vol. 25, no. 3, pp. 1368–1381, March 2016.
40. Manisha Verma, Balasubramanian Raman, "Center symmetric local binary co-occurrence pattern for texture, face and bio-medical image retrieval" Journal of Visual Communication and Image Representation, Volume 32, 2015, Pages 224–236
41. J. Kennedy, R.C. Eberhart, Swarm Intelligence, Academic Press, San Diego, CA, 2001.
42. Francisco Erivaldo Fernandes Junior, Gary G. Yen, "Particle swarm optimization of deep neural networks architectures for image classification" Swarm and Evolutionary Computation, Volume 49, 2019, Pages 62–74
43. Alfaro C, Cano-Montero J, Gómez J, Moguerza JM, Ortega F (2016) A multi-stage method for content classification and opinion mining on weblog comments. Ann Oper Res 236(1):197–213
44. Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106
45. Zhang, H., Gan, W., Jiang, B.: Machine learning and lexicon based methods for sentiment classification: a survey. In: Proceedings of IEEE Conference on Web Information System and Application (WISA), pp. 262–265 (2014)

AUTHORS PROFILE



Meenakshi Garg, received her M.Tech. from MDU, Rohtak. She is expert in image processing and optimization fields. She has published research papers in various reputed journals. Currently, she is working as an Assistant Professor in Govt. Bikram College of Commerce, Patiala.



Manisha Malhotra, received her Ph.D. from Maharishi Markandeshwar University, Mullana. She is expert in image processing, deep learning, and optimization fields. She has published more than 20 research papers in various indexed journals. Currently, she is working as an Associate Professor in Chandigarh University, Gharuan, Chandigarh.





Harpal Singh received his Ph.D. from Punjabi University, Patiala. He is expert in image processing, deep learning, and optimization fields. He has published more than 10 research papers in various indexed journals. Currently, he is working as Professor in Chandigarh Engineering College, Landran, Chandigarh.