



An Intelligent Multi-Resolutional and Rotational Invariant Texture Descriptor for Image Retrieval Systems

Shefali Dhingra^a and Poonam Bansal^b

^aUniversity School of Information, Communication and Technology, Guru Gobind Singh Indraprastha University, New Delhi, India

^bMaharaja Surajmal Institute of Technology, Guru Gobind Singh Indraprastha University, New Delhi, India
munjalshefali@gmail.com, pbansal89@gmail.com

KEYWORD

Content Based Image Retrieval (CBIR); Local Binary Pattern (LBP); Discrete Wavelet Transform (DWT); Support vector machine; Extreme learning machine

ABSTRACT

To find out the identical or comparable images from the large rotated databases with higher retrieval accuracy is the challenging task in Content based Image Retrieval systems (CBIR). Considering this problem, an intelligent and efficient technique is proposed for texture based images. In this method, firstly a new joint feature vector is created which inherits the properties of Local binary pattern (LBP) which has steadiness regarding changes in illumination and rotation and discrete wavelet transform (DWT) which is multi-resolutional and multi-oriented along with higher directionality. Secondly, after the creation of hybrid feature vector, to increase the accuracy of the system, classifiers are employed on the combination of LBP and DWT. The performance of two machine learning classifiers is proposed here which are Support Vector Machine (SVM) and Extreme learning machine (ELM). Both proposed methods P1 (LBP+DWT+SVM) and P2 (LBP+DWT+ELM) are tested on rotated Brodatz dataset consisting of 1456 texture images and MIT VisTex dataset of 640 images. In both experiments the results of both the proposed methods are much better than simple combination of DWT +LBP and much other state of art methods in terms of precision and accuracy when different number of images is retrieved. But the results obtained by ELM algorithm shows some more improvement than SVM. Such as when top 25 images are retrieved then in case of Brodatz database the precision is up to 94% and for MIT VisTex database its value is up to 96% with ELM classifier which is very much superior to other existing texture retrieval methods.

1. Introduction

With the advent of enormous digital devices such as digital cameras, mobiles, storage devices and internet there is the tremendous extension in digital image libraries. So the immense efforts have been done over the last few decades for developing a system through which indexing, annotating and searching of images can be done from stupendous databases. Traditionally this work was done by text based image retrieval systems in which the query image was represented by the textual keywords (Srivastava & Khare, 2017). This technique had proved



to be very cumbersome for very large databases such as medical where it was very difficult to memorize the keywords (Kumar, Aggarwal, Tiwari, & Singh, 2018). So to prevail over this difficulty an efficient and time saving system is required and the resolution to this was Content Based Image Retrieval (CBIR) system. CBIR systems retrieve the image automatically on the basis of low level features of the images. For retrieval of images, the CBIR systems utilizes the visual contents of images such as color, shape, texture and spatial layout.

The first and foremost step of the CBIR system is the feature extraction and the overall efficiency of the system depends upon the methods which are selected for extracting the features of the images. The generally used features are color, texture and shape (Zhao & Grosky, 2000). The method adopted depends on various factors and conditions such as illumination changes, view angle etc. Also the perception of images varies from person to person so it becomes very difficult to find out the best matching images from the huge databases. So to blown-away this problem the CBIR system finds out the visual similarity between the query and database images and rather than finding the one best image it provides the list of matching images on the basis of their ranking of similarity matching.

After the proper selection of feature extraction techniques the other main concern with these systems is of accuracy and time of retrieval. Hence the use of an intelligent classifier which is able to find out the proper class label for the query image is required for making a successful CBIR system (Puviarasan, Bhavani, & Vasanthi, 2014). Distance measurement is the next step of this system and after the classification of database images into different classes the similarity calculation of the feature vectors of query image is calculated from the images of the best matching class only. The basic block diagram of CBIR system is shown in Figure 1.

In this era of machine learning various types of classifiers are there which are used in the field of image processing. Mainly used classifiers which are able for multi-class classifications are SVM, Nearest neighbor and some advanced neural network based classifiers such as single layer perceptron, multilayer, ELM, Convolutional neural network and many more (Kastrati & Imran, 2019).

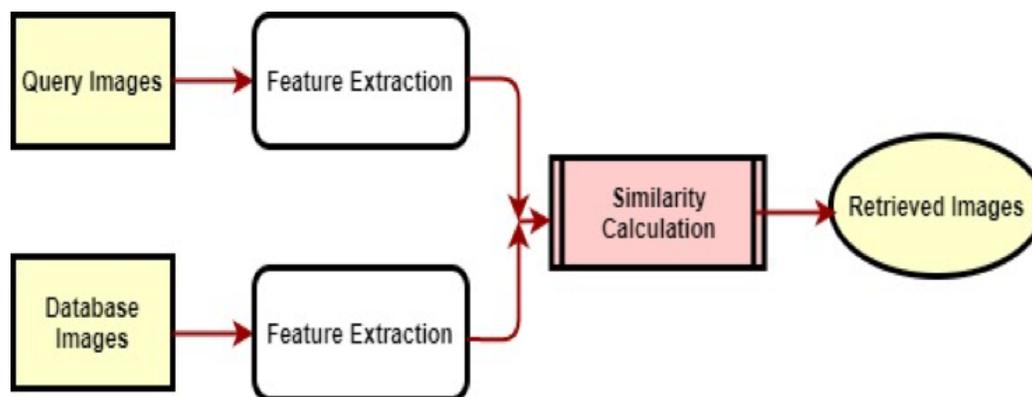


Figure 1: Block diagram of CBIR system.

The main objective of this work is to upgrade the performance of texture feature based image retrieval systems by making an intelligent and novel hybrid CBIR system which is the combination of two well known texture extraction techniques i.e. Local Binary Pattern (LBP) and Discrete Wavelet Transform along with classifier. Texture features are very dominant and decisive for CBIR systems. The chosen texture descriptors are capable of representing the texture properly and also have the property of compensating variations such as rotation, scale changes because these kinds of variations can distort the original appearance of the images. Over the last few years the intensive research is carried out in characterizing the texture pattern as it is very challenging and demanding task in the area of image processing. The LBP texture descriptor has build up very much attention due to its less computational complexity and high selectivity ability (Ricardo, Joaci, & Sá, 2017). Due to these properties of LBP technique it has been successfully employed in different applications like image analysis, pattern recognition and in CBIR systems. The second texture predictor which is used here is discrete wavelet

transform (Hemachandran, Paul, & Singha, 2012). These functions produce the optimized resolution in both frequency and spatial domains so this wavelet transform extracts the local features due to multi-orientation and multi-resolutional properties (Wang, Feng, Zhang, & Liu, 2016). Two machine learning algorithms are used here as a classifier in this proposed work first is simple supervised one and other is neural based which are SVM and ELM respectively.

The remaining part of the paper is systematically organized is as follows: In section 2 a brief survey related to this work is presented. Section 3 provides the detailed view on LBP and DWT, SVM and ELM. The following section i.e. section 4 describes in detail the proposed framework. Experimental set up and results are shown in section 5. Finally in the last section 6, the concluding summary and future work that can be explored is presented.

2. Related Work

Texture is one of the most influential low level features of an image. It provides the spatial arrangement of the visual patterns presents in the images. Various CBIR systems are proposed in literature based on different texture techniques such as DWT, Gabor Transform, Grey Level Co-occurrence Matrix (GLCM), LBP etc. Adaptive tetrotlet transform texture based retrieval technique was proposed which provides very good texture information. The best combination of tetrominoes was elected which provides the better geometry of the image at all level (Raghuwanshi & Tyagi, 2015). The new technique was proposed entitled Local tetra patterns for image indexing and image retrieval for CBIR systems in the different way(Murala, Maheshwari, & Balasubramanian, 2012).

In order to decrease the information level of the images the new technique was proposed known as Local Derivative Radial Pattern (LDRP) (Fadaei, Amirfattahi, & Ahmadzadeh, 2017). It uses multi-level coding as the replacement of binary coding in different directions which decreases the loss of information along with increase in somewhat accuracy. Tamura feature extraction technique was used in (Sreena & George, 2013) which uses the fuzzy humming distance to measure the similarity between the images. Andre Ricardo Backes *et al.* designed a method by using LBP maps technique which extracts the new and more information from the query image. After that it computes the descriptors with the help of fractal dimension (Ricardo *et al.*, 2017). In (Jain & Vailaya, 1995) the authors described a hybrid CBIR system based on texture and color. Wavelet decomposition technique was used to extract the texture feature and the characterization of texture classes with the variations in higher frequency coefficients.

Haralick recommended GLCM which gives more spatial information regarding the texture analysis by determining the co-relation between the neighbor pixels (Haralick & Shanmugam, 1973). Murala *et al.* proposed Local Tetra pattern (LTP) in which second order derivatives are calculated in both horizontal and vertical directions (Murala *et al.*, 2012). Texture descriptor based on combination of LTP and GLCM was developed known as CoALTP by integrating the feature vectors of both(Naghashi, 2018). Image retrieval by texture was described in (Kokare, Biswas, & Chatterji, 2005) in which both the dual tree complex and rotated wavelets were used which makes this method rotation invariant in twelve directions. Energy and standard deviations were used as texture features. Another color and texture based retrieval system had been proposed by employing local extreme features which were able to capture the important information from the images (Pham, 2017). The multi-resolution and rotation invariant technique based on texture retrieval was proposed in (Das, Dash, & Mukhopadhyay, 2013) and this invariance was achieved through aligning the principal texture direction with the reference axis.

LBP was proposed to extract the local properties of every pixel with its neighboring pixels and basically used to classify the texture (Prakasa, 2016). In this binary pattern technique, the magnitude of the centre pixel is compared with the neighbor pixels and finally a decimal value is created by which histogram is generated with pixel values of LBP. Despite of all these advantages LBP does not take into consideration the directional information (R, Murala, & Wu, 2014). In order to extract the edge information in the four directions Directional Local Extrema Pattern was proposed (Maheshwari & Balasubramanian, 2012). The gabor functions are very much applicable for the analysis of texture as its orientation and frequency representation resembles with the human visual systems and thus having high retrieval accuracy or performance (Hemachandran *et al.*, 2012). Di

huang *et al.* presented the survey regarding the LBP technique about its applications and analysis in the domain of image processing (D. Huang, Member, Shan, & Ardabilian, 2011).

SVM based hybrid CBIR system was designed in which color, texture and edge features are extracted from Corel database. To make the system more accurate and speedy multiclass SVM is applied (Puviarasan *et al.*, 2014). In (Ding & Xu, 2013) the concise overview about ELM and its variants such as evolutionary ELM, ordinal ELM, and symmetrical ELM are presented along with their pros and cons. The different applications of ELM such as classification, pattern recognition, regression are also discussed in detail. Chorowski *et al.* examined the performance of four machine learning classifiers which are SVM, ELM, Least Square SVM and Margin Loss ELM. It was also analyzed that how the important ideas of these classifiers can be used when merged together. The accuracies of these all classifiers are tested on different standard datasets also (Chorowski, Wang, & Zurada, 2014)

Due to the uniqueness and simplicity of ELM it has become a very leading research area in different fields. Keeping this in mind an advanced form of ELM was proposed in (Guo *et al.*, 2016) known as Ensemble of ELM.

This new ELM architecture helps in the integration of ensemble and representation learning with easy implementation and training. In(Kastrati & Imran, 2019) both parametric and nonparametric machine learning algorithms were analyzed and their performance was tested on improved concept vector space model. The experiments were carried out on various models for representation of document and seven classifiers are compared and it was observed that the proper selection of the classifier plays a major role for increasing the performance. Shenglan Liu *et al.* designed a CBIR system with relevance feedback using ELM to increase the accuracy of the system. Various relevant feedback techniques were also described in this paper (Liu, Wang, Wu, & Feng, 2015).

3. Main Contributions

A unique and effective CBIR system is proposed in this paper based on the combined approach of LBP and DWT along with two machine learning classifiers i.e. SVM and ELM. This complete approach is applied on two standard datasets i.e. Brodatz and MIT VisTex and the results obtained outperformed the other state of art methods. The LBP, DWT, SVM and ELM are described below in this section.

3.1. Local binary pattern

The LBP texture descriptor is extensively used currently due to its simplicity, dynamic and rotational invariant properties. The designed intelligent system uses LBP technique for the extraction of the texture features from the images. In this technique the pre-processing step is done in which the RGB space of the image is transformed to grey scale. After that the image is subdivided into smaller sub matrices of size 3x3 from which the feature extraction take place. All the features obtained by sub matrices are integrated to form a single feature histogram which represents the whole image (Pavithra & Sharmila, 2017).

Its working is based on the difference between the center value of the pixel and the neighbor pixel. The binary code for every pixel is produced by the step of thresholding of neighbor pixel with the centre pixel as given in Equation (1) and Equation (2) which is shown in Figure 2.

$$LBP_N = \sum_{i=0}^{N-1} f(P_i - CP) 2^i; \quad (1)$$

$$f(p) = \begin{cases} 1; & P \geq 0 \\ 0; & P < 0 \end{cases}, \quad N \text{ is total neighbouring pixels.} \quad (2)$$

$$Hist_k = \sum_{x=1}^M \sum_{y=1}^M f_2(LBP(x, y), k) \quad (3)$$

In this figure first 3 x 3 sub block is taken from the size 5 x 5. The thresholding value i.e. the centre pixel value is 34 for other 8 pixels and the code produced for this is 92. LBP's are calculated for the complete image and after that the histogram is framed which has the capability to describe the complete texture of the image.

Histogram of every bin is computed by summation of the image pixel numbers as shown in Equation (3). LBP (x, y) is the value of x's row and y's column pixel position of the image. While calculating the code for an image using LBP, the unwanted information of the edges is ignored so as to reduce the false information (Liao, Law, & Chung, 2009).

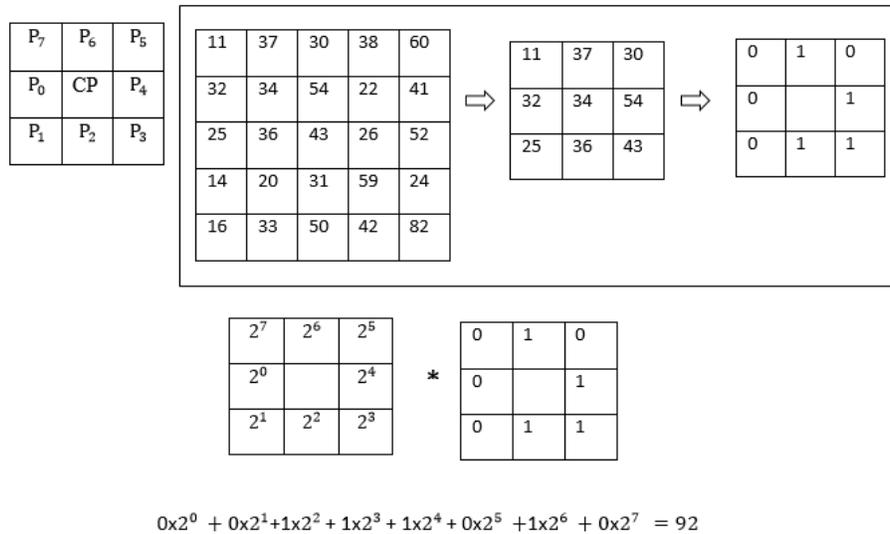


Figure 2: Calculation of Local Binary Pattern for every pixel.

3.2. Discrete Wavelet Transform

Discrete Wavelet transform is superior to the fourier transform as it is capable to analyze the components of a non stationary signal. It allows the decomposition of complicated and complex information regarding, patterns, images, speech etc to elementary forms at different scales and positions and then reconstruct the signal with high accuracy. These transforms are stand on small waves known as wavelets. These transforms can represent the images at different resolutions depending on the chosen frequencies (Alaei, Alaei, Pal, & Blumenstein, 2018). Due to this ability some important features which are generally ignored in different resolutions are taken into consideration.

For extracting the texture features from DWT the coefficient distribution of the mother wavelet is calculated. This wavelet $\psi(t)$ when translated by b and scaled by a is given in Equation 4.

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi\left(\frac{t-b}{a}\right) \quad (4)$$

DWT decomposes the image using orthogonal basis function into the set known as wavelet coefficients. These sets provide the details of four parts such as horizontal, diagonal, vertical and approximation and therefore it can provide the fine details of the image (Karthikeyan & Manikandaprabhu, 2014). The low frequency components are the approximation part while other three are the components of high frequencies. With the application of this technique an image is divided into four sub-bands and one sub-sampled band which is clearly shown in Figure 3. In this Figure LHA, HHA, HLA describes the finest details of coefficients of the wavelet while the coefficients of the coarse level are represented by sub-band LLA which means an approximate idea of an image. The sub-band LLA is further divided for getting the other level of wavelet coefficients. This is also shown in Figure 3. The process remains in continuation until the final scale is obtained. These both values of approximation and of detailed images are very useful and important for the extraction of the texture features and provide a high value of precision when image is taken at different resolutions.

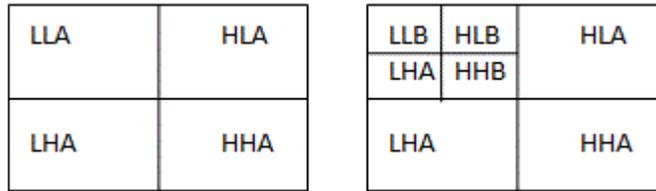


Figure 3: DWT showing its sub-band.

3.3. Support Vector Machine (SVM)

Support vector machines are the famous machine learning algorithm and works on the principle of supervised learning process. It is basically used for the classification and regression of data. These machines were initially designed for the binary classification ($N=2$) only but due to its effectiveness and various applications for multi-classification in various fields such as pattern recognition image processing etc multi-class SVM was also developed ($N>2$) by the integration of various simple binary SVM's (Chorowski *et al.*, 2014).

Decision hyper planes are constructed which predicts decision boundaries for the input space or for the high dimensionality feature space. These decision hyper planes are able to divide the positive and negative samples. At the farthest distance from the hyper plane a separator which is linear is constructed to the nearest positive and negative samples. During the training phase The SVM takes the input in the form of a matrix and labels each and every sample as positive or negative or in other words it makes two classes of the training data. This fully trained SVM is then able to classify the testing data to the belonging class (Chamasemani, 2011).

3.4. Extreme Learning Machine (ELM)

In this era of Artificial intelligence the feed forward neural networks have been extensively used in many areas due to their capabilities in approximation of the complex nonlinear mapping directly from the inputs. Moreover, these models are able to overcome the problems which are proved to be harder for classical parametric techniques. But the main issue with these networks is their slower learning speed; it may take hours to days for its training. All the important parameters in these networks are iteratively tuned by the learning algorithms which are slow gradient based (G. Huang, Member, Zhou, Ding, & Zhang, 2012).

So, to overthrow this problem a new Extreme learning algorithm is developed which is a feed forward neural network with single hidden layer as shown in Figure 4. This algorithm differs from other conventional ones such as back propagations and many other neural networks as these has to face many problems during tuning the controlling parameters manually. ELM is completely implemented automatically with least intervention of the users. Its learning speed is comparatively very much fast than other algorithms (Lu, Duan, & Wang, n.d.).

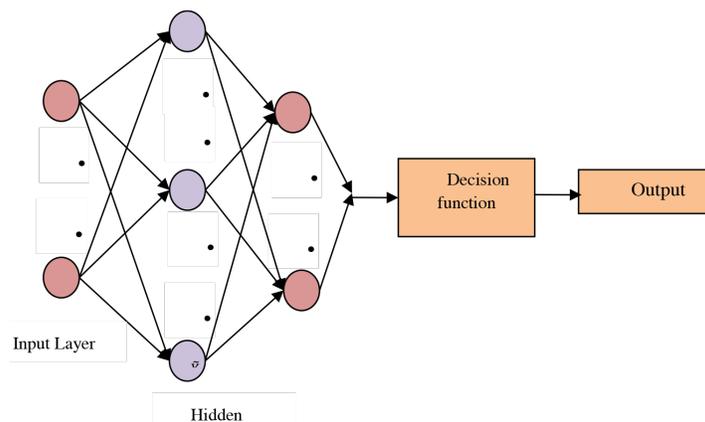


Figure 4: Basic ELM model.

In ELM, the learning parameters such as biases, weights, and input nodes are assigned inconstantly without any requirement of tuning. The output is evaluated by the inverse operation and only the thing that is required is to mention the number of hidden nodes. This algorithm has been deployed successfully in many research areas such as regression, pattern recognition and especially in the field of classification. In this work ELM has been used as a classifier for the classification of images in large databases. The activation function that is used here is Radial basis function (RBF) due to its universal approximation ability which has been already proved in the literature. The mathematical brief overview of ELM is presented below:

For X random samples (x_i, y_i) where $x_i = [x_{i1}, x_{i2} \dots x_{im}]^T \in R^n$ and $y_i = [y_{i1}, y_{i2} \dots y_{im}]^T \in R^m$ with \tilde{X} hidden nodes, the activation function $f(x)$ is modeled mathematically as

$$\sum_{i=1}^{\tilde{X}} \beta_t f_t(x_j) = \sum_{i=1}^{\tilde{X}} \beta_t f_t(w_t \cdot x_j + b_t) = O_j, \quad j = 1, 2, \dots, X \quad (5)$$

Where, $w_t = [w_{t1}, w_{t2} \dots w_{tm}]^T$, weight vector connecting the input and hidden nodes.

β Vector is the connections between hidden and output nodes and b_t is the threshold of the t th hidden node.

The Equation (5) can be written as

$H\beta = O$, here $w_t \cdot x_j$ is the inner product of w_t and x_j .

Where

$$H = H = \begin{bmatrix} f(w_1 \cdot x_1 + b_1) & \dots & f(w_{\tilde{X}} \cdot a_1 + b_{\tilde{X}}) \\ \vdots & & \vdots \\ f(w_1 \cdot x_X + b_1) & \dots & f(w_{\tilde{X}} \cdot a_X + b_{\tilde{X}}) \end{bmatrix}_{X \times \tilde{X}}, \quad H \text{ is the output matrix of hidden layer of the network.}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{X}}^T \end{bmatrix}_{\tilde{X} \times m}, \quad O = \begin{bmatrix} O_1^T \\ \vdots \\ O_X^T \end{bmatrix}_{X \times m}$$

The learning parameters x_i and y_i are nominated inconstantly without taking consideration of the input. By using simple mathematical transformation the output weights are calculated. Hence the learning speed of ELM is very fast as compared with other neural networks.

4. Proposed Framework

A smart and novel CBIR system is proposed on texture based images which is based on incorporating the properties of two leading techniques LBP and DWT to get the finer details of texture. Both these techniques are jointly complementary. The LBP technique has higher ability of discrimination and less computational complexity. It catches the detailed local appearance of the image while the DWT has the multi-resolution and multi-orientation properties. It is able to extract the information regarding shape of the image over a higher scale, thus increasing the retrieval rate. Features of the database images are extracted by both these techniques and are combined by the process of normalization. Min-Max normalization method is adopted here for merging, as this technique conserves the correspondence between all the data values.

After this, the next step that is proposed here is classification. The joint feature vector obtained by the combination of LBP and DWT is the input to the classifiers. Two classifiers are compared here i.e. ELM and SVM. Both these classifiers are firstly trained by the images of the databases. The output of these classifiers is the texture images with different classes. The comparative analyses of two systems are proposed here. Proposed method P1 is the concatenation of (LBP+DWT+SVM) and P2 is (LBP+DWT+ELM).

After the categorization process done by these classifiers the similarity matching of the specific query image is performed by its respective class only.

The steps of training and testing are shown below.

Training phase:

- Texture features by the combination of DWT and LBP are extracted from the labeled database images.

- Joint feature vectors are formed of all the images by the process of min-max normalization.
- ELM/ SVM are trained by the entire database images and are modeled as a classifier.
- Multiple classes are formed by these classifiers.

Testing phase:

- Similar joint feature vectors are formed for the query image.
- Query image is classified into the most pertinent class obtained by the classifier.
- Similarity of the query image is calculated from the images of its belonging class only not from the whole database.
- Finally images are retrieved.

This process is done with both proposed methods P1 (LBP+DWT+SVM) and P2 (LBP+DWT+ELM) with calculation of average precision and accuracy of the system by taking every image as the query image. The block diagram of the designed framework with training and testing phases is shown in Figure 5a and 5b respectively. The performances of the proposed methods are compared with other existing texture based CBIR systems.

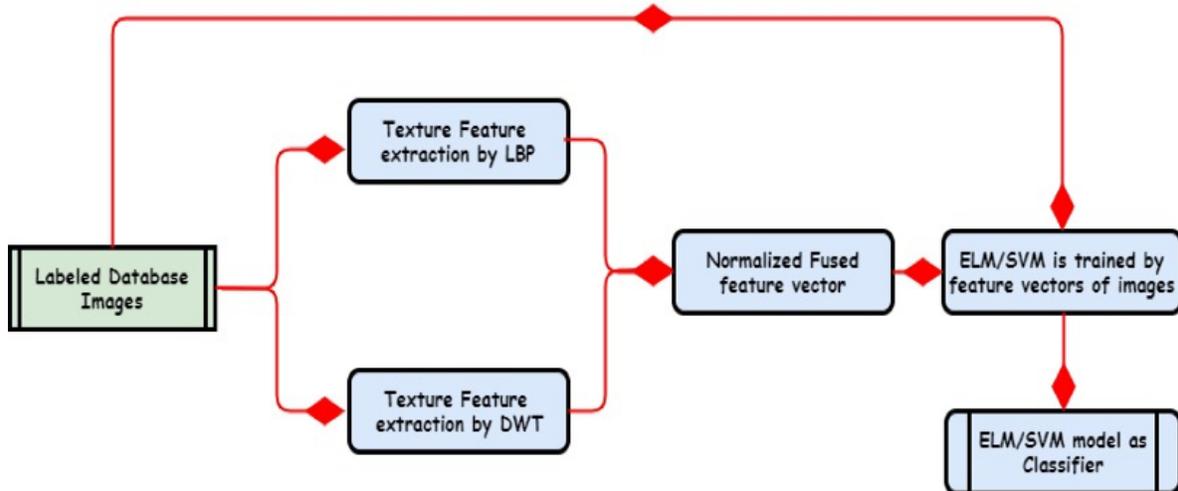


Figure 5a: Training phase of the proposed system.

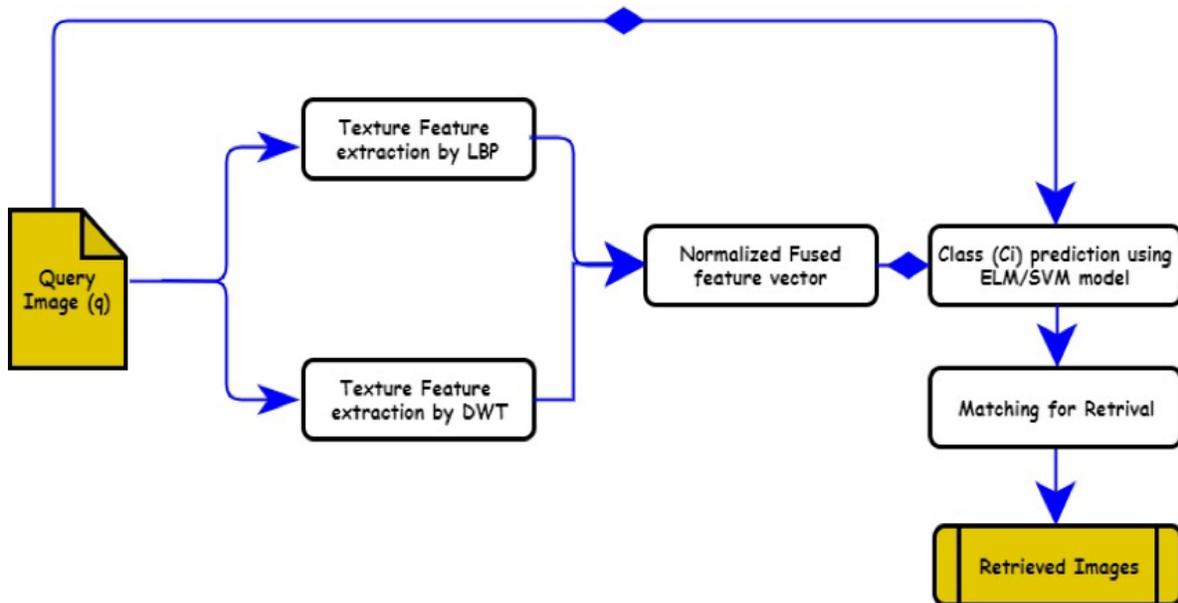


Figure 5b: Testing phase of the proposed system.

4.1. Similarity Measurement of Extracted Features

After the classification of query images into the matching classes through the above said proposed method, the matching step is performed between the images by using different distance measures which are shown below in Equations (5), (6) and (7). Lesser the distance, higher is the similarity and these distances are sorted in the increasing order in order to retrieve the similar images. Some of the significant distance measures are:

$$D_{\text{Euclidean}} = \sqrt{\sum_{i=1}^n (|I_i - D_i|)^2} \quad (6)$$

$$D_{\text{Manhattan}} = \sum_{i=1}^n |I_i - D_i| \quad (7)$$

$$D_{\text{Minkowski}} = [\sum_{i=1}^n (|I_i - D_i|)^p]^{\frac{1}{p}} \quad (8)$$

In these above equations I denote the feature vector of the query image and D denotes the feature vector extracted for every image of the database. All of these distances are tested in our experiments and it is found that Euclidean distance provides the better results because this distance tries to minimize all the errors equitably as the gradient is of consistent magnitude which is better for high dimension data.

The experiments evaluate the most important parameters of CBIR system which are precision and recall which are calculated as Equations. (9) and (10).

$$\text{Precision} = \frac{\text{No.of relevant images Retrieved}}{\text{Total No.of images Retrieved}} \quad (9)$$

$$\text{Recall} = \frac{\text{No.of relevant images retrieved}}{\text{No.of relevant images in database}} \quad (10)$$

Every image from the database is treated as the query image for the image retrieval and decisive parameters are evaluated such as precision and recall. Databases consist of different categories such as Brodatz contains

13 and MIT Vis Tex comprises of 8 categories. For each category average value of precision and recall can be figure out as Equations. (11) and (12).

$$\text{Precision}(X_i) = \frac{1}{M_i} \sum_{K=1}^{M_i} \text{Precision}(Q_a^i) \quad (11)$$

$$\text{Recall}(X_i) = \frac{1}{M_i} \sum_{K=1}^{M_i} \text{Recall}(Q_a^i) \quad (12)$$

In the above equations X_i is the number of categories in the dataset and M_i is the number of images in one category i . The symbol Q_a^i denotes the a 'th query image taken from category i .

5. Experimental Set up and Results

Two touchstone databases Brodatz and VisTex are taken for conducting the experiments. In every experiment the parameters precision and recall are evaluated by the combination of LBP and DWT technique and finally by the proposed two methods P1 (Combined with SVM) and P2 (Combined by ELM).

5.1. Experiment 1

In the 1st experiment rotated Brodatz texture database is taken which contains 13 different classes of images such as bark, brick, grass, raffia etc. having size 512 x 512 and every class consists of 7 images having orientations (0° , 30° , 60° , 90° , 120° , 150° , 200°). After that each image is subdivided into 16 smaller images and the total database now contains 1456 images each having size 128 x 128. Some sample images from this database are shown in Figure 6. Top 25, 35, 45, 55, 65 images are retrieved by taking every image as query image from the complete database.

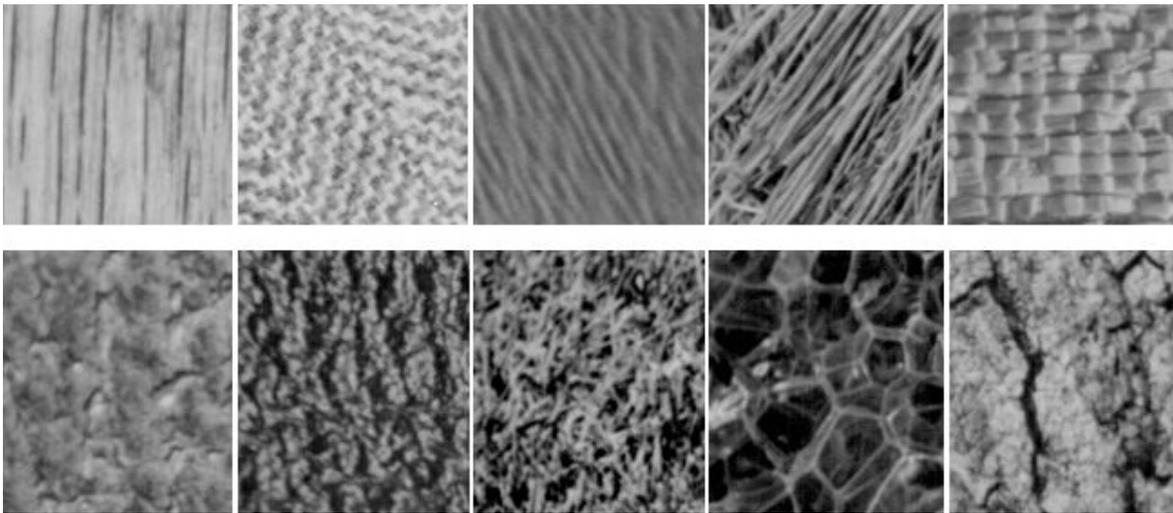


Figure 6: Sample images of Brodatz database.

Graphs are obtained for average precision and recall from the combination of DWT and LBP and also for the proposed method P1 and P2 which shows that the performance of the proposed framework is much superior to the simple combined approach of both texture techniques. However, the average precision of ELM is somewhat better than SVM. Average precision and average recall is shown in Figure 7 and Figure 8 respectively.

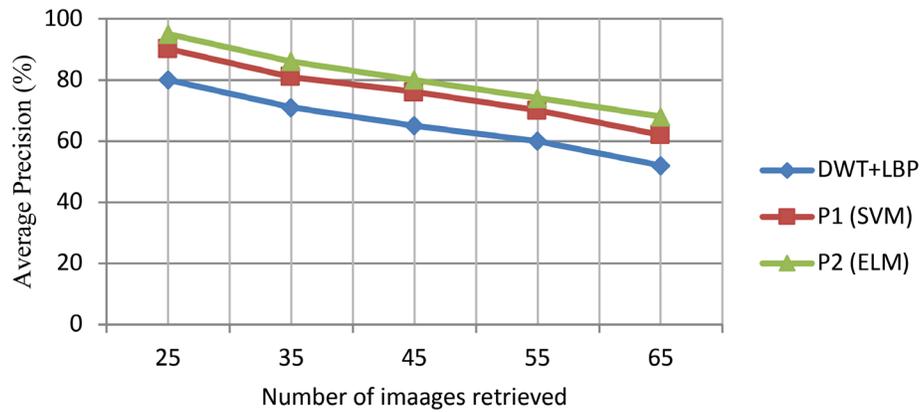


Figure 7: Comparison of average precision of all techniques.

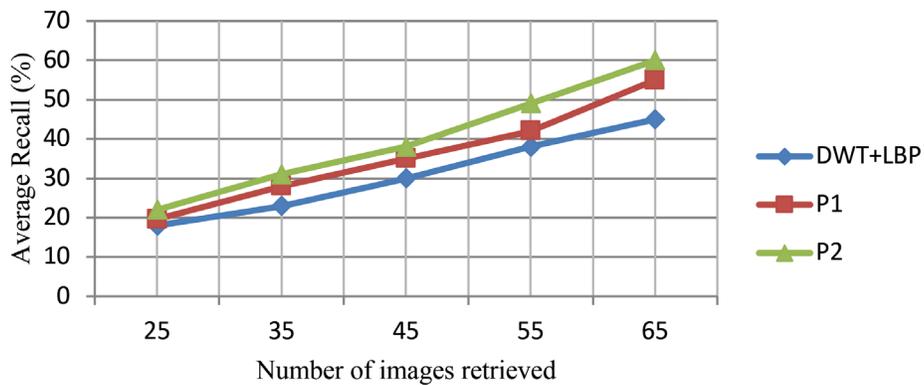


Figure 8: Comparison of average recall of all techniques.

5.2. Experiment 2

In the second experiment MIT VisTex database has been taken in which 40 different texture images are selected each of size 512 x 512. After this every image has been subdivided into 16 parts each of size 128 x 128 and the final database consists of 640 images. Similarly like the first experiment all the images are taken as the query image and top 16, 32, 48, 64 and 80 images are retrieved. Sample images of this dataset are shown in Figure 9

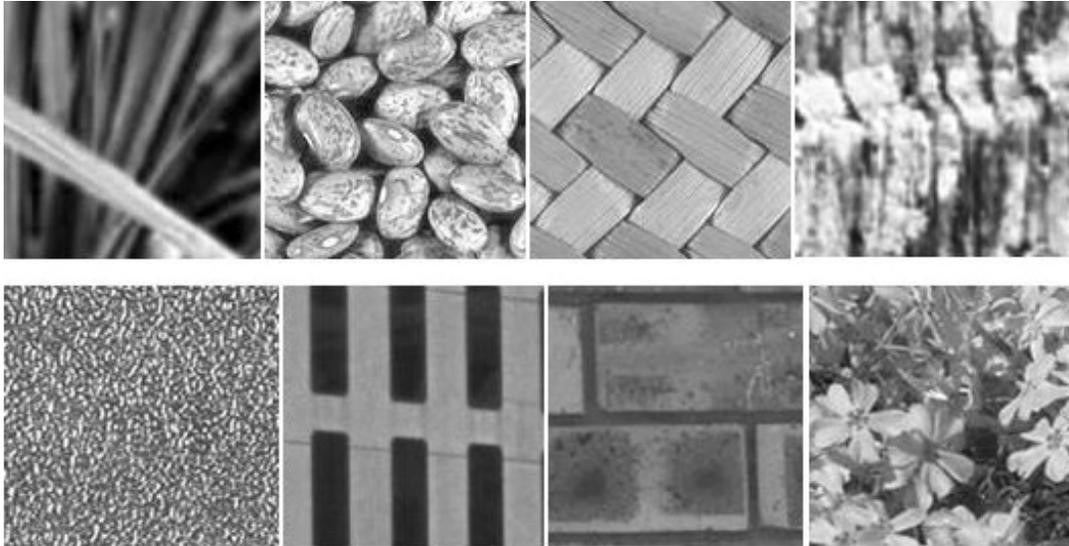


Figure 9: Sample images of Vis Tex database.

Average precision and recall are evaluated by the combination of LBP+DWT and for proposed methods P1 and P2. Comparison graphs are evaluated which are presented in Figure 10 and Figure 11 for precision and recall at different image retrievals. As per the graphs obtained, in this dataset also the results of the proposed method with ELM classifier provide the best results as compared with the others.

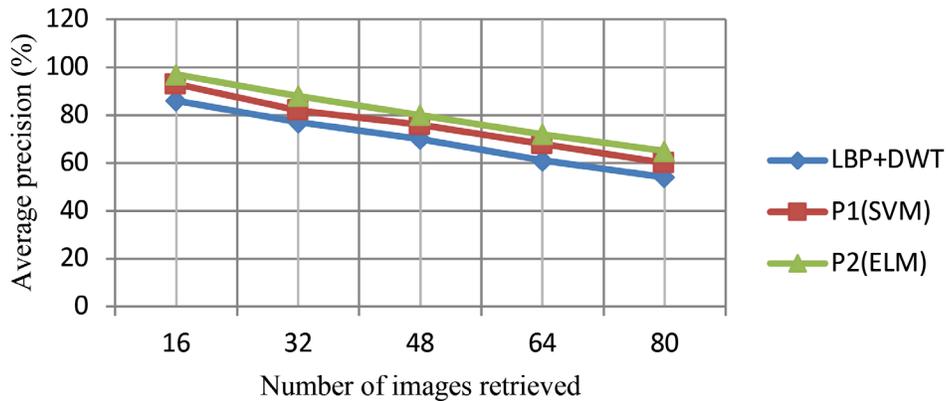


Figure 10: Comparison of average precision of all techniques.

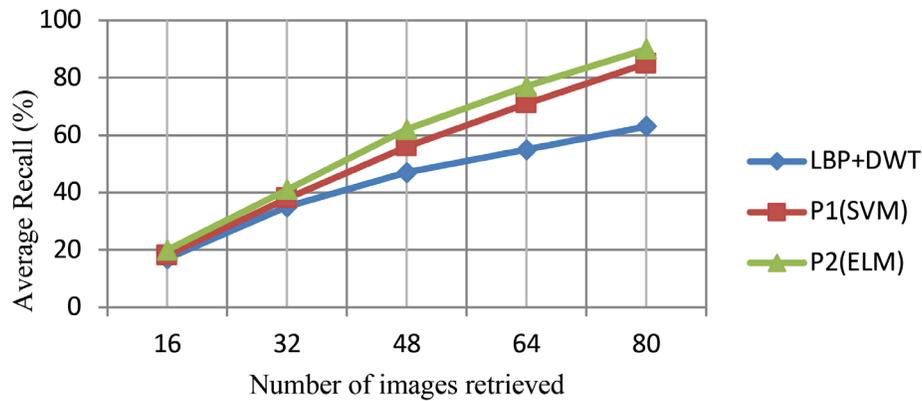


Figure 11: Comparison of average recall of all techniques.

5.2. Performance comparison of the proposed methods with other existing techniques

Three texture descriptor techniques for CBIR systems are designed in this paper. Average rate of the precision values of all the three techniques are described in Table 1.

Table 1: Precision values of all the three designed methods on euclidian distance

Average Precision of techniques (%)			
Datasets	LBP+DWT	P1(LBP+DWT+SVM)	P2(LBP+DWT+ELM)
Brodatz	86	90	94
Vistex	89	93	96

In order to authenticate the novelty in terms of precision of our proposed work, its comparison with other latest state-of-art methods is tabulated in Table 1 for both Brodatz and VisTex datasets on retrieval of top 25 and 16 images respectively.

Table 2 : Average precision comparison of proposed methods with others when number of retrieval is 25 and 16 for Brodatz and VisTex database

Database	Murala et al.	Vahid et al.	Fadaei et al.	Raghu et al.	P1	P2
Brodatz	85.3	77.94	75	86	90	94
VisTex	90.02	86.53	91.66	88	93	96

From Table 2 it is clear that the values obtained of the average precision by our proposed methods are remarkably higher than other existing texture descriptor techniques in CBIR systems.

5.3. Accuracy of the proposed methods

The accuracy of both the proposed CBIR systems i.e P1 with SVM classifier and P2 with ELM classifier is shown in Table 3. As shown in the table the accuracy of ELM is much better than the accuracy of the SVM classifier in both types of databases. The confusion matrix is created by SVM and ELM classifier by which accuracy is evaluated. For an example the matrix for brodatz database of ELM classifier is shown in Figure 12. The diagonal elements of the confusion matrix shows the relevant images retrieved per category. In the case of brodatz dataset, there are total 13 categories and 112 images per category. Clearly, from the figure it is observed that 108 images are retrieved from first category, 109 from second and so on.

```

cm =

    108     1     1     0     0     0     1     1     0     0     0     0     0
         0    109     0     1     0     2     0     0     0     0     0     0     0
         0     0    111     0     0     0     0     0     1     0     0     0     0
         1     0     0    109     0     0     0     1     0     1     0     0     0
         0     0     0     0    110     1     1     0     0     0     0     0     0
         0     0     0     0     0    112     0     0     0     0     0     0     0
         0     0     1     0     0     0    110     0     1     0     0     0     0
         0     0     0     0     0     0     0    111     0     1     0     0     0
         0     0     1     0     0     0     0     1    110     0     0     0     0
         1     0     0     0     0     0     2     0     1    108     0     0     0
         0     1     0     0     0     0     0     0     0     0    111     0     0
         0     0     0     0     0     0     0     0     1     0     0    111     0
         0     1     0     2     1     0     0     0     0     0     0     0    108
    
```

```

acc =

    98.0769
    
```

Figure 12: Confusion matrix of Brodatz database from ELM classifier.

Table 3: Accuracy comparison of proposed methods

Database	Proposed Method	Accuracy
Brodatz	P1	89%
	P2	98%
VisTex	P1	94%
	P2	99%

5.4. Image retrieval results

The image retrieval results of the proposed framework with ELM classifier for the analyzed datasets are shown in Figures 13 and Figure 14 respectively by taking random query image from any category of the databases. Top 10 images are retrieved in every case.

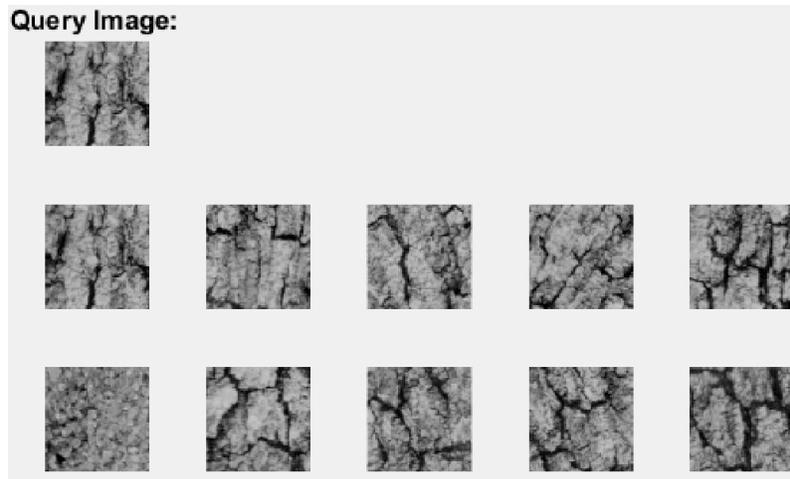


Figure 13: Retrieved top 10 images from Brodatz database by proposed ELM method.

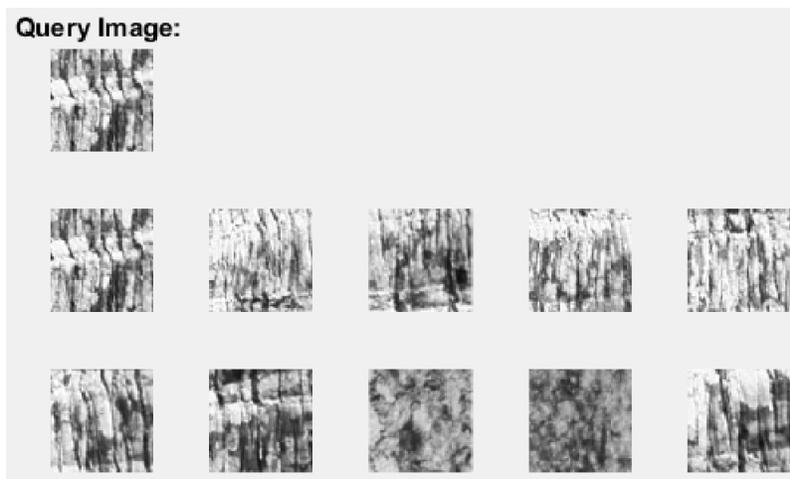


Figure 14: Retrieved top 10 images from Vis Tex database by proposed ELM method.

6. Conclusion

An intelligent, unique and hybrid CBIR system is created for texture based images to increase the overall precision and accuracy of the system. The combination of features and classifications plays the supreme role in this designed framework. For obtaining the finest details of texture, two techniques (LBP and DWT) are merged together for the creation of an efficient hybrid feature vector. After this step, the machine learning classifier is applied which produces the multiple classes of large databases. Due to this classification, the searching area of retrieval of relevant images from the query image is reduced and the system becomes more accurate and speedy. The performances of two machine learning classifiers are compared here which are SVM and ELM. This new overall system is experimented on two benchmark datasets i.e. Brodatz and MIT VisTex and similarity between the images is determined by applying Euclidean distance. The results obtained spectacles that the performance of the proposed system is superior to other texture based methods with both the classifiers and much better with ELM classifier as the learning speed of this is higher than SVM classifier.

7. References

- Alaei, F., Alaei, A., Pal, U., & Blumenstein, M. (2018). AC PT US CR. *Expert Systems With Applications*. <https://doi.org/10.1016/j.eswa.2018.12.007>.
- Chamasemani, F. F. (2011). Multi-class Support Vector Machine (SVM) classifiers - An Application in Hypothyroid detection and Classification, 353-358. <https://doi.org/10.1109/BIC-TA.2011.51>.
- Chorowski, J., Wang, J., & Zurada, J. M. (2014). Neurocomputing Review and performance comparison of SVM and ELM-based classifiers *S*, 128, 507-516. <https://doi.org/10.1016/j.neucom.2013.08.009>.
- Das, R., Dash, J. K., & Mukhopadhyay, S. (2013). Rotation invariant textural feature extraction for image retrieval using eigen value analysis of intensity gradients and multi-resolution analysis, 46, 3256-3267. <https://doi.org/10.1016/j.patcog.2013.05.026>
- Ding, S., & Xu, X. (2013). Extreme learning machine: algorithm, theory and applications, (August 2014). <https://doi.org/10.1007/s10462-013-9405-z>.
- Fadaei, S., Amirfattahi, R., & Ahmadzadeh, M. R. (2017). Local derivative radial patterns: A new texture descriptor for content-based image retrieval. *Signal Processing*, 137, 274-286. <https://doi.org/10.1016/j.sigpro.2017.02.013>.
- Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). Neurocomputing Deep learning for visual understanding: A review, 187, 27-48. <https://doi.org/10.1016/j.neucom.2015.09.116>.
- Haralick, R. M., & Shanmugam, K. (1973). Textural Features for Image Classification.
- Hemachandran, K., Paul, A., & Singha, M. (2012). Content-based image retrieval using the combination of the fast wavelet transformation and the colour histogram. *IET Image Processing*, 6(9), 1221-1226. <https://doi.org/10.1049/iet-ipr.2011.0453>.
- Huang, D., Member, S., Shan, C., & Ardabilian, M. (2011). Local Binary Patterns and Its Application to Facial Image Analysis: A Survey, (November). <https://doi.org/10.1109/TSMCC.2011.2118750>.
- Huang, G., Member, S., Zhou, H., Ding, X., & Zhang, R. (2012). Extreme Learning Machine for Regression and Multiclass Classification, 42(2), 513-529.
- Jain, A. K., & Vailaya, A. (1995). Image Retrieval using Color and Shape. *Pattern Recognition*, 29, 1233-1244. [https://doi.org/10.1016/0031-3203\(95\)00160-3](https://doi.org/10.1016/0031-3203(95)00160-3).
- Karthikeyan, T., & Manikandaprabhu, P. (2014). A Study on Discrete Wavelet Transform based Texture Feature Extraction for Image Mining, 5(5), 1805-1811.
- Kastrati, Z., & Imran, A. S. (2019). Performance Analysis of Machine Learning Classifiers on Improved Concept Vector Space Models. *Future Generation Computer Systems*. <https://doi.org/10.1016/j.future.2019.02.006>.
- Kokare, M., Biswas, P. K., & Chatterji, B. N. (2005). Complex Wavelet Filters, 35(6), 1168-1178.
- Kumar, Y., Aggarwal, A., Tiwari, S., & Singh, K. (2018). An efficient and robust approach for biomedical image retrieval using Zernike moments. *Biomedical Signal Processing and Control*, 39, 459-473. <https://doi.org/10.1016/j.bspc.2017.08.018>.
- Liao, S., Law, M. W. K., & Chung, A. C. S. (2009). for Texture Classification, 18(5), 1107-1118.
- Liu, S., Wang, H., Wu, J., & Feng, L. (2015). Incorporate Extreme Learning Machine to content-based image retrieval with relevance feedback, (March). <https://doi.org/10.1109/WCICA.2014.7052854>.
- Lu, B., Duan, X., & Wang, C. (n.d.). A Novel Approach for Image Classification Based on Extreme Learning Machine.
- Maheshwari, S. M. R. P., & Balasubramanian, R. (2012). Directional local extrema patterns: a new descriptor for content based image retrieval, 191-203. <https://doi.org/10.1007/s13735-012-0008-2>.
- Murala, S., Maheshwari, R. P., & Balasubramanian, R. (2012). Local Tetra Patterns: A New Feature Descriptor for Content-Based Image Retrieval, (May). <https://doi.org/10.1109/TIP.2012.2188809>.
- Naghashi, V. (2018). Optik Co-occurrence of adjacent sparse local ternary patterns: A feature descriptor for texture and face image retrieval. *Optik - International Journal for Light and Electron Optics*, 157, 877-889. <https://doi.org/10.1016/j.ijleo.2017.11.160>.

- Pavithra, L. K., & Sharmila, T. S. (2017). An efficient framework for image retrieval using color, texture and edge features R. *Computers and Electrical Engineering*, 0, 1-14. <https://doi.org/10.1016/j.compeleceng.2017.08.030>.
- Pham, M. (2017). Color Texture Image Retrieval Based on Local Extrema Features and Riemannian Distance. <https://doi.org/10.3390/jimaging3040043>.
- Prakasa, E. (2016). Ekstraksi Ciri Tekstur dengan Menggunakan Local Binary Pattern Texture Feature Extraction by Using Local Binary Pattern, 9(2), 45-48. <https://doi.org/10.14203/j.inkom.420>.
- Puviarasan, N., Bhavani, R., & Vasanthi, A. (2014). Image Retrieval Using Combination of Texture and Shape Features, 3(3), 5873-5877.
- R, J. V. C. I., Murala, S., & Wu, Q. M. J. (2014). Expert content-based image retrieval system using robust local patterns. *Journal Of Visual Communication And Image Representation*, 25(6), 1324-1334. <https://doi.org/10.1016/j.jvcir.2014.05.008>.
- Raghuwanshi, G., & Tyagi, V. (2015). Texture image retrieval using adaptive tetrolet transforms. *Digital Signal Processing*, 1(3), 1-8. <https://doi.org/10.1016/j.dsp.2015.09.003>.
- Ricardo, A., Joaci, J., & Sá, D. M. (2017). Neurocomputing LBP maps for improving fractal based texture classification, 266, 1-7. <https://doi.org/10.1016/j.neucom.2017.05.020>.
- Sreena, P. H., & George, D. S. (2013). Content Based Image Retrieval System with Fuzzified Texture Similarity Measurement, (Iccc), 80-85.
- Srivastava, P., & Khare, A. (2017). Utilizing multiscale local binary pattern for content-based image retrieval. <https://doi.org/10.1007/s11042-017-4894-4>.
- Wang, H., Feng, L., Zhang, J., & Liu, Y. (2016). Semantic discriminative metric learning for image similarity measurement. *IEEE Transactions on Multimedia*, 18(8), 1579-1589. <https://doi.org/10.1109/TMM.2016.2569412>.
- Zhao, R., & Grosky, W. I. (2000). From features to semantics: some preliminary results. *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on*, 2(c), 679-682 vol.2. <https://doi.org/10.1109/ICME.2000.871453>.

