

# Indian Monuments Classification using Support Vector Machine

Malay S. Bhatt<sup>1</sup>, Tejas P. Patalia<sup>2</sup>

<sup>1</sup>Rai University, Ahmedabad, India

<sup>2</sup>Dept. of Computer Engineering, V. V. P. Engineering College, Rajkot, India

---

## Article Info

### Article history:

Received Nov 15, 2016

Revised Jun 4, 2017

Accepted Jun 21, 2017

---

### Keyword:

Classification

Edge detection

Generalized co-occurrence matrix

Local binary pattern

Monument

Support vector machine

---

## ABSTRACT

Recently, Content-Based Image Retrieval is a widely popular and efficient searching and indexing approach used by knowledge seekers. Use of images by e-commerce sites, by product and by service industries is not new nowadays. Travel and tourism are the largest service industries in India. Every year people visit tourist places and upload pictures of their visit on social networking sites or share via the mobile device with friends and relatives. Classification of the monuments is helpful to hoteliers for the development of a new hotel with state of the art amenities, to travel service providers, to restaurant owners, to government agencies for security, etc.. The proposed system had extracted features and classified the Indian monuments visited by the tourists based on the linear Support Vector Machine (SVM). The proposed system was divided into 3 main phases: preprocessing, feature vector creation and classification. The extracted features are based on Local Binary Pattern, Histogram, Co-occurrence Matrix and Canny Edge Detection methods. Once the feature vector had been constructed, classification was performed using Linear SVM. The Database of 10 popular Indian monuments was generated with 50 images for each class. The proposed system is implemented in MATLAB and achieves very high accuracy. The proposed system was also tested on other popular benchmark databases.

*Copyright © 2017 Institute of Advanced Engineering and Science.  
All rights reserved.*

---

## Corresponding Author:

Malay S. Bhatt,

Rai University,

Ahmedabad, India.

E-mail : malaybhatt202@yahoo.com

---

## 1. INTRODUCTION

Content Based Image Retrieval system has become a significant research issue as plenty of image data have been generated in areas like medicine, Fashion Design, art galleries, entertainment, education, manufacturing and more. QBIC System of IBM [1], Chabot of U.C. Berkeley, Photobook of Massachusetts Institute of Technology (MIT) [2], VisualSEEK [3] and MARS [4] are popular examples of CBIR software systems. The text based retrieval system involves manual annotation of images which involves problems like the vast amount of laborious task and most importantly human perception of the image. Different person perceives the same image differently [5-7].

For many years, a tremendous amount of multimedia data in the form of images, audio and video have been generated due to availability of cost-effective electronic devices like camera, mobile or Handy cam. These multimedia data have been shared, uploaded or emailed to relatives and friends staying away to make them feel that they had not missed the precious moments. Millions of such photographs are uploaded and it is almost impossible to manually classify these pictures as per the monuments people visited. In January 2013, the India was at 3rd position with 62.6 million Facebook members [8].

Bollywood and Television industry are also considered as major sources of multimedia data (Movie stills, Posters etc.) As they release more than 2000 movies, a lot of music videos, produce serials and also organize events for 100 years [9]. It can be observed that pilot scenes and song sequences consist of the monuments in the background and thus movies promote tourism and increase the revenue of the country as a whole. Figure 1 depicts the presence of the monuments in scenes of popular movies.

Travel and tourism industry incorporate heritage, medical, business and sports tourism. The main objective of this sector is to develop and promote tourism, to maintain competitiveness of India as a tourist destination and to improve and to expand existing tourism products to ensure employment generation and economic growth [10]. The government is also promoting tourism through advertisements, campaigns and takes special interest to preserve the beauty of these monuments. Every year, lots of foreign delegates and tourists visit India too. As per India Tourism Statistics 2013, 6.97 million foreign tourists arrived with an annual growth rate of 5.9% and approximately 1145 million domestic tourists with an annual growth rate of 9.6% were observed [11].

Table 1. Bollywood Movies and Monuments

Movie Name	Release Year	Monument	Movie Name	Release Year	Monument
Ki & Ka	2016	India Gate	Fanaa	2006	Qutub Minar
Ki & Ka	2016	Gate Way of India	Jannat 2	2012	Qutub Minar
FAN	2016	Red Fort	Rang De Basanti	2006	India Gate
Tevar	2015	Taj Mahal	Rang De Bassanti	2006	Golden Temple
Jeans	1998	Taj Mahal	Rab Ne banadi Jodi	2008	Golden Temple
Jhoom Barabar Jhoom	2007	Taj Mahal	Ki & Ka	2016	India Gate
Leader	1964	Taj Mahal	Ki & Ka	2016	Gate Way of India
Mere Brother ki Dulhan	2011	Taj Mahal	FAN	2016	Red Fort
Namastey London	2007	Taj Mahal	Tevar	2015	Taj Mahal
Youngistaan	2014	Taj Mahal	Jeans	1998	Taj Mahal



Figure 1. (a) 'Fanaa' Movie (b) 'Mere Brother Ki Dulhan' Movie (c) 'Leader' Movie (d) 'Tevar' Movie (e) 'Rab Ne Bana Di Jodi' Movie (f) 'Rab Ne Bana Di Jodi' Movie (g) 'Jannat 2' Movie and (h) 'Rang De Basanti' Movie.[12]

## 2. LITERATURE SURVEY

Bhatt M.S. and Patalia T. P. [12] used Generalized Co-occurrence Matrix (GCM) obtained from HSV color space having 64 gray-levels with various distance values (3, 6, 9, 12 and 15) as input to Genetic Programming System. Genetic Programming evolved spatial descriptor using 15 Generalized Co-Occurrence matrices as terminals have been implemented with 7 functions and 2 operators. Each GCMs having size of 64x64 is considered as input. Obtained Genetic programming evolved spatial descriptor is tested on manually created Indian monuments database having four classes, namely 'Taj Mahal', 'Qutub Minar', 'Golden Temple' and 'India Gate'. Fitness function used in GP system is linear SVM with 10 fold cross validation. They obtained accuracy of 92 %.

Murala et al. [13] proposed Local Tetra Pattern (LTrP) as a new feature descriptor for Content Based Image Retrieval. LTrP uses four distinct values for encoding of information and also uses direction information. They highlighted advantages of LTrP over Local Binary Pattern (LBP), Local Ternary Pattern (LTP) and Local Derivative Pattern (LDP). Benchmark databases viz., Corel 1000 database, Brodatz texture database and MIT VisTex database were used for performance comparison.

Youness et al. [14] proposed content based image retrieval based on 2-D ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) and Gabor Filter. They used Brodatz gray scale image dataset, having 13 texture classes with 16 samples for each class, for the purpose of evaluation. They achieved an average precision of 80.19%.

Nazarloo et al. [15] applied content based image retrieval for gender classification. The face is one of the most important biometric of human and contains lots of useful information. For gender classification, they merged the Gabor Filters and Local Binary Pattern Features of the face. Self-Organized Map was used for the classification and achieved an accuracy of 92.5%.

Guo et al. [16] explained Completed Local Binary Pattern (CLBP) which is a modification to LBP. CLBP composed of Local Difference Sign Magnitude Transform (LDSMT) and center gray level value. They tested the proposed approach on CURET and Outex texture databases. It is observed that sign component preserves much better local difference as compared to magnitude component and thus outperforms conventional LBP.

Das et al. [17], [18] highlighted the importance of classified query for content based image recognition based on niblack's thresholding. They applied binarization on Red, Blue and Green planes separately based on niblack's thresholding. On the binarized images, upper mean, lower mean, upper standard deviation and lower standard deviation are calculated. For each plane 4 features are extracted and thus a feature vector of size 12 is generated. They tested the system on benchmark databases 'Wang Database' and 'OT-Scene'. Highest precision of 0.838 and recall of 0.838 was achieved using Artificial Neural Network with a Multi Layer Perceptron for 'Wang Database' while highest precision of 0.753 and recall of 0.754 are obtained using Support Vector Machine for 'OT-Scene' database.

Streater, J [19] explains the design and implementation of genetic programming system for construction of feature descriptor for skin lesion images. 6 Generalized Co-occurrence Matrices (GCMs) of size 64x64 are constructed using RGB color space with 5 inter-pixel distances for 100 images. Highest accuracy achieved was 72%. Fisher's Discriminant Ratio, naïve Bayes classifier is also used to compare the results obtained by the SVM.

Extraction of information from movies has been focused since last 10 years. Hollywood Movie Database 51 [20], YouTube [21] and the Hollywood datasets [22] are challenging datasets used widely for action recognition. Action recognition from movies is a subset of general human recognition activity.

Laptev et al. [23] highlighted limitations of the human action dataset in a controlled environment and also described the difficulties faced during the recognition of real movie actions. They identified similarities and dissimilarities between action recognition from movies with object recognition in still images. The first task accomplished with the accuracy of 60% is the automatic annotation of the human action using the movie scripts. Inaccuracy is due to script video misalignment. Classification of human-action is the main goal, achieved in the paper with 91.8% accuracy. Experiments are carried out for 8 different actions. HoG, HoF, Spatio-Temporal Bag of Features (BoF) and a combination of the above are used with the non-linear support vector machine to achieve the desired result.

Lei Chen et al. [24] depict a top-down approach based on rules for video editing and audio cues to extract dialogue and action scenes. A finite state machine with an audio-based support vector machine (SVM) classifier is applied for detection of skin type. Classifier uses three features, namely: variance of zero crossing rate, silence ratio, and harmonic ratio. The precision and recall rates achieved are 76.56% and 81.6% respectively.

Doudpota et al. [25] focused on the impact and popularity of Bollywood movies in South Asia, Middle East, UK, USA and other parts of the world. They mined song sequences from the Bollywood movies. They used in the first part, Zero Crossing Rate, Spectrum Flux and Short Time, Energy as features in Support Vector machine for binary classification of extracting segment into music and non-music. In the second phase, extracted music segments are further classified into song and non-song sequences using Probabilistic Timed Automata (Song Grammar). An experiment was carried out on 10 Bollywood movies having 74 songs and out of which 69 were successfully extracted. Recall achieved is 93.24%, while the Precision is 87.34%.

Vaudeville et al. [26] have proposed integrated color and intensity co-occurrence matrix (ICICM) for content based image retrieval. The ICICM composed of four other matrices namely,  $ICICM_{CC}$ ,  $ICICM_{CI}$ ,  $ICICM_{IC}$  and  $ICICM_{II}$ .  $ICICM_{CC}$  captures color perception of pixel  $p$  and color perception of the neighborhood of  $P$  while  $ICICM_{CI}$  captures color perception of pixel  $p$  and intensity perception of the neighbourhood of  $P$ . The other two can be described similarly. ICICM is updated based on the weight which is a function of Saturation and Intensity. The development of ICICM is based on the properties of the HSV color space. They have tested the system with a combination of various color (C) and gray level perception (I) levels on two different datasets. The first database is constructed from general purpose images obtained from International Microcomputer Software Inc. (IMSI) while the other is constructed using a crawler. A web-based application

for CBIR based on ICICM is available in the public domain for performing the experiments with images in our database as well as with externally uploaded images.

Deselaers et al. [27] performed a quantitative comparison of image features like color histograms, invariant feature histograms, Gabor feature histograms, Tamura texture feature, local feature and region based features. Correlation among these features is analyzed. They focused on two image retrieval tasks: color photographs (WANG Dataset) and medical radiographs (IRMA Dataset). Low correlation exists among region features, image features, invariant feature histogram and Gabor histograms. The combination of these will produce better image retrieval for color photographs. Invariant Feature histogram gives 15.9% error on WANG Dataset while error rate of 29.2% is identified in IRMA Dataset. Similar kind of comparison describes all remaining features. It is clear that selection of feature depends on the task at hand and the combination of positively correlated features does not improve the classification result.

Desai et al. [28] highlighted the importance of monuments classification to archaeologists in an assessment of their findings and in classification. Art galleries and museums focus on visual aspects of objects. A CBIR system based on visual shape based feature and texture feature was developed. Morphological operations were carried out for shape extraction and Gray level Co-occurrence Matrix was used for texture feature extraction. Five different classes with a total of 500 images are collected and performance was compared with Canny and Sobel edge detection approaches.

Information about Indian movies can be obtained via Indian Movie Database [29].

### 3. RESEARCH METHOD

Figure 2 provides a flow chart of a content based image retrieval system. An image query is the image file that is given as an input to the system. The features of the input are calculated. A query of the extracted features is then generated and is compared with all the other features of the image files present in the database. Based on similarity measures, the system retrieves the required image files from the database and presents it in the form of the result.

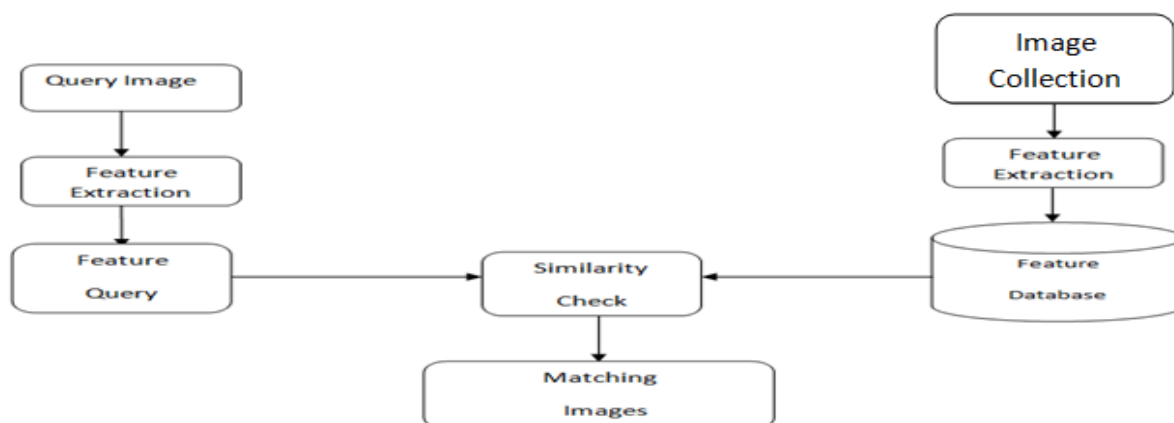


Figure.2 Content Based Image Retrieval System

#### 3.1. Preprocessing

A set of 500 input images is re-sized into 256x384. Resulted images are converted from RGB color space into Hue, Saturation and Value (HSV) color space [30]. The layers of human retina sense the light through rod cells and cone cells [31]. The gray-levels are perceived by rod cells at low-levels of illumination while at higher levels of illumination cone cells are also excited. The human perceives the color same as the HSV color space. RGB color representation is different and not as per human perception. Hue indicates the pure color, S indicates the percentage of white added in the pure color while V represents intensity. The HSV color space can be represented as a hexacone [32]. When saturation is zero, we get only shades of gray from black to white by increasing the intensity. Incident light composed of many spectral components, but causes loss of color information when saturation is low even though illumination is very high. By changing the saturation from 0 to 1, perceived color changes from shades of gray to pure color under the given hue and intensity. It is known that HSV color space has more discriminating power as compared to RGB color space. Figure 3(a) Displays RGB image while corresponding HSV image is shown in Figure 3(b).



Figure 3. (a) RGB Image and (b) HSV Image

Generalized co-occurrence matrix properties with different distance and direction are also computed and it results into 20 additional features. H, S and V planes are also extracted from the input image. The Computation of Centre Symmetric Local Binary Pattern with 16 bins and Histogram with 16 bins are carried out on each plane. Generalized co-occurrence matrix properties are also extracted from each plane. Figure 4 shows the pre-processing & Feature Vector Generation. Figure 5 covers several techniques which were merged together to generate the feature vector.

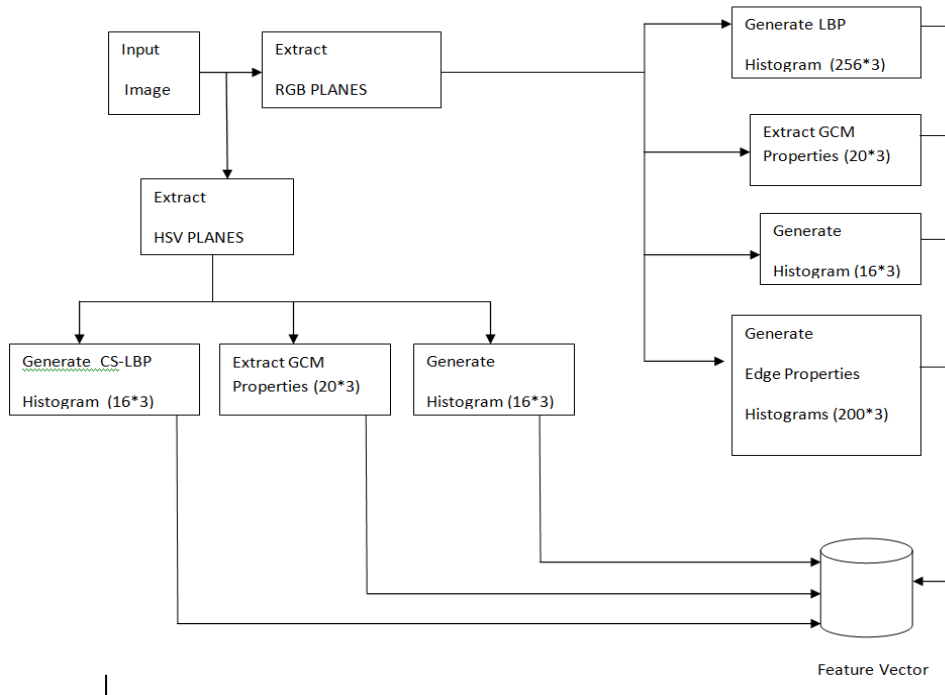


Figure 4. Pre-processing & Feature Vector Generation

LBP	CS- LBP	GCM Properties	Histogram	Edge Property Histogram
-----	---------	----------------	-----------	-------------------------

Figure 5. Feature Vector

### 3.2. Generalized Co-Occurrence Matrix

Generalized CO-Occurrence Matrix is useful to extract the texture of the image. It is represented as 4-tuple  $(i, j, d, \Theta)$ [33]. Here,  $i$  and  $j$  represent gray levels,  $d$  is the distance between pixels  $p_1$  and  $p_2$ . Graylevels of  $p_1$  and  $p_2$  are  $i$  and  $j$  respectively.  $\Theta$  is the angle between pixels  $p_1$  and  $p_2$ .

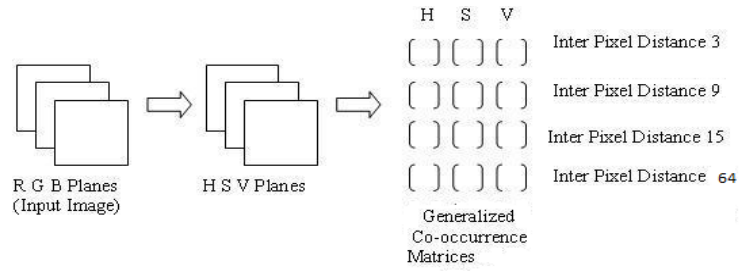


Figure 6. The Generalized Co-occurrence Matrices (GCM) of size 128x128 are calculated for inter-pixel distances 3,9, 15 and 64 in horizontal direction for H, S and V planes. (3 planes x 4 inter- pixel distance)

Figure 7(a) describes the working of GCM for a sample image of size 5x5 with 4 gray levels while Figure 7(b) contains calculated GCM in the horizontal direction with inter-pixel distance '1'.

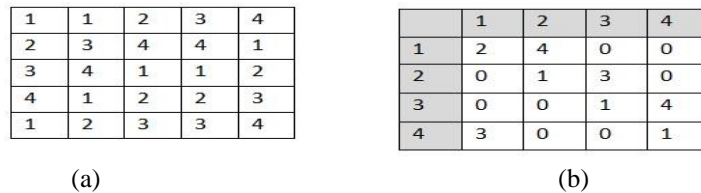


Figure 7. (a) A Sample Image (b) The GCM in horizontal Direction with inter- pixel distance '1'

Table 2 shows the generalized Co-Occurrence Matrix Properties [34].

Table 2. Generalized Co-Occurrence Matrix Properties [34]	
Property	Description
Contrast	It measures the intensity contrast between a pixel and its neighbour over the whole image.
Correlation	It indicates how correlated a pixel is to its neighbour over the whole image.
Energy	It represents the sum of squared elements in the GLCM. It is also known as uniformity.
Homogeneity	It measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

Contrast:

$$\sum_{i,j} |i - j|^2 p(i, j) \tag{1}$$

Correlation:

$$\sum_{i,j} \frac{(i-\mu_i)(j-\mu_j) p(i, j)}{\sigma_i \sigma_j} \tag{2}$$

Energy:

$$\sum_{i,j} p(i, j)^2 \tag{3}$$



Homogeneity:

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{4}$$

$p(i, j)$  represents count at position  $(i, j)$  in GLCM,  $\mu$  denotes mean and  $\sigma$  indicates the standard deviation in the above equations. Here, small, medium and large distance values are considered to capture the span of the monument in the horizontal and vertical direction. For example, span of the 'Red fort' is an almost whole image in the horizontal direction (left-to-right) with Hue, close to Red whereas 'Hawa Mahal' is spanned in both horizontal and vertical direction with Hue close to Red. Thus homogeneity and correlation properties are high for 'Red Fort' in the horizontal direction while the same properties are high in the horizontal as well as in the vertical direction in 'Hawa Mahal'.

Table 3 shows the generalized Co-Occurrence Matrix Used as Features

Number of Graylevels	Distance	Direction
128	3	Horizontal
128	9	Horizontal
128	15	Horizontal
128	64	Horizontal
128	64	Vertical

### 3.3. Local Binary Pattern and Centre-Symmetric Local Binary Pattern

The Local Binary Pattern effectively captures texture information from the local neighbourhood. Figure 8 explains working of the Local Binary Pattern and the Centre-Symmetric Local Binary Pattern.

$$LBP(x, y) = \sum_{i=0}^{n-1} s(n_c - n_i) 2^i \tag{5}$$

$$s(x) = 1 \text{ if } x \geq 0$$

$$0 \text{ otherwise}$$

Here,  $n_c$  indicates the graylevel of the centre pixel of 8-neighbourhood,  $n_i$  indicates  $i$ th pixel of the neighbourhood. The signs of the differences in a neighbourhood are interpreted as N-bit binary number resulting in  $2^N$  distinct values in the binary pattern. The LBP features are robust against illumination changes, they are very fast to compute, do not require many parameters to be set, and have high discriminative power [36]. In CS-LBP, center symmetric pairs of pixels are compared. LBP produces 256 distinct binary patterns, whereas CS-LBP generates 16 distinct binary patterns. The robustness on flat image regions is obtained by thresholding the gray level differences with a small value  $T$ . In our proposed system, histogram of CS-LBP is generated for all 3 planes of HSV image resulting in 48 ( $16 \times 3$ ) while the histogram of LBP is obtained for all 3 planes of RGB image resulting into 768 ( $256 \times 3$ ) features.

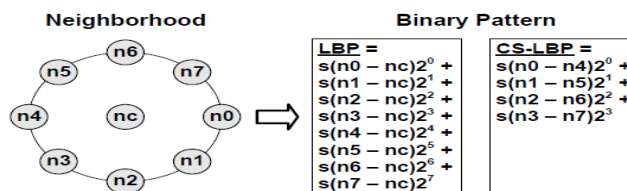


Figure 8. Local Binary Pattern and Centre Symmetric Local Binary Pattern [35]

### 3.4. Edge Histogram

Color information is obtained through histograms, an area information is added into feature vector using generalized co-occurrence matrix using different distance and direction, texture information is achieved

using LBP and CS-LBP histogram. To add the structural (behavior at the edge points) information in the feature descriptor, canny edge detector is used with threshold 0.2 so that most prominent edges are preserved. Canny edge detector consists of smoothing, finding gradients, non-maxima suppression, double thresholding and edge tracking by hysteresis [37]. For each detected edge point, 5x5 neighbourhood is considered and the mean and the standard deviation are calculated. The unique values obtained from these statistical properties vary for every image because the detected edge points are not fixed. It is observed that unique values are in the range of 2000-10000. Two Histograms with bin size 100 are generated for mean and standard deviation.

### 3.5. Fitness Function

Here, we have adopted the classification accuracy calculated by a linear SVM classifier on the training set as well as the testing set. We adopted tenfold cross-validation for which total dataset set is divided randomly into 10 equal-sized parts and perform ten repetitions of training the SVM on 9/10 of the set and testing on the remaining 1/10 [38]. The overall fitness 'Er' is the average of the tenfold cross-validation accuracy. In our case, the value of n is 10. Accuracy (i) represents the accuracy of fold I by the SVM. The fitness function is defined as follows:

$$Er = (1 - (\sum (SVM[accuracy(i)]/n)))*100 \% \quad (6)$$

## 4. RESULTS AND ANALYSIS

### 4.1. Implementation

Our goal is to classify above mentioned monuments from large repositories of photographs uploaded on the social networking websites. We have evaluated our proposed method using 64-bit MATLAB 2013a, 8GB of RAM running on the Windows 8.1 OS with i7 5th generation processor.

### 4.2. Datasets

The WANG database [39] is a subset of 1,000 images of the Corel stock photo database which have been manually selected and it forms 10 classes of 100 images each. It is shown in Figure 9. The WANG database can be considered similar to common stock photo retrieval tasks with several images from each category and a potential user having an image from a particular category and looking for similar images. The 10 classes are used for relevance estimation: given a query image, it is assumed that the user is searching for images from the same class, and therefore the remaining 99 images from the same class are considered relevant and the images from all other classes are considered irrelevant. Figure 10 shows Oliva and Torralba (OT-Scene) dataset (8 categories, 2688 images). The OT-Scene Database was also considered for the evaluation purpose [41]. Figure 11 displays the example images of Monuments dataset. The Most important task is the collection of data as no direct dataset is available for the task at hand. In our data set 10 different classes are considered, namely: 'Taj Mahal', 'Gate way of India', 'Golden Temple', 'Lotus Temple', 'India Gate', 'Qutub Minar', 'Red Fort', 'Chatra Pati Shivaji Railway Station', 'Hawa Mahal' and 'Victoria Memorial'. For each mentioned class, 50 images are collected from websites and a total of 500 images are collected. An individual tourist, family and tourist groups in various poses are considered for both training and testing. To achieve diversity, Front view (day), Near View (night), Far View, Left-View (45 degree), right-side view (45 degree), people posing (sitting or standing etc.) in front of monuments are considered.



Figure 9. The Wang Sample Database [18]





Figure 10. The OT-Scene Database [18]

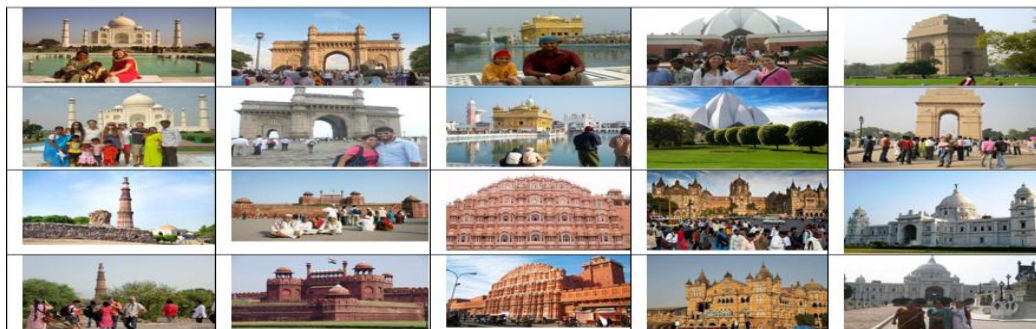


Figure 11. The Sample Images from the Monuments Database

**4.3. Experimental Results**

The performance of the system is evaluated based on Error Rate, Precision, Recall, Accuracy and F-Score [40]. Table 5 shows confusion matrix status for Wang and Monuments Dataset while, Table 6 focuses on OT-Scene dataset. Confusion Matrix for the Indian Monuments dataset is shown in Table 4.

$$\text{Precision} = \text{tp} / (\text{tp} + \text{fp}) \tag{7}$$

$$\text{Recall} = \text{tp} / (\text{tp} + \text{fn}) \tag{8}$$

$$\text{Accuracy} = (\text{tp} + \text{tn}) / (\text{tp} + \text{tn} + \text{fp} + \text{fn}) \tag{9}$$

$$\text{F-score} = 2 * ((\text{Precision} * \text{Recall}) / (\text{precision} + \text{recall})) \tag{10}$$

Here, tp indicates true positive, fp represents false positive, fn and tn are false negative and true negative respectively. The F-score is also known as harmonic mean of precision and recall.

Table 4. The Confusion Matrix (Monuments Database)

Monuments										
Taj Mahal	47	0	0	0	0	0	0	0	0	3
Gateway of India	0	41	0	0	3	0	0	0	0	6
Golden Temple	0	0	43	0	0	0	2	0	0	5
Lotus Temple	0	1	0	46	0	0	0	0	0	3
India Gate	0	1	0	0	36	1	0	0	0	12
Qutub Minaar	1	1	0	1	4	39	0	0	0	4
Red Fort	0	1	1	0	0	0	42	0	0	6
CST	0	0	0	0	0	0	1	49	0	0
Hawa Mahal	0	0	1	0	0	0	0	0	48	1
Victoria Memorial	0	0	0	0	0	0	0	0	0	50

Table 5. The Confusion Matrix Status

Monuments	Monuments Dataset				Class Label	Wang Dataset			
	Precision	Recall	Accuracy	F-Score		Precision	Recall	Accuracy	F-Score
Taj Mahal	0.97	0.94	0.99	0.95	Tribals	0.94	0.86	0.98	0.90
Gateway of India	0.91	0.82	0.97	0.86	Sea Beach	0.88	0.82	0.97	0.84
Golden Temple	0.95	0.86	0.98	0.90	Gothic Structure	0.97	0.83	0.98	0.89
Lotus Temple	0.97	0.92	0.99	0.94	Bus	1	0.99	0.99	0.99
India Gate	0.83	0.72	0.95	0.77	Dinosaur	1	1	1	1
Qutub Minaar	0.97	0.78	0.97	0.86	Elephant	0.95	0.91	0.98	0.93
Red Fort	0.93	0.84	0.97	0.88	Roses	1	0.99	0.99	0.99
CST	1	0.98	0.99	0.98	Horses	0.98	0.98	0.99	0.98
Hawa Mahal	1	0.96	0.99	0.97	Mountains	0.98	0.78	0.97	0.87
Victoria Memorial	0.55	1	0.92	0.71	Food	0.61	0.98	0.93	0.75
Average	0.91	0.88	0.97	0.88	Average	0.93	0.91	0.98	0.91

Table 6. The Confusion Matrix Status (OT-Scene Database)

OTScene Dataset	Precision	Recall	Accuracy	F-Score
Coast & Beach	1	1	1	1
Open Country	0.94	0.93	0.98	0.93
Forest	0.96	0.95	0.99	0.95
Mountain	0.97	0.91	0.99	0.94
Highway	0.97	0.93	0.99	0.95
Street	0.96	0.87	0.97	0.91
City Center	0.90	0.87	0.98	0.89
Average	0.93	0.92	0.98	0.92

Figure 12 shows the Receiver Operating Characteristic curve for Monuments Dataset. Similar curves can be easily plotted for the other benchmark databases.

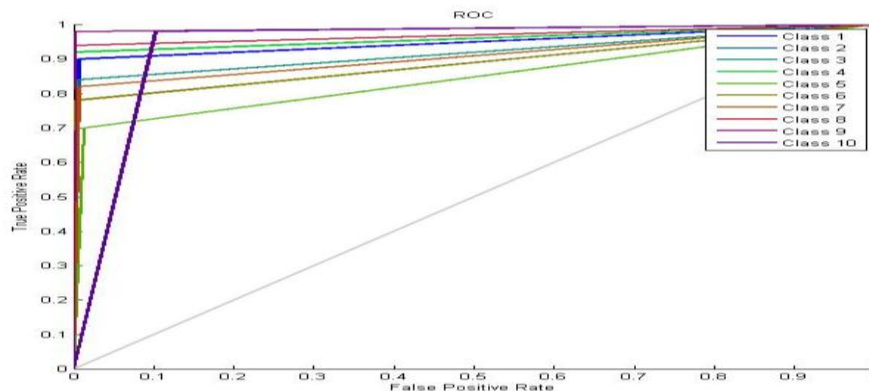


Figure 12. ROC Curve of the Monuments Dataset

5. CONCLUSION

Recently, Content Based Image Classification has generated successful applications in industries like agriculture, pharmaceutical, surveillance and many more. The tourism industry of any country plays a vital role in the economic growth of the nation. The presence of the monuments in the Bollywood movies and its impact on the tourism industry is highlighted as tourists prefer such places to visit. Feature vector is generated using Histograms, Local Binary Pattern, Generalized Co-Occurrence Matrix and Canny-Edge Detector. Ten popular Indian Monuments were considered and image database has been constructed. The

system achieved an average accuracy of 97% with high precision and recall for the Indian monument database. The proposed system also works well on the other benchmark databases.

## REFERENCES

- [1] Niblack CW, Barber R, Equitz W, Flickner MD, Glasman EH, Petkovic D, Yanker P, Faloutsos C, Taubin G, "QBIC Project: Querying Images by Content, using Color, Texture, and Shape", InIS&T/SPIE's Symposium on Electronic Imaging: Science and Technology 1993 Apr 14 (pp. 173-187). International Society for Optics and Photonics.
- [2] Pentland A, Picard RW, Sclaroff S, "Photobook: Content-based Manipulation of Image Databases", *International Journal of Computer Vision*, 1996 Jun 1:18 (3): 233-54.
- [3] Smith JR, Chang SF, "VisualSEEk: A Fully Automated Content-based Image Query System", In Proceedings of the fourth ACM international conference on Multimedia 1997 Feb 1 (pp. 87-98). ACM.
- [4] Marwaha P, Marwaha P, Sachdeva S, "Content based Image Retrieval in Multimedia Databases", *International Journal of Recent Trends in Engineering*, 2009 May.
- [5] Feng D, Siu WC, Zhang HJ, Editors, "Multimedia Information Retrieval and Management: Technological Fundamentals and Applications", Springer Science & Business Media; 2013 Apr 17.
- [6] Gudivada VN, Raghavan VV, "Content based Image Retrieval Systems", *Computer*, 1995 Sep; 28 (9): 18-22.
- [7] Kumar DA, Esther J, "Comparative Study on CBIR based by Color Histogram, Gabor and Wavelet Transform", *International Journal of Computer Applications*, 2011 Mar; 17 (3): 37-44.
- [8] [www.socialbakers.com/facebook-statistics](http://www.socialbakers.com/facebook-statistics)
- [9] Bhavan B, "Central Board of Film Certification: Annual Report", Ministry of Information and Broadcasting, Government of India. 2009.
- [10] [www.india.gov.in/topics/travel-tourism](http://www.india.gov.in/topics/travel-tourism).
- [11] [www.tourism.gov.in](http://www.tourism.gov.in)
- [12] Bhatt M. S., Patalia T. P., "Genetic Programming Evolved Spatial Descriptor for Indian Monuments Classification", In2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS) 2015 Nov 2 (pp. 131-136). IEEE.
- [13] Murala S, Maheshwari RP, Balasubramanian R, "Local Tetra Patterns: A new Feature Descriptor for Content-based Image Retrieval", *IEEE Transactions on Image Processing*, 2012 May; 21 (5): 2874-86.
- [14] Youness C, Mohammed, O, Brahim A, "New Method of Content based Image Retrieval based on 2-D ESPRIT Method and the Gabor Filters", *Indonesian Journal of Electrical Engineering and Computer Science*, 2015 Aug 1;15 (2): 313-20.
- [15] Nazarloo M, Parcham E, Pourani RA. 2014, "Gender Classification using Hybrid of Gabor Filters and Binary Features of an Image", *International Journal of Electrical and Computer Engineering*, 2014 Aug 1; 4 (4): 539
- [16] Guo Z, Zhang L, Zhang D, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification", *IEEE Transactions on Image Processing*, 2010 Jun; 19 (6): 1657-63.
- [17] Thepade S, Das R, Ghosh S, "A Novel Technique in Block Truncation Coding based Feature Extraction for Content based Image Identification", In Transactions on Computational Science XXV 2015 (pp. 55-76),Springer Berlin Heidelberg.
- [18] Das R, Thepade S, Bhattacharya S, Ghosh S, "Retrieval Architecture with Classified Query for Content based Image Recognition", *Applied Computational Intelligence and Soft Computing*, 2016 Jan 1;2016:2.
- [19] Streater J, "Genetic Programming for the Automatic Construction of Features in Skin-Lesion Image Classification", Master's thesis, University of Edinburgh. 2010.
- [20] Kuehne H, Jhuang H, Garrote E, Poggio T, Serre T, "HMDB: A Large Video Database for Human Motion Recognition", InComputer Vision (ICCV), 2011 IEEE International Conference on 2011 Nov 6 (pp. 2556-2563), IEEE.
- [21] Liu J, Luo J, Shah M, "Recognizing Realistic Actions from Videos in the wild", InComputer Vision and Pattern Recognition, 2009, CVPR 2009, IEEE Conference on 2009 Jun 20 (pp. 1996-2003), IEEE.
- [22] Marszalek M, Laptev I, Schmid C, "Actions in Context. In Computer Vision and Pattern Recognition, 2009", CVPR 2009, IEEE Conference on 2009 Jun 20 (pp. 2929-2936), IEEE.
- [23] Laptev I, Marszalek M, Schmid C, Rozenfeld B, "Learning Realistic Human Actions from Movies", InComputer Vision and Pattern Recognition, 2008, CVPR 2008, IEEE Conference on 2008 Jun 23 (pp. 1-8), IEEE.
- [24] Chen L, Rizvi SJ, Ozsu MT, "Incorporating Audio Cues into Dialog and Action Scene Extraction", In Electronic Imaging 2003, Jan 20 (pp. 252-263), International Society for Optics and Photonics.
- [25] Doudpota SM, Guha S, "Mining Movies to Extract song Sequences", InProceedings of the Eleventh International Workshop on Multimedia Data Mining 2011 Aug 21 (p. 2). ACM.
- [26] Vadivel A, Sural S, Majumdar A. K., "An Integrated Color and Intensity Co-occurrence Matrix", *Pattern Recognition Letters*, 2007 Jun 1;28(8):974-83.
- [27] Deselaers T, Keysers D, Ney H, "Features for Image Retrieval: A Quantitative Comparison", In Joint Pattern Recognition Symposium 2004 Aug 30 (pp. 228-236), Springer Berlin Heidelberg.
- [28] Desai P, Pujari J, Ayachit NH, Prasad VK, "Classification of Archaeological Monuments for different art Forms with an Application to CBIR", In *Advances in Computing, Communications and Informatics (ICACCI)*, 2013 International Conference on 2013 Aug 22 (pp. 1108-1112), IEEE.
- [29] [www.imdb.com](http://www.imdb.com)

- [30] Atlam HF, Attiya G, El-Fishawy N, "Comparative Study on CBIR based on Color Feature", *International Journal of Computer Applications*, 2013 Jan 1;78(16).
- [31] Gonzalez RC, Richard E, "Woods, Digital Image Processing", ed: Prentice Hall Press, ISBN 0-201-18075-8. 2002.
- [32] An Y, Riaz M, Park J, "CBIR based on Adaptive Segmentation of HSV Color Space", In *Computer Modelling and Simulation (UKSim)*, 2010 12th International Conference on 2010 Mar 24 (pp. 248-251), IEEE.
- [33] Waikar SM, Khanchandani KB, "Determination of Image Features for Content-based Image Retrieval using Interactive Genetic Algorithm" *IMAGE*, 2014 Mar;89(17).
- [34] <http://in.mathworks.com/help/images/ref/graycoprops.html>
- [35] Heikkilä M, Pietikäinen M, Schmid C, "Description of Interest Regions with Center-symmetric Local Binary Patterns", In *Computer Vision, Graphics and Image Processing*, 2006 (pp. 58-69), Springer Berlin Heidelberg.
- [36] Meena K, Suruliandi A, Rose RR, "Face Recognition based on Local Derivative Ternary Pattern", *IETE Journal of Research*, 2014 Jan 2;60(1):20-32.
- [37] Syam B, Victor JS, Rao YS, "Efficient Similarity Measure via Genetic Algorithm for Content based Medical Image Retrieval with Extensive Features", In *Automation, Computing, Communication, Control and Compressed Sensing (iMac4s)*, 2013 International Multi-Conference on 2013 Mar 22 (pp. 704-711), IEEE.
- [38] Han J, Pei J, Kamber M, "Data Mining: Concepts and Techniques", Elsevier; 2011 Jun 9.
- [39] Wang's Dataset <http://wang.ist.psu.edu/>
- [40] Manimala S, Hemachandran K, "Image Retrieval-based on Color Histogram and performance Evaluation of similarity Measurement", *Assam University Journal of science & Technology*, 8:94-104.
- [41] OT-Scene Dataset: <http://cvcl.mit.edu/database.htm>

## BIOGRAPHIES OF AUTHORS



**M. S. Bhatt** was born at Bhuj in Gujarat, India on 20th February, 1981. Prof. Bhatt completed Bachelor of Engineering with specialization in Computer Engineering from C.U. Shah College of Engineering & Technology, Surendranagar, Gujarat, India in the year 2002. Prof. Bhatt completed Masters in Engineering with computer engineering as specialization from Dharmsinh Desai University, Nadiad, Gujarat, India in the year 2004. Prof. Bhatt is Pursuing Ph.D. in Computer Engineering from Rai University, Ahmedabad, Gujarat. The author's major field of study covers Image Processing and Multimedia Information Retrieval. He has been working as Associate Professor in Computer Engineering Department at Dharmsinh Desai University, Nadiad, Gujarat, India since 2008. He delivered expert talks and acted as reviewer in international conferences. He has also presented papers in international journals, IEEE conferences, National Journals and National Conferences.



**Tejas P. Patalia** was born at Rajkot in Gujarat, India on 14th July, 1978. Dr. Patalia completed Bachelor of Engineering with specialization in Computer Engineering from Sarvajani College of Engineering & Technology, Surat, Gujarat, India in the year 2002. Dr. Patalia completed Masters in Engineering with computer engineering as specialization from BVM Engineering College, V. V. Nagar, Gujarat, India in the year 2009. Dr. Patalia achieved Ph.D. in Computer science and Engineering from Singhania University in the year 2012. The author's major field of study covers Wireless Sensor Network and Information Retrieval. He has been working as Head of Computer Engineering Department in V.V.P. Engineering college, Rajkot, Gujarat, India since 2003. He has received research grants from funding agencies like GUJCOST and Gujarat Technological University. He has presented many papers in reputed international journals, IEEE conferences, National Journals and National Conferences.