Performance Analysis of Efficient Content Based Image Retrieval Frameworks

[1]P. Anandababu, [2]M. Kamarasan
[1]Research Scholar, Department of Computer and Information Science, Annamalai University
[2]Assistant professor, Department of computer and Information Science, Annamalai University.

***Abstract*— Presently, the massive growth in the generation of digital images has initiated the requirement for the enhancement in the way of identifying and retrieving images from a huge dataset. A major issue exist in the process is the way to retrieve the essential images from a huge dataset with high precision with less amount of retrieval time. The choice of proper features acts as a vital part in the enhancement of the results of image retrieval models. Numerous image retrievals have been presented and exist in the literature. In this paper, a detailed comparative study between two image retrieval models namely Fused Information Feature (FIF) and color, texture and edge features (CTE) based image retrieval model. To analyze the better retrieval performance of these two models, a detailed experimentation takes place on the benchmark Corel 10-K dataset. The experimental analysis ensured that the FIF model is found to be an efficient model which is proved by attaining a maximum average precision value of 83.33% whereas the CTE model achieves slightly lower performance by obtaining an average precision value of 82.22%.**

Keywords*— *CBIR; Corel 10K, Image Retrieval; Precision

I. Introduction

In recent days, digital images become an essential part in diverse applications like medical images, satellite images, forensic images, entertainment, and so on. These distinct applications necessitate the digital images as a source for a variety of tasks such as segmentation, object identification, tracking, etc. For indexing and searching appropriate images from the exponential increase in the generation of images, an efficient image retrieval model is introduced. There are generally two ways to index and search an image from the dataset such as keywords and visual information about the images. The conventional Text-Based Image Retrieval (TBIR) model make use of keywords to retrieve the images in a semantic way. However, this model is inefficient to manage huge datasets due to the fact it make use of non-automated generation of keys. In addition, it mainly depends on the view of experts generally applied in the process of keyword creation. It mostly results in improper way of generating keywords for an image.

On the other hand, in Content-Based Image Retrieval (CBIR) system, the visual information related to an image is used to provide resemble images from massive dataset of images are employed to offer reasonable outcome from the huge dataset. From the various visual characteristics, color is an essential element in the view of a human by offering a pleasing view of surroundings. It could be utilized to recognize an object and differentiate an object from other ones. Therefore, in CBIR, color is treated as an important descriptor because of its easiness in computation and non-changing nature irrespective of translation, rotation and alteration in the viewing angles. By the utilization of global and local methods, the color details can be filtered from an image.

Many conventional global color extraction techniques utilize color histogram that offer a number of pixels under every color in image. Although the color histogram is importantly applied for labeling the color details of an image, it has errors in the representation of spatial details as it offers identical histogram definition for wholly diverse image when the color distribution is same. Resembling colors or textures could be used as a descriptor in image retrieval scheme. At present models, texture extraction is a common method which follows the local template model to mine the texture details. The texture details are mostly determined in a local way from the grayscale demonstration of an image. The objects shape is also a most important component while recognizing comparable images from the dataset.

CBIR model is depending upon the mapping among any one of the visual characteristics and related images which holds a large semantic gap. The mainly recognizable arrangement in feature level fusion is the color-texture fusion. In addition, the texture-shape and shape-color based models are also highly designed. However, they show its inefficiency to satisfy the user needs because of not having the capability of the feature fusion models to interlink particular details exist in the inquiry image with the dataset images.

A Gray Level Co-occurrence Matrix (GLCM) feature is utilized to offer precise retrieval performance in [1]. At this point, the local as well as global histograms are validated over the HSV color space image. Multi-scale edge field approach to retrieve multimedia content applied [2] Canny edge extraction as a portion of processing for obtaining the boundary of the object in diverse scales.

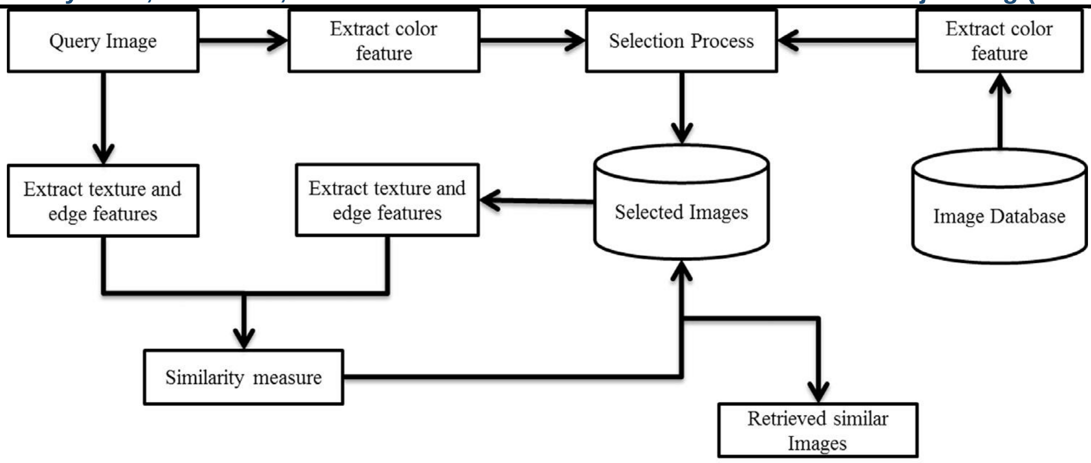

Fig. 1. Basic Flow of CTE model

Agarwal et al. [3] employed Canny edge identification for improving the retrieval performance. Liu and Yang [4] presented Color Difference Histogram (CDH) on Lab color space which is entirely varied from color histogram model. It is recommended to estimate the CDH due to the fact that it makes use of the color variation among the texture information of the image. Consequently, Canberra distance measure is applied for assess the resemblance among the inquiry and dataset images.

Numerous image retrievals have been presented and exist in the literature. In this paper, a detailed comparative study between two image retrieval models namely Fused Information Feature (FIF) [5] and color, texture and edge features (CTE) based image retrieval models [6]. To analyze the better retrieval performance of these two models, a detailed experimentation takes place on the benchmark Corel 10-K dataset. The experimental analysis ensured that the FIF model is found to be an efficient model which is proved by attaining a maximum average precision value of 83.33% whereas the CTE model achieves slightly lower performance by obtaining an average precision value of 82.22%.

II. Proposed Work

A. *Color, texture and edge(CTE) features based image retreival model*

The CTE model intends to reduce the computation load of the traditional CBIR model by minimizing the search region of identical images with no sacrifice in retrieval performance. It utilizes the process of filtering the local features like colors, textures and edges as shown in Fig. 1.

1) Color descriptor

To minimize the complexity and improve the CBIR efficiency, global color descriptor is applied in the beginning stage. It makes a comprehensive usage in diverse applications because of its uniform nature. In addition, it attains a quicker response for a particular request that is moderately high compared to local color filteration models utilized in the image. Therefore, the color moments are selected to define the color information related to the image. It provides the details of the pixel distribution of the images in two easier ways. The first order moment provides average details related to the dispersion of the pixels in the provided image and the nearness of the dispersion of the pixels regarding the mean color determined by the second order moment. In addition, the use of color features in the CTE model varies from the traditional CBIR model.

2) Texture descriptor

Texture is also an important descriptor in a CBIR system. Because of its easier design and efficient results in the domain of texture investigation, presently, LBP is commonly employed in diverse application. The CTE model utilizes a easy LBP [19] for extracting the textures. This technique is carried out on the subset of images chosen from the initial level of the retrieval procedure. Through the exploration of LBP on chosen images, RGB to grayscale conversion takes place. By the use of the threshold values, the binary depiction of that sub block is generated. Next, the LBP value of the 3 × 3 sub-block is validated in the anti-clockwise direction. At the end, the LBP value gets updated in the middle pixel location of the blocks in the image.

3) Edge descriptor

In general, edges are generated by significant variation in the intensity values of the images gathered by the edge detection techniques and it hold boundary representation of the objects exist in the image. Canny method on the chosen images is employed for representing the shape of the images at the end of the first stage. In RGB color space, every color channel is extremely interlinked to additional color channels. Grayscale details are needed to indicate the edges present in the images. However, the transformation of grayscale to RGB is impossible to generate the color images. So, the color space alteration needs to be carried out to attain the edge information from the intensity plane of an image.

B. *FIF method for image retrieval*

Feature Extraction is a major process in the CBIR model which includes the way of filtering out the image features to retrieve the images precisely. The features extracted from the inquiry images and images from the dataset are 8-explained here.

1) $8D-GLCM$ texture features

GLCM is a statistic approach employed to filter the texture feature from the images. The $GLCM\ (G[a,b])$ is derived in 8 angles with the pixel distance of 1 is utilized here. The GLCM matrix is a square of size Ng (grayscale count in the image).

Energy (E): It determines the number of repetitive pair of pixels. When the natural event of repetitive pair of pixels are high, the value of E will also be high and is defined in Eq. (1).

$$\mathrm{E} = \sum_{a=1}^{N_g}\sum_{b=1}^{N_g} G[a,b]^2 \qquad (1)$$

Contrast (C): It computes the local contrast of the image. The value is lower in case when the grayscale values of every pair of pixels is identical.

$$\mathrm{C} = \sum_{a=1}^{N_g}\sum_{b=1}^{N_g}(i-j)^2 G[a,b] \qquad (2)$$

Homogeneity (H): Determines the level of smoothness of an image. The value will be higher in case the grayscale values of every pair of pixels are identical.

$$\mathrm{H} = \sum_{a=1}^{N_g}\sum_{b=1}^{N_g}\frac{G[a,b]}{(a-b)^2} \qquad (3)$$

Correlation (C): It determines the linear dependence of grayscale levels on those of adjacent pixels

$$\mathrm{C} = \sum_{x=1}^{N_g}\frac{(a-\mu_a)(b-\mu_b)G[a,b]}{\sigma_a\sigma_b} \qquad (4)$$

where, μ_a, μ_b is the GLCM mean, σ_a, σ_b is the variance of the intensities of every reference pixel in the relationship which is related to the GLCM.

2) Geometric shape features (GSF)

The shape of the image generally represents the boundary of an object. The procedure involves in the extraction of the shape features are given below:

- Transform the pre-processed image (I) into a binary image (BI) through thresholding.
- Determine the connected components (CC) in BI.
- Generate a label matrix (L) from CC.

Determine the Convex area, Centroid, area, Orientation Euler number, Perimeter, Eccentricity, from every labeled region in the label matrix (L).

These uncomplicated geometric shape characteristics act as an important part in the description of the shape.

3) HSV color moments $(HSVCM)$ *features*

Color feature is a generally employed commonly occurring feature due to the fact it is easier to compute and is non-changeable towards scaling, rotating and translation. It is highly resistive to noises, variation in sizes of the image, direction and resolutions. The Global Descriptor is applied for reducing the complexity level and maximizing the performance while retrieving images. A color moment which symbolizes the color information of the whole image could be computed from any color space.

4) Fused features

$8\mathrm{D}-\mathrm{GLCM}$, GSF *and* $HSVCM$ features results to a 32, 8, 9 dimensional vectors respectively. These 3 attributes feature vectors are fused as follows:

$$8\mathrm{D}-\mathrm{GLCM}+\mathrm{GSF}+\mathrm{HSVAM}(\mathrm{F_{xyz}}) = \{\mathrm{f_{x1}, f_{x2}, f_{x3}} \ldots\ldots \mathrm{f_{y1}, f_{y2}, f_{y3}}, \ldots. \mathrm{f_{y8}, f_{z1}, f_{z2}, f_{z3}}, \ldots. \mathrm{f_{z9}}\} \qquad (4)$$

$$8\mathrm{D}-\mathrm{GLCM}+\mathrm{GSF}(\mathrm{F_{xy}}) = \{\mathrm{f_{x1}, f_{x2}, f_{x3}} \ldots\ldots \mathrm{f_{y1}, f_{y2}, f_{y3}}, \ldots. \mathrm{f_{y8}}\} \qquad (5)$$

$$\mathrm{GSF}+\mathrm{HSVCM}(\mathrm{F_{sc}}) = \{\mathrm{f_{y1}, f_{y2}, f_{y3}}, \ldots. \mathrm{f_{y8}, f_{z1}, f_{z2}, f_{z3}}, \ldots. \mathrm{f_{z9}}\} \qquad (6)$$

$$\mathrm{FIF}-\mathrm{IRS}(\mathrm{F_{tc}}) = \{\mathrm{f_{x1}, f_{x2}, f_{x3}} \ldots\ldots \mathrm{f_{t32}, f_{z1}, f_{z2}, f_{z3}}, \ldots\ldots \mathrm{f_{z9}}\} \qquad (7)$$

where, F is the derived fused attribute vector.

III. Performance Validation

The simulation is performed on the benchmark Corel-10K image [7] dataset which contains a collection of 10000 images correspondingly.

Fig. 2. Sample Test Images

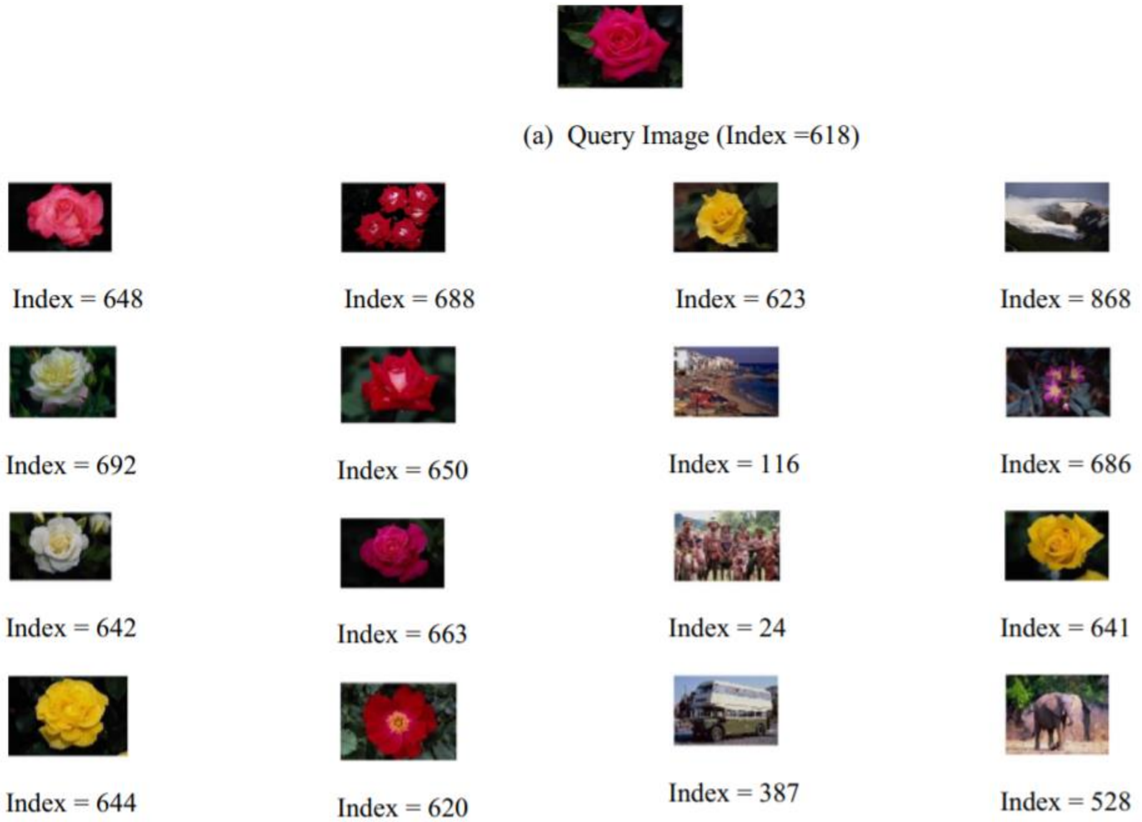

Fig. 3. Sample retrieved results for ‘Rose’ as query image on Corel Dataset

The image is provided in JPEG formation with the pixel dimensions of 126 ×187 or 187 ×126. This dataset holds diverse kinds of images starting from animals to natural images to sports to food. This dataset highly satisfies the needs to validate an image retrieval model because of its larger sizes and heterogeneity. Some sample test images are provided in Fig. 2 and it shows a sample image from every class in the dataset include bus, horse, elephant, rose, food, animal, African, beach, mountain and building.

Fig. 3 illustrates the collection of images retrieved from the applied Corel image dataset on giving a query image of 'Rose'. From the figure, it is apparently clear that upon giving a query image of index 618, numerous similar images are retrieved and some of them are only shown there.

Table 1 provides a detailed comparative study of various methods on the applied identical Corel dataset under a set of 10 classes. Under the class 'African', the maximum retrieval is attained by the FIF model with the precision value of 82 whereas the CTE shows near optimal retrieval with the precision value of 81. The compared CTS model also tried to show better retrieval with the precision value of 80.50. However, the other three methods namely smart CBIR, ENN and fusion model shows poor retrieval with the precision value of 55.50, 53 and 41.25 respectively.

Under the class 'Beaches', the maximum retrieval is attained by the fusion model with the precision value of 71 whereas the CTE and Smart CBIR shows near optimal retrieval with the precision value of 66. The compared FIF model also tried to show better retrieval with the precision value of 60. However, the other methods namely ENN and CTS models show poor retrieval with the precision value of 46 and 56 respectively.

Similarly, under the class 'Buildings', the maximum retrieval is attained by the CTE model with the precision value of 78.75 whereas the FIF model shows near optimal retrieval with the precision value of 67. The compared ENN model also tried to show better retrieval with the precision value of 59. However, the other methods namely smart CBIR, CTS and fusion models show poor retrieval with the precision value of 53.50, 48 and 46.75 respectively.

Likewise, under the class 'Buses', the maximum retrieval is attained by the CTE model with the precision value of 96.25 whereas the FIF model shows near optimal retrieval with the precision value of 95. The compared smart CBIR model also tried to show better retrieval with the precision value of 84. However, the other methods namely smart ENN, CTS and fusion models show poor retrieval with the precision value of 73, 70.50 and 59.25 respectively.

TABLE I

COMPARATIVE RESULTS OF DIFFERENT METHODS WITH RESPECT TO PRECISION

Dataset	Smart CBIR [8]	ENN [9]	CTS [10]	Fusion model [11]	CTE	FIF
African	55.50	53.00	80.50	41.25	81.00	82.00
Beaches	66.00	46.00	56.00	71.00	66.00	60.00
Buildings	53.50	59.00	48.00	46.75	78.75	67.00
Buses	84.00	73.00	70.50	59.25	96.25	95.00
Dinosaurs	98.25	99.75	100	99.50	100	100
Elephants	63.75	51.00	53.75	62.00	70.75	95.00
Flowers	88.50	76.75	93.00	80.50	95.75	100
Horses	87.25	70.25	89.00	68.75	98.75	100
Mountain	48.75	62.50	52.00	69.00	67.75	63.00
Food	68.75	70.75	62.25	29.25	77.25	71.00

Similarly, under the class 'Dinosaurs', the maximum retrieval is attained by the CTS, CTE and FIF models with the precision value of 100 whereas ENN model shows near optimal retrieval with the precision value of 99.75. The compared Fusion model also tried to show better retrieval with the precision value of 99.50. However, the smart model shows poor retrieval with the precision value of 84.

In the same way, under the class 'Elephants', the maximum retrieval is attained by the FIF model with the precision value of 95 whereas CTE model shows near optimal retrieval with the precision value of 70.75. The compared Smart CBIR model also tried to show better retrieval with the precision value of 63.75. However, the ENN, CTS and Fusion models shows poor retrieval with the precision value of 51, 53.75 and 62 respectively.

Similarly, under the class 'Flowers', the maximum retrieval is attained by the FIF model with the precision value of 100 whereas CTE model shows near optimal retrieval with the precision value of 95.75. The compared CTS model also tried to show better retrieval with the precision value of 93. However, the Smart CBIR, ENN and Fusion models shows poor retrieval with the precision value of 88.50, 76.75 and 80.5 respectively.

Likewise, under the class 'Horse', the maximum retrieval is attained by the FIF model with the precision value of 100 whereas CTE model shows near optimal retrieval with the precision value of 98.75. The compared CTS model also tried to show better retrieval with the precision value of 89. However, the Smart CBIR, ENN, and Fusion models show poor retrieval with the precision value of 87.25, 70.25 and 68.75 respectively.

TABLE II

AVERAGE PRECISION OF VARIOUS METHODS

Methods	Average Precision
Smart CBIR	71.43
ENN	66.20
CTS	67.20
Fusion model	60.91
CTE	83.22
FIF	83.33

To clearly understand the overall effective retrieval performance, an average precision analysis is made and the results are shown in Table 2 as well as Fig. 4. From this figure, it can be clearly shown that the maximum retrieval performance is attained by the FIF model with a highest precision value of 83.33.

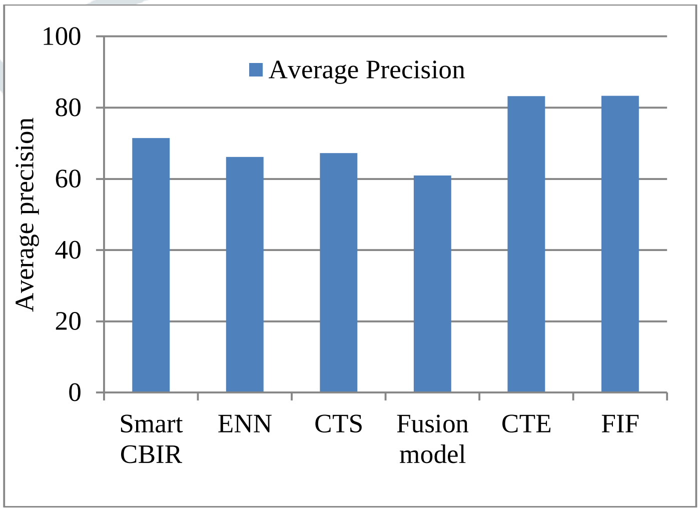

Fig. 4. Average precision analysis of various methods

Next, the near optimal performance is achieved by the CTE model with the higher precision value of 83.22. Then, the Smart CBIR model obtained manageable retrieval performance with the average precision value of 71.43. Then, the ENN and CTS model shows somewhat identical performance with the precision values of 66.20 and 67.20 respectively. However, the fusion model shows worse retrieval results with the average precision value of 60.91. These higher values attained by FIF model proved that it is an effective CBIR model over the compared methods.

IV. Conclusion

Basically, CBIR is used to provide resemble images from massive dataset of images are applied to get reasonable results from the massive dataset. CBIR model is depending upon the mapping among any one of the visual characteristics and related images which holds a large semantic gap. In this paper, a detailed comparative study between two image retrieval models namely FIF and CTE are made. The simulation is performed on the benchmark Corel-10K image dataset. From this experimental outcome, it can be clearly shown that the maximum retrieval performance is attained by the FIF model with a highest precision value of 83.33. These higher values attained by FIF model proved that it is an effective CBIR model over the compared methods.

References

[1] Yue J, Li Z, Liu L, Fu Z. Content-based image retrieval using color and texture fused features. Math Comput Model 2011;54:1121–7. doi: 10.1016/j.mcm. 2010.11.044 .

[2] M, Gabbouj M. Automatic object extraction over multiscale edge field for multimedia retrieval. IEEE Trans Image Process 2006;15(12):3759–72. doi: 10.1109/TIP.2006.881966 .

[3] Agarwal S, Verma AK, Dixit N. Content based image retrieval using color edge detection and discrete wavelet transform. In: IEEE Int Conf on Issues and Chall in Intell Comput Tech (ICICT); 2014. p. 368–72. doi: 10.1109/ICICICT.2014.6781310 .

[4] Liu GH, Yang JY. Content-based image retrieval using color difference histogram. Pattern Recognit 2013;4(1):188–98. http://dx.doi.org/10.1016/j.patcog. 2012.06.001

[5] Bella, M.I.T. and Vasuki, A., 2019. An efficient image retrieval framework using fused information feature. *Computers & Electrical Engineering*, *75*, pp.46-60.

[6] Pavithra, L.K. and Sharmila, T.S., 2018. An efficient framework for image retrieval using color, texture and edge features. *Computers & Electrical Engineering*, *70*, pp.580-593.

[7] Corel-5K and Corel-10K datasets. http://www.ci.gxnu.edu.cn/cbi/dataset.aspx .

[8] Lin, C.H., Chen, R.T. and Chan, Y.K., 2009. A smart content-based image retrieval system based on color and texture feature. *Image and Vision Computing*, *27*(6), pp.658-665.

[9] Murala, S., Maheshwari, R.P. and Balasubramanian, R., 2012. Local tetra patterns: a new feature descriptor for content-based image retrieval. *IEEE transactions on image processing*, *21*(5), pp.2874-2886.

[10] Younus ZS , Mohammad D , Saba T . Content based image retrieval using PSO and K-means clustering algorithm. Arab J Geosci 2014;8(8):6211–24

[11] Fakheri M , Sedghi T , Mahrokh GS . Framework for image retrieval using machine learning and statistical similarity matching techniques. IET Image Process 2013;7(1):1–11 .