Content-based Image Retrieval Using Color Difference Histogram in Image Textures

Armin Ajam
Department of computer engineering
Shahrood Non-profit and Non-government Higher Edu. Inst.
Shahrood, Iran
arminajam71@gmail.com

Majid Forghani
Institute of natural sciences and mathematics
Ural Federal University
Yekaterinburg, Russia
majid.forqani@gmail.com

Mohammad M. AlyanNezhadi, Hamed Qazanfari
Image Processing and Data Mining Lab,
Shahrood University of Technology
Shahrood, Iran
alyan.nezhadi@gmail.com, hamedmit@shahroodut.ac.ir

Zahra Amiri
Image Processing and Data Mining Lab,
Shahrood University of Technology
Shahrood, Iran
zahra.amiri@shahroodut.ac.ir

***Abstract*—The aim of content-based image retrieval system is finding similar images to the query image from a database based on its visual content. In this paper, a novel retrieval system based on human vision is proposed. A factor that has a high impact on the search process is a set of features which are used in. The recent studies emerged that the human eye system considers the image content, texture, and color properties more than other features. Therefore, to retrieve more precisely the images, features should be used that are close to the human eye system. In the current paper, at first, the texture is extracted from the images using the local binary patterns algorithm. After that, the color differences of two adjacent pixels with the same texture are calculated in the HSV color space. Afterward, the histogram is taken from the color difference values. The obtained features from the histogram can describe the visual content of the images in more detail. Finally, the effective features are selected based on their entropy value. The prominent advantage of the proposed method is the lack of implementation of segmentation, clustering, training, and any other method of machine learning, which requires a lot of processing and time. The method is evaluated on two standard Corel 10K and Corel 5K databases, and its retrieval rate is significantly improved compared to some recent methods.**

Keywords— Content-based Image Retrieval, Color Difference, LBP Texture, Color Difference Histogram, Entropy.

I. INTRODUCTION

Due to the rapid growth in the number of digital images, various techniques have been proposed for storing, displaying, searching, and retrieving images that are related to the well-known problem of image retrieval. In image retrieval, an image is searched through a database based on perceptual and visual features, in order to display similar images to the user. Todays, the retrieval of visual information from image databases has become an important research field. In the old methods of image retrieval, a text-based database management system had been used. However, with the rapid growth of storage and data retrieval technologies, especially the digital image database, search and retrieval in the databases have become a big challenge [1].

Since digital images are rich in content without language restrictions to facilitate information expression, the content of an image plays a certain role in search of similar images [1, 2]. Content-based image retrieval (CBIR) is considered as one of the most effective methods for accessing visual data of an image [2]. CBIR has very wide and important applications in many areas, including military affairs, medical sciences, education, architecture, justice, and agriculture. CBIR is looking for similar images to the query digital image based on their content, without the use of textual information, labels or tags. For this purpose, the most conventional way to extract content from an image is using low-level features, such as colors, textures, and image shapes. However, using only these features, the system cannot give an accurate image to describe the perceptual and visual content of the image. In fact, the computer system using low-level features cannot grasp the right content and purpose of an image. In CBIR, the difference between the computer system perception and human one from the image is called the semantic gap. In fact, the biggest challenge in this field is this gap that the researchers are attempting to reduce it. Fig. 1 demonstrates the challenge and several ways to reduce the gap.

As can be seen in Fig. 1, using only low-level features cannot express the visual and perceptual content of the image. In the literature, there are two ways to reducing the semantic gap and understanding the content of an image: a combination of low-level features, and the use of machine learning methods and object recognition. However, most semantic gap reduction methods are associated with many processes for identifying objects and understanding visual concepts. Each of these two methods has its own advantages and disadvantages. For example, methods based on combining low-level features do not require costly operations such as object identification, training, segmentation, clustering, and any machine learning process. In the current study, we attempted to build an efficient

978-1-7281-5350-6/19/$31.00 ©2019 IEEE

CBIR approach that belongs to the methods of combining low-level features.

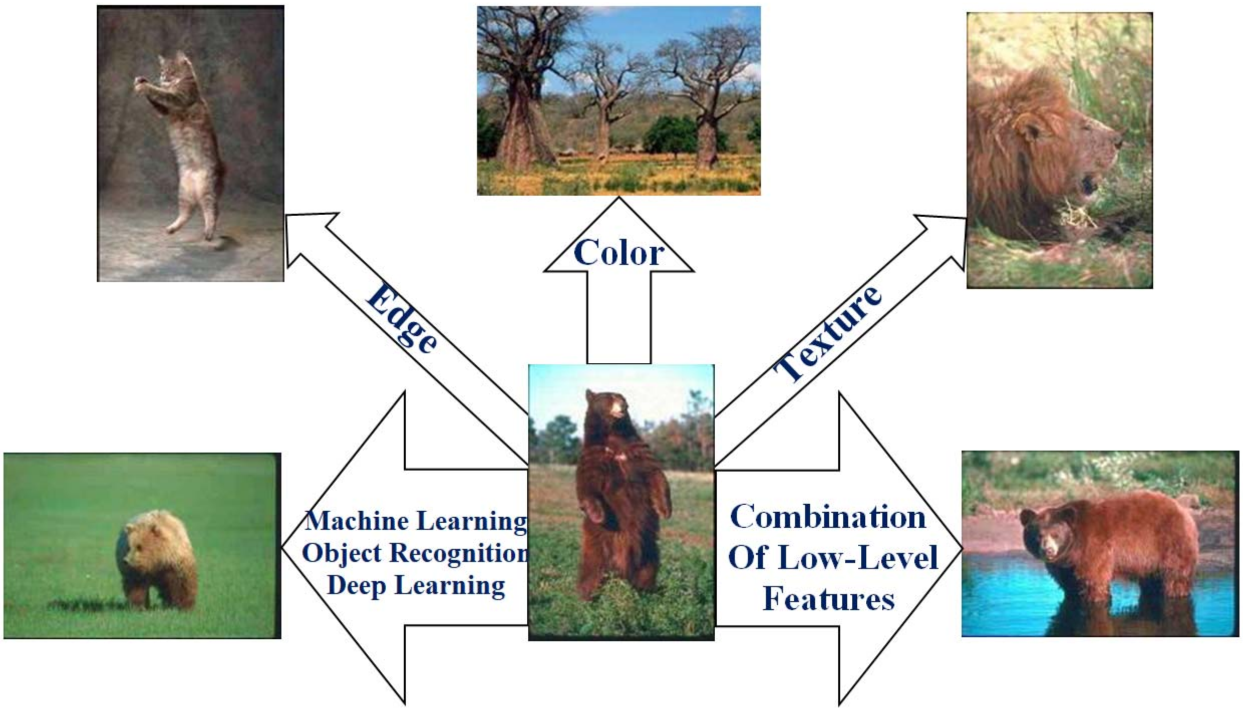

Fig. 1. Applying various methods of retrieval lead to different results in the term of content.

In addition to the two mentioned methods of semantic gap reduction, their combination is also considered for this purpose, which uses the relevance feedback along with the features. Relevance feedback is a multi-level learning method that interfaces between the user and the image retrieval system. In this method, after the first image retrieval, the user will mark the correct retrieved images. Then, using machine learning algorithms, the next retrieved images become closer to the correct images. This process repeats several times that leads to the significant improvement of image retrieval [3].

Examples of most popular machine learning methods in image retrieval are classification of different images with supervised learning methods [3, 4], the use of a Bayesian framework to score the image classes [5], image segmentation [6], deep learning methods to detect objects from different areas of the image [7] and neural networks for pattern learning from samples [8-10].

In the following, several studies which are based on the combining low-level features are briefly described. In [11], the combination of texture and color features is taking into account, while researches in [12, 13] have used the combination of color, texture, and shape features. In [14], a colorful image is first encoded into a string of letters, and then this string code serves as the image feature vector for image retrieval [14]. In [15], to extract the texture, a histogram called the microstructure descriptor (MSD) is applied. The MSD is extracted from the image in the HSV color space based on the underlying color of an edge orientation similarity of adjacent pixels. After creating a microstructure, the color values that match this structure in the image are used as microstructure histogram features in order to retrieve an image.

In [16], a CBIR approach has been presented based on a combination of colors and textures. The textural features include the block difference of inverse probabilities (BDIP), block variation of local correlation coefficients (BVLC), and the color features are extracted from the color histogram in the HSV color space. The BDIP textural feature measures the difference of local brightness in the image. By applying a 2x2 pixels window to the brightness channel, edges, and boundaries of the areas are obtained in order to make features set. The BVLC textural feature measures the level of local texture smoothness. This method is one of the fastest and most accurate CBIR approaches.

The approach presented in [17] uses the edge features and color information in the color space L*a*b*. The color difference between the adjacent pixels and the edges captures a lot of visual information and plays an important role in analyzing and understanding the image content. The color difference between two adjacent pixels, which have the same feature (for example the same texture), is called the Color Difference Histogram (CDH). This approach is based on calculating the color difference of each pixel with neighboring pixels. An overview of image retrieval techniques has been conducted in [18].

Physiological and psychological studies have shown that the human vision system is very sensitive to the color and texture of the image [19, 20]. In this paper, texture and color features are differently used to extract semantic features. The main idea is taken from the CDH proposed in [17]. One of the best algorithms to extract textural features is local binary patterns (LBP) [21]. LBP texture serves as a non-parametric and simple computational descriptor for texture analysis. In the proposed approach, at first, the image texture is extracted. Then, the difference in the color between two adjacent pixels under the same texture is calculated. The histogram of color difference is able to describe and extract the visual and perceptual content of the images in more details. Since the HSV color space is closer to the human vision system, features extracted from this space can better represent the content of the image. Finally, by calculating the Shannon entropy, the number of features is reduced and the superior features are selected. This leads to increase in the retrieval rate and decrease the computation time.

The proposed approach for analyzing the image content, despite the small number of features, improves the time and rate of image retrieval considerably higher than that of MSD [15], CDH [17], and texture and color based approaches (BDIP + BVLC + CH (BBC)) [16].

This rest of this paper is organized as follows. The proposed approach is described in more details in section II. The experiments and evaluation of the method are presented in section III. Section IV gives the conclusions and future works.

II. The Proposed Method

In this paper, a novel method for content-based image retrieval system is proposed. Fig. 2 demonstrates the overall scheme of the proposed method. As can be seen in Fig. 2, the texture of each image is extracted using the LBP algorithm for all images in dataset as well as the query image. Then the color difference of the adjacent pixels which have the same texture is computed and the histogram of these differences is calculated for the images. As mentioned, by computing the entropy of each feature, superior features are selected. Finally, the feature vector is computed for all images in dataset as well as the query image. Using the similarity criterion, the similarity between the vector of the query image and vector of each image in the database is individually computed in order to retrieve the closest images to the query image. In the following, each step of the approach is explained in more details.

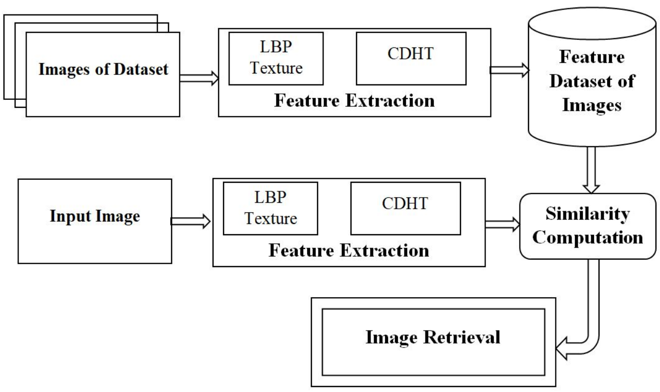

Fig. 2. The overall scheme of the proposed approach for image retrieval.

A. Computation of the image texture with LBP

One of the most powerful methods of extracting image texture is LBP. The LBP operator was introduced as an effective texture descriptor in [21]. This operator produces a binary number for target pixel with respect to the label of 3×3 neighboring pixels. These labels are obtained by thresholding the values of adjacent pixels in associated with the central pixel. When adjacent pixels larger than or equal to the value of the central pixel, the label is set to one, otherwise set to zero. Then, these labels are arranged in a circular way and form an 8-bit binary number. The example of LBP operation is demonstrated in Fig. 3. Then, by moving the window of 3×3 pixels over the image, the texture areas are obtained. After extraction of textural features from the image, obtained values are discretized into 32 levels. The number of levels, i.e. 32, was achieved in the results of experiments which is described in subsection 3.3.

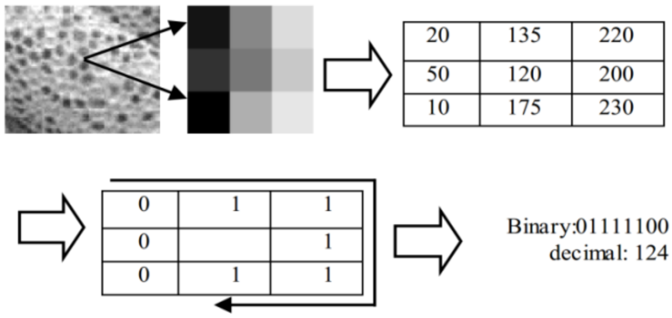

Fig. 3. The demonstration of LBP operation in texture extraction.

B. Perceptually Uniform Color Difference Histogram

As mentioned earlier, physiological and psychological studies have shown that the human vision system is very sensitive to the color and texture of the image [19,20]. The perceptually uniform color difference between image colors and also textures covers a rich variety of visual information and plays a certain role in analyzing and understanding the content of the image.

To calculate this color difference, it should be better to look for a color space that has perceptually uniform features. For this purpose, since the HSV color space is more perceptually uniform and closer to the human vision system, it was considered for feature extraction. Given the high uniformity of the HSV color space, it is also a very good choice of color space to determine the color difference between the two points. In this space, H, S and V can be calculated by standard transform from RGB to HSV [17].

In order to better expression of differences in the HSV color space, the H-channel is uniformly discretized into 15 levels. Both of V and S channels are also discretized into 5 levels. Therefore, this color space has a range of $15 * 5 * 5 = 375$ values. The choice of discretization level for each channel has experimentally been achieved. After the discretization of texture and color of each pixel in the image, feature extraction is performed from adjacent pixels which have one of below condition:

1. The pixels have the same color value.
2. The pixels have the same texture value.

The color value of a point (x,y) is denoted by C(x,y) that can take a value from set M={0,...m-1}. Similarly, the texture value of the point (x,y) using LBP operator is denoted by T(x,y) that can take a value from set N={0,…, n-1}. If the distance between two adjacent points is denoted by D, then the difference in the color and texture can be obtained as follow:

$$H_{color}(C(x,y)) = \begin{cases} \sum\sum\sqrt{(\Delta H)^2 + (\Delta S)^2 + (\Delta V)^2} \\ where\ T(x,y) = T(x',y'); \\ \max(|x-x'|,|y-y'|) = D \end{cases} \tag{1}$$

$$H_{tx}(T(x,y)) = \begin{cases} \sum\sum\sqrt{(\Delta H)^2 + (\Delta S)^2 + (\Delta V)^2} \\ where\ C(x,y) = C(x',y'); \\ \max(|x-x'|,|y-y'|) = D \end{cases} \tag{2}$$

where the dH, dS, and dV are the difference between two points in H, S, and V channels, respectively. Formula (1) expresses that if two adjacent pixels with distance D has the same texture value, their color difference is calculated by this formula and assigned to the color of the pixel (x,y) i.e. C(x.y). Thus, H_color is a vector with 375 features. Similarly, the formula (2) calculates the color difference of two adjacent pixels with distance D, which have the same color value and assigns it to texture value of pixel (x,y) i.e. T(x,y). Therefore, H_tx is a vector of 32 features. Finally, the H_color and H_tx are concatenated to make a final feature vector called H_CDHT, which has 375+32= 407 features (formula (3)).

$$H_{CDHT} = \begin{bmatrix} H_{\text{color}}(0), H_{\text{color}}(1), \ldots, H_{\text{color}}(\text{W-1}), \\ H_{\text{tx}}(0), H_{\text{tx}}(1), \ldots, H_{\text{tx}}(\text{V-1}) \end{bmatrix} \tag{3}$$

By testing various values of the discretization parameters for better image retrieval, the values of 1 and 375 and 32 have experimentally been assigned to the D and m and n.

III. Experiments and Evaluation

In this section, at first, the database and the used similarity criterion are briefly described. Then feature selection, the choice of best values for color and texture discretization, and the evaluation of the proposed approach are explained.

The proposed image retrieval system is implemented using Visual Studio 2015 (C#). This system is available online with the programming code at http://smartcbir.nph-co.ir/.

A. Image database and similarity criterion

The current paper uses two standard Corel 5K and Corel 10K databases with 5 and 10 thousand images, respectively. The classes of databases contain images with relatively difficult content. Databases can be accessed at [22].

One of the important factors in image retrieval, after the feature selection, is the similarity criterion by which the feature vectors are compared. Measurement critera scores the similarity between two images based on their features vectors. Lack of proper criterion leads to decreasing the retrieval rate, even if the extracted features are very significant and close to human perception.

In this paper, the similarity criterion that is introduced in [17] is used for image retrieval. This criterion is an improved version of the Canberra criteria and defined as follows:

$$D(T,Q) = \sum_{i=1}^{M} \frac{|T_i - Q_i|}{|T_i + u_T| + |Q_i + u_Q|} \tag{4}$$

where D is the similarity distance, T and Q are vectors for comparison, and u_T and u_Q are the average of feature vectors T and Q, respectively.

B. Entropy and Elimination of Inefficient Features

In this subsection, using Shannon entropy, features that do not affect the retrieval rate are eliminated. The larger entropy of a feature shown the more important information in the feature. The formula of entropy for vector X is presented as below:

$$En(X) = -\sum_{i=1}^{n} P(x_i)\log_2 P(x_i) \tag{5}$$

in which the value of i-th feature in vector X is denoted by P(x_i).

Applying the Shannon entropy to the proposed 407 features from Corel 10k database gives the chart demonstrated in Fig. 4. As can be seen, many features have zero entropy. This means that these features have no effect on the image retrieval rate and are redundant. So their elimination from feature vector can decrease the computation time and complexity of the algorithm. Of the 407 proposed features, 205 inefficient features are eliminated, and the remaining 202 features are assigned as final features for the proposed approach.

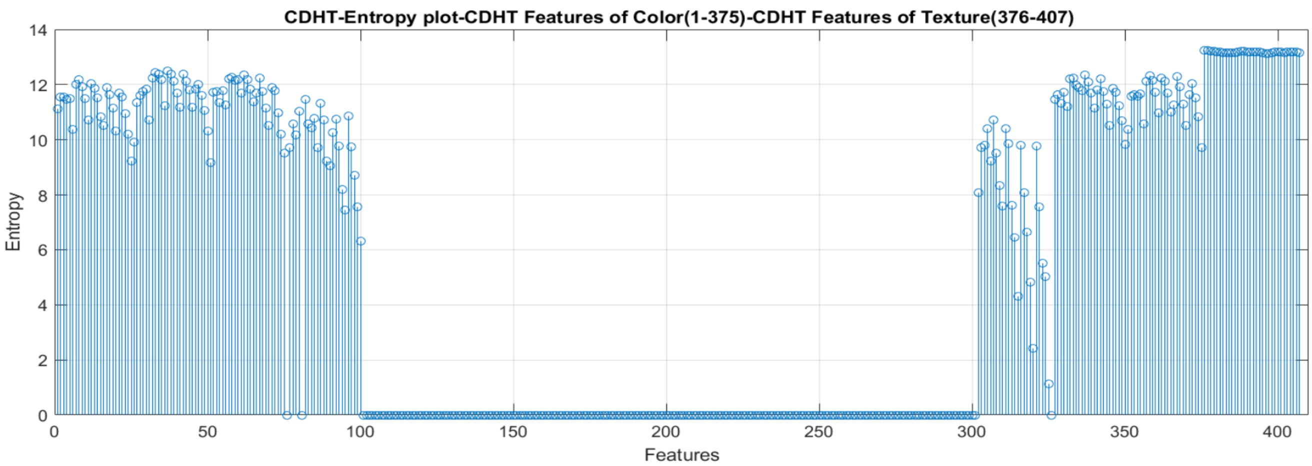

Fig. 4. Shannon Entropy of 407 Proposed Features in CDHT.

TABLE I . Image Retrieval Rate on the Corel 10k Dataset for Different Values of Color and texture discretization.

The quantization number for texture	The quantization number for color, Precision (%)										
	72	90	108	128	160	192	240	300	350	375	540
8	42.78	43.51	43.84	45.66	45.71	45.73	46.22	46.51	46.73	46.91	46.61
16	43.91	44.52	44.81	46.88	46.92	46.98	47.45	47.87	47.92	48.09	47.81
24	43.97	44.84	44.92	47.65	47.72	47.78	48.26	48.36	48.73	48.88	48.84
32	43.87	44.56	44.76	47.71	47.77	47.86	48.36	48.57	48.84	49.13	48.94
64	42.76	43.02	43.25	46.69	46.76	46.80	47.29	47.86	47.96	48.24	47.73
128	40.42	40.88	41.67	44.98	45.06	45.15	45.55	46.01	46.20	46.41	45.87
256	37.33	37.74	38.51	40.23	40.33	40.41	40.89	41.21	41.39	41.68	41.12

C. Best values for color and texture discretization

As discussed earlier, in the proposed method, there are parameters for the discretization of both color and texture values. To find the best value of these parameters, the retrieval rate improvement has been considered in associated with the various assigned values of discretization level in Table 1. In each case, the retrieval rate has been calculated over 20 percent of the Corel 10k database which was randomly selected. As shown in Table 1, the discrete levels of both texture and color have a great impact on the image retrieval rate. For example, as the color discretization level increases, the improvement in terms of retrieval rate also increases, until it reaches to 15, 6, 6 for H, S and V channels, respectively. But in case of texture discretization level, the image retrieval rate totally behaves different and does not follow any decrement or increment rule. The results indicate that the best discretization values for color are 15, 5, 5 for H, S, and V channels, respectively and 32 for texture.

D. Evaluation of the Proposed method

The features 1 to 170 (of the 202 final features) correspond to the CDHT histogram associated with the adjacent pixels with the same color value. The remaining features 171 to 202 correspond to the CDHT histogram of the adjacent pixels with the same texture value. To compare the performance of the proposed approach with those in the recent studies, the methods have been evaluated in randomly selected 20 percent of the images from each mentioned database. Table 2 demonstrates the comparison of obtained image retrieval rate for the proposed approach and three recent methods.

TABLE II. Evaluation of the Proposed Approach and Comparison of Methods

Dataset	Method			
	MSD[15]	**CDH[17]**	**BBC[16]**	**proposed method**
Corel-5K	55.92	57.23	57.5	59.51
Corel-10K	45.62	45.24	47.01	49.13

As can be seen, the proposed method in both databases has better image retrieval rates. It increases the rate by 2 percent in both databases. The results indicate that, as the human vision system uses color differences and texture of images to access their content, the proposed features in the current study have also been able to extract visual features of images and achieve improvement in terms of their perceptual content.

IV. CONCLUSION

In this paper, in order to reduce the semantic gap between the perception of the computer system and human vision system perception, a new method for image retrieval is introduced. Because the texture and color of the images have the greatest impact on the human vision system, the proposed approach uses them to extract the content of images, textures, and colors. At first, after calculating the texture of the image by the LBP operator, the color difference is calculated. The color difference is taking into account between adjacent pixels in two ways: the difference between adjacent pixels with the same texture value, and those with the same color value. Finally, the histogram is taken from the obtained values. As a feature selection procedure, Shannon entropy is applied to the set of features to remove less relevant features and decrease the feature space dimensionality. This leads to the selection of final 202 features which have a great impact on the image retrieval rate. The results indicate that the proposed approach significantly increases the rate of image retrieval. The experiments show that the provided features are able to better describe the two-dimensional space of image without any image segmentation, learning, training and clustering processes. As future work, we want to add textural features obtained by the 2D wavelet transform in order to better presentation of its content in our approach.

REFERENCES

[1] Ghosh N, Agrawal S, Motwani M. A survey of feature extraction for content-based image retrieval system. InProceedings of International Conference on Recent Advancement on Computer and Communication 2018 (pp. 305-313). Springer, Singapore.

[2] Haji MS, Alkawaz MH, Rehman A, Saba T. Content-based image retrieval: a deep look at features prospectus. International Journal of Computational Vision and Robotics. 2019;9(1):14-38.

[3] Qazanfari H, Hassanpour H, Qazanfari K. A short-term learning framework based on relevance feedback for content-based image retrieval. In2017 3rd Iranian Conference on Intelligent Systems and Signal Processing (ICSPIS) 2017 Dec 20 (pp. 136-140). IEEE.

[4] Carneiro G, Chan AB, Moreno PJ, Vasconcelos N. Supervised learning of semantic classes for image annotation and retrieval. IEEE transactions

on pattern analysis and machine intelligence. 2007 Jan 22;29(3):394-410.

[5] Yang W, Yin X, Xia GS. Learning high-level features for satellite image classification with limited labeled samples. IEEE Transactions on Geoscience and Remote Sensing. 2015 Feb 26;53(8):4472-82.

[6] Heller KA, Ghahramani Z. A simple Bayesian framework for content-based image retrieval. In2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) 2006 (Vol. 2, pp. 2110-2117).

[7] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. InProceedings of the IEEE conference on computer vision and pattern recognition 2015 (pp. 3431-3440).

[8] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. InProceedings of the IEEE conference on computer vision and pattern recognition 2015 (pp. 3431-3440).

[9] Meng X, An Y, He J, Zhuo Z, Wu H, Gao X. Similar image retrieval only using one image. Optik. 2016 Jan 1;127(1):141-4.

[10] Xu YY. Multiple-instance learning based decision neural networks for image retrieval and classification. Neurocomputing. 2016 Jan 1;171:826-36.

[11] Rao MB, Rao DB, Govardhan DA. Content based image retrieval using Dominant color and Texture features. International Journal of Computer science and information security. 2011 Feb;9:41-6.

[12] Neelima N, Reddy ES. An Efficient Multi Object Image Retrieval System Using Multiple Features and SVM. InAdvances in Signal Processing and Intelligent Recognition Systems 2016 (pp. 257-265). Springer, Cham.

[13] Singha M, Hemachandran K. Content based image retrieval using color and texture. Signal & Image Processing. 2012 Feb 1;3(1):39.

[14] Jenni K, Mandala S, Sunar MS. Content based image retrieval using colour strings comparison. Procedia Computer Science. 2015 Jan 1;50:374-9.

[15] Liu GH, Li ZY, Zhang L, Xu Y. Image retrieval based on micro-structure descriptor. Pattern Recognition. 2011 Sep 1;44(9):2123-33.

[16] Singh C, Kaur KP. A fast and efficient image retrieval system based on color and texture features. Journal of Visual Communication and Image Representation. 2016 Nov 1;41:225-38.

[17] Liu GH, Yang JY. Content-based image retrieval using color difference histogram. Pattern recognition. 2013 Jan 1;46(1):188-98.

[18] Girija O K, M.Sudheep Elayidom "Overview of Image Retrieval Techniques" International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), International Conference on Recent Trends in Computing and Communication (ICRTCC 2015) Cochin College of Engineering & Technology Vol. 4, Special Issue 1, June 2015.

[19] Livingstone MS, Hubel DH. Anatomy and physiology of a color system in the primate visual cortex. Journal of Neuroscience. 1984 Jan 1;4(1):309-56.

[20] Qazanfari H, Hassanpour H, Qazanfari K. Content-Based Image Retrieval Using HSV Color Space Features. International Journal of Computer and Information Engineering. 2019 Sep 5;13(10):537-45.

[21] Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis & Machine Intelligence. 2002 Jul 1(7):971-87.

[22] Corel Dataset, http://smartcbir.nph-co.ir/datasets.php,last accessed 2019/04/04.