JOURNAL OF COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2676-7465
http://jce.bushehr.iau.ir

Vol. 11/ No. 44/Summer 2022

Feature Extraction Framework for CBIR Systems based on Cyclic Transform Analysis & Spatial Information

Shahin Shafei, MSc [1] | Hamid Vahdati, Assistant Professor [2] | Tohid Sedghi, Assistant Professor [3,4] | Asghar Charmin, Assistant Professor [5]

[1]Department of Electrical Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran, shahin_shafei@yahoo.com

[2]Department of Electrical Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran, h_vadatii@yahoo.com

[3]Department of Electrical Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran, sedghi.tohid@gmail.com

[4]Microwave and Antenna Research center, Urmia Branch, Islamic Azad University, Urmia, Iran, t.sedghi@iaurmia.ac.ir

[5]Department of Electrical Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran, a.charmin@sut.ac.ir

Correspondence
Tohid Sedghi, Assistant Professor of Department of Electrical Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran,
Microwave and Antenna Research center, Urmia Branch, Islamic Azad University, Urmia, Iran,
Email: sedghi.tohid@gmail.com
t.sedghi@iaurmia.ac.ir

Received: 16 February 2022
Revised: 14 March 2022
Accepted: 23 April 2022

Abstract

A novel framework of feature generation for Content based image retrieval (CBIR) is proposed. This system is realized on Cyclic transform Analysis (CTA). It introduces statistical descriptors in the signals frequency domain. Then the CT of data is computed by Semi supervised algorithm (SSA) which is a simple & efficient algorithm. Presented Features are Norm-1 & energy CTA extracted from different sections of bi-frequency plane. This layout illustrates good characteristics of large-scale database. In addition, this manuscript illustrates a novel framework for generating textural and spatial information, and higher retrieval percentages. The textural features extracted with the proposed CTA utilizing first & second moments among the image tiles are so effective in the data processing. Spatial information is extracted utilizing decent field matrix (DFM). After that, moments are computed from DFM to get spatial features. The composition of the textural features and conjunction with the spatial information leads to a fantastic features matrix for retrieval. The experimental results on large-scale database guarantee the method efficiency on all classes of the database with more than 10000 images. For measuring the distance of features a simple matching system based on Minkowski & Canberra distances is introduced. The results are compared with previous scholars and the retrieval percentage is increased more than 10% in comparison with previous systems.

Keywords: Cyclic Transform Analysis, CBIR, Pattern Recognition, Retrieval.

Highlights

- A new feature extraction framework for image retrieval
- This system is designed for rotational conversion analysis
- Uses statistical descriptors in the frequency range of the signal
- The CT is calculated by a semi-regulatory algorithm, which is very simple and effective

Citation: S. Shafei, H. Vahdati, T. Sedghi, and A. Charmin, "Feature Extraction Framework for CBIR Systems based on Cyclic Transform Analysis & Spatial Information," *Journal of Communication Engineering (JCE)*, vol. 11, no. 44, pp. 1–8, 2022.

1. Introduction

Recently, the research has concentrated to fill the gap between low-level & high-level features for semantic concepts. Various systems have been introduced with different methods. Some of the methods have utilized color features and others have utilized local transform based features [1-5]. The latter technique regions the image into different segments on color features. The segments are near to the perception of humans. They are utilized as the blocks for feature calculation and similarity check. These systems are region based image retrieval systems. They have proven to be more efficient in retrieval percentage. In addition, image segmentation has still been an open research field. It is not easy to find algorithms that are equal to the human decision system. To certify inaccurate segmentation, the segment matching system presents an image similarity combining all the segments [6]. Every segment has significance in its dimensions in color image. The segment significance plays an important role in the image matching procedure. A segment is participate more than once in the matching process till its significance is met. The segmentation problems will prevent the shape analysis procedures. Shape features based on spatial information has been widely utilized for retrieval systems [7-9]. Shape descriptors are calculated from invariant moments [10,11]. The researchers studied obviously show that, in CBIR, local features play an important role in specifying the image's similarity. A scale based method are depicted as more effective than techniques based on inaccurate segmentation [11]. The goal of this manuscript is to present two methods: 1) capturing local textural information based on CTA. 2) Spatial data is reproduced from DFM. CTA are statistical parameters that are time periodic. CTA is the cross signal spectral and a frequency shifted version, which provides statistical parameters in the signals frequency domain. CTA Classification has a good characteristic and this method is applied to feature extraction to the Textural Brodatz database [1]. In addition. The robust framework is proposed to capture different image details in different dimensions. A simple matching process based on Minkowski & Canberra distance is provided on similarity check of features. Bi-level framework is utilized for textural analysis with CTA. DFM is utilized to calculate image data, capturing the object information. DFM gives outstanding results in determining the object boundaries. After that, Invariant moments are applied to DFM to introduce shape features. The features combination presents a fantastic feature matrix in CBIR applications. The experimental results are compared with different methods [2,3,5] and they are found to be encouraging.

2. Presented System

Database: Set-D & Set-Q are sets of images. Each of them contains 10000 full color images Set-Q are employed as the query images. The images in Set-D are used as the database images Database which has been utilize for testing presented system are chosen up from sites:
1) http://www.mcsh.kh.edu.tw
2) http://co25.mi.com.tw.
3) http://wang.ist.su.du/IMAGE.
Figure 1 shows some randomly chosen of the database & query.

(a) (b)
Figure 1. Database some examples. (a) Query image, (b) database images corresponding to the image in (a).

Shape Features Decent Field Matrix:
Decent Field Matrix (DFM) is a static method utilized in contour techniques. DFM is calculated as a features vectors of an edge map of images derived from transformation table. It is formulated from a force balance condition. The DFM uses a force balance condition:
The external field is referred as the DFM. The DFM is a parameter that minimizes the energy function:

$$\varepsilon = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2 \, dxdy \qquad (1)$$

When $|\nabla f|$ is large, the second term dominates the integrand, and is minimized by setting V=$|\nabla f|$.
When $|\nabla f|$ is small, the energy is dominated by the sum of squares of the partial derivatives of the vector field, yielding a slowly varying field.
The μ is a regularization parameter. It has tradeoff role between the first & second terms. The DFM gives excellent results on concavities supporting the edge pixels. The algorithm for computation is as follows:
1. Calculate the gradient map of the color image.
2. Filter out only strong edge responses.
3. Converge onto edge pixels satisfying the energy formula.
4. Calculate DFM.
After generating DFM, the following shape descriptors are computed as follows:

$$F_1 = \frac{(\mu_2)^2}{m_1};\ F_2 = \frac{\mu_3}{(\mu_2)^{3/2}};\ F_3 = \frac{\mu_4}{(\mu_2)^2} \quad (2)$$

Where

$$m_r = \frac{1}{N}\sum_{i=1}^{N}[z(i)]^r \quad (3)$$

$$\mu_r = \frac{1}{N}\sum_{i=1}^{N}[z(i) - m_1]^r \quad (4)$$

The z (i) is the set of DFM. In addition, moment's invariant to translation, scale & rotation is taken on R, G & B color images planes. A total of 9 features result from the above computations. Figure 2. Illustrates Result of Object extraction using DFM.

(a) (b)
Figure 2. (a) Main image (b) Result of Object extraction using DFM

3. Proposed algorithm (CTA & Distance Measurement)

Cyclic Transform Analysis **(**CTA) for a signal *x (n)* is defined as:

$$S_x^{\alpha}(f) = \sum_{k=-\infty}^{\infty} R_x^{\alpha}(k) e^{-j2\pi f} \quad (5)$$

Where

$$R_x^{\alpha}(k) = \lim_{N\to\infty} \frac{1}{2N+1}. \sum_{n=-N}^{N}[x(n+k)e^{-j\pi\alpha(n+k)}].[x(n)e^{j\pi\alpha n}]^* \quad (6)$$

Thus $S^{a}_{x}(f)$ is the cross spectrum of the pair of complex valued frequency-shifted signals $x(n)$ $e^{-j\pi\alpha n}$ and $x(n)$ $e^{j\pi\alpha n}$ where f is the cross spectrum frequency variable and the parameter α , called the cyclic frequency, is the relative frequency shift. Multiple algorithms for CTA have been presented:

1) Frequency smoothing method.
2) Time smoothing method.

In this paper, we present one efficient algorithm from time & frequency smoothing category namely semi supervised analysis (SSA) which is different from aforementioned methods. The fundamentals of time & frequency smoothing algorithms are explained. The complete discussion can be found at [1]. All time & frequency smoothing algorithms are based on the time smoothed cyclic analysis:

$$S_{x_T}^{\alpha}(n, f) = \frac{1}{T} < X_T(n, f + \frac{\alpha}{2}) X_T^*(n, f - \frac{\alpha}{2}) >_{\Delta t} \quad (7)$$

The physical interpretation of time & frequency smoothed period-gram is the correlation components of signal over a time span. The components $XT(n, f+\alpha/2)$ & $XT(n, f-\alpha/2)$ are the complex envelopes of band pass, narrow-band components of a signal. They are called complex demodulates. For the calculating of CTA, a data

tapering window slides over the data for a time span. At each instant the complex demodulates of the data within the window are calculated. To produce estimates of the cyclic spectrum function then, they are correlated. Correlation is done by time averaging of conjugate products over an interval of time span. To estimate the spectrum, demodulates separated in frequency by an amount α_0 and centered about a midpoint of f_0 are correlated. Calculation of the complex demodulates is expressed as:

$$X_T(n, f) = \sum_{n=-N'/2}^{N'/2} a(r)x(n-r)e^{-j2\pi f(n-r)T} \quad (8)$$

Where $a(r)$ is a data tapering window of length $T = N'Ts$. Then, the complex demodulates are correlated over a time & frequency span:

$$S_{x_T}^{\alpha_0}(n, f_0)_{\Delta t} = \sum_r X_T(r, f_1)X_T^*(r, f_2)g(n-r) \quad (9)$$

Where $g(n)$ is a data tapering window.

$$S_{X_T}^{\alpha_i+\varepsilon}(n, f_0)_{\Delta t} = \sum_r X_T(r, f_1)X_T^*(r, f_2)g(n-r)e^{-j2\pi\varepsilon rT_s} \quad (10)$$

In SSA evaluation of the sum can be simplified by discretizing the values of ε to be $\varepsilon = q\Delta\alpha$. In this case the output of the algorithm is expressed as:

$$S_{X_T}^{\alpha_i+q\Delta\alpha}(n, f_j)_{\Delta t} = \sum_r X_T(r, f_1)X_T^*(r, f_2)g(n-r)e^{-j2\pi rq/N} \quad (11)$$

In which the sum can be evaluated with an N-point estimation. Thus, point estimates with CA can be computed in blocks. For bi-frequency coverage filters are needed to produce the necessary complex demodulates. An efficient method for producing the required complex demodulates is based on a sliding time & frequency smoothing methods. The frequencies of the filters are discretized to:

$$f_k = k(f_s/N'), k = -N'/2...(N'/2)-1 \quad (12)$$

There are $(N')^2$ estimation segments. Due to symmetry, estimation of the cyclic spectrum of a single real signal requires only one quadrant of the bi-frequency plane. After estimation of CTA by proposed algorithm, feature generation and extraction from CTA is started. One dimensional signals are achieved from each of CTA estimation image by ordering of pixels row by row and column by column. For constructing the feature vector, the Norm-1 & Energy are computed separately on each sections of CTA and then feature vector is formed using these two parameter values. The energy distribution in the frequency-domain identifies a textural information, Resulting feature vectors from each partition are as follows:

$$= [f_{1_j}, f_{2_j}, ..., f_{4N_j}]^T; \text{j=1,2,…9;}\ \overline{f_j} = [E_{11_j}, ..., E_{1N_j}; E_{21_j}, ..., E_{2N_j}; \sigma_{11_j}, ..., \sigma_{1N_j}; \sigma_{21_j}, ..., \sigma_{2N_j}] \quad (13)$$

Respectively, E_{in} & σ_{inj}, are the *Norm*-1 & energy of the CTA that are computed at the different sections. The feature vector length will be equal to $M = 4N$. Since image is a real signal, one quarter of regions makes a complete estimation. For the creation of feature database, the above procedure is repeated for all the images on the database and these feature vectors are stored in the feature database. The distance among two images is calculated as $D=D_1+D_2$, where D_1 is the distance computed by simple matching layout Minkowski distance and D_2 is the distance resulting from shape comparison. Canberra distance is utilized for similarity check. The Canberra distance is:

$$canbdist = \sum_{i=1}^{d} \frac{|x_i - y_i|}{|x_i| + |y_i|} \quad (14)$$

Where x and y are the feature vectors of database and query image. For increasing the accuracy of distance measuring, a minimum cost matching method is introduced. An algorithm for finding the minimum cost matching based on highest priority (MHP) is designed. The distance matrix is computed as an adjacency matrix. The minimum distance is between tiles target. The distance and the row corresponding to tile i and column corresponding to tile j, are blocked. This will prevent tile i of query image and tile j of target image from further participating in the matching procedure. The distances, among i and other tiles of target image and, the distances between j and other tiles of query image, are ignored. This process is repeated till every tile finds a matching. The matching process complexity is reduced from n^2 to n. The minimum cost match distance between is given as follows:

$$D_{qt} = \sum_{i=1,n} \sum_{j=1,n} d_{ij} \quad (15)$$

Where d_{ij} is the best-match distance between tiles i of query image q and tile j of target image t and D_{qt} is the distance between images q and t. Figure 3 illustrates Block diagram of Image Retrieval System. In this flow graph the acronyms are as follows: F.: Features; D.B.: database M&C: Minkowski & Canberra; Alg.: Algorithm. In this flow-graph, there two line for feature generation, first line for Set-Q and second line for Set-D images. A last of

flow graph, the similar images are retrievable images for operators based on Minkowski and Canberra similarity check algorithm.

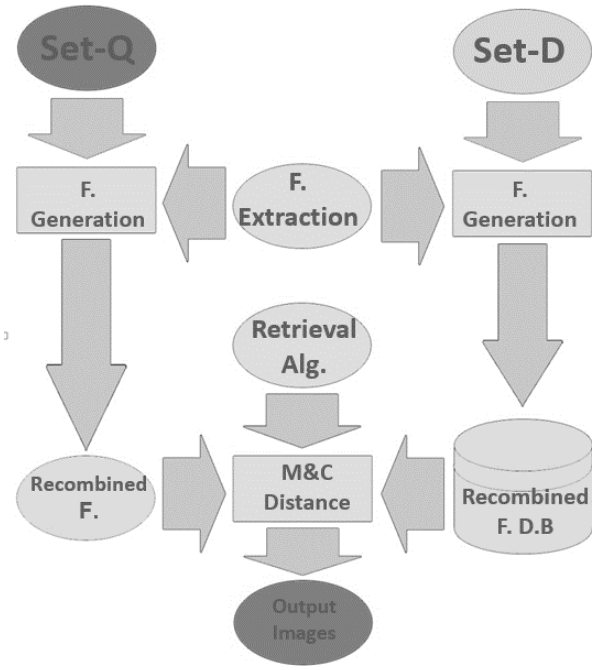


Figure3. Block diagram of proposed CBIR System

4. The Performance of Present System on Large-scale Database & Computation Cost

The proposed system performance is evaluated utilizing database. There are two images sets (Set-D & Set-Q) contains 10000 full color images. The Set-D are employed as the database images and Set-Q are used as the query. Figure 1 shows each image pair which they are randomly chosen from the same animation. Each I^{q}_{i} is utilized as the query image. For each query, the system responds to the operator, images with the shortest image matching distances. If I^{q}_{i} exists among the L database images, the system has correctly found the image. Otherwise, it has failed to find the expected image. The accuracy rate of replying a query will be explained with accuracy percentage. The experiment is to compare the retrieval accuracy of our system with [2,3,5]. Table 1 shows the result of comparison between proposed method and the other researches. The average precision is:

$$p(i) = \frac{1}{100} \sum_{1 \le j \le 100,\ ID(i)=t} 1 \qquad (16)$$

The experimental results comparison of proposed method with other standard systems reported in the scholars [2,3,5] is presented in Table 2. Their results are claimed to be better than method [2,3,5]. In comparison between our method and [5] in most of the classes the presented system has carried out better than all systems. It is clear that this method has achieved a better precision of various images than the other systems. The precision of the retrieval results with the number of returned images, and the average recall are shown in Figure 4 and Figure 5. The experimental results clearly reveal that for the first 20–100 returned images of the 10000 image database, the present method is significantly superior to the previous methods [2,3,5]. In the recall experiment, image retrieval precision increases with the number *L* of returned images.

Table 1. Comparison of accuracy percentage of retrieved image on database

	1	**3**	**5**	**10**	**20**	**30**	**100**
[2]	74.8	76.6	79.0	84.0	87.7	90.2	92.3
[3]	92.3	93.2	93.6	95.1	96.7	97.7	98.9
Present Method	93.9	94.4	94.6	95.7	97	98.1	99

The computational complexity was measured using database that contained more than 10000 images. The feature extraction and retrieval & computing time is presented at Table 3. The computing time is compared to [2] [3]. The computation was made utilizing MATLAB on a PC with Corei3 CPU and 8 GB RAM. There is a good difference between the computational cost of the proposed system and other ones. The reason is that, other methods are iterative algorithms, which makes it heavy, specifically in large databases. Figure 6 illustrates Retrieval image

from large-scale database for Query one image in different classes. It is obvious that for each query image, the results are completely true.

Table 2. Average precision percentage Comparison with proposed method and other standard retrieval systems [2,3,5]

	[2]	[3]	[5]	Present Method
Class1	47%	45%	48%	55%
Class2	35%	35%	44%	40%
Class3	35%	35%	36%	40%
Class4	60%	60%	69%	68%
Class5	95%	95%	96%	99%
Class6	25%	25%	55%	62%
Class7	65%	65%	89%	69%
Class8	65%	65%	70%	77%
Class9	30%	30%	42%	45%
Class10	48%	48%	53%	55%

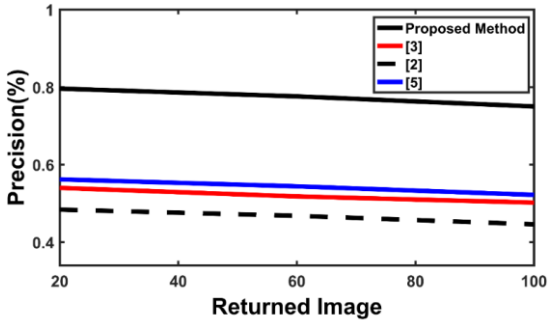


Figure 4. Average precision of different methods

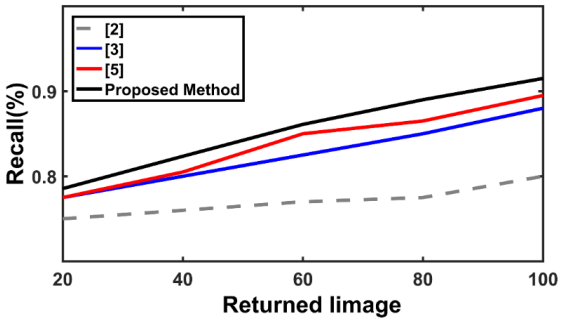


Figure 5. The average recall of these methods

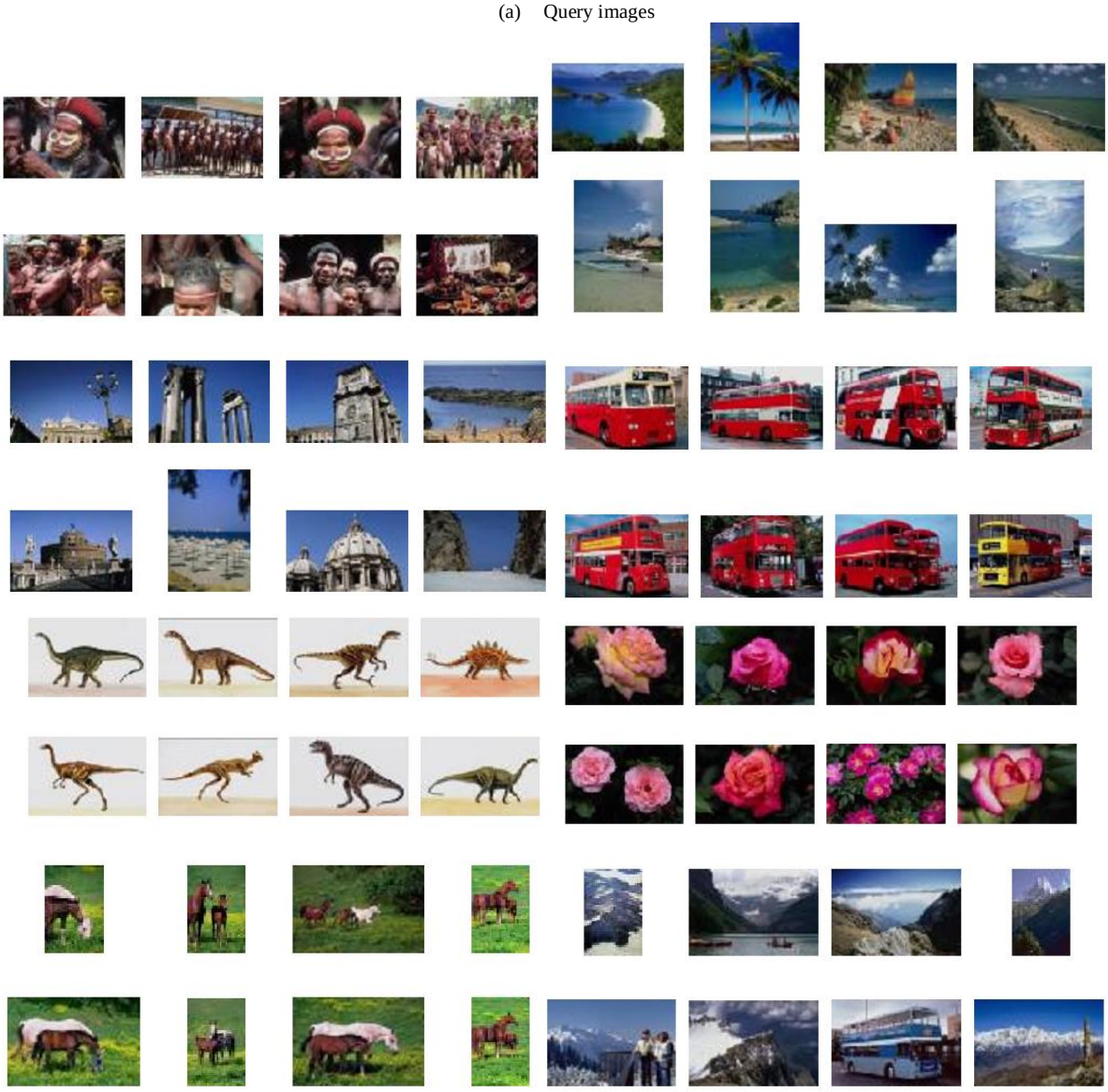
Table 3: Computing time of present method in comparison with [2,3,5]

	Feature Extraction time for 1 image (μs)	Retrieval of 100 image (ms)
[2]	1	2.4
[3]	2.3	3
Present method	≤1	2.4

5. Conclusion

A new systems, for image retrieval using textural & and spatial features within a framework in Cyclic transform domain is proposed. Presented features are generated from bi-frequency at different resolutions. Textural descriptors utilizing CTA from each of non-overlapping tiles calculate Norm-1 & energy in each section of SSA. For image similarity a simple matching layout based on Minkowski & Canberra principles constructed among the image tiles, is implemented. DFM are utilized to extract spatial data of objects in images. Then moments are applied to DFM to describe the spatial features. Textural & shape features composition provides a robust feature matrix for retrieval. The experiments using the database depicts that the proposed method increase the precious of system in comparison with the existing methods in literatures.

(a) Query images



(b)

Figure 6. Retrieval image from large-scale database for Query one image in different classes (a) Query images (b) system output images

References

[1] T. Sedghi, "A Fast and Effective Model for Cyclic Analysis and Its Application in Classification, " *Arab Journal Sci.Eng*, vol.3, pp.927–935, Sep. 2013, doi:10.1007/s13369-012-0364-5.

[2] J. Li, J. Wang, J. Z. Wiederhold, "integrated region matching for image retrieval in proceedings," *the 8th ACM International Conference on Multimedi*a, pp. 147- 156, Oct .2015, doi:10.1145/354384.354452.

[3] F .Baig, Z .Mehmood, M .Rashid, M. Arshad. Javid, A .Rehman, T. Saba, and A. Adnan, "Boosting the Performance of the BoVW Model Using SURF–Co HOG-Based Sparse Features with Relevance Feedback

for CBIR, " *Iranian Journal of Science and Technology Transactions of Electrical Engineering*, vol. 44, p.p 99–118, Aug 2020, doi:10.1186/s13640-020-00516-4.

[4] P. Wu, S. C. H. Hoi, P. Zhao, C. Miao, and Z.-Y. Liu, "Online multi-modal distance metric learning with application to image retrieval, " *IEEE Trans. Knowl. Data Eng*., vol. 28, no. 2, pp. 454-467,Feb. 2016, doi: 10.1145/2502081.2502112.

[5] M.h.Hajigholam, A.A.Raie, and K. Faez, "Using Sparse Representation Classifier (SRC) to Calculate Dynamic Coefficients for Multitask Joint Spatial Pyramid, " *Matchingranian Journal of Science and Technology Transactions of Electrical Engineering*, vol. 45, pp. 295–307, 2021,doi: 10.1007/s40998-020-00351-3.

[6] H. C. Shin , "Deep convolutional neural networks for computer-aided detection CNN architectures, dataset characteristics and transfer learning, "*IEEE Trans. Med. Imag*., vol. 35, no. 5, pp. 1285-1298, May. 2019, doi:10.1109/tmi.2016.2528162.

[7] S.Shafei, H.Vahdati, T. Sedghi, and A. Charmin, "Novel high level retrieval system based on mathematic algorithm & technique for MRI medical imaging and classification, "*Journal of Instrumentation*.,vol. 16, no. 7, pp. 1-14, July. 2021,doi:10.1088/1748-0221/16/07/P07055.

[8] F.Rahdari, E.Rashedi, M.Eftekhari, " A Multimodal Emotion Recognition System Using Facial Landmark Analysis , " *Iranian Journal of Science and Technology Transactions of Electrical Engineering*, " vol. 43, pp. 171–189, Sep. 2019, doi:10.3233/AIC-20063.1.

[9] G Hassan, KM Hosny, RM Farouk, AM Alzohairy, "An efficient retrieval system for biomedical images based on Radial Associated Laguerre Moments" *IEEE Access* vol. 8, pp. 175669- 175687 , Sep. 2020,doi: 10.1109/ACCESS.2020.3026452.

[10] Nazgol Hor, Shervan Fekri-Ershad, "Image retrieval approach based on local texture information derived from predefined patterns and spatial domain information", *International Journal of Computer Science Engineering*, Vol. 8 No.06, pp. 246-254, Dec 2019, doi:10.48550/arXiv.1912.12978.

[11] KM Hosny, RM Farouk, AM Alzohairy G Hassan, "Efficient Quaternion Moments for Representation and Retrieval of Biomedical Color Images", *Biomedical Engineering: Applications, Basis and Communications* vol. 32, pp. 1-16, Oct. 2020, doi:10.4015/S1016237220500398.