International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 7, No.6 (Nov-2018)
http://dx.doi.org/10.12785/ijcds/070604

Arabic Manuscript Content Based Image Retrieval: A Comparison between SURF and BRISK Local Features

Bayan Bagasi[1] and Lamiaa A. Elrefaei[1, 2]

[1] *Computer Science Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia*
[2] *Electrical Engineering Department, Faculty of Engineering at Shoubra, Benha University, Cairo, Egypt*

Received 16 May 2018, Revised 5 Jul. 2018, Accepted 22 Sep. 2018, Published 1 Nov. 2018

Abstract: Arabic manuscripts are worthy sources of knowledge that have been highly underutilized. Because, the vast content of the Arabic manuscript and the need of getting information from them, in a fast, efficient, and accurate way, it is essential to develop a system that supports the retrieval procedure from them. In this paper, a Content-Based Image Retrieval (CBIR) system is proposed to retrieve the Arabic manuscript images. The system has three stages: Preprocessing, feature extraction, and feature similarity matching. The features extraction techniques are the effective step for the performance of CBIR system. For this reason, we propose to apply Binary Robust Invariant Scalable Key points (BRISK) and Speeded-up Robust Feature (SURF) as features extraction techniques. The Hamming distance with BRISK and Sum of square differences (SSD) with SURF are used at the matching stage. The results of proposed system show that for SURF the average Recall is 85% and average Precision is 77%. The average time is 207.3 seconds per image. For BRISK, the average Recall is 69% and average Precision is 68%. The average time is 256.7 seconds per image. The SURF features yield the best performance for Arabic manuscript retrieval. For better time performance of the system we propose to use parallel computing as a future work.

Keywords: Arabic manuscript, Content-Based Image Retrieval (CBIR), Speeded-up Robust Feature (SURF), Binary Robust Invariant Scalable Key points (BRISK)

1. INTRODUCTION

The definition of a manuscript from Harrods's Librarians' Glossary is: "a document of any kind that written by hand or the text of music or literary composition in handwritten or typescript form, and which in that form, has not been reproduced in multiple copies" [1]. An Arabic manuscript is a handwriting document written in Arabic. These documents may contain marginal notes, signs, ink smears, etc., and these are of significant value as shown in the sample manuscript [2] in Fig. 1.

The number of the digitized documents in an image form is increasing for the historical manuscripts [3]. Hence, image retrieval could be used to retrieve Arabic manuscripts. Image retrieval is a technique whereby similar images from a dataset that are visually similar to a given query image can be retrieved. It is a generic technique that can be applied to recover any image using the features of this image. There are two types of image retrieval: the text-based image retrieval and the Content-Based Image Retrieval (CBIR). Text -based image retrieval uses a text description in a retrieval system. CBIR is an automated and efficient system, which can retrieve and rank similar images. CBIR relies on computer vision techniques to solve the problem of searching for the digital image in a large dataset.

Arabic language in a manuscript has specific features such as diacritics, decanters, ascenders, and loops or holes. Also, it has many different morphologies of handwritten words. For these reasons, CBIR is a suitable technique that could be used to retrieve Arabic manuscripts. CBIR is also known as content-based visual information retrieval and query of image content.

CBIR has two main stages: feature extraction and feature matching. The extracted features may be global or local [4]. The global features describe the visual content of the image or the global image properties such as intensity histogram, mean and standard deviation values of pixel distribution. The local features describe the content of image region or specific image properties of local image region such as edges, corners, lines, and curves [5]. The local features are used to detect objects

E-mail: {bbagasi, laelrefaei}@kau.edu.sa, lamia.alrefaai@feng.bu.edu.eg

under occlusions, and they are invariant to scale as well as rotation changes.

Figure 1. An example of Arabic manuscript [2]

In Arabic manuscript retrieval, the interest points or key point's features is a suitable local feature of the image. The most popular feature extraction techniques for a key point are SURF [6] and (BRISK) [7]. These features are used in the matching stage for image matching, similarity, and searching.

In this paper, we propose a CBIR system to retrieve Arabic manuscripts with the following contributions:

1) Collecting Arabic manuscripts image dataset. The collected dataset is classified to 29 different classes based on their sources.

2) Investigating two different local features extraction techniques, BRISK and SURF, for their suitability for Arabic manuscripts retrieval.

The rest of this paper is organized as follows: Section 2 presents the related work on Arabic manuscripts retrieval. Section 3 presents the collected Arabic manuscripts dataset. The proposed Arabic manuscripts CBIR system and its stages is explained in Section 4. Section 5 illustrates the experiments and implementation results. Finally, the conclusion and future work is drawn in section 6.

2. RELATED WORK

There are various techniques that addressed the problem of retrieving Arabic manuscripts, such as Optical Character Recognition (OCR), But the cursive nature of the Arabic script is a challenge for OCR systems [8] [9]. Another technique is word spotting that is defined as a pattern analysis task that consists of finding keywords in handwritten document images. However, word spotting technique is used in semantic content-based image retrieval [10].

The patent of Yahia and Al-Khatib [10] concerned on the semantic content-based image retrieval of Arabic Manuscripts. They focused on Latent semantic indexing (LSI) method for indexing each sub-word in a manuscript. They used segmentation process after preprocessing to segment the sub-word from the image. Otherwise, their system used novel circular polar grid feature set to extract features from sub-word and store the information in the word (image name, etc.) into the database. In the same time, they construct LSI index of the image. They used 34 pages from "Sahih Al Bukhari" to test their system, and their result is 85% of recall, and 80% of Precision. Their work is like a word spotting where a word is used as a match point.

Sari and Kefali [11] used also a word spotting technique. They proposed a system of a search engine of Arabic based on sub-word representation as ASCII codes. Each sub-word has selected a sequence of codes according to its morphological features. Loops, ascenders, decanters, up and down diacritics are the characteristics their method used. Extracting feature phase has three modules: baseline detection, median zone detection, and contour following. When the user writes a query word, the system converts it into the code and matching it with the indexing word dependent code. In their testing they used Arabic manuscripts from different sources using some query word. They got a recall of about 56.62% and a precision approximating 77.78%.

Shahab, et.al. [12] proposed a computer-aided indexing Arabic manuscript system for retrieving an image of the manuscript. They used geometric features (height of sub-word, width of sub-word, and aspect ratio (height/width), Hu's Moments (seven invariant moment descriptors), DFT of sub-word profiles, concentric circle features, and angular lines features. They used Euclidean Distance, Product of Manhattan Distance of features, and Angular separation as match similarity techniques. They apply experiments for different combinations of feature sets and similarity techniques. The experiment of the combination of the concentric circle features, the width feature, the angular line feature and the 15-point DFT have participated in at least one set of features producing a high match rate.

Herzog, et.al. [13] proposed a technique uses CBIR on a pattern from query image to find a matched image. They used Harris corner detector to detect the interesting point in the writing pattern and the query image. Then computed the combining for each interesting point with the structure tensors of all points in the neighborhood into a large feature vector characterizing strength and directionality of intensity gradient. In the matching stage they compared between the interesting point of the query and interesting point of the data construct by the probabilistic model. The target is considered a match of the query if both tests were succeeded. The experiments are conducted on Chinese and Arabic manuscripts. The result of Chinese manuscript includes many correct hits, but one hit is a false positive.

Recently, the feature extraction in CBIR have many robust keypoint techniques which are being used in a wide application of computer vision. The most prevalent methods are Scale Invariant Feature Transform (SIFT) [14] [15][16], Speed Up Robust Features (SURF) [15] [16] [17] [18], Binary Robust Independent Elementary Features (BRIEF) [15] [17] [18], and Binary Robust Invariant Scalable keypoints (BRISK) [15] [16] [17] [18].

While SIFT and SURF are extracting floating point features, BRIEF and BRISK are popular binary features extraction methods. However, an excellent performance technique depends on the domain of application [15].

In this paper, a CBIR system is proposed to retrieve Arabic manuscripts. The two local features extraction techniques, BRISK and SURF, are investigated for their suitability for Arabic manuscripts retrieval.

3. ARABIC MANUSCRIPTS IMAGE DATASET

A dataset of Arabic manuscripts images is collected from 30 books of different subjects containing 1669 images. the dataset is classified into 29 different classes depending on their sources. The images are in the same class if they are from the same book and its other editions if exist, and if they are belonging to the same writer. Each image in the dataset is in RGB Jpg format and labeled by the class number and the image number in the form Class#_Image#. This uniform naming convention facilitates the feature extraction from all images in the dataset.

Fig. 2 shows the 29 Arabic manuscripts classes, the manuscripts book title for each class, the image of the first page of the manuscript and the number of images in each class.

4. THE PROPOSED ARABIC MANUSCRIPT CBIR SYSTEM

In the proposed Arabic manuscript CBIR System, the images in the dataset and the query image are first preprocessed. Then, their SURF and BRISK local features are extracted. The features vectors of the dataset images are stored in the database. Feature similarity between the query image feature vector and the dataset images feature vectors is measured in the matching stage by Hamming distance. The output of the matching is ranked similar images. Fig. 3 shows the block diagram of the proposed system. In the following sub-sections, each stage is explained.

A. The preprocessing Stage

The first step is to transform the image in 256-gray levels. Then, the image is resized to a size of 265*265 pixels. Finally, Otsu's method is utilized to binarize the image, the primary goal of binarization is getting the text and other contents of the manuscript in the foreground [19]. Fig. 4 illustrates the preprocessing steps applied to a sample manuscript image.

Figure 2. Arabic manuscripts images dataset.

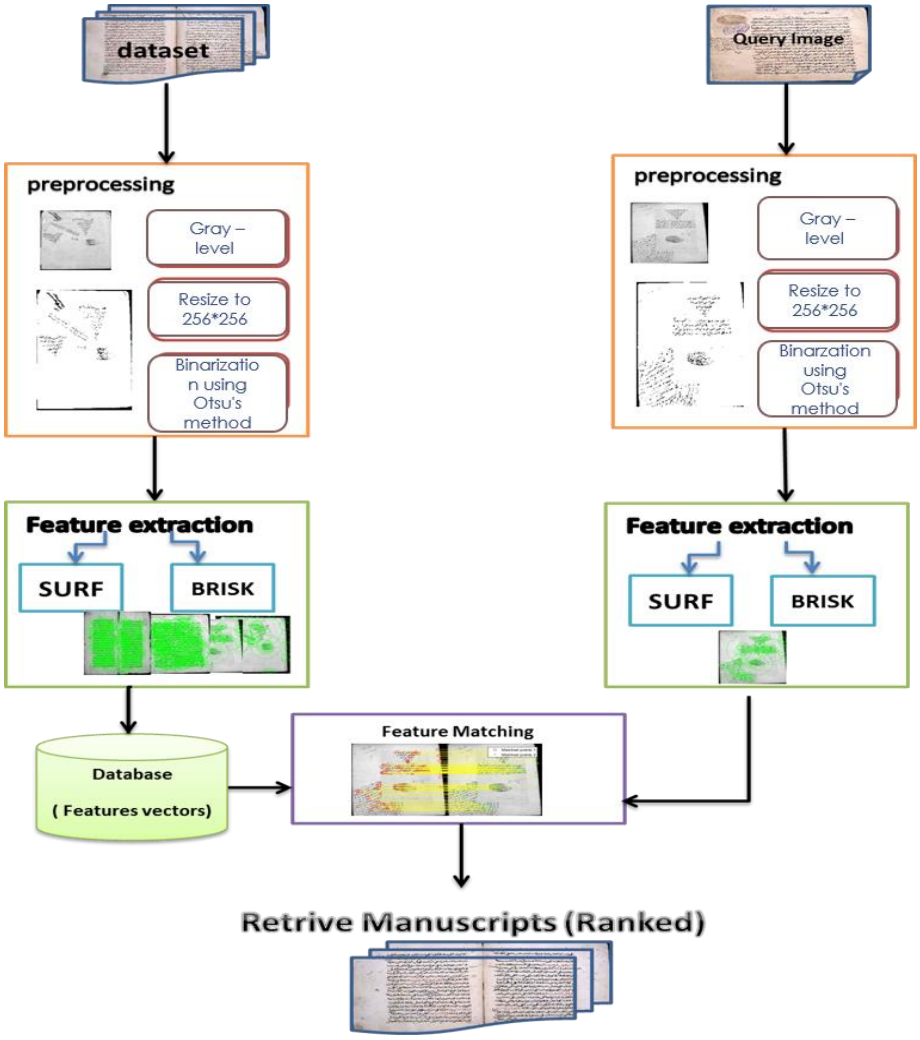


Figure 3. The Block Diagram of The Proposed Arabic Manuscript CBIR System

(a) (b) (c)

Figure 4. The preprocessing steps

(a) Original image (b) 256-gray levels image (c) Otsu's binarized image

B. *The feature extraction*

In this stage, SURF and BRISK local features are extracted from the binarized manuscript image.

SURF is proposed by Herbert Bay [6] as a novel scale- and rotation-invariant Interest point detector and descriptor. It detects the interesting point for the image then produces a set of 64- dimensional descriptors for each interesting point. There are four steps of SURF as described in [20].

BRISK is proposed by Stefan Leutenegger et.al. [7]. It has three main steps: feature detection, descriptor composition, and key point matching.

An example of the extracted features from the manuscript image shown in figure 4(c) is illustrated in Figure 5. The SURF features are shown in Figure 5(a) and they are 265 features for this image. The BRISK features are shown in Figure 5(b) and they are 284 features for this image. BRISK features are more than SURF features and covers many pieces on the manuscript.

C. *The features matching stage*

The Hamming distance is used in the proposed system for BRISK feature matching. The Hamming distance between two vectors is the number of points in which they differ [21]. However, Sum Square of difference (SDD) is sum of square differences between entries of the two descriptors that's suitable with SURF features [22]. Figure 6 shows the matched SURF keypoints between query images and sample images from the dataset. Figure 7 shows the same for the matched BRISK keypoints.

5. EXPERIMENTS AND RESULT

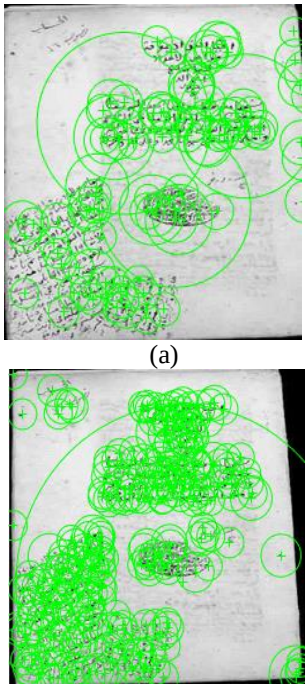
MATLAB 2017 is used as a framework of programming the proposed system. The system runs on a PC with Intel® Core ™ i7 -8550U CPU @ 1.80 GHz 1.99 GHz with Windows 10.

We have two phases: training and testing. First, the system is trained on 90 query images from the collected dataset to rank the similar images. The extracted images are sorted according to the number of similar features N, then we found the maximum number of matched features with an image, N_{max} and the minimum number of matched features which is greater than zero, N_{min}. The retrieved ranked images are images that have matched features from N_{max} to $(N_{max}+ N_{min})/2$. This range is found to give the best results based on the experiments on the 90 query images.

In the testing phase, the performance of the system is evaluated using *Recall*, Equation (1), and *Precision*, Equation (2) [21].

$$Recall = \frac{\text{The number of correct results}}{\text{The number of results that should have been returned}} \quad (1)$$

$$Precision = \frac{\text{The number of correct results}}{\text{The number of all returned results}} \quad (2)$$



(a)

(b)

Figure 5. The extracted Features

(a) SURF features (b) BRISK features.

Fig. 8 shows examples of the retrieval results using SURF and BRISK feature matching for a query image from the collected dataset.

Table 1 and Table 2 show the results of the testing phase using 10 query images with 100 images dataset size when using SURF and BRISK features respectively. Fig. 9 shows a comparison of the system performance using SURF and BRISK features. For SURF, the average Recall is 85% and average Precision is 77%. The average time is 207.3 seconds per image. For BRISK, the average Recall is 69% and average Precision is 68%. The average time is 256.7 seconds per image. Using BRISK takes more time than using SURF as BRISK features are more than SURF features, as illustrated in Fig. 5(b). From Fig. 9, it is clear that the SURF features yield the best performance for Arabic manuscript retrieval.

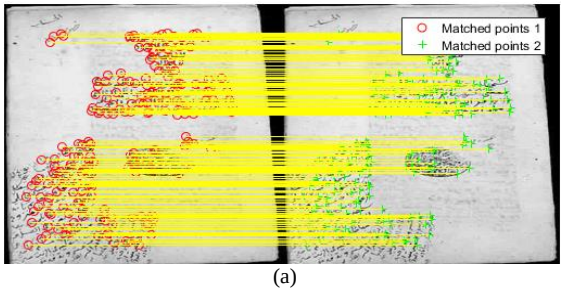
In Fig. 10, the retrieval average time is calculated with the increase of the dataset size to demonstrate the performance of the proposed system in terms of time complexity. we notice the exponential increase after 1100 images dataset size. To enhance the time, we suggest using parallel computing as a future work.

(a)

(b)

Figure 6. Sum Square of difference (SDD) with SURF feature matching

(a) all Matching points between a query image and itself
(b) all Matching points between a query image and another image from the same class



(a)

(b)

Figure 7. Hamming distance BRISK feature matching

(a) all Matching points between a query image and itself
(b) all Matching points between a query image and another image from the same class

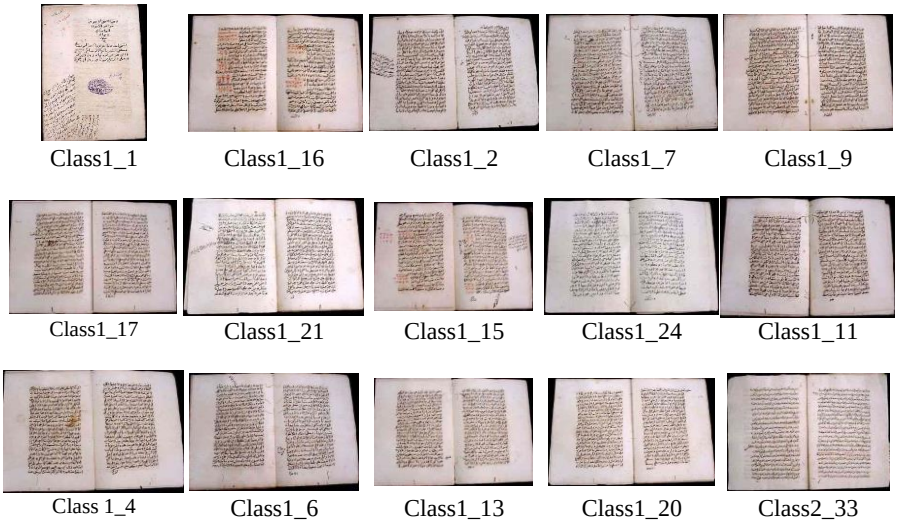
6. CONCLUISIONS AND FUTURE WORK

A CBIR system on Arabic manuscript is proposed using two different features extraction techniques SURF and BRISK. A dataset of Arabic Manuscripts is collected and classified into 29 different classes. The proposed CBIR on Arabic manuscript is trained on 90 query images to find a ranking method by sorting the retrieved images according to the number of similarity features *N*, then get N_{max} and N_{min} which is greater than zero. After that, retrieve ranked images from N_{max} to $(N_{max} + N_{min})/2$.

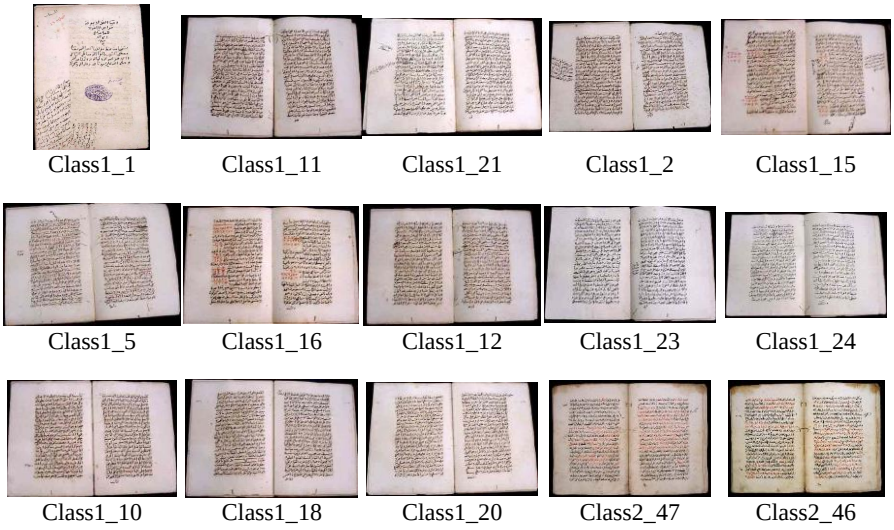
The performance of the proposed Arabic manuscript CBIR is measured in terms of the time complexity, Recall, and Precision. SURF features are found to give better results than BRISK features.

As a future work, different feature extraction methods will be investigated. Collecting more Arabic manuscript dataset. Using parallel computing to enhance the time complexity especially when the dataset size increases.

(a) Query Image (Class1_1)



(b) Retrieved images using SURF feature Matching



(c) Retrieved images using SURF feature Matching

Figure 8. The result of CBIR on Arabic Manuscripts on Class1_1 query image

TABLE I. RESULT OF CBIR IN ARABIC MANUSCRIPTS USING SURF TECHNIQUE

Query Image	Number of expected relevant images	Number of retrieved relevant images	Number of total retrieved images	Recall	Precision	Time (second)
Class4_100	8	8	10	1	0.8	187
Class4_99	8	8	10	1	0.8	184
Class4-98	8	8	10	1	0.8	190
Class1_1	24	14	15	0.58	0.933	292
Class2_25	39	25	63	0.64	0.4	200
Class2-39	39	37	54	0.94	0.69	300
Class3_64	28	26	30	0.92	0.87	190
Class4-92	8	7	10	0.87	0.7	335
Class3_70	28	25	30	0.89	0.83	307

TABLE II. RESULT OF CBIR IN ARABIC MANUSCRIPTS USING BRISK TECHNIQUE

Query Image	Number of expected relevant images	Number of retrieved relevant images	Number of total retrieved images	Recall	Precision	Time (second)
Class4_100	8	5	10	0.625	0.5	189
Class4_99	8	8	14	1	0.57	245
Class4-98	8	8	14	1	0.57	230
Class1_1	24	13	15	0.54	0.87	170
Class2_25	39	7	10	0.18	0.7	360
Class2-39	39	37	55	0.94	0.67	300
Class3_64	28	16	20	0.571	0.8	277
Class4-92	8	6	10	0.75	0.6	290
Class3_70	28	20	30	0.71	0.67	206.1
Class2_55	39	22	26	0.56	0.85	300

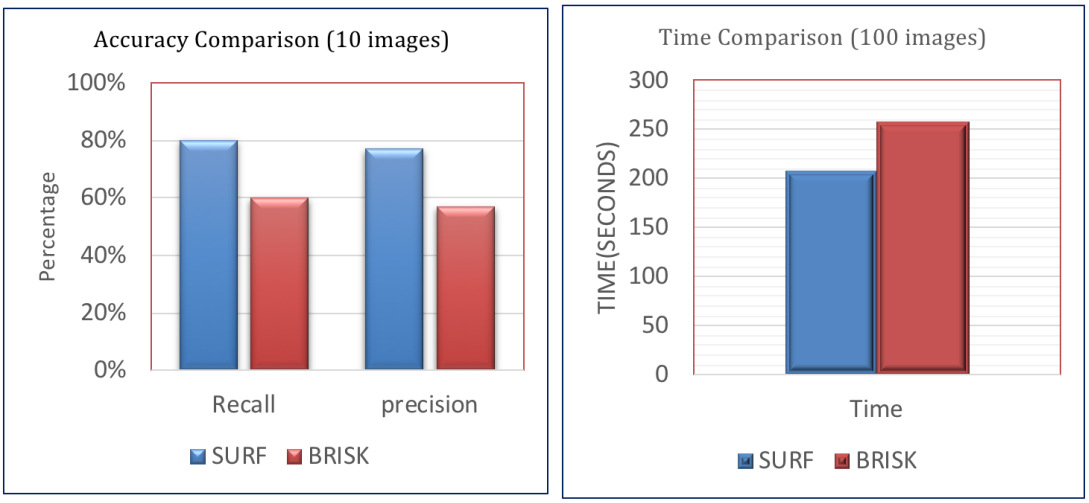


Figure 9. Comparison chart between the result of BRISK and SURF features.

The left side chart shows the accuracy's result, the right one shows the time's result.

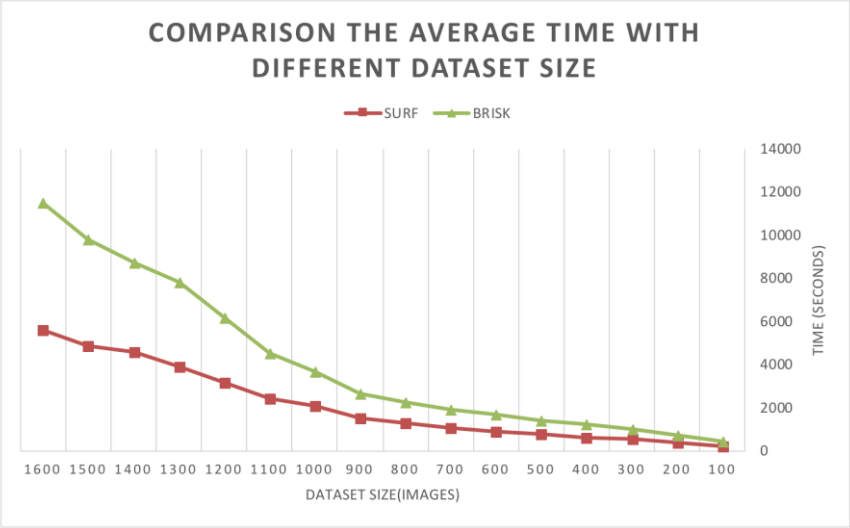


Figure 10. Average retrieval time Vs. Dataset size

REFERENCES

[1] R. Prytherch, Harrod's Librarians' Glossary and Reference Book. Ashgate Pub Ltd; 10 edition (April 6, 2005), 2005.

[2] "Manuscript Research: General, Catalogs, and Manuscripts – ghazali.org." [Online]. Available: https://www.ghazali.org/2015/09/manuscripts/. [Accessed: 02-May-2018].

[3] K. Chung and J. Bin, "A parallel architecture for feature extraction in content-based image retrieval system," *IEEE Conf. Cybern. Intell. Syst. 2004.*, vol. 1, pp. 468–473, 2004.

[4] K. V. S. S. Baboo, "Detection & Prevention of Terrorist Activities Using GIS and Remote Sensing," *Int. J. Comput. Appl.*, vol. 1, no. 7, pp. 1–6, 2010.

[5] S. A. Bakar, M. S. Hitam, and W. N. J. H. Wan Yussof, "Content-based image retrieval using SIFT for binary and greyscale images," *IEEE ICSIPA 2013 - IEEE Int. Conf. Signal Image Process. Appl.*, pp. 83–88, 2013.

[6] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," *Lect. notes Comput. Sci.*, vol. 3951, p. 405-415, 2006.

[7] S. Leutenegger, M. Chli, and Siegwart, RY 2011, "BRISK: Binary Robust invariant scalable keypoints."*Computer Vision (ICCV), 2011 IEEE International Conference.* pp. 2548-2555,2011

[8] N. El Makhfi, O. El Bannay, R. Benslimane, and N. Rais, "Search engine of ancient Arabic Manuscripts based on Metadata and XML annotations," *2011 Colloq. Inf. Sci. Technol.*, p. 10, 2011.

[9] A. Stolyarenko, N. Dershowitz, and T. Aviv, "OCR for Arabic using SIFT descriptors with online failure prediction.",2018

[10] M. Yahia and W. G. AL_KHATIB, "METHOD FOR RETRIEVAL OF ARABIC HISTORICAL MANUSCRIPTS," 2014.

[11] T. Sari and A. Kefali, "A search engine for Arabic documents,"*Francoph. sur l'Ecrit le Doc.*, pp. 97–102, 2008.

[12] S. A. Shahab, W. G. Al-Khatib, and S. A. Mahmoud, "Computer Aided Indexing of Historical Manuscripts," in *International Conference on Computer Graphics, Imaging and Visualisation (CGIV'06)*, pp. 287–295,2006.

[13] R. Herzog, A. Solth, and B. Neumann, "Using harris corners for the retrieval of graphs in historical manuscripts," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, pp. 1295–1299, 2013.

[14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[15] M. Kashif, T. M. Deserno, D. Haak, and S. Jonas, "Feature description with SIFT , SURF , BRIEF , BRISK , or FREAK ? A general question answered for bone age assessment," *Comput. Biol. Med.*, vol. 68, pp. 67–75, 2016.

[16] Ş. Işık, "A comparative evaluation of well-known feature detectors and descriptors", *Int. J. Appl. Math. Electron. Comput.*, vol. 3, no. 1, pp. 1–6, 2014.

[17] A. Patel, D. R. Kasat, S. Jain, and V. M. Thakare, "Performance Analysis of Various Feature Detector and Descriptor for Real-Time Video based Face Tracking," *Int. J. Comput. Appl.*, vol. 93, no. 1, pp. 37–41, 2014.

[18] H. Angrish, N. Kaur, and R. Chadha, "A Comparative Review and Optimization of Surf , Brisk and Freak Feature Descriptors," *Int. J. Sci. Res. Dev. IISN Online*, vol. 3, no. 06, pp. 1368–1370, 2015.

[19] M. A. Elfattah, S. Abuelenin, A. E. Hassanien, and J.-S. Pan, "Handwritten Arabic Manuscript Image Binarization Using Sine Cosine Optimization Algorithm," in *International Conference on Genetic and Evolutionary Computing*, 2016, pp. 273–280.

[20] A. Alfanindya, N. Hashim, and C. Eswaran, "Content Based Image Retrieval and Classification using speeded-up robust features (SURF) and grouped bag-of-visual-words (GBoVW)," *Proc. 2013 Int. Conf. Technol. Informatics, Manag. Eng. Environ. TIME-E 2013*, pp. 77–82, 2013.

[21] L. Kang, J. Kumar, P. Ye, and D. Doermann, "Learning Text-Line Segmentation Using Codebooks and Graph Partitioning," *2012 Int. Conf. Front. Handwrit. Recognit.*, pp. 63–68, 2012.

[22] S. Oron, T. Dekel, T. Xue, W. T. Freeman, and S. Avidan, "Best-Buddies Similarity—Robust Template Matching Using Mutual Nearest Neighbors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 8, pp. 1799–1813, 2018.

Bayan Bagasi received her B.Sc. degree in Computer Science with Honors from King Abdulaziz University in 2012. She works as web developer at Faculty of Engineering from 2012 to now. She worked as Teacher at Jazerat Al-Oloom school from 2013 to 2014. She works as Teaching Assistant from 2015 to date at Faculty of Computing and Information Technology at King Abdulaziz University, Rabigh, Saudi Arabia. Now she is working in her master's Degree at King Abdulaziz University, Jeddah, Saudi Arabia. She has a research interest in image processing, speech and natural language recognition.

Lamiaa A. Elrefaei received her B.Sc. degree with honors in Electrical Engineering (Electronics and Telecommunications) in 1997, her M.Sc. in 2003 and Ph.D. in 2008 in Electrical Engineering (Electronics) from faculty of Engineering at Shoubra, Benha University, Egypt. She held a number of faculty positions at Benha University, as Teaching Assistant from 1998 to 2003, as an Assistant Lecturer from 2003 to 2008, and has been a lecturer from 2008 to date. She is currently an Associate Professor at the faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia. Her research interests include computational intelligence, biometrics, multimedia security, wireless networks, and Nano networks. She is a senior member of IEEE.