Enabling Indexing and Retrieval of Historical Arabic Manuscripts through Template Matching Based Word Spotting

Tayyeba Faisal, Somaya AlMaadeed
Department of Computer Science and Engineering
Qatar University
Doha, Qatar
{tayyeba, s_alali}@qu.edu.qa

***Abstract*— We present a holistic segmentation-free query by example word spotting technique based on template matching. We have applied this technique to a dataset of historical Arabic handwritten manuscript images. First, the documents as well as query word images are pre-processed for separating text from the noisy background and converting to their binary equivalents. Then a pixel based approach is used for computing the similarity between the pre-processed template query word and document images by using the Correlation similarity measure. Slight variations in font sizes are tolerated by adjusting the threshold of similarity. Our robust pre-processing algorithm significantly enhances the performance of the learning-free template matching based word spotting approach. The proposed technique is simple as well as efficient as it does not involve any time-consuming learning steps. Experiments with a historical Arabic dataset yield promising results. This technique can generate locations of occurrences of query word images which is the fundamental step towards building searchable indexes for historical manuscripts.**

Keywords*— *word spotting; template matching; correlation similarity; historical; Arabic

I. Introduction

In a technically advanced world which is growing more paperless by the day, serious dedicated efforts in Pattern Recognition field are being directed towards digitization of handwritten and historical manuscripts with a view to preserve the heritage containing valuable information. In [17], Surinta reports that there are about 600 kilometers of book shelves containing government, council, and financial archives of Netherlands alone. Historical manuscripts are valuable resources representative of a culture and contain the wealth of the knowledge at the roots of its socio-economic development. Over the years, the manuscripts that have survived are already in a degraded condition and it is important to save them from repeated handling to prevent further degradation. To this end, efforts are being made to scan these documents and create digital libraries of historical document images. Historical documents handwritten in Arabic language represent rich Islamic civilization which has left a noticeable impact on the socio-economic world order. However, these documents are still largely in hard copy format or image format and hence, they have limited access to public.

Over the years, Computer Vision and Document Image Analysis researchers have developed techniques for automatic transcription (converting document images to machine readable ASCII transcripts, indexing the documents for retrieval, word recognition and word spotting within these documents. Word spotting approaches aim to spot the locations of whole words or word images inside the document images, as opposed to word recognition where it is important to map the word images to the actual textual content using handwriting recognition techniques. In degraded historical manuscripts, word recognition becomes a challenge due to large extent of degradation and ancient handwriting styles. Word spotting offers a practical alternative approach to counter these challenges as the recognition step is not needed here.

Word spotting approaches aim to provide high accuracy of retrieval of queried words within the collection with high speed and without minimum need to pre-process the document or segment the document image into word images. Word spotting approaches can be classified as segmentation-based (where the image is segmented into lines into words before query matching); and segmentation-free (where there is no need for prior segmentation and queried word images are directly matched inside the documents). Furthermore, the word spotting techniques can be categorized per the query model they follow, i.e. either *Query by String (QBS)* or *Query by Example (QBE)*. In the former, word is represented in text form for retrieval, and is converted to equivalent image representation before searching inside the document. In the latter approach of QBE, example word image is provided for retrieving relevant matches inside the document. The example word image is also called *template* image and hence the problem of spotting the locations of the word image translates to the template matching problem between the template word image and the document image. Since this method does not involve any learning (only a template word image must be extracted from the document to be analyzed), and the template matching approaches contribute well to generating locations

This research is part of the UREP grant no. QUUCP-CENG-CSE-15\16-1 (*Qatar University).*

of word images inside the documents, therefore we adopt the approach of template-based approach for word spotting for our problem domain of semi-automatic indexing of historical Arabic manuscripts to enable word retrieval.

Over the following sections, we provide an overview of the existing research in the domain of template matching based word spotting, an introduction to our proposed methodology and presentation and discussion of the results of applying our proposed system to different data sets.

II. Review of Literature

Supervised learning based word spotting approaches involve learning models for each keyword. The models are usually based on Hidden Markov Models (HMMs) for their ability to compute the probability of keywords generated by the model at test time. Also, because of the large variation in handwriting styles, features of variable lengths are used to represent the word images for better word recognition. HMMs have been used for classification of these systems because of their ability to handle features of variable length. While the HMMs can handle data of varying length, the limitation of this approach is that the keywords must be learned offline from the images and learning requires a large amount of data can only be performed offline.

There are two aspects of a word spotting system: document / word representation scheme and the matching algorithm for matching the keywords within the document. Many representation approaches have been used for the keywords as well as documents for keyword spotting scheme which distinguish one scheme from another [2]. In the early research initiated by Rath and Manmatha [22], four-dimensional profile consisting of horizontal and vertical projection profiles, upper and lower word profiles and background to ink transitions, was used for word representation. Dynamic Time Warping (DTW) algorithm proposed in [21] was for feature matching. Given the word images represented by multidimensional profile features, the DTW determines a common time axis on which the profile locations for the compared words (signals) appear at the same time. DTW can distort (or warp) the time axis, compressing or expanding it at different places to allow finding the best matches between two samples, and hence it can handle well the variations in handwriting. The problem with DTW based matching is its computational complexity, which is why it is not time-efficient. Shah and Suen [9] proposed a solution to this problem by using a different similarity measure or matching algorithm. They used a similar set of features as [22] to represent handwritten Pashto documents which are written in Arabic script. In addition to the upper, lower, left and right word profiles used in [22], the projection profiles were divided into left-diagonal and right-diagonal projection profiles in addition to vertical and horizontal profiles, considering the diagonal strokes characterizing Arabic scripts. Further they divided these projection profiles into column-wise and row-wise profiles defined on 100 zones of 10*10 pixels of grey scale Pashto words. The words were normalized before matching using the correlation similarity measure as opposed to DTW matching used in [22]. Since Urdu script is also like Arabic script, therefore similar set of profile features are used in [1] and [11] to represent Urdu words. Srihari presented another notable work [19], in which the documents are segmented into lines and words. Words are represented by a set of gradient, structural and concavity features. An index for the documents is automatically created using these features. At retrieval time, distance between query word to be spotted and the words in the document is computed using normalized cross correlation.

Other techniques recently proposed in the literature for word spotting involve training models based on a set of training data. Among the stochastic approaches following the Hidden Markov Model (HMM) is the technique proposed in [3], where lexicon-free word spotting based on learning character HMMs is performed. It is a segmentation-free approach and is robust for a multi-writer scenario. A segmentation-free query-by-string decoding scheme is proposed in [25] which employs the Bag-of-Features Hidden Markov Models (HMMs) to learn context dependent character models from a training set. In [26] images are represented through Scale Invariant Feature Transform (SIFT) descriptors aggregated into a bag of visual words model. In a later work [27], the authors improved their approach by employing query fusion methods and relevance feedback from the users. In [23] authors use sequences of features of variable length for word image description and classify them using HMMs. In an earlier work [24], Rodríguez et al. proposed the use of semi-continuous HMM (SC-HMM) for query representation. The parameters of this model are learned in an unsupervised manner. The model can be adapted online to represent the given query, hence overcoming the restriction of learning keywords offline.

In [7], Cheriet and Moghaddam present a complete segmentation-free word spotting approach which involves binarization of historical documents to remove background noise, extracting connected components (CCs) from the binarized images, and clustering the CCs into metaclasses. The CCs are skeletonized before extracting features for Self-Organizing Map (SOM) to map the CCs on to its metaclass. Each cluster is represented by a Basis CC (BCC) from which a BCC library is created. Pseudo-words consisting of all possible sequences of BCCs are stored in library. The spotting step involves generating candidate sequence for a BCC and for the sequence in pseudo-words database. Matched sequence is presented to the user as spotting result. In [12], authors integrated language models into their word spotting system by reconstructing Arabic words from the segmented Parts of Arabic Word (PAWs) or CCs based on the contextual knowledge gathered from the Arabic lexicon. Word spotting is implemented through hierarchical classifiers to recognize PAWs. Also, they included the writings by multiple authors to make their system more robust. To overcome large amount of training data required to train a language model, which, in case of historical documents, is scarcely available and computational complexity of model based systems, [8] proposes a lexicon reduction scheme for Arabic subwords or PAWs.

In [14], authors propose a QBE keyword spotting system which uses the nearest neighbor similarity matching between document and query word images represented by HOG and

LBP descriptors. In [15], a retrieval system for historical manuscripts is proposed in which the user could type in a query to obtain a match within historical document. Word spotting approach is used to create models for query words so that search can be performed both for text as well as image queries. This model representation enables retrieval that is independent of the language alphabet constraints, thus making their system omnilingual. For matching they use the technique of cohesive elastic matching. Among other approaches, [4] involves the use of word spotting method based on Generalized Hough Transform to spot the separator words in Arabic manuscripts. Their technique of using GHT has been used in the later work [5], for spotting words in Arabic manuscripts.

In more recent works, [10] and [20], a word spotting system for both QBS and QBE is presented in which words are represented as Pyramidal Histogram Of Characters (PHOC). Later, in [20], the authors propose a deep Convolutional Neural Network Architecture (CNN) trained with the PHOC word representations for word spotting in historical documents. While the proposed architecture achieves a high accuracy on a multi-writer English historical data set, the relatively longer learning time proves to be unnecessary in a single-writer scenario.

For our problem domain of semi-automating the tasks of indexation and annotation of historical manuscripts for word retrieval we adopt the approach of indexing the historical manuscripts proposed by Rath and Manmatha in [22]. Since the manuscripts are written by a single author, therefore, we do not need to employ the techniques good for multi-writer scenario. Also, since accurate word segmentation is a challenging task for historical Arabic manuscripts because of overlapped words and unclear word boundaries, therefore, we adopt the segmentation-free QBE word spotting approach to solve our problem. Since there is no prior need to segment the document into words, the problem simplifies to matching the given word example images inside the document for retrieving the word locations. As opposed to the DTW algorithm, we propose the use of Normalized Cross Correlation (NCC), inspired by [9] and [19], for template matching inside the historical Arabic manuscript page images, where the template is the example word image of important keywords extracted from the manuscript images.

In the work of [9], the database comprised of handwritten Pashto words which were size normalized and represented in the form of binary features. The binary features were matched using the correlation measure and a dissimilarity measure. In our case of historical Arabic documents, however, the matching must be performed inside the pages of the manuscript which have degraded over time, hence the background noise in the scanned images. Also, the intensity values of pixels vary widely in the scanned images of manuscripts pictured under different lighting conditions. We propose the use of NCC to cater for the intensity variations in the images. The background noise has been removed from the manuscript images as well as the query word images using various pre-processing steps. Finally, the word and manuscript images are represented as binarized images of reduced size for efficient pixel wise similarity measurement using NCC.

III. Proposed Methodology

Before presenting the details of the methodology, we present a brief description of the Normalized Cross Correlation (NCC) algorithm.

A. Details of NCC:

Normalized Cross Correlation (NCC) is used to determine the position of a pattern in an image [28]. A simple approach to locating the pattern or template in an image is perform an exhaustive search. Searching for template image over all image patches of a larger image become cross correlation of template with the larger image. The problem with cross correlation is that it gives a high score when searching for the same template in an image that has higher illumination and lower for same images pictured with low illumination. To avoid this problem, the mean value of the template is subtracted from the score. So, when the same image is pictured under different illumination, the pixels of the images are normalized before matching by subtracting the mean of the patch intensities and dividing by the standard deviation. This process is depicted in (1).

$$h[m,n] = \frac{\sum_{k,l}(g[k,l]-\bar{g})(f[m-k,n-l]-\bar{f}_{m,n})}{\left(\sum_{k,l}(g[k,l]-\bar{g})^2\sum_{k,l}(f[m-k,n-l]-\bar{f}_{m,n})^2\right)^{0.5}} \quad (1)$$

The resulting score values range from perfect match (+1) to completely non-correlated (-1). This can be seen from the fact that when normalized patches are viewed as unit vectors, then correlation becomes dot product of unit vectors, which must range between -1 and +1.

There are three reasons why we chose to use NCC as the similarity measure for our template matching problem:

- The historical manuscript images captured and scanned may have large variations in illumination. We do not want to favour high intensity pixels and punish low intensity pixels by using the distance measures for example Sum of Square Distance which do not provide for any score normalization.
- The locations where the NCC gives highest score can be easily computed, which is a highly contributing factor for automatic generation of indexes after verification of correct matches.
- Since the example word images are segmented out from the same manuscript for matching, the NCC algorithm makes an obvious choice for keyword spotting.

B. Roadmap of Template Matching-based Word Spotting

The proposed roadmap to building the indexes for historical Arabic manuscripts using template matching-based word spotting starts at acquiring the data set on which to test the proposed system. For this research, we used the HADARA80P dataset [16] which consists of high resolution tiff images of 80 pages of a scanned historical Arabic manuscript. The manuscript is hand written by one author and was published in the 9th Islamic century or 15th Gregorian century. The data set also provides 25 keyword images segmented out from random pages of the manuscript. The page images are 48-bit TIFF images with 16 bits per color channel. Each image is about 50 MB in size. Figure 1 shows a page of the original digitized image. Keyword images are segmented in the form of polygons and contain all the diacritic marks but no strokes from neighboring words. The data set is the only historical data set of its kind available freely for word spotting research [16]. There are a few complications in the images:

- Image size is too large. It must be reduced before template matching to reduce the size of the search space.
- Page images as well as keyword images have a lot of background noise which must be removed for separating the text from the background noise as well as for the accuracy of matching.
- The text is not all black. Some of the words are written in red color. This poses a problem for binarization and calls for additional thresholding measures to prevent the loss of these words during binarization.

First the images and the query words are both pre-processed before performing the actual word spotting. The following steps were carried out for pre-processing the page images and query word images.

C. Thresholding

First, the images are converted to 8-bits for representing intensity values in the range of 0-255. The resulting image has three channels. We thresholded the image such that the intensity values higher than the black text are converted to zero (black) and the text is converted to 255 (white). However, using this technique the red text was removed in the resulting binarized image. To correct this, we applied the appropriate thresholding in individual color channels and re-constructed the greyscale image into only one channel.

D. Binariztion

The adjusted threshold resulted in image containing the red text. Binarization was applied to convert the image into black and white (white text on black background). Results of binarization can be seen in Fig. 2. The above two steps reduce the original image size of 50 MB to almost 600 KB.

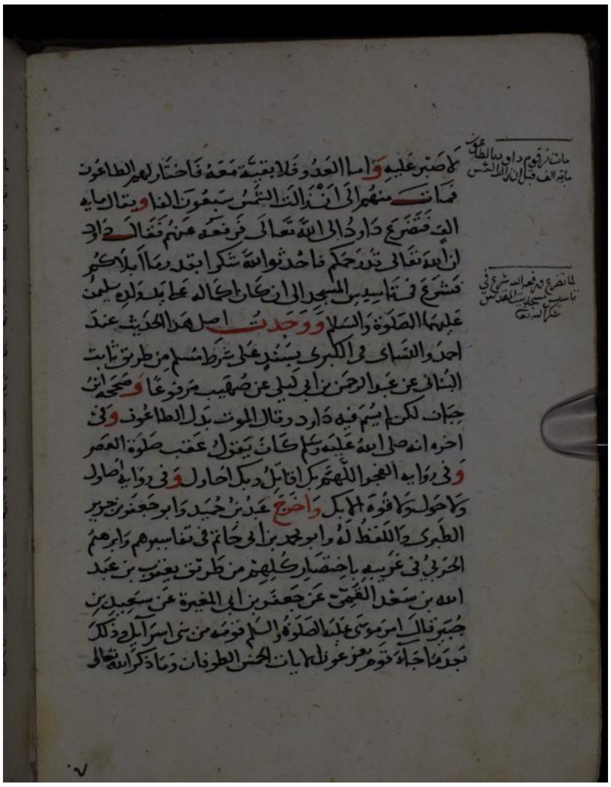


Fig. 1. Original Image

E. Noise Removal

The binarized images obtained in step B still contained some noise which was removed. The diacritics marks were also removed as they do not aid in the matching process. The same pre-processing steps were applied to the query word images.

We further obtain the region of interest containing the main body of text by using a structuring element in the form of a vertical line of size 100, and using it to dilate the connected components. After this we found the bounding box of the component with largest area which gives our region of interest (ROI). These ROIs are used in subsequent word image matching to make the matching process more efficient. The extraction of ROIs, however, becomes insignificant when side notes must be included in the search. In this stage of research, we are focusing on the main body of text for retrieval.

F. Word Image Matching

After the query word images and the page images were pre-processed and stored in similar reduced size representation, the next step is to perform the matching of the pre-processed query words inside each of the pre-processed page images using the NCC algorithm. It should be noted that the same binarization and pre-processing steps must be applied to the keyword images as the page images to maximize the consistency of results.

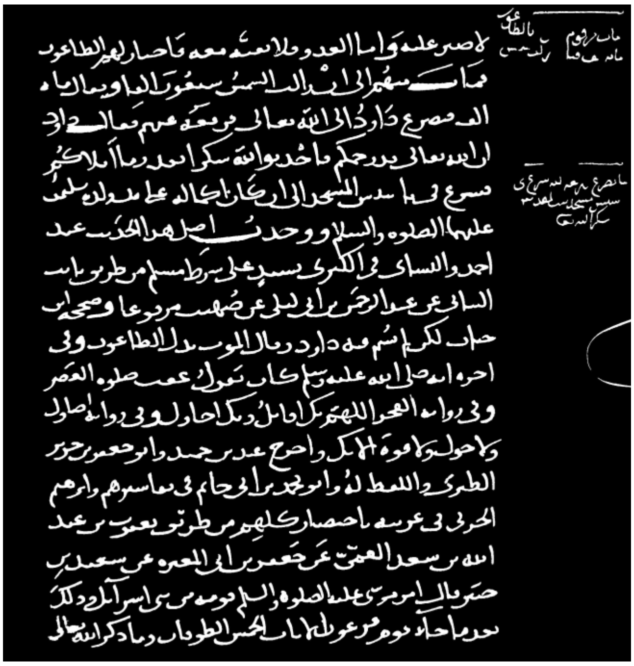
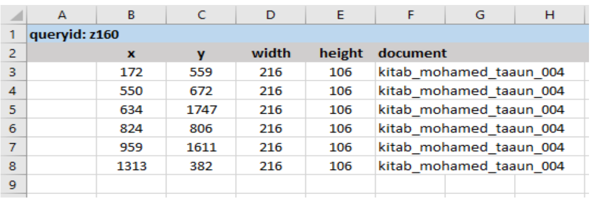


Fig. 2. Binarized Image

As mentioned in section III-A, it is easy to store the locations where NCC has high response. A local maximum is obtained for every page, based on which a global maximum is computed for the set of all document page images. The locations returned are sorted by the local maxima for each page. The locations returned are sorted by the decreasing order of local maxima for each page. In our system, locations of keywords appearing on pages with local maxima within the range of top 30% of the global maximum were retained and others were suppressed.

During image matching, relaxing the threshold for acceptable NCC responses resulted in many overlapping locations. This problem was easily overcome by finding the connected components centered around the point of high response. Large number of overlapping bounding boxes were reduced to only one bounding box per true retrieved location. The locations contain the starting row position and starting column position of the retrieved word matches.

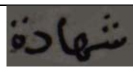
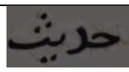
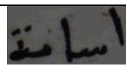
A sample index was created from the location information by storing the Query word id, and the row positions and column positions of its matched locations in a page of the document. Fig. 3 shows the prototype index. Fig. 4 shows the overall framework starting from digitized images, extraction of example words from these images and finally, index generation using the matched image locations.



	A	B	C	D	E	F	G	H
1	queryid: z160							
2		x	y	width	height	document		
3		172	559	216	106	kitab_mohamed_taaun_004		
4		550	672	216	106	kitab_mohamed_taaun_004		
5		634	1747	216	106	kitab_mohamed_taaun_004		
6		824	806	216	106	kitab_mohamed_taaun_004		
7		959	1611	216	106	kitab_mohamed_taaun_004		
8		1313	382	216	106	kitab_mohamed_taaun_004		
9								

Fig. 3. Prototype Index Generation Based on Word Spotting Results

TABLE 1. SAMPLE KEYWORDS USED FOR TESTING

Keyword Code	English Transliteration	Keyword image
Z115	At-Tauun	
Z160	Shahada	
Z131	Hadeeth	
Z168	Osama	
Z58	Intaha	

IV. EXPERIMENTS ON TEST DATASET

We applied our technique to a subset of five keyword images provided in the HADARA dataset. Table 1 shows the labels of keywords used for testing, their English transliterations, and their corresponding images. It should be noted that these images are first binarized using the same algorithm described above for page images before testing our system for spotting word locations inside binarized page images. Three keywords out of the test sample are terms which are deemed significant for the vocabulary of the manuscript, for example, word images for words like AtTauun, Hadeeth and Shahada. The other two keywords were selected randomly.

We have proposed the use of query-by-example template based word spotting and conducted experiments to determine the feasibility of its use in index generation of historical manuscript for word retrieval. The initial experiments conducted with the subset of two query words show promising results. The aim of our research is to propose and prove the usability of simple and classical template matching technique for word spotting to generate searchable indexes for single-writer historical manuscripts. Hence, full evaluation of the system against the baseline system [29] is beyond the scope of this paper. Nevertheless, we present the test results of our system for the five sample keywords in Table 2 and Fig. 5. The evaluation measures we have used are TP (True Positives), FN (False Negatives) and R (Recall). We will briefly describe each measure below.

TP: Number of correct / relevant word locations retrieved (spotted) by the system

FN: Number of correct / relevant word locations *not* retrieved or *missed* by the system.

R: Number of relevant word locations retrieved by the system out of the total number of relevant word locations. It is also defined in terms of TP and FN in (2).

$$R = TP / (TP + FN) \quad (2)$$

The choice of these performance measures is evident from the sensitive nature of index generation in which the goal is to minimize the possibility of missing a relevant word location, or in other words, maximize the recall.

V. Results and Discussion

Table 2 shows the TPs and FNs of the baseline HADARA system and our proposed system. Bar chart in Fig. 5 shows the comparison of Recall rates, in percentage, of our proposed system with the baseline system. The Recall rates are computed from TP and FN values for all keywords for both systems using (2). The preliminary results show that the performance of our system closely matches performance of baseline HADARA80P system in two keywords z115 (AtTauun) and z160 (Shahada), and surpasses that of HADARA for the other three keywords in terms of recall. It should be noted that when compiling the results, the retrieved word locations which do not cover a large area (almost 80%) of the relevant word location are not included as true positives. It is expected that when our approach is combined with additional rotation-invariant features, our proposed methodology will achieve even better recall for word location spotting.

VI. Conclusion

In this paper, we proposed the use of a template matching-based query by example word spotting system to generate indexes for historical manuscripts in a semi-automated manner. The locations generated by our system must be verified manually before using in the index generation. Our system has been tested on an Arabic Historical manuscripts dataset, namely, HADARA80P. The high recall rate for a selected subset of important query words is quite promising as a proof of feasibility of template-based approaches for indexing and retrieval of historical Arabic manuscripts while word recognition for this domain remains a challenge.

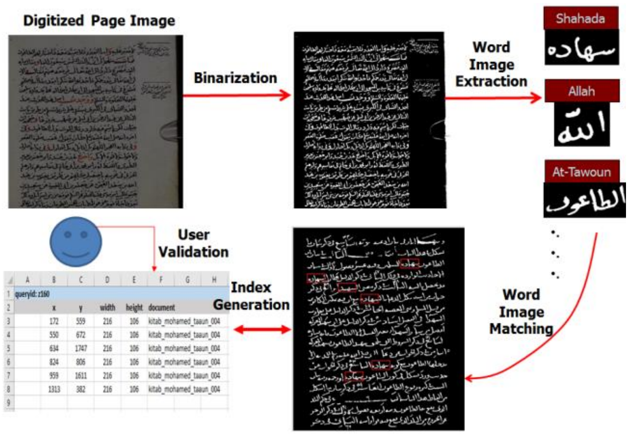


Fig. 4. Overall Framework of Index Generation using Word Spotting

TABLE 2. TP AND FN FOR SAMPLE KEYWORDS

Keyword Code	True Positives (TP) HADARA80P	False Negatives (FN) HADARA80P	True Positives (TP) Proposed System	False Negatives (FN) Proposed System
z115	109	49	106	52
z160	37	5	34	8
z131	100	74	123	51
z168	10	13	18	5
z58	14	10	19	5

VII. Future Work

In future, we wish to make use of the Scale Invariant features to cater for rotated and scaled words. It is expected that when our approach is combined with additional scale and rotation-invariant features, our proposed methodology will achieve even better results for word spotting. The matching performance will be improved by further fine tuning the pre-processing algorithm.

As a very important part of future work we will use our technique for semi-automated transcription, annotation and/or translation of words inside the manuscripts. Once these transcriptions are generated automatically, this will aid in incorporating the Query by String functionality.

Finally, the technique will be applied to the collection of manuscripts in our own library for building a full word retrieval system.

Acknowledgment

This publication was made possible by QUCP grant number QUCP-CENG-CSE-15\16-1 from Qatar University. The statements made herein are solely the responsibility of the authors.

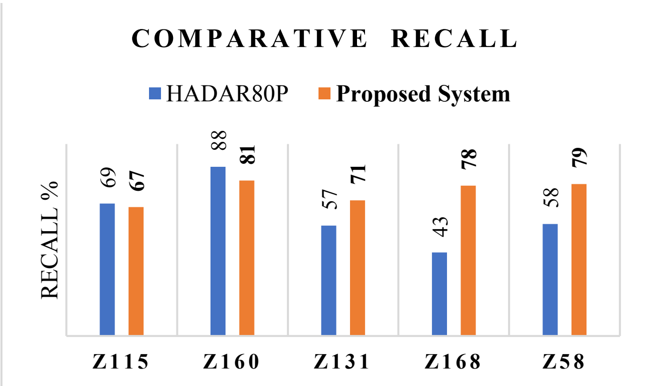


Fig. 5. Comparative Recall for Five Sample Keywords

REFERENCES

[1] A. Abidi, A. Jamil, I. Siddiqi, and K. Khurshid, “Word spotting based retrieval of Urdu handwritten documents,” Proc. - Int. Work. Front. Handwrit. Recognition, IWFHR, pp. 331–336, 2012.

[2] R. Ahmed, W. G. Al-Khatib, S. Mahmoud, and B. Rashad Ahmed, “A Survey on handwritten documents word spotting,” Int. J. Multimed. Inf. Retr., 2016.

[3] A. Fischer, A. Keller, V. Frinken, and H. Bunke, “Lexicon-free handwritten word spotting using character HMMs,” Pattern Recognit. Lett., vol. 33, no. 7, pp. 934–942, 2012.

[4] N. Aouadi, “Word Spotting for Arabic Handwritten Historical Document Retrieval using Generalized Hough Transform,” 2011, Third Int. Conf., no. c, pp. 67–71, 2011.

[5] N. Aouadi and A. K. Echi, “Prior segmentation of old Arabic manuscripts by separator word spotting,” 6th Int. Conf. Soft Comput. Pattern Recognition, SoCPaR 2014, pp.

[6] W. Pantke, A. Haak, and M. Volker, “Color Segmentation for Historical Documents Using Markov Random Fields,” pp. 151–156.31–36, 2015.

[7] M. Cheriet and R. F. Moghaddam, “A Robust Word Spotting System for Historical Arabic Manuscripts,” Guid. to OCR Arab. Scripts, pp. 453–484, 2012.

[8] H. Davoudi, M. Cheriet, and E. Kabir, “Lexicon reduction of handwritten Arabic subwords based on the prominent shape regions,” Int. J. Doc. Anal. Recognit., vol. 19, no. 2, pp. 1–15, 2016.

[9] M. I. Shah and C. Y. Suen, “Word spotting in gray scale handwritten pashto documents,” Proc. - 12th Int. Conf. Front. Handwrit. Recognition, ICFHR 2010, pp. 136–141, 2010.

[10] A. Gordo, A. Forn, E. Valveny, and J. Almaz, “Word Spotting and Recognition with Embedded Attributes,” vol. 36, no. 12, pp. 2552–2566, 2014.

[11] R. Hussain, H. A. Khan, I. Siddiqi, K. Khurshid, and A. Masood, “Keyword Based Information Retrieval System for Urdu Document Images,” Proc. - 11th Int. Conf. Signal-Image Technol. Internet-Based Syst. SITIS 2015, pp. 27–33, 2016.

[12] M. Khayyat, L. Lam, and C. Y. Suen, “Learning-based word spotting system for Arabic handwritten documents,” Pattern Recognit., vol. 47, no. 3, pp. 1021–1030, 2014.

[13] T. Konidaris, A. L. Kesidis, and B. Gatos, “A segmentation-free word spotting method for historical printed documents,” Pattern Anal. Appl., vol. 19, no. 4, pp. 963–976, 2015.

[14] A. Kovalchuk, L. Wolf, and N. Dershowitz, “A Simple and Fast Word Spotting Method,” Proc. Int. Conf. Front. Handwrit. Recognition, ICFHR, vol. 2014–December, pp. 3–8, 2014.

[15] Y. Leydier, A. Ouji, F. LeBourgeois, and H. Emptoz, “Towards an omnilingual word retrieval system for ancient manuscripts,” Pattern Recognit., vol. 42, no. 9, pp. 2089–2105, 2009.

[16] W. Pantke, M. Dennhardt, D. Fecker, V. Märgner, and T. Fingscheidt, “An Historical Handwritten Arabic Dataset for Segmentation-Free Word Spotting - HADARA80P,” Proc. Int. Conf. Front. Handwrit. Recognition, ICFHR, vol. 2014–Decem, pp. 15–20, 2014.

[17] Surinta O. Multi-Script Handwritten Character Recognition using Feature Descriptors and Machine Learning. 2016;(October). doi:10.13140/RG.2.2.14250.82888.

[18] W. Pantke, V. Margner, and T. Fingscheidt, “On evaluation of segmentation-free word spotting approaches without hard decisions,” Proc. Int. Conf. Doc. Anal. Recognition, ICDAR, pp. 1300–1304, 2013.

[19] S. N. Srihari and G. R. Ball, “Language independent word spotting in scanned documents,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 5362 LNCS, pp. 134–143, 2008.

[20] S. Sudholt and G. A. Fink, “PHOCNet: A Deep Convolutional Neural Network for Word Spotting in Handwritten Documents,” 2016.

[21] T. M. Rath and R. Manmatha, “Word Image Matching Using Dynamic Time Warping.”

[22] T. M. Rath and R. Manmatha, “Word spotting for historical documents,” Int. J. Doc. Anal. Recognit., vol. 9, no. 2–4, pp. 139–152, 2007.

[23] J. A. Rodríguez-Serrano and F. Perronnin, “A model-based sequence similarity with application to handwritten word spotting,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2108–2120, 2012.

[24] J. A. Rodríguez-Serrano and F. Perronnin, “Handwritten word-spotting using hidden Markov models and universal vocabularies,” Pattern Recognit., vol. 42, no. 9, pp. 2106–2116, 2009.

[25] L. Rothacker, M. Rusinol, and G. A. Fink, “Bag-of-features HMMs for segmentation-free word spotting in handwritten documents,” Proc. Int. Conf. Doc. Anal. Recognition, ICDAR, pp. 1305–1309, 2013.

[26] M. Rusiñol, D. Aldavert, R. Toledo, and J. Lladós, “Browsing heterogeneous document collections by a segmentation-free word spotting method,” Proc. Int. Conf. Doc. Anal. Recognition, ICDAR, pp. 63–67, 2011.

[27] M. Rusiñol and J. Lladós, “Boosting the handwritten word spotting experience by including the user in the loop,” Pattern Recognit., vol. 47, no. 3, pp. 1063–1072, 2014.

[28] J. C. Yoo and T. H. Han, “Fast normalized cross-correlation,” Circuits, Syst. Signal Process., vol. 28, no. 6, pp. 819–843, 2009.

[29] W. Pantke et al., “HADARA – A Software System for Semi-Automatic Processing of Historical Handwritten Arabic Documents,” pp. 161–166, 2013.