

Handwriting Word Recognition Based on SVM Classifier

Mustafa S. Kadhm

Computer Science Department, University of Technology/
Baghdad

Asst. Prof. Dr. Alia Karim Abdul Hassan

Computer Science Department, University of Technology/
Baghdad

Abstract—this paper proposed a new architecture for handwriting word recognition system Based on Support Vector Machine SVM Classifier. The proposed work depends on the handwriting word level, and it does not need for character segmentation stage. An Arabic handwriting dataset AHDB, dataset used for train and test the proposed system. Besides, the system achieved the best recognition accuracy 96.317% based on several feature extraction methods and SVM classifier. Experimental results show that the polynomial kernel of SVM is convergent and more accurate for recognition than other SVM kernels.

Keywords—Arabic Text; Preprocessing; Feature Extraction; SVM

I. INTRODUCTION

This Handwriting recognition is the process of converting the handwriting text images into a text file that understandable by the computer and used for many purposes. There are a lot of applications that depends on handwriting that are postal address reading for mail sorting purposes, cheque recognition and word spotting on a handwritten text page, etc. Naturally, handwriting is cursive and more difficult than printed recognition due to several factors that are the writer's style, quality of the paper and geometric factors controlled by the writing condition its very unsteady in shape and quality of tracing. Several steps taking place in handwriting recognition system, starting with preprocessing, feature extraction and classification.

To develop pattern recognition methods and systems, a lot of sample patterns is essential. In the same way as off-line character pattern databases such as IESK_arDB [1], AHDB [2], IFN/ENIT [3], KHATT [4] and so on, have been playing significant roles for off-line handwriting recognition. The purpose of this database is to enable the community to challenge the problem of object classification and recognition. Therefore, in this paper AHDB dataset has been selected for our proposed system. The dataset has the most popular Arabic words that written by many writers.

Furthermore, preprocessing is the first step in handwriting recognition systems it is helpful to reduce the variability of handwriting by correct these factors, and it will help to enhance the accuracy of segmentation and recognition methods. The second step in recognition system is the features extraction that extract a helpful information from the image text word to distinguish it from the other words. The last step of the

recognitions is the classification that make the decision to sign the text word into the correct class label. [5]

The main contribution of the work presented in this paper is the integration of using multi scale features with a SVM for handwriting word recognition.

In order to show that features for Arabic script can be learned with the HOG descriptor, we evaluate our method on the AHDB dataset. The remainder of this paper is organized as follows:

Section II gives an overview of previous publications covering part-based classifiers for handwritten word recognition. Afterwards, our work is discussed in Section III and evaluated in Section IV. A final conclusion is given in Section V.

II. RELATED WORK

The number of work for Arabic handwriting word recognition is quite limited comparing to the Latin script. The first works was given by Farah. [6] who introduced a system based on the combination of three Multi-Layer Perceptron for the recognition of Arabic words with recognition accuracy 94%. El-Hajj. [7] have used Neural Networks to combine three homogeneous HMM-based classifiers, that have different features as input , and they achieved recognition accuracy 94,44%. Alma'adeed [8] combined a rule based recognizer with a set of HMMs to recognize words in bank check of 47 words. The system tested on 4700 words and achieved 60% recognition accuracy.

Souici-Meslati [9] present a hybrid approach to the recognition of literal amounts on bank-checks. Three classifiers ran in parallel: neural networks, K-nearest neighbor, and Fuzzy K-nearest neighbor. The outputs were combined by word level score summation. 1200 words by 100 writers used for training and 3600 words for testing .The recognition accuracy was 96%. Burrow [10] applied on K-nearest Neighbors classification to each sub-word. The author scores at 74% for sub-word on correctly represented classes.

III. PROPOSED HANDWRITING RECOGNITION SYSTEM

The proposed method for handwriting word recognition has several major steps. Each of the recognition step affect the accuracy and the performance of the recognition. First of all the input images converted into grayscale it pass through several process as shown in figure1.

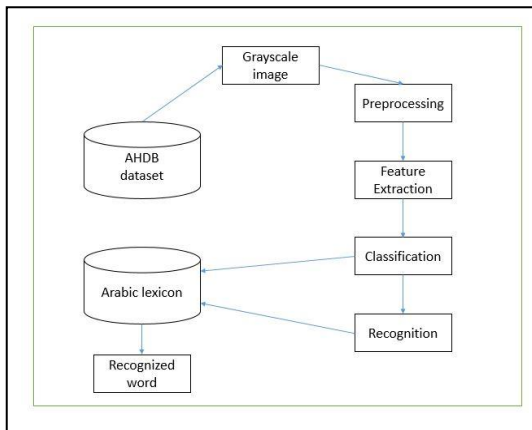


Fig. 1. Flowchart of the proposed system

The proposed system involves several steps which are; preprocessing, feature extraction, classification and recognition. Besides that, each step has its benefits for the recognition process. Here the proposed method steps described in details:

A. Preprocessing

Preprocessing is an essential step in the handwriting word recognition due to the effectiveness of this process on the recognition rate. Several steps have been taken place in the preprocessing phase that make the proposed method obtain a high accuracy.

The input to the handwriting word recognition system is a grayscale text image which has the Arabic word. During preprocessing the image converted to binary by thresholding method. The benefit of the thresholding is reducing the image diameter to make it easy to process.

In the proposed system Fuzzy C-Means clustering (FCM) in [11] has been used for thresholding purpose. After that, some noise appears due to the thresholding. 3X3 median is used to remove undesired information from the binary image as shown in figure 2.

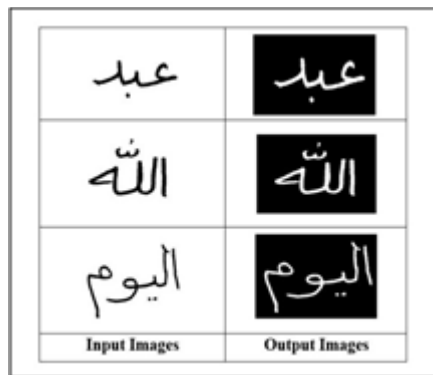


Fig. 2. Image Thresholding

Black space around the written word in an image does not help in any recognition process. So this unwanted black space around the word was eliminated. To eliminate this black space, bounding boxes were used. From each side of the binary

image, the first pixel of the written word was located. This produced four points which formed the boundaries of the bounding box. The black area around this box could then be eliminated using these four values. The elimination of black space in an image is shown in Fig. 3, below:

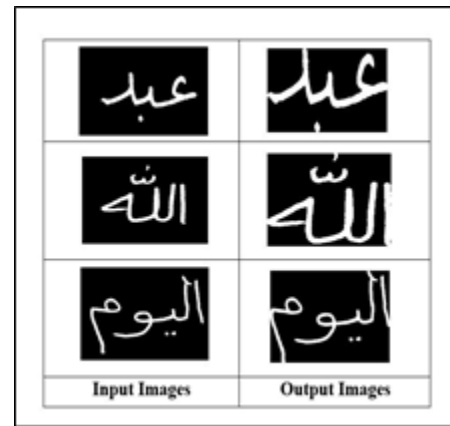


Fig. 3. Black space elimination

Moreover, image thinning is used which is the process of reducing image size by removing the redundant pixels without losing the representation of the original image. 3*3 mask is used to scan the whole image and find the 4 connected pixels. After that the unaffected pixels are eliminated from the image. This process must save the geometry and the connections between the words and the location of original word, based on border pixels removing recursively taking into account saving the geometry, location and connections. Image thinning method in [12] has been used.

The last step in preprocessing is the image normalization. AHDAB dataset has various image sizes. It is important to make all the images in the dataset the same size and make the recognition process fast. After testing several sizes (32x32, 64x64 and 128x128) the 128*128 size gave the best recognition rate. Therefore, all the dataset images are normalized into size 128*128.

B. Feature Extraction

The most important process in handwriting word recognition system is the feature extraction step. The best recognition depends on a successful feature extraction method. A lot of feature extraction methods have been proposed for recognition purpose. However, there are three main types of features that can be obtained from the character images.

1) *Structural Features*: Structural features describe the geometrical and topological characteristics of a pattern by describing its global and local properties. The structural features depend on the kind of pattern to be classified [13].

For Arabic words, the features consist of zigzag, dots, loops, end points, intersection points and strokes in many directions.

2) *Statistical Features*: Statistical features are numerical measures computed over images or regions of images. They include, but are not limited to, histograms of chain code

directions, pixel densities, moments, and Fourier descriptors [14]. Statistical features are easy to compute and text independent. In the proposed system two types of statistical feature has been used which are:

- **Connected Components Feature:** The Arabic words has different shapes. Several shapes has different numbers of connected components pixels (segments). The idea behind of the connected component is to scan the whole image from left to right to find the groups of connected pixels (8 – connected neighbors). After that, each group of the connected pixels will get a label number. Therefore, the feature that obtained from this method is the number of connected components. This method is useful in Arabic words, since there are several words has different number of connected components.
- **Zoning Features:** In zoning features the image divided into number of zones and a particular features extracted from each zone. Several features extracted in this method which increased the recognition accuracy.

First the image divided into four zones figure5 then for each zone summation of the diagonal pixels has been calculated as a feature for that zone.



Fig. 4. Divide image into 4 zones

Second, the image divided into sixteen (16) vertical and horizontal blocks figure (14) then the summation of each block pixels will be the feature of that block.

3) *Global Transformation:* The transformation schemes convert the pixels transformation of the pattern to a more compact form which reduces the dimensionality of features [15].

- **The Discrete Cosine Transform Features (DCT):** The DCT converts the pixel values of an image in the spatial domain into its elementary frequency components in the frequency domain. Given an image $f(i, j)$, its 2D DCT transform is defined as follows:

$$F(u, v) = \frac{2}{N} C(u) C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (1)$$

The inverse transform is defined by:

$$f(i, j) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) C(v) F(u, v) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (2)$$

Where

$$\alpha(u), \alpha(v) = \begin{cases} \sqrt{\frac{1}{N}} & \text{For } u, v = 0 \\ \sqrt{\frac{2}{N}} & \text{Otherwise} \end{cases} \quad (3)$$

Due to its strong capability to compress energy, the DCT is a useful tool for pattern recognition applications. The DCT can contribute to a successful pattern recognition system with classification techniques such as Support Vector Machine and Neural Network [16]. In the proposed system the DCT applied for the whole image that produced from the previous phase. The output of the DCT is an array of DCT coefficients. The features are extracted in a vector sequence by arranging the DCT coefficient in zigzag order, so that most of the DCT coefficients away from the beginning are small or zero. After testing the coefficients it found that the best number of DCT coefficients to represent the handwriting word as feature vector is the first 50 coefficients.

The DCT features extracted by the following steps:

Algorithm DCT_FEXT

- Step1: Read input image
- Step2: Compute DCT for the input image(binary image)
- Step3: Convert the DCT image into 1D array by zigzag order
- Step4: Choose the first 50 DCT coefficients as a features
- Step 5: Save the result features in 1D array
- End

- **Histogram of Oriented Gradient (HOG):** Histogram of Oriented Gradient (HOG) was first proposed by Dalal and Triggs [17] for human body detection but it is now one of the successful and popular used descriptors in computer vision and pattern recognition. HOG counts occurrences of gradient orientation in part of an image hence it is an appearance descriptor. Before applying the HOG, the binary images converted to grayscale then filtered by using proposed edge detection mask filter in figure6 which gave better recognition rate than Sobel and Roberts filters.

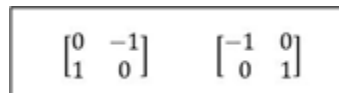


Fig. 5. proposed edge detection filter

HOG divides the input image into small square cells (here 6x6 has been used with bin=9 directions) and then computes the histogram of gradient directions or edge directions based on the central differences.

For improve accuracy, the local histograms have been normalized based on the contrast and this is the reason that HOG is stable on illumination variation. By applying this step, the total size of the feature set in the feature vector will be (6x6x9) = 324. It is a fast descriptor in compare to the SIFT and LBP due to the simple computations, it has been also shown that HOG features are successful descriptor for detection.

C. Features Normalization

An important step to make the mathematical computing simple and fast a feature normalization (scaling) has been used

to make the features ranges [0 1] by applying the following formula:

$$A' = \frac{A - \text{Min}(A)}{\text{Max}(A) - \text{Min}(A)} \quad (4)$$

Where A is an original value, A' is the normalized value.

D. Classification and Recognition

After the feature extraction, the major task is the make decision to classify the word to which class it belongs. There are various classifiers that can applied in word recognition. The most important and more effective classifier is Support Vector Machine (SVM).

1) *SVM Classifier*: Vapnik and Cortes developed SVMs [18, 19] as a statistical learning machine in the late 1990s. Within a short time, they became one of the most popular classification systems in data mining and pattern recognition applications, due to their high classification rates. Researchers successfully applied SVMs in many modern learning applications such as Optical Character Recognition (OCR), bioinformatics, document analysis, and image classification.

SVM commonly used with linear, polynomial, RBF and sigmoid kernels. A multiclass SVM classification (libsvm) has been used in the proposed system [20] with different kernels of 1) linear, 2) polynomial, 3) RBF, 4) sigmoid and it achieves a very high recognition accuracy.

The final step is the recognition which is matching the selected class by the SVM with the character ASCII and find the desired word in the Arabic lexicon.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

The proposed method is implemented using Matlab R2015a version, under windows7 64-bit Operating System, with RAM 6GB, CPU 2.50GHz core i5 and it achieved fast and effective results.

The proposed dataset has 2913 handwriting word images. Each word has 105 images written in different style. In the handwriting word recognition system 70% of the dataset used for training purpose (2044) and 30% for testing (896) and it achieved 96.317% recognition accuracy. When we increase the number of the training data, the recognition accuracy is also increase.

However, we applied the system on our proposed dataset that has 1675 word images. Besides, the dataset has many images with noise. The proposed system also achieves height recognition accuracy 96.3% with the proposed dataset. Figure 6 shows the samples of the proposed dataset.

Furthermore, in the proposed system SVM classification work with different kernels and each kernel achieved different accuracy. Besides that, there are an important parameters which make the SVM work perfectly.

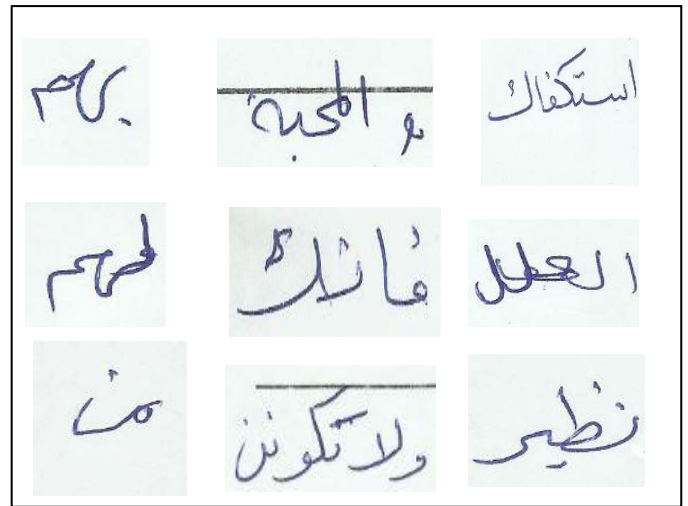


Fig. 6. proposed handwriting dataset

The most important parameters in SVM are: cost(c) and gamma (γ). After many testing of the system the best values of the parameters was $c = 8$ and $\gamma = 0.0625$.

Furthermore, different SVM kernels has been tested and the best achievement was by using SVM polynomial kernel.

In Table 1 it noted that the highest accuracy obtained by our proposed system this is due to the use of efficient features and SVM classifier. Our system achieves good accuracy compared with [6, 7, 8, 9, 10] systems which prove the performance of Polynomial kernel of SVM classifier.

TABLE I. COMPARISON RESULTS

	Classifier	Accuracy
El-Hajj	3HMMs	94.44%
Farah	3MLP	94%
Burrow	Several KNN	74%
Souici-Meslati	NN+KNN+Fuzzy	96%
Alma'adeed	HMMS	60%
Proposed System	SVM	96.317%

V. CONCLUSION

In this a paper, a proposed a high accurate handwriting word recognition system. The system use 70% of the dataset for training and 30% for testing and obtained high accuracy with SVM polynomial kernel. The high accuracy achieved by several factors starting from the efficient preprocessing stage with the use of FCM the with efficient feature extraction methods and finally with more accurate recognition classifier .Experiments, our proposed system gave best recognition accuracy than the existing systems.

TABLE II. COMPARISON BETWEEN DIFFERENT KERNELS OF SVM

SVM Kernels	Linear	Polynomial	RBF
Recognition Accuracy	92.63%	96.317%	91.5%

In addition, we are planning to apply our proposed work in future for recognition the text in documents not just word recognition. The future work need efficient segmentation method that can segment each text word without any overlapping or missing parts. Neural networks can be used for the segmentation and classification which may improve our results.

REFERENCES

- [1] M. Elzobi, A. Al-Hamadi, Z. Al Aghbari, and L. Dings, "IESK-ArDB: a database for handwritten Arabic and an optimized topological segmentation approach," International Journal on Document Analysis and Recognition, vol. 16, no. 3, pp. 295–308, 2013.
- [2] S. Al-Ma'adeed, D. Elliman and C. Higgins, "A Data Base for Arabic Handwritten Text Recognition Research," Proceedings of the 8th International Workshop on Frontiers in Handwriting Recognition (IWFHR 2002), pp. 485-489, August 2002.
- [3] S. Mozaffari, H. E. Abed, V. Maergner, K. Faez and A. Amirshahi, "IfN/Farsi-Database: a Database of Farsi Handwritten City Names," Proceedings of the 11th International Conference on Frontiers in Handwriting Recognition (ICFHR 2008), pp. 397-402, August 2008.
- [4] S. A. Mahmoud, I. Ahmad, M. Alshayeb, W. G. Al-Khatib, M. T. Parvez, G. A. Fink, V. Margner and H. E. Abed, "KHATT: Arabic Offline Handwritten Text Database," Proceedings of the 2012 International Conference on Frontiers in Handwriting Recognition (ICFHR 2012), pp. 449-454, September 2012.
- [5] Lorigo, L.M. and Govindaraju, V. (2006). Off-line Arabic Handwriting Recognition: A Survey. IEEE TRANSACTIONSON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. , 712- 724.
- [6] Aziz N., Farah N., Sellami N., and Ennaji A., " Using Diversity in Classifier set Selection for Arabic handwritten recognition" in proceedings of multi classifiers system , lecture note in CS, Berlin,pp.235-244,2010.
- [7] El-Hajj R., Mokbel C., and Likforman-Sulem L., "Combination of Arabic Handwritten words". In proceedings of the 19th International Conference on Document Analysis and Recognition,Parana,pp. 959-963,2007.
- [8] S.Alma'adeed,C.Higgins,and D.Elliman,"offline recognition of handwritten Arabic words using multiple hidden markov models",knowledge based systems ,vol.17,pp. 75-79,2004.
- [9] L. Souici-Meslati, M. Sellami, " A Hybrid NeuroSymbolic Approach for Arabic Handwritten Word Recognition ", JACIII, Journal of Advanced Computational Intelligence and Intelligent Informatics, FujiPress, Vol. 10, No. 1, pp. 17-25, 2006
- [10] Peter Burrow. Arabic handwriting recognition. Technical report, School of Informatics, University of Edinburgh., 2004
- [11] Elzobi, Moftah, Ayoub Al-Hamadi, Zaher Al Aghbari, and Dinges, Laslo (2012) viewed 5 March 2015, at. <http://www.iesk-ardb.ovgu.de/>
- [12] Niall, Gunter, Innsbruck: Optical Character Recognition. Informatics Research Institute (IRIS) at University of Salford(2011)
- [13] Vinciarelli, A. B. (2004). Offline recognition of unconstrained handwritten texts using HMMs and statistical language models. IEEE Transactions on Pattern Analysis and Machine Intelligence , 26 (6), 709–720.
- [14] Issam Bazzi ,Richard Schwartz and John Makhoul. An "Omnifont Open-Vocabulary OCR System for English and Arabic". IEE Trans. On pattern analysis and machine intelligence. Vol 21 , no. 6 , pp 495-504,1999
- [15] F.zaki, S. Elkonyaly, A. A. Elfattah, and Y. Enab. "A new technique for Arabic handwriting recognition". In processing of the 11th international conference for statistics and computer science ,(cairo , Egupt), pp. 171-180,1986
- [16] JIANG, J., WENG, Y. & LI, P. (2006) Dominant colour extraction in DCT domain. Image and Vision Computing, 24, 1269-1277.
- [17] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, 2005
- [18] Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–297.
- [19] Vapnik V. The nature of statistical learning theory. Springer; 1999.
- [20] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. ACM T. Intell. Syst. Technol., 2(3):1–27, 2011