International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 22 (2015) pp 43120-43124

Handwriting Word Recognition Based on Neural Networks

Dr. Alia Karim Abdul Hassan

Assistant Professor, Department of Computer Science, University of Technology, Iraq, Baghdad
hassanalia2000@yahoo.com

Mustafa S. Kadhm,

PG Scholar, Department of Computer Science, University of Technology Iraq, Baghdad
must.salam@yahoo.com

Abstract
Handwriting recognition is very challenging field in recent year. This paper proposed a new architecture for handwriting word recognition system Based on Neural Nets NN Classifier. The proposed work depends on the handwriting word level and it does not need for character segmentation stage. An Arabic handwriting dataset AHDB has been used for training and testing the proposed system. Besides, the system achieved best recognition accuracy 95% based on several feature extraction methods and NN classifier.

Keywords: Arabic Text, Preprocessing, Feature Extraction, NN.

Introduction

Handwriting recognition (HWR) is an active research area in artificial intelligence, pattern recognition, and computer vision. Besides, the field of text recognition achieved great success in the real world target applications especially in the e-government system, security application and other fields such as postal address reading for mail sorting, cheque recognition and word spotting on a handwritten text page. In this chapter the handwriting classifications and their general process flow has been explained. For each step a brief overview given of the handwriting techniques and methods used.

In order to develop pattern recognition methods and systems, a large amount of sample patterns are essential. In the same way as off-line character pattern databases such as IAM [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 which 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 which make the decision to sign the text word to its desired class. [5]

Proposed 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.

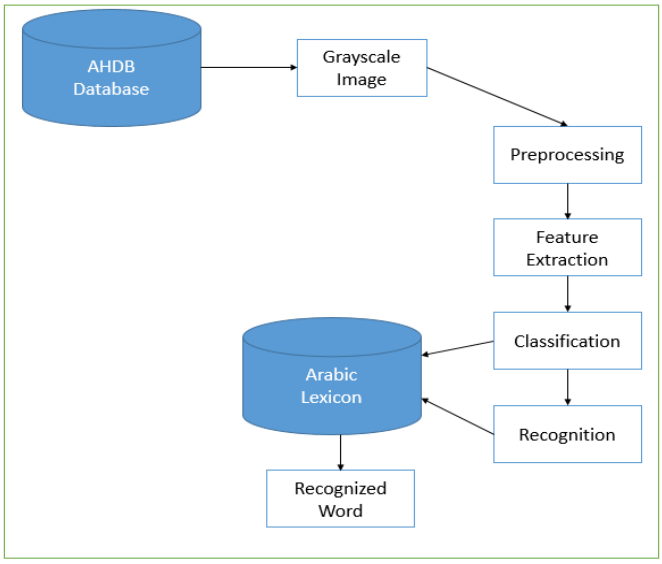

Fig.1. Flowchart of the Proposed System

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

A. Preprocessing

When the acquisition is performed, most systems include a step of preprocessing. Generally, these preprocessing are not specific to the recognition text, but are conventional preprocessing in image processing. The preprocessing designed to prepare the image of the route to the next phase of analysis. It is essentially reduce the noise superimposed data and keep as much as possible, significant information as presented. The noise may be due to the device acquisition, the acquisition conditions (lighting, incorrect document formatting ...) or yet the quality of the original document.

In the proposed system Fuzzy C-Means clustering (FCM) in [6] has been used to for thresholding purpose. After that, some noise appear due to the thresholding. 3X3 median used to remove undesired information from the binary image as shown in figure2.

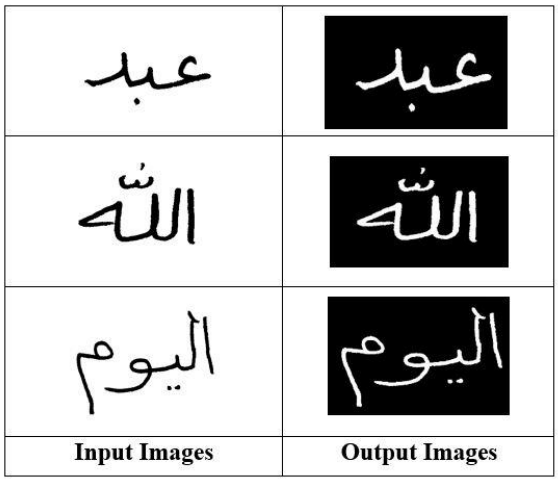

Fig.2. Image Thresholding

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:

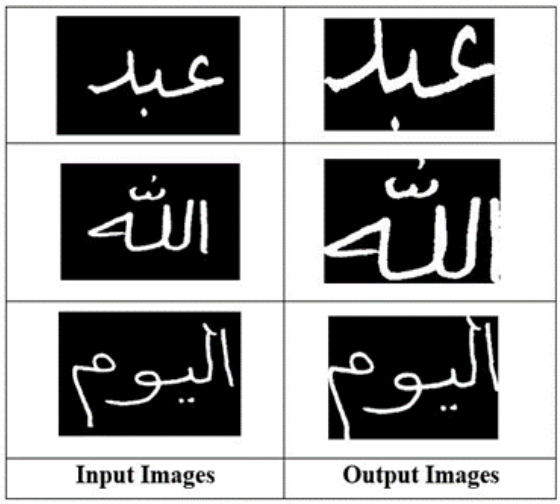

Fig.3. Black space elimination

Moreover, image thinning is used which is the process of reducing image size by remove the redundant pixels without losing the representation of the original image. 3*3 mask used to scan the whole image and find the 4 connected pixels. After that the unaffected pixels are eliminate 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 [7] has been used.

The last step in preprocessing is the image normalization. AHDAB dataset has various image sizes. It important to make all the image in the dataset in the same size and make the recognition process fast. After testing several sizes (32x32, 64x64 and 128x128) the 128*128 size gave best recognition rate. Therefore, all the dataset images normalize 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 extractions methods. A lot of feature extractions methods has been proposed for recognition purpose. However, features are generally classified into three main categories: the structural features, statistical features and global transformations.

i. Structural Features

Structural features describe the geometrical and topological features of a pattern by describing its global and local properties. The structural features depend on the kind of pattern to be classified [8].

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

ii. Statistical Features

The statistical features are extracted from the statistical distribution of pixels which describe the characteristic measurements of the input image pattern. The statistical features provide low complexity and high speed. The major statistical features can be summarized as: histograms of chain code directions, pixel densities, moments, and, Fourier descriptors [9]. 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 figure4 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.

iii. Global Transformation

The transformation schemes convert the pixels transformation of the pattern to a more compact form which reduces the dimensionality of features [10]. In this research two affective transform methods are used for extract the text features, which are:

- 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:

$$f(u,v)=\alpha(u)\alpha(v)\sum_{I=0}^{I-1}\sum_{J=0}^{J-1}f(i,j)\cos\left[\frac{(2I+1)}{2I}\right]\cos\left[\frac{(2J+1)}{2J}\right] \quad (1)$$

The inverse transform is defined by:

$$f(i,j)=\sum_{U=0}^{N-1}\sum_{V=0}^{N-1}\alpha(u)\alpha(v)f(u,v))\cos\left[\frac{(2I+1)u\pi}{2N}\right]\cos\left[\frac{(2J+1)v\pi}{2J}\right] \quad (2)$$

Where

$$\alpha(u)=\alpha(v)=\begin{cases}\frac{1}{\sqrt{2}} & for \quad u,v=0\\ \frac{2}{\sqrt{2}} & for \quad u,v\neq 0\end{cases} \quad (3)$$

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 [11].

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 [12] 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 figure5 which gave better recognition rate than Sobel and Roberts filters.

$$\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} \qquad \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$$

Fig.5. proposed edge detection filter

HOG divides the input image into small square cells (here 6×6 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 Neural Network (NN).

i. NN Classifier

A neural network consists of units (neurons), arranged in layers, which convert an input vector into some output. Each unit takes an input, applies a (often nonlinear) function to it and then passes the output on to the next layer. Generally the networks are defined to be feed-forward: a unit feeds its output to all the units on the next layer, but there is no feedback to the previous layer. Weightings are applied to the signals passing from one unit to another, and it is these weightings which are tuned in the training phase to adapt a neural network to the particular problem at hand. This is the learning phase [13].

Neural networks are relatively crude electronic networks of "neurons" based on the neural structure of the brain. They process records one at a time, and "learn" by comparing their classification of the record (which, at the outset, is largely arbitrary) with the known actual classification of the record. The errors from the initial classification of the first record is fed back into the network, and used to modify the networks algorithm the second time around, and so on for many iterations.

Neurons are organized into layers: input, hidden and output. The input layer is composed not of full neurons, but rather consists simply of the record's values that are inputs to the next layer of neurons. The next layer is the hidden layer. Several hidden layers can exist in one neural network. The final layer is the output layer, where there is one node for each class. A single sweep forward through the network results in the assignment of a value to each output node, and the record is assigned to the class node with the highest value.

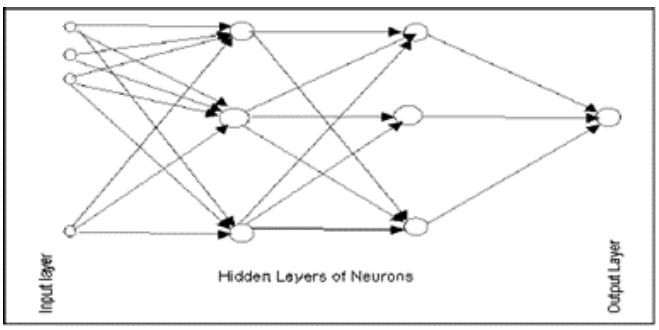

Fig.6. Neural Network Architecture

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.

Feature extraction technique is implemented and gives 405 feature points for each image accordingly. This procedure is repeated for 405x2044 samples and then stored in a feature vector of size 2044x405 which we have taken as input to our ANN as shown in Figure 7. Our the target output for ANN is 28x405, where which row represent a word.

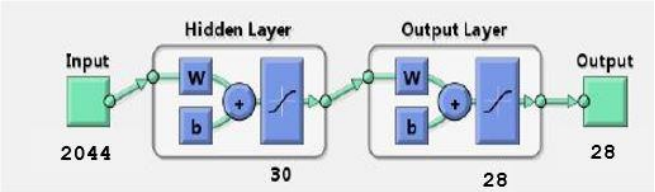

Fig.7. Neural Network Architecture

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

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 95% recognition accuracy.

Conclusion

In this paper we introduced 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 NN. 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.

References

[1] U.-V. Marti and H. Bunke, "A Full English Sentence Database for Offline Handwriting Recognition," Proceedings of the 5th International Conference on Document Analysis and Recognition (ICDAR 1999), pp. 705-708, September 1999.

[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] Mori S., Nishida H. & Yamada H. (1999) Optical Character Recognition. (JohnWiley & Sons, NY).

[6] Elzobi, Moftah, Ayoub Al-Hamadi, Zaher Al Aghbari, and Dinges, Laslo (2012) viewed 5 March 2015, at. http://www.iesk-ardb.ovgu.de/

[7] Niall, Gunter, Innsbruck: Optical Character Recognition. Informatics Research Institute (IRIS) at University of Salford(2011)

[8] V.Govindan and A.shevaprasad."Character Recognition – a review" Pattern Recognition, vol. 23, no.7 , pp 671-683 , 1990.

[9] Issam Bazzi ,Richard Schwartz and John Makhoul. An "Omnifont Open-Vocabulary OCR System for English and Arabic". IEE Trans. On pattern analysis and machine intteligance. Vol 21 , no. 6 , pp 495-504,1999

[10] 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

[11] JIANG, J., WENG, Y. & LI, P. (2006) Dominant colour extraction in DCT domain. Image and Vision Computing, 24, 1269-1277.

[12] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, 2005

[13] Jayanta Kumar Basu, Debnath Bhattacharyya, Tai-hoon Kim, "Use of Artificial Neural Network in Pattern Recognition", International Journal of Software Engineering and Its Applications, vol. 4, No. 2, pp. 24-32, 2010.