

# The Impact of Dots Representation in Recognition of Isolated Arabic Characters

**Nehad H A Hammad**

Palestine Technical College, P.O.Box 6037, Gaza, Palestine  
E-mail: nehadhh@hotmail.com

**Mohammed Elhafiz Musa**

Sudan University for Science and Technology, P.O.Box 382, Mokren, Khartoum, Sudan  
E-mail: hafiz85@hotmail.com

**Abstract**—The Arabic Optical Characters Recognition (AOCR) is one of the challenging recognition tasks nowadays, as Arabic handwriting is cursive and contains many dots. Dots are a big challenge for Arabic recognizers, as writers sometimes connect them. Moreover, Dots are prone to be considered noise. This paper proposes a new divide and conquers based approach that tries to conquer dots problems. The novelty of the proposed approach is in its feature extraction method. The extracted features are used to train Artificial Neural Network (ANN) Feed-forward. The result is interesting and shows that this method should be further investigated.

**Index Terms**—Arabic Character Recognition, Feature Extraction, Artificial Neural Network, Image Normalization, Image Resizing.

## I. INTRODUCTION

Handwriting recognition refers to the identification of written characters. The problem can be viewed as a classification problem where we need to give the input sample the appropriate class (each character is a class of infinite number of samples). The challenge of the recognition system increases with the increase of a number of classes.

One of the big challenges for AOCR is segmentation because Arabic handwriting is cursive, however, this problem is not addressed here because this paper addresses another challenge (dots). The main challenge for Arabic isolated character recognition is the recognition of the dots (its number and position). In the literature some authors use the term diacritics for these dots, however, the authors of this paper prefer dots. Dots in Arabic handwriting is very crucial as difference between the shapes of many characters is few dots. For example, letters Baa(ب), Taa(ت), and Thaa(ث) are three different letters, but they have similar body shape, they only differ in number and position of dots (one, two or three dots under or above the body of the character). Also the letter Jeem(ج), Haaa(ح) and Khaa(خ) they differ only in one dot, and the same situation happens other characters [1].

The method proposed here uses "Divide and conquer" technique to divide Arabic character into four groups based on a number of components in characters. Each character may have up to four components. Number of components is not fixed as some writers connect some components. (e.g. some writers draw dots (one, two, three) in a continuous line). Omer et. al. has explored similar ideas. They have broken the recognition into two stages. In their experiments they have used the same datasets used in this paper. However, the result achieves here in much better [14].

The connected component labeling algorithm using to determine the character member of which group, works by scanning an image, pixel-by-pixel (from top to bottom and left to right) in order to identify connected pixel

In this paper author represent a new method for handwritten characters passed in several processes start with preprocessing processes until recognition, the famous process is binarization, resizing, noise removal, thinning, feature extraction and classification for common Arabic character recognition.

This paper describes the Arabic optical character recognition (AOCR) method used a data collected from a large group of people was selected a random sample from dataset.

A multi-layer Artificial Neural Network was used to classification of Arabic handwritten characters.

## II. RELATED WORK

Ahmed Sahlol and Cheng Suen (2013) proposes new methods for handwritten Arabic character recognition which is based on novel preprocessing operations including different kinds of noise removal also different kind of features like structural, Statistical and Morphological features from the main body of the character and also from the secondary components. Evaluation of the accuracy of the selected features is made. The system was trained and tested by the back propagation neural network with CENPRMI dataset. The proposed algorithm obtained promising results as it is able to recognize 88% of our test set accurately. [11]

Ved Prakash Agnihotri (2012) proposed recognition system using neural network. Diagonal based feature

extraction is used for extracting features of the handwritten Devanagari script. After that these feature of each character image is converted into chromosome bit string of length 378. More than 1000 sample is used for training and testing purpose in this proposed work. It is attempted to use the power of genetic algorithm to recognize the character. Diagonal based feature extraction method to extract 54 features to each character. In the next step character recognize image in which extracted feature in converted into Chromosome bit string of size 378. In recognition step using fitness function in which find the Chromosome difference between unknown character and Chromosome which are store in data base. The experiment is conducted on more than 1000 characters. The testing characters are separated into data sets, the training data set and testing data set. The training set contains 904 characters and testing set contains 204 characters. The precision of offline Devanagari system is 85.78% match, 13.35% mismatch. [15]

Khaoula addakiri and mohamed bahaj (2012) represent an approach for the recognition of on-line Arabic handwritten characters. The method employed involves three phases: First, pre-processing in which the original image is transformed into a binary image. Second, training neural networks with feed-forward back propagation algorithm. Finally, the recognition of the character through the use of Neural Network techniques. The characters recognition has served as one of the principal proving grounds for neural network methods and has emerged as one of the most successful applications of this technology. The experimental results for characters' recognizer accuracy average about 83%. [12]

Alaei et al. (2010) proposed a two-stage approach for isolated handwritten Persian character recognition. They extracted features based on modified chain code directional frequencies and employed an SVM for classification. They obtained 98.1% and 96.6% recognition accuracy with 8-class and 32-class problems, respectively [4].

Desai (2010) presented a technique for Gujarati handwritten numeral recognition. the author used features abstracted from four different profiles of digits with a multilayered feed-forward neural network and achieved an approximate 82% recognition accuracy for Gujarati handwritten digit identification [5].

Sharma and Jhaji (2010) extracted zoning features for handwritten Gurmukhi character recognition. They employed two classifiers, namely k-NN and SVM. They achieved a maximum recognition accuracy of about 72.5% and 72.0% with k-NN and SVM, respectively [6].

### III. THE PROPOSED METHOD

As the novel part of the proposed method is in feature extraction, here we will try to give it full description while standard steps like preprocessing will have small summary.

#### 1. Preprocessing

The feature extraction step require a one pixel width trace, however, the input image is gray scale letter image.

Therefore standard preprocessing methods are used here to generate these traces. These are preprocessing steps are: binarization, thinning, noise removal, resizing. These steps are depicted in Fig 1, Fig 2, Fig 3 and fig. 4 respectively.

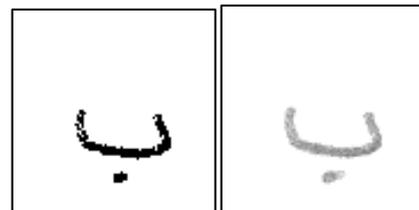


Fig.1. Binarization Step using Threshold.

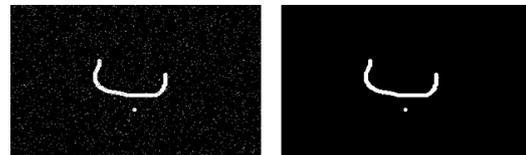


Fig.2. Noise Removing using Median Filter.

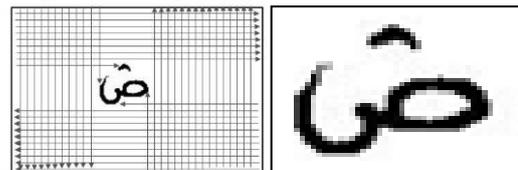


Fig.3. Binary Image Normalization(Cropping).



Fig.4. Arabic Isolated Character Thinning.

#### 2. Connected Components

A pixel  $p$  at coordinates  $(x,y)$  has adjacent neighbors whose coordinates are  $(x+1,y)$ ,  $(x-1,y)$ ,  $(x,y+1)$  and  $(x,y-1)$ . This set of 4 neighbors of  $p$  denoted  $N_4(p)$  is illustrated in Fig. 5, where the four diagonal neighbors of  $p$  have coordinates  $(x+1,y+1)$ ,  $(x+1,y-1)$ ,  $(x-1,y+1)$  and  $(x-1,y-1)$  respectively.

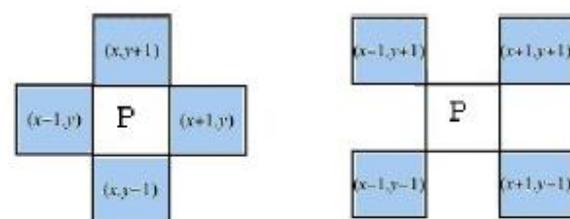


Fig.5. (a) Adjacent neighbors (b) Diagonal neighbors

Connected component extraction is an important step in this method. Most of the Arabic characters contain one or more connected components like ( ا , ب , ت , ث , ذ , خ , ي , ك , ق , غ , ض ). Fig 6. Shows the groups of characters

used in the proposed methods.

The Proposed method identifies the main body easily as it is usually the largest component. Therefore, all other component is much smaller

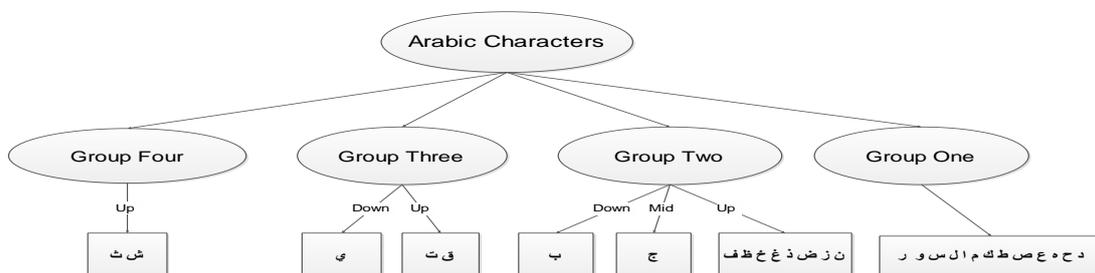


Fig.6. The main 4 groups and their constituent characters

### 3. Feature extraction

The proposed method generates a feature vector that contains 17 values 15 represent characters' contours and 2 represent groups the position of their dots (fig 8).

$$Angle(Q) = \text{Arctan}((P_n(y) - P_{n+1}(y)) / (P_n(x) - P_{n+1}(x))) \tag{1}$$

After thinning ,15 points are taken from character trace, the angles between two adjacent points of 15 points is calculated using Equation (1). To find the most appropriate number of directions 10 dirctions are proposed to tested. Table 1 illustrates these directions.

Table 1. The Selected Directions to be used in Proposed Methods

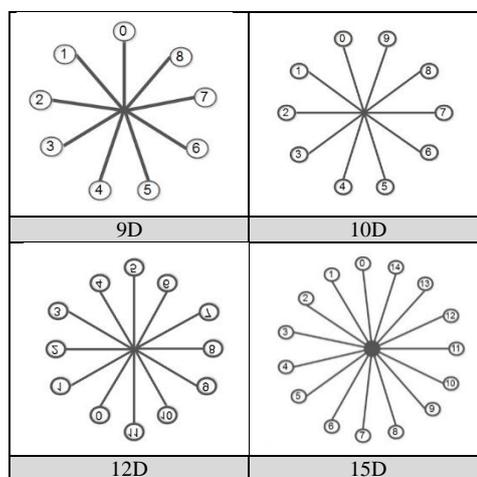
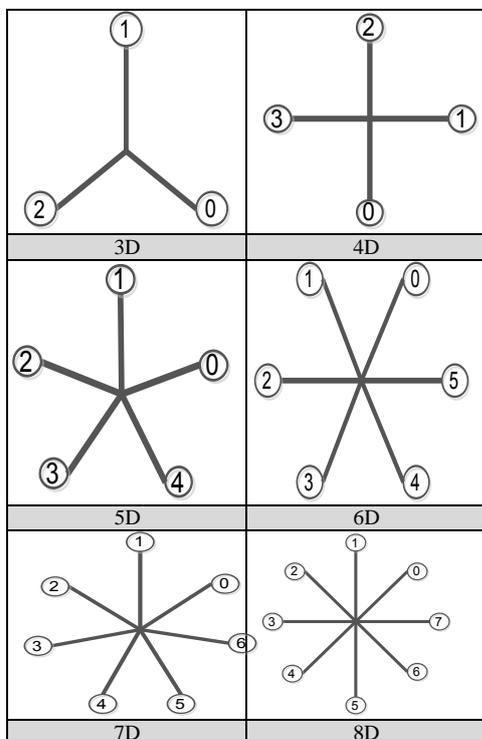


Table 2. illustrates how the first two features in feature vectors are generated (Fig. 8). The groups appear in Fig. 6, can be textually described for more illustrations as follows:

- group 1: one component characters. (Fig. 11)
- group 2: two component characters. (Fig. 13)
- group 3: three component characters. (Fig. 15)
- group 4: four component characters. (Fig. 17)

There is important subtle point we want to re-emphasize Fig. 8 presents the typical situation (typewriter letters). However as some writers connect some components (typically dots). Some letters appear in other group. For instance the letter ( ث ) in Fig 8 appears in group 4. However, this letter appear in group 2 and group 3 (see the last character in Fig 13 and Fig 15). As in group 2 the writer connects the three dots and group 3, he connects only 2.

Table 2. Illustrates these Categories and how the first two elements in the features vectors are filled.

Number of Dots	One Dot	Two Dots	Three Dots
No Dots	[00]	[00]	[00]
Level 1 (Up)	[11]	[12]	[13]
Level 2 (Mid)	[21]	-	-
Level 3 (Bottom)	[31]	[32]	-

Dot Detection process:

1. Calculate the number isolated objects in the image.
2. If a number of objects = 1 then pattern belongs to group one.
3. If number of objects  $\geq 2$  then
4. The Longest object size is character body and other is a dot(s).
5. The character dots recognized using another process to calculate a number of dots and position level.
6. The number of dots depends on dot length using thresholds and dots shape.
7. The dots level extracted when to remove all space around character after divide the character height into 3 areas on the Y axis.

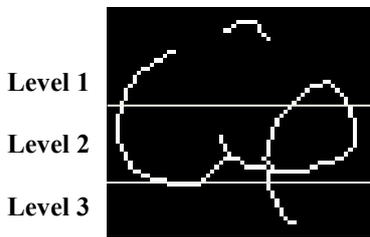


Fig.7. Character Dots Levels

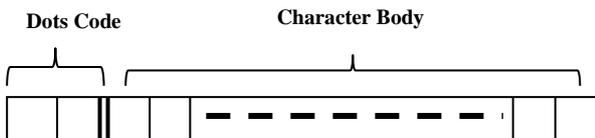


Fig. 8. Feature Vector Structure Design



Fig.9. Arabic Character Representation

4. Neural Network Architecture

The architecture of the proposed neural network has 17 inputs and 50 neurons in its output layer to identify the letters.

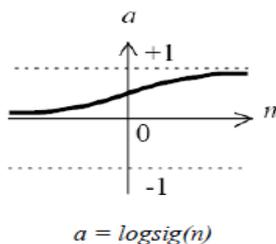


Fig.10. Log-sigmoid activation function.

The network is a two-layer network. The log-sigmoid [50] activation function at the output layer was picked because its output range (0 to 1) is perfect for learning to

output Boolean values.

The function generates outputs between 0 and 1 as the neuron's net input goes from negative to positive infinity see figure 10.

In building the network, the data was divided randomly into three categories. Training data consisted of 70% of the data. The remaining 30% of the data was assigned 15% to validation and 15% for the testing data.

After many trials of eliminating, adding and modifying features and also adjusting network hidden layers: A network of 100 neurons in hidden layer was able to predict about 96.91% of the input characters correctly.

IV. DATASET

A Datasets of Arabic isolated characters were collected from students at the Sudan University for Science and Technology (SUST). After paper scanning, a manual classification process was performed to establish the desired standard dataset [13]. The datasets contain 28 isolate characters. Each character appears in various shapes as the subjects has not been asked to write in specific way. The dataset is off-line. Each subject contributed 28 basic characters. The total number of samples for one character is 1400. The total number of samples in the dataset is 39200 ( = 28 × 1400 ). The images resolution is 88x88 pixels [13].

V. EXPERIMENTS AND RESULT

Here the results for each group is presented. As 10 directions have been tested with three numbers of points in each character ( 10,15 and 20) all results table show the results of these 30 selections (i.e. 10x3). Also for each group two result tables are presented, the first one shows the overall recognition result and the second shows the result for each character.

1. Group One

Group one contains any isolated Arabic characters with one part so the character written using one connected line, this group contains 12 characters ( , ط , ص , ه , ع , ح , د , ا , م , ل , س , و , ر ).

ا	د	ح
و	س	ل
م	ع	ه
ط	ص	ر

Fig.11. Group One Isolated Arabic Characters Handwritten

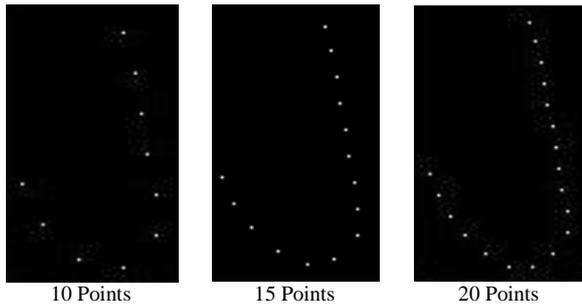


Fig.12. Group One Represented by 20,15 and 20 Points.

1.1 Accuracy Rate for Group One in Multi Directions

Table 3. Accuracy Rate for Group One Directions.

Points Number	10	15	20	Best Accuracy	
Direction	3	82.24	81.31	82.96	82.96%
	4	82.90	85.75	87.09	87.09%
	5	87.74	84.25	87.41	87.74%
	6	85.37	84.89	85.65	85.65%
	7	86.00	83.22	87.62	87.62%
	8	90.63	88.11	90.43	90.63%
	9	92.53	93.06	92.28	93.06%
	10	90.57	94.34	90.89	94.34%
	12	92.44	92.86	92.78	92.86%
	15	92.25	91.85	92.80	92.80%

Therefore, the best representation of characters to extract features in the group one is 10-directions and represent the characters using 15 points to get the recognition rate 94.34% see Table 3.

1.2 Accuracy Rate for Group One Characters

The character recognition is affected by the extent of a character complexity of character shape, the degree of curvature, crosses and overlaps. Therefore, the Heh (هـ) character is the most difficult to drawing between characters in group one and Alef(ا) is an easier character for drawing. So the character’s recognition accuracy in the group one are varies due to the foregoing factors.

Table 4. describes more detail the variation in recognition rate between characters in the group one, where represents characters using 10, 15 and 20 points.

The saad(ص),Heh(هـ) and Tah(ط) have low recognition rate than easier character drawing like Alef(ا) , Lam(ل) and Reh(ر).

These results are logical where the recognition rate inversely proportional to the difficulty of the character drawing.

Table 4. Group One Characters Recognition Accuracy Rate.

No	Character	20 Points	15 Points	10 Points	Best Accuracy	
1	ع	Ain	86.60	86.56	87.48	87.48%
2	ا	Alef	98.54	98.38	93.74	98.54%
3	د	Dal	90.42	88.58	90.27	90.42%
4	ح	Hah	86.19	86.01	86.06	86.19%
5	هـ	Heh	77.22	75.17	73.95	77.22%
6	ل	Lam	93.96	91.74	93.44	93.96%
7	م	Meem	91.10	90.69	85.33	91.10%
8	ر	Reh	90.97	89.77	92.32	92.32%
9	ص	Saad	86.70	84.33	84.85	86.70%
10	س	Seen	88.16	88.72	88.58	88.72%
11	ط	Tah	84.15	82.54	80.71	84.15%
12	و	Waw	89.24	85.67	86.59	89.24%

2. Group Two

Group two contains isolated Arabic characters with two parts so the characters are written using two connected line. Group two contains 16 characters ( , ب , ت , ث , ج , ح , خ , غ , ف , ق , ض , ذ , ز , ن , ي , ش , ط , ز ) .

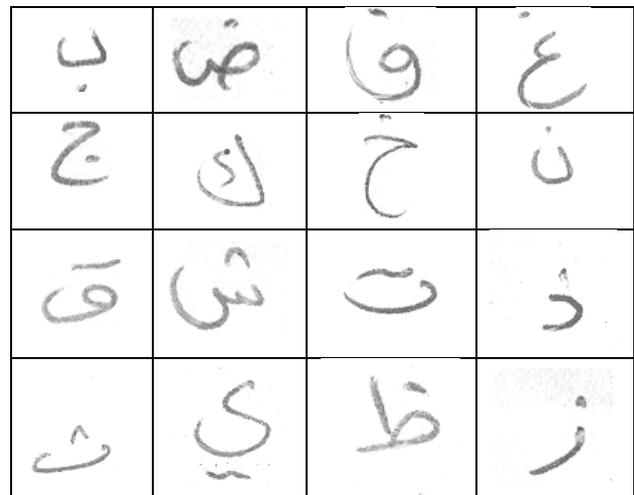


Fig.13. Group Two Isolated Arabic Characters Handwritten.

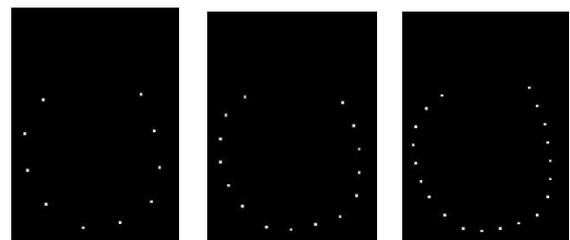


Fig.14. Arabic Character Group Two Represented by 20,15 and 20 Points.

2.1 Accuracy Rate for Group Two Multi Directions

Table 5. Accuracy Rate for Group Two Direction.

Points Number	10 Points	15 Points	20 Points	Best Accuracy	
Directions	3	93.35	93.07	<b>97.53</b>	<b>97.53 %</b>
	4	93.77	<b>94.40</b>	93.60	<b>94.40 %</b>
	5	92.95	<b>94.16</b>	91.63	<b>94.16 %</b>
	6	<b>95.04</b>	94.58	92.29	<b>95.04 %</b>
	7	92.89	<b>94.73</b>	93.32	<b>94.73 %</b>
	8	92.35	<b>95.48</b>	93.59	<b>95.48 %</b>
	9	94.45	<b>96.28</b>	94.21	<b>96.28 %</b>
	10	92.75	<b>95.86</b>	95.26	<b>95.86 %</b>
	12	94.88	<b>96.00</b>	94.53	<b>96.00 %</b>
	15	96.71	<b>97.90</b>	95.28	<b>97.90 %</b>

Therefore, the best representation of characters to extract features in the group two is 15-directions and represent the characters using 15 points to get the recognition rate 97.90% see Table 5.

2.2 Accuracy Rate for Group Two Characters

The character recognition is affected by the extent of a character complexity of character shape, the degree of curvature, crosses, overlaps, dots counts and dots position. The characters in group two have one, two or three dots used to distinguish between similar characters.

Therefore, the Zeh (ظ) character is the most difficult to drawing characters with one dot in upper like most of the character's dots in group two, but Yeh (ي) character is approximately easy for drawing, but the number of dot(s) and position leading to distinguish it from the other characters in the group two.

It turns out that the strategy of representation dots using dots number and dots location in character leading to make classification easier because most of the isolated Arabic characters have dots so some specific character has individual dots take high recognition rate such as Jeem(ج) one dot in middle, Yeh(ي) two dots in bottom, Beh(ب) one dot in bottom. So the character's recognition accuracy in the group two is varied due to the foregoing factors.

Table 6. describe more detail the variation in recognition rate between characters in the group two, where represents characters using 10, 15 and 20 points.

The Zah(ض),Zeh(ظ) and Khah(خ) have low recognition rate than characters like Yeh(ي), Beh(ب), Jeem(ج) and Kaf(ك) because these characters have special characteristics facilitating recognition process.

These results are logical where the recognition rate inversely proportional to the difficulty of the character drawing and directly proportional to the characters that contain distinct dot(s).

Table 6. Group Two Characters Recognition Accuracy Rate.

No	Character	20 Points	15 Points	10 Points	Best Accuracy
1	ب Beh	97.94	98.71	98.83	98.83%
2	ض Zah	90.57	89.65	86.38	90.57%
3	ف Feh	94.35	93.57	93.35	94.35%
4	غ Ghen	93.60	94.17	94.71	94.71%
5	ج Jeem	96.96	98.45	97.73	98.45%
6	ك Kaf	98.65	99.08	99.03	99.08%
7	خ Khah	89.54	91.02	89.19	91.02%
8	ن Noon	95.26	95.73	93.89	95.73%
9	ق Qaf	96.57	98.27	97.23	98.27%
10	ش Sheen	94.71	96.50	95.17	96.50%
11	ت Teh	94.22	95.34	94.02	95.34%
12	ث Theh	94.45	94.85	93.09	94.85%
13	ث Thah	93.96	95.51	92.89	95.51%
14	ي Yeh	98.46	99.67	99.82	99.82%
15	ظ Zeh	82.13	84.61	80.26	84.61%
16	ز Zen	91.62	92.48	92.09	92.48%

3. Group Three

Group three contains any isolated Arabic characters with three components so the character written using three connected line, This group contains five characters ( ق, ت, ش, ي, ث).

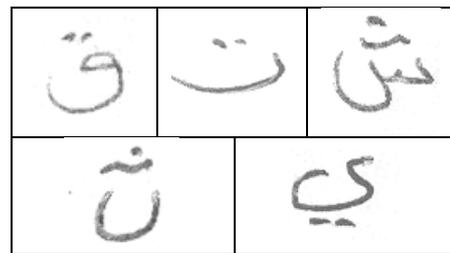


Fig.15. Group Three Isolated Arabic Characters Handwritten.

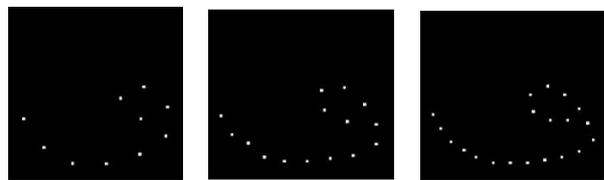


Fig.16. Arabic Character Group Three Represented by 20,15 and 20 Points.

3.1 Accuracy Rate for Group Three Multi-Directions

Table 7. Accuracy Rate for Group Three.

Points Number	10	15	20	Best Accuracy	
Direction	3	97.12	98.67	98.58	98.67%
	4	96.51	97.75	96.13	97.75%
	5	96.05	97.61	97.62	97.62%
	6	97.44	97.97	98.23	98.23%
	7	96.34	98.48	97.62	98.48%
	8	95.98	97.58	97.69	97.69%
	9	96.91	96.98	96.35	96.98%
	10	97.29	97.44	95.95	97.44%
	12	96.72	97.67	96.63	97.67%
	15	97.01	98.19	97.40	98.19%

Therefore, the best representation of characters to extract features in the group three is 3-directions and represent the characters using 15 points to get the recognition rate 98.67% see Table 7.

3.2 Accuracy Rate for Group Three Characters

Group three contain five characters Qaf(ق), Sheen(ش), Teh(ت), Theh(ث), Yeh(ي) each member have three parts one for body and other two is dots.

The character recognition is affected by the extent of a character complexity of character shape, the degree of curvature, crosses, overlaps, dots counts and dots position.

The characters in group three have two or three dots used to distinguish between similar characters like Teh(ت) and Theh(ث) have same contour and dots level but different in dots count, this dots count able to distinguish between them.

Therefore, the Sheen (ش) and Theh(ث) have low accuracy rate because they have the same dots counts and level in group three, but Yeh (ي) and Qaf(ق) have high accuracy rate because they have different dots level.

Table 8. Group Three Characters Recognition Accuracy Rate.

No	Character	20 Points	15 Points	10 Points	Best Accuracy
1	ق Qaf	99.40	99.57	99.60	99.60%
2	ش Sheen	94.19	94.96	92.69	94.96%
3	ت Teh	98.51	98.96	98.10	98.96%
4	ث Theh	92.89	94.73	91.94	94.73%
5	ي Yeh	99.96	99.90	99.97	99.97%

Table 8. describe more detail the variation in recognition rate between characters in the group three, where represents characters using 10, 15 and 20 points.

The Qaf(ق), Teh(ت) and Yeh(ي) have high recognition rate than characters like Sheen(ش) and Theh(ث) because some characters have special characteristics facilitating

recognition process.

These results are logical where the recognition rate inversely proportional to the difficulty of the character drawing and directly proportional to the characters that contain distinct dot(s).

4. Group Four

The group four is contained any isolated Arabic characters with four parts so the character written using four connected line, This group has two characters (ش, ث).



Fig.17. Group Four Isolated Arabic Characters Handwritten.

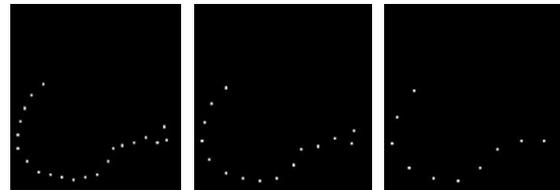


Fig.18. Arabic Character Group Four Represented by 20,15 and 20 Points.

A therefore, the best representation of characters to extract features in the group four is 3-directions and represent the characters using 10 ,15 and 20 points to get the recognition rate 100% see Table 13.

4.1 Accuracy Rate for Group Four Characters

Group Four contain two characters Sheen(ش), Theh(ث) each member have four parts one for character body and other three is dots.

The character recognition is affected by the extent of a character complexity of character shape , the degree of curvature, crosses , overlaps , dots counts and dots position.

The characters in group four have only three dots on the same level for both characters so the main factor to classified them by using characters' contour feature when the dots is similar to distinguish between them.

Therefore, the Sheen (ش) and Theh(ث) have high accuracy rate because they have different contour facilitate for the classifier to recognize between them.

Table 9. Group Four Characters Recognition Accuracy Rate.

No	Character	20 Points	15 Points	10 Points	Best Accuracy
1	ش Sheen	96.19	98.28	96.75	98.28%
2	ث Theh	95.57	97.90	95.86	97.90%

Table 9. describe more detail the variation in recognition rate between characters in the group four, where represents characters using 10, 15 and 20 points.

The Sheen(ش) and Theh(ث) have very high recognition rate because group contains only two characters so easier to classify them using character contour.

5. Proposed Approach Results Accuracy

The proposed approach accuracy includes the accuracy of four classifiers for four Arabic characters' group. The first, a neural network (NN) classifier for group one contains 12 characters the result about 94.34% recognition rate. A second neural network (NN) classifier for group two contains 16 characters the result about 97.9% recognition rate. A third neural network (NN) classifier for group three contains 5 characters the result about 98.67% recognition rate. A fourth neural network (NN) classifier for group four contains 2 characters the result about 100% recognition rate.

Total Recognition Accuracy

$$= (Group1\ Rate * Characters\ No + Group2\ Rate * Characters\ No + Group3\ Rate * Characters\ No + Group4\ Rate * Characters\ No) / (Total\ Characters\ No)$$

(2)

$$= \sum_{i=1}^4 R(i) * NoC(i) / Total\ Character\ No$$

(3)

Table 10. Total Recognition Accuracy Applied on SUST Dataset.

	Classifier1	Classifier2	Classifier3	Classifier4	Total Character
Character No	12	16	5	2	35
Accuracy Rate	94.34	97.9	98.67	100	Average Rate
Rate Balance	1132.08	1566.4	493.35	200	96.91%

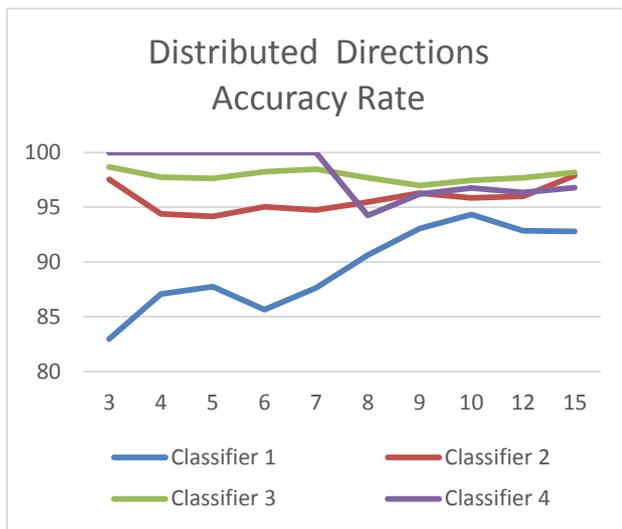


Fig.19. Arabic Characters Group Classifier Accuracy on Multi-Direction.

VI. CONCLUSION AND FURTHER WORK

Arabic character recognition methods are affected by several factors like characters' shape, the complexity of drawing the character and the amount of rotation of the character on the horizontal line. Moreover, Arabic characters contain many dots this increase the challenge of building accurate Arabic classifiers. To overcome these problems, we propose divide and conquer based recognition system the proposed method divide the Arabic characters into 4 groups. A neural network has been designed to classify each group. The result is encouraging and more accurate results could be achieved. For instance, by increasing the dataset size. And testing other powerful classifiers like SVM.

REFERENCES

- [1] Abdurazzag Ali ABURAS and Salem M. A. REHIEL, (2007), "Off-line Omni-style Handwriting Arabic Character Recognition System Based on Wavelet Compression", International Islamic University Malaysia, Electrical and Computer Engineering, PP10, 50728, 53100.
- [2] Haraty, R. and Ghaddar, C. Neuro-Classification for Handwritten Arabic Text. Proceedings ACS/IEEE International Conference on Computer Systems and Applications, 2003.
- [3] Amin, A. Recognition of Hand-Printed Characters Based on Structural Description and Inductive Logic Programming. Pattern Recognition Letters, vol. 24, pp. 3187-3196, 2003.
- [4] A. Alaei, P. Nagabhushan, U. Pal, "A new two-stage scheme for the recognition of Persian handwritten characters," in Proc. of 12th ICFHR, pp.130-135, 2010.
- [5] A. A. Desai, "Gujarati handwritten numeral optical character reorganization through the neural network," Pattern Recognition, vol. 43, no. 7, pp. 2582-2589, July 2010.
- [6] D. V. Sharma, P. Jhaji, "Recognition of isolated handwritten characters in Gurmukhi script," International Journal of Computer Applications, vol. 4, no. 8, pp. 9-17, 2010.
- [7] M. Kumar, M. K. Jindal, R. K. Sharma, "k-NN based offline handwritten Gurmukhi character recognition," in Proc. of ICIIP, pp. 1-4, 2011.
- [8] Jia Yonghong, "Digital image processing (The Second Edition)". WuHan China: Wu Han university press, pp 114-116, 2010.
- [9] Lam, Seong-Wan Lee, and Ching Y. Suen, "Thinning Methodologies-A Comprehensive Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 14, No. 9, 1992.
- [10] Kohavi, R., "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and model Selection", Proceedings Of the 15th International Conference on Artificial Intelligence (IJCAI). pp. 1137-1143, 1995.

- [11] Ahmed Sahlol and Cheng Suen, "A Novel Method for the Recognition of Isolated Handwritten Arabic Characters", Department of Computer Teacher preparation, Damietta University, Damietta, Egypt and Department of Computer Science, Concordia University, Canada, 2013.
- [12] Khaoula Addakiri and Mohamed Bahaj, "On-line Handwritten Arabic Character Recognition using Artificial Neural Network", International Journal of Computer Applications (0975– 8887) Volume 55– No.13, October 2012.
- [13] Mohamed E. M. Musa, "Arabic Handwritten Datasets for Pattern Recognition and Machine Learning", Application of Information and Communication Technologies (AICT 2011) – Baku, Azerbaijan.
- [14] Omer Balola, adnan Shaout, Mohamed E. M. Mustafa "Two stage classifier for Arabic Handwritten Character Recognition" International Journal of Advanced Research in Computer and Communication Engineering, ol. 4, Issue 12, December 2015.
- [15] Ved Prakash Agnihotri, "Offline Handwritten Devanagari Script Recognition ", International Journal of Information Technology and Computer Science (IJITCS), Vol 8, PP 37-42, 2012.

### Authors' Profiles



**Nehad H A Hammad**, Assistant Professor at Palestine Technical College, Gaza- Palestine. I obtained Bachelor degree from Islamic University of Gaza, Palestine in 2002. Master degree from Near East University, Cyprus 2005, Ph.D. degree from Omdurman Islamic University, Sudan 2015, My research interests include pattern recognition, optical character recognition and image processing.



**Mohamed Elhafiz Musa**, Associate Professor at Sudan University for Science and Technology, Sudan-Khartoum. His research interests include Artificial Neural Network, pattern recognition, data mining and image processing.

**How to cite this paper:** Nehad H A Hammad, Mohammed Elhafiz Musa, "The Impact of Dots Representation in Recognition of Isolated Arabic Characters", International Journal of Information Engineering and Electronic Business (IJIEEB), Vol.8, No.6, pp.37-45, 2016. DOI: 10.5815/ijieeb.2016.06.05