



ISSN: 0067-2904

A Design of a Hybrid Algorithm for Optical Character Recognition of Online Hand-Written Arabic Alphabets

Waleed Noori Hussein¹, Haider N.Hussain²

¹AL-Zahraa Medical College, University of Basra, Basra, Iraq

²Department of Mathematics, College of Science, University of Basra, Basra, Iraq

Abstract

The growing relevance of printed and digitalized hand-written characters has necessitated the need for convalescent automatic recognition of characters in Optical Character Recognition (OCR). Among the handwritten characters, Arabic is one of those with special attention due to its distinctive nature, and the inherent challenges in its recognition systems. This distinctiveness of Arabic characters, with the difference in personal writing styles and proficiency, are complicating the effectiveness of its online handwritten recognition systems. This research, based on limitations and scope of previous related studies, studied the recognition of Arabic isolated characters through the identification of its features and dots in view of producing an efficient online Arabic handwriting isolated character recognition system. It proposes a hybrid of decision tree and Artificial Neural Network (ANN), as against being combined with other algorithms as found in previous studies. The proposed recognition process has four main steps with associated sub-steps. The results showed that the proposed method achieved the highest performance at 96.7%, whereas the benchmark methods which are EDMS and Naeimizaghiani had 68.88% and 78.5 % respectively. Based on this, ANN has the best performance recognition rate at 98.8%, while the best rate for decision tree was obtained at 97.2%.

Keywords: Hand-written characters; Artificial Neural Network; Decision Tree; Optical Character Recognition, Online Arabic character recognition.

تصميم خوارزمية هجينة للتعرف على الحروف الابجدية العربية المكتوبة بخط اليد على الانترنت

وليد نوري حسين*¹، حيدر نوري حسين²

¹كلية الزهراء الطبية، جامعة البصرة، البصرة، العراق

²قسم الرياضيات، كلية العلوم، جامعة البصرة، البصرة، العراق

الخلاصة

نظرا لأهمية الأحرف المكتوبة بخط اليد الرقمية والمطبوعة هناك حاجة إلى التعرف التلقائي على الأحرف من خلال التعرف الضوئي على الحروف. ومن بين الأحرف المكتوبة بخط اليد، تعد اللغة العربية واحدة من تلك التي تحظى باهتمام خاص بسبب طبيعتها المميزة والتحديات الكامنة في أنظمة التعرف عليها. ان هذا التمييز بين الحروف العربية، مع اختلاف أساليب الكتابة لدى الأشخاص، يؤدي إلى تعقيد فعالية أنظمة التعرف على خط اليد على الإنترنت. في هذا البحث قمنا وبالاعتماد على ماتوصلت له الدراسات السابقة، بدراسة التعرف على الأحرف العربية المعزولة من خلال تحديد معالمه ونقاطه في ضوء إنتاج نظام

التعرف على الأحرف المعزول بخط اليد على الإنترنت. تتكون عملية التعرف المقترحة على أربع خطوات رئيسية مع الخطوات الفرعية المرتبطة بها. حيث أظهرت النتائج أن الطريقة المقترحة حققت أعلى أداء عند 96.7 % ، في حين أن الطرق القياسية وهي EDMS و Naeimzaghiani كانت 68.88 % و 78.5 % على التوالي. بناءً على ذلك ، حصلت ANN على أفضل معدل للتعرف على الأداء بنسبة 98.8 % ، في حين تم الحصول على أفضل معدل لشجرة القرار بنسبة 97.2 %.

1. Introduction

Optical character recognition (OCR) is one of the areas of applied studies of pattern recognition [1]. Pattern recognition uses machine learning models and various mathematical, statistical and heuristic techniques for the detection of patterns, and provides a computational implementation framework for its recognition systems [1, 2]. OCR is employed in document scanners mostly used in banking and postal applications. The growing relevance of the printed and digitalized hand-written characters is, however, presently necessitating the need for convalescent automatic recognition of characters [3, 4]. This is evolving into a new research area hand-written character recognition within the pattern recognition body of knowledge [5, 6]. The reasons for this novel computational attention to hand-written character recognition researches are not far-fetched. First, OCRs are fundamentally used for sensing an offline text or character recognition, while the movements of pens are sensed for online text [7, 8]. Now, the OCR systems must be improved to attend to the available text-to-digital character format in online systems. Second, hand-written character recognition is now employed and needed in automated mail sorting and check processing, and improving their performance is important [3, 9]. This is applicable to all characters, and Arabic attracts special attention. Arabic alphabets, which consist of 28 characters, form words being written from right to left [8, 10]. Hand-written Arabic characters, in OCR, demand research studies that will birth development of recognition algorithms and systems because of the distinctive characteristics of Arabic alphabets. The Arabic text is cursive and written from right to left [9, 11]. It contains dots and other small marks that can completely change its meaning if wrongly captured by automated recognition systems [10,12]. Also, Arabic letters change in structure and shapes based on their positions in the words. An Arabic letter will have distinct shapes and forms when at the beginning, middle, and the end of a word [12, 13]. The distinctiveness of Arabic characters, especially the difficulties associated with automated recognition, and complicated with its online hand-written characters, necessitated this study [14]. Arabic character recognition can be problematic and difficult, especially when it is hand-written, with varying attendant imprecisions of the inclusions and locations of its dots. Arabic characters, when scribbled by different persons could present different shapes and sizes, even when the characters are the same. This is as a result of individual preferences of writing styles [15,16,17]. To attend to this, automated character recognition must be able to recognize the components edges and dots.

This research studied the recognition of Arabic isolated characters by identifying the features and dots in view of producing efficient online Arabic handwriting isolated character recognition system. To achieve this, a hybrid algorithm of a decision tree and artificial neural network is proposed. The second section of this paper review related works. The third section presents the proposed hybrid algorithm for automated character recognition of hand-written Arabic characters, while the fourth section presents the experimental results and findings. The fifth section, which is the last section, presents the results and discussion of this study.

2. Related Works

The holistic and analytical approaches are the leading strategies in solving problems associated with character recognition [18,19, 20]. Holistic approaches entrench global attention of character image where the representation is dealt with the words rather than the individual characters [21, 22]. It is widely used for solving text recognition problems and is based on dynamic programming [23, 24]. The analytical approach, on the other hand, attends to words as a series of small-sized units. These units can be easily linked with characters to represent the vocabulary [25]. The analytical methods are further bifurcated into subclasses such as explicit segmentation and implicit segmentation [26]. In the former, the grapheme or pseudo-letter segmentation takes place prior to recognition and analytical methods and the later concurrently performs segmentation and recognition [27]. The analytical methods are the most popular and current approaches because they present a simplified method of building character recognition system, and understanding the underlying algorithms [13, 14, 28] are

some of the past studies on Arabic character recognition. Based on Hidden Markov models and its structural features extraction, Khorsheed [28] proposed a technique for handling online handwritten recognition of Arabic script, in a way that significantly extended El-Sheikh and Guindi [29] and El-Sheikh and El-Taweel [30] algorithms on Arabic handwritten characters and cursive words recognition. Khorsheed [28], presumed characters outcome from a consistent segmentation phase and referred to their positions as priori. Initial, medial, final, and isolated are their shapes names, based on the position in the word. These are further categorized into four subsets depending on the number of character's strokes. Mezghan [31], later introduced online recognition for Arabic characters using the Kohonen maps with corresponding confusion matrices. The advantage is that it prunes the characters of error-causing nodes, and to combine them consequently. Amin and Singh [8] and Amin [32], using the C4.5 machine learning system, presented a technique for the recognition of Arabic words and Chinese characters. The technique involves the sequence of digitization, pre-processing feature extraction, and classification[3]. However, they proposed an Arabic OCR system based on recognition-based segmentation which is part of Amin [32] pre-processing feature extraction. Bataineh and Sheikh [33], also presented a method for recognizing Arabic handwritten characters. Their approach used to employed global feature extraction based on a texture analysis approach using the extended EDMS namely EDM1 method. These studies, though containing a varying number of dataset and training sets, it showed the efficacy of decision tree and artificial neural network in their respective system. This study proposes a hybrid of a decision tree and artificial neural network, as against being combined with other algorithms as found in previous studies, in view of improving the online Arabic character recognition.

3. The Proposed Hybrid Algorithm

The 3×3 edge direction matrix (EDM1), in the first order relationship, was created in the first order relationship, with each cell containing a position within 0 until 315 - degree value. Secondly, the number of occurrence for each value is calculated for each value and used in determining the relationship of the scoped pixel in the edge image ledge (x,y) [1]. The algorithm is shown in the Figure-1.

```

For each pixel in Iedge (x, y)
If Iedge (x, y) is black pixel at center then
Increase number of occurrence at EDM1 (2, 2) by 1.
If Iedge (x+1, y) is black pixel at 0° then
Increase number of occurrence at EDM1 (2, 3) by 1.
If Iedge (x+1, y-1) is black pixel at 45° then
Increase number of occurrence at EDM1 (1, 3) by 1.
If Iedge (x, y-1) is black pixel at 90° then
Increase number of occurrence at EDM1 (1, 2) by 1.
If Iedge (x-1, y-1) is black pixel at 135° then
Increase number of occurrence at EDM1 (1, 1) by 1.

```

Figure 1- Algorithm for Calculating 3×3 Edge Direction Matrix [33]

Each pixel in the edge image is related with two pixels. As shown in Figure-2, for example, the scoped pixel presents 180° and 45° for X1 and X2 respectively

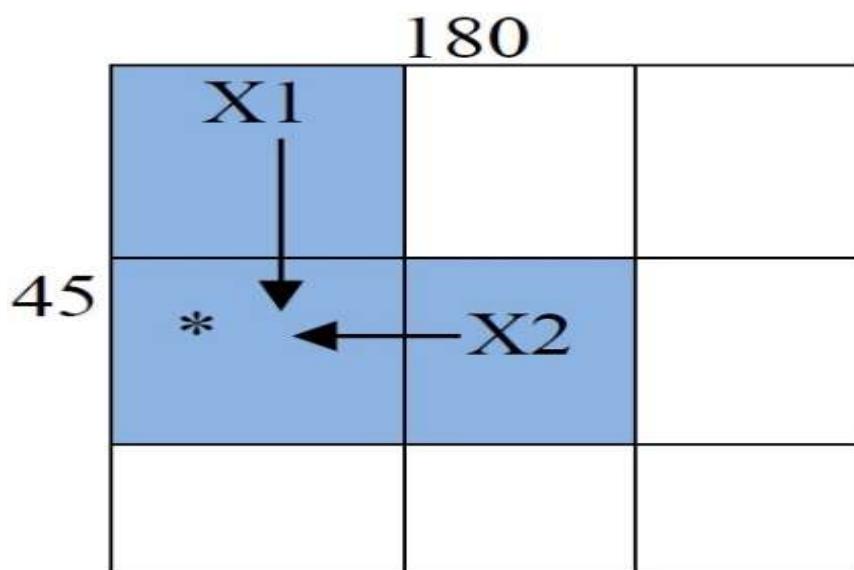


Figure 2- Two Neighboring Pixels

There are two relationships in each EDM1 pixel. Considering the relationship of the second order, one relationship only is presented by each pixel. The firstly-created is the 3×3 edge direction matrix (EDM2). Subsequently, the sorting of the EDM1 values determines the importance of the relationship of the Edge (x, y). This sorting is done in the descending order and presented in Table-1.

Table 1- The Order of importance Angle's

Order	Angle	Value
1	0	4
2	180	4
3	135	4
4	315	4
5	45	2
6	225	2
7	90	2
8	270	2

The most necessary relationship in the scoped pixel of Edge (x, y) is taken with the calculation of the number of occurrences experienced by each value in EDM2. This follows the following order: (a) when there are more angles (than one) of the same number of occurrences, it implies that the smaller angle is to be selected first, (b) the reversal is then selected next. Figure- 3 shows the algorithm of the second order of EDM2 relationship.

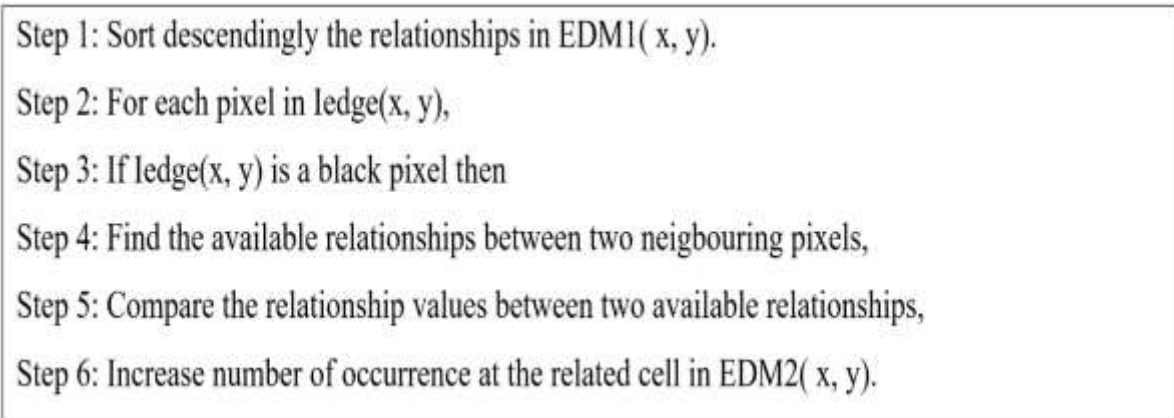


Figure 3- EDM2 Algorithm

The summary of EDM1 and EDM2 feature values are calculated by the homogeneity and edges regularity as presented.

Homogeneity: This feature is represented by the percentages of each direction to all directions that are available in the edge character as follows:

$$\text{Homogeneity}(\theta) = \text{EDM1}(x, y) / (\sum \text{EDM1}(x, y)) \tag{1}$$

Edges Regularity: This feature is represented by each real direction in EDM2 to the number of pixels scoped in the edge image as follows:

$$\text{Edges Regularity}(\theta^*) = \text{EDM2}(x, y) / \text{EDM2}(2, 2) \tag{2}$$

Where: $0^\circ, 45^\circ, 90^\circ$ and 135° are the values of θ ,

θ^* presents $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ$ and 135° , and (x, y) presents the relative position in EDM1 and EDM2.

On the other hand, in the extraction of Arabic character's features, and the associated statistical analysis, some geometrical features used which are the width and highest of the horizontal and vertical base, the number of occurrence in (H, V) projection, and width of the dot. The proposed flowchart for the geometrical features extractions presented in Figure-4.

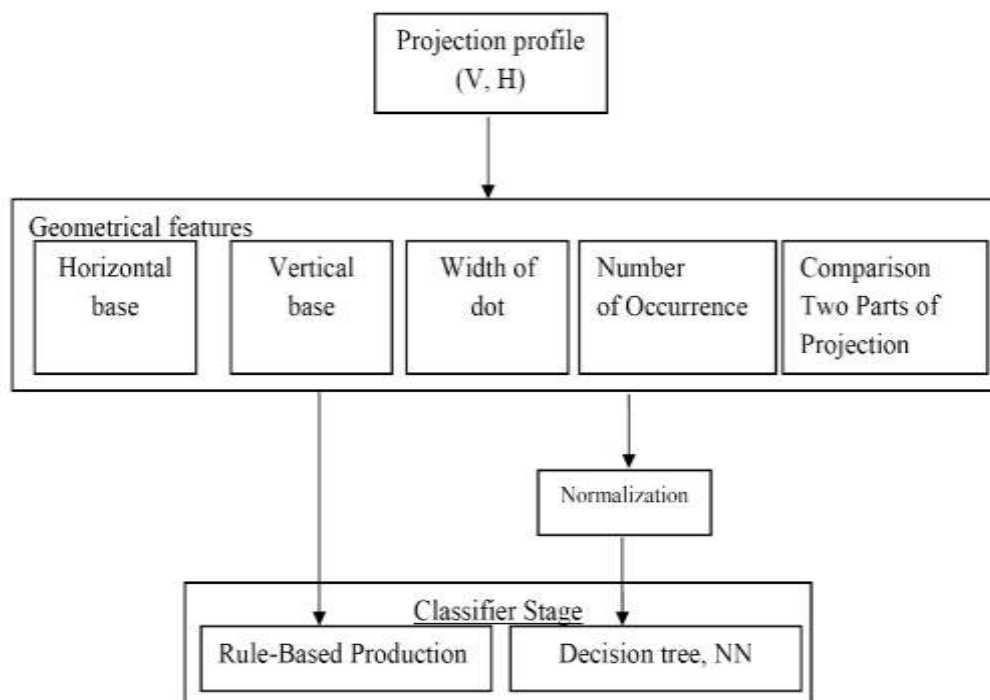


Figure 4- The proposed flowchart for Geometrical Feature Extraction

4. Experimental Results and Findings

In this study, the recognition process has four main steps with associated sub-steps. The steps are (a) Pre-processing (b) feature extraction (EDMS, Geometrical and structural Features), (c) rule-based technique for recognition, and (d) dataset collecting for testing the proposed method. In Table- 2, we presented the recognition result of each character and the total recognition rate for all characters. In this table, we showed the number of characters that completed 100% of a recognition rate for all users, and the number of characters that have a low recognition rate. Based on the results of the correction rates for each character shown in Table-2, the highest accuracy was achieved on user (1 and 4). The highest accuracy of 100% accuracy rate was achieved for most of the characters except , ط , ل , ج , ه , ف , غ , (ظ , ض , غ)

Table 2- Results for Testing Phase of All Four Users with Three Times Written

Char	User	User	User	User	Mean	STD
	1	2	3	4		
ا	1	1	1	1	1	0
ب	1	1	1	1	1	0
ت	1	1	1	1	1	0
ث	1	1	1	1	1	0
ج	1	1	1	1	1	0
ح	1	1	1	1	1	0
خ	1	1	1	1	1	0
د	1	1	1	1	1	0
ذ	1	1	1	1	1	0
ر	1	1	1	1	1	0
ز	1	1	1	1	1	0
س	1	1	1	1	1	0
ش	1	1	1	1	1	0
ص	1	1	1	1	1	0
ض	1	0.67	0.33	1	0.75	0.28
ط	1	0.67	0.67	1	0.84	0.17
ظ	1	1	0.67	0.67	0.84	0.17
ع	1	1	1	1	1	0
غ	1	1	0.67	1	0.92	0.14
ف	1	0.67	1	1	0.92	0.14
ق	1	1	1	1	1	0
ك	1	1	1	1	1	0
ل	1	1	0.67	1	0.92	0.14

م	1	1	1	1	1	0
ن	1	1	1	1	1	0
ه	0.67	1	1	1	0.92	0.14
و	1	1	1	1	1	0
ي	1	1	1	1	1	0
Total average for all characters					0.97	0.09

Table- 3 and Figure- 5 show the recognition results in terms of the characters for all users that are tested with the handwritten characters for three times. The third column in these tables represents the number of characters that were correctly recognized. The fourth column displays the number of characters that were incorrectly recognized.

Table 3- Character Recognition Rate

User name	Number of test Characters	Correctly recognized	Incorrectly recognized	Recognition rate
User 1	84	83	1	98.8%
User 2	84	81	3	96.4%
User 3	84	79	5	94.1%
User4	84	83	1	98.8%



Figure 5- Character Recognition Rate (Correctly, Incorrectly)

5. Results and discussion

5.1 Evaluation of the Online-OCR Based On Proposed Feature Extraction Technique

The dataset was split into two groups; namely, training and testing datasets. This is to evaluate the proposed feature extraction technique for online OCR. In the experiment, 60% and 70% used as a benchmark for the training dataset. It shows, according to the experimental results, that the method proposed has higher accuracy rates, when compared with EDMS and Naeimizaghiani feature extraction methods. A decision tree of 60% training dataset is shown to have the best performance

when compared with EDMS of 68.5% and Naeimizaghiani method of 79.8% when the decision tree classifier training dataset has 62%. As shown in Table- 4 and Figure-6, the proposed feature extraction method achieved the highest performance with 61% training dataset and with the artificial neural network about 98.85%. Table- 4 also shows the classification results of EDMS, Naeimizaghiani et al. (2011) and the proposed method of the training dataset split from 60% to 70%.

Table 4- The classification results of Naeimizaghiani, EDMS and proposed method

	60%	61%	62 %	63 %	64 %	65%	66%	67 %	68 %	69 %	70%
decision tree/proposed	93.85	96.55	95.2	96.4	97.5	94.9	94.74	96	97.2	94.2	97
ANN/proposed	97.76	98.85	97.6	96.4	98.7	97.4	97.36	98	97.2	95.7	97
decision tree/EDMS	68.5	65.2	64.3	65.7	66.4	64.4	63.4	68	65.8	65.3	66.3
ANN/EDMS	59.9	62.7	59	65	61.3	64.6	61	61	63.1	68.6	63.8
Decision tree/Naeimizaghiani , Abdullah	79.3	79.4	79.8	79	77	76.76	77.5	79	76.9	77.2	76
ANN/Naeimizaghiani , Abdullah	71	72.4	72	67	70.2	73.6	72.82	71	72.3	69.7	70.37

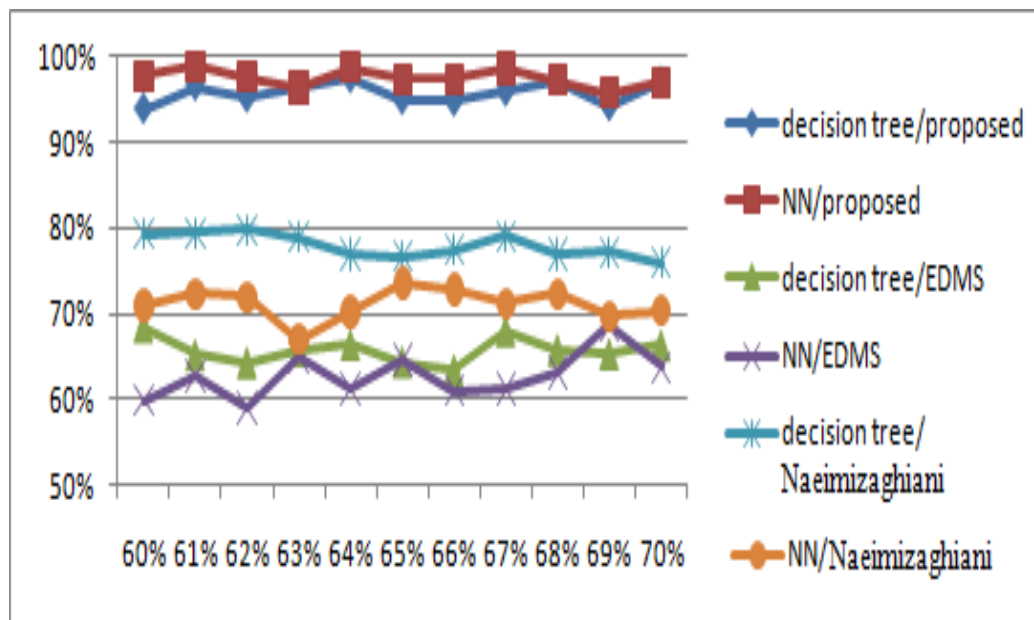


Figure 6- The classification results

The experiments on the 61% training dataset with decision tree are repeated five times in order to extract the performance of the previous experiments. This is also because the decision tree processing time is much faster than ANN, but the performance is almost similar. The average of the results is presented in Table- 5 and Figure- 7. The results show that the proposed method achieved the highest performance at 96.7%, whereas the EDMS method achieved 68.88%, and Naeimizaghiani method achieved 78.5 %.

Table 5-The five experiments and the average of the classification results

	EDMS	Naeimizaghiani	Proposed
Experiment #1	59.9	79.4	96.5
Experiment #2	74.7	77.2	96
Experiment #3	70.9	76	97
Experiment #4	73.4	81.6	97.1
Experiment #5	65.5	78.5	97.1
Average	68.88	78.5	96.7

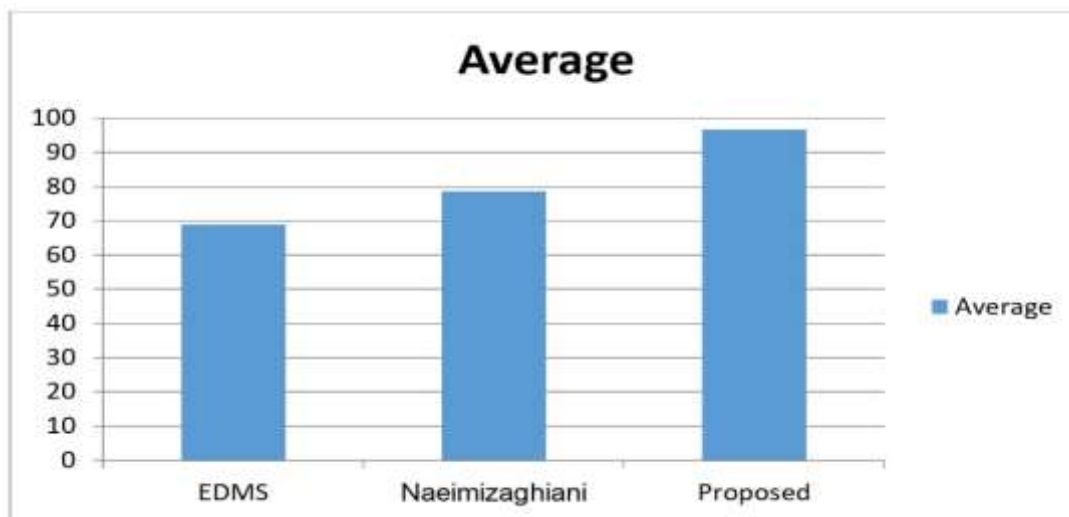


Figure 7-The average of the five experiments

The standard deviation value of 0.4, of the experiments conducted provided in Table- 6. The proposed method obtained a standard deviation lower than the EDMS method which was 6.0, and Naeimizaghiani method was 2.14. Therefore, the features of the proposed method are more effective than the other methods.

Table 6-The standard deviation for the five experiments' results

	EDMS	Naeimizaghiani	Proposed
Std. dev.	6.0	2.14	0.4

The detailed results for each class of the higher experimental results are shown in Table- 7 with a decision tree and a 61% training dataset. The proposed method has its highest accuracy with the (ا , ب , ت , ث , ج , ح , خ , ذ , ر , ز , س , ش , ص , ض , غ , ط , ظ , ق , ك , ل , م , ن , ه , و , ي) at a 100% accuracy rate. The lowest accuracy was obtained with the (د , ع , ف). For Naeimizaghiani, method shows that the highest accuracy was obtained with (ا , ب , ت , ث , ج , ح , خ , ذ , ر , ز , س , ش , ص , ض , غ , ط , ظ , ق , ك , ل , م , ن , ه , ي) at 91%. The lowest accuracy rate was obtained with the (د , ر , ز , س , ص , ض , ط , ف , ع , غ , ل , و) at 67% .

Table 7- The details results of the proposed method with decision tree classifier and 61% training dataset

Classifier technique	Decision tree		
	Proposed Method	EDMS Method	Naeimizaghiani Method
ا	1	1	1
ب	1	0.376	0.7
ت	1	0.818	1
ث	1	0.455	0.6
ج	1	0.333	0.727

ح	1	0.692	0.643
خ	1	0.875	0.929
د	0.714	0.667	0.778
ذ	1	0.778	0.846
ر	1	0.444	0.545
ز	1	0.556	0.778
س	1	0.563	0.643
ش	1	0.778	0.923
ص	1	0.636	0.7
ض	1	0.556	0.67
ع	0.933	0.308	0.667
غ	1	0.364	0.643
ط	1	0.7	0.778
ظ	1	0.7	0.8
ف	0.667	0.778	0.667
ق	1	0.889	1
ك	1	0.846	0.9
ل	1	0.833	0.57
م	1	0.667	0.875
ن	1	0.615	0.917
ه	1	0.833	0.917
و	1	0.5	0.1
ي	1	1	0.909
average	0.979	0.667	0.811

For EDMS, the highest accuracy was achieved with the (ا , ت , خ , ذ , ش , ف , ق , ك , ل , ي) at 86 % accuracy rate. The lowest accuracy was obtained with the (ب , ث , ج , ح , د , ر , ز , س , ص , ض , ع , غ ,) at 55%. Taking these results into account, the proposed method obtained an accuracy rate higher than that of EDMS and Naeimizaghiani for each class. Based on Table- 4 and Table-7 the proposed method produced the best accuracy rate compared with the EDMS method and Naeimizaghiani method using ANN and decision tree classifier it produced about 96.7% accuracy rate.

5.2 Evaluation of the Online-OCR Based on Proposed Rule-based Technique

In the experiments conducted, the evaluation of the proposed rule-based technique's performance was done through a comparison of the different classification methods used in the OCR. The artificial neural network (ANN) and decision tree classifiers are used in this study, though in combination, as against the single usage in related previous OCR studies. The best classification methods employed in handwriting recognition, from the literature review, are artificial neural network and decision tree. The decision tree and multilayer artificial neural network are applied to the earlier proposed feature extraction in finding the best performance. The dataset was split into training and testing datasets for the experiment. 60% is taken as the benchmark for the training dataset, while 40% is used for the testing dataset. The result of the experiment showed that ANN has the best performance at a recognition rate of 98.8%. The decision tree, on the other hand, was at 97.2%. As shown in Table- 8 and Figure- 8, the introduced rule-based method performs high performance by 97% recognition rates. According to the results, it is recorded that the proposed method achieved competitive performance with ANN and decision tree methods in every experiment.

Table 8- The classification results using the proposed feature extraction

Classifier technique	Decision tree	Artificial neural network	Proposed rule based
Proposed feature extraction	97.8	98.85	97

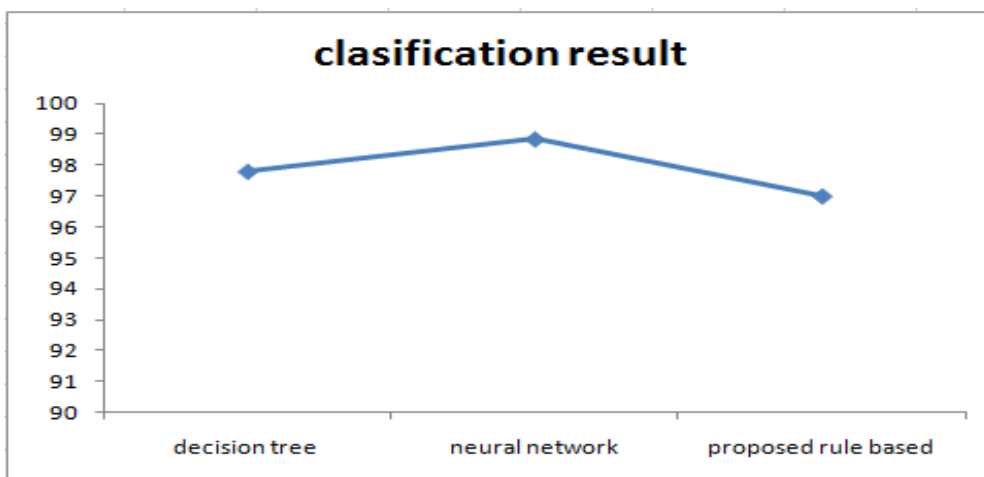


Figure 8- The classification results using the proposed feature extraction.

Table -9 shows the complete results for all character with the higher experimental results. For proposed rule based method, the highest accuracy was achieved with the (ا, ب, ت, ث, ج, ح, خ, د, ذ, ر, ز, س, ش, ص, ع, ق, ك, م, ن, و, ي, , (ط, ظ, غ, ل, ه, ض, ف). Decision tree method shows that the highest accuracy was obtained with (ا, ب, ت, ث, ج, ح, خ, ذ, ر, ز, س, ش, ص, ض, غ, ط, ظ, ق, ك, ل, م, ن, ه, و, ي), at 97.2%. The lowest accuracy rate was obtained with (د, ع, ف) at 2.7%. For ANN, the highest accuracy was achieved with the (ا, ب, ت, ث, ج, ح, خ, د, ر, ز, س, ش, ص, ض, ط, ظ, ع, غ, ف, ق, ك, ل, م, ن), which produced a 86% accuracy rate. The lowest accuracy was obtained with the (س, ذ) at 1.15%. Considering these results, the proposed method achieved a competitive accuracy rate with ANN and decision tree for each class.

Table 9-The details results of the proposed rule based method with decision tree and ANN classifiers.

Classifier technique	Proposed feature extraction		
	Decision tree	ANN	Proposed rule based
ا	1	1	1
ب	1	1	1
ت	1	1	1
ث	1	1	1
ج	1	1	1
ح	1	1	1
خ	1	1	1
د	0.714	1	1
ذ	1	0.887	1
ر	1	1	1
ز	1	1	1
س	1	0.79	1
ش	1	1	1
ص	1	1	1
ض	1	1	0.75
ع	0.933	1	1
ق	1	1	0.92
ك	1	1	0.84
ل	1	1	0.84
م	0.667	1	0.92
ن	1	1	1
ه	1	1	1
و	1	1	1
ي	1	1	0.92

م	1	1	1
ن	1	1	1
ه	1	1	0.92
و	1	1	1
ي	1	1	1
average	0.979	0.9885	0.97

Based on the experiment result, shown in Table- 9 above, the proposed method gives the competitive result and accuracy, compared with the ANN method and decision tree method after applied on the proposed feature extraction. It produced approximately 97 % accuracy rate.

6. Conclusion

In this study, a new online Arabic handwriting character recognition system is developed. The recognition system dealt with isolated Arabic character only. It uses projection profile (V, H) and Laplacian filter in pre-processing phase, geometrical and EDMS features in feature extraction phase, and rule based method in recognition phase. Based on the result obtained, the proposed methods are more efficient, when compared with related past studies. The objectives of this work which are (a) developing a new system for online Arabic character recognition, (b) proposing hybrid feature extraction technique for Arabic character recognition, and (c) proposing a rule based technique to recognize Arabic character for online Arabic handwriting character recognition system are all achieved. Finally, the implementation and evaluation of the proposed method was achieved through a comparison of the proposed feature extraction with EDMS and Naeimizaghiani feature extraction methods with ANN and decision tree classifiers.

References

1. Naeimizaghiani, M. **2011**. Character recognition based on global feature extraction, International Conference on Electrical Engineering and Informatics, Bandung, August 17-19, Bandung, Indonesia.
2. Théodorebluche, H., Ney, H. and Kermorvant C. **2015**. The LIMSI Handwriting Recognition System for the HTRtS 2014 Contest, International Conference of Document Analysis and Recognition (ICDAR), Nancy, France.
3. Cheung A., Bennamoun, M. and Bergmann, W. **2001**. An Arabic optical character recognition system using recognition-based segmentation. *Pattern recognition*, **34**(2): 215-233.
4. Zhang, H. **2011**. An Improved Scene Text Extraction Method Using Conditional Random Field and Optical Character Recognition, International Conference on Document Analysis and Recognition, September 18-21, Beijing, China .
5. Naz S. **2014**. The optical character recognition of Urdu-like cursive scripts. *Pattern Recognition*, **47**(3): 1229-1248.
6. Hamdani, M. **2009**. Combining Multiple HMMs Using On-line and Off-line Features for Off-line Arabic Handwriting Recognition, 10TH International Conference on Document Analysis and Recognition, July 26-29, Barcelona, Spain.
7. Farooq, F., Govindaraju, V. and Perrone, M. **2005**. Pre-processing methods for handwritten Arabic documents, Eighth International Conference on Document Analysis and Recognition, August 29 - September 1, Seoul , Korea
8. Amin A. and Singh, S. **1998**. Recognition of hand-printed Chinese characters using decision trees/machine learning C4. 5 system. *Pattern Analysis & Applications*, **1**(2): 130-141.
9. Ramdan, J., Omar, K., Faizul, M. and Mady, A. **2013**. Arabic Handwriting Data Base for Text Recognition, The 4th International Conference on Electrical Engineering and Informatics , Procedia Technology November ,Selangor, Malaysia.
10. Margner, V. and El-Abed, H. **2012**. *Guide to OCR for Arabic Scripts*, First Edition, Springer, London.
11. Mousa, M. and Sayed, M. **2013**. Arabic Character Segmentation Using Projection-Based Approach with Profile's Amplitude Filter, International Conference of Information and Communication Technology (ICoICT), March 20-23, Bandung, Indonesia.

12. Haboubi S. and Maddouri, S. **2011**. Identification of Arabic word from bilingual text using character features: Case of structural features, International Workshop on Frontiers in Arabic Handwriting Recognition, **3**(2): 50-54.
13. Omer, M. and Shi Long, M. **2010**. Online Arabic handwriting character recognition using matching algorithm in Computer and Automation Engineering (ICCAE), The 2nd International Conference on Computer and Automation Engineering , February 26-28, Singapore.
14. Ismail S. and Abdullah, S. **2012**. Online arabic handwritten character recognition based on a rule based approach, *Journal of Computer Science*, **8**(11): 1859-1868.
15. Bing, S., Xiaoqing, D. and Peng, L. **2013**. A Novel Baseline-independent Feature Set for Arabic Handwriting Recognition, 12th International Conference on Document Analysis and Recognition, August 25-28, Washington, DC.
16. Dehghan M., Faez, M. and Ahmed M. **2001**. Handwritten Farsi (Arabic) word recognition: a holistic approach using discrete HMM, *Pattern Recognition*, **34**(5): 1057-1065.
17. Lawgali, A. **2015**. A survey on arabic character recognition. International Journal of Signal Processing, *Image Processing and Pattern Recognition*, **8**(2): 401-426.
18. Aburas, A. and Gumah, M. **2008**. Arabic handwriting recognition: Challenges and solutions. in Information Technology, International Conference on Information Technology, August 26-19, Kuala Lumpur.
19. Ghaleb M. and George, L. **2013**. Numeral Handwritten Hindi/Arabic Numeric Recognition Method, *International Journal of Scientific & Engineering Research*, **4**(1): 229-518.
20. Eden M. **1962**. Handwriting and pattern recognition. Information Theory, IRE Transactions on information Theory, **8**(2): 160-166.
21. Hafemann, L.G., Sabourin, R. and Oliveira, L. **2017**. Offline Handwritten Signature Verification: Literature Review, Seventh International Conference on Image Processing Theory, Tools and Applications, November 28 - Decumber 1, Montreal, Canada.
22. Tappert C., Suen, C. and Wakahara, T. **1990**. The state of the art in online handwriting recognition. Pattern Analysis and Machine Intelligence, *IEEE Transactions on*, 1990. **12**(8): 787-808.
23. Parwej D. **2013**. The state of the art recognize in arabic script through combination of online and offline. *International Journal of Computer Science and Telecommunications*, **4**(3): 60-66.
24. Ahmed H. **2013**. Effective technique for the recognition of offline Arabic handwritten words using hidden Markov models, *International Journal on Document Analysis and Recognition (IJ DAR)*, **14**(9): 399-412.
25. AL-Shatnaw A., AL-Salaimh, S. and AL-Zawaideh, F. **2011**. Offline arabic text recognition an overview, *World Comput. Sci. Inform. Technol*, **1**(5): 184-192.
26. Lecolinet E. and Baret, O. **1994**. *Cursive words recognition: methods and strategies. Fundamentals in Handwriting Recognition*, Springer Berlin Heidelberg, **124**: 235-235.
27. Alnsour A. and Alzoubady, L. **2006**. Arabic handwritten characters recognized, neocognitron artificial neural network. University of Sharjah *Journal of Pure and Applied Sciences*, **3**(2): 22-27.
28. Khorsheed, M. **2003**. Recognising handwritten Arabic manuscripts using a single hidden Markov model, *Pattern Recognition Letters*, **24**(14): 2235-2242.
29. El-Sheikh T. and Guindi, R. **1988**. Computer recognition of Arabic cursive scripts. *Pattern recognition*, **21**(4): 293-302.
30. El-Sheikh T. and El-Taweel, S. **1990**. Real-time Arabic handwritten character recognition. *Pattern recognition*, **23**(12): 1323-1332.
31. Mezghani, N., Cheriet, M. and Mitiche, A. **2003**. Combination of pruned kohonen maps for on-line arabic characters recognition, the Seventh International Conference on Document Analysis and Recognition, August 6, Edinburgh, UK.
32. Amin A. **2000**. Recognition of printed Arabic text based on global features and decision tree learning techniques. *Pattern recognition*, **33**(8): 1309-1323.
33. Bataineh B. and Sheikh, A. **2013**. Arabic-Jawi Scripts Font Recognition Using First-Order Edge Direction Matrix, International Multi-Conference on Soft Computing Applications and Intelligent Systems, August 28-29, Shah Alam.