International Journal on Advanced Science Engineering Information Technology

Vol.7 (2017) No. 2
ISSN: 2088-5334

A Novel Method to Detect Segmentation Points of Arabic Words Using Peaks and Neural Network

Jabril Ramadan[#],Khairuldin Omar[#] Mohammad Faidzul[#]

[#]*School of Computer Science, Faculty of Information Science and Technology,Universiti Kebangsaan Malaysia, 43200 Bangi, Malaysia*
E-mail:Jabril@siswa.ukm.edu.my

***Abstract*— Many methods of segmentation using detection of segmentation points or where the location of segmentation points is expected before the segmentation process, the validity of segmentation points is verified by using ANNs. In this paper apply a novel method to detect correctly of location segmentation points by detect of peaks with neural networks for Arabic word. This method employs baseline and peaks identification; where using two steps to segmenting text. Where peaks identification function is applied which at the subword segment level to frame the minimum and maximum peaks, and baseline detection. Where these two steps have led to the best result through the model depends on minimum peaks attained by utilising a stroke operator with a view to extracting potential points of segmentation, and determining the baseline procedure was developed to approximate the parameters. Where this method has yielded highly accurate positive results for Arabic characters' segmentation with four kinds of handwritten datasets as AHDB, IFN-ENIT, AHDB-FTR and ACDAR. Earlier results showed that the use of EDMS to MLP_ANN gives better results than GLCM and MOMENT in different groups and gives results of EDMS features on MNN with an accuracy level of 95.09% classifier for IFN-ENIT set of data.**

***Keywords*— peaks; Arabic character; segmentation; baseline; neural network**

I. INTRODUCTION

According to [1], the strokes and curves of word characters can be identified by applying special attributes to detect the segmentation points. The validity of segmentation points is verified by using ANNs [2]. Handwritten Arabic texts can be segmented by applying a technique proposed by [3] based on ANNs. In this technique, topographic features are used to identify pre-segmentation points for every combined character blocks like holes and black pixel densities. The segmentation points are then verified by use of ANNs. Potential segmentation points are manually classified as valid or invalid, and each is fed with its characteristic features into the ANNs to be used for training [2], [4]. A method that uses Arabic handwriting recognition is developed, and it uses multi-agent technique approach to segment Arabic words [5]. This method relies on recognition for potential points segmentation validity verification. It is the best approach technique ever developed because it has resolved the inadequacies of the initial methods and achieved superior outcomes by preventing under segmentation.

The recognition method has superior performance capacity on its agents, and the artificial neural networks are correctly selected by applying the grouping rules that improve the detection of potential segmentation points [6]. It is vital to avoid over-segmentation mistakes while detecting segmentation points and characters by aiming at better results. Handwritten Arabic words can use the over segmentation algorithm whose main features are corner points, ends, and holes [3].

The segmentation points are then discriminated by being rejected or accepted by a neural network. According to [7] introduced a morphological method that applies rules analysis of words to detect segmentation points. The technique of word over segmentation by employing character's shape knowledge to avoid additional segmentation points was proposed by [8]. Contours and skeletons concept algorithm technique introduced by [9] can also be used for the detection of handwritten Arabic words segmentation points. [10] Presented a method that relies on primitive word segmentation then using a neural network to validate segmentation points guided by direction.

Baseline detection is important because most connection points linking the characters depend on the baseline to extract vital task of character segmentation points [11], [12]. In this task, the sub-word descanters have a beginning point under the baseline for them to be detected. The vertical projection is employed to detect potential segmentation points. The technique continues by applying the first estimated potential point is segmentation points (FSP) especially is the potential point is near the baseline [4]. The peaks detection points depend on many variables [13]. Therefore, peaks detection & baseline having neural network classifier is used to extract right segmentation points for sub-words and words. This happens by crossing the baseline with minimum peaks and evaluating for candidate segmentation points to avoid under and over segmentation.

Other previous works as feature-based Arabic Heuristic Segmenter (AHS) applied to determine segmentation points [10], Neuro-Heuristic Approach. [15], The strokes and curves – to identified by contours and skeletons [9], Morphological method applies rules analysis of words to detect segmentation points by [8], analyses contour for first three lower peaks of the distance map among intersection points and the chain code as the final segmentation points [14], Topographic features to identify pre-segmentation points as corner points, end points and holes [2], [3]. A multi-agent technique to segment Arabic words this method relies on recognition for potential points segmentation validity verification by [5]. Therefore, our word based on [9], [14] peaks detection & baseline having neural network classifier is used to extract right segmentation points for sub-words and words.

Arabic is a challenging task due to several reasons [16], [17]. In character segmentation, the main problem is over segmentation of the characters like Seen (س), Lam-Alif (لا), Sad (ص) and etc., segmented into more than one stage (parts) of the complete recognition system [4], [10]- [19]. The most optimal available methods applicable for over-segmentation is the one introduced by [18]. However, if the different shortcomings of the existing techniques are considered including thinning, then there is a need for improving the heuristic algorithm with a view to reducing 'bad' errors to increase the overall outcome of the desired accuracy. A more insightful method to identify the level of a component as an independent segment from a larger word character can be attained by use of linked component's height-to-width ratio. Among the existing techniques used for over-segmentation, that of [10], is clearly optimal; however, considering the various weaknesses of the available technique as Fig. 1.

II. Material and Method

The proposed method aims to solve the shortcomings of the previously outlined simple techniques. We thus developed a superior performance technique, which is largely able to mitigate all problems of over-segmentation and is specifically effective and applicable in the areas where the earlier techniques failed. Fig. 2 illustrated the steps of this method.

A. Peak Detection

Forms of argumentation: it takes a vector that contains the values to evaluate and the threshold value. The y- values are the ones to be evaluated at the top of the bar graph, and the threshold values are determined by the considered specific application. The peak detection function thus operates by checking the left and right values for each value considered. If the value considered surpasses the values to the right and left of the threshold value to the least, then the considered value qualifies as a peak. Consequently, the peak detection function should, therefore, return three number of maximum values by the produced graph provided the threshold value used is suitable. The function returns two numbers of vectors comprising all minimum and maximum points as shown in Fig. 2. If there is a failure in the peak detection function, or records zero then linked point is a single character, dots or ligature.

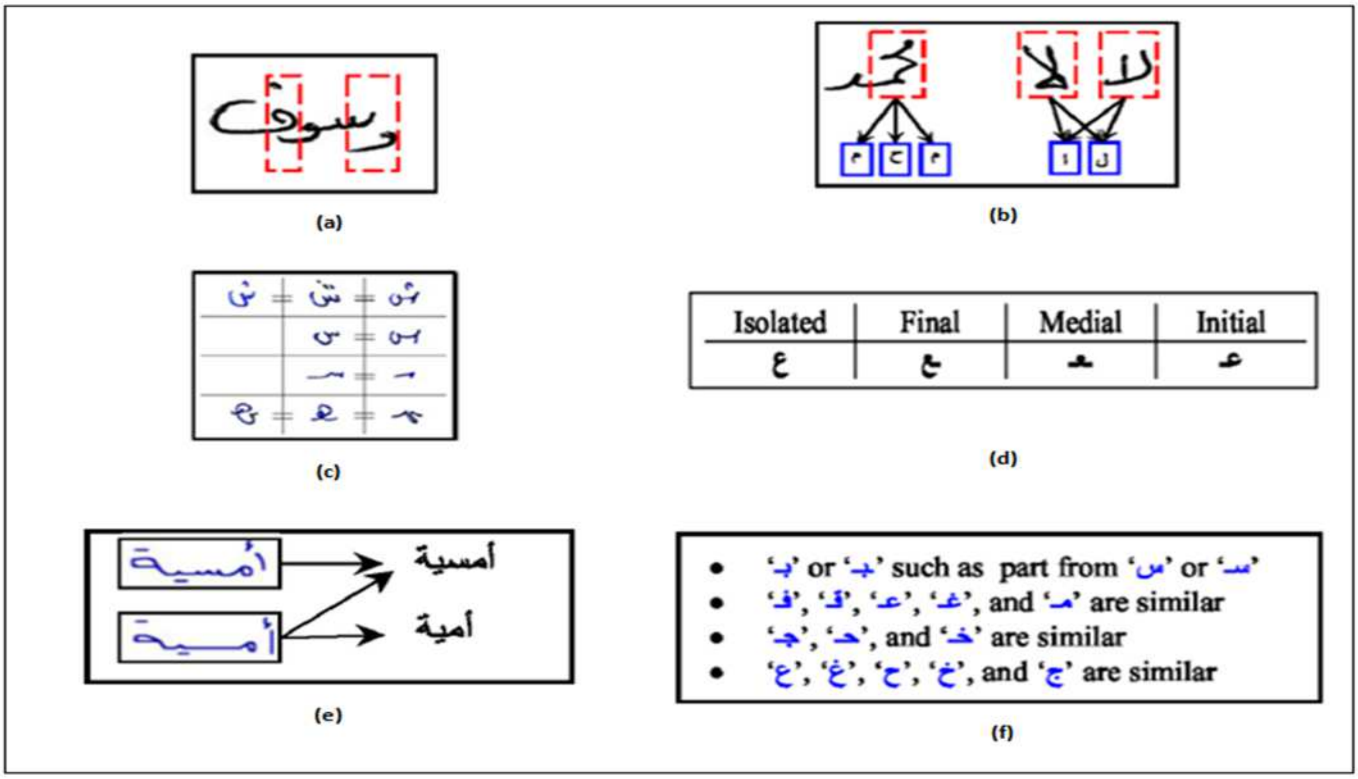
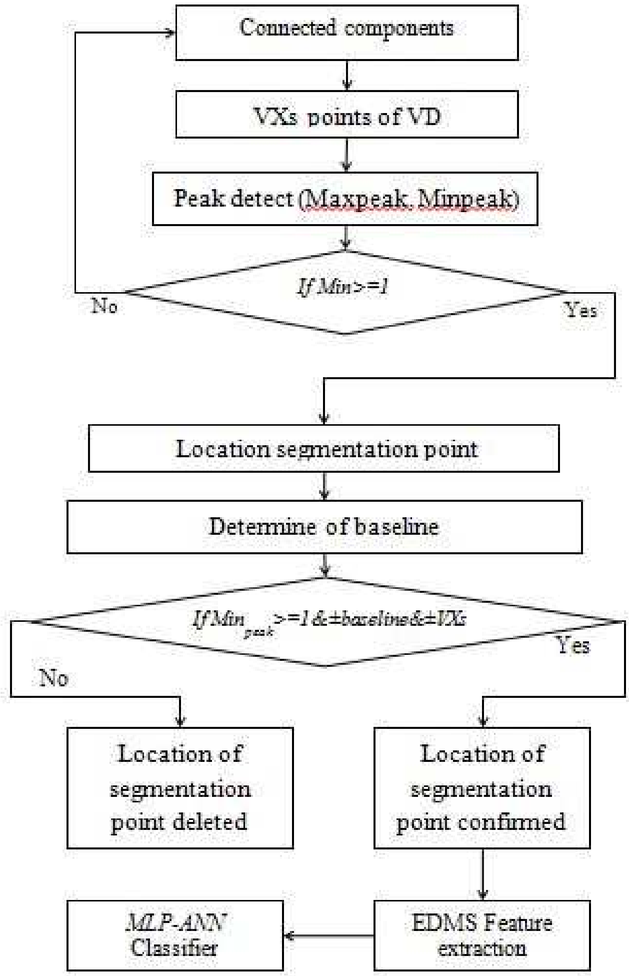

(a)
(b)
(c)
Isolated
Final
Medial
Initial
ع
ـع
ـعـ
عـ
(d)
أمسية
أمية
(e)
'ـب' or 'ـبـ' such as part from 'س' or 'سـ'
'ـفـ', 'ـقـ', 'ـعـ', 'ـغـ', and 'ـمـ' are similar
'ـجـ', 'ـحـ', and 'ـخـ' are similar
'ع', 'غ', 'ح', 'خ', and 'ج' are similar
(f)

Fig. 1 Illustrated problems of Arabic characters

B. *Evaluate the Location of Segmentation Points*

Sub-words and words minimum peaks segmentation points can be evaluated for Arabic characters by a Mathematical baseline to decide the location of the segmented character correctly, and this depends on two factors namely, namely maximum peaks that start from left to right and the baseline in the form of an equation as represented below:

```
If $Min_{peak1} > Max_{peak1}$ & $Min_{peak1} < Max_{peak2}$ then
    If $Min_{peak1 =}$ baseline or $Min_{peak1 =}$ baseline ±1 then
        $Min_{peak1 =}$ correct segmentation point
    End if
End if
```



Connected components
VXs points of VD
Peak detect (Maxpeak, Minpeak)
If Min>=1
No
Yes
Location segmentation point
Determine of baseline
If Min peak >=1&±baseline&±VXs
No
Yes
Location of segmentation point deleted
Location of segmentation point confirmed
MLP-ANN Classifier
EDMS Feature extraction

Fig. 2 The steps to determine location segmentation point

C. *EDMS Feature Extraction*

The EDMS features from global approach are applied in this study as these sub-processes utilize a statistical algorithm and identification for the features extraction.

This method applies a 3 by 3 Laplacian filter kernel matrix for extraction of edge images in which the resulting image is known as ledge (x, y). Every edge image pixel is linked to the 3 by 3 Kennel matrix and every of the 8 adjacent pixels' location a direction relationship with the middle pixel. To get important statistical features, which show data properties and distribution descriptions, 22 equations are used. These equations shoe the relationship between the text image front pixels by analysis of the occurrence values in EDM1 and EDM2 to understand the correlation, pixel regularity, edge direction, homogeneity, weights and edge regularity.

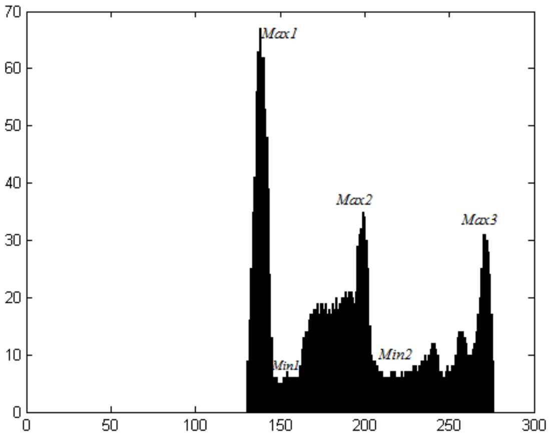


Fig. 3 Shows Maximum and Minimum peaks

D. *MLP-ANN Classifier*

This research illustrates that the MLP-ANN using BP (Backpropagation) architecture normally consists of an input layer and an output layer. The input layer comprises a total of twenty-two characteristic features that represent the nodes and are used to distinguish one of the eighteen types of fundamental shapes of the Arabic characters, which are also represented in the BPNN output layer. According to [23] all complicated relationships between the input and output layers are enclosed within the hidden layer. The input layer of BPNN comprises a total of eighteen nodes as known from the training experiments with the combination of 22 Attributes (features). These Attributes represent the Pixel_Regularity0, Pixel_Regularity45, Pixel_Regularity90, Pixel_Regularityn135, Correlation0, Correlation45, Correlation90, Correlation135, Homogeneity0, Homogeneity45, Homogeneity90, Homogeneity135, Weight, Edges' Direction, Edges_Regularit0, Edges_Regularit45, Edges_Regularit90, Edges_Regularit135, Edges_Regularit180, Edges_Regularit225, Edges_Regularit270, and Edges_Regularit315 angles.

Backpropagation Neural Network (MPLANN), which is one of the most superior and powerful classifiers is used in this work. Achieving best performance in validation set is by selecting the number of units that are hidden. The BPNN input layer has 22 features representing the nodes and is used for the identification of the 18 various Arabic characters basic shapes representing the nodes in the BPNN output layer. The hidden layer has 22 features 18 target classes as shown in Table 1.

TABLE I
THE 18 CATEGORIES OF ARABIC CHARACTERS SHAPE OUTPUT

Class. 1	Class.2	Class. 3	Class. 4	Class. 5	Class.6	Class. 7	Class. 8	Class. 9
ا	ب,بـ,ـبـ ـب	ح,حـ,ـحـ ـح	د,ـد	ر,ـر	س,سـ,ـسـ, ـس	ص,صـ, ـصـ,ـص	ط,طـ ـط	ع,عـ ـعـ,ـع
Class. 10	Class.l1	Class. 12	Class. 13	Class. 14	Class.I5	Class. 16	Class. I 7	Class.I8
ف,ق,ـق ـفـ,ـف,ـقـ,ـق,	ك,ـك كـ,ـكـ,	ل,لـ, ـلـ,ـل	م,ـم, مـ,ـمـ	ن,ـن, نـ,ـنـ	ه,ـه, هـ,ـهـ	و,ـو	ي,يـ, ـيـ,ـي	ء

E. *Network Training and Testing*

When different sets of data were analysed for performance, being split into testing and training by taking percentages ranging from 60% to 70%, the mean performance rate for all categories of characters was found to be 97.39%. The highest training performance set of data was 66% with a 97.77% accuracy level. The experiment was repeated five times, and the results were as shown in Fig. 4 Below.

Four model sets of hand written data were used for evaluating the proposed algorithm & segmentation algorithm, as well as the AHDB (Arabic handwritten database) consists of 3045 words [7], and the ACDAR database; this database was designed for a majority of segmentation experiments [8], and AHDB/FTR Arabic handwriting database comprises 497 images of the names of Libyan towns for text recognition [20], and IFN/ENIT-database [22].

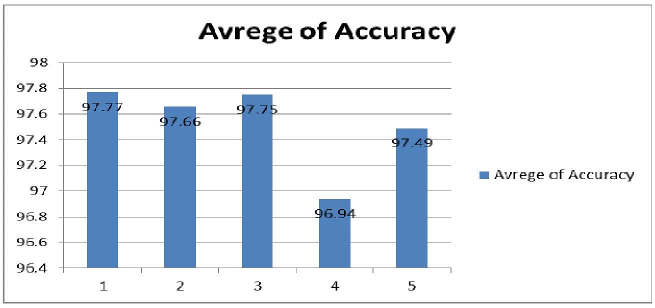


Fig. 4 The training results of five times of BPNN classifier with 60-70%

III. Results and Discussion

In this study, experiments were conducted, described the images and the results of segmentation. This experiment was conducted on selected images that presented various types of segmentation challenges. Were the problems such as disparity in text size, significant non-uniform illumination, low image quality, thin pen stroke lines and low contrast between the text and background were incorporated into the experiment. Four samples were collected from several sources for this experiment. Fig. 5(a), (b), (c) and (d) show various sizes and fonts of the handwritten text.

Table 2 illustrates the peaks with thresholding value used one image word Shaal "شعال" as Fig. 4 displays; it contains two main connected component without dots; to see how it impacts the threshold value in a number of maximum and minimum peaks.

The Table 2 shown the threshold values impacts on peak detection; as shown above is the best threshold values, it's between 8 to 18 for best results of peaks detection as maximum and minimum peaks for connected component no two, because it contains more than one character. For connected component no one the best result between 2 to 8 value of threshold because it contains one character. To refer to the Table 2, and Figure 3 to confirm those segmentation points are its correct location of the segment the characters or not before segmentation process.

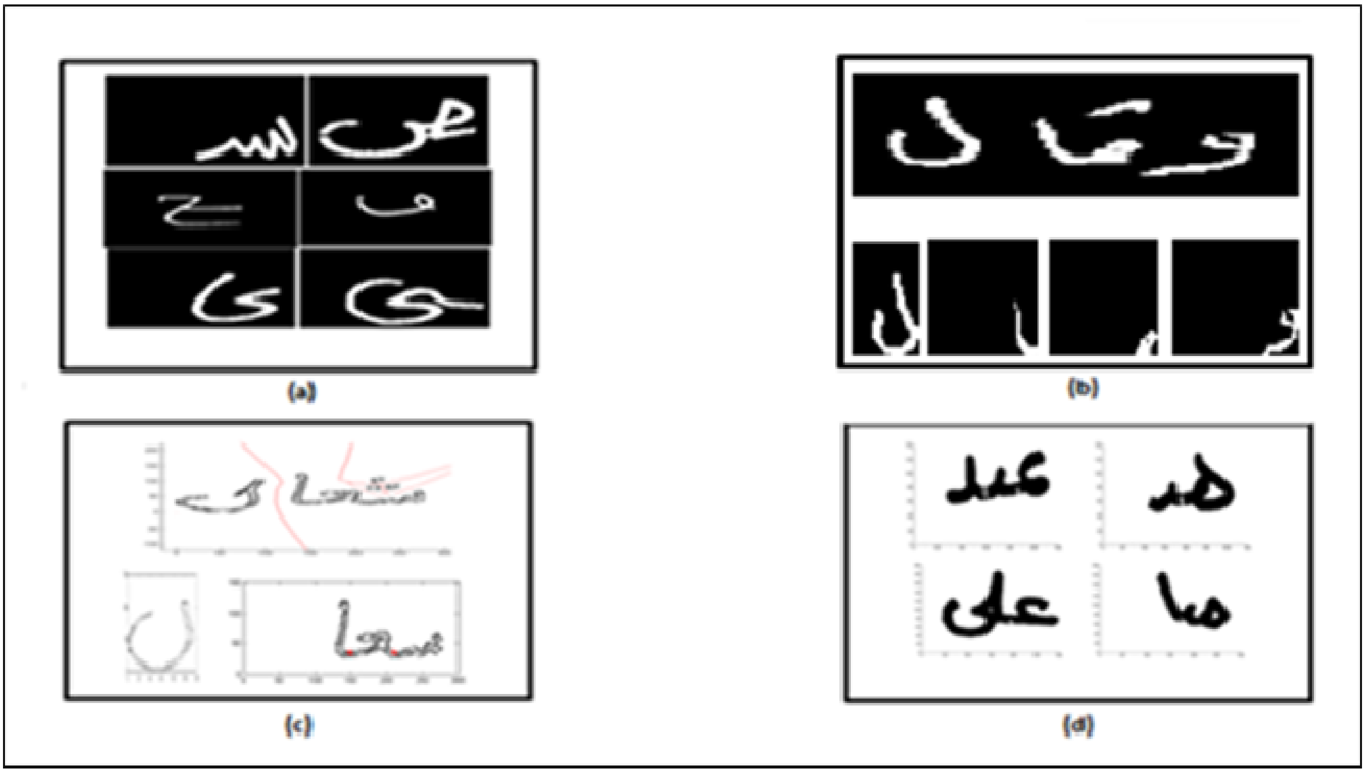


Fig. 5 Sample results for (a) from the AHDB-FTR dataset, (b) a word from the ACDAR dataset, (c) a word from the IFNENIT, and (d) words from the AHDB dataset

TABLE II
ILLUSTRATE THE PEAKS WITH THRESHOLDING VALUE FOR IMAGE "شعال"

Threshold value	Maximum peaks values (*x*,*y*)		*Minimum peaks values (x,y)*	
	CC No1	CC No2	CC No1	CC No2
2	30 66 97 112	138 126 199 82 240 79 256 68 270 81	52 33	149 55 211 56 246 58 262 59
8	30 66 97 112	138 126 199 82 270 79	52 33	149 55 211 56
18	97 112	138 126 199 82 270 81	0	149 55 211 56
28	97 112	138 126 199 82	0	149 55
38	97 112	138 126	0	0

According to [20] Table 3 illustrates the minimum peaks as the location of segmentation points of connected components after confirmed by his algorithm. As illustrated in Table 3 in minimum peaks of connected component No one the algorithm is not confirmed all points; because the connected component is one character.

TABLE III
ILLUSTRATE THE MINIMUM PEAKS AS LOCATION OF SEGMENTATION POINTS

Threshold value	Base-line value	Minimum peaks CC No1		Minimum peaks CC No2	
		CSP	Notes	CSP	Notes
2	56	52 33	Not confirmed	149 55 211 56 246 58 262 59	Not confirmed
8	56	52 33	Not confirmed	149 55 211 56	confirmed
18	56	0	Not confirmed	149 55 211 56	confirmed
28	56	0	Not confirmed	149 55	Not confirmed
38	56	0	Not confirmed	0	Not confirmed

Where the algorithm confirmed for two times in connected component no two; where the threshold value it is between 8 and 18, and not confirmed by others; which are lead to over-segmentation or under segmentation.

The experiment was carried out on various images, and there were noted disparity issues in the text size, low quality, Non-uniform illumination, low contrast and thin stroke lines between texts and background as shown in Table 4 below. Table 5 will show the comparison between MLP-ANN, and two others classifiers as Random forest (RF), RIDOR Rule (RDR) classifiers using same features (EDMS).

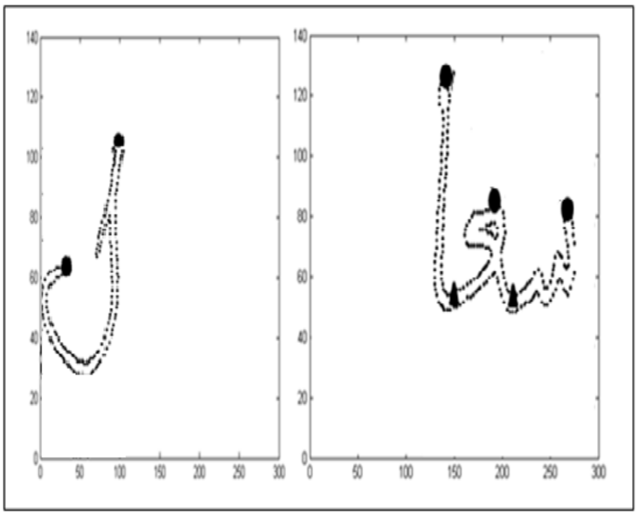


Fig. 6 Shown two main connected components without dots "شعال"

TABLE IV
EXPERIMENTAL RESULTS FROM THE PROPOSED METHOD

Classifiers	Features	DATA SETS			
		IFNENIT	ACDAR	AHDB	AHDB_FTR
MLP-ANN	EDMS	97.77%	91.10%	92.62%	90.69%

TABLE V
COMPARISON RESULTS WITH MLP-ANN ON EDMS FEATURS

Classifiers	Features	DATA SETS			
		IFNENIT	ACDAR	AHDB	AHDB_FTR
RF	EDMS	93.96%	85.64%	90.91%	89.76%
RDR	EDMS	90.17%	77.72%	89.09%	85.11%

A. Comparing Results with Others

This method should be compared with previous methods using other techniques to know the difference between them and contrast recognition results. Therefore, a comparison of this method is drawn with Al Hamad [10] method, and Abu Ain [24] methodology while utilizing similar datasets (ACDAR) in their experimentations. The scores of the classification accuracies and the average precise segmentation correspondingly are 97.77% and 89.87%.

Vertical projection histogram that utilized the text skeleton rather than using the original text image is found considerable for the identification of segmentation points in the methodology of Al-Hamad. This methodology reports the rate of average precise segmentation as 82.98%. Another methodology of the Advance Strokes Labeling Based on Direction Feature (ASLDF) is given for the segmentation of

Abu-Ain. The modified vertical projection histogram and the ASLDF are utilized for the interpretation of segmentation points as given in this methodology. With this view, the average precise segmentation accuracy rate is given as 92.45% with the validation of the recommended set of Arabic language comprising of structural-rules that further endows with segmentation points of lesser respondents.

The segmentation points are identified with the transformed vertical projection histogram and the text skeleton that are contrary to the Al-Hamad and Abu-Ain methodologies. With this view, segmentation points of a larger number of candidates are endowed as given in the neural networks and direction features through validations. The two prior methodologies are disparate whereas this methodology is utilized to detect the connected components as there is one character or more than one to reduce the time for segmentation and avoid over-segmentation. If the connected component is more than one character after determining the maximum and minimum peaks for detection, the connected components can also be used with the minimum peak for detecting the segmentation point truth from the skeleton point on the baseline for connected components.

The corroboration of segmentation points comprising of the 95.60% and 82.26% as average categorization accuracy rate is done through a neural-based categorization procedure implemented on the segmented characters of Abu-Ain and Al-Hamad correspondingly. This is done by putting every segmented character into a predefined set that already comprised of Arabic character shapes.

Table 6 indicates the comparison of the proposed characters' segmentation method results with Abu Ain and Al Hamad method for each writer with Correct, under, and over segmentation measurement criteria to obtain averages for all writers.

TABLE VI
COMPARISON OF THE PROPOSED CHARACTERS SEGMENTATION METHOD RESULTS WITH ABU AIN AND AL-HAMAD METHOD

	Correct Segmentation			Under-segmentation			Over-segmentation		
Writer	Proposed	Abu Ain	Al Hamad	Proposed	Abu Ain	Al Hamad	Proposed	Abu Ain	Al Hamad
1	95.36%	94.98%	85.34%	4.63%	3.85%	4.89%	0.2%	1.43%	9.77%
2	91.48%	90.40%	81.02%	8.51%	4.80%	4.22%	0.4%	4.80%	14.7%
3	94.36%	87.30%	82.17%	5.63%	7.14%	2.33%	0.4%	5.56%	15.5%
4	93.1%	87.64%	87.38%	6.89%	10.8%	3.24%	0.5%	1.54%	9.39%
5	89.39%	90.98%	81.79%	10.6%	6.77%	7.28%	0.3%	2.26%	10.4%
6	90.01%	93.10%	80.45%	10%	4.60%	5.67%	0.6%	2.30%	13.8%
7	88.89%	96.98%	84.29%	11.11%	2.64%	2.88%	0.2%	0.38%	12.8%
8	89.61%	91.85%	81.14%	10.38%	3.70%	8%	0.4%	4.44%	10.8%
9	90.78%	96.89%	80%	9.21%	1.56%	3.38%	0.4%	1.56%	16.6%
10	91.22%	94.35%	87.38%	8.77%	5.65%	4.32%	0.4%	0.01%	8.31%
AVG	**91.42%**	**92.45%**	**82.98%**	**8.57%**	**5.15%**	**4.6%**	**0.38%**	**2.43%**	**12.4%**

In the Table 6 comparison of the results noted the proposed accuracy i.e. less than the Abu Ain method for correct segmentation because the proposed method depends on the connected component and the ACDAR dataset have some problems for some characters like ط, ک; these are two characters whereas its more characters to the problem of suffering adjudicated usually consists of two parts in the word, although in one character. This affects the process of taking the features and recognize also. Fig. 7. shows the problems of both characters where the proposed method depends on the connected components then it will deal with these as these are the two connected components. This case led to the obstacle for features' extractions and classifiers; this affected on accuracy rate in correct segmentation and under segmentation.

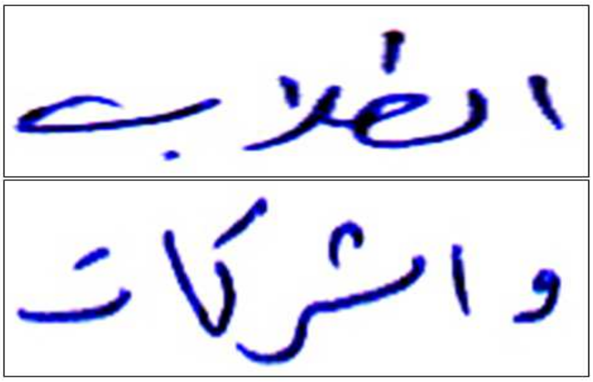


Fig. 7 The problems of suffering adjudicated like "ط", and "ک"

IV. CONCLUSION

For the better result to be realized in character recognition, there should be a perfect technique in the segmentation point detection. There should be new techniques to detect potential segmentation points by use of peaks detection, neural network and baseline to segment every word and sub-words into simpler characters. The EDMS features and MLP-ANN techniques have experimentally shown that its accuracy validity is 97.77% in regard to performance.

ACKNOWLEDGEMENTS

The authors would like to thank Prof Dr. Khairuddin Omar of the National University of Malaysia for his assistance and cooperation. This study was funded by the FRGS/1/2014/ICT07/UKM/01/1 grant entitled “Improving Segmentation of Arabic Handwriting by Determination of Neighborhood Using Voronoi Diagrams.”

REFERENCES

[1] Abdullah, Shubair A. "Off-Line Handwritten Arabic Characters Segmentation Using Slant-Tolerant Segment Features (Stsf)." Diss. USM, 2007.

[2] A. Lawgali, "A survey on arabic character recognition," International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 8, no. 2, pp. 401–426, Feb. 2015.

[3] A. Hamid and R. Haraty, "A neuro-heuristic approach for segmenting handwritten arabic text," Proceedings ACS/IEEE International Conference on Computer Systems and Applications, 2001.

[4] Zeki, A. M. (2005). The segmentation problem in arabic character recognition the state of the art. In Information and Communication Technologies, 2005. ICICT 2005. First International Conference on (pp. 11–26). IEEE.

[5] A. Elnagar and R. Bentrcia, "A multi-agent approach to arabic handwritten text segmentation," *Journal of Intelligent Learning Systems and Applications*, vol. 04, no. 03, pp. 207–215, 2012

[6] A. Elnagar and R. Bentrcia, "A recognition-based approach to segmenting arabic handwritten text," *Journal of Intelligent Learning Systems and Applications*, vol. 07, no. 04, pp. 93–103, 2015.

[7] S. Alma’adeed, C. Higgins, and D. Elliman, "Off-line recognition of handwritten arabic words using multiple hidden Markov models," Knowledge-Based Systems, vol. 17, no. 2-4, pp. 75–79, May 2004.

[8] L. Lorigo and V. Govindaraju, "Segmentation and pre-recognition of arabic handwriting," Eighth International Conference on Document Analysis and Recognition (ICDAR’05), 2005.

[9] S. Wshah, Z. Shi, and V. Govindaraju, "Segmentation of arabic handwriting based on both contour and skeleton segmentation," 10th International Conference on Document Analysis and Recognition, 2009. pp. 793-797. 26-29 July.

[10] Al Hamad and R. Abu Zitar, "Development of an efficient neural-based segmentation technique for arabic handwriting recognition," Pattern Recognition, vol. 43, no. 8, pp. 2773–2798, Aug. 2010.

[11] L. Lorigo and V. Govindaraju, "Offline arabic handwriting recognition: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 712–724, May 2006.

[12] A. Lawgali, M. Angelova, and A. Bouridane, "A framework for arabic handwritten recognition based on segmentation," International Journal of Hybrid Information Technology, vol. 7, no. 5, pp. 413–428, Sep. 2014.

[13] M.I Ghazali,Z. Harun,W.A Wan Ghopa and A. A Abbas,"Computational Fluid Dynamic Simulation on NACA 0026 Airfoil with V-Groove Riblets," International Journal on Advanced Science, Engineering and Information Technology, vol. 6, no. 4, pp. 529-533, 2016. [Online]. Available: http://dx.doi.org/10.18517/ijaseit.6.4.901.

[14] Y., Osman, "Segmentation algorithm for Arabic handwritten text based on contour analysis. Computing, Electrical and Electronics Engineering (ICCEEE)," International Conference on. 2013 pp. 447–452. 26-28 Aug. IEEE.

[15] Sari, T., Souici, L., & Sellami, M. (2002). Off-line handwritten arabic character segmentation algorithm: Acsa. In Frontiers in Handwriting Recognition, 2002. Proceedings. Eighth International Workshop on (pp. 452–457). IEEE.

[16] Merfat.M. Altawaier and Sabrina Tiun,"Comparison of Machine Learning Approaches on Arabic Twitter Sentiment Analysis," International Journal on Advanced Science, Engineering and Information Technology, vol. 6, no. 6, pp. 1067-1073, 2016. [Online]. Available: http://dx.doi.org/10.18517/ijaseit.6.6.1456.

[17] Soliman, T. H., Elmasry, M., Hedar, A. and Doss, M, “Sentiment Analysis of Arabic Slang Comments on Facebook”. InternationalJournal Of Computers & Technology, 12(5). 3470-3478, 2014.

[18] J. Ramdan, K. Omar, M. Faidzul, "A New Rule to Reconfirm Potential Segmentation Points with Vertexes Points of VDS," Middle-East Journal of Scientific Research, vol. 24 no 3, pp. 657-662, 2016.

[19] K., Mohammad, M., Ayyesh, A., Qaroush, and I., Tumar, "Printed Arabic optical character segmentation." *In SPIE/IS&T Electronic Imaging*. vol. 9399, pp. 939911-939911.2015.

[20] J., Ramdan, K. Omar, M. Faidzul, and A. Mady, "Arabic handwriting data base for text recognition," Procedia Technology, vol. 11, pp. 580–584, 2013.

[21] "Eli Billauer’s home page,". [Online]. Available: http://www.billauer.co.il. Accessed: 2013.

[22] M., Pechwitz, S., Maddouri, V., Märgner, N., Ellouze, & H. Amiri, "IFN/ENIT-database of handwritten Arabic words." In Proc. of CIFED. Citeseer, vol. 2, pp. 127–136. 2002.

[23] D., Jiménez, A., Pérez-Uribe, H., Satizábal, M., Barreto, P.Van Damme, & M., Tomassini, “A survey of artificial neural network-based modeling in agroecology” Studies in Fuzziness and Soft Computing, vol. 226, pp. 247–269, 2008.

[24] T. Abu-Ain, “Joint-landmarks baseline and advanced direction features for Arabic character segmentation,” Ph.D. dissertation, Dept. Com. Sci. IT., UKM Univ., Bangi, 2015.