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).

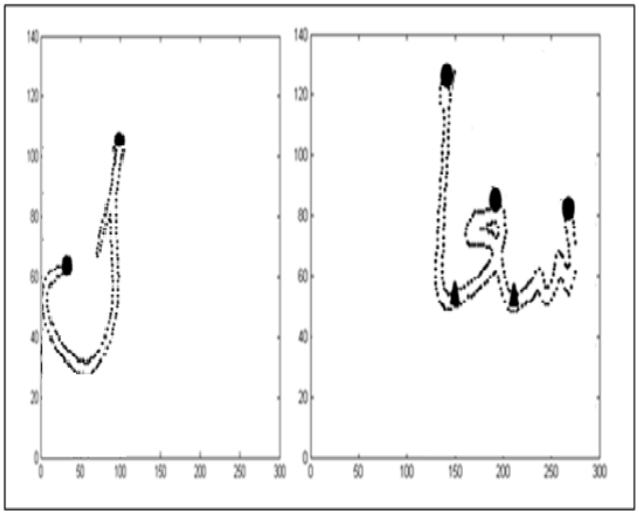

Fig. 6 Shown two main connected components without dots "شعال"

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

Writer	Correct Segmentation			Under-segmentation			Over-segmentation		
	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, ك; 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.

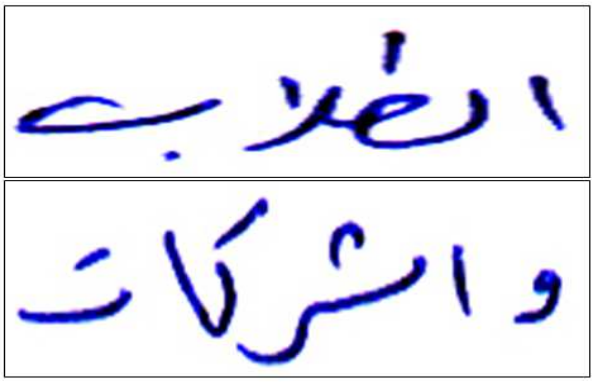

Fig. 7 The problems of suffering adjudicated like "ط", and "ك"

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.