ICDAR2017 Competition on Multi-font and Multi-size Digitally Represented Arabic Text

Fouad Slimane[1] - Rolf Ingold[1] - Jean Hennebert[1,2]
[1]DIVA: Document, Image and Voice Analysis research group , Department of Informatics
University of Fribourg (unifr), Bd de Pérolles 90, CH-1700 Fribourg, Switzerland
[2]University of Applied Science of Western Switzerland
HES-SO//FR, Bd. de Pérolles 80, 1705 Fribourg, Switzerland
{Fouad.Slimane, Rolf.Ingold}@unifr.ch, Jean.Hennebert@hefr.ch

***Abstract*—This paper describes the organisation and results of the Arabic Recognition Competition: Multi-font Multi-size Digitally Represented Text held in the context of the 14^{th} International Conference on Document Analysis and Recognition (ICDAR'2017), during November 10-15, 2017, Kyoto, Japan. This competition has used the freely available Arabic Printed Text Image (APTI) database. A first and second editions took place respectively in ICDAR'2011 and ICDAR'2013. In this edition, we propose four challenges. Six research groups are participating in the competition with thirteen systems. These systems are compared using the font, font-size, font and font-size, and character and word recognition rates. The systems were tested in a blind manner using the first 5000 images of APTI database set 6. A short description of the participating groups, their systems, the experimental setup, and the observed results are presented.**

***Keywords*-APTI Database; Arabic Text; Ultra-Low Resolution; OCR System; Competition;**

I. Introduction

The Arabic language is one of the most popular languages in the world, used natively by hundreds of millions of people in many countries. However, the printed form of the Arabic language shows complexities that are not present in latin-based or ideogram-based languages. The most prominent specificity is in its cursive nature in printed and handwritten documents. Another one is in the large variability of character representations according to the different fonts. Arabic OCR systems were therefore less studied in comparison to other languages, even though we have seen an increasing number of research works in the last decade.

Many competitions are traditionally organized at the International Conference on Document Analysis and Recognition (ICDAR)[1] [2] [3] [4] and the International Conference on Frontiers in Handwriting Recognition (ICFHR)[5] [6] [7] [8]. Their objectives are to improve the classification techniques and to compare and evaluate different techniques and systems [1], [2], [3], [4]. We successfully organised two previous competitions at ICDAR 2011 [4] and ICDAR 2013 [5] exploring the impacts of multi-fonts and multi-size digitally represented Arabic text. We have proposed here the third edition with more challenging protocols. The scientific objectives of this edition are to measure the capacity of recognition systems to identify the font and the font-size using one Arabic word images at ultra-low resolution, and the impact of font and font-size on the text recognition performances. This will be evaluated in multi-font and multi-font contexts. To our knowledge, no competition was organized before for font and font-size identification. The protocols will be defined to evaluate the capacity of recognition systems to handle different sizes and fonts using low resolution images in the aim to look for a robust approach to screen based OCR. The main difficulty is probably in the multi-font and multi-size context as differences between fonts are rather important for Arabic text.

The Arabic Printed Text Image database (APTI) is one of the most used databases by the researcher working in the field of Arabic text recognition [6]. It is used by more than hundred research groups world wide. The most interesting characteristics of APTI are:

- Very large set of images for significant benchmarking (more than 45 millions images);
- Large lexicon;
- Multi-font;
- Multi-size;
- Single word images.

Potentially less difficult than handwritten Arabic text recognition, APTI remains challenging due to the variabilities induced by the different fonts and sizes that, in some cases, change drastically the distributions of observed features. APTI is typically related to OCR and "screen-based" OCR inputs where the user grab and crop a part of the computer screen.

By proposing APTI and organizing competitions, we

[1]http://www.cvc.uab.es/icdar2009/competitions.html
[2]http://www.icdar2011.org/EN/column/column26.shtml
[3]http://www.icdar2013.org/program/competitions
[4]http://www.iapr.org/archives/icdar2015/index.html%3Fp=254.html
[5]http://www.isical.ac.in/˜icfhr2010/index.html
[6]http://www.icfhr2012.uniba.it/index.php
[7]http://www.icfhr2014.org/competitions-2/
[8]http://www.nlpr.ia.ac.cn/icfhr2016/competitions.htm

2379-2140/17 $31.00 © 2017 IEEE
DOI 10.1109/ICDAR.2017.239

Figure 1. Fonts used to generate the APTI database: (A) Andalus, (B) Arabic Transparent, (C) AdvertisingBold, (D) Diwani Letter, (E) DecoType Thuluth, (F) Simplified Arabic, (G) Tahoma, (H) Traditional Arabic, (I) DecoType Naskh, (J) M Unicode Sara

hope that we help to improve the situation on Arabic text recognition. Through this data, we give the possibility to compare different systems and algorithms.

The evaluation has been organized using a blind procedure. The testing data of the evaluation is composed by 5000 images of the unpublished set (so called *set 6* of APTI) which is kept secret for evaluation purposes. The participants were able to train and tune their systems using the public parts of APTI (Set 1, 2, 3, 4 and 5). Then, they were asked to send an executable version of their recognizer to the organizers who, in turn, arranged to run the systems against unseen images from set 6 of APTI.

The paper is organized as follows. Section 2 summarizes the main characteristics of APTI database. Section 3 is dedicated to the competition protocols. In section 4, we present the participating systems. Results are discussed in Section 5 and are followed by conclusions.

II. The APTI Database

The APTI database was developed to promote the research and development of Arabic printed word recognition systems. Available from July 2009, APTI is freely distributed to the scientific community for benchmarking purposes [9]. At the time of writing this paper, more than hundred research groups have started using it.

The APTI database was created in ultra-low resolution "72 dot/inch" with a lexicon of 113,284 different Arabic words and 10 fonts presented in Figure 1. These fonts have been selected to cover different complexity of shapes of Arabic printed characters. Different font sizes are also used in APTI: 6, 7, 8, 9, 10, 12, 14, 16, 18 and 24 points. We also used 4 different styles namely plain, *italic*, **bold** and combination of ***italic and bold***. The combination of fonts, styles and sizes guaranties a wide variability of images in the database.

[9]http://diuf.unifr.ch/diva/APTI/

Table I
Quantity of words, PAWs and characters in APTI

	Nber of Words	Nber of PAWs	Nber of Characters
	113,284	274,833	648,280
	*10 Fonts * 10 Font Sizes * 4 Font Styles		
Total	45,313,600	109,933,200	259,312,000

The total number of word images is above 45 million. Each word image in the APTI database is in grey level and fully described using an XML file containing ground truth information about the sequence of characters as well as information about its generation. All Arabic letters have a balanced distribution throughout the sets composing the database. Table I shows the total quantity of word images, Piece of Arabic Words (PAWs) and characters in APTI.

The database is divided into six comparable sets to allow for flexibility in the composition of development and evaluation partitions. For more details about APTI, we refer to [6] and [7].

III. The competition

We invited groups participating to this context to adapt their system to the APTI database and to send us executable programs of their systems.

The evaluation is reported as font, font-size, font and font-size identification rates respectively for the first three protocols and as word and character recognition rates for the fourth protocol. In this edition, we use the writing style: **Plain** and the font-sizes (**6**, **8**, **10**, **12**, **18** and **24**) used in the first and second editions and we proposed 4 protocols:

1) **First APTI Protocol for Competition:** 1^{st} **APTIPC**: This protocol aims to identify font-size based on Arabic words independently to the font. Participants in this protocol should submit one font-size recognition system for all fonts.

2) **Second APTI Protocol for Competition:** 2^{nd} **APTIPC**: This protocol aims to identify font based on Arabic words independently to the font-size. Participants in this protocol should submit one font recognition system for all font-sizes.

3) **Third APTI Protocol for Competition:** 3^{rd} **APTIPC**: This protocol aims to identify in the same time font and font-size based on Arabic words. Participants in this protocol should submit one font and font-size recognition system.

4) **Fourth APTI Protocol for Competition:** 4^{rd} **APTIPC**: This protocol uses All APTI fonts independently to the size. Participants in this protocol should submit one multi-font and multi-size text recognition system.
Fonts: **Andalus**, **Arabic Transparent**, **Advertising-Bold**, **Diwani Letter**, **DecoType Thuluth**, **Simplified Arabic**, **Tahoma**, **Traditional Aatbic**, **DecoType Naskh** and **M Unicode Sara**

IV. Participating systems

This section gives a short description of the submitted systems to the competition.

A. RDI-CU System

The RDI-CU system was submitted by Hany Ahmed, Salah Ashraf, Mahmoud Aboelazm members in the OCR researcher team at the RDI research group and Prof. Mohsen Rashwan the CEO of the RDI and professor at the Cairo University, department of Electronics and Electrical Communication Engineering.

The submitted system uses the basic Hidden Markov Models (HMM) to recognize font type, font-size and both of font and size. One system has been submitted for the first three protocols. The proposed system first locally binarizes the input images using adaptive thresholding [8] without any normalization then extracts a set of robust features which represents the distribution of foreground pixels by using an overlapping sliding window. The window size is 11 pixels in width and 10 overlapping pixels. The extracted features are then passed to the decoder. The authors used the first four sets for training; 150K images have been randomly selected from each set. In theirs experiments, 100 different classes (10 different fonts X 10 different sizes) have been considered with one state for each class. The experiments proved that choosing suitable features and suitable parameters of HMM leads to high recognition rate.

B. REGIM-LITIS Systems

The REGIM-LITIS systems are submitted by Abdelkarim Elbaati, Houcine Boubaker, Aymen Chaabouni and Adel M. Alimi, from the Research Group on the Intelligent Machines (REGIM) at Ecole Nationale d'Ingenieurs de Sfax (ENIS), University of Sfax, Tunisia and Abdellatif Ennaji from LITIS, University of Rouen, France.

The features used for the font recognition are Histograms of Oriented Gradients (HOG) calculated on the resized image of the word [9].

The font-size recognition is based on the calculation of the minimum, maximum and mean values of the different segments of the skeleton [10]. The dimensions are normalized by dividing them to the maximum value of the database.

In each system, a feed forward neural network is used as a classifier. For the first system, the neural network has 144 inputs (HOG features) and 10 outputs (font classes) and 10 hidden layers. For the second system the neural network has 9 inputs (the dimensions of the segments of the skeleton), 6 outputs (font-size classes) and 10 hidden layers. The first 1000 samples of each set are used in the learning step.

The third system is an ensemble that combines the results of the first two systems

C. SP-Curvelet-FR system

SP-Curvelet-FR system is submitted by Faten Kallel and Monji Kherallah from the university of Sfax, Tunisia.

This method consist of using the curvelet transform in different resolution levels generated by steerable pyramids for feature extraction phase and SVM for font classification.

Firstly, the authors started by spacing normalization. For each line in the text image, they deleted the word space. Then, for each incomplete text line, they extracted a line block to fill the blank space in order to create a texture block.

Secondly, they deleted the line space. For all the images, a texture block was generated by combining the normalized text copy. Next, for the feature extraction method consists in generating many images of input text with a different resolution based on steerable pyramids. The set of the new images present the same scene of the original image with different spatial resolutions. For each image, the authors calculate the curvelet features.

Finally, in the classification phase, they proposed a support vector machine (SVM) classifier. At this step they defined SVM 1vs.1 configuration to identify the font: SVM 1vs.1 which consists in a combination of 45 for 10 font classes.

D. MindGarage Font and Font-Size Recognition Systems

MindGarage systems are submitted by Vinaychandran Pondenkandath and Marcus Liwicki from MindGarage, TU Kaiserslautern, Germany.

To develop their systems, the authors use subsets from the freely available Arabic Printed Text Image (APTI) database, approximately 13 % of Set 1 for training and approximately 2.5 % of Set 4 for validation.

They use a Convolutional Neural Network based approach, and they use a ResNet18 model [11] which is pre-trained on the ImageNet dataset for the ImageNet Large Scale Visual Recognition Challenge [12] as the base model for all the tasks.

For font recognition, the authors replace the final layer with a fully connected layer with 10 outputs. Every input image is re-sized to a standard 224x224 resolution, and the network is trained to classify the font using the Adam optimizer with a learning rate of 0.0001.

In the case of font-size recognition, they re-size every image to a quarter of it's original resolution and pad it (or crop it if necessary) to the standard resolution of 224x224.

The final layer is replaced with 10 outputs, one for each font-size. The network is then trained in a similar manner as before.

For joint font and font-size recognition, they re-size every image to a standard size of 224x224, and then replace the final layer of the model with 100 outputs (10 fonts x 10 font-sizes). The network is then trained as before to identify the appropriate class.

E. MFR (Multi-Font Recognition) System

The MFR system was submitted by Asma ElAdel, Ridha Ejbali, Mourad zaied and Chokri Ben Amar; members in Research Team in Intelligent Machines unity at the National School of Engineers, university of Gabes from Tunisia.

MFR is a new system for Arabic script recognition. It is based on Convolutional Neural Networks using selected Beta filters by Adaboost algorithm.

The proposed system proceeds on three main stages: feature extraction, feature selection and classification.

In the first stage, the features are extracted, based on multiresolution analysis at different levels of abstraction, using fast wavelet transform (FWT). In order to learn the dataset, the features of the samples are extracted using Beta wavelet and analysed until the forth level.

In the second stage, the authors selected the best features, with their corresponding powers, that characterized each class (shape) of characters using Adaboost algorithm. They found that selecting 50 features, for each Font, are sufficient to discriminate between the different Font shapes. This second stage presents a very helpful phase in which, the selected features will be used in the next stage (the classification) to recognize the appropriate Font.

More details about the proposed system can be found in [13].

F. GU Font Recognition Systems

The GU_FontRecognition_A and GU_FontRecognition_B systems are submitted by Farshad Najafi a and Majid Ziaratban members of the Golestan University, Gorgan, Iran.

These systems are based on the matching of sub-images which are segmented from the words. These methods are the improved versions of the approach proposed in [14]. At first, by applying some preprocessing, the quality of images are improved, particularly for the words in small font-sizes. Then, the thickness of the base-line of an input word is calculated. The thickness value is used in our word-to-character segmentation method. Dots, small diacritics, and noise are removed [14]. The coordinates of the segmentation points are calculated by analyzing vertical histogram of the word image. Single characters and combinations of two connected characters are segmented and extracted from the word image. In the next stage, some features based on the Haar wavelet transform are extracted from the segmented sub-images [14]. The matched sample in the training set to the segmented test sub-image is found based on the extracted features. The font of the matched training sample is assigned to the test sub-image. A number of limitations have been considered for the lexicon reduction in the sub-image matching phase. The final font of an input word is recognized by voting among the fonts assigned to the sub-images of the word.

Table II
1^{st} APTIPC - RDI-CU SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	85.3	85.16	86.02	95.88	90.02	99.98	90.39
Andalus	82.32	61.66	76.90	87.38	93.3	96.68	83.04
Arabic Transparent	76.26	70.98	64.44	85.78	89.16	98.34	80.83
M Unicode Sara	91.22	74.96	86.12	93.84	96.82	99.96	90.49
Tahoma	79.06	89.22	92.98	96.06	91.06	97.4	90.96
Simplified Arabic	78.32	69.84	71.62	87.54	88.8	98.24	82.39
Traditional Arabic	56.9	44.48	63.8	72.08	72.92	95.82	67.67
DecoType Naskh	50.76	47.24	42.5	64.10	69.12	90.46	60.70
DecoType Thuluth	61.16	61.42	49.64	70.52	85.62	95.08	70.57
Diwani Letter	62.64	45.28	59.18	72.18	65.7	94.58	66.59
Mean RR	72.39	65.02	69.32	82.54	84.25	96.65	78.36

V. TESTS AND RECOGNITION RESULTS

All systems have been tested using the first 5000 images of *set 6* from the APTI database in different fonts and sizes. All participants sent us a running version of their recognition systems. The systems can be categorized in two groups depending on the operating system: RDI-CU, SP-Curvelet-FR, MFR and REGIM-LITIS, GU Font Recognition systems are developed under Microsoft Windows environment and MindGarage systems under Linux.

For each challenge the best result is marked in bold.

Three systems participate to the first APTI protocol (1^{st} APTIPC) concerning the font-size identification. Tables II, III and IV present respectively the RDI-CU, REGIM-LITIS and MindGarage system results of the first protocol. In this challenge, we try to recognize the font-size without recognizing the font or the content of the word image. The comparison of the systems based on the results shows that most systems recognize better word images generated with the sizes greater than 10. The MindGarage system shows the best results with an average of 99.67 % font-size identification rate. The RDI-CU and REGIM-LITIS systems show very good results in some tests. For example, the font-size identification rates with the Arabic word images generated with font “Advertising Bold”, font-size 18 are respectively 99.98 % and 94.96 % with RDI-CU and REGIM-LITIS systems.

The winner of the first protocol is the MindGarage System.

Table III
1^{st} APTIPC - REGIM-LITIS SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	55.52	31.42	15.18	51.96	36.78	94.96	47.64
Andalus	84.62	32.14	30.48	37.62	58.8	53.2	49.48
Arabic Transparent	76.84	38.74	31.48	57.9	62.44	81.40	58.13
M Unicode Sara	79.84	27.4	21.28	38.70	58.66	52.48	46.39
Tahoma	41.9	39.02	9.64	51	33.54	96.2	45.22
Simplified Arabic	75.96	38.44	28.9	59.72	62.12	82.98	58.02
Traditional Arabic	92.44	24.36	22.4	33.36	51.76	43.22	44.59
DecoType Naskh	64.16	21.66	16.54	36.4	39.98	72.10	41.81
DecoType Thuluth	63.9	32.98	24.42	58.36	44.72	93.6	53.00
Diwani Letter	76.66	36	23.28	54.58	52.78	74.24	52.92
Mean RR	71.18	32.22	22.36	47.96	50.16	74.44	49.72

Table IV
1^{st} APTIPC - MINDGARAGE SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	***Mean RR***
Advertising Bold	**99.96**	**100**	**100**	**99.94**	**99.94**	**100**	**99.98**
Andalus	**100**	**99.84**	**99.88**	**99.98**	**100**	**99.98**	**99.95**
Arabic Transparent	**99.94**	**99.9**	**99.88**	**99.98**	**99.98**	**100**	**99.95**
M Unicode Sara	**100**	**99.9**	**99.66**	**99.94**	**99.82**	**99.98**	**99.88**
Tahoma	**99.98**	**100**	**100**	**99.32**	**100**	**100**	**99.88**
Simplified Arabic	**99.98**	**99.92**	**99.84**	**99.98**	**99.7**	**100**	**99.90**
Traditional Arabic	**99.4**	**97.8**	**99.8**	**99.82**	**99.74**	**99.98**	**99.42**
DecoType Naskh	**98.8**	**99.32**	**97.18**	**99.62**	**99.52**	**100**	**99.07**
DecoType Thuluth	**99.9**	**95.94**	**99.98**	**99.88**	**99.98**	**100**	**99.28**
Diwani Letter	**98.04**	**99.5**	**99.4**	**99.68**	**99.84**	**100**	**99.41**
Mean RR	**99.6**	**99.21**	**99.56**	**99.81**	**99.85**	**99.994**	**99.67**

Seven systems participate to the second APTI protocol (2^{nd} APTIPC) concerning the font-family identification. Tables V, VI, VII, VIII, IX, X and XI present respectively the RDI-CU, REGIM-LITIS, SP-Curvelet-FR, MFR, MindGarage, GU_FontRecognition_A and GU_FontRecognition_B systems results of the second protocol. In this challenge, we try to recognize the font used to generate Arabic word images without recognizing the font-size or the content of the word image. All tested systems find difficulty to identify the font of images generated with the similar "Arabic Transparent" and "Simplified Arabic" fonts. However, most of the systems share good results with images generated with other fonts. The MindGarage system shows the best results with an average of 96.55 % font identification rate followed by GU_FontRecognition_A, GU_FontRecognition_B and RDI-CU systems with respectively 90.80 %, 87.87 % and 80.40 % font identification rates. The MFR, SP-Curvelet-FR and REGIM-LITIS systems show very good results with images generated respectively with "Andalus" (font identification average with MFR system 94.54 %), "Advertising Bold" (font identification average with SP-Curvelet-FR system 90.56 %) and "Diwani Letter" (font identification average with REGIM-LITIS system 94.41 %), but not with all the fonts.

The winner of the second protocol is the MindGarage System.

Table V
2^{nd} APTIPC - RDI-CU SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	75.56	94.68	97.9	98.02	99.08	99.62	94.14
Andalus	69.32	82.82	92.48	94.34	95.48	94.68	88.19
Arabic Transparent	45.26	40.38	42.1	39.74	21.58	19.48	34.76
M Unicode Sara	87.56	90.84	95.28	98.14	99.9	99.96	95.28
Tahoma	69.12	90.16	95.74	96.38	95.86	94.38	90.27
Simplified Arabic	40.72	57.44	63.5	65.76	78.48	79.96	64.31
Traditional Arabic	50.28	59.88	76.72	86.86	91.96	95.64	76.89
DecoType Naskh	56.88	67.60	79.8	84.88	92.3	89.4	78.48
DecoType Thuluth	66.5	83.66	88.62	95.94	99.26	99.7	88.95
Diwani Letter	80.82	87.36	93.18	96.66	99.32	99.26	92.77
Mean RR	64.20	75.48	82.53	85.67	87.32	87.21	80.40

Table VI
2^{nd} APTIPC - REGIM-LITIS SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	52.34	66.72	72.90	73.98	75.08	74.56	69.26
Andalus	63.32	85.16	88.48	88.86	88.3	88.12	83.71
Arabic Transparent	28.5	45.7	47.3	44.36	44.68	44.04	42.43
M Unicode Sara	47.26	72.16	74.44	73.26	74.90	74.42	69.41
Tahoma	62.34	66.78	64.88	65.06	57.7	64.90	63.61
Simplified Arabic	21.74	20.24	21.14	23.36	23.14	23.2	22.14
Traditional Arabic	51.38	52.4	50.82	53.04	49.38	47.72	50.79
DecoType Naskh	46.12	51.68	55.3	55.46	52.54	50.42	51.92
DecoType Thuluth	37.54	54.6	63.8	65.88	66.84	67.96	59.44
Diwani Letter	81.44	93.34	96.82	97.92	98.38	98.54	94.41
Mean RR	49.198	60.878	63.59	64.12	63.09	63.39	60.71

Table VII
2^{nd} APTIPC - SP-CURVELET-FR SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	74.98	89.56	91.56	92.36	97.96	96.94	90.56
Andalus	68.48	80.56	89.44	93.06	94.92	94.74	86.87
Arabic Transparent	36.04	44.82	54.98	61.58	56.7	60.14	52.38
M Unicode Sara	76.7	83	84.26	89.12	92.1	83.16	84.72
Tahoma	57.7	77.40	83.84	83.96	83.08	75.36	76.89
Simplified Arabic	14.2	22.86	30.24	25.64	44.32	34.64	28.65
Traditional Arabic	72.08	48.84	54.36	65.40	74.52	56.08	61.88
DecoType Naskh	49.28	57.56	67.54	77.38	87.74	85.02	70.75
DecoType Thuluth	70.08	81.26	86.68	88.06	90.84	88.12	84.17
Diwani Letter	71.3	79.08	86.18	91.14	96.56	94.96	86.54
Mean RR	59.08	66.49	72.91	76.77	81.87	76.92	72.34

Table VIII
2^{nd} APTIPC - MFR SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	82.28	90.72	92.96	94.5	95.2	95.4	91.84
Andalus	90.1	94.42	95.96	95.84	95.68	95.24	94.54
Arabic Transparent	38.69	48.69	44.91	44.83	40.01	41.07	43.03
M Unicode Sara	87.34	90.82	88	86.4	86.24	86.14	87.49
Tahoma	90.54	93.24	94.68	93.7	96.16	96.28	94.10
Simplified Arabic	30.73	43.37	51.15	47.69	50.35	50.49	45.63
Traditional Arabic	85.02	76.8	68.65	64.61	63.13	60.91	69.85
DecoType Naskh	64.05	70.47	73.43	72.45	71.99	71.75	70.69
DecoType Thuluth	76.58	84.66	85.5	84.22	82.2	81.34	82.42
Diwani Letter	81.14	87.6	88.34	89.06	86.4	84.92	86.24
Mean RR	72.65	78.08	78.36	77.33	76.74	76.35	76.58

Table IX
2^{nd} APTIPC - MINDGARAGE SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	***Mean RR***
Advertising Bold	**100**	**100**	**100**	**100**	**100**	**100**	**100**
Andalus	**99.94**	**100**	**100**	**100**	**100**	**100**	**99.99**
Arabic Transparent	**92.62**	**98.22**	**96.24**	**95.84**	**96.52**	**96.4**	**95.97**
M Unicode Sara	**99.98**	**100**	**100**	**100**	**100**	**100**	**100**
Tahoma	**100**	**100**	**100**	**100**	**100**	**100**	**100**
Simplified Arabic	**70.58**	**69.86**	**70.14**	**69.46**	**69.40**	**69.66**	**69.85**
Traditional Arabic	**99.76**	**99.82**	**99.96**	**99.88**	**100**	**99.96**	**99.90**
DecoType Naskh	**99.88**	**99.98**	**99.88**	**99.9**	**99.92**	**99.92**	**99.91**
DecoType Thuluth	**99.98**	**99.9**	**99.94**	**99.88**	**99.86**	**99.82**	**99.90**
Diwani Letter	**99.84**	**100**	**100**	**100**	**100**	**100**	**99.97**
Mean RR	**96.26**	**96.78**	**96.61**	**96.50**	**96.57**	**96.58**	**96.55**

Table X
2^{nd} APTIPC - GU_FONTRECOGNITION_A SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	99.06	99.36	99.58	99.54	99.78	99.9	99.54
Andalus	94.46	99.34	99.48	99.82	99.76	99.84	98.78
Arabic Transparent	66.60	77.04	74.44	81.5	78.28	82	76.64
M Unicode Sara	87.62	97.28	99.7	99.8	99.9	99.98	97.38
Tahoma	95.98	99.7	99.98	99.38	98.7	99.92	98.94
Simplified Arabic	52.82	54.18	57.82	49.24	54.76	56.08	54.15
Traditional Arabic	85.58	92.34	93.6	93.6	97.68	99.06	93.64
DecoType Naskh	88.9	94.64	95.1	96.7	97.58	98.08	95.17
DecoType Thuluth	91.66	96.2	96.46	98.52	98.84	99.24	96.82
Diwani Letter	91.92	95.6	97.66	98.08	99.28	99.34	96.98
Mean RR	85.46	90.57	91.38	91.62	92.46	93.34	90.80

Three systems participate to the third APTI protocol (3^{rd} APTIPC) concerning the font-family and the font-size identification. Tables XII, XIII,and XIV present respectively the RDI-CU, REGIM-LITIS, and MindGarage systems results of the third protocol. In this challenge, we try to recognize both the font and the size used to generate Arabic word images without recognizing the content of the word image. The mean font recognition rates are respectively 67.60 %, 29.59 % and 96.15 % for the RDI-CU, REGIM-LITIS, and MindGarage systems. The RDI-CU system shares close results to MindGarage system with for example images generated using the font "M Unicode Sara".

The winner of the third protocol is the MindGarage System.

Table XI
2^{nd} APTIPC - GU_FONTRECOGNITION_B SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	98.66	98.92	99	99.1	99.6	99.64	99.15
Andalus	90.74	98.22	98.48	99.22	98.82	99	97.41
Arabic Transparent	56.98	73.84	60.98	74.2	72.92	83.46	70.40
M Unicode Sara	81.98	96.64	99.6	99.84	99.9	99.98	96.32
Tahoma	89.56	98.04	99.82	98.84	99	99.92	97.53
Simplified Arabic	45.14	44.16	52.1	54.12	58.7	44.62	49.81
Traditional Arabic	74.06	87.06	88.2	86.68	93.4	97.74	87.86
DecoType Naskh	82	91.44	92.78	93.68	96.38	97.24	92.25
DecoType Thuluth	84.66	92.74	92.96	92.84	97.6	97.62	93.07
Diwani Letter	86.98	92.06	96.24	97.14	98.68	98.42	94.92
Mean RR	79.08	87.31	88.02	89.57	91.5	91.76	87.87

Table XII
3^{rd} APTIPC - RDI-CU SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	74.22	84.26	85.76	95.44	89.94	99.6	88.20
Andalus	62.06	57.32	75.98	86.3	92.02	94.68	78.06
Arabic Transparent	43.28	32.76	25.86	35.88	15.06	19.42	28.71
M Unicode Sara	81.64	74.08	85.98	93.78	96.82	99.96	88.71
Tahoma	67.66	86.9	92.38	95.14	89.06	94.2	87.56
Simplified Arabic	31.84	45.24	53.64	68.34	76.34	79.66	59.18
Traditional Arabic	23.92	36.96	59.84	70.78	71.96	95.54	59.83
DecoType Naskh	32.42	37.48	36.62	61.44	67.92	87.96	53.97
DecoType Thuluth	52.08	56.46	45.5	69.7	85.62	95	67.39
Diwani Letter	54.6	42.34	57.66	71.66	65.7	94.5	64.41
Mean RR	52.372	55.38	61.92	74.85	75.04	86.05	67.60

Table XIII
3^{rd} APTIPC - REGIM-LITIS SYSTEM RESULTS

System/Size	*6*	*8*	*10*	*12*	*18*	*24*	*Mean RR*
Advertising Bold	26.78	21.78	10.46	38.66	26.88	71.54	32.68
Andalus	54.24	27.98	26.94	32.80	52.08	47.36	40.23
Arabic Transparent	21.78	18.02	14.86	26.78	28.24	36.18	24.31
M Unicode Sara	37.04	22.24	17.06	29.74	45.6	41.34	32.17
Tahoma	18.8	24.14	5.58	31.24	17.56	63.84	26.86
Simplified Arabic	17.38	8.06	6.72	14.04	14.78	19.32	13.38
Traditional Arabic	47.46	12.66	11.6	19.86	28.06	22.26	23.65
DecoType Naskh	26.32	9.76	6.48	20.26	18.10	40.18	20.18
DecoType Thuluth	24.18	17.72	17.10	41.5	32.30	63.26	32.68
Diwani Letter	60.88	33.86	22.74	53.74	52.06	73.3	49.43
Mean RR	33.49	19.62	13.95	30.86	31.56	47.86	29.59

For the forth protocol (4^{th} APTIPC) about the multi-font and multi-size word recognition, we haven't received any system in this edition of the competition. It is a complex protocol which need a lot of time for training to take into account all the font, font-size and artifacts variabilities of the data. We present the state of the art best result in this protocol proposed by Siemens in the second edition of the competition (ICDAR'2013). We encourage researchers to evaluate their system in this protocol and compare their results with the Siemens system results. The presented results of the Siemens system are computed using set 6 of the APTI database. For more details about the Siemens system we refer to [5] and [15].

Table XV presents the Siemens system results of the fourth APTI protocol of the second edition of the compe-

Table XIV
3^{rd} APTIPC - MINDGARAGE SYSTEM RESULTS

System/Size	***6***	***8***	***10***	***12***	***18***	***24***	***Mean RR***
Advertising Bold	**99.96**	**100**	**100**	**99.94**	**99.94**	**100**	**99.97**
Andalus	**99.86**	**99.84**	**99.88**	**99.98**	**100**	**99.96**	**99.92**
Arabic Transparent	**89.14**	**90.86**	**67.94**	**69.86**	**69.02**	**82.1**	**78.15**
M Unicode Sara	**100**	**99.9**	**99.64**	**99.94**	**99.82**	**99.98**	**99.88**
Tahoma	**99.96**	**100**	**100**	**99.32**	**100**	**100**	**99.88**
Simplified Arabic	**75.7**	**76.60**	**95.1**	**95.52**	**94.58**	**83.66**	**86.86**
Traditional Arabic	**99.26**	**97.8**	**99.8**	**99.82**	**99.72**	**99.98**	**99.40**
DecoType Naskh	**97.84**	**99.28**	**97.14**	**99.58**	**99.48**	**100**	**98.89**
DecoType Thuluth	**99.86**	**95.86**	**99.98**	**99.88**	**99.98**	**100**	**99.26**
Diwani Letter	**97.36**	**99.5**	**99.38**	**99.54**	**99.78**	**99.9**	**99.24**
Mean RR	**95.89**	**95.96**	**95.89**	**96.34**	**96.23**	**96.56**	**96.15**

Table XV
4^{th} APTIPC - SIEMENS SYSTEM RESULTS (ICDAR 2013)

Font/Size		*6*	*8*	*10*	*12*	*18*	*24*	***Mean RR***
Advertising Bold	WRR	**99.86**	**99.96**	**99.97**	**99.97**	**99.96**	**99.95**	**99.95**
	CRR	**99.98**	**100**	**100**	**100**	**99.99**	**99.99**	**99.99**
Andalus	WRR	**98.93**	**99.88**	**99.92**	**99.91**	**99.92**	**99.76**	**99.72**
	CRR	**99.85**	**99.98**	**99.99**	**99.99**	**99.99**	**99.96**	**99.96**
Arabic Transparent	WRR	**99.57**	**99.92**	**99.99**	**99.97**	**99.99**	**99.94**	**99.90**
	CRR	**99.95**	**99.99**	**100**	**100**	**100**	**100**	**99.99**
M Unicode Sara	WRR	**95.70**	**97.63**	**97.66**	**97.75**	**97.82**	**97.68**	**97.71**
	CRR	**99.28**	**99.59**	**99.60**	**99.61**	**99.62**	**99.60**	**99.60**
Tahoma	WRR	**99.65**	**99.94**	**99.97**	**99.98**	**99.97**	**99.96**	**99.91**
	CRR	**99.96**	**99.99**	**100**	**100**	**100**	**99.99**	**99.99**
Simplified Arabic	WRR	**99.30**	**99.90**	**99.94**	**99.95**	**99.95**	**99.85**	**99.82**
	CRR	**99.90**	**99.99**	**99.99**	**99.99**	**99.99**	**99.99**	**99.98**
Traditional Arabic	WRR	**96.16**	**99.33**	**99.77**	**99.68**	**99.78**	**99.70**	**99.07**
	CRR	**99.51**	**99.92**	**99.97**	**99.96**	**99.97**	**99.95**	**99.88**
DecoType Naskh	WRR	**97.17**	**99.25**	**99.16**	**99.18**	**99.15**	**98.83**	**98.79**
	CRR	**99.61**	**99.89**	**99.41**	**99.50**	**99.87**	**99.83**	**99.69**
DecoType Thuluth	WRR	**96.35**	**99.24**	**99.92**	**99.92**	**99.44**	**99.27**	**99.02**
	CRR	**99.49**	**99.90**	**99.92**	**99.94**	**99.91**	**99.90**	**99.84**
Diwani Letter	WRR	**91.77**	**97.60**	**98.28**	**98.41**	**98.06**	**96.68**	**96.80**
	CRR	**98.70**	**99.64**	**99.72**	**99.74**	**99.68**	**99.44**	**99.49**
Mean RR	WRR	**97.64**	**99.27**	**99.46**	**99.47**	**99.40**	**99.16**	**99.07**
	CRR	**99.65**	**99.89**	**99.86**	**99.87**	**99.90**	**99.87**	**99.84**

tition (ICDAR'2013). The Siemens system evaluated with multi-font and multi-size images shares an average of 99.07 % for the word recognition rate and 99.84 % for the character recognition rate.

VI. CONCLUSIONS

This report gives an overview of the ICDAR2017 competition on multi-font and multi-size digitally represented Arabic text.

The objective of this third edition is to evaluate and compare different systems and approaches applied to font, size, font and size, and Arabic word recognition. Six groups presenting thirteen systems have participated at the three first protocols of the ICDAR2017 competition on multi-font and multi-size digitally represented Arabic text. We haven't received any system for the Fourth protocol.

MindGarage and Siemens systems show the best results respectively for font/size/font and Size and Word recognition compared to state-of-the art results using APTI database [16].

The MindGarage system is the winner in the first three protocols of this third edition of the competition.

VII. AKNOWLADGMENTS

We would like to thank all participants who helped us to succeed this third edition: ICDAR2017 competition on multi-font and multi-size digitally represented Arabic text.

REFERENCES

[1] H. E. Abed, V. Margner, M. Kherallah, and A. M. Alimi, "Icdar 2009 online arabic handwriting recognition competition," *ICDAR*, pp. 1388–1392, 2009.

[2] V. Margner and H. El Abed, "Arabic handwriting recognition competition," *ICDAR*, vol. 2, pp. 1274–1278, 2007.

[3] V. Margner and H. E. Abed, "Icdar 2009 arabic handwriting recognition competition," *ICDAR*, pp. 1383–1387, 2009.

[4] F. Slimane, S. Kanoun, H. Abed, A. Alimi, R. Ingold, and J. Hennebert, "Icdar 2011 - arabic recognition competition: Multi-font multi-size digitally represented text," *ICDAR*, pp. 1449–1453, 2011.

[5] F. Slimane, S. Kanoun, H. E. Abed, A. M. Alimi, R. Ingold, and J. Hennebert, "Icdar2013 competition on multi-font and multi-size digitally represented arabic text," *ICDAR*, pp. 1433–1437, 2013.

[6] F. Slimane, R. Ingold, S. Kanoun, A. M. Alimi, and J. Hennebert, "A new arabic printed text image database and evaluation protocols," *ICDAR*, pp. 946–950, 2009.

[7] ——, "Database and evaluation protocols for arabic printed text recognition," *DIUF-University of Fribourg - Switzerland*, 2009.

[8] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *CVPR*, vol. 1, pp. 886–893 vol. 1, 2005.

[10] A. Elbaati, M. Kherallah, A. Ennaji, and A. M. Alimi, "Temporal order recovery of the scanned handwriting," *ICDAR*, pp. 1116–1120, 2009.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CVPR*, pp. 770–778, 2016.

[12] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "Imagenet large scale visual recognition challenge," *IJCV*, vol. 115, no. 3, pp. 211–252, 2015.

[13] A. ElAdel, R. Ejbali, M. Zaied, and C. B. Amar, "Fast deep neural network based on intelligent dropout and layer skipping," *IJCNN*, pp. 897–902, 2017.

[14] M. Ziaratban and F. Bagheri, "Farsi font recognition based on the fonts of the text samples extracted by som," *JMCS*, vol. 15, no. 1, pp. 40–56, 2015.

[15] S. F. Rashid, M.-P. Schambach, J. Rottland, and S. von der Nüll, "Low resolution arabic recognition with multidimensional recurrent neural networks," *ICDAR-MOCR*, pp. 6:1–6:5, 2013.

[16] F. Slimane, "Multi-font and multi-size low-resolution arabic text recognition using apti-database," in *Document Analysis and Text Recognition: Benchmarking State-of-the-Art Systems*, S. in Machine Perception and A. I. (SMPAI), Eds. World Scientific Publishing Co, 2016, ch. 9.