Online Arabic Handwriting Recognition Competition

Monji Kherallah, Najiba Tagougui, Adel M. Alimi

University of Sfax,
Research Group on Intelligent Machines (REGIM)
Sfax,Tunisia
{monji.kherallah,najiba.tagougui,adel.alimi}@ieee.org

Haikal El Abed, Volker M¨argner

Technische Universitaet Braunschweig,
Institute for Communications Technology (IfN),
Braunschweig, Germany
{v.maergner,elabed}@tu-bs.de

Abstract—**Arabic script presents a challenge complexity and variability for handwriting recognition. The first on line Arabic Database called ADAB is known as a standard benchmark in the ICDAR competition of 2009. This paper describes the Online Arabic handwriting recognition competition held at ICDAR 2011. 3 groups with 5 systems are participating in the competition. The systems were tested on known data (sets 1 to 4) and on two test datasets which are unknown to all participants (set 5 and set 6). The systems are compared on the most important characteristic of classification systems, the recognition rate. Additionally, the relative speed of every system was compared. A short description of the participating groups, their systems, the experimental setup, and the performed results are presented.**

Keywords-On line Handwriting; ADAB-database; systems; recognition rate.

I. INTRODUCTION

In few last years, the handwriting analysis and recognition is a paramount subject of the researchers interest. The validation of the works done in this area was successfully established thank to the databases use. Two sorts of databases are considered. One interests the on line studies like UNIPEN and the other interests the off line studies like (CEDAR, IRONOFF, NIST, IFN/ENIT, etc.). All these databases are important for the research community in order to test new ideas and algorithms and to perform benchmarks and thereby measure progress and general tendencies. Large databases were developed for the handwriting recognition in Latin scripts. In contrast, very few databases have been developed for the Arabic script and fewer have become publicly available.

On line recognition of the cursive Arabic handwritten words, aims to contribute in the evolution of on line Arabic handwriting recognition research. Since 2009 the freely available (ADAB data base) is used by some groups all over the world to develop on line Arabic handwriting recognition systems. This database was the basis for the competition of ICDAR'2009 for systems that are specialized in on line recognition of the cursive Arabic handwritten words. This

ICDAR'2011 competition uses as a next step the same background of the ADAB database but now with and extended collected data of freely written words. These sets are unknown to all participants. Note that the writers of the set 5 are adults, whereas set 6 it consists of a collection made by young school students (between 9 and 13 years old).

A comparison and discussion of different algorithms and recognition methods should give a push in the field of on line Arabic handwritten word recognition.

Our paper is written as follow: The next section describes the evaluation process. Section 3 gives the details of ADAB-database. Section 4 presents the participants and their systems description. Section 5 deals with results and discussion. The last section announces the winner of this competition and the future prospects.

II. EVALUATION PROCESS

The object is to run each Arabic handwritten word recognizer (trained on a part of version 2.0 of the ADAB-database) on an already published part of the ADAB-database and on a test set not included in the published part. The recognition results on word level of each system are compared on the basis of correct recognized words, i.e. there correspondent consecutive Numeric Character References (NCR). A dictionary can be used in the recognition process. A recognizer may return up to 10 candidates for each classification that not only the first ranked result can be used for comparison but also the correct result between the 5 or 10 candidates will be used for comparison. The evaluation process of all systems will be released in our laboratory REGIM: Group of Research on Intelligent Machines. We run the recognizer (called myrec) by invoking it from the command line as follows: myrec input.txt output.txt. Fig. 1 presents an example of the input file which is just a list of relative paths to each *.inkml online trace to be recognized.

The output file should have one line as result for each input file. Each line should show the name of the online

trace file that was recognized, followed by the responses (sequence of NCRs code) for that file. Each response is given as a pair of values: the text, followed by the confidence.

Fig. 2 presents an example of the output file, the first line shows that for the file word/1.inkml the recognizer has produced two word hypotheses (بــوذر and نبّر) with confidences of 1.0 and 0.3 respectively.
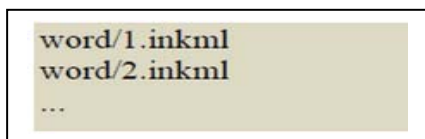
```
word/1.inkml
word/2.inkml
...
```

Figure 1. Input file

```
word/1.inkml UFE91 UFEEE U0630 U0631 1.0 UFEE7 UFE92 U0651 UFEAE 0.3
word/2.inkml U062A U0627 U0644 U0629 0.7 U0645 U0627 U0631 U062B 0.4 U0634 U064A U0627
U0628 0.2
```
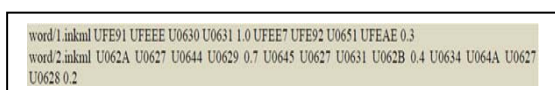
Figure 2. Output file

III.   ADAB-DATABASE DESCRIPTION

The database ADAB (Arabic DAtaBase) was developed to advance the research and development of Arabic on-line handwritten text recognition systems. This database is developed in cooperation between the Institut fuer Nachrichtentechnik (IfN) and the Research group on intelligent Machines (REGIM). The database consists of 20575 Arabic words handwritten by more than 170 different writers, most of them selected from the narrower range of the National school of Engineering of Sfax (ENIS). ADAB database is developed by a special tool for the collection of data and verification of the ground truth, which will be available for other groups for the collection of their own data in the same form of the ADAB database. These tools give the possibilities to record the on-line written data, to save some writer information, to select the lexicon for the collection, and re-write and correct wrong written text. Ground truth was added to the text information automatically from the selected lexicon and verified manually. The ADAB database is freely available for non commercial research (www.regim.org) [1]. Our aim was to collect a database of handwritten town names written in a similar quality as on a Mobile Phone with a digital input device. The collection process starts when the writer clicks on start bottom. The collection tool generates a town name randomly from 937 Tunisian town/village names, the writer must write the displayed word as it is shown in Fig.3. A pre-label will be automatically assigned to each file. It consists of the postcode in a sequence of Numeric Character References which will be stored in the UPX file format. An InkML file including trajectory information and a plot image of the word

trajectory are also generated. Additional information about the writer can also be provided.

The ADAB-database is divided to 6 sets. Details about the number of files, words, characters, and writers for each set 1 to 6 are shown in Table 1.
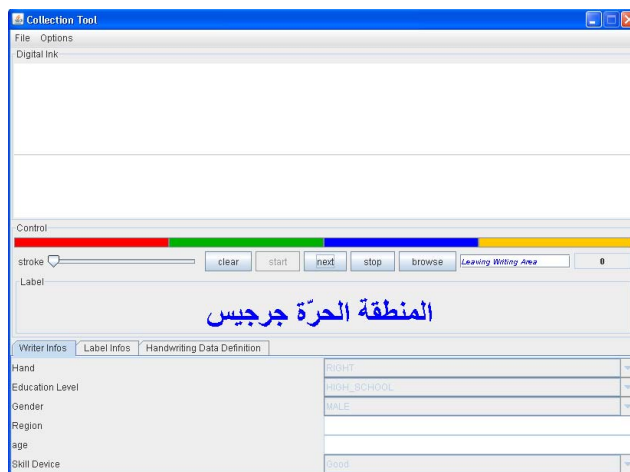

Figure 3. ADAB's collection tool

TABLE I.   ADAB SETS

Set	Files	Words	Characters	Writers
1	5037	7670	40500	56
2	5090	7891	41515	37
3	5031	7730	40544	39
4	4417	6786	35832	25
5	1000	1551	8189	6
6	1000	1536	8110	3
Sum	21575	33164	174690	166

IV.   PARTICIPATING SYSTEMS

The recognition process is divided into pre-processing steps and subsequent classification. In this section, we present a brief description of the systems submitted to the competition. Each system description has been provided by the system's authors and edited (summarized) by the competition organizers.

A.   VisionObjects

This systems is submitted by the VisionObjects company. They have built a cursive Arabic handwriting recognition system for this competition based on MyScript handwriting recognition technology. The overall system follows the following concepts:

- Use of a modular and hierarchical recognition system
- Use of soft decisions (often probabilistic) and deferred decisions by means of considering concurrent hypotheses in the decision paths

- Use of complementary information at all stages of the recognition process, and
- Use of global optimization criteria, making sure that the recognizer is trained in order to perform optimally on all levels.

The processing chain of the recognizer starts out with some of the usual preprocessing operations, such as ink smoothing and reference line detection. Then the on-line handwriting is pre-segmented into strokes and sub-strokes. The general idea is to over-segment the signal and let the recognizer decide later on where the boundaries between characters and words are. Here, specific techniques for processing diacritical marks have been employed to assure the proper association of letters and their diacritical marks. This segmentation stage is followed by the feature extraction stage. Feature sets use a combination of on-line and off-line information at various resolutions, including some higher level structural features. The feature sets are processed by a set of character classifiers, which use Neural Networks and other pattern recognition paradigms. The total number of characters classes is 150, which corresponds to the number of Arabic letters multiplied by the number of different shapes for each letter (initial, medial, final and isolated), plus some other symbols encountered in Tunisian cities like digits or the Latin letter 'V'. All the information accumulated in the various processing steps is then processed by dynamic programming on the word and sentence level in order to generate character, word, and sentence level candidates with corresponding confidence scores. A global discriminate training scheme on the word level with automatic learning of all classifier parameters and meta-parameters of the recognizer is employed for the overall training of the recognizer. For the recognition process, a lexicon containing around 1000 Tunisian city names is employed. We have designed the recognizer according to two different criteria. The first system (VisionObjects-1) provides the best accuracy whereas the second system (VisionObjects-2) is faster in exchange for a somewhat lower accuracy.

B. AUC-HMM

Two systems are submitted by Hesham Eraqi, Hany Ahmed and Sherif Abdelazeem working in American University of Cairo (AUC).
These systems are based on Hidden Markov Model (HMM). Beside the common use of the off-line features for the HMM-based Arabic recognition systems, the authors add the use of on-line features and combination of the two approaches. Delayed strokes are a well-known problem in on-line handwriting recognition due to its varying writing order among different writers. This problem is solved by removing these delayed strokes by using a new delayed strokes detection approach that makes use of the baseline information and the shape of the strokes, while the baseline detection method used in our system is based on horizontal projection. Removing delayed strokes has also led to the ability to combine some of the Arabic characters that share

the same primary stroke and are only distinguishable by their delayed strokes into one class (HMM model), which increased the recognition rate. The interesting thing is that the detection of the delayed strokes has not only increased the performance of the HMMs and the whole system significantly, but it is also used in the newly developed algorithm for lexicon reduction [2] that depends on the formation of the delayed strokes sequence and how far it is from the expected delayed strokes sequence of each lexicon entry.

C. FCI-CU-HMM

Two systems are submitted by Ibrahim Hosney, Sherif Abdou, Aly Fahmy and Mostafa Shahin working in Faculty of Computers and Information, Cairo University (FCI-CU).
These systems present a new approach to online Arabic handwriting recognition system based on Hidden Markov models (HMM). The HMM is a flexible tool that can search all the possible segmentation hypotheses for a word to find the optimum one, the one with highest match with the training data that the model has seen before [3]. Each letter is represented by a set of states. The recognizer has a set of phases. The first phase is preprocessing for the input including smoothing, re-sampling and interpolation [4, 5]. This is followed by a rearrangement of delayed strokes by detecting and inserting them in its proper location [6]. The second phase is extracting the features from the output of the previous phase; the set of features used here include chain code, curvature, vertical position from baseline and slope [5, 7, 11]. Finally a Viterbi decoder is used to recognize the input handwriting sample. To train the system, we followed a set of steps. Initially a single re-estimation of the parameters of the set of HMMs, using linear transforms, is employed. This is followed by writer adaptation training using constrained maximum likelihood linear regression [8]. After that we tie states within tri-grapheme sets in order to share data and thus be able to make robust parameter estimates. Finally Gaussian PDFs are converted into Mixture Gaussian PDFs for the recognition process; the system is using a dictionary of around 1000 Tunisian city names [1, 9, 10, 12]. The only difference between two system versions is that the second version is based on context-dependent graphemes with tied state mixtures while the first version is based on mono-grapheme models.

V. EVALUATION RESULTS AND DISCUSSION

For the evaluation process we plan to evaluate the 6 systems by 3 levels of evaluation. The first one consists of the testing by using 3 subsets extracted randomly from the sets 1, 2, 3 and 4 (see table 2). The second level consists to testing the systems using set 5 and set 6 (see table 3). These sets are unknown to all participants. Note that the writers of the set 5 are adults, whereas set 6 it consists of a collection made by young school students (between 9 and 13 years

old). In a third level, we are testing the rapidity of the different systems (See table 4). In fact, the processing performance of the systems was compared on two data subsets t and t1.

A. Recognition Results and discussion

As a first level of evaluation, the comparison of the systems based on the results of subsets 1 to 4, which are part of the training set, shows 3 systems with a recognition rate better than 80 % on sets 1 to 4. As it is shown in table 2, the recognition rate in all systems is limited between 68.2 % and 99.53 %. In the second level we consider the sets 5 and 6. As it is shown in table 3 the recognition rate is limited between 60.28 % and 98.97%.

B. Speed Tests (sets 1 and 6)

The average processing time per name on the two tests sets 1 and 6 respectively is shown in Table 4. A substantial difference in speed can be observed. The slowest system is about 137 times slower than the fastest one. An average processing time of 81.4 ms per name image with set 6 is a very good result.

VI. CONCLUSION

The competition results show that Online Arabic handwriting recognition systems made a remarkable further progress. Most of the participating systems show a very high accuracy and some also a very high speed. Details and specific features of the systems cannot be presented in this short paper. All systems are based on HMM approach. Note that the HTK or HMM are known and used largely in speech recognition. In this competition it was demonstrated that the HMM tool is also powerful in the field of handwriting recognition. The system of Vision objects 1 is the winner of this competition. This system represents the shortest average processing time.

REFERENCES

[1] H. El Abed, M. kherallah, V. Mârgner and A. M. Alimi, "on-line Arabic handwriting recognition competition ADAB database and participating systems", International Journal on Document Analysis and Recognition, Vol. 14, Num. 1, pp 15-23, 2010.

[2] H. Eraqi and S. Abdelazeem, "On-line Arabic Handwritten Personal Names Recognition System based on HMM," International Conference on Document Analysis and Recognition (ICDAR 2011).

[3] S. J. Young et al., "The HTK Book", Entropics Cambridge Research Lab, Cambridge, UK, 1995.

[4] B. Q. Huang, Y. B. Zhang, M. Tahar Kechadi. "Preprocessing Techniques for Online Handwriting Recognition". In Collection of Intelligent Text Categorization and Clustering, pp. 25~45, 2009.

[5] S. Jaeger, S. Manke, J. Reichert, and A. Waibel, "Online Handwriting Recognition: The NPen++ Recognizer," International Journal on Document Analysis & Recognition, vol. 3(3), pp. 169-180, 2001.

[6] S. Abdou, A. Fahmy, I. Hosney and I. Mostafa. "Artificial Tutor For Arabic Handwriting Training". In Proceedings of the Second International Conference on Arabic Language Resources and Tools, 2009.

[7] E. Tapia, "Understanding Mathematics: A System for the Recognition of On-Line Handwritten Mathematical Expressions". Freie Universität Berlin, Fachbereich Mathematik u. Informatik, 2004

[8] M. J. F. Gales and P. Woodland. "Speech recognition the HTK way: Tutorial session". In Proceeding of ICASSP, 2006.

[9] Haikal El Abed, Volker Märgner, Monji Kherallah and M.Adel Alimi: Online Arabic Handwriting Recognition competition. ICDAR'09.1388-1392, 2009

[10] M. Kherallah, L. Hadded, A. Mitiche, A. M. Alimi. "On-Line Recognition Of Handwritten Digits Based On Trajectory And Velocity Modelling". International journal of Pattern Recognition Letter. Vol. 29. pp. 580-594,2008.

[11] M. Kherallah., F. Bouri., And A. M. Alimi. "On-Line Arabic Handwriting Recognition System Based On Visual Encoding And Genetic Algorithm." Engineering Applications of Artificial Intelligence. Vol. 22 . pp 153–170, 2009.

[12] M. Hamdani, H. El-Abed, M. Kherallah and A. M. Alimi: Combining Multiple HMMs using On-line and Off-line Features for Off-line Arabic Handwriting Recognition.ICDAR'2009. pp 201-205, 2009.

TABLE II. FIRST LEVEL OF EVALUATION

Systems	Subset 1			Subset 2			Subset 3			Subset 4		
	Top 1	Top 5	Top 10	Top 1	Top 5	Top 10	Top 1	Top 5	Top 10	Top 1	Top 5	Top 10
AUC-HMM1	96.55	98.7	98.7	95.97	98.72	98.72	97.39	98.49	98.49	68.2	70.7	70.7
AUC-HMM2	96.2	98.2	98.3	95.87	98.5	98.5	97.39	98.09	98.09	68.2	70.3	70.5
FCI-CU-HMM1	79.1	92.6	94.4	87.3%	95.8	96.6	89.27	97.57	98.37	89.4	96.7	97.5
FCI-CU-HMM2	89.2	96.8	97.4	92.7	96.8	97.5	94.68	98.98	99.18	93.4	97.8	98.1
V-O 1	99.24	99.24	99.24	99.16	99.43	99.43	98.76	98.98	98.98	98.56	98.98	98.98
V-O2	99.01	99.29	99.29	99.33	99.51	99.53	99.00	99.27	99.31	98.05	98.17	98.41

TABLE III.   SECOND LEVEL OF EVALUATION

Systems	Set 5			Set 6		
	Top 1	Top 5	Top 10	Top 1	Top5	Top 10
AUC-HMM1	60.58	69.2	69.78	63.6	67.4	67.8
AUC-HMM2	60.28	68.89	69.50	63.4	66.7	67.1
FCI-CU-HMM1	62.06	81.71	85.51	66.06	83.7	87.21
FCI-CU-HMM2	67.3	83.2	85.82	71.2	87.5	89.2
V-O 1	98.89	99.18	98.18	98.45	98.97	98.97
V-O2	98.02	98.13	98.13	98.11	98.55	98.55

TABLE IV.   EXECUTION TIME OF SYSTEMS

Systems	Set 1(1000 samples)	Set 6 (1000 samples)
AUC-HMM1	120000 s	10560 s
AUC-HMM2	13890 s	11220 s
FCI-CU-HMM1	1117 s	1328 s
FCI-CU-HMM2	1093 s	1712 s
V-O 1	74.590 s	81.451 s
V-O2	271.654 s	265.236 s