International Conference on Frontiers in Handwriting Recognition (ICFHR 2010) – Competitions Overview

Haikal El Abed and Volker Märgner
Technische Universität Braunschweig
Institute for Communications Technology (IfN)
Braunschweig, Germany
Email: elabed@tu-bs.de and v.maergner@tu-bs.de

Michael Blumenstein
Science, Environment, Engineering and Technology Group
Gold Coast Campus, Griffith University
Queensland, Australia
Email: M.Blumenstein@griffith.edu.au

Abstract—The great success and high number of participants in pattern recognition related competitions last years show an important improvement of recognition and classification approaches. This success is unconceivable without the availability of huge datasets of real world data. We have invited for proposals for competitions to be held in the framework of the 12th International Conference on Frontiers in Handwriting Recognition (ICFHR2010). These competitions should aim at evaluating the performance of algorithms and methods for a particular task of Handwriting Recognition. Eight different teams composed of more than one group have submitted their proposals. The subjects of these propositions cover the field of research of handwriting recognition from pre-processing over handwritten document analysis to handwriting text/word recognition. These competitions represent an overview of current research topics and frontiers in handwriting document analysis and recognition. Only 5 competitions have received enough participants (we have defined the threshold to 3 systems) to present their evaluation at the ICFHR 2010.

Section (II to IX) a short description of each competition is presented. These descriptions include the organization steps. This rate is used to measure the effective potential of comparing the quality of recognition systems. For the ICFHR 2010 competition we have defined a so called "Participation rate" ($P_r$), as a measure between the number of groups have mentioned registered at a competition and the number of groups, which have submit their systems for the evaluation steps. This rate is used to measure the effective potential of a competition.

This overview paper is organized as follow: in each Section (II to IX) a short description of each competition is presented. These descriptions include the organization teams, a short description of the objectives of the competition and a list of participating groups, if any.

II. ARABIC HANDWRITING RECOGNITION COMPETITION

The Arabic handwriting recognition competition is organized by Volker Märgner and Haikal El Abed from the Technische Universität Braunschweig, Institute for Communications Technology (IfN), Braunschweig, Germany.

This Arabic Handwriting Recognition Competition aims to bring together researchers working on Arabic handwriting recognition. Since 2002 the freely available IfN/ENIT-Database is used by many groups all over the world to develop Arabic handwriting recognition systems. This database was the basis for the last years competitions [1], [2], [3] for systems recognizing Arabic handwritten words. For this year, 11 groups have mentioned their interests to participate at the competition, finally only 4 groups with 6 systems have submitted their systems for the evaluation phase. Two groups participate for the first time at the Arabic handwriting recognition competition. A list of the participating system...
is presented in Table I. The participation rate of this competition is \( P_r = \frac{36}{11} = 36.36\% \).

### III. BANGLA HANDWRITING RECOGNITION COMPETITION

The Bangla handwriting recognition competition was proposed by Dr. Utpal Garain from the Indian Statistical Institute, Kolkata, India and Mr. Nibaran Das from the Jadavpur University, Kolkata, India. Bangla is the second most popular language and script in the Indian subcontinent and among the top ten popular languages/scripts in the world. As far as recognition of Bangla handwriting is concerned, there are several research contributions as reported in the literature but in almost all studies, recognition accuracies had been reported based on non-standard, private datasets.

Therefore, the aim of this competition was to bring prospective researchers/groups working in this area together and to provide them with a platform for evaluating the performance of their systems, comparing one system with others using a common database in order to document state-of-the-art of the respective research. The competition was restricted to recognition of off-line handwritten Bangla (i) isolated digits (10 classes), (ii) isolated basic characters (vowels and consonants, altogether 49 classes) and (iii) a set of frequently used isolated 150 compound (conjuncts of two or more consonants) characters. Training data for digit and basic character classes were prepared by Mr. Ujjwal Bhattacharya and his team in Indian Statistical Institute, Kolkata; whereas training data for compound character classes and the whole test dataset were contributed by Mr. Nibaran Das and his group of Jadavpur University, Kolkata.

Six groups from different countries registered themselves for participating in this competition and receiving the training data. Unfortunately, only one group submitted their system and others informed that they couldnt finish. The system submitted from Mr. Philippe Dreuw and Mr. Tobias Gass of RWTH Aachen University, Department of Computer Science 6, Germany, was named as RWTH-OCR. The system makes use of a k-NN (nearest neighbor) classification technique for recognition. The distance function employs a constrained energy minimisation based method, details of which have been published in the 20th Int. Conf. on Pattern Recognition (ICPR) 2010, Istanbul, Turkey. A two-dimensional warping is found for each training image to a given test image. The classes of the most similar training images then determine the classifier output by maximum vote. The warping is obtained by optimizing a pixel labelling in a Markov-random-field (MRF), where a label assigns a warped position for each pixel. Testing of RWTH-OCR shows that the classifiers are computationally very expensive. On a todays workstation of moderate resources, the recognition of a digit takes about 15 seconds, whereas time requires for recognizing an alphabet is approximately 75 seconds. It took more than 2 hours to classify 500 test samples of digits. The recognition accuracies achieved by the classifiers are also not impressive. However, the participants lately identified some bug in their code and they do hope that the corrected code would show better performance.

#### A. Organizers Impressions

Holding of this competition reveals an important aspect that we perhaps continue to organize this particular competition track for few more times to attract more research groups from pattern recognition and machine learning communities, in general and handwriting recognition community, in particular, in order to achieve a state-of-the-art solution of the problem. Moreover, the organizers, in future, may decide whether registration fees can be introduced to keep casual participants from joining such competition. It is experienced that many groups showed initial interest, took training data but finally did not submit any system. Many among these were possibly not that serious (many might have participated to get the training data and use it in their future research) about the competition. A concept of registration fees may help a bit in this regard.

### IV. FORENSIC SIGNATURE VERIFICATION COMPETITION (4NSigComp2010)

#### A. Scenario 1

The 4NSigComp2010 – Forensic Signature Verification Competition: Scenario 1 (Detection of Simulated and Disguised Signatures) is organized by:

- Marcus Liwicki and Muhammad Imran Malik from the German Research Center for AI (DFKI), Germany
- C. Elisa van den Heuvel from the Netherlands Forensic Institute, The Hague, The Netherlands
- Bryan Found from the Victoria Police & La Trobe University, Melbourne, Australia

This competition scenario aims at a performance comparison of several automated systems for the task of signature verification. The systems have to rate the probability of authorship and non-authorship of signatures. In particular they have to detect skilled forgeries (simulated and disguised
signatures) from genuine signatures of a reference writer. Furthermore, the results will be compared to forensic handwriting examiners (FHEs) on the same tasks. As such, to the best of the authors knowledge, this scenario will be the first attempt in literature to relate system performances to the performance of FHEs who gave their opinion on exactly the the same signatures.

The organizers of this competition have received 6 different submissions from 5 groups (Details about the participating systems are listed in Table II). All of them ran smoothly and gave good results especially if no disguised signatures were considered. Considering the results of the 4NSigComp2010 and the importance of forensic handwriting verification the organizers can say that computer scientists should also focus on disguised signatures, since it is a crucial aspect in real FHEs’ casework. For a next competition at ICFHR the organizers plan to use a larger test set to investigate the diversity of the recognizers more thoroughly. Regarding simulations systems produced quite good results. Regarding genuine signatures, large and diverse test sets where signatures are produced by the different authors under various different psychological and physical conditions may also yield interesting results.

An interesting observation of this contest is that the performance of the automated systems is not so far away from human decisions. A more detailed analysis will be performed in future to directly assess the strengths and weaknesses of several classifiers. Also, in regard to the fact that FHEs cannot make use of thresholds but need to provide the court the likelihood of the two competing hypothesis, in a next competition the organizers will reference existing system scores into a forensic system using within-source and between-source variabilities according to the Bayesian approach.

### B. Scenario 2

The 4NSigComp2010 – Forensic Signature Verification Competition: Scenario 2 is organized by:

- Michael Blumenstein is affiliated with the School of Information and Communication Technology at Griffith University in Queensland, Australia.

The objective of this competition (4NSigComp2010) was to ascertain the performance of automatic off-line signature verifiers to evaluate recent technology developments in the areas of document analysis and machine learning. The competition focused on the second scenario, which aimed at performance evaluation of off-line signature verification systems on a newly-created large dataset that comprises genuine, simulated signatures produced by unskilled imitators or random signatures (genuine signatures from other writers).

After the competition announcement, 15 teams showed their interest in participating in the competition. Of those 15 teams, 7 submitted their programs: 3 from academia, 1 from a governamental institution and the remainder from industry. Two groups submitted several programs, therefore the competition evaluated 10 different systems. The teams are from 7 different countries: India, Canada, Turkey, U.S.A, Austria, Spain and France. Table III shows the participants groups with the identification of the programs submitted. The participation rate of this competition is \( P_r = \frac{7}{15} = 46.67\% \).

Scenario 2 of the 4NSigComp2010 competition had very positive participation with ten systems being submitted and evaluated on a newly created large database of off-line signatures. Overall, the systems performed well in terms of speed and accuracy. The top ranking system attained an overall error of 8.94%. This result interestingly correlates with the top ranking accuracy achieved in a previous signature verification competition at ICDAR 2009 [4].
V. QUANTITATIVE EVALUATION OF BINARIZATION ALGORITHMS OF IMAGES OF HISTORICAL DOCUMENTS WITH BLEEDING NOISE COMPETITION

The Quantitative Evaluation of Binarization Algorithms of Images of Historical Documents with Bleeding Noise Competition is organized by:

- Ergina Kavallieratou from the University of the Aegean, Greece
- Rafael Dueire Lins from the Universidade Federal de Pernambuco (UFPE), Brazil
- Roberto Paredes from the PRHLT Group, Universidad Politécnica de Valencia, Spain

The evaluation and comparison of binarization algorithms proved to be difficult task since there is no objective way to compare the results. Several review papers have tried to compare binarization algorithms by using the precision and recall analysis of the resultant words in the foreground or by evaluating their effect on end-to-end character or word recognition performance in a complete archive document recognition system utilizing OCR. Every work, that performed comparison, presented some very interesting conclusions. However, the problem is that in all cases, they try to use results from ensuing tasks in document processing hierarchy, in order to survey the algorithm performance. Although in many cases this is the objective goal, it is not always possible. Such is the case of historical documents, where their quality, in many cases obstructs the recognition, and sometimes even the word segmentation, this way of evaluation can be proved problematic.

On the other hand, we need different evaluation technique, since the processing of historical documents is one of the hardest cases and binarization can be required for removing the noise and facilitate their appropriate presentation. The ideal way of evaluation should be able to decide, for each pixel, if it finally has succeeded the right color (black or white) after the binarization. This is an easy task for a human observer but very difficult to do it automatically for all the pixels of several images.

The proposed method includes the experimentation on document archives made by constructing noisy images, using techniques of image mosaicing, and combining old blank historical document pages with noise-free pdf documents. After the application of the binarization algorithms to the synthetic images, it is easy to evaluate the results by comparing the resulted image with the original document as ground truth image. This way, we are able to count the exact amount of the remaining wrong pixels either on background or on foreground.

A list of participating systems is given in Table IV.

VI. HANDWRITTEN HISTORICAL DOCUMENT RECOGNITION COMPETITION

The first Handwritten Historical Document Recognition Competition was proposed by Nicolás Serrano, Oriol Ramos Terrades and Alfons Juan from the PRHLT group Instituto Tecnológico de Informática Univ. Politécnica de Valencia, Cami de Vera, Valencia, Spain.

Summary lack of data for training statistical models is one of the main difficulties that researchers has to face when transcribing old handwritten documents. This competition aims to evaluate different methods to extract information from a real and complete historical document when few data is available. Thus, two different scenarios are proposed to participants on the same database: handwritten text recognition and word spotting. The proposed document collection is the RODRIGO database, which has recently been released and is fully annotated. The RODRIGO database corresponds to a manuscript from 1545 entitled Historia de España del arzobispo Don Rodrigo, and it is completely written in old Castilian (Spanish) by a single author. It is a 853-page bound volume describing chronicles from the Spanish history. Most pages only contains a single text block of nearly calligraphic handwriting on well-separated lines.

Three groups have mentioned their interests to participate at this competition, however neither of them have submitted their systems for the evaluation steps. The participation rate of this competition is \( P_r = \frac{0}{3} = 0\% \).

VII. HANDWRITING SEGMENTATION COMPETITION

The Handwriting Segmentation Competition is organized by Basils Gatos and Nikolaos Stamatopoulos and Georgios Louloudis from the Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, National Center for Scientific Research "Demokritos", Agia Paraskevi, Athens, Greece.

This competition represent the third contest in this series, after the competitions organized at the ICDAR 2007 [5] and at ICDAR 2009 [6]. The objective of the competition is to use well established evaluation practices and procedures in order to record recent advances in off-line handwriting segmentation. Two new benchmarking datasets, one for text line and one for word segmentation, were created in order

---

Table IV

<table>
<thead>
<tr>
<th>S-ID</th>
<th>Group (Corresponding Author)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Synchro1,2,3]</td>
<td>Synchronedia Laboratory for multimedia communication in telepresence of Ecole de technologie supérieure (ETS), University of Quebec (Reza Farrahi Moghaddam)</td>
</tr>
<tr>
<td>[Nifisoft]</td>
<td>NifSoft, Saint-Etienne, France (Abdel³cili Hassaine)</td>
</tr>
<tr>
<td>[TEI]</td>
<td>Technological Educational Institution (TEI) of Athens, Greece (Anguelos Nicolaou)</td>
</tr>
<tr>
<td>[LSIS]</td>
<td>LSIS, UMR CNRS 6168 South University of Toulon-Var, France (Thibault Lelore)</td>
</tr>
</tbody>
</table>
to test and compare recent algorithms for handwritten document segmentation in realistic circumstances. Handwritten document images were produced by many writers in several languages (English, French, German and Greek). This

For this competition, 19 groups have mentioned their interest to participate. However, only five research groups have participated in the competition with seven different algorithms (two participants submitted two algorithms each). The list of the participating groups is presented in Table V. The participation rate of this competition is \( P_r = \frac{5}{19} = 26.32\% \).

VIII. ON-LINE ARABIC HANDWRITING RECOGNITION COMPETITION

The On-Line Arabic Handwriting Recognition Competition is the second competition for on-line Arabic script. The first competition in this series [7], [8] was organized at the ICDAR 2009 conference in Barcelona, Spain. The ICFHR 2010 edition of the on-line Arabic competition is organized by

- Monji Kherallah and Adel M. Alimi from the REGIM: Research Group on Intelligent Machines, University of Sfax, National School of Engineers (ENIS), Sfax, Tunisia
- Haidak El Abed and Volker Märgner from the Technische Universität Braunschweig, Institut fuer Nachrichtentechnik (IfN), Braunschweig, Germany
- Ioannis Pratikakis from The Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, National Center for Scientific Research “Demokritos”, Agia Paraskevi, Athens, Greece and from the Department of Electrical and Computer Engineering Democritus University of Thrace, Xanthi, Greece
- Basilis Gatos and Konstantinos Ntiogiannis from the Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, National Center for Scientific Research ”Demokritos”, Agia Paraskevi, Athens, Greece.

Document image binarization is an important step in the document image analysis and recognition pipeline. Therefore, it is imperative to have a benchmarking dataset along with an objective evaluation methodology in order to capture the efficiency of current document image binarization practices. This competition represent the second one after the DIBCO 2009 organized in conjunction with ICDAR09 [9]. The general objective of the H-DIBCO 2010 is to record recent advances in handwritten document image binarization using established evaluation performance measures. We will create a benchmarking dataset that is representative of the potential problems which are challenging in the binarization process and use a common evaluation platform in order to test and compare the submitted algorithms for handwritten document image binarization.

A total of 31 distinct expressions of interest have been received. Finally, the competition comprises 17 algorithms that have been submitted by the groups listed in Table VI (The system ID, S-ID, denotes the order of expressing the interest to participate in the competition). The participation rate of this competition is \( P_r = \frac{17}{31} = 51.61\% \).

X. CONCLUSION

Testing recognition systems with large identical datasets is crucial for performance evaluation. Another challenge comes from their complexity, because they consist of many specialized parts solving very diverse tasks. The competitions organized at the 12th International Conference on Frontiers in Handwriting Recognition (ICFHR 2010) show the importance of such evaluation contests to make an important improvement in a research field. Some of these research fields are interesting for different groups and some other research fields are in the start phase and just a couple of groups are working on it. It is interesting to see the variation of the “Participation rate” between the different competitions, from about 0% to 100%.

ACKNOWLEDGMENT

The ICFHR 2010 Competition Chairs would like to thank ALL competition organizers for the time and effort spend to make the competition successful and ALL competition participants.
Table VI

GROUPS PARTICIPATING AT THE H-DIBCO 2010 COMPETITION

<table>
<thead>
<tr>
<th>S-ID</th>
<th>Group (Corresponding Author(s))</th>
</tr>
</thead>
<tbody>
<tr>
<td>[4]</td>
<td>National University of Singapore &amp; Institute for Infocomm Research, Singapore (B. Su, S. Lu, C.L. Tan)</td>
</tr>
<tr>
<td>[5]</td>
<td>Computer Science department, Ben-Gurion University, Israel (I. Bar-Yosef, K. Kedem, I. Dinstein)</td>
</tr>
<tr>
<td>[6]</td>
<td>LSIS, UMR CNRS 6168 South University of Toulon-Var, France (Thibault Lelore and F. Bouchara)</td>
</tr>
<tr>
<td>[8]</td>
<td>EPITA Research and Development Laboratory (LRDE), Le Kremlin-Bicetre, France (T. Geraud, G. Lazzara)</td>
</tr>
<tr>
<td>[9]</td>
<td>Synchromedia Laboratory, École de technologie supérieure, Montréal, Québec, Canada (R.F. Moghaddam and M. Cheriet)</td>
</tr>
<tr>
<td>[10]</td>
<td>EPITA Research and Development Laboratory (LRDE), Le Kremlin Bicetre cedex France &amp; MINES ParisTech, CMM-Centre de morphologie mathématique, Mathématiques et Systèmes, Fontainebleau, France (J. Fabrizio, B. Marcotegui)</td>
</tr>
<tr>
<td>[12]</td>
<td>Institute of Space Technology, Pakistan (K. Khurshid)</td>
</tr>
<tr>
<td>[14]</td>
<td>CMM, Mines ParisTech, France (J. Hernandez)</td>
</tr>
<tr>
<td>[16]</td>
<td>NifiSoft, Saint-Etienne, France (A. Hassaine)</td>
</tr>
<tr>
<td>[18]</td>
<td>Brigham Young University, USA (O. Nina)</td>
</tr>
<tr>
<td>[20]</td>
<td>Synchromedia Laboratory, École de technologie supérieure, Montréal, Québec, Canada (D. Rivest-Henault, R.F. Moghadam and M. Cheriet)</td>
</tr>
<tr>
<td>[21]</td>
<td>SMCC, Jadavpur University, Kolkata, India (A.F. Mollah)</td>
</tr>
<tr>
<td>[24]</td>
<td>Smith College, MA, USA (N.R. Howe)</td>
</tr>
<tr>
<td>[30]</td>
<td>Jean Monnet University St. Etienne, France (S. Karaoglu)</td>
</tr>
<tr>
<td>[31]</td>
<td>Technological Educational Institute (TEI) of Athens, Greece (A. Nikolau)</td>
</tr>
</tbody>
</table>

REFERENCES


