

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 for systems recognizing Arabic handwritten words. Whereas these competitions were used on a fixed lexicon with 973 Tunisian town names, this ICFHR competition uses as a next step the same background of the IfN/ENIT-Database but now with new collected data of freely written text of about five lines. The fourth competition takes place at the **12th International Conference on Frontiers in Handwriting Recognition (ICFHR2010), November 16-18, 2010, Kolkata, India.**

Evaluation Process:

The objective is to run each Arabic handwritten word recognizer, trained on the IfN/ENIT-Database, version 2.0 and on a set of the new collected data. 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 following tasks are part of evaluation:

- The recognition results on word level of each system are compared on the basis of correct recognised words / respective there dedicated ZIP(Post)-Codes. A dictionary can be used and should include all 937 different Tunisian town/village names.
- The systems are tested on the new data, which includes segmentation and recognition of words from IfN/ENIT-lexicon and words out of this lexicon.

Running a Recognizer:

We run your recognizer (called myrec) by invoking it from the command line as follows:
"myrec dataset.txt output.txt",

- dataset.txt: The dataset is just a list of relative paths to each binary *.tif or *.bmp image to be recognized.
- output.txt: The output file should have one line for each input image. Each line should show the name of the image file that was recognized, followed by the responses (corresponding reference codes) for that image. Each response is given as a pair of values: the text, followed by the confidence. The following example shows that for image word/1/1.tif the recognizer has produced three word hypotheses: Code 1000, 2000 and 3000, with confidences of 1.0, 0.8 and 0.4 respectively.

word/1/1.tif 1000 1.0 2000 0.8 3000 0.4

Important Dates:

Deadline for submission of systems:

May 01, 2010

Volker Märgner and **Haikal El Abed** (v.maergner@tu-bs.de and elabed@tu-bs.de)

Institute for Communications Technology (IfN)
Braunschweig Technical University



www.ifn.ing.tu-bs.de