ICDAR 2011 Robust Reading Competition
Challenge 1: “Reading Text in Born-Digital Images (Web and Email)”

http://www.cvc.uab.es/icdar2011competition/

Dimosthenis Karatzas, Sergi Robles Mestre, Joan Mas, Farshad Nourbakhsh, Partha Pratim Roy
ICDAR 2011 Robust Reading Competition

Two Challenges:

• Challenge 1: “Reading Text in Born-Digital Images (Web and Email)”
• Challenge 2: “Reading Text in Scene Images”

Born Digital Images

• Low-Resolution
• Digitally Created Text
• Compression
• Anti-Aliasing

Real Scene Images

• High-Resolution
• Captured Text
• Illumination Artefacts
• Perspective
The Challenge

Images are important **information carriers**. In electronic documents (e-mails, Web pages) images are used to carry **semantically important text**: Headers, Logos, Titles, Captions, ...

A great part of this information is **nowhere else** other than in the images!
The Challenge
Objectives (I)

Create a representative dataset of born-digital images

- Easy to collect, but what is representative?
- Two main types of born-digital documents: Web and Email
- Only English language for the time being

<table>
<thead>
<tr>
<th>HTML Documents Analysed</th>
<th>412</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web Pages</td>
<td>315</td>
</tr>
<tr>
<td>Email (spam)</td>
<td>22</td>
</tr>
<tr>
<td>Email (ham)</td>
<td>75</td>
</tr>
</tbody>
</table>

Top 3 Internet Languages

- Internet and Informatics: 43.2%
- Commerce: 29.4%
- Science and Technology: 18.9%
- News: 8.5%
- Other languages: 4.0%

Source: Internet World Stats - www.internetworldstats.com/languages.htm
Based on 1,463,632,361 estimated Internet users for 2Q 2008
Copyright © 2008, Miniwatts Marketing Group
Objectives (II)

Create new tools for ground truthing

- From born-digital document management to GT management
- From text localisation and transcription to pixel level segmentation
Objectives (III)

Offer a qualitative performance evaluation framework

- Final OCR results are not informative
- But the pipeline is not strictly defined
- A number of independent tasks were therefore planned
- Qualitative performance evaluation for each task

Series of steps and intermediate goals towards text extraction ...

Colour Segmentation

Text Extraction Characters

Text Extraction Words / Lines

Text Localisation
Structure

The challenge was organised over 3 tasks:

Task 1 – **Text Localization**

**Objective**: To obtain a rough estimation of text areas in the image, in terms of bounding boxes corresponding to parts of text (words or text lines)

Task 2 – **Text Segmentation**

**Objective**: Pixel-level separation of text from the background

Task 3 – **Word Recognition**

**Objective**: Assuming known word bounding boxes, to obtain the correct text transcriptions
Datasets and Ground Truth

**New Datasets** created for this challenge

- Images extracted from different types of HTML documents (Web pages, spam and ham emails)
  - Minimum image size: 100 x 100
- Word images (cut-out) provided separately
  - Minimum word size: 3 characters
- Different ground truth provided for the three tasks
  - Task 1: Bounding box positions of individual words
  - Task 2: Pixel-level classification to text / non-text
  - Task 3: Word images with transcription

Datasets in Numbers

<table>
<thead>
<tr>
<th>HTML Documents Analysed</th>
<th>412</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Dataset (Full Images)</td>
<td>420</td>
</tr>
<tr>
<td>Test Dataset (Full Images)</td>
<td>102</td>
</tr>
<tr>
<td>Training Dataset (Word Images)</td>
<td>3583</td>
</tr>
<tr>
<td>Test Dataset (Word Images)</td>
<td>918</td>
</tr>
</tbody>
</table>
Open mode participation: Participants run their own algorithms and provided results.
Participating Methods

Participation in Numbers

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Visits of the Web Site</td>
<td>692</td>
</tr>
<tr>
<td>Registered Users</td>
<td>37</td>
</tr>
<tr>
<td>Submissions (Task 1)</td>
<td>6</td>
</tr>
<tr>
<td>Submissions (Task 2)</td>
<td>3</td>
</tr>
<tr>
<td>Submissions (Task 3)</td>
<td>1</td>
</tr>
</tbody>
</table>

Submissions Received

<table>
<thead>
<tr>
<th>Submitted Methods</th>
<th>Task 1: Text Localization</th>
<th>Task 2: Text Segmentation</th>
<th>Task 3: Word Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>TH-TextLoc / TH-OCR (China)</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>TDM_IACAS (China)</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>OTCYMIST (India)</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>SASA (USA)</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>TextHunter (Pakistan/France)</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Textorter (Pakistan)</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Results: Task 1 – Text Localization

Performance Evaluation Methodology

- Methodology proposed by Wolf and Jolion [1]
- Takes into account both
  - Bounding Box area overlapping and
  - Precision at the level of detection counts
- Possible to create meaningful cumulative results over many images
- Ways to deal with one-to-many and many-to-one cases
- Set up to penalise over-segmentation (words split to parts), but no under-segmentation (group of words detected as text line)

Results: Task 1 – Text Localization

Baseline Method

• We used a commercial OCR package (ABBYY OCR SDK v.10) to obtain text localization and word recognition (see task 3) results
• Factory default parameters used, enabling the option for low-resolution images
## Results: Task 1 – Text Localization

### Text Localization Results (%)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Submitted Methods – Task 1</th>
<th>Recall</th>
<th>Precision</th>
<th>Harmonic Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Textorter</td>
<td>69.62</td>
<td>85.83</td>
<td>76.88</td>
</tr>
<tr>
<td>2</td>
<td>TH-TextLoc</td>
<td>73.08</td>
<td>80.51</td>
<td>76.62</td>
</tr>
<tr>
<td>3</td>
<td>TDM_IACAS</td>
<td>69.16</td>
<td>84.64</td>
<td>76.12</td>
</tr>
<tr>
<td>4</td>
<td>OTCYMIST</td>
<td>75.91</td>
<td>64.05</td>
<td>69.48</td>
</tr>
<tr>
<td>5</td>
<td>SASA</td>
<td>65.62</td>
<td>67.82</td>
<td>66.70</td>
</tr>
<tr>
<td>6</td>
<td>Text Hunter</td>
<td>57.76</td>
<td>75.52</td>
<td>65.46</td>
</tr>
<tr>
<td>N/A</td>
<td>Baseline Method</td>
<td>70.32</td>
<td>84.25</td>
<td>76.66</td>
</tr>
</tbody>
</table>
Results: Task 1 – Text Localization

Text Localization Results (Precision / Recall)

The diagram shows a scatter plot comparing different text localization methods across precision and recall. The methods compared include Textorter, TDM_IACAS, Baseline, TH-TextLoc, Text Hunter, SASA, OTCYMIST. Each method is represented by a different symbol on the plot.
Results: Task 1 – Text Localization

Observations

• Commercial system works at par with submitted methods
• Detailed analysis shows qualitative difference: baseline system more efficient in rejecting images it could not treat
Results: Task 2 – Text Segmentation

Performance Evaluation Methodology

- Primary scheme used is the framework proposed by Clavelli et al [2]
- If measures the degree to which morphological properties of the text are preserved, not simply the number of misclassified pixels
- As a secondary evaluation scheme we implemented standard pixel level precision and recall (compatibility with other results)

## Results: Task 2 – Text Segmentation

### Text Segmentation Results (%) – Primary Evaluation Scheme

<table>
<thead>
<tr>
<th>Rank</th>
<th>Submitted Methods – Task 2</th>
<th>Well Segmented</th>
<th>Merged</th>
<th>Lost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OTCYMIST</td>
<td>65.96</td>
<td>15.44</td>
<td>18.59</td>
</tr>
<tr>
<td>2</td>
<td>Textorter</td>
<td>58.73</td>
<td>32.53</td>
<td>08.73</td>
</tr>
<tr>
<td>3</td>
<td>SASA</td>
<td>42.71</td>
<td>10.70</td>
<td>46.57</td>
</tr>
</tbody>
</table>

### Text Segmentation Results (%) – Secondary Evaluation Scheme

<table>
<thead>
<tr>
<th>Rank</th>
<th>Submitted Methods – Task 2</th>
<th>Precision</th>
<th>Recall</th>
<th>Harmonic Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OTCYMIST</td>
<td>80.99</td>
<td>71.13</td>
<td>75.74</td>
</tr>
<tr>
<td>2</td>
<td>Textorter</td>
<td>65.20</td>
<td>62.50</td>
<td>63.82</td>
</tr>
<tr>
<td>3</td>
<td>SASA</td>
<td>71.93</td>
<td>54.78</td>
<td>62.19</td>
</tr>
</tbody>
</table>
Results: Task 2 – Text Segmentation

Observations

• Both metrics used agree in the ranking, nevertheless there is qualitative differences pointed out by the primary metric
• The main problem of “Textorter” seems to be that of under-segmentation
Results: Task 3 – Word Recognition

Performance Evaluation Methodology

• Edit distance (normalised to the length of the ground truth transcription)
• Equal weights for deletions, additions, substitutions
• Also report statistics on correctly recognised words

Baseline method

• Used ABBYY OCR (v.10) on word images
• No pre-processing, option for “low resolution” images turned on
Results: Task 3 – Word Recognition

Word Recognition Results

<table>
<thead>
<tr>
<th>Rank</th>
<th>Submitted Methods – Task 3</th>
<th>Total Edit Distance</th>
<th>Correctly Recognized Words (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>TH_OCR</td>
<td>189.9</td>
<td>61.54</td>
</tr>
<tr>
<td>N/A</td>
<td>Baseline Method</td>
<td>232.8</td>
<td>63.40</td>
</tr>
</tbody>
</table>

Observations

- Edit distance peak of baseline at 1 (all characters changed, indicative of an empty response)
- Baseline seems to have a good rejection criterion
- 48.8% words recognized by both
What’s Next

The competition is going on – entering a **continuous mode**

**Datasets** are freely available (on Competition Web site and soon on TC11)

**Performance evaluation** available online – automatic and real-time!

Watch this space – **more resources coming**:

- Full Ground Truth (XML)
- New datasets (including real scenes and colour documents)
- Online GT management framework
- Ground truthing tools

http://www.cvc.uab.es/icdar2011competition/
ICDAR 2011 Robust Reading Competition

**Challenge 1:** "Reading Text in Born-Digital Images (Web and Email)"

### Submit results

**Task 1: Text Localisation**

<table>
<thead>
<tr>
<th>Submit</th>
<th>Participant</th>
<th>folder</th>
<th>method</th>
<th>Validated</th>
<th>True Pos</th>
<th>num Det</th>
<th>Recall</th>
<th>Precision</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>173</td>
<td>01-Jun-11 17:26</td>
<td>21 Deepak Kumar</td>
<td>icdar_2011 POTOCRST</td>
<td>OK (102)</td>
<td>1015</td>
<td>1012</td>
<td>75.91%</td>
<td>68.85%</td>
<td>89.39%</td>
</tr>
<tr>
<td>160</td>
<td>01-Jun-11 14:07</td>
<td>30 Mohamed Seddik Rifi</td>
<td>icdar_2011 Text Hunter</td>
<td>OK (102)</td>
<td>1015</td>
<td>1012</td>
<td>73.75%</td>
<td>71.52%</td>
<td>65.46%</td>
</tr>
<tr>
<td>165</td>
<td>01-Jun-11 15:59</td>
<td>18 Zhangheng</td>
<td>icdar_2011 OpenOCR</td>
<td>OK (102)</td>
<td>1015</td>
<td>1012</td>
<td>73.34%</td>
<td>67.92%</td>
<td>70.94%</td>
</tr>
<tr>
<td>161</td>
<td>05-Jun-11 10:34</td>
<td>14 Chen Fang</td>
<td>icdar_2011 TiF-Textlet</td>
<td>OK (102)</td>
<td>1015</td>
<td>1012</td>
<td>72.84%</td>
<td>82.39%</td>
<td>70.94%</td>
</tr>
<tr>
<td>157</td>
<td>05-Jun-11 11:04</td>
<td>22 gao</td>
<td>icdar_2011 Tentator</td>
<td>OK (102)</td>
<td>1015</td>
<td>1012</td>
<td>73.21%</td>
<td>67.52%</td>
<td>70.94%</td>
</tr>
<tr>
<td>131</td>
<td>05-Jun-11 18:35</td>
<td>29 qyi</td>
<td>icdar_2011 Tentator</td>
<td>OK (102)</td>
<td>1015</td>
<td>1012</td>
<td>73.21%</td>
<td>68.52%</td>
<td>72.42%</td>
</tr>
<tr>
<td>115</td>
<td>05-Jun-11 16:36</td>
<td>39 Joem</td>
<td>icdar_2011 Baseline</td>
<td>OK (102)</td>
<td>1015</td>
<td>1012</td>
<td>73.21%</td>
<td>68.52%</td>
<td>72.42%</td>
</tr>
<tr>
<td>115</td>
<td>05-Jun-11 09:08</td>
<td>17 Samou Shao</td>
<td>icdar_2011 Tentator</td>
<td>OK (102)</td>
<td>1015</td>
<td>1012</td>
<td>73.21%</td>
<td>68.52%</td>
<td>72.42%</td>
</tr>
</tbody>
</table>

**Task 2: Text Segmentation**

<table>
<thead>
<tr>
<th>Submit</th>
<th>Participant</th>
<th>folder</th>
<th>method</th>
<th>Validated</th>
<th>Wells</th>
<th>Merged</th>
<th>Lost</th>
<th>Recall</th>
<th>Precision</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>175</td>
<td>03-Jun-11 19:25</td>
<td>21 Deepak Kumar</td>
<td>icdar_2011 POTOCRST</td>
<td>OK (102)</td>
<td>65.34</td>
<td>15.44</td>
<td>23.22</td>
<td>86.89</td>
<td>71.21</td>
<td>75.34</td>
</tr>
<tr>
<td>159</td>
<td>02-Jun-11 20:05</td>
<td>32 gao</td>
<td>icdar_2011 Tentator</td>
<td>OK (102)</td>
<td>65.34</td>
<td>15.44</td>
<td>23.22</td>
<td>86.89</td>
<td>71.21</td>
<td>75.34</td>
</tr>
<tr>
<td>135</td>
<td>05-Jun-11 19:29</td>
<td>25 qyi</td>
<td>icdar_2011 Tentator</td>
<td>OK (102)</td>
<td>65.34</td>
<td>15.44</td>
<td>23.22</td>
<td>86.89</td>
<td>71.21</td>
<td>75.34</td>
</tr>
<tr>
<td>117</td>
<td>02-Jun-11 17:54</td>
<td>49 Joem</td>
<td>icdar_2011 Baseline</td>
<td>OK</td>
<td>65.34</td>
<td>15.44</td>
<td>23.22</td>
<td>86.89</td>
<td>71.21</td>
<td>75.34</td>
</tr>
</tbody>
</table>

**Task 3: Word Recognition**

<table>
<thead>
<tr>
<th>Submit</th>
<th>Participant</th>
<th>folder</th>
<th>method</th>
<th>Validated</th>
<th>Total Word</th>
<th>Correctly Recognised Words</th>
<th>details</th>
</tr>
</thead>
<tbody>
<tr>
<td>148</td>
<td>06-Jun-11 03:26</td>
<td>24 Chenfang</td>
<td>icdar_2011</td>
<td>OK (102)</td>
<td>20.73</td>
<td>61.85</td>
<td>details</td>
</tr>
<tr>
<td>125</td>
<td>06-Jun-11 10:20</td>
<td>39 Joem</td>
<td>icdar_2011 Baseline</td>
<td>OK (102)</td>
<td>20.73</td>
<td>61.85</td>
<td>details</td>
</tr>
</tbody>
</table>

*"HuPerText" Project (TIN2009-34998)*
### What's Next

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Clubcard</td>
<td>Amazon</td>
<td>Reduced mortgage rates</td>
<td>Lenders</td>
<td>Current accounts</td>
<td>Free UK MMA</td>
<td>&quot;Play60 With Me!&quot;</td>
<td>Standard Chartered</td>
<td>&quot;QuickLock&quot; Project</td>
<td>&quot;WePac&quot; Project</td>
<td>&quot;Reach Us and Get Your Free&quot;</td>
<td>&quot;The Month of Zupa&quot;</td>
<td>Researcher Voices</td>
<td>&quot;Why do you тысячи&quot;</td>
<td>&quot;It's easier to have a professional site...&quot;</td>
<td>&quot;Contact us for more information&quot;</td>
<td>&quot;16-20 June: Test period&quot;</td>
<td>&quot;16 June: Submission of results&quot;</td>
<td>&quot;16 June: Test dataset available&quot;</td>
<td>&quot;16 June: Test period&quot;</td>
<td>&quot;16 June: Submission of results&quot;</td>
<td>&quot;16 June: Test dataset available&quot;</td>
<td>&quot;16 June: Test period&quot;</td>
<td>&quot;16 June: Submission of results&quot;</td>
<td></td>
</tr>
</tbody>
</table>

**Performance Metrics**

- **Image 1:**
  - well segment: 88.35%
  - recall: 86.65%
  - precision: 90.62%

- **Image 2:**
  - well segment: 84.54%
  - recall: 86.65%
  - precision: 78.87%

- **Image 3:**
  - well segment: 84.54%
  - recall: 86.65%
  - precision: 78.87%

- **Image 4:**
  - well segment: 84.54%
  - recall: 86.65%
  - precision: 78.87%

- **Image 5:**
  - well segment: 84.54%
  - recall: 86.65%
  - precision: 78.87%
What’s Next

ICDAR 2011 Robust Reading Competition
Challenge 2: "Reading Text in Born-Digital Images (Web and Email)"

Results Summary

Task 2 - Text Segmentation
Method: ICDAR 2011 Texttoter

Results:
- Well segmented: 0.38 % (mean: 16.41 %)

Original Image

Detected Image

---

Important Dates

- 05/03/2011: Text Sets Now Available!
- 05/03/2011: Submissions Is Now Open
- 06/04/2011: Submissions Guidelines Updated
- 05/03/2011: Training Datasets Now Available!

"Kuery'sWeb" Project (TIN2009-04998)
Thanks for participating!

Task 1 – Text Localization Winner: “Textorter”, S. Tehsin and A. Masood, Military College of Signals, National University of Science and Technology, Pakistan

Task 2 – Text Segmentation Winner: “OTCYMIST”, D. Kumar and A.G. Ramakrishnan, Medical Intelligence and Language Engineering Laboratory, Indian Institute of Science, Bangalore, India

Task 3 – Word Recognition Mention: “TH_OCR”, C. Yang, C. Liu and X. Ding, Department of Electronic Engineering, Tsinghua University Beijing, China

The competition site is open!
Register to download the datasets and upload new results

http://www.cvc.uab.es/icdar2011competition/