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

Centre de Visió per Computador

MINISTERIO DE CIENCIA E INNOVACIÓN

UAB Universitat Autònoma de Barcelona

Generalitat de Catalunya

Parc de Recerca UAB

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

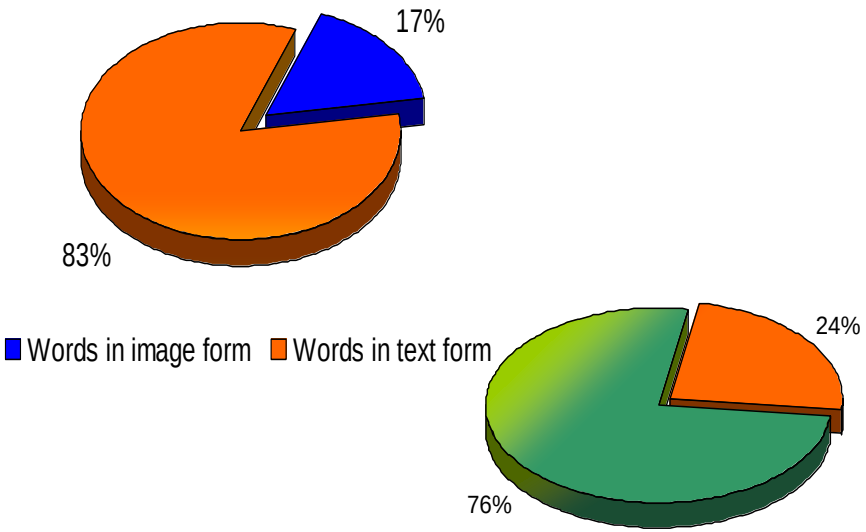
- Low-Resolution
- Digitally Created Text
- Compression
- Anti-Aliasing

Real Scene Images

- High-Resolution
- Captured Text
- Illumination Artefacts
- Perspective

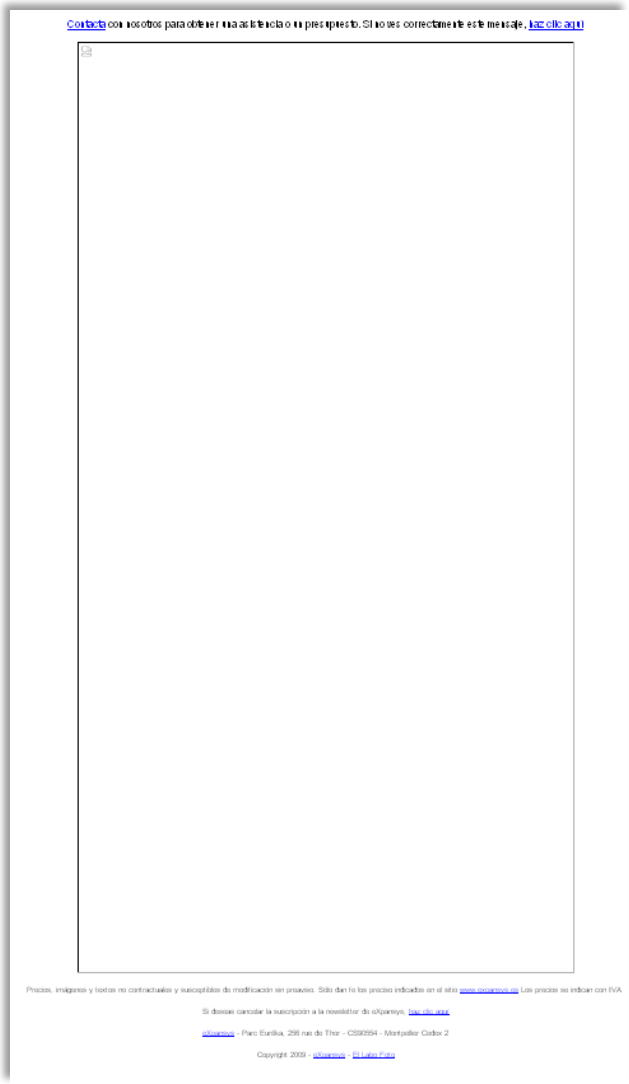
The Challenge

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

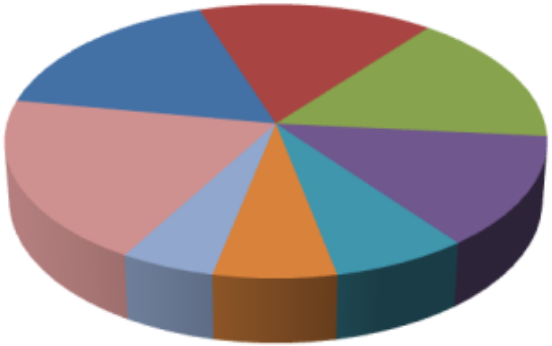


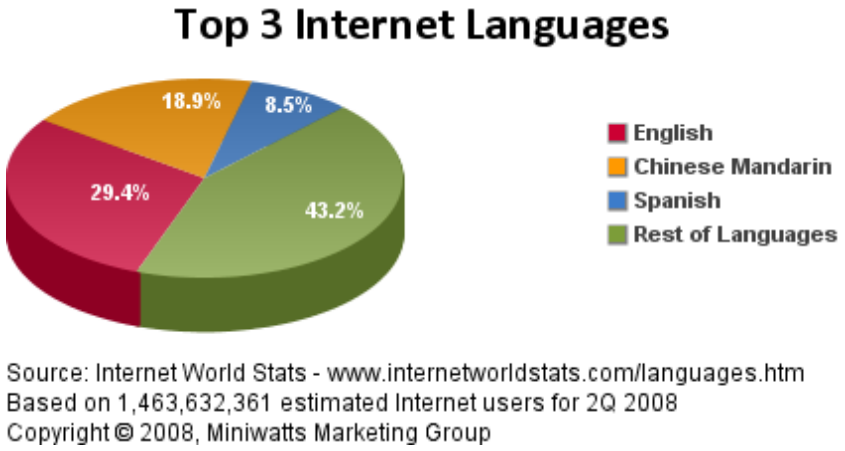
Objectives (I)

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

HTML Documents Analysed	412
Web Pages	315
Email (spam)	22
Email (ham)	75

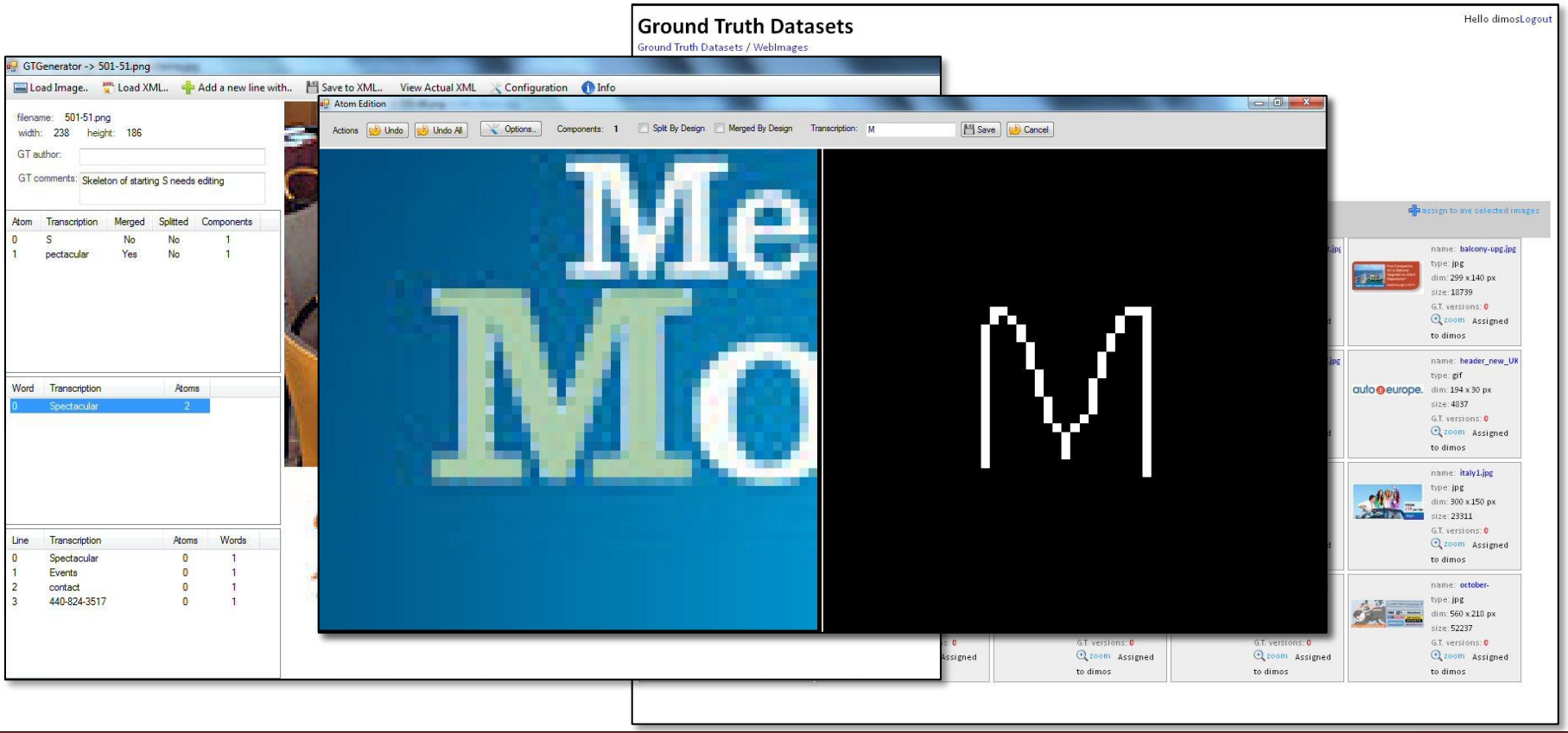




Objectives (II)

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

HTML Documents Analysed	412
Training Dataset (Full Images)	420
Test Dataset (Full Images)	102
Training Dataset (Word Images)	3583
Test Dataset (Word Images)	918

Timeline

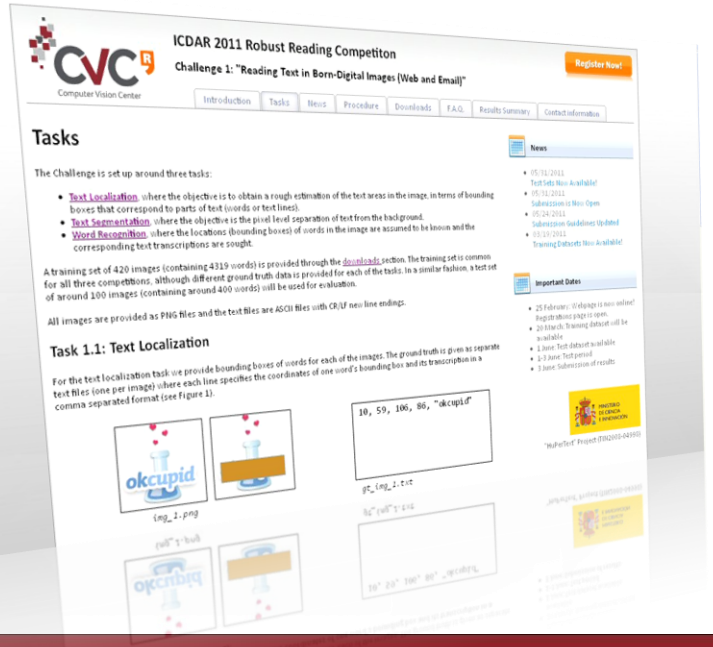
Web Site Online
Registrations Open
25/2

Training Set
Online
20/3

Test Set
Available
1/6

Results Submission
Deadline
3/6

Feb | Mar | Apr | May | Jun



Open mode participation:
Participants run their own algorithms and provided results.

Participating Methods

Participation in Numbers

Visits of the Web Site	692
Registered Users	37
Submissions (Task 1)	6
Submissions (Task 2)	3
Submissions (Task 3)	1

Submissions Received

Submitted Methods		Task 1: Text Localization	Task 2: Text Segmentation	Task 3: Word Recognition
TH-TextLoc / TH-OCR	(China)	X		X
TDM_IACAS	(China)	X		
OTCYMIST	(India)	X	X	
SASA	(USA)	X	X	
TextHunter	(Pakistan/France)	X		
Textorter	(Pakistan)	X	X	

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)

1. C. Wolf and J.M. Jolion, “Object Count / Area Graphs for the Evaluation of Object Detection and Segmentation Algorithms”, IJDAR 8(4), pp. 280-296, 2006

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

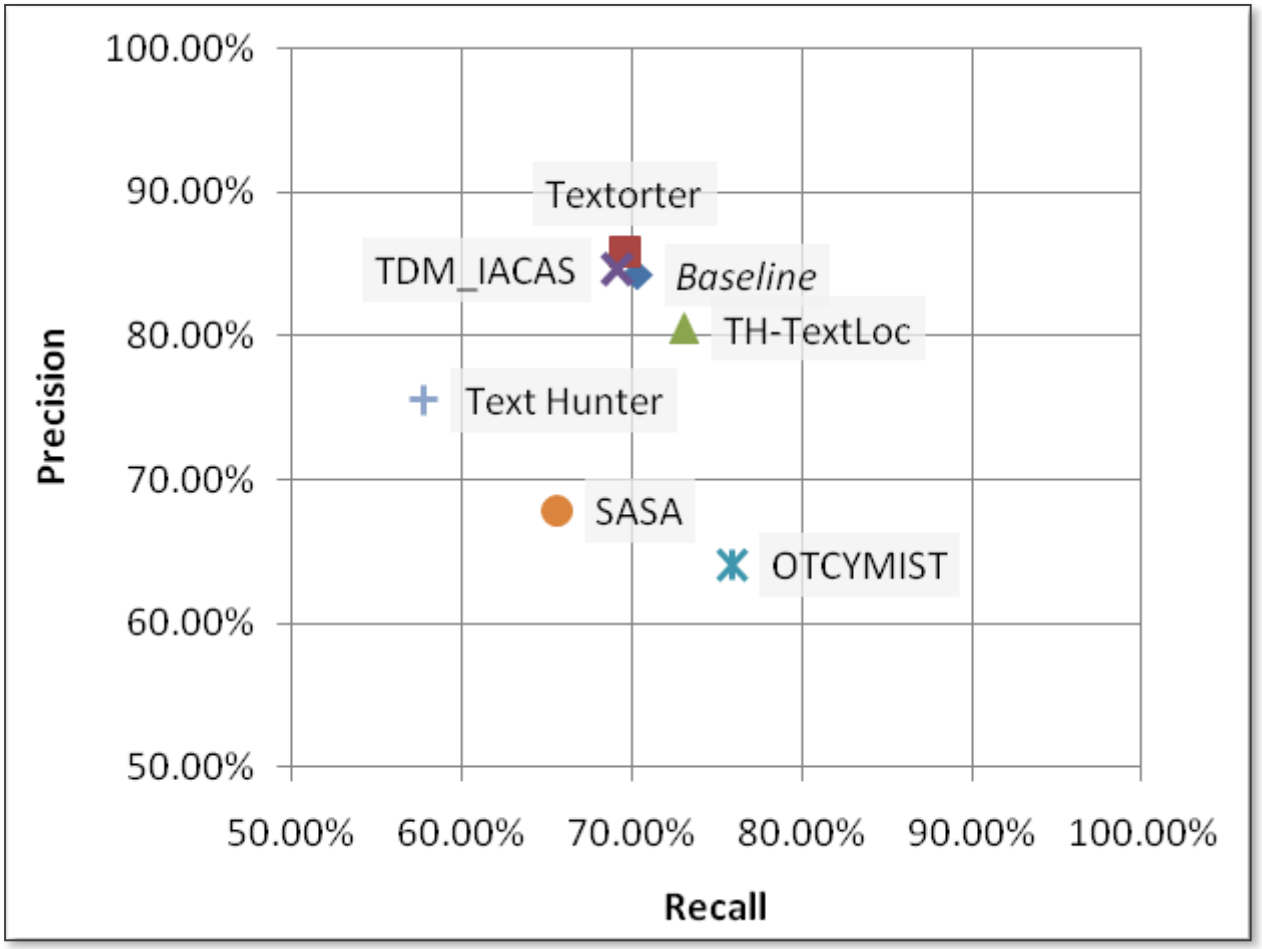
Results: Task 1 – Text Localization

Text Localization Results (%)

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

Results: Task 1 – Text Localization

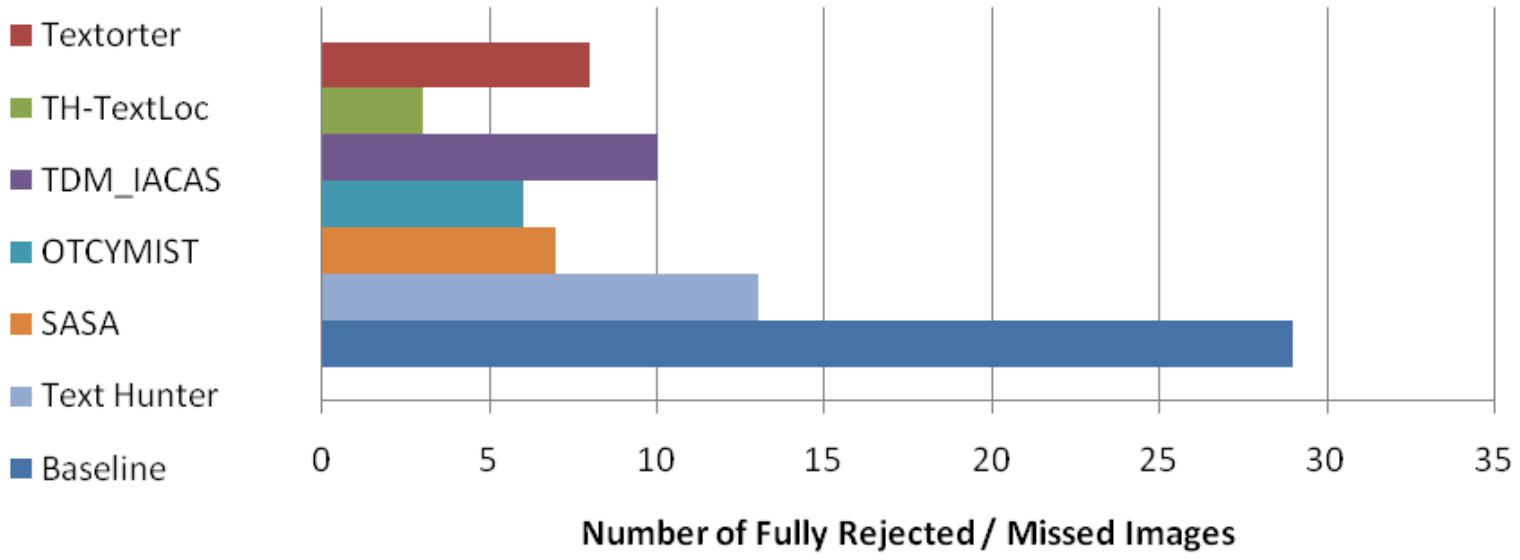
Text Localization Results (Precision / Recall)



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)

(a) (b) (c)

2. A. Clavelli, D. Karatzas and J. Lladós “A Framework for the Assessment of Text Extraction Algorithms on Complex Colour Images”, DAS 2010, pp. 19-28, 2010

Results: Task 2 – Text Segmentation

Text Segmentation Results (%) – Primary Evaluation Scheme

Rank	Submitted Methods – Task 2	Well Segmented	Merged	Lost
1	OTCYMIST	**65.96**	15.44	18.59
2	Textorter	**58.73**	32.53	08.73
3	SASA	**42.71**	10.70	46.57

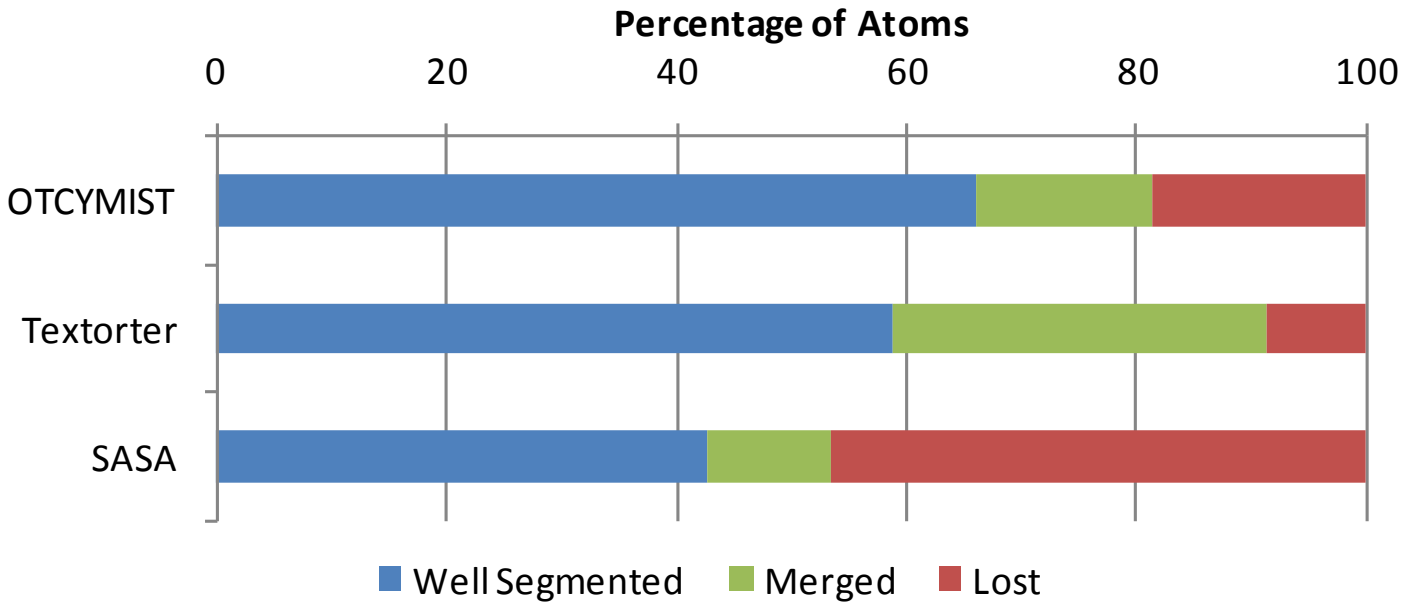
Text Segmentation Results (%) – Secondary Evaluation Scheme

Rank	Submitted Methods – Task 2	Precision	Recall	Harmonic Mean
1	OTCYMIST	80.99	71.13	**75.74**
2	Textorter	65.20	62.50	**63.82**
3	SASA	71.93	54.78	**62.19**

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

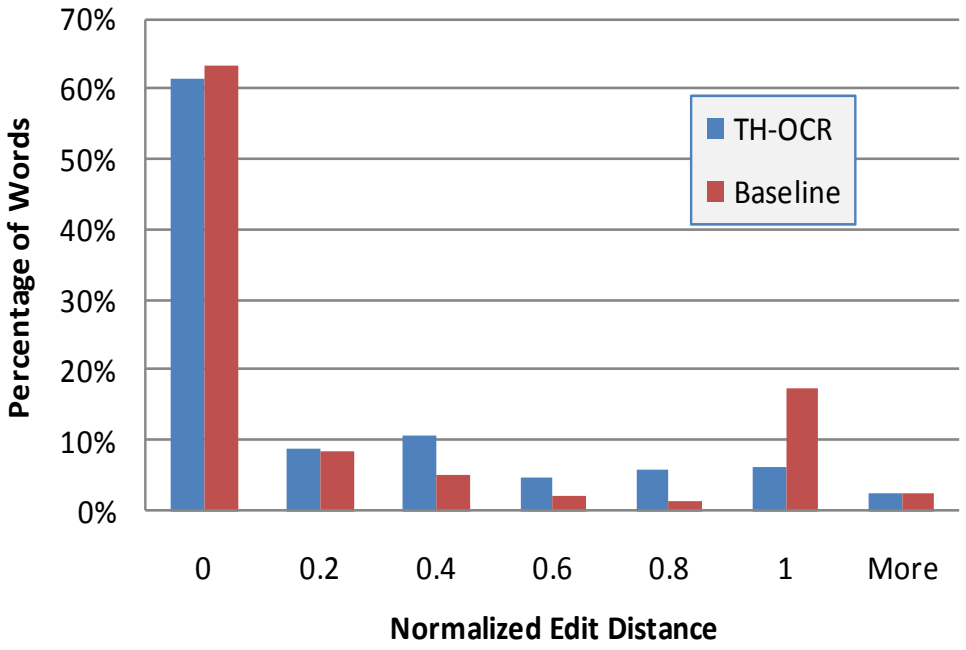
Results: Task 3 – Word Recognition

Word Recognition Results

Rank	Submitted Methods – Task 3	Total Edit Distance	Correctly Recognized Words (%)
N/A	TH_OCR	**189.9**	61.54
N/A	*Baseline Method*	***232.8***	*63.40*

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

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

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

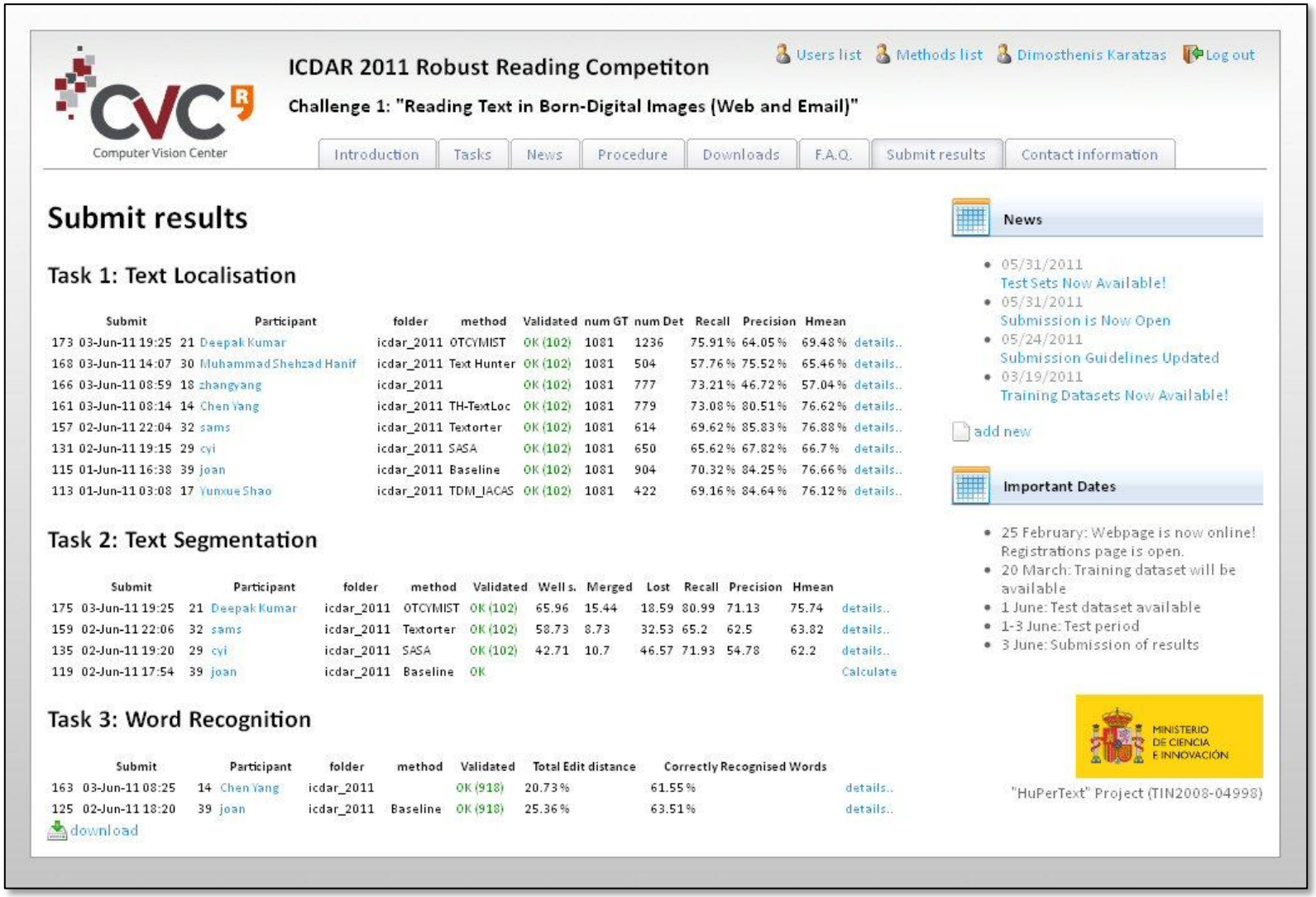
Datasets are freely available (on Competition Web site and soon on TC11)

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

What's Next



Computer Vision Center

ICDAR 2011 Robust Reading Competiton

Users list | Methods list | Dimosthenis Karatzas | Log out

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

Introduction | Tasks | News | Procedure | Downloads | F.A.Q. | Submit results | Contact information

Submit results

Task 1: Text Localisation

Submit	Participant	folder	method	Validated	num GT	num Det	Recall	Precision	Hmean	
173 03-Jun-11 19:25	21 Deepak Kumar	icdar_2011	OTCYMIST	OK (102)	1081	1236	75.91%	64.05%	69.48%	details..
168 03-Jun-11 14:07	30 Muhammad Shehzad Hanif	icdar_2011	Text Hunter	OK (102)	1081	504	57.76%	75.52%	65.46%	details..
166 03-Jun-11 08:59	18 zhangyang	icdar_2011		OK (102)	1081	777	73.21%	46.72%	57.04%	details..
161 03-Jun-11 08:14	14 Chen Yang	icdar_2011	TH-TextLoc	OK (102)	1081	779	73.08%	80.51%	76.62%	details..
157 02-Jun-11 22:04	32 sams	icdar_2011	Textorter	OK (102)	1081	614	69.62%	85.83%	76.88%	details..
131 02-Jun-11 19:15	29 cyi	icdar_2011	SASA	OK (102)	1081	650	65.62%	67.82%	66.7%	details..
115 01-Jun-11 16:38	39 joan	icdar_2011	Baseline	OK (102)	1081	904	70.32%	84.25%	76.66%	details..
113 01-Jun-11 03:08	17 Yunxue Shao	icdar_2011	TDM_IACAS	OK (102)	1081	422	69.16%	84.64%	76.12%	details..

Task 2: Text Segmentation

Submit	Participant	folder	method	Validated	Wells	Merged	Lost	Recall	Precision	Hmean	
175 03-Jun-11 19:25	21 Deepak Kumar	icdar_2011	OTCYMIST	OK (102)	65.96	15.44	18.59	80.99	71.13	75.74	details..
159 02-Jun-11 22:06	32 sams	icdar_2011	Textorter	OK (102)	58.73	8.73	32.53	65.2	62.5	63.82	details..
135 02-Jun-11 19:20	29 cyi	icdar_2011	SASA	OK (102)	42.71	10.7	46.57	71.93	54.78	62.2	details..
119 02-Jun-11 17:54	39 joan	icdar_2011	Baseline	OK							Calculate

Task 3: Word Recognition

Submit	Participant	folder	method	Validated	Total Edit distance	Correctly Recognised Words	
163 03-Jun-11 08:25	14 Chen Yang	icdar_2011		OK (918)	20.73%	61.55%	details..
125 02-Jun-11 18:20	39 joan	icdar_2011	Baseline	OK (918)	25.36%	63.51%	details..

download

News

- 05/31/2011 Test Sets Now Available!
- 05/31/2011 Submission is Now Open
- 05/24/2011 Submission Guidelines Updated
- 03/19/2011 Training Datasets Now Available!

add new

Important Dates

- 25 February: Webpage is now online! Registrations page is open.
- 20 March: Training dataset will be available
- 1 June: Test dataset available
- 1-3 June: Test period
- 3 June: Submission of results

MINISTERIO DE CIENCIA E INNOVACIÓN

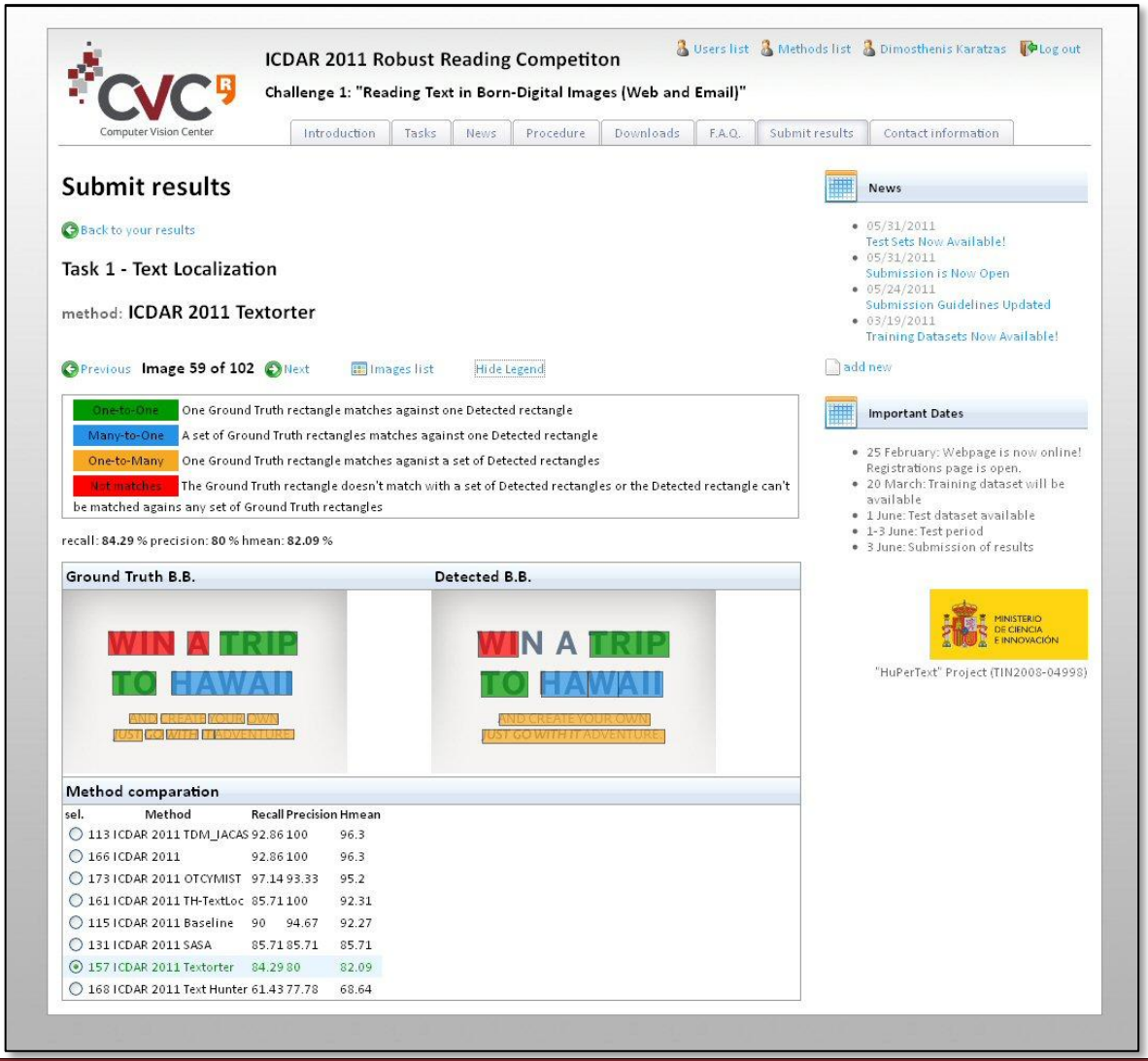
"HuPerText" Project (TIN2008-04998)

Centre de Visió per Computador

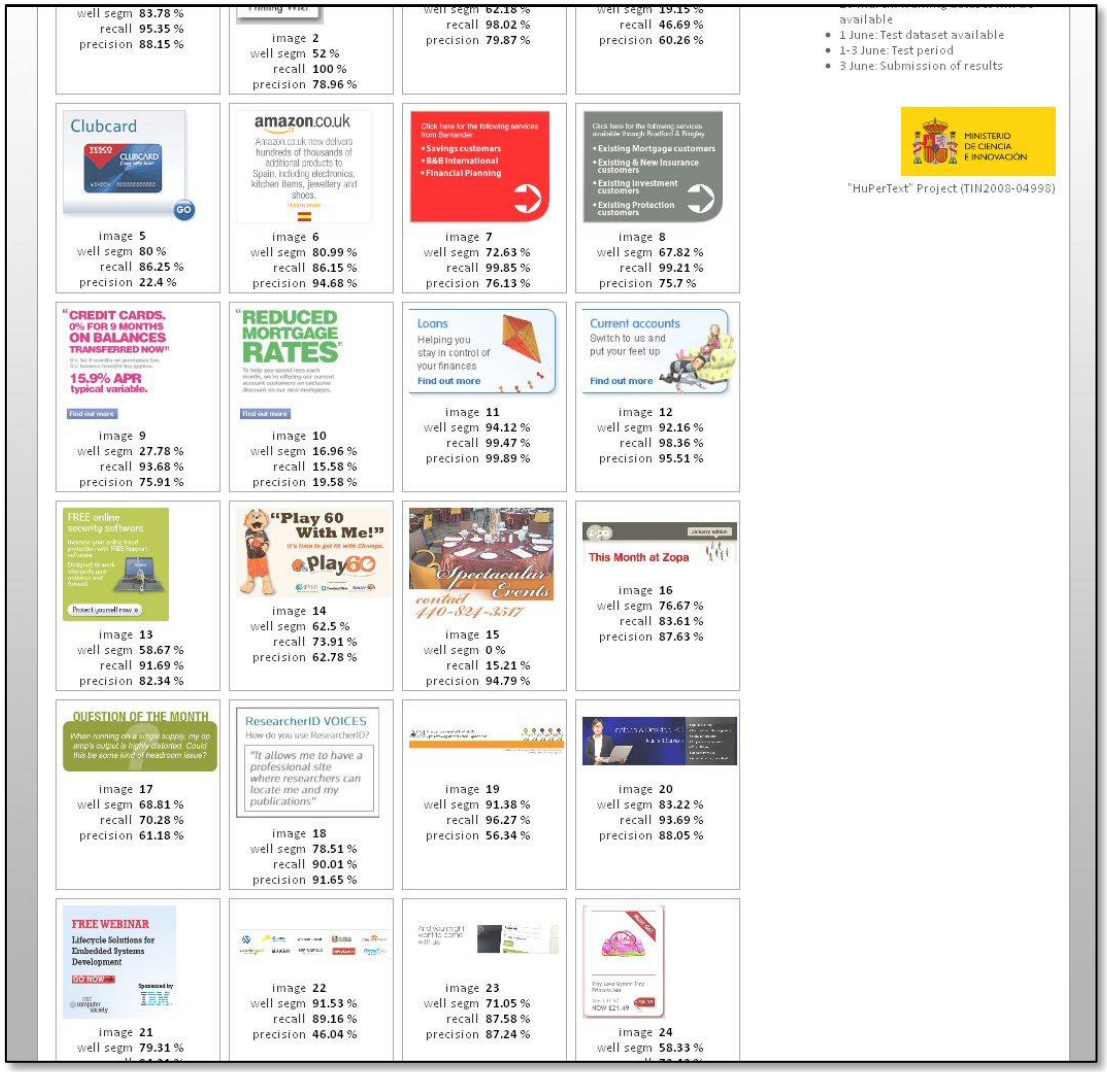
What's Next


image 5
recall 33.33 %
precision 50 %
hmean 40 %
image 6
recall 76.19 %
precision 100 %
hmean 86.49 %
image 7
recall 85.71 %
precision 100 %
hmean 92.31 %
image 8
recall 100 %
precision 100 %
hmean 100 %
image 9
recall 58.62 %
precision 88.89 %
hmean 70.65 %
image 10
recall 69.23 %
precision 85.71 %
hmean 76.6 %
image 11
recall 83.33 %
precision 100 %
hmean 90.91 %
image 12
recall 92.31 %
precision 100 %
hmean 96 %
image 13
recall 80 %
precision 100 %
hmean 88.89 %
image 14
recall 66.67 %
precision 90 %
hmean 76.6 %
image 15
recall 20 %
precision 80 %
hmean 32 %
image 16
recall 71.43 %
precision 100 %
hmean 83.33 %
image 17
recall 88 %
precision 100 %
hmean 93.62 %
image 18
recall 86.96 %
precision 100 %
hmean 93.02 %
image 19
recall 75.24 %
precision 98.46 %
hmean 85.3 %
image 20
recall 83.33 %
precision 100 %
hmean 90.91 %
image 22
recall 31.58 %
precision 66.67 %
image 23
recall 86.67 %
precision 100 %

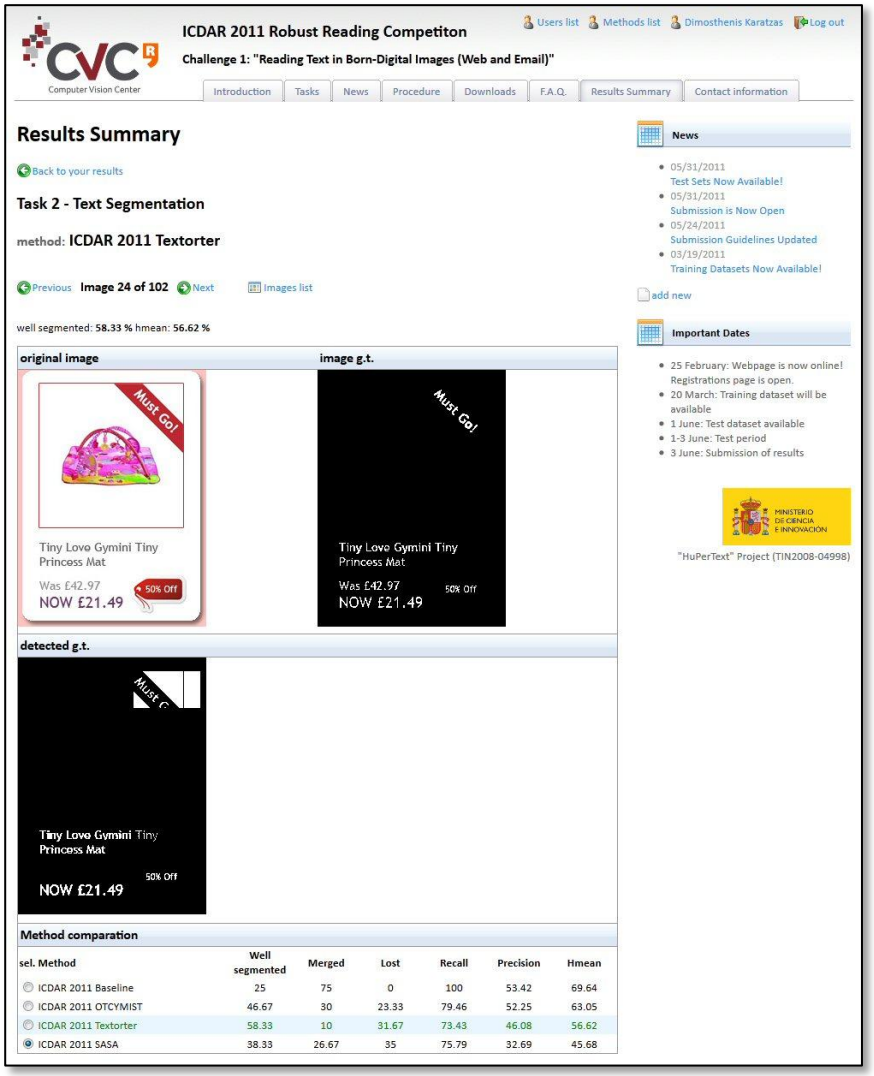
What's Next


ICDAR 2011 Robust Reading Competiton
Challenge 1: "Reading Text in Born-Digital Images (Web and Email)"
Users list
Methods list
Dimosthenis Karatzas
Log out
Computer Vision Center
Introduction
Tasks
News
Procedure
Downloads
F.A.Q.
Submit results
Contact information
Submit results
Back to your results
Task 1 - Text Localization
method: ICDAR 2011 Textorter
Previous
Image 59 of 102
Next
Images list
Hide Legend
One-to-One
One Ground Truth rectangle matches against one Detected rectangle
Many-to-One
A set of Ground Truth rectangles matches against one Detected rectangle
One-to-Many
One Ground Truth rectangle matches aganist a set of Detected rectangles
Not matches
The Ground Truth rectangle doesn't match with a set of Detected rectangles or the Detected rectangle can't be matched agains any set of Ground Truth rectangles
recall: 84.29 % precision: 80 % hmean: 82.09 %
Ground Truth B.B.
Detected B.B.
WIN A TRIP TO HAWAII
AND CREATE YOUR OWN JUST GO WITH IT ADVENTURE!
Method comparation
sel. Method Recall Precision Hmean
113 ICDAR 2011 TDM_IACAS 92.86 100 96.3
166 ICDAR 2011 92.86 100 96.3
173 ICDAR 2011 OTCYMIST 97.14 93.33 95.2
161 ICDAR 2011 TH-TextLoc 85.71 100 92.31
115 ICDAR 2011 Baseline 90 94.67 92.27
131 ICDAR 2011 SASA 85.71 85.71 85.71
157 ICDAR 2011 Textorter 84.29 80 82.09
168 ICDAR 2011 Text Hunter 61.43 77.78 68.64
News
05/31/2011
Test Sets Now Available!
05/31/2011
Submission is Now Open
05/24/2011
Submission Guidelines Updated
03/19/2011
Training Datasets Now Available!
add new
Important Dates
25 February: Webpage is now online! Registrations page is open.
20 March: Training dataset will be available
1 June: Test dataset available
1-3 June: Test period
3 June: Submission of results
MINISTERIO DE CIENCIA E INNOVACIÓN
"HuPerText" Project (TIN2008-04998)

What's Next


well segm 83.78 %
recall 95.35 %
precision 88.15 %
image 2
well segm 52 %
recall 100 %
precision 78.96 %
recall 98.02 %
precision 79.87 %
recall 46.69 %
precision 60.26 %
available
1 June: Test dataset available
1-3 June: Test period
3 June: Submission of results
MINISTERIO DE CIENCIA E INNOVACIÓN
"HuPerText" Project (TIN2008-04998)
Clubcard
image 5
well segm 80 %
recall 86.25 %
precision 22.4 %
amazon.co.uk
image 6
well segm 80.99 %
recall 86.15 %
precision 94.68 %
Savings customers
B&B International
Financial Planning
image 7
well segm 72.63 %
recall 99.85 %
precision 76.13 %
Existing Mortgage customers
Existing & New Insurance customers
Existing Investment customers
Existing Protection customers
image 8
well segm 67.82 %
recall 99.21 %
precision 75.7 %
"CREDIT CARDS. 0% FOR 9 MONTHS ON BALANCES TRANSFERRED NOW"
15.9% APR typical variable.
image 9
well segm 27.78 %
recall 93.68 %
precision 75.91 %
"REDUCED MORTGAGE RATES"
image 10
well segm 16.96 %
recall 15.58 %
precision 19.58 %
Loans
Helping you stay in control of your finances
Find out more
image 11
well segm 94.12 %
recall 99.47 %
precision 99.89 %
Current accounts
Switch to us and put your feet up
Find out more
image 12
well segm 92.16 %
recall 98.36 %
precision 95.51 %
FREE online security software
image 13
well segm 58.67 %
recall 91.69 %
precision 82.34 %
"Play 60 With Me!"
Play60
image 14
well segm 62.5 %
recall 73.91 %
precision 62.78 %
Spectacular Events
contact 440-824-3517
image 15
well segm 0 %
recall 15.21 %
precision 94.79 %
This Month at Zopa
image 16
well segm 76.67 %
recall 83.61 %
precision 87.63 %
QUESTION OF THE MONTH
image 17
well segm 68.81 %
recall 70.28 %
precision 61.18 %
ResearcherID VOICES
How do you use ResearcherID?
"It allows me to have a professional site where researchers can locate me and my publications"
image 18
well segm 78.51 %
recall 90.01 %
precision 91.65 %
image 19
well segm 91.38 %
recall 96.27 %
precision 56.34 %
image 20
well segm 83.22 %
recall 93.69 %
precision 88.05 %
FREE WEBINAR
Lifecycle Solutions for Embedded Systems Development
Sponsored by IBM
image 21
well segm 79.31 %
image 22
well segm 91.53 %
recall 89.16 %
precision 46.04 %
image 23
well segm 71.05 %
recall 87.58 %
precision 87.24 %
image 24
well segm 58.33 %

What's Next


ICDAR 2011 Robust Reading Competiton
Challenge 1: "Reading Text in Born-Digital Images (Web and Email)"
Users list
Methods list
Dimosthenis Karatzas
Log out
Computer Vision Center
Introduction
Tasks
News
Procedure
Downloads
F.A.Q.
Results Summary
Contact information
Results Summary
Back to your results
Task 2 - Text Segmentation
method: ICDAR 2011 Textorter
Previous
Image 24 of 102
Next
Images list
well segmented: 58.33 % hmean: 56.62 %
original image
image g.t.
detected g.t.
Must Go!
Tiny Love Gymini Tiny Princess Mat
Was £42.97
NOW £21.49
50% Off
Method comparation
sel. Method
Well segmented
Merged
Lost
Recall
Precision
Hmean
ICDAR 2011 Baseline 25 75 0 100 53.42 69.64
ICDAR 2011 OTCYMIST 46.67 30 23.33 79.46 52.25 63.05
ICDAR 2011 Textorter 58.33 10 31.67 73.43 46.08 56.62
ICDAR 2011 SASA 38.33 26.67 35 75.79 32.69 45.68
News
05/31/2011
Test Sets Now Available!
05/31/2011
Submission is Now Open
05/24/2011
Submission Guidelines Updated
03/19/2011
Training Datasets Now Available!
add new
Important Dates
25 February: Webpage is now online! Registrations page is open.
20 March: Training dataset will be available
1 June: Test dataset available
1-3 June: Test period
3 June: Submission of results
MINISTERIO DE CIENCIA E INNOVACION
"HuPerText" Project (TIN2008-04998)

What's Next


component
word 733
gt component
detection component
edit dist. 0
normalised 0
evaluation.
word 734
gt evaluation.
detection evaluation.
edit dist. 0
normalised 0
Deadline!
word 735
gt Deadline!
detection
edit dist. 9
normalised 1
Rarely
word 736
gt Rarely
detection
edit dist. 6
normalised 1
Asked
word 737
gt Asked
detection Asked
edit dist. 0
normalised 0
Questions
word 738
gt Questions
detection WKBHMlM
edit dist. 9
normalised 1
sensonor
word 739
gt sensonor
detection sensonor
edit dist. 0
normalised 0
SHARE
word 740
gt SHARE
detection SHARE
edit dist. 0
normalised 0
THIS
word 741
gt THIS
detection THIS
edit dist. 0
normalised 0
CONTENT
word 742
gt CONTENT
detection CONTENI
edit dist. 1
normalised 0.14
MORE...
word 743
gt MORE...
detection MORE...
edit dist. 0
normalised 0
30%
word 744
gt 30%
detection
edit dist. 3
normalised 1
off*
word 745
gt off*
Amsterdam
word 746
gt Amsterdam
Apartments
word 747
gt Apartments

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/