Albert Gordo

Lluís Gómez

xerox

Centre de Visió per Computador

Centre de Visió per Computador

Agenda

1. Introduction
2. Scene-text Localization [60 minutes]
   - Theoretical introduction [20 minutes]
   - Practical [30 minutes]
   - Questions [10 minutes]
3. Scene-text Recognition [60 minutes]
   - Theoretical introduction [30 minutes]
   - Practical [20 minutes]
   - Questions [10 minutes]
4. Scene-text Understanding [50 minutes]
   - Theoretical introduction [30 minutes]
   - Practical [20 minutes]
   - Questions [10 minutes]
5. Closing Remarks

Centre de Visió per Computador

Scene text understanding tasks

LOCALIZATION

RECOGNITION

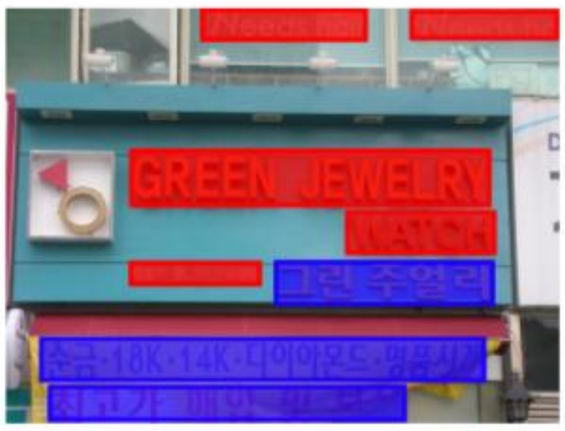
END-TO-END

Scene text understanding tasks

LOCALIZATION

SEGMENTATION

SCRIPT IDENTIFICATION



RECOGNITION

END-TO-END

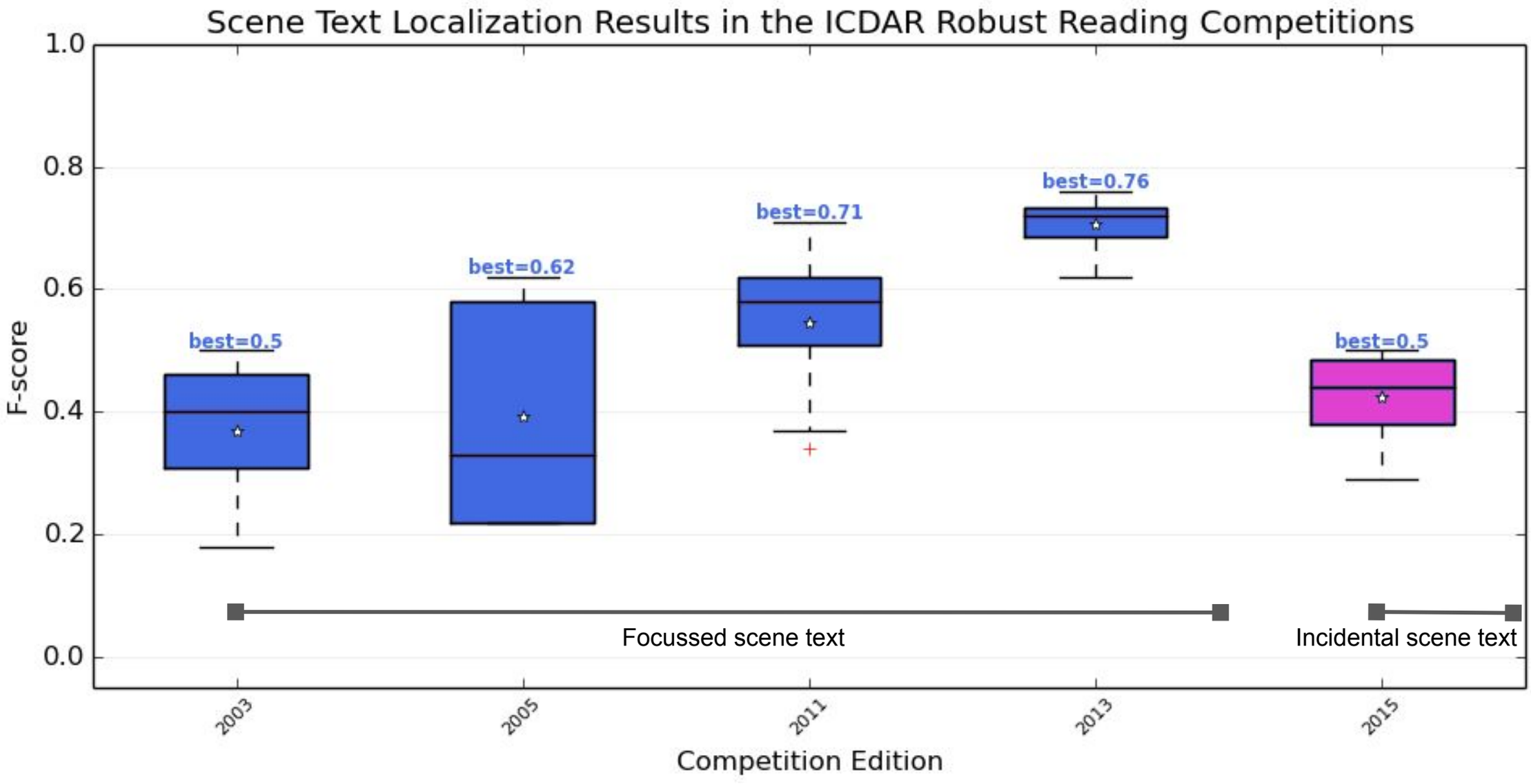
Scene text has a huge intra-class variability

Scene text has a huge intra-class variability

ICDAR 2003 “Focussed text” dataset **[Lucas ICDAR03]**

ICDAR 2015 “Incidental text” dataset **[Karatzas ICDAR15]**

Research progress in the last decade



Main approaches to scene text localization

Sliding window

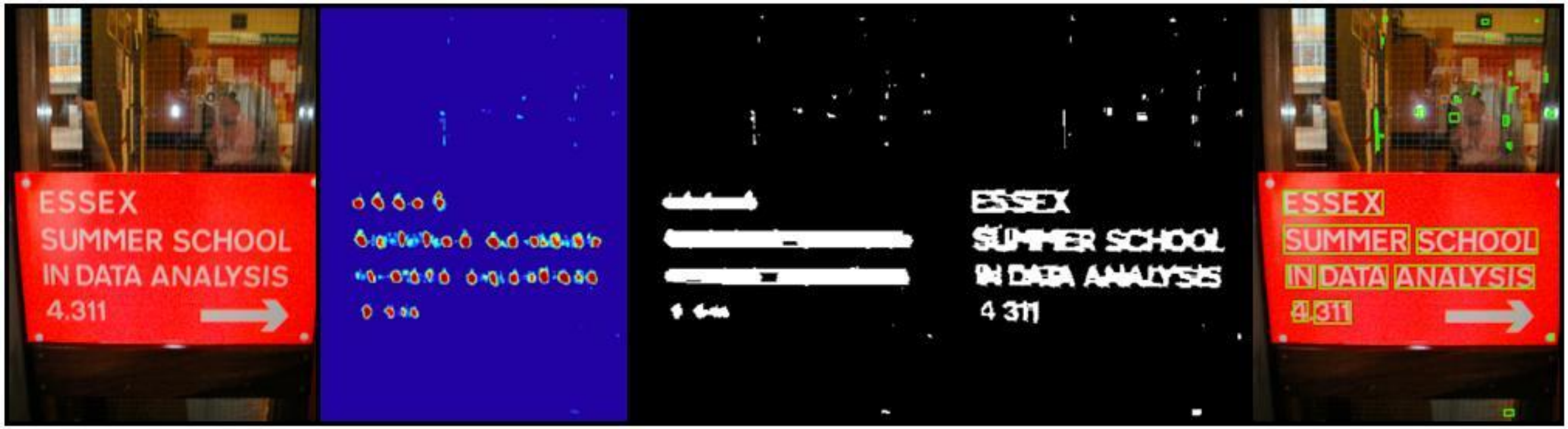
[Coates ICDAR11]

Connected components

[Neumann ACCV10]

Sliding window based text localization

Deep CNN classifiers



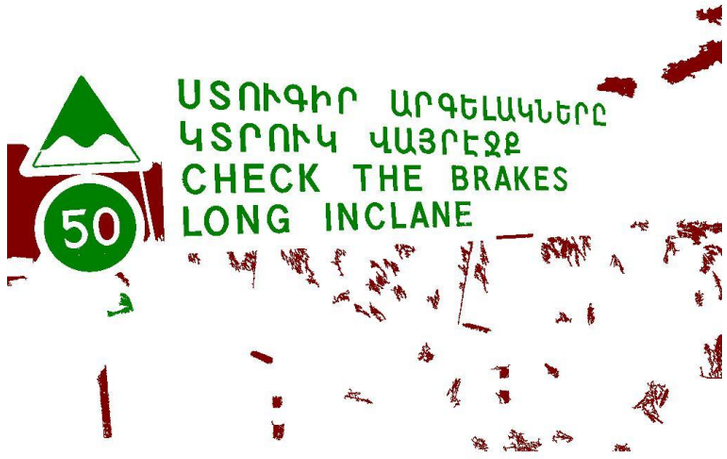
[Jaderberg ECCV14]

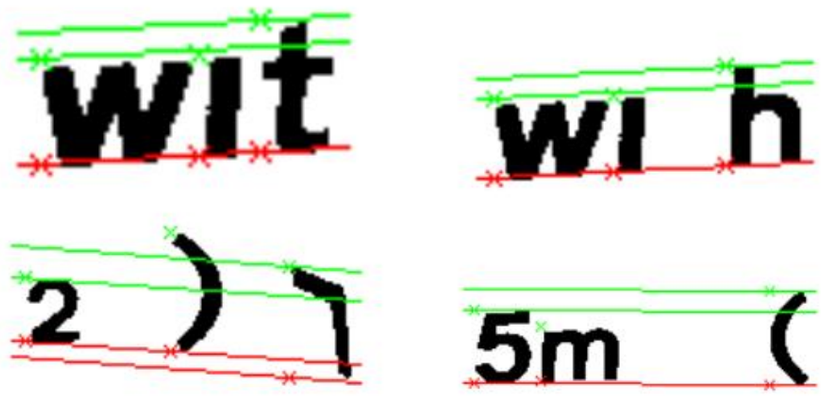
Hand-crafted features (before deep CNNs)

HOG, HaarLike, LBP, ….

Connected components based methods

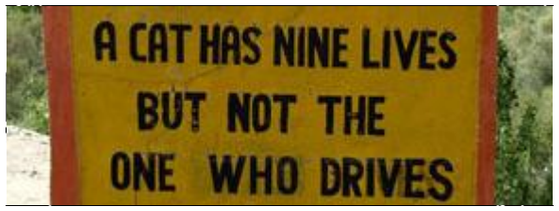
Maximally Stable Extremal Regions (MSER)

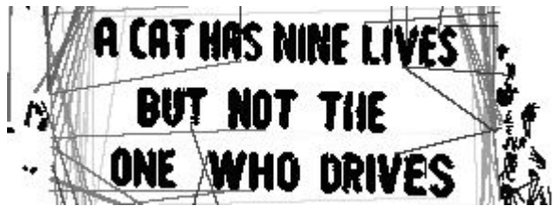




[Neumann ECCV10, Neumann ICDAR11]

Stroke Width Transform (SWT)





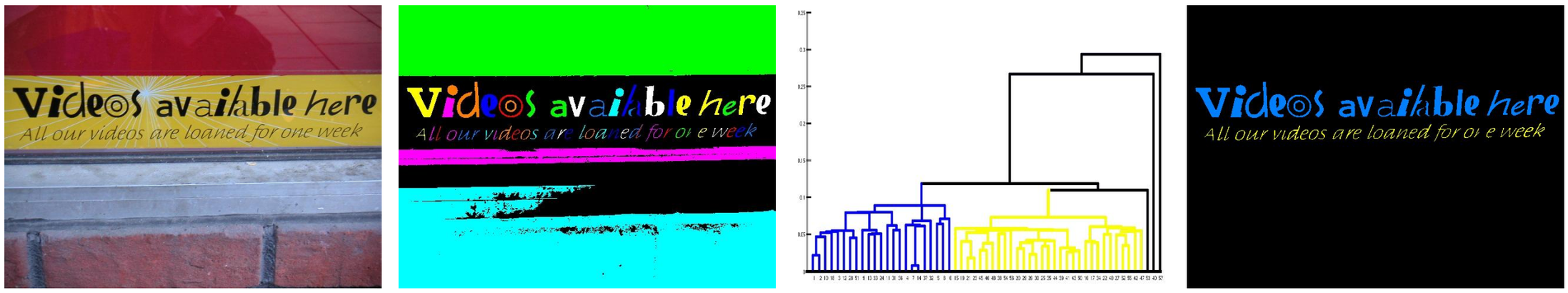
[Epshtein CVPR10]

Character candidates grouping strategies

Font type model heuristics

[Neumann ICDAR11]

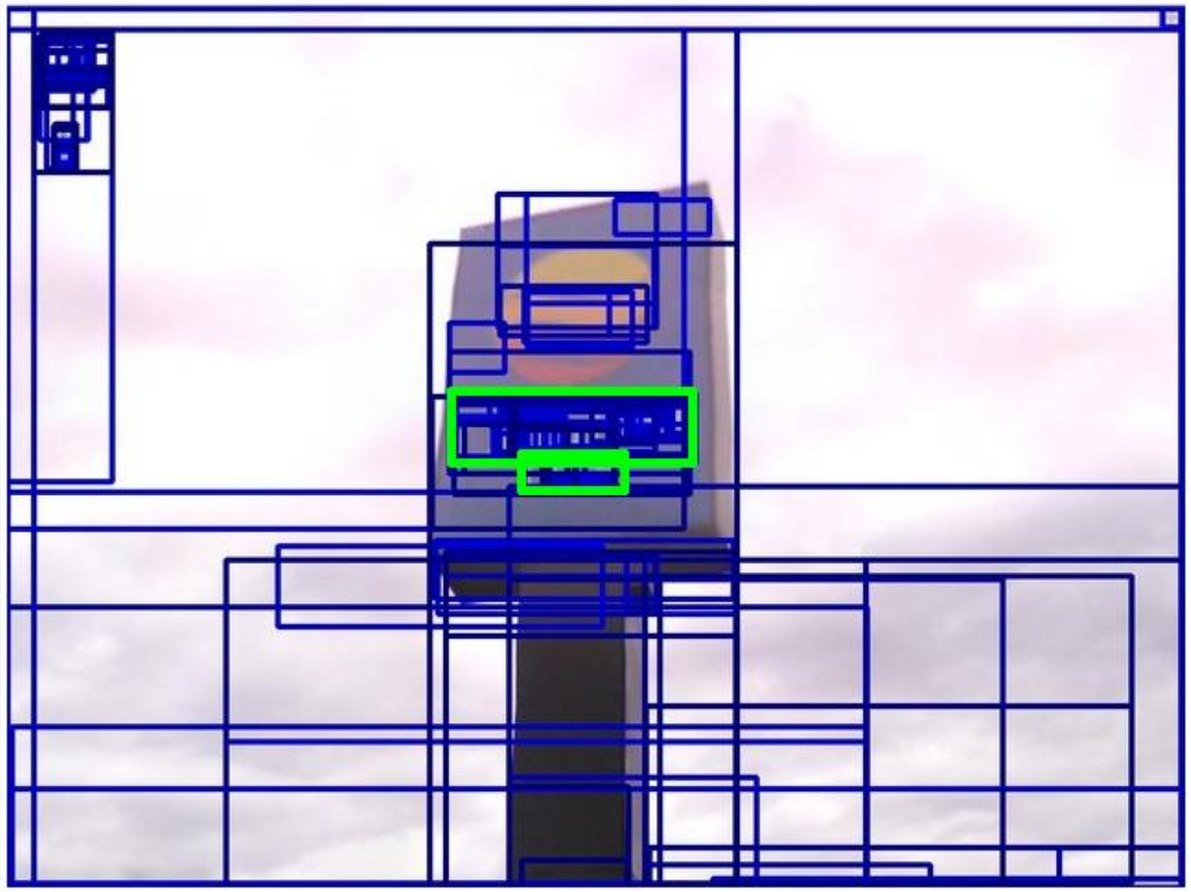
Bottom-up agglomerative grouping by similarity and proximity



[Gomez ICDAR2013, Yin TPAMI2014]

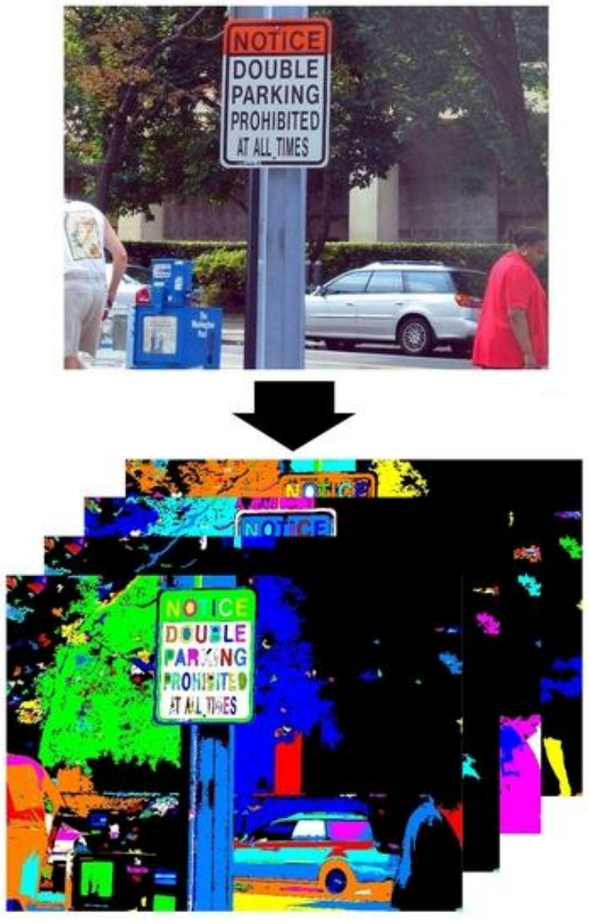
Object Proposals for end-to-end pipelines

Exhaustive search vs. object proposals

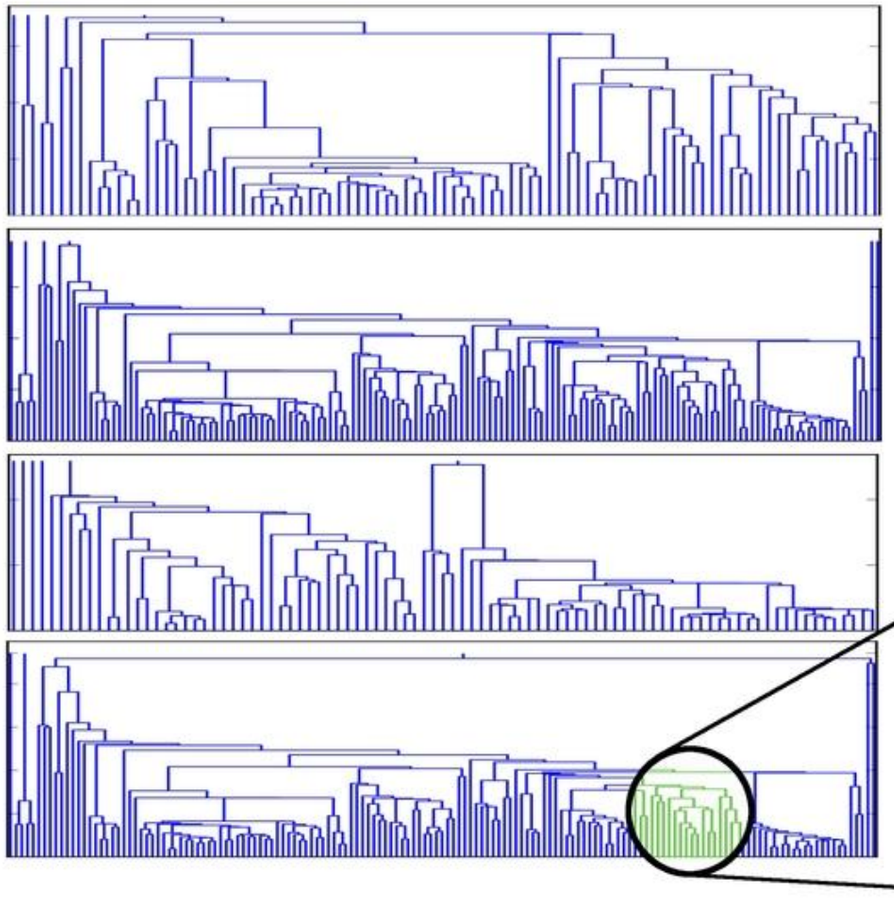


Object Proposals for end-to-end pipelines

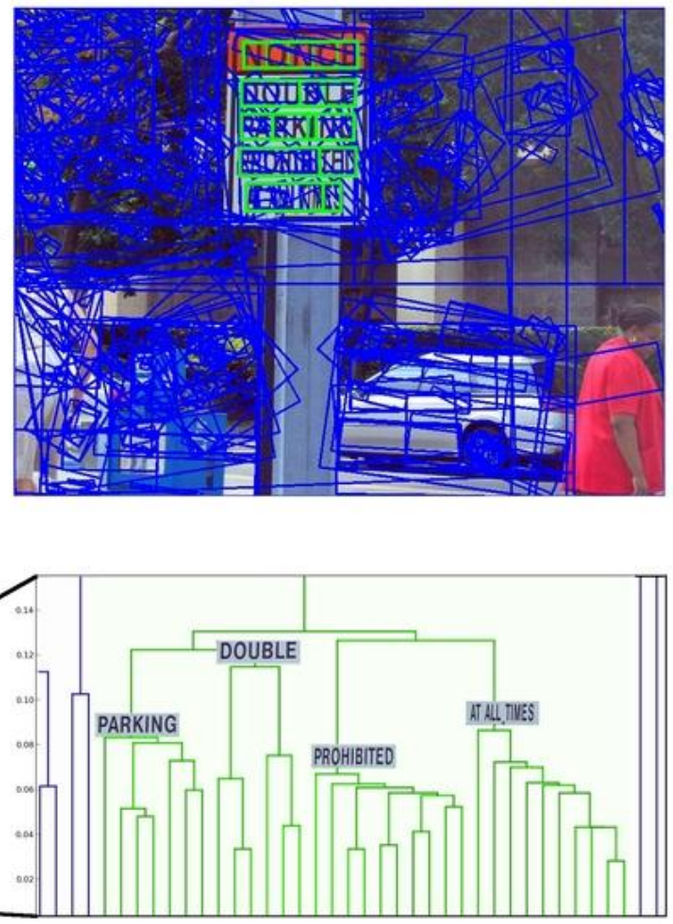
TextProposals: text specific selective search



Diversification (I) extract regions from different color channels and scales



Diversification (II) agglomerate regions using different similarity/proximity metrics.



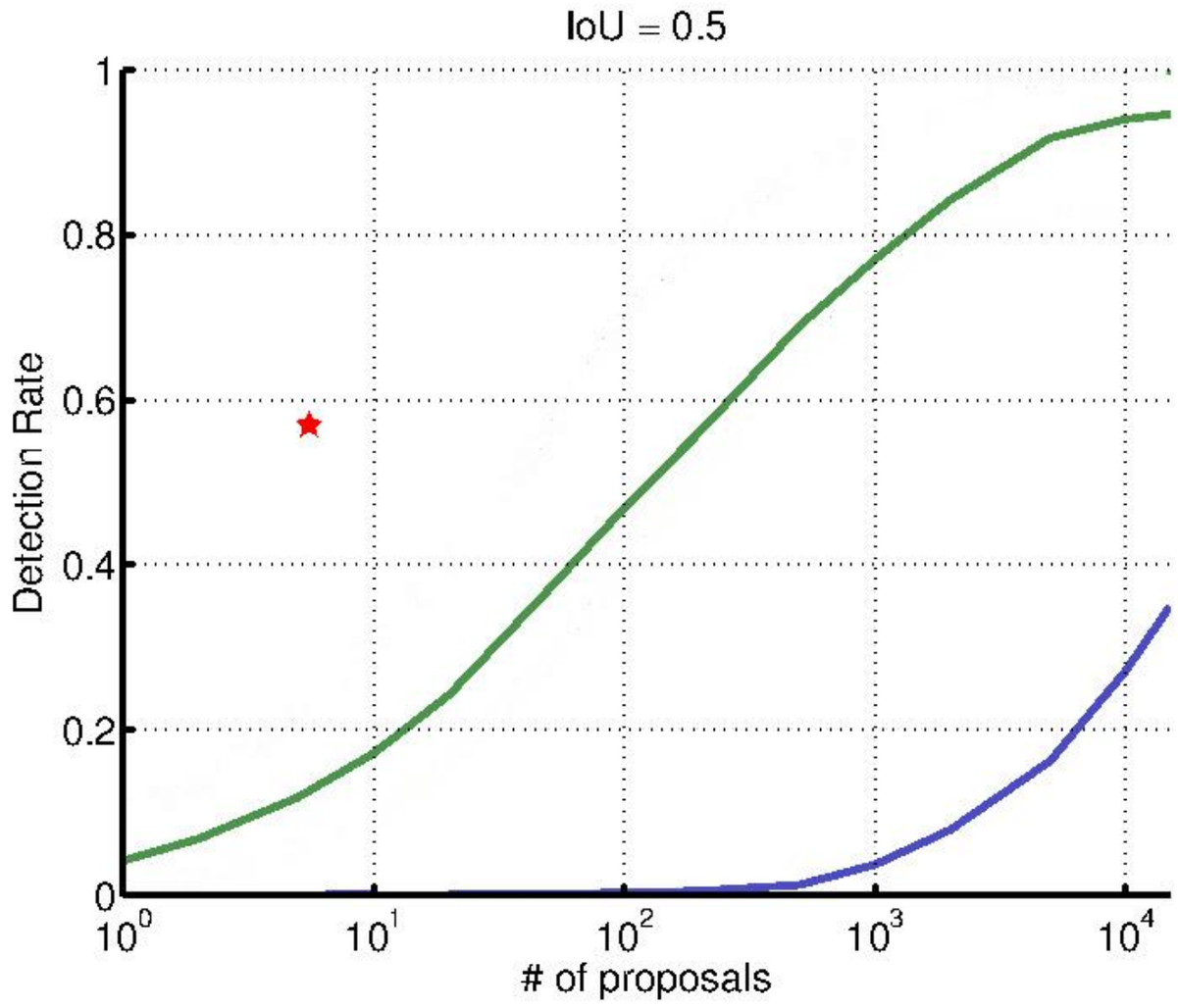
Inference: apply holistic word recognition methods.

[Gomez ICDAR15, Gomez 2016]

Object Proposals for end-to-end pipelines

ICDAR2015 Incidental Text dataset - Localization rates (recall)

TextProposals vs. class-independent proposals vs. traditional text detector



[Gomez ICDAR15, Gomez 2016]

OpenCV text module

text. Scene Text Detection and Recognition

The opencv_text module provides different algorithms for text detection and recognition in natural scene images.

http://docs.opencv.org/3.0-beta/modules/text/doc/text.htmlhttp://docs.opencv.org/3.0-beta/modules/text/doc/text.html

Other publicly available algorithms

[Wang ICCV11] http://vision.ucsd.edu/~kai/grocr/

[Minnetto CVIU13] http://www.dainf.ct.utfpr.edu.br/~rminetto/projects/snoopertext/

[Epshtein CVPR10] http://libccv.org/doc/doc-swt/

[Gomez ICDAR13] https://github.com/lluisgomez/text_extraction

[Busta ICCV15] https://github.com/MichalBusta/FASText

others ...

Centre de Visió per Computador

Tutorial demos and code

- Installing the OpenCV text module is easy. (Use the Source, Luke!)
- MSER demo
- Class Specific Extremal Regions
- Region Grouping
- Text Proposals