



# Automatic Text Localisation in Scanned Comic Books

Christophe Rigaud<sup>1</sup>, **Dimosthenis Karatzas**<sup>2</sup>, Joost Van de Weijer<sup>2</sup>, Jean-Christophe Burie<sup>1</sup>, Jean-Marc Ogier<sup>1</sup>

<sup>1</sup> Laboratory L3i, University of La Rochelle, Avenue Michel Crépeau 17042 La Rochelle, France

<sup>2</sup> Computer Vision Center, Universitat Autnoma de Barcelona, E-08193 Bellaterra (Barcelona), Spain  
{christophe.rigaud, jean-marc.ogier, jean-christophe.burie}@univ-lr.fr, {dimos, joost}@cvc.uab.es

## Presentation

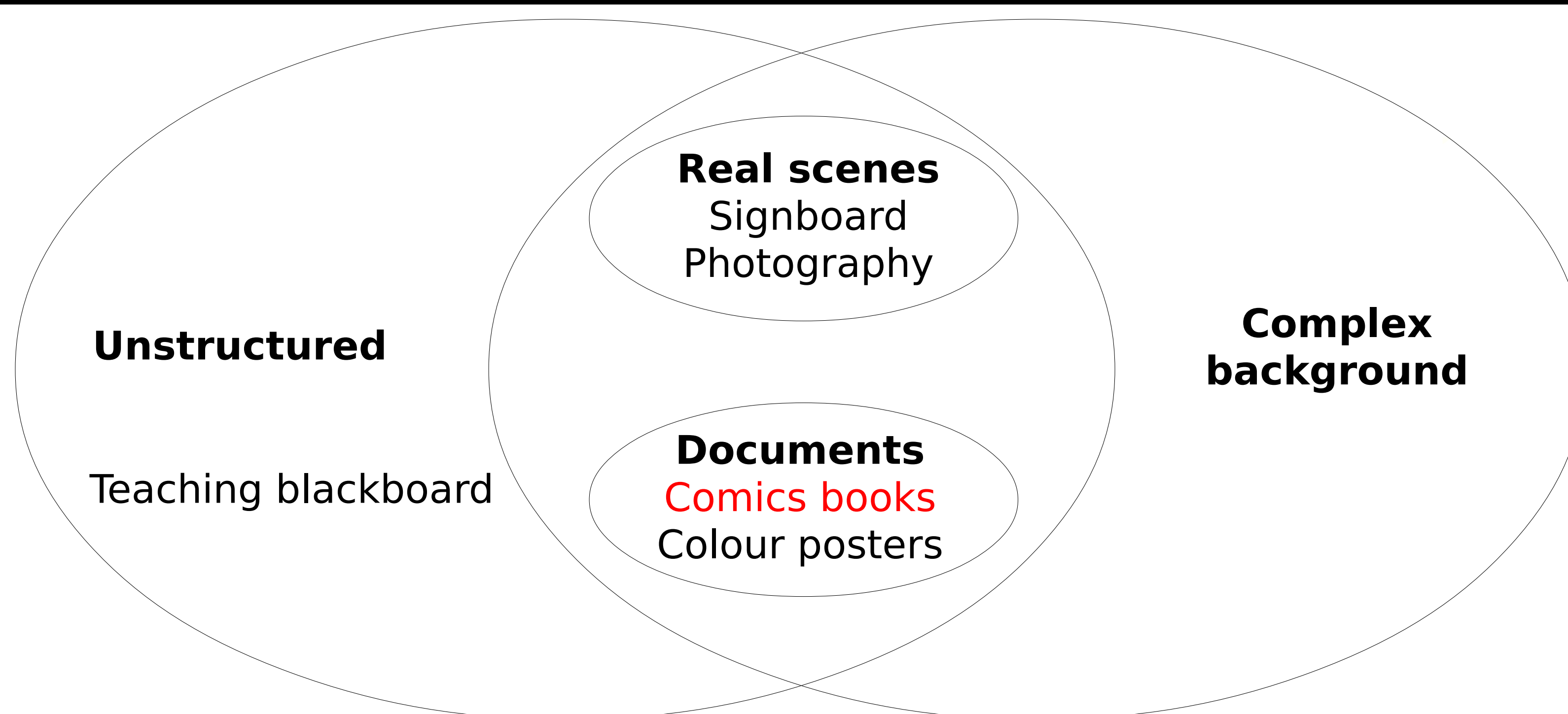
### Keywords

Comic content understanding

Text / graphic separation

Text lines localisation

Benchmark dataset



## Experiments

### Dataset

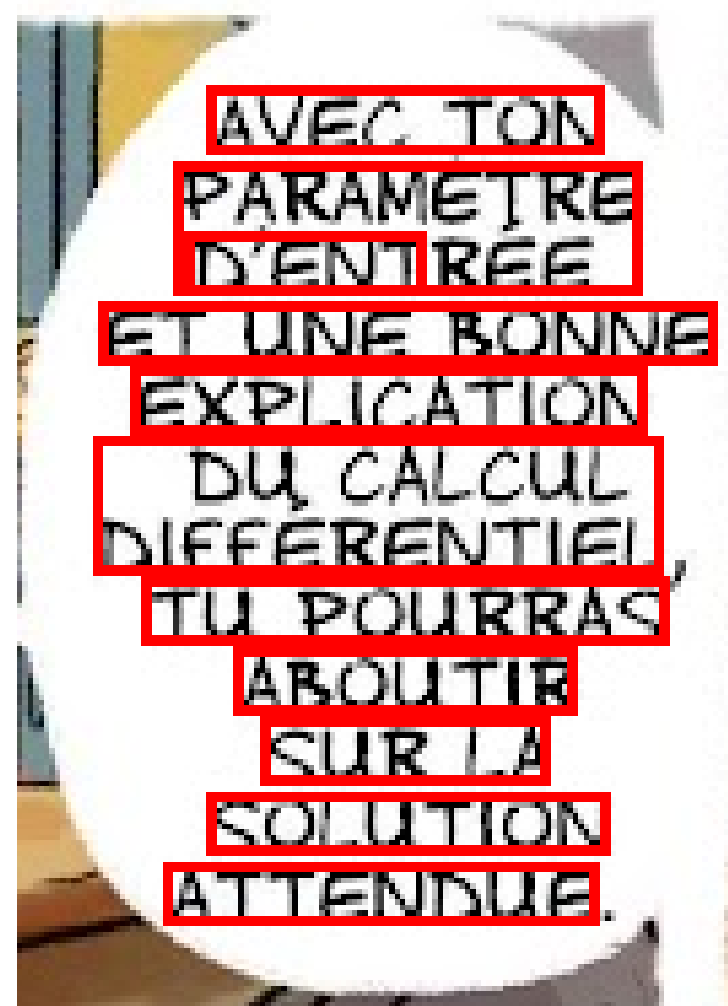
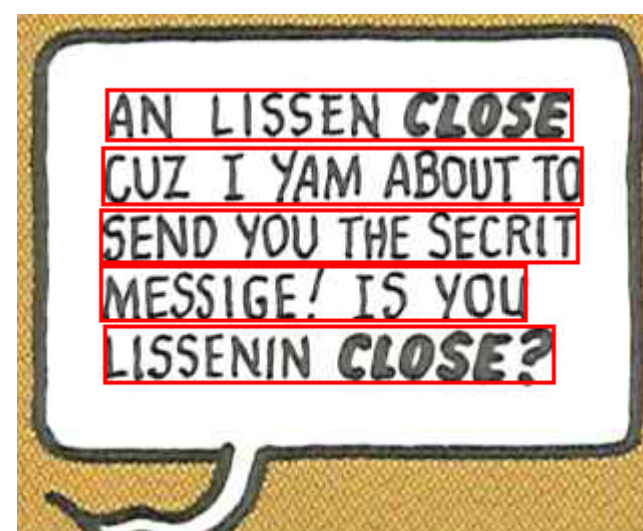
image 1 recall 79.26% precision 84.22% fmean 79.73%	image 2 recall 69.23% precision 73.61% fmean 71.42%	image 3 recall 69.23% precision 80.4% fmean 74.81%	image 4 recall 66.67% precision 61.2% fmean 63.9%	image 5 recall 84.1% precision 84.71% fmean 84.4%
image 6 recall 86.11% precision 80.54% fmean 83.3%	image 7 recall 86.43% precision 83.78% fmean 85.1%	image 8 recall 86.11% precision 74.14% fmean 79.8%	image 9 recall 86.67% precision 86.67% fmean 86.67%	image 10 recall 89.23% precision 76% fmean 82.2%
image 11 recall 100% precision 80.46% fmean 87.97%	image 12 recall 87.47% precision 81.67% fmean 89.56%	image 13 recall 100% precision 82.35% fmean 90.32%	image 14 recall 81.3% precision 85.45% fmean 83.3%	image 15 recall 86.85% precision 86.85% fmean 86.85%
image 16 recall 79.23% precision 89.26% fmean 84.55%	image 17 recall 87.47% precision 76.47% fmean 79.27%	image 18 recall 86.7% precision 88.24% fmean 87.47%	image 19 recall 75.13% precision 77.24% fmean 76.17%	image 20 recall 86.85% precision 86.85% fmean 86.85%

eBDtheque [1]

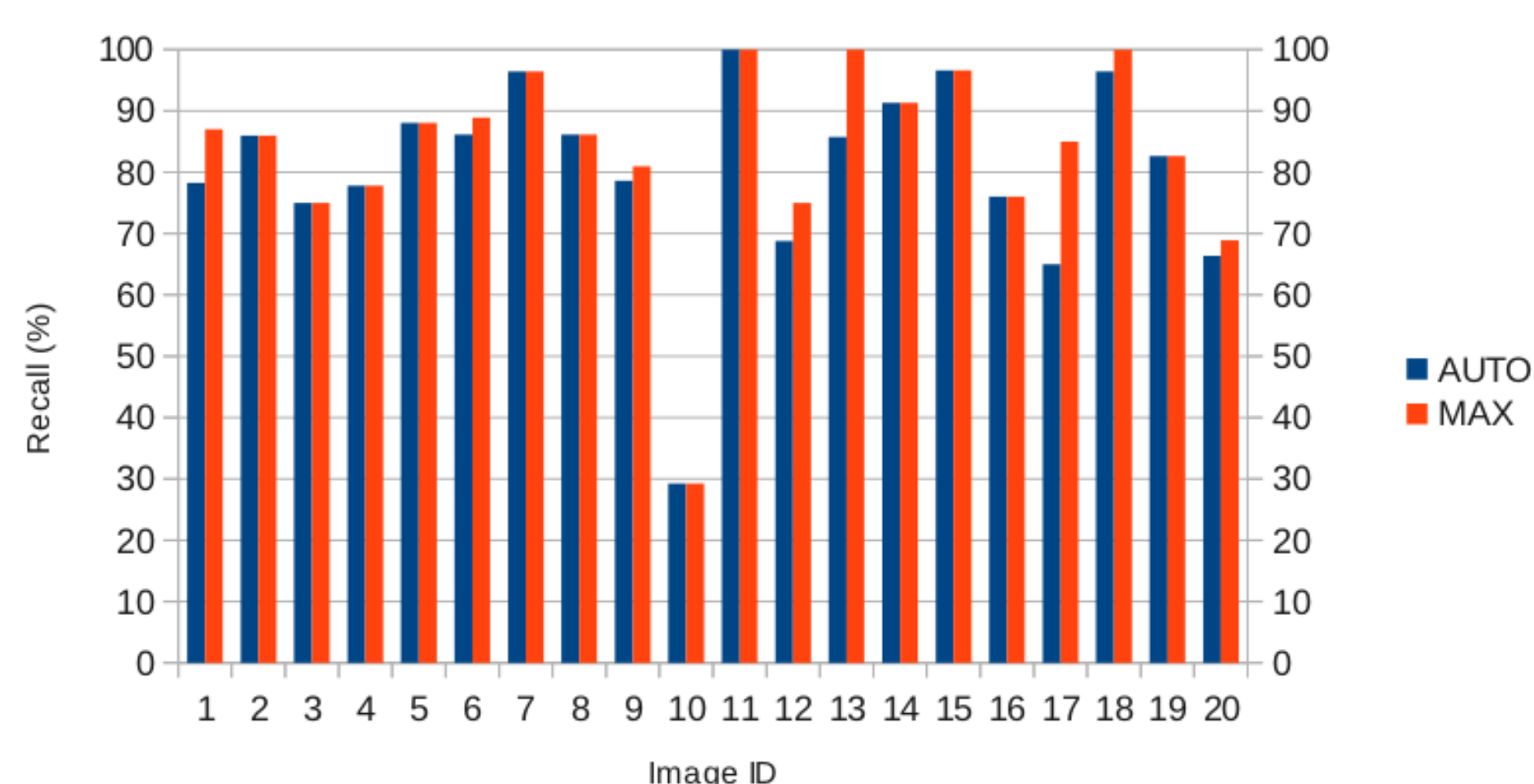
- 1700 text lines
- 100 to 300 DPI
- Text line level

### Results

ET PUIS LEUR CERVEAU, COMME LEURS YEUX  
SONT ENCORE IMMATURES. ILS N'ENREGISTRENT  
PAS GRAND CHOSE...



### Evaluation



## Contributions

### Minimum Connected Component Thresholding (MCCT)

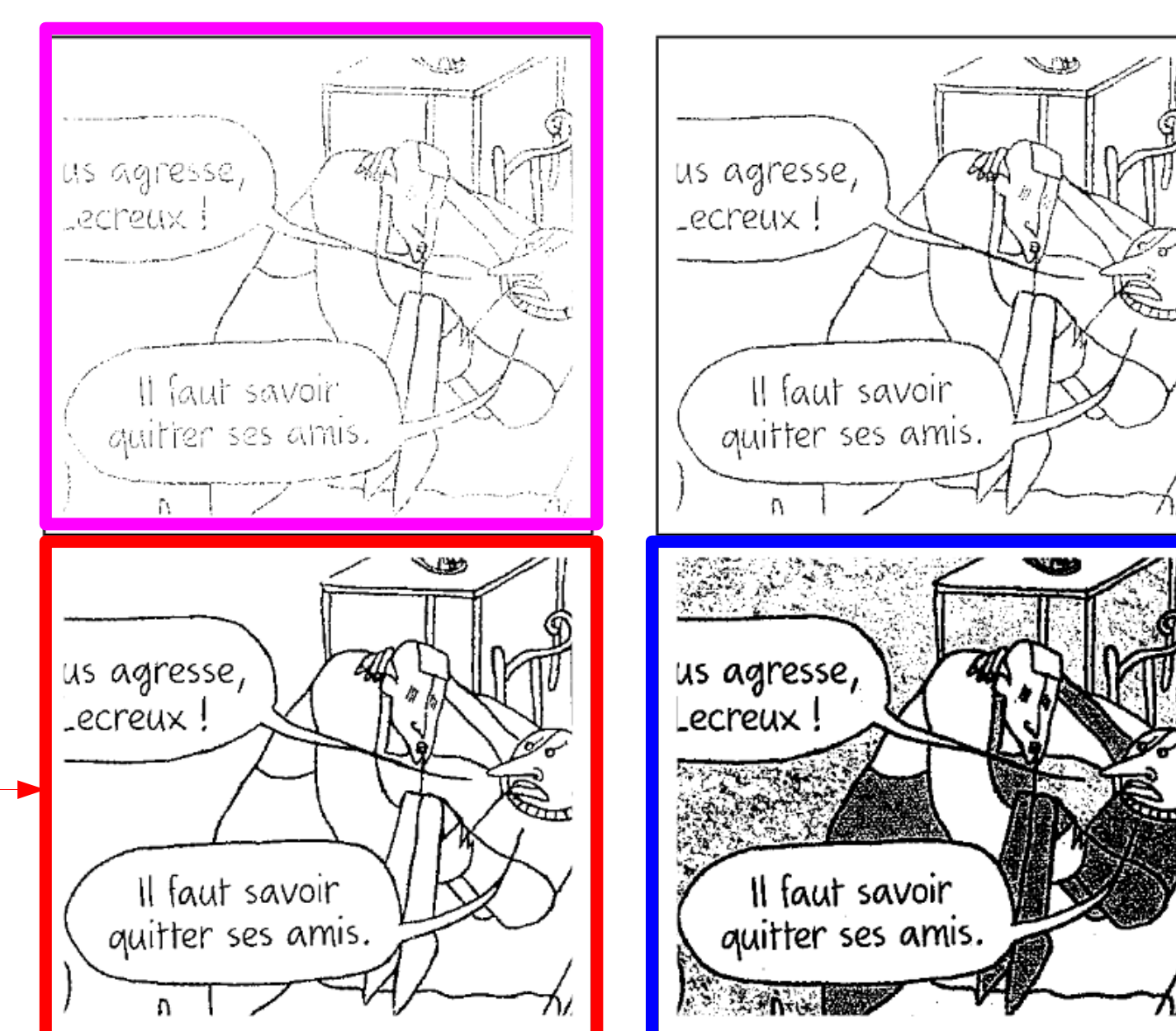
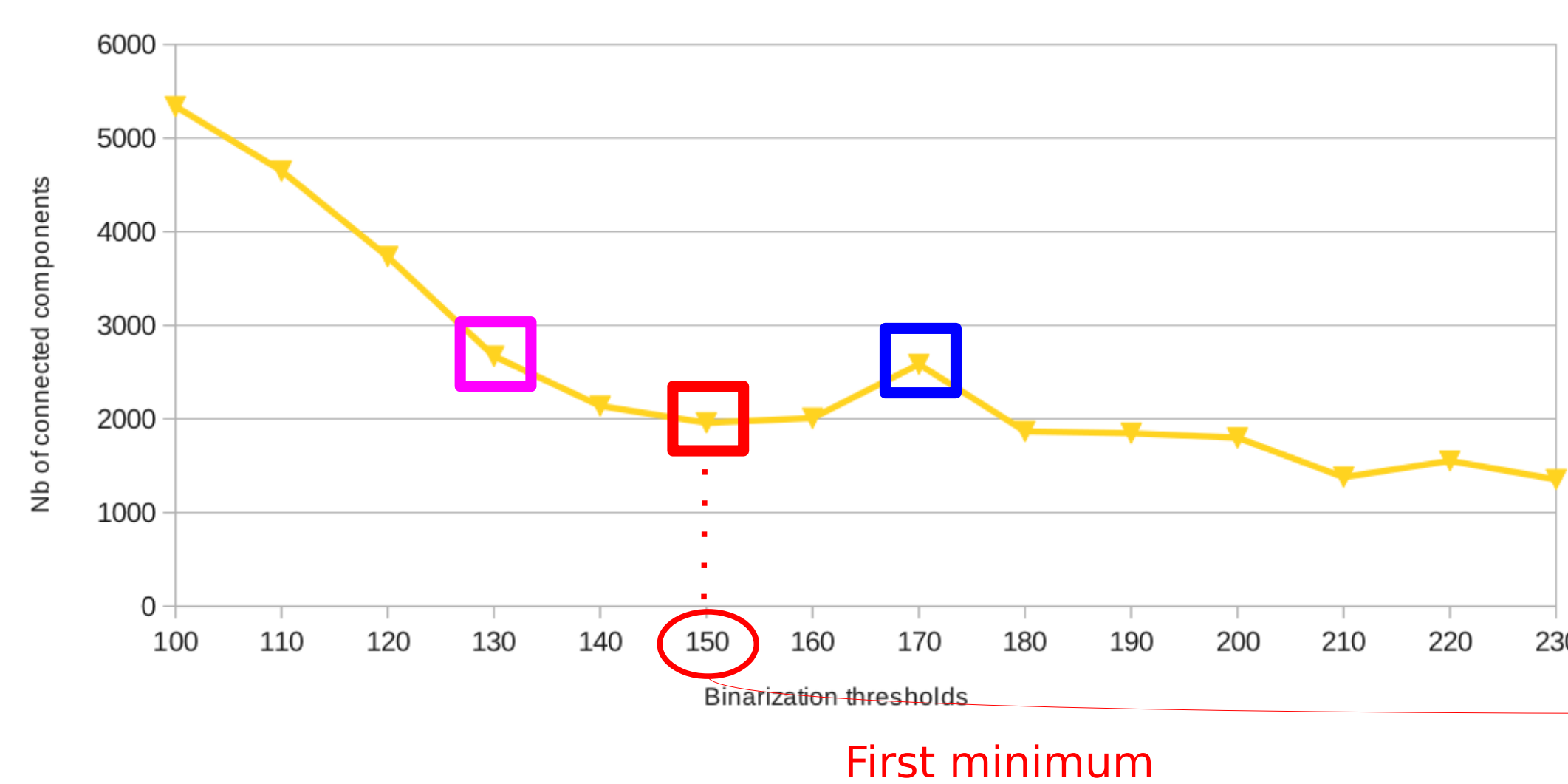
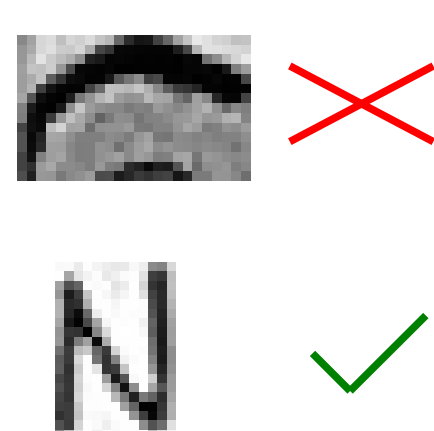


Image credit [4]

### Text / graphic separation

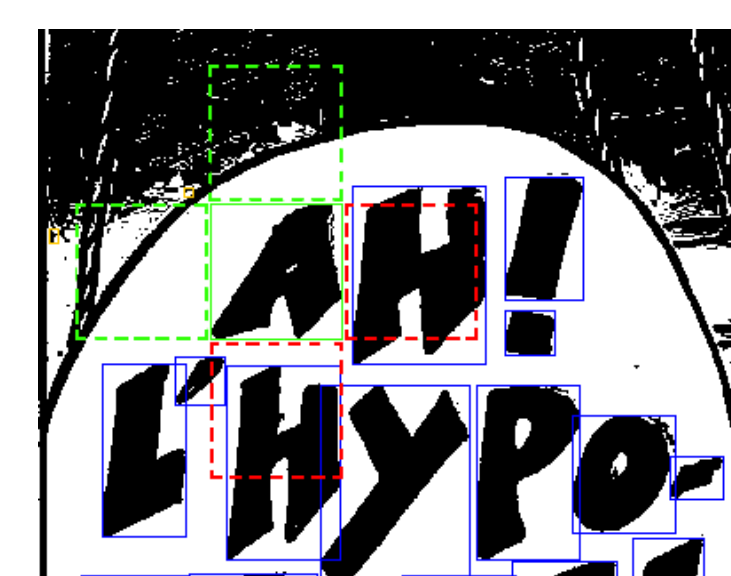
Contrast



Black on white

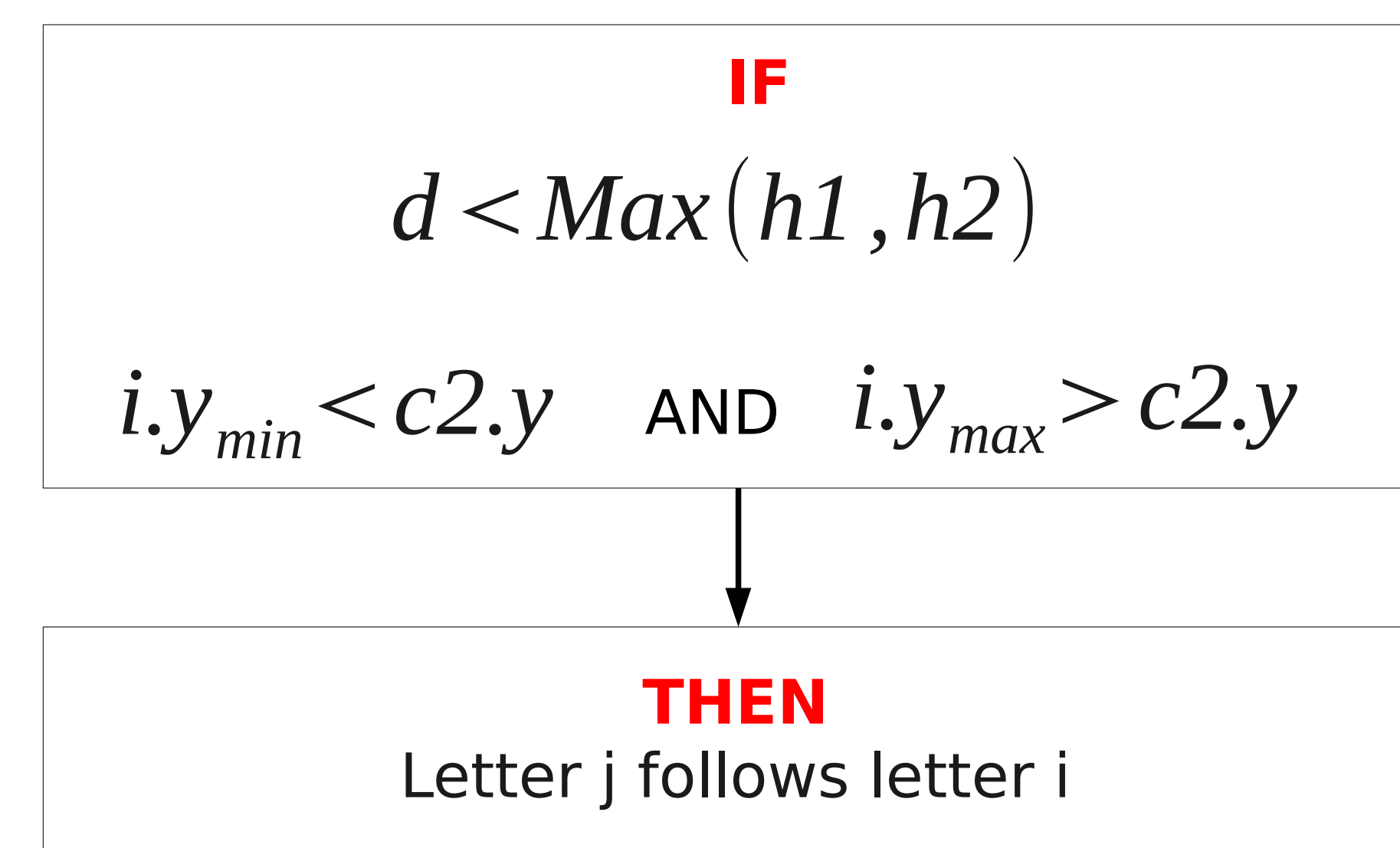
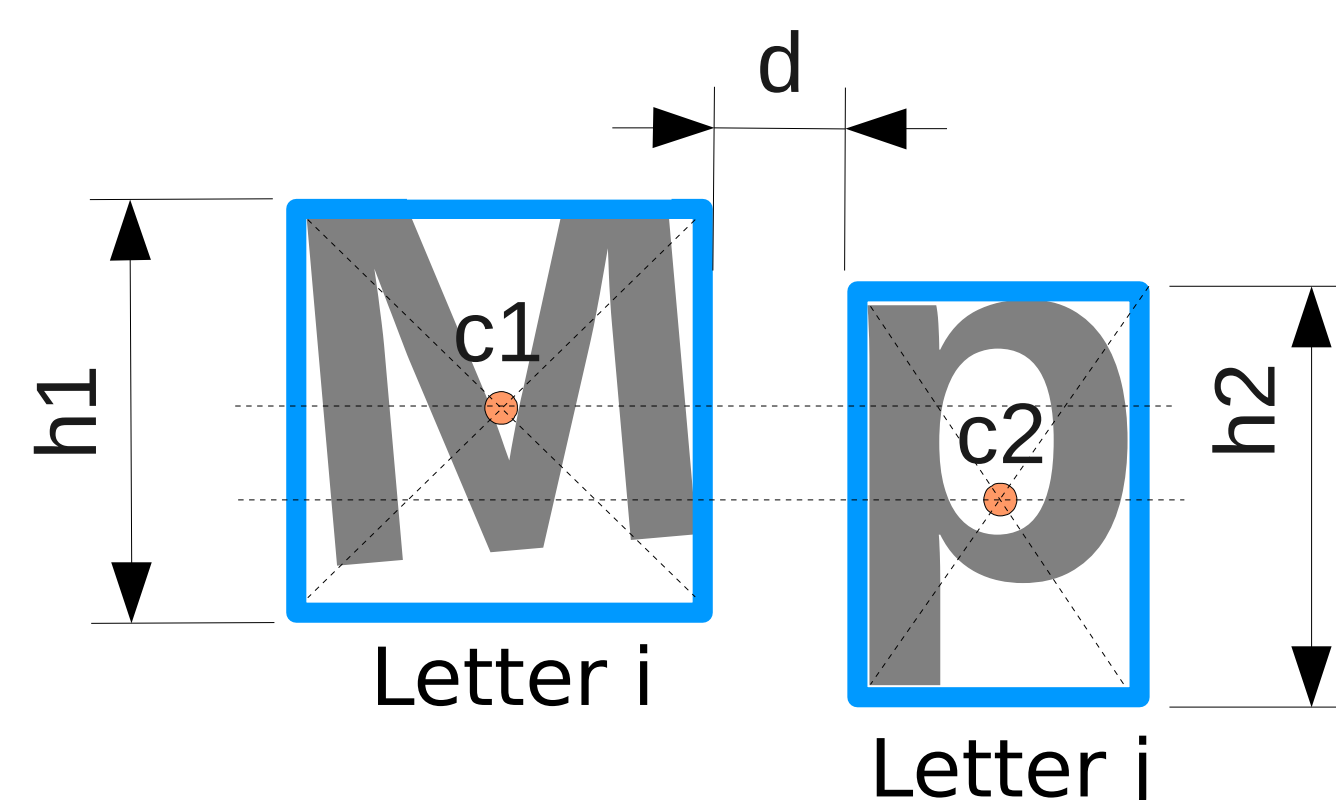


Neighbourhood



Inclusion

### Text line localisation



## Applications

### Localization

Image compression  
OCR training  
Retargeting & reflowing

### Localization + OCR

Automatic translation  
Speech synthesis

## Conclusion & Perspectives

Comics are specific drawing  
Consider other type of text  
What about speech balloon?

## References

- [1] eBDtheque database, website: <http://ebdtheque.univ-lr.fr>
- [2] Neumann, L. and Matas, J. Real-time scene text localization and recognition. *Computer Vision and Pattern Recognition (CVPR)*, pages 1485–1490. 2012
- [3] Tombre, K., Tabbone, S., Plissier, L., Lamiroy, B., and Dosch, P. Text/graphics separation revisited. In *Workshop on Document Analysis Systems (DAS)*, pages 200–211. Springer-Verlag. 2002
- [4] Roudier, N. Les terres creusees, Acte sur BD. Actes Sud. 2011

Segment.	Text/graphic sepa.	R (%)	P (%)
(Neumann and Matas, 2012)[2]		12.56	30.19
Colour	Proposed	15.69	6.92
Proposed	(Tombre et al., 2002)[3]	74.18	61.25
Otsu	Proposed	75.14	64.14
Proposed	Proposed	<b>75.82</b>	<b>76.15</b>