



CBDAR 2011

# NEOCR: A Configurable Dataset for Natural Image Text Recognition

Robert Nagy, Anders Dicker, Klaus Meyer-Wegener

22.09.2011

- **motivation**
- **related work**
- **NEOCR dataset**
- **experiments**
- **conclusion**

- **applications for OCR in natural images**
  - help for visually impaired
  - mobile applications
  - object classification
  - image annotation
  - driving assistant systems

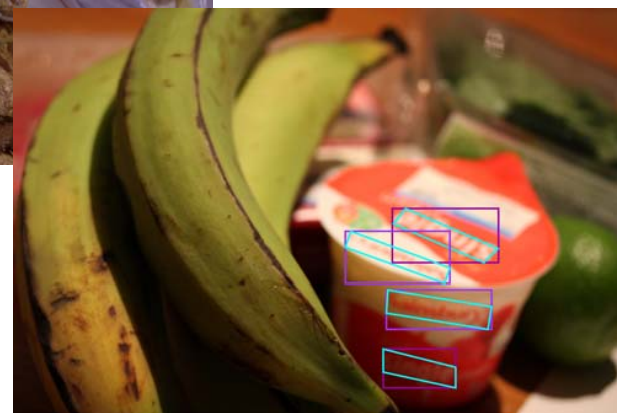
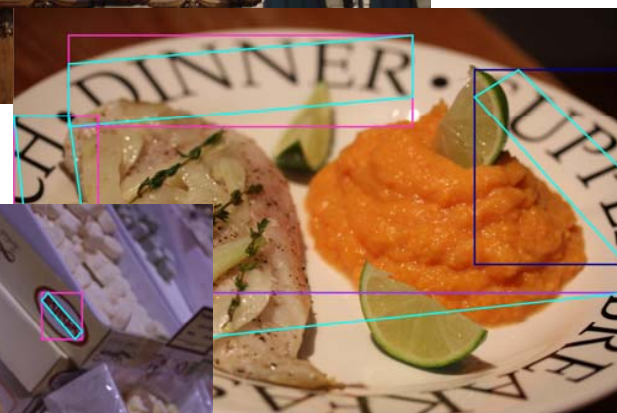
- **background**
- **blurredness**
- **camera position**
- **character arrangement**
- **colors**
- **contrast**
- **font size**
- **font type (diversity in document)**
- **font type (in general)**
- **noise**
- **number of lines**
- **occlusion**
- **rotation**
- **surface**

- **current datasets for OCR in natural images**
  - ICDAR 2003 (Robust Reading) Dataset
  - Chars74K
  - Street View Text Dataset
  - Microsoft Text Detection Database
  
- **limitations of current datasets**
  - bounding boxes parallel to the axes
  - not all text visible in the images annotated
  - limited scope of natural image variations
  - no additional metadata

- **comparison of natural image text recognition datasets**

<b>Dataset</b>	<b># images</b>	<b># boxes</b>	<b>avg. # char/box</b>
ICDAR 2003	509	2263	6.15
Chars74K	312	2112	6.47
MS Text DB	307	1729	10.76
Street View Text	350	904	6.83
<b>NEOCR</b>	<b>659</b>	<b>5238</b>	<b>17.62</b>

- 659 images, 5238 textfields
- distortion quadrangles
- rich annotation (XML)
  - optical,
  - geometrical and
  - typographical characteristics



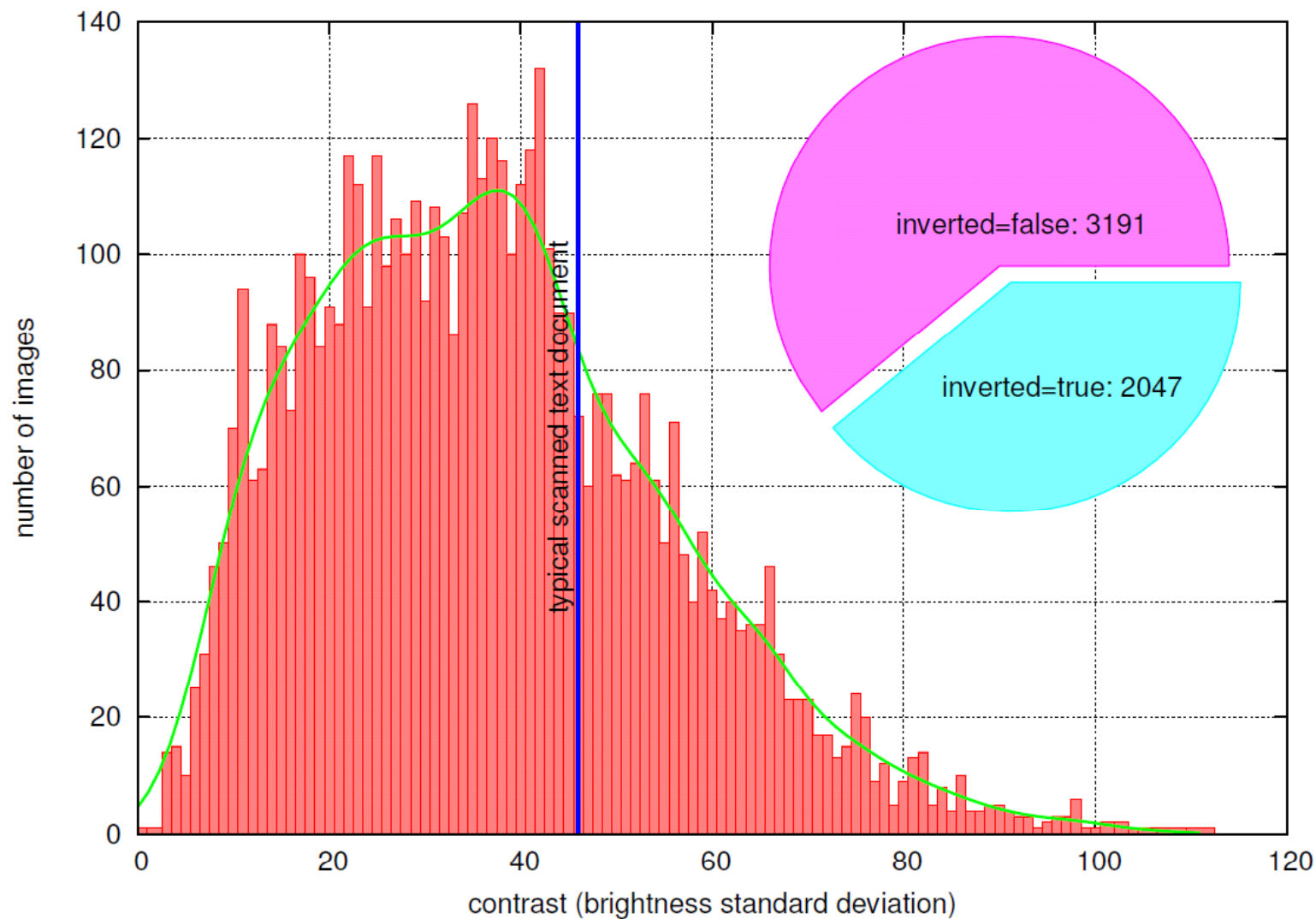
**distorted**

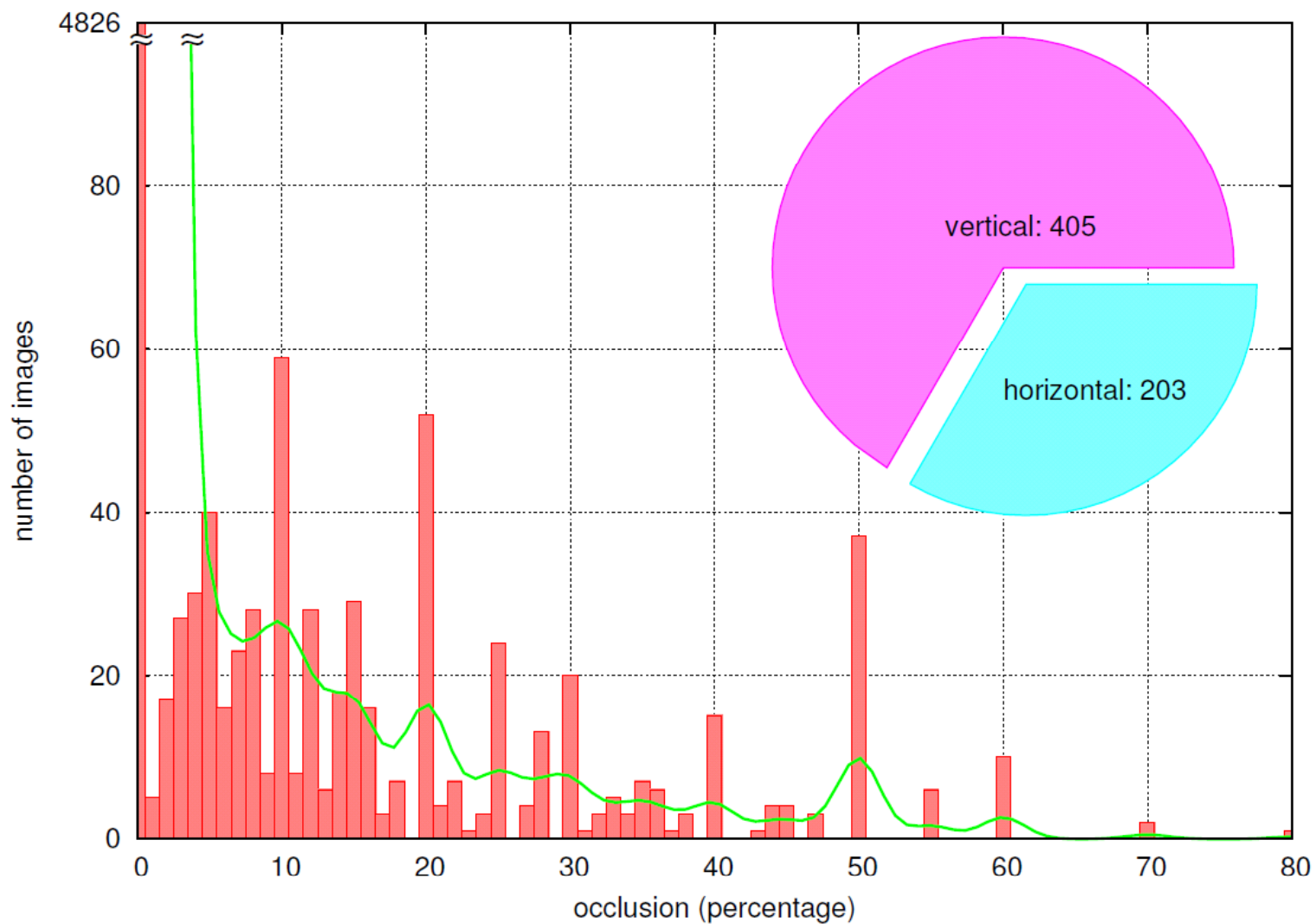


**straightened**

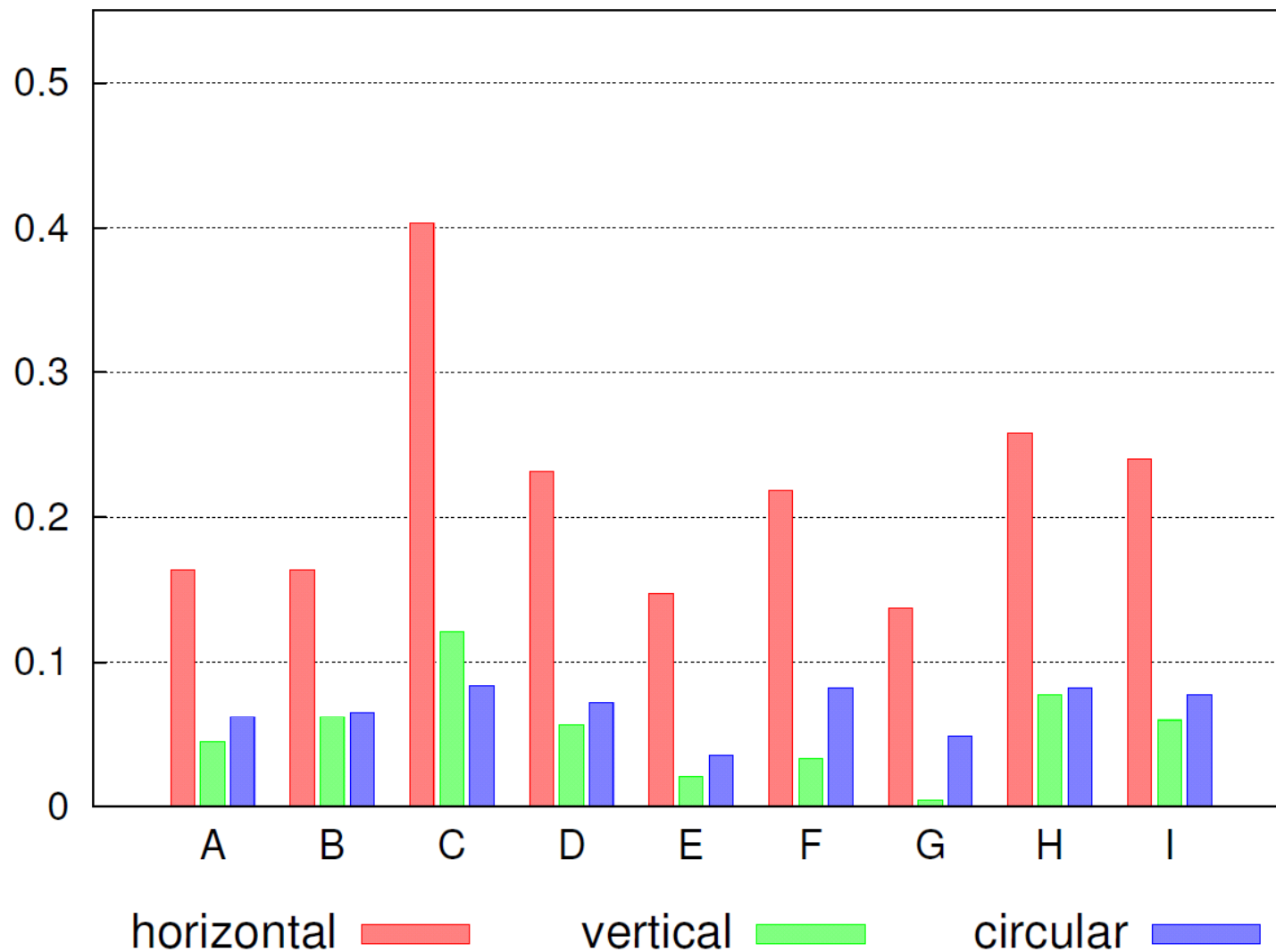


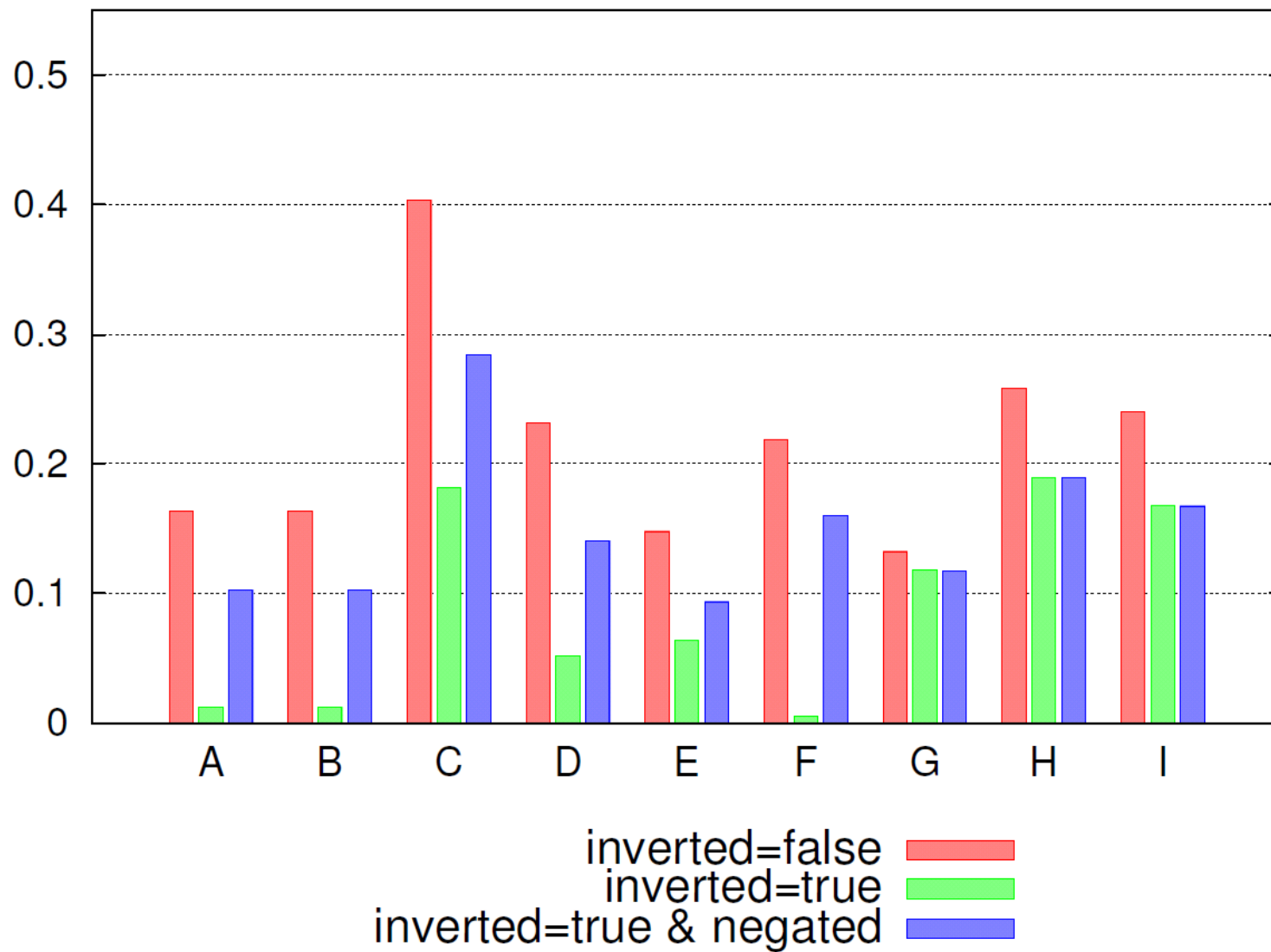






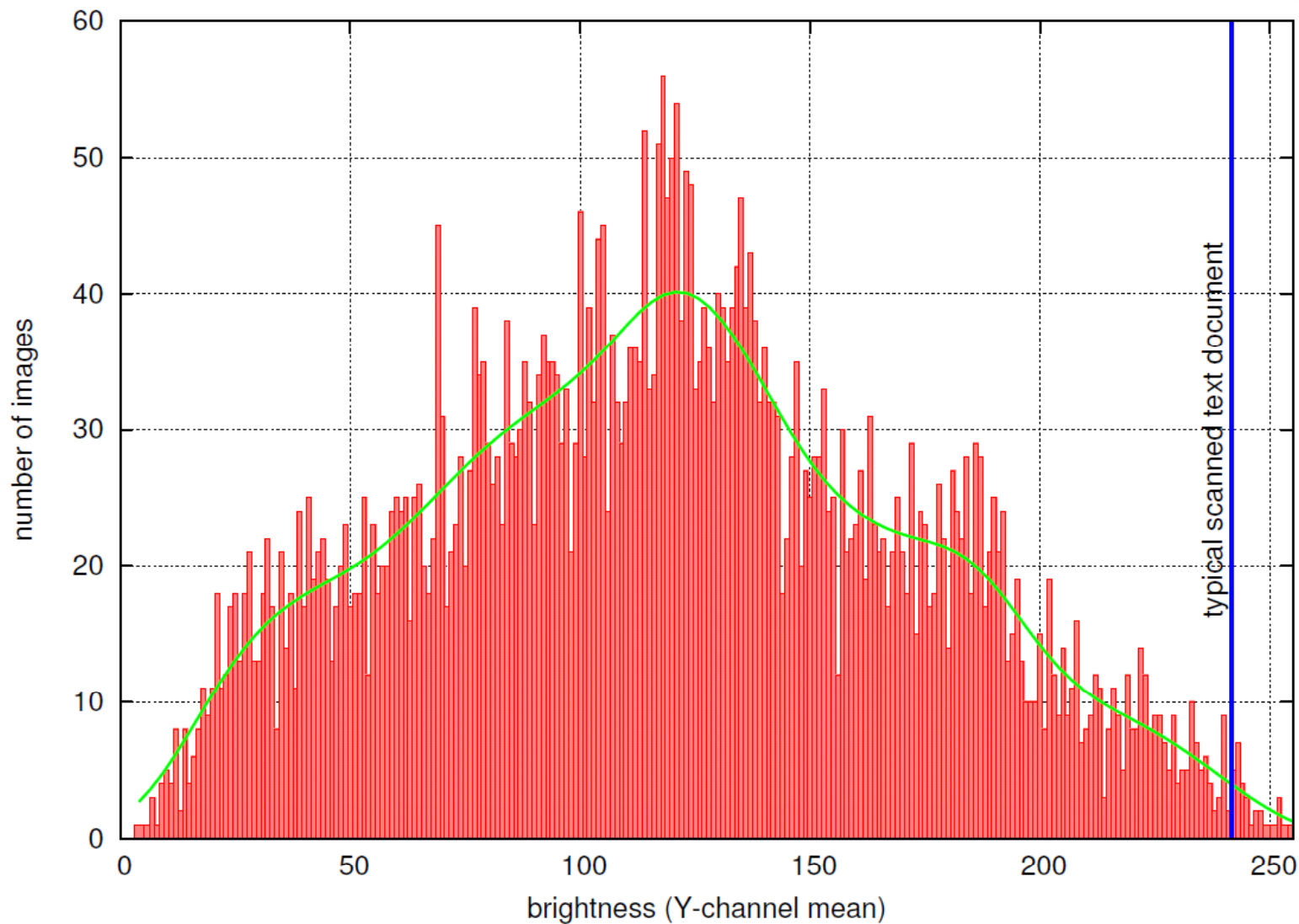
- **configurable dataset**
- **special test scenarios, e.g.:**
  - character arrangement
  - contrast / inversion
  - distortion
  - impact of vocabularies
  - combinations of these and more
- **results using popular open source and commercial OCR tools**
  - only bounding boxes were used
  - Levenshtein (edit-)distance
  - similarity:  $s(x_n, y_m) = 1 - \frac{d(x_n, y_m)}{\max(n, m)}$



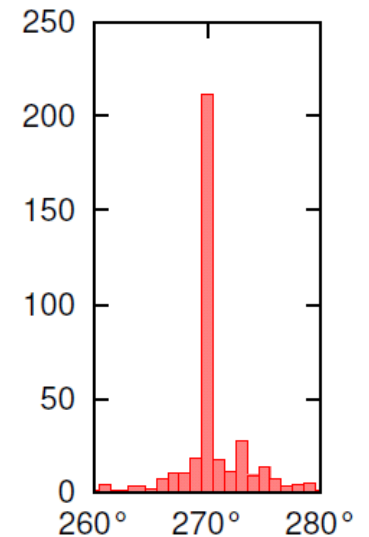
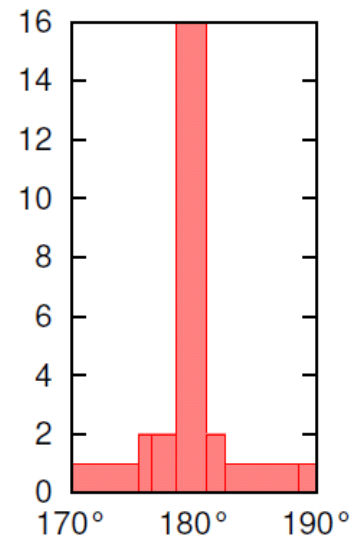
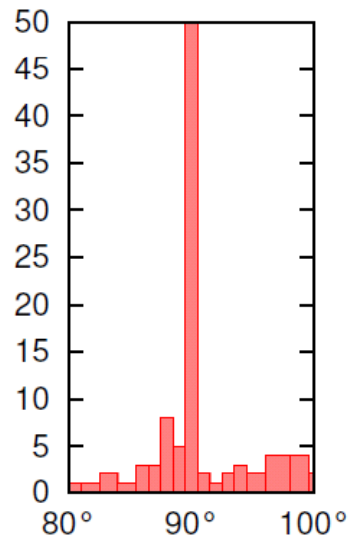
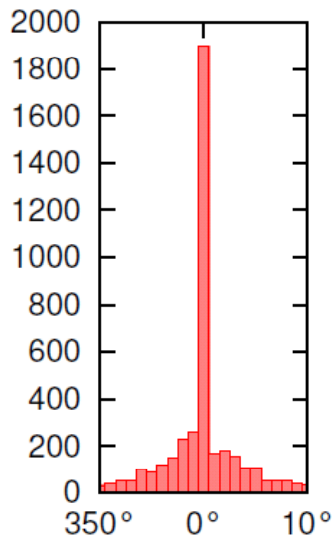
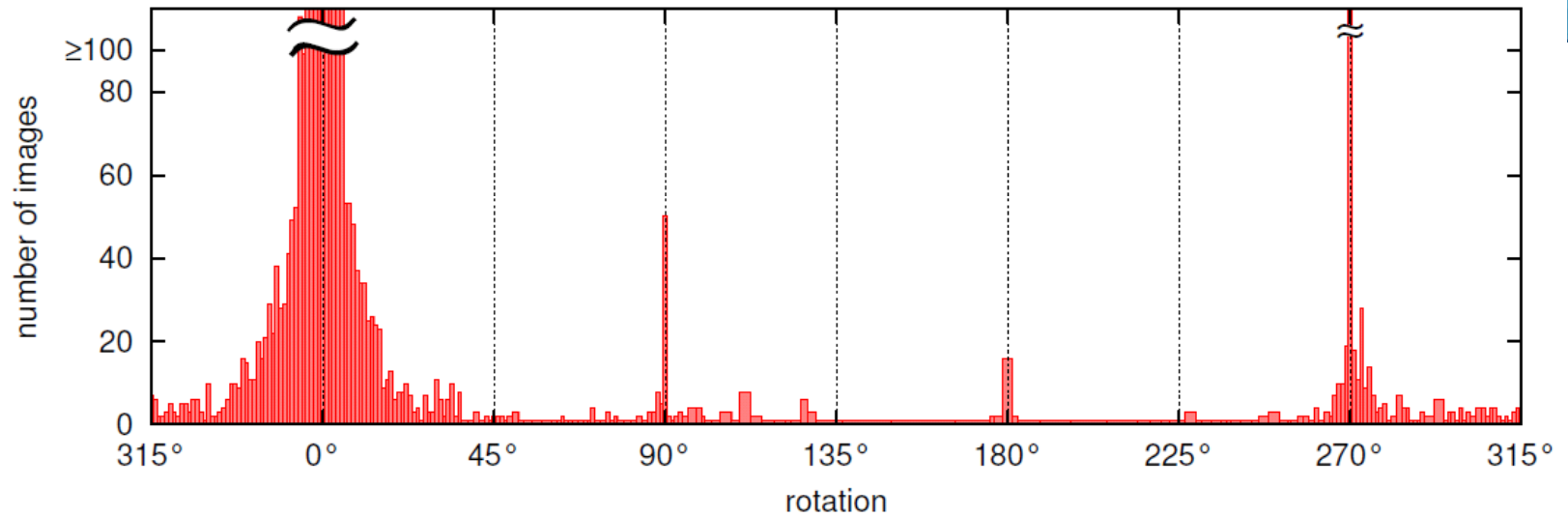


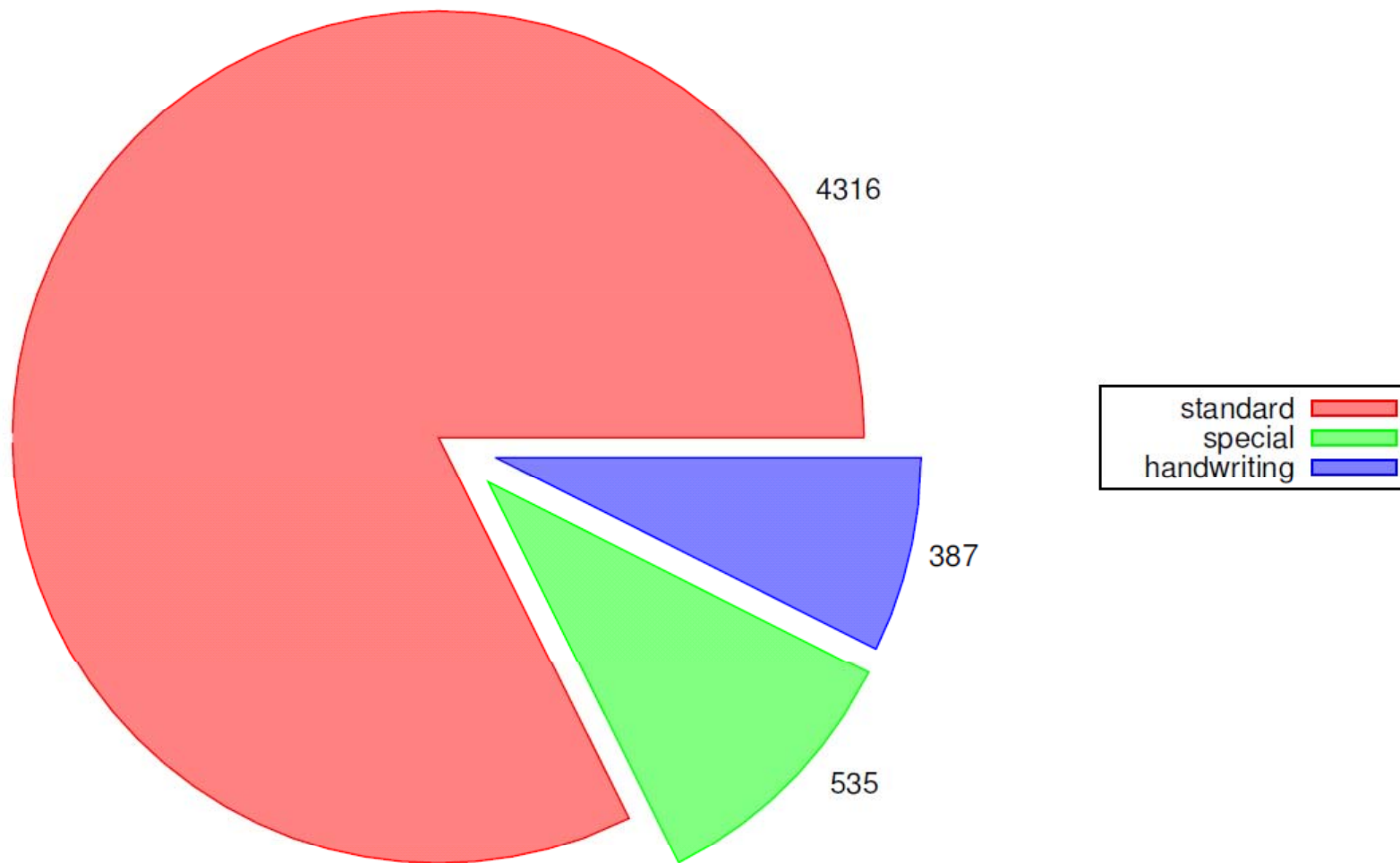
- **NEOCR dataset**
  - all visible textfields annotated
  - rich additional metadata
  - distortion quadrangles
- **future work**
  - extend dataset with further images
  - open access to adapted annotation tool

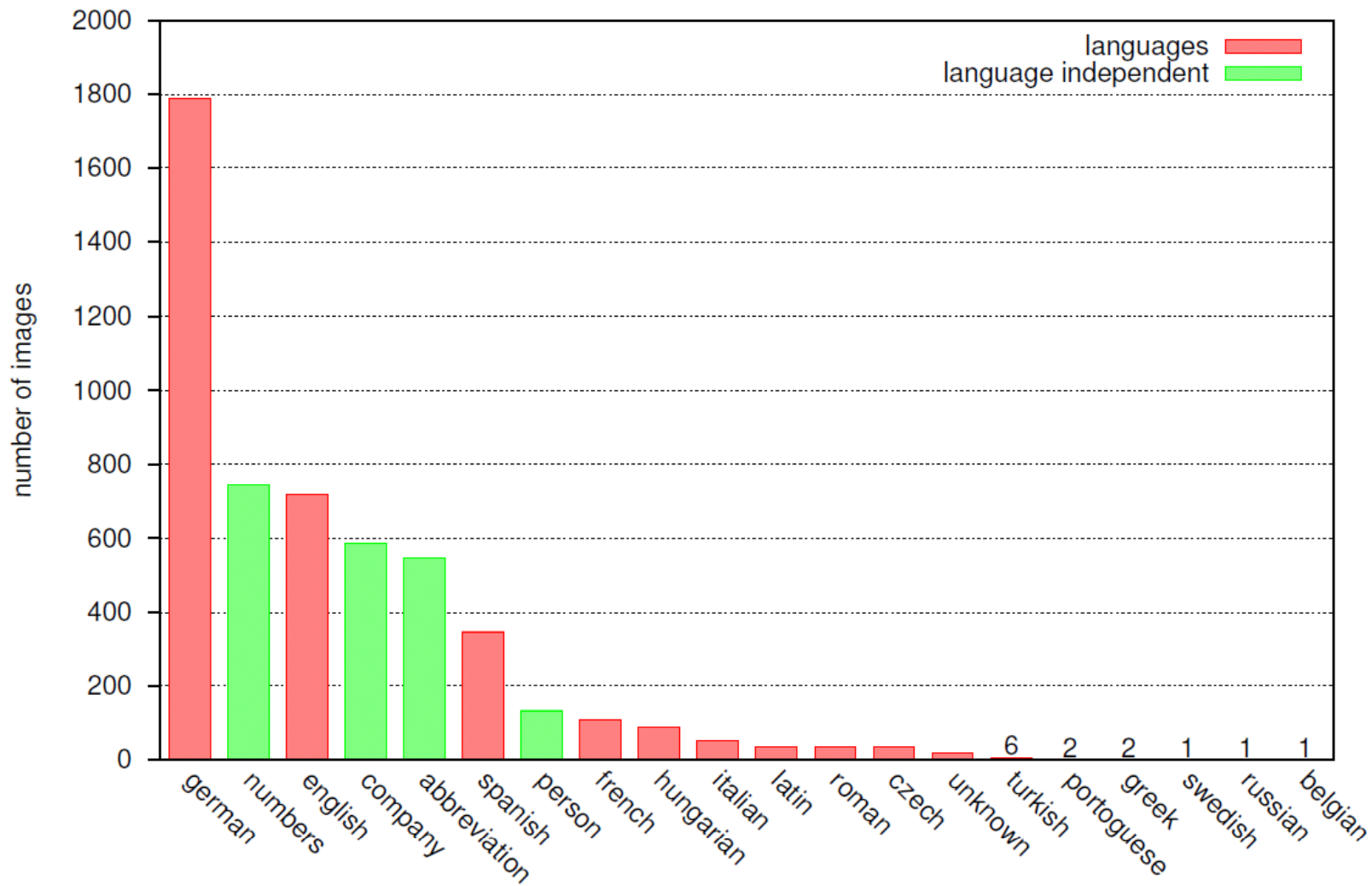
- **NEOCR Dataset.**  
<http://www6.cs.fau.de/research/projects/pixtract/neocr>
- **Chars74K Dataset.**  
<http://www.ee.surrey.ac.uk/CVSSP/demos/chars74k>
- **ICDAR 2003 Robust Reading Dataset.**  
<http://algoval.essex.ac.uk/icdar/Datasets.html>
- **Microsoft Text Detection Database.**  
[http://research.microsoft.com/en-us/um/people/eyalofek/text\\_detection\\_database.zip](http://research.microsoft.com/en-us/um/people/eyalofek/text_detection_database.zip)
- **Street View Text Dataset.** <http://vision.ucsd.edu/~kai/svt>



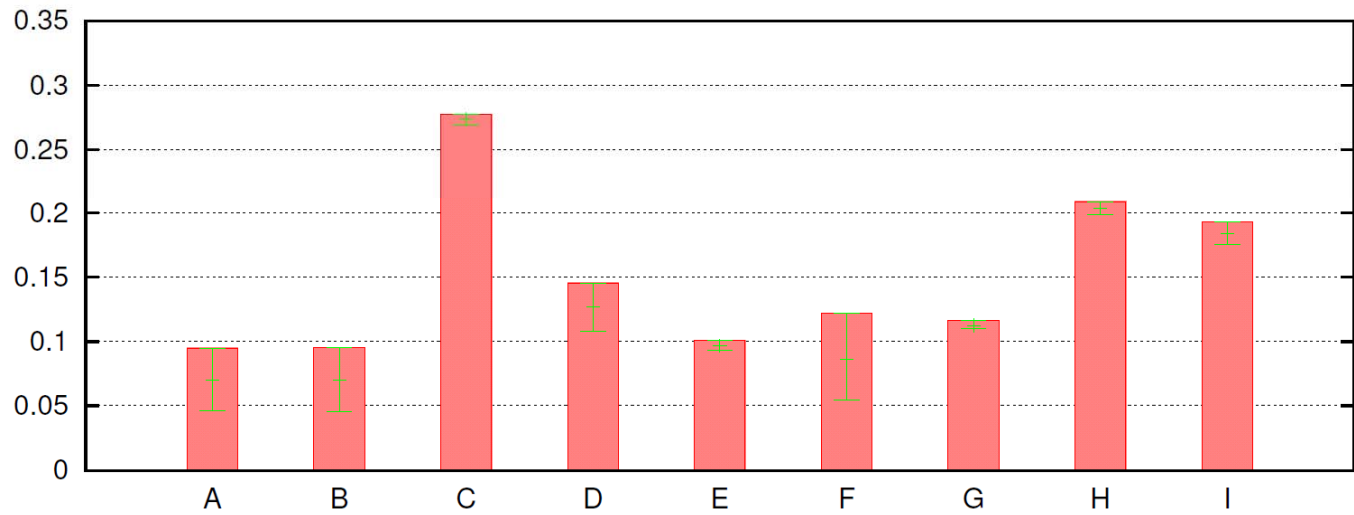








**distorted**



**straightened**

