* *Pintura de Rajib Roy, Case Investigation - 2012*

Digital Forensics

MO447 / MC919

Prof. Dr. Anderson Rocha

Microsoft Research Faculty Fellow
Affiliate Member, Brazilian Academy of Sciences
Reasoning for Complex Data (Recod) Lab.

anderson.rocha@ic.unicamp.br
http://www.ic.unicamp.br/~rocha

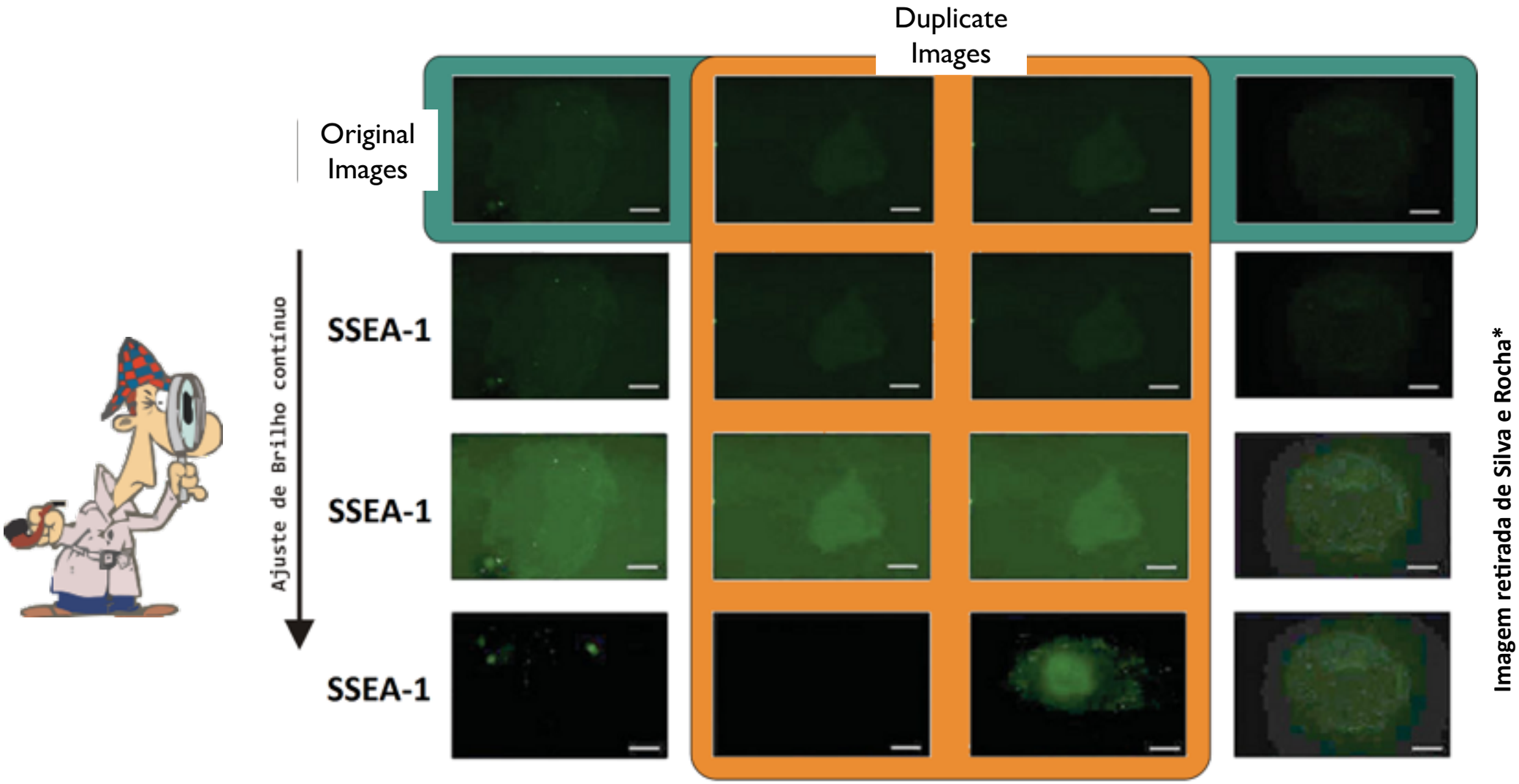
Reasoning for Complex Data (RECOD) Lab.
Institute of Computing,
University of Campinas (Unicamp)

Av. Albert Einstein, 1251 – Cidade Universitária
CEP 13083-970 • Campinas/SP – Brasil

RECOD

Motivation

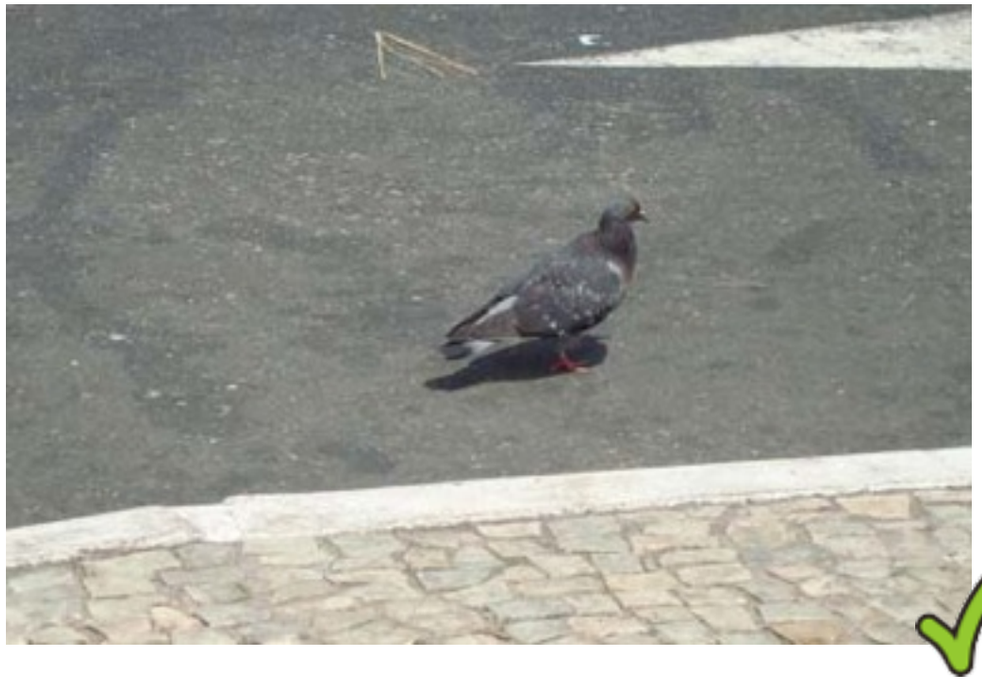
Frauds in Science



* *Silva EA, Rocha A. Digital document forensics: beyond the human vision. Saúde, Ética & Justiça. 2011;16(1):1-8.*

Cloning or Copy-Pasting

- Effortless
- How?
  i. Copy one image segment
  ii. Apply a transformation
  iii. Paste the transformed segment in a different image region
  iv. Apply global operation to the image.
- What's the potential?




MCAT Online
Science Review
FREE!
SOS


MCAT Online
Science Review
FREE!
SOS

Rotation

Scaling

Horizontal Mirroring

Vertical Mirroring

Smoothing

Global Operations

- Compression
- Noise adding
- Relighting
- etc

//

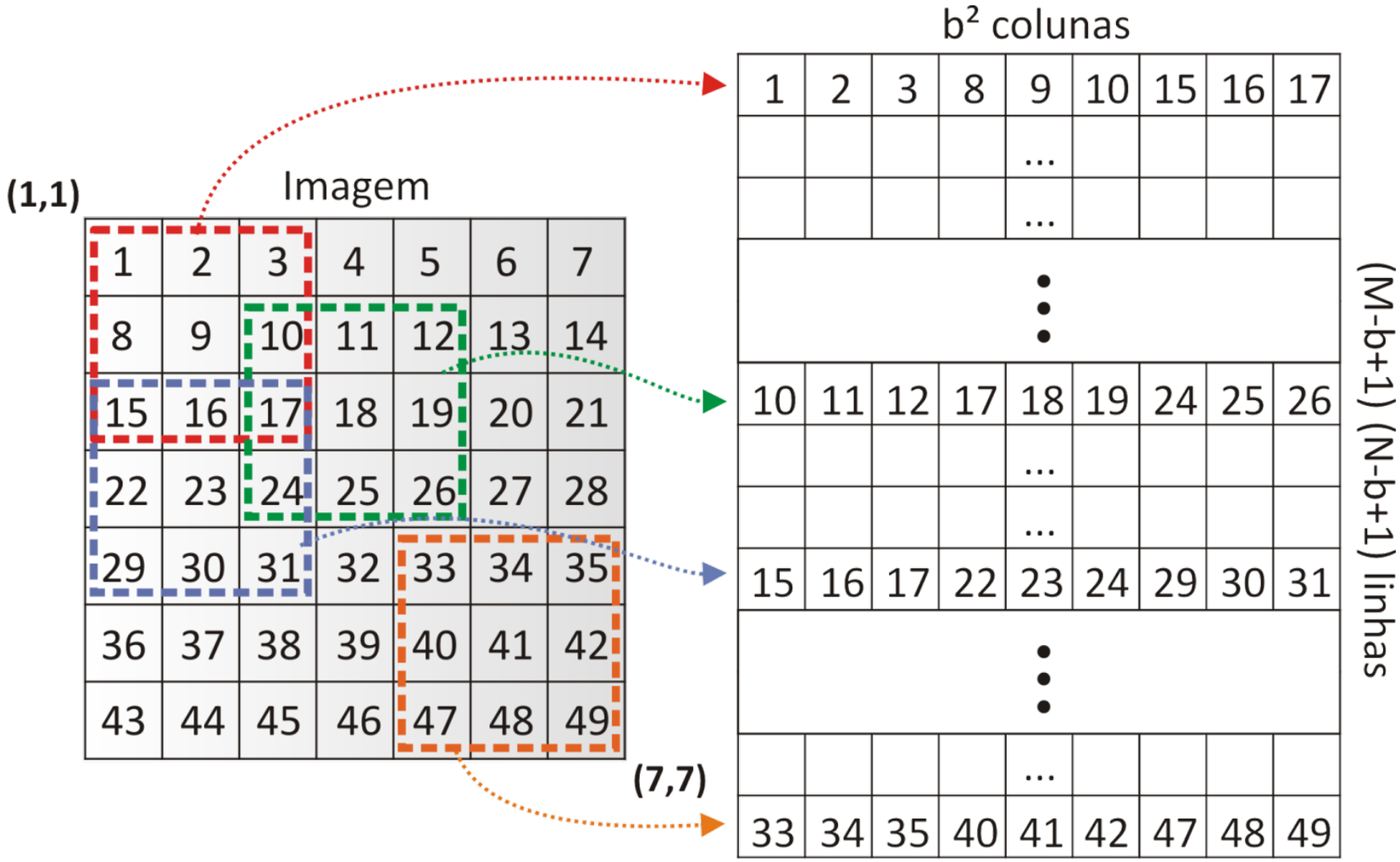
Existing Solutions

Existing Solutions

- Exact Matching by Fridrich et al. [1]
- This strategy is the building block of virtually all methods in the literature
- Extended to Robust Matching via DCT transform of blocks (Robust to compression)
- Popescu and Farid [2], PCA transformation of blocks (Robust to compression)

[1] *Jessica Fridrich, David Soukal, and Jan Lukas. Detection of copy-move forgery in digital images. In Digital Forensic Research Workshop (DFRWS), Cleveland, USA, 2003.*
[2] Alin C. Popescu. Statistical Tools for Digital Image Forensics. PhD thesis, Dept. of Computer Science { Dartmouth College, Hanover, USA, December 2004.

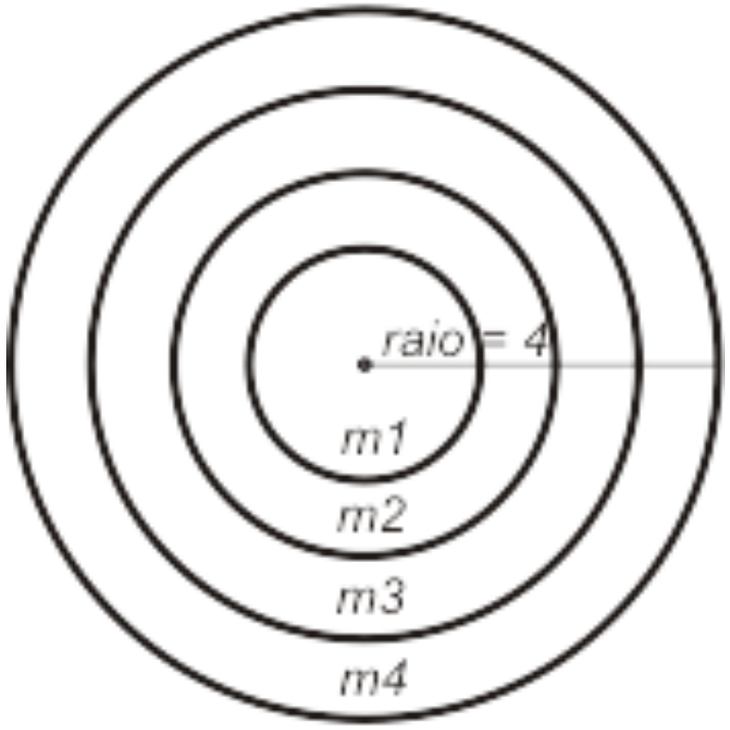
Exact-Matching Solution [1]



[1] *Jessica Fridrich, David Soukal, and Jan Lukas. Detection of copy-move forgery in digital images. In Digital Forensic Research Workshop (DFRWS), Cleveland, USA, 2003.*

Evolution of existing methods

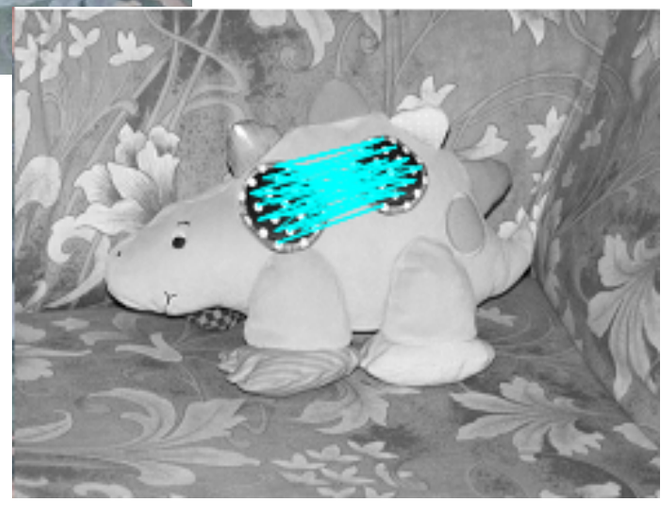
- Circular blocks by Wang et al. [1]
- Each block described as the average value of pixels inside the block



[1] J*unwen Wang, Guangjie Liu, Hongyuan Li, Yuewei Dai, and Zhiquan Wang. Detection of image region duplication forgery using model with circle block. In IEEE Intl. Conference on Multimedia Information Networking and Security (MINES), pages 25-29, 2009.*

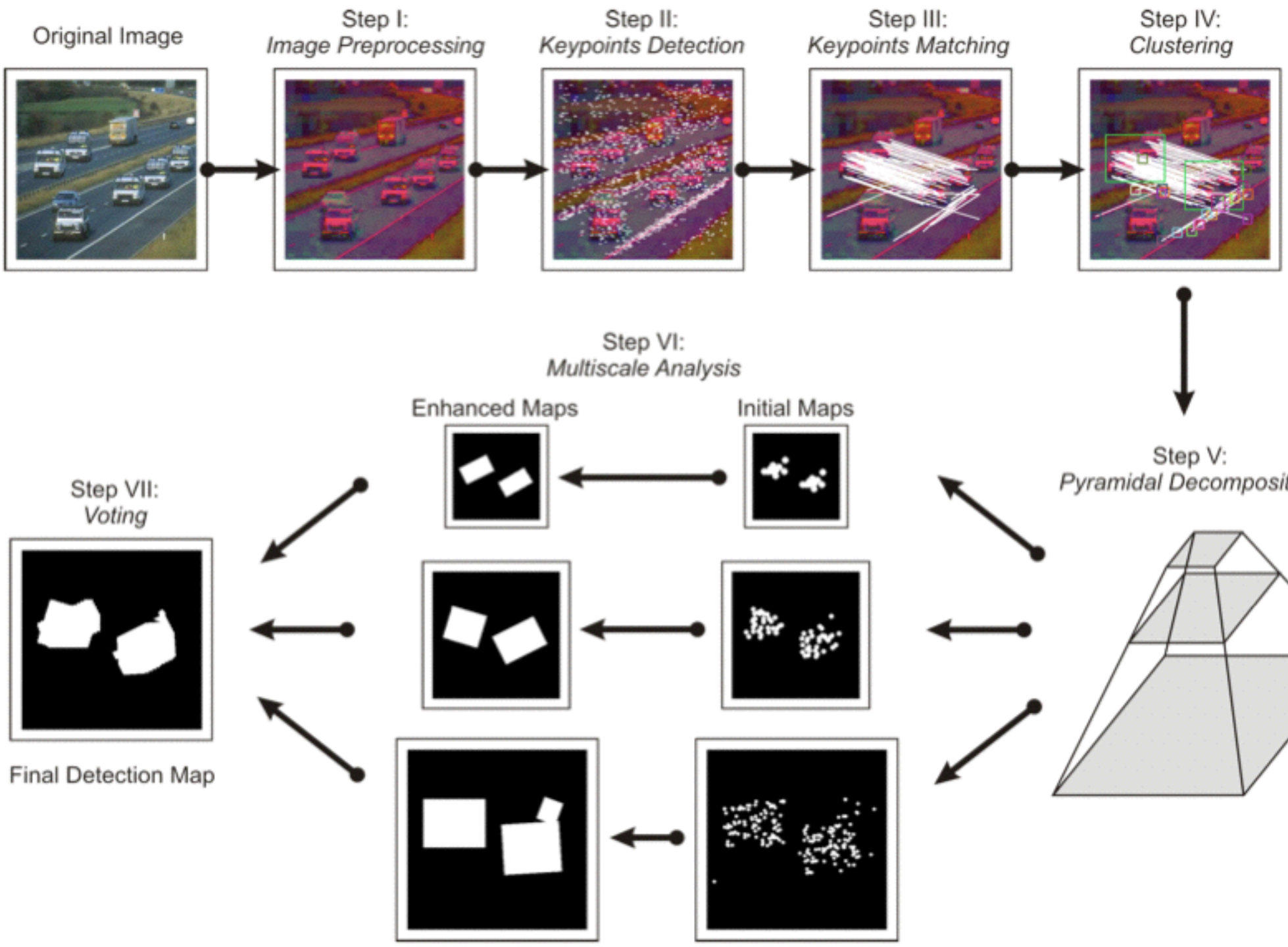
Evolution of existing methods

- Interest points by Xu et al. [1]
- Speeded Up Robust Features
- Points robust to scale, rotation, smoothing
- Problems with small regions



[1] *Bo Xu, Junwen Wang, Guangjie Liu, Hongyuan Li, and Yuewei Dai. Image copy-move forgery detection based on surf. In IEEE Intl. Conference on Multimedia Information Networking and Security (MINES), pages 889-892, 2010.*

Going deeper into copy-move forgery detection: exploring image telltales via multi-scale analysis and voting processes


Original Image
Step I:
Image Preprocessing
Step II:
Keypoints Detection
Step III:
Keypoints Matching
Step IV:
Clustering
Step VI:
Multiscale Analysis
Enhanced Maps
Initial Maps
Step V:
Pyramidal Decomposition
Step VII:
Voting
Final Detection Map

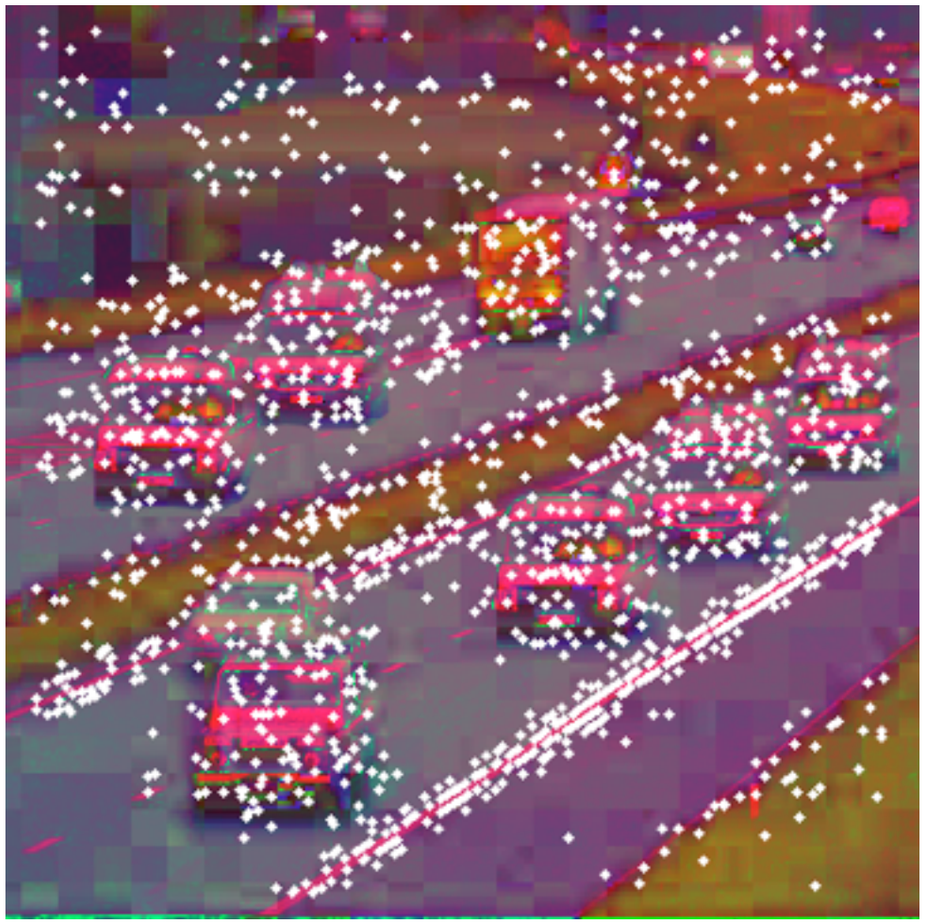
Step 1: Pre-Processing

- HSV conversion
  - Reduces false positives in regions with high homogeneity (e.g., paved regions, blue sky, textureless regions etc.);
  - When evaluating regions on the HSV domain, it is possible to highlight the difference between naturally similar regions which are not really cloned, preventing the occurrence of some mismatches.
- Normalization
  - Channels with distinct *Dynamic Ranges* (S,V ϵ [0, 255] e H ϵ [0, 179]).

$$p'_H = \frac{p_H - min(HSV)}{max(HSV) - min(HSV)}$$

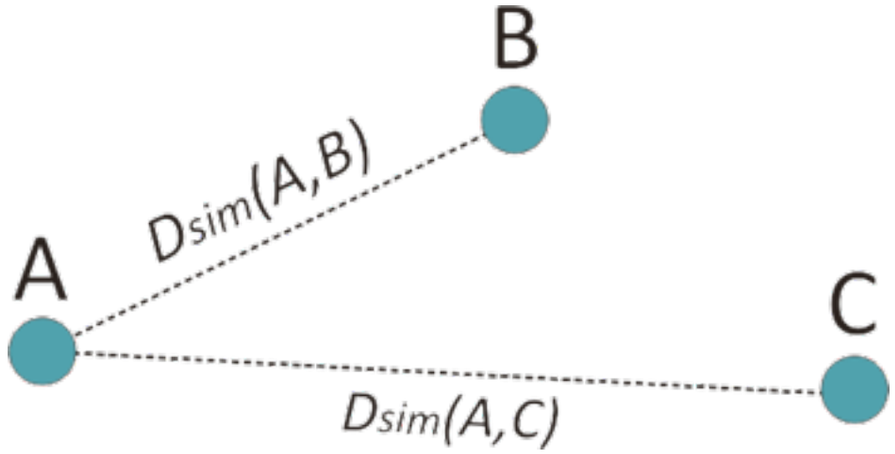
Step 2: Keypoint Detection

- SURF

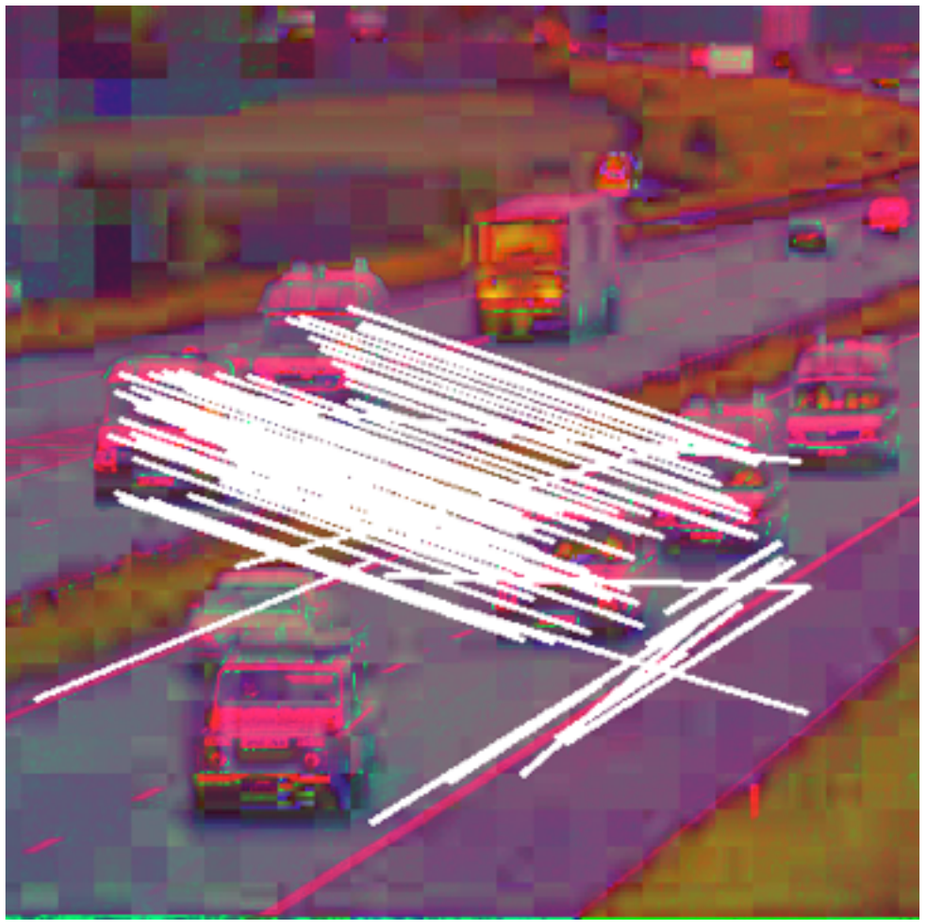


Step 3: Keypoint Matching

- We decide the similarity score between keypoints using L2.
- *Nearest Neighbor Distance Ratio* (NNDR) heuristic

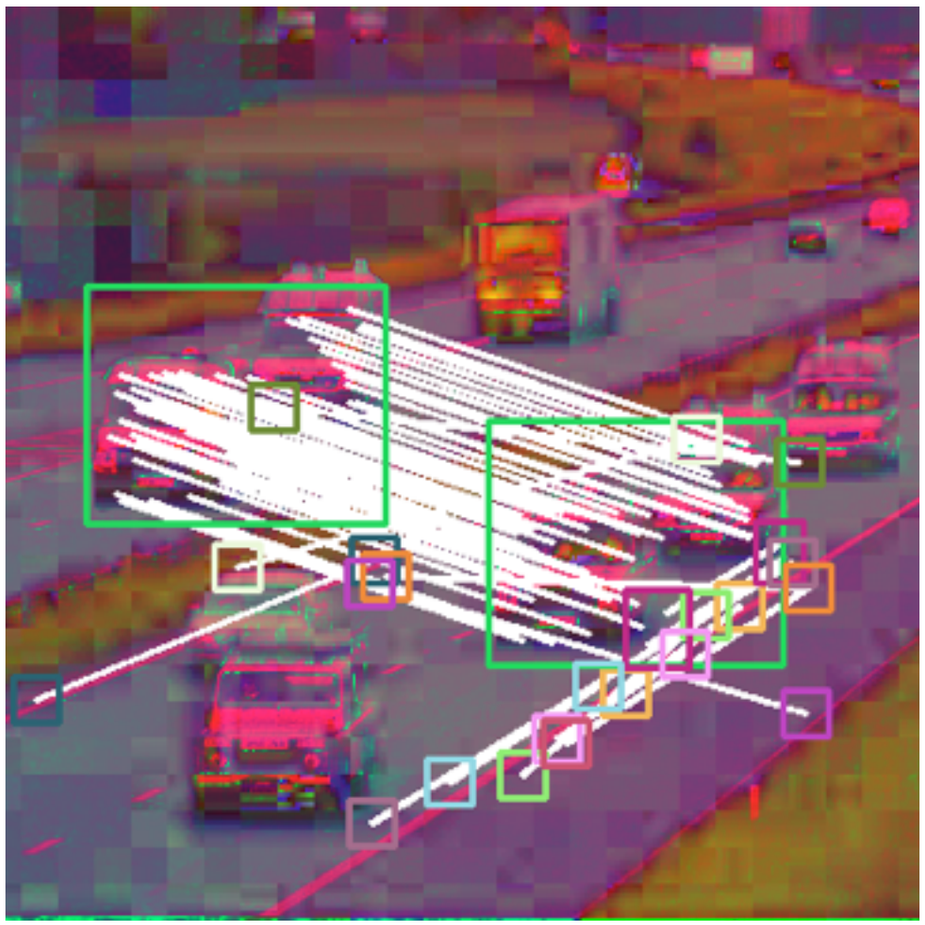


$$score = \frac{D_{sim}(A, B)}{D_{sim}(A, C)}$$



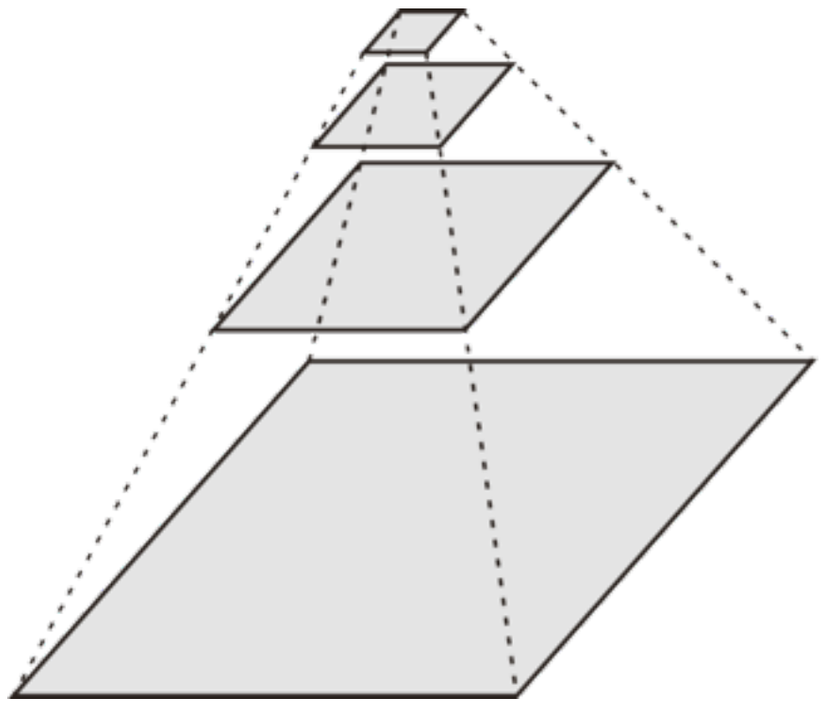
Step 4: Grouping/Clustering

- Reduces the search space for duplicate segments
- Geometric constraints:
  - Physical proximity of the points;
  - Angular consistency



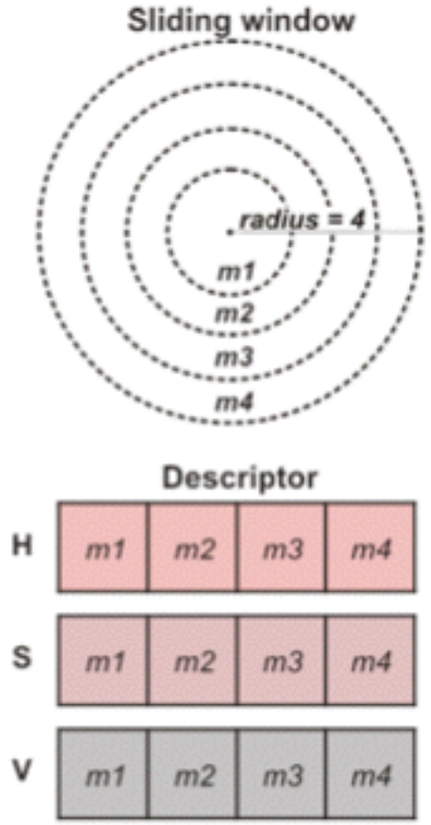
Step 5: Pyramidal Decomposition

- Representing an image in the scale-space
- Successive reductions in 25%.

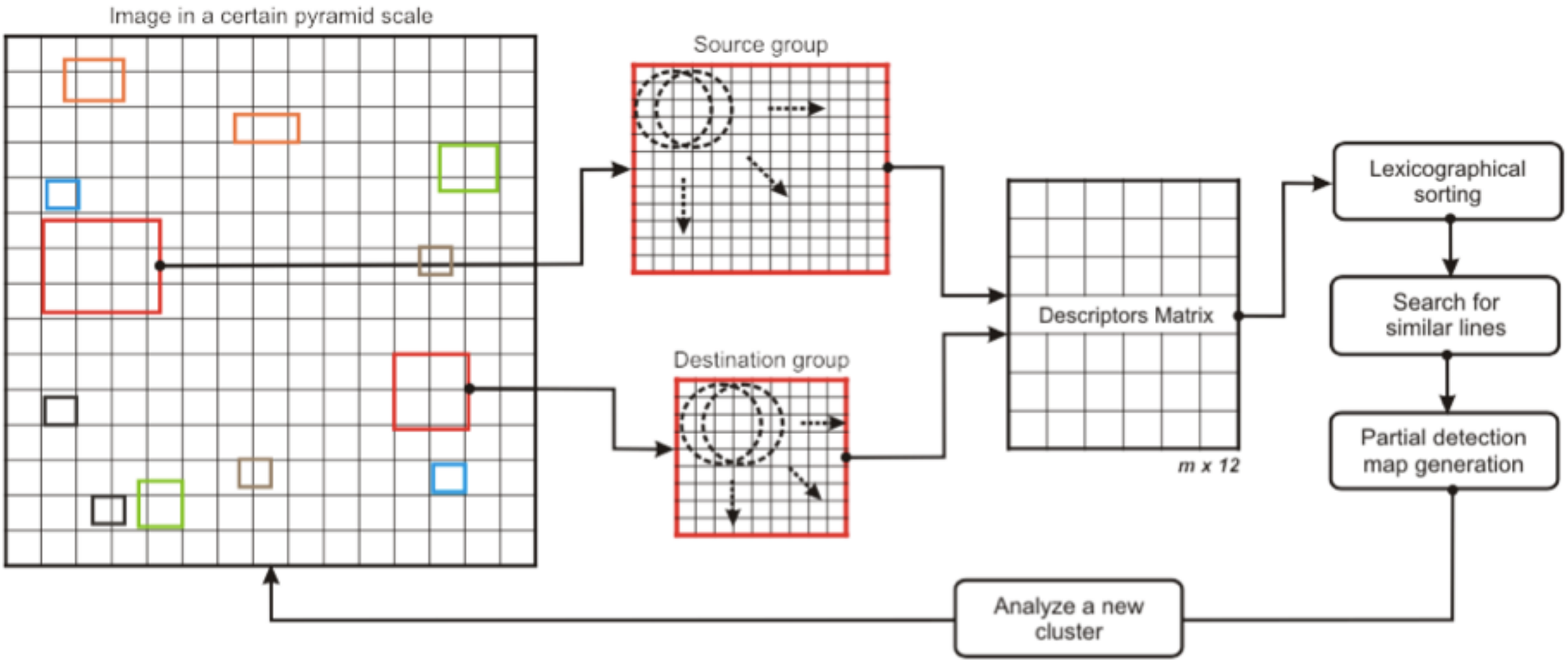


Step 6: Multiscale analysis

- Block comparison with a few differences:
  - Intra-group inspection (source and target);
  - Group re-scaling proportional to the image reductior
  - Sliding window and descriptor based on Wang et al.;
  - Block normalization
  - Improved block masks (convex hull)



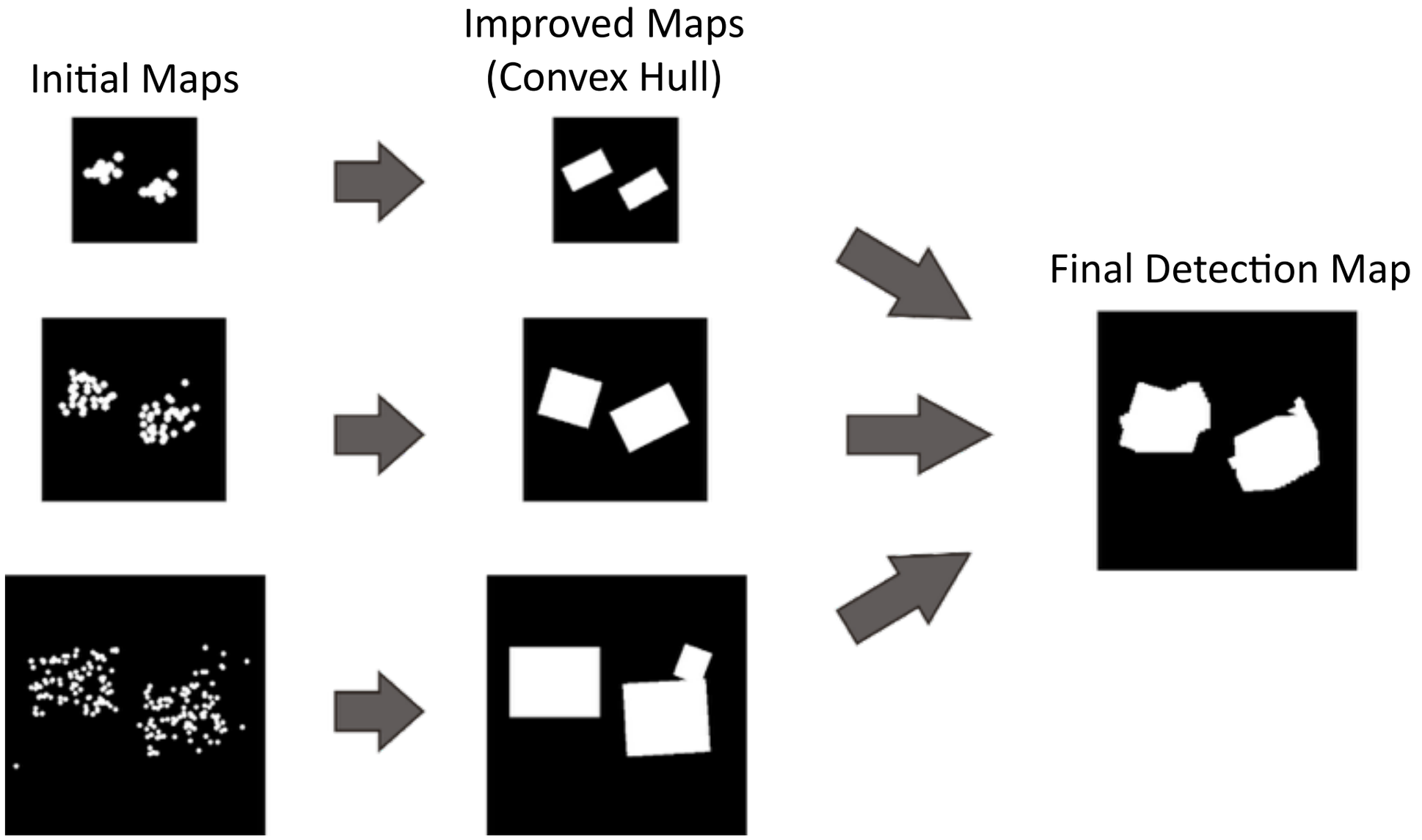
Step 6 overview



Step VI overview. In a given pyramid level, the red group is analyzed (source and destination) and its blocks descriptors are stored in a matrix to be lexicographically sorted. Then, a search for similar matrix lines is performed, which yields a partial detection map for the red group. These steps are repeated for all groups, with several updatings of the partial detection map. When all clusters have been examined, the final detection map (for that specific image) is generated. This whole process is applied separately to all scales in the pyramid.

Step 7: Voting

- Only segments compatible across scales are considered cloned



Experiments and Validation

Dataset

- Image dataset comprising 108 cloning examples:
  - Simple copy & pasting
  - Rotations
  - Scaling
  - Complex copy & pasting (all operations combined)
- All images converted to JPEG in 70, 80 and 90 quality factors

Evaluation Metrics

- *True Positive Rate* (TPR)

$$TPR = \frac{|TP|}{|R_{clone}|}$$

- *False Positive Rate* (FPR)

$$FPR = \frac{|FP|}{|R_{normal}|}$$

- *Accuracy* (ACC)

$$ACC = \frac{TPR + (1 - FPR)}{2}$$

Qualitative Analysis

Simple Cloning

Rotation -30 deg

Scaling +19%

Rotation -18 deg
Scaling -10%

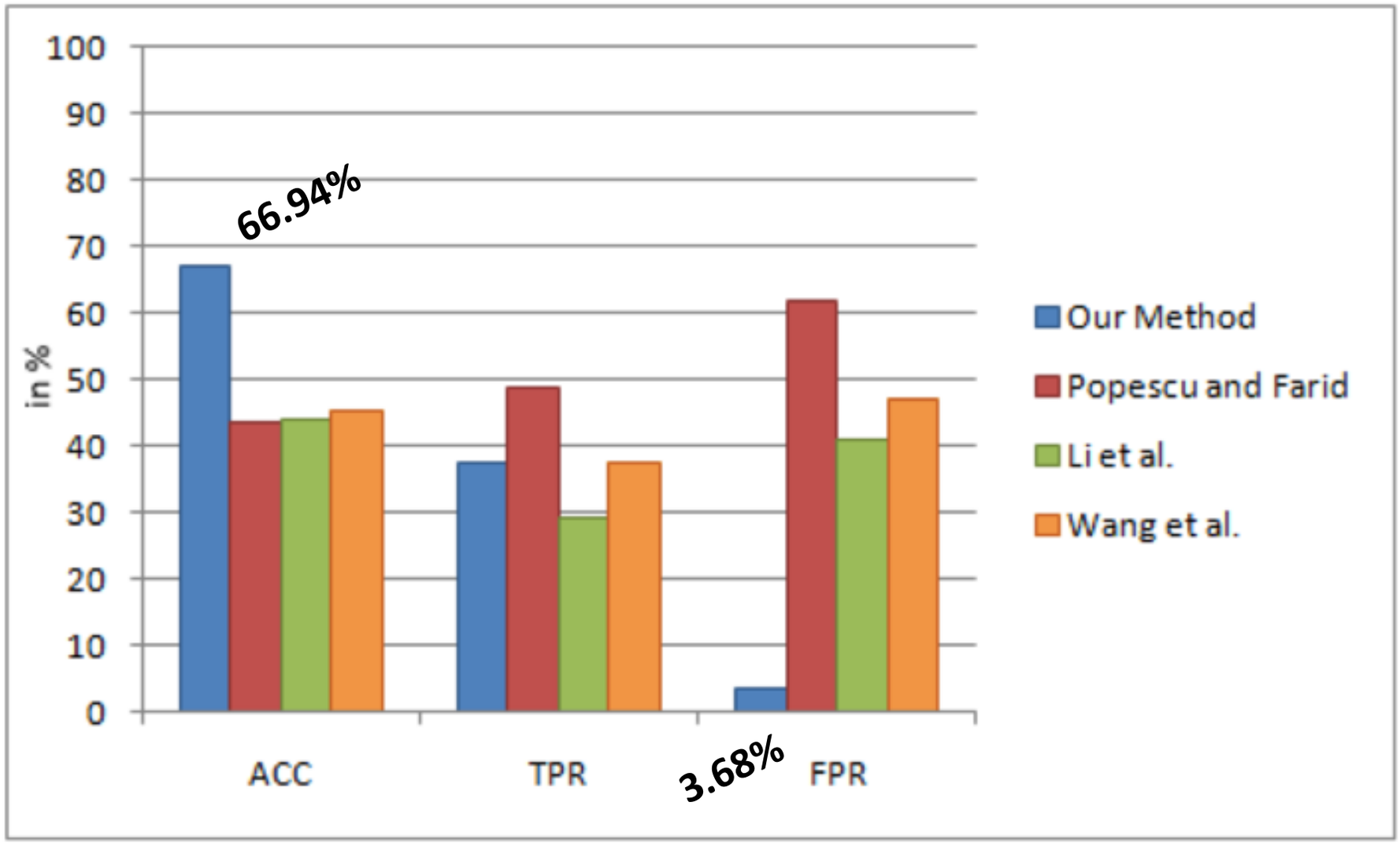
Limitations

- Regions too small
- Very homogeneous regions
  - Keypoint location has problems
- Not robust to mirroring
  - The method does not deal with line matching crossings in a robust way
  - Changes in the orientation of the matching lines can increase the number of groups

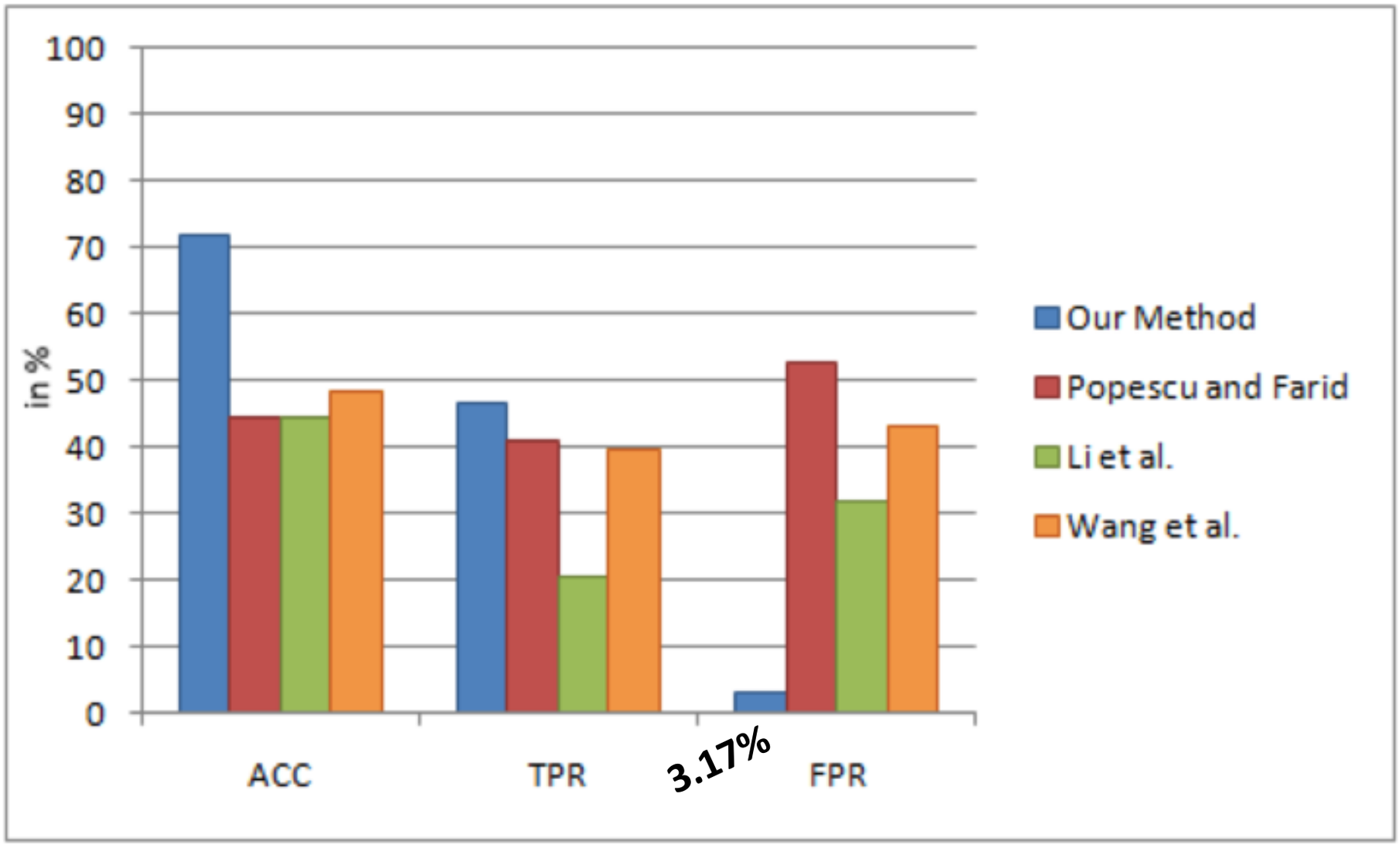
Quantitative Evaluation

- Comparison to three methods in the literature:
  - Popescu e Farid
  - Wang et al.
  - Li et al.
- Experiments with the proposed dataset

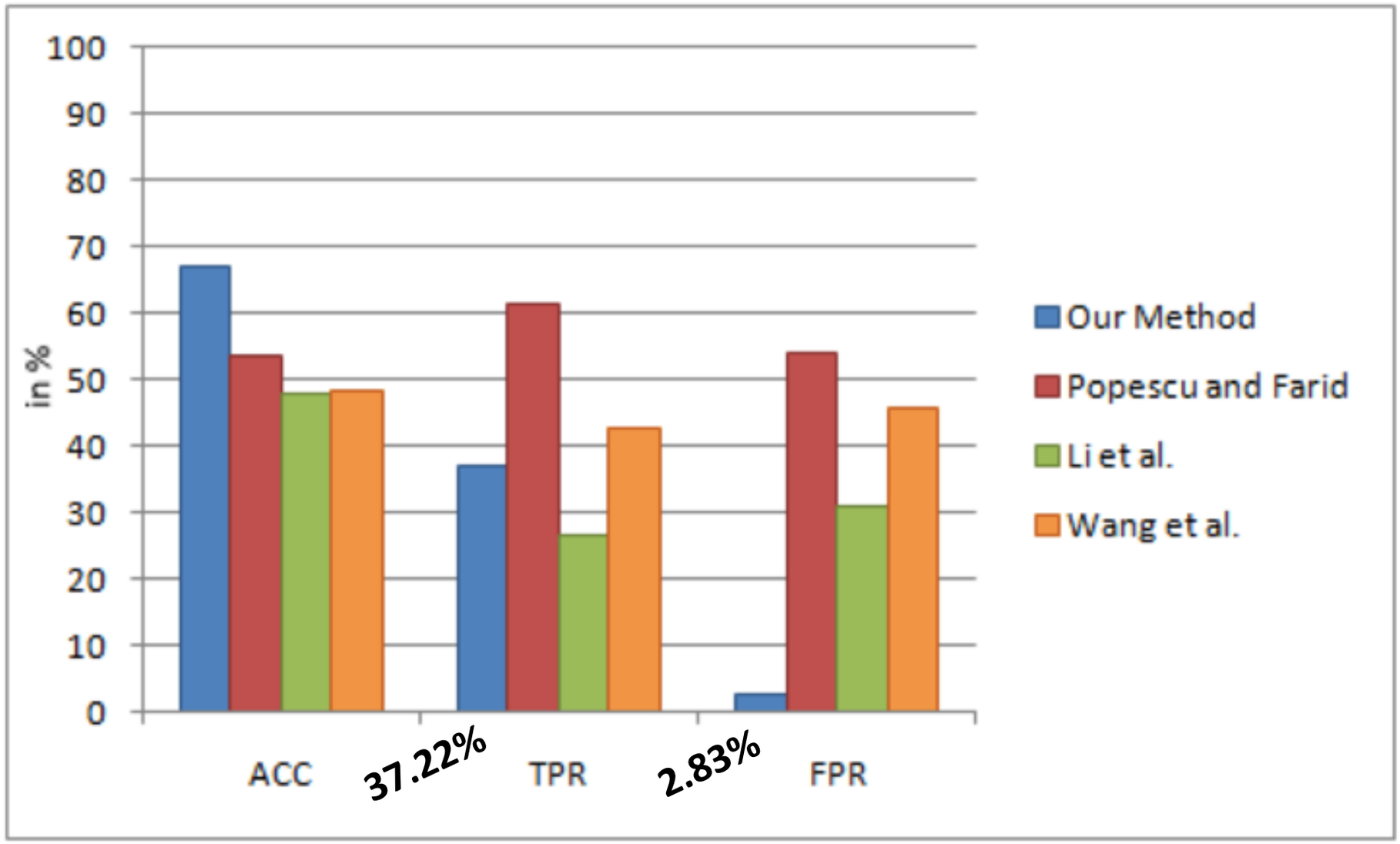
JPEG (factor 70)



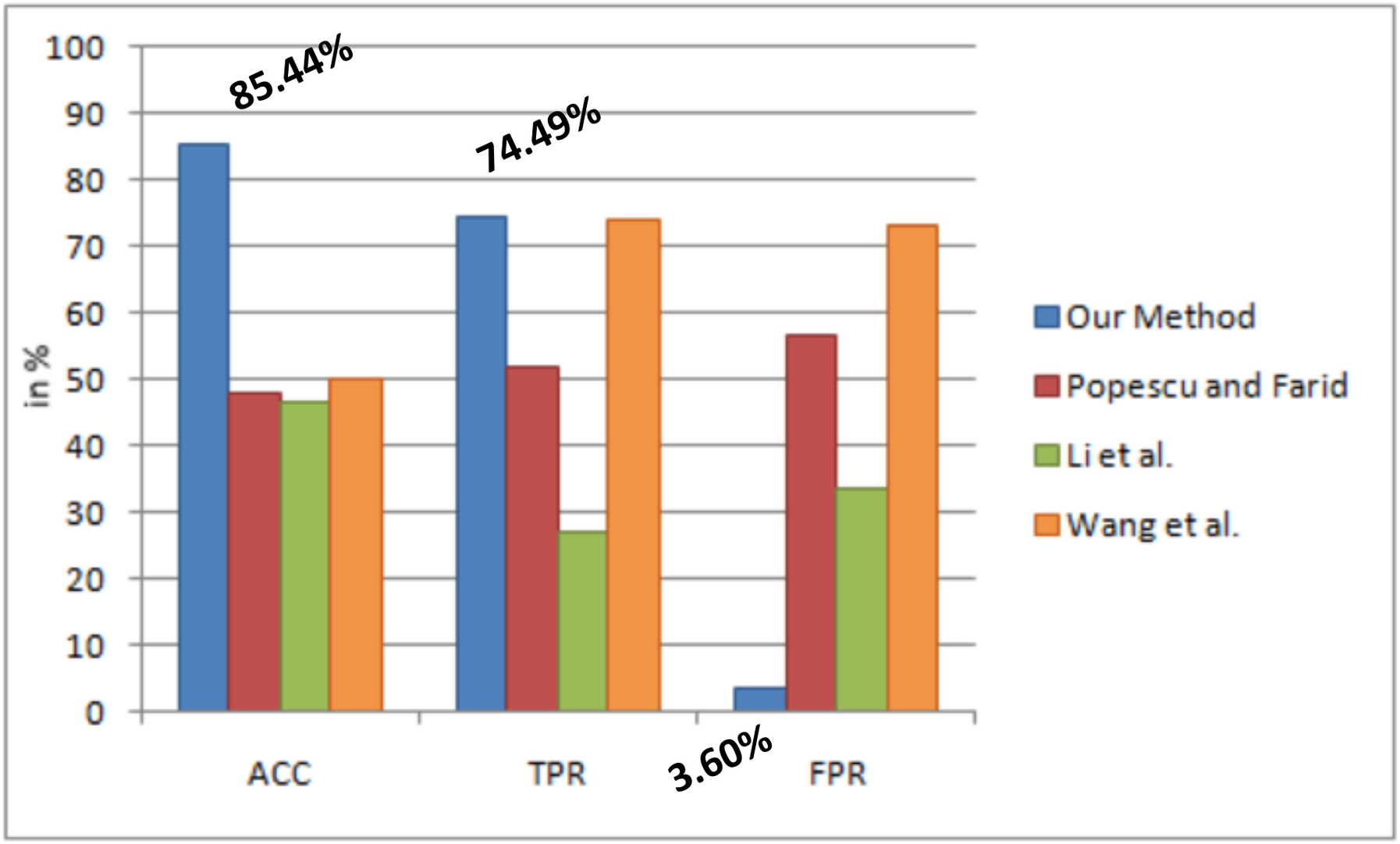
JPEG (factor 80)



JPEG (factor 90)



No Compression



Conclusions

Conclusions

- Interesting results obtained
  - (ACC de 85.44% e FPR de 3.68%)
  - Deals with combined operations
- Future investigations
  - Incorporate other image transformations
  - Alternative ways for describing the blocks
  - Image segmentation
  - *Content-Aware Fill* detection