

Motivation

Current state of the art solutions for object detection use components and parts on top of the now classic HOG+linear SVM combo [1].
Is this the best base method ?

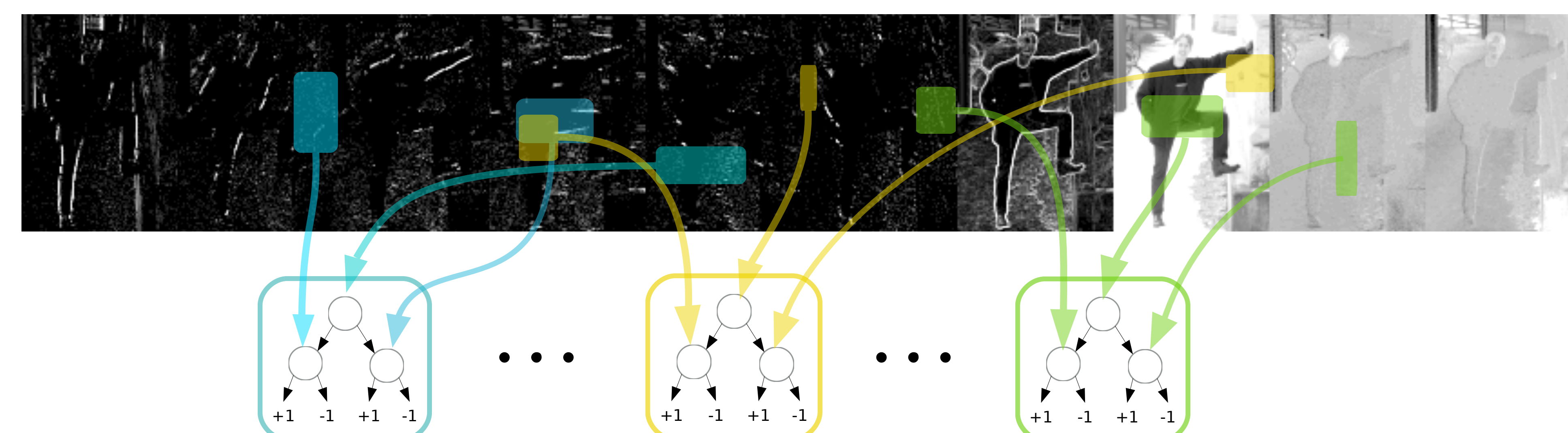
How far can we go
evaluating only one rigid template
per candidate detection ?

Common methods to improve pedestrian detection:

- Non linear SVM
- Sophisticated features
- Geometric priors
- Motion information
- Deformable models
- Deeper architectures

**We use
none of them**

We build upon: HOG & Color + Adaboost [2, 3]

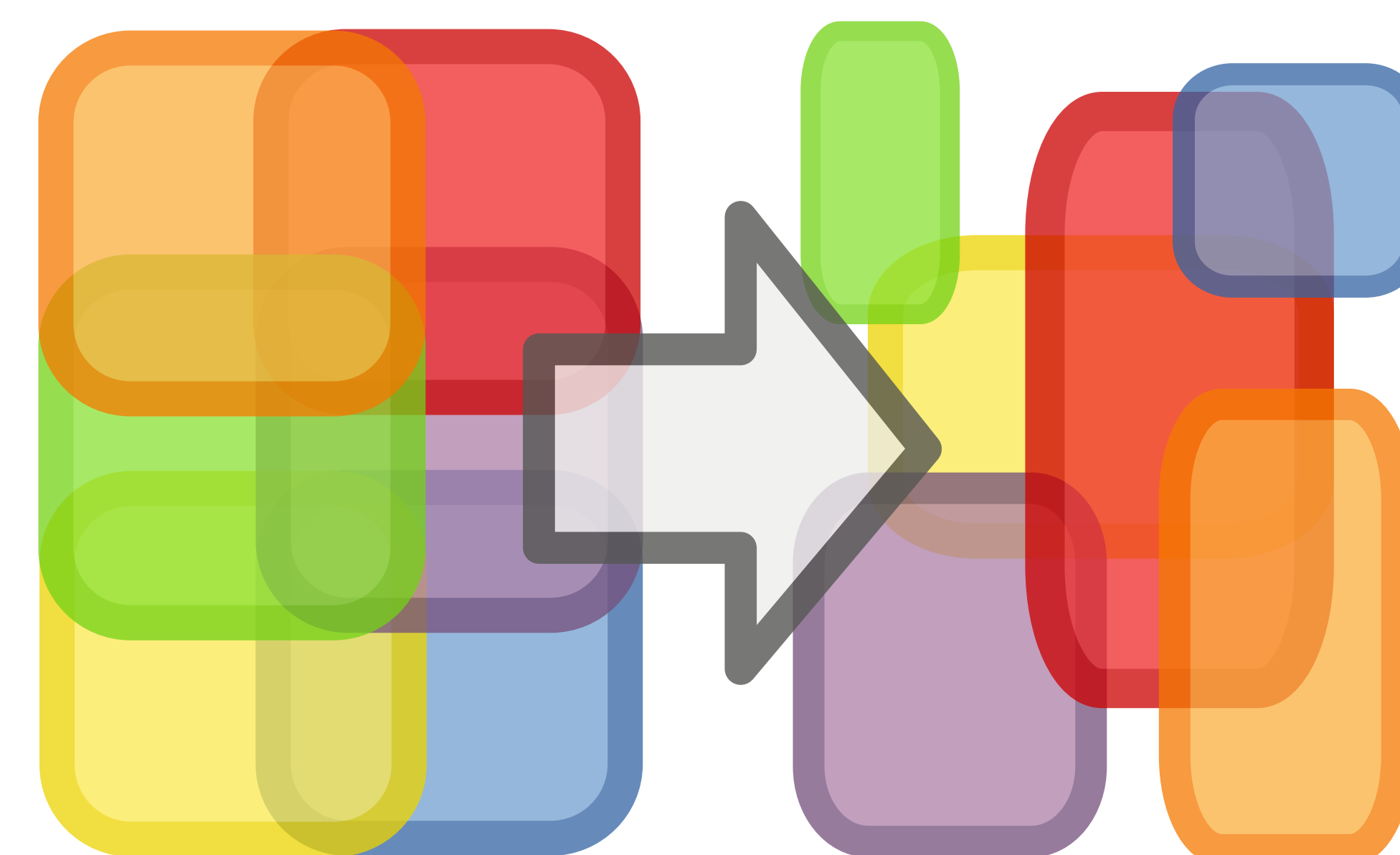


References

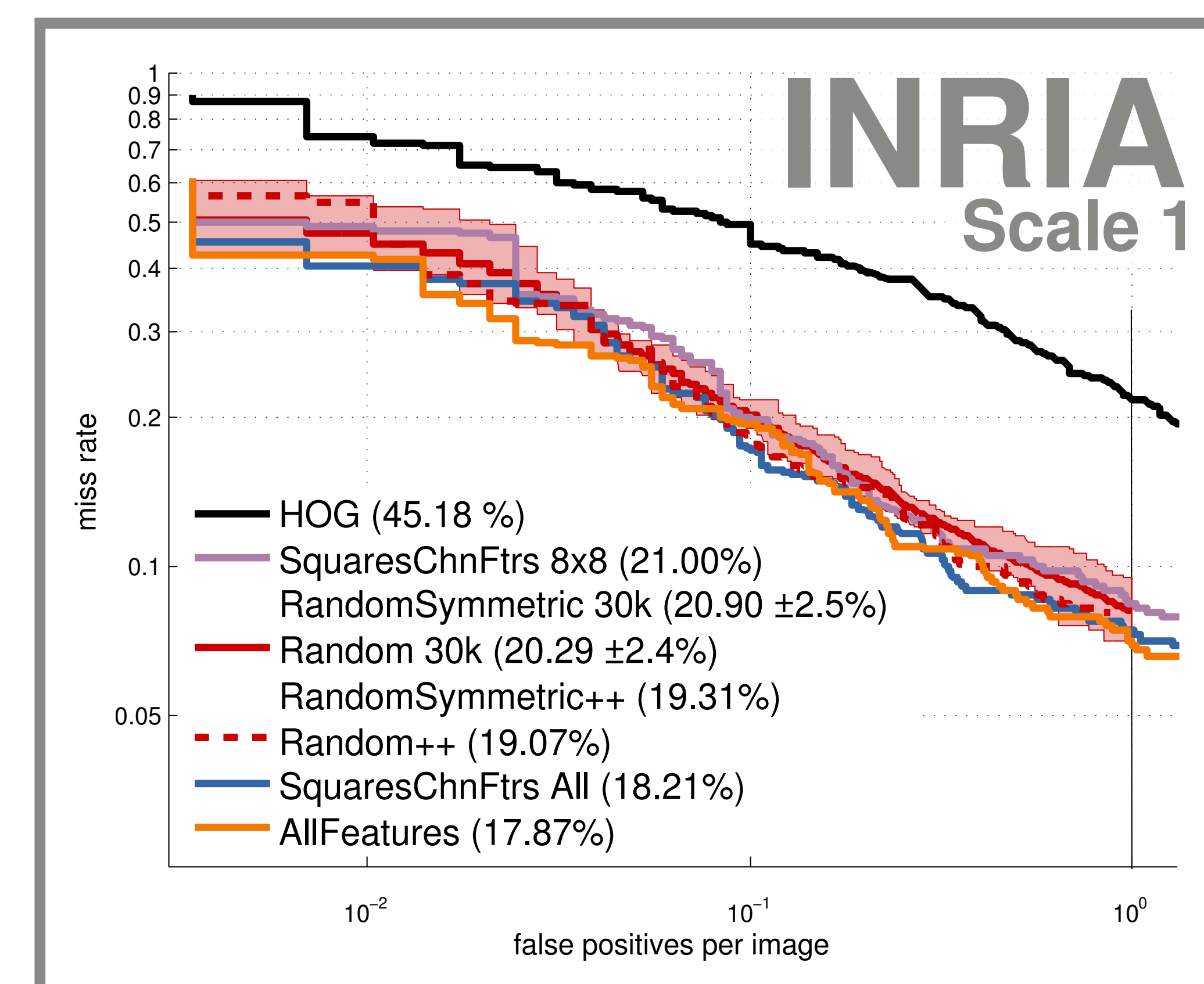
- [1] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. CVPR 2005
- [2] J. Shotton, J. Winn, C. Rother, and A. Criminisi. TextonBoost. ECCV 2006
- [3] P. Dollár, Z. Tu, P. Perona, and S. Belongie. Integral channel features. BMVC 2009
- [4] R. Benenson, M. Mathias, R. Timofte, and L. Van Gool. Pedestrian detection at 100 frames per second. CVPR 2012
- [5] A. Rizzi, C. Gatta, and D. Marini. A new algorithm for unsupervised global and local color correction. PRL, 2003

Three insights to improve detection quality

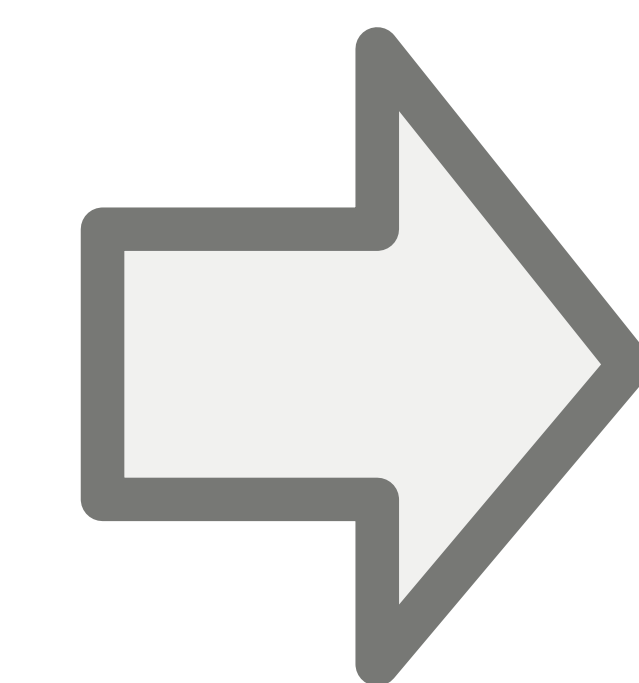
Irregular cells are better than a regular grid



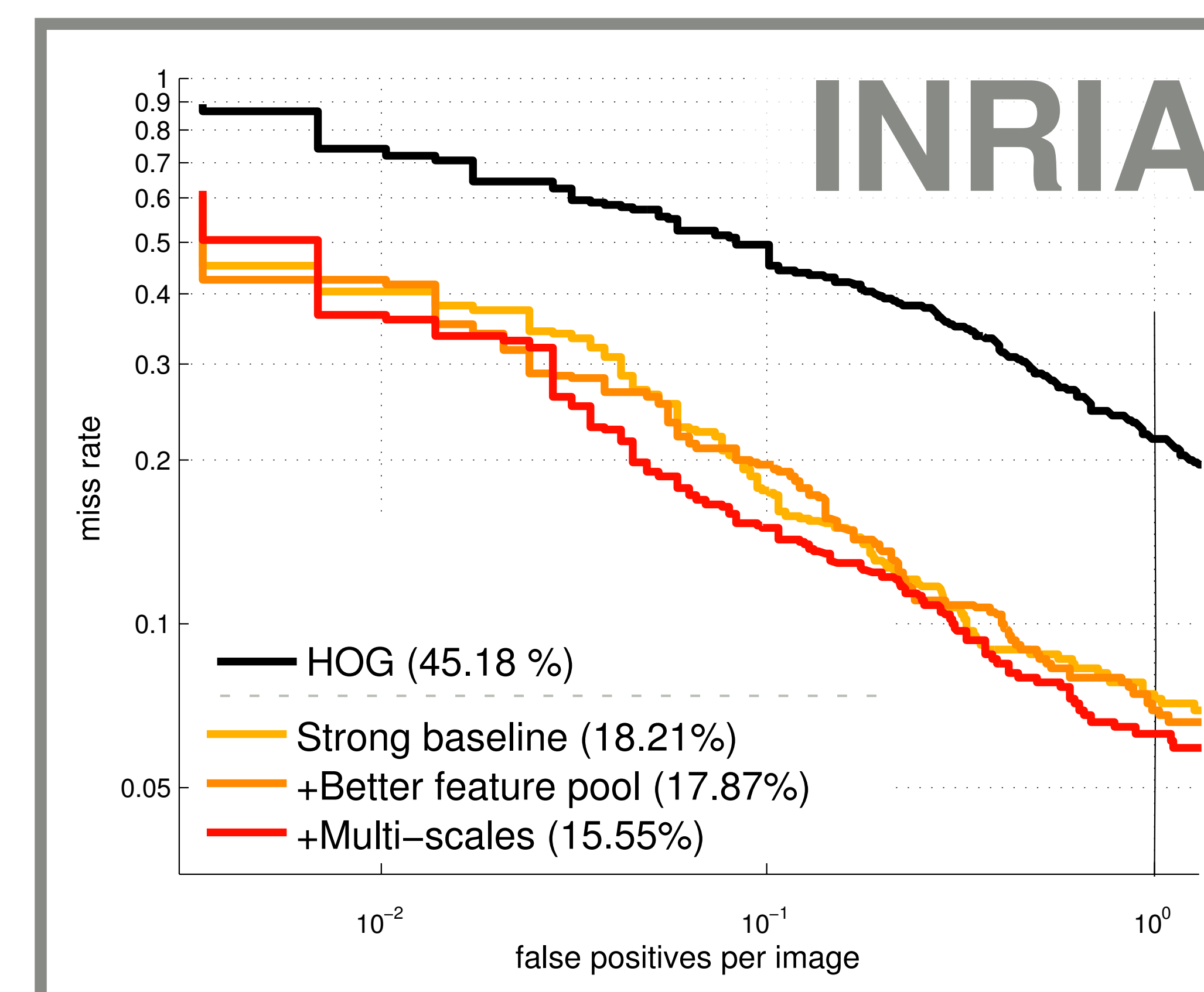
Better representation
of the class shape



Multi-scales models are better than single scale models



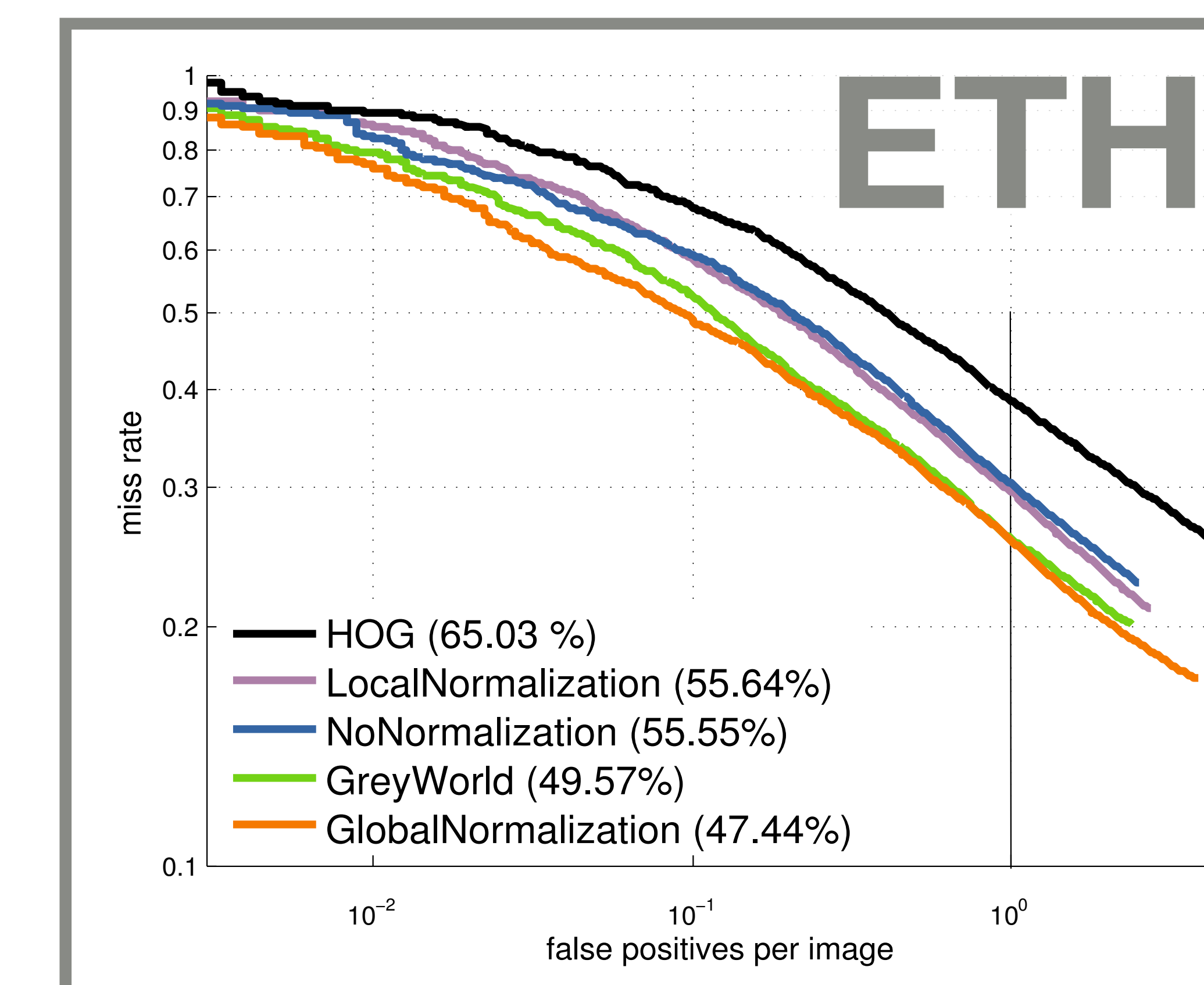
We exploit better
the available pixels



Global normalization is better than local normalization

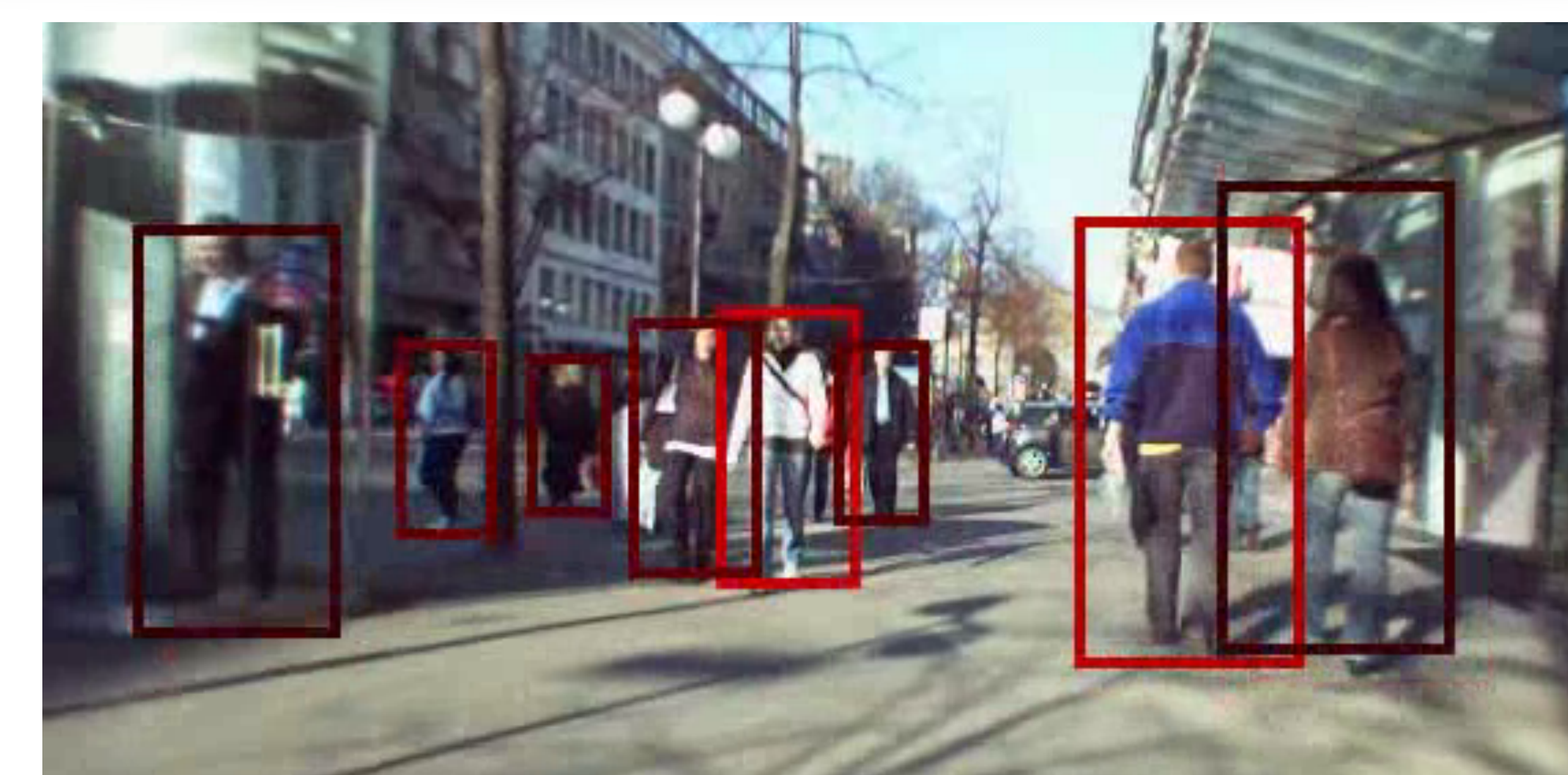


Most effective choice
in our architecture



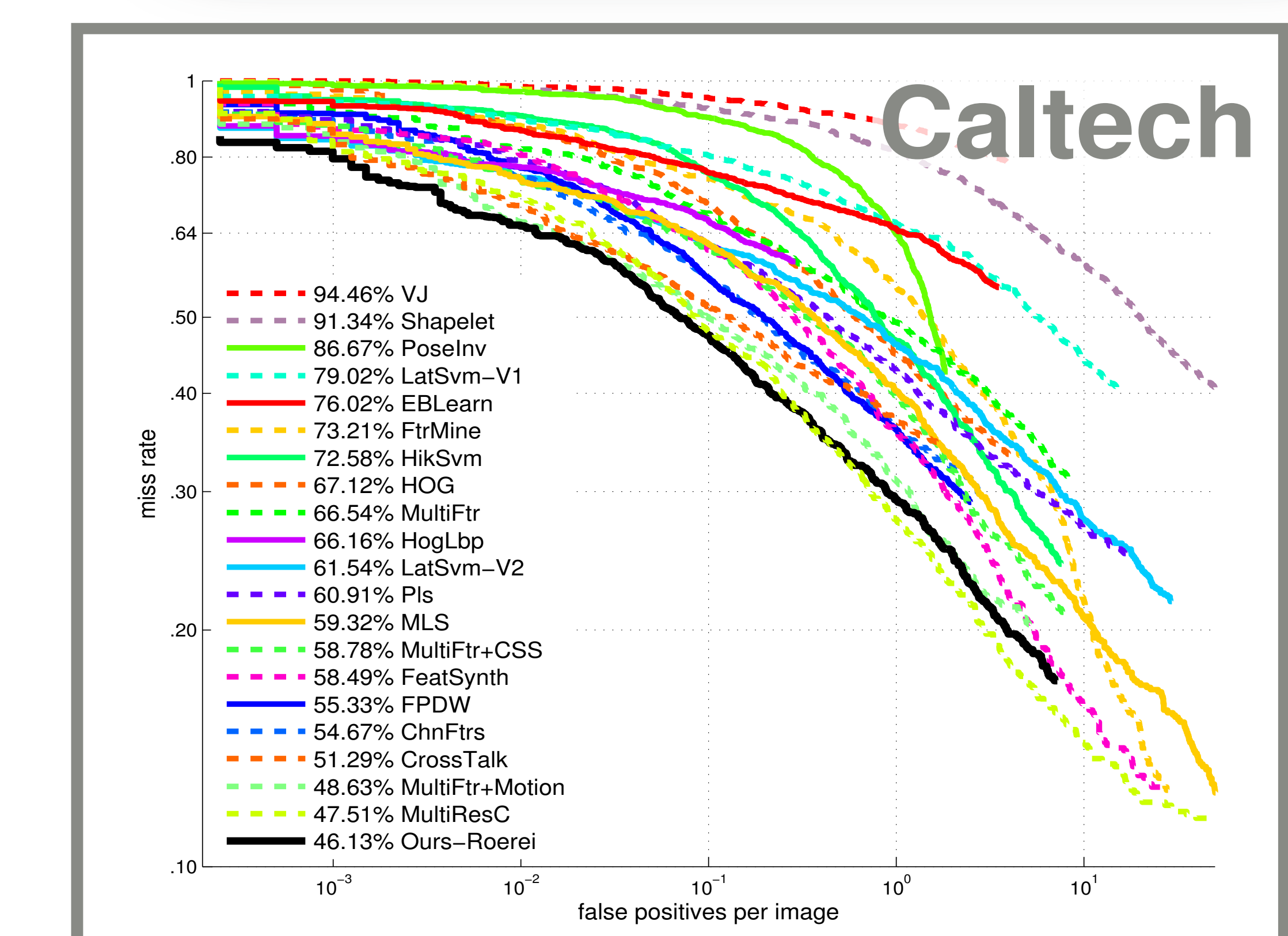
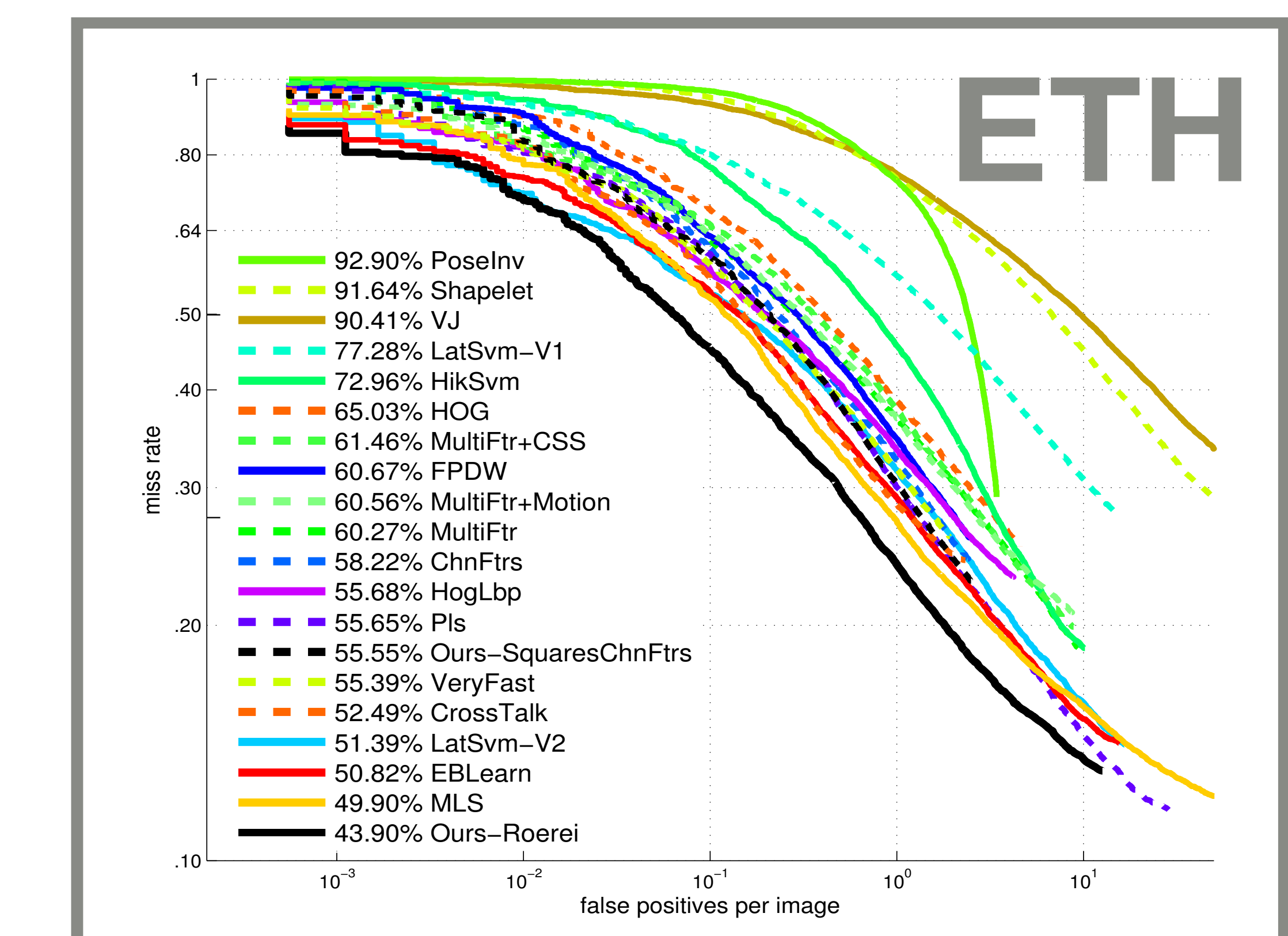
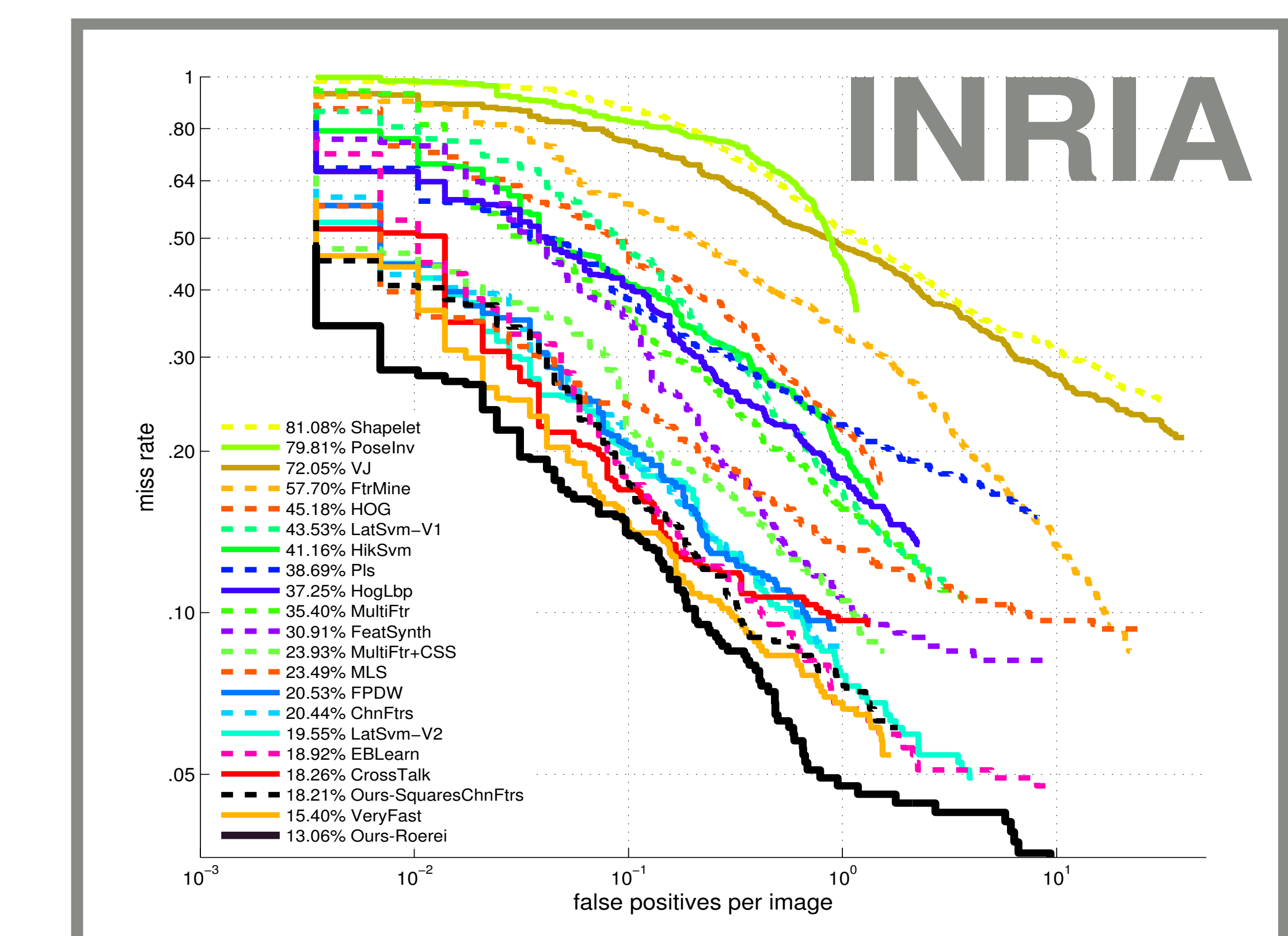
Some ideas that have not yet paid off for us:

- Stronger weak learners
- Better boosting method
- Larger training set



Detector aspect	Average miss-rate	
	INRIA	ETH
Strong baseline	18.21%	55.55%
+ AllFeatures	17.87%	55.50%
+ Multi-scales	15.55%	53.17%
+ GlobalNormalization	13.06%	43.90%
= Roerei detector	13.06%	43.90%
HOG+SVM	45.18%	65.03%
Previous best, VeryFast/MLS	15.40%	49.90%

We obtain top results



HOG is dead, long live HOG!
2005 [1] 2013