It Takes Two to Tango: Cascading Off-the-shelf Face Detectors

THE UNIVERSITY OF QUEENSLAND AUSTRALIA ADVANCED SURVEILLANCE

Siqi Yang, Arnold Wiliem and Brian C. Lovell
The University of Queensland, Australia

The False Positive Problem

- The existing face detectors, including deep learning based methods, still generate false alarms (false positives)
- Face detection is an initial step of many facial analysis tasks, including facial landmark localisation and face recognition. The unexpected false positives will affect the accuracy and speed of the subsequent tasks
- Any method aimed to reduce false positives has the potential to improve all existing methods' performance

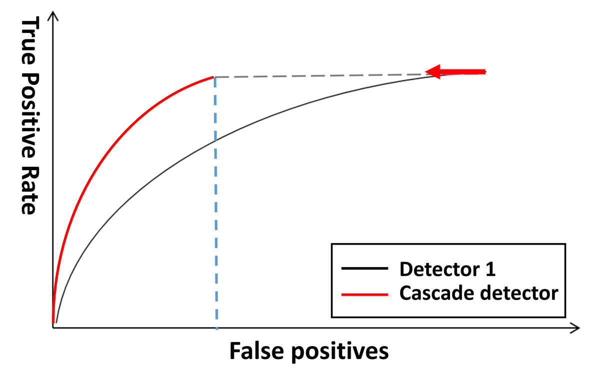
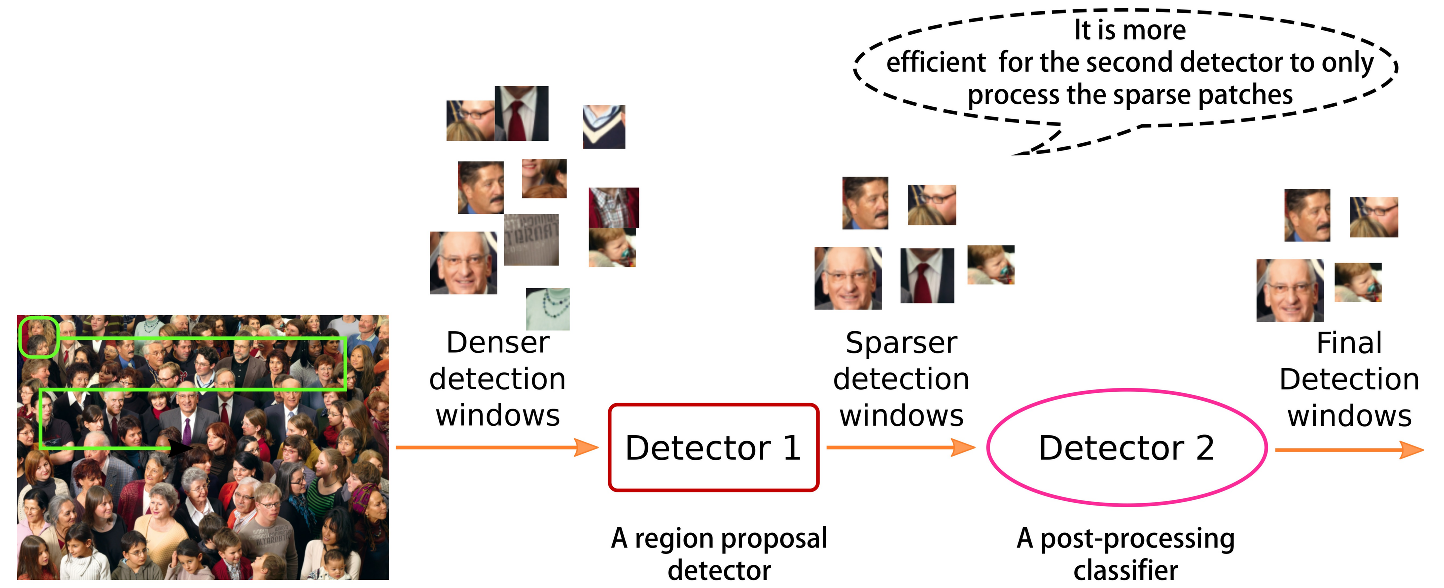


Figure 1: By cascading a second detector, a large number of false positives can be removed while the recall is well maintained. As a result, at a low number of false positives, the true positive rate can be increased significantly

Cascading Off-the-shelf Face Detectors

- The effort to train a new face detection model is enormous
- Solution: We propose to cascade two pre-trained face detectors in a two-stage framework:



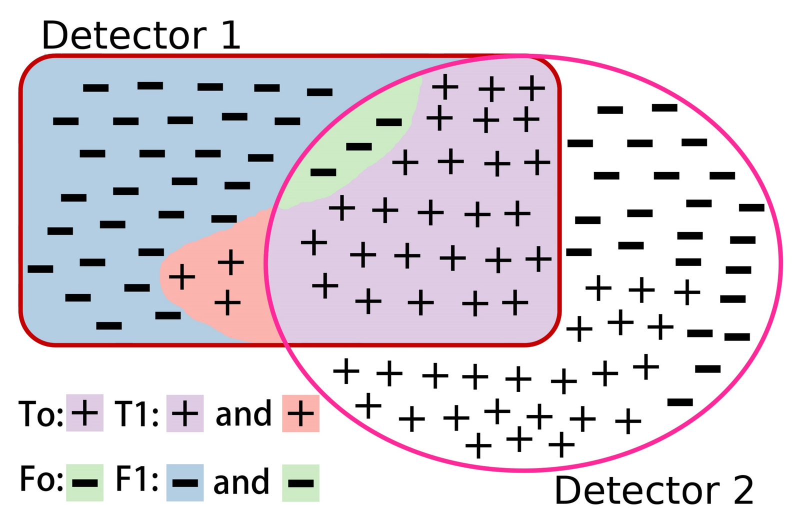
- The computational complexity of fusion-based detectors (placed in parallel) increase linearly according to the number of detectors and the overall running time is constrained by the slowest detector
- **Two questions arise:**
  1. **which two detectors should be cascaded?**
  2. **which order should they be cascaded?**

Correlation and Diversity Metrics

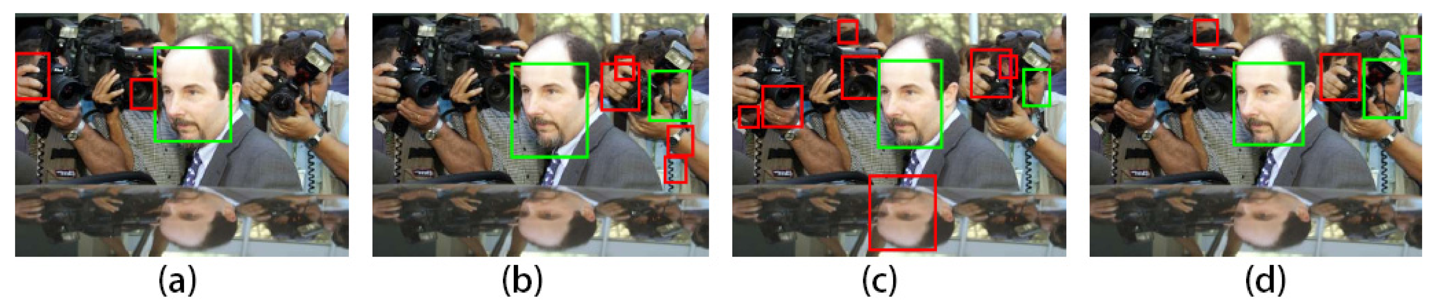
- We define the **correlation** of overlapping true positives and **diversity** of overlapping false positives:

$$c^T_{2\to 1} = \frac{|\mathcal{T}_o|}{|\mathcal{T}_1|},$$

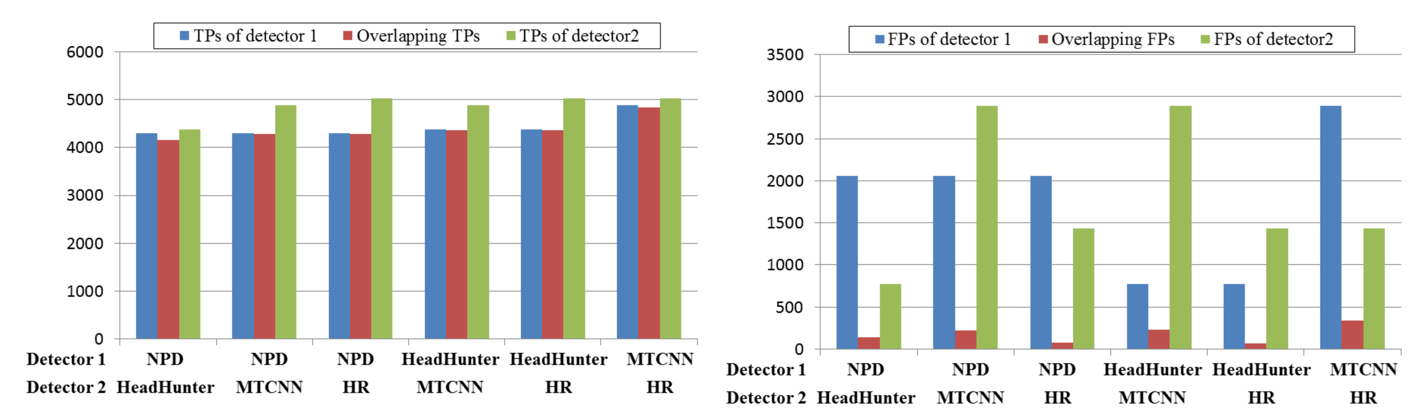
$$d^F_{2\to 1} = 1 - \frac{|\mathcal{F}_o|}{|\mathcal{F}_1|}.$$



- Detections from four different face detectors on the FDDB dataset [1]: (a) NPD [2], (b) HeadHunter [3], (c) MTCNN [4] and (d) HR [5]. Green: true positives, red: false positives



- The distribution of the overlapping true positives (left) and false positives (right) between four face detectors. Only a small number of false positives are detected by both detectors, whereas a majority of true positives overlap.



Proposed Cascade Properties

1. A high correlation of TPs, $c^T_{2\to 1} \approx 1$,
2. A high diversity of FPs, $d^F_{2\to 1} \approx 1$,
3. Faster detector in the first stage to achieve an overall fast speed

Experiments

- Runtime analysis on the FDDB dataset [1]:

Method	CPU time (SPF*)			TPR (FPPI#=0.1)
	1st stage	2nd stage	total time	
VJ [7]	0.271	-	0.271	0.462
NPD [2]	0.678	-	0.678	0.801
NPD-HeadHunter	0.678	988	988.678	0.810
NPD-MTCNN	0.678	0.073	0.751	**0.841**
NPD-HR	0.678	2.678	3.356	**0.841**
HeadHunter [3]	1961	-	1961	0.834
HeadHunter-NPD	1961	0.404	1961.404	0.819
HeadHunter-MTCNN	1961	0.116	1961.116	**0.889**
HeadHunter-HR	1961	3.648	1964.648	**0.889**
MTCNN [4]	0.355	-	0.355	0.919
MTCNN-NPD	0.355	0.220	0.575	0.843
MTCNN-HeadHunter	0.355	456	456.355	0.882
MTCNN-HR	0.355	3.496	3.851	**0.930**
HR [5]	17.687	-	17.687	**0.943**
HR-NPD	17.687	0.170	17.857	0.839
HR-HeadHunter	17.687	794	811.687	0.886
HR-MTCNN	17.687	0.076	17.763	0.930

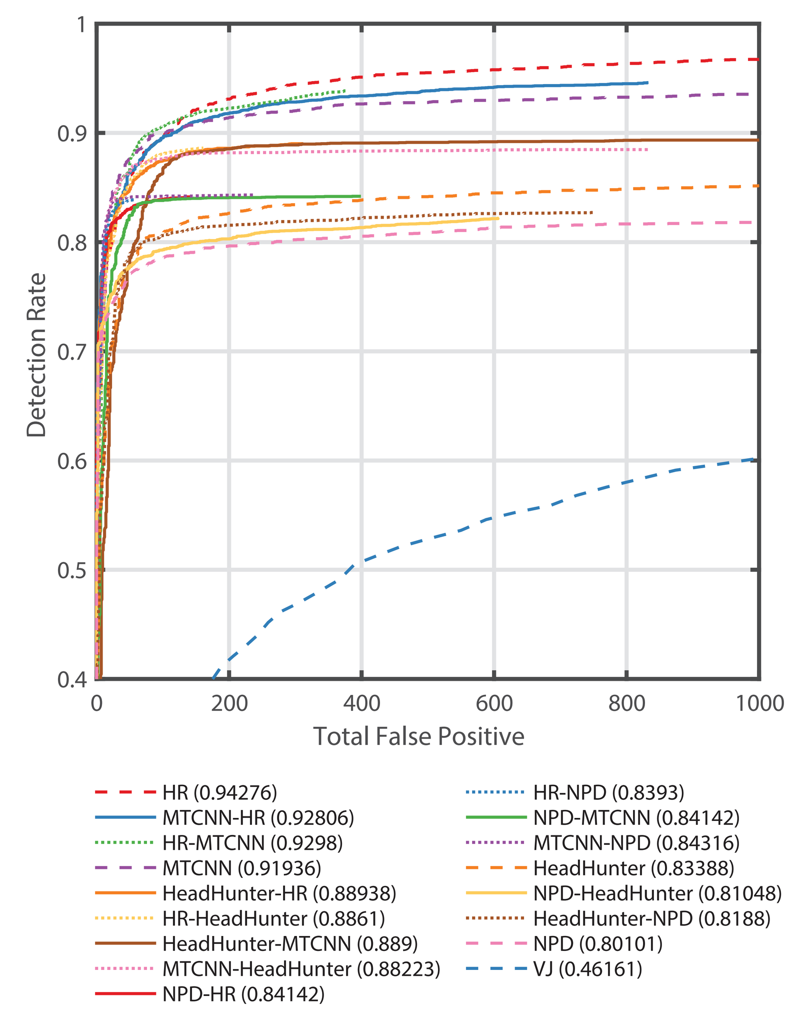
*SPF–Seconds Per Frame # FPPI–False Positives Per Image, TPR–True Positive Rate

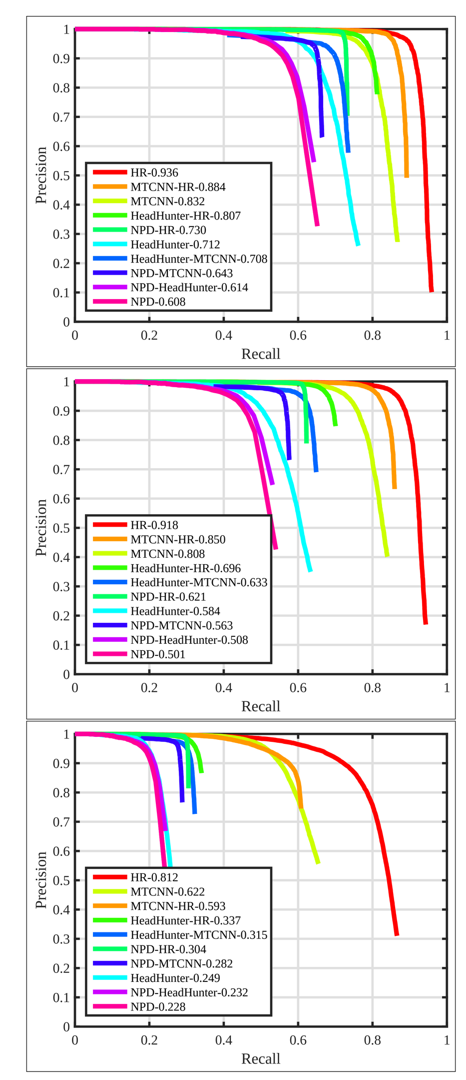
- Comparison of our proposed framework and the state-of-the-art face detector:

Method	CPU time (SPF*)	TPR (FPPI#=0.1)
HR [5]	17.687	**0.943**
MTCNN-HR (ours)	**3.851**	0.930

*SPF–Seconds Per Frame # FPPI–False Positives Per Image, TPR–True Positive Rate

- Comparisons on FDDB dataset [1] (left) and WIDER FACE validation set [6] (right):





Conclusions

- We propose three essential cascade properties that guide us in determining the efficacy of the cascaded detector
- Experimental results show our framework is able to remove a large number of false positives with an insignificant loss of true positive rate
- We found a pair of face detectors that achieves significantly lower false positive rate with competitive detection rate, which is five times faster than the state-of-the-art detector described in [5]

[1] Vidit Jain and Erik G Learned-Miller. Fddb: A benchmark for face detection in unconstrained settings. *UMass Amherst Technical Report*, 2010.
[2] Shengcai Liao, Anil K Jain, and Stan Z Li. A fast and accurate unconstrained face detector. *PAMI*, 38(2):211–223, 2016.
[3] Markus Mathias, Rodrigo Benenson, Marco Pedersoli, and Luc Van Gool. Face detection without bells and whistles. In *ECCV*, 2014.
[4] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, 2016.
[5] Peiyun Hu and Deva Ramanan. Finding tiny faces. In *CVPR*, 2017.
[6] Shuo Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang. Wider face: A face detection benchmark. In *CVPR*, 2015.
[7] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *CVPR*, 2001.

Acknowledgement: This work has been partly funded by Sullivan Nicolaides Pathology, Australia and the Australian Research Council (ARC) Linkage Projects Grant LP160101797. Arnold Wiliem is funded by the Advance Queensland Early-Career Research Fellowship. We thank Dr Shaokang Chen for the useful discussions.