



Real-Time Rotation-Invariant Face Detection with Progressive Calibration Networks

CVPR 2018

Xuepeng Shi, Shiguang Shan, Meina Kan, Shuzhe Wu, Xilin Chen

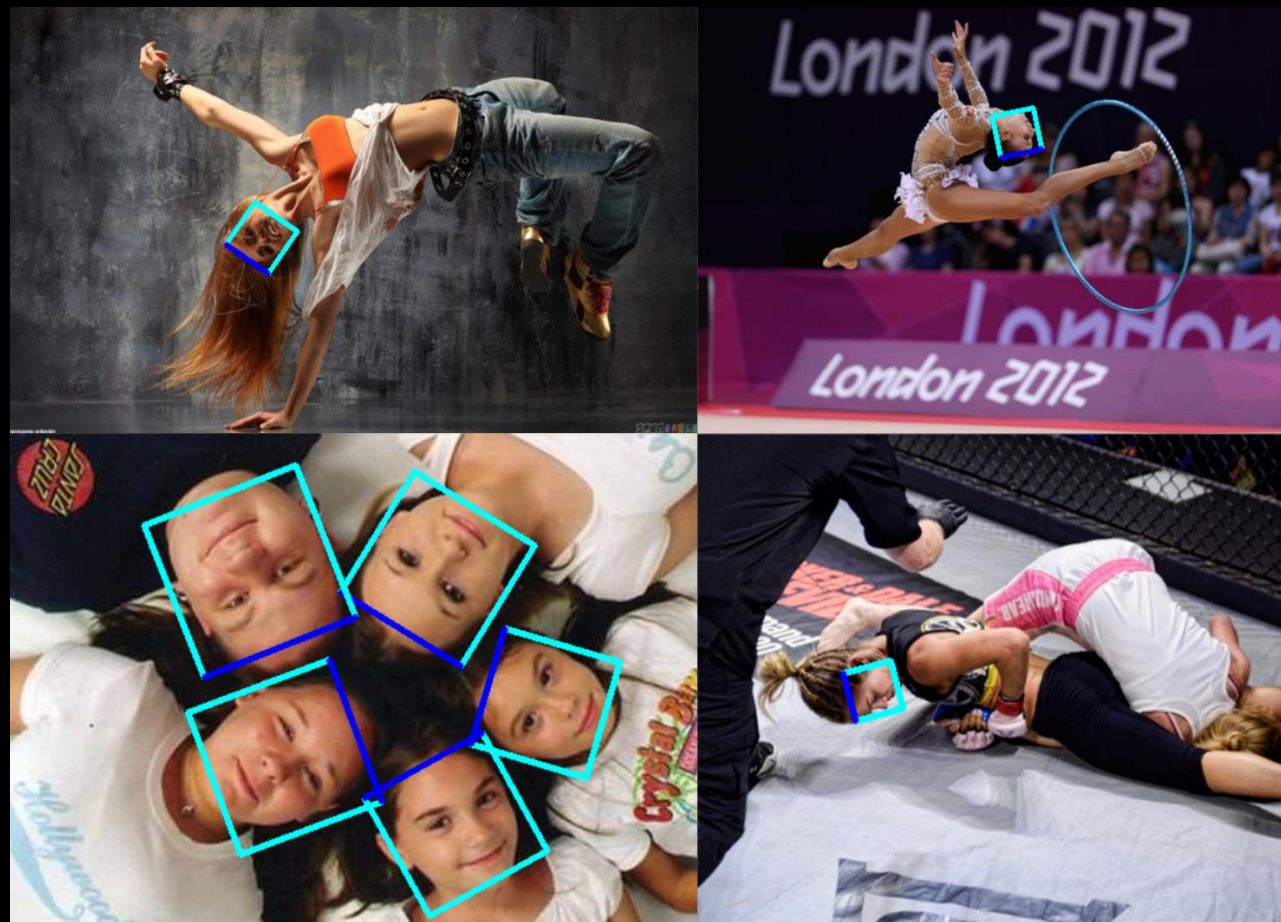
Rotation-Invariant Face Detection



Full rotation-in-plane (RIP) angles



Large appearance variations



Existing Methods [1/3]

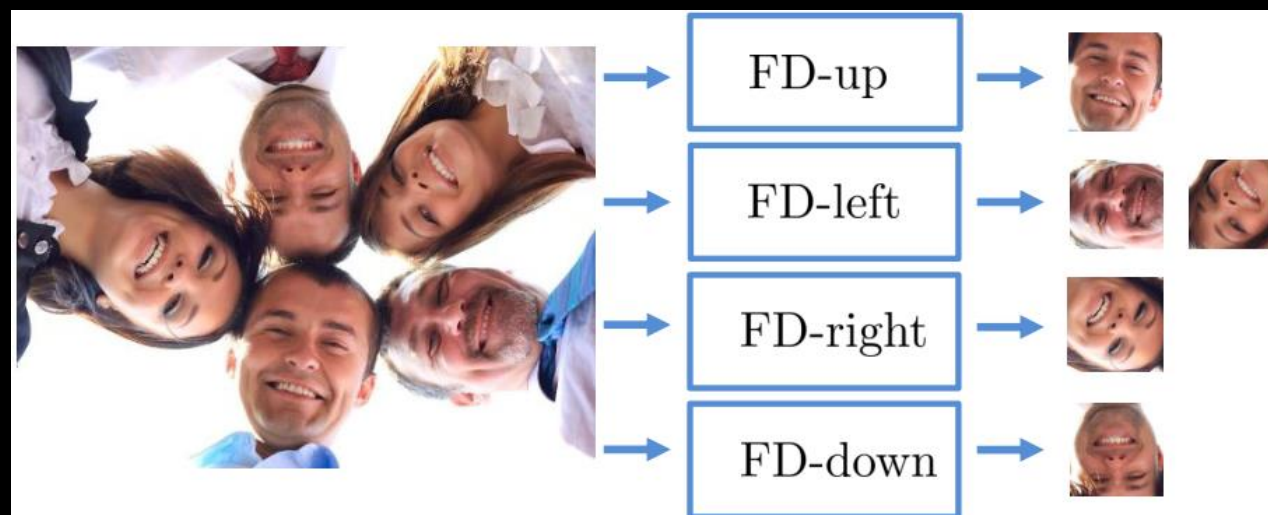


- Data Augmentation
 - To ensure the accuracy, a large neural network is usually required



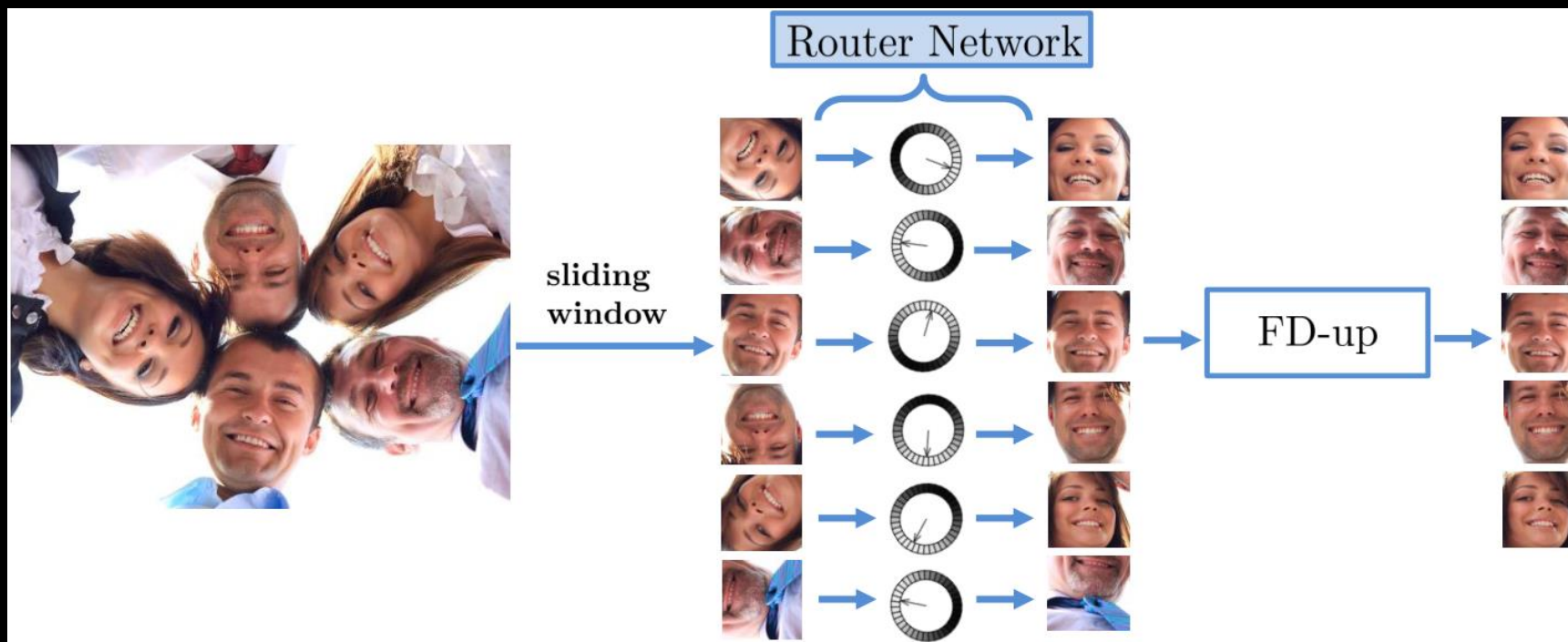
Existing Methods [2/3]

- Divide-and-Conquer [Huang, TPAMI 2007]
 - The overall time-cost largely increases
 - More false alarms are easily introduced



Existing Methods [3/3]

- Rotation Router [Rowley, CVPR 1998]
 - Precisely estimating the RIP angles of faces is quite challenging





Motivation and Goal

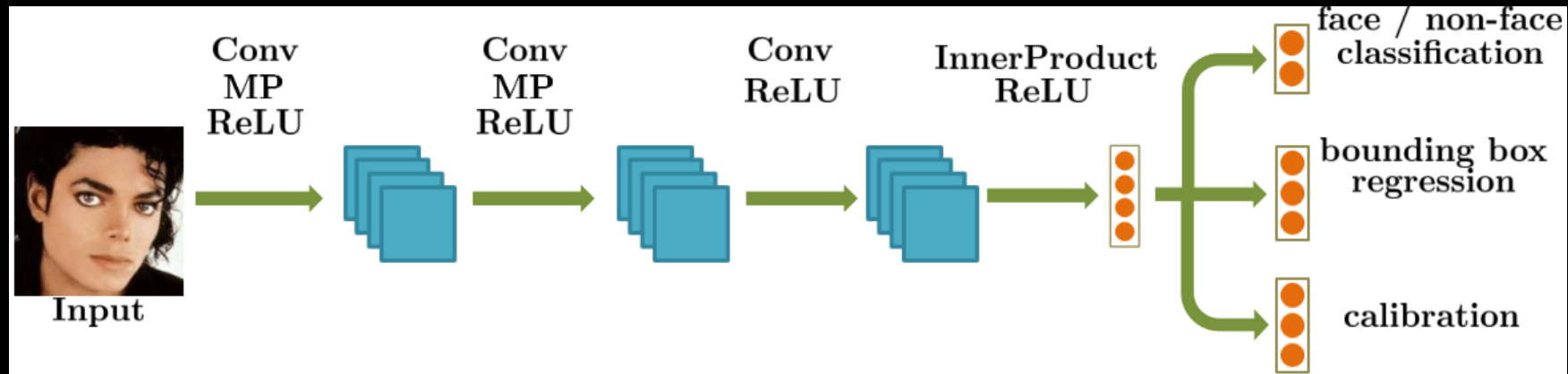
- Most existing methods compromise with speed or accuracy to handle the large RIP variations
- Our PCN aims at accurate rotation-invariant face detection with low time-cost

Progressive Calibration Networks [1/2]



- A sequence of multi-task CNNs
 - Face / non-face classification + bounding box regression + calibration

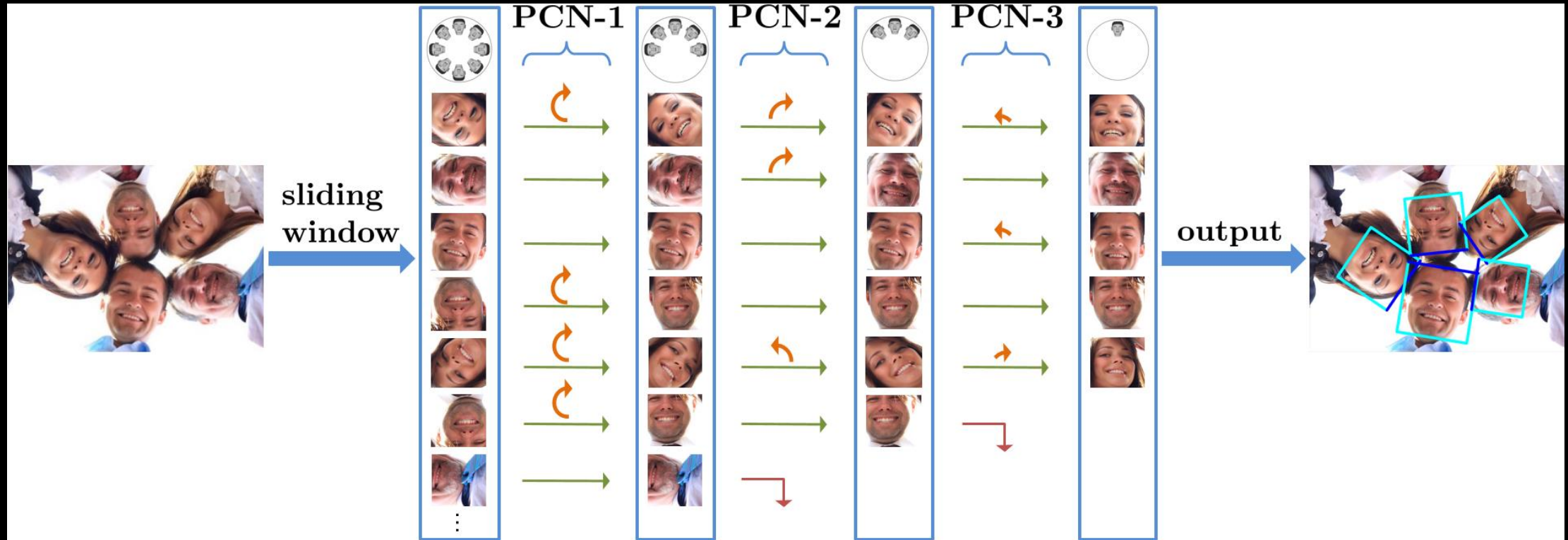
$$\min_F L = L_{classification} + \lambda_{bbox_reg} \cdot L_{bbox_reg} + \lambda_{calibration} \cdot L_{calibration}$$



Progressive Calibration Networks [2/2]



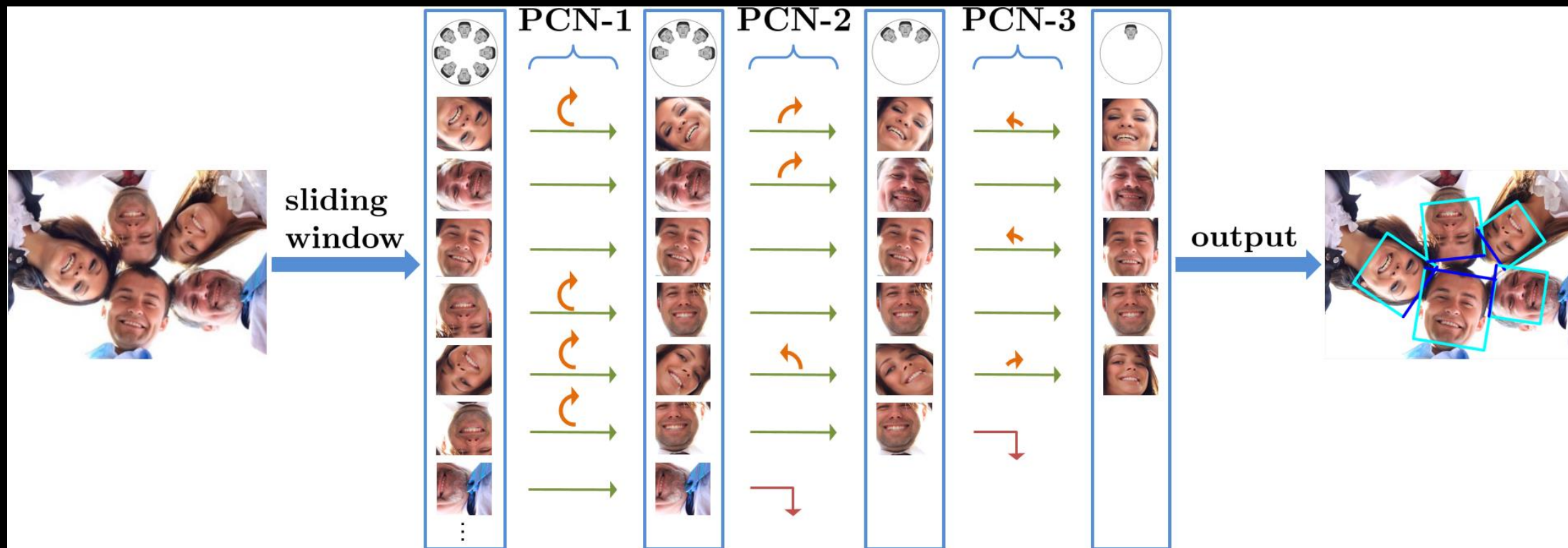
- Calibrates the RIP orientation to upright progressively
- Only coarse calibrations are conducted in early stages





Two Cascade Structures in PCN [1/2]

- Cascade classification [Li, CVPR 2015]
 - Significantly improve the speed of detection



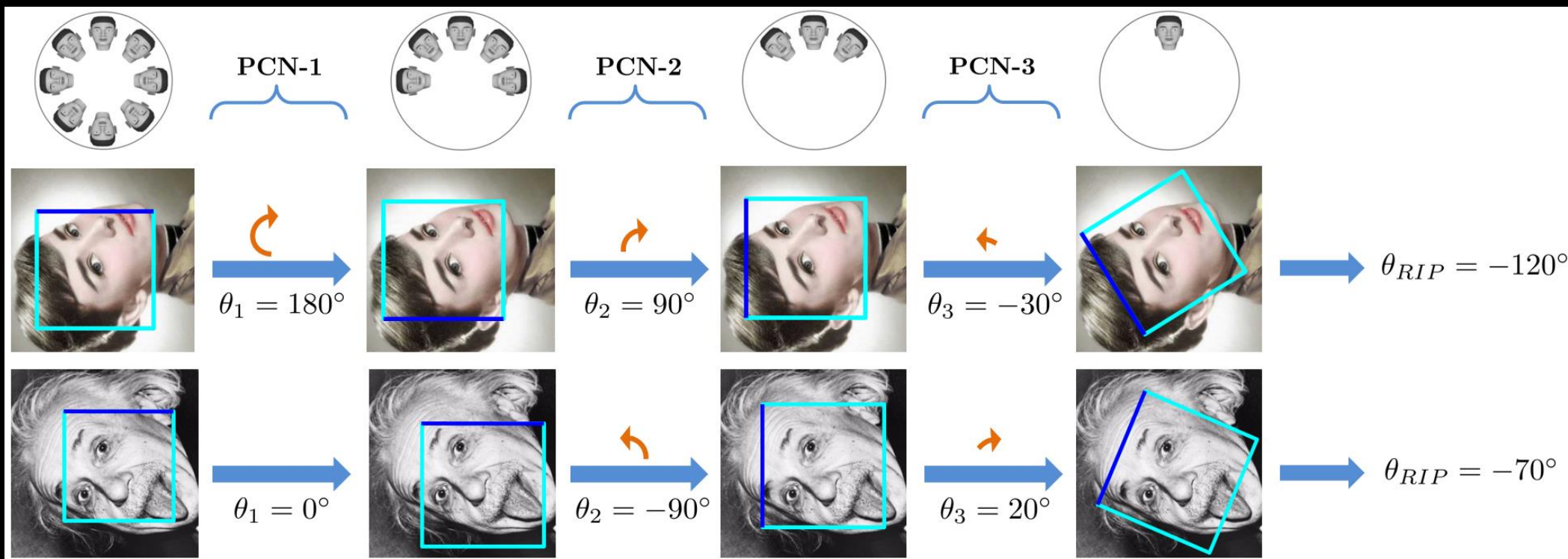


Two Cascade Structures in PCN [2/2]

- Cascade calibration

- Binary classification + ternary classification + regression

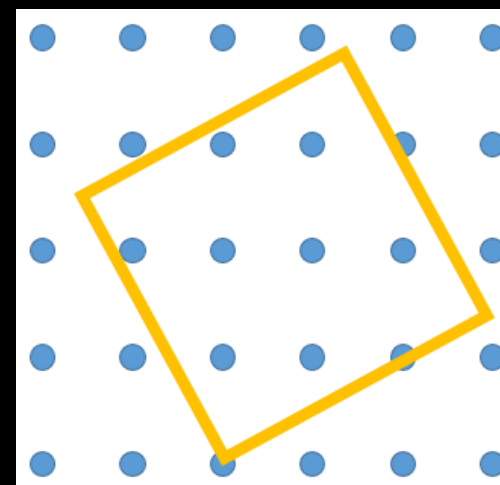
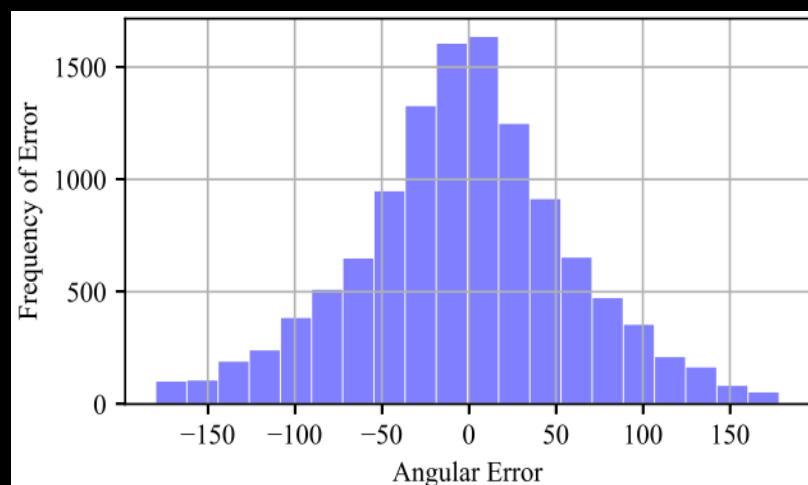
$$\theta_{RIP} = \theta_1 + \theta_2 + \theta_3, \theta_1 = 0^\circ / 180^\circ, \theta_2 = -90^\circ / 0^\circ / 90^\circ$$





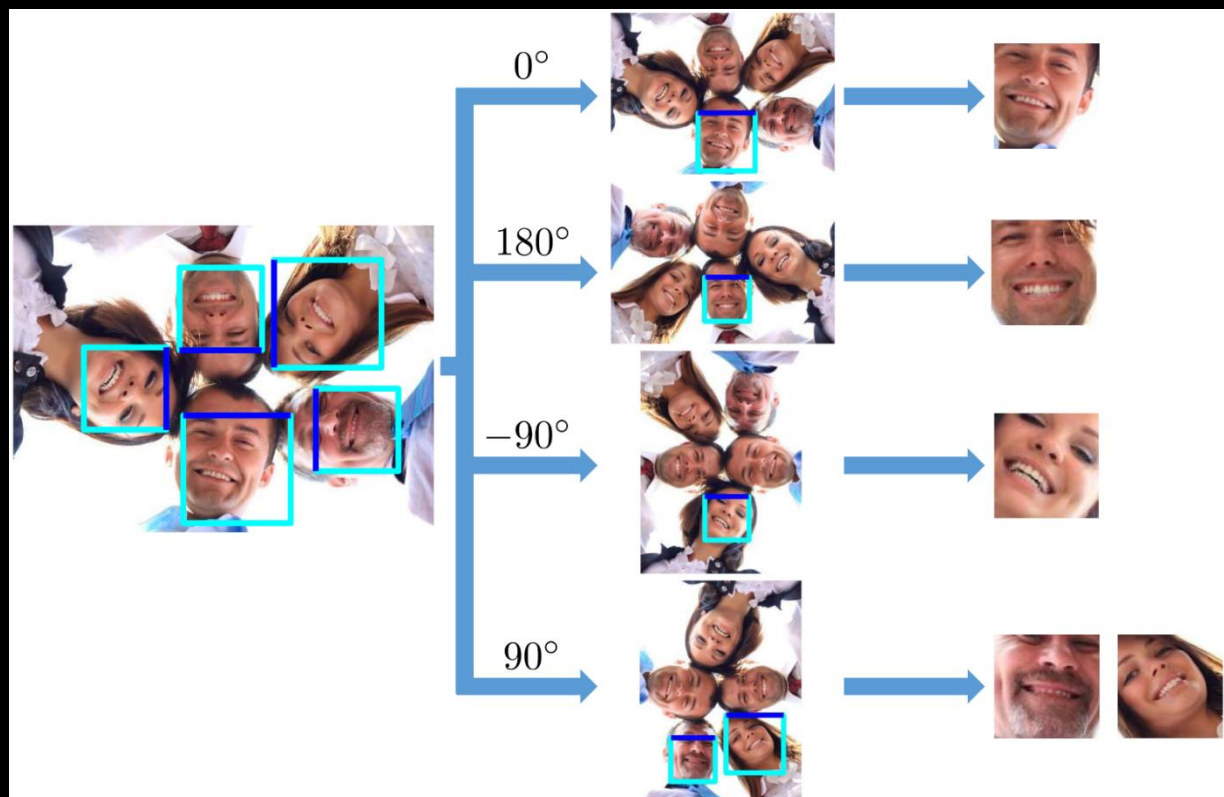
Cascade Calibration [1/2]

- Why cascade?
 - Easier
- Why coarse classification?
 - More accurate
 - More efficient
(Interpolation VS Flipping)



Cascade Calibration [2/2]

- Flip original image three times in advance (1ms - 2ms)
- Reduce the time-cost of calibration dozens of times





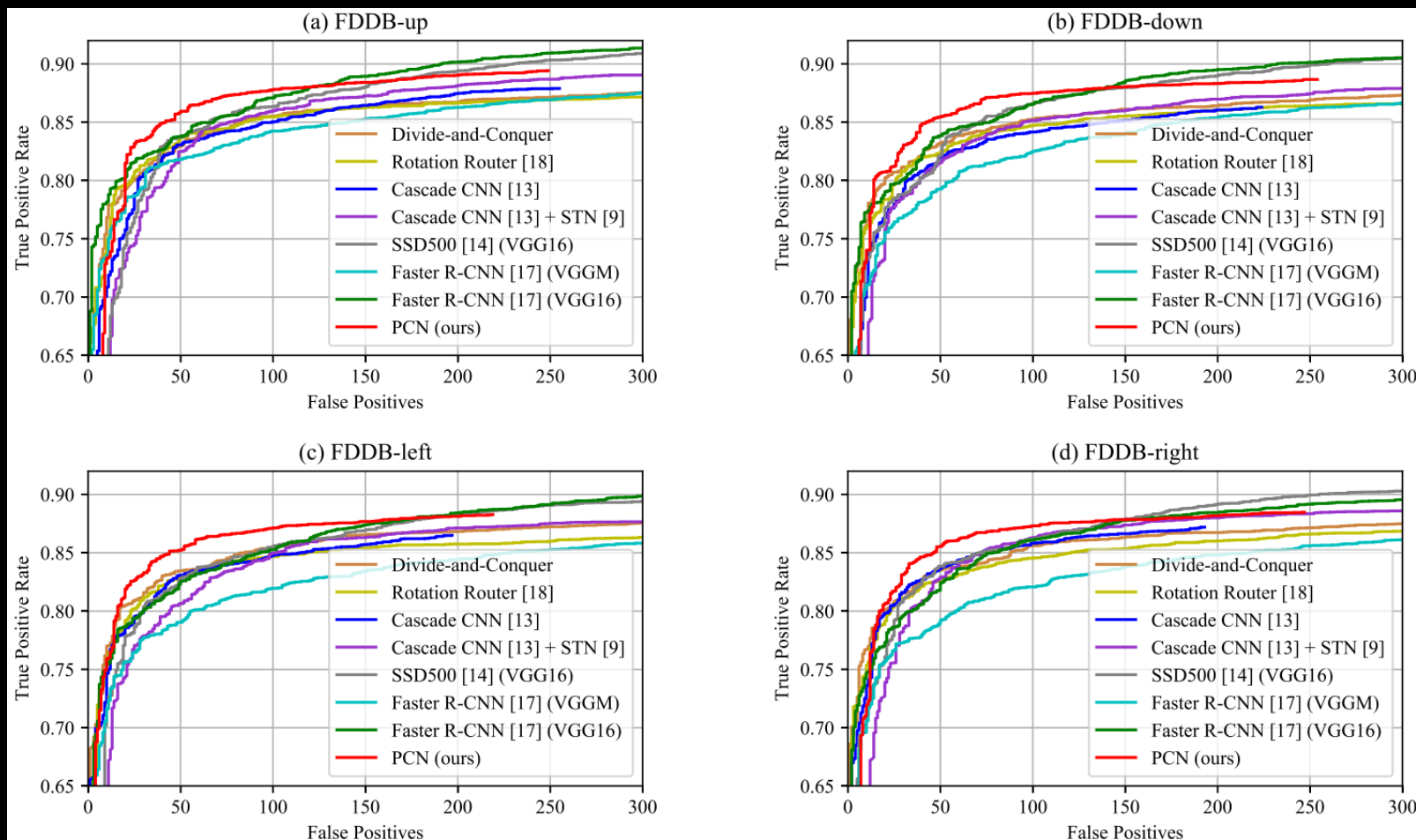
Methods for Comparison

- Data Augmentation
 - Faster RCNN (VGGM, VGG16)
 - SSD500 (VGG16)
 - RFCN (ResNet-50)
 - Cascade CNN
- Divide-and-Conquer
- Rotation Router

Results [1/3]



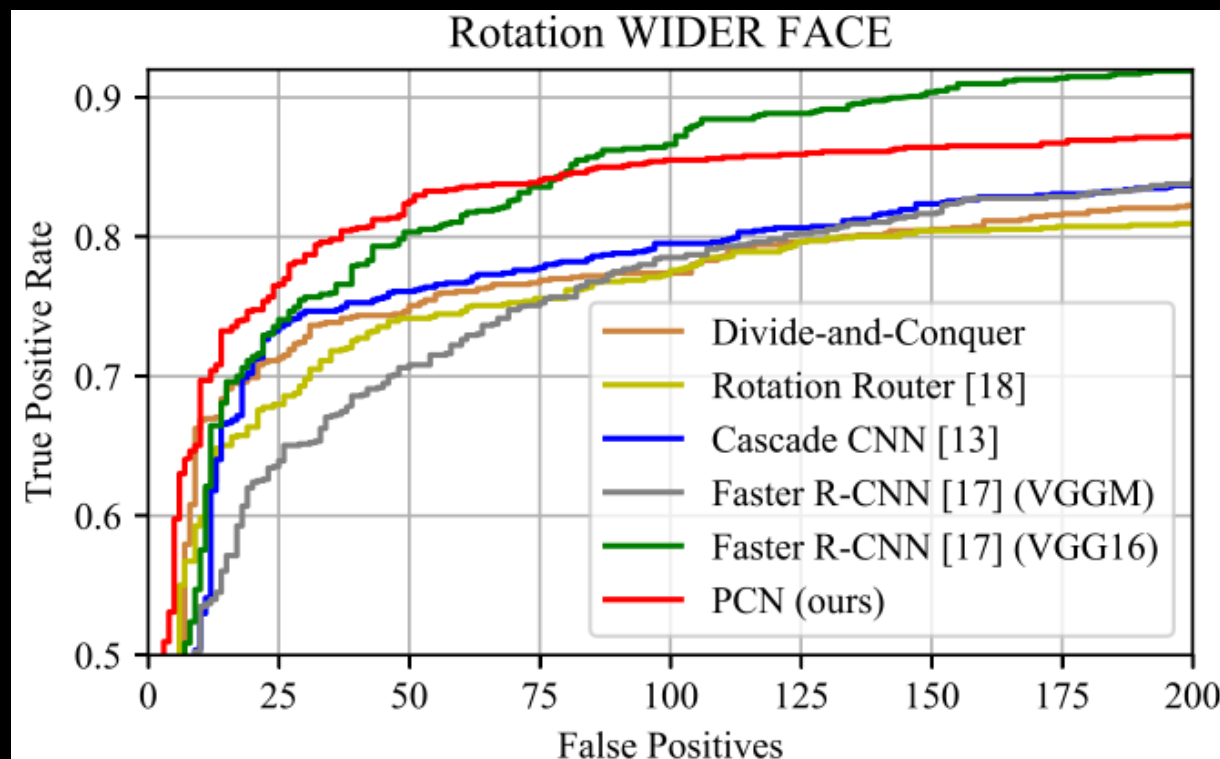
- Rotate the Fddb images by -90° , 90° , and 180° respectively





Results [2/3]

- A subset of the WIDER FACE test set containing rotated faces





Results [3/3]

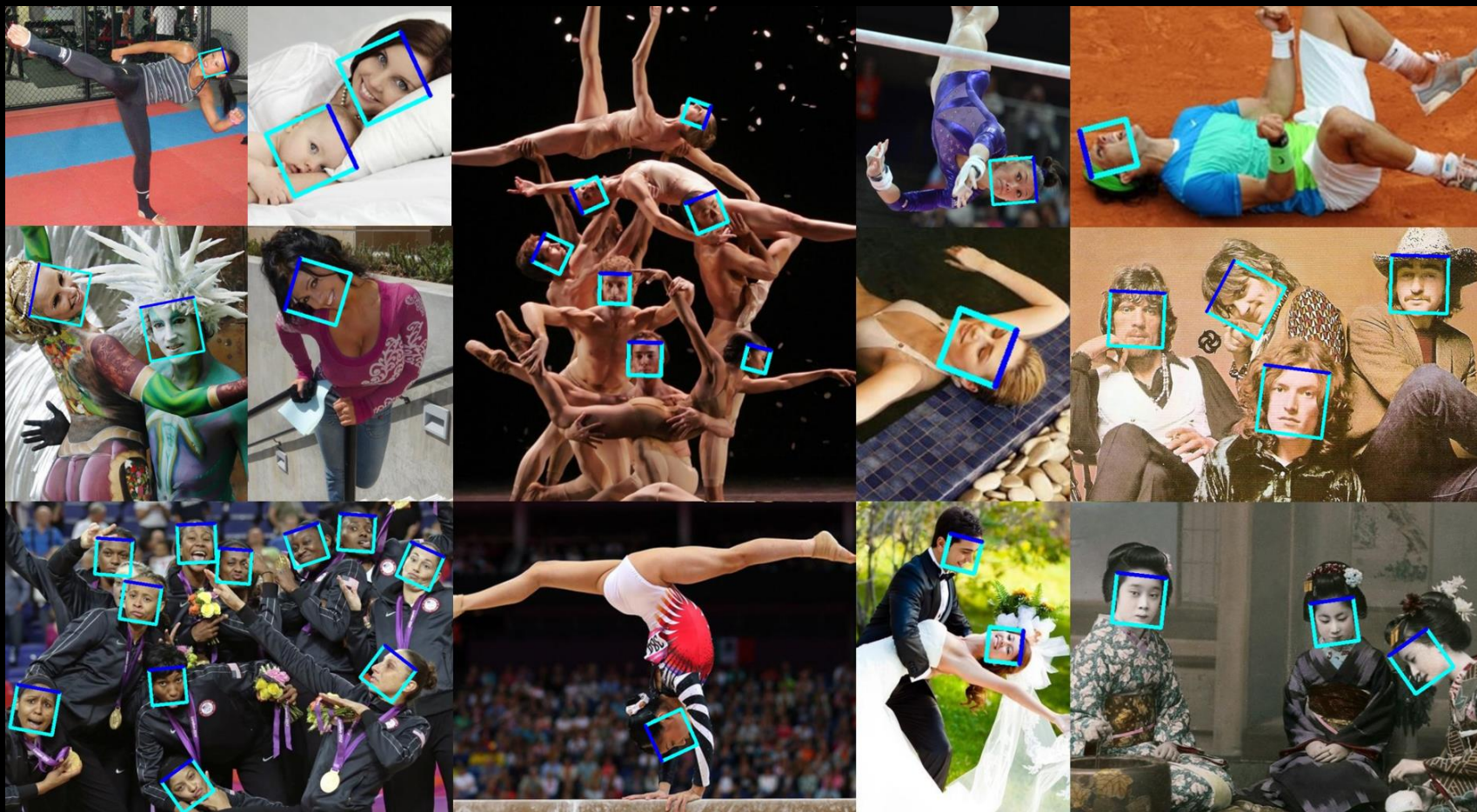
- 640 × 480 VGA images with 40 × 40 minimum face size
- A desktop computer with 3.4GHz CPU, GTX Titan X
- PCN can run in real-time on both CPU and GPU

Method	Recall rate at 100 FP on Fddb					Speed		Model Size
	Up	Down	Left	Right	Ave	CPU	GPU	
Divide-and-Conquer	85.5	85.2	85.5	85.6	85.5	15FPS	20FPS	2.2M
Rotation Router [18]	85.4	84.7	84.6	84.5	84.8	12FPS	15FPS	2.5M
Cascade CNN [13]	85.0	84.2	84.7	85.8	84.9	31FPS	67FPS	4.2M
Cascade CNN [13] + STN [9]	85.8	85.0	84.9	86.2	85.5	16FPS	30FPS	4.7M
SSD500 [14] (VGG16)	86.3	86.5	85.5	86.1	86.1	1FPS	20FPS	95M
Faster R-CNN [17] (VGGM)	84.2	82.5	81.9	82.1	82.7	1FPS	20FPS	350M
Faster R-CNN [17] (VGG16)	87.0	86.5	85.2	86.1	86.2	0.5FPS	10FPS	547M
R-FCN [2] (ResNet-50)	87.1	86.6	85.9	86.0	86.4	0.8FPS	15FPS	123M
PCN (ours)	87.8	87.5	87.1	87.3	87.4	29FPS	63FPS	4.2M

Progressive Calibration Networks



- Fast, accurate, robust





Thanks!