

# 基于卷积神经网络的鲁棒人脸检测

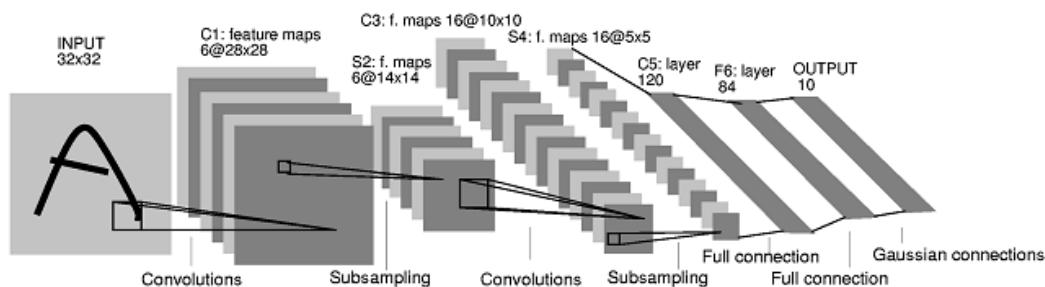
中国科学院 计算技术研究所 VIPL 研究组

邬书哲

2018.4.26

# 主要内容

- CVPR 2018论文分享
  - 旋转不变人脸检测：PCN
- 增强人脸检测鲁棒性的策略
  - Coarse-to-Fine
  - 多尺度
  - 上下文
  - Anchor适配



# 旋转不变人脸检测：PCN

Xuepeng Shi, Shiguang Shan, Meina Kan, Shuzhe Wu and Xilin Chen. *Real-Time Rotation-Invariant Face Detection with Progressive Calibration Networks*. CVPR, 2018.

# 旋转不变人脸检测：PCN

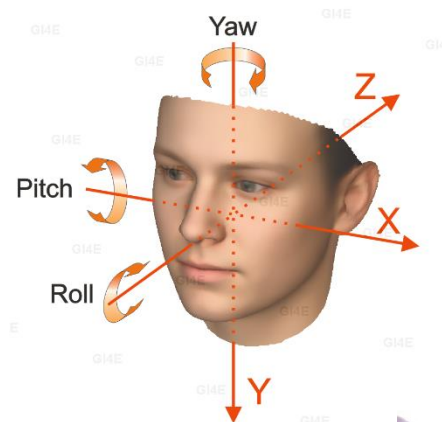
- 头部姿态变化

- 旋转

- 平面内旋转：Roll
    - 平面外旋转：Yaw, Pitch
    - 平面内和平面外旋转组合

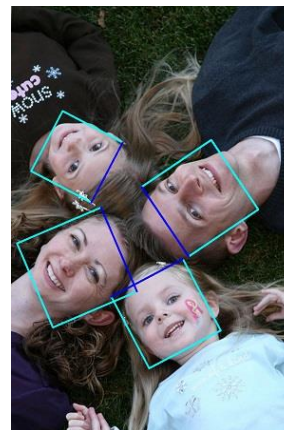
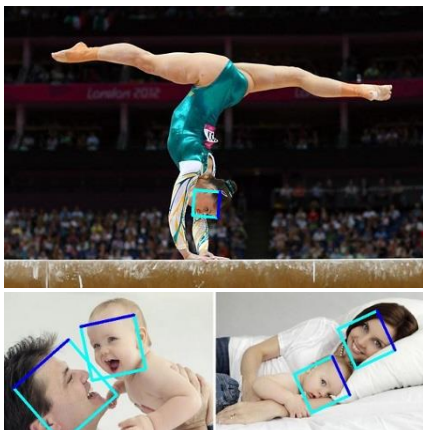
- 头部姿态变化造成人脸表观发生变化

- 五官被部分遮挡
    - 五官的相对位置发生变化
    - 不同人脸之间的（类内）差异增大，给人脸分类/检测带来困难



# 旋转不变人脸检测：PCN

- 旋转不变人脸检测 *Rotation-Invariant*
  - 人脸在平面内任意旋转
    - 以竖直人脸的方向为 $0^\circ$ ，在平面内允许旋转范围： $[-180^\circ, 180^\circ]$
  - 实际场景
    - 手持设备拍摄方向的不确定性：手机，平板
    - 拍摄角度：俯拍，航拍
    - 特殊运动：体操，跳水



# 旋转不变人脸检测：PCN

- 现有的人脸检测方法

- 三类方法

- Cascade CNN: 速度最快，精度相对较低

- MT-CNN, ICS

- Faster R-CNN: 速度较慢，精度较高

- Face R-CNN, Face R-FCN, FDNet

- SSD: 速度较快，精度较高

- SSH, S<sup>3</sup>FD, FAN, PyramidBox

- 特点

- 通用目标检测方法 + 人脸数据

- 一般借助于高复杂度模型对人脸表观进行建模，检测速度难以保证

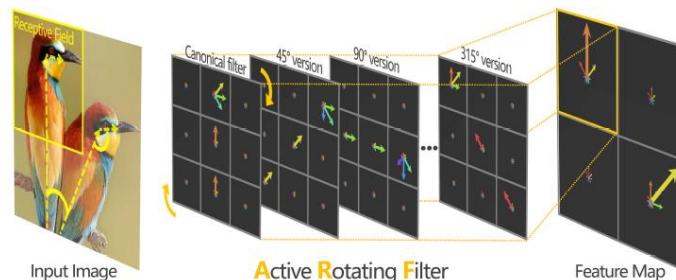
- 设计上比较通用，不针对特定问题，如姿态

# 旋转不变人脸检测：PCN

- 现有的处理平面内旋转的方法

- 学习旋转不变特征

- 数据增广：人工生成不同旋转角度的人脸样本，学习单个旋转鲁棒分类器



- 旋转卷积核：Oriented Response Network (CVPR 2017)
      - 正则化：约束不同旋转角度的人脸具有相同的特征
      - 问题：需要更高复杂度的模型来处理复杂的类内变化

- 分治法

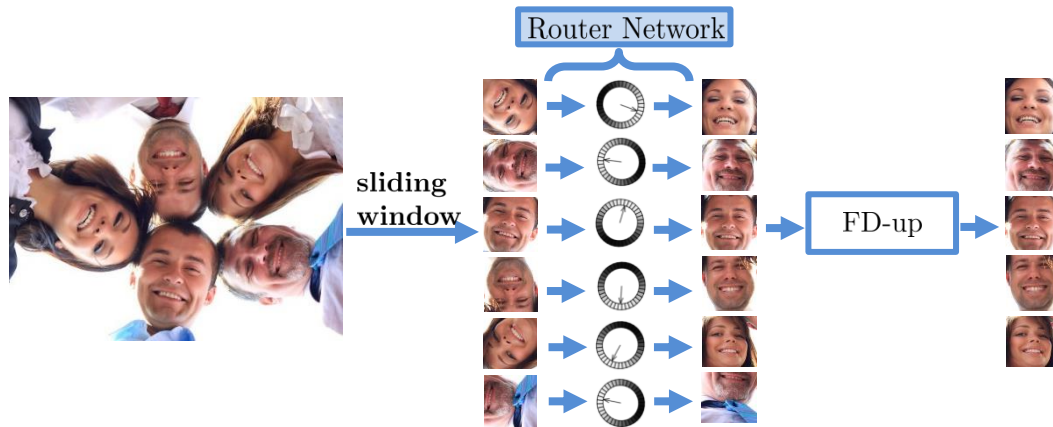
- 对样本空间进行划分，在对应的各个子集上分别训练模型
    - 问题：检测时间随着模型数量线性增长

# 旋转不变人脸检测：PCN

- 现有的处理平面内旋转的方法

- 学习旋转不变特征
- 分治法
- 进行旋转校准

- **Rotation Router**: 先预测人脸角度，然后进行旋转校正，将旋转人脸校正为垂直人脸，再进行人脸/非人脸的二分类
- 问题：角度预测本身就非常困难





# 旋转不变人脸检测：PCN

- 现有的处理平面内旋转的方法
  - 学习旋转不变特征
  - 分治法
  - 进行旋转校准
  - 小结
    - 需要增加额外的模型或者显著提升模型复杂度来保证检测精度
    - 缺乏高效而鲁棒的旋转校准方式
    - 在检测精度和速度之间难以进行很好地权衡



# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*
  - Cascade CNN
    - 速度快，能够在CPU以及手机等设备上运行
    - 精度尚可，改进后在部分场景下也能获得较高的精度



- PCN
  - 基于Cascade CNN针对旋转人脸进行改进
  - 基本思路：将Coarse-to-fine的思想进一步应用于旋转校准
    - 逐步校准人脸的旋转角度
    - 粗略校准和精细校准相结合

# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 各级校准设计

- 最直接的校准方式：每一级都增加一个校准任务

- 矛盾：简单的模型面对的样本分布最为复杂
      - 严重的误差累积，无法控制校准“路径”（转错/转过头）
      - 每一级的校准都需要考虑全平面360°旋转的人脸

- 最直接的校准任务：预测旋转角度

- 角度是连续值，取值范围包含无穷多个值
      - 常用的损失函数要求精确预测，限制过于严格
      - 有限样本只能覆盖离散点
      - 无法精确标定角度，即使采用人工增广数据的方式（初始角度未知）
      - 无法保证样本标签的一致性
      - 实现困难：每个人脸校准的角度都不同，而任意角度的旋转难以高效率地实现

# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 各级校准设计：Coarse-to-Fine

- 逐步校准人脸的旋转角度

- 粗略校准和精细校准相结合

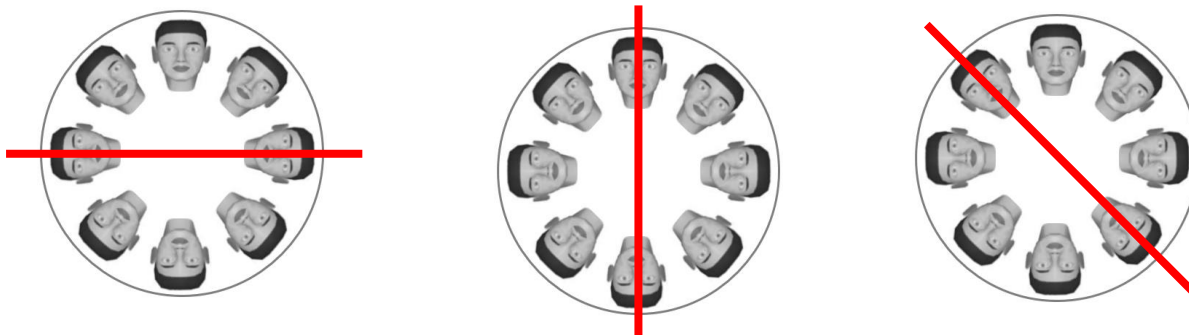
- 第一级：上下翻转校准

- 输入： $[-180^\circ, 180^\circ] \Rightarrow$  输出： $[-90^\circ, 90^\circ]$

- 二分类问题：任务相对简单，小模型能够获得高校准精度，因此可以基本保证第一级之后人脸的旋转范围减半

- 思考：二分类是否有其它设计？

- 上下分，左右分，斜着分，.....

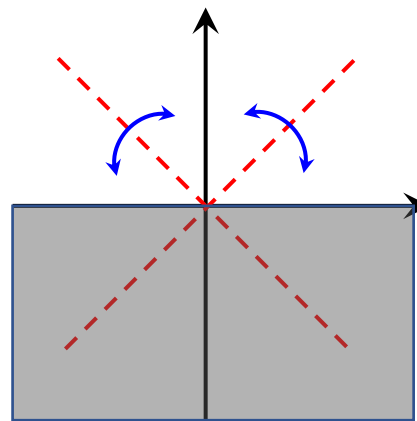


# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 各级校准设计：Coarse-to-Fine

- 逐步校准人脸的旋转角度
    - 粗略校准和精细校准相结合
    - 第一级：上下翻转校准
    - 第二级：左右翻转校准
      - 输入： $[-90^\circ, 90^\circ] \Rightarrow$  输出： $[-45^\circ, 45^\circ]$
      - 三分类问题
        - 向右翻转： $[-90^\circ, -45^\circ]$
        - 向左翻转： $[45^\circ, 90^\circ]$
        - 不翻转： $[-45^\circ, 45^\circ]$



# 旋转不变人脸检测：PCN

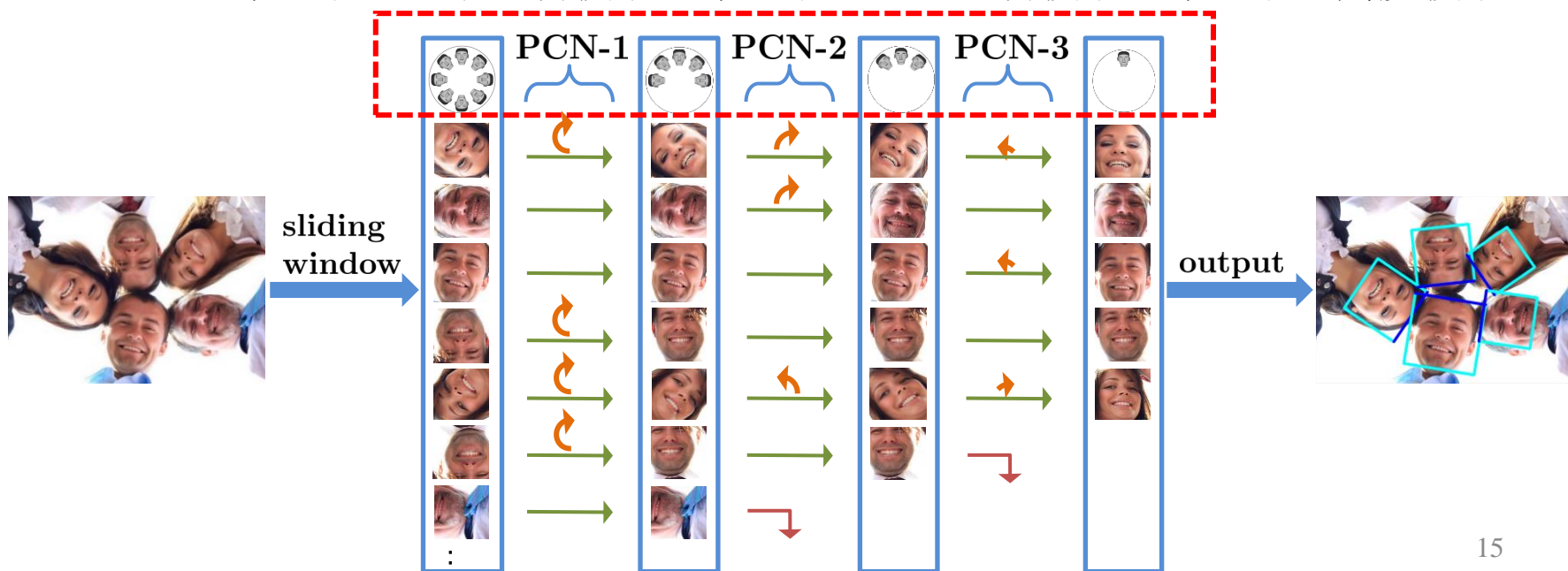
- 渐进校准网络 *Progressive Calibration Networks, PCN*
  - 各级校准设计：Coarse-to-Fine
    - 逐步校准人脸的旋转角度
    - 粗略校准和精细校准相结合
    - 第一级：上下翻转校准
    - 第二级：左右翻转校准
    - 第三级：角度校准
      - 输入： $[-45^\circ, 45^\circ] \Rightarrow$  输出： $0^\circ$
      - 回归问题
        - 第三级网络建模能力较强
        - 输入人脸的旋转角度范围比较小
        - 最后一级：不再需要顾虑误差累积的问题

# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 各级校准设计：Coarse-to-Fine

- 逐步校准人脸的旋转角度
- 粗略校准和精细校准相结合
- 第一级：上下翻转校准，第二级：左右翻转校准，第三级：角度校准

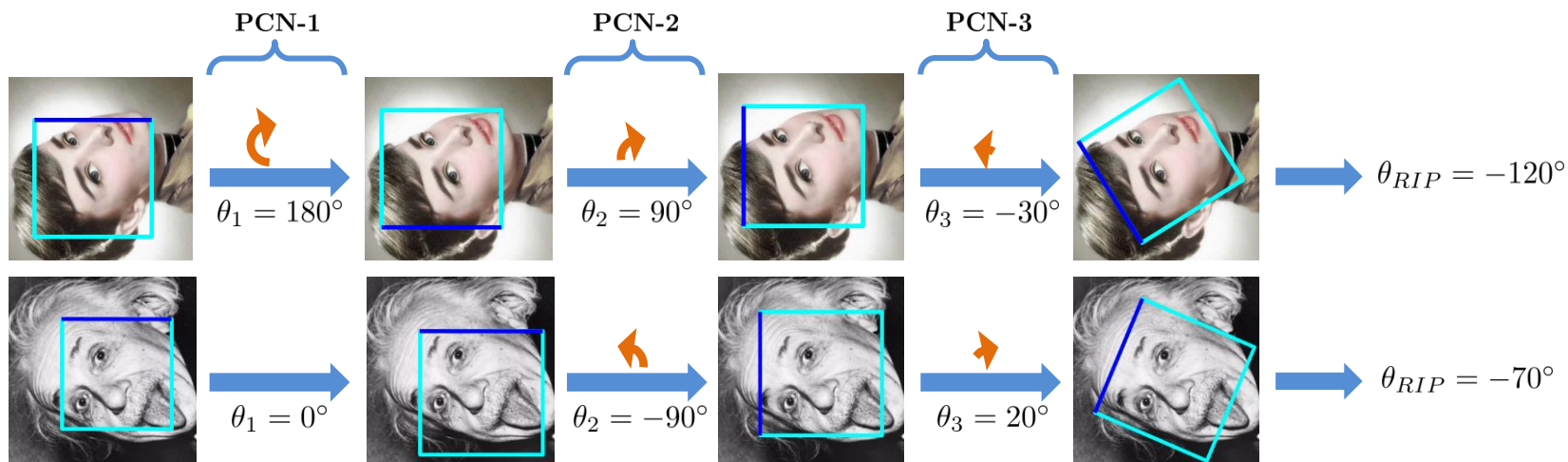


# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 各级校准设计：Coarse-to-Fine

- 逐步校准人脸的旋转角度
    - 粗略校准和精细校准相结合
    - 第一级：上下翻转校准，第二级：左右翻转校准，第三级：角度校准



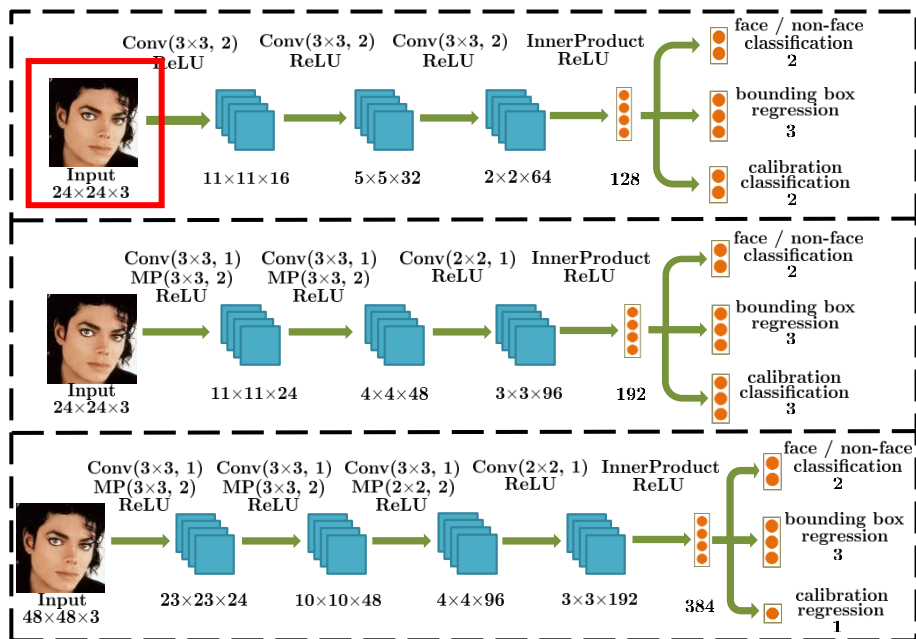


# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 网络结构

- 延续了Cascade CNN的部分设计：输入分辨率逐级增高
  - 为了更好地进行校准，第一级采用了 $24 \times 24$ 大小的输入
- 对网络做了进一步精简以提升速度



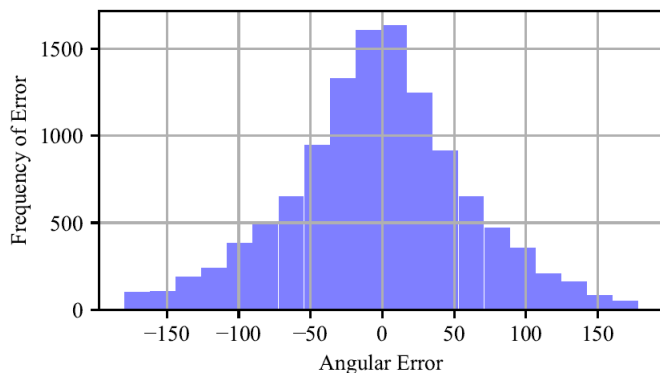
# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 实验结果

- 实验一：旋转预测

- Rotation-Router：一步到位（直接预测角度）



- PCN：逐步预测

- 分类准确率：第一级95%，第二级96%

- 第三级角度误差：8°

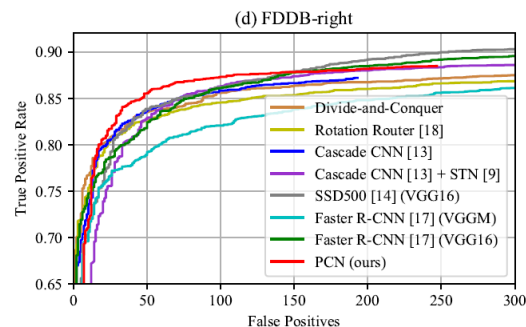
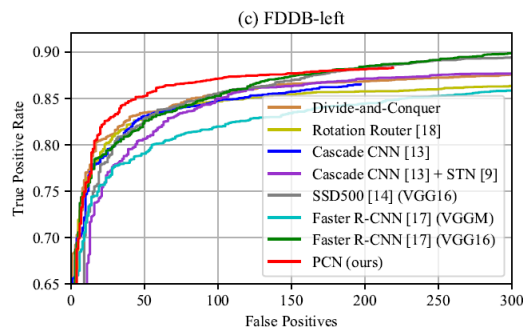
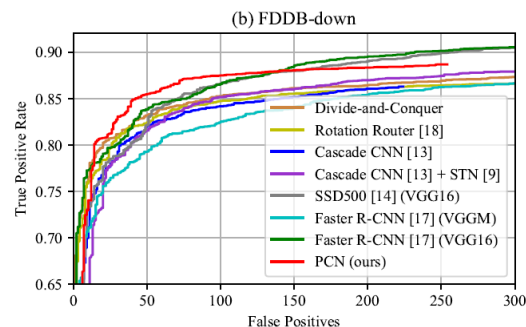
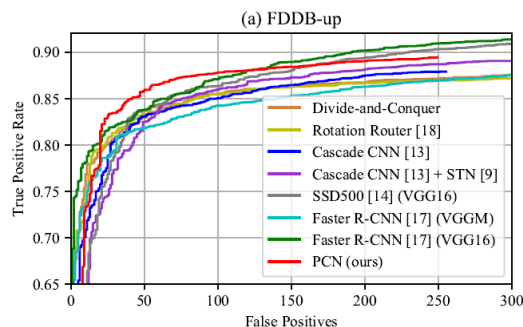
# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 实验结果

- 实验二：旋转不变人脸检测，Multi-Oriented FDDB

- 训练集：WIDER FACE训练集
- 测试集：FDDB，旋转四个方向



# 旋转不变人脸检测：PCN

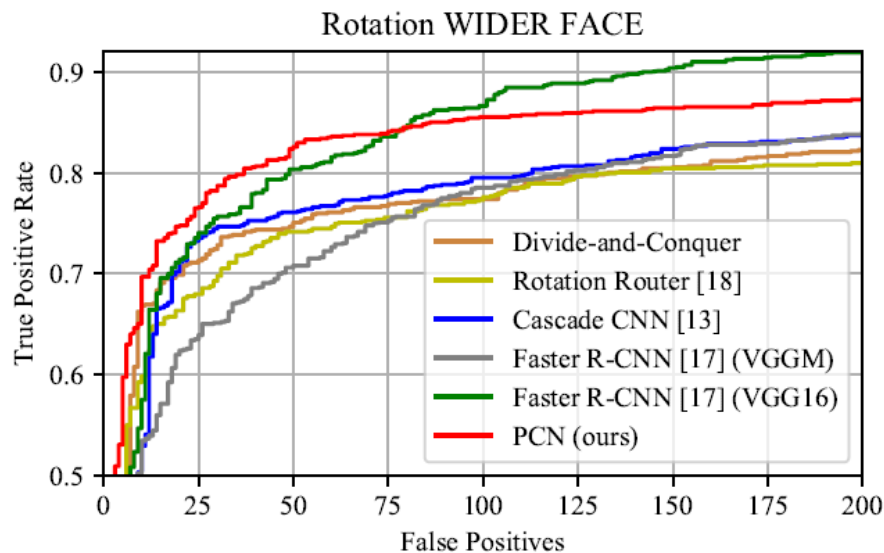
- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 实验结果

- 实验二：旋转不变人脸检测，WIDER FACE旋转人脸子集

- 测试集：WIDER FACE测试集中包含旋转人脸的图像

- 共370张图像，987张旋转人脸，手工标注



# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 实验结果

- 不同检测器在精度、速度、模型大小等方面的比较

- PCN显著优于分治法和Rotation-Router，以及采用相同模型和设置基于数据增广训练的Cascade CNN (Baseline)
- 在误检数目控制在合理的范围内，PCN的召回率超过了采用更高复杂度网络的SSD、Faster R-CNN、R-FCN等检测器（基于旋转增广的数据训练）

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	<b>2.2M</b>
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	<b>31FPS</b>	<b>67FPS</b>	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)	<b>87.8</b>	<b>87.5</b>	<b>87.1</b>	<b>87.3</b>	<b>87.4</b>	29FPS	63FPS	4.2M

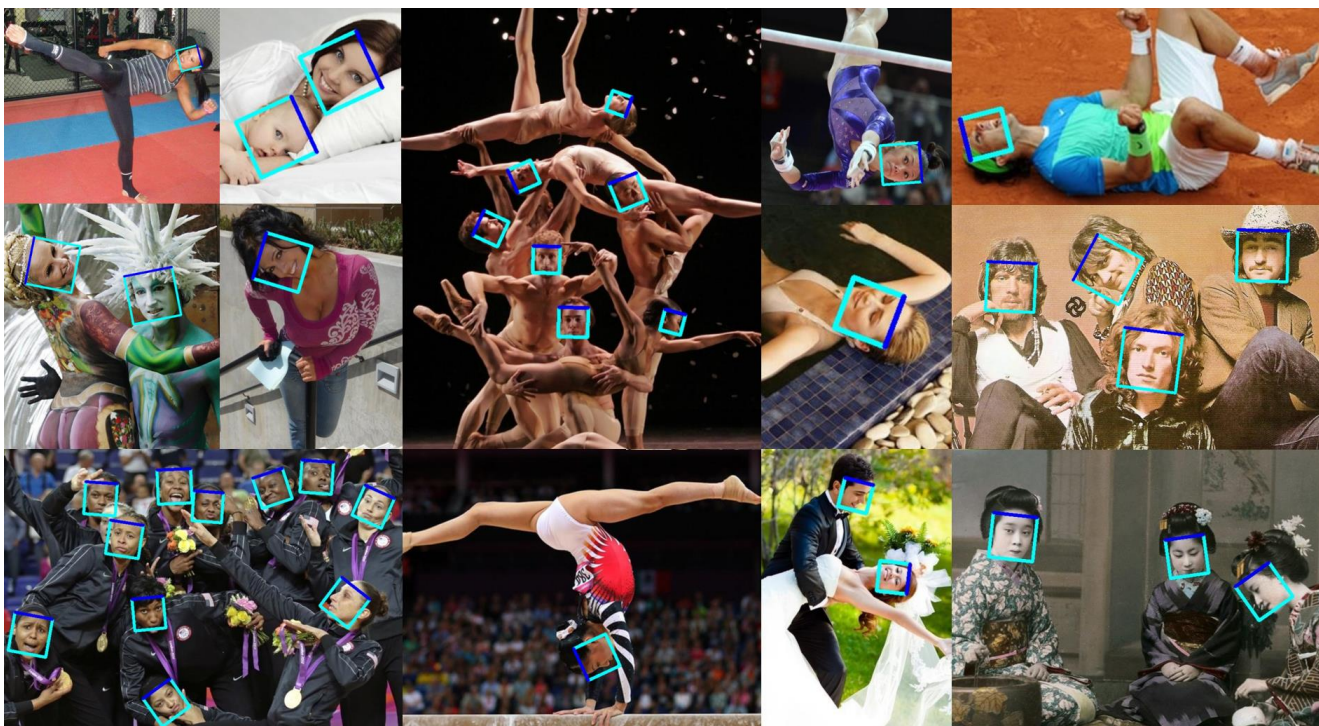
# 旋转不变人脸检测：PCN

- 渐进校准网络 *Progressive Calibration Networks, PCN*

- 小结：Coarse-to-Fine

- 规划校准路径，逐步进行旋转角度的校准
    - 先采用分类的方式进行粗校准，最后再预测连续角度精确校准

arXiv: <https://arxiv.org/abs/1804.06039>  
demo: <https://github.com/Jack-CV/PCN>



# 增强人脸检测鲁棒性的策略

Coarse-to-Fine • 多尺度 • 上下文 • Anchor适配

# 增强人脸检测鲁棒性的策略

## • 策略一：Coarse-to-Fine

- 基本思路：将解决问题的路径拆解成多段，设立中继目标

- 大量用于级联结构模型
- 级联分类器：把负样例逐步排除
- PCN：把旋转人脸逐步转正
- 人脸特征点定位：CFAN (ECCV 2014)
  - 先在低分辨率图像上做粗略搜索，确定大致区域
  - 再在高分辨率图像上做精细搜索，定准每一个点的位置

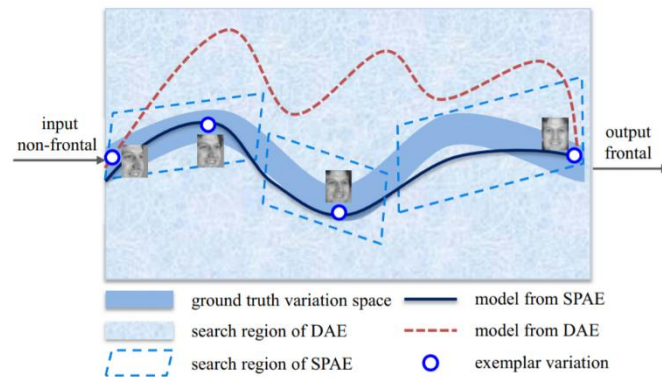


- 跨姿态人脸识别：SPAE (CVPR 2014), ReST (ICCV 2017)

- 将大姿态人脸逐步校正为准正面人脸
- 插入若干虚拟姿态作为中转校正目标

- 图像分割：FCN (CVPR 2015)

- 由低分辨率的分割结果逐步过渡到高分辨率





# 增强人脸检测鲁棒性的策略

- 策略二：多尺度

- 基本思路

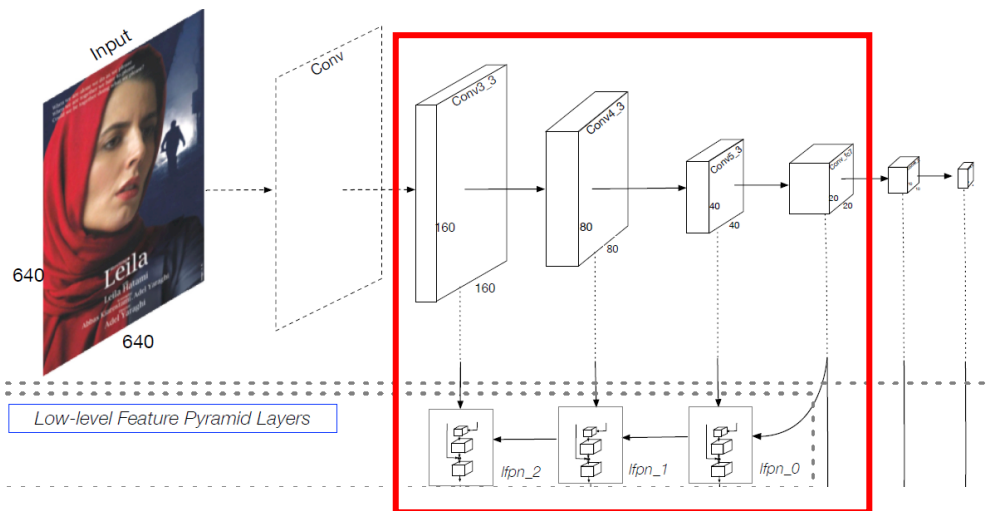
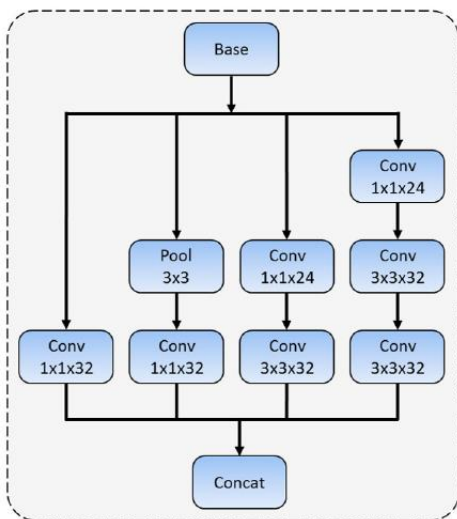
- (1) 尺度互补性：在不同的尺度上提取信息并进行融合

- 多尺度特征互补

- FaceBox (IJCB 2017): 采用Inception结构提取特征

- TinyFace (CVPR 2017): 融合ResNet不同Stage输出的特征图

- PyramidBox (arXiv 1803.07737): 采用Top-Down连接融合不同尺度的特征图



# 增强人脸检测鲁棒性的策略

- 策略二：多尺度

固定大小滑动窗口+图像金字塔

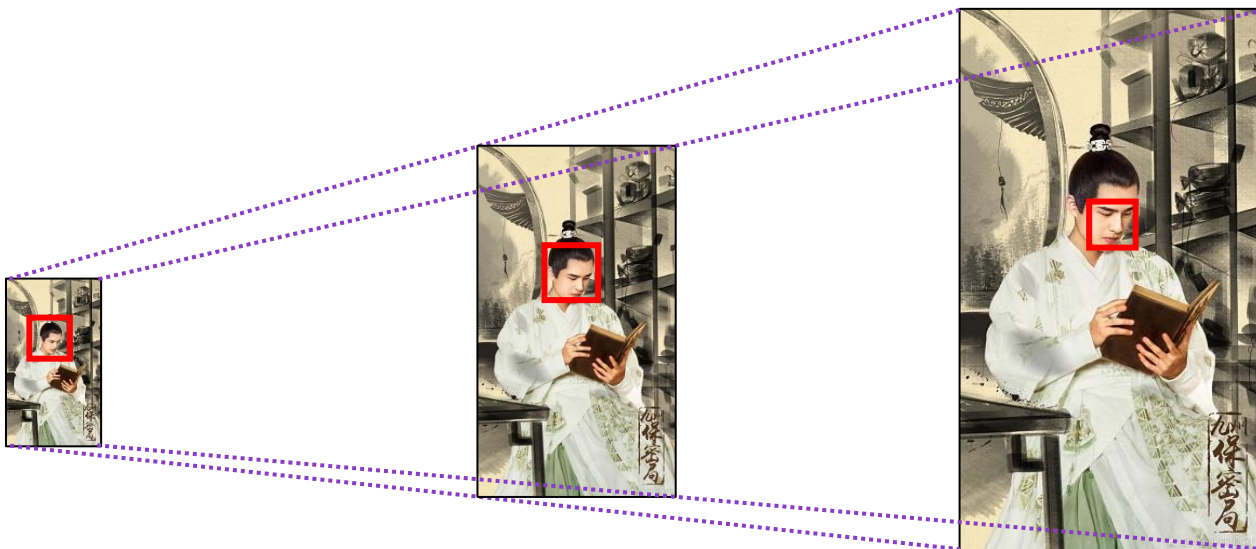
- 基本思路

- (1) 尺度互补性：在不同的尺度上提取信息并进行融合

- 多尺度Anchor互补

- PyramidBox (arXiv 1803.07737)：在不同的尺度下检测同一张脸，融合结果

- 不同尺度的Anchor分别对应于：脸部，头部，头肩



# 增强人脸检测鲁棒性的策略

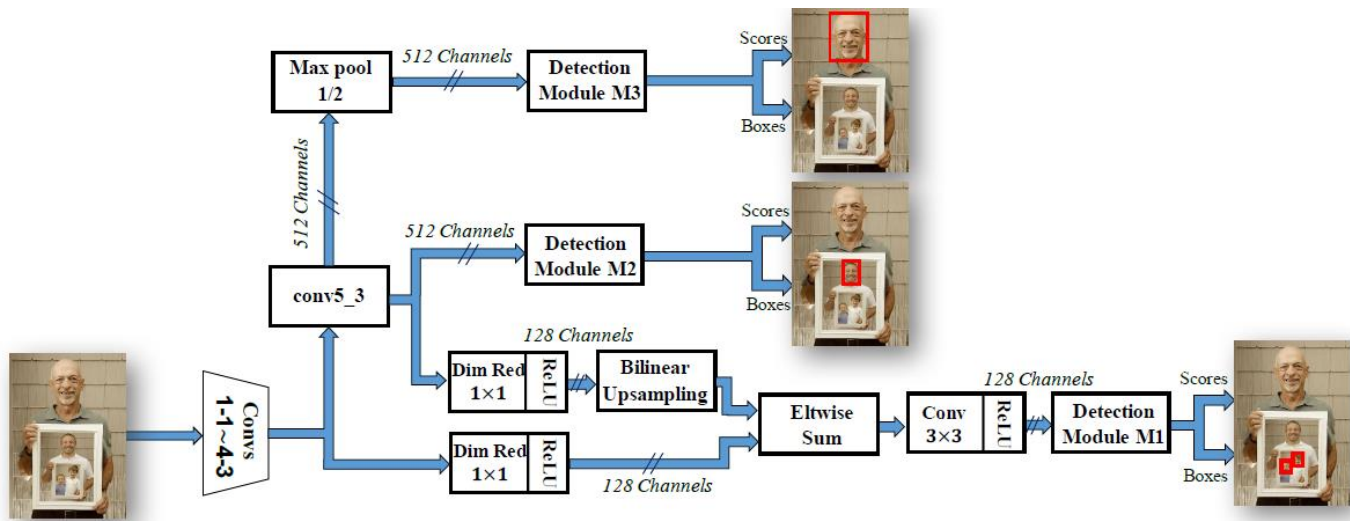
- 策略二：多尺度

- 基本思路

- (2) 尺度一致性：在合适的尺度上操作相应的对象

- Anchor、人脸尺度、神经元感受野大小一致

- SSH (ICCV 2017), S<sup>3</sup>FD (ICCV 2017), ScaleFace (arxiv 1706.02863), PyramidBox (arXiv 1803.07737): 在不同分辨率的特征图上设置不同大小的Anchor, 分别匹配大小相近的人脸



# 增强人脸检测鲁棒性的策略

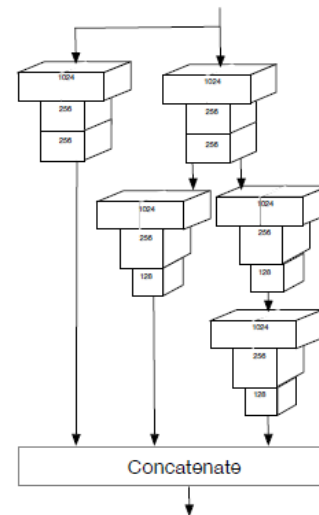
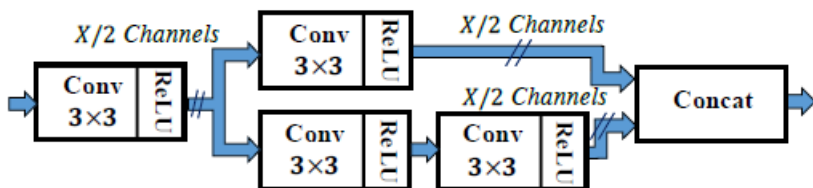
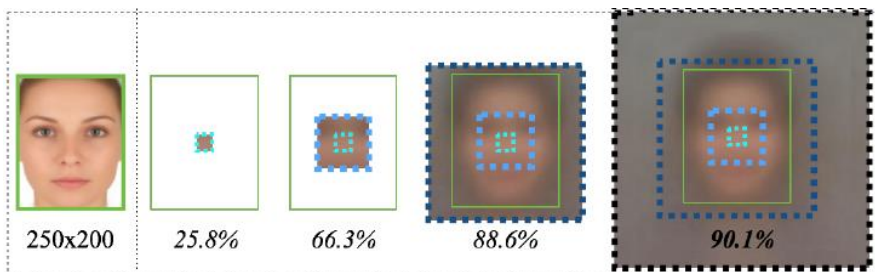
- 策略三：上下文

- 基本思路：利用人脸之外的信息

- (1) 朴素上下文：更大的感受野，更多的特征

- 比较直接粗暴，用来补充信息，但并不对补充的信息本身做特定的约束和设计，一般只能从结果进行解释
- 可以看成是多尺度特征融合

- TinyFace (CVPR 2017), SSH (ICCV 2017), PyramidBox (arXiv 1803.07737)



# 增强人脸检测鲁棒性的策略

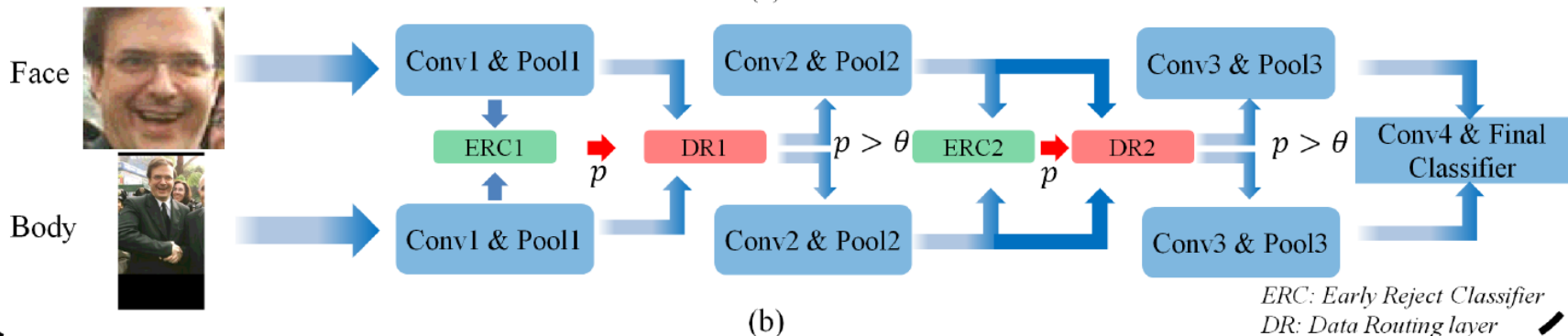
- 策略三：上下文

- 基本思路：利用人脸之外的信息

- (2) 物体上下文：头肩信息，人体信息

- 根据人脸的特殊性显式定义上下文的内容

- ICS (ICCV 2017): 利用人体信息来推测人脸信息, 利用了MS COCO上标注了人体关节点的数据
- CMS-RCNN (arXiv 1606.05413), PyramidBox (arXiv 1803.07737): 假定头肩/人体和人脸的相对位置和比例等固定 (利用了一定的先验知识, 但假设实际上并不成立, 只是人脸检测确实并不要求精确的头肩/人体信息)



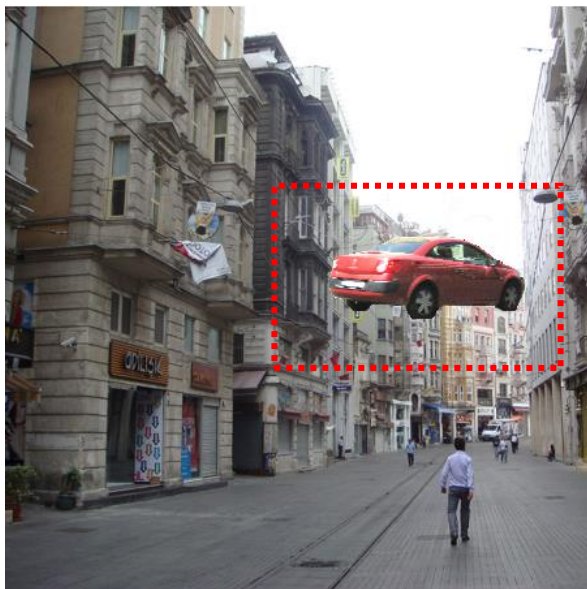
# 增强人脸检测鲁棒性的策略

- 策略三：上下文

- 基本思路：利用人脸之外的信息

- (3) 关系上下文：场景，互动，各种关联关系

- 在人脸检测（单类目标）中使用较少，多见于通用目标检测相关的工作



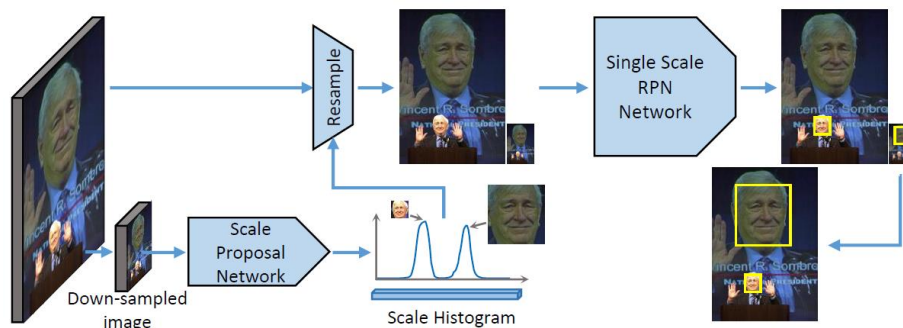
# 增强人脸检测鲁棒性的策略

## • 策略四：Anchor适配

### • 基本思路：更多和更合适的Anchor

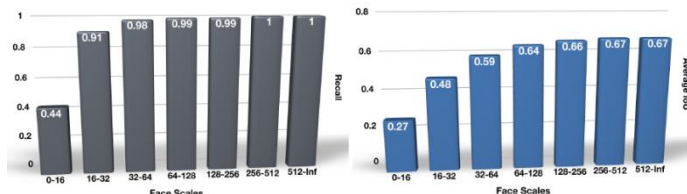
#### (1) 尺度一致性：Anchor，人脸，神经元感受野

- 补充：SAFD (CVPR 2017)，多尺度输入+单尺度模型



#### (2) 增大排布密度：更多的Anchor可以增大匹配机会和提高匹配质量

- Single-shot detector vs Two-stage detector
- FaceBox (IJCB 2017), S<sup>3</sup>FD (ICCV 2017): 保证单位面积上不同尺度Anchor的数量相同；降低匹配的交并比阈值（质量换数量）
- EMO (CVPR 2018): 具体分析了匹配质量的问题



# 增强人脸检测鲁棒性的策略

- 小结
  - Coarse-to-Fine
    - 将解决问题的路径拆解成多段，设立中继目标
  - 多尺度
    - 尺度互补性，尺度一致性
  - 上下文
    - 朴素上下文，物体上下文，关系上下文
  - Anchor适配
    - 尺度一致性，增大排布密度



# 谢谢！



邬书哲

中国科学院 计算技术研究所 VIPL研究组

[shuzhe.wu@vip1.ict.ac.cn](mailto:shuzhe.wu@vip1.ict.ac.cn)

论文: <https://arxiv.org/abs/1804.06039>

Demo程序: <https://github.com/Jack-CV/PCN>



更多分享请关注极市平台公众号



更多视觉算法合作请联系小助手