

人脸检测



主讲人 **张士峰**

中国科学院自动化研究所
模式识别国家重点实验室



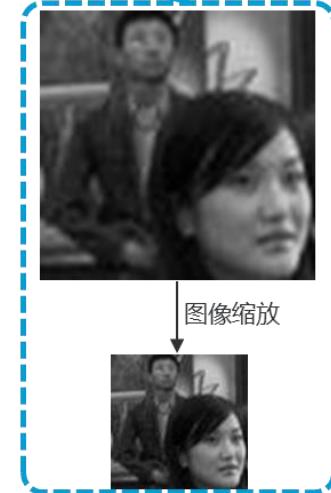


内容回顾：传统人脸检测算法

- 利用手工特征+分类器，以滑窗方式在图像金字塔上遍历所有位置和大小，进行人脸检测



灰度化



输出结果

人脸结果

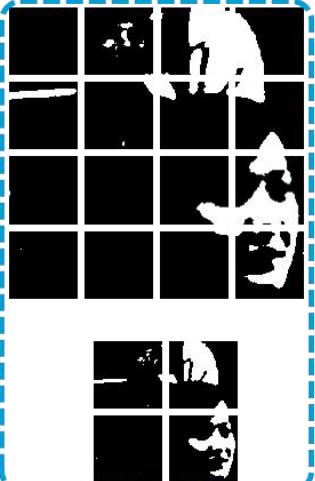


AdaBoost
分类器2

4个人脸候选



特征提取



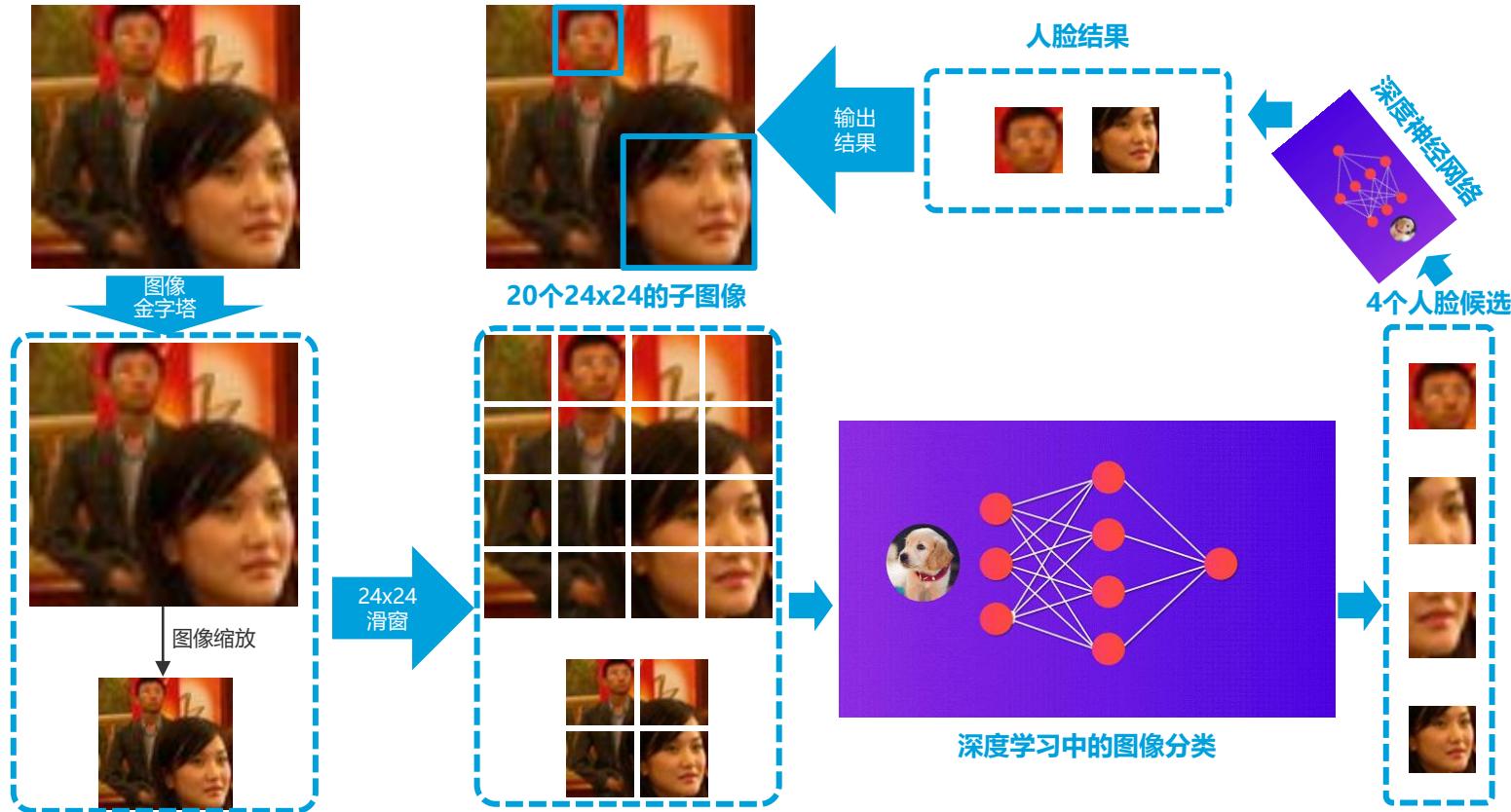
AdaBoost
分类器1





内容回顾：深度学习早期人脸检测算法

- 在传统人脸检测算法的流程中，把**手工设计的特征和分类器**变成**深度学习中的特征和分类器**





内容回顾：深度学习早期人脸检测算法

CascadeCNN

+关键点分支

MTCNN

+角度预测分支

PCN

滑窗	滑窗操作遍历所有的位置
金字塔	图像金字塔遍历所有的大小
级联	3个级联的阶段，人脸数量从多到少，人脸难度从易到难，网络结构从简单到复杂
深度学习	利用深度学习中的神经网络进行特征提取+分类器+矫正器+其他
优点	在简单场景下，有着满足实际需求的精度，具备CPU实时的速度
缺点	<p>局部最优：三个独立的阶段，容易取得局部最优而非全局最优</p> <p>训练繁琐：训练不是端到端的，每个阶段单独处理，非常繁琐</p> <p>检测速度：检测速度与图像上人脸的数量高度相关</p> <p>检测精度：在复杂的场景下，检测精度不能够满足性能的需求</p>
总结	深度学习早期人脸检测算法的代表，开创了深度学习时代下，人脸检测的一个派系，很多实际场景中都在使用该类型的算法





内容回顾：深度学习后期人脸检测算法

- 深度学习后期人脸检测算法：对**通用物体检测算法**进行**相应改进**，应用于**人脸检测领域**

高效率的人脸检测算法

- 基础网络为专门设计的轻量级的网络结构
- 在实际场景中，检测大于30个像素的人脸，有着满足需求的检测精度
- 能够在资源受限的前端设备（CPU、ARM、FPGA等）上实时的运行
- 追求检测速度和检测精度的平衡，满足实用性

高精度的人脸检测算法

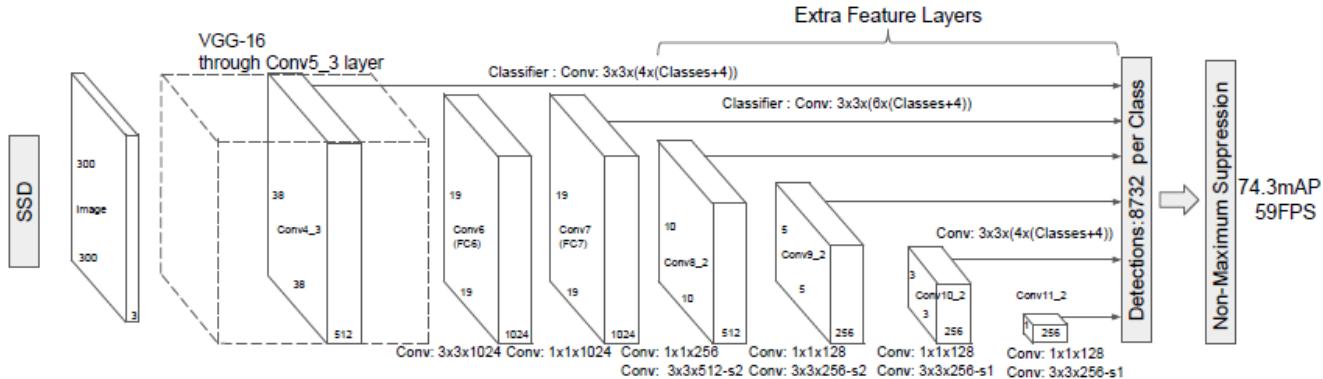
- 基础网络为重量级的VGG16或ResNet-50/101/152等
- 在复杂场景下，有着非常高的检测精度，非常小的人脸也能检测
- 可以在高性能的GPU设备上实时的运行
- 追求极致的检测精度，检测速度可以不考虑





内容回顾：高效率的深度学习后期人脸检测算法

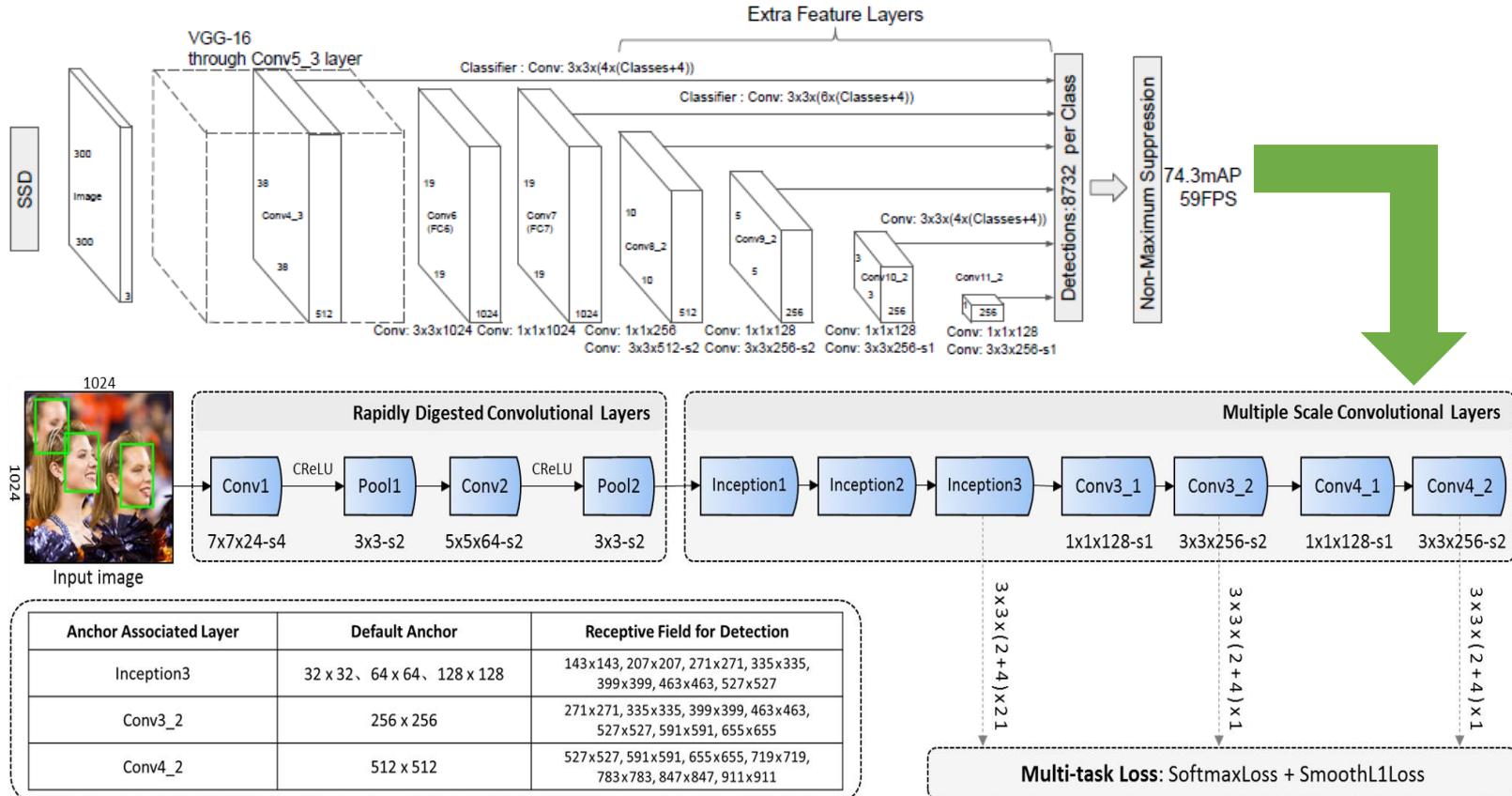
- 高效率的深度学习后期人脸检测算法几乎都是基于通用物体检测算法SSD进行的改进





内容回顾：高效率的深度学习后期人脸识别算法FaceBoxes

- 高效率的深度学习后期人脸识别算法几乎都是基于通用物体检测算法SSD进行的改进





目录

-  **人脸检测概述**
-  **传统人脸检测算法**
-  **深度学习早期人脸检测算法**
-  **深度学习后期人脸检测算法**





深度学习后期人脸检测算法

高效率的人脸检测算法

- 基础网络为专门设计的轻量级的网络结构
- 在实际场景中，检测大于30个像素的人脸，有着满足需求的检测精度
- 能够在资源受限的前端设备（CPU、ARM、FPGA等）上实时的运行
- 追求检测速度和检测精度的平衡，满足实用性

高精度的人脸检测算法

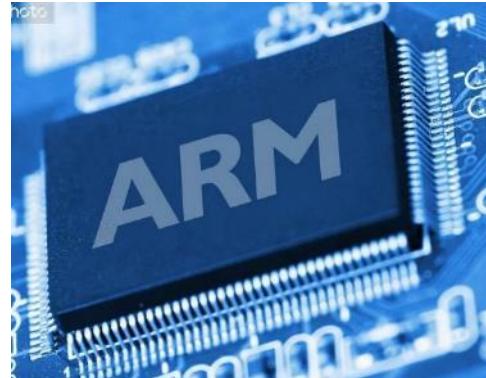
- 基础网络为重量级的VGG16或ResNet-50/101/152等
- 在复杂场景下，有着非常高的检测精度，非常小的人脸也能检测
- 可以在高性能的GPU设备上实时的运行
- 追求极致的检测精度，检测速度可以不考虑





深度学习后期人脸检测算法: 高效率 vs 高精度

- 高效率算法: 运行在**计算资源受限**的前端设备, 如CPU、ARM、FPGA



- 高精度算法: 运行在**计算资源充足**的后端设备, 如英伟达显卡2080Ti、RTX





深度学习后期人脸检测算法: 高效率 vs 高精度

- 高效率算法: 由于资源受限, 只需要在**正常场景**下满足精度需求 (人脸 > 30个像素, 背景比较简单)



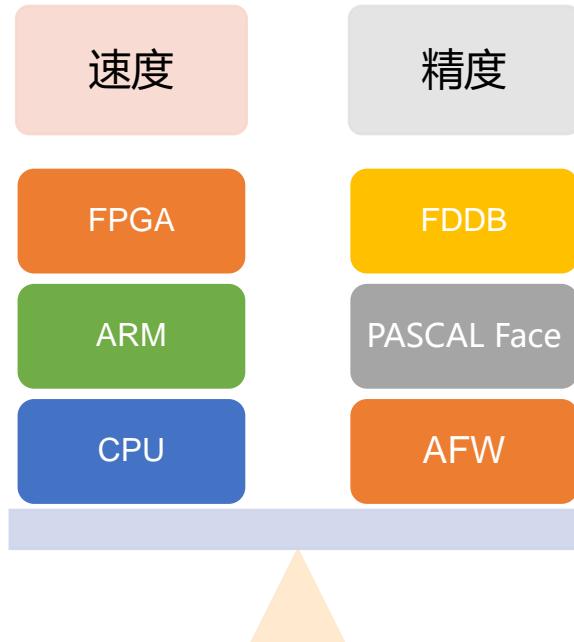
- 高精度算法: 由于资源充足, 需要在**所有场景**下都满足精度需求 (任何尺度人脸, 背景非常复杂)



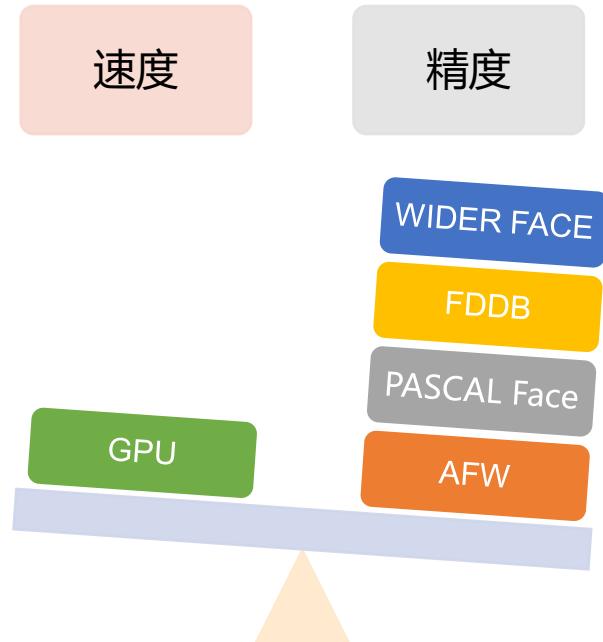


深度学习后期人脸检测算法: 高效率 vs 高精度

高效率算法: 实用性

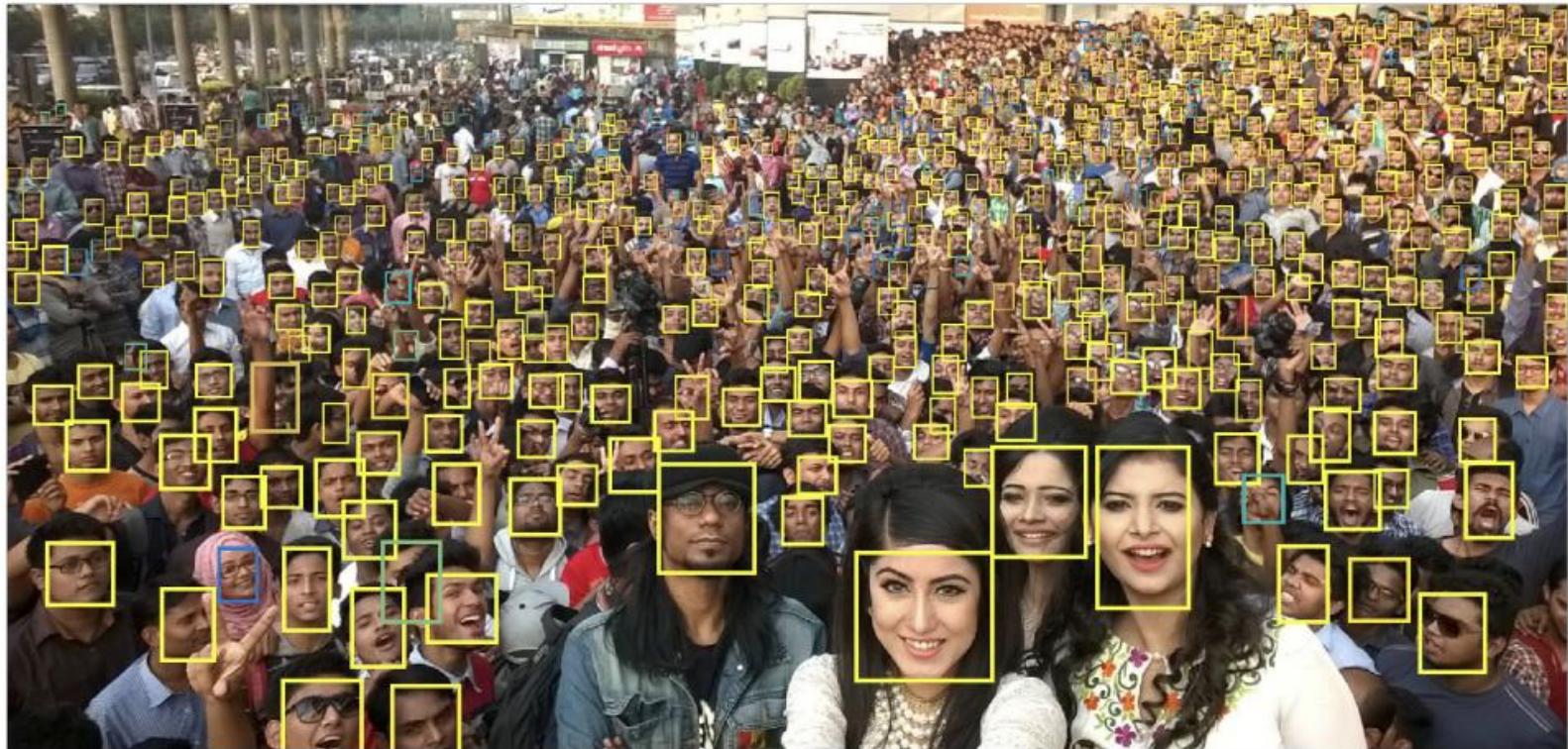


高精度算法: 竞争性





物体检测的难点问题分析



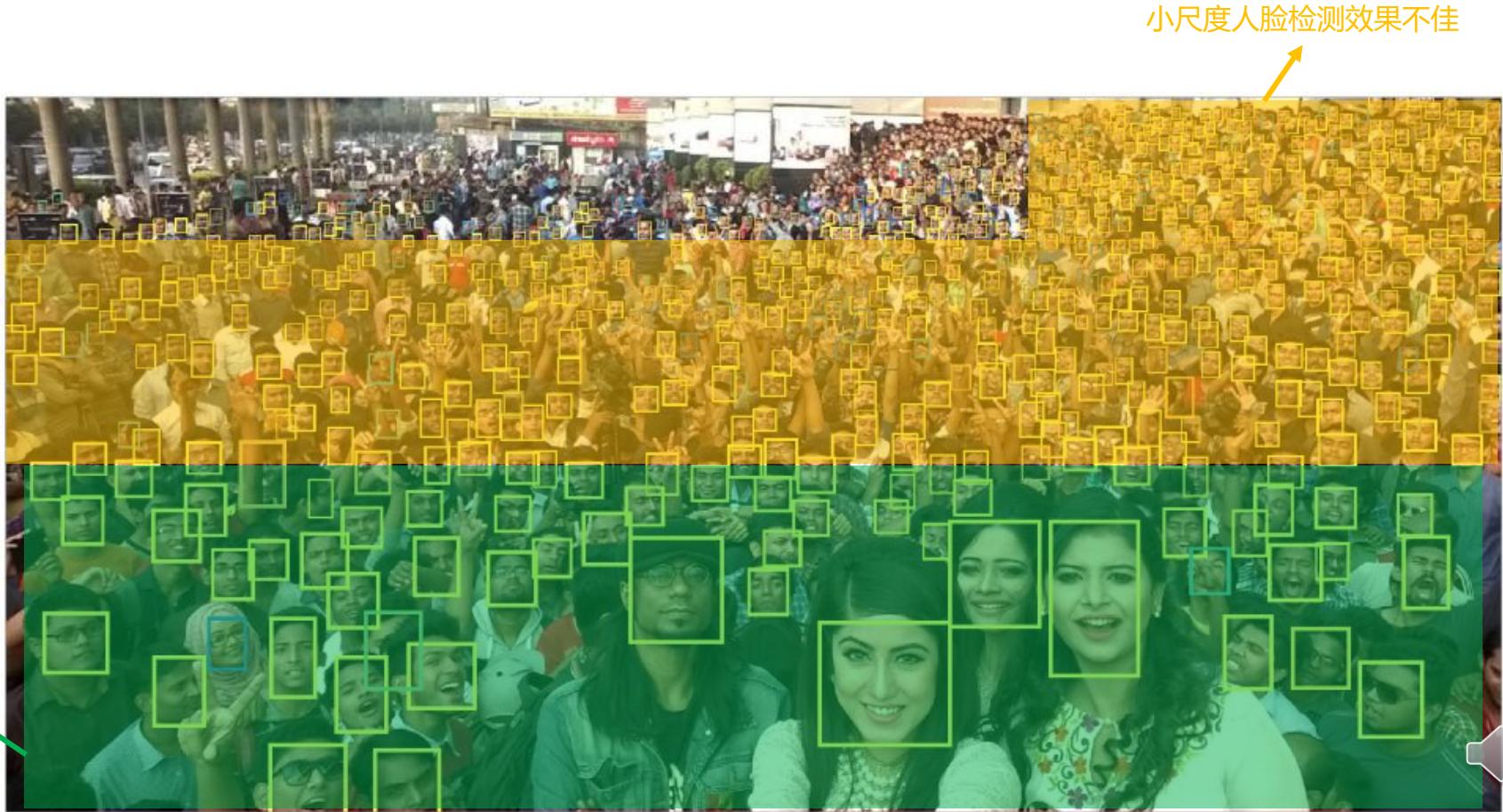


物体检测的难点问题分析





物体检测的难点问题分析

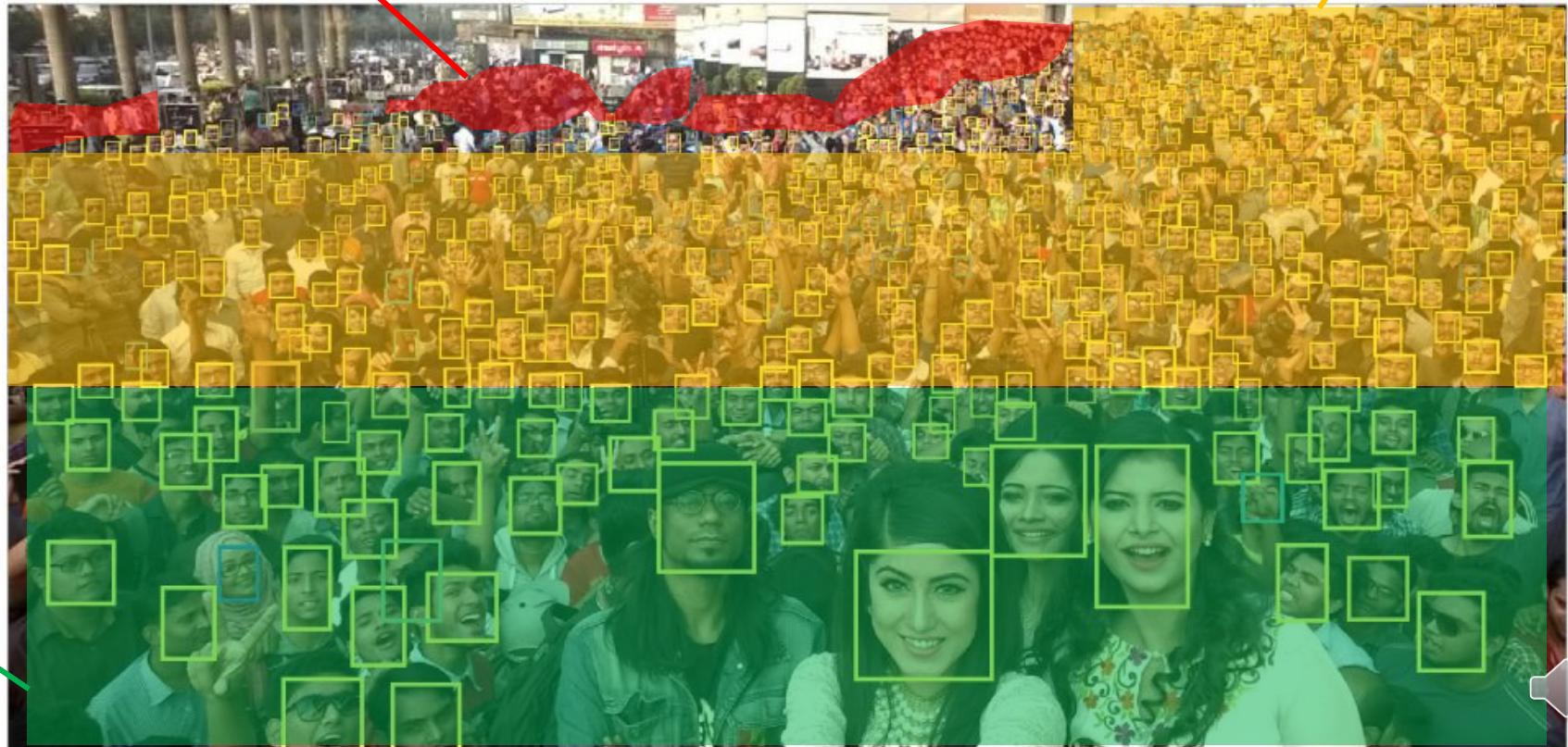




物体检测的难点问题分析

极小尺度人脸几乎都漏检

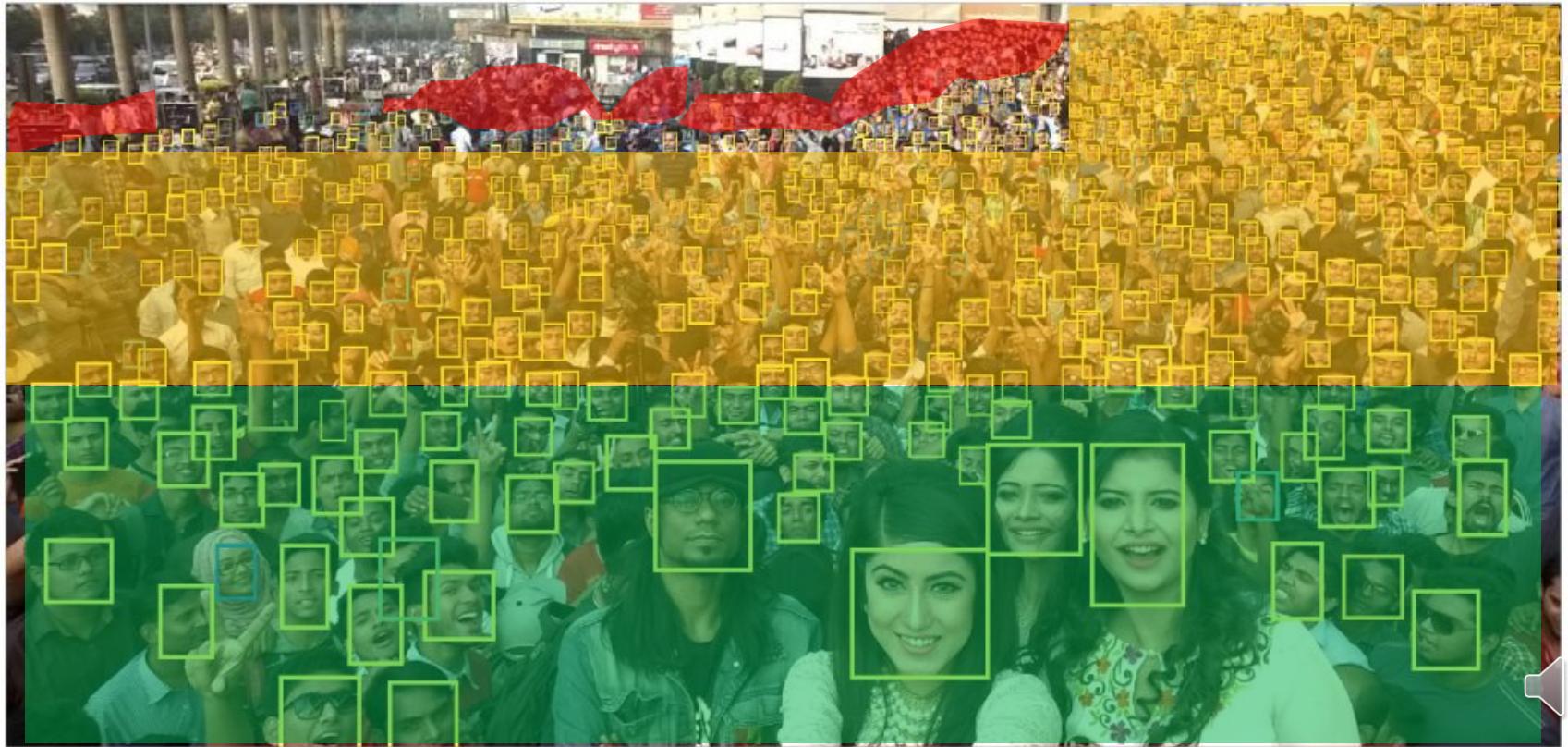
小尺度人脸检测效果不佳





物体检测的难点问题分析

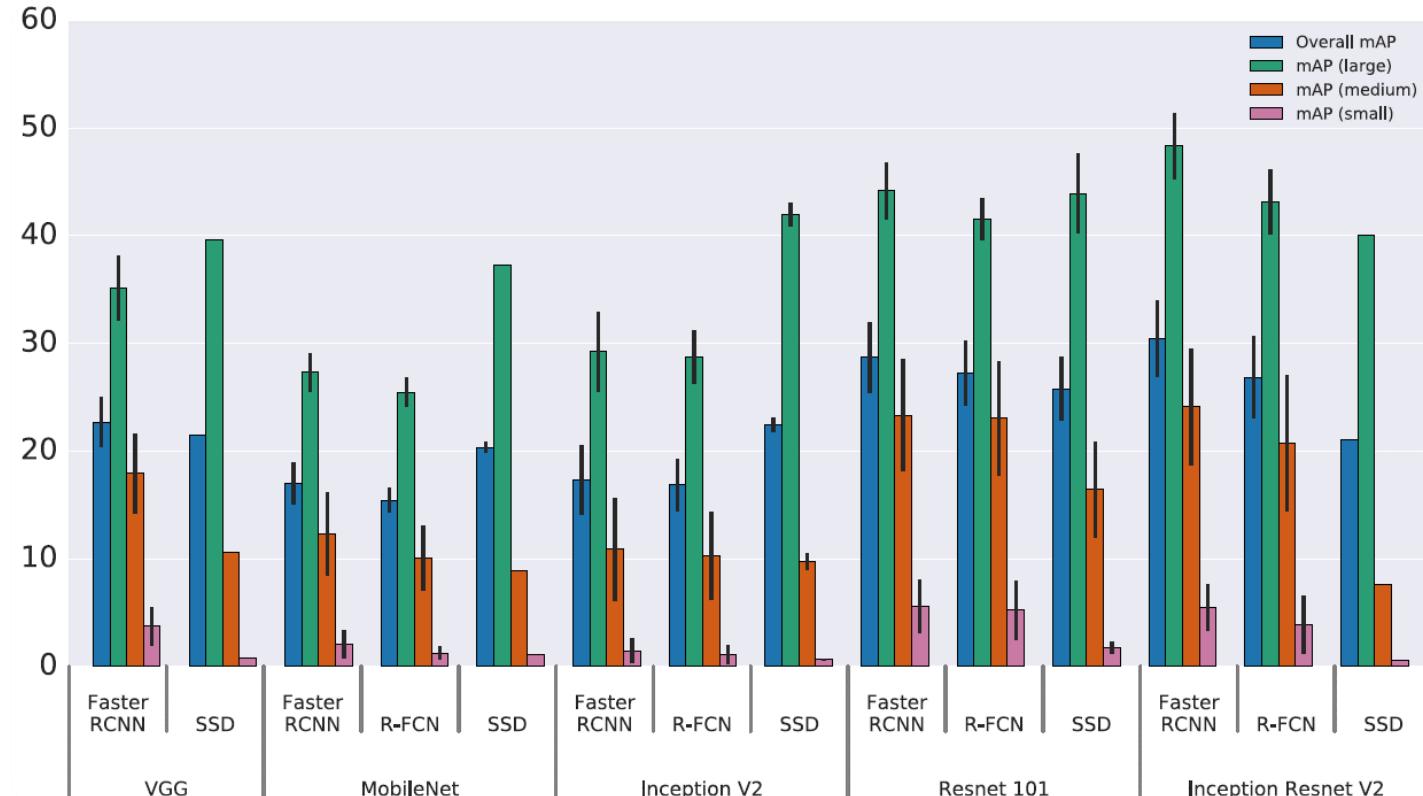
- 小尺度物体检测是一个难点问题





难点问题：小尺度物体检测

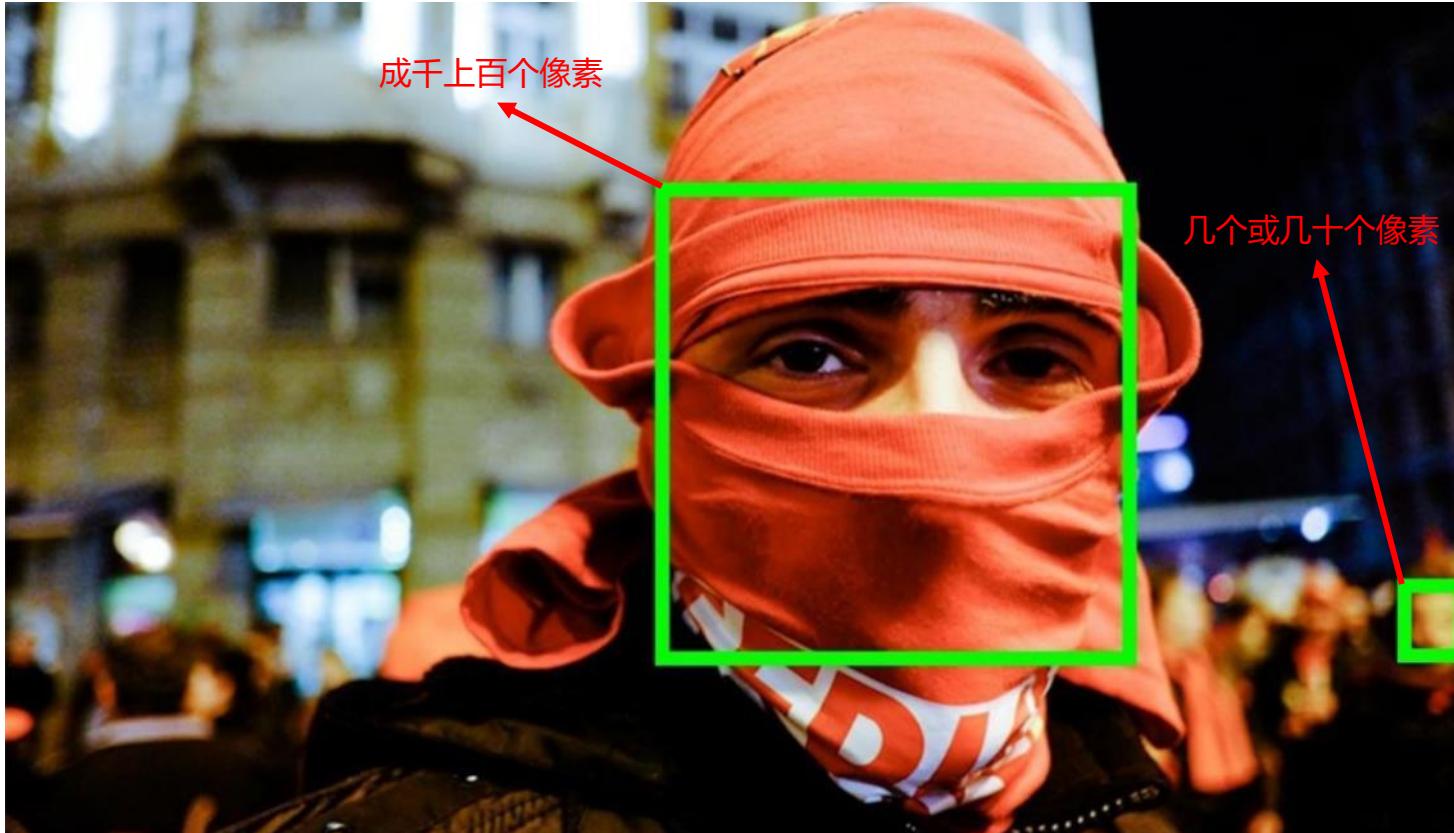
- 随着物体尺度的变小，各个检测算法基于各个网络，检测精度都会急剧下降





小尺度物体为何难？

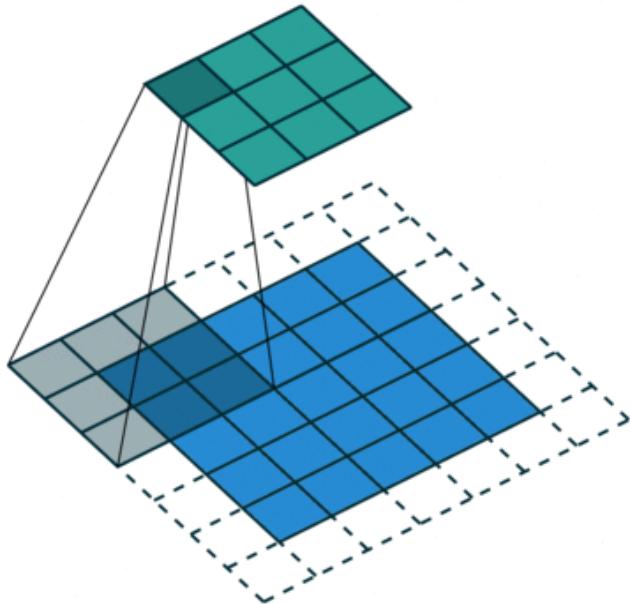
- 小尺度物体的像素个数（原始信息）非常少



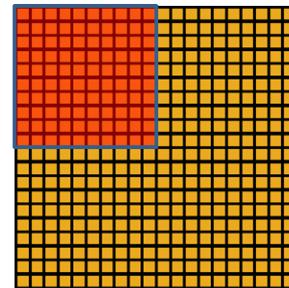


小尺度物体为何难？

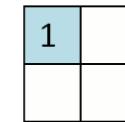
- 卷积网络中存在下采样操作（卷积或池化），导致小尺度物体的信息容易丢失



下采样倍数为2的卷积操作



Convolved
feature



Pooled
feature

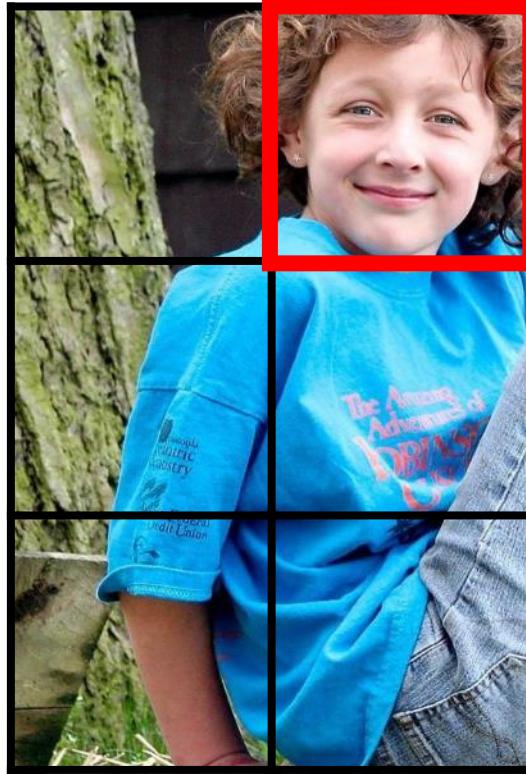
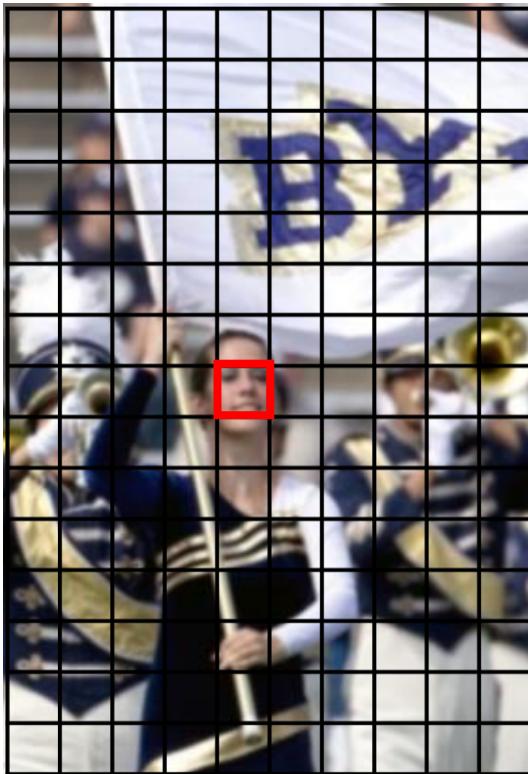
下采样倍数为10的池化操作





小尺度物体为何难？

- 小尺度物体的搜索空间巨大





小尺度物体之人脸

- 小尺度物体在人脸检测任务中很常见
- 在WIDER FACE数据集中，50%人脸的高度只有10~50个像素
- 2017年前，WIDER FACE中小人脸上的检测精度（AP）只有60%+

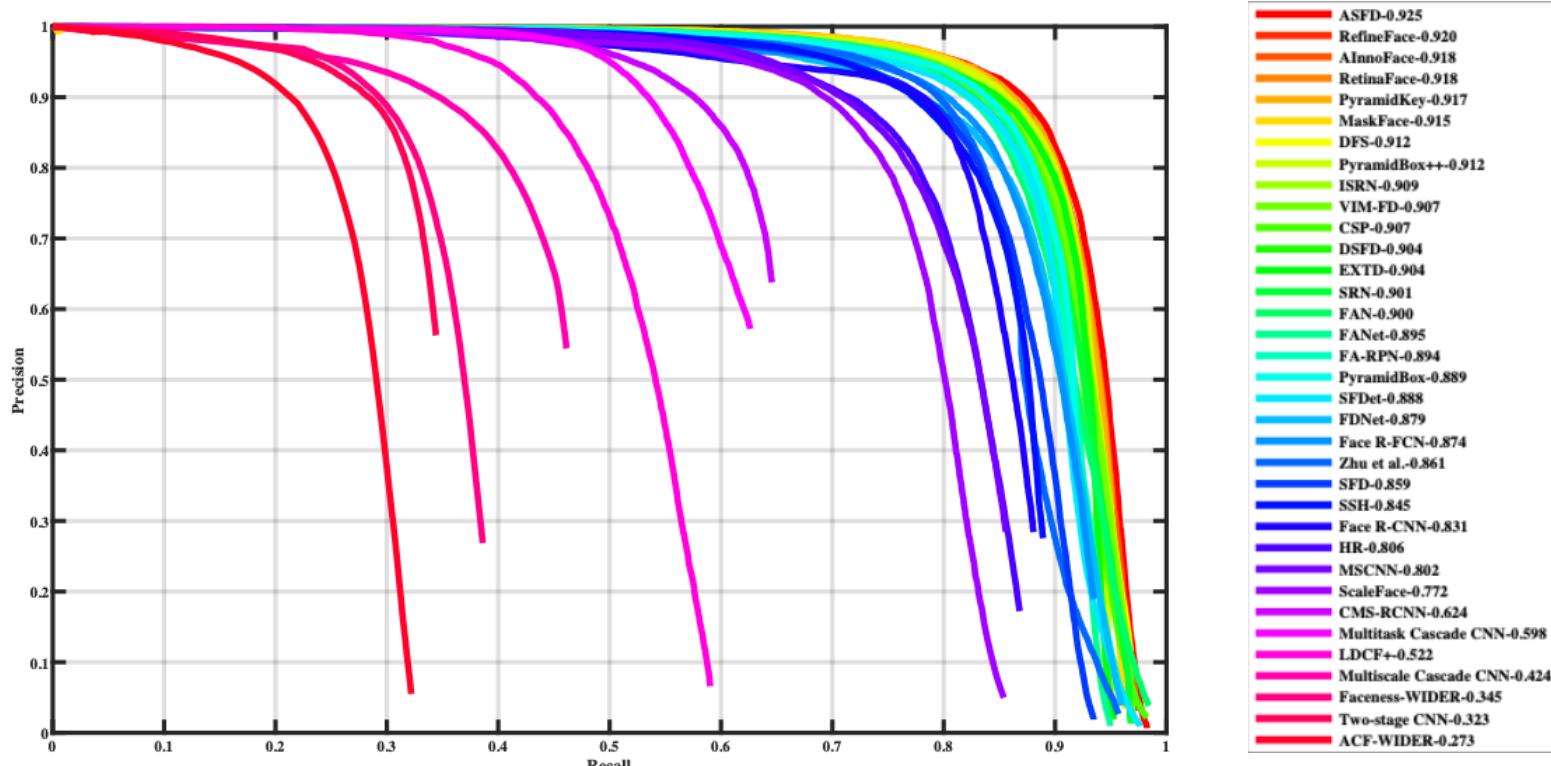
Dataset	Training		Testing		Height			Properties		
	#Image	#Face	#Image	#Face	10-50 pixels	50-300 pixels	<1	Occlusion labels	Event labels	Pose labels
AFW [39]	-	-	0.2k	0.47k	12%	70%	18%	-	-	✓
FDDDB [13]	-	-	2.8k	5.1k	8%	86%	6%	-	-	-
PASCAL FACE [32]	-	-	0.85k	1.3k	41%	57%	2%	-	-	-
IJB-A [14]	16k	33k	8.3k	17k	13%	69%	18%	-	-	-
MALF [34]	-	-	5.25k	11.9k	N/A	N/A	N/A	✓	-	✓
WIDER FACE	16k	199k	16k	194k	50%	43%	7%	✓	✓	✓





小尺度人脸检测：研究进展

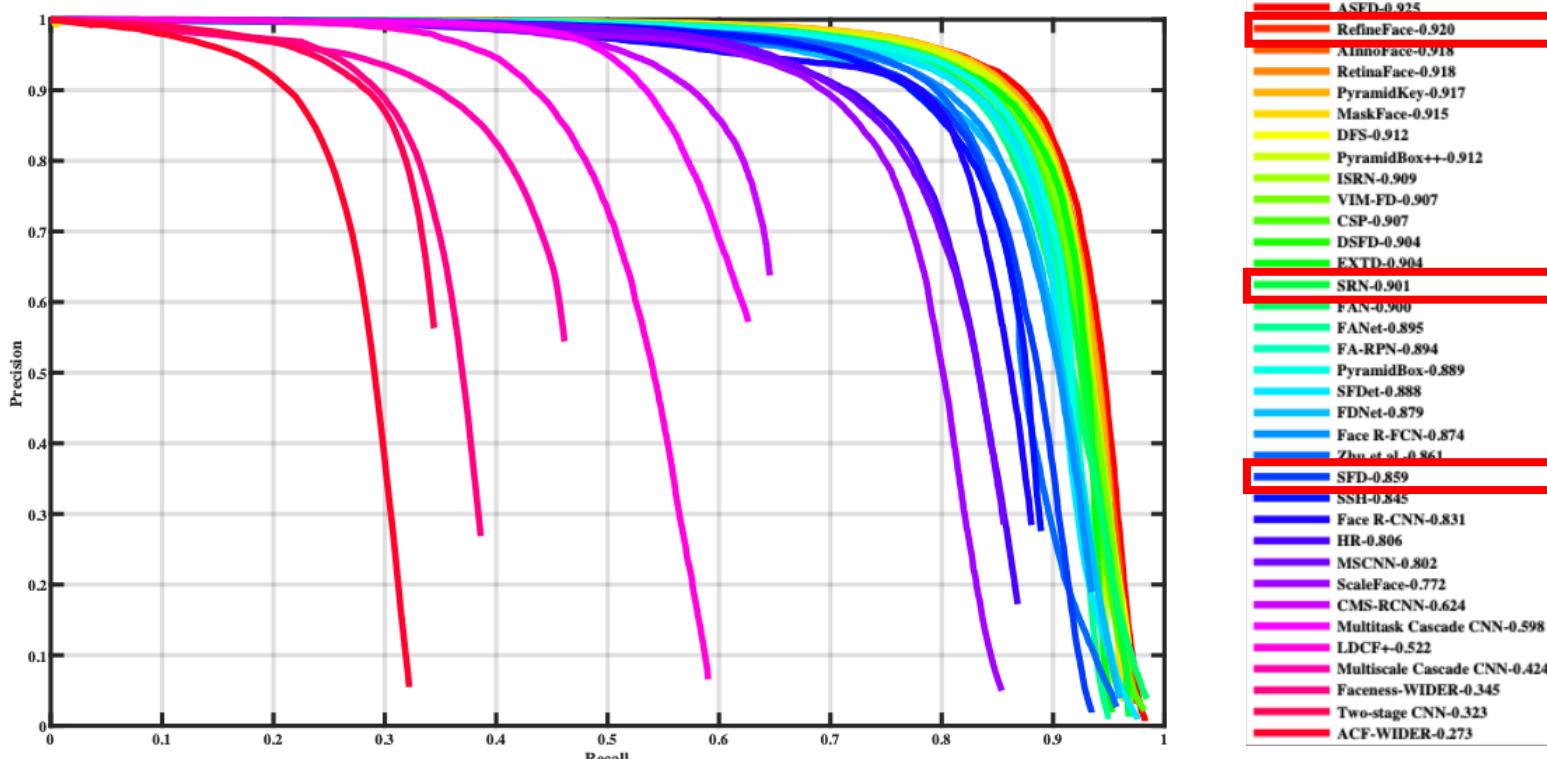
- 针对小尺度人脸的特点，在框架策略、数据增广、损失函数等方面进行针对性改进
- 自从2016年出现到目前为止，WIDER FACE的Hard子集上的性能已经从60%+刷到了92.5%





小尺度人脸检测：研究进展

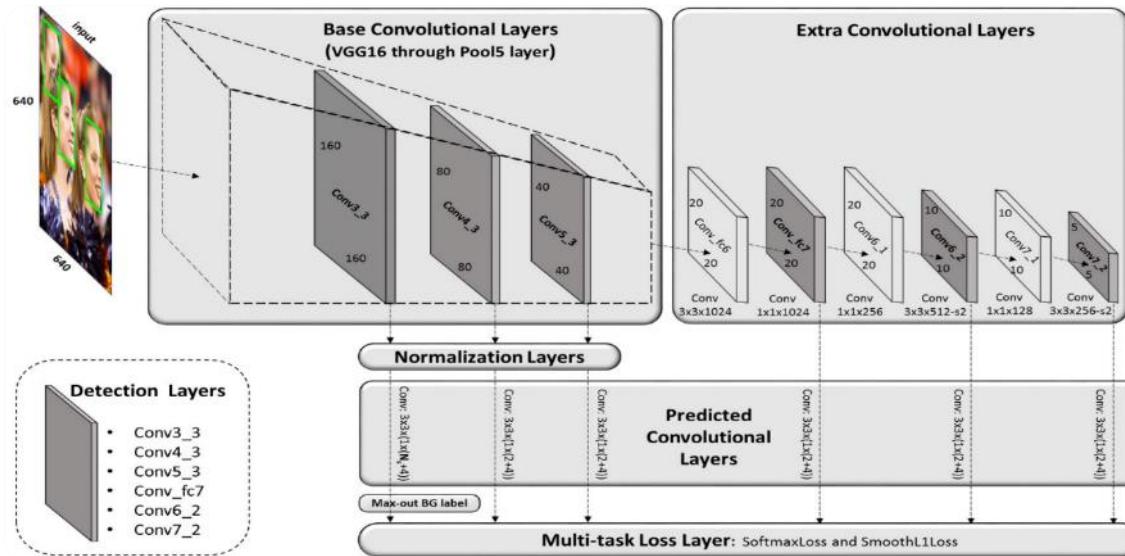
- 针对小尺度人脸的特点，在框架策略、数据增广、损失函数等方面进行针对性改进
- 自从2016年出现到目前为止，WIDER FACE的Hard子集上的性能已经从60%+刷到了92.5%





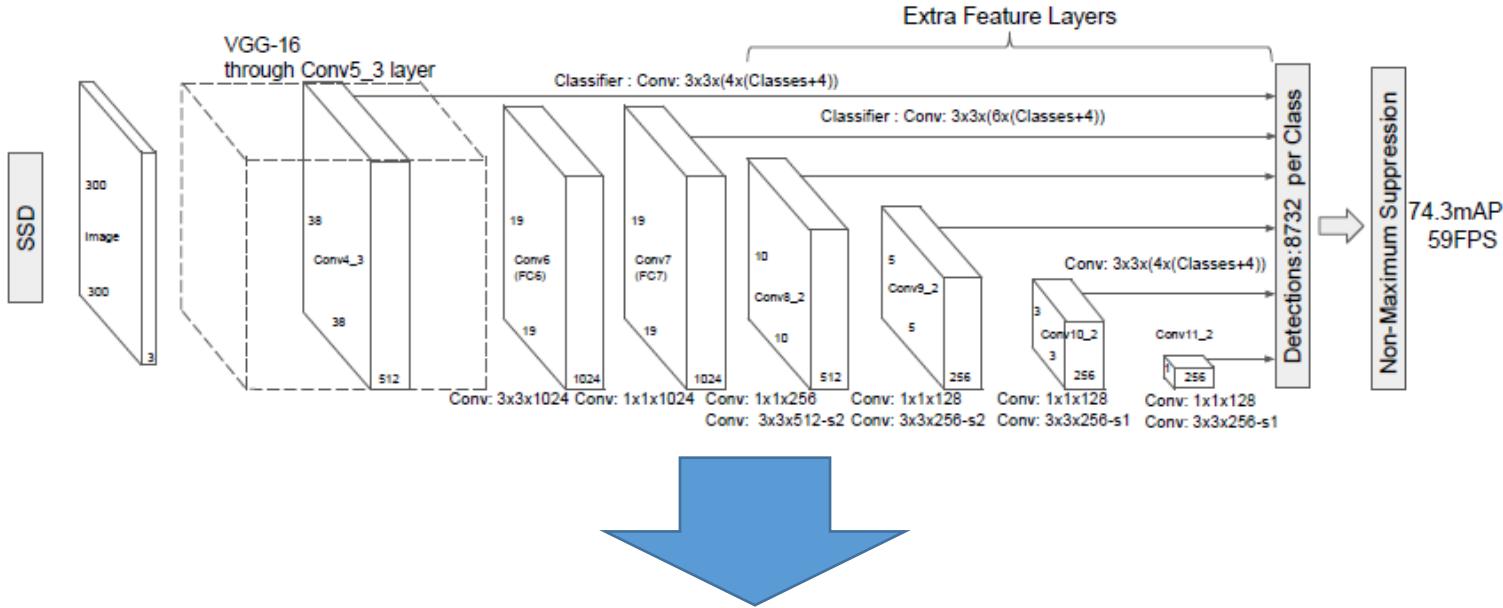
高精度的人脸检测算法：S³FD

- 基于通用物体检测算法SSD所设计的一个高精度人脸检测算法
- 主要改进点是：①尺度上公平的检测框架；②尺度补偿的锚框匹配策略；③背景标签输出最大化操作
- 发表时在WIDER FACE上取得最好结果，对于640x480分辨率的图像，能够在GPU上实时运行





高精度的人脸检测算法S³FD：SSD效果不佳

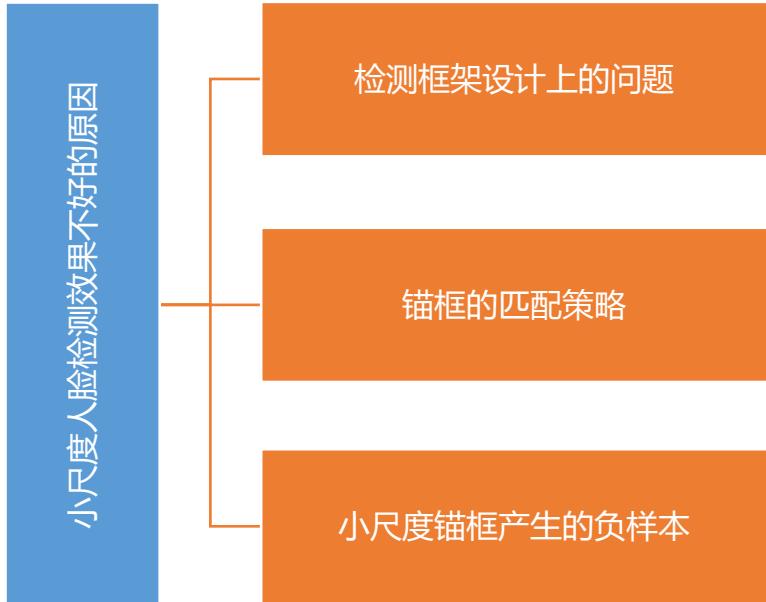


- 对通用物体检测算法SSD进行简单改进（数据增广，锚框比例和尺度等）来做人脸检测任务
- 使用WIDER FACE数据集的训练集进行训练，并在WIDER FACE数据集的验证集进行测试
- 小尺度人脸检测效果不理想，在WIDER FACE验证集的Hard子集中只有50%左右的精度



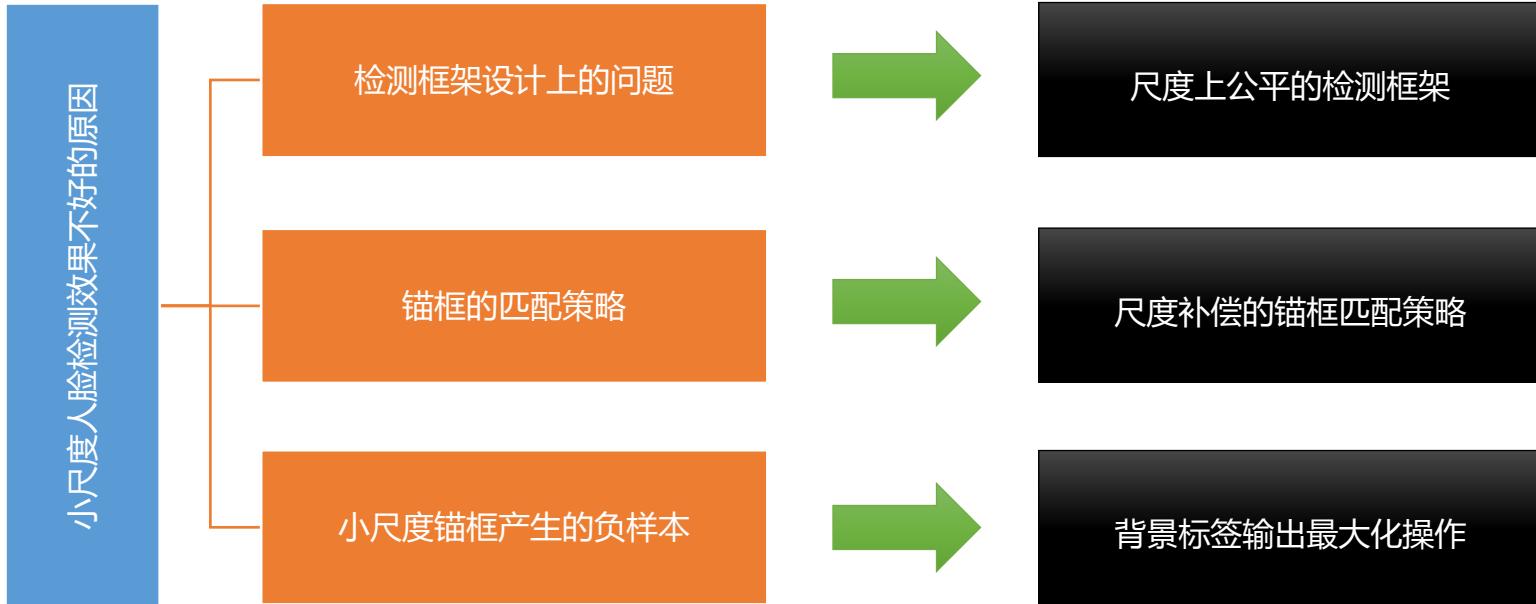


高精度的人脸检测算法S³FD：背后原因





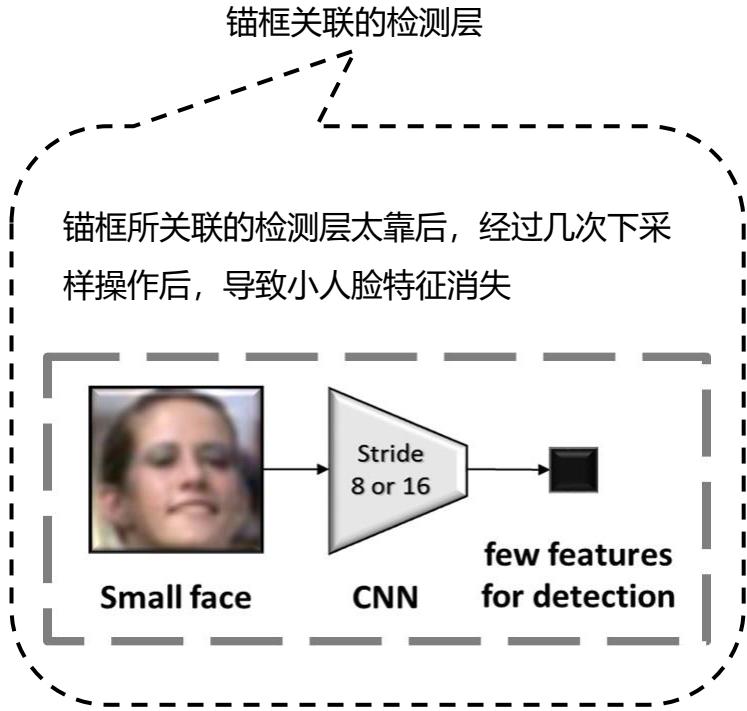
高精度的人脸检测算法S³FD：解决方案





高精度的人脸检测算法S³FD：背后原因①

- 检测框架设计上的问题



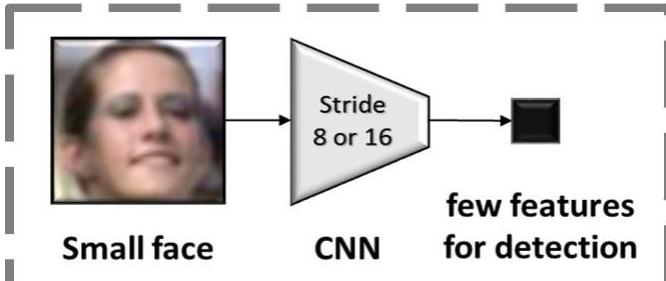


高精度的人脸检测算法S³FD：背后原因①

■ 检测框架设计上的问题

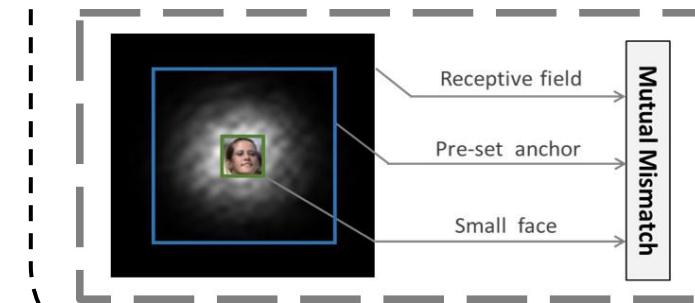
锚框关联的检测层

锚框所关联的检测层太靠后，经过几次下采样操作后，导致小人脸特征消失



锚框大小

锚框大小设计的不合适，导致锚框大小、小人脸尺度、有效感受野大小互不匹配





高精度的人脸检测算法S³FD：解决方案①

- 尺度上公平的检测框架

Detection Layer	Stride	Anchor	RF
conv3_3	4		48
conv4_3	8		108
conv5_3	16		228
conv_fc7	32		340
conv6_2	64		468
conv7_2	128		724

* Detection Layer指检测层，Stride指下采样倍数，RF指感受野





高精度的人脸检测算法S³FD：解决方案①

- 尺度上公平的检测框架

Detection Layer	Stride	Anchor	RF
conv3_3	4		48
conv4_3	8		108
conv5_3	16		228
conv_fc7	32		340
conv6_2	64		468
conv7_2	128		724

* Detection Layer指检测层，Stride指下采样倍数，RF指感受野





高精度的人脸检测算法S³FD：解决方案①

- 尺度上公平的检测框架

Detection Layer	Stride	Anchor	RF
conv3_3	4		48
conv4_3	8		108
conv5_3	16		228
conv_fc7	32		340
conv6_2	64		468
conv7_2	128		724

* Detection Layer指检测层，Stride指下采样倍数，RF指感受野





高精度的人脸检测算法S³FD：解决方案①

- 尺度上公平的检测框架

Detection Layer	Stride	Anchor	RF
conv3_3	4	16	48
conv4_3	8	32	108
conv5_3	16	64	228
conv_fc7	32	128	340
conv6_2	64	256	468
conv7_2	128	512	724

* Detection Layer指检测层，Stride指下采样倍数，RF指感受野





高精度的人脸检测算法S³FD：解决方案①

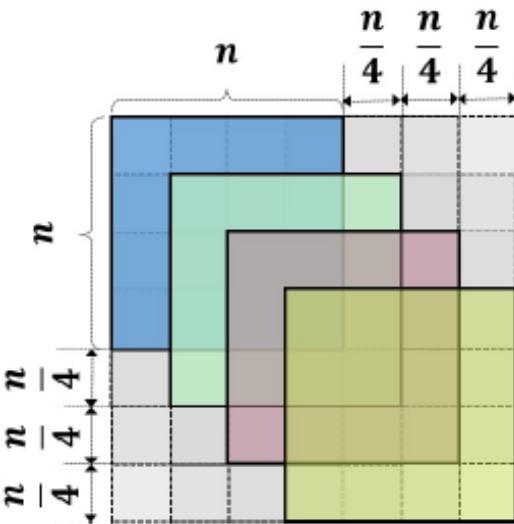
- 尺度上公平的检测框架

Detection Layer	Stride	Anchor	RF
conv3_3	4	16	48
conv4_3	8	32	108
conv5_3	16	64	228
conv_fc7	32	128	340
conv6_2	64	256	468
conv7_2	128	512	724

* Detection Layer指检测层，Stride指下采样倍数，RF指感受野

- 等比间隔铺设原则：不同尺度锚框，应有相同的铺设密度

等比铺设间隔原则





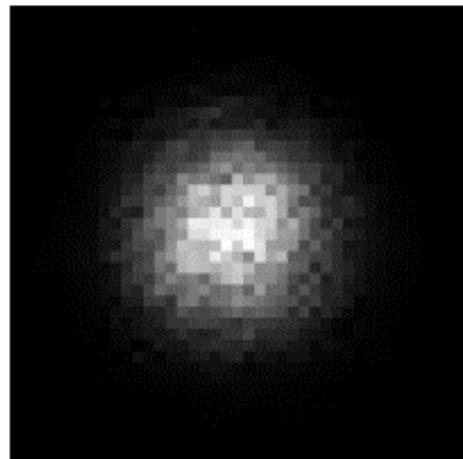
高精度的人脸检测算法S³FD：解决方案①

- 尺度上公平的检测框架

Detection Layer	Stride	Anchor	RF
conv3_3	4	16	48
conv4_3	8	32	108
conv5_3	16	64	228
conv_fc7	32	128	340
conv6_2	64	256	468
conv7_2	128	512	724

* Detection Layer指检测层，Stride指下采样倍数，RF指感受野

有效感受野理论



* 有效感受野是理论感受野中间的20%~40%

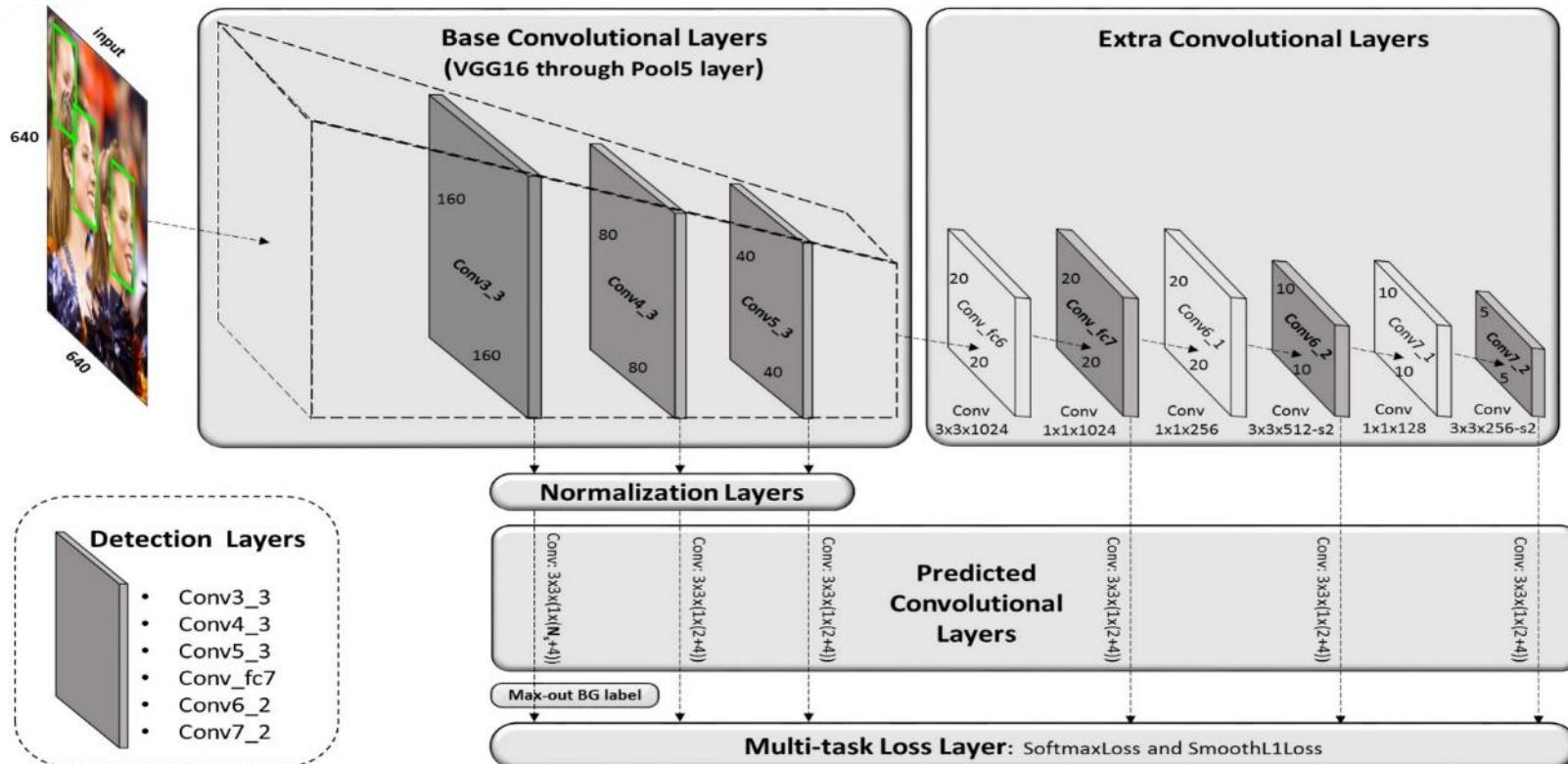
- 等比间隔铺设原则：不同尺度锚框，应有相同的铺设密度
- 有效感受野理论：锚框的大小应跟有效感受野相匹配





高精度的人脸检测算法S³FD：解决方案①

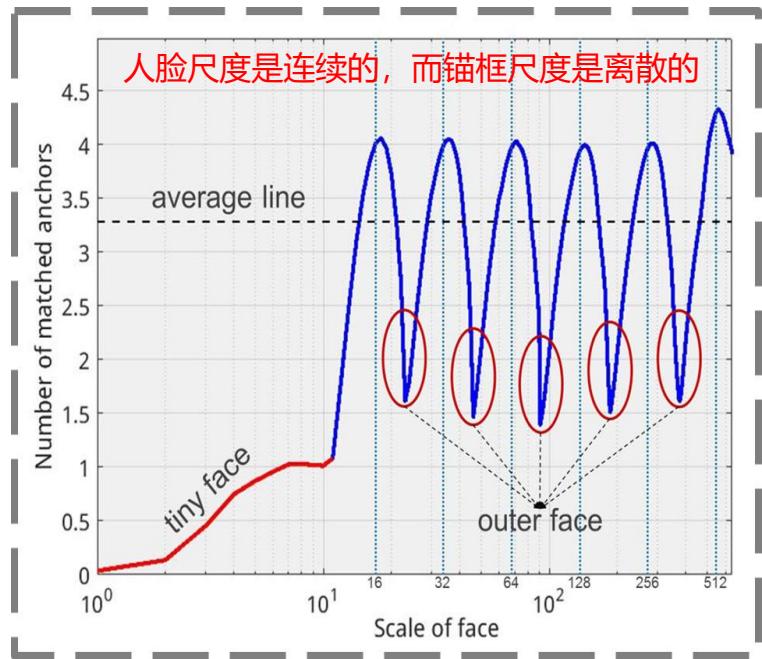
■ 尺度上公平的检测框架





高精度的人脸检测算法S³FD：背后原因②

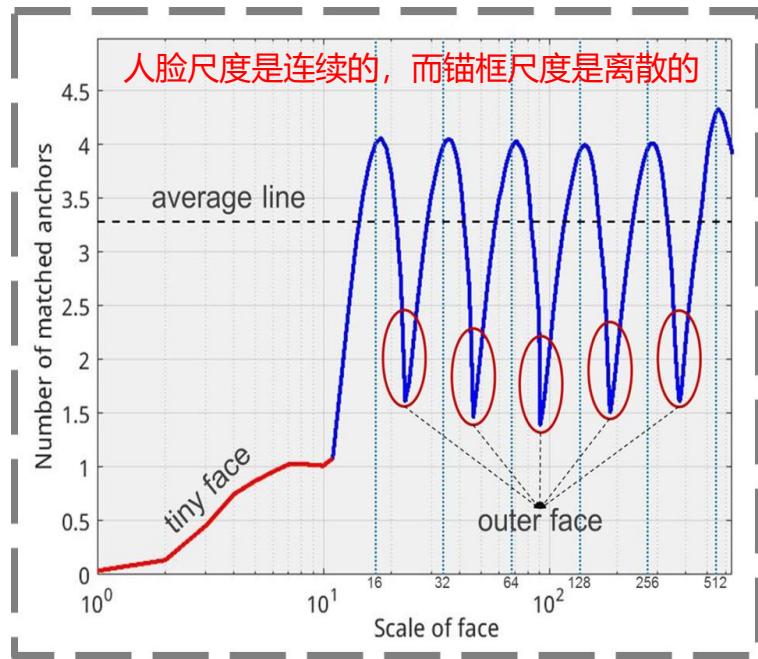
- 锚框匹配策略的问题



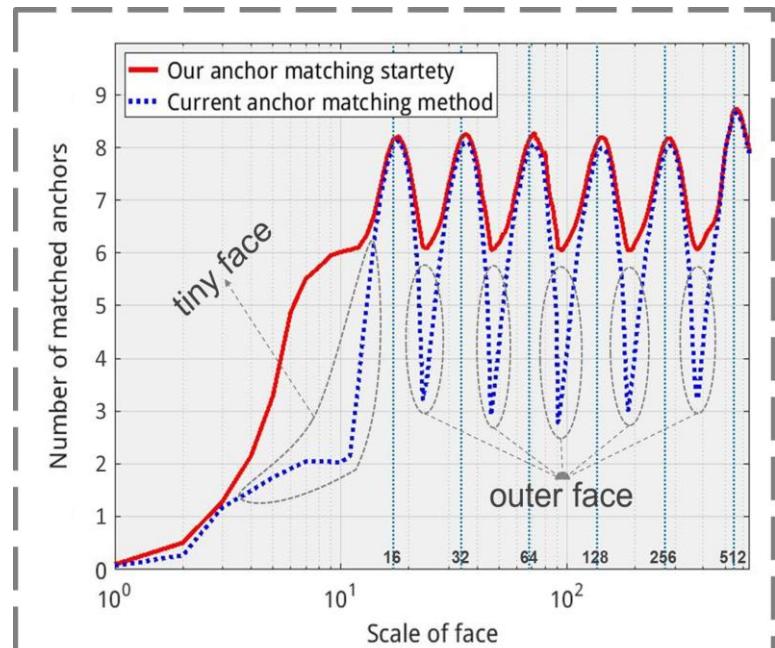


高精度的人脸检测算法S³FD：解决方案②

■ 尺度补偿的锚框匹配策略



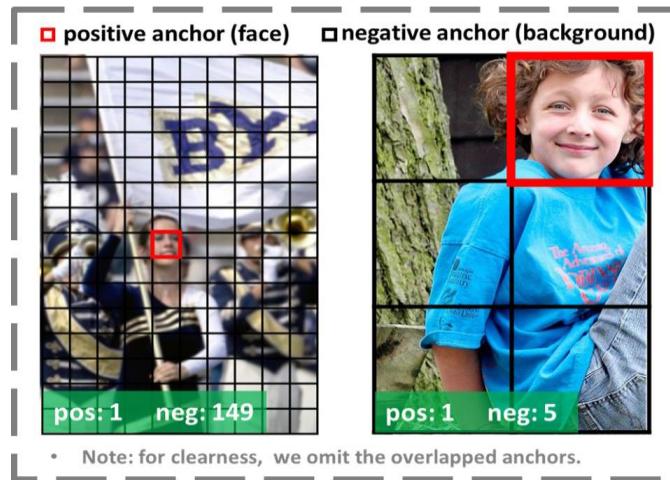
尺度补偿的锚框匹
配策略





高精度的人脸检测算法S³FD：背后原因③

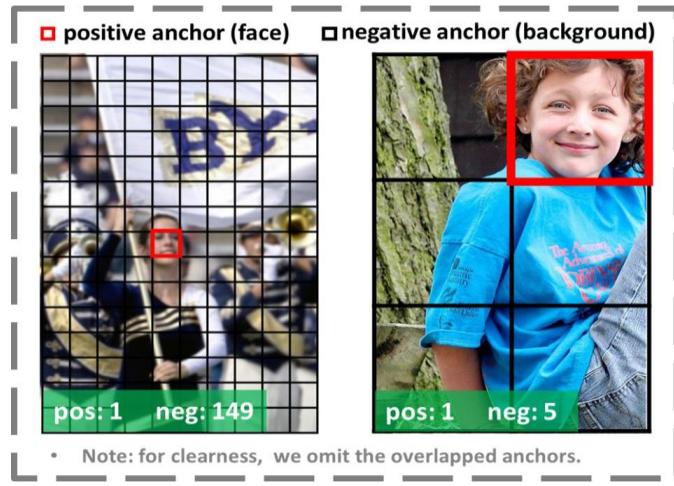
- 小尺度锚框产生的负样本





高精度的人脸检测算法S³FD：背后原因③

- 小尺度锚框产生的负样本



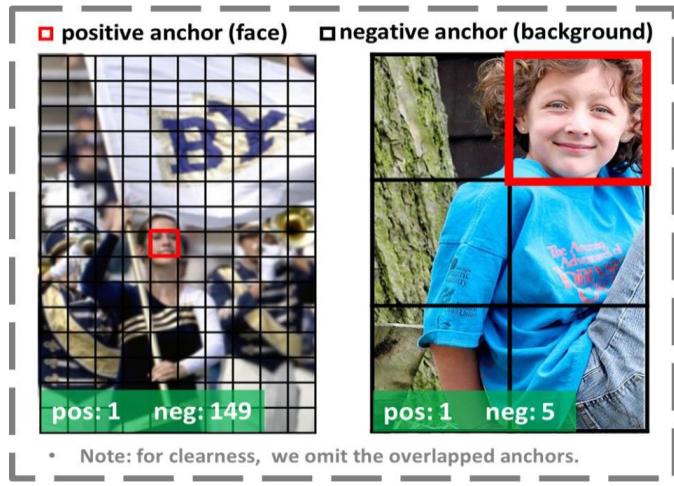
Position	Scale	Number	Percentage (%)
conv3_3	16	25600	75.02
conv4_3	32	6400	18.76
conv5_3	64	1600	4.69
conv_fc7	128	400	1.17
conv6_2	256	100	0.29
conv7_2	512	25	0.07





高精度的人脸检测算法S³FD：解决方案③

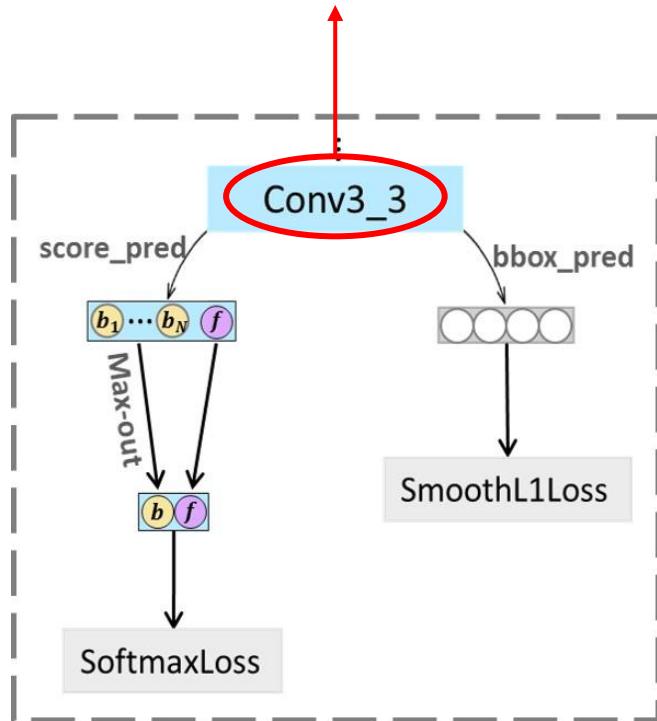
- 背景标签输出最大化操作



背景标签输出
最大化操作

Position	Scale	Number	Percentage (%)
conv3_3	16	25600	75.02
conv4_3	32	6400	18.76
conv5_3	64	1600	4.69
conv_fc7	128	400	1.17
conv6_2	256	100	0.29
conv7_2	512	25	0.07

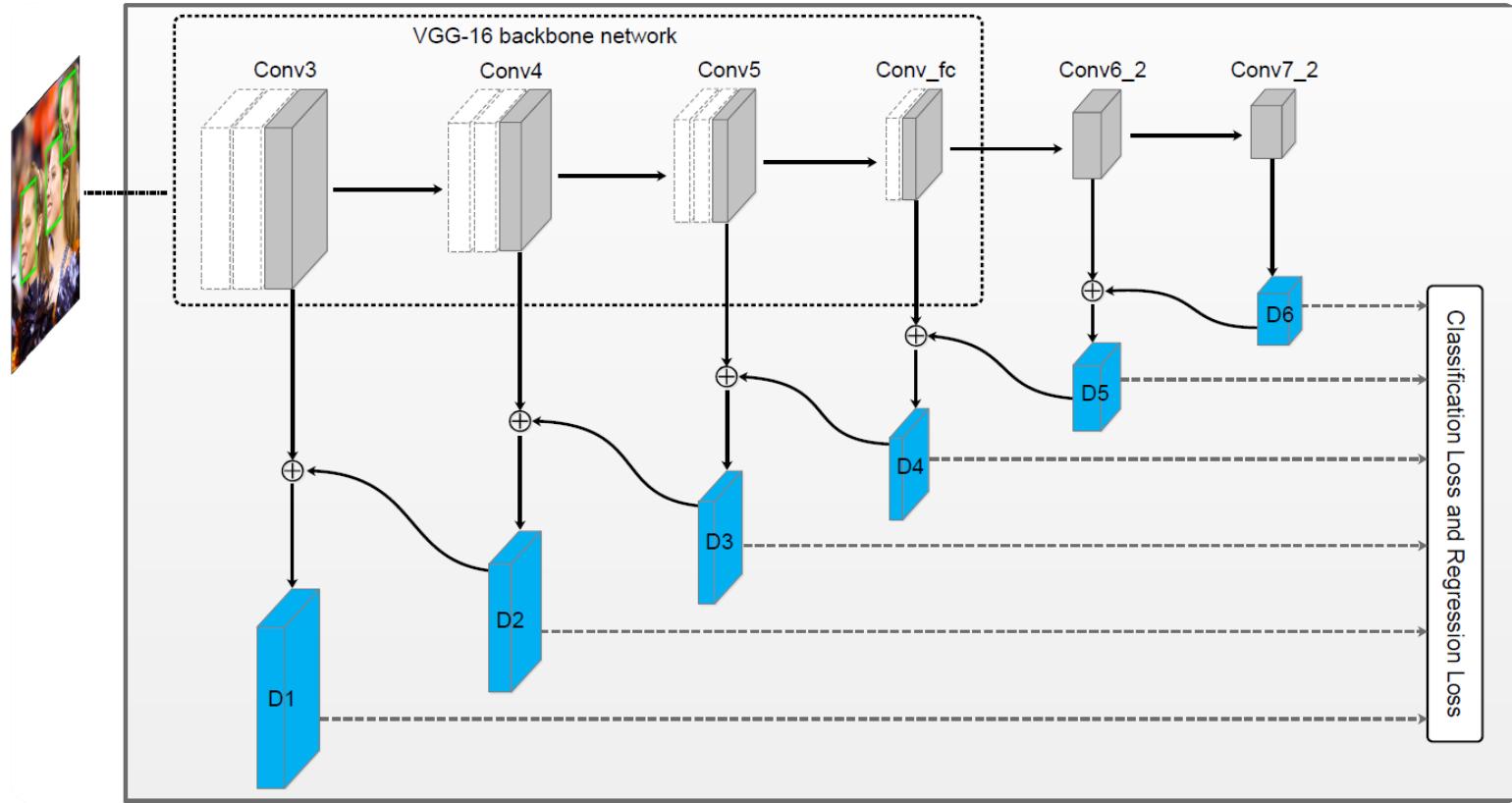
只在Conv3_3上使用
因为它关联的锚框最多，最容易产生虚检





高精度的人脸检测算法S³FD：期刊改进①

- 框架上：使用特征金字塔





高精度的人脸检测算法S³FD：期刊改进②

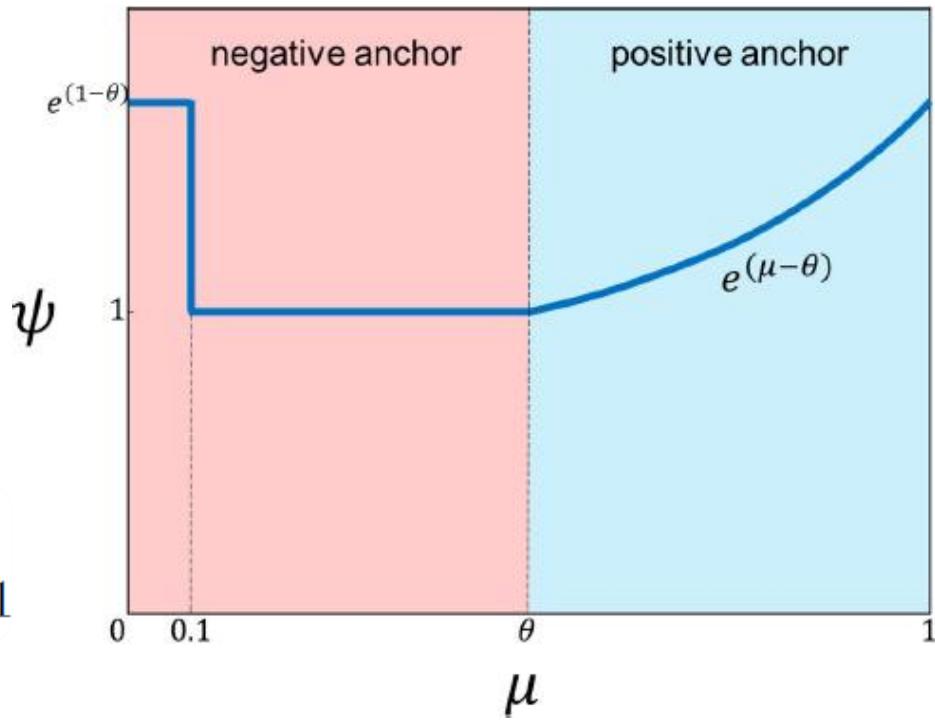
■ 训练样本加权

- 正样本: IoU越大, 权重越大

$$\psi(p_i^*) = \exp(\mu_i - \theta)$$

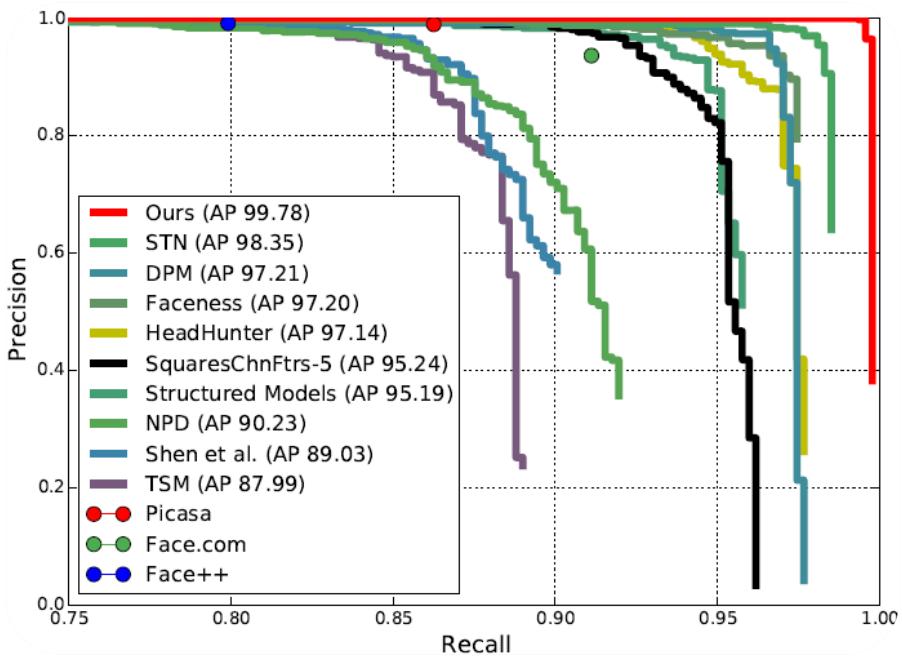
- 负样本: 加大属于背景区域的负样本的权重

$$\psi(p_i^*) = \begin{cases} 1, & \theta > \mu_i > 0.1 \\ \exp(1 - \theta), & \mu_i \leq 0.1 \end{cases}$$

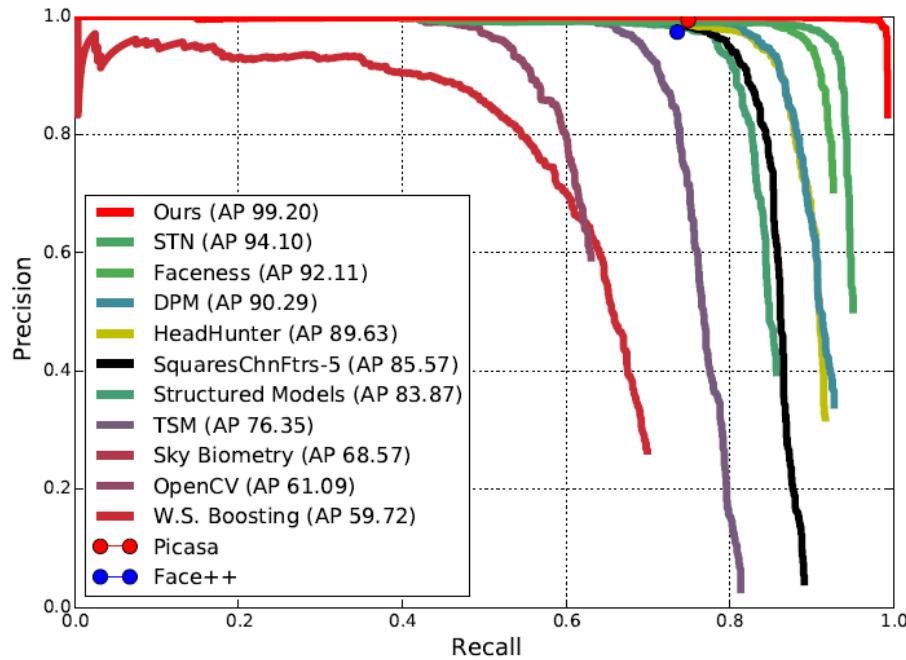




高精度的人脸检测算法S³FD：检测精度



AFW

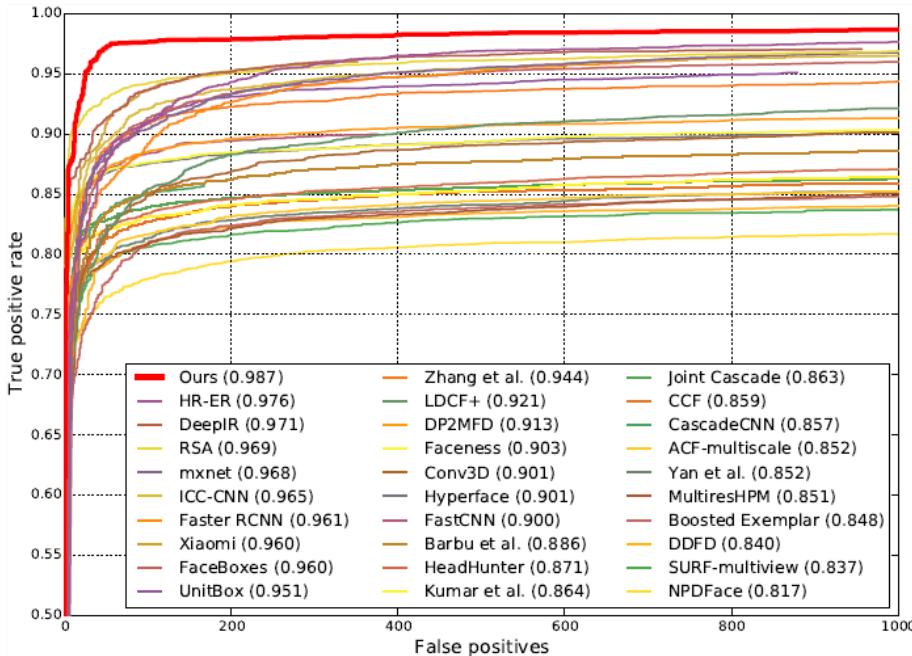


PASCAL Face

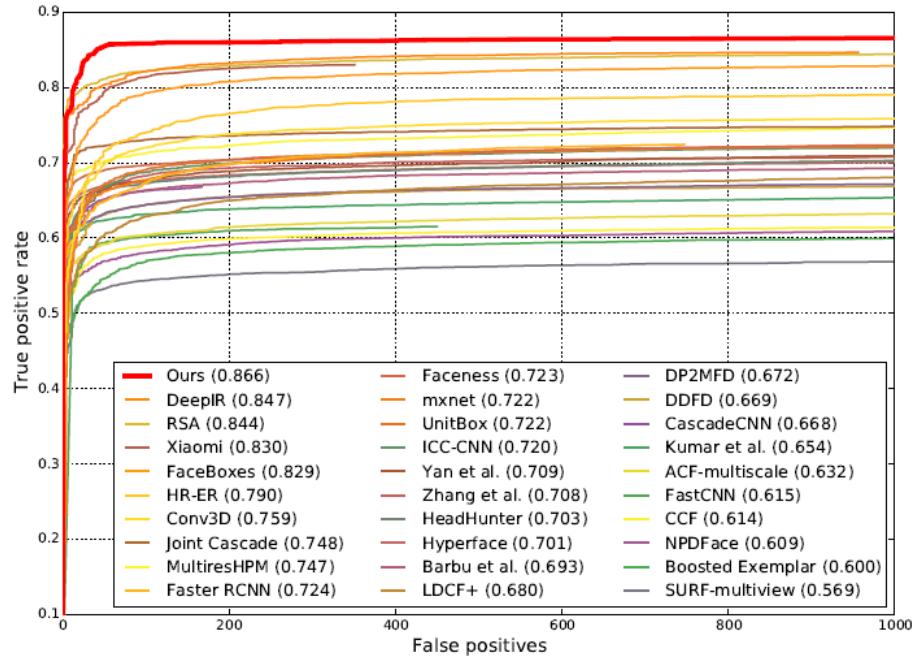




高精度的人脸检测算法S³FD：检测精度



FDDB离散



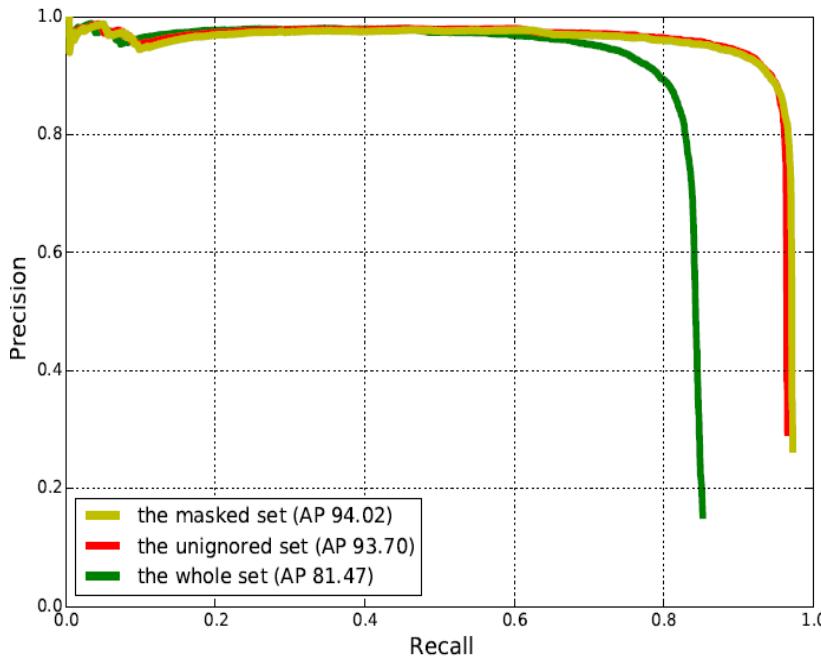
FDDB连续





高精度的人脸检测算法S³FD：检测精度

Methods	Whole set	Masked set	Unignored set
TSM [91]	-	-	41.6
HeadHunter [49]	-	-	50.9
HPM [16]	-	-	60.0
MTCNN [85]	-	-	60.8
LLE-CNNs [15]	-	-	76.4
FAN [73]	-	76.5	88.3
AOFD [7]	77.7	80.8	88.4
Ours	81.5	94.0	93.7

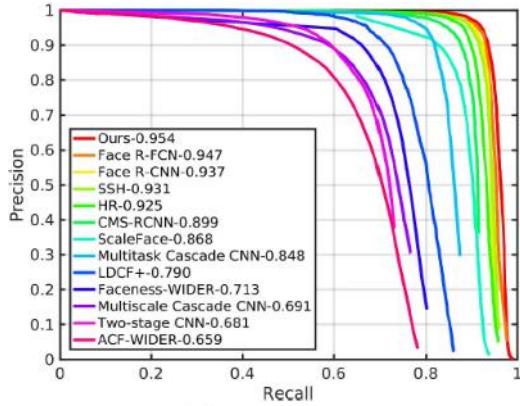


MAFA

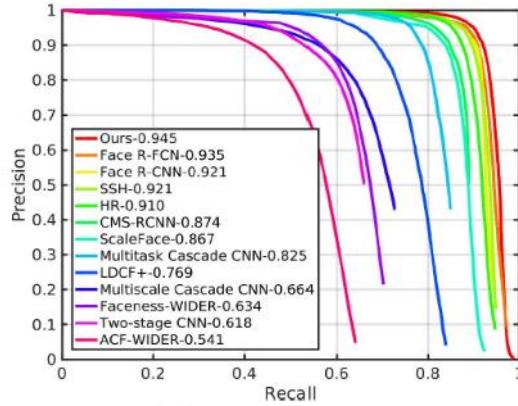




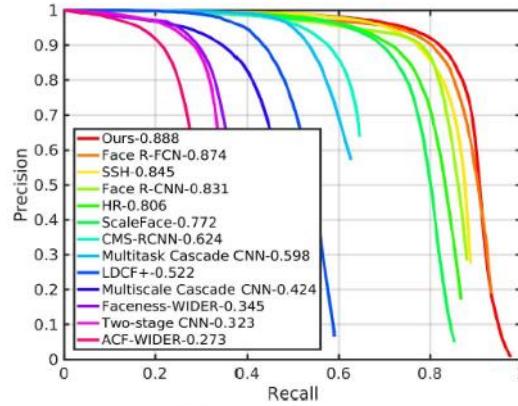
高精度的人脸检测算法S³FD：检测精度



(a) Val: Easy



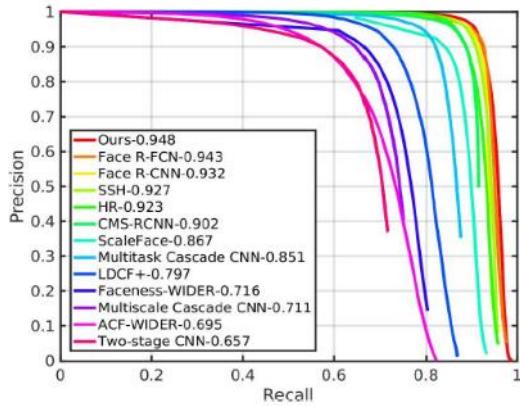
(b) Val: Medium



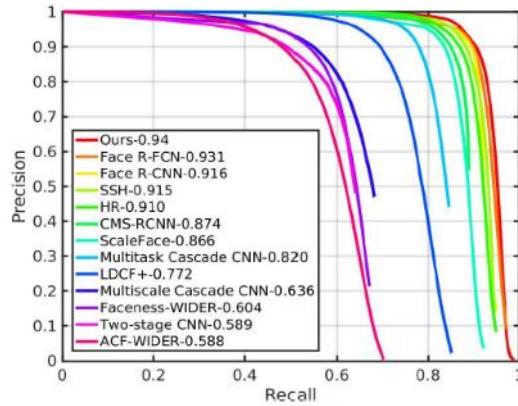
(c) Val: Hard

验证集

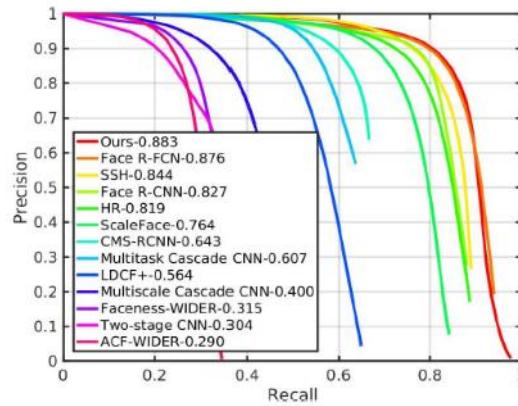
- 95.4 (Easy)
- 94.5 (Medium)
- 88.8 (Hard)



(d) Test: Easy



(e) Test: Medium



(f) Test: Hard

测试集

- 94.8 (Easy)
- 94.0 (Medium)
- 88.3 (Hard)





高精度的人脸检测算法S³FD：检测效果





高精度的人脸检测算法S³FD：检测效果

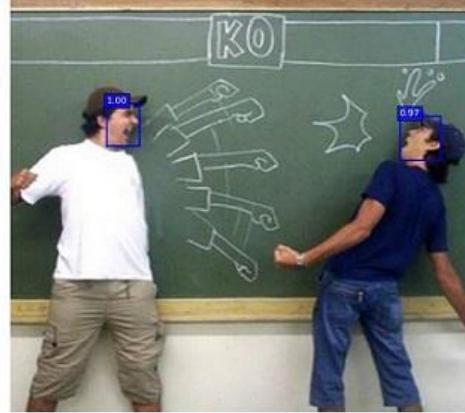
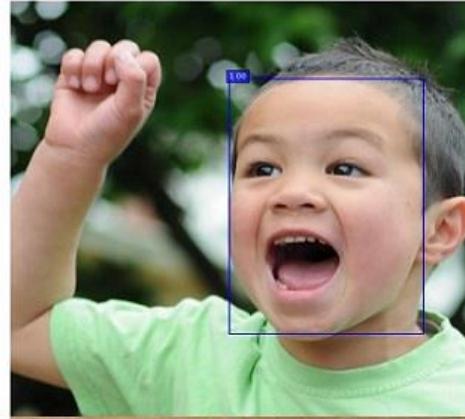
Pose



Occlusion



Expression





高精度的人脸检测算法S³FD：检测效果

Makeup



Illumination



Blur



XINHUANET



高精度的人脸检测算法S³FD：检测速度

- 在英伟达显卡Titan X (Pascal)上进行速度的测试
- N_a是非极大值抑制 (NMS) 之前保留的检测结果
- N_b是非极大值抑制 (NMS) 之后保留的检测结果

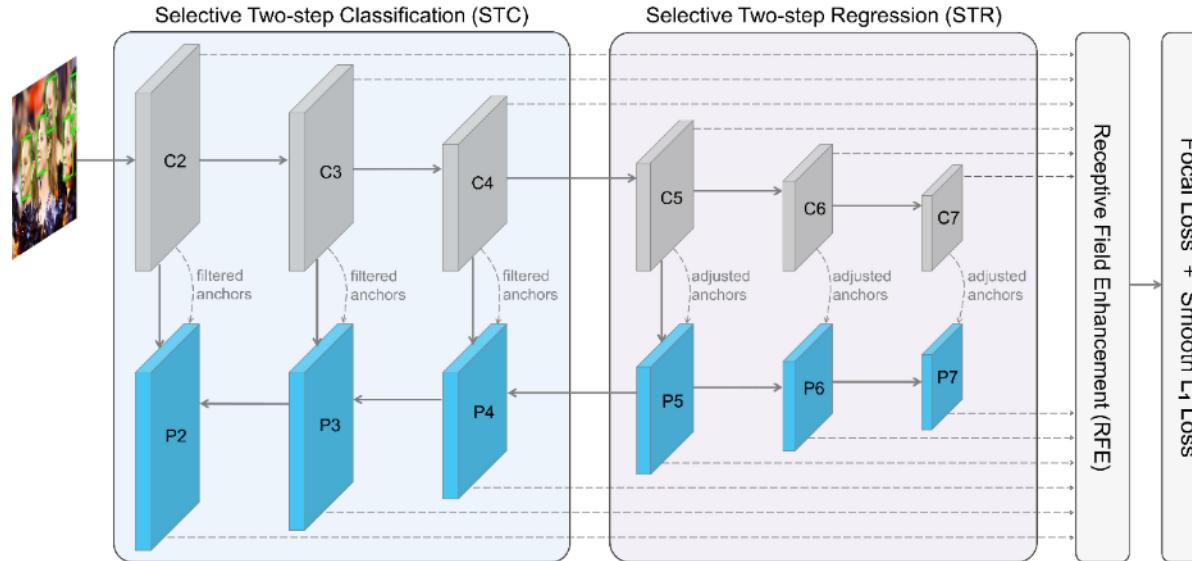
Resolution	Batch size	N_a	N_b	FPS
640×480	1	400	200	36
640×480	1	200	100	37
640×480	8	400	200	44
640×480	8	200	100	45





高精度的深度学习后期人脸检测算法：SRN

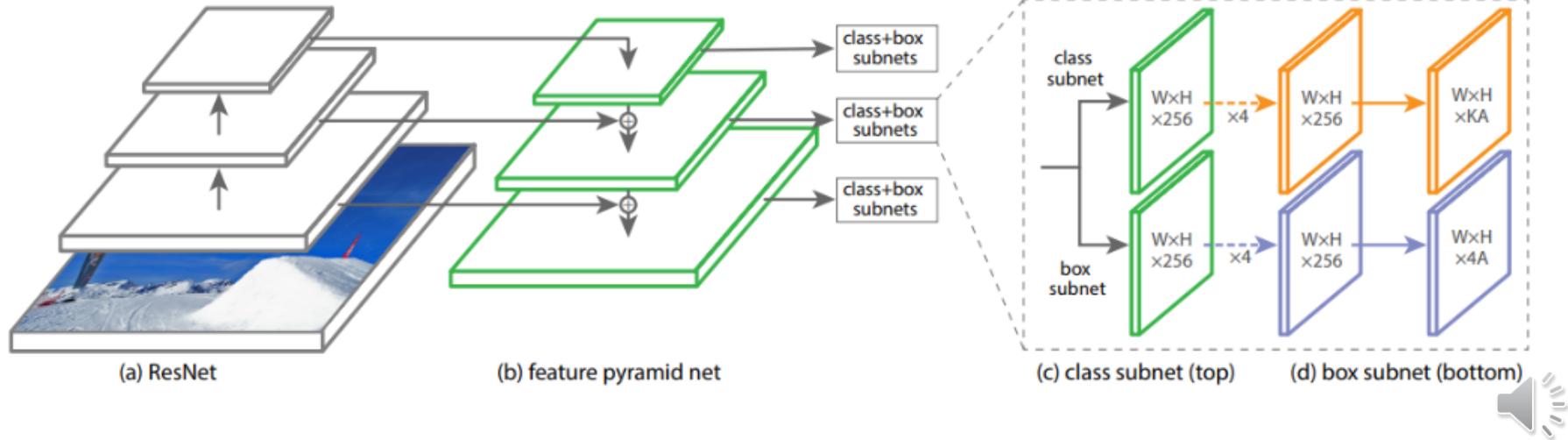
- 基于通用物体检测算法RefineDet和RetinaNet所设计的一个高精度人脸检测算法
- 主要改进点是：①选择性二阶段分类；②选择性二阶段回归；③感受野增强模块
- 发表时在WIDER FACE上取得最好结果，对于640x480分辨率的图像，能够在GPU上实时运行





高精度的人脸检测算法SRN：初始动机

- 通用物体检测算法RetinaNet开源代码，它在MS COCO数据集上性能比SSD好很多
- 尝试基于RetinaNet通用物体检测算法来做人脸检测这一特定任务
- 借鉴S³FD的锚框设计和数据增广，训练了一个RetinaNet人脸检测算法
- 基于RetinaNet训练的人脸检测算法存在两个：1) 召回效率不够高；2) 边框位置不够准

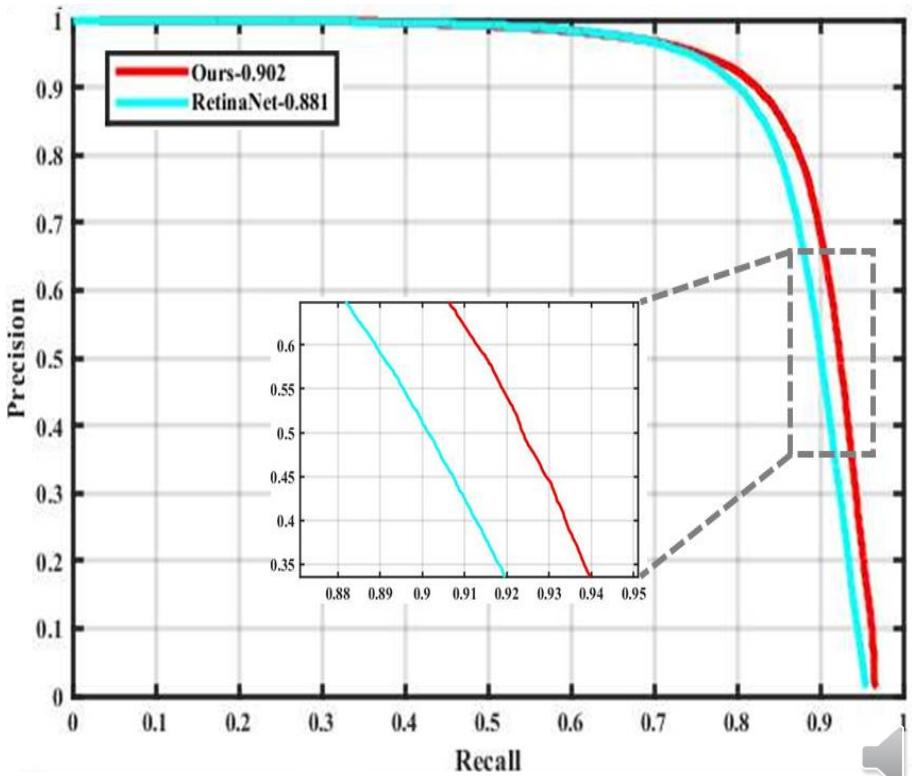




高精度的人脸检测算法SRN：初始动机

■ 召回效率不够高

- 基于RetinaNet的人脸检测器，在高召回率时，虚检的数量较多，需要减少
- 虽然平均精度 (AP) 很高，但是在高召回率 (Recall Rate) 时，精度 (Precision) 不够高
- 如右图蓝色曲线所示，RetinaNet人脸检测器在召回率为90%时，检测结果的精度只有~50%
- 换句话说，召回90%的人脸时，输出的检测结果有一半是虚检

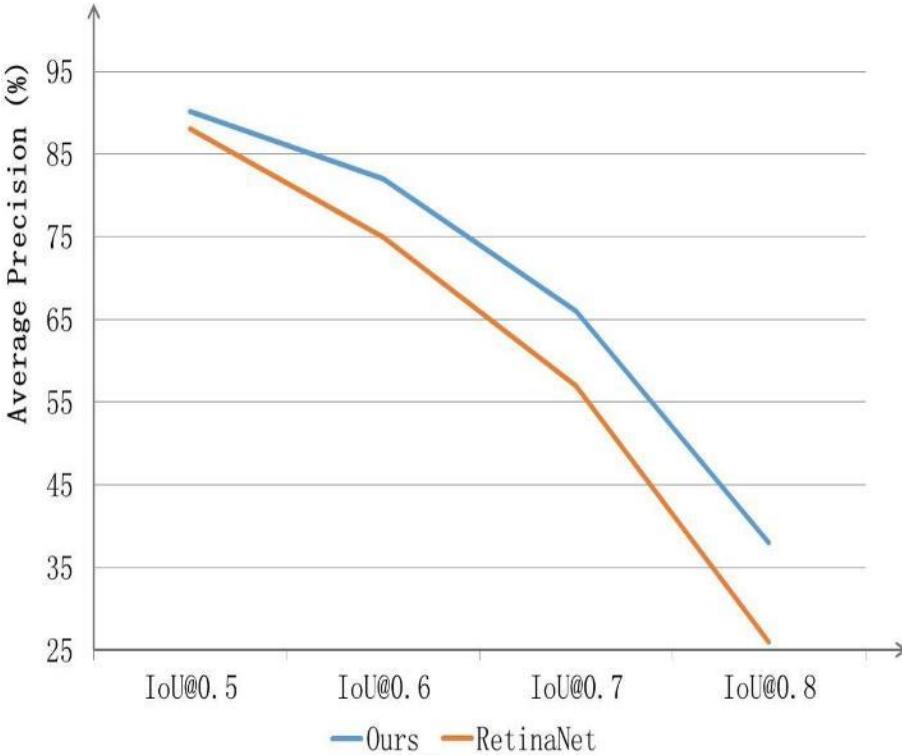




高精度的人脸检测算法SRN：初始动机

■ 边框位置不够准

- 基于RetinaNet的人脸检测算法，检测结果的边框位置还准备有待提高
- 参考MS COCO的评价指标，利用不同的IoU阈值得到对应的平均精度（AP）
- 当IoU的阈值越大时，对检测结果框的位置要求就会变得越高
- 如右图橙色曲线所示，随着IoU阈值的变高，平均精度（AP）急剧下降





高精度的人脸检测算法SRN：背后问题

召回效率不够高

小尺度人脸会加重正负样本不平衡的问题

RetinaNet的单阶段分类不能很好的解决

边框位置不够准

RetinaNet是单阶段法，只具备单阶段回归

单阶段回归不能得到非常精准的边框位置





高精度的人脸检测算法SRN：解决方案

召回效率不够高

小尺度人脸会加重正负样本不平衡的问题

RetinaNet的单阶段分类不能很好的解决



借鉴RefineDet
引入选择性二阶段分类

边框位置不够准

RetinaNet是单阶段法，只具备单阶段回归

单阶段回归不能得到非常精准的边框位置

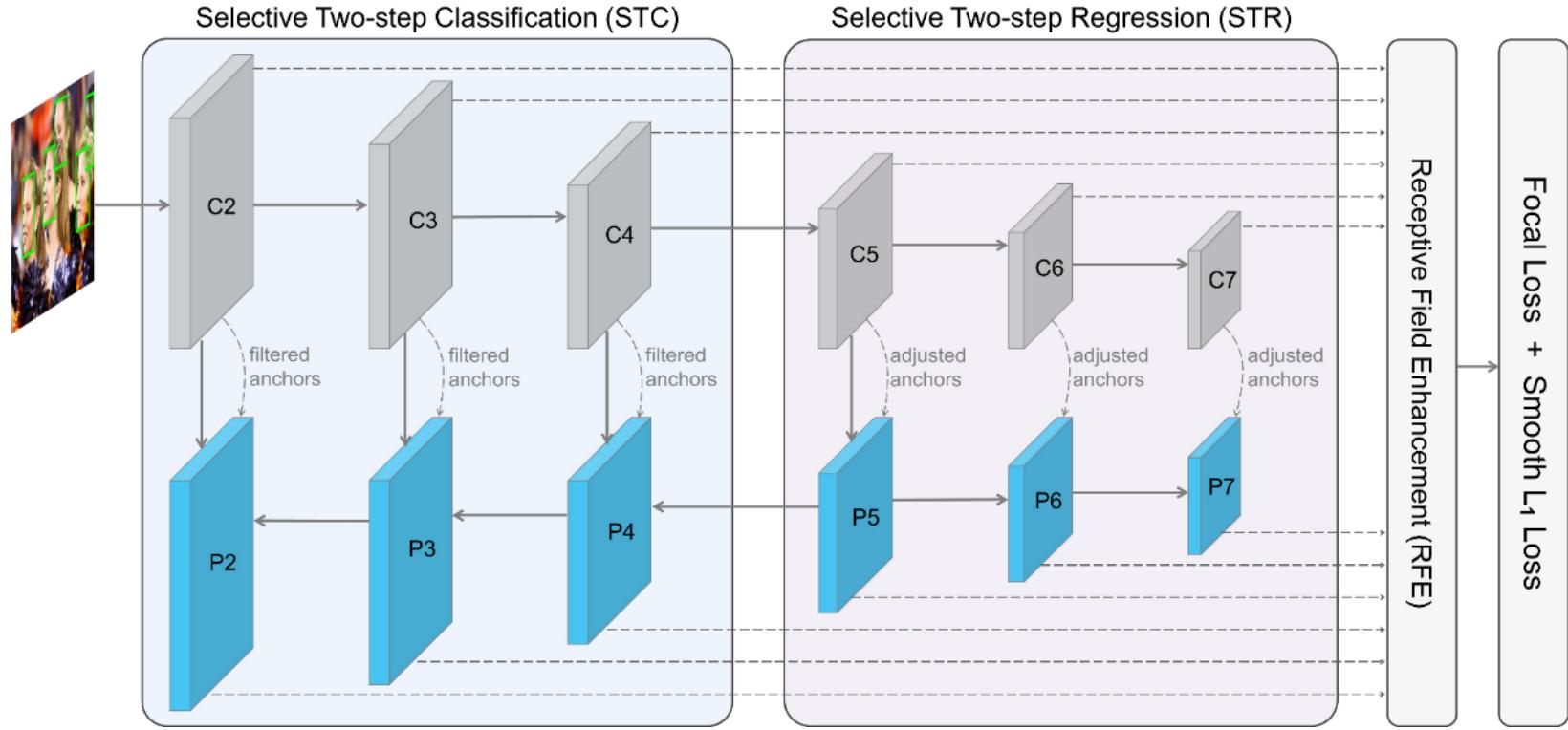


借鉴RefineDet
引入选择性二阶段回归





高精度的人脸检测算法SRN：整体框架

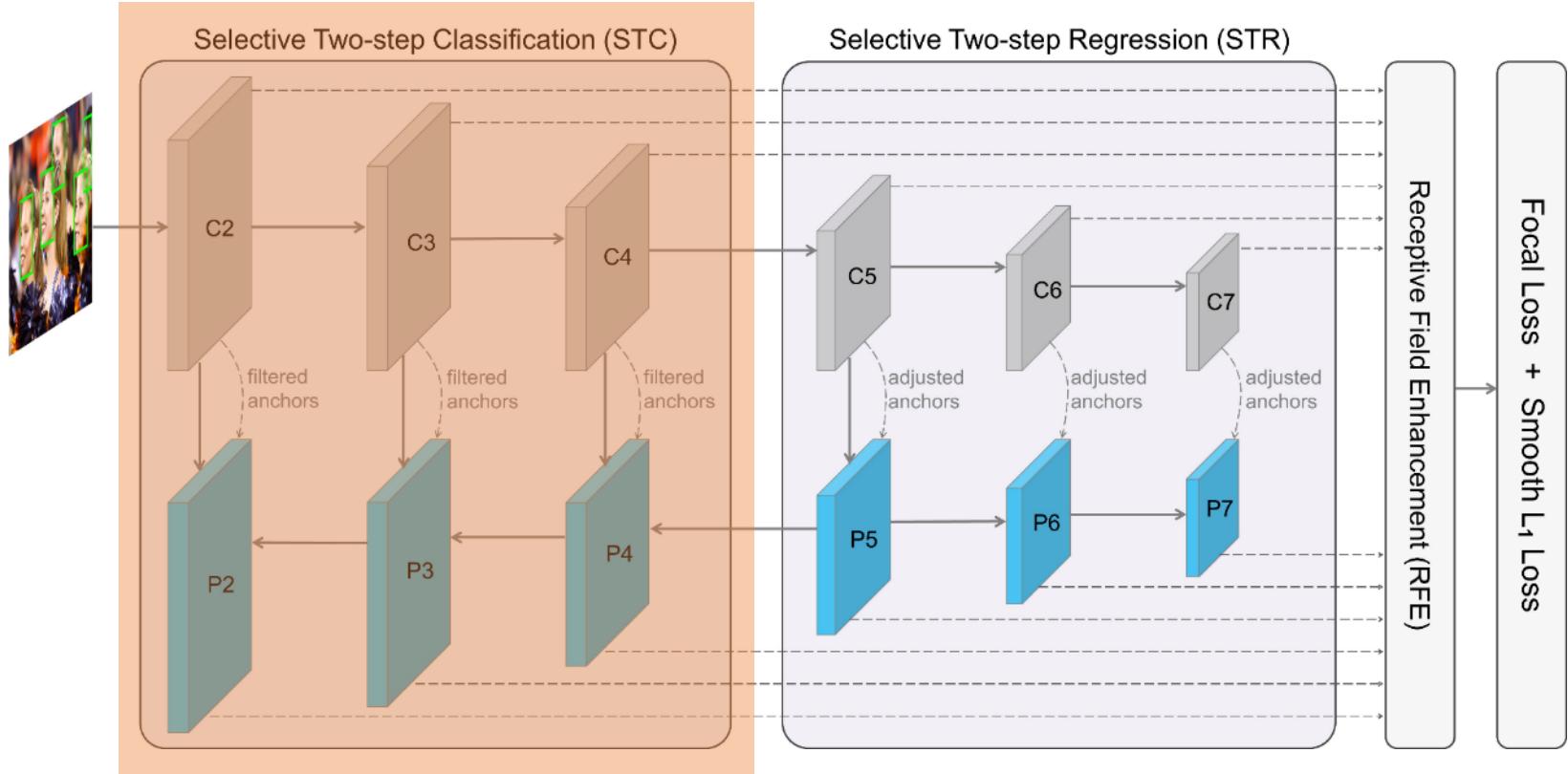


■ SRN人脸检测算法：选择性二阶段分类 (STC)、选择性二阶段回归 (STR)、感受野增强模块 (RFE)





高精度的人脸检测算法SRN：整体框架



■ SRN人脸检测算法：选择性二阶段分类 (STC)、选择性二阶段回归 (STR)、感受野增强模块 (RFE)





高精度的人脸检测算法SRN：选择性二阶段分类

- 为了检测小尺度人脸，需要密密麻麻地铺设大量的小尺度锚框
- 只有少量的小尺度锚框能够匹配到小尺度人脸成为正样本，剩下的几乎全是负样本
- 所以检测小尺度人脸，会导致正负样本的比例变得更加的失衡，容易产生大量虚检
- 应用RefineDet的二阶段分类来缓解这一问题，但是效果提升不明显
- 为了分析原因，在每个检测层上单独使用二阶段分类，效果如下

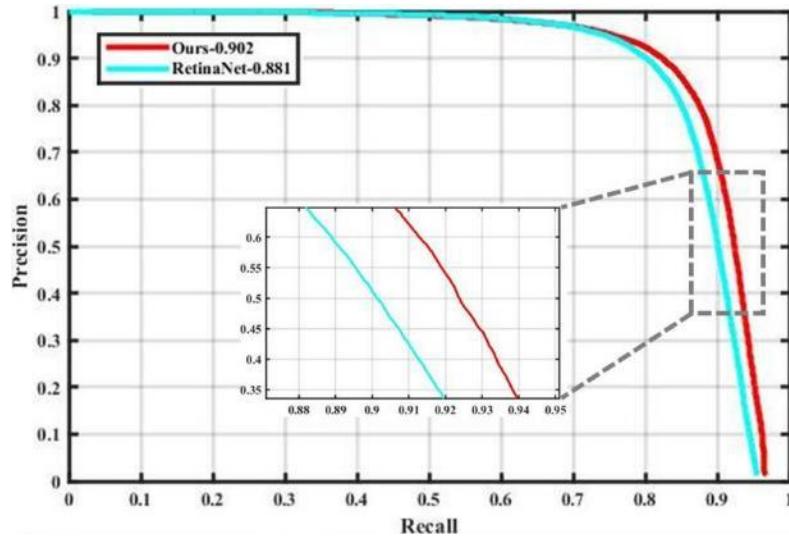
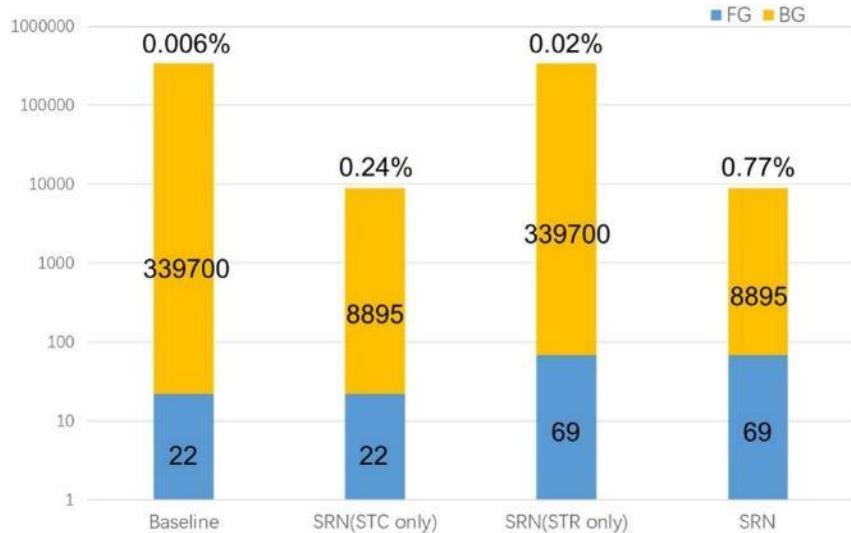
STC	B	P2	P3	P4	P5	P6	P7
<i>Easy</i>	95.1	95.2	95.2	95.2	95.0	95.1	95.0
<i>Medium</i>	93.9	94.2	94.3	94.1	93.9	93.7	93.9
<i>Hard</i>	88.0	88.9	88.7	88.5	87.8	88.0	87.7

- 并不是在所有检测层上使用二阶段分类都能得到提升，故选择性的仅在最有效的P2, P3, P4上执行





高精度的人脸检测算法SRN：选择性二阶段分类

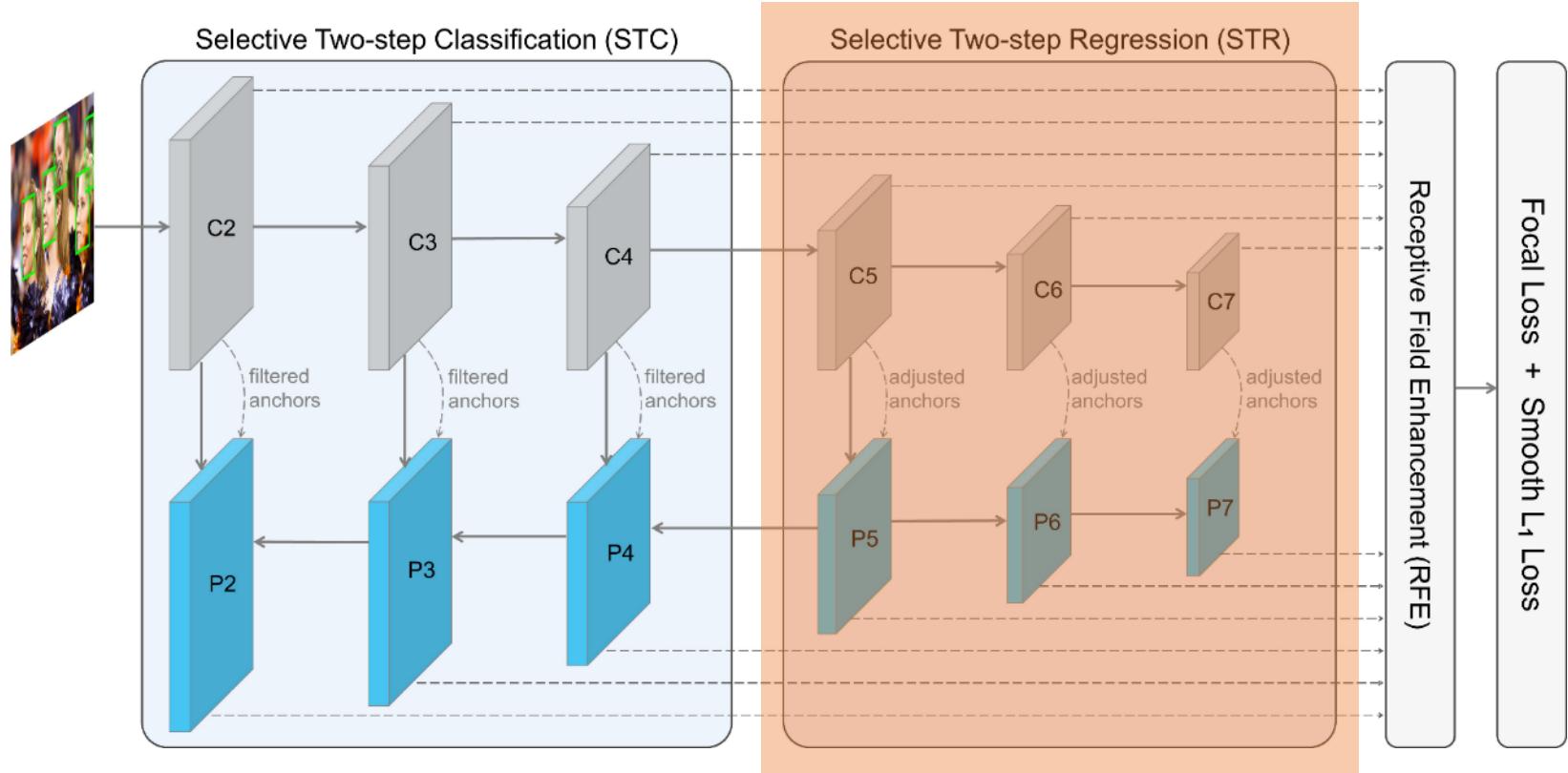


- 如左上图所示，选择性二阶段分类可以把正负样本比例提升约114倍
- 如右上图所示，在召回率为91%，选择性二阶段分类能够把精度提升约20%





高精度的人脸检测算法SRN：整体框架



■ SRN人脸检测算法：选择性二阶段分类 (STC)、选择性二阶段回归 (STR)、感受野增强模块 (RFE)





高精度的人脸检测算法SRN：选择性二阶段回归

- 使检测结果框的位置变得更加精确是一个具有挑战性的问题
- 单阶段法RetinaNet依赖于单阶段回归来得到检测结果
- 导致检测结果框不够精准，使得当评测的IoU阈值变大时，精度急剧下降
- 应用RefineDet的二阶段回归来缓解这一问题，但是效果没有提升反而下降
- 为了分析原因，在每个检测层上单独使用二阶段回归，效果如下

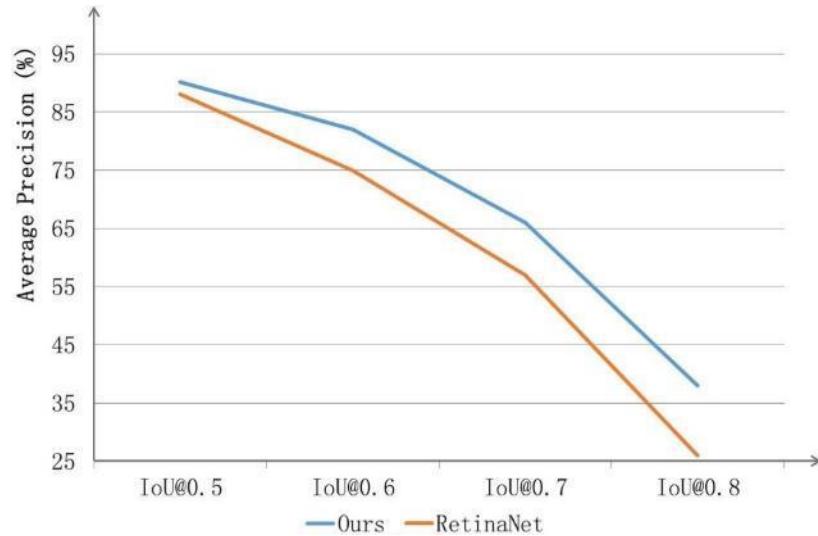
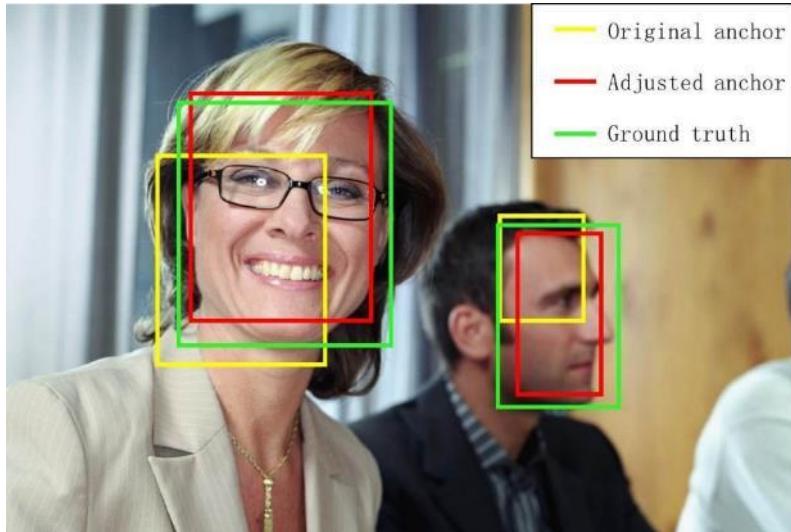
STR	B	P2	P3	P4	P5	P6	P7
<i>Easy</i>	95.1	94.8	94.3	94.8	95.4	95.7	95.6
<i>Medium</i>	93.9	93.4	93.7	93.9	94.2	94.4	94.6
<i>Hard</i>	88.0	87.5	87.7	87.0	88.2	88.2	88.4

- 并不是在所有检测层上使用二阶段回归都能得到提升，故选择性的仅在有提升的P5, P6, P7上执行





高精度的人脸检测算法SRN：选择性二阶段回归

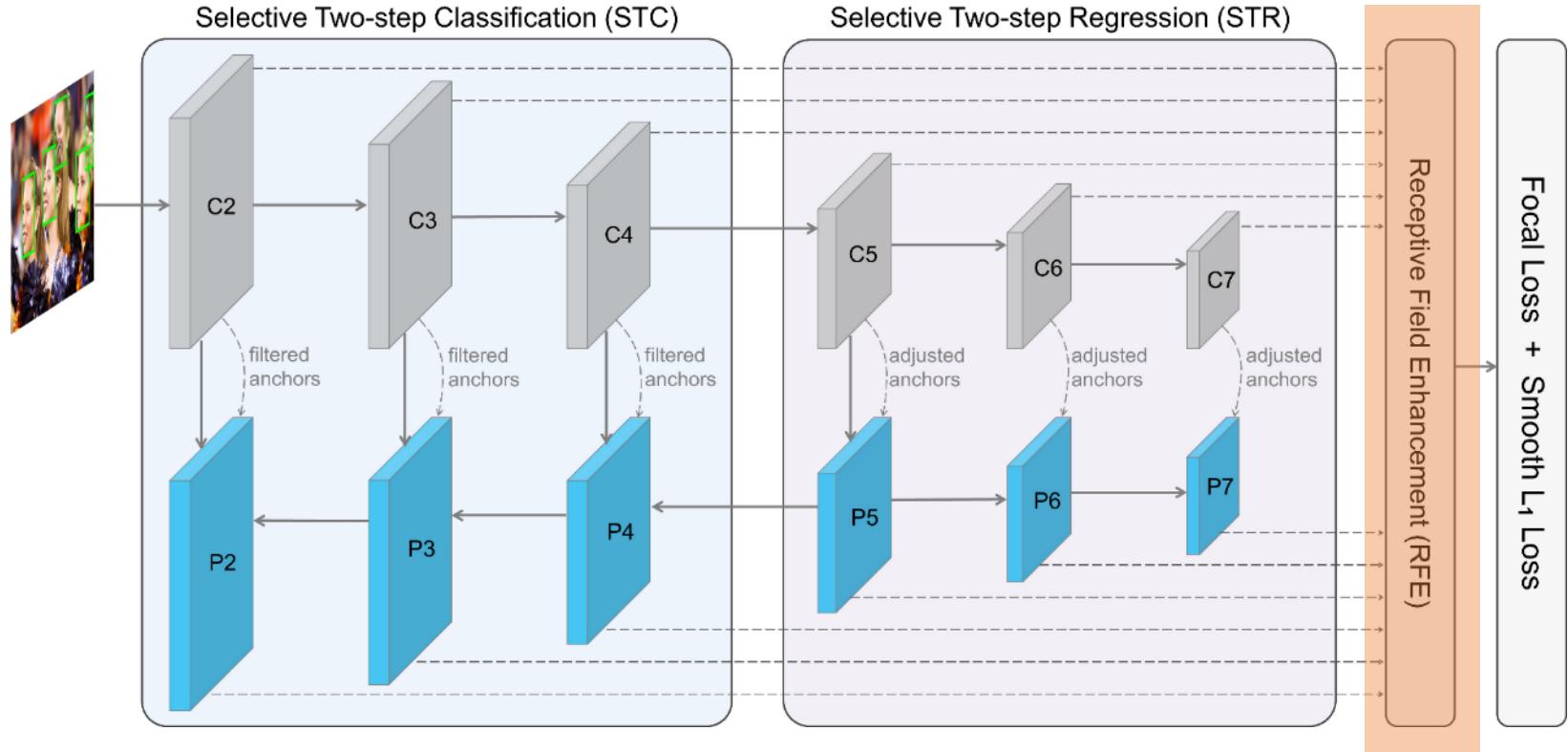


- 如左上图所示，选择性二阶段回归为后续的回归器提供更好的初始位置，从而得到更精准的边框位置
- 如右上图所示，随着IoU阈值的变大，两个曲线的差距越大，即选择性二阶段回归带来的提升越明显





高精度的人脸检测算法SRN：整体框架



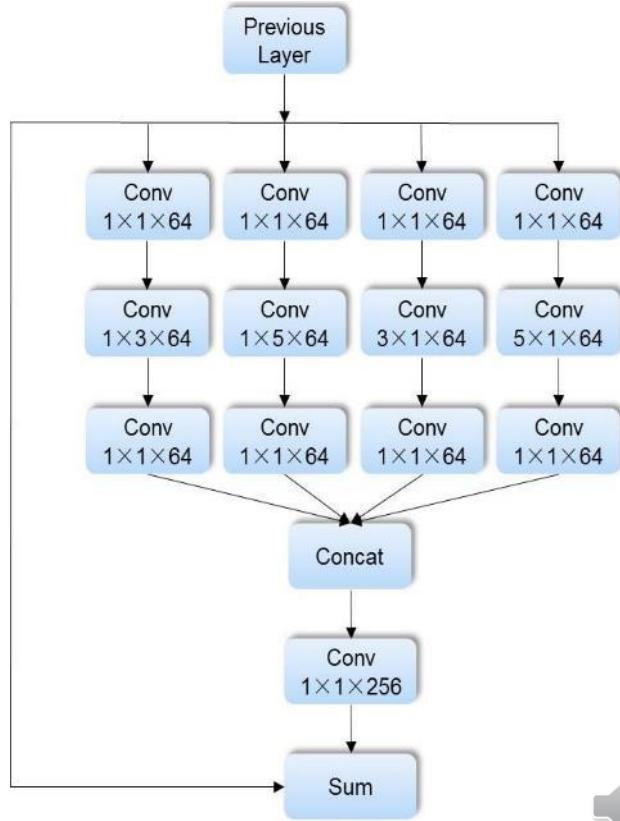
■ SRN人脸检测算法：选择性二阶段分类 (STC)、选择性二阶段回归 (STR)、感受野增强模块 (RFE)





高精度的人脸检测算法SRN：感受野增强模块

- 当前深度卷积网络的感受野都是正方形的
- 但是需要检测的人脸并不全是正方形的
- 即网络感受野和待检测人脸之间存在着不适应性
- 该不适应性在一定程度上会影响检测的性能
- 感受野增强模块借鉴Inception
- 使用不同的分支提供不同形状的感受野，来提高检测精度





高精度的人脸检测算法SRN：训练细节

- 基础网络: ResNet-50 + 6级FPN
- 损失函数: sigmoid focal loss + smooth L1 loss
- 数据增广: 与S³FD保持一致, 即颜色抖动, 随机裁剪, 随机水平翻转
- 锚框设计: 检测层上的每个特征关联2个尺度(2, 2√2)锚框, 比例都是1.25:1
- 优化算法SGD + 0.9 momentum + 0.0001 weight decay + batch size 32
- 初始学习率为0.01, 在100和120次epoch时分别衰减10倍, 总共训130个epoch





高精度的人脸检测算法SRN：验证改进

- 选择性二阶段分类 (STC) 带来0.2% (Easy)、 0.5% (Medium)、 1.4% (Hard)
- 选择性二阶段回归 (STR) 带来0.8% (Easy)、 0.9% (Medium)、 0.8% (Hard)
- 感受野增强模块 (RFE) 带来0.3% (Easy)、 0.3% (Medium)、 0.1% (Hard)

Component	SRN				
	STC	STR	RFE		
STC	✓			✓	
STR		✓	✓	✓	
RFE				✓	
<i>Easy subset</i>	95.1	95.3	95.9	96.1	96.4
<i>Medium subset</i>	93.9	94.4	94.8	95.0	95.3
<i>Hard subset</i>	88.0	89.4	88.8	90.1	90.2





高精度的人脸检测算法SRN：检测精度

Dataset	Criterion	Value			
AFW	Average Precision (AP)	99.87			
PASCAL Face	Average Precision (AP)	99.09			
FDDB	True Positive Rate @ False Positive=1000	98.8			
WIDER FACE	Average Precision (AP)	Validation	96.4 (Easy)	95.3 (Medium)	90.2 (Hard)
		Test	95.9 (Easy)	94.9 (Medium)	89.7 (Hard)





高精度的人脸检测算法SRN：检测效果

Blur



Expression



Illumination



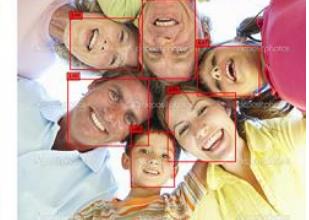
Makeup



Occlusion



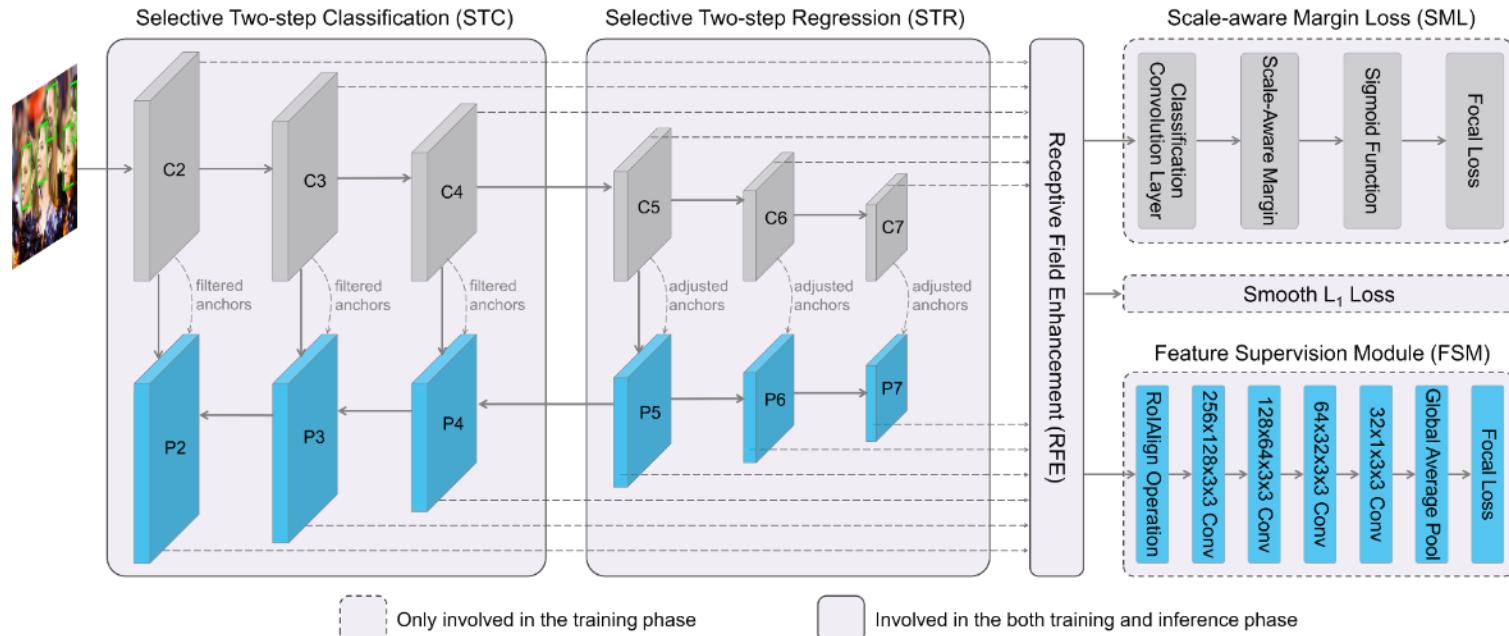
Pose





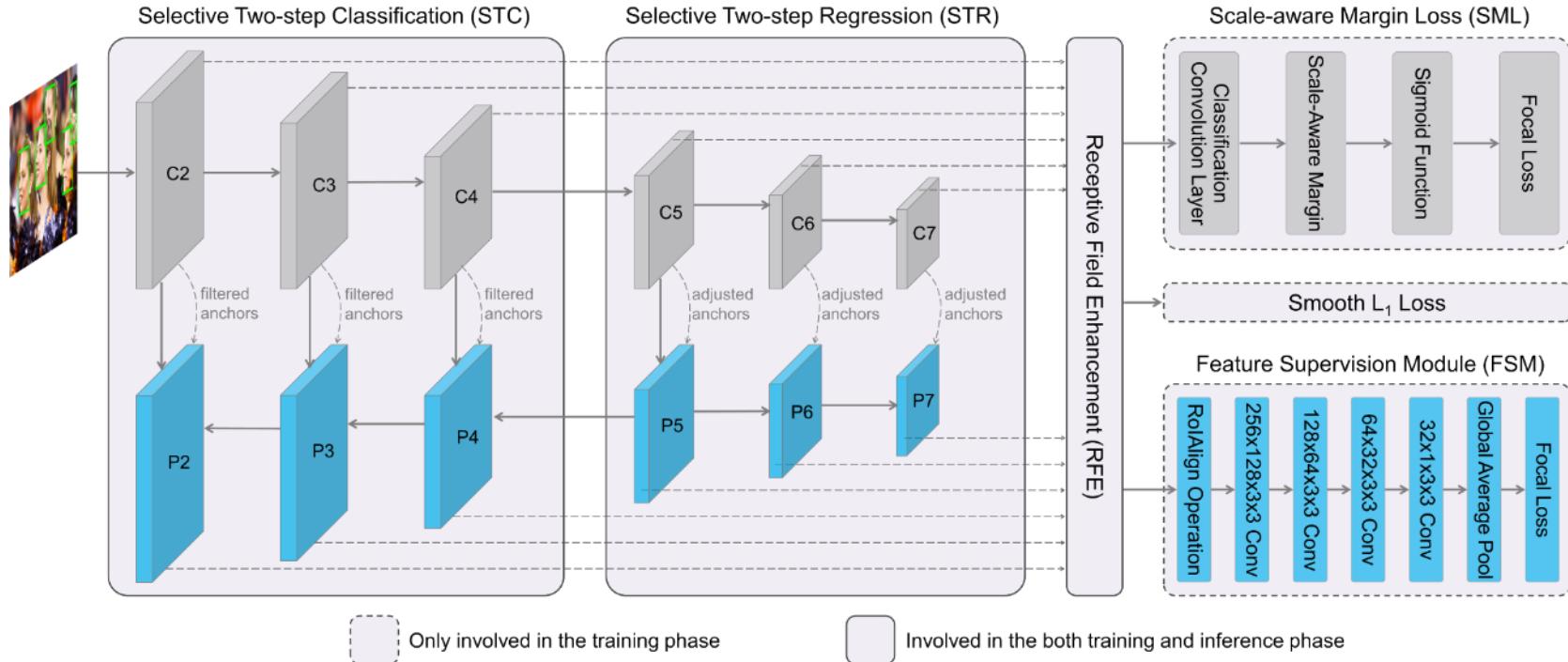
高精度的深度学习后期人脸检测算法：RefineFace

- SRN是基于通用物体检测算法RefineDet和RetinaNet所设计的一个高精度人脸检测算法
- RefineFace是对SRN的改进，在不增加耗时的前提下，提升了检测精度
- 主要改进点是：①尺度敏感margin的损失函数；②特征监督模块





高精度的人脸检测算法RefineFace：整体框架

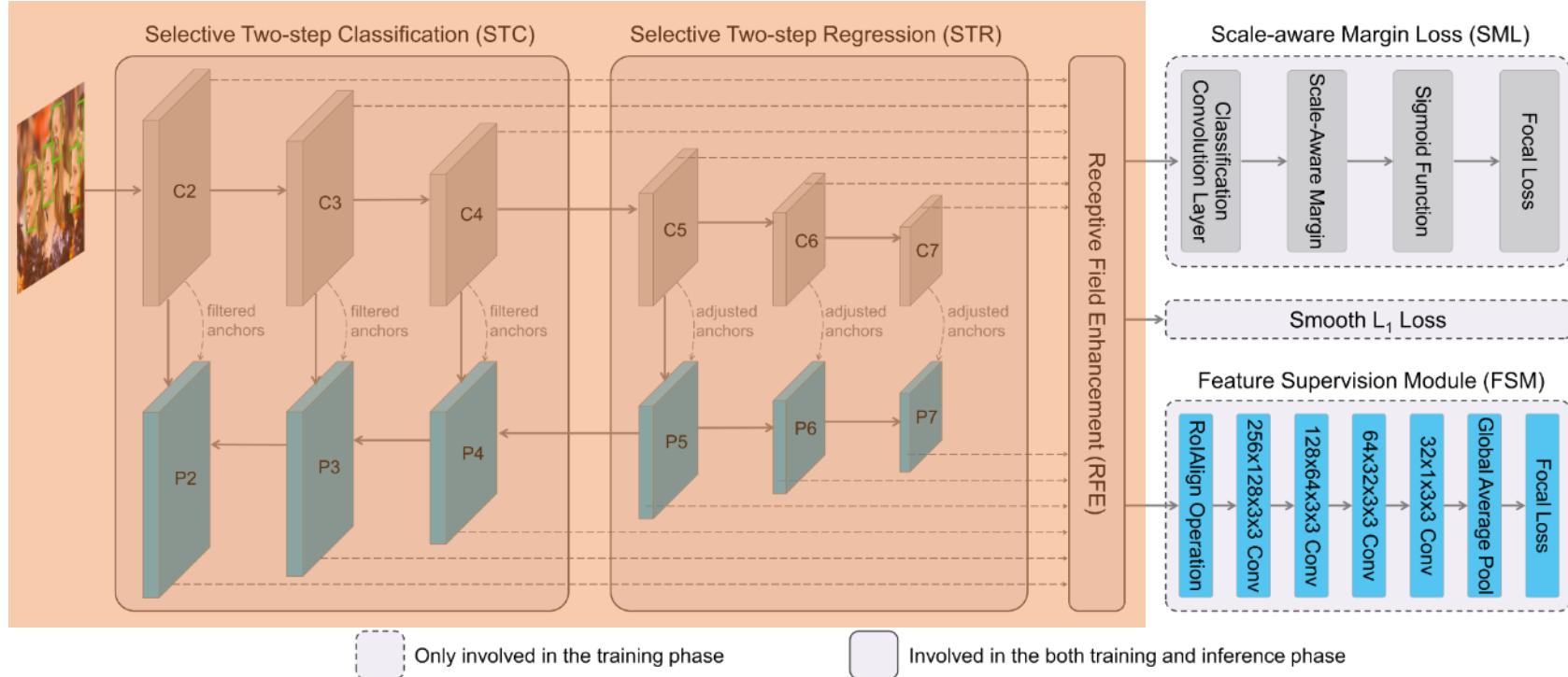


■ RefineFace人脸检测算法 = SRN人脸检测算法 + 尺度敏感margin的损失函数 + 特征监督模块





高精度的人脸检测算法RefineFace：整体框架

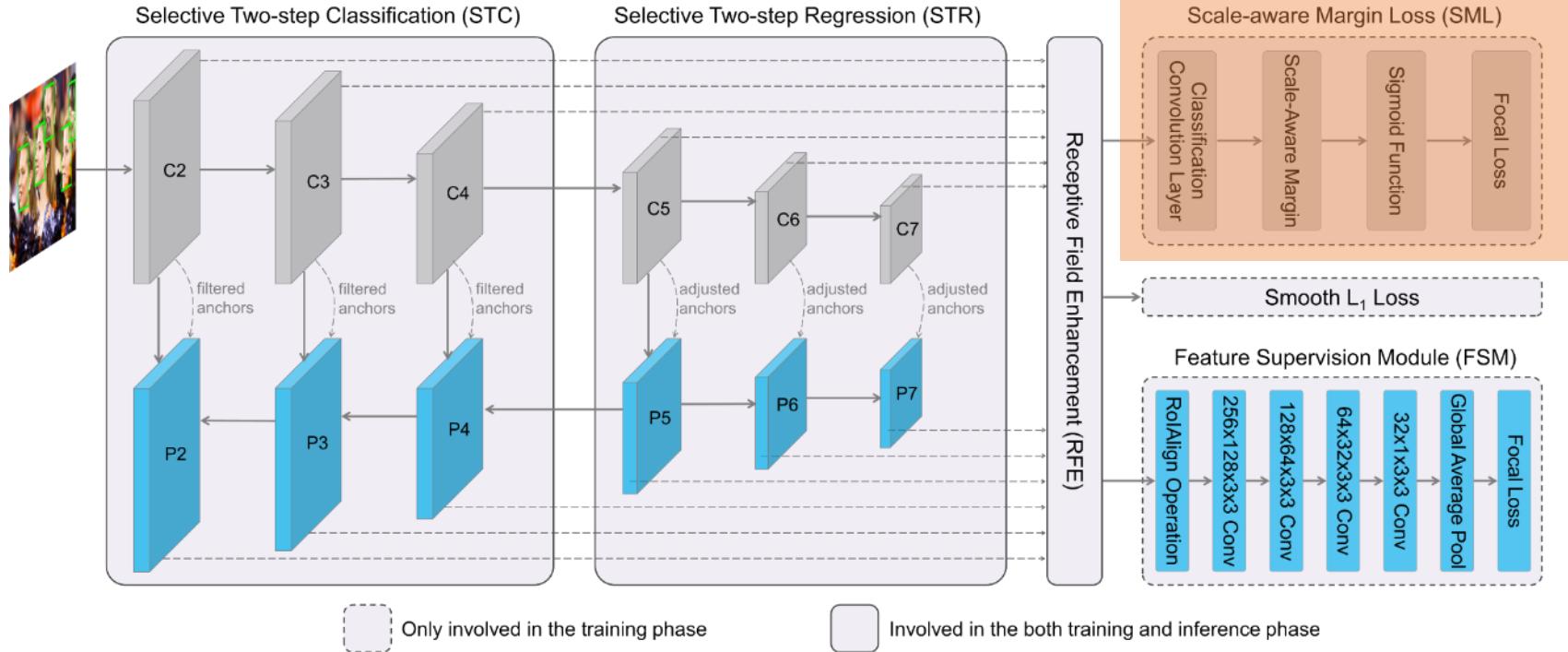


■ RefineFace人脸检测算法 = **SRN人脸检测算法** + 尺度敏感margin的损失函数 + 特征监督模块





高精度的人脸检测算法RefineFace：整体框架



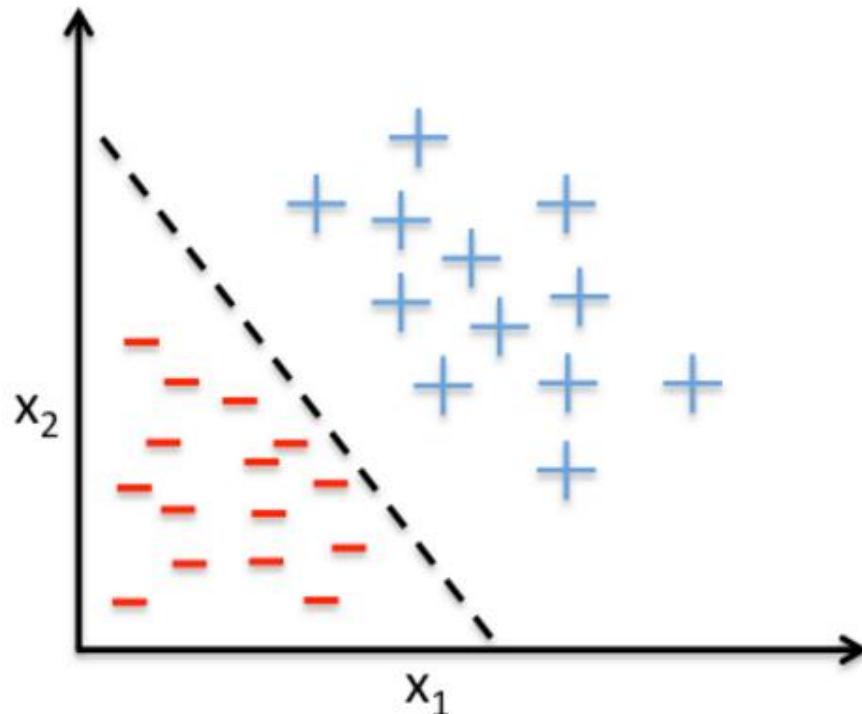
■ RefineFace人脸检测算法 = SRN人脸检测算法 + **尺度敏感margin的损失函数** + 特征监督模块





高精度的人脸检测算法RefineFace：尺度敏感margin

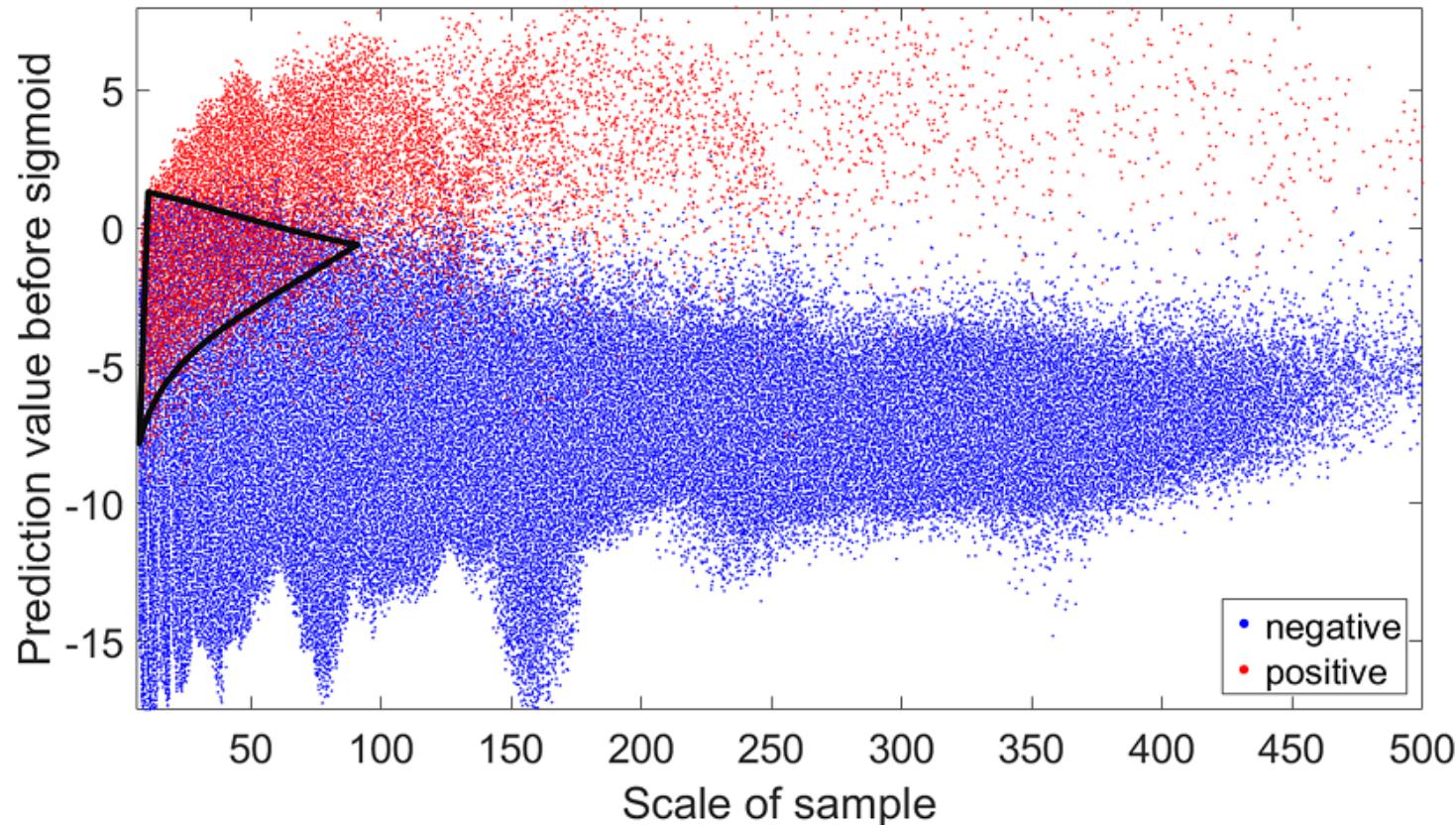
- 人脸检测 = 人脸分类 + 人脸回归
- 人脸分类是一个二分类任务：人脸 vs 非人脸（背景）





高精度的人脸检测算法RefineFace：尺度敏感margin

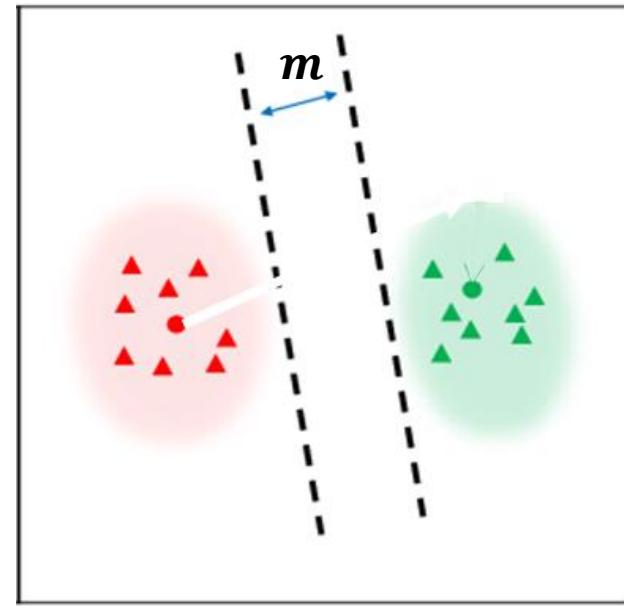
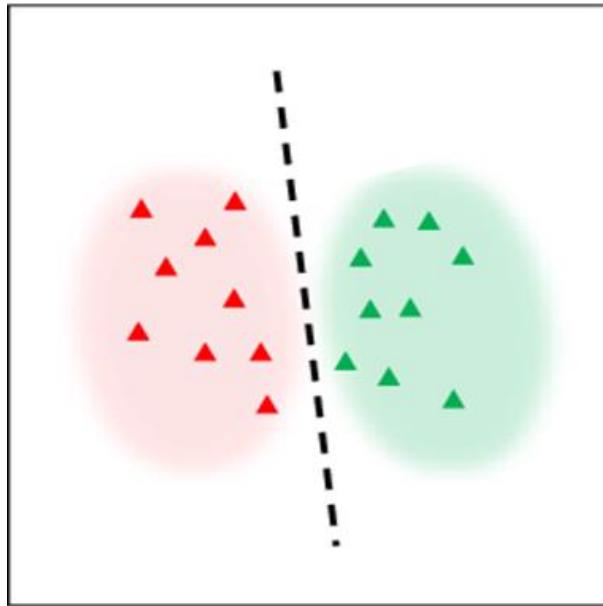
- 人脸检测中人脸二分类的情况，尺度越大的人脸与背景分的越开，尺度较小的人脸无法与背景分开





高精度的人脸检测算法RefineFace：尺度敏感margin

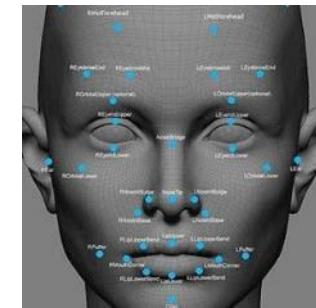
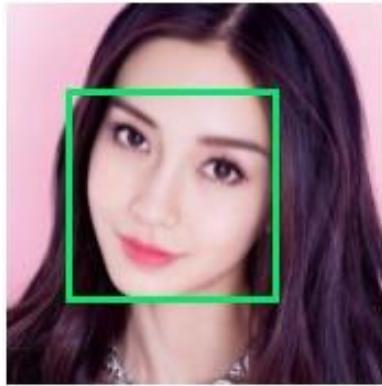
- 人脸识别中，为了让两个类别之间分的更开，常用的一种方式是在分类损失函数中加一个不同形式的margin





高精度的人脸检测算法RefineFace：尺度敏感margin

- 但是在人脸识别中，在分类之前会把所有样本都会归一化到相同大小，即不存在尺度的问题



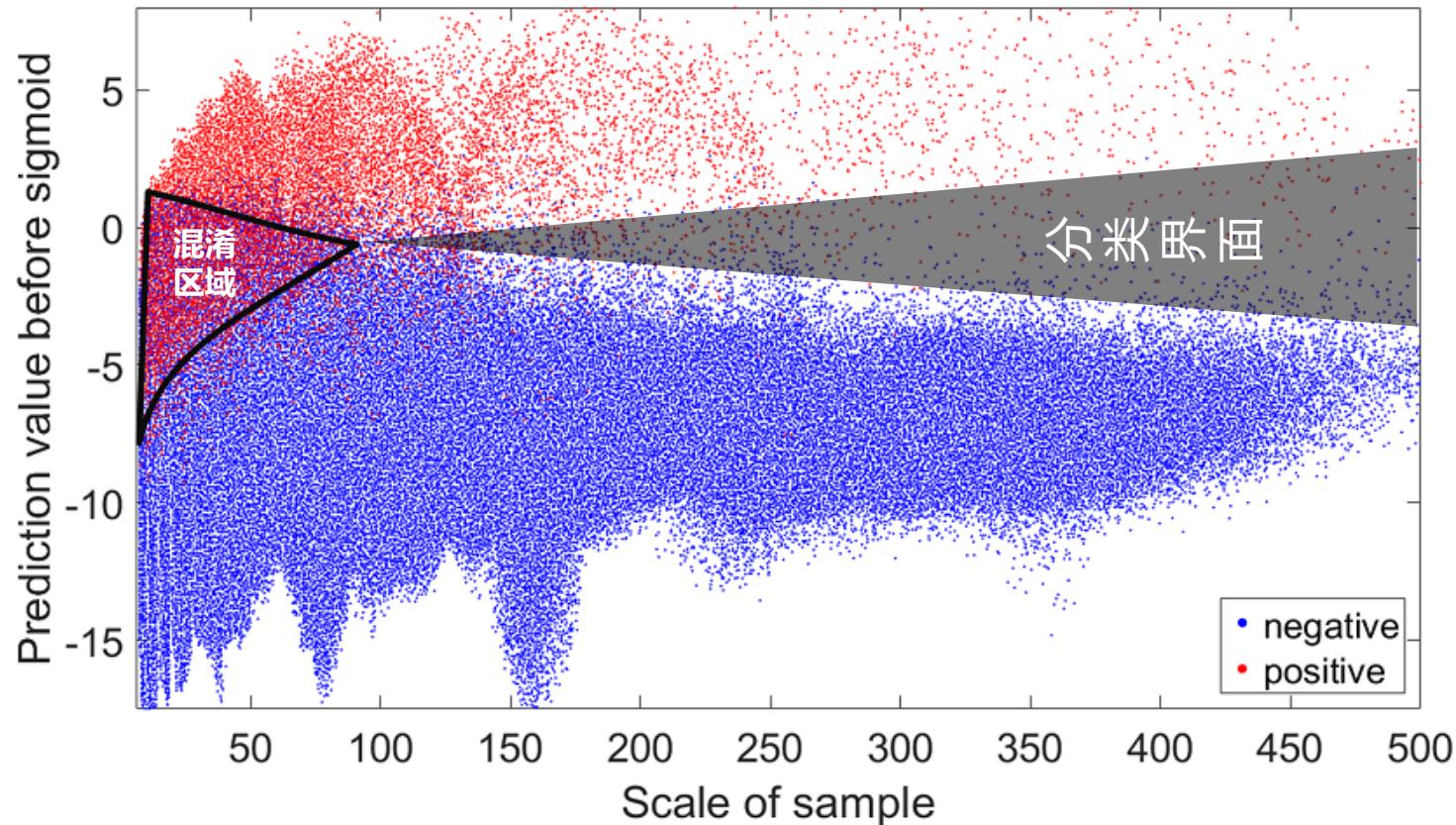
固定大小的正面人脸





高精度的人脸检测算法RefineFace：尺度敏感margin

- 而人脸检测中人脸二分类的情况存在着**尺度问题**，尺度越小分类界面越模糊

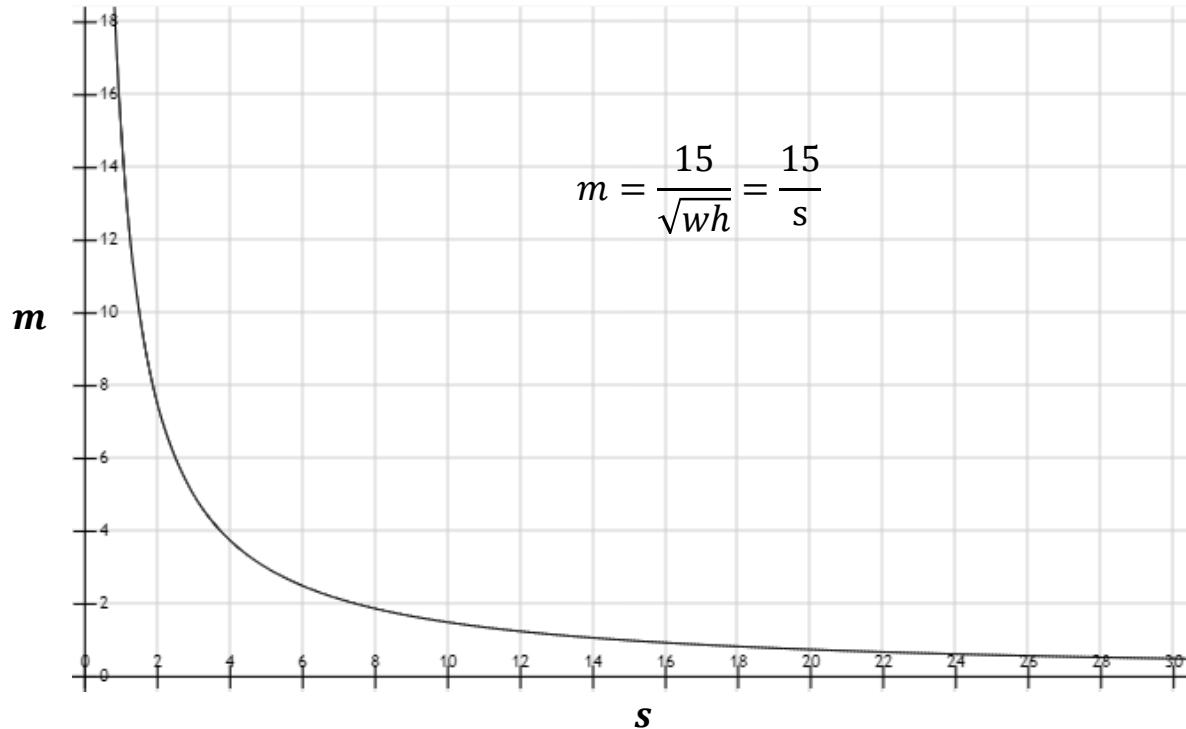




高精度的人脸检测算法RefineFace：尺度敏感margin

- 提出尺度敏感的margin策略，尺度越小的人脸，对应的margin越大

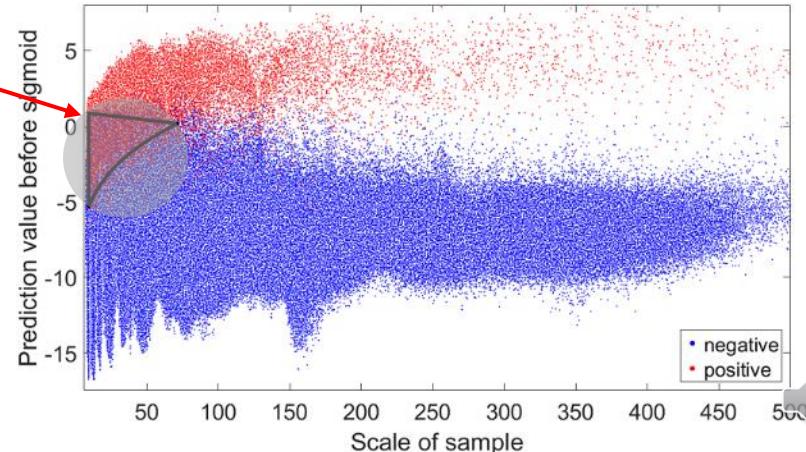
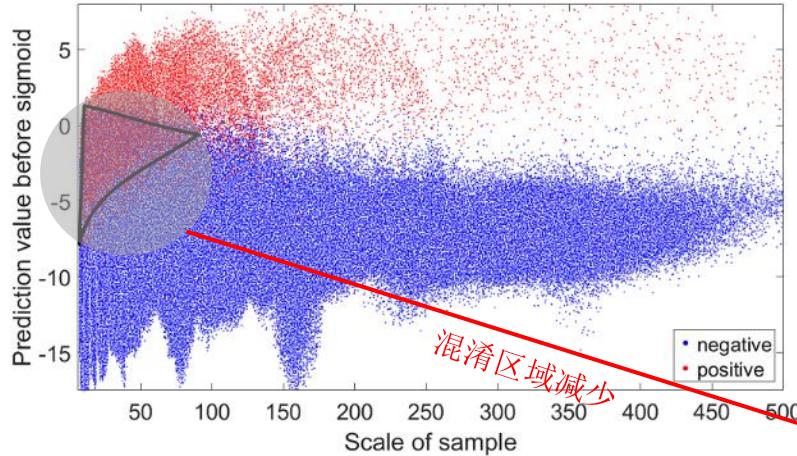
$$y = \text{sigmoid}(x - m), \text{ 其中 } m = \frac{15}{\sqrt{wh}}, h \text{ 和 } w \text{ 是正样本（锚框）的长和宽}$$





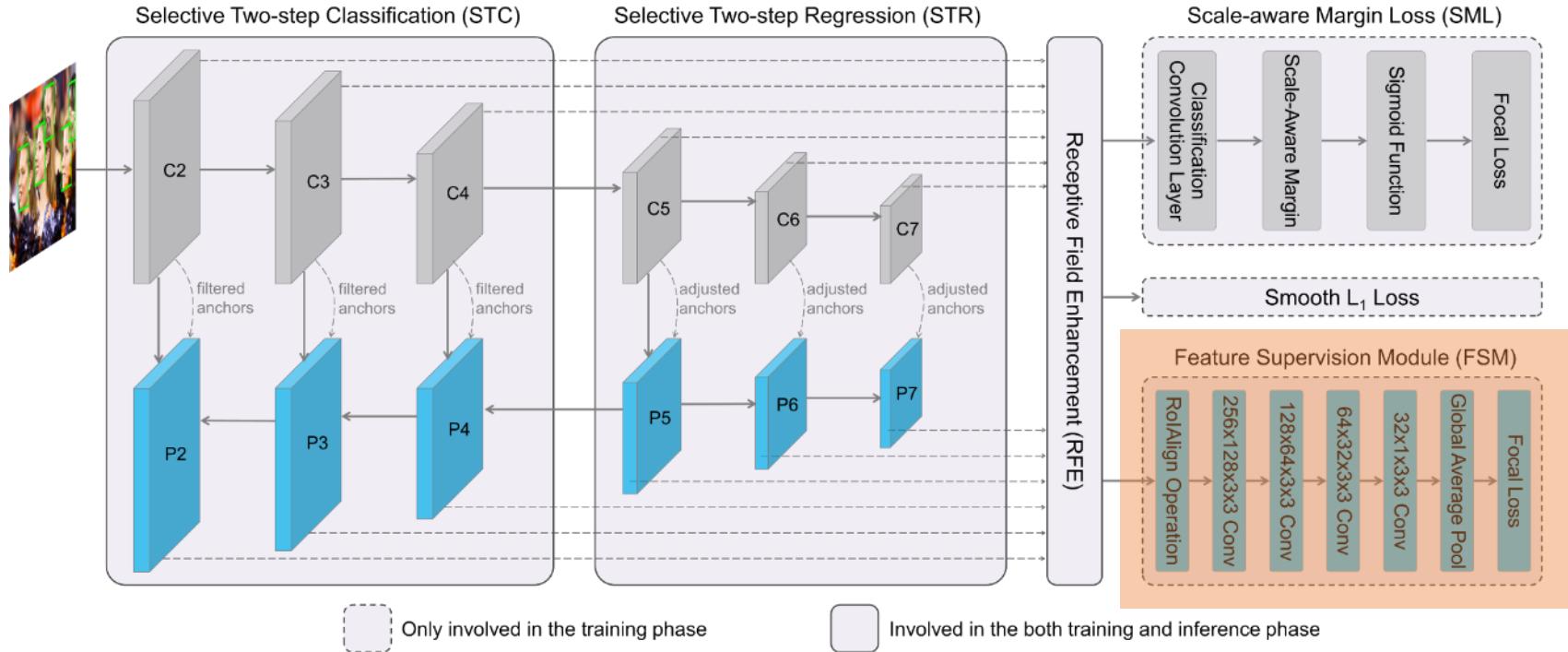
高精度的人脸检测算法RefineFace：尺度敏感margin

- 尺度敏感的margin策略，尽量让小尺度的人脸能够从复杂的背景中区分出来





高精度的人脸检测算法RefineFace：整体框架



■ RefineFace人脸检测算法 = SRN人脸检测算法 + 尺度敏感margin的损失函数 + 特征监督模块





高精度的人脸检测算法RefineFace：特征监督模块

- 单阶段法中，原始图像上的一个锚框A关联于检测层上的一个特征T
- 对这个锚框A的预测（分类和回归）都是使用它对应的特征T进行的
- 特征T在原图上有一个对应的感受野，即与特征T有关的那些像素组成的区域
- 所以对锚框A进行预测所使用的像素信息是来自于特征T对应的感受野
- 如右图所示，绿色的框是一个锚框，蓝色的框是对这个锚框进行预测时所使用的原始像素信息
- 从图中可以看出，对**绿色锚框**进行预测所使用的**蓝色区域信息**包含很多来自背景的干扰信息，对预测产生不好的影响



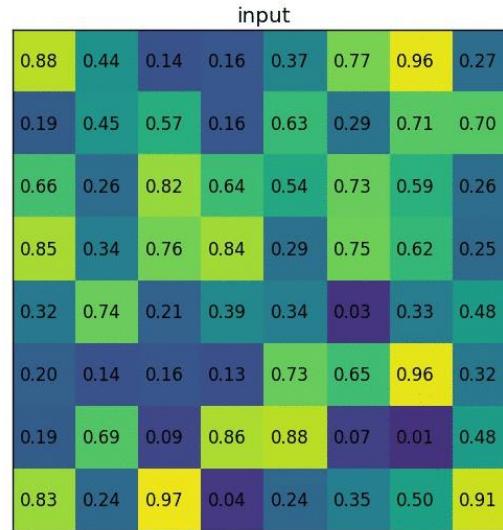
blue region information





高精度的人脸检测算法RefineFace：特征监督模块

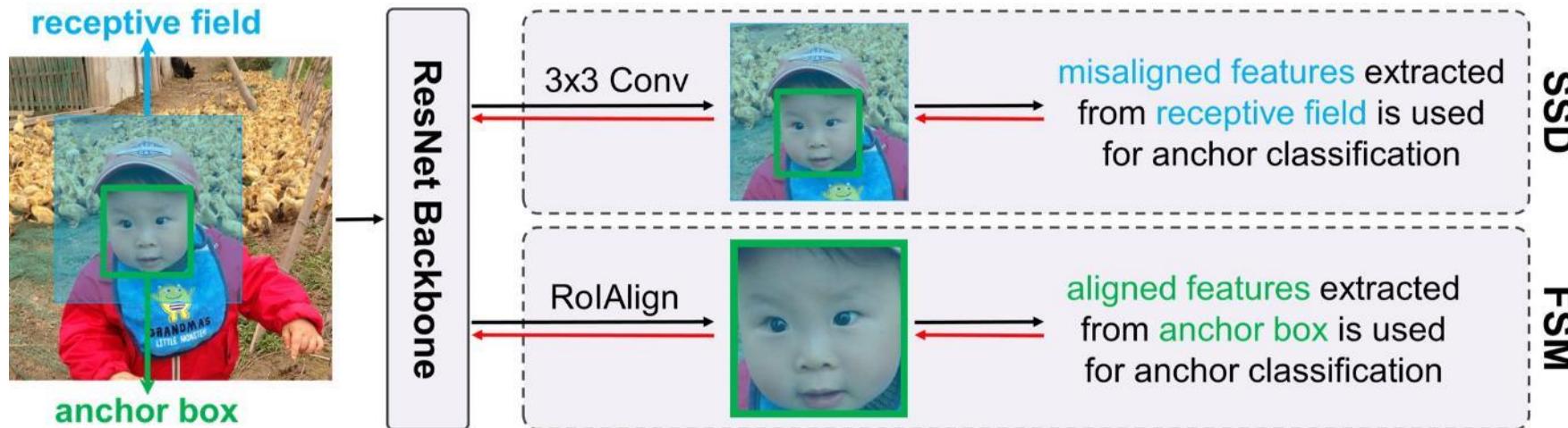
- 多阶段法中，第二阶段会使用RoIPooling，根据候选区域的精准位置，从检测层上抠出对应的特征
- RoIPooling得到的特征跟候选区域完全对应，即只包含候选区域内的信息，不包含候选区域外的任何信息
- RoIPooling操作能够实现特征校准，排除来自背景的干扰信息，集中于候选区域内的特征对其进行预测





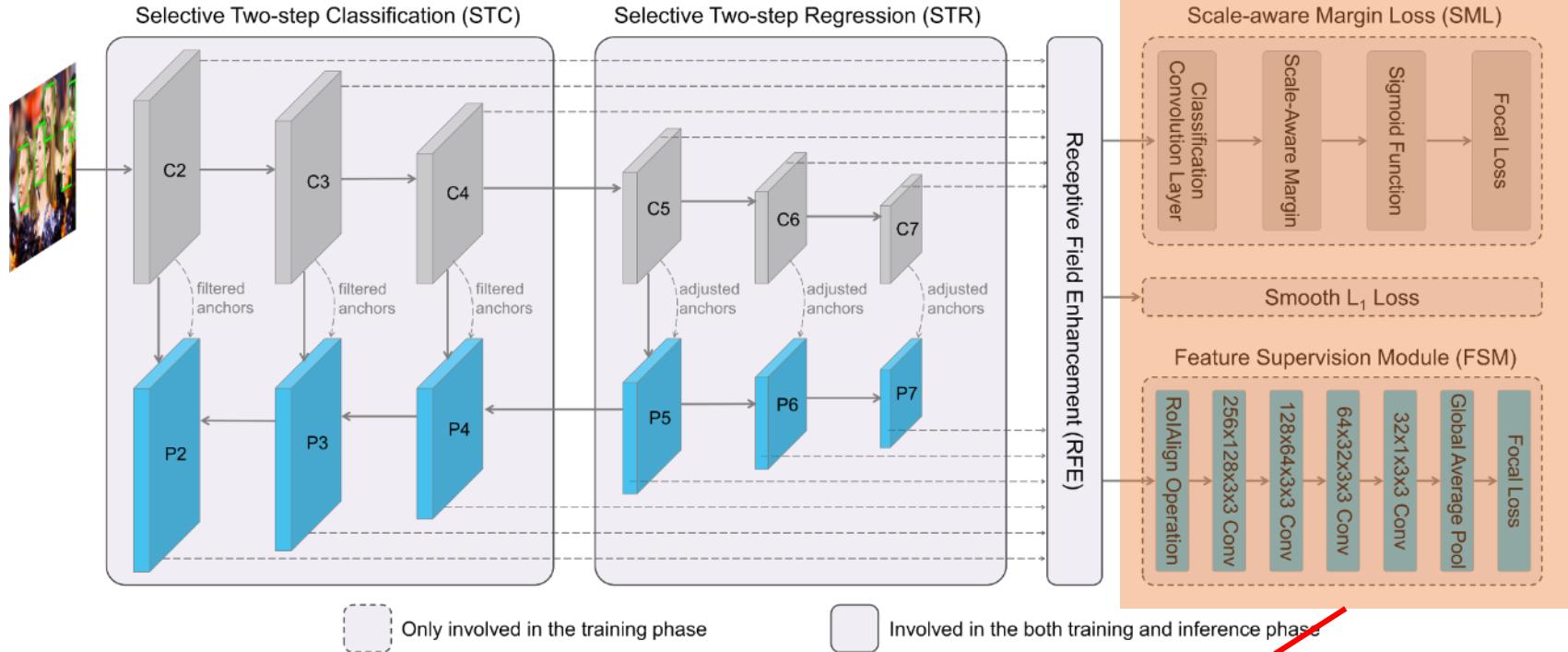
高精度的人脸检测算法RefineFace：特征监督模块

- 鉴于此，在**训练过程中**引入特征监督模块，在原有的单阶段人脸检测后，添加一个RoIAlign操作来监督特征学习
- RoIAlign是RoIPooling的升级版，**改进了RoIPooling操作中取整行为带来的信息偏差**，使得抠取出的特征更加精准
- 在训练阶段中，对于检测结果利用RoIAlign抠取特征进行一个额外的二分类任务来学习更加精准的特征表达





高精度的人脸检测算法RefineFace：整体框架

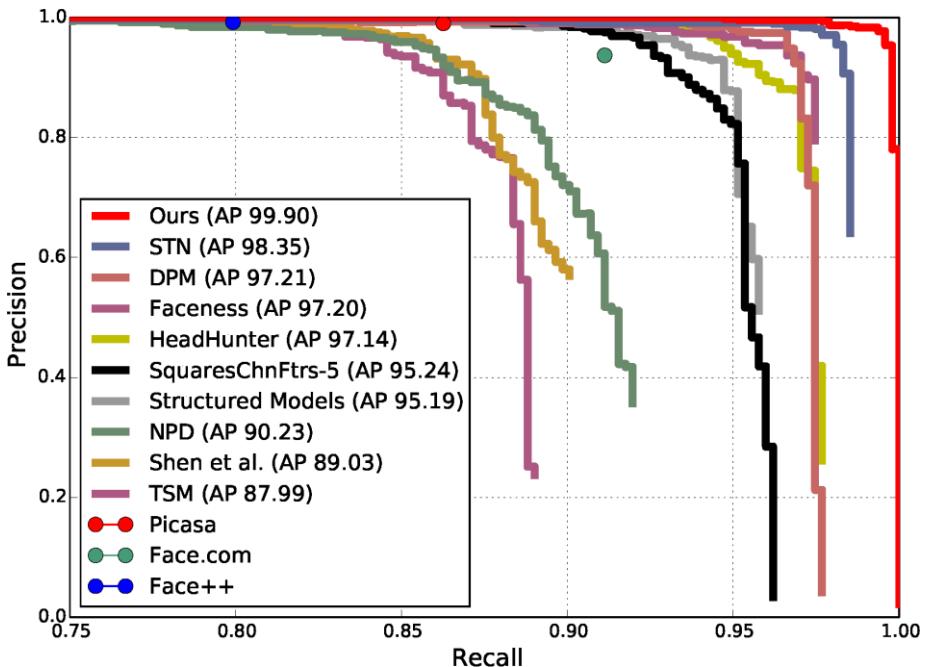


只出现在训练阶段，测试阶段不参与，因此不会引入额外的计算

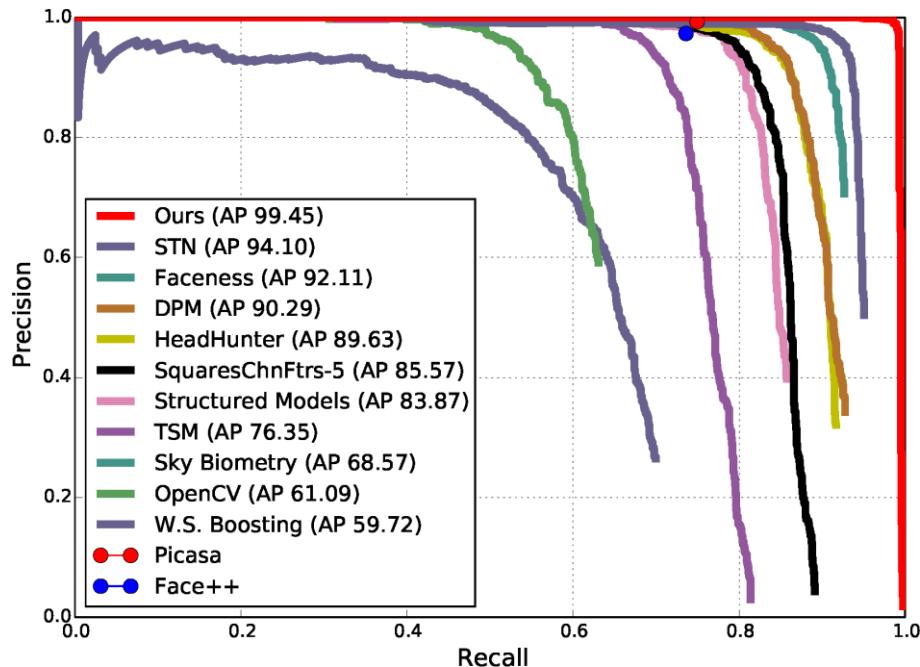




高精度的人脸检测算法RefineFace：检测精度



AFW

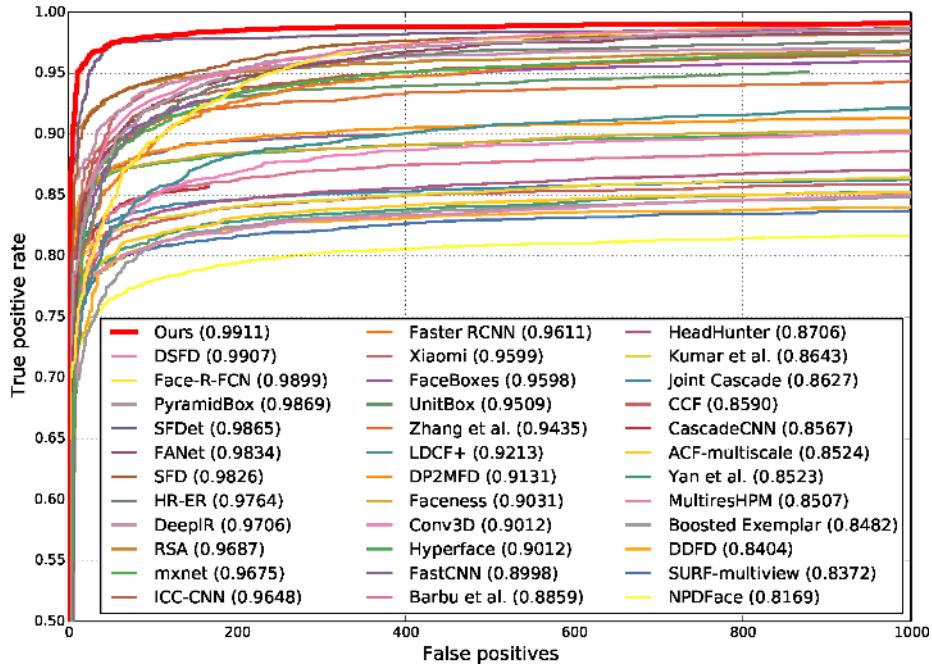


PASCAL Face





高精度的人脸检测算法RefineFace：检测精度



Methods	Whole set	Masked set	Unignored set
TSM [9]	-	-	41.6
HeadHunter [19]	-	-	50.9
HPM [75]	-	-	60.0
MTCNN [25]	-	-	60.8
LLE-CNNs [63]	-	-	76.4
FAN [45]	-	76.5	88.3
AOFD [83]	81.3	83.5	91.9
SFDet [67]	81.5	94.0	93.7
Ours	83.9	96.2	95.7

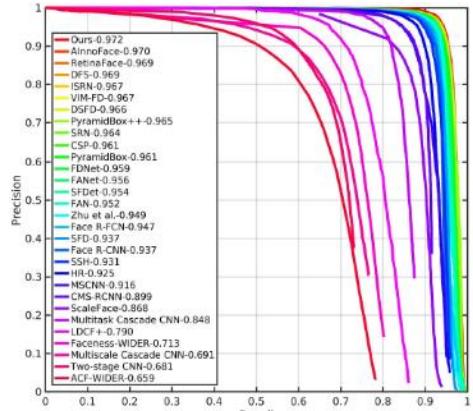
FDDB

MAFA

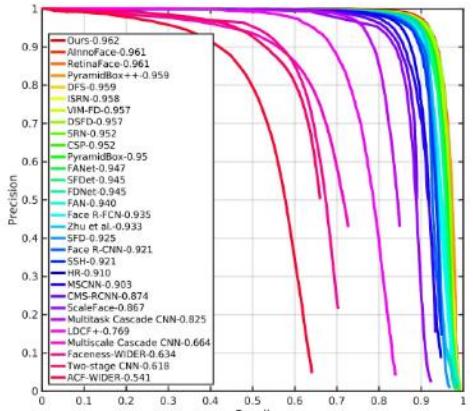




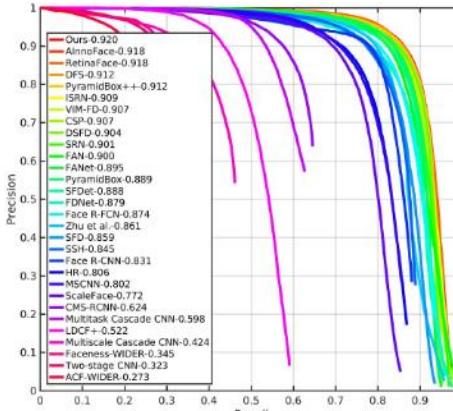
高精度的人脸检测算法RefineFace：检测精度



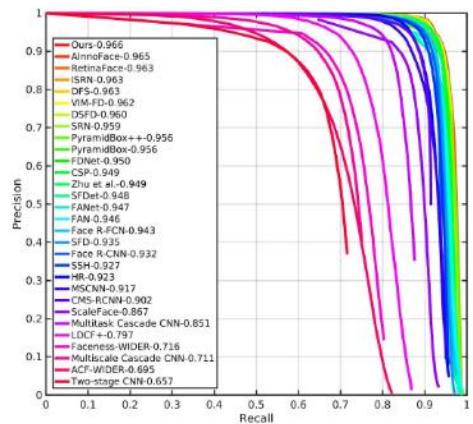
(a) Val: Easy



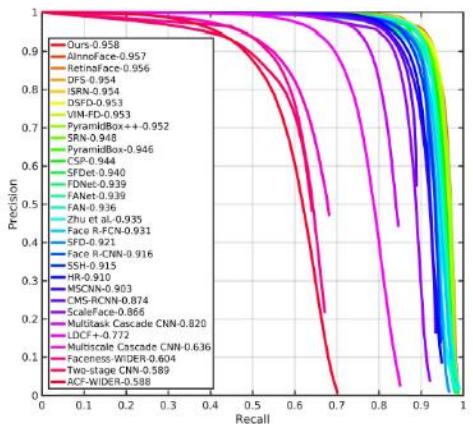
(b) Val: Medium



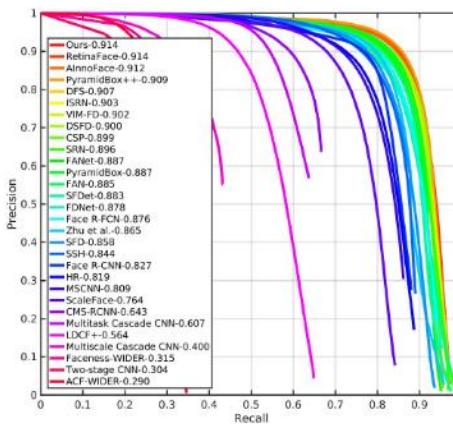
(c) Val: Hard



(d) Test: Easy



(e) Test: Medium



(f) Test: Hard

验证集

- 97.2 (Easy)
- 96.2 (Medium)
- 92.0 (Hard)

测试集

- 96.6 (Easy)
- 95.8 (Medium)
- 91.4 (Hard)





高精度的人脸检测算法RefineFace：检测速度

- 在英伟达显卡GTX 1080 Ti上进行速度的测试
- 使用ResNet不同深度的4个网络作为基础结构
- 速度是针对VGA图像（640x480）进行测试的
- 精度是在WIDER FACE不同子集上进行测试的

Backbone	Speed		Accuracy (%)		
	FPS	ms	Easy	Medium	Hard
ResNet-18	37.3	26.8	96.3	95.1	90.2
ResNet-50	28.5	35.1	96.9	95.9	91.1
ResNet-101	22.8	43.9	97.1	96.1	91.6
ResNet-152	17.7	56.6	97.2	96.2	92.0





高精度的深度学习后期人脸检测算法：总结

- 高精度算法：运行在**计算资源充足**的后端设备，如英伟达显卡2080Ti、RTX



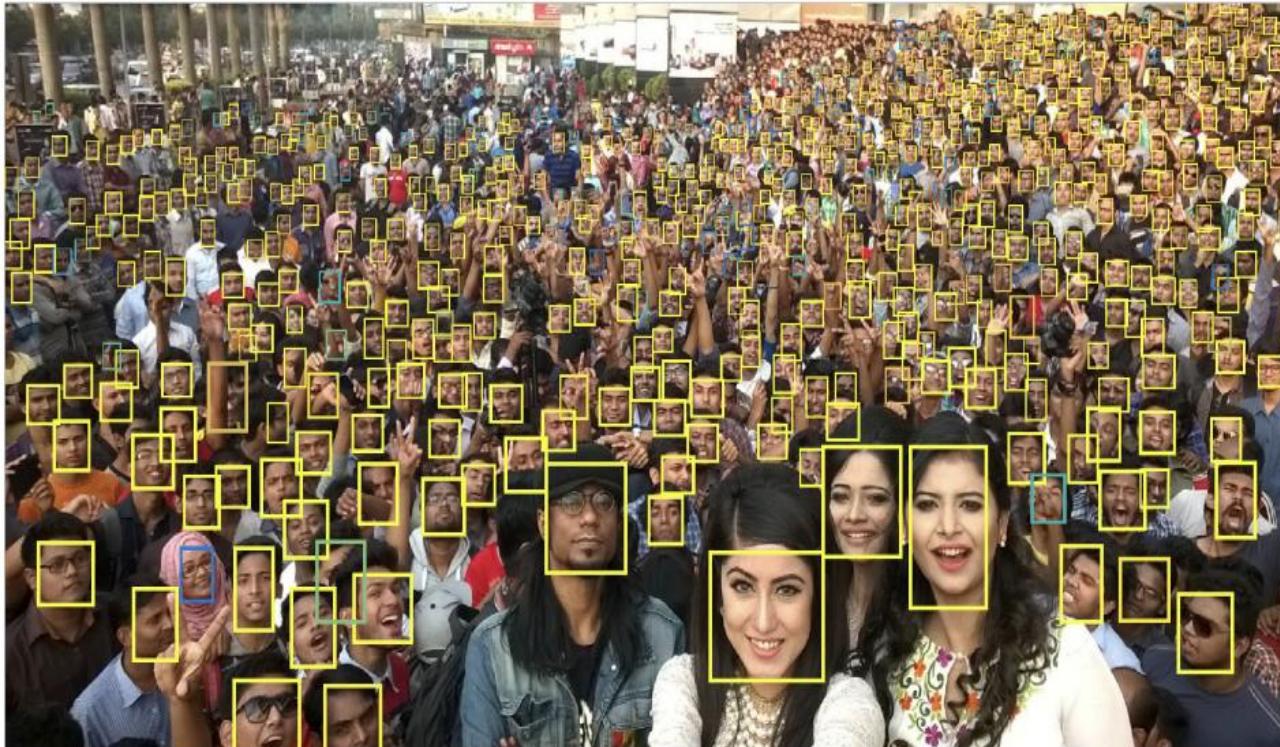
- 高精度算法：由于资源充足，需要在**所有场景**下都满足精度需求（任何尺度人脸，背景非常复杂）





高精度的深度学习后期人脸识别算法：总结

- 在数据增广、特征增强、锚框设计、锚框匹配、锚框预测、损失函数等方面进行改机
- 解决小尺度人脸识别这个痛点问题，在各个场景取得非常高的精度，具备很强的竞争性



竞争性

速度

精度

WIDER FACE

FDDB

PASCAL Face

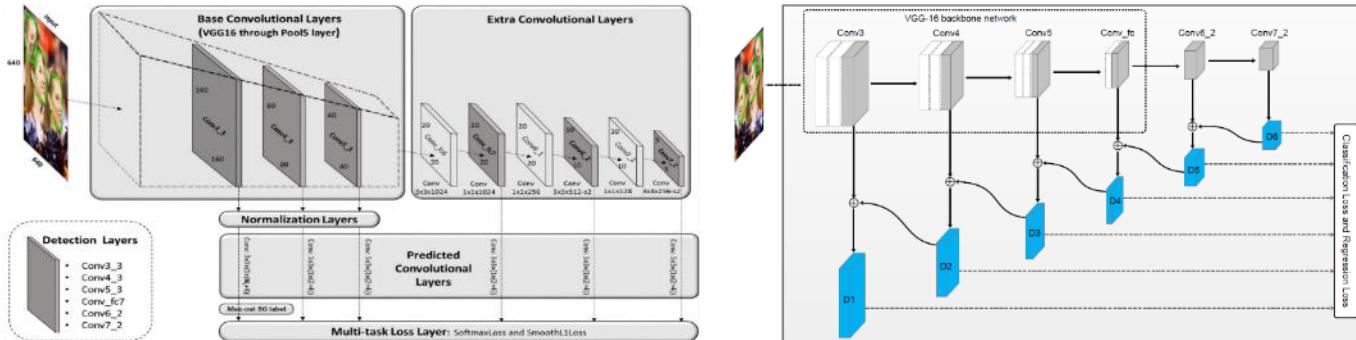
AFW

GPU



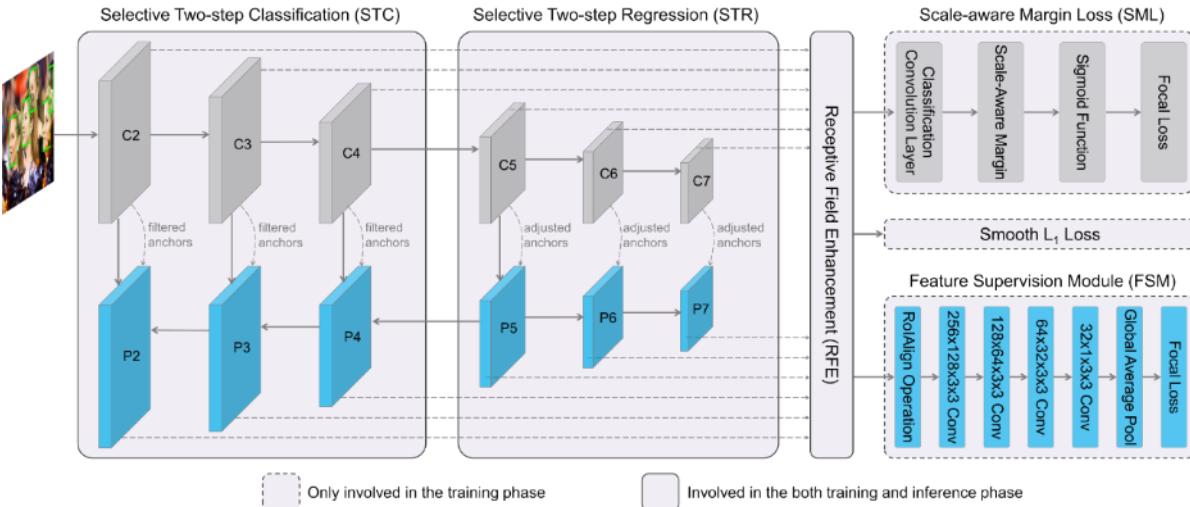


高精度的深度学习后期人脸检测算法：总结



S³FD & SFDet

- ① 尺度上公平的检测框架
- ② 尺度补偿的锚框匹配策略
- ③ 背景标签输出最大化操作



SRN & RefineFace

- ① 选择性二阶段分类
- ② 选择性二阶段回归
- ③ 感受野增强模块
- ④ 尺度敏感的margin损失函数
- ⑤ 特征监督模块





课程作业

■ WIDER FACE数据集上人脸检测算法的调研报告

1. WIDER FACE数据集的排行榜单 http://shuoyang1213.me/WIDERFACE/WiderFace_Results.html
2. 总结WIDER FACE数据集自从2016年11月20日出现后，它榜单上的检测精度的发展历程曲线
3. 归纳WIDER FACE数据集上各个算法都是基于哪些算法进行改进的
4. 整理WIDER FACE数据集上哪些算法有官方代码，哪些算法有第三方实现代码，哪些算法没有代码





结语

感谢聆听 !
Thanks for Listening !

