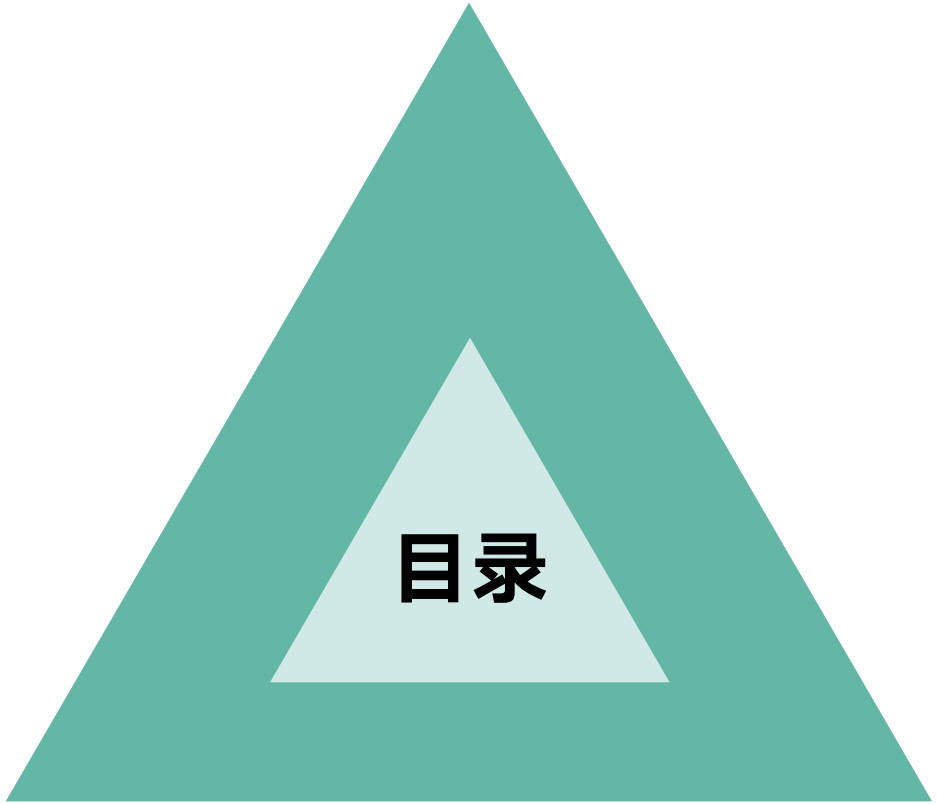
Accelerating Proposal Generation Network for Fast Face Detection on Mobile Devices

制作人：以若



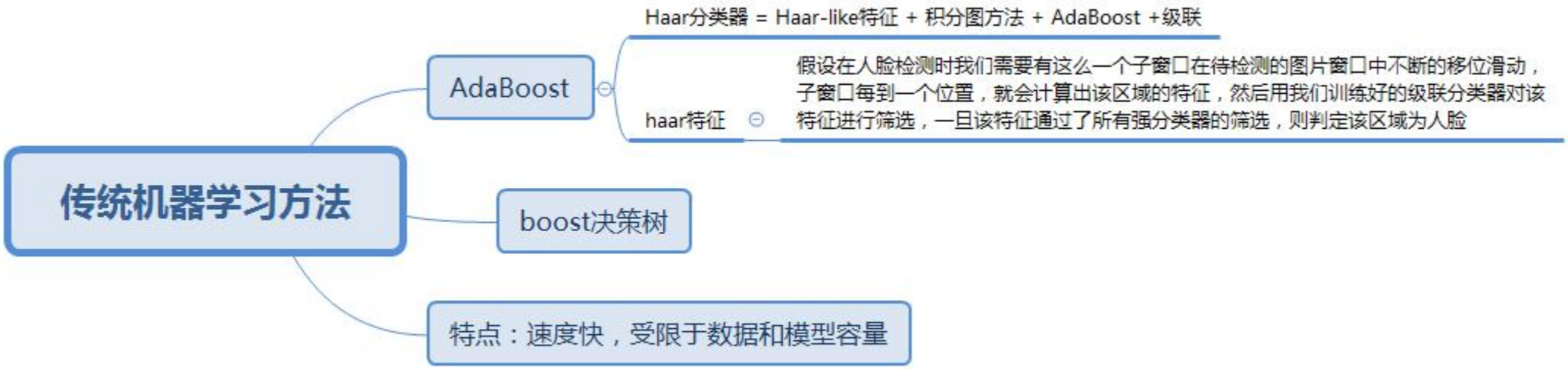
01 Introduction

一、Introduction

- Face detection

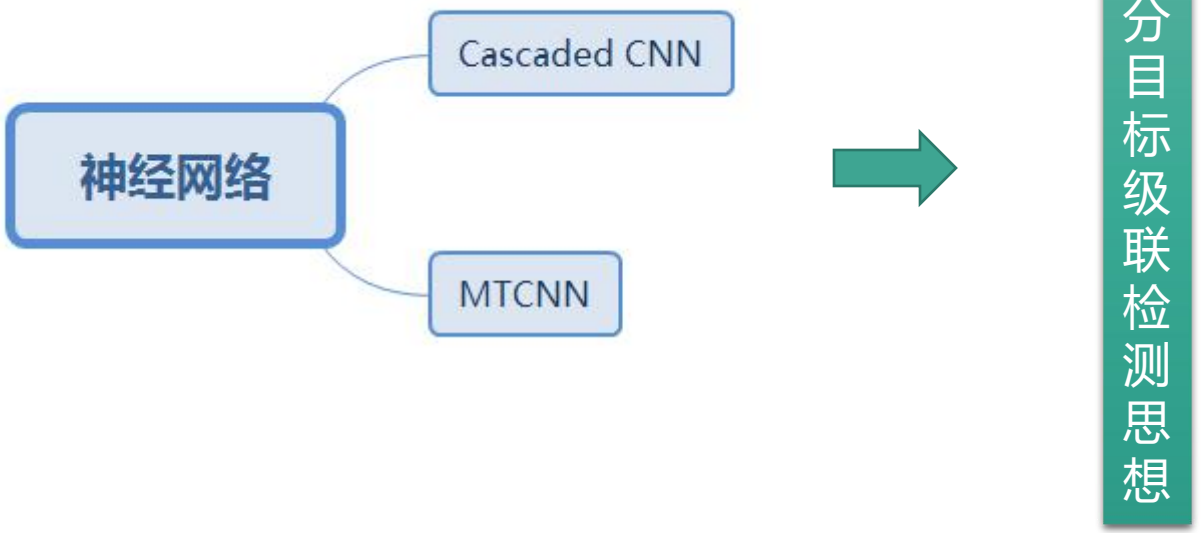
一、Introduction

- State-of-art machine learning



一、Introduction

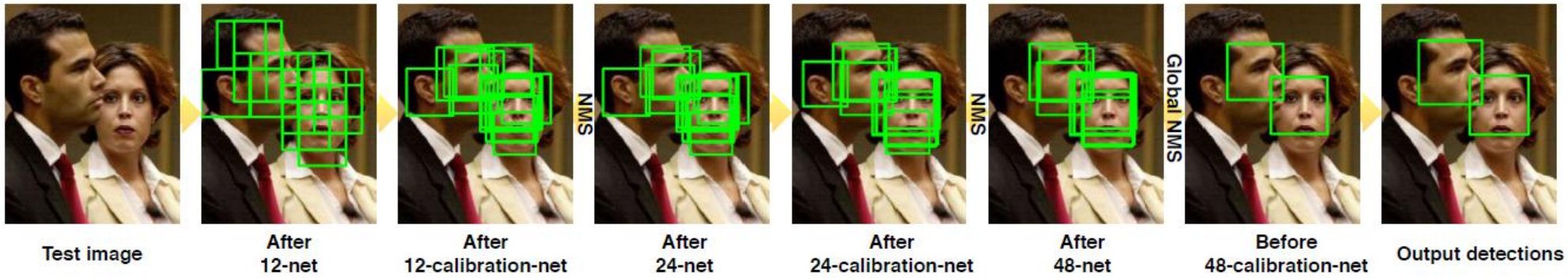
- State-of-art deep learning

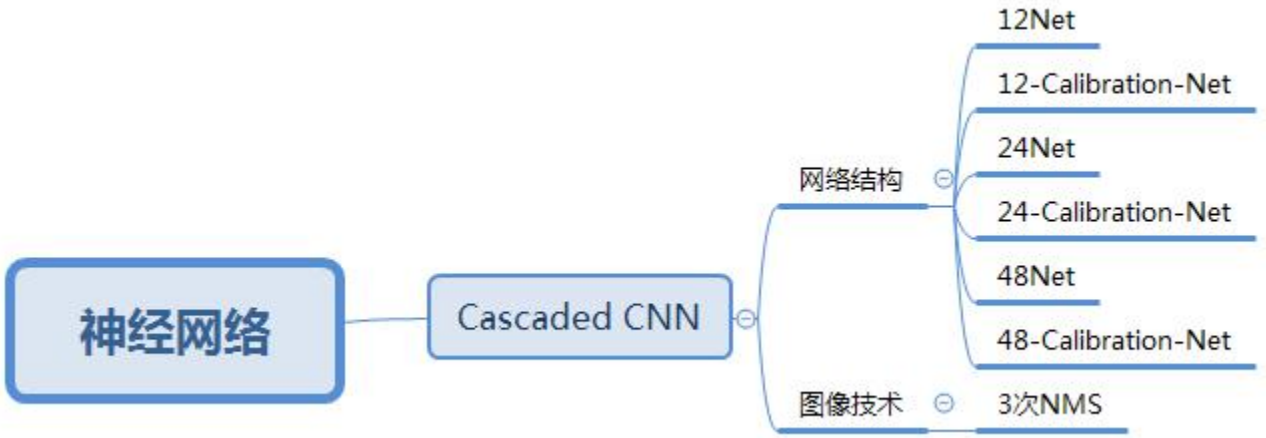


一、Introduction

- State-of-art deep learning --- Cascaded CNN

Cascaded CNN Pipeline

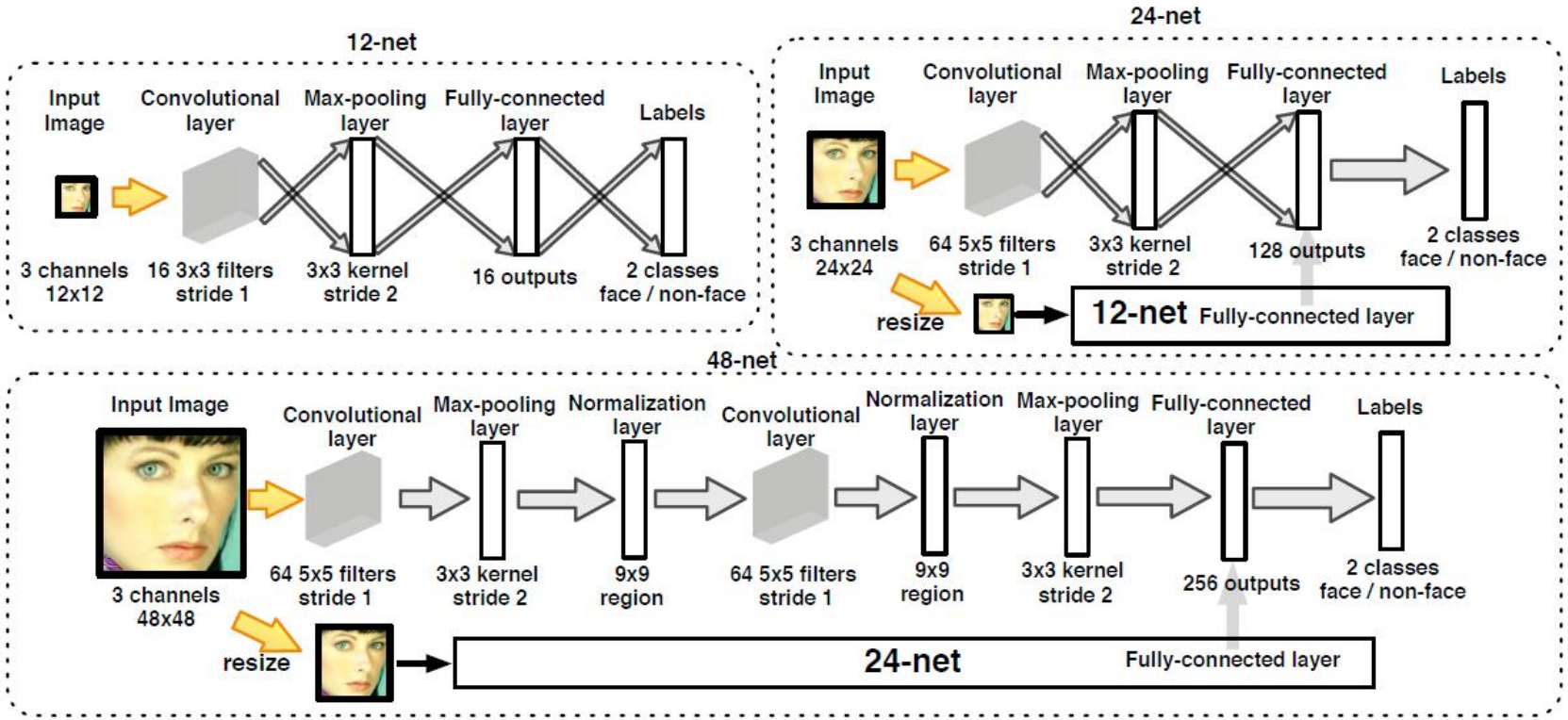




一、Introduction

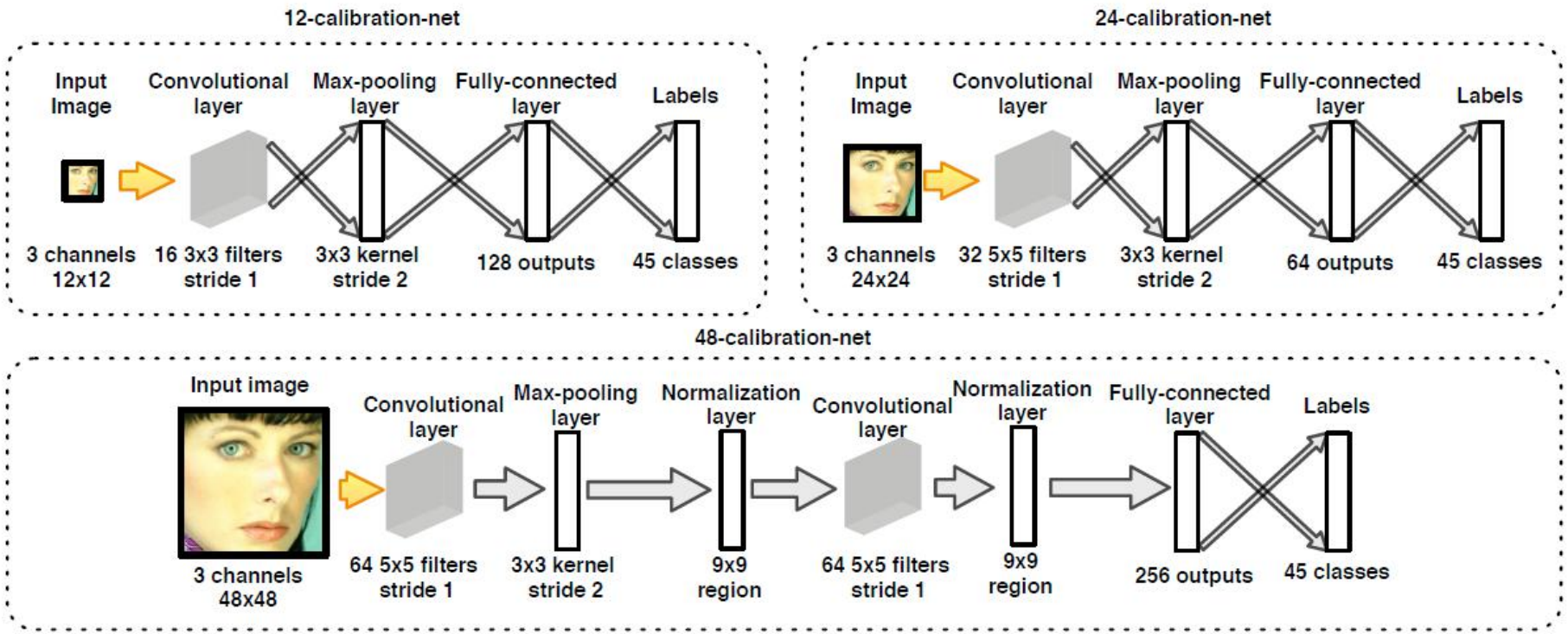
- State-of-art deep learning --- Cascaded CNN

Face / No Face Classifier



一、Introduction

- State-of-art deep learning --- Cascaded CNN



Bounding Box Calibration

一、Introduction

- State-of-art deep learning --- Cascaded CNN

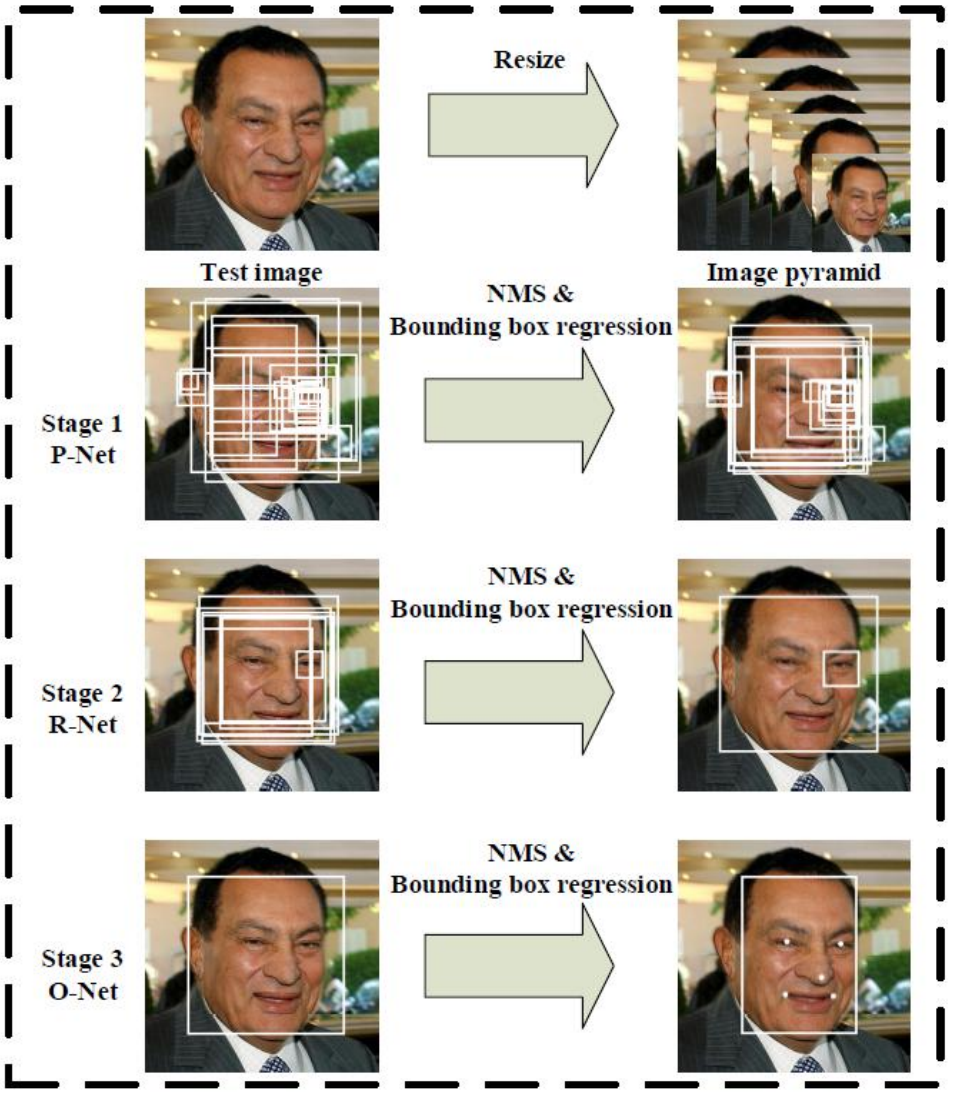
Non-maximum suppression (NMS)

- 所有bbox取最大score
- 筛选出所有与之iou比较大的bbox
- 再迭代地选择次大的score
- 再继续筛选
- 直至所有bbox筛选完毕

一、Introduction

- State-of-art deep learning

Multi-Task Cascaded CNN



一、Introduction

- State-of-art deep learning -- MTCNN

Multi-Task Cascaded CNN Pipeline

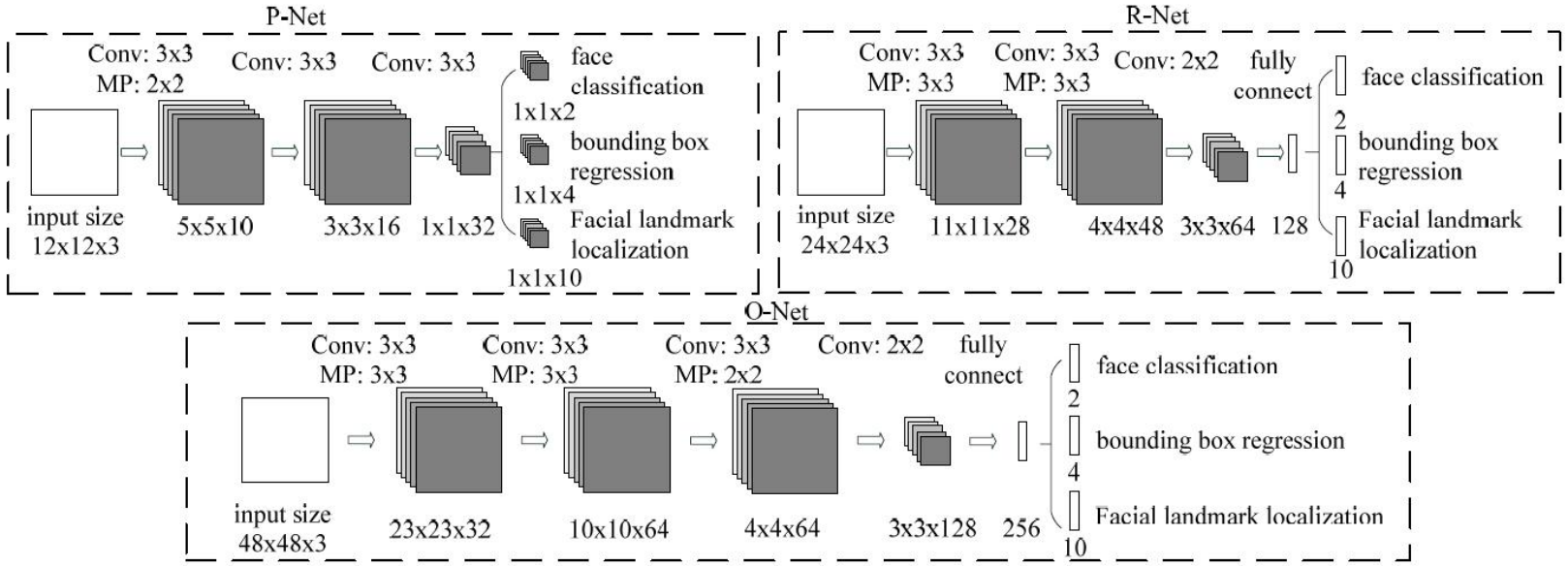


Fig. 2. The architectures of P-Net, R-Net, and O-Net, where "MP" means max pooling and "Conv" means convolution. The step size in convolution and pooling is 1 and 2, respectively.

- 每个网络同时做人脸分类+校准+landmark定位的输出；
- 相当于一个网络做三个任务；
- 训练时分类+定位都有loss、测试时只有o-net输出landmark);

一、Introduction

- State-of-art deep learning -- MTCNN

Multi-Task Cascaded CNN Loss

Loss function

face

$$L_i^{det} = -(y_i^{det} \log(p_i) + (1 - y_i^{det})(1 - \log(p_i))) \quad (1)$$

box

$$L_i^{box} = \left\| \hat{y}_i^{box} - y_i^{box} \right\|_2^2$$

landmark

$$L_i^{landmark} = \left\| \hat{y}_i^{landmark} - y_i^{landmark} \right\|_2^2$$

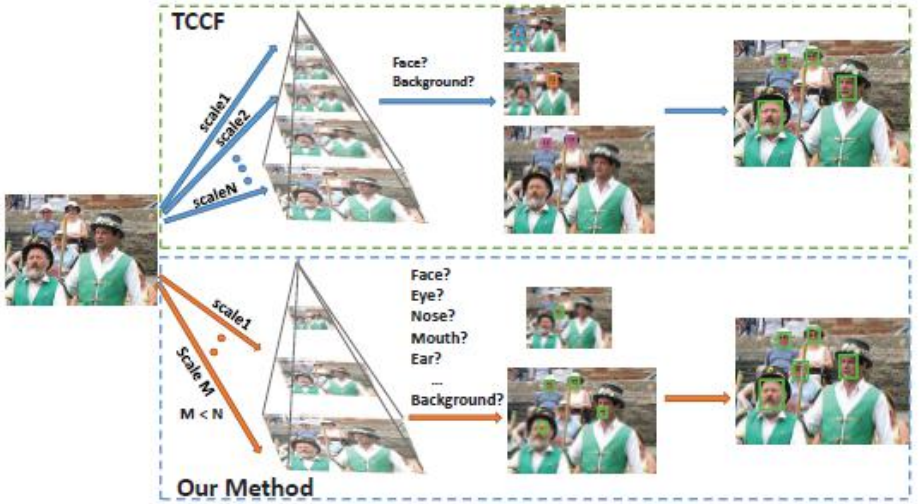
02 Structure

一、Structure

- Structure

Spotlight

- 抓住全局和局部面部特征去减少输入图片金字塔的level;
- 满足手机设备的存储和速度;
- 使用局部特征可以快速检测到脸的位置，相比于全脸检测来说;



一、Structure

- Structure

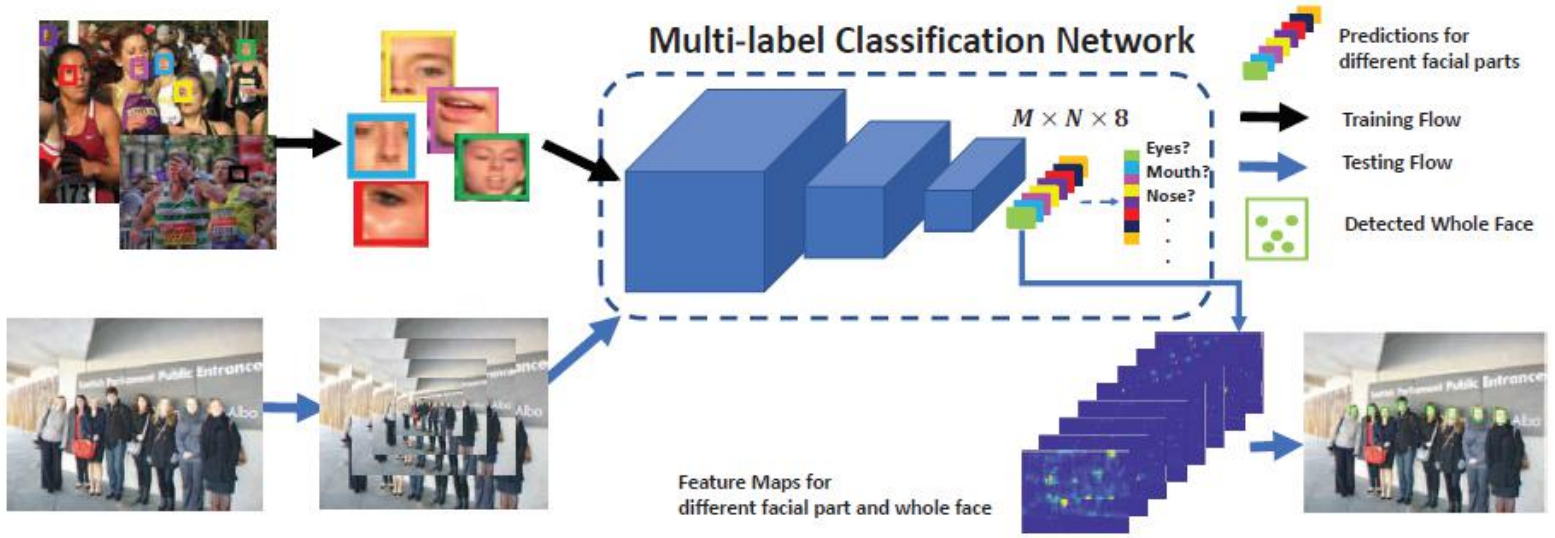


Table 1. Proposal network architecture

Layer	Kernel size	Output size
Input		12×12×3
Conv1	3×3	12×12×16
Pool1	3×3	6×6×16
Conv2	3×3	4×4×32
Conv3	3×3	2×2×32
Conv4	2×2	1×1×64
Conv5	1×1	1×1×8

寻找局部最优值 → 使用模板生成bounding box → 将box组合 → R-Net O-Net

step1 step2 step3 final

03 Experiments

一、Structure

- Structure

Table 2. Comparisons of model size with state-of-the-art networks.

Work	Model size
CEDN [20]	1.1GB*
DDFD [10]	233MB*
HR [14]	98.9MB
LCDF+ [9]	2.33MB
MTCNN [2]	1.9MB
Nested [19]	1.6MB*
Ours	1.96MB

Table 3. Comparisons of detection performance with MTCNN[2] on WIDER-face validation set [13] with different scale factors.

Scale factor		0.79	0.50	0.25
Easy	MTCNN [2]	0.836	0.817	0.755
	Ours	0.844	0.842	0.826
Medium	MTCNN [2]	0.809	0.798	0.744
	Ours	0.809	0.805	0.794
Hard	MTCNN [2]	0.622	0.600	0.529
	Ours	0.603	0.568	0.519

Speed（Matlab Code）
MTCNN： 40.1ms
本文： 39.1ms

THANKS