

# 物体检测总结展望



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：课程大纲

## 第1章：科研经验分享（1节课）

- 讲师自我介绍
- 科研经验分享
- 课程安排简介

## 第2章：物体检测概述（1节课）

- 物体检测研究背景
- 物体检测发展脉络
- 物体检测常用数据集
- 物体检测评价指标

## 第3章：通用物体检测（5节课）

- 物体检测环境配置
- 通用物体检测概述
- 基于锚框的检测算法
- 无需锚框的检测算法
- 物体检测算法的对比总结
- 实用检测算法的研究思路

## 第4章：人脸检测（3节课）

- 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

## 第5章：行人检测（2节课）

- 行人检测概述
- 传统行人检测算法
- 深度学习早期行人检测算法
- 深度学习后期行人检测算法

## 第6章：检测总结展望（1节课）

- 物体检测发展总结
- 物体检测未来展望



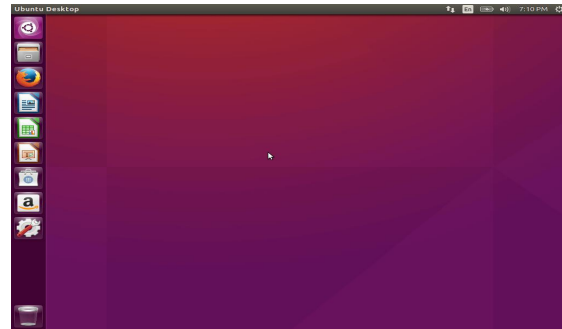
## 内容回顾：硬件软件



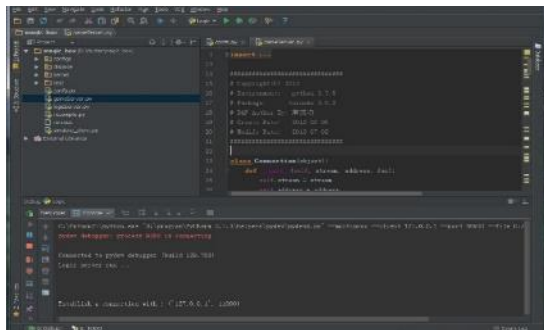
服务器



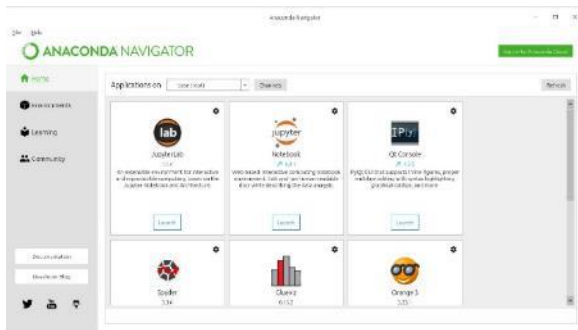
NVIDIA显卡



LINUX操作系统



代码调试软件PyCharm



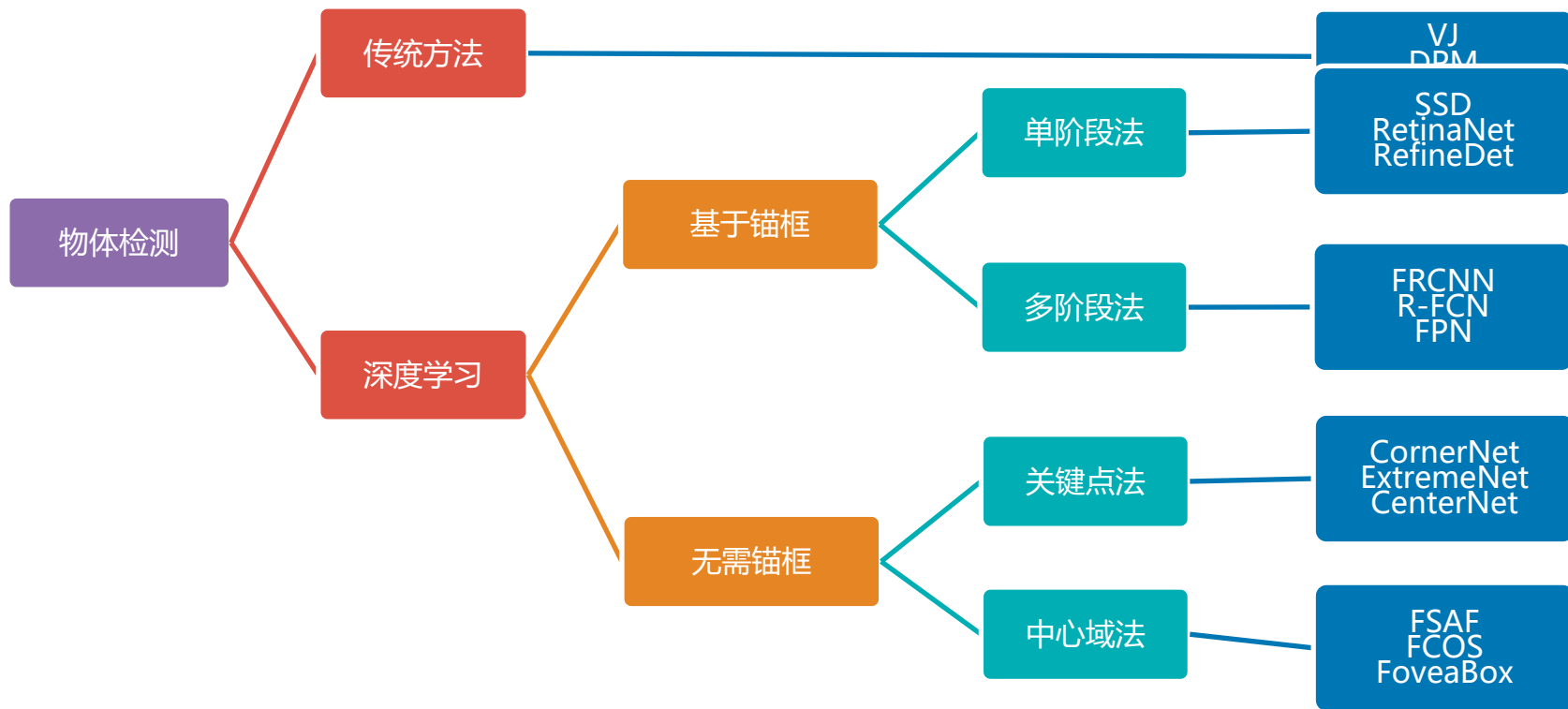
环境管理软件Anaconda



物体检测平台Detectron2



## 内容回顾：物体检测的派系





## 内容回顾：物体检测常用数据集

### 通用物体检测数据集

- PASCAL VOC
- MS COCO
- OpenImages
- LVIS

### 人脸检测数据集

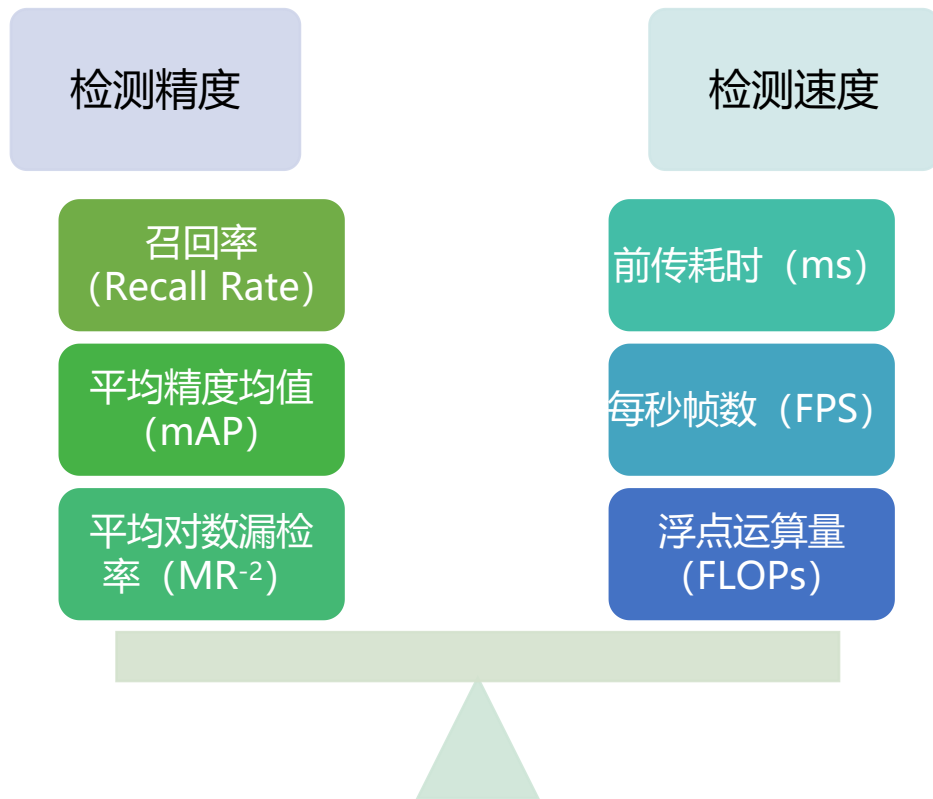
- AFW
- PASCAL FACE
- MALF
- MAFA
- FDDB
- WIDER FACE

### 行人检测数据集

- Caltech-USA
- CityPersons
- CrowdHuman
- WiderPerson
- EuroCityPersons



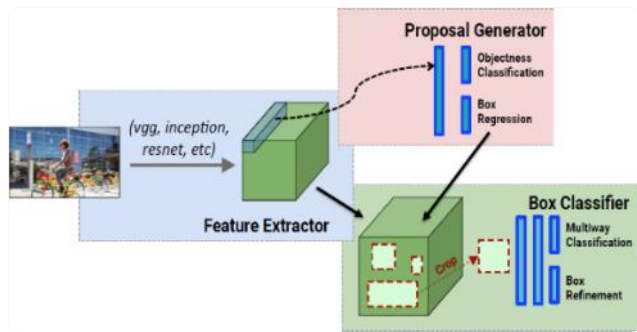
## 内容回顾：物体检测评价指标



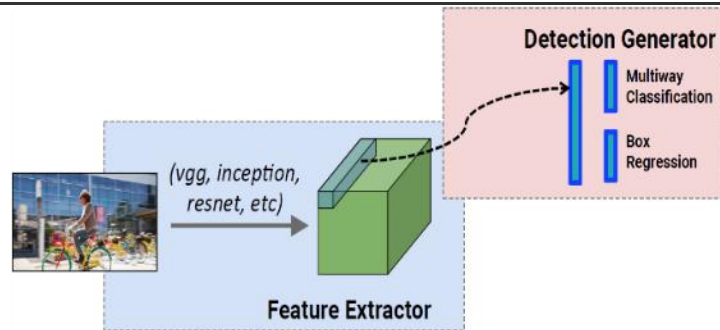


## 内容回顾：通用物体检测

### 基于锚框

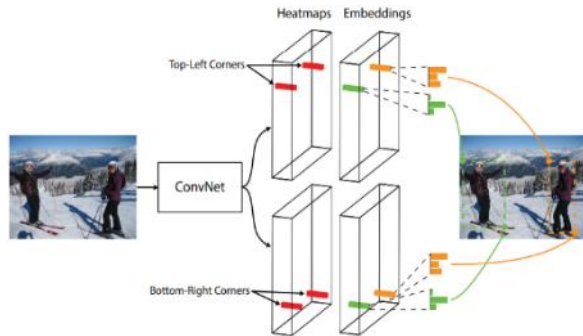


多阶段法

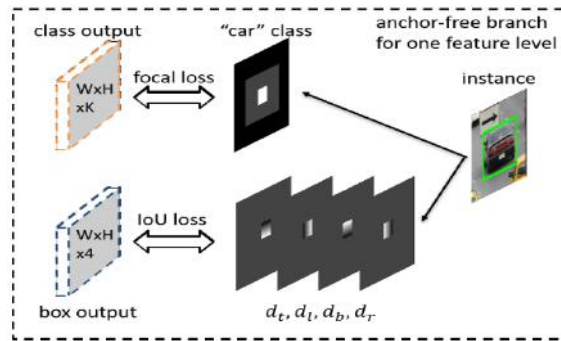


单阶段法

### 无需锚框



关键点法

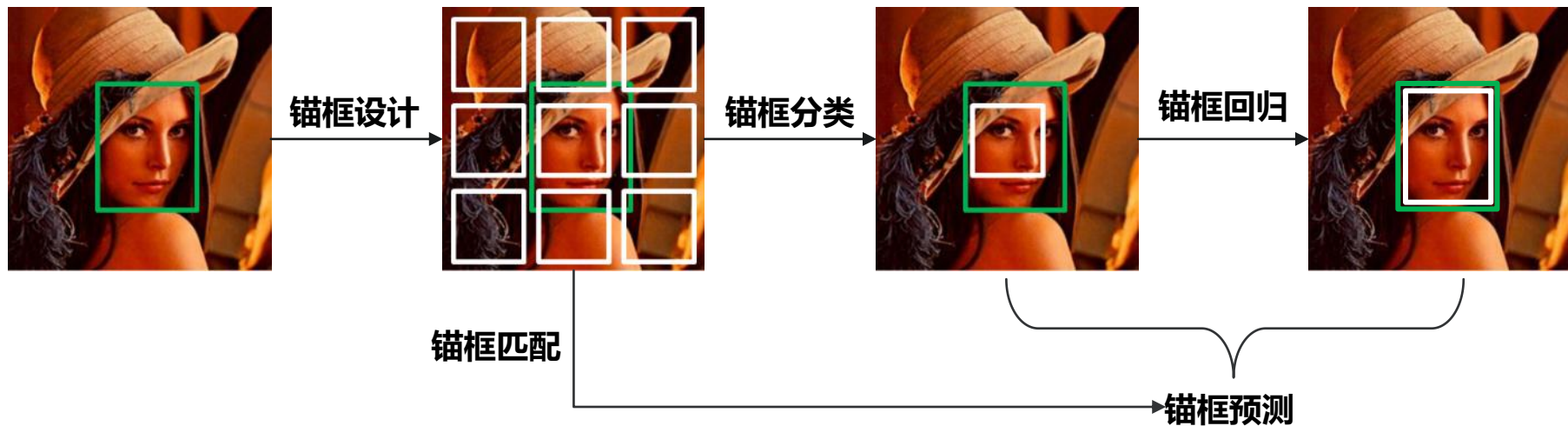


中心域法



## 内容回顾：通用物体检测

### 基于锚框的单阶段法 (SSD/RetinaNet)

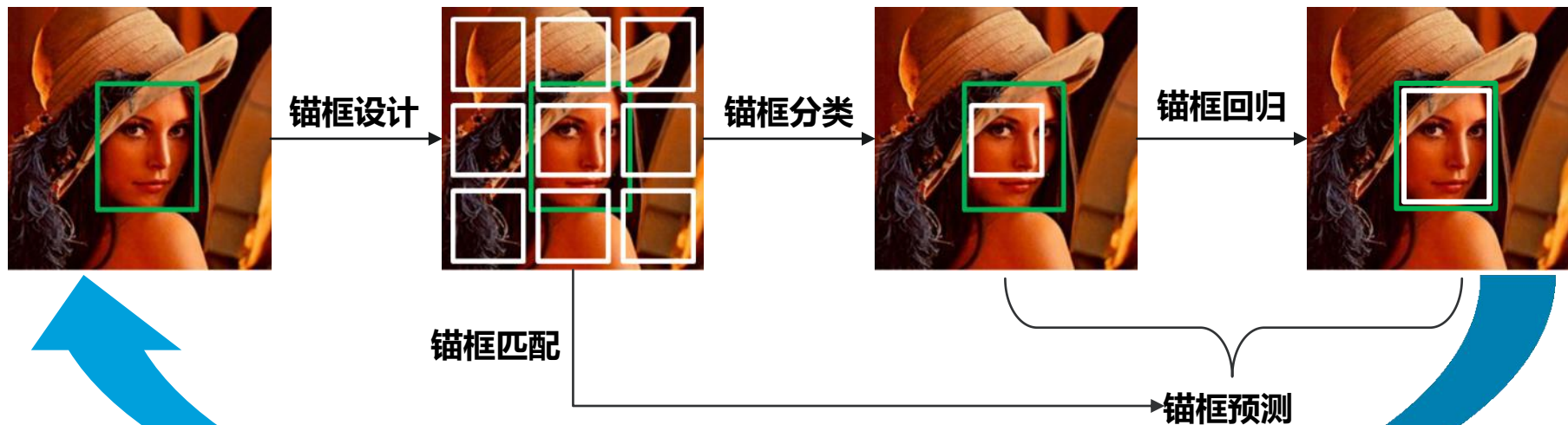






## 内容回顾：通用物体检测

### 基于锚框的多阶段法（Faster R-CNN）

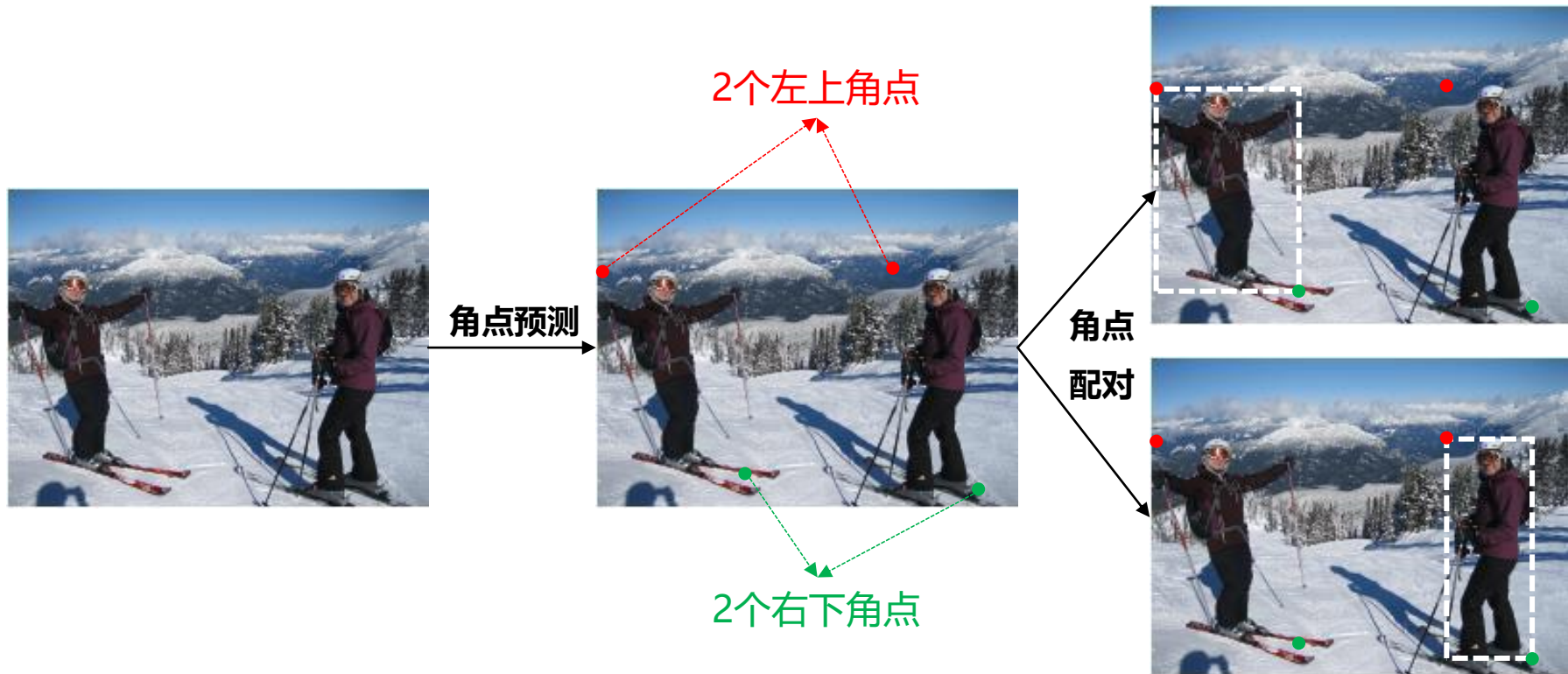


级联地重复这个过程



## 内容回顾：通用物体检测

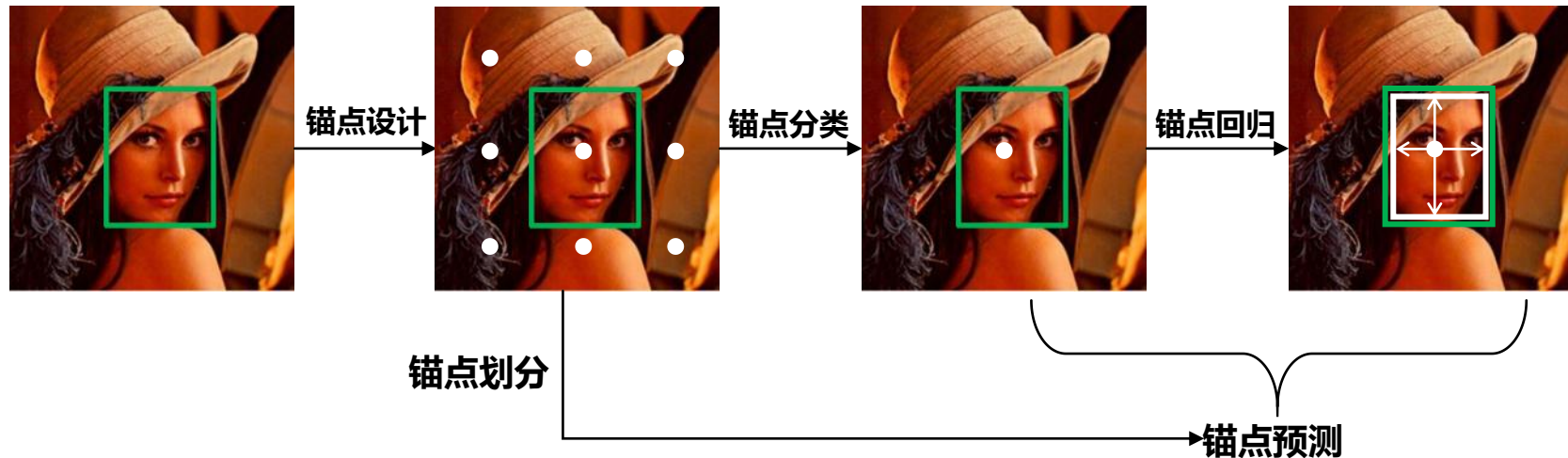
### 无需锚框的关键点法 (CornerNet)





## 内容回顾：通用物体检测

### 无需锚框的中心域法（FCOS）





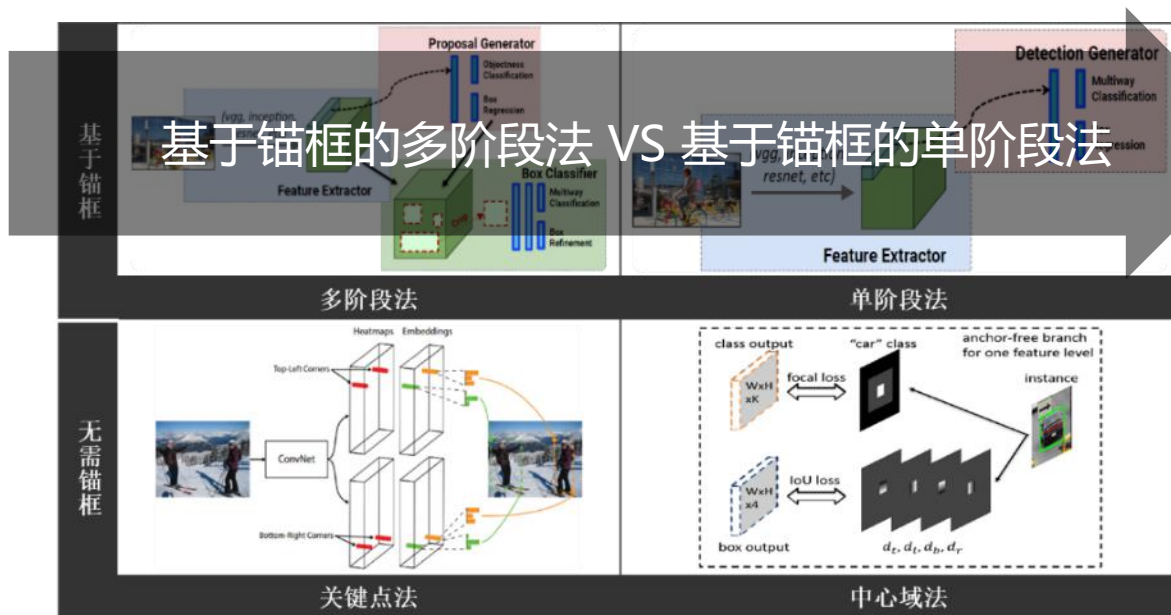
## 内容回顾：通用物体检测

基于锚框		多阶段法	单阶段法
相同点	检测思想	铺设的锚框为检测起点，对锚框的类别和位置进行矫正	
	检测起点	铺设的锚框	
	检测结果	矫正的锚框	
不同点	难点问题之一	小尺度物体	正负样本的平衡
	锚框矫正次数	$\geq 2$ 次	1次
	检测精度	较高	较低
	检测速度	较慢	较快
无需锚框	关键点法		中心域法
算法动机	移除掉锚框，减少超参数，增加灵活性		
算法思想	先检测关键点，再进行配对来框定物体		铺设锚点替代锚框来检测物体
算法优点	全新的检测流程，为检测带来了新的思路		减少超参数，简化计算
算法难点	不同关键点之间的配对问题		正负样本的划分问题
计算速度	流程比较复杂，速度相对较慢		流程比较简单，速度相对较快
检测精度	精度能达到甚至超过基于锚框的单阶段法		



## 内容回顾：通用物体检测

- 对比不同类型的检测算法，探索两者之间的本质区别
- 取之长补己短的改进思路，提出面向实用的全新算法



① 二阶段的分类

② 二阶段的回归

③ 二阶段的特征

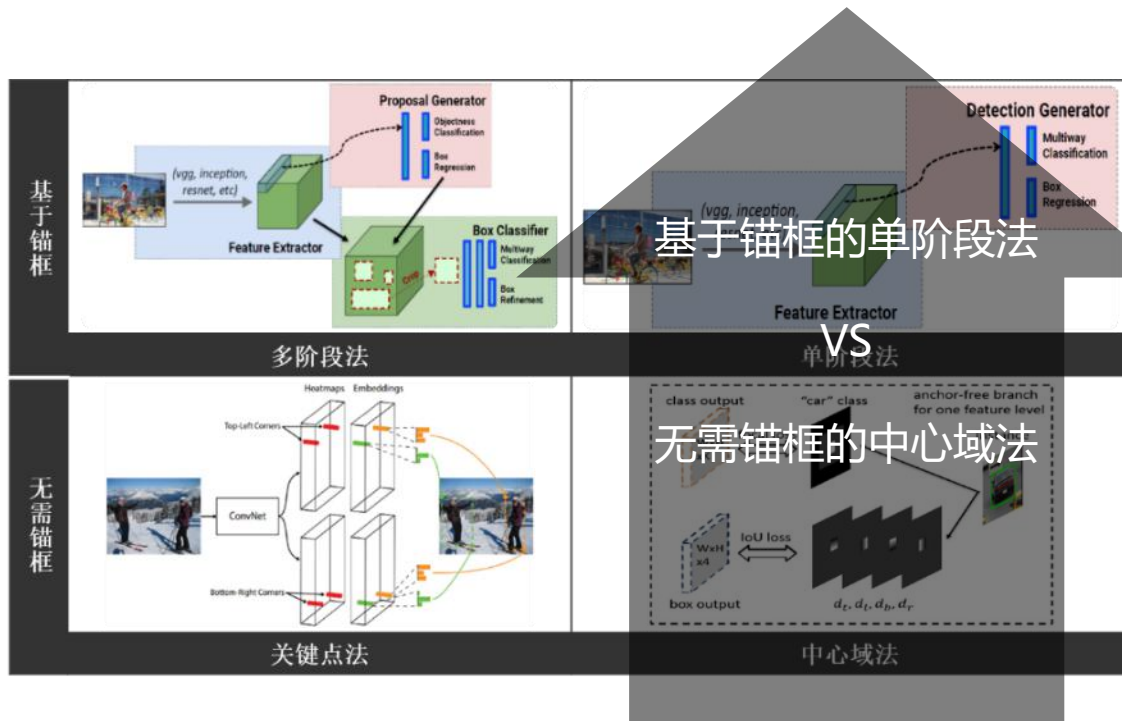
④ 特征校准

RefineDet算法



## 内容回顾：通用物体检测

- 对比不同类型的检测算法，探索两者之间的本质区别
- 取之长补己短的改进思路，提出面向实用的全新算法



① 正负样本定义

② 回归起始状态

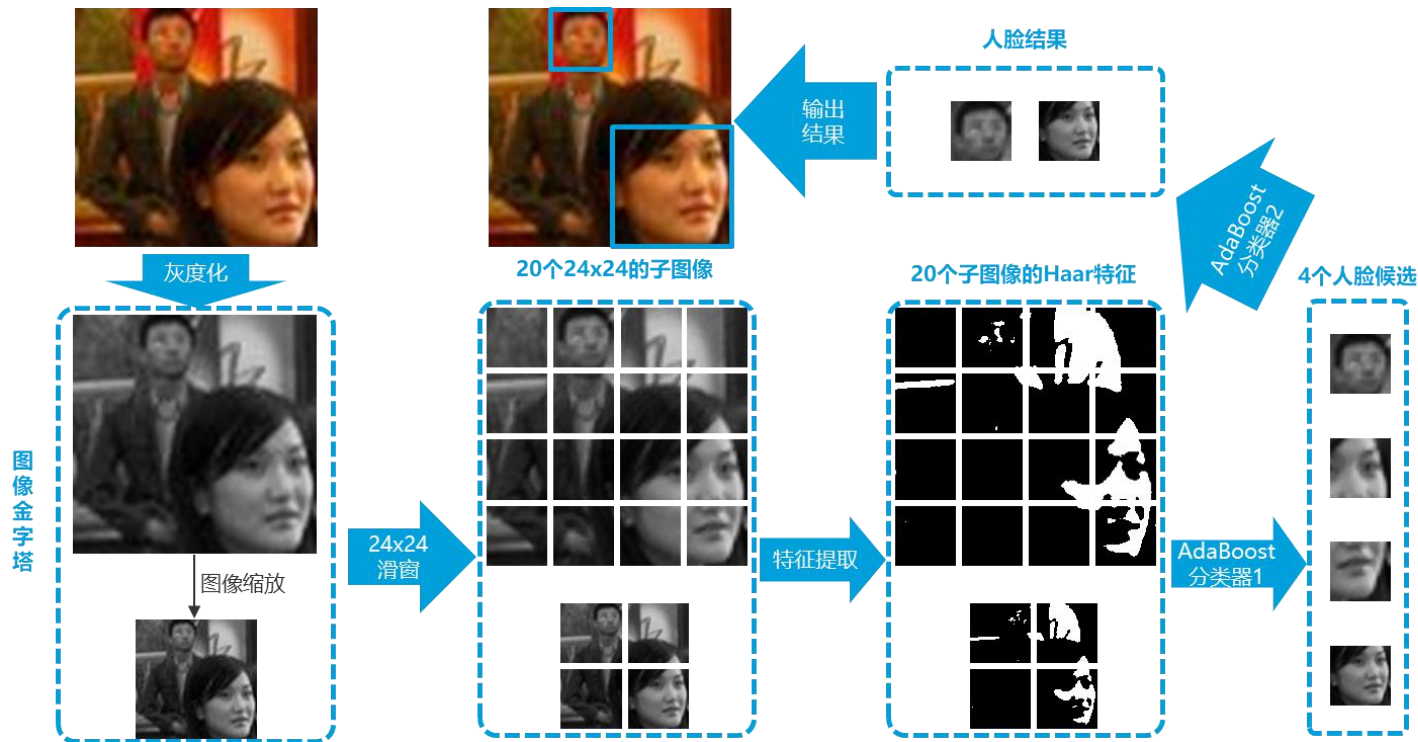
③ 每个位置样本数量

ATSS算法



## 内容回顾：人脸检测

- 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

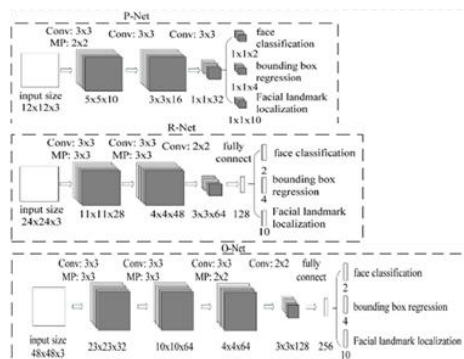
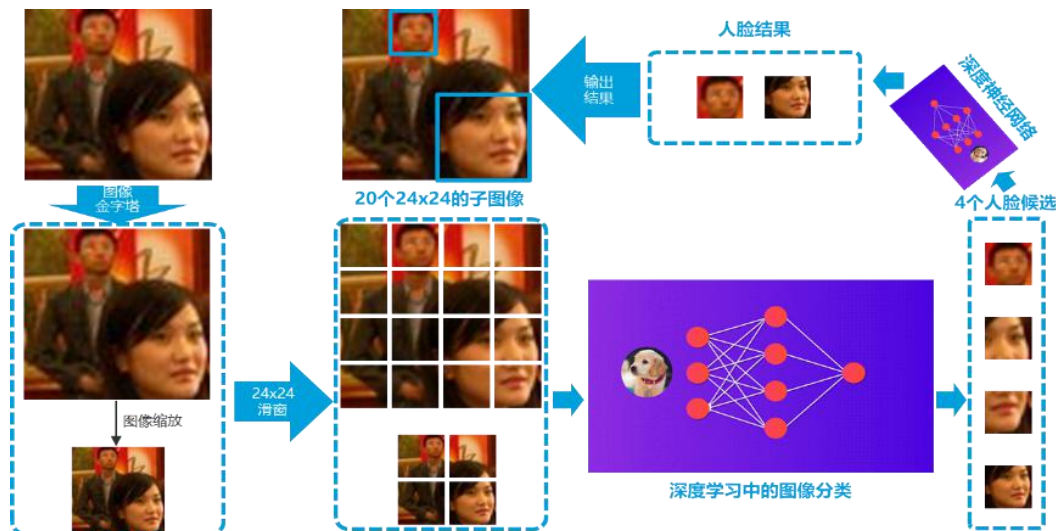




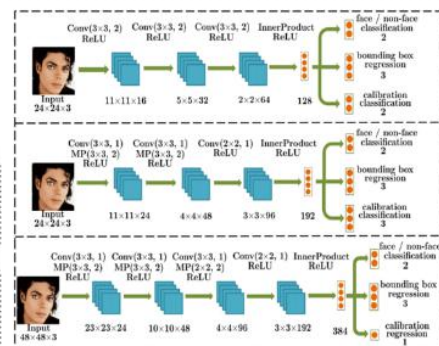


# 内容回顾：人脸检测

- 深度学习早期人脸检测算法在传统算法的流程中，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**

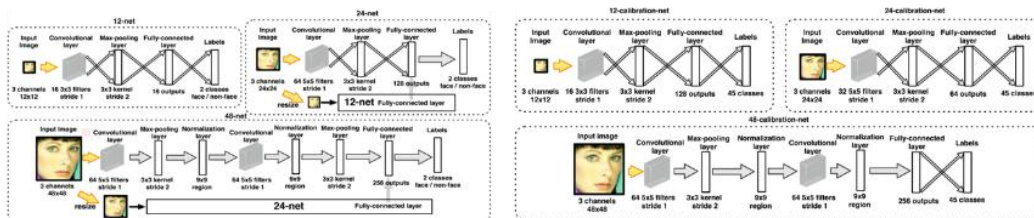


MTCNN



PCN

CascadeCNN







# 内容回顾：人脸检测

## 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

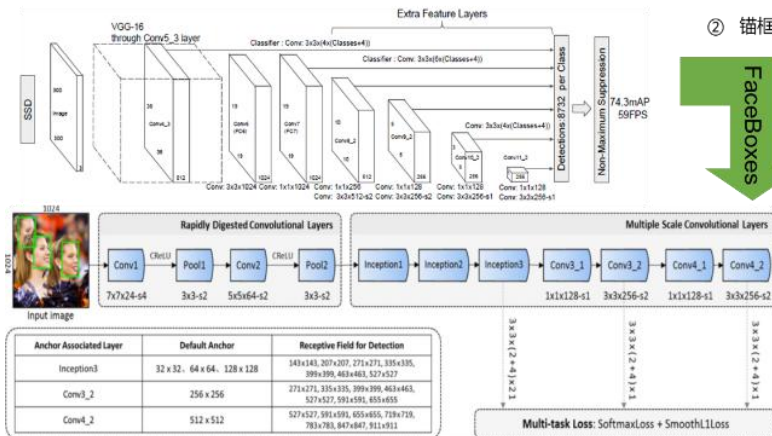
### 高效率的人脸检测算法

- 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

① 轻量级网络结构

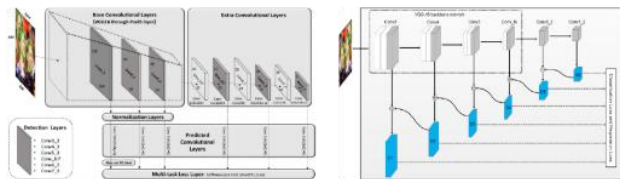
② 锚框密集化操作

FaceBoxes



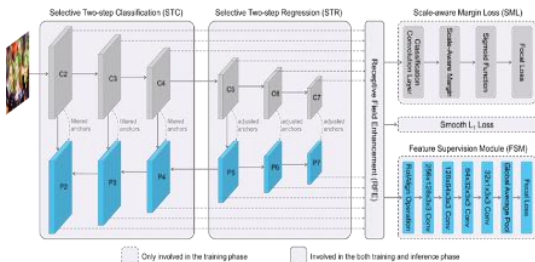
### 高精度的人脸检测算法

- 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



S<sup>3</sup>FD & SFDet

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



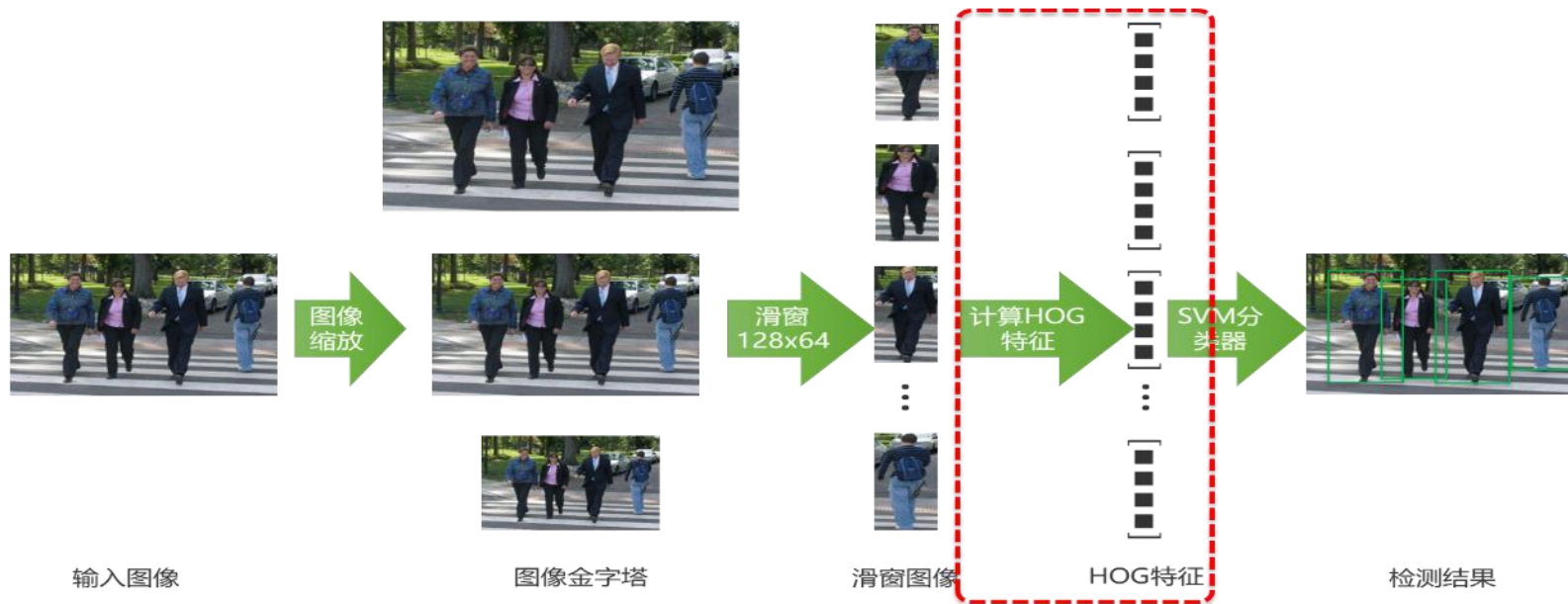
SRN & RefineFace

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



## 内容回顾：行人检测

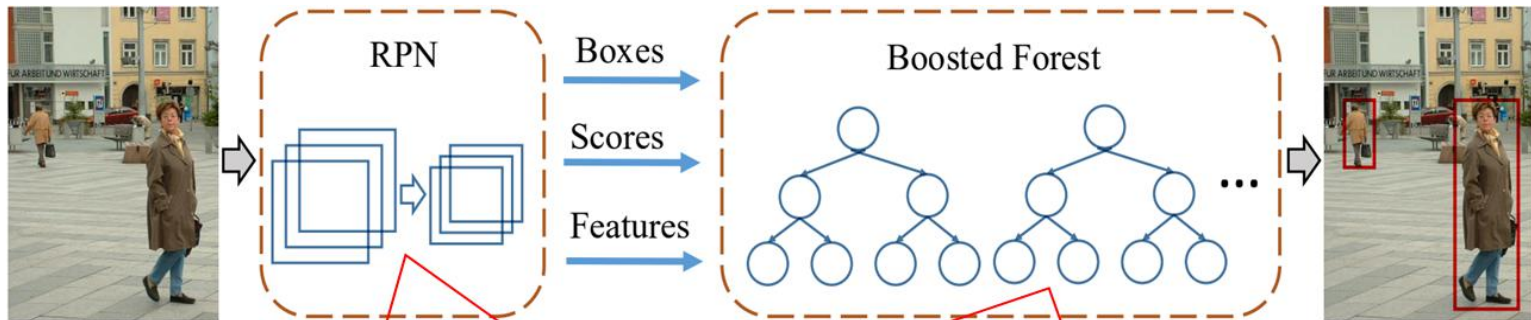
- 传统行人检测算法：利用手工特征+分类器，以滑窗方式在图像金字塔上遍历所有位置和大小，进行行人检测





## 内容回顾：行人检测

- 深度学习早期行人检测算法：传统检测算法与深度学习相结合



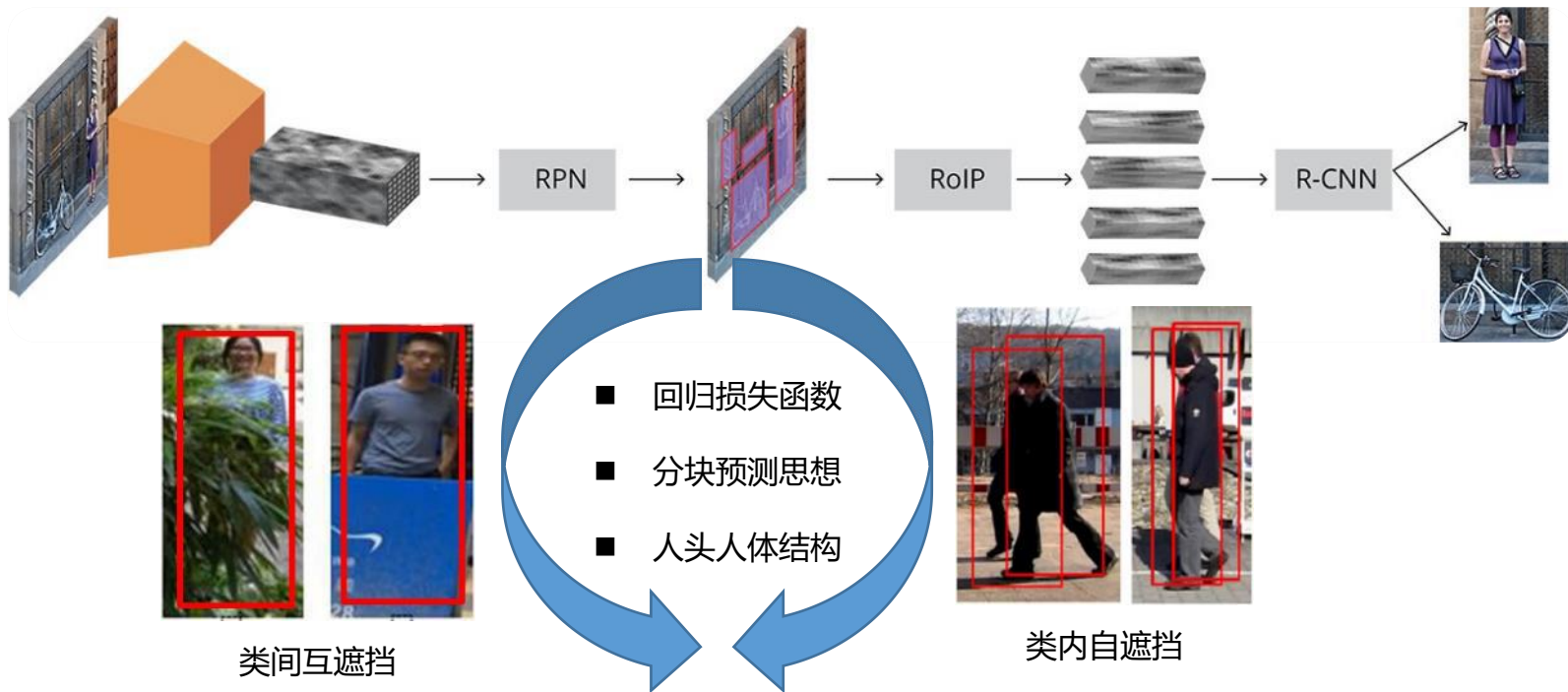
- **问题：**小尺度的人可利用的特征太少，不利于后续的分类和回归
- **方案：**从更浅的、分辨率更高的特征层上来进行特征扣取，从不同分辨率的特征层上扣取特征并进行融合，去掉下采样 + 空洞卷积，来增加特征层的分辨率

- **问题：**行人检测中，分类错误主要是难负样本的混淆，即把背景分为行人，而Faster R-CNN中第二阶段的Fast R-CNN不能很好的处理这些难负样本
- **方案：**采用级联的Boosted Forest (BF) 来替换Fast R-CNN，级联BF输入RoI Pooling后的特征，并挖掘困难负样本，CNN特征比手工特征更加高效，更具表现力



## 内容回顾：行人检测

- 深度学习后期行人检测算法：基于**通用物体检测算法Faster R-CNN**，针对**遮挡问题**进行**相应改进**



行人检测算法

Adapted FasterRCNN

Repulsion Loss

OR-CNN

JointDet



## 目录



物体检测发展总结



物体检测未来展望



## 目录



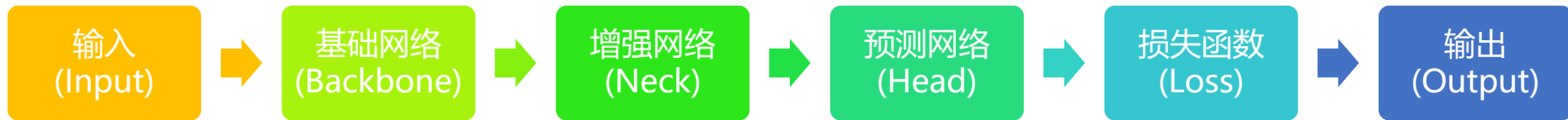
物体检测发展总结



物体检测未来展望

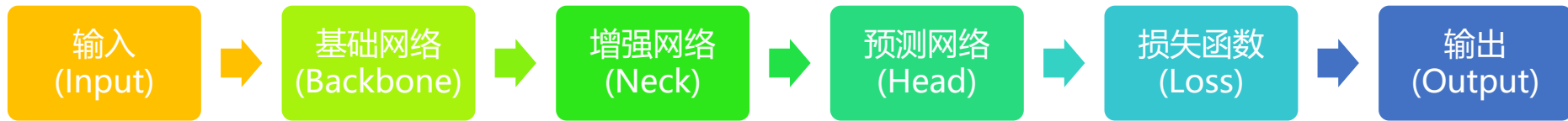


## 物体检测发展总结





# 物体检测发展总结



## 数据增广

- Random Erase
- CutOut
- MixUp
- CutMix
- AutoAug

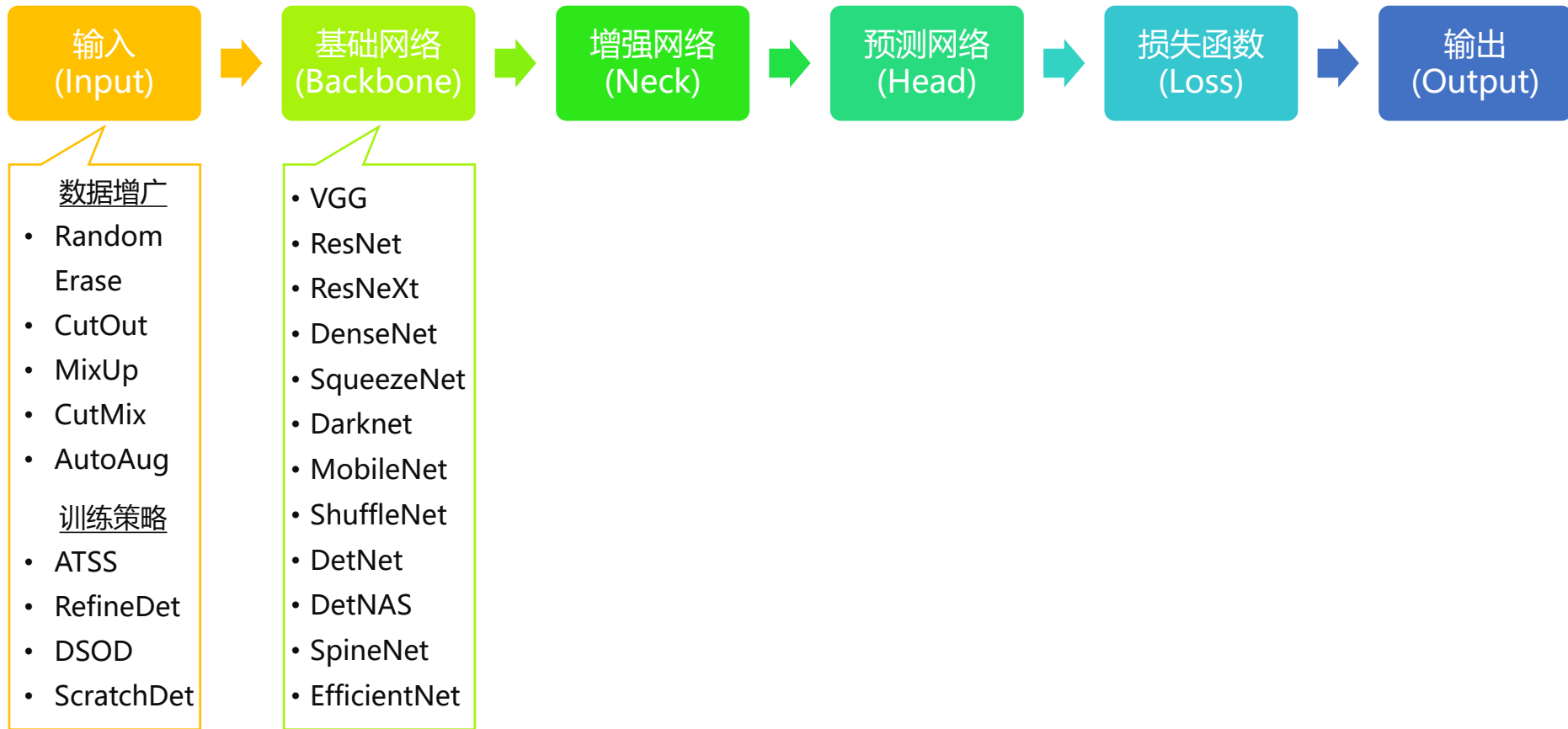
## 训练策略

- ATSS
- RefineDet
- DSOD
- ScratchDet



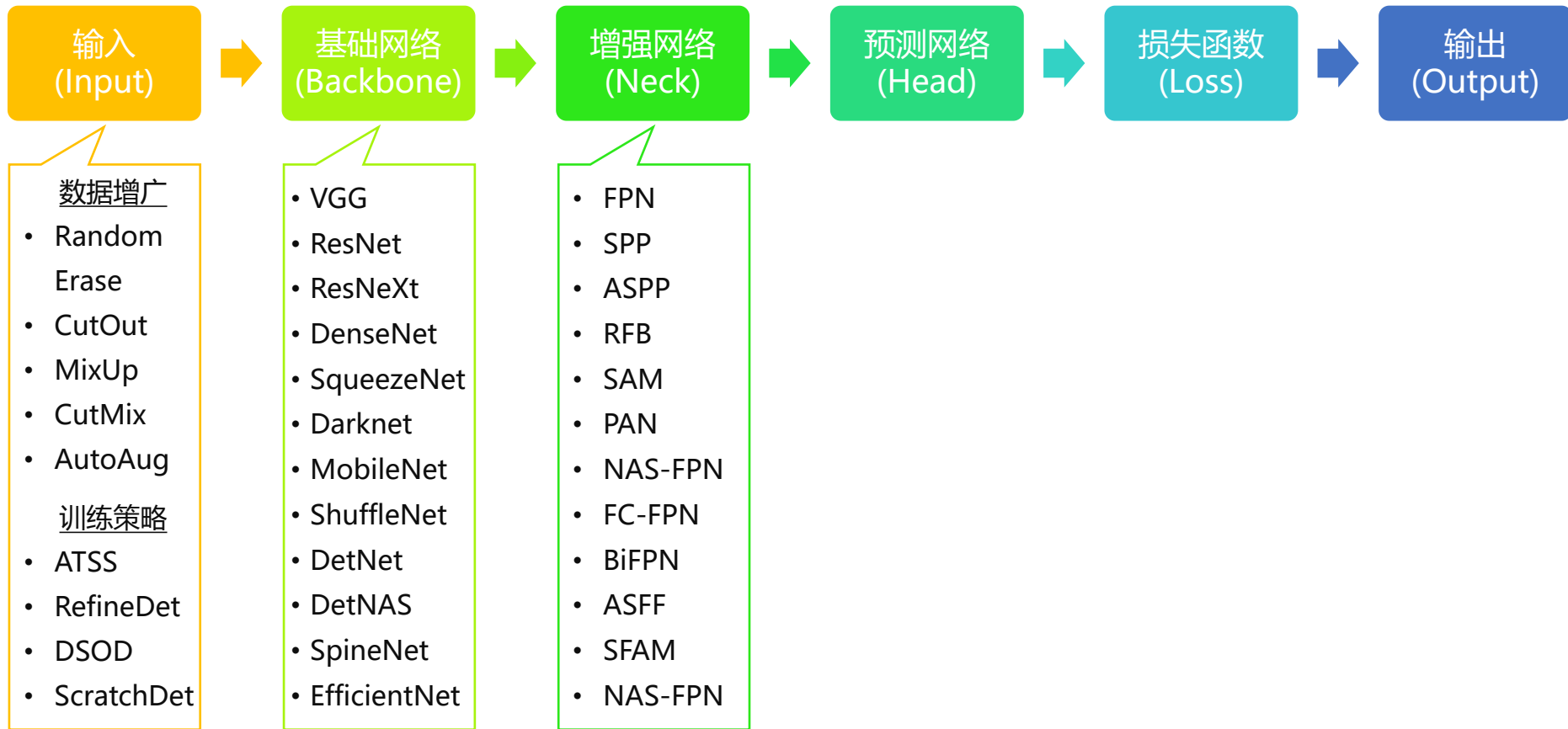


# 物体检测发展总结



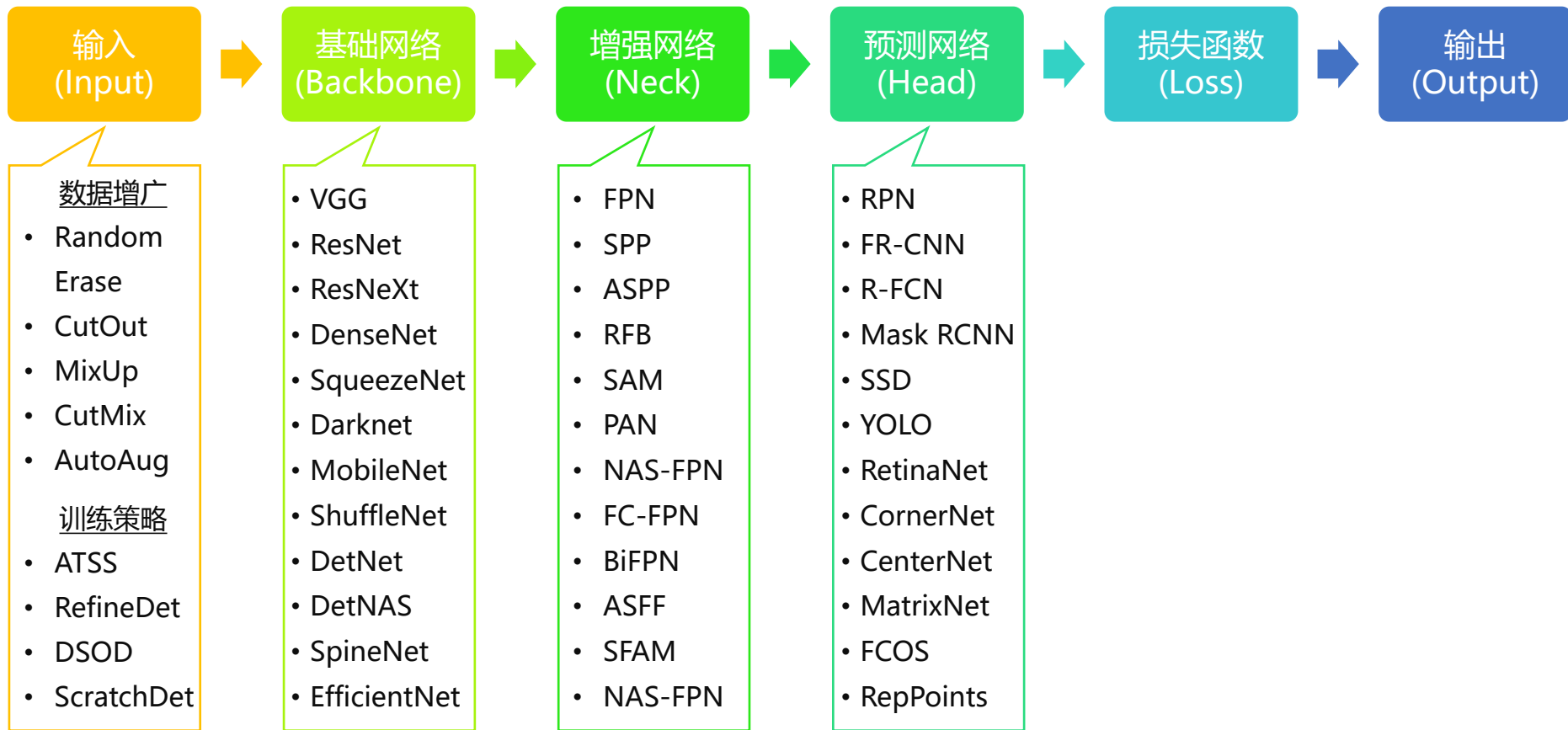


# 物体检测发展总结



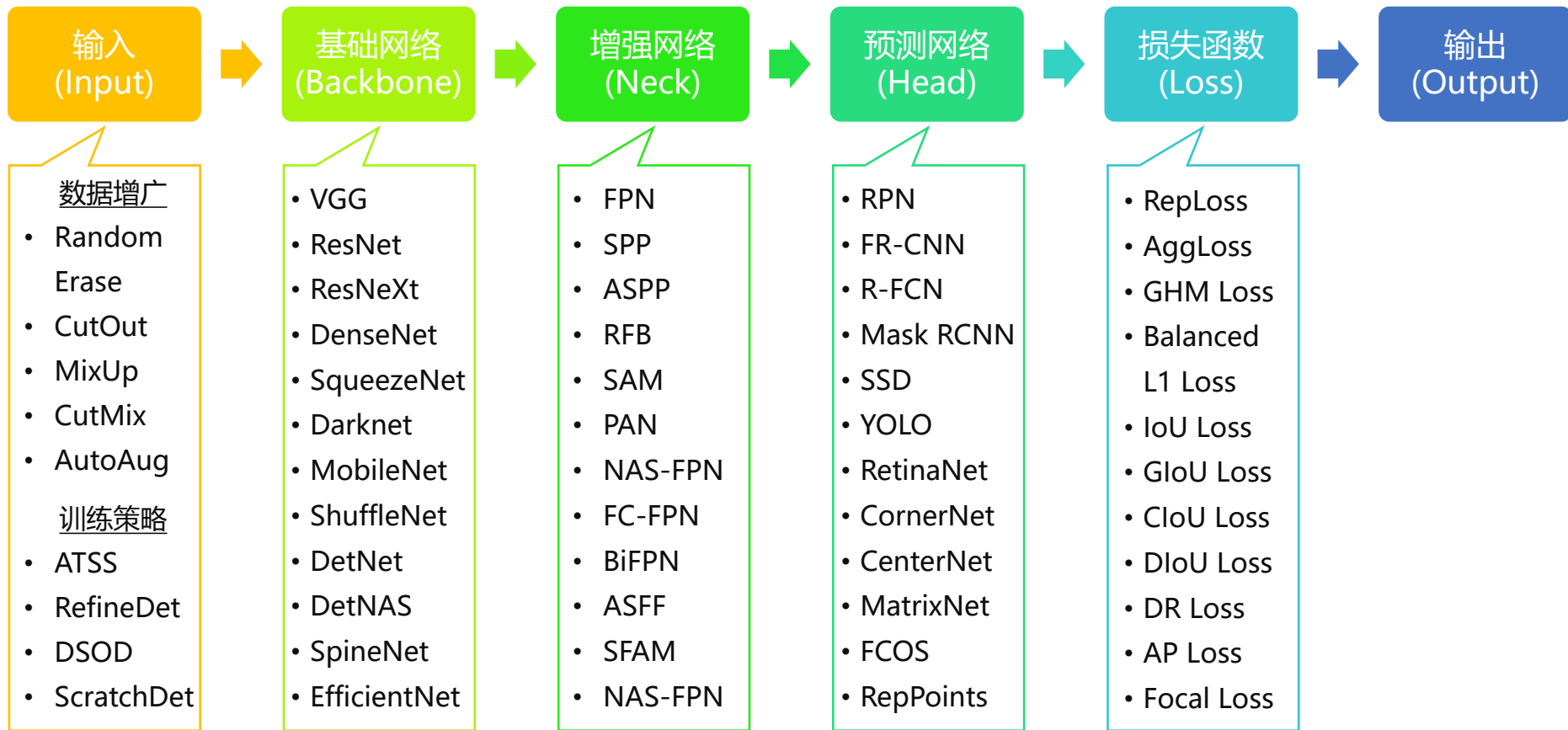


# 物体检测发展总结



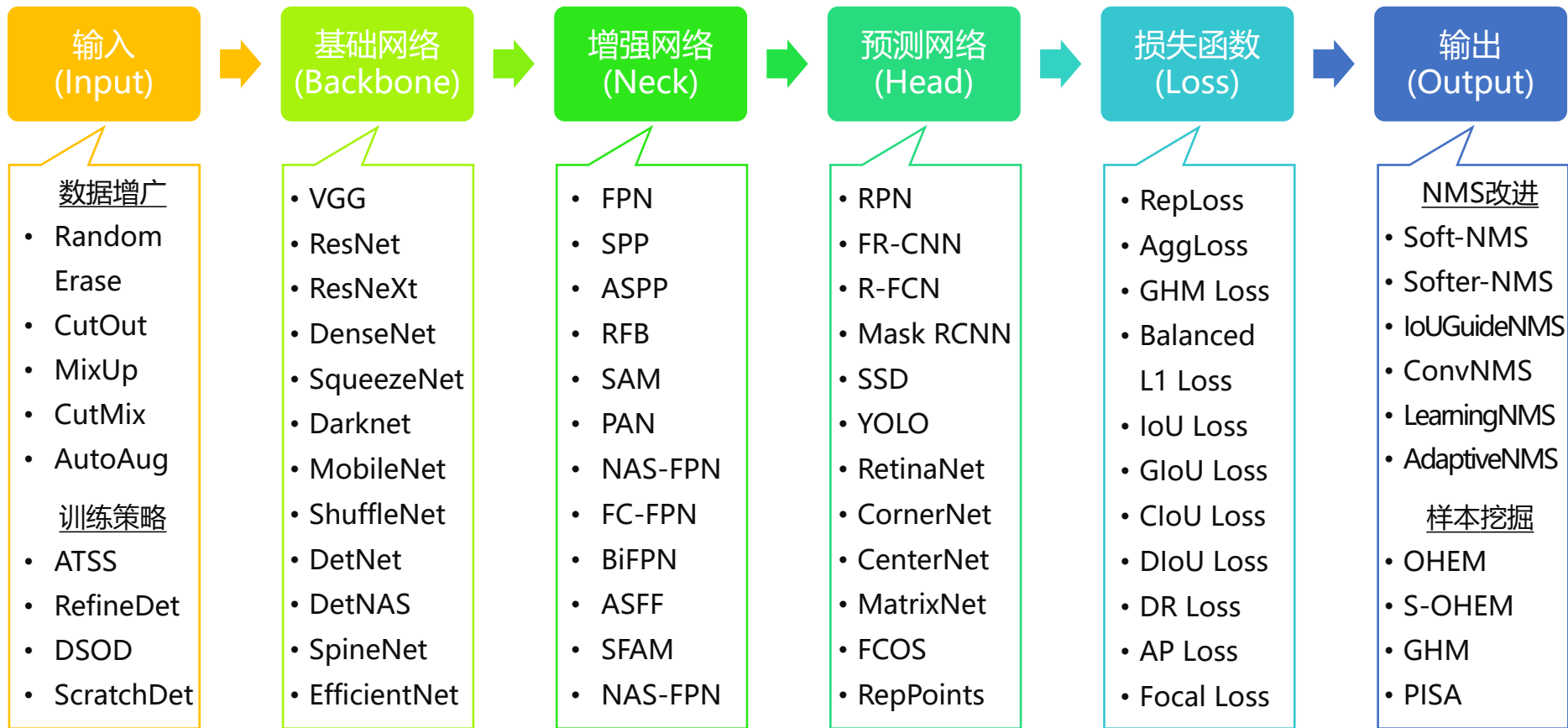


# 物体检测发展总结





# 物体检测发展总结





## 目录



物体检测发展总结



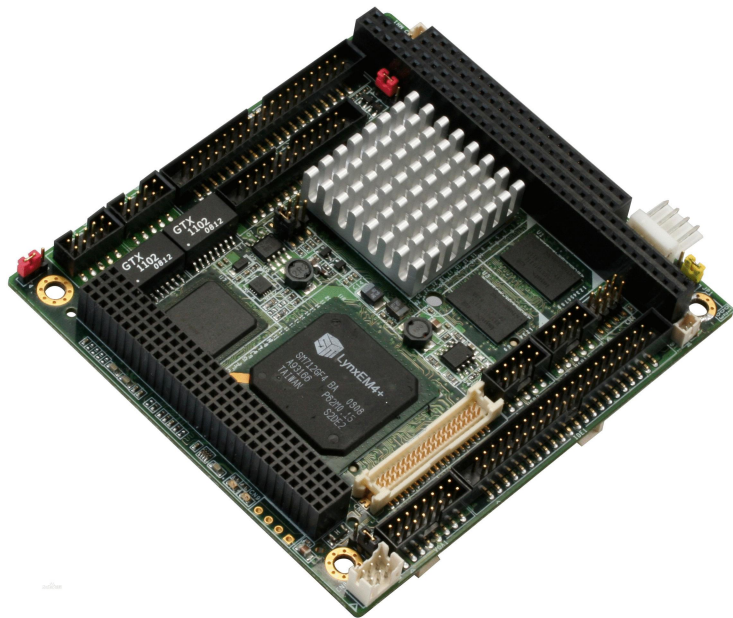
物体检测未来展望



# 物体检测未来展望

## ■ 更快的速度

- 满足嵌入式设备等的速度，利用知识蒸馏、剪枝、量化等技术加速





# 物体检测未来展望

## ■ 更快的速度

- 满足嵌入式设备等的速度，利用知识蒸馏、剪枝、量化等技术加速

## ■ 更高的精度

- 解决特定的错误，如小尺度、遮挡







## 物体检测未来展望

### ■ 更快的速度

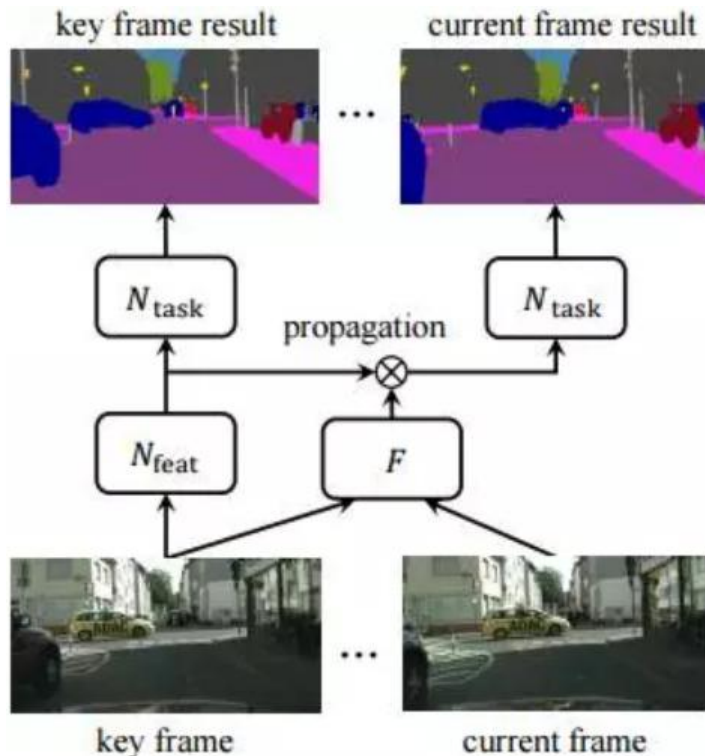
- 满足嵌入式设备等的速度，利用知识蒸馏、剪枝、量化等技术加速

### ■ 更高的精度

- 解决特定的错误，如小尺度、遮挡

### ■ 视频物体检测

- 利用帧间信息的冗余性来加速
- 利用帧间信息的连续性来提升精度





## 物体检测未来展望

### ■ 更快的速度

- 满足嵌入式设备等的速度，利用知识蒸馏、剪枝、量化等技术加速

### ■ 更高的精度

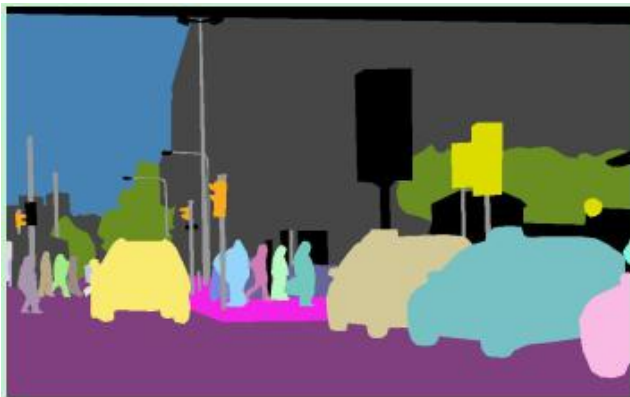
- 解决特定的错误，如小尺度、遮挡

### ■ 视频物体检测

- 利用帧间信息的冗余性来加速
- 利用帧间信息的连续性来提升精度

### ■ 物体检测+其他任务

- 检测 + 分割：实例分割、全景分割
- 检测 + 无监督/半监督/自监督/迁移学习等新问题





# 物体检测未来展望

## ■ 更快的速度

- 满足嵌入式设备等的速度，利用知识蒸馏、剪枝、量化等技术加速

## ■ 更高的精度

- 解决特定的错误，如小尺度、遮挡

## ■ 视频物体检测

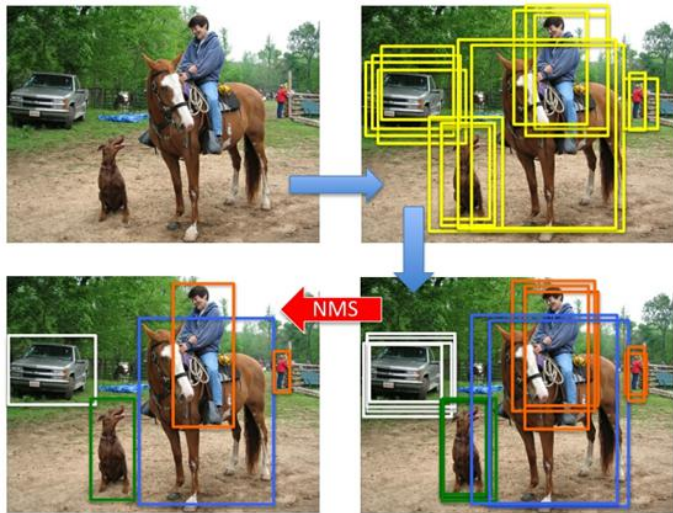
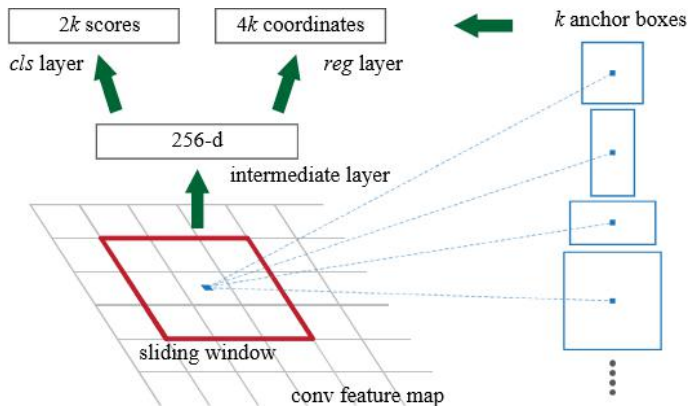
- 利用帧间信息的冗余性来加速
- 利用帧间信息的连续性来提升精度

## ■ 物体检测+其他任务

- 检测 + 分割：实例分割、全景分割
- 检测 + 无监督/半监督/自监督/迁移学习等新问题

## ■ 物体检测AutoML

- 算法层面：网络结构、锚框设计、NMS
- 数据层面：数据增广、样本挖掘、数据筛选等





结语

感谢各位聆听!

Thanks for Listening

