



北京交通大学

图像处理与机器学习

Digital Image Processing and Machine Learning

主讲人：黄琳琳

电子信息工程学院



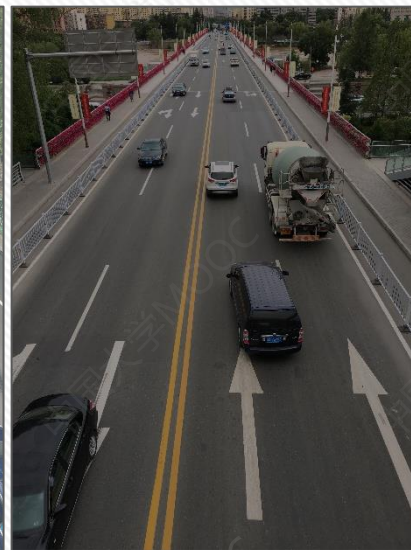
第九章 应用实例

基于深度学习的 交通车辆检测与识别

Vehicle detection and recognition based on
deep learning

研究背景及意义

- 人工智能、智能交通飞速发展
- 车辆检测和识别成为关键技术
- 深度学习方法逐渐成为主流

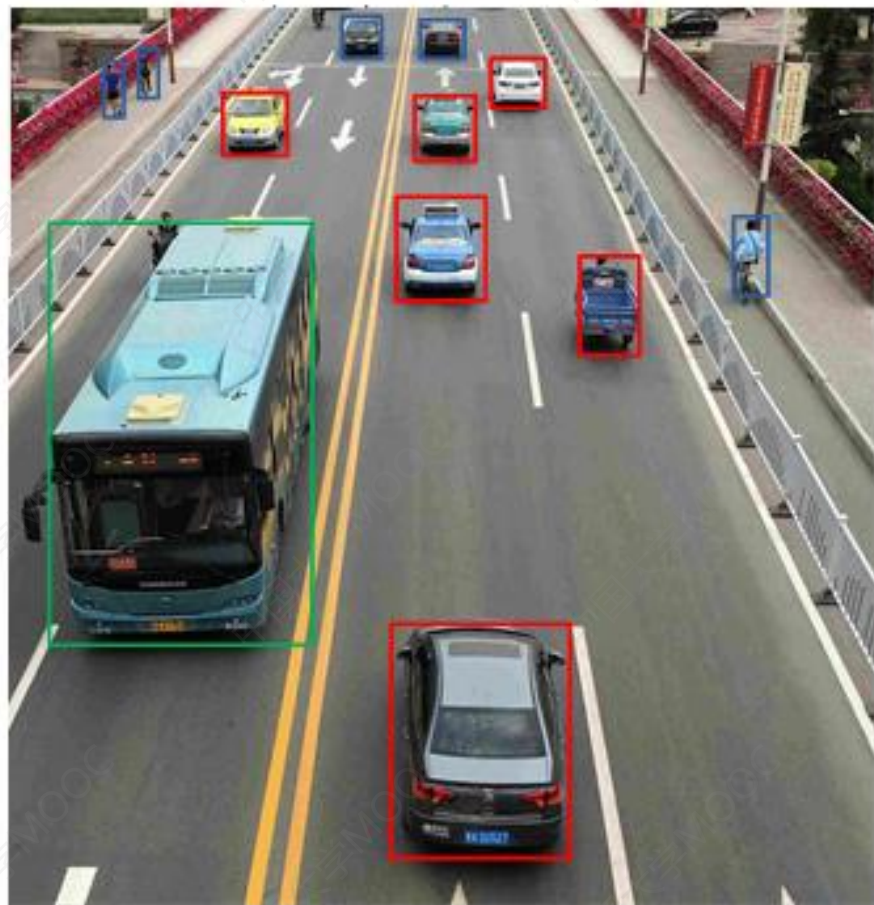


研究背景及意义

- 交通环境复杂多变
- 车辆尺度及属性多样性
- 检测速度难以满足实际需求

可分离反向连接网络

(Separable Reverse Connected network, SRC)



研究内容及方法

- ▽ 检测识别框架对比研究
- ▽ 基于SRC的车辆检测与识别算法
- ▽ 车辆检测识别优化算法

▼ 检测识别框架对比研究

▫ 两级网络

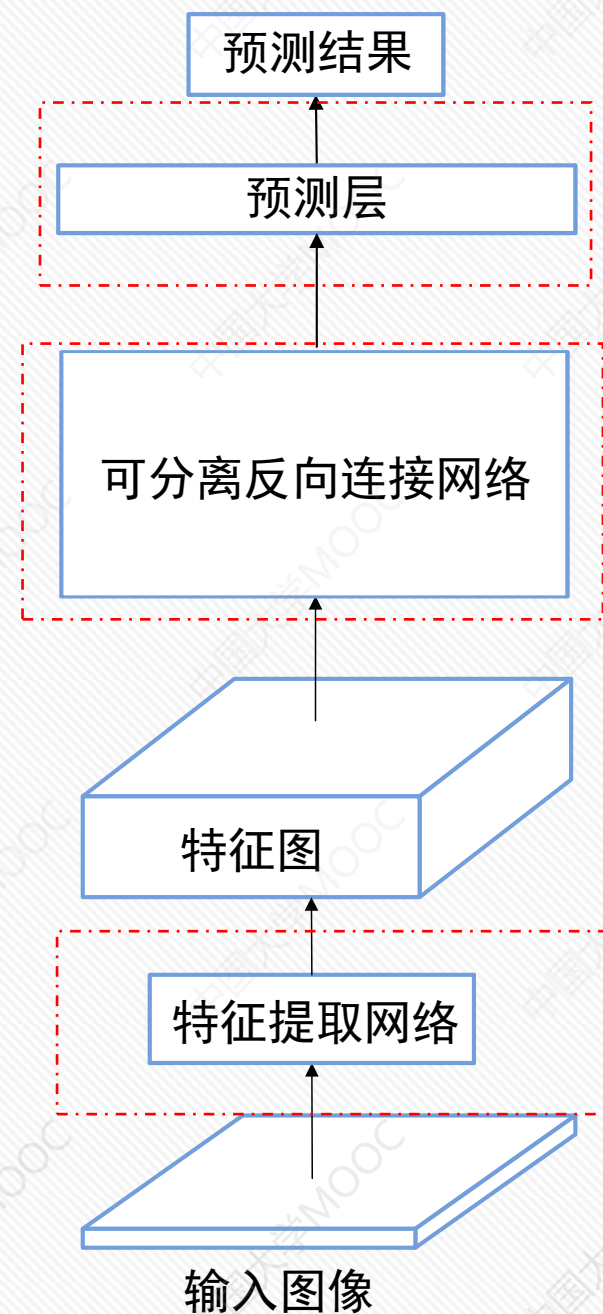
- 候选框生成网络（RPN）+ 检测识别网络
- 网络结构较复杂，检测精度高，检测速度慢

▫ 单级网络

- 同时预测物体选框和类别
- 网络结构较简单，精度稍低，检测速度快

▽ 基于SRC的车辆检测与识别算法

- 特征提取网络
- 可分离反向连接网络（SRC）
 - 可分离卷积：提升检测速度
 - 反向连接网络：提升多尺度检测能力
- 预测层



▫ 对比实验

两级网络

网络结构	特征提取
Faster R-CNN	ZF
Faster R-CNN	VGG
Faster R-CNN	ResNet101
R-FCN	VGG
R-FCN	ResNet50
R-FCN	ResNet101
FPN	VGG
FPN	ResNet101

单级网络

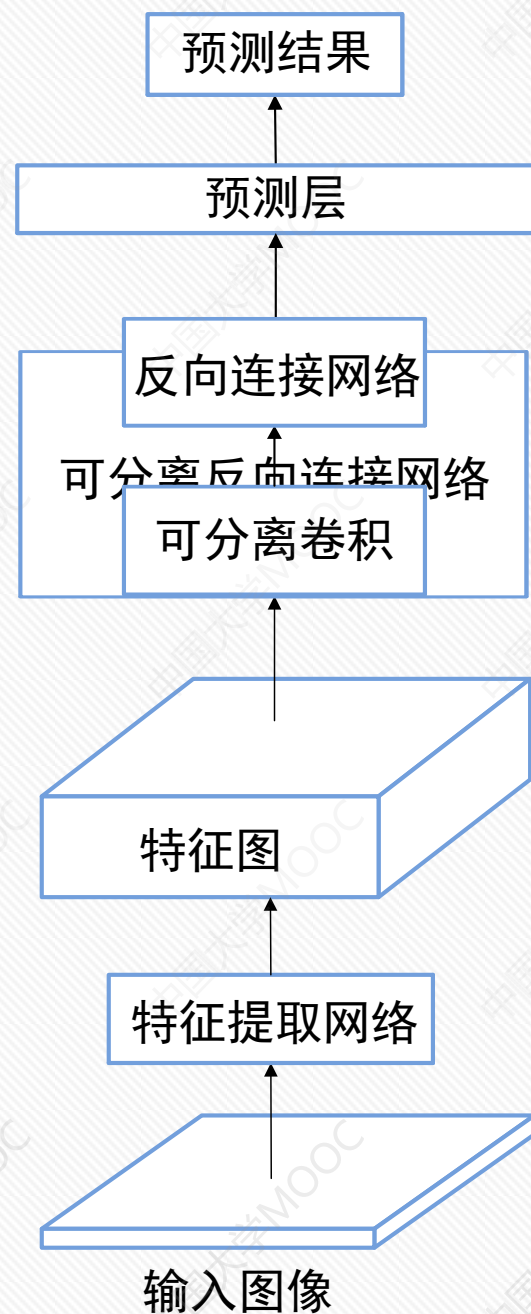
网络结构	特征提取
SSD 300	VGG
SSD 512	VGG
YOLO	GoogLeNet
YOLOv2	darknet

选取检测性能较好的网络

为高性能车辆检测识别奠定基础

▽ 基于SRC的车辆检测与识别算法

- 特征提取网络
- 可分离反向连接网络（SRC）
 - 可分离卷积：提升检测速度
 - 反向连接网络：提升多尺度检测能力
- 预测层



▪ 可分离卷积（Separable Convolution, SC）

◦ 控制卷积层的参数量

$$k \times k \rightarrow k \times 1, 1 \times k \quad k = 7$$

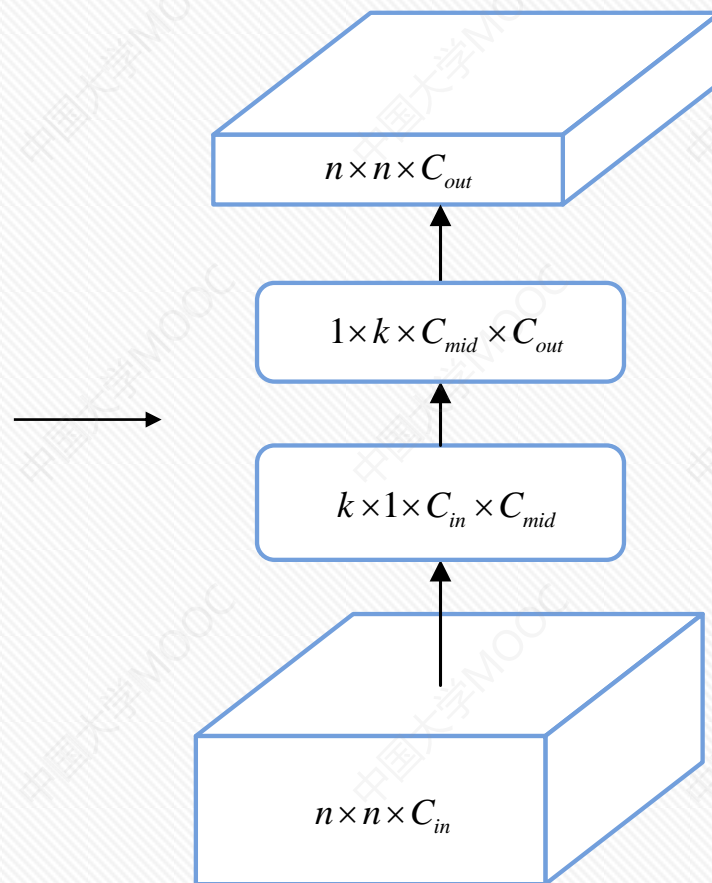
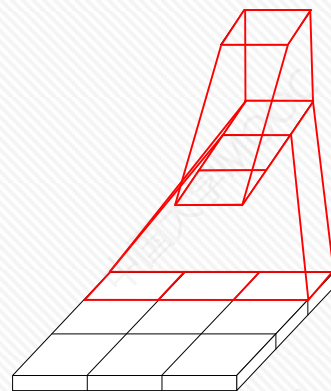
提取稀疏特征

◦ 控制特征图的通道数

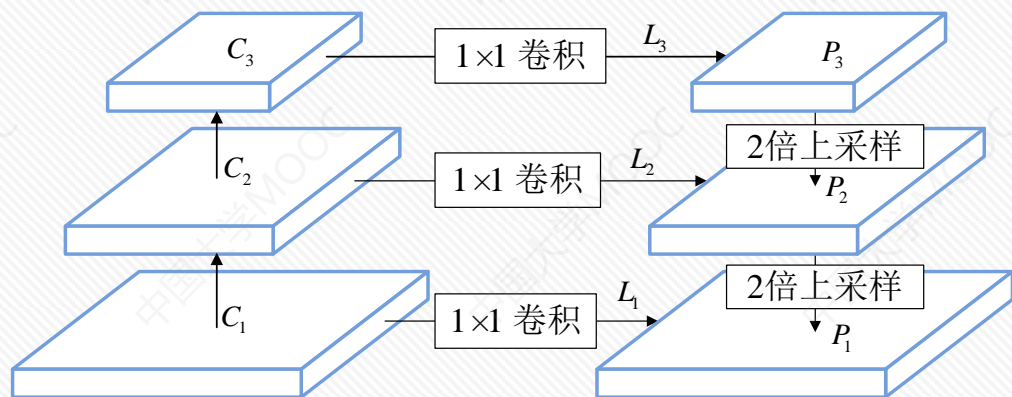
$$C_{mid} = 64, 128 \quad C_{out} \leq \frac{1}{2} C_{in}$$

减少过拟合

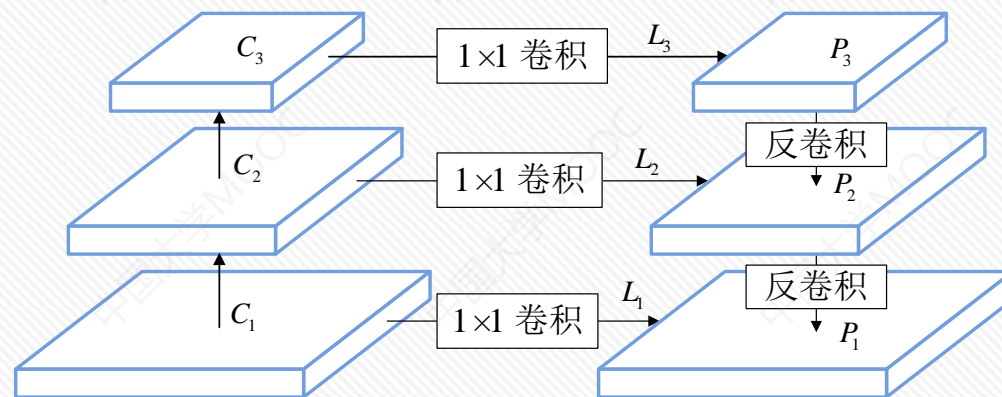
增加特征表达效率，提升检测速度



▪ 反向连接网络（Reverse Connected network, RC）

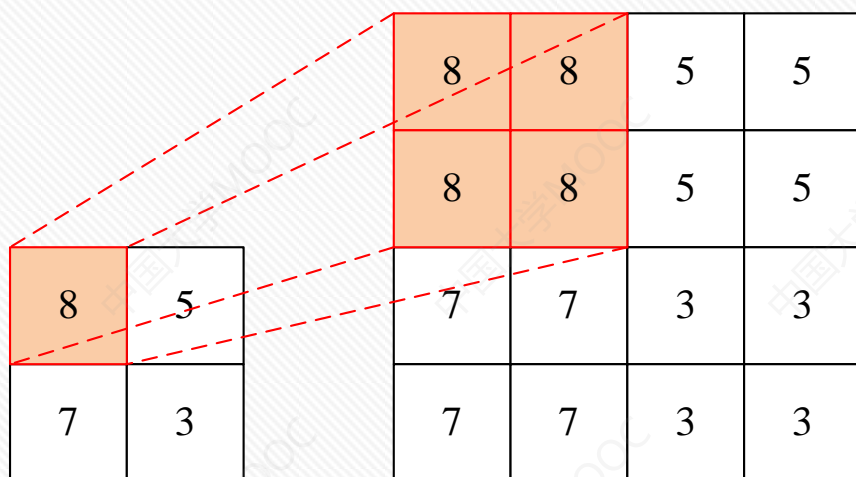


FPN中的反向连接网络



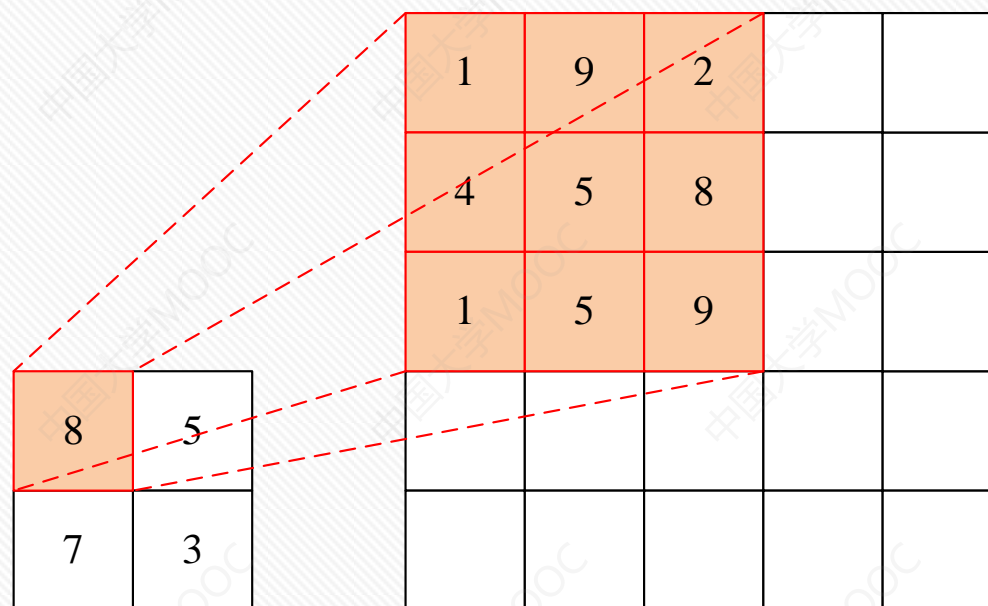
本文采用的反向连接网络RC

上采样 (Up-Sampling)



$stride = 2$

反卷积 (De-convolution)

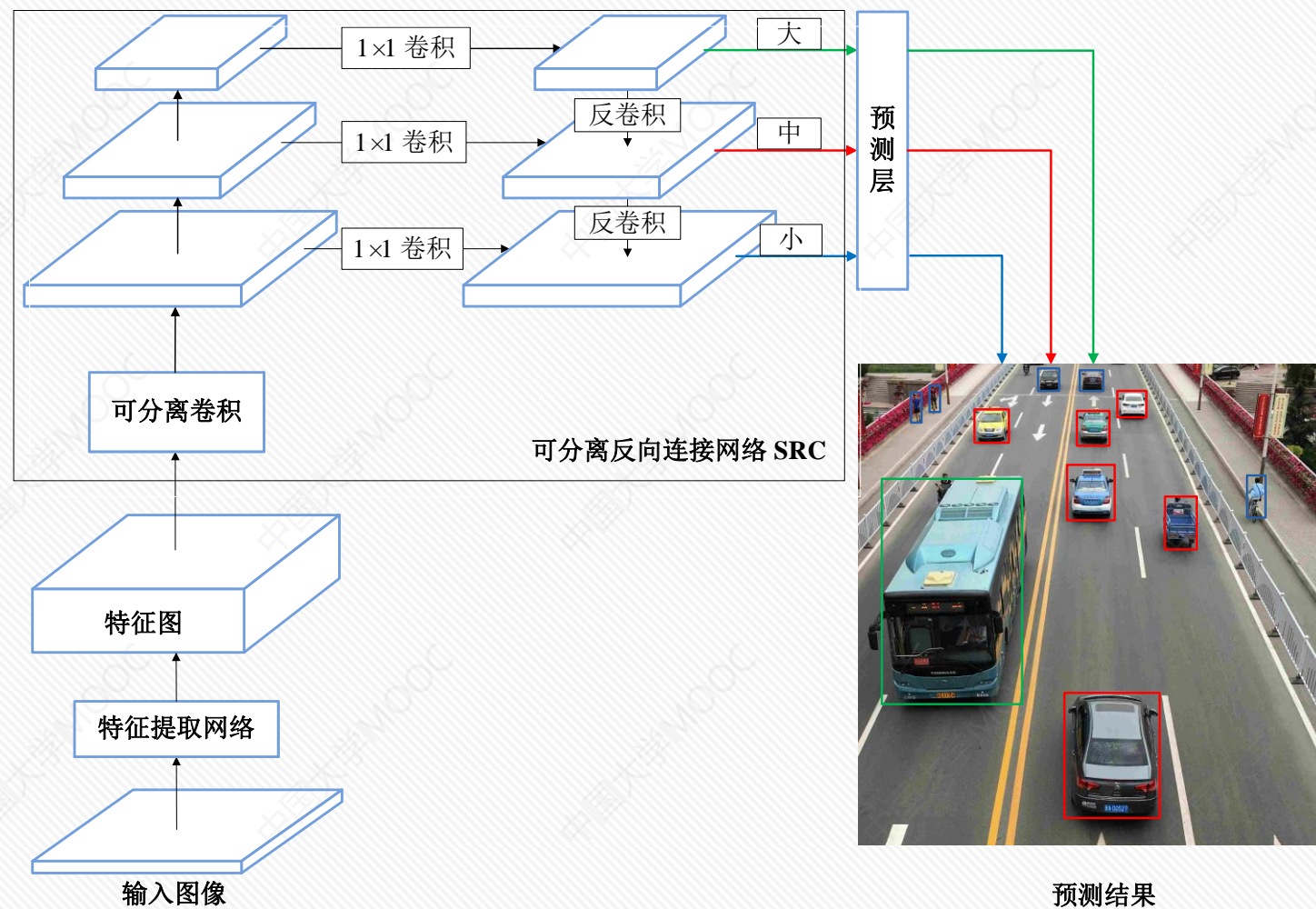


$k = 3, stride = 2$

提取更丰富的深层语义信息，提升多尺度特征表达能力

▣ 多尺度预测

◦ 分层预测：大、中、小



▽ 车辆检测识别优化算法

▫ 训练策略优化算法

- 多尺度训练 (Multi-scale Training)

使网络学习到样本的多尺度特征表示

- 在线难例挖掘 (Online Hard Example Mining, OHEM)

有效改善训练过程中正负样本不均问题

▣ 网络模型优化算法

◦ 剪枝（Pruning）

阈值操作

0.22	-0.38	0.33	0.09
0.13	-0.14	-0.48	0.12
-0.31	0.92	2.94	-0.13
0.49	0.14	0.23	3.17

→ 阈值为0.5

0	0	0	0
0	0	0	0
0	0.92	2.94	0
0	0	0	3.17

稀疏压缩

index	0	1	...	8	9	10	...
diff				8	1	1	
value				0	0.92	2.94	

多次“剪枝+微调”，不损失太多关键特征，减少网络参数规模

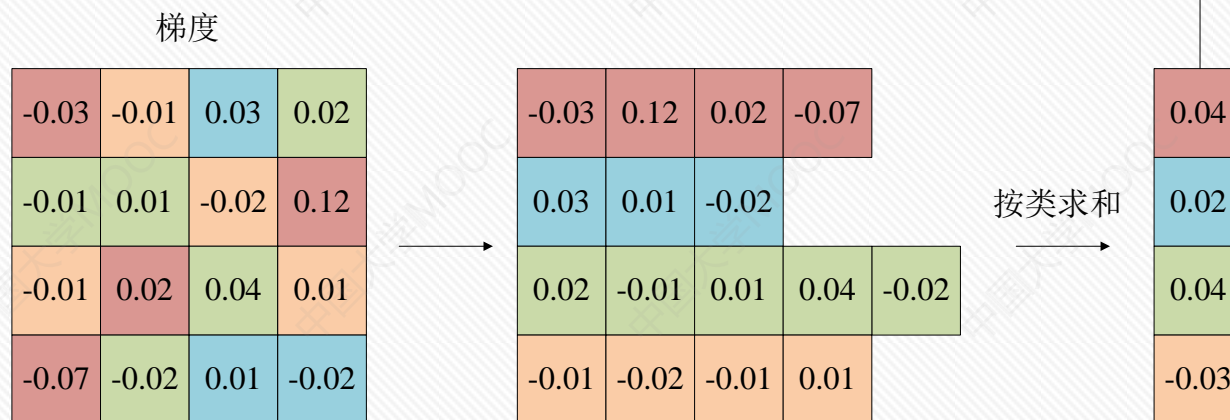
量化 (Quantization)



将32bit权重以类中心和2bit索引存储

各类梯度之和更新类中心

减少网络中的浮点数运算



实验结果及分析

▽ 数据集

需要兼具多尺度和多类别的车辆信息

数据集	车辆类别	训练集图像	测试集图像
Pascal VOC	3	16551	4952
MS COCO	4	118287	5000

▽ 评价指标

数据集	指标	含义
Pascal VOC	AP	所有 $\text{IoU} > 0.5$ 的选框中各类物体的平均精确度
	mAP	Pascal VOC 中 20 类待识别物体 AP 的平均值
	VmAP	Pascal VOC 中车辆类别的 mAP
	FPS	每秒处理图片数量
MS COCO	VmAP	MS COCO 中车辆类别的 mAP
	AP	IoU 处于 $[0.5:0.95]$ 之间所有物体的 mAP
	VAP_S	所有像素面积小于 32×32 车辆的 mAP
	VAP_M	所有像素面积大于 32×32 且小于 96×96 车辆的 mAP
	VAP_L	所有像素面积大于 96×96 车辆的 mAP

▽ 实验结果

▣ 检测识别框架对比研究

○ 两级网络 Pascal VOC

网络结构	特征提取	VmAP(%)	FPS
Faster R-CNN	ZF	69.1	18.0
Faster R-CNN	VGG	78.8	7.0
Faster R-CNN	ResNet101	80.6	5.0
R-FCN	ResNet50	80.1	4.9
R-FCN	ResNet101	83.3	3.3
FPN	VGG	81.7	4.4
FPN	ResNet101	83.8	4.0

○ 两级网络 MS COCO

网络结构	特征提取	VAR _S (%)	VmAP(%)
Faster R-CNN	ZF	5.7	25.3
Faster R-CNN	VGG	7.7	28.7
Faster R-CNN	ResNet101	7.8	31.4
R-FCN	ResNet50	7.6	29.1
R-FCN	ResNet101	7.2	33.8
FPN	VGG	10.9	32.3
FPN	ResNet101	11.1	34.6

FPN 小尺度车辆检测能力突出，检测精度较高

◦ 单级网络 Pascal VOC

网络结构	特征提取	摩托(%)	轿车(%)	公交(%)	mAP(%)	VmAP(%)	FPS
SSD300	VGG	80.6	80.8	81.1	78.6	79.3	46
SSD512	VGG	80.9	85.6	84.9	81.2	82.4	19
YOLO	GoogLeNet	71.3	55.9	68.3	63.0	64.6	45
YOLOv2	darknet	83.4	76.5	79.8	79.4	80.3	81

◦ 单级网络 MS COCO

网络结构	特征提取	$VAR_S(\%)$	$VAR_M(\%)$	$VAR_L(\%)$	AP(%)	VmAP(%)	FPS
SSD300	VGG	5.3	23.2	39.6	23.2	25.1	46
SSD512	VGG	8.0	28.9	40.6	23.7	29.8	19
YOLO	GoogLeNet	4.3	19.3	31.2	19.8	23.3	45
YOLOv2	darknet	8.7	22.4	35.5	23.4	28.2	81

YOLOv2 对小尺度车辆检测率高，检测速度有明显优势

▪ 基于SRC的车辆检测识别

◦ 两级网络

网络结构	特征提取	摩托(%)	轿车(%)	公交(%)	mAP(%)	VmAP(%)	FPS
Faster R-CNN	ResNet101	81.1	79.6	83.1	79.9	80.6	5.0
FPN	ResNet101	83.4	87.1	85.5	83.0	83.8	4.0
SC	ResNet101	79.1	77.4	80.9	77.1	78.9	12.2
RC	ResNet101	88.3	90.8	87.8	87.8	88.4	3.2
SRC	ResNet101	83.9	86.7	85.2	83.1	84.0	10.6

检测速度提升
证明可分离卷积有效性

网络结构	特征提取	$VAR_S(\%)$	$VAR_M(\%)$	$VAR_L(\%)$	AP(%)	VmAP(%)	FPS
Faster R-CNN	ResNet101	7.8	27.3	37.9	24.9	31.4	5.0
FPN	ResNet101	11.1	29.5	38.7	26.2	34.6	4.0
SC	ResNet101	7.6	27.0	36.8	24.1	30.3	12.2
RC	ResNet101	16.7	34.1	44.1	30.6	39.8	3.2
SRC	ResNet101	13.2	29.6	38.4	33.5	34.9	10.6

小尺寸车辆检测率提升
证明反向连接网络有效性

相比FPN有一定优势
证明SRC有效性

。单级网络

网络结构	特征提取	摩托(%)	轿车(%)	公交(%)	mAP(%)	VmAP(%)	FPS
YOLOv2	darknet	83.4	76.5	79.8	79.2	80.3	81
SC	darknet	79.0	75.5	77.6	76.2	78.4	88
RC	darknet	84.9	85.6	87.0	82.1	85.8	18
SRC	darknet	83.8	84.8	85.3	84.0	84.5	34

车辆检测识别精度提升

网络结构	特征提取	$VAR_S(\%)$	$VAR_M(\%)$	$VAR_L(\%)$	AP(%)	VmAP(%)	FPS
YOLOv2	darknet	8.7	22.4	35.5	23.4	28.2	81
SC	darknet	7.0	21.2	32.8	21.6	25.6	88
RC	darknet	14.3	34.0	43.9	31.7	36.6	18
SRC	darknet	11.3	28.6	40.9	30.6	31.7	34

小尺度车辆
检测识别精度提升

小结：SRC充分结合了SC、RC的优势，有效提升了多尺度车辆检测识别性能

车辆检测识别优化算法

两级网络

网络结构	特征提取	摩托(%)	轿车(%)	公交(%)	mAP(%)	VmAP(%)	FPS
FSRC	ResNet101	83.9	86.7	85.2	83.1	84.0	10.6
FSRC+	ResNet101	85.4	88.1	85.9	85.0	85.2	10.6
FSRC++	ResNet101	87.9	89.4	87.0	86.7	87.1	10.6
FSRC++:A	ResNet101	87.1	88.9	86.8	86.2	86.7	28
FSRC++:B	ResNet101	86.9	88.0	86.4	85.5	86.0	35

网络结构	特征提取	$VAR_S(\%)$	$VAR_M(\%)$	$VAR_L(\%)$	AP(%)	VmAP(%)	FPS
FSRC	ResNet101	13.2	29.6	38.4	33.5	34.9	10.6
FSRC+	ResNet101	13.5	30.1	39.6	36.1	36.8	10.6
FSRC++	ResNet101	15.1	32.4	42.3	38.0	38.5	10.6
FSRC++:A	ResNet101	14.9	31.8	42.0	37.9	38.2	28
FSRC++:B	ResNet101	14.1	30.2	41.4	36.7	37.7	35

FSRC : 两级网络SRC
FSRC+ : 多尺度训练
FSRC++ : 多尺度训练 + OHEM
A、B : 网络模型优化程度

训练策略优化
进一步提升检测精度

网络模型优化
有效提升检测速度

◦ 单级网络

网络结构	特征提取	摩托(%)	轿车(%)	公交(%)	mAP(%)	VmAP(%)	FPS
YSRC	darknet	83.8	84.8	85.3	84.0	84.5	34
YSRC+	darknet	85.1	84.7	87.6	85.5	86.0	34
YSRC++:A	darknet	80.2	82.1	84.4	80.9	82.2	68
YSRC++:B	darknet	77.8	79.4	82.4	78.6	79.9	95

网络结构	特征提取	VAR _S (%)	VAR _M (%)	VAR _L (%)	AP(%)	VmAP(%)	FPS
YSRC	darknet	11.3	28.6	40.9	30.6	31.7	34
YSRC+	darknet	13.8	33.7	41.5	31.8	34.8	34
YSRC++:A	darknet	10.6	26.4	39.1	29.0	29.4	68
YSRC++:B	darknet	8.1	21.3	36.9	25.9	26.7	95

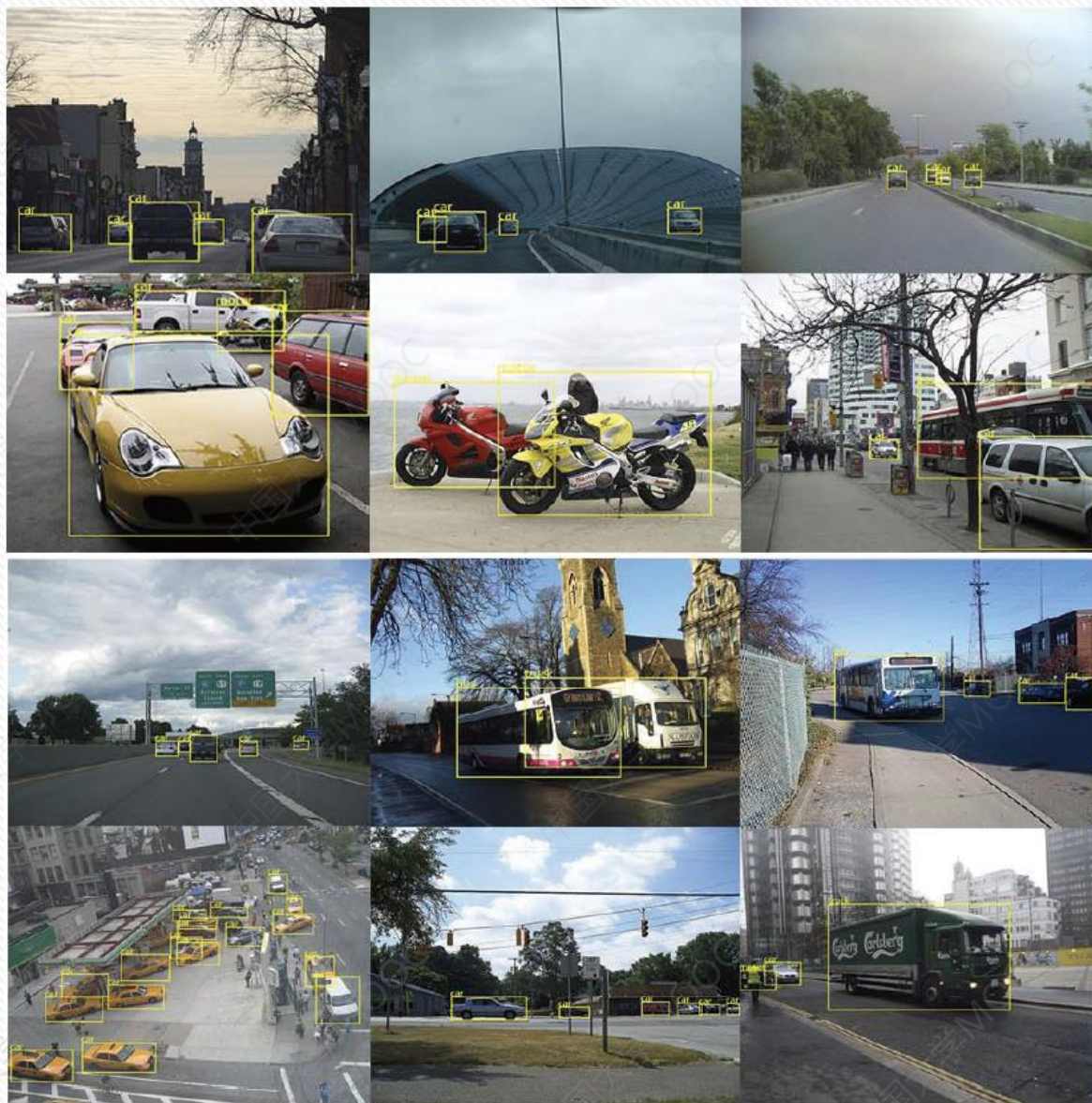
YSRC : 单级网络SRC
 YSRC+ : 多尺度训练
 A、B : 网络模型优化程度

多尺度训练
 提升检测识别精度

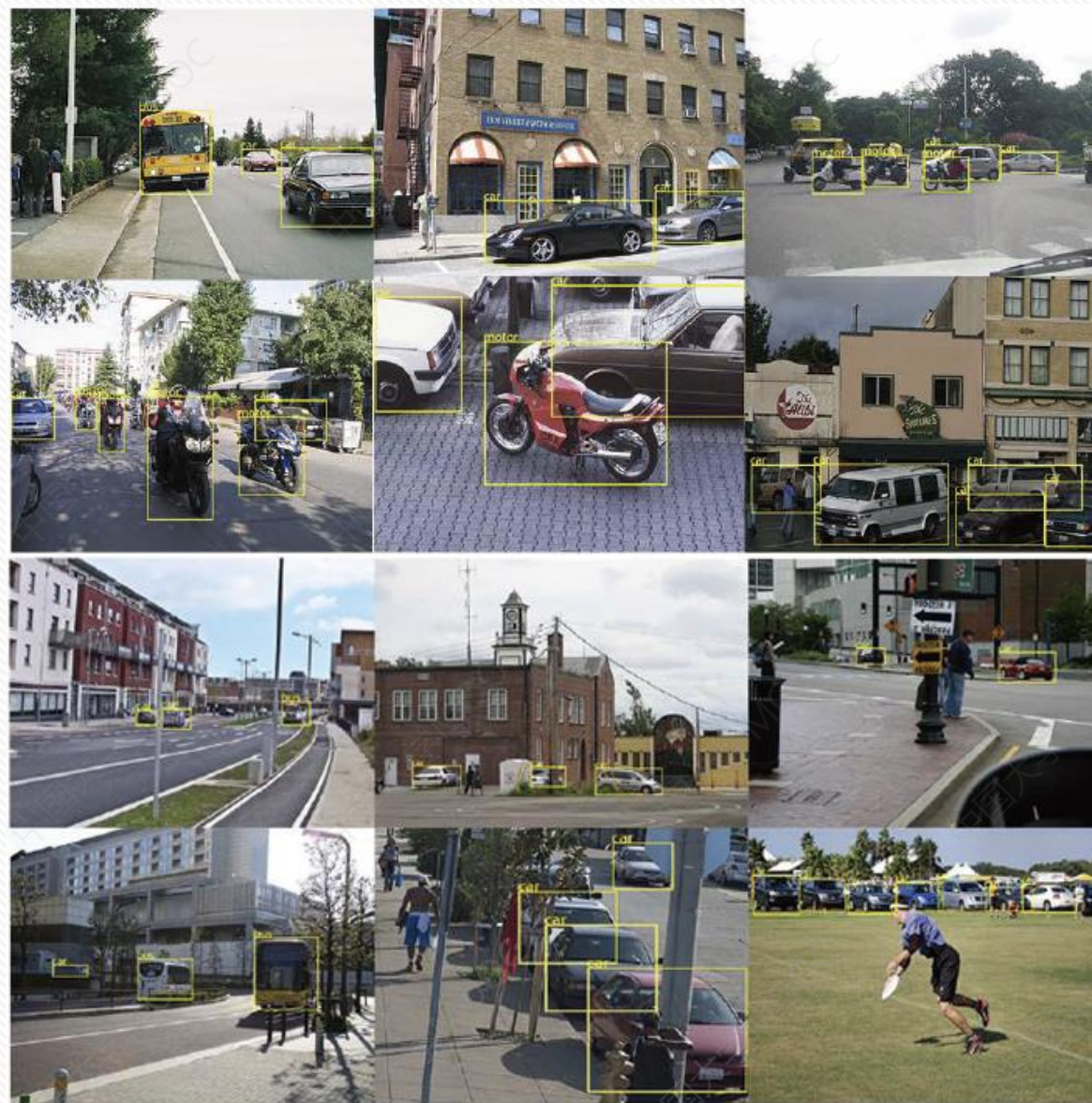
网络模型优化程度增加
 检测精度随之下降

小结: 训练策略优化提升车辆检测识别精度,
 网络模型优化应用在两级网络中效果显著, 保持精度提升速度

两级网络SRC检测结果



单级网络SRC检测结果





谢 谢

本课程所引用的一些素材为主讲老师多年的教学积累，来源于多种媒体及同事和同行的交流，难以一一注明出处，特此说明并表示感谢！