

# 基于 DeR-FCN 模型的车辆检测算法

王 玲, 李厚博, 王 鹏<sup>+</sup>, 孙爽滋

(长春理工大学 计算机科学技术学院, 吉林 长春 130022)

**摘 要:** 针对复杂城市环境下天气、光照、目标尺度以及车辆之间的遮挡等因素影响带来的车辆检测精度较差问题, 提出一种改进区域全卷积网络的车辆检测算法 (DeR-FCN)。通过特征级联的方式, 跨层连接融合车辆底层细节特征和高层语义特征; 使用维度分解区域提议网络, 获得更加精准的区域候选框; 在预测阶段采用软化非极大值抑制的方法, 输出更加准确的检测结果。为验证算法的有效性, 在 KITTI 和 PASCAL VOC 数据集, 使用 DeR-FCN 算法和常用的车辆检测算法进行对比实验, 实验结果表明, DeR-FCN 算法的检测精度高于其它方法。

**关键词:** 深度学习; 车辆检测; 区域全卷积网络; 维度分解区域提议网络; 软化非极大值抑制

**中图分类号:** TP391.4 **文献标识码:** A **文章编号:** 1000-7024 (2020) 10-2927-07

**doi:** 10.16208/j.issn1000-7024.2020.10.038

## Vehicle detection algorithm using DeR-FCN model

WANG Ling, LI Hou-bo, WANG Peng<sup>+</sup>, SUN Shuang-zi

(School of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China)

**Abstract:** There are some affecting factors such as weather, illumination, target scale and occlusion among vehicles in complex urban environments, they affect detection accuracy. To resolve the problem, a vehicle detection algorithm basing on regional full convolutional network was proposed, named as DeR-FCN. The underlying layer features and high-level semantic features of the hybrid vehicle were connected by means of feature cascading. The dimensional decomposition region proposal network was used to get a more accurate area candidate box. In the prediction stage, softer non-maximum suppression was used to output more accurate detection results. KITTI and PASCAL VOC data sets were used to verify the validity of the proposed algorithm. Experimental results show that DeR-FCN algorithm has higher detection accuracy than other methods.

**Key words:** deep learning; vehicle detection; R-FCN; DeRPN; soft-NMS

## 0 引 言

传统的车辆检测方法通常将特征提取算法与分类器结合进行检测<sup>[1-4]</sup>, 这类方法严重依赖人工特征提取, 针对复杂多变的环境模型泛化能力较差。随着深度学习的发展, 人们发现卷积神经网络可以自主地提取图像特征, 对于复杂场景的适应性更强。Ross Girshick 提出目标检测算法 R-CNN<sup>[5]</sup>, 将 CNN 应用于目标检测, 使得检测精度得到大幅度提升。但通过 Selective Search 提取的每一个候选框依次进入 CNN 网络, 存在冗余特征提取、冗余存储和计算复杂度大的问题。针对上述问题, Fast R-CNN<sup>[6]</sup>使用卷积特征

共享、多任务损失函数的方式, 显著提高了检测的速度与精度; Faster R-CNN<sup>[7]</sup>用 RPN 网络取代 Selective Search, 并整合区域建议与卷积神经网络, 不仅速度得到大大提高而且还获得了更加精确的结果, 真正实现了端到端的目标检测框架; R-FCN<sup>[8]</sup>算法引入全卷积操作, 将位置敏感得分图加入 FCN, 充分利用分类任务的平移不变性及检测任务的平移可变性, 解决网络中全连接层计算复杂度较高的问题, 大大提升了检测的效率。上述检测框架主要由区域建议和检测网络两部分组成, 为典型的基于区域建议的目标检测算法。还有研究者提出将分类和检测任务统一的基于回归的目标检测算法如 YOLO<sup>[9]</sup>、SSD<sup>[10,11]</sup>、

收稿日期: 2019-07-10; 修订日期: 2020-07-17

基金项目: 吉林省科技发展计划技术攻关基金项目 (20190302118GX)

作者简介: 王玲 (1979-), 女, 吉林通化人, 博士, 讲师, CCF 会员, 研究方向为深度学习、图像处理; 李厚博 (1995-), 男, 山东德州人, 硕士研究生, 研究方向为深度学习、图像处理; 十通讯作者: 王鹏 (1974-), 男, 内蒙古包头人, 博士, 教授, CCF 会员, 研究方向为数据库与数据挖掘; 孙爽滋 (1972-), 女, 吉林长春人, 硕士, 副教授, 研究方向为数字图像处理、模式识别。

©1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

E-mail: 635461179@qq.com

YOLOv2<sup>[12]</sup>等。基于回归的目标检测算法与基于区域建议的目标检测算法相比,其在通用目标检测数据集上的检测速度得到了大幅度提高,但也造成了检测精度的下降。

综上所述,R-FCN的全卷积操作不仅适用于不同尺寸图像的输入,显著减少了算法计算复杂度,而且在检测精度和效率上较为平衡,但仍无法满足实际生活中对车辆目标检测高准确率的要求。本文对R-FCN的车辆检测算法进行改进,通过特征级联的方式有效结合底层特征的位置及边缘信息和高层特征的语义信息;引入维度分解区域提议(DeRPN)<sup>[13]</sup>网络代替RPN网络,使用灵活的锚链机制获得更加精准的候选框;预测阶段使用软化非极大值抑制算法(soft-NMS)解决重叠目标漏检问题。这样,提出的DeR-FCN可以有效优化实际场景中车辆的漏检及误检,显著提高模型检测精度。

## 1 相关研究

### 1.1 R-FCN 检测框架

R-FCN模型首先通过RPN网络在特征图上提取候选区域,然后使用RoI子网络对候选区域进行分类识别,从而实现目标检测。R-FCN总体架构如图1所示。

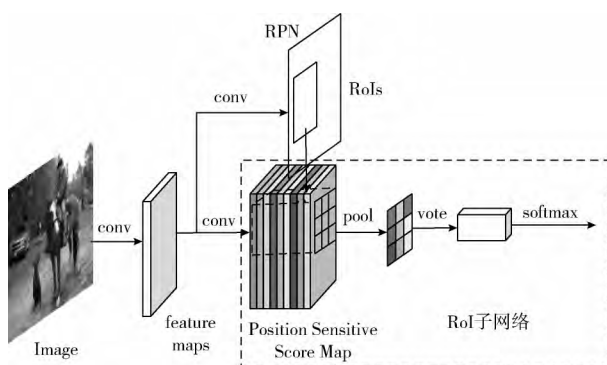


图1 R-FCN 模型架构

#### 1.1.1 RPN 网络

RPN网络的主要功能是提取候选区域。RPN使用默认大小为 $128 \times 128$ 、 $256 \times 256$ 、 $512 \times 512$ ,长宽比为 $1:1$ 、 $1:2$ 、 $2:1$ 的9种固定尺寸的锚框,在特征图上通过滑动窗口产生搜索框。对每个搜索框进行分类与回归操作,然后通过非极大值抑制法将高度重叠的区域剔除掉,解决候选区域过多带来的冗余问题,最后将剩余的候选区域用于RoI子网络训练。

#### 1.1.2 RoI 子网络

将每个RoI分割为 $k^2$ 个部分,每个部分映射到一张Score Map上,最终得到 $k^2$ 个Score Map,分类的种类有 $C$ 个类,背景为1个类,所以每一个Map通道数为 $C+1$ ,最终得到通道数为 $k^2(C+1)$ 的输出层。然后通过RoI pool即位置敏感池化操作,针对上面的其中一个Score Map

执行池化操作,重新排列成大小为 $k^2$ 的得分图。在计算 $k^2$ 个得分图时,假设每个RoI区域大小为 $w \times h$ ,则每个小区域(bin)大小为 $w/k \times h/k$ ,每个bin得分计算如式(1)所示

$$r_c(i, j | \Theta) = \sum \frac{1}{n} z_{i,j,c}(x + x_0, y + y_0 | \Theta) \quad (1)$$

公式各参数详情请参见文献[8]。

然后对 $k^2$ 个得分图进行投票,生成 $C+1$ 维向量,利用softmax函数计算每一类的置信度。

### 1.2 深度残差网络

2015年何凯明等在图像识别大赛上提出一种卷积神经网络,命名为深度残差网络(ResNet)<sup>[14]</sup>,该网络通过采用残差学习的思想对多层的残差映射进行拟合来解决传统网络的退化问题,残差单元如图2所示。

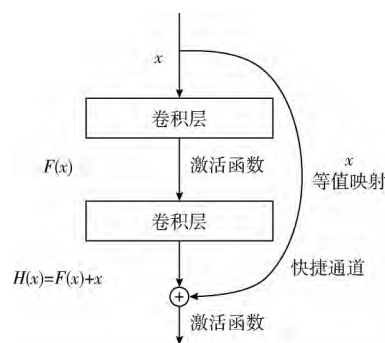


图2 残差单元

$x$ 为当前残差块的入口, $F(x)$ 表示残差单元第二层的输出, $H(x)$ 为当前残差块的期望映射,通过使用激活函数来缩短学习周期。

从图2中可以看出,残差单元通过“快捷通道”的方式,直接将输入 $x$ 传递到输出目标值 $H(x)$ ,为实现 $H(x)$ 与 $x$ 之间的恒等映射,即 $H(x) = x$ ,将学习目标改变为目标值 $H(x)$ 和 $x$ 的差值,即残差 $F(x) = H(x) - x$ ,因此训练目标就是要将 $F(x)$ 结果逼近于0,实现随着网络加深,准确率不下降。

## 2 改进的 R-FCN 模型的车辆检测

本文提出的DeR-FCN车辆检测方法实现架构如图3所示。

主要由3部分构成,第一部分是由5个卷积模块conv1~conv5组成的特征提取部分,用于提取图片的特征;第二部分是以DeRPN区域建议网络为主的区域建议部分,用于生成区域候选框;第三部分是由RoI位置敏感得分图及softmax分类器组成的RoI子网络部分,用于车辆目标分类和定位。

### 2.1 深度特征融合

图像的信息在网络不断的卷积操作中被逐步提取,卷

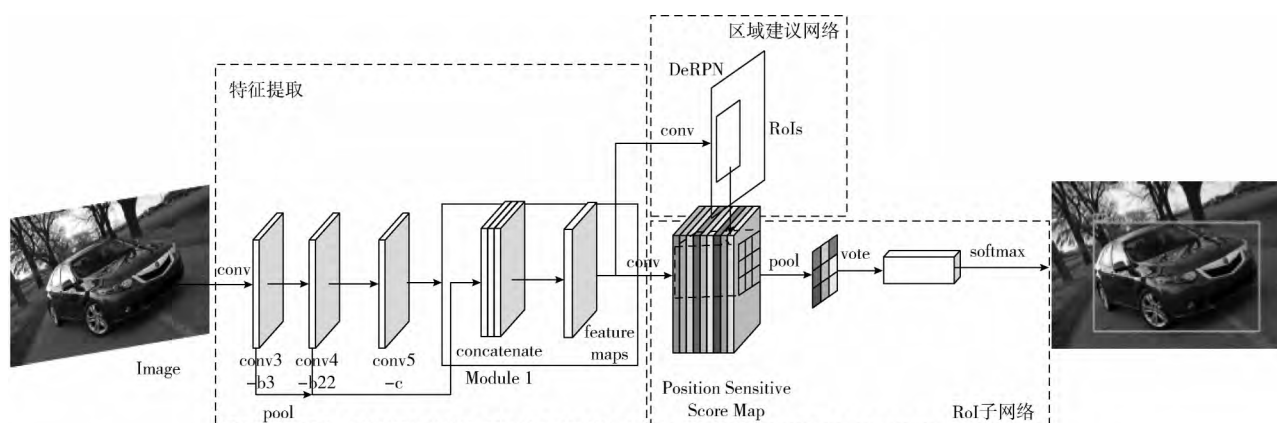


图3 DeR-FCN 模型架构

积得到的特征图尺寸也逐渐变小,这意味着网络对原图像的感受野逐渐变大,因此网络的底层特征提取器更好地抓住了图像的位置和细节信息,但由于经过的卷积较少,提取到的特征语义信息较低,噪声较多;网络的高层特征提取器更好地抓住了图像的语义信息,但提取到的特征分辨率很低,对细节的感知能力较差。为解决特征图中包含信息不充分的问题,DeR-FCN 模型以 ResNet 101 为基础网络,去掉网络中原有的全局平均池化层和最后的全连接层,只使用卷积层计算特征图,并使用特征融合模块将包含丰富位置及细节信息的底层特征和包含语义信息的高层特征进行拼接得到最终的特征图。DeR-FCN 模型有一个特征融合模块,即 Conv3-b3、Conv4-b22 与 Conv5-c 的特征融合模块,融合模块结构如图4所示。

为了将 Conv3-b3、Conv4-b22 与 Conv5-c 的特征图进行融合,本文选用不同尺度特征直接级联预测的方式(Concat)进行特征融合。假设 Conv3-b3 输出尺度为  $28 \times 28 \times 512$ ,则 Conv4-b22 与 Conv5-c 的输出尺度分别为  $14 \times 14 \times 1024$  和  $14 \times 14 \times 2048$ ,所以进行特征级联操作之前需要对 Conv3-b3 的特征图进行下采样,模块 Conv3-b3 首先通过卷积核大小为  $2 \times 2$  的平均下采样层,输出通过激活函数层(ReLU)得到  $14 \times 14 \times 512$  的特征图。Conv3-b3 经过下采样层的输出与 Conv4-b22 及 Conv5-c 的输出依据层次前后进行特征级联操作,合并之后传入 ReLU 层得到  $14 \times 14 \times 3584$  的特征图。然后添加一个卷积核为  $1 \times 1$  通道数为 1024 的卷积层,降低输入特征图的维度,最后在一个 ReLU 层之后实现融合功能。

## 2.2 维度分解的候选区域提取方法

区域建议方法作为两阶段检测算法的重要组成部分,对最终检测结果有重要影响。近年来基于区域的检测方法绝大多数都使用 RPN 来生成区域提议。但 RPN 中采用的锚箱非常敏感,限制了对变体对象的适应性。在一些特殊的目标检测模型中直接使用 RPN 作为区域提议方法,不能产生应有的效果,且性能会显著缩减,所以许多用于特殊

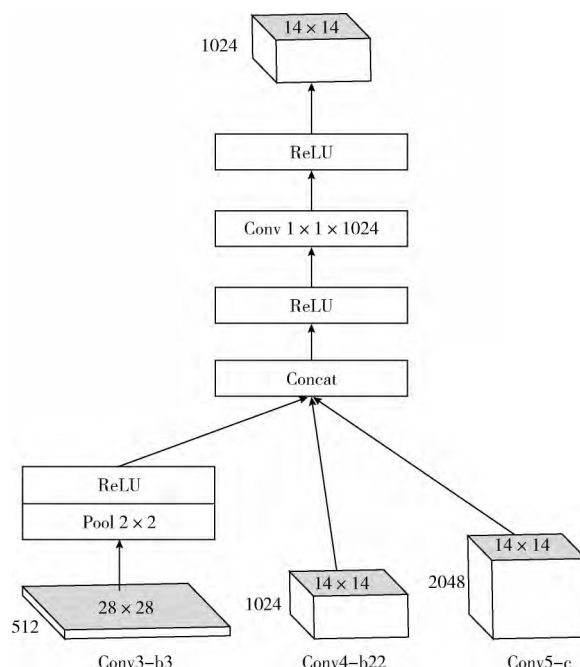


图4 特征融合模块

场景检测的方法强调手动锚箱的设计<sup>[15,16]</sup>。但是,手动设置锚箱很麻烦,很难保证最佳性能。尽管 Redmon 和 Farhadi 提出了一种基于 K 均值聚类的算法来预先计算锚箱,但改进仍然有限。因此,本文采用维数分解的思想来解决锚箱敏感的问题,通过维度分解机制,使用灵活的锚边,生成具有高召回率和更精准的区域提案,从而提高模型的检测精度,图5为 DeRPN 示意图。

DeRPN 网络将 Conv4-b22 模块输出的特征图作为输入,输出大小不一的矩形区域候选框。DeRPN 网络的操作流程如图6所示,主要包含了维度分解和维度重组两个步骤。

维度分解引入锚边(anchor strings)机制,让目标的宽高独立地与锚边进行匹配,以寻求最佳的回归参考,如

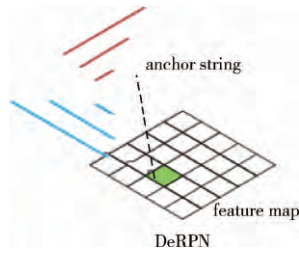


图 5 DeRPN

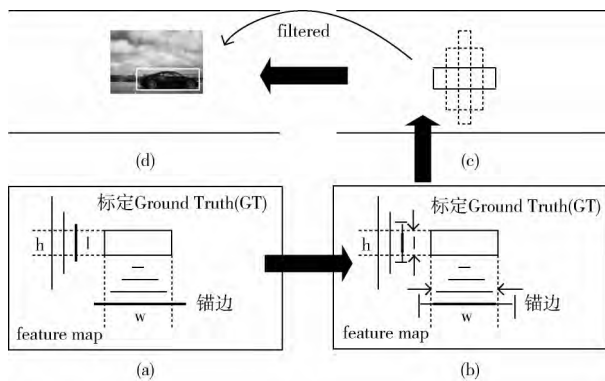


图 6 DeRPN 流程

图 6 (a) 所示, 粗线代表匹配良好的锚边。匹配最佳锚边实现如式 (2) 所示

$$M_j = \{i | \arg \min_i |\log e_j - \log a_i| \cup \{i, i+1 | \left| \frac{e_j}{a_i} - \sqrt{q} \right| \leq \beta\} \} (i = 1, 2, 3, \dots, N) \quad (2)$$

其中,  $M_j$  表示第  $j$  个目标的匹配锚边的索引集,  $e_j$  是目标的宽度或高度,  $a_i$  是第  $i$  个锚边。

接下来对锚边进行分类和回归操作, 如图 6 (b) 所示, 实线代表分类回归后高概率的锚边, 锚边回归预测如式 (3) 所示

$$\begin{aligned} t^w &= W_r^w x + b_r^w \\ S_w(x, w) &= \Psi(t^w, S_a^w(x_a, w_a)) \\ t^h &= W_r^h x + b_r^h \\ S_h(y, h) &= \Psi(t^h, S_a^h(y_a, h_a)) \\ P_s^w &= \sigma(W_c^w x + b_c^w) \\ P_s^h &= \sigma(W_c^h x + b_c^h) \end{aligned} \quad (3)$$

然后将所预测的线段经过维度重组恢复成二维的候选区域框, 如图 6 (c) 所示。DeRPN 网络采用了像素组合的维度重组算法, 可以精确地召回检测目标, 实现如式 (4) 所示

$$\begin{aligned} B(x, y, w, h) &= f(S_w(x, w), S_h(y, h)) \\ P_B &= g(P_s^w, P_s^h) \end{aligned} \quad (4)$$

最后过滤掉重叠的以及得分较低的边界框, 从而得到高概率目标区域候选框, 如图 6 (d) 所示。

### 2.3 软化非极大值抑制

R-FCN 模型在预测阶段会产生大量冗余候选框, 加入非极大值抑制算法 (NMS) 能够准确剔除重叠的候选框, 保留目标最佳坐标, 提高模型的平均精度。NMS 算法在目标与目标相距较远的情况下有着较好的检测定位效果, 但当不同的目标在重叠区域出现时, 则导致对目标的检测失败, 从而影响模型的检出率。因此 DeR-FCN 引入 Soft-NMS<sup>[17]</sup> 代替传统 NMS 算法。Soft-NMS 算法如式 (5) 所示

$$S_i = \begin{cases} s_i & IOU(A, b_i) < N \\ s_i(1 - IOU(A, b_i)) & IOU(A, b_i) \geq N \end{cases} \quad (5)$$

Soft-NMS 算法各参数详情请参见文献 [17]。

Soft-NMS 算法对 IoU 大于阈值的窗口, 进行得分加权处理, 并非直接置为 0, 从而有效避免漏检重叠区域目标, 并且 Soft-NMS 算法对剔除冗余检测结果有较好的效果, 有效提升了模型的检测精度。

在模型训练过程中, DeR-FCN 图像输入尺寸设置为 600、850、1100, 每张图像被随机分配一种尺寸输入到网络进行训练, 从而实现模型的多尺度训练, 增强模型对不同尺度目标的特征提取能力及模型的泛化能力。同时为解决模型学习样本不均衡的问题, 加入在线难例挖掘算法, 挑选出训练过程中损失值最大的  $k$  个目标区域作为难例加入网络进行训练, 计算损失, 并反向传播累计梯度来更新整个网络的参数, 从而提高网络对数据集中难样本或与正样本相似的混淆负样本的识别效果。

## 3 实验与结果分析

### 3.1 实验平台

实验运行环境见表 1。

表 1 实验运行环境

类别	版本号
操作系统	Ubuntu16.04
CPU	Xeon E5-2698 v4
GPU	NVIDIA Tesla V100
显存	32 G
深度学习框架	Caffe
脚本语言	python 2.6
CUDA 版本	CUDA 9.0
cuDNN 版本	cuDNN 7.4

### 3.2 实验数据集

为验证本文改进算法在真实场景中的实用性, 选用包含高速公路、城市道路和乡村道路等真实场景图像的 KITTI 数据集, 人为对原数据集进行筛选, 保留包含 Car, Van, Truck 和 Tram 类别标签的 6820 张图像作为实验数

据,同时根据实验框架需求将其转化为 PASCAL VOC2007 数据集格式,其中 80%划分为训练集,20%划分为测试集。并且为了验证 DeR-FCN 针对不同物体检测的有效性,本文在开源的数据集 PASCAL VOC2012 上对 DeR-FCN 与常用的目标检测算法进行对比评估,该数据集包括 20 个不同的对象类别,以及 17 125 张图片。实验前随机将数据的 60%划分为训练集,40%划分为测试集。

### 3.3 模型评价指标

本文使用 AP 和 mAP 作为模型的评价指标,AP 由精度和召回率两部分组成。

(1) 精度 P (precision) 指被判定为正例中真正的正例样本的比重,精度计算如式 (6) 所示

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

(2) 召回率 R (recall) 指被判定的真正的正例样本占总的正例样本的比重,召回率计算如式 (7) 所示

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

其中,TP 为模型预测正确的正样本个数,FP 为模型预测错误的负样本个数,FN 为模型预测错误的正样本个数。

首先以每一类的精度和召回率为坐标绘制曲线,得到该类的 R-P 曲线,然后计算每一类 R-P 曲线与坐标轴之间面积得到该类的 AP,最后再对所有类别的 AP 求平均值,即可求得 mAP。

### 3.4 模型训练参数

实验模型训练参数见表 2。

表 2 模型训练参数

参数名	参数值
base_lr	0.001
max_iters	110 000
lr_policy	"step"
gamma	0.1
momentum	0.9
weight_decay	0.0005
stepsize	80 000

DeR-FCN 模型在 KITTI 数据集上训练时总共迭代 110 000 次,经过 80 000 次迭代学习率降为原来的 0.1,模型训练的损失收敛情况如图 7 所示。

从图 7 中可以看出,模型训练损失曲线随迭代次数的增加呈下降趋势,当模型训练次数达到 100 000 次之后,损失值总体下降趋势趋于平缓,不再有大跨度下降,至此模型训练结束。

### 3.5 实验结果与分析

#### 3.5.1 不同模型在 KITTI 数据集上检测效果对比

利用基于回归的检测算法 YOLOv2、SSD 以及基于区

域建议的检测算法 Faster R-CNN、R-FCN、DeR-FCN 对 KITTI 中的测试集图像进行检测,各算法的类别检测结果如图 8 所示。

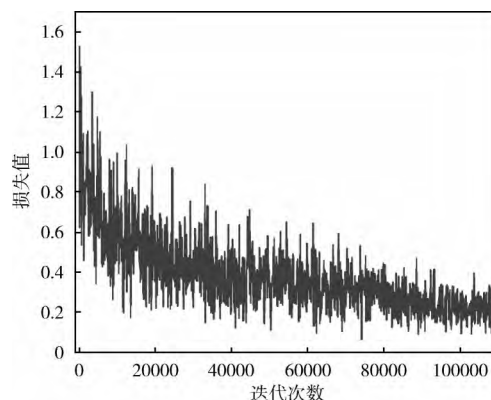


图 7 损失曲线

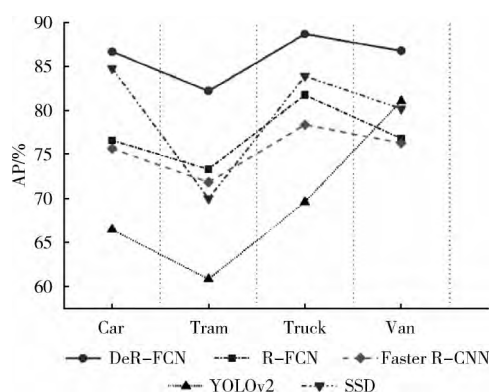


图 8 不同模型的检测准确率对比

从图 8 中可以看出,DeR-FCN 算法在各类别上的检测精度均高于其它算法,且与原算法相比各类别的检测精度都有很大的提升,这是因为 DeR-FCN 网络结构融合了低层特征的位置和边缘信息和高层特征的语义信息,使得模型在保持原有检测精度的前提下对小目标检测具有更强的适应性;引入 DeRPN 网络提高了模型对变体对象的适应性,获得更加贴合目标的候选框;预测阶段使用软化非极大值抑制算法有效避免重叠区域车辆漏检;训练阶段使用在线难例挖掘算法有效提高模型对难识别样本或与正样本相似的混淆负样本的识别效果,从而提高了模型的检出率。

表 3 是 Faster R-CNN、YOLOv2、SSD、R-FCN 与本文提出的 DeR-FCN 算法在 KITTI 数据集进行测试的实验结果。从表中可以看到,本文提出的 DeR-FCN 获得了 86.06% 的 mAP 和 0.184 s/张的检测速度,在精度方面优于所有其它算法,但在检测速度方面逊色于基于回归的检测模型 YOLOv2 和 SSD,与原 R-FCN 算法相比本文提出的方法检测精度提高了 8.96 个百分点,由于改进后算法网络结构更加复杂每张图像检测时长增加了 0.008 s,但也基本



满足实时检测的要求。

表 3 不同模型在 KITTI 数据集上检测结果对比

算法	mAP/%	AP/%				检测速度/s
		Car	Van	Truck	Tram	
YOLOv2	69.48	66.45	60.81	69.57	81.07	0.039
SSD	79.70	84.77	69.96	83.88	80.17	0.098
Faster R-CNN	75.51	75.63	71.81	78.34	76.26	0.213
R-FCN	77.10	76.57	73.30	81.74	76.79	0.176
DeR-FCN	86.06	86.65	82.18	88.66	86.75	0.184

图 9 显示了 DeR-FCN 方法在 KITTI 数据集上检测的效果。

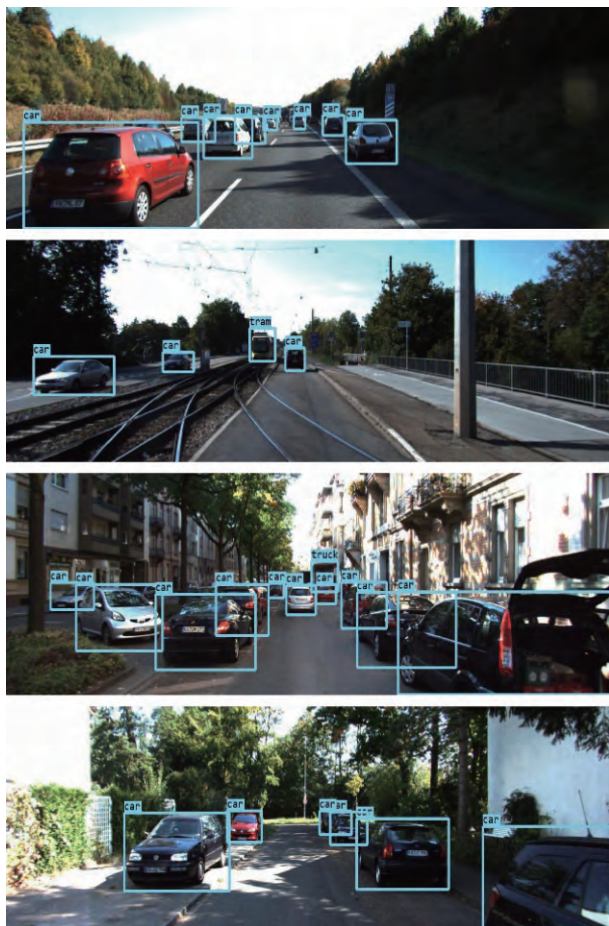


图 9 模型检测结果展示

从图 9 中可以看出,本文提出的模型对复杂环境中遮挡车辆及小目标车辆检测具有较好的鲁棒性。

### 3.5.2 不同模型在 PASCAL VOC 数据集上检测效果对比

DeR-FCN 算法与其它算法在 PASCAL VOC 数据集进行训练与测试的实验结果见表 4。从表中可以看到,本文改进后的 DeR-FCN 对 20 个不同物体类别的 mAP 为 71.5%,与改进前的 R-FCN 模型相比提高了 2.6 个百分点,实验结

果表明,改进后的 DeR-FCN 模型适用于一般物体的目标检测。

表 4 不同模型在 PASCAL 数据集上检测结果对比

算法	mAP/%
YOLOv2	66.1
SSD	63.6
Faster R-CNN	66.4
R-FCN	68.9
DeR-FCN	71.5

## 4 结束语

针对复杂城市环境下车辆检测遮挡、小目标漏检、误检等精度低下问题,本文提出了 DeR-FCN 算法,考虑特征提取阶段不同卷积层的特征图具有不同图像信息的特点,加入深度特征融合模块,并使用维度分解区域提议网络生成更加精准的区域候选框,同时在预测阶段引入软化非极大值抑制算法,显著提高模型对遮挡目标的检出率。在公共数据集上的实验结果表明,本文提出的 DeR-FCN 车辆检测算法在检测精度上具有较为明显的优势,且适用于其它物体的目标检测。未来考虑在保证模型检测精度的基础上,裁剪优化基础网络,提升网络的实时性能,以便更好应用于现实生活中。

## 参考文献:

- [1] Wen X, Shao L, Fang W, et al. Efficient feature selection and classification for vehicle detection [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25 (3): 508-517.
- [2] Shujuan S, Zhize X, Xingang W, et al. Real-time vehicle detection using Haar-SURF mixed features and gentle AdaBoost classifier [C] //The 27th Chinese Control and Decision Conference. IEEE, 2015: 1888-1894.
- [3] YIN Yu, ZHENG Hong, GAO Tingting, et al. Algorithm of license plate recognition based on joint HOG feature [J]. Computer Engineering & Design, 2015, 36 (2): 476-481 (in Chinese). [殷羽, 郑宏, 高婷婷, 等. 基于联合 HOG 特征的车牌识别算法 [J]. 计算机工程与设计, 2015, 36 (2): 476-481.]
- [4] LING Yongguo, HU Weiping. Vehicle detection under complicated background [J]. Computer Engineering & Design, 2016, 37 (6): 1573-1578 (in Chinese). [凌永国, 胡维平. 复杂背景下的车辆检测 [J]. 计算机工程与设计, 2016, 37 (6): 1573-1578.]
- [5] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation

- [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [6] Girshick R. Fast r-cnn [C] //Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [7] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks [C] //Advances in Neural Information Processing Systems, 2015: 91-99.
- [8] Dai J, Li Y, He K, et al. R-fcn: Object detection via region-based fully convolutional networks [C] //Advances in Neural Information Processing Systems, 2016: 379-387.
- [9] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [10] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector [C] //European Conference on Computer Vision. Springer, 2016: 21-37.
- [11] FEI Wei, XU Yanming, ZHU Yongying, et al. The target detection method of aerial photography images with improved SSD [J]. Journal of Software, 2019, 30 (3): 738-758 (in Chinese). [裴伟, 许晏铭, 朱永英, 等. 改进的SSD航拍目标检测方法 [J]. 软件学报, 2019, 30 (3): 738-758.]
- [12] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.
- [13] Xie L, Liu Y, Jin L, et al. DeRPN: Taking a further step toward more general object detection [J]. arXiv Preprint arXiv: 1811.06700, 2018.
- [14] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [15] WANG Lin, ZHANG Hehe. Application of Faster R-CNN model in vehicle detection [J]. Journal of Computer Application, 2017, 38 (3): 666-670 (in Chinese). [王林, 张鹤鹤. Faster R-CNN模型在车辆检测中的应用 [J]. 计算机应用, 2017, 38 (3): 666-670.]
- [16] ZHAO Zhenbing, CUI Yaping, QI Yincheng, et al. Detection method of insulator in aerial inspection image based on modified R-FCN [J]. Computer Science, 2019, 46 (3): 159-163 (in Chinese). [赵振兵, 崔雅萍, 戚银城, 等. 基于改进的R-FCN航拍巡线图像中的绝缘子检测方法 [J]. 计算机科学, 2019, 46 (3): 159-163.]
- [17] Bodla N, Singh B, Chellappa R, et al. Soft-NMS-improving object detection with one line of code [C] //Proceedings of the IEEE International Conference on Computer Vision, 2017: 5561-5569.