

文章编号: 2095-2163(2020)07-0097-05

中图分类号: TP391

文献标志码: A

# 基于改进的 Faster RCNN 模型在车辆类型检测中的应用

魏相站, 邵丽萍, 周 骅

(贵州大学 大数据与信息工程学院, 贵阳 550025)

**摘要:** 针对车辆类型检测受物体遮挡以及车辆重叠等影响, 导致车辆类型的检测难度大的问题, 本文提出了一种基于改进 Faster RCNN 模型的检测方法, 在特征提取网络中嵌入卷积模块的注意力机制模块结构, 使得特征提取网络可以重点关注与目标相关的有用信息, 并弱化其它的无用信息, 还引用 Soft-NMS 算法优化 NMS 算法, 减少重叠或相邻目标漏检和错检问题。测试结果表明, 与未改进的 Faster RCNN 模型做对比, 改进后 Faster RCNN 模型的 MAP 值由 84% 提升至 89%, 证明使用该方法使检测精度有一定的提升。

**关键词:** Faster RCNN; 特征提取; 注意力机制模块; Soft-NMS

## Application of improved Faster RCNN model in vehicle type detection

WEI Xiangzhan, SHAO Liping, ZHOU Hua

(College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

**[Abstract]** Vehicle type detection is affected by object occlusion and vehicle overlap, which makes vehicle type detection difficult. This paper proposes a detection method based on the improved Faster RCNN model, which embeds the Convolutional Block Attention Module (CBAM) structure of the convolution module in the feature extraction network, so that the feature extraction network can focus on useful information related to the target and weaken other useless information. It also refers to the Soft-NMS algorithm optimization the NMS algorithm to reduce the missed detection and false detection problems of overlapping or adjacent targets. The test results show that compared with the unimproved Faster RCNN model, the MAP value of the improved Faster RCNN model is increased from 84% to 89%, which proves that the detection accuracy can be improved to a certain extent by using this method.

**[Key words]** Faster RCNN; feature extraction; Convolutional Block Attention Module; Soft-NMS

## 0 引言

随着计算机技术在交通系统中的广泛应用, 使得传统的交通系统变得更智能化。智能汽车作为智能化交通系统的一个重要分支, 成为研究的前沿和热点问题, 车辆检测是智能汽车领域的一个重要应用, 它对提高自动驾驶的安全性起着关键作用。

国内外学者基于车辆检测方法的研究提出了许多算法, 主要分为机器学习方法和深度学习方法。传统的目标检测算法大多采用人工特征提取算法来得到目标的特征描述信息, 如刘杰提出一种利用 SIFT 算法进行雷达 SAR 图像的匹配方法<sup>[1]</sup>; 王俭等提出了一种基于 SURF 算法的遥控器缺陷自动检测方法<sup>[2]</sup>; 李子彦等提出了一种基于局部 HOG 特征提取及识别方法<sup>[3]</sup>, 这些人工特征提取方法在不同检测目标的泛化能力很差。由于环境的复杂多变, 如光照变化、拍摄角度、图像清晰度等, 都会对特征

提取产生各种影响, 特征提取意愿差, 降低识别对象的检测率, 难以应用于日常生活中。

随着深度学习技术在人工智能领域的普及, 尤其是卷积神经网络在图像处理模块中广泛应用, 与传统的人工特征提取算法不同, 卷积神经网络以原始数据作为输入, 通过大量的训练得到原始数据中的目标特征信息, 具有较高的灵活性和泛化能力。2014年, Girshick 等人提出了 R-CNN 模型, 该模型采用了选择性搜索方法, 在图像处理模块的目标检测方向得到了显著的效果<sup>[4]</sup>; 随后 He 等提出共用特征卷积层以及空间金字塔池化对 RCNN 模型进行改进得到 SPP-Net 模型<sup>[5]</sup>; 2015年 Girshick 等借鉴 SPP-net 模型提出 Fast RCNN 模型<sup>[6]</sup>; 2017年 Ren 等提出 RPN 结构结合 Fast RCNN 模型, 进而提高检测速度得到 Faster RCNN 模型<sup>[7]</sup>; 2016年 Liu 等提出一种直接预测目标类别和目标框的多目标检

**基金项目:** 贵州大学培育项目(黔科合平台人才[2017]5788-60); 贵州大学引进人才培育项目(贵大人基合字[2015]53号)。

**作者简介:** 魏相站(1995-), 男, 硕士研究生, 主要研究方向: 图像处理; 邵丽萍(1996-), 女, 硕士研究生, 主要研究方向: 通信与信息系统;

周 骅(1978-), 男, 博士, 副教授, 硕士生导师, 主要研究方向: 深度学习、物联网安全性。

**通讯作者:** 周 骅 Email: zhouhua97@gmail.com

**收稿日期:** 2020-04-30

测 SSD 算法模型,该模型相比较 Faster RCNN 模型的检测速度大幅度提升<sup>[8]</sup>;2018 年 Redmon 等提出一种新的端到端的检测模型 YOLOV3,该模型结合新的 darknet53 网络结构,使得检测速度进一步提高<sup>[9]</sup>。Faster RCNN 模型将目标检测 4 个基本操作融合到一个深度网络框架之中,完全实现端到端的目标检测,检测精度相对于其他网络模型都要好。因此,本文选取 Faster RCNN 网络作为车型检测的基础框架。目前已经有很多基于 Faster RCNN 的改进与应用,例如张毅等提出一种基于 ResNeXt 模型改进的 Faster RCNN 检测算法的交通标志的检测<sup>[10]</sup>;汤伟等提出一种基于改进 VGG16 网络的 Faster RCNN 水面漂浮垃圾检测模型<sup>[11]</sup>。由于目标检测的原理大致相同,但面对复杂的环境因素影响,就会加大目标检测的难度,本文考虑到车辆易受光照、遮挡等环境因素影响,提出一种以 Faster RCNN 为基础模型,在特征提取网络中嵌入 CBAM 结构,使得特征提取网络可以更多地提取有用信息,同时针对目标物体重叠检测问题,优化非极大值抑制算法,进一步提高了车型检测的精度。

## 1 相关基础结构

### 1.1 CBAM 结构

CBAM(Convolutional Block Attention Module)于 2018 年提出,它是一个轻量级的通用模块,可以无缝的嵌入到任何卷积神经网络架构中,并且可以与基本的卷积神经网络一起进行端到端的训练<sup>[12]</sup>。CBAM 结构由通道注意力模块和空间注意力模块 2 部分组成,其模型结构如图 1 所示。

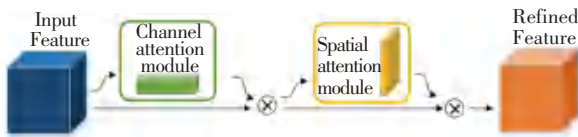


图 1 卷积注意力模块

Fig. 1 Convolutional block attention module

通道注意力模块主要关注的是目标特征信息,对输入的特征信息分别进行最大池化和平均池化操作,之后使用 2 个全连接层对图像特征进行提取,经过 sigmoid 激活后将提取后的特征信息重新作为输入图像特征的权重与输入图像特征进行内积运算,生成空间注意力模块所需要的输入特征信息。其结构如图 2 所示,该模块的计算如公式(1)、(2)所示。

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (1)$$

$$M_c(F) = \sigma(W_1(W_0(F_{avg}))) + W_1(W_0(F_{max}))) \quad (2)$$

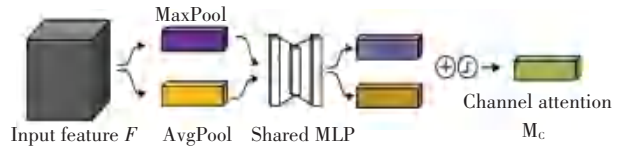


图 2 通道注意力模型

Fig. 2 Channel Attention Module

空间注意力模块主要关注的是目标位置信息,对输入的特征进行最大池化和平均池化操作,通过 concat 函数将 2 个特征信息合并,再进行卷积操作,生成一个单通道特征图,经过 sigmoid 激活后将提取后的特征信息作为输入维度权重与本模块的输入做乘法运算,得到卷积注意力机制模块的最终特征图。其结构如图 3 所示,该模块的计算如公式(3)、(4)所示。

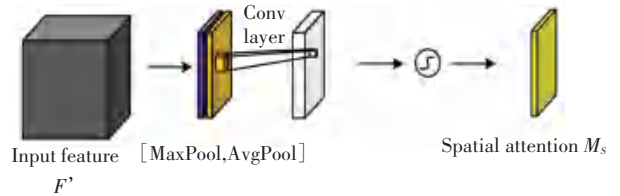


图 3 空间注意力模型

Fig. 3 Spatial Attention Module

$$M_s(F) = \sigma(f([AvgPool(F); MaxPool(F)])) \quad (3)$$

$$M_s(F) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (4)$$

### 1.2 Faster RCNN 模型原理

Faster RCNN 目标检测模型的 4 个基本操作:特征提取、候选区域生成、位置精修和分类,被融合到一个深度网络框架之中,框架结构如图 4 所示。首先,对输入模型的图像经过卷积神经网络进行特征提取;其次,将提取的特征图进行区域建议网络中滑动窗操作,取得区域建议以及区域得分,区域得分经过非极大抑制操作得到前  $N$  个得分高的区域建议,再进行 ROI 池化操作,通过该操作可以得到区域建议特征;最后,进行全连接操作得到此区域的类别分数以及此区域坐标信息。

区域建议网络(RPN)是一个全卷积网络,用来生成高品质的区域建议框,其与检测网络共同使用同一个特征图,解决了原有的选择性搜索方法的速度瓶颈问题,使得此模型的目标检测速度有了显著的提高。使用一个滑动窗口在特征图上滑动,获得一个向量,将这个向量送入 2 个平行的全连接层(分类层和位置回归层)得到类别信息和位置信息。每个滑动窗口的中心,对应  $k$  个 anchor,每个 anchor 对应一种尺寸和长宽比,相应地在每一个滑动窗位置

得到预测的候选域, 在位置回归层得到该候选域的坐标编码, 在分类层得到该候选域为目标或非目标的概率。

在 Faster RCNN 模型中, 特征提取网络在提取图像特征时会存在特征信息提取不充分的问题, 不能重点关注与目标相关的特征信息, 也会提取一些与目标相关度不大的特征信息, 使得此模型的检测准确度降低; 在区域得分中采用 NMS 算法, 最大问题就是将重叠值大于设定的重叠阈值的检测框移除, 在此状况下, 如果一个真实目标出现在重叠区域中, 模型将会漏检进而使得该模型的检测精度降低。

## 2 Faster RCNN 检测模型的优化

### 2.1 特征提取网络的改进

Faster RCNN 模型对图像信息进行特征提取时, 特征提取网络会关注图像的全部特征信息, 不能重点关注与目标相关的特征信息, 同时还会提取一些无用的特征信息, 使得检测精度下降。为了应对在特征图中很多无用的特征信息输入到特征图中, 本文在 VGG16 特征提取网络中每个卷积块后面嵌入 CBAM 结构, 不仅增大有效通道的权重, 减少无效通道的权重, 还增加有效位置信息的权重, 减少无效位置信息的权重。特征提取网络结构如图 5 所示。

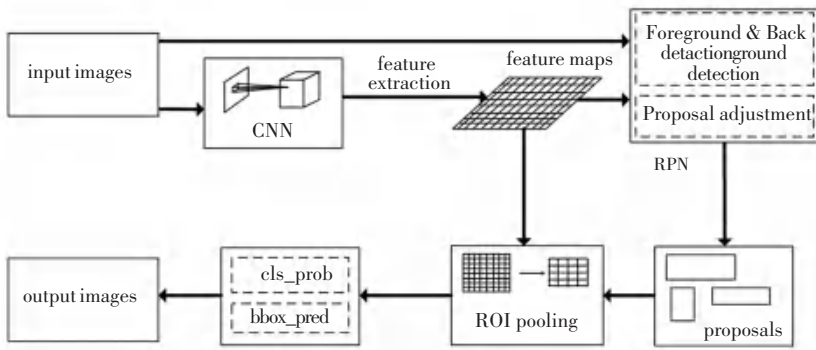


图 4 Faster RCNN 结构图

Fig. 4 The framework of Faster RCNN

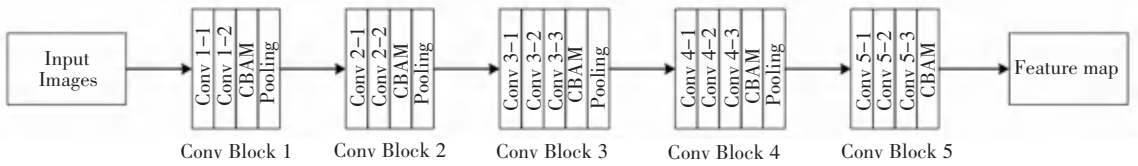


图 5 特征提取网络图

Fig. 5 Feature extraction network graph

### 2.2 非极大值抑制算法的改进

传统的 NMS 算法依靠分类器得到多个候选框, 将得到的候选框按照得分大小排序, 选择出最大分数的候选框  $M$ , 将剩余的候选框  $b_i$  依次与候选框  $M$  进行对比, 如果二者的重叠面积大于设定阈值, 就将此候选框的检测分数置为 0, NMS 算法处理方法可以通过分数重置函数 (5) 来表达:

$$s_i = \begin{cases} s_i, & iou(M, b_i) < N_t; \\ 0, & iou(M, b_i) \geq N_t. \end{cases} \quad (5)$$

其中,  $N_t$  为重叠阈值;  $s_i$  为第  $i$  个预测目标窗口对应的  $iou$  分类得分;  $M$  为当前得分最大的目标窗口;  $iou$  是预测目标窗口和实际标记窗口的交叠率。

针对 NMS 算法存在的将相邻检测框重叠值大于重叠阈值的检测分数强制归零问题, 采用线性

Soft-NMS 算法代替传统的 NMS 算法。在此算法中, 基于重叠值的高低为相邻检测框设置一个衰减函数, 而不是彻底将其分数归零。如果一个检测框  $b_i$  与最大分数的检测框  $M$  重叠值很高, 该衰减函数会让重叠值高的检测框的检测分数变得很低; 而假设一个检测框  $b_i$  与最大分数的检测框  $M$  重叠值不是很高, 此检测框的检测分数不会受到太大影响。Soft-NMS 算法处理方法可以通过分数重置函数 (6) 来表达:

$$s_i = \begin{cases} s_i, & iou(M, b_i) < N_t; \\ s_i(1 - iou(M, b_i)), & iou(M, b_i) \geq N_t. \end{cases} \quad (6)$$

其中,  $N_t$  为重叠阈值;  $s_i$  为第  $i$  个预测目标窗口对应的  $iou$  分类得分;  $M$  为当前得分最大的目标窗口;  $iou$  是预测目标窗口和实际标记窗口的交叠率。



### 3 实验分析

#### 3.1 实验环境与数据集

硬件环境配置: Ubuntu16.04 操作环境, 128GB 运行内存, 处理器 Intel 至强金牌 Xeon Gold 6130 两颗主频 2.1GHz, 显卡为 NVIDIA Quadro P5000。软件环境采用 PyTorch 深度学习框架, 在 CUDA8.0 架构平台上并行计算, 调用深度学习计算库 CuDNN 7.0.4 加速运算。

数据集: 通过网络采集得到 1 200 张车辆图像, 对图像进行旋转以及增加对比度操作, 使得数据集数量扩充为 3 000 张, 对所有数据集使用 labelImg 标定, 只标定其中 4 类不同的车型 (car、truck、van、bus), 制作成 VOC 数据集格式, 选择其中 2 100 张图像作为训练集, 600 张作为验证集, 剩余 300 张作为测试集。

在 Faster RCNN 模型训练中, 训练网络的初始学习率设置为 0.001, 权重衰减参数为 0.000 5, 模型先训练 40 000 次, 让学习率增大 0.1 倍, 再训练 25 000 次。在训练过程中, 模型的损失值会随着迭代次数的增加而减少, 直到处于稳定状态, 此时模型的识别精度也处于稳定状态, 甚至更好的状态。由图 6 可以看出, 此时的损失值曲线几乎处于稳定状态, 表明模型已经收敛。

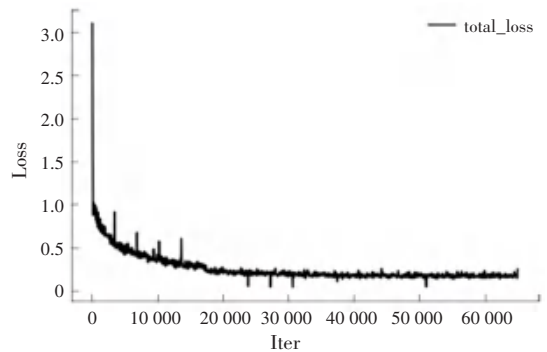


图 6 损失曲线图

Fig. 6 Curve of loss

#### 3.2 实验结果与分析

##### 3.2.1 非极大值抑制算法对识别结果的影响

非极大值抑制算法对去除冗余检测窗口、提升准确精度有着重要作用。本文采用 Soft-NMS 算法对 Faster RCNN 模型进行优化, 通过优化前后的 Faster RCNN 模型对测试集的 300 张图像进行测试, 测试结果见表 1。比较不同非极大值抑制算法的 MAP 值, 可以看出传统的 NMS 算法模型和 Soft-NMS 算法模型的 MAP 值由 84% 提高到 87%, 提高了 3 个百分点。因此, 不直接移除重叠值较高的检测框, 采用将重叠值较高的检测框的分类置信度分数降低的方法, 可以使得模型检测精度有一定提高。

表 1 不同非极大值抑制算法的识别结果

Table 1 Recognition results of different non-maximum suppression algorithms

Method	Car (AP)	Truck (AP)	Van (AP)	Bus (AP)	MAP
Faster RCNN	0.813	0.885	0.807	0.874	0.84
Faster RCNN+Soft-NMS	0.859	0.897	0.826	0.886	0.87

##### 3.2.2 嵌入 CBAM 结构的特征提取网络对识别结果的影响

在传统的 Faster RCNN 模型的特征提取网络基础上, 在特征提取网络的卷积块后嵌入 CBAM 结构, 通过对测试集测试得到的测试结果见表 2。由表 2 可知嵌入 CBAM 结构特征提取网络的 Faster RCNN 模型的 MAP 值由 84% 提升至 86%, 本文在嵌入 CBAM 结构的同时加入 Soft-NMS 算法的 MAP

值达到 89%, MAP 值提高了 5 个百分点, 但是相应的检测速度就比传统的 Faster RCNN 模型慢了很多。结果表明, 本文提出的结合 CBAM 结构的特征提取网络和优化非极大值抑制算法是有效的, 可以一定程度的提高检测精度。图 7 为改进后的 Faster RCNN 模型检测结果, 与传统的 Faster RCNN 模型相比, 可以检测出更多的遮挡目标以及重叠的目标。

表 2 模型改进后的识别结果

Tab. 2 Recognition results after model improvement

Method	Car (AP)	Truck (AP)	Van (AP)	Bus (AP)	MAP	Time/ms
Faster RCNN	0.813	0.885	0.807	0.874	0.84	191
Faster RCNN+CBAM	0.832	0.890	0.825	0.882	0.86	360
Faster RCNN+CBAM+Soft-NMS	0.891	0.907	0.891	0.890	0.89	385