



基于改进R-FCN的交通标志检测

喻清挺¹, 喻维超², 喻国平¹

(1.南昌大学信息工程学院,南昌 330031; 2.国家电网南昌供电公司,南昌 330077)

摘要:为在交通标志检测过程中同时满足精度和速度的需求,建立一种改进的基于区域全卷积网络(R-FCN)的交通标志检测模型。通过K-means聚类算法对数据集进行分析,选择合适的锚点框。对特征提取网络ResNet101进行结构简化,只使用前25层来提取特征,以缩短检测时间。在模型中引入可变形卷积和可变形位置敏感RoI池化层,以提高模型对交通标志的感应能力。模型训练过程中使用在线困难样本挖掘策略从而减少简单样本数量。在交通标志检测数据集GTSDB上的实验结果表明,该模型对交通标志位置信息较敏感,AP₅₀和AP₇₅指标分别达到97.8%和94.7%,检测时间缩至48 ms,检测精度与速度优于Faster R-CNN、R-FCN等模型。

关键词:交通标志;区域全卷积网络;ResNet101网络;可变形卷积;可变形位置敏感RoI池化

开放科学(资源服务)标志码(OSID):



中文引用格式:喻清挺,喻维超,喻国平.基于改进R-FCN的交通标志检测[J].计算机工程,2021,47(12):285-290,298.
英文引用格式:YU Q T, YU W C, YU G P. Traffic sign detection based on improved R-FCN[J]. Computer Engineering, 2021, 47(12): 285-290, 298.

Traffic Sign Detection Based on Improved R-FCN

YU Qingting¹, YU Weichao², YU Guoping¹

(1. Information Engineering School, Nanchang University, Nanchang 330031, China;

2. State Grid Nanchang Electric Power Company, Nanchang 330077, China)

[Abstract] To meet both the precision and speed requirements in the process of traffic sign detection, this paper proposes an improved traffic sign detection model based on improved Region-based Fully Convolutional Network (R-FCN). By analyzing the data set through the K-means clustering algorithm, the appropriate anchor box is selected. Then the structure of the feature extraction network, ResNet101, is simplified, and only the first 25 layers are used for feature extraction to reduce the detection time. The deformable convolution and the deformable position-sensitive RoI pooling layer are introduced into the model to improve the ability of the model to sense traffic signs. In the training process, the online hard example mining strategy is used to reduce the number of simple samples. The experimental results on the GTSDB dataset for traffic sign detection show that the improved model is more sensitive to traffic sign location information. Its AP₅₀ reaches 97.8%, and the AP₇₅ reaches 94.7%. The model also reduces the detection time to 48 ms, displaying a higher accuracy and speed than Faster R-CNN, R-FCN and other models.

[Key words] traffic sign; Region-based Fully Convolutional Network (R-FCN); ResNet101; deformable convolution; deformable position-sensitive RoI pooling

DOI: 10.19678/j.issn.1000-3428.0060093

0 概述

交通标志检测是自动驾驶和辅助驾驶的重要组成部分,交通标志中含有速度限制、路面状况等信息,是安全驾驶的重要一步,因此,对交通标志进行研究具有重要的理论和现实意义。研究人员已经提出多种交通标志检测方法,在深度学习模型提出以

前,交通标志检测方法^[1-3]大体分为基于交通标志颜色形状的视觉信息方法^[4-5]、基于特征提取和机器学习的识别方法^[6-8]等。

自从AlexNet^[9]在ImageNet大赛夺冠后,多种深度学习模型被提出并应用于交通标志检测任务。文献[10]设计一种端到端的全卷积神经网络,其在Tsinghua-Tencent 100K数据集检测中表现出强大的

基金项目:江西省重点研发计划项目(20161BBE50089)。

作者简介:喻清挺(1996—),男,硕士研究生,主研方向为深度学习、图像处理;喻维超,工程师、硕士;喻国平(通信作者),教授。

收稿日期:2020-11-24 修回日期:2021-01-12 E-mail:1131376488@qq.com

鲁棒性。文献[11]对Faster R-CNN^[12]中的区域建议网络(RPN)接收域进行修改,使模型能够更好地适应交通标志这种小目标。文献[13]通过使用多个 1×1 的卷积核,在降低模型计算复杂度的同时提高了检测的实时性。文献[14]提出一种基于多尺度区域的卷积神经网络,其在RPN中融入上下文信息,能准确检测出较小的交通标志。

由于模型结构设计存在差异,导致各种目标检测模型的侧重点不同:一阶检测器(如YOLO^[15]、SSD^[16]等)检测速度快,但对交通标志这种小目标并不敏感,精度较低,容易产生误判;二阶检测器(如Faster R-CNN、FPN^[17]等)精度高,但速度较慢,对于交通场景这种变化较快的情况,无法及时将信息传递给汽车。为此,本文基于区域全卷积网络(R-FCN)^[18]框架提出一种改进的交通标志检测算法。通过对R-FCN模型的特征提取网络进行简化,以在不影响准确率的情况下提高模型检测速度。引入K-means聚类算法、在线困难样本挖掘(OHEM)策略和可变形卷积技术,从而提高模型检测精度。

1 交通标志检测算法框架

1.1 R-FCN算法

R-FCN算法由经典Faster R-CNN算法演变而来,两者同为二阶段检测器。R-FCN算法先通过特征提取网络提取图像特征并输出特征图,再经过区域建议网络在输出的特征图上提取感兴趣区域(RoI)。

R-FCN与Faster R-CNN在RPN之前的步骤相同,两者最大的不同在于对RoI的分类与回归。Faster R-CNN通过一组全连接层完成目标分类和回归,每个被RPN提取出来的RoI都需要经过全连接层进行计算,这将消耗大量的时间;R-FCN通过额外添加的2组卷积,得到2张分别为类别和位置的得分图,从而对目标进行预测,即将全连接层替换成卷积层进行分类和回归计算,这在一定程度上加快了检测速度。R-FCN检测框架如图1所示。

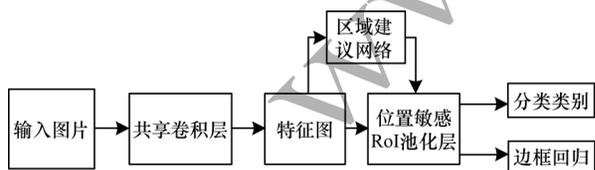


图1 R-FCN检测框架

Fig.1 R-FCN detection framework

1.2 可变形卷积

传统的卷积操作通常基于 3×3 、 5×5 等形状规则的卷积核进行特征提取,然而这种方式具有一定的局限性,只能提取规则区域内的特征,无法根据物体

的形状提取到更细致的特征。可变形卷积^[19]具有学习空间几何形变的能力,可以更好地适应不同空间。可变形卷积可以根据目标的形状进行卷积位置调整,使卷积的区域更符合物体形状特征,具体实现过程为:增加一组 $2N$ (N 为卷积核大小)个卷积核,产生尺寸与输入特征图相同、通道数为 $2N$ 的特征图,称为偏置域(offset field),从offset field上可得到每个像素点的偏置矩阵(offsets)。可变形卷积结构如图2所示。

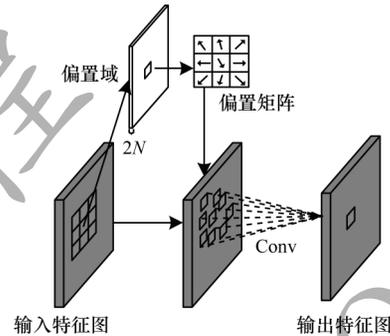


图2 可变形卷积结构

Fig.2 Deformable convolution structure

传统卷积的计算方式是与特征图进行固定区域的卷积计算,对于特征图上的每个点位置 p_0 ,计算方式如下:

$$y(p_0) = \sum_{p_n \in R} w(p_n)x(p_0 + p_n) \tag{1}$$

其中: x 为输入特征图; y 为输出特征图; p_0 为卷积计算点; R 表示预采样偏移,假设使用 3×3 卷积核,则 $R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$; p_n 为 R 中位置的枚举。

可变形卷积将学习到的偏置矩阵offsets与预采样偏移 R 相加,得到新的采样偏移量,则输入特征图上每个点 P_0 进行可变形卷积的表达式如下:

$$y(p_0) = \sum_{p_n \in R} w(p_n)x(p_0 + p_n + \Delta p_n) \tag{2}$$

其中: Δp_n 为偏置矩阵offsets对应于常规卷积网络的偏移量。

1.3 特征提取网络

R-FCN通过残差网络(ResNet)对图片进行特征提取,ResNet^[20]被广泛应用于分类、检测等任务。通过增加跳线连接,ResNet能够缓解梯度消失和梯度爆炸问题,ResNet结构如图3所示。假设在不增加残差前,网络需要学习的是 $H(x) = F(x)$,这样一个恒等映射在神经网络层数增加时将变得很难拟合;在加入残差结构后,神经网络学习的是 $H(x) = F(x) + x$,即将恒等映射作为网络的一部分,则网络真正学习的是残差函数 $F(x) = H(x) - x$,这比原先的恒等映射更易拟合。

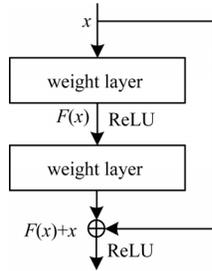


图 3 ResNet 结构

Fig.3 ResNet structure

本文使用经过预训练的 ResNet101, 在其 conv4_x 模块中融入可变形卷积, 具体结构如图 4 所示。在残差结构中加入额外的 72 个 256 维的卷积核, 产生 4 组 offsets, 通过这些 offsets 的偏置量指引, 对输入特征进行可变形卷积操作从而获得输出特征。

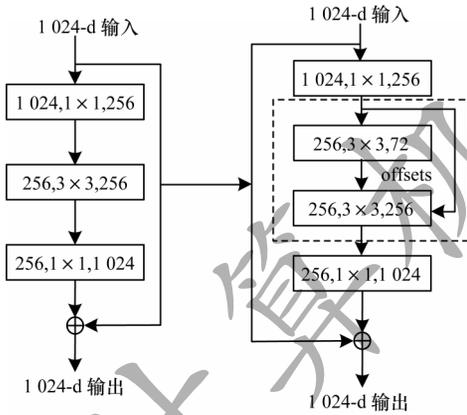


图 4 可变形 ResNet 结构

Fig.4 Deformable ResNet structure

1.4 可变形位置敏感 RoI 池化

R-FCN 在位置敏感得分图上进行位置敏感 RoI 池化, 位置敏感得分图实质上也是特征图, 但其具有确切的含义。位置敏感 RoI 池化会将 RPN 层提取的 RoI 分成 $k \times k$ 个子区域 (bins), 假设目标的总类别数为 C , 则位置敏感得分图的层数为 $k^2(C+1)$, 每一层分别代表某类物体的某个部位在整个图像中的响应值。位置敏感得分图的层数受目标种类数影响, 当目标种类增加时, 得分图的通道数也增加, 这会带来大量参数, 增加了检测时间, 但是本文只有 3 类交通标志的情况, 适合进行 R-FCN 位置敏感 RoI 池化设计。

相比于原始的位置敏感 RoI 池化, 可变形位置敏感 RoI 池化通过卷积在位置敏感得分图上产生一个额外的偏置域, 这个偏置域的尺寸与得分图相同, 通道数是得分图的两倍, 对应于位置敏感得分图上每个点在 X 与 Y 方向的偏移量。可变形位置敏感 RoI 池化如图 5 所示。

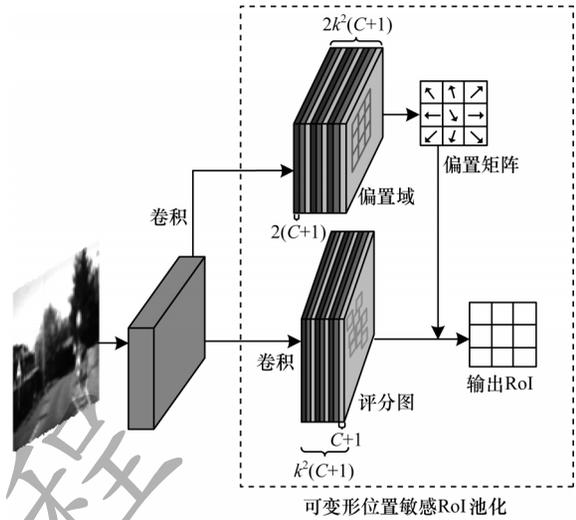


图 5 可变形位置敏感 RoI 池化

Fig.5 Deformable position-sensitive RoI pooling

将被 RPN 提取出来的 RoI 区域对应的 offsets 加入可变形位置敏感 RoI 池化中, 计算公式如下:

$$y(i, j) = \sum_{p \in \text{bin}(i, j)} x(p_0 + p + \Delta p_{ij}) / n_{ij} \quad (3)$$

其中: x 为输入特征图; y 为输出的 RoI; bin 为由 RoI 划分的子区域; n_{ij} 是每个 bin 区域中的像素数; p_0 为 RoI 左上角的坐标; p 为 bin 区域内像数点的枚举; Δp 为偏移量 offset。

1.5 在线困难样本挖掘

经过特征提取网络输出的特征图会在 RPN 的作用下产生大量的候选区域 (RoIs), 这些 RoIs 中正样本和负样本分布并不平衡, 简单地设置正负比例将这些 RoIs 送入后续分类与回归训练, 模型将无法关注困难样本 (预测时与真值标签误差较大的样本, 这类样本常使得模型产生误判现象)。在线困难样本挖掘策略^[21]会从这些 RoIs 中筛选出困难样本并送入后续训练中, 具体操作为: 设置 2 个相同的 RoI 子网络, 2 个子网络结构相同而功能不同, 根据功能不同可将其分为只读子网络和可读可写子网络, 只读子网络只负责前向计算, 输出 RoIs 的损失, 然后将这些 RoIs 的损失值从大到小排序, 并通过非最大值抑制 (NMS) 算法挑选出前 K 个 RoIs, 将其送入可读可写子网络中进行参数反向传播, 从而完成网络更新。

1.6 改进的网络结构

本文交通标志检测方法主要基于 R-FCN 网络结构, 为了加快检测速度, 在基于 ResNet101 的预训练模型上进行删减实验, 最后只保留前 25 层来实现图像特征提取, 并将 RoI 区域中 bin 的个数从 7×7 变为 3×3 。以上 2 个措施有效提升了模型的检测速度, 且并未因为参数减少而影响准确率。

上述措施虽然能提升模型检测速度,但是无法提高模型的准确率。为此,本文引入可变形卷积核和可变形位置敏感RoI池化,总共使用2个可变形卷

积核,分别置于RPN两侧,将可变形位置敏感RoI池化层代替原始位置敏感RoI池化层,改进后的网络结构如图6所示。

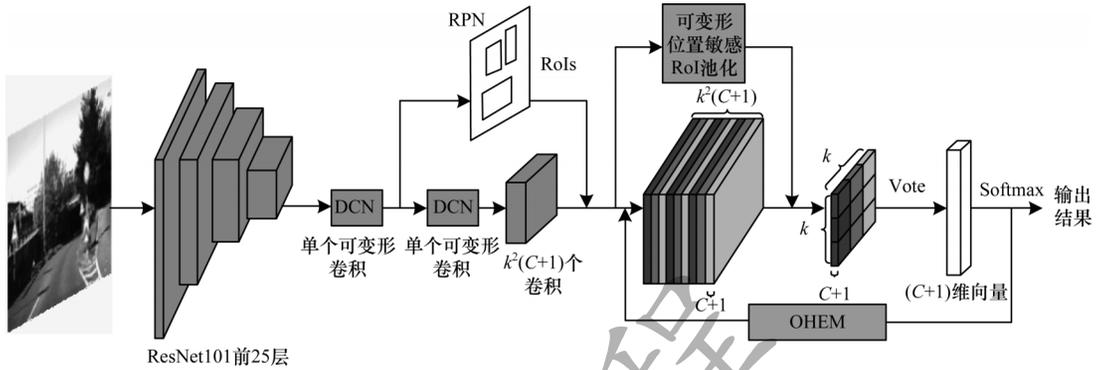


图6 改进的R-FCN网络结构

Fig.6 Improved R-FCN network structure

2 实验结果与分析

2.1 实验数据集

本文实验采用德国交通标志检测数据集(German Traffic Sign Detection Benchmark, GTSDDB)^[22]。GTSDDB发布于2013年的国际神经网络会议(IJCNN2013),被用于当年的交通标志检测大赛。该数据集包含900张高分辨率的道路交通图像,其中,训练集为600张,测试集为300张,图像大小均为800像素×1360像素。图像中交通标志的数量为0~6个,多为小目标,分辨率在16像素×16像素~128像素×128像素之间,形状大多数为三角形或圆形。GTSDDB中将43种交通标志分为四大类,即禁令标志、指示标志、警告标志、其他标志。根据当年比赛的规定,本文只对前3种标志进行检测,本文实验训练集总共包含666个标志,测试集包含273个标志。图7所示为该数据集的部分图片及其检测结果。

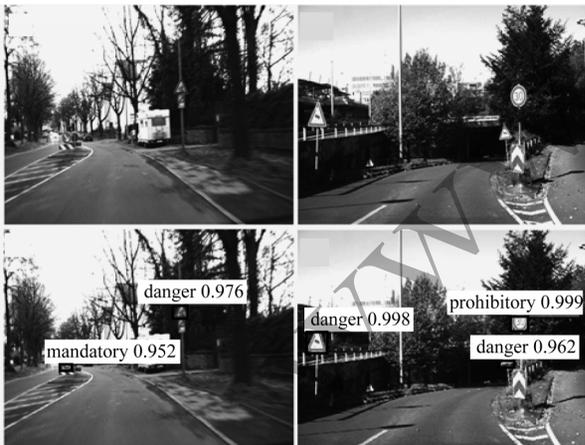


图7 GTSDDB数据集部分样本及其检测结果

Fig.7 Some samples of the GTSDDB dataset and their detection results

2.2 锚点框维度聚类

通过对数据集进行分析可知,交通标志的面积远小于原R-FCN的锚点框大小,使用原始的锚点将

会使模型难以拟合那些过小的交通标志,因此,需要对交通标志数据集进行聚类分析。

K-means聚类方法用欧氏距离定义损失函数,但在目标检测算法中更适合采用预测框和anchor box的面积重叠度IOU作为度量距离,则新的度量标准公式为:

$$d(B_{\text{box}}, C_{\text{centroid}}) = 1 - \text{IOU}(B_{\text{box}}, C_{\text{centroid}}) \quad (4)$$

其中: B_{box} 表示目标框尺寸; C_{centroid} 表示聚类中心框尺寸。通过改变聚类中心点的个数得到不同的平均IOU值,结果如图8所示。

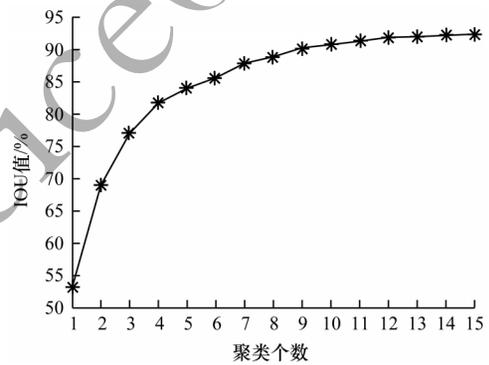


图8 K-means聚类折线

Fig.8 K-means clustering polyline

从图8可以看出,当聚类个数为9时曲线逐渐平缓。此外,通过聚类可知交通标志长宽比的标注基本为1:1。因此,本文选定9个锚点框,并且将锚点框具体大小设置为(22, 22)、(27, 27)、(32, 32)、(36, 36)、(40, 40)、(46, 46)、(56, 56)、(70, 70)、(100, 100)。

2.3 实验评价

本文实验平台设置:CPU为Inter Core i7-8700,内存16 GB,GPU为NVIDIA GeForce RTX 2070,显存8 GB,操作系统为Ubuntu 16.04,深度学习框架为MXNET。

实验采用ImageNet数据集上训练好的ResNet101,其参数不参与更新,总共训练15个epoch,初始学习率为0.001,在训练10个epoch后将学习率调整为0.0001。

本文采用目标检测中常见的 AP_{50} 、 AP_{75} 、 AP 这 3 个评价指标进行评测: AP_{50} 、 AP_{75} 分别表示 IOU(预测目标与实际目标的重叠比)大于 0.5 和大于 0.75 时正确检测出目标的概率; AP 为 IOU[0.5:0.05:0.95] 共 10 个不同阈值下检测出目标的概率的平均值。使用上述 3 个指标能更好地体现可变形网络对交通标志的适应能力, 使得评价更为全面。

为了选择适合本文的预训练模型, 首先利用以 ResNet101 和 ResNet50 为特征提取网络的 R-FCN 在原模型下进行实验, 结果如表 1 所示。

表 1 特征提取网络 ResNet101 与 ResNet50 的对比

Table 1 Comparison of feature extraction networks

ResNet101 and ResNet50			
特征提取网络	AP_{50}	AP_{75}	AP
ResNet101	96.4	88.8	67.4
ResNet50	96.4	76.5	63.2

ResNet101 相比于 ResNet50 最大的区别是 conv4_x 模块多了 51 层, 但从表 1 可以看出, 这额外的 51 层卷积层并没有使 ResNet101 相比 ResNet50 取得太大优势, 两者的性能相差不大。因此, 对于本文的 3 类交通标

志检测任务, 选用总体偏优的 ResNet101 作为特征提取网络, 为了加快检测速度, 将 conv4_x 模块进行简化, 只保留其前 3 层参数, 即整个特征提取网络只剩下 35 层。不同层数网络的实验结果如表 2 所示。

表 2 不同层数网络的识别性能对比

Table 2 Comparison of recognition performance of networks with different layers

层数	$AP_{50}/\%$	$AP_{75}/\%$	AP/%	时间/ms
101	96.4	88.8	67.4	108
35	96.5	86.7	67.8	68

从表 2 可以看出, 对特征提取网络 ResNet101 进行层数删减, 并没有对检测结果产生太大影响, 对 ResNet101 中 conv4_x 模块的简化仅造成了微小的性能变化, 说明前 25 层提取出来的特征图已经包含足够的交通标志特征信息。

在获得简化的特征提取网络后, 本文在简化模型上进行实验, 通过聚类锚点, 在模型中加入可变形卷积和可变形位置敏感 ROI 池化, 且训练过程中采用在线困难样本挖掘方法, 实验结果如表 3 所示。

表 3 不同组合方法的性能对比结果

Table 3 Performance comparison results of different combination methods

实验组号	层数压缩	K-means	OHEM	可变形卷积	$AP_{50}/\%$	$AP_{75}/\%$	AP/%	时间/ms
1	否	否	否	否	96.4	88.8	67.4	108
2	是	否	否	否	96.5	86.7	67.8	66
3	是	是	否	否	97.1	80.0	65.1	48
4	是	是	否	是	97.2	93.7	74.5	48
5	是	是	是	是	97.8	94.7	76.3	48

从表 3 可以看出: 第 5 组实验的 AP_{50} 准确率相对第 1 组实验提升了 1.4 个百分点, 达到 97.8%, 而 AP_{75} 准确率达到 94.7%, 即本文算法相比原算法提升了 6 个百分点, 总体 AP 提升了将近 9 个百分点; 在检测时间方面, 本文算法相较原始 R-FCN(ResNet101) 减少了 60 ms, 速度提升高于 2 倍。上述结果有效证明了本文保留前 25 层网络进行实验的可行性以及 K-means 聚类后锚点框的适应性。

交通标志信息标注采用人工标注方式, 其标注框都是正方形, 加入可变形卷积和可变形位置敏感 RoI 池化, 可以提取出更多关于交通标志形状的信息, 更加符合交通标志圆形、三角形以及矩形的特征, 使得在更严格的 AP_{75} 和 AP 评价指标下准确率明显提升。通过添加 OHEM 技术, 可以对困难样本进行训练, 模型的准确率进一步提升。

2.4 模型对比

在目标检测领域, 研究人员已经提出众多优秀的目标检测模型, 但这些模型都是在 VOC、COCO 数据集上进行实验, 这 2 个数据集种类多、数据量大, 模型为了适应数据集, 其结构会设计得比较庞大。这些模型虽然可以直接用于交通标志分类检测, 但其结构并不能完全契合交通标志检测数据集, 因此, 需要进行适应性改进。

表 4 所示为 9 种模型在本文数据集上的对比结果。其中: Faster R-CNN、R-FCN、FPN 的特征提取网络都为经过预训练的 ResNet101, 其参数不参与更新, 训练与检测均为原尺寸; YOLO 系列特征提取网络为 DarkNet, 尺寸则为 608×608; CenterNet 特征提取网络为 DLA-34, 尺寸为 512×512; RetinaNet 和 Cascade RCNN 的特征提取网络为 ResNet50-FPN, 尺寸为原尺寸。各模型的锚点框大小均未进行调整。

表 4 9 种模型检测性能对比

Table 4 Comparison of detection performance of nine models

模型	$AP_{50}/\%$	$AP_{75}/\%$	AP/%	时间/ms
Faster R-CNN	94.5	50.1	52.8	120
R-FCN	96.4	88.8	67.4	108
FPN	96.9	89.3	72.1	185
YOLOv3	92.4	67.8	55.9	41
YOLOv4	97.4	87.4	60.1	45
CenterNet	92.9	81.5	70.2	36
RetinaNet	93.7	89.7	74.6	81
Cascade RCNN	95.5	92.8	77.2	97
本文模型	97.8	94.7	76.3	48

从表4可以看出:FPN因为其多尺度的设计,AP和 AP_{75} 指标明显高于Faster R-CNN与R-FCN这种单尺度模型,证明通过多尺度图像特征金字塔可以有效检测出交通标志这种小目标,但FPN使用的特征提取网络为ResNet101,速度较慢;本次实验中Cascade RCNN选择参数较少的ResNet50,并在其上使用FPN架构,使用了多尺度和多阈值的Cascade RCNN的AP指标取得最高值,达到77.2%,但 AP_{50} 指标仅有95.5%,低于未改进的R-FCN;原始R-FCN由于位置敏感得分图的设计,适合于交通标志这种有特定形状的目标,且本文交通标志只有3类,位置敏感得分图层数少,这在一定程度上加快了检测速度,因此,原始R-FCN在速度与精度2个方面都取得了较好的结果;使用512尺寸的CenterNet具有最快的速度,中心点的设计也契合交通标志这种对称物体,因此,其整体表现较好;使用原尺寸的RetinaNet的AP和 AP_{75} 指标同样较高;YOLOv3与YOLOv4^[23]则在608尺寸时和本文所提模型的检测速度接近,且YOLOv4的准确率明显高于YOLOv3,其 AP_{50} 指标最接近本文模型,但 AP_{75} 和AP评价指标与本文模型有较大差距。综上,可变形卷积具有有效性,在速度接近的情况下,本文模型在更严格的评价指标上依然具有优势。

3 结束语

针对交通标志检测问题,原始R-FCN模型能取得较好的准确率,但是速度较慢,难以满足实时性的要求。为此,本文通过模型简化、K-means算法锚点框聚类,并将可变形卷积和可变形位置敏感RoI池化加入到简化模型中,在训练过程中使用在线困难样本挖掘策略,从而有效检测交通标志信息。实验结果表明,该改进模型能够在提升检测速度的同时保证检测精度,且准确率高于Faster R-CNN、R-FCN等模型,其在精度与速度上取得了较好的平衡。后续将为模型选择如EfficientNet、Desnet等更适合交通标志检测的特征提取网络,从而进一步提升模型的检测速度与精度。

参考文献

- [1] RUTA A, LI YO M, LIU X H. Towards real-time traffic sign recognition by class-specific discriminative features [EB/OL]. [2020-10-02]. <https://bura.brunel.ac.uk/bitstream/2438/14127/1/FullText.pdf>.
- [2] GAO X, HONG K, PASSMORE P, et al. Colour vision model-based approach for segmentation of traffic signs[J]. *Eurasip Journal on Image & Video Processing*, 2008(1): 1-7.
- [3] 王方石, 王坚, 李兵, 等. 基于深度属性学习的交通标志检测[J]. *吉林大学学报(工学版)*, 2018, 48(1): 319-329. WANG F S, WANG J, LI B, et al. Deep attribute learning based traffic sign detection[J]. *Journal of Jilin University (Engineering and Technology Edition)*, 2018, 48(1): 319-329. (in Chinese)
- [4] LOY G, BARNES N. Fast shape-based road sign detection for a driver assistance system [C]//*Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Washington D. C., USA: IEEE Press, 2004: 70-75.
- [5] 朱双东, 张懿, 陆晓峰. 三角形交通标志的智能检测方法[J]. *中国图象图形学报*, 2006, 11(8): 1127-1131. ZHU S D, ZHANG Y, LU X F. Intelligent approach for triangle traffic sign detection [J]. *Journal of Image and Graphics*, 2006, 11(8): 1127-1131. (in Chinese)
- [6] XIE Y, LIU L F, LI C H, et al. Unifying visual saliency with HOG feature learning for traffic sign detection [C]//*Proceedings of 2009 IEEE Intelligent Vehicles Symposium*. Washington D. C., USA: IEEE Press, 2009: 24-29.
- [7] ZAKLOUTA F, STANCIULESCU B. Real-time traffic sign recognition using spatially weighted HOG trees [C]//*Proceedings of 2011 International Conference on Advanced Robotics*. Washington D. C., USA: IEEE Press, 2011: 61-66.
- [8] BARÓ X, ESCALERA S, VITRIÀ J, et al. Traffic sign recognition using evolutionary Adaboost detection and forest-ECOC classification [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2009, 10(1): 113-126.
- [9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [EB/OL]. [2020-10-02]. <https://web.cs.ucdavis.edu/~yjee/teaching/ecs289g-winter2018/alexnet.pdf>.
- [10] ZHU Z, LIANG D, ZHANG S H, et al. Traffic-sign detection and classification in the wild [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2016: 2110-2118.
- [11] ZHANG Z, ZHOU X L, CHAN S X, et al. Faster R-CNN for small traffic sign detection [C]//*Proceedings of CCF Chinese Conference on Computer Vision*. Berlin, Germany: Springer, 2017: 155-165.
- [12] REN S, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(6): 1137-1149.
- [13] ZHANG J M, HUANG M T, JIN X K, et al. A real-time Chinese traffic sign detection algorithm based on modified YOLOv2 [J]. *Algorithms*, 2017, 10(4): 127-129.
- [14] LIU Z G, DU J, TIAN F, et al. MR-CNN: a multi-scale region-based convolutional neural network for small traffic sign recognition [J]. *IEEE Access*, 2019, 7: 57120-57128.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2016: 779-788.
- [16] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//*Proceedings of European Conference on Computer Vision*. Berlin, Germany: Springer, 2016: 21-37.

(上接第 290 页)

- [17] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA ; IEEE Press, 2017: 2117-2125.
- [18] DAI J F, LI Y, HE K M, et al. R-FCN: object detection via region-based fully convolutional networks [EB/OL]. [2020-10-02]. <https://arxiv.org/pdf/1605.06409.pdf>.
- [19] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks [C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C. , USA ; IEEE Press, 2017: 764-773.
- [20] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA ; IEEE Press, 2016: 770-778.
- [21] SHRIVASTAVA A, GUPTA A, GIRSHICK R. Training region-based object detectors with online hard example mining [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA ; IEEE Press, 2016: 761-769.
- [22] HOUBEN S, STALLKAMP J, SALMEN J, et al. Detection of traffic signs in real-world images; the German traffic sign detection benchmark [C]//Proceedings of 2013 International Joint Conference on Neural Networks. Washington D. C. , USA ; IEEE Press, 2013: 1-8.
- [23] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. [2020-10-02]. <https://arxiv.org/pdf/2004.10934.pdf>.

编辑 吴云芳