

基于改进 Anchor-Free 模型的交通标志检测算法

吕卫, 梁芷茵, 褚晶辉*

天津大学电气自动化与信息工程学院, 天津 300072

摘要 交通标志检测是自动驾驶系统的一项重要功能,当前先进的交通标志检测器大多采用 Anchor-Based 网络模型,根据锚框遍历所有潜在的目标位置。为了减少锚框带来的计算开销和过多的超参数设置,提出了一种基于编码-解码结构的 Anchor-Free 交通标志检测算法。为了增加解码模块的特征表征能力,在解码模块中引入残差增强分支。为了高效地提取和利用多尺度特征,设计了特征融合子网络,提升对多尺度目标的检测能力,并使用 Ghost 轻量化模块提取多尺度特征图,不显著引进运算量。在 Tsinghua-Tencent 100K 数据集上进行验证,所提算法实现了 92.5% 的召回率和 90.3% 的准确率,模型的参数量和模型大小分别为 1.61×10^7 和 64.4 Mbit。实验结果表明,与主流目标检测算法相比,所提算法的检测精度较高,计算开销较低,在综合性能上具有优越性。

关键词 机器视觉; 交通标志检测; Anchor-Free; 残差结构; 多尺度特征融合

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.2415002

Traffic Sign Detection Algorithm Based on Modified Anchor-Free Model

Lü Wei, Liang Zhiyin, Chu Jinghui*

School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

Abstract Traffic sign detection is an essential function of autonomous driving systems, and most modern traffic sign detectors are anchor-based, traversing potential object locations based on anchors. To solve the problems of heavy computing costs and the need to set several hyperparameters in anchor-based models, we propose an anchor-free traffic sign detection algorithm based on an encoder-decoder structure. We introduce a residual augmentation branch in the decoder module in this study to improve feature expression ability during the decoding process. To improve the ability to detect multiscale traffic signs, we propose a multiscale feature fusion subnetwork to effectively extract and use multiscale features. A Ghost lightweight module is adopted by the multiscale feature extraction module, which indistinctively increases the computational cost. On the Tsinghua-Tencent 100 K dataset, our approach achieved a recall of 92.5% and an accuracy of 90.3%, while the model's parameter amount and model size are approximately 1.61×10^7 and 64.4 Mbit, respectively. The experimental results show that the proposed algorithm outperforms the mainstream object detection algorithms in terms of precision, computing cost, and overall performance.

Key words machine vision; traffic sign detection; Anchor-Free; residual structure; multiscale feature fusion

1 引言

交通标志包含丰富的交通信息,如车速限制提示、道路状况提示、行人状况提示等,是道路上的重要设施之一,可帮助驾驶员更好地了解实时道路信息,从而减少交通拥堵和交通事故发生的次数。交通标志检测也是自动驾驶系统中的一项重要功能,由于交通标志具有鲜明的颜色特征和形状特征,传统的检测方法通常使用低级特征,如颜色、纹理和几何特征来检测交通标志。de la Escalera 等^[1]利用交通标志鲜明的颜色特征,

在颜色空间中设置阈值分割感兴趣颜色信息,来提取图像中的感兴趣区域。部分研究者通过引入手工特征,如方向梯度直方图(HOG)^[2-3]和尺度不变特征变换(SIFT)^[4],进一步提高了交通标志检测性能。然而,这些手工设计的特征在真实交通路况的复杂背景下难以获得理想的交通标志检测结果。随着计算机技术的蓬勃发展,深度学习成为解决复杂计算机视觉问题最有效的方法之一,也为交通标志检测带来了新的思路。Faster R-CNN^[5]、YOLO^[6]、SSD^[7]等网络模型在目标检测领域取得了出色的结果。Cao 等^[8]通过改进 Faster

收稿日期: 2021-09-01; 修回日期: 2021-10-12; 录用日期: 2021-10-27

通信作者: *cjh@tju.edu.cn

R-CNN中候选区域提取阶段的损失函数来优化模型的训练效果,并在分类阶段加入多尺度特征融合机制提高最终的分类效果。Zhang等^[9]将空间池化金字塔和特征增强路径引入YOLOv3中,增加模型对空间特征的利用能力。Jin等^[10]通过在SSD引入特征融合机制和通道增强机制增强对交通标志的特征表达能力。上述目标检测网络采用Anchor-Based目标检测模型,该模型通过在特征图上设置密集的锚框,回归得到目标的边界。Anchor-Based目标检测模型具有较高的召回率,但在计算锚框时也会花费大量的计算资源。同时,锚框超参数(如锚框的数量、大小和宽高比)的设置对目标检测性能有较大影响,因此在使用Anchor-Based网络模型时,需要根据不同的数据集手动调整锚框的超参数。针对当前目标检测Anchor-Based网络模型的不足,Anchor-Free网络模型被提出,该模型不依赖于预设的锚点,直接预测目标的类别与位置。CenterNet^[11]是具有代表性的Anchor-Free检测模型之一,通过预测目标的中心点位置和宽高得到目标的检测框。虽然现有的检测模型已经在大幅度交通标志检测中取得了较好的检测结果,但依然存在计算开销较大、在复杂背景下对小尺度交通标志检测效果不佳的问题。

本文提出了一种基于改进Anchor-Free网络模型的算法,该算法可有效解决现有交通标志检测模型在多尺度交通标志检测任务中的几个主要问题。首先,为解决现有交通标志算法计算开销较大的问题,所提算法采用编码-解码结构,使用ResNet-18^[12]作为特征提取网络,通过逐点预测取代了遍历大量预设锚框的检测方式,提高了预测效率。其次,现有编码-解码结构在解码过程中使用简单的解码模块,难以得到高质

量的输出热力图。为解决这一问题,所提算法在解码过程中引进残差分支,突出包含交通标志的通道域和空间域的感兴趣信息,可得到更高质量的高分辨率特征图,提高了检测性能。最后,为解决真实交通场景下的交通标志尺度变化较大、远距离的交通标志尺度较小而导致的检测性能不佳的问题,所提算法在特征提取网络后加入多尺度特征融合子网络,该子网络采用轻量化模块进一步提取多尺度特征和增强语义信息,并融合低层的细节信息和高层的语义信息,在不引过多参数的同时保持并利用多分辨率特征,提高了对多尺度交通标志的特征表征能力。

2 基本原理

2.1 基于中心点检测的Anchor-Free网络模型

基于中心点预测的网络模型已被证明是交通标志检测的有效方法^[13-14],由中心点预测和尺度预测组成。该模型不依赖预设锚框,因此无需手工设定锚框的超参数,可以有效适应不同真实场景下的交通标志目标。所提基于Anchor-Free网络的交通标志检测算法模型如图1所示。网络模型采用编码-解码结构,包含1个用于提取特征的主干网络、1个用于保持多尺度特征的多尺度特征融合子网络、1个用于提升特征图分辨率的解码子网络和3个分别负责中心点分类、中心点偏移量预测和边框预测的子网络。设原始图像的宽和高分别为 W 和 H ,所提模型采用ResNet-18作为特征提取网络,对预处理后的图像 $I \in \mathbf{R}^{W \times H \times 3}$ 进行初步特征提取,引入多尺度特征融合子网络提取并融合主干网络的多尺度特征。模型使用解码模块逐级提升特征图分辨率得到高分辨率输出,并将多尺度特征融合子

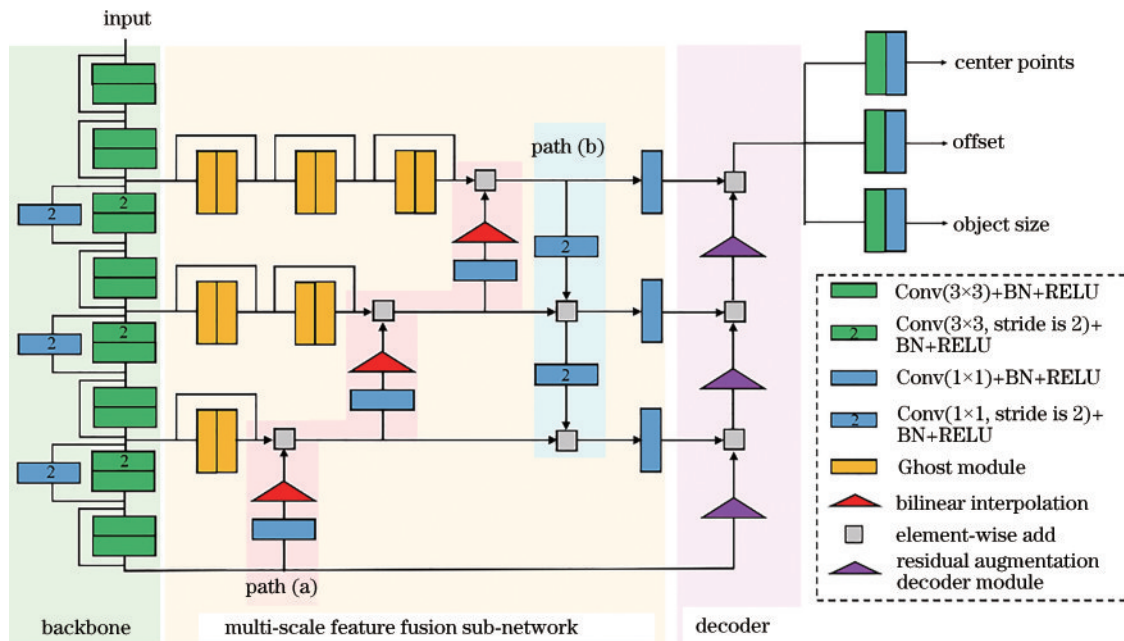


图1 所提基于Anchor-Free模型的网络结构
Fig. 1 Proposed structure based on Anchor-Free model

网络的输出作为解码子网络的增强。用 R 代表输出特征图对应原始图像的下采样步长值, n 代表解码模块最后一级的输出通道值, 最终的输出特征图 $I' \in [0, 1]^{\frac{w}{R} \times \frac{h}{R} \times n}$ 。解码子网络的输出特征图分别经过 3 组 3×3 卷积提升特征图通道数, 再分别经过 1×1 卷积调整通道得到中心点预测热力图 \hat{Y} 、中心点偏移预测图 \hat{O} 和尺度预测图 \hat{S} 。

对于每一个交通标志目标, 真实中心点在图像中的坐标可以用 p 表示。为了适应网络的输出, 真实的中心点在步长为 R 的下采样图像的坐标 $\tilde{p} = \left\lfloor \frac{p}{R} \right\rfloor$ 。设定 σ_p 为目标自适应尺度标准差, 待检测的交通标志类别数目为 c , 使用高斯核 $Y_{xyc} = \exp\left[-\frac{(x - \tilde{p}_x)^2 + (y - \tilde{p}_y)^2}{2\sigma_p^2}\right]$ 将中心点按照类别映射到真值热力图上, 其中中心点由于量化而导致的偏移量表示为 $O_p = \frac{p}{R} - \tilde{p}$, 第 k 个交通标志的真值框 (ground truth) 可以用坐标表示为 $(x_1^k, y_1^k, x_2^k, y_2^k)$, 计算得到交通标志的尺度 $s_k = (x_2^k - x_1^k, y_2^k - y_1^k)$ 。

在训练阶段, 使用 focal 损失^[15] 作为中心点的损失函数:

$$L_c = \frac{-1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \ln \hat{Y}_{xyc}, & Y_{xyc} = 1 \\ (1 - Y_{xyc})^\beta (\hat{Y}_{xyc})^\alpha \ln (1 - \hat{Y}_{xyc}), & Y_{xyc} \neq 1 \end{cases} \quad (1)$$

式中: α 和 β 为超参数, 参考文献[11], 分别设置为 2 和 4; N 为图像中的目标个数, 用于归一化。

使用 L1 损失作为中心点偏移量训练的损失函数:

$$L_o = \frac{1}{N} \sum_p |\hat{o}_p - o_p| \quad (2)$$

使用 L1 损失作为目标尺度训练的损失函数:

$$L_s = \frac{1}{N} \sum_{k=1}^N |\hat{s}_{p_k} - s_k| \quad (3)$$

最后, 总的损失函数可以表示为

$$L_d = L_c + \lambda_o L_o + \lambda_s L_s, \quad (4)$$

式中: λ_o 和 λ_s 分别表示中心点偏移量和目标尺度损失函数的权重系数, 参考文献[11], λ_o 设置为 1, λ_s 设置为 0.1。

2.2 残差增强解码模块

网络输出热图依赖于解码子网络的最后一级输出, 因此增强解码模块的通道和空间特征表达能力对检测的准确性至关重要。原始的解码模块由 1 个上采样模块和 1 个 3×3 卷积层组成, 分别用于提高特征图分辨率和减少通道数。改进的解码模块如图 2 所示, 本实验组加入残差分支将增强的通道和空间信息补充到原始解码分支来增强解码模块中的特征表示能力。受文献[16]启发, 在残差分支中, 每级解码模块的输入特征首先经过全局池化操作, 再经过 1×1 卷积, 得到自适应通道信息权重, 并与输入特征相乘。而后经过 1×1 卷积层减少残差分支的通道数, 感兴趣的通道信息在经过 1×1 卷积层时得到更好保留。考虑到解码模块采用的上采样插值引起的混叠效应^[17], 使用空间域掩模对上采样后的输入特征图进行处理。空间域掩模聚合通道信息, 自适应强调前景信息并抑制背景信息。最后, 残差分支生成后, 通过求和运算与原分支合并, 得到解码模块输出。所提残差增强解码模块与原始的解码模块相比, 在解码过程中更充分地保留和突出输入特征中的感兴趣特征信息, 最终提高了解码子网络输出热图质量。

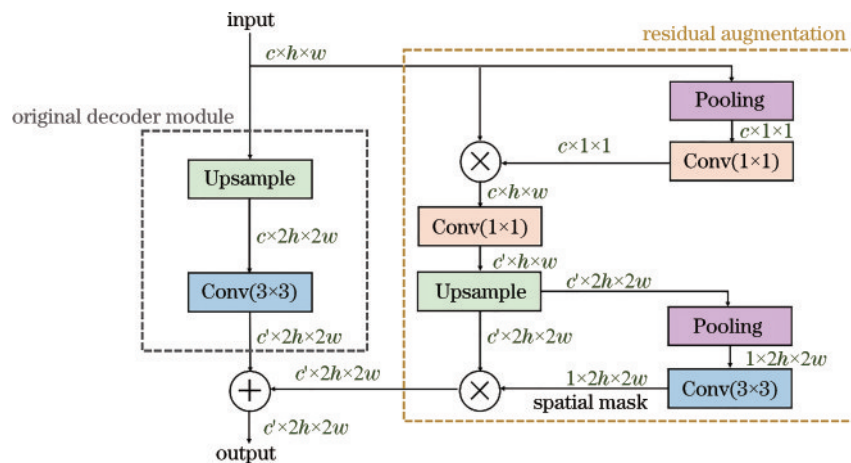


图 2 残差增强解码模块

Fig. 2 Residual augmentation decoder module

2.3 多尺度特征融合子网络

2.3.1 多尺度特征融合子网络设计

传统的编码-解码网络在编码子网络与解码子网

络的同尺度特征图之间插入跳层结构, 以利用主干网络中的多尺度特征^[18]。主干网络中浅层特征具有丰富的细节信息, 深层特征具有较强的语义信息, 利用主干

网络中产生的多尺度特征能获得更准确的定位结果。然而,在传统跳层结构中,解码子网络的最后一层输出特征图与最浅层的编码子网络特征图直接连接,这样的设计会导致输出特征图掺杂浅层特征图的背景噪声信息,降低了最终检测热图质量。因此,本实验组在编码和解码子网络之间引入多尺度特征融合子网络,如图 1 所示。首先,该子网络根据主干网络多级特征图的特征层次使用不同个数的特征提取模块进行特征的进一步提取,保持特征图分辨率不变,加深网络层数,减少浅层特征中背景噪声信息的影响。其次,在特征提取网络中,深层特征具有较强的语义信息,浅层特征具有较丰富的细节信息。受到文献[19]的启发,逐级引入从深层特征图到浅层特征图的特征融合通路,用于增强较浅层特征图的语义信息,如图 1 中的 path(a),而后引入从浅层特征图到深层特征图的特征融合通路,如图 1 中的 path(b),用于增强较深层特征图的定位信息。将含有多尺度特征信息的多分辨率特征用于增强解码路径,增加解码特征的多样性,最终提升对交

通标志的定位和分类效果。

2.3.2 多尺度特征提取的轻量化实现

在特征融合子网络中,使用普通卷积模块提取多尺度特征会在网络中引进较高的计算开销。近期的一些工作,如文献[20]和文献[21]证明了网络中同级特征之间一些特征图具有较高的相似性,主要体现在相似的背景信息和相似的轮廓信息等。考虑到这些相似性导致的特征冗余,在多尺度融合子网络中,采用 Ghost bottleneck^[20]对多尺度特征进行提取以减少计算成本。Ghost bottleneck 由 Ghost module 组成, Ghost module 的结构如图 3 所示。在 Ghost module 中,输入特征经过卷积层得到通道数为 a 的内在特征图 F ,通过内在特征图,生成输出特征图 F' :

$$f'_{ij} = \Phi_{i,j}(f_i), \quad \forall i = 1, \dots, a, j = 1, \dots, b, \quad (5)$$

式中: f_i 表示内在特征图 F 中第 i 个通道的特征; f'_{ij} 表示由 f_i 生成的第 j 个输出 Ghost 特征; $\Phi_{i,j}$ 表示线性运算,用于生成 f'_{ij} 。对于每一个 f_i , 经过线性运算后,能够生成一个或多个 Ghost 特征 $\{f'_{ij}\}_{j=1}^b$ 。

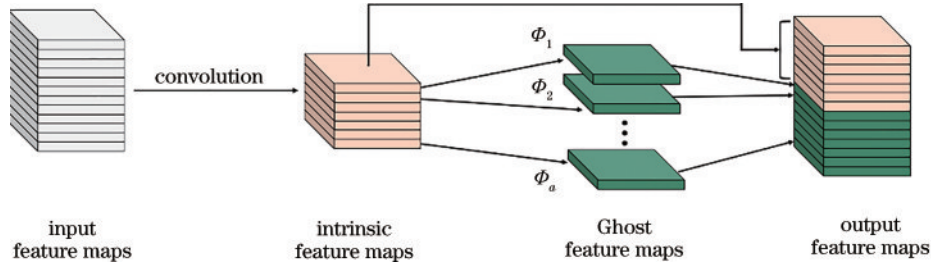


图 3 Ghost module 结构

Fig. 3 Structure of Ghost module

Ghost bottleneck 结构如图 4 所示。第 1 个 Ghost module 用于扩展通道数量,利用输入特征生成高通量特征图。第 2 个 Ghost module 用于减少通道,压缩多通道特征信息,并使通道数与跳层连接通道数匹配。

本实验组在多尺度融合子网络中使用 Ghost bottleneck 模块,进一步提取从主干网络中得到的中间特征,并减少由于中间特征图内部的特征冗余引进的计算开销。

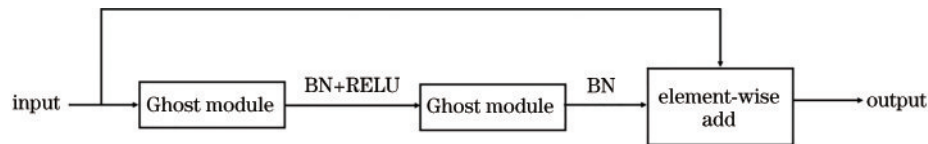


图 4 Ghost bottleneck 结构

Fig. 4 Structure of Ghost bottleneck

3 实验结果分析

3.1 交通标志检测数据集

使用交通标志数据集 Tsinghua-Tencent 100 K (TT100K)^[22]进行训练和验证。TT100K 数据集的图像采集自中国街景全景图,分辨率为 2048 pixel×2048 pixel,图像与真实驾驶环境下驾驶员视角相似,包含 6105 张训练集图像和 3071 张验证

集图像,包含真实驾驶环境下不同亮度、不同场景及不同遮挡程度的情况。遵循文献[22]的设置,在训练和验证时,忽略实例出现频次少于 100 的类别,使用数据集中出现频次大于 100 的 45 类交通标志。图 5 为上述 45 个类别交通标志的实例分布,该数据集在每个类别的交通标志实例数量上存在明显的不均匀分布,使得检测模型在数据集上取得良好的检测效果具有挑战性。

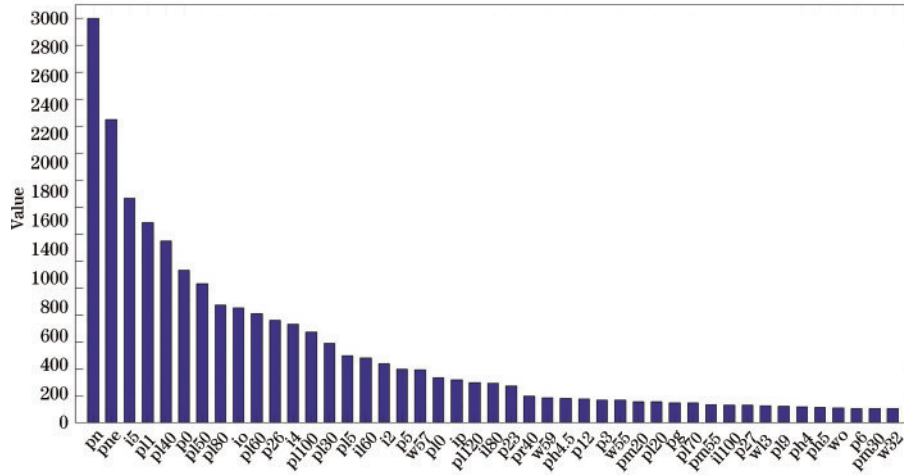


图 5 TT100K 中出现频次大于 100 的交通标志频次分布图

Fig. 5 Frequency distribution of traffic signs with frequency greater than 100 in TT100K

3.2 实验设置

使用 NVIDIA 1080 Ti 显卡, 在 Pytorch0.4.1 框架下搭建网络模型。在训练阶段, 根据交通标志实例位置, 将 TT100K 训练集中原始高分辨率图像裁剪为 $512 \text{ pixel} \times 512 \text{ pixel}$ 的子图, 对整个网络参数进行随机初始化, 使用 Adam 作为训练优化器, 初始学习率为 0.00125, 批次设置为 8。在验证阶段, 使用 TT100K 验证集中原始高分辨率图像来评估实验结果。根据文献[22]采用的衡量标准, 使用召回率 R 及

准确率 A 作为评价指标, 将交并比 (IoU) 阈值设置为 0.5。同时, 在消融实验中增加了平均精确率 (AP) 的对比。

3.3 实验结果分析

将所提方法与当前先进的目标检测方法 Faster R-CNN^[5]、YOLOv3^[23]、YOLOv4^[24]、文献^[22]中的方法、CenterNet^[11]中以 ResNet-101 为主干的模型及以 Hourglass 为主干的模型进行了对比, 实验结果如表 1 所示。

表 1 不同检测方法在 TT100K 数据集上的性能对比

Table 1 Comparison results of detection performance for different sizes of traffic signs on TT100K dataset

Method		All / %	Small / %	Medium / %	Large / %	Params / 10^6	Model size / Mbit	FLOPs / 10^9
Faster R-CNN ^[5]	R	86	77	91	92	134	518	1267
	A	82	73	88	86			
YOLOv3 ^[23]	R	92	87	95	91	62	237	802
	A	85	80	89	82			
YOLOv4 ^[24]	R	92	87	95	92	64	214	728
	A	88	81	92	86			
Zhu et al ^[22]	R	91	87	94	88	81	418	
	A	88	82	91	91			
CenterNet-Res101 ^[11]	R	91	85	95	94	53	214	768
	A	89	78	94	92			
CenterNet-Hourglass ^[11]	R	94	89	96	94	191	765	4672
	A	90	83	95	93			
Proposed method	R	93	88	96	95	16	64	320
	A	90	83	95	93			

采用的衡量指标除了召回率及准确率外, 还有模型参数量 (Params)、模型大小及 floating point of operations (FLOPs), 认为真值框大小在 $(0, 32 \times 32)$ 区间的实例是小目标, 在 $[32 \times 32, 96 \times 96]$ 区间的实例是中目标, 不小于 96×96 的实例是大目标。所提方法的模型参数量、模型大小和 FLOPs 分别为 $1.6 \times$

10^7 、64 Mbit 和 3.2×10^{11} 。在计算开销小于其他检测器的情况下, 考虑所有尺度的交通标志, 所提方法实现了 92% 的召回率和 90% 的准确率。对于同样使用 Anchor-Free 网络模型的检测器, 所提方法的模型大小和参数量远小于使用 Hourglass 主干的 CenterNet 模型, 代价仅是准确率降低了 1 个百分点, 并且保持了在

不同尺度目标上的检测性能。实验结果表明,所提方法在多尺度目标检测和计算开销上具有很大的优势。

为了验证所提改进模块的有效性,以采用 ResNet-18 主干网络的编码-解码模型为基准(Base),分别加入残差增强解码模块(RADM)、多尺度特征融合子网络(MFFN),对比检测结果的召回率、准确率及 AP,结果如表 2 所示。通过在解码模块中引入残差增强分支,模型的参数量和模型大小相比基准模型分别增加了 0.3×10^6 和 0.8 Mbit,而召回率、准确率和 AP 有显著提升,分别为 1.5 个百分点、1.4 个百分点和 2.1 个百分点。实验结果表明,残差增强的解码模块相比原始解码模块在解码过程中具有更好的

特征表征能力,能够产生质量更高的输出特征图,提升最终检测性能。在模型中加入多尺度融合子网络后,模型比基准模型的召回率、准确率和 AP 分别提升了 2.8 个百分点、2.2 个百分点和 2.9 个百分点。由于采用 Ghost 轻量化模块作为多尺度特征提取模块,多尺度融合子网络仅比基准模型增加了 0.4×10^6 的参数量。同时加入残差增强解码模块和多尺度融合子网络后,所提模型相对于基准模型在召回率、准确率和 AP 上分别有 3.2 个百分点、3.1 个百分点和 3.8 个百分点的提升,相对于基准模型仅增加了 4.5% 的参数量,充分说明了所提模块在引入较小的计算开销下实现了较大的性能提升。

表 2 不同模块在 TT100K 数据集上的消融实验

Table 2 Results of ablation experiments of various components on TT100K dataset

Model	R / %	A / %	AP / %	Params / 10^6	Model Size / Mbit	FLOPs / 10^9
Base	89.3	87.2	89.8	15.4	62.2	295.4
Base+RADM	90.8	88.6	91.9	15.7	63.0	296.9
Base+MFFN	92.1	89.4	92.7	15.8	63.6	318.4
Base+RADM+MFFN	92.5	90.3	93.6	16.1	64.4	320.1

所提模型与其他检测模型的精确率-召回率(P-R)曲线如图 6 所示。从图中可以看出,所提模型可以取得和 CenterNet-Hourglass 模型相近的性能。在漏检率较低的情况下,所提模型相对于基准模型在小、中、大等 3 个尺度目标的检测精确率都有提升。

图 7 为所提模型在 TT100K 测试集中的部分可视化结果,展示的图像包含了视角变化、远距离、近距离和目标密集出现等情况。第 1 行图像为交通标志检测结果,第 2 行图像为与第 1 行图像对应的中心点热力图,为便于观察,将图像中存在交通标志的区域放大后

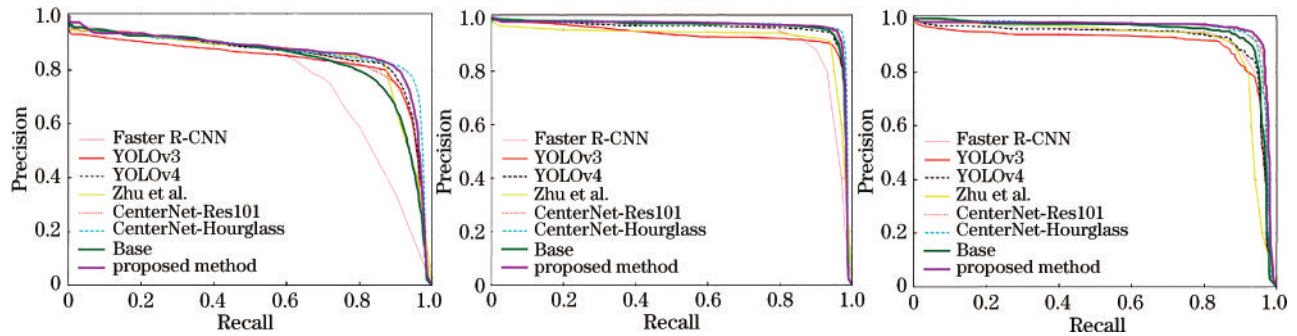


图 6 不同尺度下交通标志的 P-R 曲线

Fig. 6 P-R curves for traffic signs of different sizes

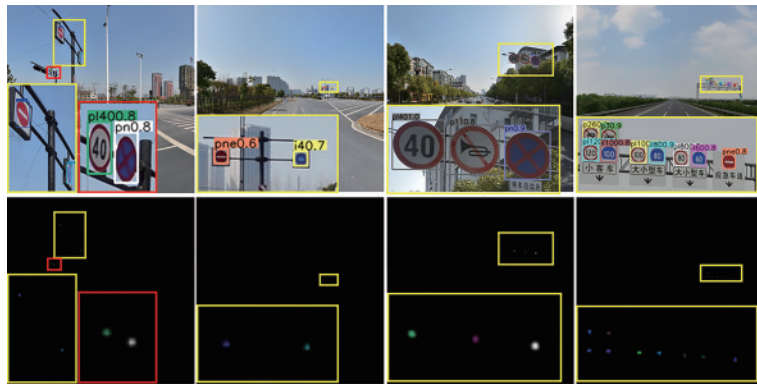


图 7 可视化结果

Fig. 7 Visualization results

置于图像左下角。可视化结果表明,在道路交通背景下,所提模型能够准确地检测交通标志。

4 结 论

提出了一种基于编码-解码结构的改进 Anchor-Free 模型的交通标志检测算法。首先,为了增强解码器中的特征表达并获得更高质量的输出特征图,在原始编码模块中引入残差增强分支,在通道变化过程中自适应提高感兴趣通道的权重,并减少上采样带来的混叠效应。其次,为了提取和融合多尺度特征,设计了多尺度特征融合子网络,充分利用主干网络中的多尺度特征。考虑到同级别特征之间具有较高的冗余性,使用 Ghost 模块对中间特征进行进一步提取,Ghost 模块使用内在特征图,经过线性操作可还原高通道特征图,以减少计算开销。同时,在特征融合子网络中使用信息融合通路充分融合语义信息和细节信息,进一步提升交通标志检测性能。实验结果表明,与目前先进的目标检测方法相比,所提方法在较小的计算开销下,实现了较高的检测精度,具有优越的综合性能。

参 考 文 献

- [1] de la Escalera A, Moreno L E, Salichs M A, et al. Road traffic sign detection and classification[J]. IEEE Transactions on Industrial Electronics, 1997, 44(6): 848-859.
- [2] Zaklouta F, Stanculescu B. Real-time traffic-sign recognition using tree classifiers[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(4): 1507-1514.
- [3] Yao C, Wu F, Chen H J, et al. Traffic sign recognition using HOG-SVM and grid search[C]//2014 12th International Conference on Signal Processing (ICSP), October 19-23, 2014, Hangzhou, China. New York: IEEE Press, 2014: 962-965.
- [4] Li L M, Li L, Tong R Q, et al. Feature detection and matching for traffic sign images[C]//Proceedings of the 10th World Congress on Intelligent Control and Automation, July 6-8, 2012, Beijing, China. New York: IEEE Press, 2012: 4628-4632.
- [5] Ren S Q, 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 and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [7] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer Vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.
- [8] Cao C Q, Wang B, Zhang W R, et al. An improved faster R-CNN for small object detection[J]. IEEE Access, 2019, 7: 106838-106846.
- [9] Zhang H B, Qin L F, Li J, et al. Real-time detection method for small traffic signs based on Yolov3[J]. IEEE Access, 2020, 8: 64145-64156.
- [10] Jin Y M, Fu Y S, Wang W Q, et al. Multi-feature fusion and enhancement single shot detector for traffic sign recognition[J]. IEEE Access, 2020, 8: 38931-38940.
- [11] Zhou X Y, Wang D Q, Krahenbuhl P, et al. Objects as points[EB/OL]. (2019-04-16)[2021-04-05]. <https://arxiv.org/abs/1904.07850>.
- [12] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [13] Wei L J, Xu C, Li S Q, et al. Traffic sign detection and recognition using novel center-point estimation and local features[J]. IEEE Access, 2020, 8: 83611-83621.
- [14] Ding T H, Feng K L, Li T P, et al. An improved anchor-free detection method for traffic sign detection [C]//2021 2nd International Symposium on Computer Engineering and Intelligent Communications (ISCEIC), August 6-8, 2021, Nanjing, China. New York: IEEE Press, 2021: 348-351.
- [15] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [16] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [17] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [18] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [19] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 8759-8768.
- [20] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1577-1586.
- [21] Zhang Q L, Jiang Z Q, Lu Q S, et al. Split to be slim: an overlooked redundancy in vanilla convolution[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, July 11-17, 2020, Yokohama, Japan. California: International Joint

- Conferences on Artificial Intelligence Organization, 2020: 3195-3201.
- [22] Zhu Z, Liang D, Zhang S H, et al. Traffic-sign detection and classification in the wild[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2110-2118.
- [23] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2021-04-05]. <https://arxiv.org/abs/1804.02767>
- [24] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2021-04-03]. <https://arxiv.org/abs/2004.10934>.