

基于改进 YOLOv5 的车辆属性检测 *

刘俊, 钟国韵, 黄斯雯, 刘麒麟

(东华理工大学 信息工程学院, 江西 南昌 330013)

摘要: 车辆属性检测是一个基础任务, 其属性检测结果可以被应用到很多下游的交通视觉任务。提出了一种基于 YOLOv5 的车辆属性检测改进算法。针对检测目标较小的问题, 加入了卷积注意力模块, 让网络模型把更多的注意力放在小目标对象上; 针对数据集样本种类较少的问题, 改进了 YOLOv5 的马赛克数据增强方式; 使用自门控激活函数 Swish, 起到抑制噪声、加快收敛速度并提升模型鲁棒性的作用。此外, 还在公开车辆数据集 VeRi-776 的基础上进行了详细的车辆属性标注, 构建了一个车辆属性数据集。实验结果表明, 改进后的算法比原始 YOLOv5 的平均精确率提升了 4.6%, 能够准确地检测到车辆图像的通用属性, 可以供下游任务使用。

关键词: 车辆属性; 目标检测; YOLOv5

中图分类号: TP391; TP183

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.222802

中文引用格式: 刘俊, 钟国韵, 黄斯雯, 等. 基于改进 YOLOv5 的车辆属性检测[J]. 电子技术应用, 2022, 48(7): 19-24, 29.

英文引用格式: Liu Jun, Zhong Guoyun, Huang Siwen, et al. Vehicle attribute detection based on improved YOLOv5[J]. Application of Electronic Technique, 2022, 48(7): 19-24, 29.

Vehicle attribute detection based on improved YOLOv5

Liu Jun, Zhong Guoyun, Huang Siwen, Liu Qilin

(School of Information Engineering, East China University of Technology, Nanchang 330013, China)

Abstract: Vehicle attribute detection is a basic task, which can be applied to many downstream traffic vision tasks. This paper presents an improved vehicle attribute detection algorithm based on YOLOv5. Aiming at the problem of small target detection, this paper adds the convolution attention module to make the network model pay more attention to the small target object. Aiming at the problem of less sample types of the dataset, this paper improves the mosaic data enhancement method of YOLOv5. The self-gated activation function Swish is used to suppress noise, accelerate convergence speed, and improve the robustness of the model. In addition, this paper also makes a detailed vehicle attribute labeling based on the public vehicle dataset VeRi-776, and constructs a vehicle attribute dataset. The experimental results show that the average accuracy of the improved algorithm is 4.6 % higher than that of the original YOLOv5, which can accurately detect the general attributes of vehicle images and can be used for downstream tasks.

Key words: vehicle attribute; object detection; YOLOv5 algorithm

0 引言

目前计算机视觉研究者们在逐渐探索车辆图像数据处理的落地应用, 以助力智慧交通。车辆属性检测就是其中一个基础的计算机视觉任务, 主要检测车辆的车灯、车牌、车辆 logo 等车辆属性区域位置, 其结果可以被应用到很多下游的交通视觉任务。例如利用检测到的车灯属性来判定车辆是否变道打转向灯; 车牌属性可以作为车牌 OCR 识别的输入, 还可结合整体的车辆属性, 来实现车辆重识别等^[1]。

由于早期显卡的显存和计算能力的限制, 神经网络无法设计得很深, 导致当时目标检测的研究还是偏向于

传统的图像处理, 主要可以分为区域选择、特征提取和分类三步。区域选择一般是通过在图像上进行逐块像素的遍历, 来找到与目标匹配的区域; 特征提取则是依据研究员的相关先验知识, 如待检测目标的形状、纹理、明暗颜色等, 从上一步获取的区域中提取特征, 代表算法有尺度不变特征变换(SIFT)^[2]和方向梯度直方图(HOG)^[3]; 最后是训练分类器将特征分类, 主要利用将待分类数据的特征向量映射到高维空间, 以实现将不同类别的数据分开, 经典算法有支持向量机(SVM)^[4]和 AdaBoost^[5]。

深度学习^[6]开始流行后, 基于卷积神经网络的方法逐渐在计算机视觉领域占据主导地位。具体到目标检测任务, 检测的实时性不断提高, 速度能够超过 25 f/s, 准确率也在逐渐上升, 已经可以满足落地使用的要求。现

* 基金项目: 国家自然科学基金(61662002)

有的目标检测算法包括两类:One-Stage 和 Two-Stage^[7]。One-Stage 类的算法是一个端到端的网络直接对物体进行分类以及回归出 Bounding box 的坐标,现在使用最广泛的是 YOLO 系列算法^[8-10]和 SSD;Two-Stage 类的算法分为两步,需要先选出最有可能的区域作为候选区,然后再对这些区域进行修正,这种算法的准确率比前者要高一些,代表算法是 Faster R-CNN。

Zhao Yanzhu 等人在 SSD 的基础上,增加了选择更多的底层候选区域,并且将损失函数换成了 focal loss 以提高车辆局部属性区域的检测^[11],但是工程性、易用性和准确率还有待提高。本文采用最新的 YOLOv5 对车灯、后视镜等车辆的局部零件区域位置进行检测,由于它们相对较小,为了让模型把更多的注意力放在小目标对象上,加入了卷积注意力模块;由于构建的数据集样本种类相对较少,为了增强模型的泛化能力,改进了 YOLOv5 的马赛克数据增强方式;最后,本文把激活函数改为自门控激活函数 Swish,以抑制噪声并且加速网络训练。

1 YOLOv5 网络简介

YOLOv5 的网络架构整体上与 YOLOv4 非常接近,由主干网络、颈部网络和输出三部分组成,具体结构图如图 1 所示。

主干网络是用来进行提取图像底层特征的,主要包括 Focus 结构和跨阶段层次结构(CSP)^[12]。Focus 结构中主要是一个切片操作,具体流程为:对图像在每个

通道进行均匀划分,将划分后每个图像块对应位置的像素值拼成一个新的通道。所以操作完成后会增加很多新通道,这样可以让模型学习到更多图像的局部特征。考虑到计算瓶颈和内存成本,CSP 模块通过先把底层特征图分为两部分,然后通过跨阶段层次结构再将它们合并的方式,既大大降低了计算量,又保证了准确率^[13]。

颈部网络是利用特征金字塔的架构,将多个下采样(图 1 网络结构图中的 4、6、9 号位置)得到的不同尺度的特征图,通过上采样的方式将高层特征信息流通并融合起来。输出模块会输出相对于原图的 8、16、32 倍下采样的 3 个尺度特征图,采用多个尺度是因为考虑到检测的物体大小不一,这样不会造成小目标的检测率过低。输出的预测结果包括 Bounding box 的坐标、坐标的置信度分数和预测为某类别的概率。

与之前的版本相比,在网络架构方面一个很大的改进是,YOLOv5 通过调节深度倍数(depth multiple)和宽度倍数(width multiple)两个参数,可以得到不同大小的网络,这样便于根据自己具体的需求来调整网络参数的大小。

2 改进 YOLOv5 算法

2.1 CBAM 注意力模块

人类每天都要接收大量的视觉信息,其实它们已经超过了大脑所能处理的量,但是人类在进化中学会了将

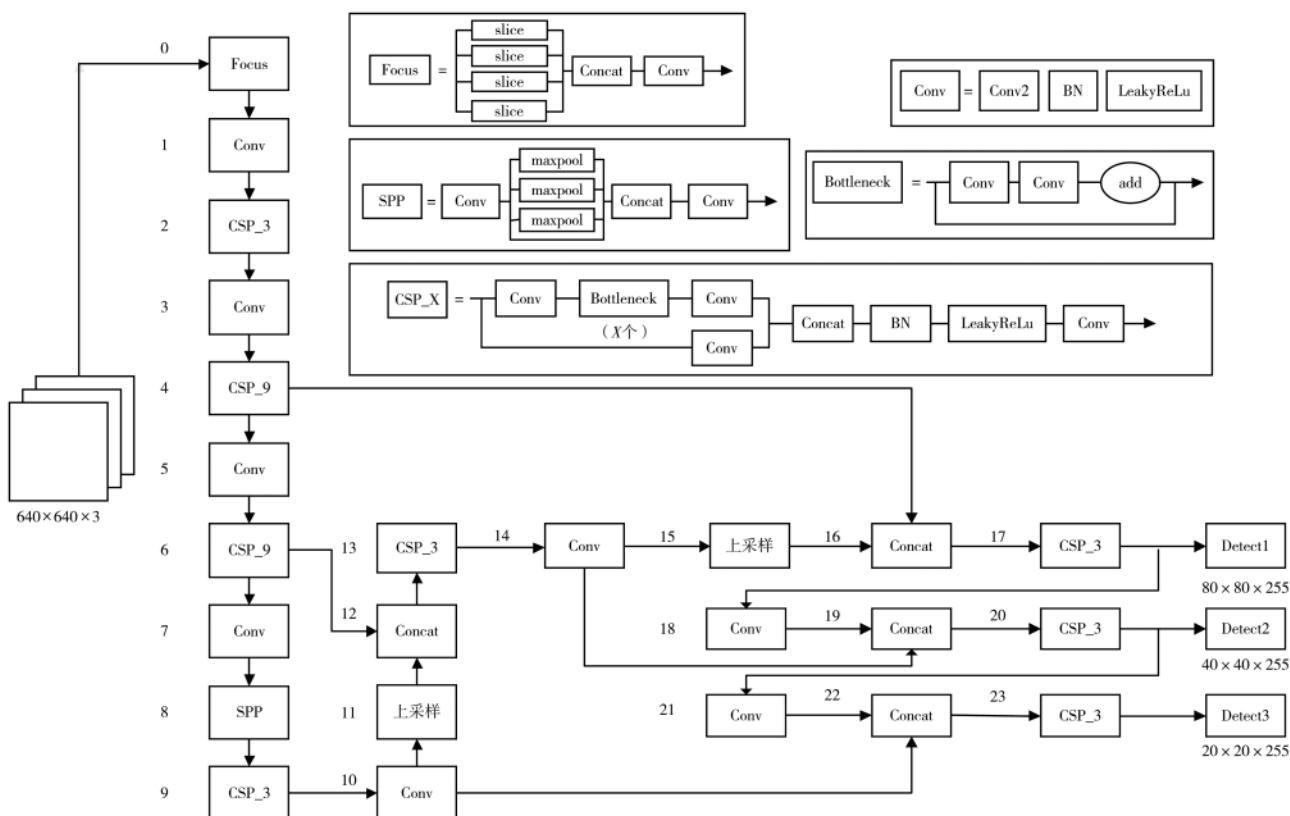


图 1 YOLOv5 网络结构

有限的注意力更多地分配到感兴趣区域。例如人类的眼睛会自动关注视野内运动或者发光的物体。这些往往是输入到大脑中信息量比较多的区域。网络模型也是如此,需要把有限的注意力放在信息量大的地方。

对于输入图像,车辆的通用目标只占整体的较小部分,大部分为背景信息,在经过多次的卷积操作后,会产生大量冗余信息,逐渐将小目标对象的信息淹没,导致网络检测不到部分小目标,如年检标志、后视镜等。因此,本文将卷积块注意力模块(CBAM)嵌入到YOLOv5中,它通过结合通道和空间的注意力计算分配,让网络更多地关注到应该关注的小目标对象。图2是CBAM模块的结构,输入其中的特征图依次通过通道和空间注意力模块,分别进行通道与空间上的注意力计算^[14]。

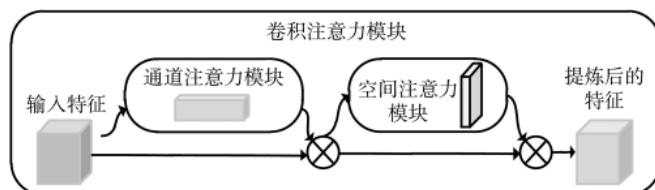


图2 卷积注意力模块

在通道注意力模块(结构如图3所示)中,其操作具体如式(1)所示:

$$M_c(F) = \text{Sigmoid}(\text{MLP}(\text{AvgPool}(F) + \text{MLP}(\text{MaxPool}(F))) \quad (1)$$

其中,MLP()表示通过一个多层次感知机; F 表示输入的特征; AvgPool()表示平均池化操作; MaxPool()表示最大池化操作; Sigmoid()表示将数据输入 Sigmoid 函数中,将其映射到 0~1 之间。

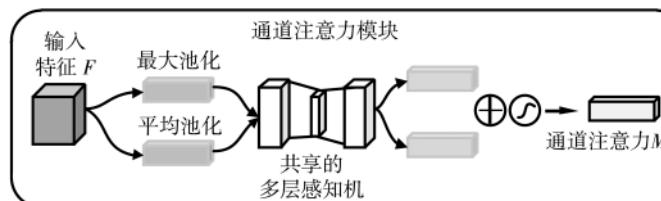


图3 通道注意力模块

在空间注意力模块(结构如图4所示)中,其具体操作如式(2)所示:

$$M_s(F) = \text{Sigmoid}(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)]) \quad (2)$$

其中, $f^{7 \times 7}$ 表示 7×7 的卷积操作。

虽然池化操作会造成一定程度的细粒度信息的丢失,但是这两个模块均使用了 Global Max Pooling 和 Global Average Pooling,这样可以减少信息的丢失,而且丰富了网络能够学习到的信息。

本文把CBAM插入到YOLOv5的颈部,具体方式为在图1所示YOLOv5网络结构图中标识的4、6、9号3个CSP模块内部的最后加上CBAM模块,这样能更好地提

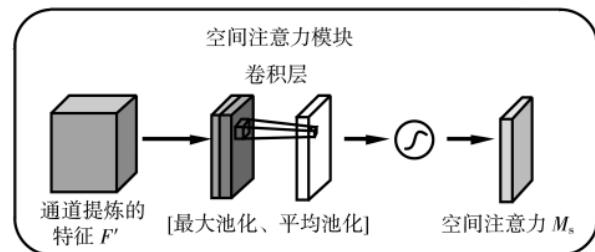


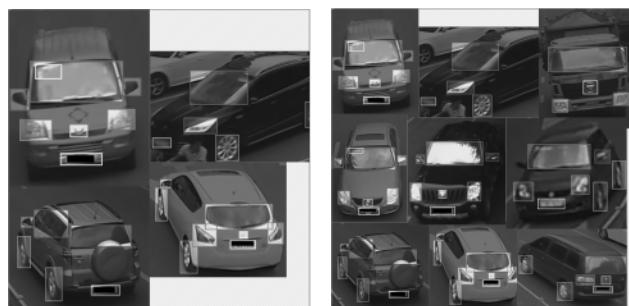
图4 空间注意力模块

升模型检测小目标的性能。

2.2 改进马赛克数据增强

马赛克(Mosaic)数据增强是YOLOv5中很重要的一种数据增强方式。即在载入一张图片时,再额外从数据集中任意选择3张图片,先通过随机裁剪,再顺时针拼接在一张图片上,最后缩放为设定的输入尺寸,并作为新的样本传入模型。这样既丰富了待检测目标的背景,又增加了每次训练每批次中小目标的数量,实现了不同尺度目标间的平衡。

由于本文构建的车辆通用目标数据集中只有7个类别,不同的车辆总共只有776,与一些标准数据集(例如含有80个类别,超过20万张图片的COCO数据集)相比,数据过少,因此需要对现有的数据进行数据增广以提高模型的泛化能力。又因为原始马赛克数据增强方式是随机裁剪的,会有很大概率将目标裁剪,使输入模型的样本只有背景;另外原始图片本身尺度也不一致,这样会使拼接后的图片有较多黑白的边界,从而会导致模型训练大量无用的特征信息,影响了模型收敛的速度,如图5(a)所示。因此,本文改进了马赛克数据增强方式。在载入图片时,再额外选择8张图片进行拼接,再按包围图片的最小矩形面积裁剪,然后对图片增加随机旋转、平移、缩放、视角变换操作。如图5(b)所示,改进后的马赛克数据增强方式比改进前一方面减少了很多空白的边界,减少了无用信息,加快模型收敛,提高了训练效率;另一方面会形成大量的不同小目标,极大地丰富了车辆通用目标数据集,显著地改善了模型检测小尺寸目标(如车标、后视镜等)的性能^[15]。



(a)原始马赛克增强图

(b)改进后的马赛克增强

图5 马赛克增强方式对比

2.3 Swish 激活函数

激活函数在神经网络中有着至关重要的功能,它可以为网络增强非线性表达能力,让神经网络能够拟合任意的复杂函数,在YOLOv5网络架构中的卷积模块(Conv)中使用了带泄露的ReLU激活函数(Leaky ReLU),其表达式如式(3)所示:

$$f(x) = \max(\alpha, x) \quad (3)$$

其中, α 为常数,一般设置为 0.01。

由于 Leaky ReLU 在不同区间函数不一样,因此结果不一致,导致无法为正负输入值提供一致的关系预测。Ramachandran 等^[16]提出了一种效果更好的激活函数,称为 Swish,其表达式如式(4)所示,图像如图 6 所示。

$$f(x) = x \cdot \text{sigmoid}(x) \quad (4)$$

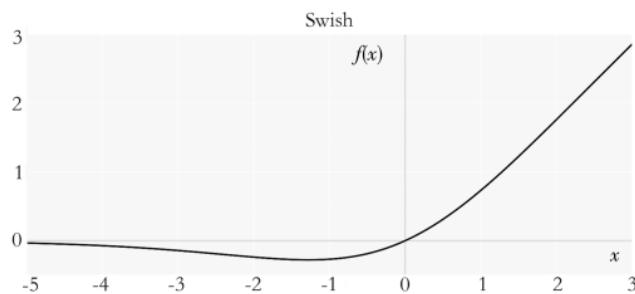


图 6 Swish 激活函数图像

Swish 函数有以下特点:没有上边界,因此不会出现梯度饱和的现象;有下边界,当 x 越靠近左半轴,纵坐标的值越小,甚至接近于 0,如果 x 值是 -5,那么经过激活之后的值接近于 0,那么就可以一定程度上过滤掉一部分信息,起到正则化的效果;非单调,处处连续且可导,更容易训练。本文在 YOLOv5 网络架构中引入 Swish 激活函数代替原有的 Leaky ReLU。

3 模型训练过程

3.1 实验数据集

由于缺乏高质量开源的车辆属性数据集,本文在公开车辆数据集 VeRi-776^[17]的基础上进行车辆属性标

注,共 776 个不同车辆,49 357 张图片,进行了 7 类标注,具体包括车标、车轮、挡风玻璃、车牌、车灯、后视镜、年检标志。本实验通过在线标注平台 Roboflow 进行标注,然后导出为 YOLOv5 所使用的格式,标注示例如图 7 所示。

3.2 训练过程

实验环境的具体配置为 Linux Ubuntu18.04.7 LTS, CPU 为 Intel® Xeon® Gold 5118, 显卡用的是 16 GB 显存的 Tesla T4, 深度学习框架选择的是 PyTorch, 版本为 1.10。采用 YOLOv5 4 种模型中最小的 YOLOv5s 模型以方便后续训练,因为它的参数量小,推理速度快。将标注好的数据集的 60% 作为训练集,20% 作为验证集,20% 作为测试集进行实验。初始学习率为 0.01, 终止学习率为 0.2, 批处理量(batch size) 为 64, 训练轮数(epochs) 为 300 轮, 训练和测试图像通过等比缩放并用填充像素“0”的方式,调整为 640×640 尺寸。

对于锚框,不是采用 YOLOv5s 默认的锚框大小,而是使用通过 K-means 聚类算法得到最适合本实验的训练数据集的锚框设置。

3.2.1 迁移学习

为了增强模型的鲁棒性和加快训练的速度,利用迁移学习,在基于 COCO 数据集上训练好的模型基础上对网络模型进行微调。具体地,在训练时将类别数量修改为 7,并冻结住 backbone 中的前 5 层,这样可以利用预训练模型在 COCO 数据集上学习到的大量识别底层信息的能力,并且加快训练速度。

3.2.2 评价指标

本文采用精确率(Precision)、召回率(Recall)和平均精确率(mAP)作为模型性能评价的相关指标。精确率用来衡量模型检测的是否准确,召回率则是用来判断模型检测结果是否全面。单个类别的精确率(AP)是通过 PR 曲线以及坐标轴所围成的面积得到的。mAP 是相对所有类别的平均值,通常认为模型检测输出的 Bounding box 与 Ground Truth 的 IOU=0.5 时的检测结果是正确的,此时计算 mAP 值,就表示为 mAP@0.5。其中 IOU 为交并

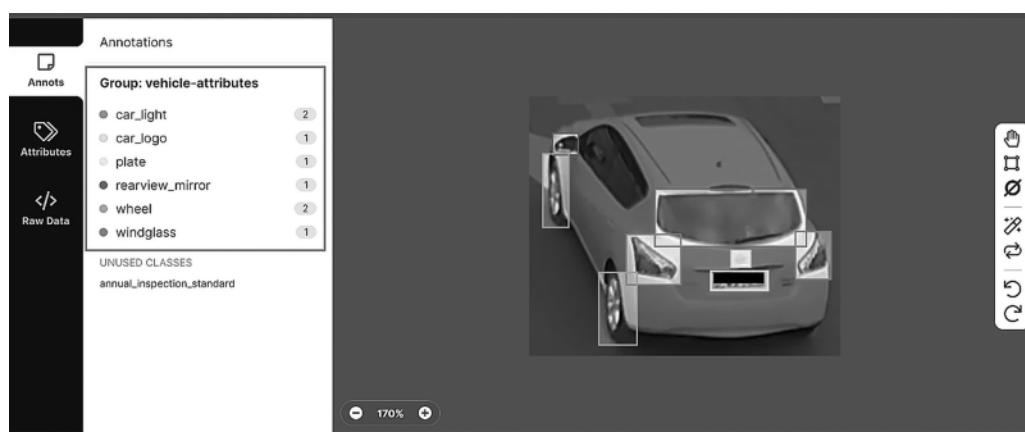


图 7 标注示例

比,是计算 mAP 的重要函数,见式(5):

$$IOU = \frac{C \cap G}{C \cup G} \quad (5)$$

其中,C、G 分别为预测框和真实框,分子为两框交集的面积,分母为两框并集面积。

式(6)~式(8)分别是精确率、召回率和单个类别的精确率的计算公式:

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

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

$$AP = \int_0^1 p(r)dr \quad (8)$$

其中,TP 表示预测为正类的正样本,FP 表示预测为正类的负样本,FN 为预测为负类的正样本。

本文不是直接求 PR 曲线的积分结果,而是采用一个平滑处理。也就是曲线上的每一个点,精度值取它右边精度的最大值。用公式描述为:

$$P_{smooth}(r) = \max_{r' \geq r} P(r') \quad (9)$$

$$mAP = \frac{\sum_{i=1}^c AP_i}{C} \quad (10)$$

其中,C 为类别数,AP_i 表示第 i 个类别的平均精确率。

训练时的损失变化曲线如图 8 所示,可以看到相比原 YOLOv5 模型,改进的 YOLOv5 最终收敛时的损失值更小、收敛速度也更快,其中原 YOLOv5 在 230 次迭代后才开始收敛,改进的 YOLOv5 在 180 次迭代就开始收敛了。

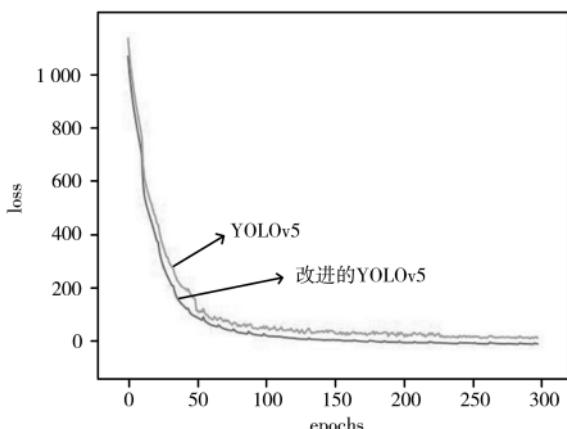


图 8 模型训练时的损失变化曲线

模型训练时的平均精度变化曲线如图 9 所示,可以看到改进的 YOLOv5 模型更快地收敛,精度表现更平稳,并且精度更高。

4 实验结果与分析

训练完成后,得到模型权重文件,然后将测试图片输入网络中进行推理,检测结果示例如图 10 所示,图 11 是注意力可视化图^[18],表 1 是车辆各属性的检测精度。

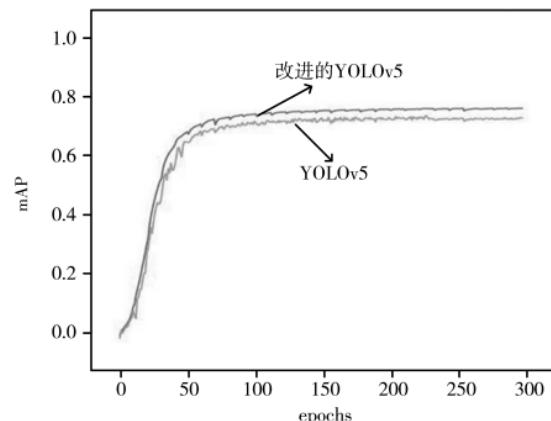


图 9 模型训练时的平均精度变化曲线



图 10 检测结果可视化



图 11 注意力图可视化

表 1 车辆各属性的检测精度

	车标	车轮	挡风玻璃	车牌	车灯	后视镜	年检标志	mAP
	66.4	76.1	85.6	82.8	79.2	74.6	55.8	74.4

可以看到,改进后的 YOLOv5 网络能够准确地检测到车辆各属性的位置,有些小目标如年检标志,由于拍摄距离较远,比较模糊,导致检测精度较低。根据注意力图可视化,可以发现网络能够更好地关注车辆本身的细节区域。

4.1 对比分析

4.1.1 检测效果对比

为了清晰直观地观察到本文改进算法的优越性,将其与原 YOLOv5 的部分检测结果进行可视化并进行对比,如图 12 所示。可以看到,原 YOLOv5 在图 12(a)中将车辆左边的后视镜这个小目标漏掉了,而图 12(b)中显示出本文改进后的 YOLOv5 模型检测到了;图 12(c)是原 YOLOv5 模型的定位结果,其输出的边界框包含了部分检测目标外的无效区域,但在图 12(d)中,可以看到本文改进后的算法检测的边界框与待检测目标重合更好,即精度更高。

4.1.2 不同网络模型的对比

为了准确评估本文改进算法的有效性,将其与 YOLOv5、YOLOv4、YOLOv3、SSD、Faster R-CNN 算法进行对比,最终得到各个模型的性能效果如表 2 所示。其中 YOLOv5 采用的是 YOLOv5s,所以推理速度极快。可以看到,最新的 YOLOv5 相比之前的目标检测经典算法在精度和检测

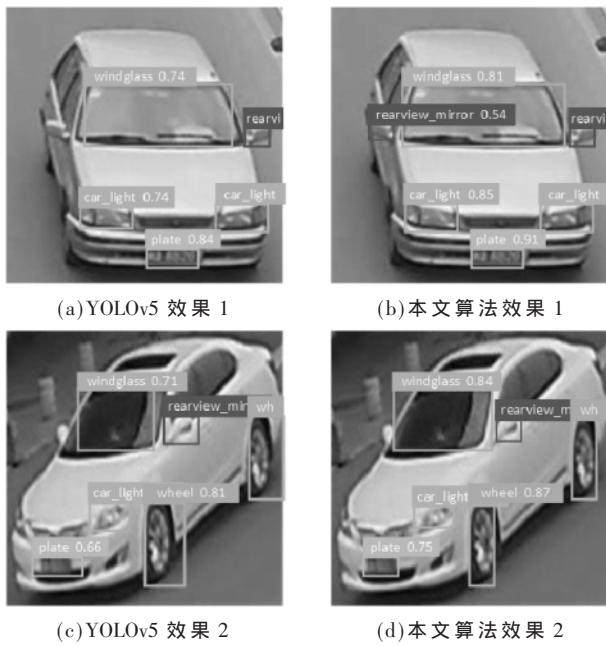


图 12 改进 YOLOv5 与原 YOLOv5 的检测效果对比

表 2 不同网络模型的对比

模型	mAP/%	FPS
Faster R-CNN	68.1	51
SSD	66.4	68
YOLOv3	67.3	76
YOLOv4	68.7	84
YOLOv5	69.8	105
本文改进的 YOLOv5	74.4	100

速度上都有提升,特别是速度方面,提升明显。本文改进的 YOLOv5 模型,相比原 YOLOv5,在轻微降低检测速度的前提下,mAP 提高了 4.6%,更具实际应用价值。

4.2 消融实验

本文在 YOLOv5 模型的基础上,增加了注意力卷积模块、改进了马赛克数据增强方式、改进了激活函数,为了评估不同模块的改动和不同模块组合对于算法性能优化的程度,本文详细地进行了消融实验,表 3 为消融实验数据。在实验环境配置一致的情况下,实验结果表明,每个改进都能提升模型的准确率,其中 CBAM 的贡

表 3 消融实验

CBAM	马赛克数据增强	Swish 激活函数	mAP/%
✓			69.8
	✓		72.5
		✓	71.1
✓	✓		70.4
✓		✓	73.8
✓	✓	✓	73.1
✓	✓	✓	71.7
✓	✓	✓	74.4

献较为明显,mAP 提升了 2.7%,3 个模块的组合应用对于最终的识别精度优化效果最好。

5 结论

本文将 YOLOv5 应用到车辆属性检测领域,提出了一种改进的 YOLOv5 算法。但是目前,自然道路场景下的车辆属性检测识别依旧面临很多问题,如监控摄像头采集的图像质量不佳;容易受雾、不良光照干扰,从而丢失目标纹理细节,对识别造成困难。对于一些车辆属性(如车标、年检标志等),由于本文是基于 VeVi-776 数据集标注的,样本较少,检测效果有待增强,后续还需要不断收集标注更多的数据。后续会基于本文的模型,构建 Pipeline,进行下游任务的研究与实现。例如在道路监控视频场景下,增加选取最优帧的逻辑,选出质量高的一帧图像,结合检测到的车牌位置,再进行车牌 OCR 识别;还有利用整体的车辆属性,进行车辆重识别,实现车辆轨迹还原等。

参考文献

- [1] 柴江云. 交通监控中车辆属性的识别方法研究[D]. 西安: 西安理工大学, 2020.
- [2] LOWE D G. Object recognition from local scale-invariant features[C]//Proceedings of IEEE International Conference on Computer Vision, 1999.
- [3] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//IEEE Computer Society Conference on Computer Vision & Pattern Recognition. IEEE, 2005.
- [4] CHEN P H, LIN C J, SCHÖLKOPF B. A tutorial on v-support vector machines[J]. Applied Stochastic Models in Business and Industry, 2005, 21(2): 111–136.
- [5] FREUND Y, SCHAPIRA R E. Experiments with a new boosting algorithm[C]//Machine Learning, 1996: 148–156.
- [6] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back propagating errors[J]. Nature, 1986, 323(6088): 533–536.
- [7] 方路平, 何杭江, 周国民. 目标检测算法研究综述[J]. 计算机工程与应用, 2018, 54(13): 11–18, 33.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Computer Vision & Pattern Recognition. IEEE, 2016: 779–788.
- [9] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Computer Vision & Pattern Recognition. IEEE, 2017: 7263–7271.
- [10] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv e-prints arXiv: 1804.02767, 2018.
- [11] ZHAO Y, SHEN C, WANG H, et al. Structural analysis of attributes for vehicle re-identification and retrieval[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, PP(99): 1–12.
- [12] 张文龙, 南新元. 基于改进 YOLOv5 的道路车辆跟踪算法[J]. 广西师范大学学报(自然科学版), 2022, 40(2): (下转第 29 页)

- of X-Ray Science and Technology, 2020, 28(6): 200740.
- [4] CAO Z T, DUAN L X, YANG G W, et al. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures[J]. BioMed Central, 2019, 19(1): 51.
- [5] 木拉提·哈米提, 周晶晶, 严传波, 等. 新疆地方性肝包虫 CT 图像的灰度直方图特征提取与分析[J]. 科技导报, 2012, 30(6): 66-70.
- [6] 周晶晶, 木拉提·哈米提, 严传波, 等. 新疆地方性肝包虫 CT 图像纹理特征的分类研究[J]. 科技通报, 2013, 29(1): 42-46, 53.
- [7] 严传波, 木拉提·哈米提, 李莉, 等. 新疆地方性肝包虫 CT 图像检索算法比较[J]. 生物医学工程学杂志, 2013, 30(5): 942-945.
- [8] 茹仙古丽·艾尔西丁. 基于深度学习的肝囊型包虫病 CT 图像分类方法研究[D]. 乌鲁木齐: 新疆医科大学, 2020.
- [9] 刘志华, 李丰军, 严传波. 卷积神经网络在肝包虫病 CT 图像诊断中的应用[J]. 电子技术应用, 2019, 45(11): 17-20.
- [10] WU M, YAN C B, WANG X R, et al. Automatic classification of hepatic cystic echinococcosis using ultrasound images and deep learning[J]. J. Ultrasound Med., 2021, 41: 163-174.
- [11] 娜迪亚·阿卜杜迪克依木. 肝囊型包虫病 CT 图像病灶检测方法研究[D]. 乌鲁木齐: 新疆医科大学, 2020.
- [12] 杜燕, 李书兵. 超声对肝包虫病的筛查及诊断价值[J]. 临床超声医学杂志, 2019, 21(12): 957-958.
- [13] 阮激扬. 基于 YOLO 的目标检测算法设计与实现[D]. 北京: 北京邮电大学, 2019.
- [14] 刘文亚, 蒋奕, 王健. 肝包虫病影像学诊断专家共识[J]. 临床肝胆病杂志, 2021, 37(4): 792-797.
- [15] 邵党国, 邓阳阳, 相艳, 等. 基于自适应高斯滤波的超声斑点降噪[J]. 数据采集与处理, 2017, 32(4): 746-753.
- [16] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// IEEE Computer Society. IEEE Computer Society, 2013.
- [17] GIRSHICK R. Fast R-CNN [C]// 2015 IEEE ICCV, 2015: 1440-1448.
- [18] Ren Shaoqing, He Kaiming, Ross G, 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.
- [19] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv e-prints arXiv: 1804.02767, 2018.
- [20] 李兴怡, 岳洋. 梯度下降算法研究综述[J]. 软件工程, 2020, 23(2): 1-4.
- [21] 蒋文斌, 彭晶, 叶阁焰. 深度学习自适应学习率算法研究[J]. 华中科技大学学报(自然科学版), 2019, 47(5): 79-83.
- [22] 李克俭, 陈少波, 李万琦. 基于深度学习的交通标志检测系统仿真[J]. 软件导刊, 2020, 19(9): 31-34.
- [23] 杨晓玲, 江伟欣, 袁浩然. 基于 yolov5 的交通标志识别检测[J]. 信息技术与信息化, 2021(4): 28-30.
- [24] Wang Yijun, Zhou Pengyu, Zhong Wenya. An optimization strategy based on hybrid algorithm of Adam and SGD[C]// Proceedings of 2018 2nd International Conference on Electronic Information Technology and Computer Engineering (EITCE 2018), 2018: 4.

(收稿日期: 2021-07-20)

作者简介:

王正业(1996-), 男, 硕士研究生, 主要研究方向: 医学图像处理。

卡迪力亚·库尔班(1995-), 女, 硕士研究生, 主要研究方向: 医学信息处理。

严传波(1970-), 通信作者, 男, 教授, 主要研究方向: 生物信息处理、数据库应用, E-mail: ycbksy@126.com。



扫码下载电子文档

(上接第 24 页)

- 49-57.
- [13] 曾杨吉, 刘自红, 蔡勇, 等. 基于 YOLOv5 的姿态交互球类陪练机器人[J]. 电子技术应用, 2022, 48(1): 76-79.
- [14] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision(ECCV), 2018: 3-19.
- [15] 朱晓慧, 钱丽萍, 傅伟. 图像数据增强技术研究综述[J]. 软件导刊, 2021, 20(5): 230-236.
- [16] RAMACHANDRAN P, ZOPH B, LE Q V. Swish: a self-gated activation function[J]. arXiv preprint arXiv: 1710.05941, 2017.
- [17] Liu Xinchen, Liu Wu, Mei Tao, et al. PROVID: progressive and multimodal vehicle reidentification for large-scale urban

surveillance[J]. IEEE Trans. Multimedia, 2018, 20(3): 645-658.

- [18] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization[J]. International Journal of Computer Vision, 2020, 128(2): 336-359.

(收稿日期: 2022-03-29)

作者简介:

刘俊(1996-), 男, 硕士研究生, 主要研究方向: 计算机视觉。

钟国韵(1979-), 通信作者, 男, 博士, 副教授, 主要研究方向: 计算机视觉, E-mail: gyzhong@ecut.edu.com。

黄斯雯(2000-), 女, 本科, 主要研究方向: 机器学习。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所