

基于改进的 YOLOv5 的户外垃圾检测识别

陈胜选, 王爱民*

东南大学仪器科学与工程学院, 江苏 南京 210096

摘要 随着垃圾污染问题日益严重, 垃圾自动检测识别具有越来越重要的应用价值。改进了 YOLOv5 算法, 提升了对户外复杂背景下垃圾的检测性能, 收集了 6 个类别的户外常见垃圾的图片, 建立了一个背景复杂的垃圾图片数据集, 提出了一种简单、高效的方法用于生成图片中垃圾目标物的简易真值热力图。基于 YOLOv5 网络, 以真值热力图为量化标准, 设计并实验得出一种用于生成垃圾目标物预测热力图的分支结构。随后, 将预测热力图送回 YOLOv5 的 backbone 结构, 增加目标检测网络前向传播过程中特征图的空间注意力权重, 以提高整个目标检测网络的性能, 改进后的网络仅增加了少量参数, 生成了效果可观的预测热力图, 垃圾检测的性能得到较大提升。

关键词 机器视觉; 图像处理; 目标检测; 垃圾识别; 神经网络训练; 热力图

中图分类号 TP391.4; X705

文献标志码 A

DOI: 10.3788/LOP230472

Outdoor Garbage Detection Based on Improved YOLOv5

Chen Shengxuan, Wang Aimin*

School of Instrument Science and Engineering, Southeast University, Nanjing 210096, Jiangsu, China

Abstract Owing to the increasing severity of garbage pollution, automatic garbage detection has become significantly important in practice. The detection mechanism of YOLOv5 is improved in this study to achieve better performance in outdoor garbage detection against a complicated background. Moreover, here, a garbage dataset is constructed comprising six garbage image types collected in a complex background; subsequently, a simple yet efficient method is proposed to generate ground truth heat maps of garbage objects presented in the images. We treat the corresponding heat maps as a quantization standard and then obtain a branch structure based on YOLOv5 by conducting experiments to generate predicted heat maps. Subsequently, the predicted heat maps are sent back to the backbone structure of YOLOv5 to increase the spatial attention weights of the feature maps in the training process to improve the performance of the entire target detection network. Only a few parameters are added to the improved network, which generates proper predicted heat maps and the performance of garbage detection has been greatly improved.

Key words machine vision; image processing; object detection; garbage recognition; neural network training; heat map

1 引言

随着现代经济的快速发展, 人们的生活质量越来越高, 生活垃圾产生量也在迅速增加, 环境隐患日益突出。利用垃圾检测和分类可以实现资源的循环利用, 实现垃圾的无害化和资源化。近年来, 国内外许多学者都将深度学习应用于垃圾检测与分类。Ma 等^[1]提出了增强单发多框检测器(SSD)^[2]算法 L-SSD, 其相较于 SSD 算法增加了一个轻量且高效的特征融合模块, 并使用 ResNet-101^[3]作为 backbone 结构。L-SSD 垃圾检测的平均精度均值(mAP)为 83.48%。文灿华

等^[4]基于 Faster RCNN^[5]算法, 通过特异层微调、困难样本增强训练等方法改进算法, 并以 MobileNet_v1^[6]为 backbone 结构, 其所提算法使得垃圾检测的 mAP 提升到 94.90%。Feng 等^[7]提出了一个以 MobileNet 为 backbone 结构来提高 Mask RCNN^[8]性能模型。该模型可应用于嵌入式设备, 并且垃圾检测的 mAP 可达 93%。Yu 等^[9]提出了一种基于 YOLOv4^[10]目标检测网络并结合自学习策略的垃圾检测算法, 其 mAP 可达 95.36%, 但实验数据集很小。Yan 等^[11]提出了一种将剪枝策略集成到 YOLOv5 模型中的垃圾分类检测方法。该方法在一定程度上改善了垃圾分类耗时费力的

收稿日期: 2023-01-09; 修回日期: 2023-01-29; 录用日期: 2023-02-22; 网络首发日期: 2023-03-09

通信作者: *wangam@seu.edu.cn

问题,但实验数据集图片数量较少。此外,以上研究使用的均为自建数据集,且图片背景都较为简单。Huang 等^[12]使用以 MobileNet_v1 为 backbone 结构的目标检测算法 RetinaNet^[13],提出一种双向融合特征金字塔方法(garFPN)。在华为生活垃圾检测数据集上进行训练验证,该方法的 mAP 最高为 76.6%。Liu 等^[14]通过改进通道注意力机制 SENet^[15]模块提出了一种跨信道交互注意力机制(ECA),并将其引入到 YOLOv3^[16]的残差单元中,在“华为垃圾分类挑战杯”垃圾图像数据集上进行实验。最终,改进方法的 mAP@0.5(交并比阈值为 0.5 时的 mAP)和 mAP@0.5:0.95(交并比阈值在 0.5~0.95 之间每隔 0.05 取一次 mAP,并将这些 mAP 再取平均)分别比原 YOLOv3 算法提升了 0.72% 和 1.07%。

为了提高对多类垃圾的检测性能,本文改进了一阶段的 YOLOv5 目标检测算法。目前,垃圾检测研究使用的垃圾数据集图片背景相对简单,公开的垃圾数据集也相对较少。因此,自制了一个图片背景相对复杂的垃圾数据集,提出了一种简单的方法用于生成垃圾目标物的简易真值热力图,并在 YOLOv5 网络的开

始部分增加一个用于生成预测热力图的网络分支。将预测热力图作为空间注意力权重送回 YOLOv5 网络,改变特征图的空间权重以提升算法性能。

2 数据获取方法和算法

2.1 数据集建立

数据集选取户外空间日常出现的 6 个类别垃圾(塑料瓶、金属罐、小纸盒、口罩、塑料包装袋、纸类)。目前,已有的垃圾公开数据集相对较少,且拍摄于自然非平整地面的背景较为复杂的垃圾图片也非常少。最终,只在开源的 taco^[17]数据集中找到了 200 张符合要求的图片。因此,数据集中的绝大多数图片都是通过手机相机从户外实拍采集,所有图片都拍摄于草地、树丛、沟槽等地方,图片的背景信息都较为复杂。数据集中的图片总数为 9870 张,目标总数为 10188 个(塑料瓶 1851 个、金属罐 1754 个、小纸盒 1680 个、口罩 1510 个、塑料包装袋 1793 个、纸类 1600 个)。数据集中的 8220 张图片作为训练集,1650 张图片作为验证集,即训练集与验证集的图片数量比约为 5:1,各类别图片按此比例随机分配。数据集中的典型图片样本如图 1 所示。



图 1 各类别图片典型样本

Fig. 1 Typical samples of various types of pictures

2.2 YOLOv5 算法

YOLOv5 网络结构以及其中各模块的具体结构如图 2 所示,其中 YOLOv5 网络分为前后的 backbone 和 head 两部分结构。为了在图 2 中更好地展示 YOLOv5 的 C3 模块,使用 C3b1/2_n 表示 C3 模块,其中 b1/2 表示 C3 模块中使用的是 BottleNeck 1 或 BottleNeck 2 模块,*n* 表示 C3 模块中 BottleNeck 模块重复的次数。例如 C3b1_3 表示 C3 模块,该模块中使用

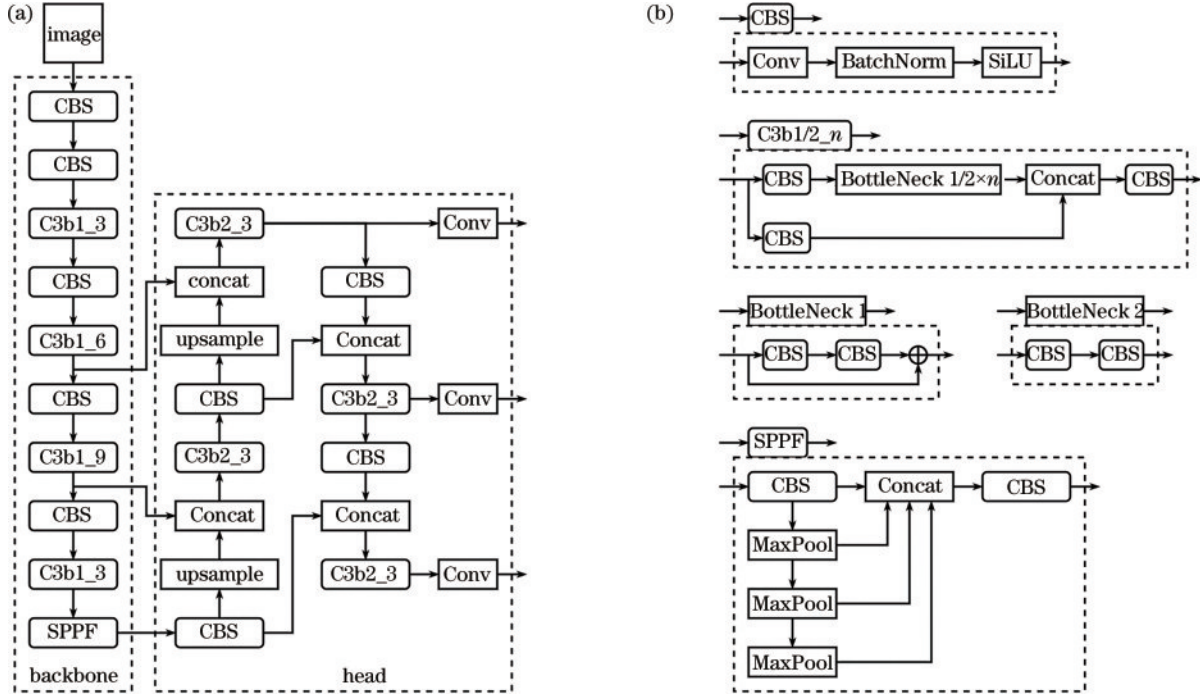


图 2 YOLOv5 网络结构及各模块结构图。(a) 网络结构图; (b) 模块结构图

Fig. 2 Structure diagrams of YOLOv5 network and main models. (a) Network structure diagram; (b) module structure diagram

YOLOv5 的损失值由分类损失 L_{cls} 、置信度损失 L_{obj} 、定位损失 L_{loc} 等 3 部分组成,即 YOLOv5 损失值 L_{YOLOv5} 可表示为

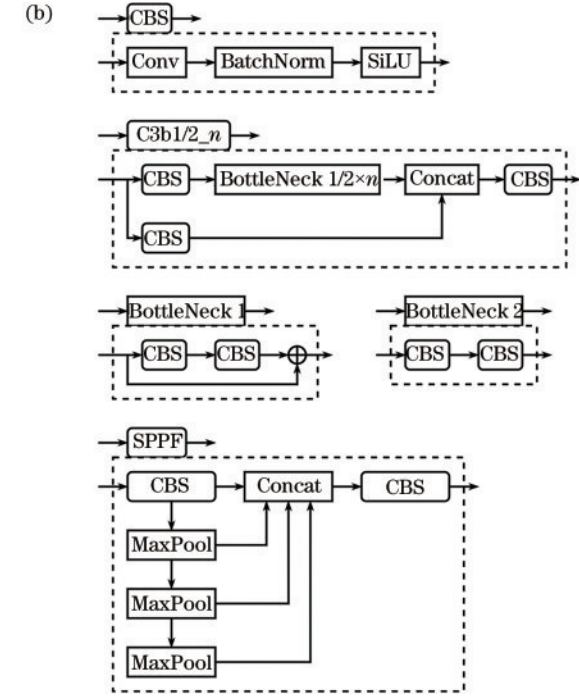
$$L_{YOLOv5} = \lambda_1 L_{cls} + \lambda_2 L_{obj} + \lambda_3 L_{loc}, \quad (1)$$

式中: λ_1 、 λ_2 、 λ_3 为平衡系数。YOLOv5 使用的数据增强方法有 Mosaic 方法、随机平移、缩放、水平翻转以及随机调整图片的色度、饱和度和明度等。其中, Mosaic 方法基于 YOLOv4, 随机将 4 张训练图片拼在一起组合成一张新的图片,增加了数据的多样性,也增加了单个图像的目标个数,有利于网络的训练。

2.3 算法改进方法

为生成目标物体热力图,先使用标注软件 Labelme 给训练集图片中的垃圾目标物标注一种新的折线标签,如图 3(a) 所示。该标签简单表示了物体的中心结构形态,使用简形线表示该标签。将图片中的所有简形线映射成相应的连续点,如图 3(b) 所示。将每根简形线上的每个点进行高斯模糊,其中,高斯核的宽度与简形线对应目标物体的 bounding box 短边长度成正比。高斯模糊后,生成简易的目标热力图,如图 3

BottleNeck 1 模块,并且 BottleNeck 1 模块重复 3 次。YOLOv5 模型从小到大依次分为 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x,通过程序中 depth_multiple 和 width_multiple 这两个参数来控制模型深度和各层输出通道数。本实验都是基于 YOLOv5s 模型,其每个模块的输出通道数为原 YOLOv5 的 1/2, C3 模块中的 BottleNeck 模块重复次数为原 YOLOv5 的 1/3。



(c)、图 3(d) 所示。

得到简易热力图后,进一步提出基于 YOLOv5s 网络的改进策略,如图 4 所示。图 2 已详细展示网络结构,图 4 中通过简略方式表示 YOLOv5s。

使用 YOLOv5s 网络中 backbone 的前几层输出构成一个特征金字塔(FPN)网络分支。该 FPN 分支输出一个单通道特征图,即预测热力图,用于预测图片中垃圾目标的空间位置和形状。并将预测热力图作为图片的空间注意力权重送回 YOLOv5s 的 backbone 中,用于在网络前向传播过程中提高数据中可能被认为是垃圾目标物所在位置的空间注意力。研究了预测热力图返回位置的不同(图 4 中的 P1、P2、P3 等位置)对改进算法性能的影响。

将已生成的简易热力图作为该 FPN 分支输出的预测特征图的真值,使用均方误差(MSE)损失函数计算其损失值,因此热力图的损失值 L_{hmap} 计算方式为

$$L_{hmap} = \text{MSE}(y, y') = \frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2, \quad (2)$$

式中: N 为一个 batch 中图片的张数; y 、 y' 分别为一个

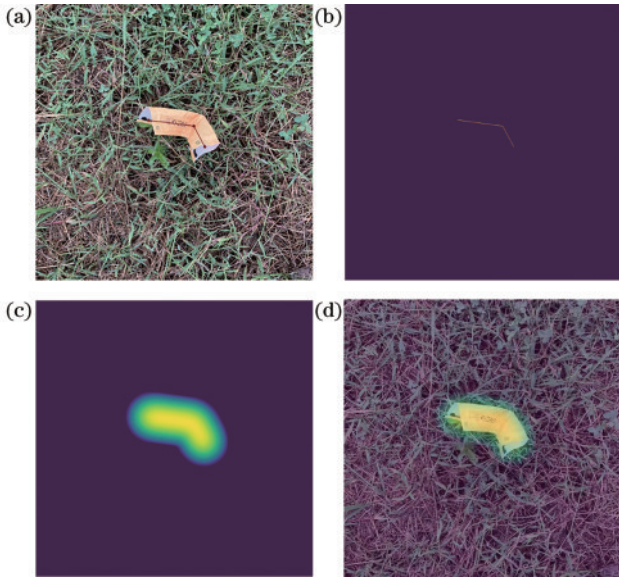


图 3 生成简易热力图过程。(a)折线标签;(b)折线标签映射成连续的点;(c)(d)简易热力图
Fig. 3 Process of generating a simple heat map. (a) Broken line label; (b) broken line map to continuous points; (c)(d) simple thermodynamic diagram

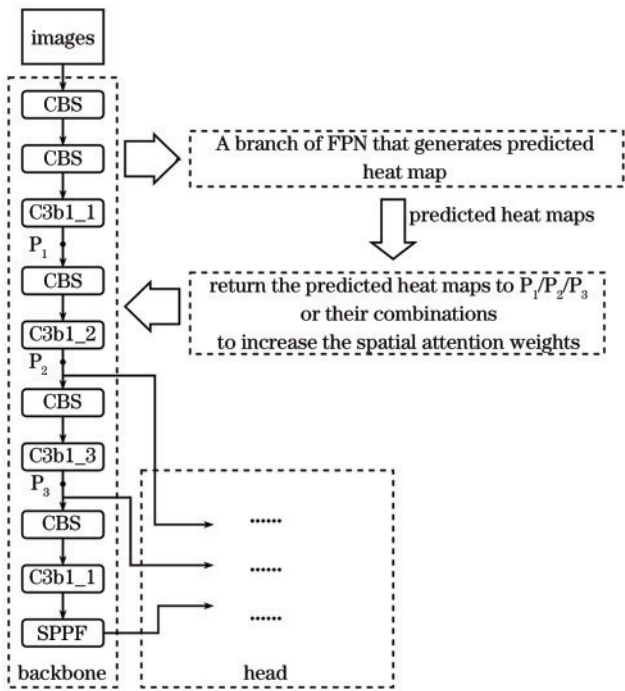


图 4 基于 YOLOv5s 网络的改进方法
Fig. 4 An improved method based on YOLOv5s network

batch中所有图片的真值热力图和预测热力图。使用一个新的优化器单独优化该 FPN 分支的参数。设 YOLOv5s 网络在返回位置的特征图为 x , 为增加权重, 预测热力图 y' 返回到这些位置后执行的操作为

$$x' = x * (1 + y'), \quad (3)$$

式中: $*$ 为特征图相乘; x' 为增加空间注意力权重后的特征图, 并送入 YOLOv5s 后续网络继续进行训练。

学习率采用余弦下降策略, 初始学习率设为 0.01, 最小学习率设为 0.0001。原 YOLOv5s 网络参数和 FPN 网络分支参数通过两个优化器分别单独优化, 两个优化器的优化算法均采用随机梯度下降 (SGD) 算法。

3 结果与讨论

3.1 设计 FPN 分支具体结构并对比性能

YOLOv5 网络自身有着非常良好的检测性能。建立的复杂背景垃圾数据集在原 YOLOv5s 网络中训练验证的 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别为 0.857 和 0.731, 而 CBS 模块和 C3 模块起到了非常重要的作用。因此设计的 FPN 分支结构全由 CBS 和 C3 模块组成, 降低了设计的复杂性。基于 YOLOv5s 前 3 层的输出组成的 FPN 分支大致结构如图 5 所示, 其中, k 为模块的卷积核大小, c 为模块的输出通道数, 未标识卷积核大小的模块, 其卷积核大小均为 1×1 , 并且所有模块的卷积步距均为 1。图中的 Module 1 和 Module 2 会设置成 CBS 和 C3 模块的不同组合以进行实验对比, 进而得出更优的 FPN 分支结构。该组实验中, 预测热力图的返回位置均为 $P_1 + P_2$ 。

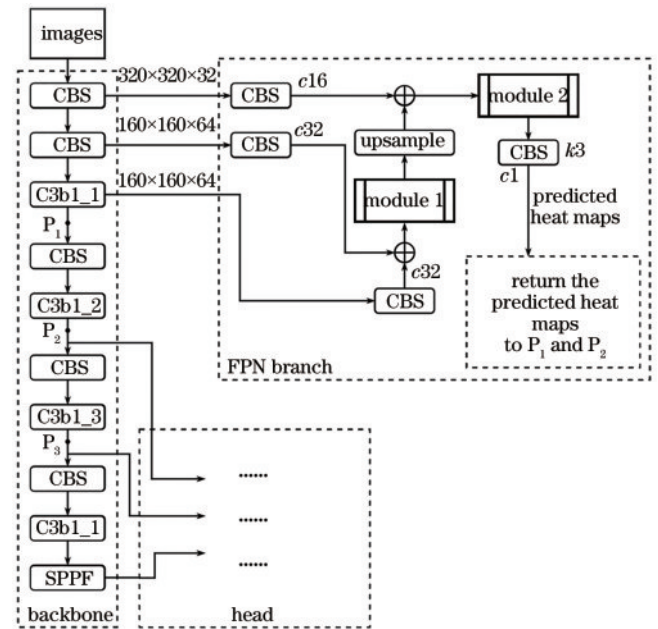


图 5 FPN 分支大致结构
Fig. 5 General structure of FPN branch

具有代表性的对比实验结果, 如表 1 所示。其中, Module 1 和 Module 2 中所有模块的卷积核大小均为 1×1 , 卷积步距均为 1, c 为对应模块的输出通道数。由表 1 可知, 当 Module 1 为 CBS 模块, Module 2 为 C3 模块时, 即将 C3 模块设置于所有浅层信息相加后的位置, 实验中网络的检测性能提升较好, 将此种组合下的 FPN 分支结构称为 FPN_b1。

表 1 FPN 分支结构中不同模块组合的实验性能对比

Table 1 Comparison of experimental performance of different module combinations in FPN branch structure

Module 1	Module 2	Elapsed time of each epoch /min	mAP@0.5	mAP@0.5:0.95
CBS(c16)		2.27	0.858	0.737
CBS(c16)	CBS(c8)	2.29	0.853	0.736
C3b2_1(c16) + CBS(c8)		2.26	0.859	0.737
CBS(c16)	C3b2_1(c8)	2.37	0.863	0.744
C3b2_1(c16) + CBS(c8)	C3b2_1(c4)	3.07	0.858	0.737

3.2 预测热力图返回位置测试对比

基于图 4 的改进思路,尝试将 FPN_b1 分支生成的预测热力图分别返回到 YOLOv5s 的 backbone 结构中的不同位置以增加数据的空间注意力,从而改进网络性能。

其中, P_1 点是 FPN_b1 分支生成的预测热力图能够返回 YOLOv5s 网络最靠前的位置; P_2 、 P_3 位置的特征图是 YOLOv5s 网络内部 FPN 结构的浅层信息,因此在 P_2 、 P_3 位置增加空间注意力可以同时影响到 YOLOv5s 网络中间的 Concat 模块的输入,使网络中数据的空间权重更加充分的更新,故选取 P_1 、 P_2 、 P_3 这 3 个位置作为实验测试位置。对比实验结果如表 2 所示。

由表 2 可知,网络中增加空间注意力的位置应该尽量靠前,返回位置 P_1+P_2 的组合使得改进后的网络检测性能的提升尤为明显。

表 2 预测热力图返回不同位置的实验结果

Table 2 Experimental results of predicted heat maps return to different positions

Network	Return position	mAP@0.5	mAP@0.5:0.95
YOLOv5s+	P_1	0.860	0.739
	P_2	0.862	0.736
	P_3	0.859	0.733
	P_1+P_2	0.863	0.744
	P_2+P_3	0.860	0.738
	$P_1+P_2+P_3$	0.860	0.736

3.3 不同分支结构检测性能对比

还尝试了几种不同的生成预测热力图的分支结构,如图 6 所示。其中,多列卷积神经网络(MCNN)^[18]直接由图片生成预测图。图 6(b)、图 6(c)中设定的 FPN 分支中的模块输出通道数为多次实验后得出的最优参数。为了突出展示分支结构,图 6 将 YOLOv5s 网络部分进一步简化。

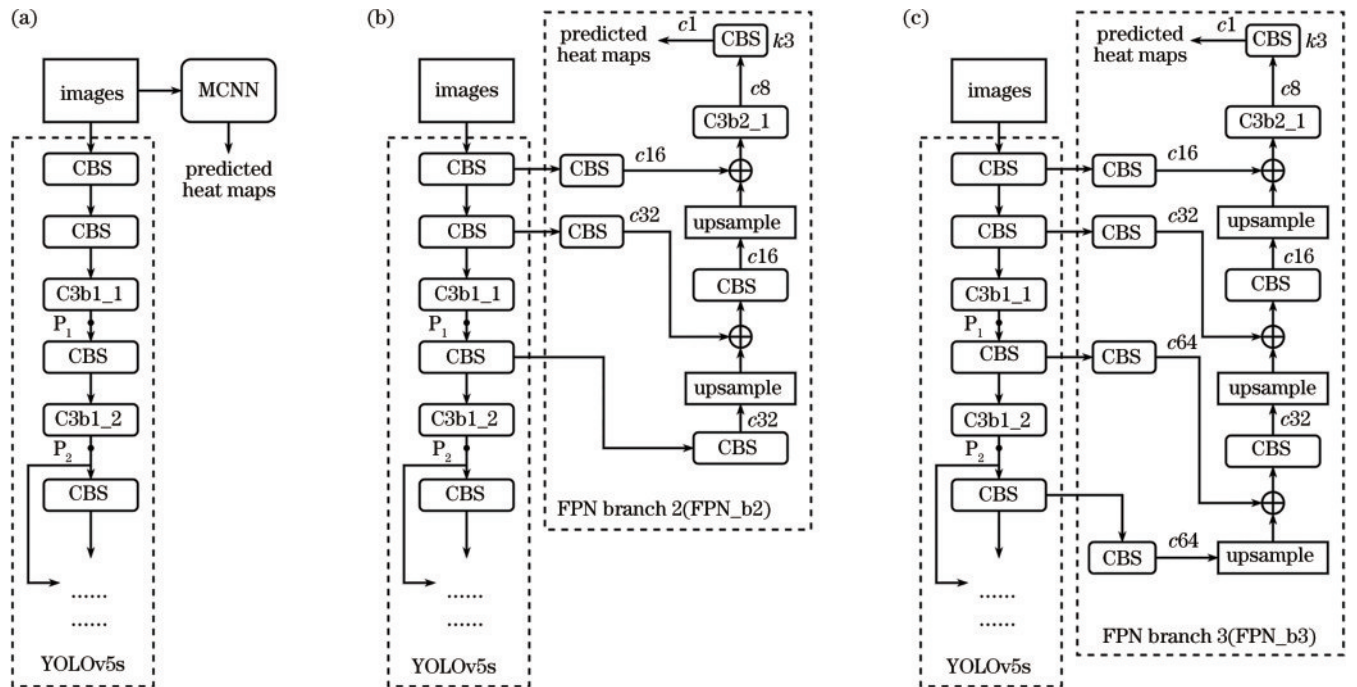


图 6 其他分支结构。(a)MCNN;(b)FPN_b2;(c)FPN_b3

Fig. 6 Other branch structures. (a) MCNN; (b) FPN_b2; (c) FPN_b3

算法改进的同时增加了分支结构的参数量,为了验证检测性能的提升主要来自算法的改进而非数量的提升,改变 YOLOv5 算法的模型大小,即调整程序中 depth_multiple 和 width_multiple 这两个参数来增大网络深度和每层的输出通道数并进行实验对比。YOLOv5 网络每层的输出通道数必须是 8 的倍数,因此

YOLOv5s 网络参数量增大最少量是每层输出通道数只增加 8,每层输出通道数只增加 8 的 YOLOv5s 网络定义为 YOLOv5s_plus。此外,还将收集的垃圾数据集分别在 SSD 和 Faster RCNN 目标检测算法上进行训练验证,实验结果如表 3 所示。

表 3 不同网络及分支结构的检测性能对比

Table 3 Comparison of detection performance of different networks and branch structures

Network	Parameters increment	Return position	mAP@0.5	mAP@0.5:0.95
YOLOv5s_plus	387664		0.860	0.738
YOLOv5s+FPN_b1	5610	P_1+P_2	0.863	0.744
YOLOv5s+FPN_b2	7658	P_2+P_3	0.860	0.738
YOLOv5s+FPN_b3	30442	P_3	0.863	0.740
YOLOv5s+MCNN	128224	P_1+P_2	0.860	0.739
SSD			0.777	0.623
Faster RCNN			0.781	0.648

由图 6 可知,FPN_b2 的输入已经超过了 P_1 的位置,因此 FPN_b2 生成的预测热力图不能返回到 P_1 位置,故返回到 P_2+P_3 ,同理将 FPN_b3 返回位置设为 P_3 。由表 3 可知,FPN_b1 分支结构的优越性在于,其相较于其他网络结构的参数量增加最少,对改进后的网络检测性能的提升却是最优的。这说明 FPN 结构用于生成热力图的优越性;YOLOv5s 网络最前面的 3 个模块输出的

特征层更适合作为 FPN 分支的输入;且 P_1+P_2 的返回位置也更适用于增加数据的空间注意力。

3.4 最终改进结果

以 FPN_b1 作为网络分支, P_1+P_2 为预测热力图返回位置,与原 YOLOv5s 网络相比,改进后的 YOLOv5s 网络中每个垃圾类别的召回率均有所提升,每个垃圾类别的召回率均有所提升,具体数据如表 4 所示。

表 4 YOLOv5s 网络与改进的 YOLOv5s 网络实验数据对比

Table 4 Comparison of experimental data of YOLOv5s network and improved YOLOv5s network

Network	Recall						mAP@0.5	mAP@0.5:0.95	Detection speed / (frame·s ⁻¹)
	Plastic bottle	Can	Carton	Mask	Plastic wrapper	Paper			
YOLOv5s	0.760	0.702	0.798	0.925	0.672	0.815	0.857	0.731	~36
YOLOv5s + FPN_b1	0.772	0.724	0.810	0.943	0.709	0.822	0.863	0.744	~34

预测热力图增加了特征层中被认为可能是垃圾目标物的空间位置和形状权重后,垃圾目标物能够被更加准确地识别。同时,垃圾检测的 mAP 也得到了提升。由表 4 可知,垃圾检测的 mAP@0.5 提升了 0.6 个百分点,mAP@0.5:0.95 提升了 1.3 个百分点,且检测速率并未发生明显变化。

训练过程中热力图的损失如图 7 所示。经过训练后 FPN_b1 生成的预测热力图如图 8 所示。虽然真值目标热力图形状简单,但这种高效的生成真值热力图的方法节省了大量的人力和时间成本,网络分支据此得出的预测热力图效果良好。

原 YOLOv5s 网络和改进后的 YOLOv5s 网络训练后对测试图片中的目标进行检测的部分结果如图 9(a)、图 9(b) 所示。由图 9 可知,原 YOLOv5s 网络将有些塑料包装袋识别成塑料瓶或小纸盒,而改进后的 YOLOv5s 网络对目标的类别预测得更加准确;此

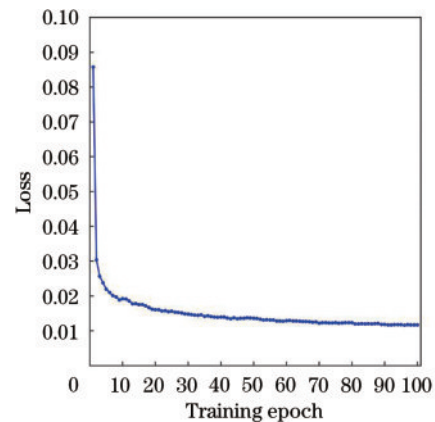


图 7 热力图损失

Fig. 7 Loss of heat map

外,原 YOLOv5s 网络有时会将背景中的树叶等元素识别成垃圾目标,而改进后的网络能更好地区分图片中的垃圾目标和背景。



图 8 预测热力图

Fig. 8 Predicted heat maps



图 9 YOLOv5s 网络改进前后检测性能对比。(a)原 YOLOv5s 网络;(b)改进后的 YOLOv5s 网络

Fig. 9 Detection performance comparison of YOLOv5s network before and after improvement. (a) Original YOLOv5s network;

(b) improved YOLOv5s network

4 结 论

提出了一种基于 YOLOv5s 目标检测网络的改进

思路,先后生成了图片中垃圾目标物的真值热力图和预测热力图。将预测热力图返回到原 YOLOv5s 网络的不同位置、对比不同的网络分支结构等实验,得出合

适有效的生成预测热力图的分支结构。分支结构的输入及增加空间注意力的位置选择对改进网络的性能提升非常重要。实验结果表明,所提方法生成的简易热力图可以作为垃圾目标物的预测热力图的量化标准,且生成的预测热力图效果良好。只增加少量参数的情况下,改进后的 YOLOv5s 网络的垃圾检测性能在 mAP@0.5 和 mAP@0.5:0.95 上分别提升了 0.6 个百分点和 1.3 个百分点,垃圾目标能够被更加准确地识别。

参 考 文 献

- [1] Ma W, Wang X, Yu J. A lightweight feature fusion single shot multibox detector for garbage detection[J]. IEEE Access, 2020, 8: 188577-188586.
- [2] 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.
- [3] 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 (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [4] 文灿华, 李佳, 董雪. 基于 Faster RCNN 的生活垃圾智能识别[J]. 激光与光电子学进展, 2020, 57(20): 201014. Wen C H, Li J, Dong X. Intelligent domestic garbage recognition based on faster RCNN[J]. Laser & Optoelectronics Progress, 2020, 57(20): 201014.
- [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] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2022-12-30]. <https://arxiv.org/abs/1704.04861>.
- [7] Feng J W, Tang X Y, Jiang X J, et al. Garbage disposal of complex background based on deep learning with limited hardware resources[J]. IEEE Sensors Journal, 2021, 21(18): 21050-21058.
- [8] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2980-2988.
- [9] Yu L, Pan G Y, Li M N. Garbage detection algorithm based on deep learning[C]//2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), October 15-17, 2021, Xi'an, China. New York: IEEE Press, 2021: 469-473.
- [10] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2022-12-30]. <https://arxiv.org/abs/2004.10934>.
- [11] Yan X B, Yang Y, Feng L, et al. A garbage classification method based on improved YOLOv5[C]//2022 International Conference on Networks, Communications and Information Technology (CNCIT), June 17-19, 2022, Beijing, China. New York: IEEE Press, 2022.
- [12] Huang Y T, Wen Q, Li Y Q. Lightweight domestic garbage detection method based on improved FPN[C]//2021 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), May 14-16, 2021, Guangzhou, China. New York: IEEE Press, 2021: 370-374.
- [13] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2999-3007.
- [14] Liu B H, Wang X P. Garbage detection algorithm based on YOLO v3[C]//2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), February 25-27, 2022, Changchun, China. New York: IEEE Press, 2022: 784-788.
- [15] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [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: 7132-7141.
- [16] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2022-12-30]. <https://arxiv.org/abs/1804.02767>.
- [17] Proença P F, Simões P. TACO: trash annotations in context for litter detection[EB/OL]. (2020-03-16)[2022-12-30]. <https://arxiv.org/abs/2003.06975>.
- [18] Zhang Y Y, Zhou D S, Chen S Q, et al. Single-image crowd counting via multi-column convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 589-597.