

控制与决策

Control and Decision

基于AC-YOLO的路面落叶检测方法

缪燕子, 张宗伟, 王贺升, 代伟, 赵忠祥, 王啸林, 杨春雨, 史延诺

引用本文:

缪燕子, 张宗伟, 王贺升, 代伟, 赵忠祥, 王啸林, 杨春雨, 史延诺. 基于AC-YOLO的路面落叶检测方法[J]. 控制与决策, 2023, 38(7): 1878–1886.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.1759>

您可能感兴趣的其他文章

Articles you may be interested in

[改进金字塔和跳跃连接的YOLOv5目标检测网络](#)

YOLOv5 object detection network with improved pyramid and skip connection

控制与决策. 2023, 38(6): 1730–1736 <https://doi.org/10.13195/j.kzyjc.2021.1414>

[基于注意力特征融合的无人机多目标跟踪算法](#)

UAV multi-target tracking algorithm based on attention feature fusion

控制与决策. 2023, 38(2): 345–353 <https://doi.org/10.13195/j.kzyjc.2021.1098>

[全局与局部图像特征自适应融合的小目标检测算法](#)

Small object detection algorithm based on adaptive fusion of global and local image features

控制与决策. 2023, 38(4): 935–943 <https://doi.org/10.13195/j.kzyjc.2021.1800>

[基于ResNet34_D改进YOLOv3模型的行人检测算法](#)

Pedestrian detection based on developed YOLOv3 with ResNet34_D

控制与决策. 2022, 37(7): 1713–1720 <https://doi.org/10.13195/j.kzyjc.2021.0136>

[多目标小尺度车辆目标检测方法](#)

Multi-target and small-scale vehicle target detection method

控制与决策. 2021, 36(11): 2707–2712 <https://doi.org/10.13195/j.kzyjc.2020.0635>

基于 AC-YOLO 的路面落叶检测方法

缪燕子[†], 张宗伟, 王贺升, 代伟, 赵忠祥, 王啸林, 杨春雨, 史延诺

(中国矿业大学 信息与控制工程学院, 江苏 徐州 221000)

摘要: 随着城市绿化程度的不断提高, 落叶清理任务变得更加复杂繁重. 针对落叶形状多变、大小不一、背景复杂、分布不均的特点, 提出一种融合 Attention-Context (AC) 网络和 YOLOv3 的落叶检测算法 (AC-YOLO), 解决现有模型对落叶漏检、误检的问题, 实现快速、准确地识别检测路面落叶. 针对小目标落叶易发生漏检的问题, 提出 AC 网络结构, 将不同层次的特征映射融合作为小目标的上下文信息, 同时引入自注意力机制来抑制复杂背景和底层噪声带来的影响, 提升小目标落叶检测能力; 其次, 采用 Mish 激活函数替换 Leaky ReLU 以提升模型的泛化能力, 提高落叶检测准确度; 最后, 考虑到落叶堆叠情况对清理机器人的工作效率有影响, 提出非极大值融合算法 (non-maximum fusion, NMF) 来融合密集落叶预测框, 从而通过更少的导航点解决密集落叶的检测问题, 同时提升落叶检测清理的效率. 实验结果表明, 基于 AC-YOLO 的检测算法对落叶检测的覆盖率 (cover) 达到 95%, 检测速度达到 53 帧/s, 可以完成实际应用环境中的落叶检测任务, 实现对落叶的高效率、智能化清理.

关键词: 落叶检测; YOLOv3; 小目标检测; 上下文信息; 注意力机制; Mish 激活; 非极大值融合

中图分类号: TP391.4

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.1759

引用格式: 缪燕子, 张宗伟, 王贺升, 等. 基于 AC-YOLO 的路面落叶检测方法 [J]. 控制与决策, 2023, 38(7): 1878-1886.

Detection method of fallen leaves on road based on AC-YOLO

MIAO Yan-zi[†], ZHANG Zong-wei, WANG He-sheng, DAI Wei, ZHAO Zhong-xiang, WANG Xiao-lin, YANG Chun-yu, SHI Yan-nuo

(School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221000, China)

Abstract: With the continuous improvement of urban greening, it is more complex and heavy to clean the fallen leaves. According to the characteristics of variable shape, complex background and uneven distribution of fallen leaves, we propose a fallen leaf detection algorithm, which integrates an attention context (AC) network and YOLOv3, named AC-YOLO, to detect fallen leaves on the road quickly and accurately. To solve the problem that small leaves are difficult to detect, we propose the AC network. We use different feature levels as the context information of small leaves and attention mechanism to suppress the influence of complex background and bottom noise, so as to improve the ability of small leaves detection. In addition, we use the Mish activation function to replace Leaky ReLU to enhance the generalization ability of the model and improve the accuracy of leaf detection. Finally, considering that the stacking of fallen leaves has an impact on the work efficiency of the cleaning robot, we propose the non-maximum fusion to fuse the dense fallen leaves detection boxes, which promotes the detection of dense fallen leaves and improves the work efficiency by reducing goal nodes. Experiment results show that the cover of leaf detection based on the AC-YOLO algorithm is 95% and the detection speed is 53 frames per second, which can complete the leaf detection task in the practical application and realize the efficient and intelligent leaves cleaning.

Keywords: fallen leaves detection; YOLOv3; small target detection; context information; attention mechanism; Mish activation; non-maximum fusion

收稿日期: 2021-10-13; 录用日期: 2022-03-28.

基金项目: 国家自然科学基金项目 (61976218); 徐州市重点科技项目 (KC19072); 中央高校基本科研业务费专项资金项目 (2020ZDPY0303); 江苏省研究生科技创新项目 (KYCX21_2262, SJCX21_1034, SJCX21_0992).

责任编辑: 张文安.

[†]通讯作者. E-mail: myz@cumt.edu.cn.

0 引言

随着目前城市绿化程度的不断提高,落叶清理工作也变得更加复杂繁重^[1-2]。不同于其他路面垃圾,落叶具有季节性特征,需要不断重复清扫,特别是绿化率高的园区环境和非机动车道,只能依靠人工清理,费时费力,效率低下。近年来,随着人工智能和目标检测技术的迅速发展,研究落叶智能检测方法从而实现落叶智能清理,是环卫领域的一条全新途径^[3]。因此,研究基于深度学习的落叶检测方法,实现对落叶的智能识别、定位是关键的技术问题,同时对于智能环卫车的推广应用也具有重要意义。

目前,基于深度学习的目标检测方法大致可以分为基于区域建议的两阶段算法^[4-7]和基于回归的单阶段算法^[8-12]。其中,单阶段算法YOLOv3因其更快的运算速度和更好的鲁棒性得到了广泛应用^[13-17]。但是,落叶的背景复杂,分布不均,且标签中存在大量小目标落叶,给检测带来了困难。所以,要实现落叶的准确检测,首先必须解决复杂背景下小目标落叶的检测问题。现有的基于深度学习的小目标检测方法按照思路的不同,可分为多尺度预测^[18]、超级分辨率^[19-20]、上下文信息融合^[21]、数据增强、新主干网络和训练方法^[22]共5种方法。其中,特征金字塔网络(feature pyramid networks, FPN)^[23]取得了较大进展。此方法主要在于分治策略的应用^[24],采用多尺度的方式对目标进行探测,提升了小目标的检测效率,但以YOLOv3为代表的算法在COCO数据集中对于小目标的检测精度也仅有17%左右,仍不能满足实际检测要求。

落叶分布不均也是落叶检测的难点之一。大量落叶聚集的状态使落叶之间存在严重遮挡问题,特征不明显,导致大量漏检情况。现有的针对密集目标检测的方法有很多,例如Soft-NMS^[25]、Repulsion Loss^[26]等。Soft-NMS是一种后端处理算法,通过降低与高置信度预测框重合面积超过阈值的预测框的置信度的方式提升了密集目标的检测效果。Repulsion Loss使预测框远离相邻的但不是预测同一目标的预测框,在密集人群检测方面取得了较好效果。但由于落叶过于密集,现有的密集目标检测算法并不能满足密集落叶的检测需求。而且,密集的落叶会传递出更多的导航点,需要落叶清理机器人付出更多的算力,降低了清扫的工作效率。所以,如何在准确检测密集落叶的基础上,提升落叶清理机器人的工作效率也是亟待解决的问题。

本文在YOLOv3模型的基础上,提出一种基于

AC-YOLO的路面落叶检测方法,用于解决以上复杂情况下的落叶检测任务。针对小目标落叶较多的特点,首先构建一种新的AC网络结构,通过多尺度层的连接将不同层次的特征作为上下文信息,配合CBAM自注意力算法替换原FPN结构,提升小目标落叶的检测精度;其次,针对落叶背景环境复杂的特点,使用Mish激活函数增强模型泛化能力,进一步提升落叶检测的准确度。最后,针对落叶分布不均匀,密集堆叠落叶难以检测的问题,同时考虑到落叶机器人实际工作方式是线扫的形式,提出非极大值融合算法(non-maximum fusion, NMF)替换非极大值抑制算法(non-maximum suppression, NMS),在减少导航点的同时解决密集状态落叶的检测问题。本文提出的落叶检测方法,在保证检测精度的前提下,通过精简网络结构,减少网络参数,提高网络运行速度,满足落叶智能检测的实时性要求。

1 AC-YOLO

本文主要研究路面落叶检测问题,是实现智能落叶清理的基础工作。由落叶清理机器人通过机载摄像头采集路面图片,将图片进行预处理获得608×608的图片,通过Darknet卷积神经网络提取网络特征。为了解决落叶检测中的复杂问题,采用多尺度检测方法在4倍、8倍和16倍降采样层检测落叶。通过AC网络融合上下文信息,抑制复杂环境的影响,提升落叶检测的准确率和召回率。最后,使用非极大融合算法的处理预测框对落叶进行识别和定位,并将位置信息上传,由机器人通过路径规划完成落叶扫除任务。

1.1 AC网络

经过对落叶数据集的分析,落叶标签大多数均小于原图的0.1,如图1所示,即实际的落叶中以小目标居多^[27]。Darknet-53较深的网络层数提高了模型的特征提取能力^[8],然而随着网络层数的加深,高层特征会产生特征消失的问题,尤其是小目标特征信息。虽然YOLOv3算法中使用特征金字塔网络FPN提高了小目标的检测精度,但是与大中目标检测相比,其对于小目标的检测精度依然不满足要求。所以,为了得到更适合小目标落叶检测的深度学习网络,降低落叶中小目标的漏检率,本文首先减少YOLOv3模型的特征提取网络层数,将原网络的32倍降采样层舍弃,加入4倍采样层,获取更丰富的小目标纹理信息,如图2所示。

为了给小落叶目标提供充足的上下文信息,并减弱复杂背景带来的影响,本文提出一种融合注意力机

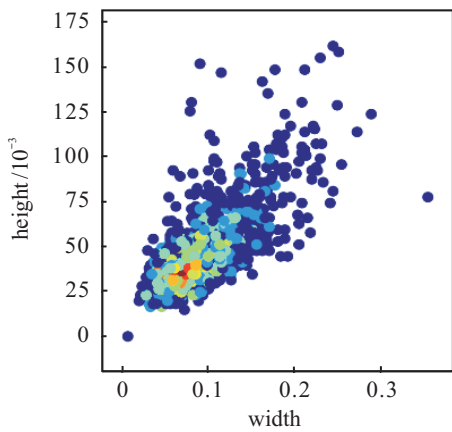


图1 落叶目标真实框大小统计图

制和上下文信息的AC网络结构,如图3所示. AC网络结构将目标特征层的更高层次与被注意力模块强化的低层次的特征映射相融合,生成新的包含目标上下文信息的特征组合. 例如,当采用8倍降采样层 P_3 检测目标时,其上下文特征来自4倍降采样层 P_2 和16倍降采样层 P_4 . 在特征拼接之前,对 P_2 特征层进行卷积下采样,使其与目标特征层 P_3 具有相同的空间大小,通道数设置为目标特征层的1/2. 对 P_4 特征层进行反卷积,获得与 P_3 相同的尺度,通道数设置为 P_3 的1/2. 最后,对目标特征与上下文特征进行叠加处理,获得强化之后的特征信息 P_5 .

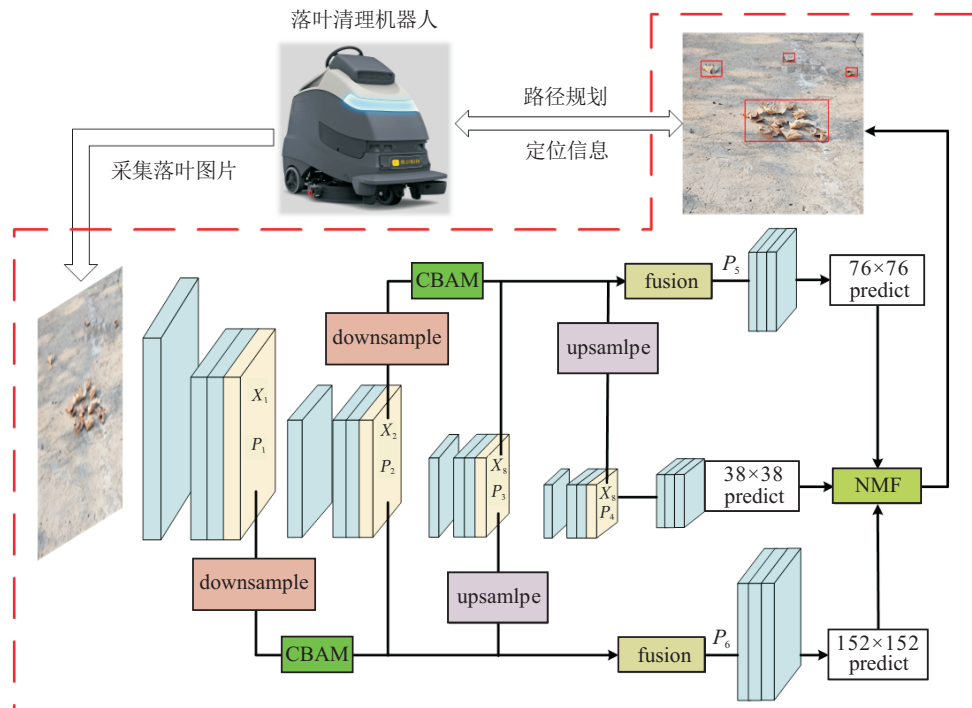


图2 AC-YOLO结构图

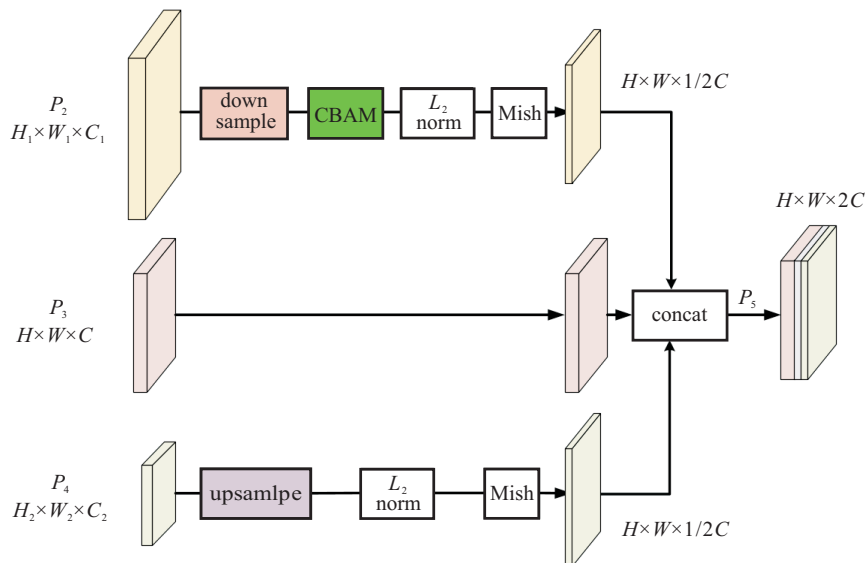


图3 特征融合

图3中,4倍降采样特征的加入,在丰富小目标纹理信息的同时,也带入了更多的背景纹理信息和噪声,对检测效果产生消极的影响.同时,落叶检测多为室外环境,背景复杂多变,导致模型的误检率较高.所以,在低层特征下采样之后加入注意力机制模块(convolutional block attention module, CBAM)^[28-29],以减弱背景带来的消极影响,让模型更加关注目标物本身. CBAM网络结构如图4所示.

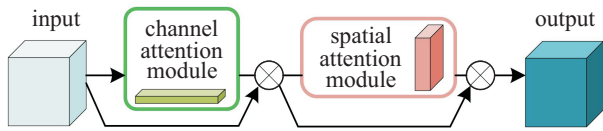


图4 CBAM结构

在YOLOv3网络中,在主干网络特征输出后会有5个连续的3×3和1×1卷积层.一般而言,高层卷积中重复卷积的操作可以处理类别较多的情况.但对于路面落叶检测而言仅有一类,这意味着可以通过降低高层卷积层数提高模型的识别效果.因此,本文将连续的5层卷积层减少为3层,以提高落叶检测效果.

1.2 非极大值融合

为了同时解决密集落叶的检测和清扫工作效率的问题,使算法更好地配合落叶机器人实现落叶清扫效果,本文提出一种非极大值融合算法(non-maximum fusion, NMF)取代原始YOLOv3中非极大值抑制算法(non-maximum suppression, NMS),作为落叶检测模型的预测框后端处理算法.

设A、B为预测模型得到的预测框,其参数分别为 $[x_{a1}, y_{a1}, x_{a2}, y_{a2}, s_{a1}, s_{a2}]$ 和 $[x_{b1}, y_{b1}, x_{b2}, y_{b2}, s_{b1}, s_{b2}]$.其中: $[x_1, y_1, x_2, y_2]$ 为所对应预测框的定位参数, s_1 和 s_2 分别为位置得分和分类得分.首先根据落叶机器人的清扫覆盖面积,对具有较大 s_1 的预测框进行放大处理以获取更好的清扫效果.当两预测框相交时,NMF将融合预测框. NMF可用下式表示:

$$\text{nmf} \left(\begin{bmatrix} \delta x_1 \\ \delta y_1 \\ \delta x_2 \\ \delta y_2 \\ s_1 \\ s_2 \end{bmatrix}, \begin{bmatrix} x_1 \\ y_1' \\ x_2 \\ y_2' \\ s_1' \\ s_2' \end{bmatrix} \right) = \begin{bmatrix} \min(x_1, x_1') \\ \min(y_1, y_1') \\ \max(x_2, x_2') \\ \max(y_2, y_2') \\ \max(s_1, s_1') \\ \max(s_2, s_2') \end{bmatrix}. \quad (1)$$

如图5所示,经过放大 δ 倍的预测框与其他预测框存在相交时,通过NMF得到包括更多落叶在内的更大预测框.这不仅能够解决密集落叶检测问题,提高算法对落叶的覆盖率,而且能通过更少的导航点减少算力,提高落叶清扫机器人的工作效率.

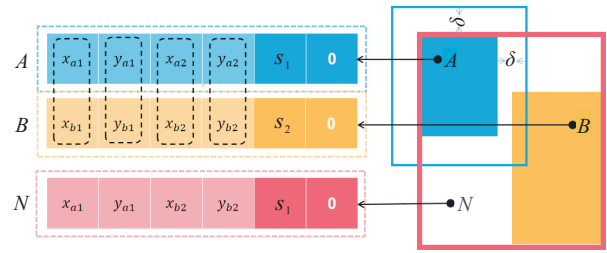


图5 非极大值融合

值得一提的是,当路面落叶密度过大而导致落叶连续分布时,NMF会得到一个超大的预测框,甚至超过落叶清理机器人的覆盖面积.为此,本文引入一个新的可根据落叶清理机器人覆盖面积进行调节的参数 l .当预测框宽高参数为 $w, h (w > h)$ 时,若 $w > l$,则对预测框进行均分处理,其过程可用下式表示:

$$(w', h') = \frac{(w, h)}{\left[\frac{w}{l} + 1 \right]}, \quad (2)$$

其中 (w', h') 表示新的预测框宽高.

1.3 Mish激活

为了使模型具有更好的泛化性能,提高在复杂情景下落叶检测的准确度,采用Mish激活函数替换YOLOv3中的Leaky ReLU激活函数. Mish激活函数可用下式表示:

$$f(x) = x \tanh(\text{softplus}(x)) = x \tanh(\ln(1 + e^x)). \quad (3)$$

如图6所示,Mish和Leaky ReLU激活函数都具有上无界下有界的特性且正值收敛速度相当.激活函数上无界的特性可以保证模型训练时避免梯度饱和,加快模型训练速度;有下界且有较小负值,可以保证正则化效果的实现和网络梯度流的稳定.但是,相比于Leaky ReLU激活函数, Mish激活函数具有更好的非线性特征,使模型具有更好的泛化能力,可以提高模型预测落叶的准确性.

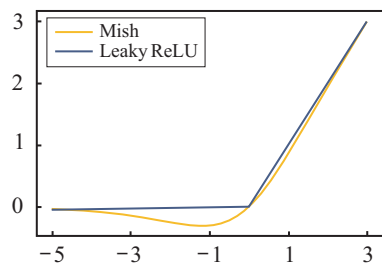


图6 Mish和Leaky ReLU激活函数

2 实验结果与分析

2.1 构建落叶数据集

选择训练数据集和图像标定是目标检测过程中两个至关重要的步骤.由于目前缺乏公开通用落叶

数据集,本文构建并扩充数据集训练和验证落叶检测算法.考虑到落叶检测的影响因素,本文从不同季节、不同环境、遮挡和落叶形状角度,拍摄搜集落叶图片2300张.使用LabelImage将数据集中图片按照VOC2007数据集格式整理、标注.同时,考虑到室外复杂环境和光照的影响,对数据集进行对比度和亮度调整,并使用翻转等数据增强方法进一步扩充数据集,增强训练模型的泛化能力.扩充后数据集图片数量为6900张.其中:6000张用于训练,900张用于测试.

2.2 Anchor参数优化

YOLOv3借鉴先验框Anchors来预测边框坐标,Anchors是几组宽高固定的先验框.在目标检测过程中,先验框的大小与目标的匹配程度直接影响模型检测的速度和精度.因此,根据落叶数据集中落叶标签的特点设定相匹配的Anchor参数组合尤为重要.本文所用数据集中的落叶目标多为小目标,YOLOv3原本的先验框大小不再满足检测需求,需要重新聚类优化Anchor参数.考虑到K-means算法在初始聚类中心的选择上具有很大随机性,而这种特点会对聚类结果产生消极影响,因此,本文选用随机性较小的K-means++算法^[30]进行聚类计算.使用K-means++算法对自制的落叶数据集重新聚类.通过多次聚类对比,当先验框的组合大于9组时,有冗余先验框组合出现,故9组先验框为落叶数据集的最优组合.检测尺度分配如表1所示.

表1 先验框组合

目标特征层	先验框组合
×4	(20, 10) (26, 15) (31, 22)
×8	(37, 30) (42, 22) (52, 26)
×16	(55, 19) (70, 32) (94, 48)

2.3 评价指标

考虑到在实际应用中对落叶检测速度和准确率的要求,实验采用精确率(precision)、召回率(recall)、平均准确率均值(mAP@0.5)和帧率(FPS)作为评价指标.

1) 精确率(precision)和召回率(recall)分别为

$$P_{\text{precision}} = \frac{TP}{TP + FP}, \quad (4)$$

$$P_{\text{recall}} = \frac{TP}{TP + FN}. \quad (5)$$

其中:TP表示检测到的目标类别与真实目标类别一致的样本数量,FP表示检测到的目标类别与真实类别不一致的样本数量,FN表示真实目标存在但未被

检测出的样本数量.本文所用落叶数据集,小目标较多,因而召回率的提高更为重要.

2) 平均准确率均值(mAP@0.5)为

$$AP = \int_0^1 P(R)dR, \quad (6)$$

$$mAP = \sum_{i=1}^N AP_i / N, \quad (7)$$

其中N表示目标类别的数量.因本文专注于落叶检测,故N=1,mAP@0.5即为AP50.

3) 落叶检测覆盖率(cover).

因为落叶清扫任务是按导航点清扫所有落叶,并不需要详细分类,特别是针对密集落叶不需要按片定位,所以目标检测中常用的评价指标平均准确率(mAP)不适用于评价落叶检测的任务完成率.为了更好地评价算法对于路面落叶的检测覆盖程度,本文提出适用于落叶检测的评价指标——检测覆盖率(cover),即

$$\text{cover@0.5} = \frac{\text{Num}(\text{Singleiou} \geq 0.5)}{TP + FN}, \quad (8)$$

$$\text{Singleiou} = \frac{S_{\text{union}}}{S_T}. \quad (9)$$

其中:Num表示符合条件的预测框数量, S_{union} 表示真实框与预测框的相交面积, S_T 表示真实框面积表示.如图7所示,在一张图片上, T_1 、 T_2 、 T_3 、 T_4 为真实框,P为预测框,可得 $\text{Singleiou}(T_1) < 0.5$, $\text{Singleiou}(T_2, T_3, T_4) > 0.5$,则 $\text{Num}(\text{Singleiou} \geq 0.5) = 8$,得 $\text{cover@0.5} = 88.9\%$.

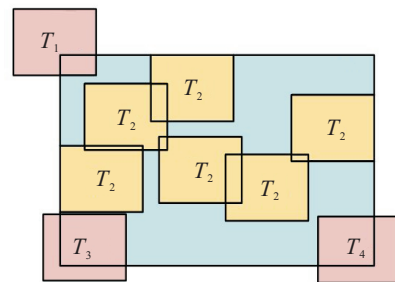


图7 落叶检测覆盖率

4) 帧率(frames per second, FPS).

帧率是指目标检测网络每秒中能够检测的图片数量,用来评价网络的检测速度,帧率的大小能够反映出网络是否满足实时性要求.

2.4 实验分析

实验基于LINUX平台通过Python语言实现,应用深度学习框架Pytorch搭建网络模型,硬件环境为NVIDIA GeForce RTX 2080Ti.对比实验的硬件配置环境相同.图8为YOLOv3和AC-YOLO的训练损失对比.

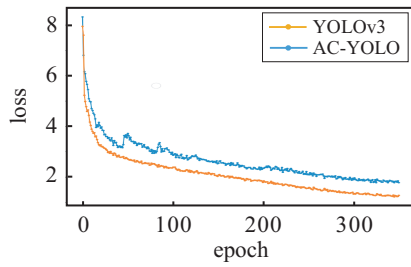


图8 YOLOv3和AC-YOLO训练损失

训练中得到YOLOv3算法参数总量为 6.15237×10^7 . 相比于FPN, AC网络检测参数量减少了 6×10^5 . 此外, 32倍降采样的舍弃使AC-YOLO参数总量降至 2.00257×10^7 , 相比于原模型参数量大大减少. 训练得到YOLOv3与AC-YOLO的权重文件分别为233 M和76 M. 较少的参数量保证了AC-YOLO模型占据更小的内存, 能在有限的算力情况下实现更高的效率.

2.4.1 定量对比实验

为了验证本文方法可以有效提高小目标检测的效率, 选取COCO评价指标中的平均准确率mAP@0.5、AP_s和AR_s在PASCAL VOC 2007数据集上进行对比实验, 输入图片大小为416×416. 从表2

可以看出, 在保证检测精度的基础上, AC-YOLO在小目标的准确度和召回率上都取得了比YOLOv5和超分辨率方法更好的效果, 这说明AC-YOLO可以有效提升对小目标落叶的检测效果.

表2 VOC 2007数据集实验结果 %

算法	mAP@0.5	AP _s	AR _s
YOLOv5	70.0	10.5	21.9
SR ^[19]	77.0	10.2	—
AC-YOLO	66.9	11.6	27.0

为了进一步验证本文方法的有效性, 选取目标检测算法评价指标精确率、召回率和mAP@0.5进行消融对比实验. 为了更好地体现模块效果, 预测框后端处理算法使用NMS. 从表3可以看出, AC网络结构的加入使模型的召回率和准确度都得到明显提升, 尤其是召回率的显著提升, 说明AC结构有效提升了网络对于小落叶目标的提取与检测能力. 同时, 表3第4、第5行的结果表明, 相比于Leaky ReLU激活函数, Mish激活函数可使模型的泛化能力增强, 进一步提升落叶检测模型的检测性能.

表3 消融实验

算法	Context module	CBAM	Mish	precision/%	recall/%	mAP@0.5/%
YOLOV3	—	—	—	75.4	82.0	80.5
				83.2	91.9	85.9
AC-YOLO-NMS	✓	✓	—	84.0	91.5	87.8
	✓	✓	✓	85.2	91.6	89.3

为了体现本文算法在落叶检测方面的优越性能, 将AC-YOLO算法与Faster R-CNN、YOLO系列算法一起在落叶数据集上进行训练、测试, 并选取落叶检测覆盖率cover和检测速度FPS两项指标对各算法进行横向对比. 表4展示了Faster R-CNN、YOLO系列与AC-YOLO算法的对比实验结果.

表4 不同目标检测算法结果对比

算法	cover/%	FPS/(帧/s)
Faster R-CNN	77.8	20.5
YOLOv3	82.5	42
YOLOv4	87.2	36.5
YOLOv5	90.2	36
AC-YOLO	95.2	53

从表4中可以看出, 两阶段目标检测算法Faster R-CNN的落叶覆盖率为77.8%, 且检测速度较低, 这主要是由于两阶段目标检测算法在检测过程中候选区域生成网络(RPN)生成包含待检物体的候选框会

增加网络模型的计算时间. 使用单阶段目标检测算法YOLO系列算法检测落叶时, 检测速度得到了很大提升, 且检测性能优良, YOLOv4和YOLOv5分别获得了87.2%和90.2%的落叶覆盖率; 而AC-YOLO算法较YOLOv4与YOLOv5算法对落叶目标检测在落叶覆盖率上又进一步得到提升, 在AC网络和NMF的作用下提升了对小目标落叶和密集落叶的检测效果, 落叶检测覆盖率达到95.2%, 同时检测速度达到53帧/s, 在保证落叶检测覆盖率的前提下具有更好的实时性能.

2.4.2 定性对比实验

为了验证本文算法在实际环境中的检测效果, 使用园区落叶清扫车搭载摄像头进行检测结果可视化实验. 图9为Darknet-53特征提取网络的部分通道特征图可视化.

图9(a)为路面落叶实拍图片, 大小为608×608像素. 红框内为目标落叶, 约为30~80像素. 图9(b)、(c)、

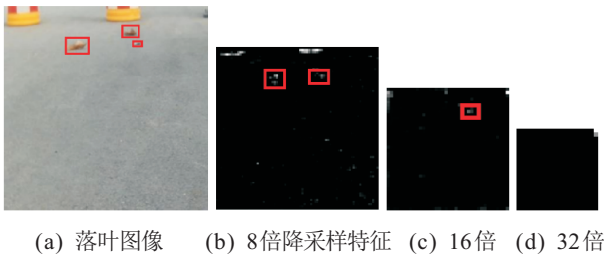


图9 可视化特征通道

(d)依次为8倍、16倍、32倍降采样特征图.在8倍特征图中可以看到落叶的纹理信息,16倍降采样特征图中也含有部分纹理信息以及语义信息.但在32倍降采样特征图中,由于落叶目标较小,特征图中已经基本失去了目标特征信息.这也说明了使用4倍、8倍、16倍降采样特征进行多尺度检测的效果比32倍降采样层更好的原因.

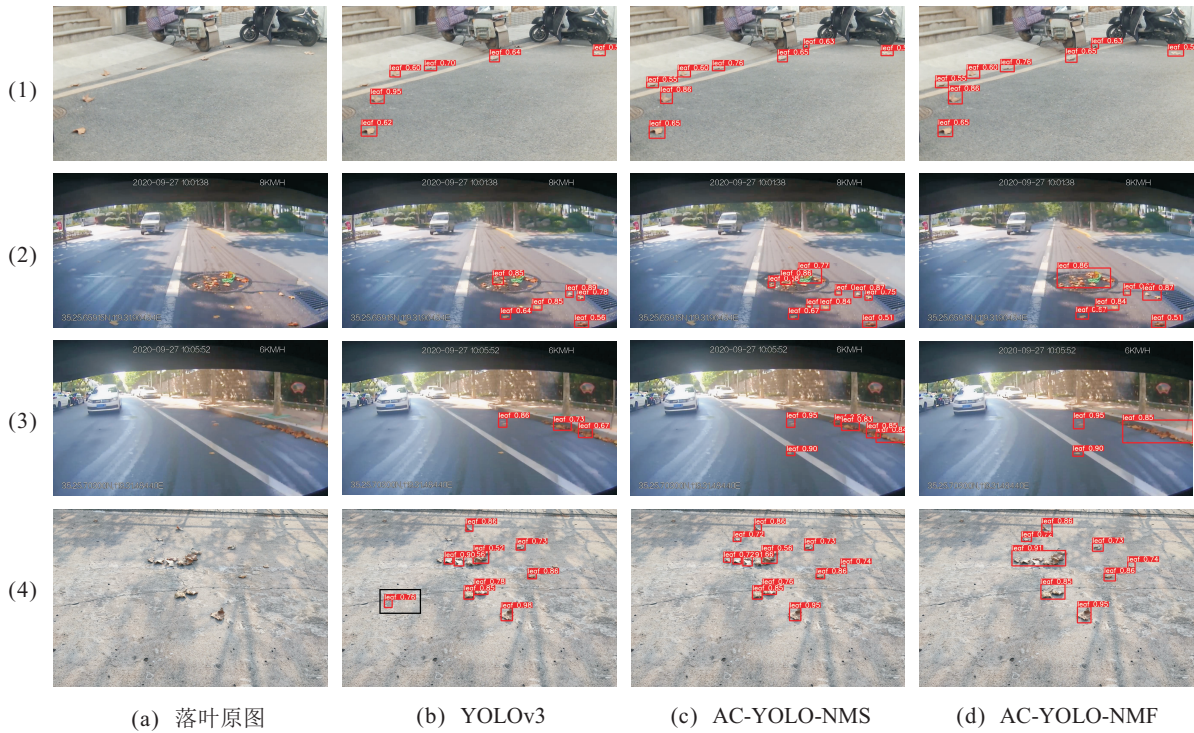


图10 YOLOv3与AC-YOLO定性对比实验

图10为在真实落叶场景中YOLOv3算法和AC-YOLO算法检测的对比实验.

AC网络:对比图10(b)和图10(c)的可视化效果图,AC网络有效提升了落叶检测效率.在落叶较大、遮挡不严重的简单环境中,两种算法均能准确地检测出落叶所处位置,如图10(1)所示.对于较小的落叶目标,YOLOv3算法出现大量漏检情况,而AC-YOLO算法则能实现对小目标落叶的精准预测.这是因为YOLOv3提取到的小目标特征信息不足,而AC-YOLO降低采样层次,并使目标特征层融合了上下文信息,提供了足够的小落叶特征,所以在小落叶检测上取得了更好的效果.由于复杂背景的影响,YOLOv3算法所得预测框的置信度较低,存在误检情况,例如图10(b)(4)中黑色标注框;而AC-YOLO算法在注意力机制和Mish激活函数的作用下,重要特征层的权重增加,弱化了背景的消极影响,模型泛化能力增强,预测准确度相比原算法有整体提高.实验结果表明,AC-YOLO算法针对落叶检测具有更好的准

确率和召回率.

NMF:对比图10(c)和图10(d)的预测框,说明NMF算法能够有效解决落叶分布不均的问题.在面对密集堆叠的落叶时,AC-YOLO算法能够检测出边缘与表面的部分落叶.通过NMF算法调节放大具有最大置信度的预测框,捕捉融合其周边预测框,将未检测到的落叶囊括在内,融合效果如图10(d)所示.图10(c)(3)中,落叶导航点共有6个,而在融合操作后,导航点降为3个,但囊括了更多落叶.当落叶不是密集聚拢时,NMF算法得到的预测框较NMS所得到的稍大,保证了在稀疏情况下对落叶的准确检测;当落叶密集堆积时,通过NMF融合预测框,AC-YOLO可以将一些未检测到的落叶包含在预测框内,并且提供更少的导航点,提升落叶机器人的清扫效率.当落叶机器人具有较大的工作面时,可以调大 δ 的取值,使高置信度框捕捉更多附近的目标,实现更高的落叶覆盖率.当融合入的预测框过大而超过落叶清理机器人的覆盖面积时,NMF会根据预设参数 l 对预测框进行

均分处理以适应机器人清洁工作。

3 结论

为了解决复杂背景下落叶检测漏检率和误检率较高的问题,本文提出了基于AC-YOLO的落叶检测算法。首先,使用K-means++算法对落叶数据集进行聚类分析,获得最优Anchor参数;其次,融合自注意力特征和上下文特征信息获得AC网络,并调用Mish激活函数增强模型泛化能力,提升了小落叶的检测精度;最后,使用NMF算法融合预测框,提升了落叶检测的覆盖率,减少了导航目标点。

实验结果表明,AC-YOLO显著提升了落叶检测准确率,且检测速度满足实时性要求,为落叶智能清理提供了一种新的思路,推动了环卫车“电动化”“智能化”的创新发展。

参考文献(References)

- [1] 雒占福, 张金, 刘娅婷, 等. 2000-2017年中国城市绿化水平的时空演变及其影响因素研究[J]. 干旱区地理, 2020, 43(2): 481-490.
(Luo Z F, Zhang J, Liu Y T, et al. Spatiotemporal evolution of urban greening in China and the affecting factors between 2000 and 2017[J]. Arid Land Geography, 2020, 43(2): 481-490.)
- [2] 周亮, 车磊, 周成虎. 中国城市绿色发展效率时空演变特征及影响因素[J]. 地理学报, 2019, 74(10): 2027-2044.
(Zhou L, Che L, Zhou C H. Spatio-temporal evolution and influencing factors of urban green development efficiency in China[J]. Acta Geographica Sinica, 2019, 74(10): 2027-2044.)
- [3] 张慧, 王坤峰, 王飞跃. 深度学习在目标视觉检测中的应用进展与展望[J]. 自动化学报, 2017, 43(8): 1289-1305.
(Zhang H, Wang K F, Wang F Y. Advances and perspectives on applications of deep learning in visual object detection[J]. Acta Automatica Sinica, 2017, 43(8): 1289-1305.)
- [4] Dong P L, Wang W M. Better region proposals for pedestrian detection with R-CNN[C]. 2016 Visual Communications and Image Processing (VCIP). Chengdu, 2016: 1-4.
- [5] Wang X L, Shrivastava A, Gupta A. A-fast-RCNN: Hard positive generation via adversary for object detection[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 3039-3048.
- [6] 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.
- [7] 王殿伟, 杨旭, 韩鹏飞, 等. 复杂背景下全景视频运动小目标检测算法[J]. 控制与决策, 2021, 36(1): 249-256.
(Wang D W, Yang X, Han P F, et al. Panoramic video motion small target detection algorithm in complex background[J]. Control and Decision, 2021, 36(1): 249-256.)
- [8] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J/OL]. 2018, arXiv: 1804.02767.
- [9] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, 2017: 6517-6525.
- [10] 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. Las Vegas, 2016: 779-788.
- [11] Ryu J, Kim S. Chinese character detection using modified single shot multibox detector[C]. The 18th International Conference on Control, Automation and Systems (ICCAS). PyeongChang, 2018: 1313-1315.
- [12] Tan M X, Pang R M, Le Q V. EfficientDet: Scalable and efficient object detection[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020: 10778-10787.
- [13] 刘杰, 朱旋, 宋密密. 改进YOLOv2的端到端自然场景中文字符检测[J]. 控制与决策, 2021, 36(10): 2483-2489.
(Liu J, Zhu X, Song M M. End-to-end Chinese character detection in natural scene based on improved YOLOv2[J]. Control and Decision, 2021, 36(10): 2483-2489.)
- [14] Xu Z F, Jia R S, Sun H M, et al. Light-YOLOv3: Fast method for detecting green mangoes in complex scenes using picking robots[J]. Applied Intelligence, 2020, 50(12): 4670-4687.
- [15] Wang J T, Xiao W, Ni T W. Efficient object detection method based on improved YOLOv3 network for remote sensing images[C]. The 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD). Chengdu, 2020: 242-246.
- [16] 柳长源, 王琪, 毕晓君. 多目标小尺度车辆目标检测方法[J]. 控制与决策, 2021, 36(11): 2707-2712.
(Liu C Y, Wang Q, Bi X J. Multi-target and small-scale vehicle target detection method[J]. Control and Decision, 2021, 36(11): 2707-2712.)
- [17] Ouyang L C, Wang H L. Aerial target detection based on the improved YOLOv3 algorithm[C]. The 6th International Conference on Systems and

- Informatics (ICSAI). Shanghai, 2019: 1196-1200.
- [18] Guo C X, Fan B, Zhang Q, et al. AugFPN: Improving multi-scale feature learning for object detection[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020: 12592-12601.
- [19] Noh J, Bae W, Lee W, et al. Better to follow, follow to be better: Towards precise supervision of feature super-resolution for small object detection[C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, 2019: 9724-9733.
- [20] Zhou P, Ni B B, Geng C, et al. Scale-transferrable object detection[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 528-537.
- [21] Zhu Y S, Zhao C Y, Guo H Y, et al. Attention CoupleNet: Fully convolutional attention coupling network for object detection[J]. IEEE Transactions on Image Processing, 2019, 28(1): 113-126.
- [22] Wang T C, Anwer R M, Cholakkal H, et al. Learning rich features at high-speed for single-shot object detection[C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, 2019: 1971-1980.
- [23] 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. Honolulu, 2017: 936-944.
- [24] Chen Q, Wang Y M, Yang T, et al. You only look one-level feature[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, 2021: 13034-13043.
- [25] Bodla N, Singh B, Chellappa R, et al. Soft-NMS — Improving object detection with one line of code[C]. 2017 IEEE International Conference on Computer Vision. Venice, 2017: 5562-5570.
- [26] Wang X L, Xiao T T, Jiang Y N, et al. Repulsion loss: detecting pedestrians in a crowd[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, 2018: 7774-7783.
- [27] 刘颖, 刘红燕, 范九伦, 等. 基于深度学习的小目标检测研究与应用综述[J]. 电子学报, 2020, 48(3): 590-601.
- (Liu Y, Liu H Y, Fan J L, et al. A survey of research and application of small object detection based on deep learning[J]. Acta Electronica Sinica, 2020, 48(3): 590-601.)
- [28] Xie L M, Huang C B. A residual network of water scene recognition based on optimized inception module and convolutional block attention module[C]. The 6th International Conference on Systems and Informatics (ICSAI). Shanghai, 2019: 1174-1178.
- [29] Chen Y T, Zhang X Z, Chen W N, et al. Research on recognition of fly species based on improved RetinaNet and CBAM[J]. IEEE Access, 2020, 8: 102907-102919.
- [30] Chakraborty S, Nagwani N K. Analysis and study of incremental K -means clustering algorithm[J]. Communications in Computer and Information Science, 2011, 169: 338-341.

作者简介

缪燕子(1981—), 女, 教授, 博士生导师, 从事自主移动机器人主动嗅觉、机器视觉、智能感知与信息融合等研究, E-mail: myz@cumt.edu.cn;

张宗伟(1996—), 男, 硕士生, 从事机器视觉、深度学习等研究, E-mail: zongwei_cumt@163.com;

王贺升(1980—), 男, 教授, 博士生导师, 从事视觉伺服、医疗机器人、无人驾驶、计算机视觉、自适应控制等研究, E-mail: wanghesheng@sjtu.edu.cn;

代伟(1984—), 男, 教授, 博士生导师, 从事人工智能驱动的复杂流程工业过程建模与运行优化控制、迁移学习、联邦学习、增量学习等研究, E-mail: weidai@cumt.edu.cn;

赵忠祥(1982—), 男, 博士生, 从事深度学习、迁移学习及数字孪生等研究, E-mail: 163zxx163@163.com;

王啸林(1998—), 男, 硕士生, 从事机器视觉、深度学习等研究, E-mail: 857187477@qq.com;

杨春雨(1979—), 男, 教授, 博士生导师, 从事智能系统控制理论、矿山装备优化控制、机器人导航与控制等研究, E-mail: chunyuyang@cumt.edu.cn;

史延诺(1998—), 男, 硕士生, 从事SLAM、深度学习等研究, E-mail: 1492841357@qq.com.