

基于 YOLOv5 的害虫识别研究

黄丽珊

(华南师范大学 数学科学学院, 广东 广州 510631)

摘要: 针对害虫图像数据中小目标误检、漏检、类别不平衡及特征提取能力不足等问题, 提出了一种改进的基于 YOLOv5 的害虫检测模型。首先, 该算法通过伪标签技术缓解标签数量不足带来的问题; 其次, 通过增加一个 4 倍的下采样层, 调整损失函数增强对少数类别的感知能力; 再次, 通过修改目标框回归公式解决训练过程中梯度消失的问题, 提升小目标的检测精度; 最后, 利用虫情测报灯采集的图像数据进行实例验证。实验结果表明, 该害虫检测模型具有较好的预测效果。

关键词: 害虫图像; 小目标检测; 伪标签; 损失函数; YOLOv5

中图分类号: TP391.4 文献标识码: A 文章编号: 1003-9767 (2022) 15-194-04

Pest Identification Based on YOLOv5

HUANG Lishan

(School of Mathematical Sciences, South China Normal University, Guangzhou Guangdong 510631, China)

Abstract: To solve the problems of false detection, missed detection, category imbalance and insufficient feature extraction ability of small targets in pest image data, an improved pest detection model based on YOLOv5 is proposed. Firstly, the algorithm uses pseudo-label technique to deal with insufficient labels. Secondly, a 4-fold down-sampling layer and focal loss function are introduced to catch minority categories better. Thirdly, the target box regression formula is modified to solve the problem of gradient disappearance in the training process and improves the accuracy of small targets detection. Finally, the image data collected by the pest detection lamp is used for experiment. The experimental results show that the pest detection model has a good prediction effect.

Keywords: pest image; small target detection; pseudo-label; loss function; YOLOv5

0 引言

虫情测报灯是一种虫情信息采集的智能设备, 可以实现自动诱集、杀虫、虫体分散以及拍照等作业, 并实时将虫情信息上传至云平台。由于农田害虫的多样性和信息类型的复杂性, 传统人工分类工作费事、费力、成本高、易出错, 基于图像识别与计算机视觉的害虫检测技术应运而生, 在实例分割、图像识别和目标检测等领域应用广泛, 特别在自动化害虫识别方面。

1 模型构建

近年来, 神经网络由于其优越的性能, 被广泛用于目标检测任务。文章选择了在速度和精度方面均有良好表现的 YOLOv5 架构。针对害虫数据集中标签缺失的问题, 利用伪

标签技术将 YOLOv5 模型改进为半监督学习模型, 充分利用未标注的图片信息。针对类别不平衡的问题, 引入了 Focal Loss 函数加强检测少数类。此外, 对于小目标检测的问题, 在 YOLOv5 的骨干网络中新增一个 4 倍下采样层来提取小目标信息。为了解决训练过程中梯度消失等问题, 本文还修改了目标框回归公式, 用于提升模型的检测精度^[1]。下面从伪标签技术、特征提取器、损失函数和目标框回归 4 个方面进行详细介绍。

1.1 YOLOv5 简介

YOLOv5 是一种单阶段目标检测算法。检测时无须使用候选框网络, 而是采用预定义候选框框定图像中的目标, 再对边框位置坐标进行回归, 最终输出检测信息, 大大降低

收稿日期: 2022-07-05

作者简介: 黄丽珊 (2001—), 女, 广东湛江人, 本科在读。研究方向: 数据挖掘与人工智能。

了模型的复杂度,加快了检测速度,使实时目标检测成为可能。图1展示了YOLOv5s的网络架构^[2]。其中,对流边界层(Convective Boundary Layer, CBL)是YOLO系列模型的基础模块之一,包含卷积层Conv、批量标准化模块BN和激活函数LeakyRELU。

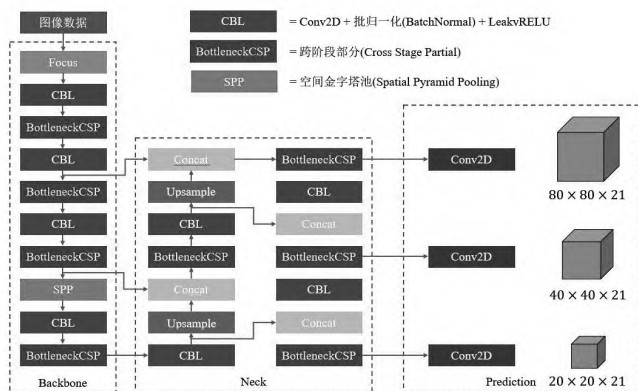


图1 YOLOv5 结构图

1.2 伪标签技术预测方法介绍

原始的YOLOv5模型是一种有监督的机器学习方法,训练模型时需要大量数据。若数据集存在严重的标签缺失问题,则会降低检测模型的鲁棒性。因此,通过伪标签技术预测未标注的数据,并输入检测模型,将原来的有监督学习过程转换为半监督学习过程,进而借助无标签数据提升有监督过程中的模型性能,其流程如图2所示^[3]。伪标签技术的算法过程如下:使用有标记的数据训练有监督模型M;使用有监督模型M预测无标签数据,得到预测概率P;根据预测概率P筛选高置信度样本;使用有标记数据和选出的伪标签数据训练新模型M';将模型M替换为M',重复以上步骤直到模型的效果不再提升。

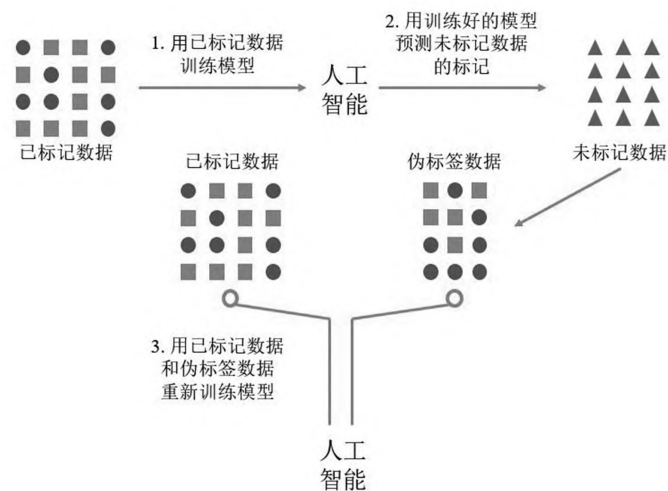


图2 伪标签算法示意图

1.3 小目标检测

小目标存在覆盖面积占比较小、分辨率较低、位置不准确以及特征不明显等特点,常常是目标检测的难点之一。

目标检测领域对小目标通常有两种定义:一种是相对定义,指定小目标占原图比例的阈值;另一种是绝对定义,如COCO数据集将尺寸小于 32×32 像素的目标定义为小目标^[4]。这里参考相对定义,将占原图的比例小于0.25%的目标定义为小目标。

原始的YOLOv5骨干网络通过8倍、16倍和32倍3种下采样方式从不同尺度提取目标特征,然后将特征图输入特征融合网络。根据特征金字塔网络的思想^[5],小目标的位置信息容易在多次卷积中丢失,因此难以检测,而浅层卷积虽然无法得到丰富的语义信息,但可以提供比较准确的位置信息。因此,这里在YOLOv5原始的骨干网络上增加一个4倍下采样层,通过特征融合网络将其结果与原有的下采样结果融合得到新的特征图,如图3所示。新尺度的特征图感受野较小,可以提供较多的位置信息,提升小目标的检测效果。

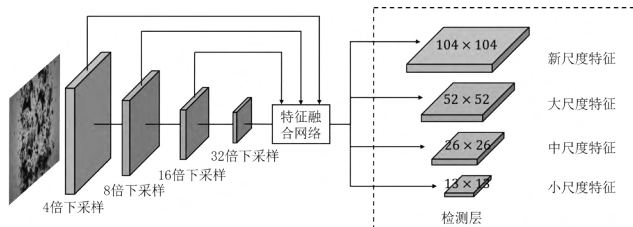


图3 采样过程

1.4 Focal Loss

数据存在类别不平衡的问题会导致样本数少的类别分类难度较高。尽管可以通过重采样的方式平衡类别,但数据集中包含的有效信息并不会因此增加。另外,原始的YOLOv5损失函数只考虑了类别损失、置信度损失和定位损失3种情况,并没有特别关注类别不平衡的问题。因此,在原始模型的基础上引入Focal Loss作为损失函数来减少类别不平衡带来的影响^[6]。

多分类的Focal Loss函数表达式为

$$FL(p_i) = -(1-p_i)^\gamma \log(p_i) \quad (1)$$

$$p_i = \begin{cases} \hat{p}, & y=1 \\ 1-\hat{p}, & y>1 \text{ 或 } y<1 \end{cases} \quad (2)$$

式中: t 为目标类别; γ 为常数; \hat{p} 为预测概率大小; y 为标签值。当 $\gamma=0$ 时,Focal Loss与普通的交叉熵函数 $CE(p_i)=-\log(p_i)$ 相同。相比交叉熵损失函数,当 γ 大于1时,Focal Loss增加了难分样本的权重,使损失函数倾向于难分样本,从而提高难分样本的准确度。

1.5 目标框回归

目标框回归的目的是寻找一种映射关系,使得候选目标框的映射无限接近真实目标框。在YOLOv5中,锚框给出了宽和高的初始值,回归的是目标真实的宽和高与初始值的偏移量,其原理如图4所示^[7]。其中,先验框用虚线表示,预测框用实线表示。先验框通过平移缩放得到预测框。

回归时,原始图片按照特征图的尺寸被划分成 $S \times S$ 个

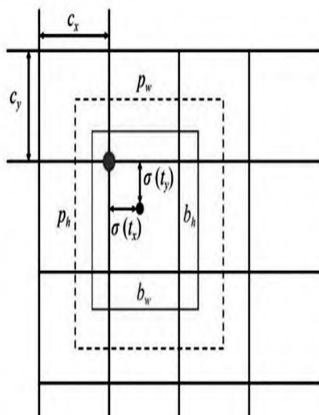


图4 目标框回归示意图

网格单元, 每个网格单元产生多个预测框。预测时, 由真实框中目标中心所在的网格预测该目标, 包括坐标信息和置信度。

目标框的坐标预测计算公式为

$$b_x = 2\sigma(t_x) - 0.5 + c_x \quad (3)$$

$$b_y = 2\sigma(t_y) - 0.5 + c_y \quad (4)$$

$$b_w = p_w [2\sigma(t_w)]^2 \quad (5)$$

$$b_h = p_h [2\sigma(t_h)]^2 \quad (6)$$

$$P_r(object) \times IOU(b, object) = \sigma(t_o) \quad (7)$$

式中: t_x 、 t_y 、 t_w 和 t_h 分别为偏移量; σ 为 Sigmoid 激活函数, 用于将偏移量 t_x 、 t_y 、 t_w 和 t_h 分别映射到 $[0,1]$; c_x 和 c_y 分别为单元网格中相对图片左上角的偏移量; p_w 和 p_h 分别为先验框的宽和高; b_x 、 b_y 、 b_w 和 b_h 分别为预测框的中心点坐标以及宽和高; $\sigma(t_o)$ 为预测框的置信度。对 $\sigma(t_o)$ 设定阈值, 过滤置信度较低的预测框, 再使用非极大值抑制算法可以从剩下的预测框中筛选最终的预测框^[8]。

2 实验过程与结果

2.1 数据来源与数据处理

本次实验选择从泰迪杯竞赛官网下载的数据, 共包含 3105 张图片。其中: 1637 张图片标注为不含虫子, 占图片总数的 54.30%; 576 张图片标注害虫的位置和类别信息, 用于训练目标检测模型; 802 张图片没有标注, 需要进行检测。含有标记的虫子类别如表 1 所示。

表1 变量关系表

单位: 个

虫子名称	标记数量	虫子名称	标记数量	虫子名称	标记数量
紫条尺蛾	247	粘虫	150	线委夜蛾	120
蟋蟀	90	甜菜夜蛾	87	甜菜白带夜蛾	59
水螟蛾	40	石蛾	27	歧角螟	21
棉铃虫	20	蝼蛄	20	黄足猎蝽	18
黄毒蛾	18	褐飞虱属	17	瓜绢野螟	15
干纹冬夜蛾	14	甘蓝夜蛾	12	二化螟	7
二点委夜蛾	6	豆野螟	5	地老虎	5
稻纵卷叶螟	4	稻螟蛉	4	大螟	4
大黑鳃金龟	4	草地螟	3	—	—
八点灰灯蛾	1	白背飞虱	1	—	—

图 5 是目标大小与整张图片大小的比例关系。其中, 大部分样本的宽和高占原图的比例都小于 5%, 即整个目标占原图的比例小于 0.25%, 给本次害虫检测工作带来了极大困难。因此, 将样本定义为小目标, 并利用前文提到的方法进行针对性处理。

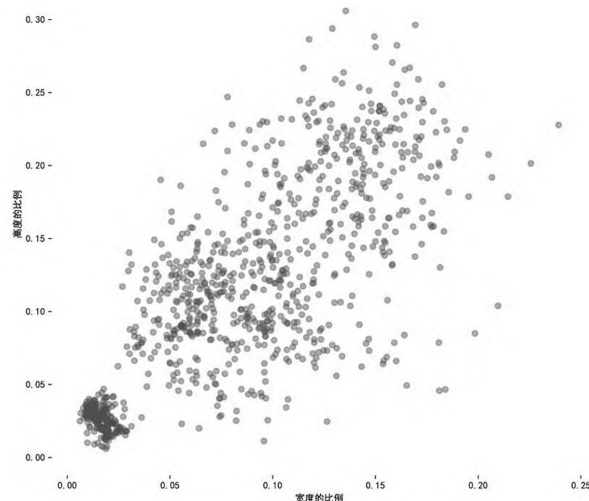


图5 检测目标与图片大小比例关系图

此外, 针对类别不平衡和数据量不足的问题, 采用重采样和 Mosaic 数据增强的方法。Mosaic 数据增强的算法框架如下: 从数据集中每次随机选取 4 张图片; 分别对 4 张图片进行裁剪、翻转、色域变化等操作; 按照左上、左下、右下、右上的顺序摆放变换后的图片; 利用矩阵截取对应的区域, 并拼接成一张新的图片。拼接后图片上含有目标检测框等内容, 且不同的图片之间会存在明显的分割线。如果原来的目标检测框超出分割线, 那么需要对超出的部分进行边缘处理。Mosaic 数据增强过程如图 6 所示。

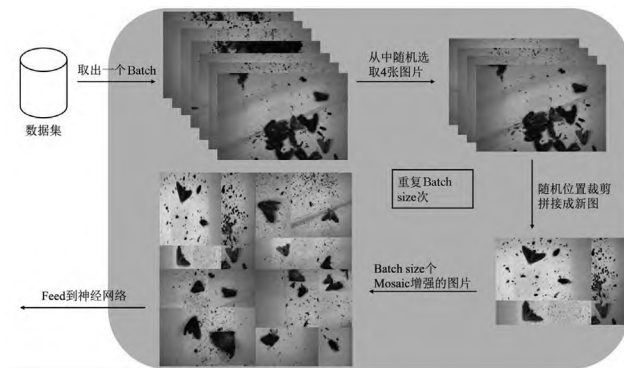


图6 Mosaic 数据增强过程

2.2 实验环境

本次实验使用 NVIDIA GeForce RTX 3090 显卡进行运算, 显存大小为 24 GB, 中央处理器 (Central Processing Unit, CPU) 配置为 AMD EPYC 7302, Pytorch 版本为 1.8.1, Python 语言环境为 3.8。

2.3 评价指标

实验结果采用平均精度均值 (mAP) 来评估模型的性能。

平均精度与精确率和召回率相关。精确率指预测正确的正例占所有预测为正例的比例；召回率指预测正确的正例占实际正例的比例。 mAP 值越大，表示算法检测效果越好，识别精度越高。

2.4 结果分析

2.4.1 伪标签结果

利用伪标签技术预测 1637 张标注为不含虫子的图片, 得到其中 558 张图片的标签、置信度和预测坐标。从中筛选置信度高于 0.95% 的 32 张图片作为伪标签样本, 与带有标签的样本一起输入神经网络进行训练, 筛选结果如图 7 所示。

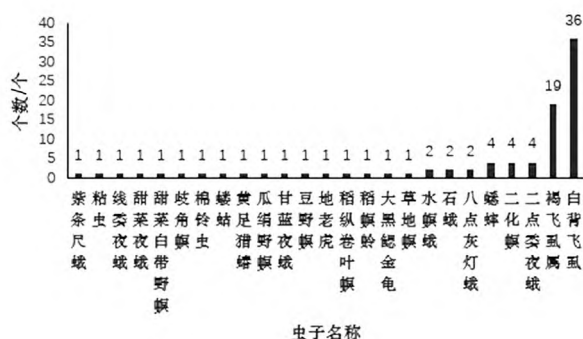


图7 伪标签结果

由图 7 可知, 预测结果含大量的白背飞虱和褐飞虱属, 约占伪标签样本总数的 61.1%, 而在原始数据中, 这两种害虫约占样本总数的 26.57%。白背飞虱和褐飞虱属体长最大只有 5 mm, 在数据集属于小目标数据, 容易漏检。但是, 此模型能捕捉如此多的小目标, 说明针对小目标的改进策略是十分有效的。

2.4.2 各类模型结果

本实验迭代批量大小为 16, 总迭代次数为 150 次, 优化器采用 Adamw^[9-10]。训练模型时采用动态学习率, 使用 Warmup 方法预热模型, 减缓初始阶段模型对小批量数据的过拟合现象, 避免模型振荡带来的不稳定。在 Warmup 阶段, 偏置层的学习率由 0.001 下降至 0.000 1, 其他的参数学习率由 0 增加至 0.000 1, Warmup 结束之后, 采用余弦退火算法更新学习率。模型之间的比较结果如表 2 所示。从表 2 可以看出, 使用 YOLO 模型越大, 检测能力越高。

表 2 评价指标结果

模型	精确率	召回率	$mAP@0.5$
YOLOv5s	38.82%	27.13%	25.39%
YOLOv5m	40.88%	28.44%	26.26%
YOLOv5x	41.52%	29.53%	27.95%
CBNetv2	41.01%	28.90%	26.70%

3 结语

在害虫检测任务上,利用YOLOv5算法可以有效提升网络的检测性能,其中通过训练一个网络给无标签的数据打上伪标签的方法带来的增益较大。对于小目标识别难问题,通过在YOLOv5原始的骨干网络上增加一个4倍下采样来提升小物体的检测能力。对于数据分布不平衡问题,采用对数量少的虫子重采样和引入Focal Loss来缓解。最后实测阶段,本文对比了YOLOv5s、YOLOv5m、YOLOv5x以及CBNetv2的检测结果说明改进的YOLOv5模型的有效性。通过观察伪标签结果,发现该模型非常容易将细小的杂物判断为尺寸小的害虫,说明在检测小目标方面具有较强能力,即不容易漏检,但是检测精度还有待提高。

参考文献

- [1] 郭磊,王邱龙,薛伟,等.基于改进YOLOv5的小目标检测算法[J].电子科技大学学报,2022,51(2):251-258.
- [2] ZHOU F,ZHAO H,NIE Z.Safety helmet detection based on YOLOv5[C]//2021 IEEE International Conference on Power Electronics,Computer Applications (ICPECA),2021:6-11.
- [3] 何颖,陈丁号,彭琳.基于改进YOLOv5模型的农林害虫目标检测算法研究[J].中国农机化学报,2022,43(4):106-115.
- [4] LIN T Y,MAIRE M,BELONGIE S,et al.Microsoft COCO:common objects in context[C]//Computer Vision-ECCV 2014,2014:740-755.
- [5] 李衡霞,龙陈锋,曾蒙,等.一种基于深度卷积神经网络的油菜虫害检测方法[J].湖南农业大学学报(自然科学版),2019,45(5):560-564.
- [6] 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.
- [7] 刘浏.基于深度学习的农作物害虫检测方法研究与应用[D].合肥:中国科学技术大学,2020:25.
- [8] NEUBECK A,VAN GOOL L.Efficient non-maximum suppression[C]//18th International Conference on Pattern Recognition(ICPR' 06),2006:850-855.
- [9] 韩瑞珍,何勇.基于计算机视觉的大田害虫远程自动识别系统[J].农业工程学报,2013,29(3):156-162.
- [10] 陈俭.基于卷积神经网络和度量学习的害虫检测方法研究[D].杭州:浙江大学,2021:26.