第42卷 第3期 华 中 农 业 大 学 学 报 Vol.42 No.3
2023年 5月 Journal of Huazhong Agricultural University May 2023,133～142

刘志,翟瑞芳,彭万伟,等.融合注意力机制的Cascade R-CNN田间害虫检测方法[J].华中农业大学学报,2023,42(3):133-142.
DOI:10.13300/j.cnki.hnlkxb.2023.03.016

融合注意力机制的Cascade R-CNN田间害虫检测方法

刘志[1],翟瑞芳[1],彭万伟[2],陈珂屹[3],杨万能[3]

1.华中农业大学信息学院,武汉 430070; 2.上海云农信息科技有限公司,上海 201299;
3.华中农业大学植物科学技术学院,武汉 430070

摘要 为解决测报灯采集图像中害虫依赖人工识别及统计结果可靠性低和准确性差的问题,本研究提出一种改进型Cascade R-CNN田间害虫检测算法。该算法以Cascade R-CNN为基础框架,采用ResNeSt-50作为主干网络,融合了跨通道注意力机制;使用统一目标检测头(unifying object detection heads with attentions,Dy-Head),并融合尺度感知、空间位置感知和任务感知。此外,采用简单复制-粘贴(simple copy-paste,SCP)方法进行了数据增强。研究共采集到20类害虫总计1 500张图像,制作了符合MS COCO格式(microsoft common objects in context 2017, MS COCO 2017)的测报灯田间害虫数据集。结果显示,本研究提出的方法的F_1分数(F_1-score)达到了86.2%。当交并比(intersection over union ,IoU)为0.5时,其F_1-分数与经典Cascade R-CNN、Faster R-CNN和YOLOv4相比,分别提升了2.8、5.8和8.2个百分点。表明该方法满足测报灯害虫检测任务对判别能力和实时性的要求,实现了害虫的高精度自动识别与计数,可直接应用于田间害虫检测。

关键词 深度学习;测报灯;害虫识别;Cascade R-CNN;精准检测;注意力机制;绿色防控

中图分类号 TP391.41;S126 **文献标识码** A **文章编号** 1000-2421(2023)03-0133-10

田间害虫以作物的组织为食,尤其是在功能叶的大部分光合作用尚未进行时吸食,导致作物严重减产[1]。由于害虫具有强繁殖能力和高破坏力,大规模虫害暴发会造成巨大损失。因此,在害虫尚未大量聚集时,及时采取防治措施,才能有效避免虫害的发生。传统的害虫检测通过人工统计完成,检测效率、可靠性与准确性均相对较低[2]。虫情测报灯能够实时采集田间害虫的RGB图像,使得害虫的自动化检测成为可能。因此,更多的研究者将精力集中在害虫的自动化检测上[3]。

早期的害虫自动化检测主要采用模式识别方法,将特征组合成特征向量后,根据特征相似度设计分类器[4-5]。张红涛等[6]在分离B通道灰度图像的基础上,提取多个特征进行归一化处理,然后再进行SVM分类,实现了昆虫性别的自动判别。潘梅等[7]提出了一种基于机器视觉的茶园害虫智能识别算法,该方法将图像依次进行HSV空间转换、阈值分割、SIFT特征提取、K-means聚类、SVM处理,得到了准确率较高的茶尺蠖分类模型。荆晓冉[8]选择7个形态学特征作为原始特征,组成特征矩阵后,利用改进的BP神经网络进行训练,得到了高精度的斜纹夜蛾分类器。尽管这些模式识别类算法已经获得了较好的分类性能,但它们具有明显的局限性,主要表现在受限于特定的害虫对象,泛化能力不强、鲁棒性差,在大规模推广时往往受限。

近年来,深度学习技术在计算机视觉领域取得重大突破[9-10],尤其是在目标检测任务中,卷积神经网络的应用已经取得显著的成效[11-13]。同时,越来越多的研究人员开始利用卷积神经网络来识别田间害虫[14-15],通过分析不同场景的需求,使用和改进各种算法,实现对害虫的高精度检测[16-17]。如姚青等[18]提出一种基于改进CornerNet的水稻灯诱飞虱自动检测方法,通过抑制去除冗余检测框,解决了水稻中飞虱自动检测存在的严重误检和漏检问题。张诗雨等[19]在建立粮仓害虫图像数据集SGI-6后,使用聚类算法改进Faster R-CNN模型的区域提案网

收稿日期:2023-02-13
基金项目:国家自然科学基金联合基金项目(U21A20205);中央高校基本科研业务费专项(2662022JC004)
刘志,E-mail:1451097369@qq.com
通信作者:翟瑞芳,E-mail:rfzhai@mail.hzau.edu.cn

络，实现了粮仓环境下害虫的检测识别。张博等[20]在识别算法中融合特征金字塔，并通过改进的YOLOv3深度卷积神经网络，提升了小目标害虫的检测效果。然而，在上述研究中，所涉及的害虫种类往往较为有限，且实验场景的设定主要集中在相对理想化的简化环境中进行。

本研究以测报灯在田间实际环境下采集的害虫为研究对象，针对图像中害虫种类多、密度大、面积占比相对偏小、可辨别性较弱等特性，提出一种融合注意力机制的Cascade R-CNN田间害虫检测算法，在使用简单复制-粘贴（simple copy-paste，SCP）[21]方法进行数据增强的基础上，采用ResNeSt-50[22]提取多维度特征，融合统一目标检测头（unifying object detection heads with attentions，DyHead）[23]，以期实现害虫的低误差自动识别与计数，为测报灯的自动化、智能化处理奠定基础。

1 材料与方法

1.1 图像的采集

为建立算法所用的数据集，本研究采用上海云农智联生产的智能虫情测报灯作为采集装置，该装置结构如图1 A所示。测报灯采用灯管作为诱虫光源，害虫被光源吸引后撞击玻璃屏，然后落入烤虫仓内。在烤虫仓内，害虫经过远红外加热处理被杀灭。装置内的2个烤虫仓可以确保害虫被完全杀死，害虫随后落入接虫仓，相机拍摄害虫产生图像后自动上传到相应服务器上。图像分辨率采用2 112像素×1 568像素。图像采集地点分别位于湖北省恩施市蓝焙茶叶病虫害观测场、广西南宁市隆安县水稻病虫害数字化观测试验场和中科院黄河三角洲研究中心玉米观测场，中科院黄河三角洲研究中心玉米观测场实际采集环境如图1 B、C所示。

图1 测报灯结构（A）和拍摄环境（B、C）

Fig.1 Structure diagram (A) and shooting environment of telemetering lamp (B, C)

图像数据于2018年4月15日至2020年5月20日采集，共采集到20种害虫图像数据，具体害虫信息如图2所示。拍摄时，摄像头以距离拍摄平板25 cm的高度进行拍摄，每15 min采集1次图像并自动上传至服务器，共获取了10 000幅图像。

1.2 数据集的制作

随机选择1 500幅图像建立高质量的数据集，使用开源工具Labelme对害虫图片进行人工标注，昆虫种类标签和图2中序号一一对应。

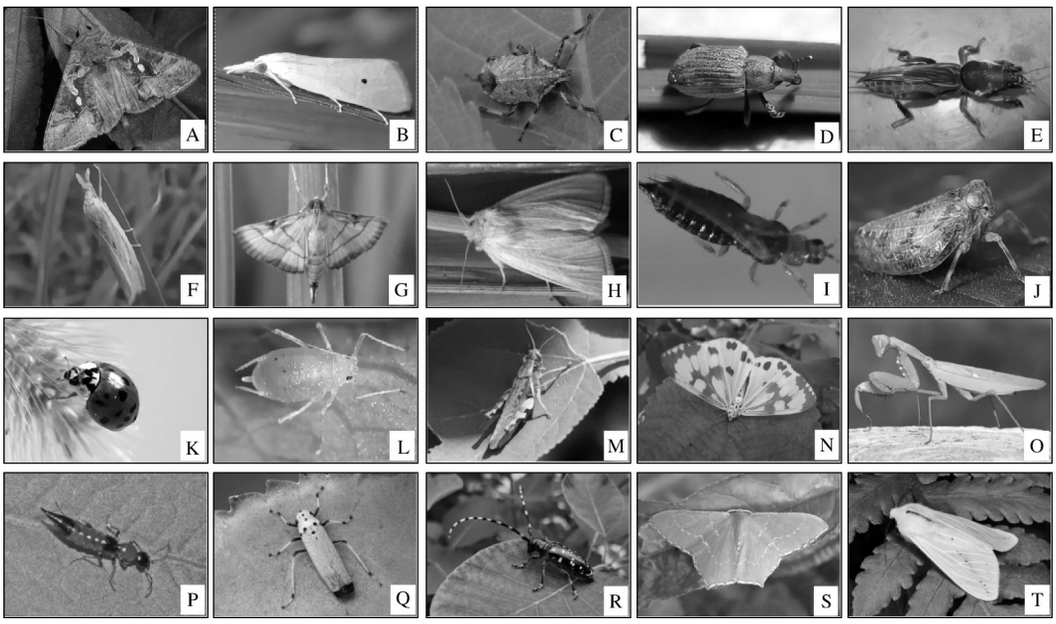
标注完成的所有图片制作成MS-COCO 2017格式的数据集，按8∶1∶1的比例随机将数据集中所有图片分配为训练集、验证集和测试集。其中，训练集用于对模型算法进行训练，验证集用于对模型超参数进行调整，测试集则用于对模型的识别能力做最终的评估。本研究采用COCO数据集检测评估指标对算法进行评估，即IoU（交并比，intersection over union，IoU）阈值为0.5的精确率（precision）进行评估。

1.3 SCP数据增强

在模型的训练过程中，为了提高模型的泛化能力，本研究引入SCP数据增强方案[21]。SCP方法在训练过程中将1张图片中的害虫实例直接复制粘贴到另一张图像中，从而生成新的训练样本。如图3所示，SCP创造出了场景更加复杂的新数据以显著提升全监督方式训练模型的性能，提高了模型的泛化能力。

1.4 改进型Cascade R-CNN田间害虫检测模型

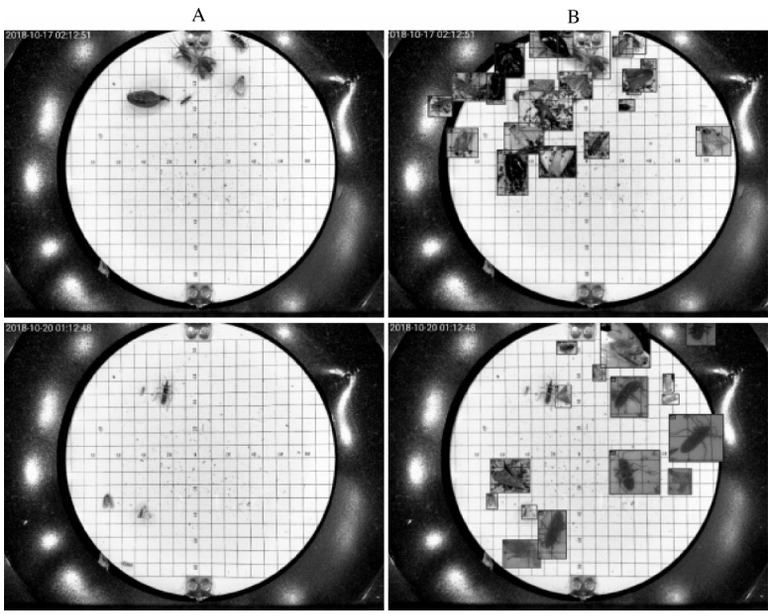
深度学习检测算法至今已发展出基于锚框（anchor-based）和非锚框（anchor-free）两大技术分支。其中，基于锚框的技术在图像检测领域得到了更广泛的应用。基于锚框分为单阶段模型、双阶段模型



A：水稻金翅叶蛾 *Cnaphalocrocis medinalis*；B：水稻三化螟 *Scirpophaga incertulas*；C：稻蝽 *Scotinophara lurida*；D：稻象甲 *Echinocnemus squameus* Billberg；E：蝼跍 *Gryllotalpa spps*；F：水稻二化螟 *Cnaphalocrocis medinalis*；G：稻纵卷叶螟 *Cnaphalocrocis medinalis*；H：水稻大螟 *Sesamia inferens*；I：稻蓟马 *Stenchaetothrips biformis*；J：稻飞虱 *Nilaparvata lugens Stal*；K：七星瓢虫 *Coccinella septempunctata*；L：蚜虫 *Aphididae*；M：蝗虫 *Locusta migratoria manilensis* Meyen；N：粉蝶灯蛾 *Nyctemera adversata*；O：螳螂 *Mantodea*；P：隐翅虫 *Oxytelus batiuculus*；Q：稻叶蝉 *Nephotettix cincticeps*；R：天牛 *Cerambycidae*；S：青尺蛾 *Nehipparchus hypoleuca*；T：灯蛾 *Artiae caja*.

图 2　20 类害虫图像

Fig.2　Images of 20 types of pests



A：原始图像 Original images；B：简单复制-粘贴 Simple copy-paste.

图 3　SCP 数据增强结果

Fig.3　SCP data enhancement results

和多阶段模型。随着阶段数的提高，锚框检测算法的精度也随之提升。多阶段模型的代表网络 Cascade R-CNN[24]在 COCO 数据上的平均精确率（average precision）达 42.8%，高于双阶段模型 Faster RCNN（faster region convolutional neural networks）[9]的 34.9% 和单阶段模型 RetinaNet[25]的 39.1%。从检测精度的角度出发，本研究选择多阶段模型中的 Cascade R-CNN 作为基础算法。

为适应田间复杂环境下的害虫识别需求，本研究在Cascade R-CNN的基础框架之上，提出一种融合注意力机制的Cascade R-CNN，使用ResNeSt-50作为骨干网络来提取害虫图像特征。在不增加额外计算开销的情况下，通过引入统一目标检测头(unifying object detection heads with attentions，DyHead)，提高害虫目标检测模型头部的表达能力。此外，在模型训练阶段，引入简单复制-粘贴(simple copy-paste，SCP)数据增强方法，以提高模型的特征学习效果和泛化能力。改进型Cascade R-CNN结构如图4所示。

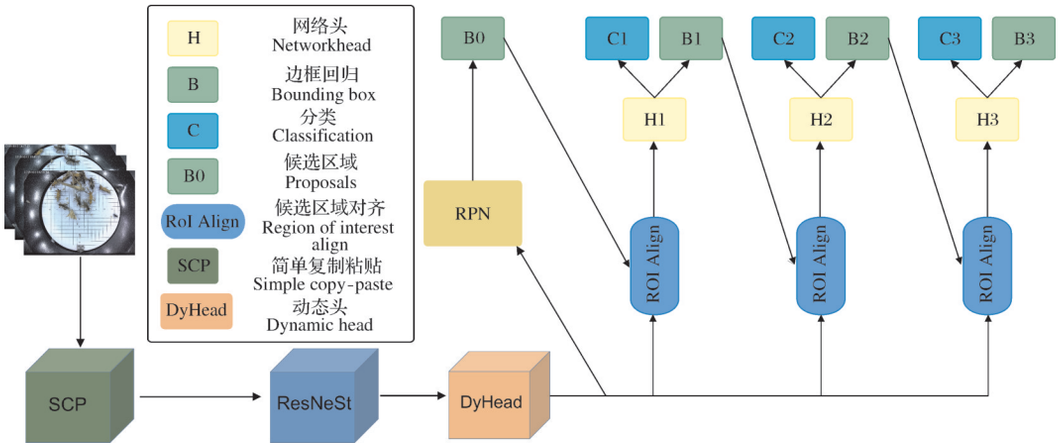


图4 改进型Cascade R-CNN结构

Fig.4 Improved Cascade R-CNN structure diagram

1.5 Cascade R-CNN检测模型

Cascade R-CNN在双阶段目标检测模型Faster R-CNN的基础上，通过在第二阶段ROI(region of interest)目标检测部分级联多个不同IoU阈值的检测网络，可以对预测结果不断优化，以提高目标识别的准确度。与普通级联网络不同的是，Cascade R-CNN级联结构中的每个检测模型都基于不同IoU阈值的正负样本训练得到，前一个检测网络输出的边界框作为下一个检测网络的输入。这样级联后的检测器接收的正负样本的IoU阈值越高，可获得的样本的质量越高，使得检测可以获得高质量的正样本来进行训练识别，提高了网络对目标信息的获取能力。随着输入的检测器接收到的正负样本IoU阈值不断上升，每个阶段的检测器都专注于检测某个IoU范围内的感兴趣目标候选区域，从而提高检测效果。

在图4中，本研究使用的Cascade R-CNN级联网络位于区域生成网络(region proposal network，RPN)后部，级联的3个检测器IoU阈值分别为0.5、0.6、0.7。害虫图片输入Cascade R-CNN检测网络后，经过ResNeSt-50提取特征图，然后将特征图输入到2个分支：一个分支输入到候选区域对齐模块，另一个分支输入到RPN网络，通过RPN网络预测得到图中的B0，即特征图中的特征坐标。在第一阶段，根据坐标提取局部特征图输入到H1中，进行分类和回归得到C1和B1；在第二阶段，根据B1的坐标在特征图上提取局部特征，并将提取的局部特征输入到H2进行分类和回归，得到C2和B2；在第三阶段，根据B2的坐标在特征图上提取局部特征，然后将提取到的特征输入到H3进行分类和回归，得到C3和B3作为最终结果。

1.6 ResNeSt-50骨干网络

在田间害虫的识别中，需要识别多个种类的害虫目标。ResNeSt网络在ResNet网络结构的基础上对卷积部分进行了重构，引入了拆分注意力机制(split attention)，能够在不改变整体特征通道数和输出特征图大小的情况下，使网络提取到更丰富的害虫特征信息。为了提高算法模型提取不同种类害虫目标特征信息的能力，本研究使用ResNeSt-50作为Cascade R-CNN中用于提取图像特征的骨干网络，在ImageNet数据集上，其准确率达到80.64%，分别高于经典的网络ResNet-50和ResNeXt-50的76.15%和77.77%[22]。

在ResNeSt模块中对害虫特征图的通道维数进行统计，并将通道维数分为k个簇(cardinal)。每个簇内部根据通道维数再分为r组，对每组的输入张量数据依次使用1×1卷积核和3×3卷积核进行处理，随后再将所有组输出的张量数据送入拆分注意力机制(split attention)处理。最后进行拼接(concatenate)，将k个簇的特征通道进行融合，输出的数据使用1×1卷积核进行卷积。其结构如图5所示。

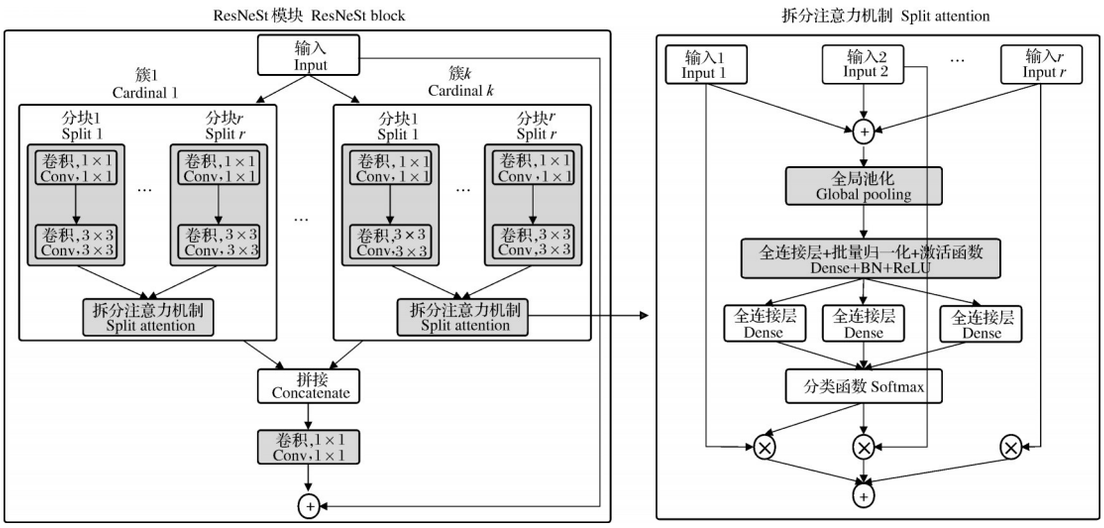


图5　ResNeSt结构

Fig.5　ResNeSt structure diagram

1.7　DyHead结构

目前主流的深度学习模型在骨干网络提取到足够的特征信息之后，采用特征金字塔先将提取到的特征送入模型头部，再进行进一步的分类和回归，得到最终的检测结果。在田间环境多种类害虫检测研究中，为了达到高精度检测的目的，需克服以下挑战。首先，在采集的图像中，由于害虫种类较多且处于不同发育阶段，使得多个尺度不一的目标出现在1张图像上，因此检测算法需要具有尺度感知能力。其次，在特征金字塔不同尺度的特征图中，目标通常以不同的形状和位置出现，检测算法需要感知目标在不同空间位置之间的关系。为了不增加额外计算开销，本研究引入DyHead来提高害虫目标检测模型头部的表示能力，DyHead结构如图6所示。

与多数深度学习检测算法中对特征金字塔的处理方式不同，DyHead使用注意力机制来统一目标检测头部，通过在尺度感知的特征层之间、空间感知的空间位置之间以及任务感知的输出通道内连贯地结合多种自注意力机制(self-attention)。DyHead在害虫图像特征的层序(level)维度处理上，学习多个语义层之间的相对重要性，根据害虫目标的尺度大小在适当的特征层次上增强特征。空间感知(spatial-aware)注意力模块只在害虫图像特征的空间维度上处理，在空间位置上学习不同害虫特征的区别。任务感知(task-aware)注意力模块处理通道上的害虫特征数据，根据害虫特征在不同卷积核的卷积计算输出差别，引导不同的特征通道分别识别不同的害虫种类。DyHead可以将检测器在COCO数据集上的平均精确率(average precision)提高1.2%～3.2%[23]。

1.8　损失函数

由于级联结构的引入，训练阶段中的损失函数在Faster R-CNN损失函数的基础上，引入了递归函数，损失总体表达式如式(1)和式(2)所示。

$$L(x^t,\boldsymbol{g})=L_{\mathrm{cls}}\left(h_t(x^t),y^t\right)+\lambda\left[y^t\geqslant 1\right]L_{\mathrm{loc}}\left(f_t(x^t,\boldsymbol{b}^t),\boldsymbol{g}\right)\quad(1)$$

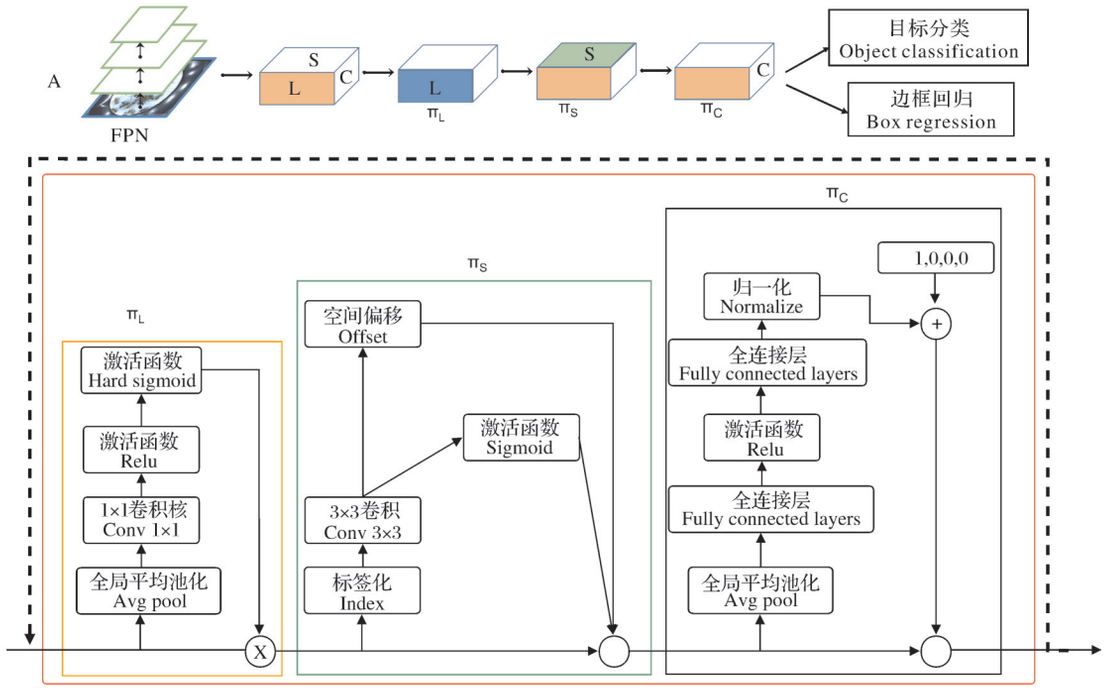
$$\boldsymbol{b}^t=f_{t-1}\left(x^{t-1},\boldsymbol{b}^{t-1}\right)$$

$$L_{\mathrm{cls}}(p_i,y)=\sum_{u\in \mathrm{U}}L_{\mathrm{cls}}\left(p_i{}^u,y^u\right)\quad(2)$$

在式(1)中，t表示级联结构中所处的级联结构次序数，在本研究中t值为4[24]。L_{cls}为经典的交叉熵损失函数，输入p_i为t阶段时物体类别的预测概率值。y为在t阶段时的样本标签，值为0表示负样本，值为1表示正样本。x为网络中输入的张量数据，$\boldsymbol{b}$为输入图片中网络预测的目标边界框，$\boldsymbol{g}$为输入图片中的目标真实标注，[·]是指示函数，λ取值1。式(2)中$\boldsymbol{U}$为IoU阈值的集合，$\boldsymbol{U}=\{0.5,\ 0.55,\ \cdots,\ 0.75\}$。$p_i^u$和$y^u$为不同IoU阈值时的类别预测概率值和样本标签。

1.9　软硬件平台及模型训练

本研究中所用操作系统为Ubuntu 20.04，深度学习模型框架为Pytorch 1.7.1，使用Python 3.7.9作为整体系统软件编写语言。改进型Cascade R-CNN算法实验所用服务器硬件平台配置为：处理器Intel(R) Core(TM)i9-9900k，主频为3.6 GHz，内存32 GB，显



A:DyHead方法示例 An illustration of DyHead approach; B:DyHead模块详细设计 Detailed design of DyHead.

图6 DyHead原理图

Fig.6 Schematic diagram of DyHead

卡NVIDIA GeForce RTX 3090(24 GB)。

本试验采用MS COCO2017格式的田间害虫数据集。根据迁移学习方法,改进型Cascade R-CNN使用在ImageNet上经过预训练的权重文件,在训练集中对网络进行训练,通过测试集测试模型的效果,比较不同算法的性能。

使用随机梯度下降法(SGD)训练模型,权值衰减系数为0.000 05,动量参数为0.9,基于单RTX 3090显卡的batch size大小为3。学习率为0.002 5,设定训练迭代(epochs)为36,在训练过程中采用动态机制调整学习率,初始学习率设为0.001,在前500次迭代中线性增加至0.002 5后保持不变,在第28至第34期减少为初始值的1/10。

1.10 模型评估方法

采用召回率(recall,R)、精确率(precision,P)和F_1分数(F_1-score,F_1)等指标衡量模型性能,取值区间为[0,1]。在召回率、精确率和F_1分数的计算公式(式3、式4和式5)中,T_P表示正样本被正确识别为正样本,F_P表示负样本被错误识别为正样本,F_N表示正样本被错误识别为负样本。

$$P=\frac{T_P}{T_P+F_P} \tag{3}$$

$$R=\frac{T_P}{T_P+F_N} \tag{4}$$

$$F_1=\frac{2T_P}{2T_P+F_P+F_N} \tag{5}$$

2 结果与分析

2.1 损失函数

图7中展示了训练过程中损失函数的变化情况。从图7可以看出,模型在经过5 000次迭代后基本稳定,模型参数得到充分地学习,在后续的训练中并未出现大的波动。

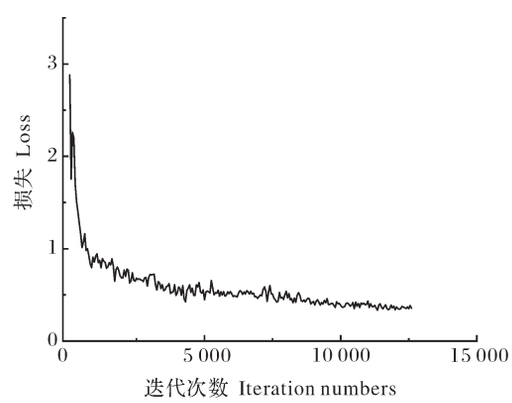


图7 训练过程损失函数变化曲线

Fig.7 Loss curves of training

2.2　消融实验

为了验证算法中改进部分的有效性，本研究分别针对 SCP、ResNeSt-50 和 DyHead 完成了消融实验，结果如表 1 所示。从表 1 可知，SCP 的引入使得模型的精确率提高了 1.6 个百分点，显示 SCP 创造出了场景更加复杂的新数据，算法模型对于本研究场景下害虫识别的泛化能力得到了提升。将骨干网络从 ResNet-50 更换为 ResNeSt-50 之后，其精确率、召回率和 F_1 分数分别提高了 0.9、0.7 和 0.8 个百分点。试验结果表明，骨干网络的改进使得模型能够更有效地提取害虫图像特征，最终提高模型的检测精度。DyHead 的引入使得模型的检测精确率、召回率和 F_1 分数分别提高了 1.0、1.1 和 1.0 个百分点。这说明在不增加额外计算开销的情况下，DyHead 的引入使得检测模型头部能够更有针对性地表示害虫目标的特征，进而提升模型的检测精度。图 8 展示了 Cascade R-CNN 和改进型 Cascade R-CNN 的实际识别效果。从图 8 中可以看出，改进型 Cascade R-CNN 改善了害虫密集分布情形下原有 Cascade R-CNN 存在漏检和误检的问题。

表 1　消融实验数据

Table 1　Pest identification results　　%

模型 Model	特征提取网络 Backbone	精确率 Precision	召回率 Recall	F_1 分数 F_1-score
Cascade R-CNN	ResNet-50	88.6	78.7	83.4
Cascade R-CNN	ResNeSt-50	89.7	79.4	84.2
Cascade R-CNN＋SCP	ResNet-50	90.2	80.5	85.1
Cascade R-CNN＋DyHead	ResNet-50	89.6	79.8	84.4
Improved Cascade R-CNN	ResNeSt-50	91.4	81.6	86.2

由表 1 可知，改进型 Cascade R-CNN 的精确率、召回率和 F_1 分数比原始的 Cascade R-CNN（特征提取网络为 ResNet-50）分别提高了 2.8、2.9 和 2.8 个百分点。试验结果表明改进的 Cascade R-CNN 相较于原始模型更好地适应了害虫检测任务，检测性能得到了提升和优化。

2.3　对比试验

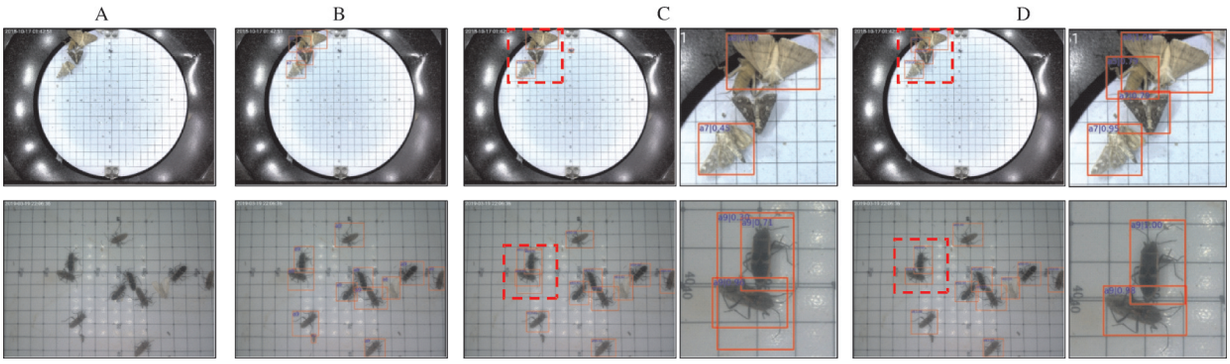
在本研究建立的测报灯害虫数据集中，除了消融试验外，分别利用改进型 Cascade R-CNN、Cascade R-CNN、Faster R-CNN、YOLOv4[26]、YOLOX[27] 和 Cascade RPN[28] 进行训练并评估识别精度。表 2 为不同模型的对比试验数据。本研究将改进型 Cascade R-CNN 与双阶段经典模型 Faster R-CNN 和单阶段经典模型 YOLOv4 以及现行的 YOLOX 和 Cascade RPN 进行对比。所有网络的训练与评估均在本研究害虫数据集中完成。从表 2 数据可知，改进后的 Cascade R-CNN 的精确率、召回率和 F_1 分数比 Faster R-CNN 分别提高了 5.6%、5.9% 和 5.8%，相较于 YOLOv4 精确率、召回率和 F_1 分数分别提高了 7.8、8.5 和 8.2 个百分点。在与现行模型比较中，改进后的 Cascade R-CNN 的精确率、召回率和 F_1 分数比 YOLOX 分别提高了 5.0、4.3 和 4.6 个百分点，比 Cascade RPN 分别提高了 4.2、3.5 和 3.8 个百分点。试验结果表明与其他代表性模型相比，改进的 Cascade R-CNN 获得了最优性能，且对不同种类的害虫检测均具有较高的鲁棒性和精确度。改进的 Cascade R-CNN 达到采集装置中害虫检测的精度要求。

表 2　不同模型对比试验数据

Table 2　Comparing experimental data of different models　　%

模型 Model	特征提取网络 Backbone	精确率 Precision	召回率 Recall	F_1 分数 F_1-score
Faster R-CNN	ResNet-50	85.8	75.7	80.4
YOLOv4	DarkNet53	83.6	73.1	78.0
YOLOX	DarkNet53	86.4	77.3	81.6
Cascade RPN	ResNet-50	87.2	78.1	82.4
Improved Cascade R-CNN	ResNeSt-50	91.4	81.6	86.2

图 8 展示了 Cascade R-CNN 和改进型 Cascade R-CNN 的实际识别效果，图 8 中包含了 2 组害虫图像的检测结果对照和相应的识别细节对照。从图 8C、8D 中看出，当害虫分布密集，存在粘连、遮挡和覆盖时，原始 Cascade R-CNN 模型会存在漏检和误检，而改进型 Cascade R-CNN 则在一定程度上解决了此类问题。原有模型的检测性能通过本研究工作得到了提升。综上，本研究提出的改进型 Cascade R-



A:原始图像 Original images; B:真实标签 True label; C:Cascade R-CNN; D:Cascade R-CNN+SCP+ResNeSt-50+DyHead.

图8 害虫检测效果对比图

Fig.8 Image of the pest detection effect comparison

CNN实现了检测类别多样化的目标,满足了业界对于田间害虫检测的精度需求。

3 讨 论

本研究针对田间害虫检测问题,在Cascade R-CNN模型基础上针对多目标图像特征学习进行改进,提出了一种融合注意力机制的Cascade R-CNN害虫检测方法。该方法采用ResNeSt-50作为骨干网络提取害虫图像特征,引入DyHead以增强害虫目标检测的模型头部表示能力;通过应用SCP提升模型的泛化能力。本研究提出的改进型Cascade R-CNN在F_1-score上达到了86.2%,与经典的Cascade R-CNN、Faster R-CNN和YOLOv4相比分别提升了2.8、5.8和8.2个百分点。这一结果表明所提出的方法可满足田间害虫检测的精度要求,可为后续田间害虫精准检测装备的研究奠定基础。

在实际应用场景中,害虫密集分布会导致虫体堆积和遮挡问题,给识别带来一定的困难。为了应对这一挑战,后续研究可以探索更先进的模块或算法,以提高模型在复杂场景下的性能。此外,为了进一步提升模型的泛化能力,可以考虑在数据集中增加更多种类害虫的图像数据。在采集装置研发方面,针对虫体堆积和遮挡问题,可为测报灯添加预警装置,在虫量超出设定阈值时产生反馈信号并自行调节拍摄频率,以获取遮挡较少的灯诱害虫图像,从而降低图像处理难度,以提升模型在实际应用中的稳定性,为农业生产带来实质性的益处。

参考文献 References

[1] 盛承发.间接性害虫为害与作物产量损失的关系Ⅰ.食叶害虫[J].应用生态学报,1993,4(2):192-197.SHENG C F.Relationship of crop yield to feeding injury by indirect insect and mite pests. Ⅰ.Leaf eating insect pests[J].Chinese journal of applied ecology,1993,4(2):192-197 (in Chinese with English abstract).

[2] 李改完,王艳,冀晓燕.基层病虫测报工作存在问题及对策[J].现代农村科技,2011(7):4-5.LI G W,WANG Y,JI X Y.Problems and countermeasures of grass-roots pest forecasting work[J].Modern agricultural science and technology,2011(7):4-5 (in Chinese).

[3] DING W G,TAYLOR G.Automatic moth detection from trap images for pest management[J].Computers and electronics in agriculture,2016,123:17-28.

[4] 杨红珍,张建伟,李湘涛,等.基于图像的昆虫远程自动识别系统的研究[J].农业工程学报,2008,24(1):188-192.YANG H Z,ZHANG J W,LI X T,et al.Remote automatic identification system based on insect image[J].Transactions of the CSAE,2008,24(1):188-192 (in Chinese with English abstract).

[5] 张建伟,王永模,沈佐锐.麦田蚜虫自动计数研究[J].农业工程学报,2006,22(9):159-162.ZHANG J W,WANG Y M,SHEN Z R.Novel method for estimating cereal aphid population based on computer vision technology[J].Transactions of the CSAE,2006,22(9):159-162 (in Chinese with English abstract).

[6] 张红涛,刘迦南,谭联,等.基于计算机视觉的棉铃虫成虫雌雄自动判别研究[J].环境昆虫学报,2019,41(4):908-913.ZHANG H T,LIU J N,TAN L,et al.Study on utomatic discrimination of male and female imagoes of *Helicoverpa armigera* (Hübner) based on computer vision[J].Journal of environmental entomology,2019,41(4):908-913 (in Chinese with English abstract).

[7] 潘梅,李光辉,周小波,等.基于机器视觉的茶园害虫智能识别系统研究与实现[J].现代农业科技,2019(18):229-230,233.PAN M,LI G H,ZHOU X B,et al.Research and implementation of intelligent recognition system for tea garden pest based on machine vision[J].Modern agricultural science and technology,2019(18):229-230,233 (in Chinese with English abstract).

[8] 荆晓冉.基于图像的害虫自动计数与识别系统的研究[D].杭

州：浙江大学，2014.JING X R.Study on automatic pest counting and identification system based on image[D]. Hangzhou：Zhejiang University，2014（in Chinese with English abstract）.

[9] REN S Q，HE K M，GIRSHICK R，et al. Faster R-CNN：towards real-time object detection with region proposal networks [C]//Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1.December 7 - 12，2015，Montreal，Canada.New York：ACM，2015：91 - 99.

[10] QIAO S Y，CHEN L C，YUILLE A.DetectoRS：detecting objects with recursive feature pyramid and switchable atrous convolution[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition（CVPR），June 20-25，2021，Nashville，TN，USA.Nashville：IEEE，2021：10208-10219.

[11] WANG Z B，WANG K Y，LIU Z Q，et al. A cognitive vision method for insect pest image segmentation[J].IFAC-PapersOnLine，2018，51(17)：85-89.

[12] SUN Y，LIU X X，YUAN M S，et al.Automatic in-trap pest detection using deep learning for pheromone-based *Dendroctonus valens* monitoring [J]. Biosystems engineering，2018，176：140-150.

[13] LIU L，WANG R J，XIE C J，et al.PestNet：an end-to-end deep learning approach for large-scale multi-class pest detection and classification[J].IEEE access，2019，7：45301-45312.

[14] 甘雨，郭庆文，王春桃，等.基于改进EfficientNet模型的作物害虫识别[J].农业工程学报，2022，38(1)：203-211.GAN Y，GUO Q W，WANG C T，et al.Recognizing crop pests using an improved EfficientNet model [J]. Transactions of the CSAE，2022，38(1)：203-211（in Chinese with English abstract）.

[15] 张善文，许新华，齐国红，等.基于可形变VGG-16模型的田间作物害虫检测方法[J].农业工程学报，2021，37(18)：188-194.ZHANG S W，XU X H，QI G H，et al.Detecting the pest disease of field crops using deformable VGG-16 model[J].Transactions of the CSAE，2021，37(18)：188-194（in Chinese with English abstract）.

[16] 鲍文霞，吴德钊，胡根生，等.基于轻量型残差网络的自然场景水稻害虫识别[J].农业工程学报，2021，37(16)：145-152.BAO W X，WU D Z，HU G S，et al.Rice pest identification in natural scene based on lightweight residual network[J].Transactions of the CSAE，2021，37(16)：145-152（in Chinese with English abstract）.

[17] 蔡润基，江方湧，郑涛涛，等.深度模型融合数据合成机制的长尾目标识别[J/OL].华中农业大学学报：1-10[2023-02-13].http://kns.cnki.net/kcms/detail/42.1181.S.20230112.1916.002.html.CAI R J，JIANG F Y，ZHENG T T，et al. Synthetic samples combined model-based recognition of long-tailed target[J/OL].Journal of Huazhong Agricultural University，2023：1-10[2023-02-13]. http://kns. cnki. net/kcms/detail/42.1181.S.20230112.1916.002.html（in Chinese with English abstract）.

[18] 姚青，吴叔珍，蒯乃阳，等.基于改进CornerNet的水稻灯诱飞虱自动检测方法构建与验证[J].农业工程学报，2021，37(7)：183-189.YAO Q，WU S Z，KUAI N Y，et al.Automatic detection of rice planthoppers through light-trap insect images using improved CornerNet[J].Transactions of the CSAE，2021，37(7)：183-189（in Chinese with English abstract）.

[19] 张诗雨，夏凯，杜晓晨，等.一种基于聚类特征的Faster R-CNN粮仓害虫检测方法[J].中国粮油学报，2020，35(4)：165-172.ZHANG S Y，XIA K，DU X C，et al.A faster R-CNN method for insect detection in granary based on clustering feature[J].Journal of the Chinese cereals and oils association，2020，35(4)：165-172（in Chinese with English abstract）.

[20] 张博，张苗辉，陈运忠.基于空间金字塔池化和深度卷积神经网络的作物害虫识别[J].农业工程学报，2019，35(19)：209-215.ZHANG B，ZHANG M H，CHEN Y Z.Crop pest identification based on spatial pyramid pooling and deep convolution neural network[J].Transactions of the CSAE，2019，35(19)：209-215（in Chinese with English abstract）.

[21] GHIASI G，CUI Y，SRINIVAS A，et al.Simple copy-paste is a strong data augmentation method for instance segmentation [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition（CVPR），June 20-25，2021，Nashville，TN，USA.Nashville：IEEE，2021：2917-2927.

[22] ZHANG H，WU C R，ZHANG Z Y，et al.ResNeSt：split-attention networks[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops（CVPRW），June 19-20，2022.New Orleans，LA，USA：IEEE，2022：2735-2745.

[23] DAI X Y，CHEN Y P，XIAO B，et al.Dynamic head：unifying object detection heads with attentions[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition（CVPR），June 20-25，2021.Nashville，TN，USA：IEEE，2021：7369-7378.

[24] CAI Z W，VASCONCELOS N.Cascade R-CNN：delving into high quality object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition，une 18-23，2018，Salt Lake City，UT，USA.Salt Lake City：IEEE，2018：6154-6162.

[25] 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：IEEE，2017：2999-3007.

[26] BOCHKOVSKIY A，WANG C Y，LIAO H Y M.YOLOv4：optimal speed and accuracy of object detection[DB/OL]. arXiv，2020：2004.10934.https://doi.org/10.48550/arXiv.2004.10934.

[27] GE Z，LIU S T，WANG F，et al.YOLOX：exceeding YOLO series in 2021[DB/OL].arXiv，2021：2107.08430.https://doi.org/10.48550/arXiv.2107.08430.

[28] VU T，JANG H，PHAM T X，et al.Cascade RPN：delving into high-quality region proposal network with adaptive convolution [DB/OL]. arXiv，2019：1909.06720. https://doi. org/10.48550/arXiv.1909.06720.

Field pest detection method based on improved Cascade R-CNN by incorporating attention mechanism

LIU Zhi[1], ZHAI Ruifang[1], PENG Wanwei[2], CHEN Keyi[3], YANG Wanneng[3]

1.*College of Informatics*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*;
2.*Shanghai Yunnong Information Technology Co.*, *Ltd.*, *Shanghai* 201299, *China*;
3.*College of Plant Science and Technology*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*

Abstract In order to address the challenges of manual identification of pests in images collected by light traps, as well as the low reliability and poor accuracy of statistical results, this study proposes an improved Cascade R-CNN algorithm for field pest detection. The algorithm is based on the Cascade R-CNN framework and uses ResNeSt-50 as the backbone network, incorporating cross-channel attention mechanisms to obtain feature maps more conducive to pest detection. A unifying object detection head with attentions (DyHead) is used, incorporating scale awareness, spatial position awareness, and task awareness to improve the performance of the detection head. Additionally, the simple copy-paste (SCP) method is employed for data augmentation to enhance the model's detection capabilities in complex scenarios. A total of 1 500 images of 20 pest categories were collected, and a monitoring lamp field pest dataset compliant with the microsoft common objects in context (MS COCO 2017) format was created. The results show that the F_1-score of the proposed method reaches 86.2%. When the intersection over union (IoU) is set to 0.5, the F_1-score increases by 2.8, 5.8, and 8.2 percentages compared to the classic Cascade R-CNN, Faster R-CNN, and YOLOv4, respectively. The results shows that the proposed method meets the requirements of discriminative ability and real-time performance for monitoring lamp pest detection tasks, achieving high-precision automatic identification and counting of pests, and can be directly applied to field pest detection.

Keywords deep learning; telemetering lamp; pest identification; Cascade R-CNN; accurate detection; attention mechanism; green prevention and control

(责任编辑:边书京)