

基于改进 Mask R - CNN 的苹果园害虫识别方法

王金星^{1,2} 马 博¹ 王 震^{1,3} 刘双喜^{1,4} 慕君林¹ 王云飞¹

(1. 山东农业大学机械与电子工程学院, 泰安 271018; 2. 山东省农业装备智能化工程实验室, 泰安 271018;
3. 山东农业大学园艺科学与工程学院, 泰安 271018; 4. 山东省园艺机械与装备重点实验室, 泰安 271018)

摘要: 针对基础卷积神经网络识别苹果园害虫易受背景干扰及重要特征表达能力不强问题,提出一种基于改进 Mask R - CNN 的苹果园害虫识别方法。首先,基于 Haar 特征方法对多点采集得到的苹果园害虫图像进行迭代初分割,提取害虫单体图像样本,并对该样本进行多途径扩增,得到用于深度学习的扩增样本数据集。其次,对 Mask R - CNN 中的特征提取网络进行优化,采用嵌入注意力机制模块 CBAM 的 ResNeXt 网络作为改进模型的 Backbone,增加模型对害虫空间及语义信息的提取,有效避免背景对模型性能的影响;同时引入 Boundary 损失函数,避免害虫掩膜边缘缺失及定位不准确问题。最后,以原始 Mask R - CNN 模型作为对照模型,平均精度均值作为评价指标进行试验。结果表明,改进 Mask R - CNN 模型平均精度均值达到 96.52%,相比于原始 Mask R - CNN 模型,提高 4.21 个百分点,改进 Mask R - CNN 可精准有效识别苹果园害虫,为苹果园病虫害绿色防控提供技术支持。

关键词: 苹果园; 害虫识别; 深度学习; 注意力机制; 损失函数; Mask R - CNN

中图分类号: TP391 文献标识码: A 文章编号: 1000-1298(2023)06-0253-11

OSID:



Pest Identification Method in Apple Orchard Based on Improved Mask R - CNN

WANG Jinxing^{1,2} MA Bo¹ WANG Zhen^{1,3} LIU Shuangxi^{1,4} MU Junlin¹ WANG Yunfei¹

(1. College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China

2. Shandong Agricultural Equipment Intelligent Engineering Laboratory, Taian 271018, China

3. College of Horticulture Science and Engineering, Shandong Agricultural University, Taian 271018, China

4. Shandong Provincial Key Laboratory of Horticultural Machinery and Equipment, Taian 271018, China)

Abstract: Aiming at the problem that the basic convolutional neural network is vulnerable to background interference and the expression ability of important features is not strong in apple orchard pest recognition, an apple orchard pest recognition method based on improved Mask R - CNN was proposed. Firstly, based on Haar feature method, the apple orchard pest images collected from multiple points were iteratively preliminarily segmented, the single pest image sample was extracted, and multi-channel amplification on the sample was performed to obtain the amplified sample data for deep learning. Secondly, the feature extraction network in Mask R - CNN was optimized, and the ResNeXt network embedded in the attention mechanism module CBAM was used as the Backbone of the improved model, which increased the extraction of pest space and semantic information by the model, and effectively avoided the influence of background on performance of the model. At the same time, the Boundary loss function was introduced to avoid the problem of missing edge of pest mask and inaccurate positioning. Finally, the original Mask R - CNN model was used as the control model, and the mean average precision (mAP) was used as the evaluation index to conduct experiments. The results showed that the mean average precision of the improved Mask R - CNN model reached 96.52%. Compared with the original Mask R - CNN model, the mean average precision was increased by 4.21 percentage points. The results showed that the improved Mask R - CNN can accurately and effectively identify pests in apple orchards. The research result can provide technical support for green control of apple orchard pests and diseases.

Key words: apple orchard; pest identification; deep learning; attention mechanism; loss function; Mask R - CNN

收稿日期: 2022 - 09 - 26 修回日期: 2022 - 11 - 24

基金项目: 国家自然科学基金项目(32071908)和财政部和农业农村部:国家现代农业(苹果)产业技术体系项目(CARS - 27)

作者简介: 王金星(1970—),男,教授,博士生导师,主要从事智能农业装备及技术研究,E-mail: jinxingw@163.com

0 引言

我国是世界上最大的苹果生产国和消费国,根据国家统计局《中国统计年鉴 2021》显示,我国苹果种植面积达 $1.264\ 6 \times 10^7\ \text{hm}^2$,苹果年产量共计 $4.406\ 6 \times 10^7\ \text{t}^{[1]}$ 。苹果产业的高质量发展对于提高果农经济收入,促进乡村振兴意义重大。由于苹果生长周期长、生态系统相对稳定,导致虫害种类众多,严重制约着果品产量和质量^[2]。因此,对苹果园害虫进行准确识别并制定合适的防治措施,有利于减少农药用量、实现绿色防控,促进苹果产业的可持续发展。目前我国大部分苹果产区多采用人工巡视或信息素诱捕的方式监测害虫,然后再根据植保专家或果农经验进行人工判断,存在效率低、劳动强度大、易受主观性影响、费时费力且无法满足现代农业生产需求等问题^[3]。

随着计算机技术不断发展,机器视觉技术在农业领域得到广泛应用^[4]。传统图像处理技术主要通过提取害虫的颜色、纹理及形状等特征进行害虫识别。竺乐庆等^[5]采用颜色直方图和双树复小波变换的方法提取图像颜色、纹理特征,实现对鳞翅昆虫的识别。WANG 等^[6]运用了支持向量机和人工神经网络模型提取昆虫图像特征,实现了对不同类别昆虫的分类。LARIOS 等^[7]通过结合 Haar 特征和随机森林的方法来提取石蝇颜色及图像平面梯度特征,并利用支持向量机(SVM)实现分类识别。传统图像处理技术虽然可实现害虫自动识别,但特征提取过程复杂,易受到环境因素干扰,泛化能力差。

近年来,深度学习以其强大学习能力和特征提取能力在农业图像分类和信息感知中的应用愈发广泛。林相泽等^[8]提出一种基于迁移学习和 Mask R-CNN 的稻飞虱图像分类方法,通过迁移学习在 ResNet50 框架上进行预训练学习,并基于 Mask R-CNN 对存在粘连和重合的昆虫进行分类试验,结果表明该方法能够快速有效识别稻飞虱。RONG 等^[9]提出一种基于 Mask R-CNN 的田间黄板害虫识别计数方法,通过改进特征金字塔网络,将语义信息和定位信息更有效地结合,解决了害虫识别计数不准确的问题。邓朝等^[10]提出一种基于 Mask R-CNN 的马铃薯晚疫病量化评价方法,通过生成分割掩膜计算病斑面积与叶片总面积之比 K ,从而根据 K 值对病害的严重程度级别进行量化评价。STOREY 等^[11]提出一种基于 Mask R-CNN 的苹果园叶锈病检测方法,通过对比 3 种主干网络的检测性能,确定以 ResNet50 网络为主干的 Mask R-CNN 模型在检测较小叶锈病时有较好的性能。综上所述,国内外学者基于 Mask R-CNN 对病

虫害识别已开展相关研究,但由于不同种类害虫形态和纹理特征相近,在一定程度上影响模型识别性能。因此,在特征提取网络中添加注意力机制模块如 SENet、SCSE、CBAM 等^[12-16],能够以无监督的方式来定位害虫判别性区域,进一步提高模型对害虫特征的提取能力,提高识别网络精度。

为进一步提高苹果园害虫在自然环境下的识别精度,本文提出一种基于改进 Mask R-CNN 的苹果园害虫识别方法,采用嵌入注意力机制模块 CBAM 的 ResNeXt 网络作为改进模型的 Backbone,同时引入 Boundary 损失函数,以期提升模型的特征提取能力以及害虫边缘检测的精度。

1 材料与方法

1.1 图像采集

研究所用图像由两部分构成:①由济南祥辰科技有限公司与山东农业大学共同研发的虫情监测装备采集得到。②主要由山东省济南市济阳区山东省农业科学院害虫库提供。

虫情监测装备为双摄图像采集装备,型号为 XC-CP006,如图 1 所示。该装备主要由诱虫灯、接虫板、第 1 摄像头、第 2 摄像头与控制单元等构成。

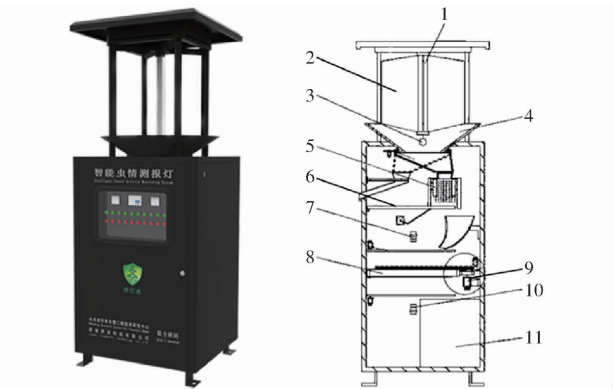


图 1 XC-CP006 型虫情监测装备

Fig.1 XC-CP006 insect monitoring equipment

1. 诱虫灯 2. 撞击板 3. 性诱芯 4. 接虫漏斗 5. 杀虫单元
6. 电控翻转底板 7. 第 1 摄像头 8. 接虫板 9. 驱动机构
10. 第 2 摄像头 11. 收集单元

作业时,诱虫灯及性诱剂诱捕害虫,害虫首先进入高压和高热杀虫单元被杀灭,然后害虫通过接虫漏斗滑落至透明接虫板,控制单元控制第 1 摄像头和第 2 摄像头对害虫正面与背面进行拍摄,最后通过 5G 模块上传到害虫图像云平台。其中,第 1 摄像头及第 2 摄像头为 1 200 万像素,摄像头距标定板拍摄距离均为 300 mm。为保证试验所采集样本的多样性,虫情监测装备分布在山东省内 18 个试验站点,试验站点分布及采集的害虫图像如图 2 所示,虫情监测点位置如表 1 所示。

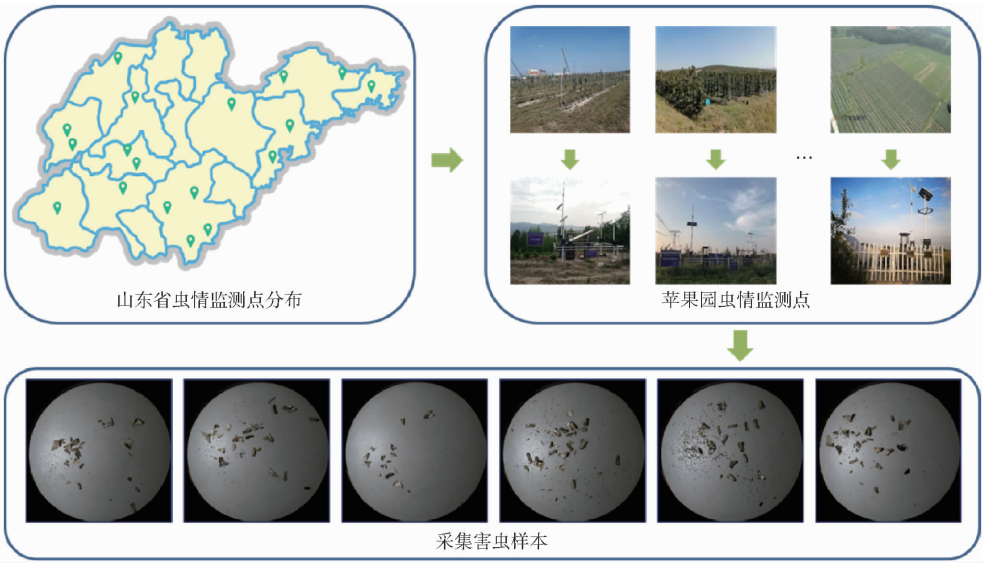


图 2 试验站点分布及采集的害虫图像

Fig. 2 Test site and collected pest images

表 1 济南祥辰科技有限公司虫情监测点位置
Tab. 1 Ji'nan Xiangchen insect monitoring point

地址	经纬度
威海市文登市第一观测场	121°98'E,37°09'N
青岛市黄岛区董家河村	120°19'E,35°91'N
青岛市即墨区学农基地观测场	120°73'E,36°50'N
烟台莱山观测场	121°41'E,37°41'N
烟台莱州 1 号观测场	120°02'E,37°33'N
潍坊市寒亭林业站	119°21'E,37°77'N
临沂市郯城县归义农场	118°41'E,34°67'N
临沂市沂水县四十里堡镇	118°72'E,35°68'N
临沂市平邑县流峪镇苗泉	117°61'E,35°38'N
临沂市临沭县青云镇苹果园	118°66'E,35°03'N
济宁市曲阜市尼山观测场	117°18'E,35°51'N
济南市济阳区太平街道	116°95'E,36°97'N
泰安市岱岳区徂徕山监测站	117°24'E,36°04'N
泰安市泰山区山东农业大学南校监测站	117°16'E,36°16'N
菏泽市牡丹区安兴镇汇芳监测站	115°67'E,35°36'N
聊城市阳谷县魏庄监测站	115°89'E,36°22'N
聊城市东昌府区田庄苗圃	116°06'E,36°43'N
德州市乐陵市观测场	117°26'E,37°75'N

为丰富苹果园害虫训练样本数量,便于后续更好地实现苹果园害虫的精准识别、分类,在山东省农业科学院试验示范基地采用得力-15165 型高拍仪拍摄第 2 部分害虫图像,其害虫样本库取自省内 30 个试验监测点,汇总到试验示范基地后,进行现场识别分类。采集的样本如图 3 所示。

1.2 单体样本获取

选取所采集苹果园害虫样本中的棉铃虫、黏虫和桃蛀螟作为此次试验样本集。虫情监测装备所拍摄的图像为随机散落的多个害虫图像,为防止数据过拟合、提高卷积神经网络的训练效率及更好地提

取目标特征,基于 Haar 特征方法获取苹果园害虫单体样本。Haar 特征值反映图像灰度变化情况,它分为边缘特征、线性特征、中心特征和对角线特征,如图 4 所示。

基于 Haar 特征方法获取单体害虫样本主要通过 Haar 特征分类器实现。首先,将苹果园害虫划分为正样本、与害虫相近的背景划分为负样本。其次,Haar 特征分类器通过对正负样本特征的学习及训练,得到一个描述检测虫体 Haar 特征值的分割模型。最后,通过得到矩形框坐标点来分割提取出单体害虫样本,如图 5 所示。若有漏分割样本,人工区分后再加入 Haar 特征器进行迭代训练,从而获得更高的分割精度和泛化能力。

1.3 数据集构建

1.3.1 样本扩增

为丰富试验数据集、更好地提取苹果园害虫特征及提高模型泛化能力,采用数据增强技术对苹果园害虫数据集进行样本扩增,分别对单体害虫图像进行高斯滤波、色度、椒盐噪声、对比度以及翻转不同角度处理。扩增后图像如图 6 所示。

1.3.2 多边形标注

采用 Labelme 中多边形标注工具,标注苹果园害虫轮廓,避免复杂背景对虫体的影响,标注后生成 Json 标签文件,经转换后生成 mask、label_viz、label_names 等文件,标注后生成的文件如图 7 所示。

1.3.3 数据集划分

为增强苹果园害虫识别模型的鲁棒性,按照不同比例分别对棉铃虫、黏虫以及桃蛀螟图像进行扩增。苹果园害虫样本集按照比例 7:2:1 划分为训练集、测试集和验证集,三者之间无重叠。其中,训练

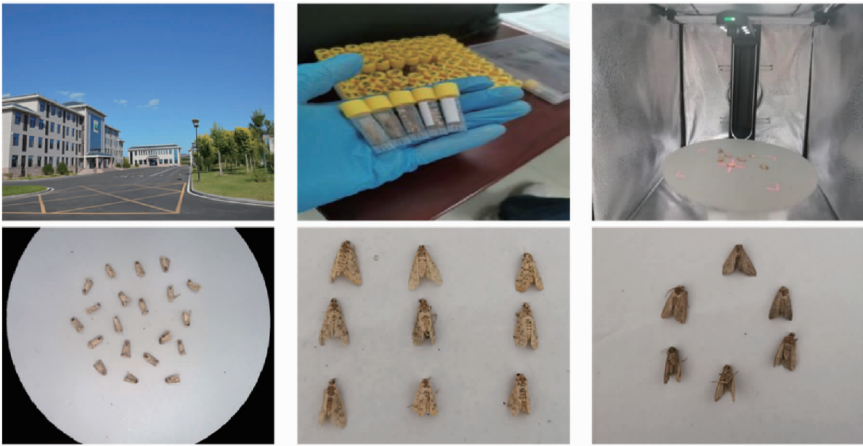


图 3 示范基地采集样本

Fig. 3 Sample collection at demonstration base

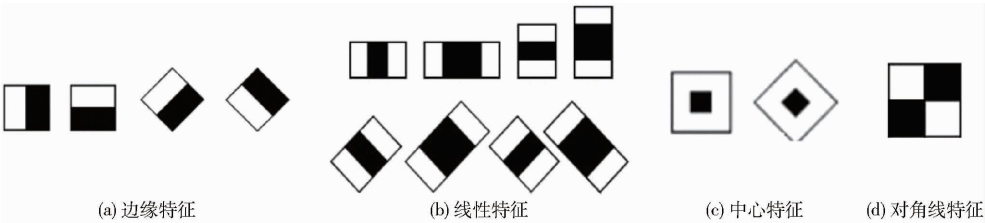


图 4 Haar 特征

Fig. 4 Haar features

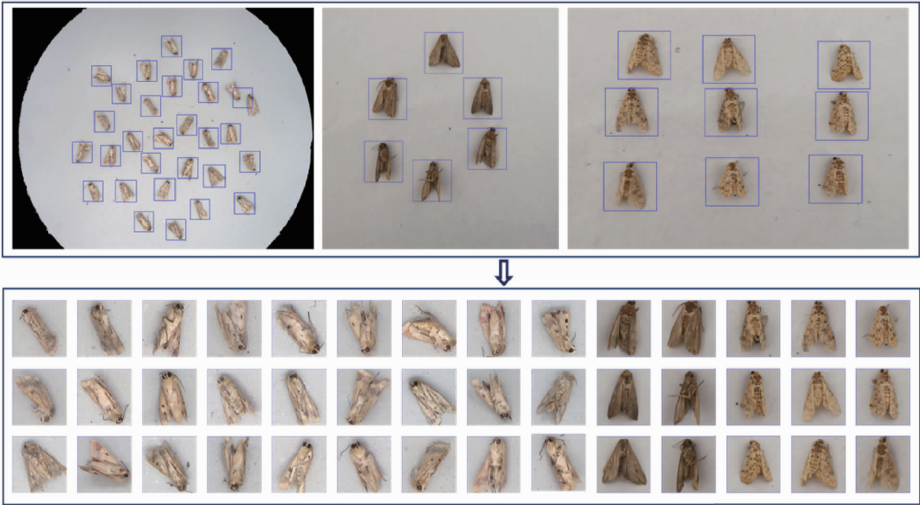


图 5 单体害虫分割图像

Fig. 5 Single pest segmentation image



图 6 单体样本数据扩增后图像

Fig. 6 Images after amplification of single sample data

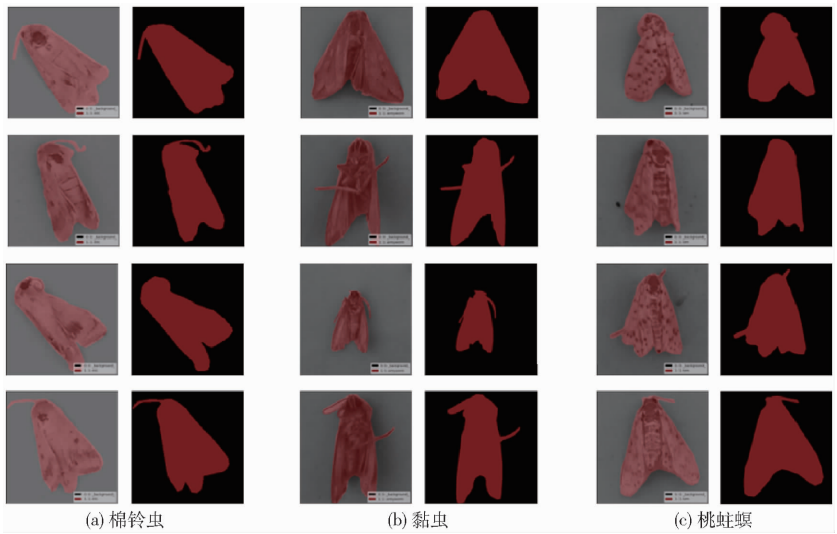


图 7 Json 标签文件

Fig. 7 Json tag file

集用于训练模型,验证集用于优化超参数,测试集用于评估模型性能。数据集划分结果如表 2 所示。

表 2 苹果园害虫数据集

Tab. 2 Apple orchard pest data set

幅

害虫种类	标注标签	原始数目	扩增后数目		
			训练集	测试集	验证集
棉铃虫	mle	2 320	6 496	1 856	928
黏虫	armyworm	1 925	6 737	1 925	962
桃蛀螟	tzm	1 654	5 789	1 654	827

2 基于改进 Mask R - CNN 的苹果园害虫识别模型

2.1 苹果园害虫识别模型基本网络架构

苹果园害虫识别模型是在 Faster R - CNN 的基础上将感兴趣池化层更换为感兴趣区域对齐层 (Region of interest align layer, RoI Align), 并且添加一个以全卷积神经网络 (Fully convolutional neural networks, FCN) 为主的掩膜分支进行像素级分割^[17-19], 改进 Mask R - CNN 的苹果园害虫识别模

型架构如图 8 所示, 主要是由特征提取网络 (Backbone)、区域建议网络 (Region proposal network, RPN)、感兴趣区域对齐层以及目标检测和分割 4 部分组成^[20-23]。

2.2 Mask R - CNN 模型改进

2.2.1 ResNeXt 模块

Mask R - CNN 采用残差网络 ResNet 作为骨干特征提取网络^[24], 而传统模型均通过加深或加宽网络结构提高识别精度, 但是随着超参数数量的增加, 模型的设计难度和计算开销也会增加。ResNeXt 模块可以在不明显增加参数量级的前提下, 用一种平行堆叠相同拓扑结构的 blocks 代替原来 ResNet 的 3 层卷积的 block, 从而提高平均识别精度^[25], 改进 Mask R - CNN 模型采用 ResNeXt 模块作为骨干特征提取网络, ResNeXt 在结构上的改进使其具有更强的表征能力, ResNet 和 ResNeXt 结构如图 9 所示。

由图 9 可知, ResNeXt 是把 ResNet 的单个卷积改成了多支路的卷积, 输入的害虫图像送入各个支路分别进行卷积操作, 然后将各个支路输出的特征

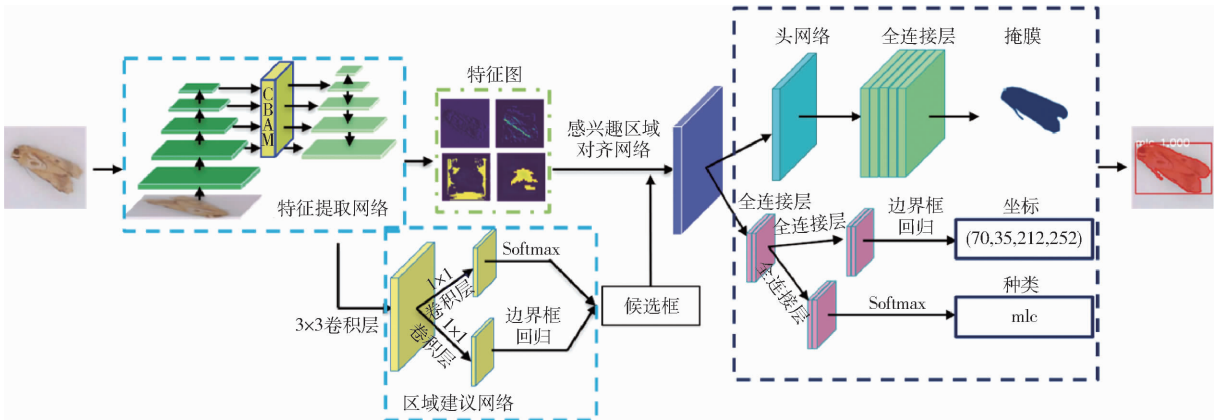


图 8 改进 Mask R - CNN 模型架构

Fig. 8 Improving Mask R - CNN model architecture

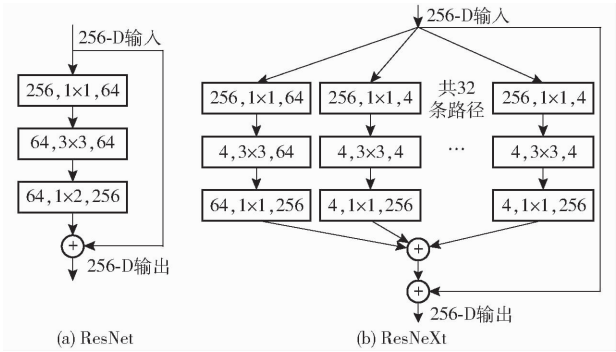


图 9 ResNet 和 ResNeXt 结构
Fig. 9 ResNet and ResNeXt structures

图进行维度拼接,最终得到提取的输出结果^[26-28]。

2.2.2 Boundary 损失函数

Mask R-CNN 的损失函数 L_{loss} 定义为分类、边界框回归以及分割掩模损失之和,其损失函数表达式为

$$L_{loss} = L_{cls} + L_{box} + L_{mask} \tag{1}$$

式中 L_{cls} ——分类损失值
 L_{box} ——边界框回归损失值
 L_{mask} ——分割损失值

Mask R-CNN 的 L_{mask} 采用的是平均二值交叉熵损失函数,对苹果园害虫预测时,会出现边缘信息不够完整的情况,造成掩膜缺失以及影响模型识别精度。因此,本研究对 L_{mask} 部分的损失函数进行优化,引入 Boundary 损失函数^[29]对害虫边界进行像素级评估。Boundary 损失函数使用边界匹配度监督网络的 loss,和真实框边界吻合的像素点的 loss 记为 0,不吻合的点,根据其距边界的距离评估 loss。为了以可微分的方式计算 2 个边界距离 ($\partial G, \partial S$),边界损失使用边界上的积分而不是区域上的不平衡积分来减轻高度不平衡分割的困难,具体计算过程和最终得到的边界框损失函数式为

$$\text{Dist}(\partial G, \partial S) = \int_{\partial G} \|q - \partial s(p)\|^2 dp \approx$$
$$2 \int_{\Delta_s} D_c(p) dp =$$
$$2 \left(\int_{\Omega} \phi_c(p) s(p) dp - \int_{\Omega} \phi_c(p) g(p) dp \right) \tag{2}$$

$$L_{BD} = \int_{\Omega} \phi_c(p) S_{\theta}(p) dp \tag{3}$$

式中 G ——真实框区域
 q ——属于整个边界区域任意一点, $q \in \Omega$
 S ——预测待分割区域
 Δs ——真实框和预测框两个轮廓之间的区域
 Ω ——整个边界区域
 $D_c(p)$ ——真实框的距离图
 $\phi_c(p)$ ——边界水平集,若 $q \in G$,则 $\phi_c = -D_c(q)$

$s(p)、g(p)$ ——二元指标函数
 $S_{\theta}(p)$ ——网络的 Softmax 概率输出

2.2.3 注意力模型模块

为更好提取目标特征,在特征提取网络 ResNeXt 上添加卷积注意力模块^[30] (Convolutional block attention module, CBAM),该卷积注意力模块包含 2 个独立的子模块,通道注意力模块 (Channel attention module, CAM) 和空间注意力模块 (Spatial attention module, SAM),可以将其融入到各种卷积神经网络中去进行端到端的训练^[31],CBAM 结构如图 10 所示。

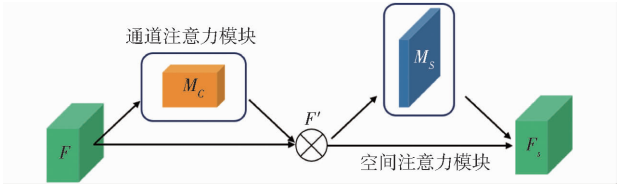


图 10 注意力机制 CBAM 整体网络架构
Fig. 10 Attention mechanism CBAM overall network architecture

通道注意力模块结构如图 11 所示,图中输入是一幅 $H \times W \times C$ (H 为高度, W 为宽度, C 为通道数) 的特征图 F ,首先分别通过全局平均池化和全局最大池化 2 种方式提取不同的特征信息,得到 2 个维度为 $1 \times 1 \times C$ 的特征图,再经过共享多层感知机 (Multi-layer perceptron, MLP) 得到 2 个特征向量,其神经元个数分别为 C/R (R 为减少率) 和 C ,其次对 2 个特征向量进行逐个元素 (element-wise) 相加累积操作,然后经过 sigmoid 激活函数得到权重系数 M_c ,最后将权重系数和最初的特征图进行相乘,得到 SAM 模块需要的输入特征 F' 。具体表达式为

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) =$$
$$\sigma(W_1(W_0(F_{\text{avg}}^C)) + W_1(W_0(F_{\text{max}}^C))) \tag{4}$$

式中 σ ——sigmoid 函数
MLP——多层感知机
AvgPool——平均池化
MaxPool——最大池化
 $W_0、W_1$ ——输入共享的 MLP 权重
 F_{avg}^C ——平均池化后的通道描述特征
 F_{max}^C ——最大池化后的通道描述特征

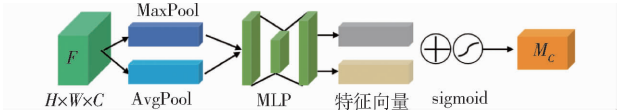


图 11 通道注意力模块 CAM 结构图
Fig. 11 CAM structure diagram of channel attention module

空间注意力模块结构如图 12 所示,图中首先将

维度为 $H \times W \times C$ 的特征图 F' 经过全局平均池化和全局最大池化得到 2 个 $H \times W \times 1$ 的特征图,并将这 2 个特征图基于通道拼接在一起。然后,经过卷积核 7×7 的卷积运算以及 sigmoid 激活操作,得到特征图权重系数 M_s ,最后,用权重系数和特征 F' 相乘并缩放后,最终得到所需要的特征。具体表达式为

$$M_s(F') = \sigma(f^{7 \times 7}([\text{AvgPool}(F'); \text{MaxPool}(F')])) = \sigma(f^{7 \times 7}([F_{\text{avg}}^S; F_{\text{max}}^S])) \quad (5)$$

式中 $f^{7 \times 7}$ —— 7×7 的卷积层
 F_{avg}^S —— 平均池化后的空间描述特征
 F_{max}^S —— 最大池化后的空间描述特征

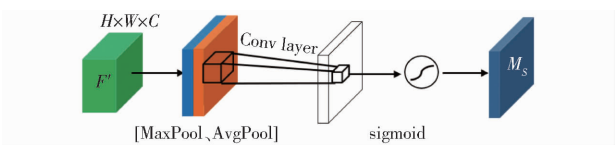


图 12 空间注意力模块 SAM 结构图

Fig. 12 SAM structure diagram of spatial attention module

3 试验与结果分析

3.1 试验环境及超参数设置

试验在运行环境为 Windows 10 (64 位) 操作系统、处理器为 Intel(R) Xeon(R) Silver4210R 的计算机上进行,并配备 NVIDIA GeForce RTX 2080Ti GPU 加速试验进程。使用 Python 3.6 版本以 TensorFlow (1.13.2) 为后端的 Keras 框架 (2.1.5 版本),以及配备 CUDA10、cuDNN7.4.1.5 对改进 Mask R - CNN 模型进行训练。

为确定模型最佳的训练参数,通过验证集在相同环境下、不同参数的模型性能对比,来设置模型训练参数。改进 Mask R - CNN 的超参数设置如表 3 所示。

表 3 模型超参数设置

Tab.3 Model super parameter setting

参数	数值
初始学习率	0.000 01
最小图像维度	384
最大图像维度	2 816
锚框	(16, 32, 64, 128, 256)
动量因子	0.9
置信度	0.7
迭代轮数	200

3.2 试验评价指标

为验证模型识别效果,采用平均精度均值 (Mean average precision, mAP) 作为苹果园害虫识别模型的评价指标, mAP 表示各类别害虫平均精度 (AP) 的平均值。其中, AP 为识别单类害虫的平均

精度,以精确率 P (Precision) 为纵坐标,召回率 R (Recall) 为横坐标,绘制 $P - R$ (Precision - Recall) 曲线,曲线与坐标轴围成的面积为 AP。

3.3 特征图分析

改进 Mask R - CNN 选用特征提取网络 ResNeXt 结合特征金字塔网络 (Feature pyramid network, FPN) 构建多尺度特征金字塔模型,融合注意力机制的特征金字塔网络如图 13 所示,考虑到计算冗余及内存占用量,故没有将 Conv1 包含在特征金字塔网络中。

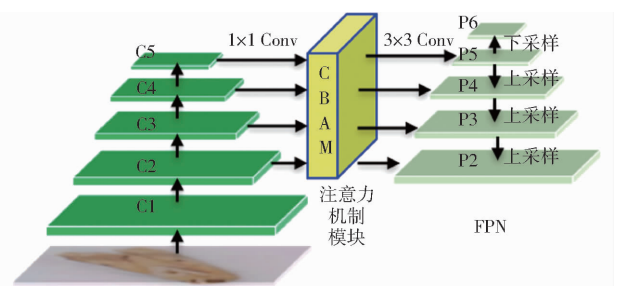


图 13 融合注意力机制的特征金字塔网络

Fig. 13 Feature pyramid network integrating attention mechanism

首先卷积层 (Conv1 ~ Conv5) 通过对输入图像进行自下而上的特征图提取,将输出的特征图表示为 C1、C2、C3、C4、C5,然后再进行 1×1 的卷积融合改变特征图的维度,并通过注意力机制模块 CBAM 自动判别害虫区域以此提高模型的特征提取能力,其次通过自上而下的上采样,与侧边特征层进行卷积融合,采用 3×3 的卷积对每个融合结果进行卷积以消除上采样的混叠效应,最后融合卷积结果得到 P2 ~ P5,同时与顶层特征图 P6 结合,实现对特征的多尺度融合。

为更好理解卷积神经网络对苹果园害虫的特征提取过程,使用交互式笔记本 (jupyter notebook) 软件对模型部分卷积层的特征提取效果进行可视化分析。选取特征提取网络 ResNeXt 中 C2 ~ C5 层以及特征金字塔 P2 ~ P5 层进行分析,特征图层如图 14 所示。由图 14 可以看出,经过特征提取网络中 Conv2 层输出的特征映射和原图相似,能够显示出害虫的清晰轮廓,但是随着模型网络的加深,越来越多的细节信息消失,输出的特征也越来越抽象。而如图 14 所示,随着特征金字塔 (FPN) 对卷积层 C2 ~ C5 的多尺度融合,即把更关注于细节信息的低层特征和更关注于深层语义信息的高层网络进行自上而下的侧边连接,生成不同层次多尺度特征图 P2 ~ P5,由图中也可以看出,经过特征金字塔网络,使得各个尺度下的害虫特征都包含丰富的语义信息。

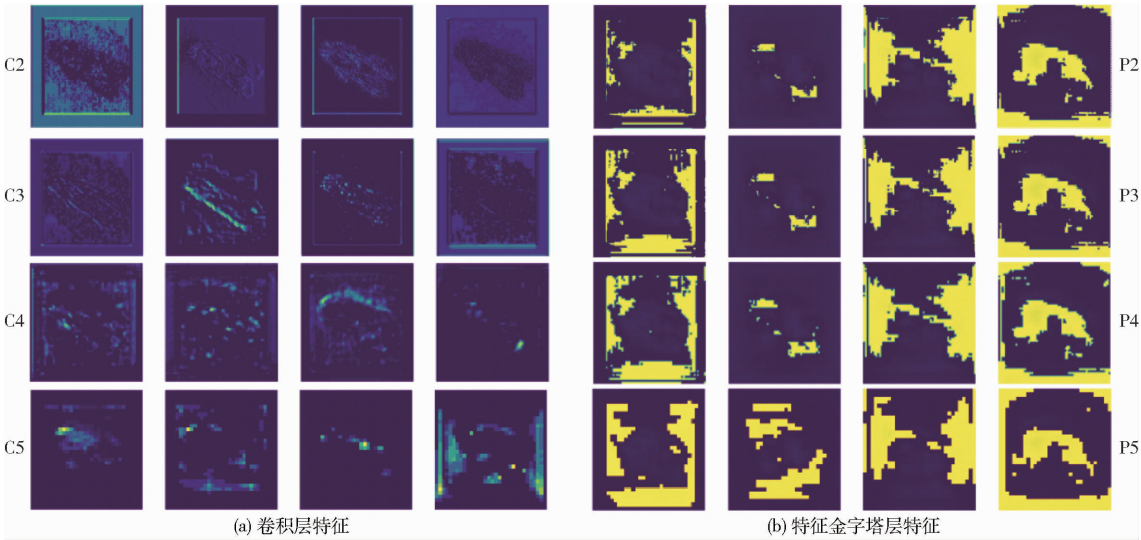


图 14 各卷积层特征提取效果

Fig. 14 Feature extraction effect of each convolution layer

3.4 定性试验

苹果园害虫棉铃虫、黏虫以及桃蛀螟的识别结果如图 15 所示。可以看出本文提出的改进 Mask R-CNN 模型能够较好地识别出苹果园害虫,同时害虫的掩膜边缘信息也被完整地分割出来,苹果园害虫图像边界和细节更为平滑,结果表明本文提出的改进算法具有较强的鲁棒性和泛化能力。

3.5 苹果园害虫识别系统

采用 PyQt5 搭建识别模型用户界面 (GUI),并且可以通过 PyQt5 中的 Qt Designer 来编写 UI 界面。Qt Designer 的设计符合模型-视图-控制器 (MVC) 的架构,实现了视图和逻辑的分离并且容易开发,设计的界面如图 16 所示。通过本地相册上传一幅图像后,系统将会调用改进 Mask R-CNN 模型和运行

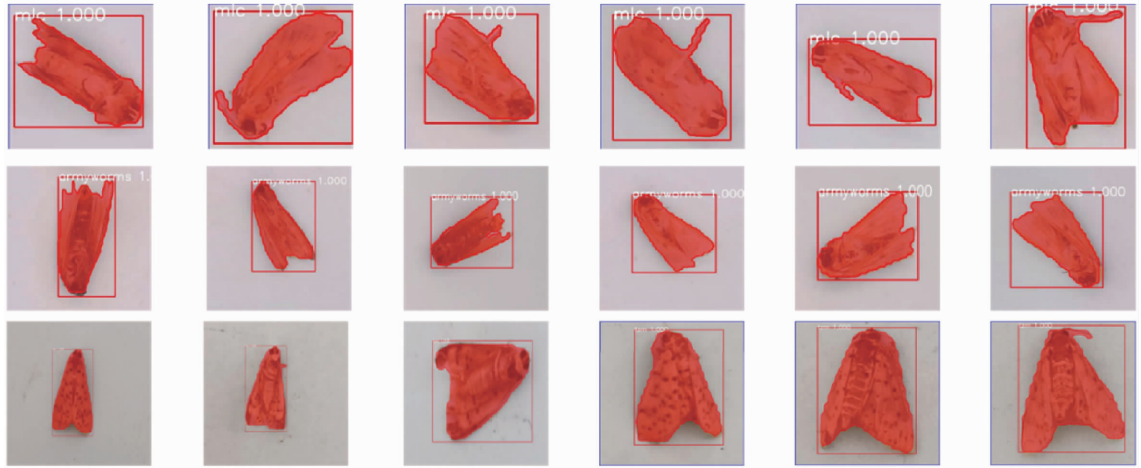


图 15 定性试验识别结果

Fig. 15 Qualitative test identification results

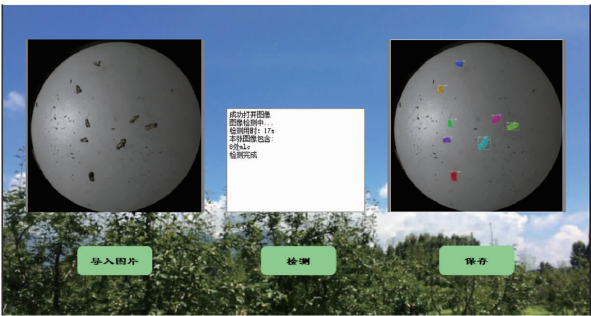


图 16 模型可视化界面

Fig. 16 Model visualization interface

环境,并将识别的结果返回给系统。最终,界面展示出所识别的信息,包括识别种类、识别时间以及识别数量,并且可以将识别结果 temp.png 保存到指定位置。本文所述部署操作简单、快捷,便于在实际生产中应用。

3.6 模型对比试验

3.6.1 不同骨干网络对比试验

为进一步验证骨干网络 ResNeXt 在苹果园害虫识别模型上的有效性,基于 Mask R-CNN 的框架,分别对骨干网络 ResNet50、ResNet101、ResNeXt50、

ResNeXt101 和 ResNeXt152 进行试验验证,同时对各个骨干网络的参数量进行计算,试验结果如表 4 所示。

表 4 不同骨干网络对比试验

Tab. 4 Comparative test of different backbone networks

骨干网络	平均精度均值/%	参数量
ResNet50	88. 72	$5. 60 \times 10^7$
ResNet101	92. 31	$7. 30 \times 10^7$
ResNeXt50	91. 39	$7. 10 \times 10^7$
ResNeXt101	93. 18	$8. 90 \times 10^7$
ResNeXt152	89. 52	$1. 32 \times 10^8$

由表 4 可知,选用 ResNeXt101 作为骨干网络识别苹果园害虫效果最佳。ResNeXt 在结构上的改进相对于 ResNet 增加了网络宽度,有更强的特征提取能力,因此在相同网络层数的基础上,ResNeXt50 和 ResNeXt101 比 ResNet50 和 ResNet101 的平均识别精度高。ResNeXt152 相对于 ResNeXt101 网络深度更深,理论上特征提取能力也相对较强,但网络太深会带来参数量过大及梯度消失等问题,影响网络识别精度。因此,选择 ResNeXt101 网络作为骨干网络来提取苹果园害虫特征,同时本文后续试验均采用 ResNeXt101 网络。

3. 6. 2 不同损失函数对比试验

为对比 Boundary 损失函数在识别苹果园害虫方面的优越性,基于 Mask R - CNN 的框架,分别使用 Smooth L1 损失函数、交叉熵损失函数 (Cross entropy) 和 Focal 损失函数进行对比试验。由于 Boundary 损失函数可有效改善害虫掩膜边缘缺失及定位不准确问题,所以本试验引入 IoU_{Mask} 对掩膜的质量进行评估,通过计算苹果园害虫预测的掩膜区域和人工标记区域的交集与并集之比来量化所预测的掩膜质量,从而进一步评判苹果园害虫定位的精确程度,试验结果如表 5 所示。

表 5 不同损失函数对比试验

Tab. 5 Comparative test of different loss functions

损失函数	平均精度均值/%	$IoU_{Mask}/\%$
Smooth L1	89. 31	85. 62
Cross entropy	89. 62	83. 28
Focal	92. 93	86. 53
Boundary	93. 14	92. 25

由表 5 可知,Boundary 损失函数在平均识别精度和掩膜分割质量上都占有优势。究其原因,Smooth L1 损失函数只是针对预测框的偏移量进行损失计算,无法准确描述预测框与真实框之间的位置关系;Cross entropy 损失函数没有考虑到苹果园害虫边缘像素的邻域信息,细粒度信息效果不好;

Focal 损失函数旨在解决类别不平衡的问题,侧重点是根据样本分辨的难易程度给样本对应的损失添加权重,更加聚焦于困难样本,所以 Focal 损失函数和 Boundary 损失函数的平均识别精度差别不大,但存在边缘掩膜信息分割不完整现象, IoU_{Mask} 比 Boundary 损失函数低 5. 72 个百分点。

3. 6. 3 不同注意力机制对比试验

为进一步验证注意力机制模块 CBAM 的性能,在相同的试验条件下,将 SENet、Channel attention 和 Spatial attention 嵌入改进 Mask R - CNN 的特征提取网络 ResNeXt 中,通过对每种注意力机制的对比试验,得到网络模型的识别精度,分别为 94. 08%、93. 82%、93. 68%,而 CBAM 的识别精度为 94. 82%。

通过对比试验可知,与其他注意力机制相比,在特征提取网络中融合注意力机制模块 CBAM 平均识别精度最高。究其原因,SENet 仅对特征图通道信息添加注意力模块,而忽略空间维度的特征信息。Channel attention 和 Spatial attention 也分别仅在通道维度和空间维度中提取单维度特征信息,易造成特征信息提取不够全面的问题。注意力机制模块 CBAM 是先通过通道注意力对原始图像的全局特征信息进行提取,再通过空间注意力对全局特征中的重要信息进行提取,提高模型对重要特征的提取能力;并且从模型结构分析可知,注意力机制模块 CBAM 融合最大池化层和平均池化层,使提取到的目标高层信息更加丰富,从而提高模型的平均识别精度。

3. 6. 4 不同网络模型对比试验

为验证本文所述改进 Mask R - CNN 模型在识别苹果园害虫上的优势,分别使用目前较为成熟的卷积神经网络模型 Faster R - CNN、YOLACT、YOLO v3、SSD 对苹果园害虫数据集进行试验验证,同时引入帧速率 (Frame per second, FPS) 评估不同模型的检测速度,试验结果如表 6 所示。

表 6 不同网络模型对比试验

Tab. 6 Comparison test of different network models

模型	平均精度均值/%	帧速率/($f \cdot s^{-1}$)
Faster R - CNN	89. 72	7
YOLACT	85. 69	25
YOLO v3	81. 22	29
SSD	82. 13	22
Mask R - CNN	92. 31	4
改进 Mask R - CNN	96. 52	4

由表 6 可知,改进 Mask R - CNN 模型的平均识别精度最高,达到 96. 52%,虽然在检测速度上不占

优势,但本文只是针对静态苹果园害虫进行识别,对检测速度没有过高的要求。Faster R-CNN 和 Mask R-CNN 模型同属于两阶段算法,比属于单阶段算法的 YOLO v3 和 SSD 模型的平均精度均值高。Mask R-CNN 较 Faster R-CNN 的平均精度均值提升 2.59 个百分点,改进 Mask R-CNN 较 Faster R-CNN 的平均精度均值提升 6.8 个百分点,究其原因,Mask R-CNN 在 Faster R-CNN 的基础上增添掩膜分支,并将 RoI Pooling 层替换为 RoI Align 层,提升了识别精度。YOLO v3 和 SSD 识别速度较快,但平均识别精度较低,YOLACT 作为单阶段实例分割算法,较 YOLO v3 和 SSD 模型的平均精度均值分别高 4.47、3.56 个百分点,但仍低于本文所述改进 Mask R-CNN 模型。

3.6.5 公开数据集测试试验

为进一步验证本文所述改进 Mask R-CNN 模型在识别不同害虫数据集上的性能,采用 Pest24 公开害虫数据集^[32]进行验证。Pest24 害虫数据集有 24 个害虫类别,共包含 25378 幅图像,其数据特点是目标害虫尺寸较小、密集分布且存在粘连、不同种类害虫较为相似。选用 Pest24 数据集中棉铃虫、黏虫、草地螟、斜纹夜蛾和甜菜夜蛾 5 种常见的苹果园害虫进行试验验证,Mask R-CNN 和改进 Mask R-CNN 的平均精度均值分别为 68.23% 和 74.75%。

通过试验可知,本文所述改进 Mask R-CNN 模型较原始 Mask R-CNN 模型的平均精度均值提高 6.52 个百分点。由于 Pest24 害虫数据集中不同害虫间形状和颜色相似,存在种间的细粒度差异,而且 5 种害虫的相对尺寸较小,容易出现误识别和漏识别现象,而改进 Mask R-CNN 模型能够对不同害虫区别性部位特征进行强化,提升模型对害虫特征的提取能力,从而进一步提高了模型在复杂环境下对害虫识别的适用性。

3.6.6 消融试验

为验证改进 Mask R-CNN 模型不同模块对特征提取网络性能的影响,同时验证各个模块结构的有效性,将改进模型分别与采用 ResNeXt101 作为 Backbone、引入 Boundary 损失函数、添加注意力机制模块 CBAM 以及原始 Mask R-CNN 模型进行对比试验,试验结果如表 7 所示。

表 7 各个模块对网络性能的影响

Tab.7 Impact of each module on network performance

Mask R-CNN	改进模块			平均精度 均值/%
	ResNeXt	Boundary 损失函数	CBAM	
√				92.31
√	√			93.18
√		√		93.14
√			√	94.82
√	√	√	√	96.52

注:√表示在网络中添加该模块。

由表 7 可知,在 Mask R-CNN 的基础框架之上,采用 ResNeXt 网络代替传统的 ResNet 网络,模型的平均精度均值提升 0.87 个百分点;只引入 Boundary 损失函数,模型的平均精度均值提升 0.83 个百分点;只融合注意力机制模块 CBAM,模型的平均精度均值提升 2.51 个百分点;而同时融合 ResNeXt 网络、Boundary 损失函数以及注意力机制模块 CBAM,模型的平均精度均值提升 4.21 个百分点,达到 96.52%。消融试验结果表明,本文所提出的基于改进 Mask R-CNN 的苹果园害虫识别模型设计合理,能充分发挥各模块的优势,可有效提升苹果园害虫识别模型的性能。

4 结论

(1)针对果园环境下识别苹果园害虫易受背景信息干扰及边缘特征表达能力不强的问题,提出一种基于改进 Mask R-CNN 的苹果园害虫识别方法,在 Mask R-CNN 基本框架下,采用嵌入注意力机制模块 CBAM 的 ResNeXt 网络作为改进模型的骨干特征提取网络,提高模型对害虫特征的提取能力,抑制背景对模型性能的影响;同时,引入 Boundary 损失函数,避免害虫掩膜边缘缺失及定位不准确问题,实现对苹果园害虫精准有效识别。

(2)在自定义数据集上,基于改进 Mask R-CNN 的苹果园害虫识别模型的平均识别精度为 96.52%,相较于原始 Mask R-CNN,平均精度均值提升 4.21 个百分点;在 Pest24 公开数据集上,改进 Mask R-CNN 的平均精度均值提升 6.52 个百分点,表明该模型设计合理,可为苹果园害虫精准防控提供理论和技术支持。

参 考 文 献

[1] 国家统计局. 中国统计年鉴[J]. 北京:中国统计出版社,2021.

[2] 国栋,杨久涛,于玲雅,等. 山东省苹果病虫害发生防控现状及绿色防控对策[J]. 现代农业科技,2022(5):70-71,85.

[3] 翟肇裕,曹益飞,徐焕良,等. 农作物病虫害识别关键技术研究综述[J]. 农业机械学报,2021,52(7):1-18.

ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1-18. (in Chinese)

[4] LI Y F, WANG H X, DANG L M, et al. Crop pest recognition in natural scenes using convolutional neural networks[J]. Computers and Electronics in Agriculture, 2020, 169: 105174.

[5] 竺乐庆,张真,张培毅. 基于颜色直方图及双树复小波变换(DTCWT)的昆虫图像识别[J]. 昆虫学报,2010,53(1):91 - 97.
ZHU Leqing, ZHANG Zhen, ZHANG Peiyi. Image identification of insects based on color histogram and dual tree complex wavelet transform (DTCWT)[J]. Acta Entomologica Sinica, 2010, 53(1): 91 - 97. (in Chinese)

[6] WANG J N, LIN C T, JI L Q, et al. A new automatic identification system of insect images at the order level[J]. Knowledge-Based Systems, 2012, 33: 102 - 110.

[7] LARIOS N, SORAN B, SHAPIRO L G, et al. Haar random forest features and SVM spatial matching kernel for stonefly species identification[C]//20th International Conference on Pattern Recognition, 2010: 2624 - 2627.

[8] 林相泽,朱赛华,张俊媛,等. 基于迁移学习和 Mask R - CNN 的稻飞虱图像分类方法[J]. 农业机械学报,2019,50(7): 201 - 207.
LIN Xiangze, ZHU Saihua, ZHANG Junyuan, et al. Rice planthopper image classification method based on transfer learning and Mask R - CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7): 201 - 207. (in Chinese)

[9] RONG M X, WANG Z Z, BAN B, et al. Pest identification and counting of yellow plate in field based on improved Mask R - CNN[J]. Discrete Dynamics in Nature and Society, 2022, 2022: 1913577.

[10] 邓朝,纪苗苗,任永泰. 基于 Mask R - CNN 的马铃薯叶片晚疫病量化评价[J]. 扬州大学学报(农业与生命科学版), 2022, 43(1): 135 - 142.
DENG Zhao, JI Miaomiao, REN Yongtai. Quantitative evaluation of potato late blight disease based on Mask R - CNN[J]. Journal of Yangzhou University (Agricultural and Life Science Edition), 2022, 43(1): 135 - 142. (in Chinese)

[11] STOREY G, MENG Q G, LI B H. Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture[J]. Sustainability, 2022, 14(3): 1458.

[12] 章权兵,胡姗姗,舒文灿,等. 基于注意力机制金字塔网络的麦穗检测方法[J]. 农业机械学报,2021,52(11): 253 - 262.
ZHANG Quanbing, HU Shanshan, SHU Wencan, et al. Wheat spikes detection method based on pyramidal network of attention mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(11): 253 - 262. (in Chinese)

[13] 陈娟,陈良勇,王生生,等. 基于改进残差网络的园林害虫图像识别[J]. 农业机械学报,2019,50(5): 187 - 195.
CHEN Juan, CHEN Liangyong, WANG Shengsheng, et al. Pest image recognition of garden based on improved residual network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 187 - 195. (in Chinese)

[14] LORIS N, GIANLUCA M, FABIO P. Insect pest image detection and recognition based on bio-inspired methods[J]. Ecological Informatics, 2020, 57: 101089.

[15] 曾伟辉,唐欣,胡根生,等. 基于卷积块注意力胶囊网络的小样本水稻害虫识别[J]. 中国农业大学学报,2022,27(3): 63 - 74.
ZENG Weihui, TANG Xin, HU Gensheng, et al. Rice pests recognition with small number of samples based on CBAM and capsule network[J]. Journal of China Agricultural University, 2022, 27(3): 63 - 74. (in Chinese)

[16] 姚青,姚波,吕军,等. 基于双线性注意力网络的农业灯诱害虫细粒度图像识别研究[J]. 中国农业科学,2021,54(21): 4562 - 4572.
YAO Qing, YAO Bo, LÜ Jun. Research on fine-grained image recognition of agricultural light-trap pests based on bilinear[J]. Scientia Agricultura Sinica, 2021, 54(21): 4562 - 4572. (in Chinese)

[17] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R - CNN[C]//International Conference on Computer Vision, 2017: 2980 - 2988.

[18] 杨长辉,王卓,熊龙烨,等. 基于 Mask R - CNN 的复杂背景下柑橘树枝干识别与重建[J]. 农业机械学报,2019,50(8): 22 - 30, 69.
YANG Changhui, WANG Zhuo, XIONG Longye, et al. Identification and reconstruction of citrus branches under complex background based on Mask R - CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(8): 22 - 30, 69. (in Chinese)

[19] DORRER M G, ALEKHINA A E. Normalization of data for training and analysis by the Mask R - CNN model using the K-means method for a smart refrigerator's computer vision[J]. Journal of Physics: Conference Series, 2021, 1889(2): 022076.

[20] WANG D D, HE D J. Fusion of Mask R - CNN and attention mechanism for instance segmentation of apples under complex background[J]. Computers and Electronics in Agriculture, 2022, 196: 106864.

[21] 王辉,韩娜娜,吕程序,等. 基于 Mask R - CNN 的单株柑橘树冠识别与分割[J]. 农业机械学报,2021,52(5): 169 - 174.
WANG Hui, HAN Na'na, LÜ Chengxu, et al. Recognition and segmentation of individual citrus tree crown based on Mask R - CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(5): 169 - 174. (in Chinese)

[22] 邓颖,吴华瑞,朱华吉. 基于实例分割的柑橘花朵识别及花量统计[J]. 农业工程学报,2020,36(7): 200 - 207.
DENG Ying, WU Huarui, ZHU Huaji. Recognition and counting of citrus flowers based on instance segmentation[J]. Transactions of the CSAE, 2020, 36(7): 200 - 207. (in Chinese)

[23] 李昊玥,陈桂芬,裴傲. 基于改进 Mask R - CNN 的奶牛个体识别方法研究[J]. 华南农业大学学报,2020,41(6): 161 - 168.
LI Haoyue, CHEN Guifen, PEI Ao. Research on individual recognition of dairy cows based on improved Mask R - CNN[J]. Journal of South China Agricultural University, 2020, 41(6): 161 - 168. (in Chinese)

3943. (in Chinese)

[33] 郑利芳, 吴三鼎, 党廷辉. 不同施肥模式对春玉米产量、水分利用效率及硝态氮残留的影响[J]. 水土保持学报, 2019, 33(4): 221 – 227.
ZHENG Lifang, WU Sanding, DANG Tinghui. Effect of different fertilization modes on spring maize yield, water use efficiency and nitrate nitrogen residue[J]. Journal of Soil and Water Conservation, 2019, 33(4): 221 – 227. (in Chinese)

[34] 杨世琦, 王永生, 韩瑞芸, 等. 宁夏引黄灌区秸秆还田对麦田土壤硝态氮淋失的影响[J]. 生态学报, 2015, 35(16): 5537 – 5544.
YANG Shiqi, WANG Yongsheng, HAN Ruiyun, et al. Effect of nitrate leaching of winter wheat field based on straw application in the Yellow River irrigation area of Ningxia[J]. Acta Ecologica Sinca, 2015, 35(16): 5537 – 5544. (in Chinese)

[35] QIU S, JU X, LI L, et al. Nitrate transformation and N₂O emission in a typical intensively managed calcareous fluvaquent soil: a 15-nitrogen tracer incubation study[J]. Communications in Soil Science and Plant Analysis, 2015, 46(14): 1763 – 1777.

[36] 隋鹏祥, 张文可, 梅楠, 等. 不同秸秆还田方式对春玉米产量、水分利用和根系生长的影响[J]. 水土保持学报, 2018, 32(4): 255 – 261.
SUI Pengxiang, ZHANG Wenke, MEI Nan, et al. Effects of different straw returning methods on spring maize yield, water use and root growth[J]. Journal of Soil and Water Conservation, 2018, 32(4): 255 – 261. (in Chinese)

[37] 宋佳丽, 王鸿雁, 张淑香, 等. 施肥和秸秆还田对黑土农田土壤剖面养分分布的影响[J]. 生态学杂志, 2021, 41(1): 108 – 115.
SONG Jiali, WANG Hongyan, ZHANG Shuxiang, et al. Effect of fertilization and straw return on nutrient distribution along soil profile in black soil cropland[J]. Chinese Journal of Ecology, 2021, 41(1): 108 – 115. (in Chinese)

[38] NAFI E, WEBBER H, DANSO I, et al. Soil tillage, residue management and site interactions affecting nitrogen use efficiency in maize and cotton in the Sudan Savanna of Africa[J]. Field Crops Research, 2019, 244:107629.

[39] 孙晓, 景建元, 吕慎强, 等. 不同缓/控释尿素在黄土台塬区春玉米的减量施用效果[J]. 中国生态农业学报, 2017, 25(6): 848 – 855.
SUN Xiao, JING Jianyuan, LÜ Shenqiang, et al. Effect of different rates of slow/controlled release urea on nitrogen content in spring maize in loess highlands[J]. Chinese Journal of Eco-Agriculture, 2017, 25(6): 848 – 855. (in Chinese)

[40] 武鹏, 杨克军, 王玉凤, 等. 缓释尿素对土壤和玉米植株氮素及干物质和产量的影响[J]. 干旱地区农业研究, 2018, 36(5): 94 – 101.
WU Peng, YANG Kejun, WANG Yufeng, et al. The effects of slow-release urea on soil and plant nitrogen content, dry matter and yield of maize[J]. Agricultural Research in the Arid Areas, 2018, 36(5): 94 – 101. (in Chinese)

(上接第 263 页)

[24] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[J]. CoRR, 2015; abs/1512.03385.

[25] SAINING X, ROSS G, PIOTR D, et al. Aggregated residual transformations for deep neural networks[J]. CoRR, 2016; abs/1611.05431.

[26] 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 Wenxia, WU Dezhaoh, HU Gensheng, 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)

[27] 何自芬, 黄俊璇, 刘强, 等. 基于非对称混洗卷积神经网络的苹果叶部病害分割[J]. 农业机械学报, 2021, 52(8): 221 – 230.
HE Zifen, HUANG Junxuan, LIU Qiang, et al. High precision identification of apple leaf diseases based on asymmetric shuffle convolution[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(8): 221 – 230. (in Chinese)

[28] CHENG X, ZHANG Y H, CHEN Y Q, et al. Pest identification via deep residual learning in complex background[J]. Computers and Electronics in Agriculture, 2017, 141: 351 – 356.

[29] HOEL K, JIHENE B, CHRISTIAN D, et al. Boundary loss for highly unbalanced segmentation[J]. Medical Image Analysis, 2020, 67: 101851.

[30] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[J]. Springer, Cham, 2018, 11072(64): 3 – 19.

[31] 章权兵, 胡姗姗, 舒文灿, 等. 基于注意力机制金字塔网络的麦穗检测方法[J]. 农业机械学报, 2021, 52(11): 253 – 262.
ZHANG Quanbing, HU Shanshan, SHU Wencan, et al. Wheat spikes detection method based on pyramidal network of attention mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(11): 253 – 262. (in Chinese)

[32] WANG Q J, ZHANG S Y, DONG S F, et al. Pest24: a large-scale very small object data set of agricultural pests for multi-target detection[J]. Computers and Electronics in Agriculture, 2020, 175:105585.