

郑果,姜玉松,沈永林.基于改进YOLOv7的水稻害虫识别方法[J].华中农业大学学报,2023,42(3):143-151.  
DOI:10.13300/j.cnki.hnlkxb.2023.03.017

## 基于改进YOLOv7的水稻害虫识别方法

郑果<sup>1</sup>,姜玉松<sup>2</sup>,沈永林<sup>3</sup>

1. 重庆三峡学院生物与食品工程学院,重庆 404000; 2. 重庆文理学院园林与生命科学学院,重庆 402100;  
3. 中国地质大学(武汉)国家GIS工程研究中心,武汉 430074

**摘要** 为解决水稻害虫体型小且不同类型害虫外观差异小、同类型害虫不同生长过程中外观差异大导致水稻害虫难以识别的问题,将卷积注意力特征金字塔模块引入图像识别网络YOLOv7。以湖北省鄂州市水稻种植基地为样本采集点,构建一个具有挑战性的大规模水稻虫害数据集;根据样本分布特点进行数据增强,引入随机噪声、Mixup、Cutout等数据增强方法,使深度学习模型从更深的维度学习害虫判别力视觉特征;将MobileNetv3作为主干网络,对YOLOv7网络进行改进,并构建基于特征金字塔的多尺度神经网络模型,提升小个体害虫的识别精度。试验结果显示,基于改进YOLOv7的水稻虫害检测平均准确率为85.46%,超越YOLOv7、EfficientNet-B0等网络。改进YOLOv7模型大小为20.6 M,检测速度为92.2 帧/s,检测速度是原始YOLOv7算法的5倍以上。结果表明,该方法能用于实现水稻虫害远程实时自动化识别。

**关键词** 智慧农业; 害虫识别; 深度学习; 卷积神经网络; 空间注意力; 图像视觉; 虫害监测

**中图分类号** S431.9 **文献标识码** A **文章编号** 1000-2421(2023)03-0143-09

水稻虫害是威胁水稻粮食安全生产的主要原因之一。由于我国水稻种植面积广,地域差异大,气候条件变化复杂,虫害呈现基数大、种类多、发展快的特点,给水稻虫害防治工作带来巨大挑战。常见水稻病虫害威胁以稻飞虱、卷叶螟、二化螟、稻曲病、稻瘟病、纹枯病等“三虫三病”为主<sup>[1-2]</sup>。据统计,我国每年因虫害造成水稻减产达到5%以上<sup>[1]</sup>。因此,及时准确地掌握水稻不同生长时期虫害类型和覆盖区域,既能快速采取针对性治理措施,降低虫害对水稻生产造成的经济损失,又能避免盲目用药造成生态环境的污染。传统水稻虫害监测主要依赖于专家通过抽样观测水稻叶片或根茎表面害虫外观特征进行判别,这种以人工方式为主的主观判别方法准确率与专家的经验密切相关,同时判别过程较为繁琐、耗时,难以满足实际生产中大规模、快速虫害监测的应用需求。

近年来,基于机器视觉的水稻病虫害自动识别技术引起科研工作者普遍关注。此类方法的基本原理是首先构建图像视觉特征表示模型,然后将该模型在带有标签的训练数据集中进行训练,确定算法

模型参数,最后在测试数据集中验证算法性能。从视觉特征表示角度看,基于机器视觉的水稻病虫害识别方法大致可以分为两类:基于手工特征<sup>[3-6]</sup>和基于深度学习特征<sup>[7-9]</sup>。手工特征指人们根据图像的颜色、纹理、边缘等像素分布特点,构建多种视觉特征表达模型,例如颜色直方图<sup>[5]</sup>、局部二值模式<sup>[6]</sup>等。手工特征由于特征表示简单,仅对浅层视觉进行表达,因此,在实际应用中,采用多种手工特征组合进行视觉表达。例如,马鹏鹏等<sup>[3]</sup>针对图像特征选择和样本失衡等问题,根据5种目标害虫体型大小,将水稻灯诱昆虫分成大型和小型昆虫,研究全局视觉特征、局部视觉特征以及二者融合对水稻目标害虫识别结果的影响。鲍文霞等<sup>[5]</sup>采用滑动窗口法提取水稻虫害图像的HSV颜色特征和SILTP纹理特征,采用深度语义分割U-Net网络去除复杂背景,引入对数据具有更好区分性的椭圆型度量模型,提取水稻虫害图像空间结构和语义特征,对虫害图像特征的潜在关系进行建模。杨颖等<sup>[6]</sup>针对水稻病虫害类型所需样本多、处理设备要求高等问题,提出基于方向梯度直方图特征和局部二值模式特征加权融合的

收稿日期: 2022-09-27

基金项目:国家自然科学基金面上项目(42271397)

郑果, E-mail: zhengguo19840@163.com

通信作者: 姜玉松, E-mail: jysong@126.com

方法提取水稻病虫害图像视觉特征。深度学习特征指通过深度神经网络模型对图像建立视觉特征表达。广泛应用的神经网络模型为深度卷积神经网络,该模型通过堆叠卷积层、池化层、激活层,最优通过全连接层进行图像分类。该模型的显著优点是特征表达能力强,网络模型通过底层卷积层可以提取局部浅层视觉特征,高层卷积层可以提取丰富的全局语义特征。与手工特征不同,该模型特征表达是从图像像素中学习具有判别力的特征,其特征表达与训练数据密切相关,而手工特征表达与训练数据无关。例如,黄双萍等<sup>[7]</sup>通过Inception模块堆叠构建深度神经网络建立水稻穗瘟病图像特征表达,利用多尺度卷积核提取不同尺度穗瘟病斑分布式特征并进行级联融合,提出基于深度卷积神经网络GoogLeNet模型的水稻穗瘟病检测方法。谭云兰等<sup>[8]</sup>针对水稻图像病斑交界特征复杂多变问题,采用深度卷积神经网络模型,使用数据集扩增技术,运用参数微调方法对网络进行优化,实现自然场景下采集的常见8类水稻病害图像识别。范春全等<sup>[9]</sup>针对数据量小和数据种类不够丰富导致水稻病虫害识别性能下降问题,构建了包含16种水稻病虫害,共计2万余张图像的完整水稻病虫害识别数据集,并以ResNet50作为主干网络,设计并验证该水稻病虫害识别模型。

以上关于水稻病虫害识别的研究主要侧重病害图像视觉特征表达,而针对虫害识别研究偏少。由于水稻害虫具有个体较小、不同类别害虫外观差异小(即不同类型害虫外观具有相似性,例如螟虫和卷叶螟、二化螟等)、相同类别害虫外观差异大(即同类型害虫在不同生长时期,外观差异较大,例如螟虫在幼虫、成虫外观变化较大)的特点导致实际大田水稻虫害精确识别非常困难。因此,本研究针对上述问题,通过对大田水稻生长过程进行长期观测,建立水稻害虫影像数据集,采用随机噪声、Mixup<sup>[10]</sup>、Cut-out<sup>[11]</sup>等图像融合数据增强方法对数据集进行扩展,并引入混合注意力模块和特征金字塔模块对YOLOv7进行改进,提出一种轻量级水稻虫害识别模型,旨在为水稻害虫远程自动识别提供参考。

## 1 材料与方法

### 1.1 试验数据采集

试验数据采集于湖北省农业科学院鄂州市水稻

实验基地(湖北省鄂州市杜山镇东港村),水稻种植面积约33 hm<sup>2</sup>。为保证数据具有普适性,数据采集地点选择实验基地中自然种植的1 hm<sup>2</sup>左右的大田。采用佳能EOS 6D型数码单反相机(镜头参数为17~40 mm f/4L USM和100 mm f/2.8L IS USM),对水稻害虫实体进行近距离(约10~50 cm)拍摄,图像分辨率为5 472像素×3 648像素。拍摄时采用光圈优先模式和防抖模式,保证图像中害虫实体清晰度。为了更好地研究水稻害虫完整生命周期内外观特征分布特点,采集时间为2021年4月中旬至10月中旬、07:00至19:00时间段,选择在早稻和晚稻种植周期内,按秧苗期、分蘖期、拔节期、孕穗期、抽穗期、扬花授粉期和灌浆期等不同时段进行均匀采样。经过数据清洗,将失焦、抖动、曝光等因素导致成像质量差的图像去掉。数据集包含灰飞虱、稻瘿蚊、稻蓟马、稻飞虱、螟虫、稻纵卷叶螟等6种害虫图像,样本数量分别为3 052、2 956、3 126、3 019、3 126、2 927张,共计为18 206张。此外,没有包含任何害虫的图像作为背景类,样本数量为4 500张。样本实例如图1所示。图1的第1至第6行分别为灰飞虱、稻瘿蚊、稻蓟马、稻飞虱、螟虫、稻纵卷叶螟的样本实例。

### 1.2 图像处理与标注

深度卷积神经网络模型运行所需的计算资源和存储资源与输入图像的分辨率成正比关系,当前经典的神经网络模型输入图像分辨率为224像素×224像素,例如ImageNet图像识别数据集。水稻数据集中图像分辨率为5 472像素×3 648像素,直接作为训练数据,深度学习模型需要的计算和存储资源超过现有深度学习系统承受极限。因此,以图像中害虫对象为中心,使用MATLAB视觉库对原始图像进行裁剪和归一化处理,将图像分辨率降为600像素×600像素。采用Labelme标注工具对样本图像进行类别标注,标注格式为PASCAL VOC2007,并对标注结果进行校正,减少因标注错误引入的噪声。为了解决因遮挡导致深度模型识别性能下降问题,数据集中遮挡害虫图像数量占数据集30%。根据遮挡率(图像中对象被遮挡部分与对象实际面积比值)差异对遮挡数据集进行划分,其中严重遮挡(遮挡率50%~75%)占10%,中等遮挡(遮挡率30%~50%)占20%,轻微遮挡(遮挡率低于30%)占70%。

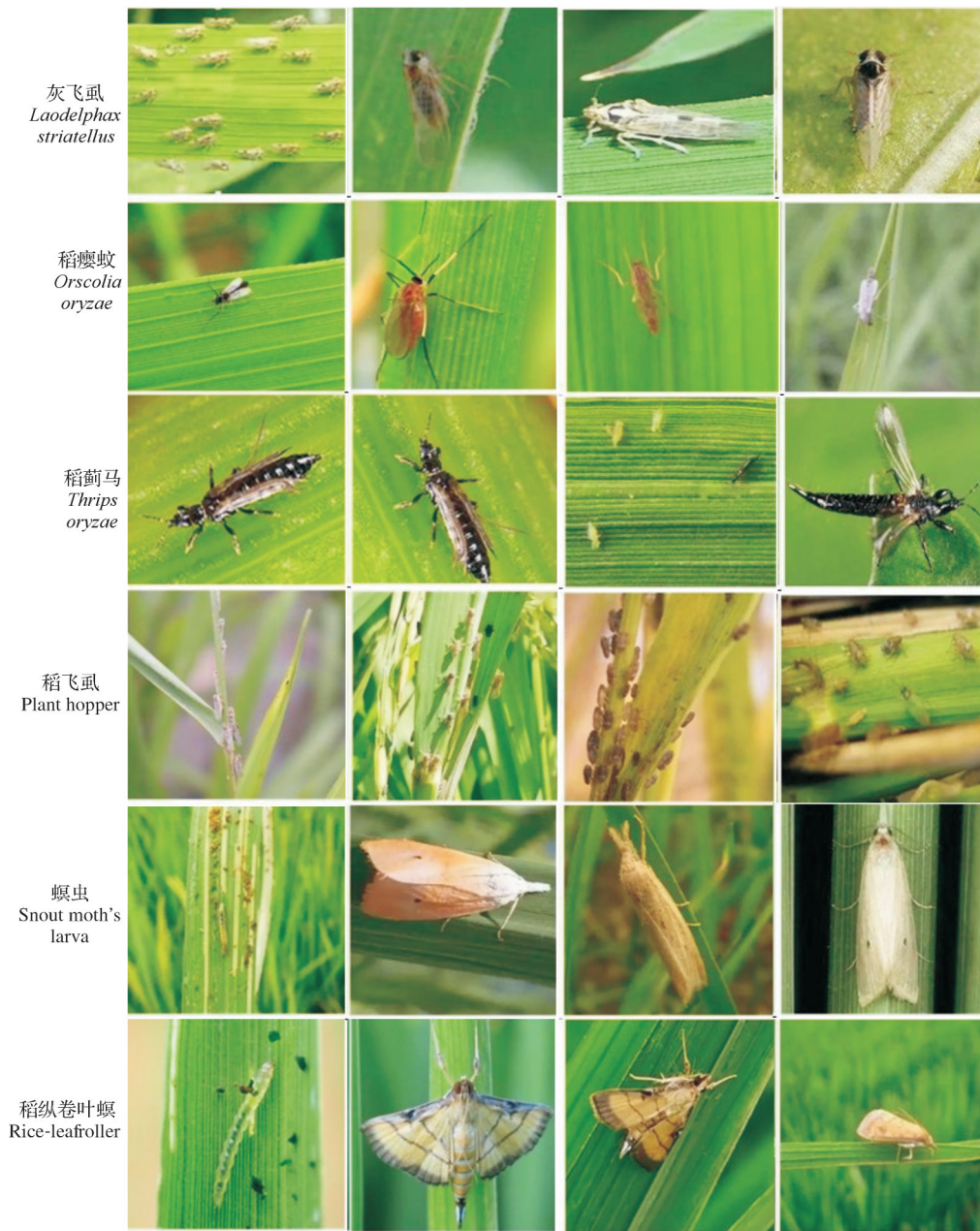


图 1 水稻害虫数据集

Fig.1 Rice pests dataset

### 1.3 数据增强

所有图像数据进行裁剪和归一化预处理,保证害虫在图像中心,每张图像分辨率为 500 像素×500 像素。虽然在数据采集时尽量保证水稻 6 类害虫样本数量相同,但不同生长时期的害虫样本数量差距较大,例如灰飞虱和稻飞虱幼虫样本数量仅为 200 张左右,而螟虫幼虫样本数量达 807 张。此外,害虫处于水平和垂直方向的样本居多,其他角度的样本数量很少。不同尺度的害虫样本数量也差别较大。如果害虫识别模型直接在该数据集下训练,

可能导致其在样本数量较少的类别上性能表现较差。

为了解决这种样本数据不均衡导致模型性能下降的问题,本研究采用多种数据增强方法,将样本较少的害虫类别图像进行扩充。首先,通过翻转(水平和垂直 2 个方向)、平移、随机裁剪、旋转、缩放(按照 10%、20%、30% 比例放大)、加高斯噪声、Mixup 和 Cutout 图像融合等方法将原始图像转化为 10 张,使得所有类别害虫图像数量达到 8 000 张,然后,将所有数据进行归一化,使其分辨率为 300 像素×300 像

素。为了减少数据集划分对试验的影响,本研究采用随机划分方式,将增广数据集按照7:1:1的比例,分成训练集、验证集和测试集。采用交叉验证方式,

随机重复10次。每次划分后,模型进行训练和测试,最后取10次试验结果的平均值作为最后试验结果。数据增强效果如图2所示。

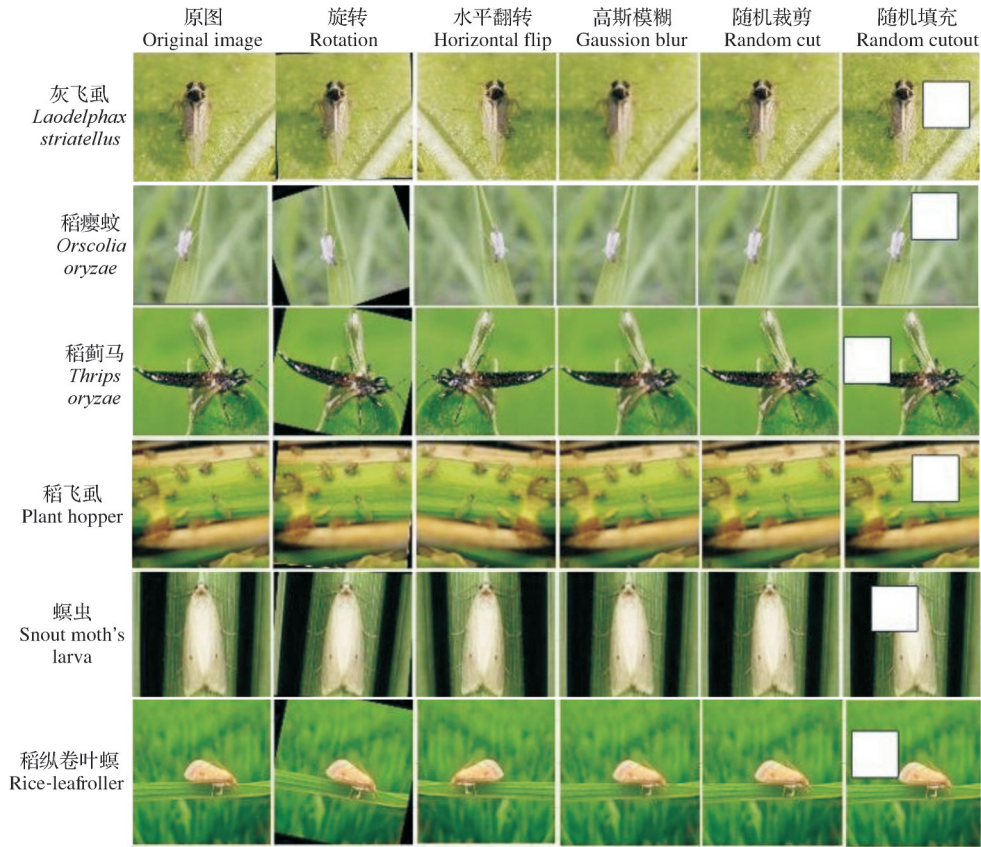


图2 增强后的样本图像

Fig.2 Enhanced sample image

#### 1.4 YOLOv7 目标识别算法改进

YOLOv7<sup>[12]</sup>目标识别算法是在YOLOv1<sup>[13]</sup>的基础上经过一系列改进后提出的。YOLOv7保持了YOLO系列算法的实时性特点,其准确度超过了所有已知的实时目标检测器,在ImageNet和MS COCO等大规模图像识别数据集上具有最高的平均准确率。Wang等<sup>[12]</sup>提出E-ELAN模块,采用组卷积扩展计算块的通道和基数,通过控制梯度路径长度,有效地学习图像特征,加速网络收敛。该模型对计算层的所有计算块采用相同的组参数和通道乘数,预测一个包含边界框、置信度和类别的三维张量编码,实现目标定位和分类。

1)卷积块注意力模型。通过观察数据集中害虫图像,发现部分害虫图像中害虫个体小、数量大,且在图像中分布位置差异较大,为了使卷积神经网络集中在害虫分布区域学习鲁棒的特征表示,本研究在YOLOv7网络后端引入注意力模型。在人类大脑

视觉皮层中,注意力机制能够更快捷和高效地分析复杂场景信息。视觉注意力机制是一种通过设计特殊的神经网络结构,让主干深度神经网络在特征学习时把注意力聚集在图像重要区域的方法。本文引入一种新的轻量级注意力模型CBAM(convolutional block attention module)<sup>[14]</sup>,由通道注意力(channel attention module, CAM)和空间注意力(spatial attention module, SAM)2个子模块组成,如图3所示。

在通道注意力子模块中,假设输入特征图为 $F(H \times W \times C)$ ,其中, $H$ 、 $W$ 、 $C$ 分别为特征图的高、宽和通道数,分别经过基于水平和垂直方向的全局最大池化操作和全局平均池化操作之后,生成2个 $1 \times 1 \times C$ 的特征图;再将其分别输入到1个2层的全连接神经网络,即共享的多层感知机,其中第1层神经元个数为 $C/r$ ( $r$ 为压缩率),激活函数为ReLU;第2层神经元个数为 $C$ ,共享多层感知机输出的2个特征向量进行元素加和运算,再经过sigmoid激活函数,得

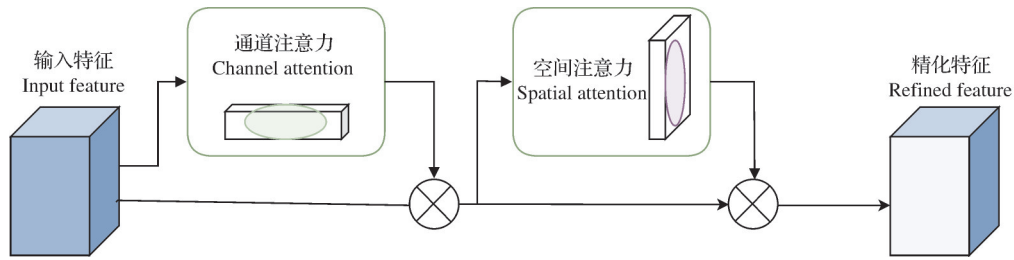


图3 卷积块注意力模型

Fig.3 Convolutional block attention module

到通道注意力特征,最后将该特征与输入特征进行元素乘法运算,其结果作为空间注意力子模块的输入。

在空间注意力子模块中,首先将上述通道注意力子模块产生的特征图通过基于通道的全局最大值池化和全局平均值池化操作,分别得到2个 $H \times W \times 1$ 的特征图;然后将这2个特征图进行基于通道的拼接操作,再通过 $7 \times 7$ 卷积操作,获得 $H \times W \times 1$ 的特征图;最后,该特征图通过sigmoid激活生成空间通道特征,并与输入特征图进行元素级乘法运算,得到最终特征。

2)替换主干网络。为了实现在移动设备上部署水稻害虫识别算法,使用轻量级主干网络MobileNetv3<sup>[15]</sup>,替换YOLOv7原始骨干网络DarkNet-53,不仅可以提高算法识别精度,还能有效降低模型计算代价和存储空间。MobileNetv3在MobileNetv2<sup>[16]</sup>基础上进行改进,融合深度可分离卷积结构、基于NetAdapt<sup>[17]</sup>的网络结构搜索、SE通道注意力网络等核心模块,大幅减少模型参数数量,同时提高了模型的特征表达能力。

3)预训练模型。深度学习模型是多层神经网络,参数数量超过百万,模型训练过程是利用后向传播将误差从高层向底层逐层传递,以误差最小化为约束条件,确定最优参数数值。当训练数据集规模较小时,模型中部分参数不是最优解,会导致模型泛化能力差。为解决这个问题,深度学习模型通常会设置默认参数,在大规模数据集上进行充分训练,然后将收敛之后的模型参数作为初始参数再在目标数据集上进行再训练,微调模型参数。本研究也采用预训练方式,改进的YOLOv7模型在ImageNet数据集进行预训练,将训练好的模型参数作为初始参数,在水稻病虫害数据集上进行再次训练。

4)试验环境。试验平台硬件部分为一台包含Intel i7 10700k处理器,32 G DDR4内存,2 T硬盘和2张GTX 3060 12 G显存GPU卡的深度学习服务器;

软件部分为英伟达公司发布的深度学习加速库CUDA Toolkit 9.0和CUDNN V10.0;Tensorflow-GPU 1.8.0深度学习框架。

5)评价标准。试验结果评价指标为平均准确率和召回率。假设单个类别识别的精确度(precision,  $P$ )、召回率(recall,  $R$ )、平均精度(average precision,  $P_A$ )、平均精度均值(mean average precision,  $P_{mA}$ )计算如下:

$$P = T_p / (T_p + F_p) \quad (1)$$

$$R = T_p / (T_p + F_n) \quad (2)$$

$$P_A = \int_0^1 P(R) dR \quad (3)$$

$$P_{mA} = \frac{1}{n} \sum_{i=1}^n P_{A_i} \quad (4)$$

其中, $T_p$ 是表示样本预测标签为正且实际的标签也为正的样本数量; $F_p$ 表示样本预测标签为正且实际的标签为负的样本数量; $F_n$ 表示样本预测标签为负而实际标签为正的样本数量; $P_A$ 是以召回率为横坐标、准确率为纵坐标绘制 $P-R$ 曲线,曲线与坐标轴围成的面积,衡量模型对某类害虫识别的效果。

另一个重要评价指标是检测速度,即每秒内检测图片数量,采用帧/s(frame per second, fps)表示,检测速度是实现实时检测的基础。

6)对比算法。为了验证不同主干网络对YOLOv7算法识别精度的影响,本研究选择当前主流的主干网络作为测试基准,分别为Darknet-53、EfficientNetB0<sup>[18]</sup>、ResNet50和MobileNetv3,其中Darknet-53为原始YOLOv7的主干网络,MobileNetv3为本研究选择的替代主干网络;为了验证本研究提出算法模型的有效性,选择当前主流的水稻害虫识别算法ADM\_Ret<sup>[18]</sup>和EML<sup>[5]</sup>作为对比测试方法。

## 2 结果与分析

### 2.1 模型训练

本研究提出的模型在ImageNet数据集上进行图像识别预训练,然后将预训练模型在水稻害虫数据

集下进行 fine-tune 训练。初始学习率为 0.000 1, 迭代次数为 2 000, 每迭代 100 次后, 学习率退化为原来的 10%。采用批量训练的方法将训练集划分为多个批次, 每个批次训练 100 张图片, 即 batch\_size=100。模型训练损失和准确率变化如图 4 所示。

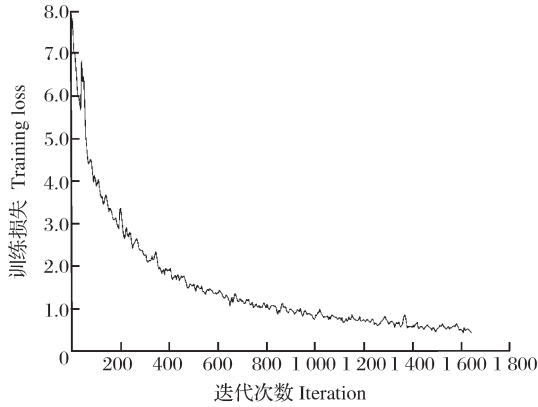


图4 模型训练损失曲线图

Fig.4 Model training loss graph

## 2.2 消融实验

针对水稻害虫不同生长期外观形态变化大、个体较小导致虫害识别精度下降的问题, 本研究提出通用数据增强、图像融合增强、预训练等方法对数据集进行扩展和训练参数迁移。基于改进 YOLOv7 算法, 分别进行 4 组消融实验验证所提出方法的有效性。实验结果如表 1 所示。

表 1 消融实验

Table 1 Ablation experiment

试验序号 Test No.	预训练 Pre-training	数据增强 Data augment	图像融合 Image mixup	$P_{mA}/\%$
1	—	—	—	78.82
2	✓	—	—	80.73
3	✓	✓	—	82.62
4	✓	✓	✓	85.29

从表 1 中可以看出, 未对数据集进行增强操作、图像融合增强、预训练等处理步骤时, 水稻害虫平均识别准确率  $P_{mA}$  为 78.82%, 而依次采取预训练、常规数据增强和图像融合增强后, 水稻害虫平均识别准确率分别提升了 1.91、3.80、6.47 个百分点。由此可见, 模型预训练、数据增强和图像融合增强等操作对水稻害虫识别的性能提升具有重要作用。

## 2.3 骨干网络性能对比

YOLOv7 算法应用 4 个不同骨干网络模型 Darknet-53、EfficientNetB0、ResNet50 和 MobileNetv3 分

别在本数据集下训练测试。模型平均精确度、检测速度和模型大小如表 2 所示。

表 2 不同骨干网络的模型性能对比

Table 2 Model performance comparison of different backbone networks

骨干网络 Backbone network	$P_{mA}/\%$	检测速度/ (帧/s) Detection speed	模型大小/M Model size
YOLOv7-Darknet-53	80.25	16.30	130.62
YOLOv7-ResNet-50	82.46	12.89	89.82
YOLOv7-EfficientNet-B0	81.75	15.82	63.86
YOLOv7-MobileNetv3	85.29	92.21	20.63

从表 2 可以看出, 不同骨干网络对模型性能具有重大影响。与原始 YOLOv7-Darknet-53 相比, 本研究替换的 MobileNetv3 骨干网络提升了模型的识别精度,  $P_{mA}$  达到 85.29%, 比当前最优主干网络 EfficientNet-B0 提高 3.5 个百分点; 模型大小降低为原来 1/6, 仅为 20.6 M, 模型识别速度提升了 5 倍以上。

## 2.4 前沿算法对比

为测试本研究方法的性能, 将近年来基于机器视觉的水稻虫害识别前沿算法与本方法进行性能对比实验。为保证实验客观和公平, 所有算法在相同的训练集和测试集下进行训练和测试, 训练参数以算法测试结果最优为准。试验结果如表 3 所示。

表 3 与前沿算法的性能对比

Table 3 Performance comparison of our method vs the state-of-art methods

识别方法 Recognition methods	$P_{mA}/\%$	检测速度/(帧/s) Detection speed
ADM_Ret	83.69	23.86
EML	76.56	125.49
YOLOv7-MobileNetv3	85.29	92.21

从表 3 可以看出, ADM\_Ret、EML 算法在 6 类水稻害虫数据集上的平均识别精度均值分别为 83.69% 和 76.56%, 而本研究方法的平均识别精度均值达到 85.29%, 高于对比算法。由于 ADM\_Ret 算法采用 RetinaNet 网络模型, 模型深度和宽度不及本研究采用的 YOLOv7 模型, 且 YOLOv7 汲取了前者的优点, 在网络结构设计、样本均衡和困难样本挖掘以及代价函数等方面进行改进, 因此, 后者识别性能更好。EML 算法采用传统底层视觉特征, 即颜色和纹理特征作为害虫视觉特征表达, 引入椭圆度量学习提

升特征判别能力,在少量样本的数据集上性能表现较好,在大规模数据集上性能表现较差。但因为特征表达模型简单,时间复杂度较低,因此,该算法检测速度最快,达到 125.49 帧/s。

### 2.5 模型预测结果

本研究提出的模型在水稻害虫识别数据集上对 6 类害虫正确分类与错误分类的详细结果如图 5 所示。归一化混淆矩阵将子类预测结果进行归一化,颜色越深表明概率值越高。

真实标签 True label	预测标签 Predict label					
	灰飞虱 <i>Laodelphax striatellus</i>	稻飞虱 <i>Orscolia oryzae</i>	稻飞虱 <i>Plant hopper</i>	螟虫 <i>Snout moth's larva</i>	稻蓟马 <i>Thrips oryzae</i>	稻纵卷叶螟 <i>Rice-leaf roller</i>
灰飞虱 <i>Laodelphax striatellus</i>	0.852	0.034	0.027	0.038	0.014	0.035
稻飞虱 <i>Orscolia oryzae</i>	0.015	0.886	0.032	0.016	0.020	0.031
稻飞虱 <i>Plant hopper</i>	0.029	0.011	0.864	0.023	0.039	0.034
螟虫 <i>Snout moth's larva</i>	0.024	0.033	0.029	0.833	0.036	0.045
稻蓟马 <i>Thrips oryzae</i>	0.025	0.031	0.036	0.037	0.845	0.026
稻纵卷叶螟 <i>Rice-leaf roller</i>	0.017	0.021	0.034	0.031	0.016	0.881

图 5 归一化混淆矩阵

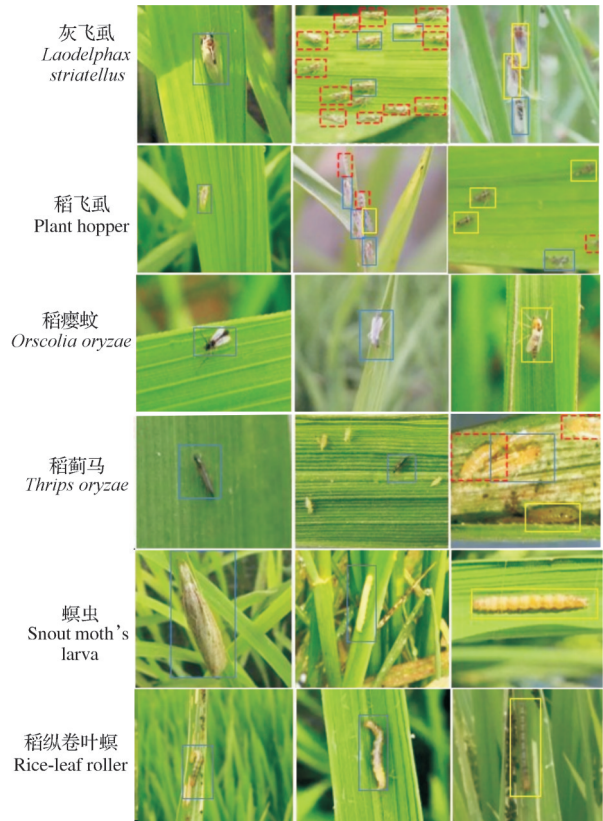
Fig.5 Normalized confusion matrix

图 5 中第一行表示标签为“灰飞虱”的样本,模型预测为“灰飞虱”标签的概率为 0.852,预测为“稻飞虱”“稻飞虱”“螟虫”“稻蓟马”“稻纵卷叶螟”的概率分别为 0.034、0.027、0.038、0.014 和 0.035。图 6 列举了水稻害虫识别结果。

从图 6 可以看出,当图像中仅包含单个实例时,害虫外观清晰可见,算法识别准确率高;当个体密集出现时,由于拍摄距离远,个体较小导致细节特征损失,容易出现漏检或错误识别;幼虫期实例识别由于外观相似容易出现识别错误。

## 3 讨论

姚青等<sup>[19]</sup>采用 ResNeXt101 作为主干网络,将特征金字塔网络结构引入到 RetinaNet 网络模型,再进行水稻害虫识别。鲍文霞等<sup>[5]</sup>采用传统底层视觉特征,即颜色和纹理特征作为害虫视觉特征表达,引入椭圆度量学习提升特征判别能力。前者采用深度学习特征对害虫图像进行建模,该模型的主干网络层数达到 101 层,在本研究构建的数据集上,平均准



图中蓝色实线框表示害虫识别为正确的实例,黄色实线框表示识别错误实例,红色虚线框表示漏检实例。The blue solid line box, the yellow solid line box and the red dashed line box indicate the correct instance, the identification error, and the missed detection, respectively.

图 6 水稻害虫识别结果

Fig.6 Results of rice pest recognition

确率为 83.69%,但该网络模型对外观相似的害虫识别效果不佳。后者采用多种手工特征融合对害虫图像进行建模,由于手工特征表达能力不及深度学习特征,所以,该方法在本文构建的数据集上,平均准确率仅为 76.56%。此外,这 2 种方法计算复杂度高,不适合在计算资源受限的设备上部署应用。本文提出的基于改进 YOLOv7 的深度卷积神经网络模型,通过替换轻量级骨干网络、引入注意力机制和特征金字塔融合模块,有效提高了模型的判别力和计算速度。与上述算法比较,本研究采用的 YOLOv7 模型,在网络结构设计、样本均衡和困难样本挖掘以及代价函数等方面进行改进后,水稻害虫识别平均准确率达到 85.29%。在实际应用场景下,受摄像机成像条件制约,水稻虫害图像分辨率变化较大,而卷积神经网络模型对低分辨率目标识别精度较低,因此,未来研究工作将围绕低分辨率水稻害虫识别展开。

因此,本研究提出的方法在大田水稻害虫识别任务上,其平均准确率和召回率比原始YOLOv7模型具有明显优势。该方法能够应用于大田水稻病虫害远程自动化诊断,提高农业生产智能化水平,可为后续深入研究提供参考。

## 参考文献References

- [1] 全国农业技术推广服务中心. 2021年全国粮食作物重大病虫害发生趋势预报[J]. 中国植保导刊, 2021, 41(1): 37-39, 42. National Agricultural Technology Extension Service Center. Forecast of occurrence trend of major diseases and insect pests of grain crops in 2021[J]. China plant protection, 2021, 41(1): 37-39, 41(in Chinese).
- [2] 刘芹, 杨俊杰, 袁浩, 等. 2021年湖北省农作物主要病虫害发生趋势[J]. 湖北植保, 2021(2): 51-54. LIU Q, YANG J J, YUAN H, et al. Occurrence trend forecast of the major diseases and insect pests of main crops in Hubei Province in 2021[J]. Hubei plant protection, 2021(2): 51-54 (in Chinese with English abstract).
- [3] 马鹏鹏, 周爱明, 姚青, 等. 图像特征和样本量对水稻害虫识别结果的影响[J]. 中国水稻科学, 2018, 32(4): 405-414. MA P P, ZHOU A M, YAO Q, et al. Influence of image features and sample sizes on rice pest identification[J]. Chinese journal of rice science, 2018, 32(4): 405-414 (in Chinese with English abstract).
- [4] 景晨. 基于无人机平台和图像处理的水稻感染稻纵卷叶螟虫害自动识别技术[J]. 农业开发与装备, 2018(11): 98-100. JING C. Automatic identification technology of rice pests infected with *Cnaphalocrocis medinalis* based on UAV platform and image processing [J]. Agricultural development and equipments, 2018(11): 98-100 (in Chinese).
- [5] 鲍文霞, 邱翔, 胡根生, 等. 基于椭圆型度量学习空间变换的水稻虫害识别[J]. 华南理工大学学报(自然科学版), 2020, 48(10): 136-144. BAO W X, QIU X, HU G S, et al. Identification of rice pests based on space transformation by elliptic metric learning[J]. Journal of South China University of Technology (natural science edition), 2020, 48(10): 136-144 (in Chinese with English abstract).
- [6] 杨颖, 文小玲, 章秀华. 基于方向梯度直方图和局部二值模式混合特征的水稻病虫害识别方法研究[J]. 河南农业大学学报, 2021, 55(6): 1089-1096. YANG Y, WEN X L, ZHANG X H. Study of diseased rice image recognition based on histograms of oriented gradients and local binary pattern hybrid features[J]. Journal of Henan Agricultural University, 2021, 55(6): 1089-1096 (in Chinese with English abstract).
- [7] 黄双萍, 孙超, 齐龙, 等. 基于深度卷积神经网络的水稻穗瘟病检测方法[J]. 农业工程学报, 2017, 33(20): 169-176. HUANG S P, SUN C, QI L, et al. Rice panicle blast identification method based on deep convolution neural network [J]. Transactions of the CSAE, 2017, 33(20): 169-176 (in Chinese with English abstract).
- [8] 谭云兰, 欧阳春娟, 李龙, 等. 基于深度卷积神经网络的水稻病害图像识别研究[J]. 井冈山大学学报(自然科学版), 2019, 40(2): 31-38. TAN Y L, OUYANG C J, LI L, et al. Image recognition of rice diseases based on deep convolutional neural network[J]. Journal of Jinggangshan University (natural science), 2019, 40(2): 31-38 (in Chinese with English abstract).
- [9] 范春全, 何彬彬. 基于迁移学习的水稻病虫害识别[J]. 中国农业信息, 2020, 32(2): 36-44. FAN C Q, HE B B. Identification of rice diseases and insect pests using transfer learning [J]. China agricultural informatics, 2020, 32(2): 36-44 (in Chinese with English abstract).
- [10] ZHANG H Y, CISSE M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization [DB/OL]. arXiv, 2017: 1710.09412 [2022-09-27]. <https://doi.org/10.48550/arXiv.1710.09412>.
- [11] DEVRIES T, TAYLOR G W. Improved regularization of convolutional neural networks with cutout [DB/OL]. arXiv, 2017: 1708.04552 [2022-09-27]. <https://doi.org/10.48550/arXiv.1708.04552>.
- [12] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [DB/OL]. arXiv, 2022: 2207.02696 [2022-09-27]. <https://doi.org/10.48550/arXiv.2207.02696>.
- [13] 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 (CVPR). June 27-30, 2016, Las Vegas, NV, USA. Las Vegas: IEEE, 2016: 779-788.
- [14] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [M] // Computer Vision - ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [15] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27 - November 2, 2019, Seoul, Korea (South). Seoul: IEEE, 2020: 1314-1324.
- [16] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018, Salt Lake City, UT, USA. Salt Lake City: IEEE, 2018: 4510-4520.
- [17] YANG T J, HOWARD A, CHEN B, et al. NetAdapt: platform-aware neural network adaptation for mobile applications [M] // Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 289-304.
- [18] TAN M X, LE Q V. EfficientNet: rethinking model scaling for convolutional neural networks [C] // International Conference on Machine Learning. [S.l.]: [s.n.], 2019: 6105-6114.
- [19] 姚青, 谷嘉乐, 吕军, 等. 改进RetinaNet的水稻冠层害虫为害状自动检测模型[J]. 农业工程学报, 2020, 36(15): 182-188.



YAO Q, GU J L, LÜ J, et al. Automatic detection model for pest damage symptoms on rice canopy based on improved Reti-

naNet[J]. Transactions of the CSAE, 2020, 36(15): 182-188 (in Chinese with English abstract).

## Recognition of rice pests based on improved YOLOv7

ZHENG Guo<sup>1</sup>, JIANG Yusong<sup>2</sup>, SHEN Yonglin<sup>3</sup>

1. *College of Biology and Food Engineering, Chongqing Three Gorges University, Chongqing 404000, China;*
2. *College of Landscape and Life Sciences, Chongqing University of Arts and Sciences, Chongqing 402100, China;*
3. *National Engineering Research Center for Geographic Information System, China University of Geosciences(Wuhan), Wuhan 430074, China*

**Abstract** Because rice pests are usually small in size, some different types of pests have similar appearance, and the same type of pests have different appearance in different growth processes, it is very difficult to identify rice pest types. We improved YOLOv7 neural network by introducing convolutional block attention module and feature pyramid module and constructed a challenging rice pest dataset, which is collected from rice planting base of Ezhou City in Hubei province to recognize rice pests. According to the characteristics of sample distribution, data enhancement was carried out, and random noise, Mixup, Cut-out and other data enhancement methods were introduced to make the deep learning model learn the visual features of pest discrimination from a deeper dimension. Taking MobileNetv3 as the backbone network, the YOLOv7 network was improved, and a multi-scale neural network model based on feature pyramid was constructed to improve the identification accuracy of small individual pests. The results showed that the average accuracy rate of rice pest detection based on the improved method is 85.46%, surpassing the networks such as YOLOv7 and Efficient Net-B0. The size of the improved YOLOv7 model is 20.6 M, and the detection speed is 92.2 frames/s, which is more than 5 times that of the original YOLOv7 algorithm. The results indicate that this method can be applied for remote automatic recognition of rice pests.

**Keywords** smart agriculture; pest recognition; deep learning; convolutional neural network; spatial attention; image vision; pest monitoring

(责任编辑:边书京)