

基于改进 YOLOv5s 的自然场景下生姜叶片病虫害识别

兰玉彬^{1,2,3}, 孙斌书^{1,3}, 张乐春^{1,3}, 赵德楠^{1,3}

(1. 山东理工大学农业工程与食品科学学院, 淄博 255049; 2. 山东省农业航空智能装备工程技术研究中心, 淄博 255049;
3. 山东理工大学生态无人农场研究院, 淄博 255049)

摘要: 为提高自然环境下生姜叶片病虫害的识别精确率, 提出一种基于改进 YOLOv5s 的生姜叶片病虫害识别模型。建立了田间不同自然环境条件下的生姜叶片病虫害数据集, 为保证模型在田间移动设备上流畅运行, 实现网络模型的轻量化, 在 YOLOv5s 中引入 GhostNet 网络中的 Ghost 模块和 Ghost BottleNeck 结构。同时, 为避免生姜叶片病虫害图像小目标特征丢失的情况, 增强图像特征提取, 加入 CA 注意力机制模块, 提升生姜叶片病虫害的识别准确率和定位精确度。改进后的模型参数量、计算量和权重文件大小分别为 YOLOv5s 模型的 52.0%、50.6% 和 55.2%, 对生姜叶片病虫害识别平均精度均值达到了 83.8%。与 Faster-RCNN、SSD、YOLOv4、YOLOv5s、Tea-YOLOv5s 等算法相比, 平均精度均值分别提高 37.6、39.1、22.5、1.5、0.7 个百分点, 将改进后的目标检测模型部署在 Jetson Orin NX 开发板上, 并使用 TensorRT、Int8 量化和 CUDA 等方法对检测模型加速, 加速后的模型检测速度为 74.3 帧/s, 满足实时检测的要求, 测试结果显示, 改进后的模型减少了漏检、误检的情况, 并且对目标定位更加精准, 适用于自然环境下生姜叶片病虫害的精准识别, 为后续生姜机械自动化施药作业提供技术理论支持。

关键词: 图像识别; 深度学习; 目标检测; 轻量化; 病虫害; 生姜叶片

doi: 10.11975/j.issn.1002-6819.202310124

中图分类号: S24

文献标志码: A

文章编号: 1002-6819(2024)-01-0218-07

兰玉彬, 孙斌书, 张乐春, 等. 基于改进 YOLOv5s 的自然场景下生姜叶片病虫害识别[J]. 农业工程学报, 2024, 40(1):

218-224. doi: 10.11975/j.issn.1002-6819.202310124 <http://www.tcsae.org>

LAN Yubin, SUN Binshu, ZHANG Lechun, et al. Identifying diseases and pests in ginger leaf under natural scenes using improved YOLOv5s[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(1): 218-224. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202310124 <http://www.tcsae.org>

0 引言

生姜是世界上重要的经济作物之一, 在中国已有 2 500 多年的栽培历史, 具有较高的食用和药用价值^[1]。2022 年中国生姜种植面积约为 36.87 万 hm², 年产量 1 219 万 t^[2], 为姜农带来极大经济效益。生姜在夏天易遭病虫害影响而减产, 其有黄尖病、炭疽病、姜瘟病、姜螟(玉米螟)、甜菜夜蛾幼虫等病虫害, 其中黄尖病和甜菜夜蛾幼虫最为常见^[3]。目前, 生姜叶片病虫害主要依靠人工施药防治, 劳动力短缺、机械智能化程度低的问题急需解决。对生姜叶片病虫害进行精准防治是减少农药使用, 提高产量的关键, 脱离人工施药促进智能施药机械发展是降低成本、提升生姜田间管理效率的重要手段。对于智能生姜施药车的设计, 首要任务是解决田间自然环境下对生姜叶片病虫害精准识别与定位的问题。

近年来, 随着图像处理检测算法的不断迭代更新, 国内外研究学者开始将其广泛应用于农业病虫害识别领域^[4-5]。RANGARAJAN 等^[6]使用改进的 AlexNet 和

VGG16 模型对番茄病害图像进行训练, 其模型缺陷在于检测精度不佳。蒋丰千等^[7]为实现对生姜病虫害的识别使用 GoogLeNet 模型中的 Inception 结构层, 与优化后的 LeNet-5 神经网络进行特征融合, 并在模型中引入新的 Swish 激活函数、BN 层以及 Gabor 滤波器, 该模型不仅训练时需要大量数据, 而且对生姜病害目标定位不够精准。白雪松等^[8]提出基于隐式语义数据增强算法的 CBAM-Res2Net50 模型, 在 AI Challenger 2018 农作物病虫害数据集上进行验证, 该模型一定程度上解决了由于数据不足导致的过拟合现象, 但模型参数量较大, 对硬件计算力要求较高, 不适用于移动端的部署。李子茂等^[9]在 YOLOv3 中引入 Inception 模块对叶片定位, 形成掩模图像, 并结合 Faster R-CNN 对月季叶片病虫害进行检测, 虽能够有效消除复杂背景对识别结果产生的影响, 但应用过程操作繁琐, 并且 YOLOv3 模型过大, 检测速度较慢, 不满足实时检测的要求。樊相鹏等^[10]为建立高效率的葡萄叶部病害检测系统, 引入迁移学习机制, 利用公开数据集对 VGG16 模型预训练, 在不改变前 13 层参数和权重的前提下对全连接层和分类层改进后使用新数据集对模型微调, 虽然能够完成葡萄叶部病虫害的识别, 但训练成本过大, 且不适用于小目标病虫害的检测。王卫星等^[11]针对荔枝病虫害检测提出了 YOLOv4-GCF 模型, 虽然此模型识别率较高, 但仍存在模型占用内存较大和检测速度慢的问题。

综合分析, 上述学者在目标检测技术存在检测精度

收稿日期: 2023-10-18 修订日期: 2023-12-20

基金项目: 山东省引进顶尖人才“一事一议”专项经费资助项目(鲁政办字[2018]27号); 宁夏回族自治区重点研发计划项目(项目编号: 2023BCF01051)

作者简介: 兰玉彬, 法国欧洲科学、艺术与人文学院院士, 博士, 教授, 博士生导师, 研究方向为精准农业航空应用技术, 现代农业装备等。
Email: ylan@sdu.edu.cn

不佳、定位精度低、模型计算量大、小目标识别效果差以及难以在移动设备部署等问题，本文以田间自然环境下生姜叶片病虫害为研究对象，避免后期在移动设备部署时因兼容性差而出现软件层面的问题，故选用 YOLO^[12-15] 系列中 TensorRT 高性能算子库兼容性好且综合性能较强的 YOLOv5s 模型为基础，针对上述问题进行优化，设计一种适用于移动设备部署的改进 YOLOv5s 的生姜叶片病虫害识别模型，为生姜智能化施药机械提供技术支持，并将改进后的模型与 Faster R-CNN^[16]、SSD^[17]、YOLOv4^[13] 等网络模型对比，以此评估模型的性能和效果。

1 材料与方法

1.1 生姜叶片病虫害图像采集

本文选取生姜叶片较为常见且特征明显的黄尖病和甜菜夜蛾幼虫两种病虫害作为研究对象，在相关专家指导下于山东省潍坊市寒亭区生姜种植基地进行图像数据拍摄，拍摄日期为 2023 年 8 月上旬至 9 月上旬期间，拍摄时间为 8:30 至 17:00，拍摄设备为小米 10 Ultra，在自然条件下进行拍摄，采集不同环境时的图像，其中包括晴天、阴天、雨后、顺光、逆光、近距离、远距离、俯角、仰角等多种情况，尽可能增加图像多样性，提高模型泛化能力，得到生姜叶片病虫害图像共 2 696 张，分辨率为 4 000×3 000 像素，为防止神经网络出现过拟合现象，从拍摄的图片中选出部分样本进行翻转、镜像、亮度调节、噪声添加及模糊处理等方法对数据集进行扩充，后对数据进行筛选，删除亮度过高、噪声过多、过度模糊等与实际场景下相差较大的图像，保留 3 848 张图像制作数据集。其中包含生姜黄尖病病害图像 1 977 张，

甜菜夜蛾幼虫虫害图像 1 268 张，同时存在两种病虫害图像 603 张，部分采集图像如图 1 所示。



图 1 生姜叶片病虫害

Fig.1 Diseases and insect pests of ginger leaf

1.2 数据集构建

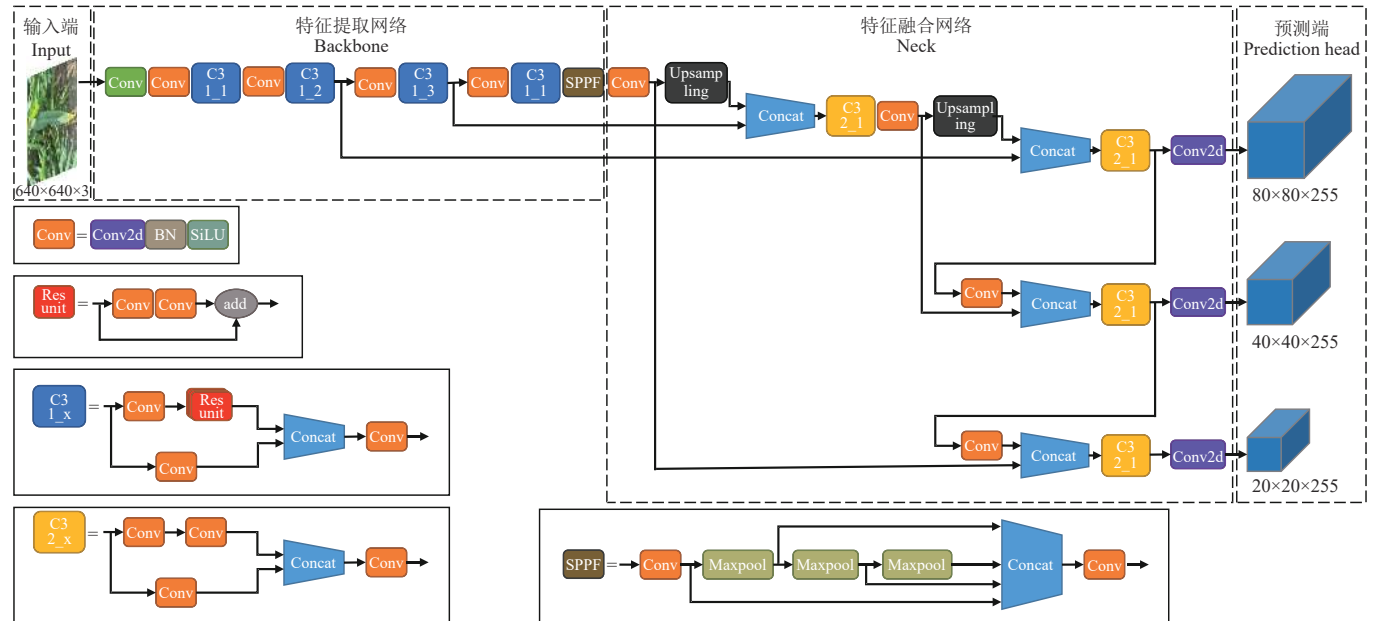
利用 YOLO 格式数据集在线标注工具 Make Sense 对数据集中的 3 848 张图像进行人工标注，标注时对两种病虫害进行分类标记，以获得图像中目标病虫害的类别和位置信息。标注完的信息以 txt 文件的形式保存，完成生姜病虫害数据集的构建。本研究按照 7:2:1 的比例将数据集中的图像随机划分为训练集、验证集和测试集，其中训练集 2 693 张，验证集 769 张，测试集 386 张。

2 生姜病虫害识别算法及改进

2.1 YOLOv5s 网络模型

YOLOv5 为一阶段目标检测算法，其性能良好，是 YOLO 算法中训练时间和推理速度最优的模型^[18]，其中，YOLOv5s 网络的深度和特征图的宽度较小^[19]，相较于其他版本推理速度快，最适合部署于嵌入式设备。

YOLOv5s 的网络结构分为 4 部分，即输入端 Input、特征提取网络 Backbone、特征融合网络 Neck 和预测端 Prediction Head，其网络结构如图 2 所示。



注：Conv 模块为 Conv2d+BN+SiLU，Conv2d 为卷积，BN 表示批量归一化，SiLU 为激活函数，Upsampling 为上采样操作，Res unit 为残差模块，C3_{1_x} 由 x 个残差组件集成的卷积结构，C3_{2_x} 由 2x 个 Conv 模块集成的卷积结构，add 为张量相加，Concat 为张量拼接，SPPF 为空间金字塔池化结构，Maxpool 为最大池化操作。

Note: Conv module is Conv2d+BN+SiLU, Conv2d is convolution, BN stands for batch normalization, SiLU is activation function, Upsampling is upsampling operation, Resunit is residual module, C3_{1_x} convolutional structure integrated by x residual components, C3_{2_x} convolutional structure integrated by 2x Conv modules, add is tensor addition, Concat is tensor concatenation, SPPF is spatial pyramid pooling structure, and Maxpool is a maximum pool operation.

图 2 YOLOv5 网络结构

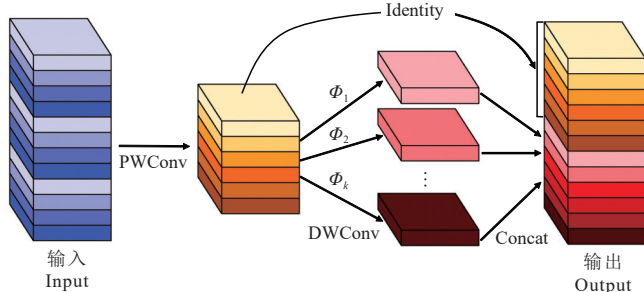
Fig.2 The network structure of YOLOv5

输入端的 Mosaic 数据增强随机选取 4 张图片, 通过随机裁剪、缩放、分布进行拼接, 以此丰富数据集, 从而增强模型的泛化性; 通过自适应锚框自动计算生姜病虫害数据集的最佳锚框位置。图片通过输入端送入神经网络后, 首先在 Backbone 部分进行特征提取, 获得 3 个有效特征层后, Neck 部分的特征金字塔网络通过自顶向下的上采样方式增强语义信息; Neck 部分的路径聚合网络采用自底向上的下采样方式加强定位信息, 两个网络分别从不同的路径进行特征融合, 从而获得含有丰富信息的特征图。Prediction Head 部分包含 3 个预测分支, 利用提取到的特征信息对不同尺寸的目标进行预测, 获得预测目标的类别、置信度及其位置信息。

2.2 YOLO5s 网络改进

2.2.1 引入 GhostNet 轻量级网络

在实际应用的自然场景下需要将模型部署在移动设备上, 移动设备相较于电脑端硬件计算力较低, 为解决硬件计算力低, 模型运行缓慢的问题, 本研究在 YOLOv5s 模型的基础上引入 GhostNet^[20] 网络, 将除第一个卷积外的普通卷积层替换为 Ghost 模块, 将原网络中的 Resunit 替换为 Ghost BottleNeck, 即原网络中的 Conv 模块和 C3 模块改为 GhostConv 和 C3Ghost, 实现模型的轻量化处理。GhostNet 提出了由 Ghost 模块叠加而成的 Ghost BottleNeck 结构, 在保持精度的同时, 减少模型的计算和储存成本。其中, Ghost 模块代替普通卷积^[21] 实现高效的网络设计, Ghost 模块是一种特征重用机制, 可在神经网络中引入更多的轻量级子网络, 这些子网络与主干网络通过并行连接, 使用共享卷积核来提高计算效率。Ghost 模块的操作流程分为三步: 首先, 选择通道数减半卷积核大小为 1×1 的 PWConv 卷积^[22] 对输入图像进行一次卷积操作进行特征整合, 生成输入特征层的浓缩特征图。然后, 对上一步输出的浓缩特征图使用 DWConv 卷积^[23] 进行廉价操作生成冗余特征图。最后, 将前两步生成的浓缩特征图和冗余特征图进行拼接操作生成输出特征图, 具体结构如图 3 所示。



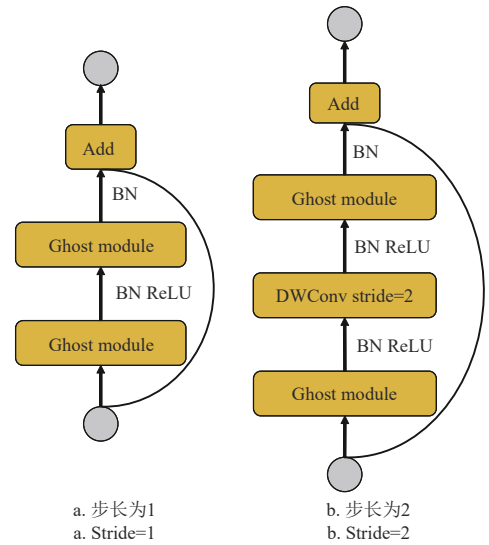
注: PWConv 表示点卷积; DWConv 表示深度卷积; Φ 表示廉价操作; Identity 表示本体特征图。
Note: PWConv stands for point-wise convolution; DWConv stands for depth-wise convolution; Φ represents the cheap operation; Identity represents the intrinsic feature map.

图 3 Ghost 模块

Fig.3 The Ghost module

利用 Ghost 模块的性能优势, 对两个 Ghost 模块进行叠加构建 Ghost BottleNeck 结构。如图 4 所示, 主干

由两个 Ghost 模块串联而成, 其中第一个 Ghost 模块的作用为增加特征维度, 扩张通道数, 第二个 Ghost 模块将通道数降低至与输入通道数一致, 然后通过 shortcut 和输入进行连接得到最终的输出, 这样 Ghost BottleNeck 结构的输入输出维度相同, 可以方便的嵌入神经网络当中; 当步长为 2 时, 在主干通路的两个 Ghost 模块之间添加一个步长为 2 的 DWConv 卷积层, 这样可使输出特征图大小降为输入特征图的 $1/2$ 。Ghost BottleNeck 结构中 Ghost 模块输出每层之后都应用了 BN^[24] 操作, 除最后一个 Ghost 模块之后不使用 ReLU^[25] 函数, 其他层都使用了 ReLU 函数。



注: ReLU 为激活函数。
Note: ReLU is the activation function.

图 4 Ghost BottleNeck 结构

Fig.4 The Ghost BottleNeck structure

2.2.2 添加 CA 注意力机制

自然场景下的生姜叶片生长密集, 并且病虫害特征较小, 为使网络模型在识别检测时能够更准确的定位病虫害位置信息及特征, 本文选择将 CA^[26] 注意力机制模块添加在 Neck 网络中的每个 C3 2_1 模块后。CA 注意力机制模块可增强移动网络学习特征的表达能力, 对网络中的任意中间特征张量进行转化后输出同样尺寸的张量。实现过程如图 5 所示。

CA 注意力机制为获取图像高度和宽度上的注意力并对精确位置编码, 将输入特征图在高度和宽度两个方向分别进行全局平均池化, 获得高度和宽度两个方向的特征图。首先使用尺寸为 $(H, 1)$ 和 $(1, W)$ 的池化核 (pooling kernel) 分别沿水平方向和垂直方向对每个通道进行编码, 如式 (1) (2) 所示:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} |x_c(h, i)| \quad (1)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j \leq H} |x_c(j, w)| \quad (2)$$

式中 $z_c^h(h)$ 、 $z_c^w(w)$ 分别为第 c 个通道输出的高度特征图和宽度特征图, $x_c(h, j)$ 和 $x_c(j, w)$ 分别表示输入特征图的第

c 通道沿 H 方向和 W 方向的输入。在获得两个方向的特征图后对其进行拼接操作，并将拼接后的特征图送入卷积核大小为 1×1 的共享卷积层中，使其维度降为原来的 C/r ，然后特征图经过批量归一化处理，再利用非线性激活函数得到形为 $C/r \times 1 \times (W+H)$ 的空间信息在水平和垂直两个方向的中间特征图 f 。计算式如下：

$$f = \delta(F_1([z^h, z^w])) \quad (3)$$

式中 δ 为非线性激活函数， F_1 为卷积变换函数，“ $;$ ”表示沿空间维度的拼接操作。随后，将特征图 f 沿空间维度切分为两个单独张量 f^h 和 f^w ，分别利用两个卷积核为 1×1 卷积使 f^h 和 f^w 的通道数与输入图像相同，后经 Sigmoid 激活函数分别得到特征图在高度方向的注意力权重 g^h 和宽度方向的注意力权重 g^w ，如式 (4) (5) 所示：

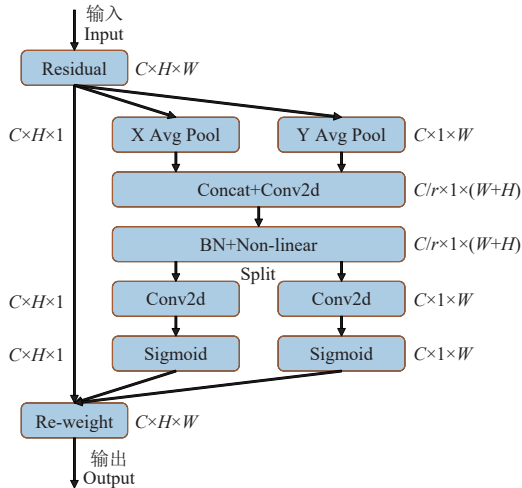
$$g^h = \sigma(F_h(f^h)) \quad (4)$$

$$g^w = \sigma(F_w(f^w)) \quad (5)$$

式中 σ 为 Sigmoid 激活函数， F_h 和 F_w 为卷积操作。为了降低模型复杂度和较少计算量，在此通常适当缩减通道数，然后对输出进行扩展得到 g^h 和 g^w 。最终，使用注意力权重 g^h 和 g^w 与输入图像进行乘法加权计算，得到在宽度方向和高度方向带有注意力权重的特征图，计算式如式 (6)：

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

式中 y_c 和 x_c 分别为第 c 个通道的输入和输出， $g_c^h(j)$ 和 $g_c^w(j)$ 分别表示第 c 通道中沿 H 方向和 W 方向的注意力权重。



注：C 为图像通道数，H 为图像高度，W 为图像宽度，r 为下采样缩减比，X Avg Pool 为水平方向平均池化，Y Avg Pool 为垂直方向平均池化，Non-linear 为非线性激活函数，Sigmoid 为激活函数，Re-weight 为权重获取。
Note: C is the number of image channels, H is the image height, and W is the image width, r is the down-sampling reduction ratio, X Avg Pool is the horizontal average pooling, Y Avg Pool is the vertical average pooling, Non-linear is the nonlinear activation function, Sigmoid is the activation function, Re-weight is the weight acquisition.

图 5 CA 注意力机制模块示意图

Fig.5 Schematic diagram of the CA attention mechanism module

2.3 训练环境与方法

本试验使用的计算机平台主要参数为：Intel Core i9-

13900k CPU，主频 3.0 GHz，运行内存 128GB，Nvidia GeForce RTX 4090 24GB GPU。试验在 Windows 操作系统上进行，采用 PyTorch 深度学习框架对模型进行搭建、训练以及评估。

训练参数设置：图像输入尺寸为 640×640 ，批次设置为 16，多线程设置为 12，选用随机梯度下降法优化器，训练轮次为 500，初始学习率为 0.01，动量参数和权重衰减参数分别设置为 0.937 和 0.0005。

2.4 评价指标

试验结果采用准确率 P (precision, %)、召回率 R (recall, %)、平均精确度 AP (average precision, %)、平均精度均值 mAP (mean Average Precision, %) 来衡量模型识别检测的准确度；采用帧速率 FPS (frames per second, 帧/s) 作为模型识别速度的评价指标。其中准确率 P 表示预测正确的样本数占预测为正样本数的比例，如式 (7) 所示：

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (7)$$

召回率 R 表示预测正确的样本数占实际为正样本数的比例，如式 (8) 所示：

$$R = \frac{T_p}{T_p + F_n} \times 100\% \quad (8)$$

平均精度均值 mAP 为平均精确度 AP 的均值，平均精确度 AP 即为 P - R 曲线的面积，如式 (9) 所示：

$$mAP = \frac{\sum_{i=1}^N \int_0^1 P(R) dR}{N} \times 100\% \quad (9)$$

式中 T_p 表示图片中正确检测叶片病虫害的样本数； F_p 表示图片中被误测为叶片病虫害的样本数； F_n 表示图片中未被正确检测的叶片病虫害样本数； N 表示类别数量，本文中研究的生姜叶片病虫害为两种，故 $N=2$ 。

3 结果与分析

3.1 Ginger-YOLOv5s 消融试验性能对比

本研究在测试集上通过消融试验以验证本文基于 YOLOv5s 对模型轻量化优化策略的可行性。试验结果如表 1 所示，由表可知，使用 GhostNet 网络中的 Ghost 模块和 Ghost BottleNeck 分别替换原网络中的卷积模块和 Resunit 模块，可以在 mAP 和精确度略微下降的情况下，大幅减少模型的参数量和计算量以及权重文件的大小，并使召回率提升，分析原因在于 GhostNet 中的轻量化模块使用了 DWConv 卷积，相较于普通卷积计算量降低，同时增加网络宽度和深度，实现了网络轻量化，保证了识别精度；从试验中还可看出，在网络中添加 CA 注意力机制模块后，相比于原 YOLOv5s 网络能够明显提升模型的精确度、召回率和平均精度均值，分别提高 2.3、2.7 和 1.6 个百分点，此时网络结构相交原始网络更复杂，参数量、计算量和权重文件内存占用量也随之增加。

表 1 YOLOv5s 消融试验性能对比

Table 1 Performance comparison of YOLOv5s ablation tests							
Ghost 网络 GhostNet network	CA 注意力机制 Coordinate attention	参数量 Parameters/ ×10 ⁶ M	计算量 Floating point operations FLOPs/G	权重大小 Model size/ MB	精确度 Precisi P/%	召回率 Recall R/%	平均精度均值 Mean average precision mAP/%
—	—	7.24	16.6	14.1	79.2	81.3	82.3
—	√	7.28	16.7	14.2	81.5	84.0	83.9
√	—	3.67	8.2	7.62	77.6	81.2	81.6
√	√	3.76	8.4	7.79	80.5	83.6	83.8

注：“—”表示不使用此模块；“√”表示使用此模块。
Note: “—” indicates that this module is not used; “√” indicates the use of this module.

综合消融试验，在平衡模型大小和识别精度的前提下，本研究基于 YOLOv5s 利用 GhostNet 网络和 CA 注意力机制模块改进的 Ginger-YOLOv5s 模型，在生姜病虫害的检测识别方面，参数量（3.76×10⁶M）、计算量（8.4GFLOPs）和权重文件大小（7.79MB）分别为原模型的 52.0%、50.6% 和 55.2%，并且精确率、召回率和平均精度均值均有提升，相比原 YOLOv5 模型分别提升 1.3、2.3 和 1.5 个百分点，可在满足模型轻量化的同时提高模型对生姜病虫害的识别精度。

3.2 不同目标检测模型性能对比

将基于改进 YOLOv5s 的 Ginger-YOLOv5s 模型与目前主流的目标检测模型 Faster-RCNN、SSD、YOLOv4、YOLOv5s 以及王梦妮等^[27]提出的 Tea-YOLOv5s 进行对比，结果如表 2 所示。Ginger-YOLOv5s 模型的平均精度均值比其他 5 种模型分别高出 37.6、39.1、22.5、1.5、0.7 个百分点，帧速率相比 Faster-RCNN、SSD、YOLOv4、Tea-YOLOv5s 分别提升 56.7、40.1、25.7、11.1 帧，单阶段检测网络模型 SSD 模型识别精度最低，双阶段检测模型 Faster-RCNN 推理速度最慢，YOLOv4 的平均精度均值和推理速度较于 Faster-RCNN、SSD 两者中的最大值皆有所提高，但仍无法满足生姜病虫害的实时检测。改进后的 Ginger-YOLOv5s 模型的精确度、召回率和平

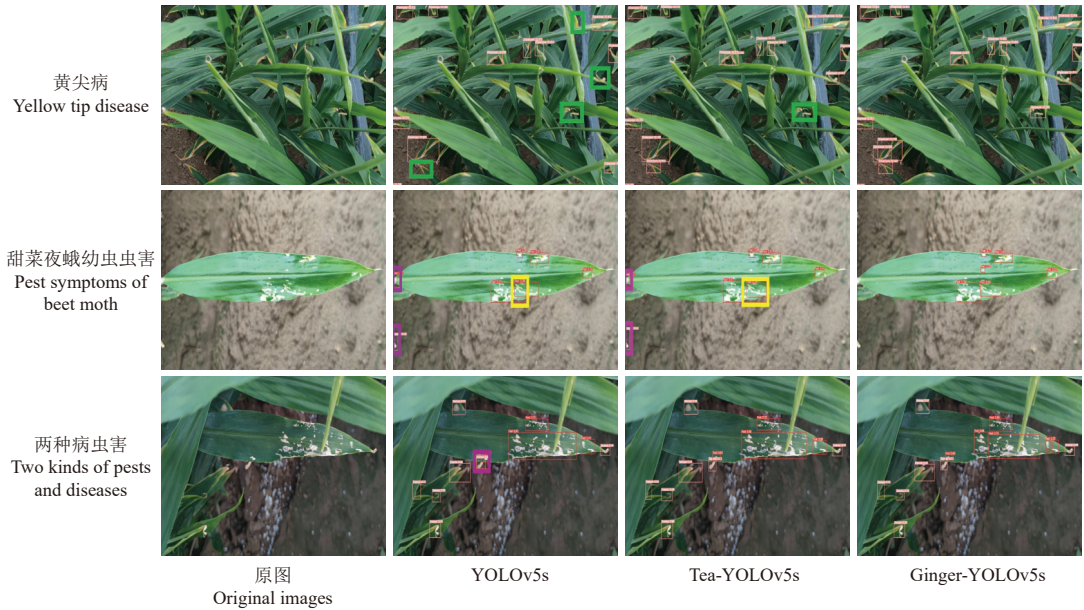
均精度均值分别提高到了 80.5%、83.6% 和 83.8%，虽然推理速度低于原 YOLOv5s 模型，但仍满足对于生姜叶片病虫害实时检测的要求。

表 2 不同模型识别效果对比

Table 2 Comparison of recognition effect of different models							
模型 Models	参数量 Parameters/ ×10 ⁶ M	计算量 FLOPs/ G	权重大小 Model size /MB	P/%	R/%	mAP/ %	帧速率 Frames per second FPS/(帧·s ⁻¹)
Faster-RCNN	137.10	370.2	111.5	45.4	45.8	46.2	19.8
SSD	26.29	62.8	93.3	43.1	44.2	44.7	36.4
YOLOv4	64.36	60.5	251.8	60.8	61.5	61.3	50.8
YOLOv5s	7.24	16.6	14.1	79.2	81.3	82.3	90.6
Tea-YOLOv5s ^[27]	7.33	16.9	14.7	80.4	82.3	83.1	65.4
本文 This paper	3.76	8.4	7.79	80.5	83.6	83.8	76.5

3.3 Jetson Orin NX 部署试验分析

由表 2 可知 Faster-RCNN、SSD 以及 YOLOv4 模型的 mAP 值较低，不满足生姜叶片病虫害识别的要求，因此不再对其进行部署验证。为测试 Ginger-YOLOv5s 模型移植到试验样机实时检测生姜叶片病虫害的效果，将 Ginger-YOLOv5s 与原 YOLOv5s 模型分别部署在 Jetson Orin NX 开发板上进行对比，两种模型检测 3 处相同自然场景下的生姜叶片病虫害，如图 6 所示。



注：图中绿色粗体矩形为漏检标记，玫红色粗体矩形为误检标记，黄色粗体矩形为定位误差标记。
Note: In the figure, the green bold rectangle is the missed detection mark, the rose red bold rectangle is the false detection mark, the yellow bold rectangle is the positioning error mark.

图 6 YOLOv5s 改进前后的生姜病虫害检测效果对比
Fig.6 Comparison of the detection effect of ginger diseases and pests before and after YOLOv5s improvement

由测试图像可知, 生姜叶片病虫害目标较小, 且生姜叶片生长茂密。YOLOv5s 和 Tea-YOLOv5s 模型在复杂场景下检测黄尖病时存在漏检的问题, 对于虫害的检测时出现对图像背景误检和定位不精准的情况 (图中已标记), 并且其置信度较 Ginger-YOLOv5s 模型略低。经过综合比较, Ginger-YOLOv5s 模型对生姜病虫害图像进行检测时拥有更高的置信度, 且目标定位更加准确, 模型鲁棒性更强, 具有更好的检测效果。同时在模型部署时对其使用 C++ 重写检测代码, 利用 TensorRT 高性能算子对模型加速, 并进行 Int8 量化处理, 再使用 CUDA 改写预处理部分进而改为多线程处理, 避免模型由电脑端移植到移动端因算力降低而出现检测速率大幅下降的情况。经过加速处理后, Ginger-YOLOv5 模型在 Jetson Orin NX 开发板上的检测速度达到 74.3 帧/s, 依旧满足实时检测的效果。

4 结 论

1) 本研究提出一种基于 YOLOv5s 改进的目标检测模型, 使用 GhostNet 网络中的 Ghost 模块和 Ghost BottleNeck 结构替换了原模型中的卷积层和 C3 模块, 并且引入 CA 注意力机制模块, 在减少模型参数数量和计算量的同时, 降低了模型权重文件内存占用量, 实现了模型的轻量化, 并且精确率、召回率和平均精度均值达到了 80.5%、83.6% 和 83.8%, 相比原模型分别提高 1.3、2.3 和 1.5 个百分点, 说明改进后的模型在实现模型轻量化的同时仍具有较高的检测精度。

2) 改进后的轻量化模型与其他主流目标模型进行试验对比, 结果表明, 本研究提出的 Ginger-YOLOv5s 模型相较于 Faster-RCNN、SSD、YOLOv4、Tea-YOLOv5s 模型平均精度均值分别提高 37.6、39.1、22.5、0.7 个百分点; 帧速率分别提升 56.7、40.1、25.7、11.1 帧, 并通过与原 YOLOv5s 在自然场景下的识别效果比较, 分析得出改进后的模型在复杂情况下对生姜病叶片虫害的检测优势明显。

3) Ginger-YOLOv5s 模型在自然场景下识别精度、鲁棒性均有提升, 并且在移动端部署时对其进行加速处理, 经验证平均检测帧率为 74.3 帧/s, 解决了算力降低检测速度变慢的问题, 可以满足智能生姜施药车实时识别检测要求, 保证了机器的作业效率, 为模型在生姜施药车上的迁移部署提供技术支持。

【参 考 文 献】

- [1] 胡臻钰. 中国生姜栽培史考述[J]. 蔬菜, 2023(5): 40-47.
- [2] 王宁, 李维华, 焦伟, 等. 我国生姜全程机械化技术装备研究现状及发展趋势[J]. 农业装备与车辆工程, 2023, 61(5): 47-51, 74.
- [3] NAIR K P. Turmeric (*Curcuma Longa* L.) and Ginger (*Zingiber Officinale* Rosc.)-World's Invaluable Medicinal Spices: The Agronomy and Economy of Turmeric and Ginger[M]. Springer International Publishing, 2019.
- [4] 邹修国. 基于计算机视觉的农作物病虫害识别研究现状[J]. 计算机系统应用, 2011, 20(6): 238-242.
- [5] 陈兵旗, 吴召恒, 李红业, 等. 机器视觉技术的农业应用研究进展[J]. 科技导报, 2018, 36(11): 54-65.
- [6] RANGARAJAN A K, PURUSHOTHAMAN R, RAMESH A. Tomato crop disease classification using pre-trained deep learning algorithm[J]. *Procedia Computer Science*, 2018, 133: 1040-1047.
- [7] 蒋丰千, 李旻, 余大为, 等. 基于 Caffe 的生姜病害识别系统研究与设计[J]. 中国农机化学报, 2019, 40(1): 126-131.
- [8] 白雪松, 吴建平, 景文超, 等. 基于卷积神经网络的农作物病虫害检测研究[J]. 计算机技术与发展, 2022, 32(12): 200-205.
- [9] 李子茂, 刘恋冬, 夏梦, 等. 基于深度学习的月季多叶片病虫害检测研究[J]. 中国农机化学报, 2021, 42(8): 169-176.
- [10] 樊湘鹏, 许燕, 周建平, 等. 基于迁移学习和改进 CNN 的葡萄叶部病害检测系统[J]. 农业工程学报, 2021, 37(6): 151-159.
- [11] 王卫星, 刘泽乾, 高鹏, 等. 基于改进 YOLO v4 的荔枝病虫害检测模型[J]. 农业机械学报, 2023, 54(5): 227-235.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only lookonce: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 779-788.
- [13] ALEXEY B, WANG C, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2023-10-18]. <https://arxiv.org/abs/2004.10934>.
- [14] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hawaii: IEEE, 2017: 7263-7271.
- [15] REDMON J, FARHADI A. YOLOv3: An incremental improvement[EB/OL]. (2018-04-08) [2023-10-18]. <https://arxiv.org/abs/1804.02767>.
- [16] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, Santiago: IEEE, 2015: 1440-1448.
- [17] LIU W, DRAGOMIR A, DUMITRU E, et al. SSD: Single shot multibox detector[C]//European Conference Computer Vision, Amsterdam: Springer, 2016: 20-37.

- [18] 白强, 高荣华, 赵春江, 等. 基于改进 YOLOV5s 网络的奶牛多尺度行为识别方法[J]. *农业工程学报*, 2022, 38(12): 163-172. BAI Qiang, GAO Ronghua, ZHAO Chunjiang, et al. Multi-scale behavior recognition method for dairy cows based on improved YOLOV5s network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(12): 163-172. (in Chinese with English abstract)
- [19] 尚钰莹, 张倩如, 宋怀波. 基于 YOLOv5s 的深度学习在自然场景苹果花朵检测中的应用[J]. *农业工程学报*, 2022, 38(9): 222-229. SHANG Yuying, ZHANG Qianru, SONG Huaibo. The application of deep learning based on YOLOv5s in the detection of apple flowers in natural scenes[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(9): 222-229. (in Chinese with English abstract)
- [20] HAN K, WANG Y, TIAN Q, et al. GhostNet: More features from cheap operations[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle: IEEE, 2020: 1580-1589.
- [21] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]// European Conference Computer Vision, Zurich: Springer, 2014: 818-833.
- [22] ZHAO H, ZHANG Y, LIU S, et al. Psanet: Point-wise spatial attention network for scene parsing[C]// Proceedings of the European Conference on Computer Vision, Munich: Springer, 2018: 267-283.
- [23] LI G, YUN I, KIM J, et al. Dabnet: Depth-wise asymmetric bottleneck for real-time semantic segmentation[EB/OL]. (2019-10-01) [2023-10-18]. <https://arxiv.org/abs/1907.11357>.
- [24] SANTURKAR S, TSIPRAS D, ILYAS A, et al. How does batch normalization help optimization? [EB/OL]. (2018-05-29) [2023-10-18]. <https://arxiv.org/abs/1805.11604>.
- [25] AGARAP A F. Deep learning using rectified linear units (ReLU) [EB/OL]. (2019-02-07) [2023-10-18]. <https://arxiv.org/abs/1803.08375>.
- [26] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Virtual: IEEE, 2021: 13713-13722.
- [27] 王梦妮, 顾寄南, 王化佳, 等. 基于改进 YOLOv5s 模型的茶叶嫩芽识别方法[J]. *农业工程学报*, 2023, 39(12): 150-157. WANG Mengni, GU Jinan, WANG Huajia, et al. Method for identifying tea buds based on improved YOLOv5s model[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(12): 150-157. (in Chinese with English abstract)

Identifying diseases and pests in ginger leaf under natural scenes using improved YOLOv5s

LAN Yubin^{1,2,3}, SUN Binshu^{1,3}, ZHANG Lechun^{1,3}, ZHAO Denan^{1,3}

(1. College of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo 255049, China;

2. Shandong Provincial Engineering Technology Research Center for Agricultural Aviation Intelligent Equipment, Zibo 255049, China;

3. Research Institute of Ecological Unmanned Farm, Shandong University of Technology, Zibo 255049, China)

Abstract: Ginger diseases and pests have posed a serious threat to the yield in recent years. However, the artificial and mechanical application cannot fully meet the large-scale production at present, due to the slow overall progress and the low degree low of intelligence. This study aims to develop the intelligent application equipment of ginger for the high efficiency and accuracy of the intelligent vehicle. A lightweight model was proposed to realize the high-performance deployment of the ginger leaf disease and pest detection on mobile terminals using improved YOLOv5s. Ghost module of GhostNet was selected to replace the convolutional layers in the original YOLOv5s neural network, except the first layer. Ghost BottleNeck was used to replace the Resunit residual component in the original C3 concentrate-comprehensive convolution block. The lightweight of the network model was obtained to reduce the number of parameters and the amount of calculation. At the same time, the memory consumption was reduced in the model weight file. CA attention mechanism module was added after the C3 block in the feature fusion network, in order to improve the recognition and positioning accuracy. The reason was that the lightweight of the model caused the feature loss, when the neural network was used to extract the features of the image. The experimental results show that the number of parameters of the improved YOLOv5s model was 3.76×10^6 M, which was 52.0% of the original. The computational complexity was 8.4GFLOPs, which was 50.6% of the original. The size of the weight file was 7.79MB, which was 55.2% of the original. The average precision and average precision reached 80.5% and 83.8%, respectively, which were 1.3 and 1.5 percentage points higher than those of the original model. The improved model was greatly reduced the number of parameters, calculation amount and weight file size for the high detection accuracy, compared with Fast-RCNN, SSD, YOLOv4, YOLOv5s and Tea-YOLOv5s target detection models. The missed and false detection of image targets were also reduced, compared with the YOLOv5s model. And the improved network model was required less hardware conditions. The performance of the improved model was verified on mobile terminals. The Ginger-YOLOv5s model was deployed on the Jetson Orin NX development board, where the detection code was rewritten in C++. The model was accelerated using TensorRT high-performance operator, Int8 quantization processing, CUDA rewriting preprocessing and multi-thread processing. The final FPS reached 74.3, which was fully met the requirements of operation efficiency in the application machinery for the real-time detection of ginger leaf diseases and pests. The finding can provide the technical support for the migration and deployment of the model on the ginger application vehicle.

Keywords: image recognition; deep learning; object detection; lightweight; diseases and pests; ginger leaves