

施 杰,熊凯祥,李 志,等. 基于轻量化改进 YOLOv8 模型和边缘计算的玉米病虫害检测系统[J]. 江苏农业学报, 2025, 41(2): 313-322.

doi: 10.3969/j.issn.1000-4440.2025.02.012

基于轻量化改进 YOLOv8 模型和边缘计算的玉米病虫害检测系统

施 杰^{1,2}, 熊凯祥¹, 李 志¹, 陈立畅¹, 唐秀英¹, 杨琳琳^{1,2}

(1. 云南农业大学机电工程学院, 云南 昆明 650201; 2. 云南省作物生产与智慧农业重点实验室, 云南 昆明 650201)

摘要: 为实现玉米病虫害的原位准确检测与识别, 本研究设计了一套基于边缘计算的玉米病虫害智能检测系统。该系统基于 YOLOv8 模型并进行改进, 具体改进方法包括: 采用高效视觉网络(EfficientViT)作为主干网络, 以降低计算量; 在特征融合网络中引入幻影卷积(GhostConv), 进一步减轻计算负担; 在 C2f 模块中引入空间通道重建卷积(SCConv), 以增强特征提取性能; 将损失函数替换为具有动态非单调聚焦机制的损失函数(WIoU), 以提高模型的识别精度。同时, 本研究设计了基于边缘计算的病虫害检测系统上位机、下位机架构, 并将该轻量化模型部署到 Jetson orin nano 边缘计算设备上。系统采用 Pyside6 开发系统可视化界面, 除具备识别与训练功能外, 还集成了基于大模型技术的 AI 专家库, 可以实现对病虫害的智能化诊断。通过自建的玉米病虫害数据集对改进模型 YOLOv8-EGCW 进行检验。结果表明, 与原始模型 YOLOv8m 相比, 改进模型 YOLOv8-EGCW 的精确度、召回率和平均精度均值分别提升了 0.4 个百分点、1.6 个百分点和 1.2 个百分点, 参数量和模型大小大幅减少, 单张图像检测时间缩短。建立的玉米病虫害检测系统测试结果显示, 准确率达到 93.4%, 检测速度达 1 s 25 帧。表明该系统能够满足边缘计算环境下玉米病虫害原位检测的需求。

关键字: 玉米; 病虫害检测系统; YOLOv8 模型; 轻量化改进; 边缘计算

中图分类号: S435.13 **文献标识码:** A **文章编号:** 1000-4440(2025)02-0313-10

Maize pest and disease detection system based on lightweight improved YOLOv8 model and edge computing

SHI Jie^{1,2}, XIONG Kaixiang¹, LI Zhi¹, CHEN Lichang¹, TANG Xiuying¹, YANG Linlin^{1,2}

(1. Faculty of Mechanical and Electrical Engineering, Yunnan Agricultural University, Kunming 650201, China; 2. The Key Laboratory for Crop Production and Smart Agriculture of Yunnan Province, Kunming 650201, China)

Abstract: In order to achieve in-situ accurate detection and identification of maize pests and diseases, this study designed an intelligent detection system for maize pests and diseases based on edge computing. The system was improved based

收稿日期: 2024-10-30

基金项目: 云南省重大科技专项(202302AE090020); 云南省农业基础研究联合专项(202401BD070001-069); 云南省作物生产与智慧农业重点实验室开放课题

作者简介: 施 杰(1981-), 男, 云南昆明人, 博士, 副教授, 硕士研究生导师, 主要从事智慧农业、智能农业装备研究。(E-mail) km_shijie@126.com

通讯作者: 杨琳琳, (E-mail) 29545343@qq.com

on the YOLOv8 model with specific improvement methods, including adopting the Efficient Vision Transformer (EfficientViT) as the backbone network to reduce the computational load, introducing Ghost Convolution (GhostConv) into the feature fusion network to further reduce the computational burden, introducing Spatial-Channel Convolution (SCConv) into the C2f module to enhance the feature extraction performance, and replacing the loss function with

the Wise Intersection over Union (WIoU) loss function that had a dynamic non-monotonic focusing mechanism to improve the recognition accuracy of the model. At the same time, this study designed the upper and lower computer architectures of the pest and disease detection system based on edge computing and deployed this lightweight model to the Jetson Orin Nano edge computing device. The system used Pyside6 to develop a visual interface. In addition to the recognition and training functions, it also integrated an AI expert library based on large-model technology, which could realize intelligent diagnosis of pests and diseases. The improved model YOLOv8-EGCW was tested using a self-built maize pest and disease dataset. The results showed that compared with the original YOLOv8m model, the precision, recall rate, and mean average precision of the improved model YOLOv8-EGCW increased by 0.4 percentage points, 1.6 percentage points, and 1.2 percentage points, respectively. The number of parameters and the model size were greatly reduced, and the detection time for a single image was shortened. The test results of the established corn pest and disease detection system indicated that the accuracy rate reached 93.4% and the detection speed reached 25 frames per second. These results indicated that the system could meet the requirements of in-situ detection of maize pests and diseases in the edge computing environment.

Key words: corn; pest and disease detection system; YOLOv8 model; lightweight improved model; edge computing

2023 年,中国玉米总产量达到 $2.888\ 4\times 10^8\ \text{t}^{[1]}$,玉米不仅是主要粮食作物,也是重要的工业原料和饲料来源。然而,玉米在生长过程中易遭受病虫害侵袭,病虫害严重影响其产量和品质。因此,开发一套玉米病虫害检测系统对于保障粮食安全、提高粮食产量具有重要意义。

传统病虫害检测主要依靠人工判断和图像处理技术,如 Dong 等^[2]提取形态学、颜色和纹理特征,结合最近邻距离方法识别黄瓜叶片病害,识别准确率超过 96%;戴宁^[3]利用基于 K 均值聚类的图像分割技术提取虫害形状特征,识别准确率达到 95.7%;Jothiaruna 等^[4]提出了一种基于高级综合颜色特征 (ACCF) 和区域生长方法的病害斑点分割算法,在处理非均匀光照和复杂背景条件下的叶片图像时,分割准确率可达 87%;蒋龙泉等^[5]通过融合叶片图像的颜色、纹理等特征,并利用支持向量机 (SVM) 对特征进行分类,识别准确率超过 90%。尽管图像处理能够较为准确地识别出病虫害,但其泛化能力弱,特征提取能力有限,难以满足实际生产需求。

随着深度学习技术的发展,其强大的特征提取能力有效解决了传统技术在特征选择上的局限性。Wang 等^[6]提出了一种基于聚合深度残差网络的玉米病害检测方法,通过调整卷积核规模提高检测精度,在实际检测中准确率达到 90.22%;陈啟昇等^[7]设计了基于 ResNet 模型的植物病虫害智能检测系统,该系统采用服务器模式运行,通过 web 端采集图像,由服务器端完成分析处理;Bi 等^[8]提出了一种基于 MobileNet 的玉米叶片病害检测模型,该模型采

用了高效通道注意力模块 (EAC),同时引入扩张卷积,识别精度可达 98.2%;徐会杰等^[9]基于 YOLOv3 模型,将输出特征图像与下采样特征图像拼接,并利用多尺度检测层对玉米病虫害进行检测和识别;李硕^[10]开发了一种基于 RFBNet 模型的桃树病虫害检测系统,并将训练后的系统部署至 Android 移动设备上进行检测。在病虫害识别中,深度学习可以自动提取特征,即使在复杂环境下也能表现出较强的鲁棒性,但其计算量较大,对硬件资源要求较高。

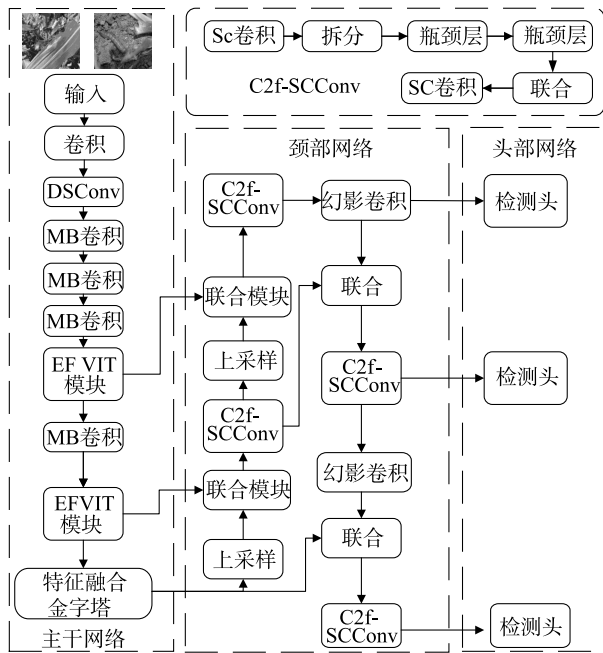
随着边缘计算技术的发展,将病虫害检测模型部署到边缘计算设备上,能够更加便捷高效地识别病虫害。如 Chen 等^[11]在 Nvidia Jetson TX2 上部署了基于 YOLOv3-tiny 的果园虫害识别模型,通过无人机进行虫害识别并上传位置信息,用于记录果树生长状况;刘忠等^[12]在 Nvidia Jetson Xavier NX 上部署了基于改进 YOLOv8 的荷叶病虫害检测模型,其推理速度达到 1 s 27 帧。将计算和数据处理从中央服务器转移到边缘设备上,不仅提高了响应速度和效率,还降低了传输成本和能耗,更符合实际生产需求。针对目前玉米病虫害检测模型参数量大、配套系统不够完善、边缘计算应用不足等问题,本研究拟改进 YOLOv8 模型,并结合边缘计算技术构建一种玉米病虫害检测系统,旨在实现对玉米病虫害的原位高效检测。

1 材料与方法

1.1 YOLOv8 模型的改进策略

YOLOv8 是一种单阶段目标检测模型,作为 YOLO 系列模型的主流版本,被广泛应用于农业目

标检测中。该模型主要由主干网络、颈部网络和头部网络 3 部分组成,其核心思想是将输入图像划分为多个网格,并在每个网格中预测物体的边界框和类别信息。YOLOv8 采用了先进的网络结构,能够在短时间内处理大规模图像数据,但其对硬件性能的要求较高。为使 YOLOv8 能够流畅地部署在边缘计算设备上,本研究对 YOLOv8 模型进行了轻量化改进,具体改进方法包括:使用更高效视觉网络(EfficientViT)^[13]替换原主干网络;将颈部网络的标准卷积替换为轻量化的幻影卷积^[14];在跨阶段特征融合模块(C2f)中融合空间通道重建卷积(SCConv)^[15]以增强检测效果;引入具有动态非单调聚焦机制的损失函数(WIOU)^[16]以优化边界框回归精度。将改进后的模型命名为 YOLOv8-EGCW,其整体结构如图 1 所示。



DSConv:深度可分离卷积;MB卷积:移动翻转瓶颈卷积;EF VIT 模块:轻量化高效视觉网络模块;C2f-SCConv:融合了空间通道重建卷积的跨阶段特征融合模块。

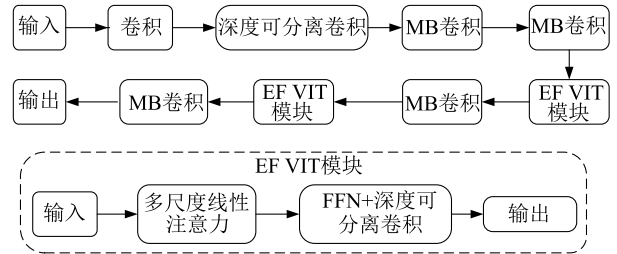
图1 YOLOv8-EGCW 模型结构

Fig.1 YOLOv8-EGCW model structure

1.1.1 主干网络的改进 YOLOv8 中的主干网络通过深层网络结构提升了图像特征提取能力,但也增加了计算复杂性,延长了模型的训练和推理时间。为了解决这一问题,本研究将 YOLOv8 的主干网络替换为轻量级的 EfficientViT 网络,以加快训练和推

理速度。EfficientViT 采用多尺度线性注意力机制取代了传统的 Softmax 注意力机制,同时引入深度可分离卷积 (Depthwise separable convolution, DSConv),显著减少了模型的参数量和计算复杂度。

图 2 为 EfficientViT 网络结构,首先通过多尺度线性注意力模块从输入信息中提取多尺度全局特征,并将不同尺度的特征进行融合。为了增强局部特征的提取能力,引入移动翻转瓶颈卷积 (Mobile inverted bottleneck convolution, MBConv)。MBConv 基于 DSConv 对输入通道进行扩宽和缩放,实现了对局部特征的高效提取。EfficientViT 实现了全局与局部特征的高效提取以及轻量化设计,在边缘计算设备等资源受限的平台上表现出了卓越的性能。



MB 卷积:移动翻转瓶颈卷积;EF VIT 模块:轻量化高效视觉网络模块;FFN:前馈神经网络。

图2 EfficientViT 网络结构

Fig.2 EfficientViT network structure

1.1.2 幻影卷积的引入 在颈部网络中,采用 GhostConv 替换原有的标准卷积,以实现颈部网络的轻量化处理。GhostConv 的核心思想是通过少量卷积核提取输入图像的特征,再对提取的特征图像进行简单的线性变化,最终通过特征拼接生成新的特征图像。该方法显著减少了对非关键特征的计算量,在降低计算量的同时还能够有效保持模型的性能,从而提升整体效率。幻影卷积计算公式如下:

$$Y' = Xf' \quad (1)$$

$$y_{i,j} = \varphi_{i,j}(y'_{i,j}) \quad (2)$$

式中, Y' 为原始特征值; X 为输入特征图像; f' 为卷积核; $y_{i,j}$ 为第 i 个原始特征图像生成的第 j 个特征图像; $\varphi_{i,j}$ 为第 i 个原始特征图像生成的第 j 个恒等隐射; $y'_{i,j}$ 为第 i 个通道特征图像。

1.1.3 C2f 模块的改进 在颈部网络设计中,将空间和通道重构卷积(SCConv)与跨阶段特征融合模块(C2f)融合成为 C2f-SCConv 模块。相较于传统的 C2f 模块,C2f-SCConv 模块具有更强的特征提取能

能力。SCConv 通过自矫正操作,自适应地调节特征空间与通道之间的相关性,使网络能够捕捉到更丰富的空间和通道信息,从而生成更具判别力的特征图像。SCConv 的内部操作为:将输入特征 X 通过空间重构单元(Spatial reconstruction unit, SRU)进行重构,得到空间细化特征 X^w ;然后利用通道重构单元(Channel reconstruction unit, CRU)对空间细化特征 X^w 进一步处理,得到通道细化特征 Y ,最后输出特征图像。通过 SRU 和 CRU, SCConv 可以降低特征的空间和通道计算量,提高特征的传递效率,同时增强模型对复杂数据结构的识别精度,使其能够灵活处理不同输入特征。虽然 C2f-SCConv 模块增加了空间重构单元与通道重构单元,但其计算复杂度较低,适合部署在计算资源有限的边缘计算设备上。

1.1.4 边框损失函数的优化 在目标检测任务中,通常使用交并比损失函数(IoU)来衡量预测框与真实框的重叠程度。然而, IoU 在处理不同尺度和长宽比的目标时存在一定局限性,尤其是在预测框与真实框不重叠的情况下,难以准确描述框与框之间的距离关系。具有动态非单调聚焦机制的损失函数(WIoU)作为一种新型损失函数,在计算过程中引入了中心点距离惩罚项,即通过计算预测框和真实框中心点之间的欧氏距离来优化损失函数。当预测框与真实框不重叠时,该惩罚项可以准确描述两者之间的相对位置,从而提升对非重叠框的优化效果。同时, WIoU 通过动态聚焦机制提升了模型的泛化能力,该机制考虑到训练数据中可能存在低质量标注,通过适当降低对这些样本的惩罚权重,减少其对模型训练的干扰,从而显著提升模型的泛化性能,计算公式如下:

$$L_{\text{Wiou}} = r \times R_{\text{Wiou}} \times L_{\text{IoU}} \quad (3)$$

$$\text{其中 } R_{\text{Wiou}} = \exp \left[\frac{(x - x_{\text{gt}})^2 + (y - y_{\text{gt}})^2}{(W_{\text{g}}^2 + H_{\text{g}}^2)} \right] \quad (4)$$

$$L_{\text{IoU}} = 1 - \text{IoU} \quad (5)$$

$$\beta = \frac{L_{\text{IoU}}^*}{L_{\text{IoU}}} \quad (6)$$

$$r = \frac{\beta}{\delta \alpha^{(\beta - \delta)}} \quad (7)$$

式中, r 为聚焦调整系数; α 和 δ 为超参数; β 为离群度; W_{g} 和 H_{g} 分别为最小封闭框的宽度和高度; x 、 y 为预测框的宽度和高度; x_{gt} 、 y_{gt} 为真实框的宽度和

高度; IoU 为预测框与真实框的交并比; L_{IoU}^* 为不参与反向传播的普通框损失值; \bar{L}_{IoU} 为增量的滑动平均值。

WIoU 能够显著提升目标定位的准确性。在玉米病虫害检测任务中,由于拍摄角度的差异,病虫害特征在大小和形态上会发生明显变化。引入 WIoU 能够有效增强模型对多尺度特征的检测能力,使其更精确地定位病虫害区域并优化边界框的回归精度,从而进一步提升模型的鲁棒性和检测性能。

1.2 模型训练平台构建

1.2.1 试验数据集 如图 3 所示,玉米的生长环境复杂多样,为确保病虫害检测模型检测结果的准确性,本研究综合了多种数据来源构建试验数据集,包括玉米病虫害的公开数据、实地拍摄数据和网络收集数据等。由于数据集来源广泛,图片的格式和分辨率存在较大差异。为保证试验的严谨性与规范性,对图片进行筛选,剔除重复和无用的图片,并对数据集进行了归一化处理。

使用 LabelImg 软件对数据集中的图片进行数据标注,标注信息包括锈病、黏虫、斑病和灰飞虱。标注结果以 VOC 格式的 XML 文件保存,并进一步转换为 YOLOv8 所需的 YOLO 格式 TXT 文件。每个标注文件中包含病虫害的类别信息及其在图像中的位置信息。如图 4 所示,针对采集到的图像质量参差不齐的问题,对标注好的图像进行了数据增强处理,具体包括变暗、镜像、添加高斯噪声等。数据增强不仅有效扩充了数据集的规模,还显著提升了模型的泛化能力。最后,将数据集按照 8 : 2 的数量比例划分为训练集和验证集。

1.2.2 训练环境与训练参数 硬件配置包括 Intel (R) I5 10400CPU 处理器, NVIDIA GeForce GTX1060 显卡(6 G 显存),操作系统为 Window10,软件环境为 Python 3.10.13, Torch 2.1.2, CUDA 12.3。根据文献[17]与试验测试,设置训练超参数(Batch size)为 16、初始化学率为 0.01、迭代次数为 100 次,并使用 GPU 进行训练。

为评价轻量化改进模型 YOLOv8-EGCW 的检测速度、准确性和鲁棒性,利用相同数据集分别对 YOLOv8-EGCW、YOLOv8、YOLOv5、SSD 和 Faster-RCNN 模型进行对比。选取平均精度均值(mAP)、精确度(P)、召回率(R)、参数量、模型大小与检测速度作为评价指标对模型进行性能评估。计算公式如下:

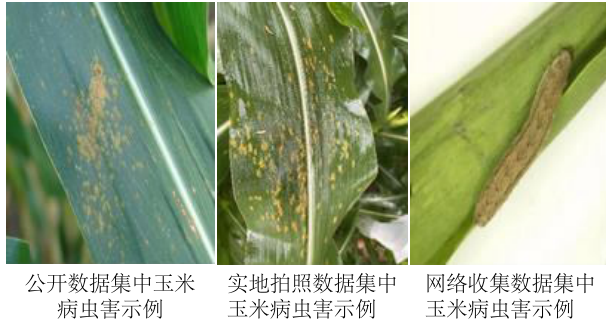


图3 试验数据集中图像示例

Fig.3 Images in the experimental dataset

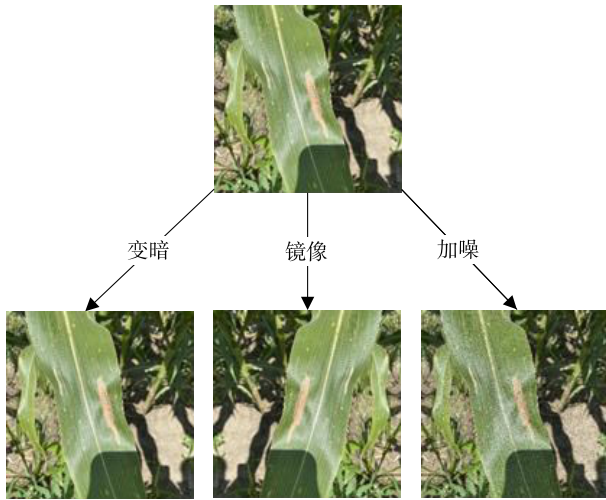


图4 数据集中数据增强

Fig.4 Data enhancement in dataset

$$P = \frac{TP}{TP+FP} \times 100\% \quad (8)$$

$$R = \frac{TP}{TP+FN} \times 100\% \quad (9)$$

$$mAP = \frac{TP \sum_{K=1}^K AP(K)}{K} \quad (10)$$

式中, TP 为检测正确的样本数; FP 为检测错误的样本数; FN 为检测遗漏的样本数; K 为检测类别数; AP 为平均精度。

2 结果与分析

2.1 各模块对比

为验证各模块对模型性能的有效性,以原始模型 YOLOv8m 作为基线模型,引入改进模块,并通过消融试验评估各模块的贡献。采用精确度、召回率、平均精度均值、模型大小作为评价指标,结果如表 1

和图 5 所示。消融试验结果表明,将 YOLOv8 原始主干网络替换为 EfficientViT 后,模型的总体精度略有下降,但模型大小减少了 36.8%,轻量化效果显著。在此基础上,在颈部网络中将普通卷积替换为幻影卷积(GhostConv),进一步减小了模型体积。在颈部网络中加入 C2f-SCConv 模块后,模型 P 值和 mAP 值分别提高了 0.1 个百分点和 0.2 个百分点。WIoU 通过引入动态聚焦机制和中心点距离惩罚,优化了边界框回归的准确性,模型 mAP 值和 R 值分别提高了 0.3 个百分点和 0.7 个百分点。综上,通过引入 EfficientViT、GhostConv、C2f-SCConv 和 WIoU 损失函数,在实现轻量化的同时,显著提升了模型精确度(P 值)、召回率(R 值)等多项性能指标。

表 1 各模块对比

Table 1 Comparison of each module

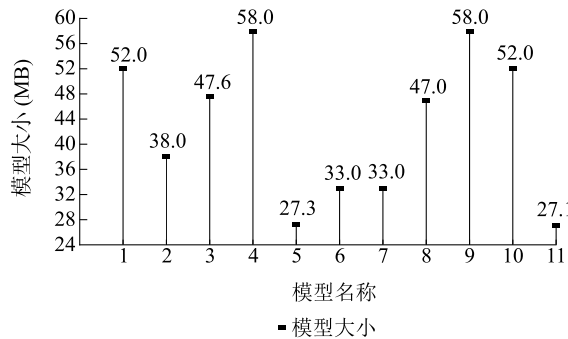
模型名称	精确度 (%)	召回率 (%)	平均精度均值 (%)
YOLOv8m	94.8	89.7	94.4
YOLOv8m+EfficientViT	88.3	87.5	91.7
YOLOv8m+GhostConv	89.3	88.6	92.8
YOLOv8m+C2f-SCConv	94.9	89.5	94.6
YOLOv8m+EfficientViT+GhostConv	92.6	82.7	92.2
YOLOv8m + EfficientViT + GhostConv + C2f-SCConv	94.3	88.8	93.5
YOLOv8m+EfficientViT+WIoU	89.1	87.4	92.1
YOLOv8m+GhostConv+WIoU	90.4	90.0	93.0
YOLOv8m+C2f-SCConv+WIoU	89.4	89.3	93.7
YOLOv8m+WIoU	94.5	90.4	94.7
YOLOv8m + EfficientViT + GhostConv + C2f-SCConv+WIoU	95.2	91.3	95.6

YOLOv8m: YOLOv8 模型中的 m 参数量版本; EfficientViT: 轻量化的高效视觉网络; GhostConv: 幻影卷积; C2f-SCConv: 融合了空间通道重建卷积的跨阶段特征融合模块; WIoU: 具有动态非单调聚焦机制的损失函数。

2.2 模型改进前后性能对比

模型检测过程主要包括预处理、推理和后处理 3 个阶段。其中,推理阶段是最耗时的部分,推理速度主要受模型参数量、模型大小和设备计算能力的限制。通过轻量化改进,模型的计算量和体积显著降低。如表 2 所示,相较于原始模型,改进后的 YOLOv8-EGCW 模型检测时间缩短了 28.8 ms,检测单张图片仅耗时 35.7 ms。这一优化提高了模型的检测效率。

在目标检测中,置信度是衡量模型预测结果可靠程度的重要指标。置信度的数值越接近 1,表明模型的预测结果越可靠。如图 6 所示,改进后的 YOLOv8-EGCW 模型对检测目标的置信度显著高于



1: YOLOv8m; 2: YOLOv8m+EfficientViT; 3: YOLOv8m+GhostConv; 4: YOLOv8m+C2f-SCConv; 5: YOLOv8m+EfficientViT+GhostConv; 6: YOLOv8m+EfficientViT+GhostConv+C2f-SCConv; 7: YOLOv8m+EfficientViT+WIoU; 8: YOLOv8m+GhostConv+WIoU; 9: YOLOv8m+C2f-SCConv+WIoU; 10: YOLOv8m+WIoU; 11: YOLOv8m+EfficientViT+GhostConv+C2f-SCConv+WIoU。YOLOv8m: YOLOv8 模型中的 m 参数量版本; EfficientViT: 轻量化的高效视觉网络; GhostConv: 幻影卷积; C2f-SCConv: 融合了空间通道重建卷积的跨阶段部分特征融合模块; Wiou: 具有动态非单调聚焦机制的损失函数。

图 5 模型大小消融试验结果

Fig.5 Model size ablation experiment results

原始模型。YOLOv8-EGCW 模型在检测精度、置信度和检测速度上的全面提升,能够更好地满足实际应用中高精度和高可靠性检测的需求。

表 2 模型检测速度对比

Table 2 Comparison of model detection speeds

模型	预处理 (ms)	推理 (ms)	后处理 (ms)	总时间 (ms)
YOLOv8m	1.3	61.6	1.6	64.5
YOLOv8m+EfficientViT+GhostConv+C2f-SCConv+WIoU	0.4	33.8	1.5	35.7

YOLOv8m: YOLOv8 模型中的 m 参数量版本; EfficientViT: 轻量化的高效视觉网络; GhostConv: 幻影卷积; C2f-SCConv: 融合了空间通道重建卷积的跨阶段部分特征融合模块; Wiou: 具有动态非单调聚焦机制的损失函数。

2.3 不同模型的性能对比

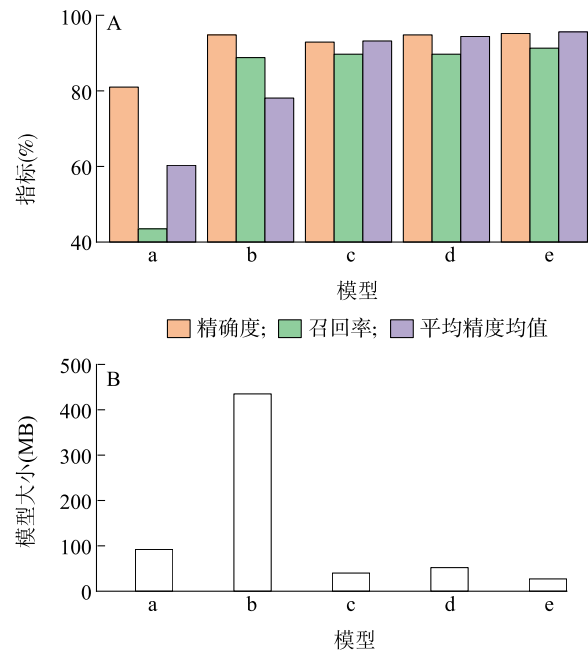
在相同训练参数下,对 YOLOv8-EGCW、YOLOv8、YOLOv5、SSD^[18] 和 Faster-RCNN^[19] 进行对比。试验结果如图 7 所示, YOLOv8-EGCW 模型的平均精度均值 (mAP)、召回率 (R) 和精确度 (P) 均高于其他模型。与精度较低的 SSD 模型相比, YOLOv8-EGCW 模型的 P 值提高 14.2 个百分点, mAP 值提高 35.3 个百分点。YOLOv8-EGCW 模型大小仅为 Faster-RCNN 的 6.2%。尽管 YOLOv8-EGCW 的模型大小和参数量大幅减少,但其平均精度均值、召回率和精确度等关键性能指标仍表现较好,充分证明了改进方法的有效性。



YOLOv8: 第 8 代 YOLO 检测模型; YOLOv8-EGCW: 本研究改进的检测模型。

图 6 模型检测置信度

Fig.6 Model detection confidence levels



a: SSD; b: Faster-RCNN; c: YOLOv5; d: YOLOv8; e: YOLOv8-EGCW。SSD: 目标检测模型; Faster-RCNN: Faster 系列目标检测模型; YOLOv5: 第 5 代 YOLO 检测模型; YOLOv8: 第 8 代 YOLO 检测模型; YOLOv8-EGCW: 本研究改进的 YOLOv8 检测模型。

图 7 不同模型性能对比

Fig.7 Performance comparison of different models

2.4 基于边缘计算的玉米病虫害检测系统设计

2.4.1 基于边缘计算的下位机设计

2.4.1.1 边缘计算技术架构 边缘计算作为云计算或服务器端的延伸和补充,具有轻量化、性能强、体积小特点,在玉米病虫害原位检测中发挥重要作用。一方面,边缘计算能够对采集到的数据进行本地检测和处理,不仅分担了服务器的计算负载,还显著提升了原位检测的实时性。另一方面,边缘计算节点

可以将监测数据存储在本地,再通过通信模块上传至服务器,便于对检测数据进行统一管理和集中分析。这种分布式的检测模式能够有效降低网络延迟,提升资源利用效率。如图 8 所示,以英伟达 Jetson orin

nano 作为核心设备构建边缘计算节点,该节点能够对玉米病虫害信息进行采集,并部署深度学习模型对数据进行处理和保存,最后通过通信模块上传结果。

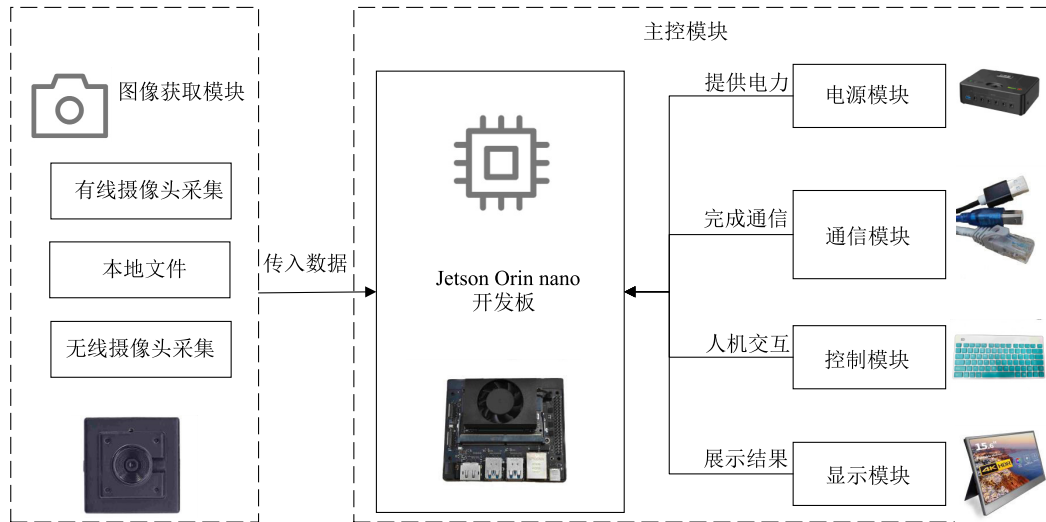


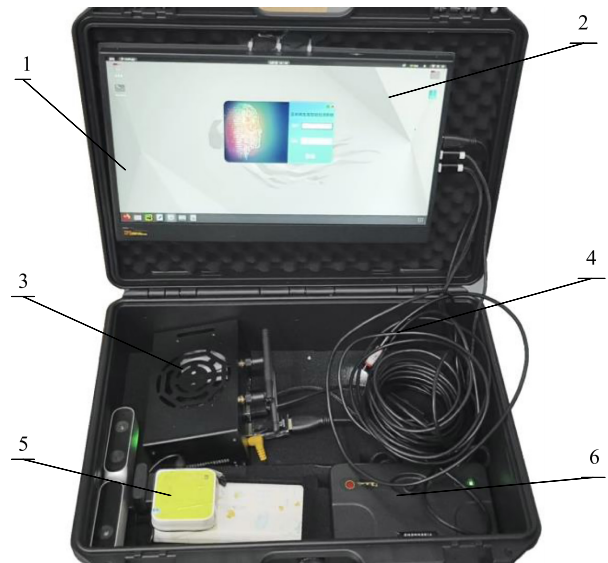
图 8 边缘计算技术架构

Fig.8 Edge computing technology architecture

2.4.1.2 下位机设计 如图 9 所示,下位机系统由图像获取模块和主控模块组成。其中图像获取模块通过多种图像采集设备获取玉米不同部位和不同角度的病虫害图像数据,以确保数据的多样性和全面性。主控模块包括开发板、电源模块、控制模块、显示屏和通信模块,主要负责数据存储、检测分析和结果输出等功能。采集到的图像数据首先在本地进行存储,随后通过主控模块进行实时检测和分析。检测结果可通过显示屏直接输出,还可以利用通信模块将数据上传至服务器,实现数据的远程管理和监控。

采集模块通过采集程序调用摄像头模块获取玉米病虫害图像,无线摄像头模块采用奥比中光公司的 Orbbec 摄像头,无线数据传输模块为搭载了 OpenWrt 系统的 TP-LINK 微型路由器;有线摄像头模块采用 S-YUE 公司的 AF16V1 型号 RGB 相机。系统的核心组件为 Jetson orin nano 开发板,该开发板搭载 Linux 操作系统,配备 6 核 ARM 架构 CPU 和具有 1 024 个 CUDA 核心安培架构 GPU,具备强大的并行计算能力,能够部署至深度学习模型并执行推理任务。

2.4.2 基于边缘计算的上位机设计 图 10 为系统上位机的工作流程示意图。系统上位机在获取病虫害图像数据后,首先对图像进行预处理,随后进行识



1:装置外壳;2:显示模块;3:核心开发板;4:有线摄像头模块;
5:无线摄像头模块;6:电源模块。

图 9 下位机示意图

Fig.9 Schematic diagram of the lower computer

别检测,最终输出并保存检测结果。上位机界面基于 Python 语言开发,结合了 Pytorch、OpenCV 和 Py-Side6 等数据库。其中,Pytorch 数据库用于深度学习模型的训练与推理,OpenCV 数据库用于读取传

感器信息和图像预处理,PySide6 数据库用于构建可视化功能界面。

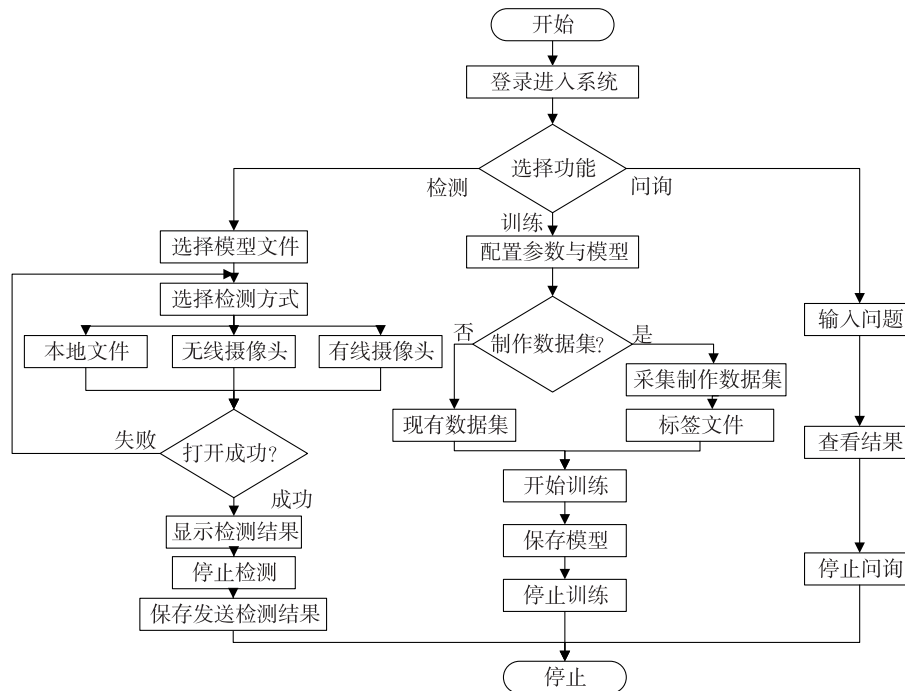


图 10 上位机工作流程示意图

Fig.10 Working process of the upper computer

套接字(Socket)通信流程如图 11 所示。系统完成检测后,检测结果被保存至指定的 TXT 文件中。为便于对检测信息进行集中管理,系统采用基于 Socket 的通信机制,将边缘设备中保存的检测结果传输至服务器端。通过在相同 IP 地址下使用不同的端口号,实现数据传输的加密和隔离,从而增强了通信过程的安全性。Socket 通信的建立遵循服务器监听、客户端请求和连接确认 3 个关键步骤,确保了通信过程的稳定性、数据传输的有序性及信息的完整性。

将改进后的模型 YOLOv8-EGCW 部署在边缘计算装置上,依托 Jetson orin nano 的推理性能,实现对玉米病虫害的快速检测。为充分发挥边缘计算设备的性能,采用 TensorRT 对模型进行加速部署。由于模型依赖于大型深度学习框架,检测过程中涉及大量对不同网络层函数的调用,增加了计算负担。通过 TensorRT 优化模型结构,可以有效减少调用操作,从而提升推理速度。TensorRT 加速机制在于将检测网络中的不同层进行融合,形成更宽的复合层。例如,将卷积层、归一化层和激活层整合为一个 CBR(Conv+BN+ReLU)层,从而减少网络的深度和宽度,减少内部函数的调用次数。这种优化设计不

仅简化了计算流程,还显著提升了检测速度。边缘端模型部署流程如下:首先,将基于 PyTorch 框架训练完成的模型文件转换为 ONNX 格式;然后使用 TensorRT 对病虫害检测模型进行分析和映射;通过 Builder 创建模型张量,加载模型并进行填充和分析,确定输出张量;生成模型引擎文件并执行序列化操作;通过反序列化操作生成最终的推理引擎文件,并在程序中调用该文件进行推理,从而实现病虫害检测模型的高效部署。

系统软件基于 Pyside6 数据库开发,并针对 Linux 系统进行适应性优化。如图 12 所示,软件主要功能包括模型检测、模型训练、大模型询问等。在模型检测功能中,系统通过相机模块采集图像,并利用模型检测病虫害信息,检测结果以 TXT 文件形式保存,同时通过 Socket 通信模块上传至服务器进行后续分析。模型训练支持对 YOLOv8-EGCW 模型进行再训练,以进一步提升模型的鲁棒性和适应性。AI 专家系统功能基于百度 ERNIE 4.0 大语言模型(LLM),能够对检测结果进行智能化解读并提供决策支持,帮助用户高效应对病虫害问题。

2.4.3 系统性能测试结果 在系统性能测试中,随

机选取 200 幅病害图像进行测试,测试结果如表 3 所示,系统共检测到病虫害目标 248 个,而通过人工标注方法确定的测试集病虫害总数为 265 个,计算

得到检测精度为 93.4%。此外,系统对图像的识别速度为 1 s 25 帧。结果表明,该系统具有较高的检测精度和识别速度,能够满足实际应用需求。

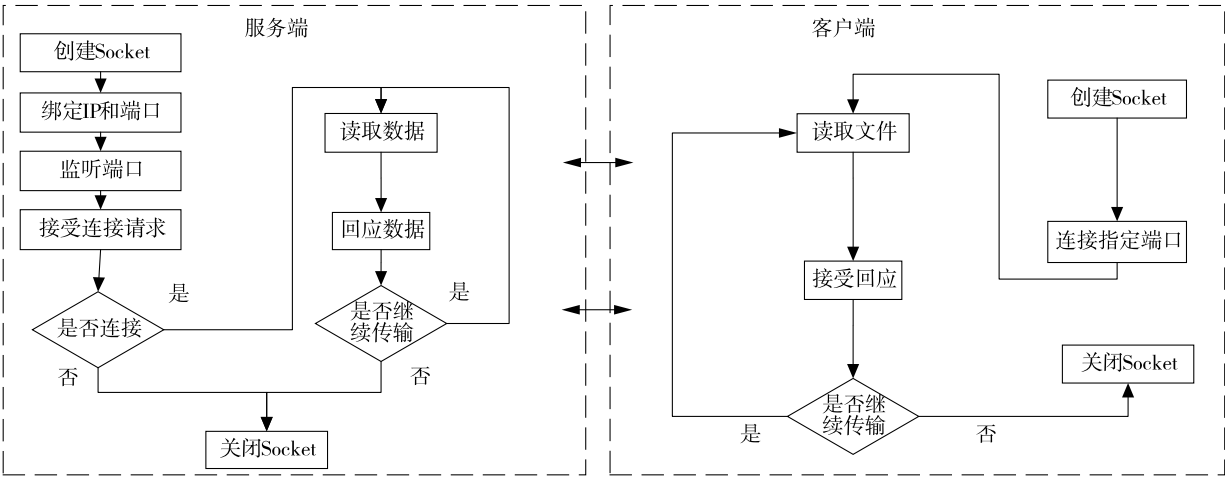


图 11 套接字(Socket)通信流程示意图
Fig.11 Socket communication process



图 12 系统功能界面示意图
Fig.12 Schematic diagram of system functions

表 3 系统性能

Table 3 Device testing

实际病虫害 总数	检测到病虫 害数	准确率 (%)	识别速度 (帧,1 s)
265	248	93.4	25

3 讨论与结论

本研究提出了一种基于 YOLOv8 轻量化改进模

型和边缘计算的玉米病虫害检测系统。首先,在模型轻量化改进方面,采用 EfficientViT 替换原始主干网络,并引入幻影卷积,有效降低了模型计算复杂度,同时保持了优异的特征提取能力。试验结果表明,改进后的 YOLOv8-EGCW 模型的精确率、召回率和平均精度均值均高于原始 YOLOv8 模型,且模型参数量大幅减少,显著降低了模型对硬件资源的需求。特别是在边缘计算环境下,改进后的模型功耗更低,充分体现了轻量化模型在边缘计算设备中的

应用潜力。此外,相较于单一检测模型,本模型实现了病害与虫害的同时检测,具有更强的泛化能力和实用价值。

其次,本研究将 YOLOv8-EGCW 模型部署至 Jetson orin nano 边缘设备,并利用 TensorRT 进行加速。试验结果表明,系统检测速达到 1 s 25 帧,而原始系统检测速度仅为 1 s 12 帧。边缘计算的应用不仅有效分担了云计算的负载,还实现了田间环境的原位检测,显著降低了网络延迟,增强了检测系统的实时性。最后,本研究还集成了基于百度ERNIE 4.0 大语言模型的 AI 专家系统,旨在为病虫害防治提供智能化决策支持。这一集成提高了系统的多应用场景适应性和智能化水平。

然而,本研究也存在一定的局限性。首先,数据集制作过程依赖于人工标注,这不仅增加了整体工作量,而且在处理大规模数据时,可能因疲劳或疏忽而导致数据准确性下降,从而影响整体检测性能。未来的研究将重点关注以下几个方面:构建基于无监督学习的检测模型,通过自动学习图像特征以降低数据标注成本;引入多光谱等传感器技术,实现多源数据融合检测,从而提升复杂背景下的识别能力;设计智能移动底盘并结合太阳能发电技术,实现全天候无人自动化检测。通过这些改进,系统将更好地满足智能农业发展的实际需求,为农业病虫害防治提供更高效的技术支持。

参考文献:

- [1] 刘石磊. 2023 年我国玉米市场运行情况与 2024 年行情展望[J]. 粮食与饲料工业,2024(2):1-3.
- [2] DONG P X, WANG X D. Recognition of greenhouse cucumber disease based on image processing technology[J]. Open Journal of Applied Sciences,2013,3(1):27-31.
- [3] 戴 宁. 图像处理在玉米虫害识别系统中的应用研究[J]. 牡丹江大学学报,2019,28(2):104-107.
- [4] JOTHIARUNA N, SUNDAR K J A, KARTHIKEYAN B, et al. A segmentation method for disease spot images incorporating chrominance in Comprehensive Color Feature and Region Growing[J]. Computers and Electronics in Agriculture,2019,165:104934.
- [5] 蒋龙泉,鲁 帅,冯 瑞,等. 基于多特征融合和 SVM 分类器的植物病虫害检测方法[J]. 计算机应用与软件,2014,31(12):186-190.
- [6] WANG J X, YU H Y, SUI Y Y, et al. Research on identification of corn disease occurrence degree based on improved ResNeXt network[J]. International Journal of Pattern Recognition and Artificial Intelligence,2022,36(2):1-15.
- [7] 陈啟昇,谢 浩,张纬东,等. 基于深度学习的植物病虫害智能检测系统研究[J]. 农村经济与科技,2021,32(18):50-52.
- [8] BI C G, XU S Z, NAN H, et al. Identification method of corn leaf disease based on improved Mobilenetv3 Model[J]. Agronomy, 2023,13:300.
- [9] 徐会杰,黄仪龙,刘 曼. 基于改进 YOLOv3 模型的玉米叶片病虫害检测与识别研究[J]. 南京农业大学学报,2022,45(6):1276-1285.
- [10] 李 颀. 桃树病害和害虫图像检测系统的研究与实现[D]. 泰安:山东农业大学,2022.
- [11] CHEN C J, HUANG Y Y, LI Y S, et al. Identification of fruit tree pests with deep learning on embedded drone to achieve accurate pesticide spraying[J]. IEEE Access,2021,9:21986-21997.
- [12] 刘 忠,卢安舸,崔 浩,等. 基于改进 YOLOv8 的轻量化荷叶病虫害检测模型[J]. 农业工程学报,2024,40(19):168-176.
- [13] LIU X Y, PENG H W, ZHEN N X, et al. Efficientvit: Memory efficient vision transformer with cascaded group attention[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Jersey:IEEE Computer Society, 2023.
- [14] HAN K, WANG Y H, TIAN Q, et al. Ghostnet: more features from cheap operations[C]//IEEE. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. New Jersey: IEEE Computer Society,2020.
- [15] LI J F, WEN Y, HE L H. Seconv: spatial and channel reconstruction convolution for feature redundancy[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Jersey:IEEE Computer Society,2023.
- [16] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. (2023-04-08)[2024-07-21]. <https://doi.org/10.48550/arXiv.2301.10051>.
- [17] 李 茂,肖洋轶,宗望远,等. 基于改进 YOLOv8 模型的轻量化板栗果实识别方法[J]. 农业工程学报,2024,40(1):201-209.
- [18] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//ECCV. 14th European Conference on Computer Vision. Berlin: Springer,2016.
- [19] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2016,39(6):1137-1149.

(责任编辑:成纾寒)