

基于改进型轻量化的YOLOv5玉米病害检测

张立伟¹, 肖国锋¹, 高东浩², 艾山江·阿卜杜拉², 张庆莉^{2*}

¹新疆农业大学计算机信息与工程学院, 新疆 乌鲁木齐

²新疆农业大学数理学院, 新疆 乌鲁木齐

收稿日期: 2025年5月15日; 录用日期: 2025年6月14日; 发布日期: 2025年6月23日

摘要

本文提出了一种基于改进型轻量化YOLOv5的玉米叶片病害检测方法, 通过采用GhostNet模块替换骨干网络并结合GSConv模块优化颈部网络, 显著降低了模型的参数量和计算复杂度。实验结果表明, 改进后的YOLOv5s-Ghost-GSConv-seg在保持较高检测精度的同时, 参数量减少37.5%, 计算量降低20.8%, 推理速度提升15.6%。该模型通过多尺度特征融合设计, 有效平衡了检测精度与实时性, 为农业智能化中的轻量化目标检测提供了高效解决方案。

关键词

实例分割, 病害检测, YOLOv5, 轻量化

Corn Disease Detection Based on Improved Lightweight YOLOv5

Liwei Zhang¹, Guofeng Xiao¹, Donghao Gao², Aishanjiang·Abudula², Qingli Zhang^{2*}

¹College of Computer Information and Engineering, Xinjiang Agricultural University, Urumqi Xinjiang

²College of Mathematics and Physics, Xinjiang Agricultural University, Urumqi Xinjiang

Received: May 15th, 2025; accepted: Jun. 14th, 2025; published: Jun. 23rd, 2025

Abstract

This paper proposes an improved lightweight YOLOv5-based method for detecting maize leaf disease detection by introducing the GhostNet module to replace the backbone network and optimizing the neck network with the GSConv module, significantly reducing model parameters and computational complexity. Experimental results demonstrate that the enhanced YOLOv5s-Ghost-GSConv-seg

*通讯作者。

文章引用: 张立伟, 肖国锋, 高东浩, 艾山江·阿卜杜拉, 张庆莉. 基于改进型轻量化的YOLOv5玉米病害检测[J]. 计算机科学与应用, 2025, 15(6): 99-109. DOI: 10.12677/csa.2025.156161

maintains high detection accuracy while achieving a 37.5% reduction in parameters, a 20.8% decrease in computational cost, and a 15.6% improvement in inference speed. The multi-scale feature fusion design effectively balances detection accuracy and real-time performance, providing an efficient solution for lightweight object detection in smart agriculture applications.

Keywords

Segmentation, Disease Detection, YOLOv5, Lightweight

Copyright © 2025 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

计算机视觉与深度学习的快速发展推动了农业智能化的变革[1]，通过高精度果实识别、机器人自主采摘和基于多模态数据的产量预测技术，显著提升生产效率、降低损耗，并缓解劳动力短缺问题，为精准农业和全链条自动化奠定基础。

玉米作为我国三大主要粮作物之一，在保障国家粮食安全中发挥着战略支柱作用。根据国家统计局数据显示，2024 年我国玉米种植面积达 44740.7 千公顷，玉米总产量突破 29491.7 万吨，占我国谷物总产量的 45.2%。然而，玉米生长周期中易受叶枯病、叶斑病等病虫害威胁，传统病虫害检测依赖人工田间巡查，存在效率低、误判率高、时效性差等问题。

近年基于深度学习的目标检测在农业自动化领域得到了广泛应用，如赵佳悦提出的 Dense-YOLOv3 番茄病害检测模型[2]、杨国国等人结合 GrabCut 算法进行害虫目标的定位和分割[3]、彭红星等人提出改进 SSD (Single Shot MultiBox Detector)深度学习水果检测模型在自然环境下平均检测精度达 88.4% [4]等。当前主流算法可分为两类：第一类为基于候选区域的两阶段检测算法，典型代表包括 Faster R-CNN、Mask R-CNN 等，其通过区域建议网络(RPN)生成候选框再分类回归，精度较高但计算复杂度大；第二类为基于回归的单阶段检测算法，如 YOLO 系列和 SSD 系列，通过端到端方式直接预测目标位置与类别，在速度与精度间取得更好平衡[5]。其中，YOLO 算法凭借其高效的检测性能，在农业场景中展现出显著优势，但 YOLO 在农业场景中仍然存在如复杂背景干扰、病害形态多样性以及边缘设备部署限制等多种挑战。

YOLO 系列与传统机器学习、深度学习模型的比较如表 1，实际应用决策树如图 1，由此分析 YOLOv5 在目标检测中展现出显著优势，使其成为玉米病虫害快速检测的首选解决方案。

Table 1. Comparison of the YOLO series with traditional machine learning and deep learning models

表 1. YOLO 系列与传统机器学习、深度学习模型的比较

指标	YOLO 系列	传统方法	深度学习模型
检测速度	极快(实时)	慢(非实时)	慢(5~10 FPS)
检测精度	高(85% + mAP)	低(环境敏感)	最高(两阶段)
硬件成本	中(需边缘 GPU)	低(树莓派即可)	高(需服务器 GPU)
数据需求	中(迁移学习优化)	低(规则驱动)	高(需大量标注)
适用场景	实时监测	实验室简单检测	高精度病理分析

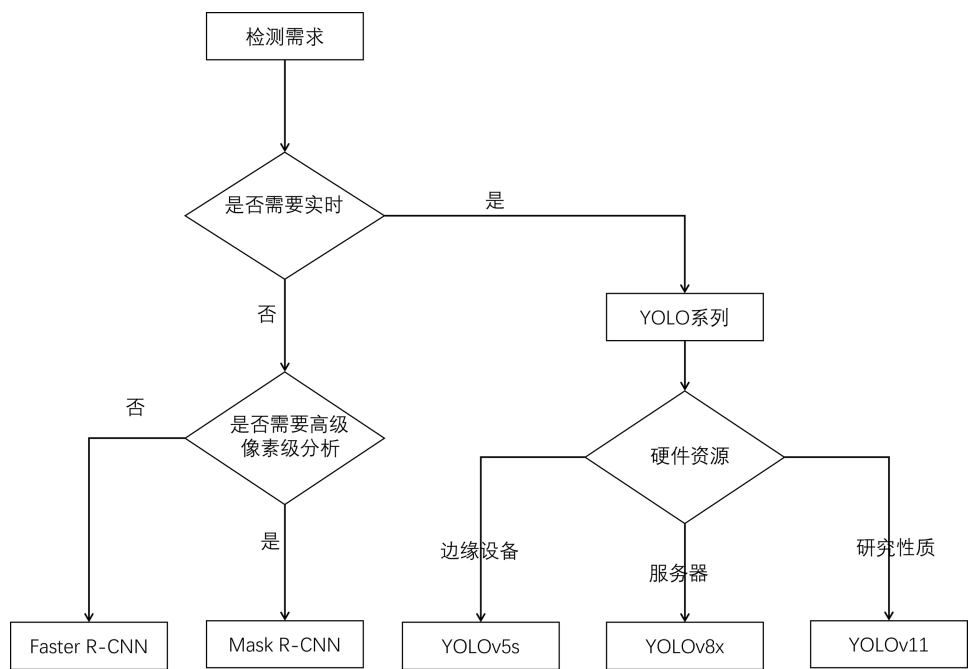


Figure 1. Practical application decision tree
图 1. 实际应用决策树

2. 数据集制作

玉米枯叶病是由真菌或细菌感染引起的常见病害，主要表现为叶片出现黄褐斑、条斑或干枯，严重时整株叶片焦枯、植株早衰，导致光合作用减弱、籽粒发育不良，减产可达 30%以上，因此选择该病为研究对象，其病害样本图像如图 2，共 1000 张，数据来源于 kaggle 开源数据集玉米枯叶病数据集。

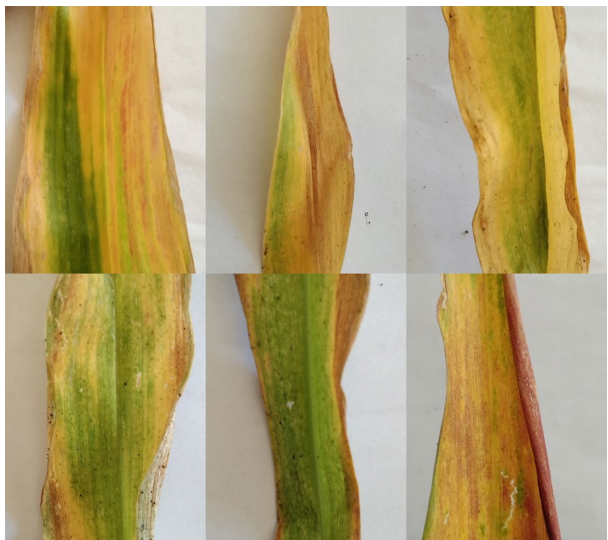


Figure 2. Sample of leaf blight disease
图 2. 枯叶病样本

为实现数据集的扩展和多样性，提高模型训练的鲁棒性和增强泛化能力，首先对枯叶病图片进行翻转、噪声、变亮、变暗等数据增强处理，然后采用 labelme 工具对数据增强后的图片进行特征标记，标记

完成后将保存的 JSON 文件转换为 YOLOv5 所需的 txt 格式，最后划分训练集：验证集：测试集比例为 16:3:1。

3. YOLOv5 模型

YOLOv5 延续了 YOLO 系列的单阶段检测思路，将目标检测任务简化为回归问题，直接在特征图上预测边界框和类别概率。网络结构分为 Backbone (骨干网络)、Neck (特征融合层)和 Head (检测头)，YOLOv5 网络结构如图 3 所示[6]。

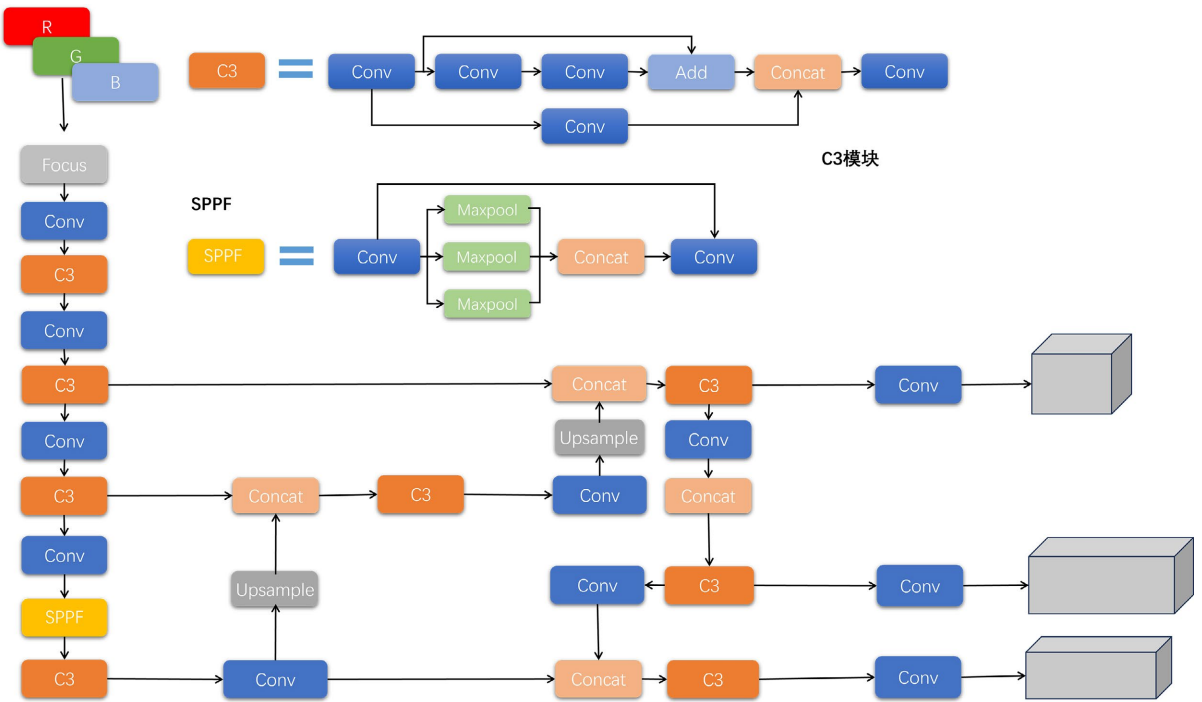


Figure 3. YOLOv5 network structure
图 3. YOLOv5 网络结构

3.1. 骨干网络(Backbone)

Backbone 主要负责对输入图像进行特征提取，将原始的输入图像转化成多层特征图用于后续的目标检测任务，在 YOLOv5 中使用的是 CSPDarknet53 或 ResNet 骨干网络，能够在保证较高检测度的同时尽可能减少计算量和内存占用，Backbone 中的主要结构有 Conv 模块、C3 模块和 SPP 模块。

Conv 模块是卷积神经网络中的基础模块，主要由卷积层、BN 层和激活函数组成，如图 4。卷积层通过滑动窗口和卷积核提取输入特征的局部空间信息，其超参数(如卷积核大小、步长和填充方式)决定了输出特征的大小和感受野。BN 层在卷积层后对特征进行批归一化，通过计算每个通道的均值和方差进行标准化，再通过可学习的仿射变换恢复数据分布，从而加速训练并提升模型泛化能力。激活函数(如 ReLU、sigmoid 等)则引入非线性变换，增强模型对复杂数据分布的拟合能力。三者协同工作，构成高效的特征提取单元。

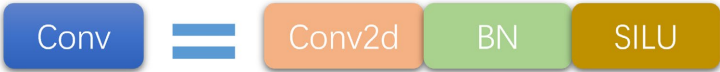


Figure 4. Conv model structure
图 4. Conv 模块结构

C3 模块是 YOLOv5 网络中的关键组件，主要用于增强网络的深度和感受野，提升特征提取能力。该模块由三个 Conv 块组成：第一个 Conv 块采用步幅 2 的 3×3 卷积，将特征图尺寸减半以扩大感受野并降低计算量，同时结合 BN 层和 LeakyReLU 激活函数确保训练稳定性；后续两个 Conv 块保持步幅 1，通过 3×3 卷积在不改变特征图尺寸的情况下进一步深化特征表征，保留局部细节信息。其中，第一个 Conv 块后接的 Bottleneck 结构(由两个 Conv 块组成)可高效压缩和扩展通道维度，平衡计算效率与特征表达能力，如图 5。C3 模块通过这种分层设计，兼顾全局语义信息与局部细节特征，显著提升了目标检测任务中对多尺度目标的识别与定位精度。

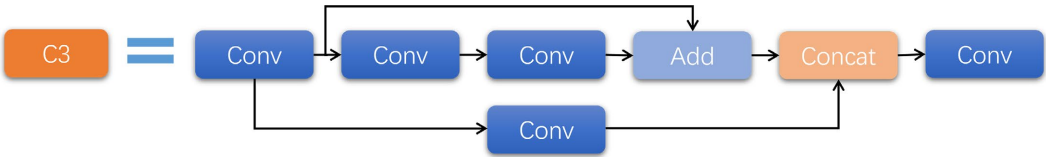


Figure 5. C3 model structure
图 5. C3 模块结构

SPP 模块是卷积神经网络中用于增强特征表达能力的关键结构，如图 6，它通过并行执行多种尺度的池化操作(如 1×1 、 5×5 、 9×9 等)来提取多级感受野特征，有效捕捉不同尺度的空间信息。该模块首先对输入特征图进行多尺度池化获得不同大小的特征图，然后将这些特征拼接融合，最后通过全连接层压缩为固定维度的特征向量。这种设计不仅突破了网络对输入尺寸的限制，还显著提升了模型的空间不变性和多尺度特征提取能力，使其在目标检测和图像识别任务中表现出更强的鲁棒性和准确性。

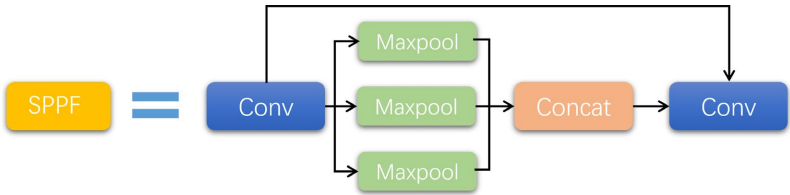


Figure 6. SPPF model structure
图 6. SPPF 模型结构

3.2. 颈部网络(Neck)

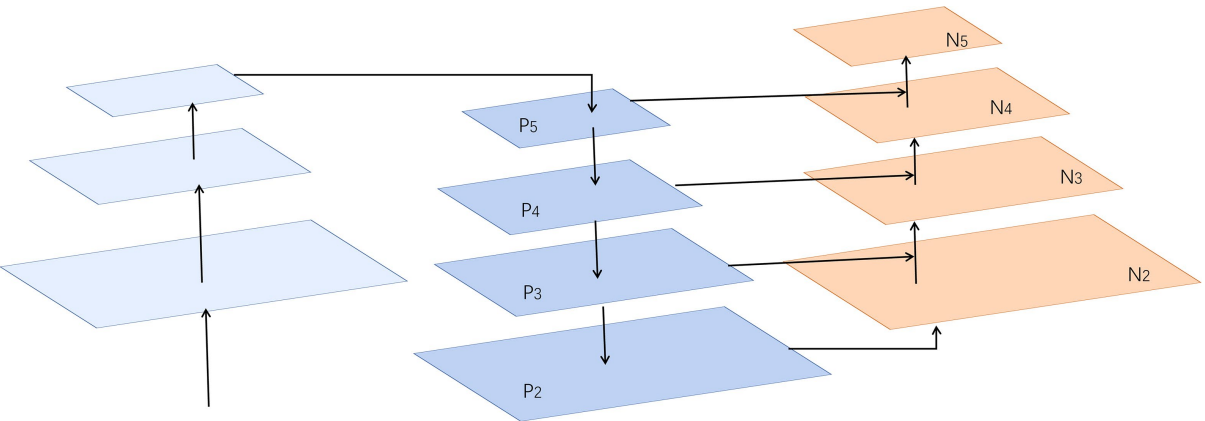


Figure 7. Neck module structure
图 7. Neck 模块结构

YOLOv5 的颈部网络(Neck)采用 FPN + PAN 的双向多尺度特征融合架构,如图 7,通过自顶向下(FPN)和自底向上(PAN)的双向路径,将骨干网络提取的深层语义特征与浅层细节特征进行跨层级融合。该结构以 C3 模块为基础单元,结合上采样、下采样和特征拼接操作,构建了覆盖不同感受野的特征金字塔(P3~P5),既能增强小目标的检测能力,又能提升大目标的定位精度。这种设计显著提高了模型对多尺度目标的适应性,在保持轻量化的同时实现了精度与速度的平衡,使其成为 YOLOv5 高效检测能力的核心组件。

3.3. 检测头(Head)

YOLOv5 的检测头(Head)是一个多尺度目标检测模块,如图 8,通过融合特征金字塔的不同层级特征进行预测。它由三部分组成: 1) Anchors (通过 K-means 聚类预定义的多种比例边界框); 2) Classification 分支(卷积层 + Softmax 分类); 3) Regression 分支(卷积层 + 坐标回归)。检测头采用卷积层进行特征处理,通过预测层输出边界框和类别,最后使用 NMS 过滤冗余检测。关键技术包括 GloU Loss、Mish 激活函数和多尺度训练,通过 Conv 模块降维后实现跨层级特征融合提升检测精度。

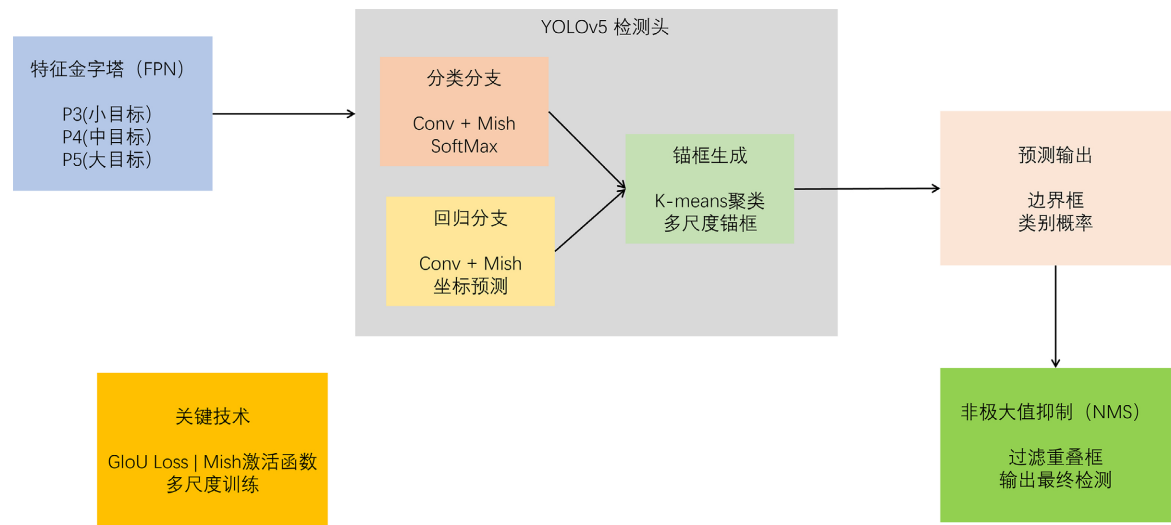


Figure 8. Head module structure
图 8. Head 模块结构

4. 改进 YOLOv5 模型

4.1. 改进 GhostNet 模块代替模型主干网络[7]

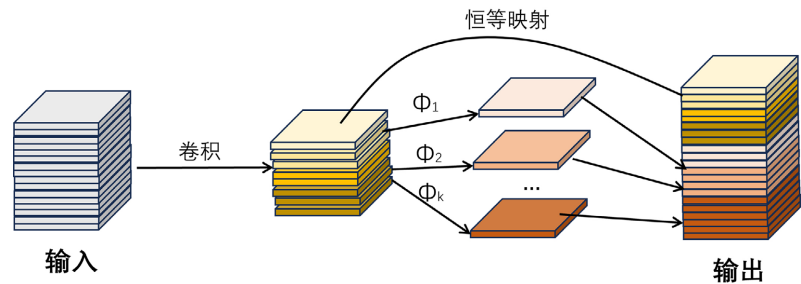


Figure 9. Ghost module
图 9. Ghost 模块

Ghost Moule 是一种模型压缩的方法,核心思想是将一个卷积层分解成两个较小的卷积层,其中一个

卷积层被称为 Ghost 卷积层, 如图 9, 它只使用原始卷积层的一小部分通道来计算输出, 而另一个卷积层则被称为剩余卷积层, 它使用剩余的通道来计算输出。在保证网络精度的同时减少网络参数和计算量, 从而提升计算速度和降低延时。这种分解方式可以大大减少计算量和参数数量, 同时保持网络的精度。

Ghost Bottlenecks 是由 Ghost Module 组成的瓶颈结构, 可分为主干部分和残差部分, 如图 10。

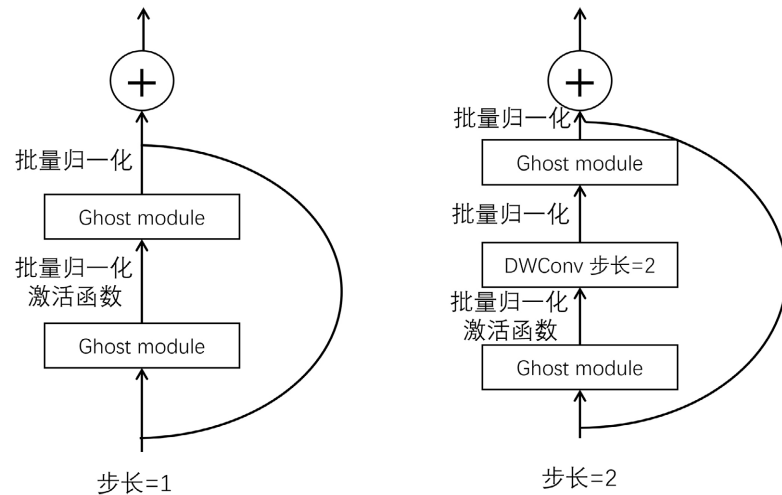


Figure 10. Ghost bottleneck structure with different strides
图 10. 不同步长的 Ghost 瓶颈结构

步长为 1 时, 主干部分用两个 Ghost Module (GM) 串联组成, 其中第一个 GM 扩大通道数, 第二个 GM 将通道数降低到与输入通道数一致; 残差部分与 ResNet 一样。由于步长为 1, 因此不会对输入特征层的高和宽进行压缩, 其功能为加深网络的深度。步长为 2 时, 主干部分的两个 GM 之间加入了一个步长为 2 的 Deepwise 卷积, 可以将特征图高和宽进行压缩, 使其大小降为输入的 1/2; 在残差部分, 也会添加一个步长为 2×2 的深度可分离卷积和 1×1 的普通卷积, 以保证 Add 操作可以对齐。由于步长为 2, 因此会对输入特征层的高和宽进行压缩, 其功能为改变输入特征层的形状。

4.2. 检测头融入 GSConv 模块[8]

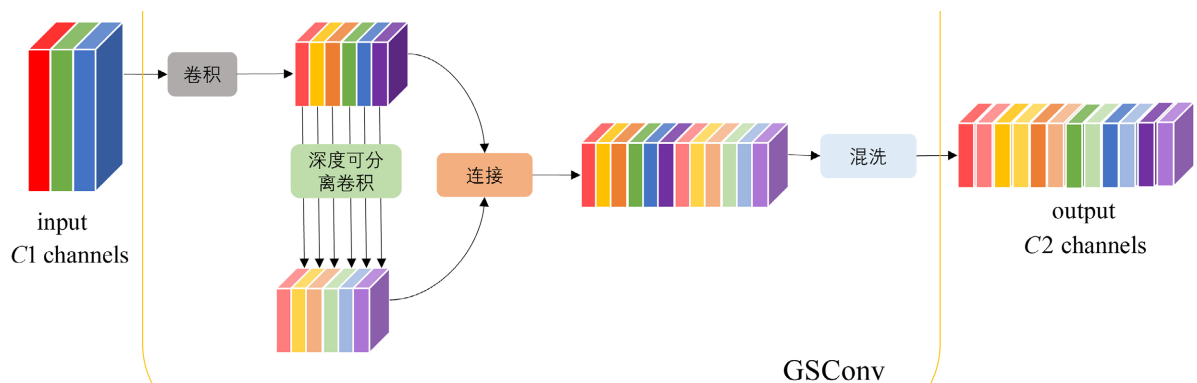


Figure 11. GSConv module structure
图 11. GSConv 模块结构

GSConv (Global Sliding Convolution) 是一种轻量化卷积设计, 通过通道分割策略在单层内融合标准卷积(SC)和深度可分离卷积(DSC)。它将输入通道划分为两组, 分别用 SC 提取细节特征和 DSC 降低计算

量,最后拼接结果。随着网络加深,GSConv 动态调整两组通道的比例,以平衡特征提取能力和计算效率,如图 11。H Li 等人[9]在研究无人驾驶时发现由大量深度可分离卷积建成的轻量级模型也无法达到实时监测足够的精度,因此引入了一种新的轻量级卷积技术 GSConv,在模型精度和速度之间实现了权衡。

优化改进后的 YOLOv5 网络结构如图 12 所示:

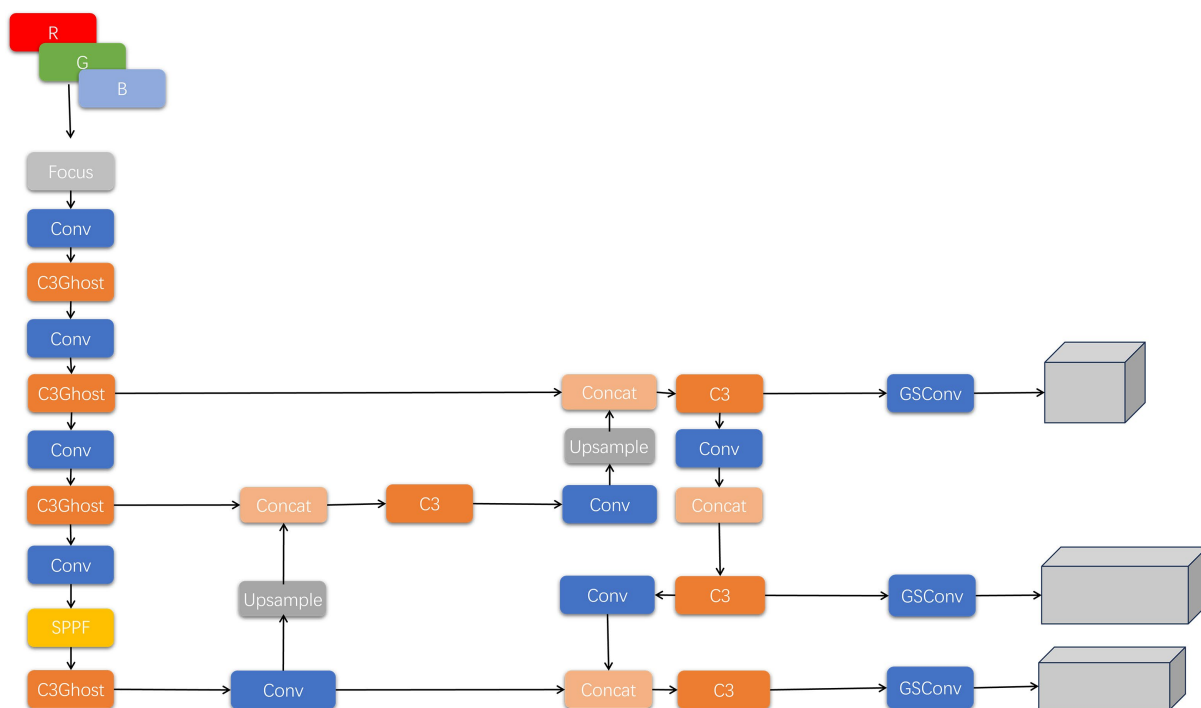


Figure 12. The improved network structure of YOLOv5

图 12. YOLOv5 改进后网络结构

5. 性能分析与结论

5.1. 模型训练

PyTorch 框架 2.0.0, CUDA 版本 11.8, 选用 YOLOv5s-seg、YOLOv5m-seg, YOLOv5l-seg 进行实例分割训练,改进后的 YOLO 模型在原有文件的基础上添加 Ghost 模块和 GSConv 模块,并对配置文件进行修改,模型预训练权重文件及相应配置文件在 Ultralytics 官网获得,迭代次数 epoch 设为 100, batch_size 设 16, imagesz 为 640 像素进行训练。

Backbone 部分以 $640 \times 640 \times 3$ 输入开始,通过四次 3×3 卷积, $\text{stride} = 2$,逐步下采样至 $20 \times 20 \times 1024$ 特征图,其间交替插入 C3Ghost 模块,分别使用 3、6、9、3 次不同深度的特征增强,最后经 SPPF 层增强感受野。

Head 部分通过多尺度双向特征融合,高层特征 P5 经 GSConv 压缩通道后上采样与中层 P4 拼接,经 C3 模块融合后继续上采样与浅层 P3 交互,同时底层特征 P3 通过下采样与中层特征二次融合,中层特征再下采样与高层特征二次交互,形成双向跨尺度特征金字塔。

YOLOv5-seg 模型的训练框架采用全链路优化设计,使用 SGD 优化器,配合动态权重衰减和参数分组策略,学习率调度集成余弦退火与线性衰减双模式,迭代学习率从 0.1 线性攀升,动量从 0.8 渐增,训练过程采用 AMP 混合精度与梯度裁剪保障稳定性,结合 EMA 权重平均提升模型鲁棒性,动态调整机制

包括自动锚框验证、损失函数参数层级归一化和类别权重平衡,配合多尺度训练与分布式同步批归一化,形成兼顾训练效率与精度的完整优化体系。

5.2. 评价指标

采取精确率、召回率、平均精度均值 mAP 、浮点运算数 $FLOPs$ 、模型参数数量和模型在 RTX3060 监测的帧率 FPS 作为评价指标来评估模型整体性能。精确率($Precision$)、召回率($Recall$)、平均精度均值(mAP)计算如式(1)所示[10]。

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN}, \quad mAP = \frac{\sum_{i=1}^C \int_0^1 P(Recall) dRecall}{C} \quad (1)$$

参数说明: C 为目标检测任务中的类别数, TP 预测为正例, 实际为正例, 预测成功(True), FP 预测为正例, 实际为负例, 预测错误(False), FN 预测为负例, 实际为正例, 预测错误(False)。

5.3. 实验结果

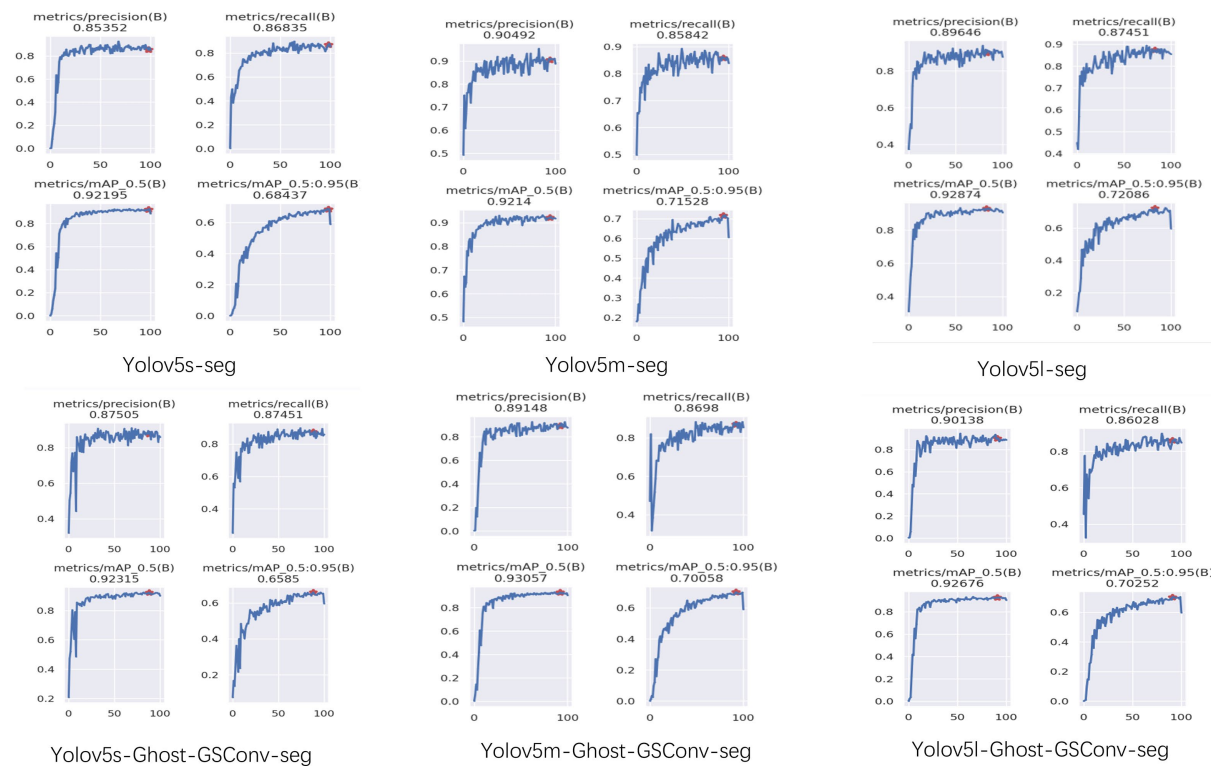


Figure 13. Model training results

图 13. 模型训练结果

Table 2. Performance of different models in instance segmentation tasks

表 2. 不同模型在实例分割任务中的表现

model	mAP	Precision	recall	F1	GFLOPs	Params (M)	FPS
YOLOv5s	0.921	0.854	0.868	0.86	25.9	7.2	45
YOLOv5m	0.921	0.905	0.858	0.85	70.2	21.2	32
YOLOv5l	0.928	0.896	0.875	0.88	147.0	46.5	22

续表

YOLOv5s-Ghost-GSConv	0.923	0.875	0.875	0.86	20.5	4.5	52
YOLOv5m-Ghost-GSConv	0.917	0.891	0.870	0.86	45.9	15.8	38
YOLOv5l-Ghost-GSConv	0.927	0.901	0.860	0.87	81.9	34.7	26

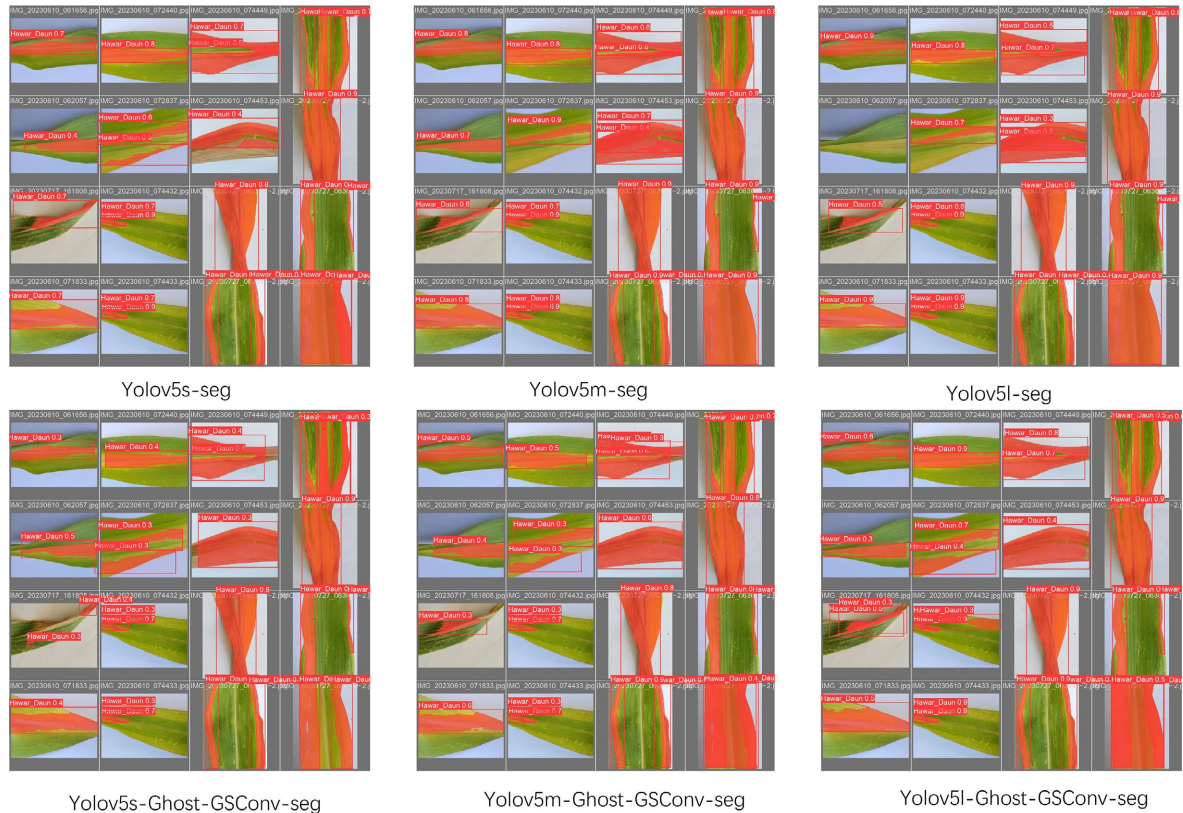


Figure 14. Different model segmentation effects
图 14. 不同模型分割效果

为对比所提方法与原模型的性能差异，将实例分割模型 YOLOv5s-seg, YOLOv5m-seg, YOLOv5l-seg, YOLOv5s-Ghost-GSConv-seg, YOLOv5m-Ghost-GSConv-seg, YOLOv5l-Ghost-GSConv-seg 进行对比实验，结果如图 13 和表 2 所示。不同模型的检测效果如图 14 所示。

5.4. 结论

本研究通过引入 GhostNet 和 GSConv 模块对 YOLOv5 模型进行轻量化改进，在玉米叶片病害检测任务中取得了显著成效。实验结果表明，改进后的 YOLOv5s-Ghost-GSConv-seg 模型在保持原有检测精度 (mAP 0.923) 的同时，参数量从 7.2 M 降至 4.5 M (降幅 37.5%)，计算量从 25.9 GFLOPs 减少至 20.5 GFLOPs (降幅 20.8%)，推理速度从 45 FPS 提升至 52 FPS (增幅 15.6%)，不仅显著降低了计算资源需求，还提高了实时检测性能，特别适合部署在资源有限的农业边缘设备上，为田间实时病害监测提供了高效可靠的解决方案，同时该方法的轻量化设计思路也为其他农作物病害检测任务提供了有益借鉴。

基金项目

新疆农业大学校级大学生创新训练项目，《基于 YOLO 算法的农作物叶片病虫害的检测》，

dxscx2025388。

参考文献

- [1] 何东健, 张海亮, 宁纪锋, 等. 农业自动化领域中计算机视觉技术的应用[J]. 农业工程学报, 2002(2): 171-175.
- [2] 赵佳悦. 基于卷积神经网络的番茄病害检测研究[D]: [硕士学位论文]. 济南: 山东师范大学, 2020.
- [3] 杨国国, 鲍一丹, 刘子毅. 基于图像显著性分析与卷积神经网络的茶园害虫定位与识别[J]. 农业工程学报, 2017, 33(6): 156-162.
- [4] 彭红星, 黄博, 邵园园, 等. 自然环境下多类水果采摘目标识别的通用改进 SSD 模型[J]. 农业工程学报, 2018, 34(16): 155-162.
- [5] 张阳婷, 黄德启, 王东伟, 等. 基于深度学习的目标检测算法研究与应用综述[J]. 计算机工程与应用, 2023, 59(18): 1-13.
- [6] Redmon, J., Divvala, S., Girshick, R., *et al.* (2016) You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>
- [7] 陈继平, 陈永平, 谢懿, 等. Ghost-YOLO: 轻量化口罩人脸检测算法[J]. 信号处理, 2022, 38(9): 1954-1964.
- [8] Wang, J., Qin, C., Hou, B., *et al.* (2024) LCGSC-YOLO: A Lightweight Apple Leaf Diseases Detection Method Based on LCNet and GSConv Module under YOLO Framework. *Frontiers in Plant Science*, **15**, Article ID: 1398277. <https://doi.org/10.3389/fpls.2024.1398277>
- [9] Li, H., Li, J., Wei, H., *et al.* (2024) Slim-Neck by GSConv: A Lightweight-Design for Real-Time Detector Architectures. *Journal of Real-Time Image Processing*, **21**, Article No. 62. <https://doi.org/10.1007/s11554-024-01436-6>
- [10] 丁展博, 徐健, 朱耀麟, 等. 基于轻量化 YOLO 模型的花椒簇检测研究[J]. 中国农机化学报, 2025, 46(4): 120-125+353.