

GD-RiceNet: 轻量级水稻叶病检测网络

洪佳乐

重庆科技大学数理科学学院, 重庆

收稿日期: 2026年9月14日; 录用日期: 2026年10月5日; 发布日期: 2026年10月20日

摘要

针对复杂田间环境中水稻叶部病害目标尺度差异大、背景干扰严重且资源受限等问题, 本文借鉴水稻叶脉的重复纹理特性, 构建轻量级检测网络GD-RiceNet。叶脉纹理的局部自相似性与Ghost模块的轻量化思路相契合, 据此, 该网络以GhostHGNetV2为骨干减少标准卷积计算, 颈部引入Slim-neck实现低开销多尺度特征融合, 同时采用DySample动态上采样与DyHead动态检测头缓解轻量化带来的空间细节损失与特征选择不足。在公开田间水稻病害数据集上, GD-RiceNet的mAP@50和mAP@50:95分别达57.3%和35.3%, 参数量6.56 M, 计算量12.7 GFLOPs, 推理速度169 FPS。相比YOLO11s, 参数量和计算量分别降低30.4%和40.9%, mAP@50和mAP@50:95分别提升0.6和1.4个百分点, 推理速度由185 FPS降至169 FPS。结果表明, GD-RiceNet在检测精度与模型效率之间取得了较好平衡, 具备面向农业边缘设备部署的潜力。

关键词

水稻病害检测, 轻量级网络, 特征融合, 边缘计算, YOLO11

GD-RiceNet: A Lightweight Network for Rice Leaf Disease Detection

Jiale Hong

School of Mathematical and Physical Sciences, Chongqing University of Science and Technology, Chongqing

Received: September 14, 2026; accepted: October 5, 2026; published: October 20, 2026

Abstract

Rice leaf disease detection in complex field environments is challenged by large variations in lesion scale, strong background interference, and limited edge-computing resources. The local self-similarity of leaf vein textures aligns with the lightweight design principle of the Ghost module. Accordingly, the network uses GhostHGNetV2 as its backbone to reduce standard convolutional computation and

incorporates Slim-neck into its neck for low-cost multiscale feature fusion. It also employs DySample dynamic upsampling and the DyHead dynamic detection head to mitigate the loss of spatial detail and inadequate feature selection associated with lightweight design. Under a unified data split and shared test protocol, GD-RiceNet achieves 57.3% mAP@50 and 35.3% mAP@50:95 on a public field rice disease dataset, with 6.56 M parameters, 12.7 GFLOPs, and 169 FPS under the shared test conditions. Relative to YOLO11s, GD-RiceNet reduces parameters and computation by 30.4% and 40.9%, respectively, while improving mAP@50 and mAP@50:95 by 0.6 and 1.4 percentage points; throughput decreases from 185 to 169 FPS. These results indicate a favorable accuracy-efficiency trade-off under the present dataset and test conditions. The model therefore warrants further validation on independent datasets and agricultural edge devices.

Keywords

Rice Disease Detection, Lightweight Network, Feature Fusion, Edge Computing, YOLO11

Copyright © 2026 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

水稻是全球重要的粮食作物之一，也是全球一半以上人口的主食，其稳定生产对粮食安全和农业可持续发展具有重要意义[1]。Conde 等人[2]的综述表明，稻瘟病、白叶枯病等病害在全球主要稻作区广泛发生，并可显著影响产量。Mukherjee 等人[3]进一步总结了稻瘟病、白叶枯病和褐斑病等水稻叶部病害的机器视觉识别研究，指出及时、准确的识别是病害管理的重要前提。

在此背景下，病害的及时识别成为降低产量损失和提高植保效率的关键环节。当前田间监测仍较多依赖人工巡查，其效率和一致性容易受到作业规模、人员经验和现场条件的影响。Zahoor 等人[4]指出，化肥和农药的不合理使用会通过径流和淋溶造成水体污染，这凸显了以精准识别支持按需防治的必要性。Saleem 等人[5]的综述表明，随着深度学习技术的发展，基于视觉感知的植物病害自动识别已成为重要研究方向。

在深度学习广泛应用之前，早期植物病害识别多采用传统机器学习方法。Yao 等人[6]先分割水稻病斑并提取形状与颜色纹理特征，再用支持向量机(SVM)识别水稻白叶枯病、纹枯病和稻瘟病。此类流程依赖人工特征，其鲁棒性受采集条件和特征设计影响。Singh 等人[7]构建的 PlantDoc 包含非实验室环境下的植物病害图像，体现了复杂背景和自然成像条件对模型泛化的现实挑战。

相比之下，深度学习方法能够从图像中自动学习多层次特征，已逐渐成为农业视觉检测的重要工具。以 Faster R-CNN [8]为代表的两阶段检测方法通过区域候选网络生成候选区域，再完成分类和定位。这类方法通常具有较强的特征表达能力，但候选区域生成和后续精细分类会增加计算开销，在强调实时性的农业监测场景中存在一定限制[9]。

与此同时，真实田间环境中的水稻病害检测仍面临多方面挑战。Cao 等人[10]指出，稻瘟病斑面积小、边界不清，特征传递过程中的小目标信息损失会增加检测难度。Liu 等人[11]提出 Swin Transformer，以分层结构和移位窗口实现多尺度视觉建模，并将自注意力计算量控制为相对图像尺寸线性增长。Nugroho 等人[12]在 ARM Cortex-M 微处理器上评估水稻病害卷积神经网络，进一步表明边缘部署需要在准确率、存储和计算资源之间取得平衡。

为降低部署成本,研究者提出了多种轻量化网络结构。Li 等人[13]提出 GSConv 和 Slim-neck,在保持通道交互能力的同时降低实时检测器的计算成本。另一方面, Lin 等人[14]提出的 FPN 通过自顶向下路径和横向连接融合高层语义与高分辨率特征,为小目标的多尺度表示提供了经典方案。因此,如何在轻量化的同时尽可能保持病斑边缘、纹理及多尺度语义信息,是农业检测模型需重点考虑的问题。

在上述研究趋势中, YOLO (You Only Look Once) [15]系列单阶段检测算法将目标分类和位置回归统一在端到端框架中,在检测速度和精度之间取得了较好的平衡。Kiratiratanapruk 等人[16]已将 YOLOv3 等检测器用于自然田间水稻病害图像的定位与分类。YOLO11s 具有相对较低的参数规模,适合作为本文轻量化改进的基础框架。

在轻量化与特征增强的具体技术上, DySample [17]通过学习采样偏移量实现内容自适应上采样, DyHead [18]通过尺度、空间和任务三个维度的注意力调节检测特征; Ghost 卷积[19]利用廉价线性变换生成与本征特征相关的辅助特征,以减少部分重复计算; ShuffleNetV2 [20]则通过通道重排等方式改善轻量网络中的信息流动。值得注意的是,水稻叶片具有平行叶脉和重复纹理,这些结构容易产生相关性较高的视觉响应,本文将这种结构相关性作为采用 Ghost 类轻量化特征生成方式的直观动机。

基于上述分析,本文在 YOLO11s 基础上构建面向资源受限场景的轻量级水稻病害检测网络 GD-RiceNet。本文的主要工作如下:

- (1) 以 GhostHGNetV2 作为骨干网络,借助 Ghost 卷积的轻量化特征生成机制减少标准卷积中的冗余计算。
- (2) 构建“特征压缩-动态补偿”组合结构,通过 GhostHGNetV2 和 Slim-neck 降低计算开销,结合 DySample 和 DyHead 恢复轻量化过程中的空间细节损失。
- (3) 在真实田间水稻病害数据集上进行实验验证,结果表明 GD-RiceNet 仅包含 6.56 M 参数量和 12.7 GFLOPs 计算量,相较 YOLO11s 减少 30.4%的参数量和 40.9%的计算量,同时达到 57.3%的 mAP@50 和 35.3%的 mAP@50:95,为农业边缘智能设备部署提供了一种有效方案。

2. 方法

2.1. 整体网络设计

YOLO 系列检测算法将目标定位和类别预测统一为端到端学习过程,具有结构清晰和推理速度较快等特点[15]。Kiratiratanapruk 等人[16]在自然田间水稻病害图像上比较了 Faster R-CNN、RetinaNet、YOLOv3 和 Mask R-CNN,验证了 YOLO 模型在该场景中的适用性。本文选择 YOLO11s 作为基础检测框架,主要考虑其参数规模相对较小,同时保留了多尺度特征提取和检测能力,便于在此基础上进一步研究复杂田间环境下模型精度与计算复杂度之间的平衡。

水稻病害图像具有其特殊性:叶片表面连续分布的平行叶脉和重复纹理易导致标准卷积产生冗余响应,而田间背景中的叶片交叠、土壤和水面反射进一步加剧了这一问题;同时小面积病斑在逐层下采样后容易出现细节弱化。因此,本文将减少特征冗余和保持多尺度病斑信息作为网络改进的主要目标。

针对上述问题, GD-RiceNet 采用“特征压缩-动态补偿”的整体设计思路。首先,以 GhostHGNetV2 替换 YOLO11s 的原始骨干结构,利用 Ghost 特征生成机制降低部分卷积计算;随后,在 Neck 中引入 Slim-neck,以 GSConv 和 VoV-GSCSP 完成多尺度特征融合,以较低开销实现不同层级特征的信息交互;最后,使用 DySample 替代固定上采样方式恢复空间细节,并采用 DyHead 对多尺度检测特征进行动态调节,增强模型对不同尺度病斑的适应性。

GD-RiceNet 的整体结构如图 1 所示。骨干网络输出的多尺度特征进入 Slim-neck 进行融合, DySample

用于上采样路径中的空间分辨率恢复, DyHead 接收不同尺度的融合特征并输出最终检测结果。该结构将计算压缩部署于骨干和颈部, 将特征补偿集中于上采样和检测阶段, 通过“压缩-补偿”的分阶段协同寻求检测性能与计算开销的平衡。

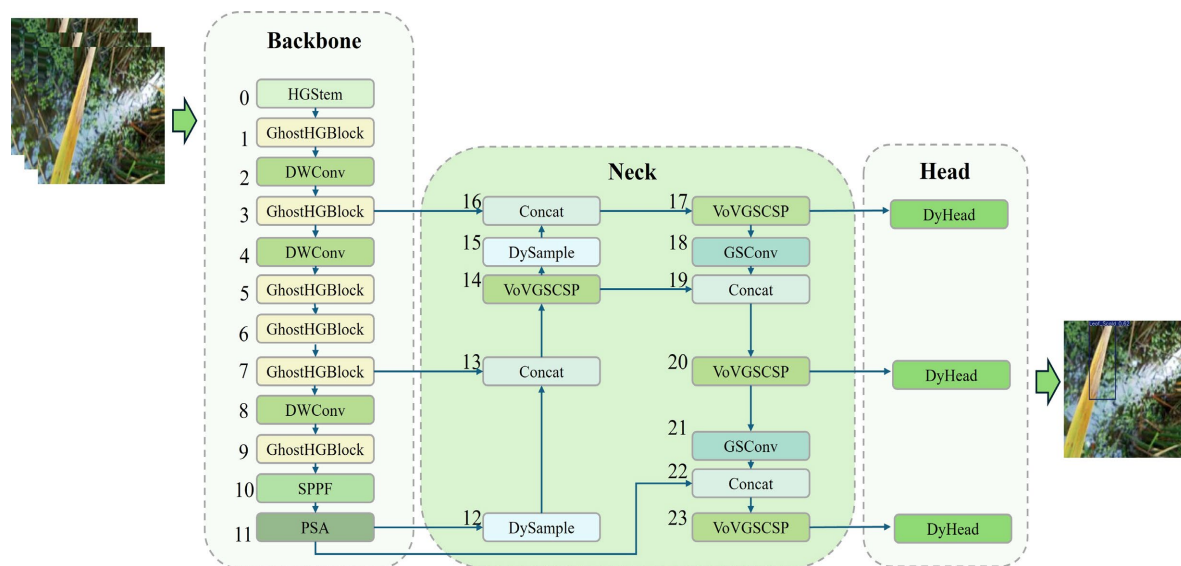


Figure 1. Overall architecture of GD-RiceNet

图 1. GD-RiceNet 整体网络架构

2.2. 基于叶脉结构先验的轻量化特征提取

水稻叶片图像中平行叶脉和重复纹理具有较高的视觉相关性, 标准卷积在相似区域可能产生较多冗余响应。相比之下, 病害区域常伴随颜色、纹理和边界的局部变化, 其视觉特征与正常组织存在差异。基于这一观察, 本文将叶片结构相关性作为轻量化特征提取的直观启发, 采用 GhostHNetV2 以减少部分重复计算。图 2 给出了标准卷积与 GD-RiceNet 设计思路的概念性对比。

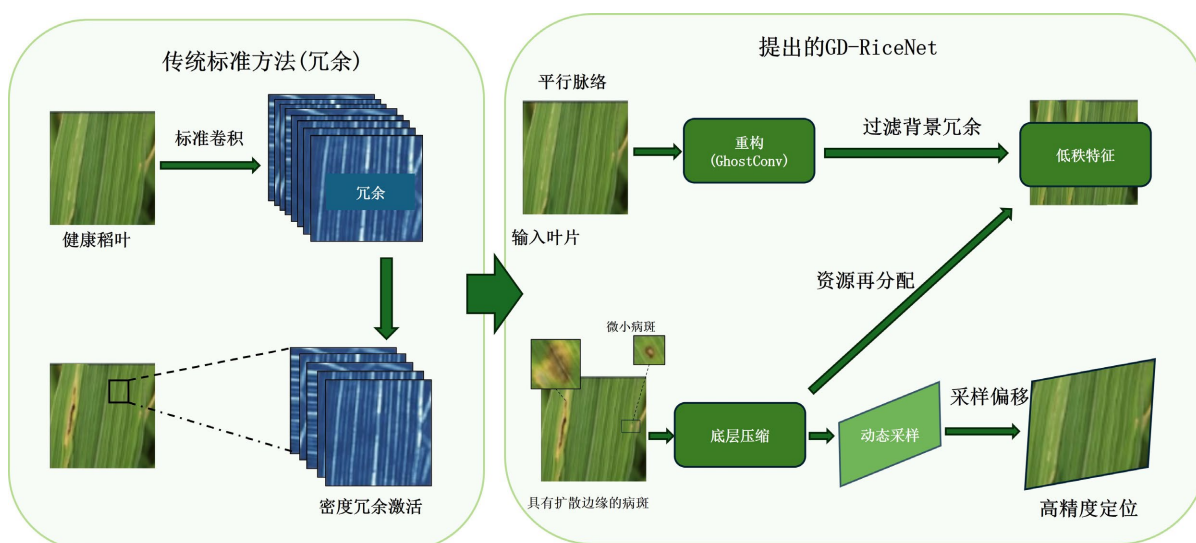


Figure 2. Comparison between standard convolution and GD-RiceNet

图 2. 标准卷积方法与 GD-RiceNet 的概念对比

Ghost 卷积[19]先通过标准卷积生成少量本征特征,再利用低成本线性变换生成辅助特征,以较低开销获得完整输出。水稻叶片重复纹理在局部区域呈现的高度自相似性,与 Ghost 卷积通过廉价变换生成相关特征的轻量化思路高度契合,因此本文采用 GhostHGNetV2 作为特征提取骨干。

在 GD-RiceNet 中, Ghost 卷积被嵌入 GhostHGBlock,用于替代部分标准卷积操作。该设计旨在减少参数量和计算量并保持特征表达能力,其对背景响应的抑制效果由后续消融实验和 Grad-CAM 可视化进行辅助讨论。

本文的 GhostHGNetV2 以 PaddlePaddle 视觉团队公开的 PP-HGNetV2 架构[21]为原型,将其中部分卷积替换为 Ghost 卷积并构造 GhostHGBlock;骨干网络及 GhostHGBlock 结构如图 3 所示。HGStem 首先完成输入图像的初始特征提取,随后由多个 GhostHGBlock 和下采样单元逐级生成层次化特征。图 3(a)中的末端分类结构为 PP-HGNetV2 原型的完整形式,本文检测网络实际使用送入 Neck 的多尺度中间特征。Ghost 卷积首先生成本征特征:

$$Y' = X * f' + b \quad (1)$$

其中, X 表示输入特征, f' 表示卷积核参数, b 表示偏置项, Y' 表示生成的本征特征图。与直接生成全部输出通道的标准卷积相比, Ghost 卷积仅生成部分本征特征,其输出通道数满足 $m < n$, n 表示标准卷积期望生成的输出通道数量。

随后,通过低成本线性变换操作 Φ 进一步生成与本征特征相关的 Ghost 特征:

$$y_{ij} = \Phi_{ij}(y'_i) \quad (2)$$

其中, Φ_{ij} 表示第 i 个本征特征经过第 j 个线性变换后得到的辅助特征。最终,将本征特征与 Ghost 特征进行拼接,形成完整的输出特征表示。

通过上述特征生成机制, GhostHGNetV2 能够减少部分标准卷积所需的参数量和计算量,并为后续多尺度特征融合提供较为紧凑的输入表示。

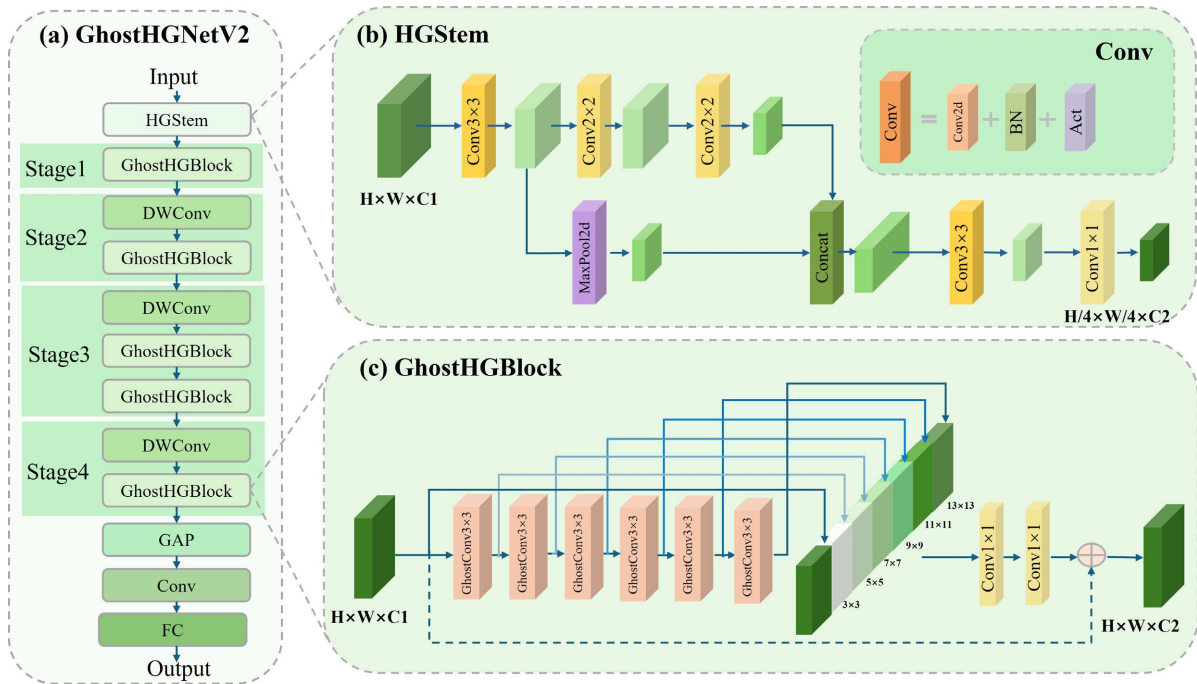


Figure 3. GhostHGNetV2 backbone and GhostHGBlock module: (a) backbone hierarchy; (b) HGStem; (c) GhostHGBlock
图 3. GhostHGNetV2 骨干与 GhostHGBlock 模块: (a) 骨干层次; (b) HGStem 组件; (c) GhostHGBlock 结构

2.3. 基于 Slim-Neck 的低开销多尺度特征融合

轻量化骨干网络完成特征提取后, 还需在不同尺度之间传递和融合语义信息。然而, 传统 Neck 通常包含较多标准卷积操作, 在丰富特征表达的同时也会增加参数量和计算量。为降低特征融合阶段的开销, 本文采用 Slim-neck 连接 GhostHGNetV2 骨干与动态检测头, 以更轻量的结构实现多尺度特征交互。

本文的 Slim-neck 以 GSConv 和 VoV-GSCSP 替代 YOLO11s 中部分 C3k2 结构。GSConv [13] 结合标准卷积(SC)和深度可分离卷积(DSC), 在通道交互与计算效率之间取得平衡, 其标准卷积分支为跨通道信息交互保留了补充通路。

如图 4(a)所示, GSConv 通过标准卷积分支提取特征, 同时利用深度可分离卷积分支降低计算开销, 随后将两路特征沿通道维度拼接。该结构的设计目标是减少 Neck 中的计算量, 同时尽可能保留病斑边缘和纹理变化等局部细节信息, 其实际效果由后续累计消融实验进行检验。

GSConv 引入通道混洗(Channel Shuffle)操作以改善分支间信息交换, 该设计遵循 ShuffleNetV2 [20] 的轻量化思想, 通过重排通道缓解分组计算造成的信息隔离(图 4(b))。

在此基础上, 本文集成 VoV-GSCSP 模块, 以 GSBottleneck 作为主要计算单元, 并通过一次聚合融合不同阶段的特征。该模块在控制计算开销的同时促进特征复用, 不改变本文检测任务的最终输出形式。

GSConv 与 VoV-GSCSP 共同构成 Slim-neck, 完成多尺度特征融合并为后续 DySample 和 DyHead 提供特征输入。根据表 1 消融结果, 引入 Slim-neck 后模型复杂度进一步降低, 检测指标基本持平。

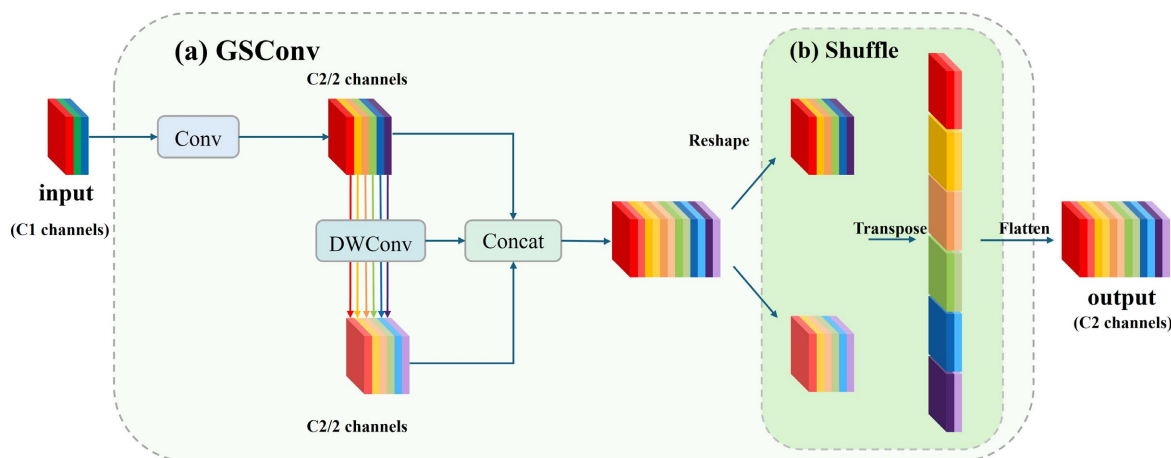


Figure 4. GSConv hybrid operator and shuffle operation: (a) GSConv design; (b) channel shuffle

图 4. GSConv 混合算子与 Shuffle 操作: (a) GSConv 设计; (b) 通道混洗

2.4. 基于动态感知的自适应定位

GhostHGNetV2 和 Slim-neck 有效降低了模型的参数量与计算量, 但轻量化过程也可能在一定程度上削弱空间细节和跨尺度信息的表达。对于面积较小、边界较模糊的病斑, 这种信息弱化效应可能更为明显。为此, 本文在上采样和检测头阶段分别引入 DySample 与 DyHead, 作为对轻量化特征的动态补充。

传统上采样采用固定采样位置和插值规则, 缺乏内容自适应能力。DySample [17] 通过学习采样偏移量动态调整重采样位置, 本文将其用于替代固定上采样以改善高分辨率特征重建中的局部空间信息保持。

具体而言, DySample 根据原始采样网格 G 和网络学习得到的偏移量 Δ 构建动态采样位置:

$$S = G + \Delta \quad (3)$$

其中, S 表示动态调整后的采样位置。随后, 根据新的采样位置对输入特征 X 进行重采样, 得到增强后

的高分辨率特征 X' :

$$X' = \text{grid_sample}(X, S) \quad (4)$$

该机制使采样位置随输入内容变化,为局部结构重建提供更灵活的采样方式。DySample 的实际作用依据表 1 中 YOLO11s-B 与 YOLO11s-C 之间的数值变化进行讨论。

在完成上采样之后,本文采用 DyHead [18] 统一动态检测头替换基础检测头,以完成多尺度病害目标的分类与定位。DyHead 从尺度、空间和任务三个维度调节特征响应,能够在不同分辨率特征和不同检测任务之间进行动态选择,增强模型对多样化目标的适应能力。

如图 5 所示, DyHead 将输入表示为三维特征张量 F , 并依次进行尺度感知、空间感知和任务感知调节:

$$W(F) = \pi_c(\pi_s(\pi_L(F) \cdot F) \cdot F) \cdot F \quad (5)$$

其中, π_L 、 π_s 和 π_c 分别表示尺度感知、空间感知和任务感知注意力函数。该多维调节机制旨在改善跨尺度特征交互和检测阶段的特征选择质量。

综上, GD-RiceNet 在骨干和 Neck 阶段进行计算压缩, 在上采样和检测头阶段采用动态特征处理, 形成“特征压缩 - 动态补偿”的组合结构, 其有效性在第 4 节中通过累计消融、对比实验和可视化结果进行系统验证。

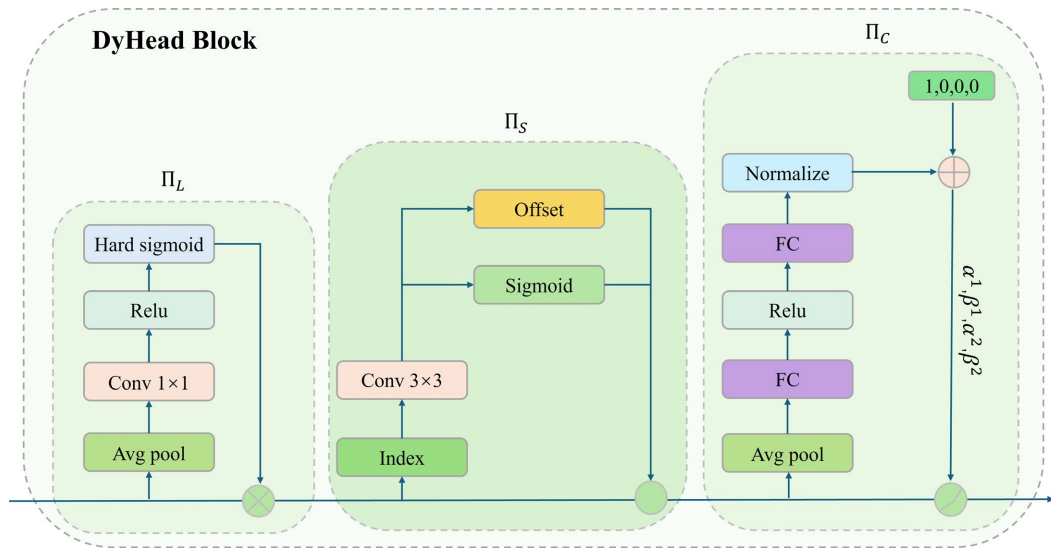


Figure 5. Architecture of the unified dynamic detection head (DyHead)

图 5. 统一动态检测头(DyHead)架构

3. 数据集及评价指标

3.1. 数据集

本文采用公开水稻叶斑病害标注数据集[22]进行实验。该数据集包含 3567 张图像和 5466 个实例标注, 图像均采集于实际田间环境, 涵盖 7 类水稻叶部病害和 1 类健康叶片样本, 不同类别在颜色、纹理、目标尺度和空间分布方面存在明显差异。

如图 6 所示, 数据集中包含多种水稻叶部病害表现, 同时存在土壤背景、水面反射、叶片遮挡和光照变化等典型田间干扰因素。本文基于该数据集评价不同模型在同一数据划分条件下的检测结果。在数

据增强之前, 本文按照 7:2:1 的比例随机划分数据集, 其中训练集包含 2497 张图像, 验证集包含 713 张图像, 测试集包含 357 张图像。所有对比模型均使用相同的数据划分进行训练和测试, 以避免划分差异对结果公平性的影响。

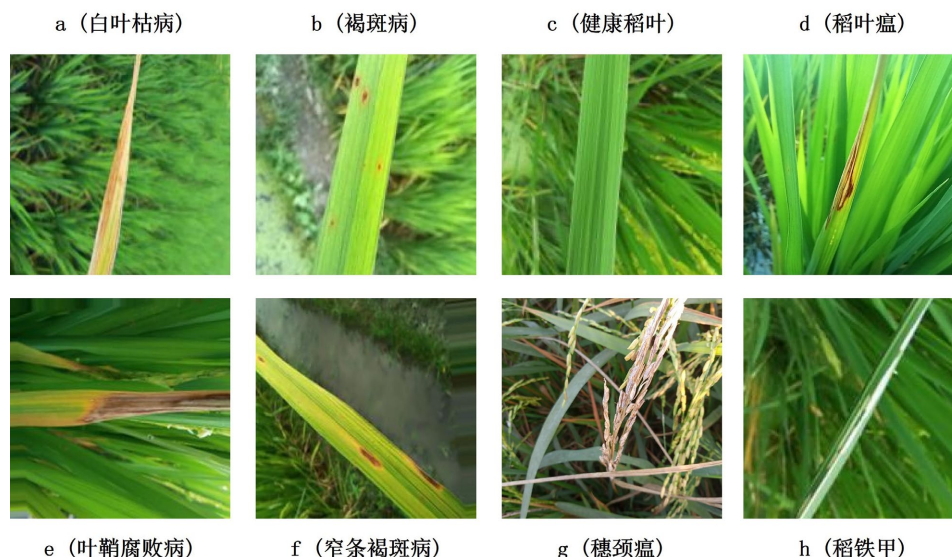


Figure 6. Representative samples of rice leaf diseases under complex field conditions
图 6. 复杂田间环境下水稻叶部病害的代表性样本

3.2. 实验环境与训练参数设置

为保障实验的公平性与可复现性, 全部实验均基于 Intel Core i5-12600KF CPU 与 NVIDIA GeForce RTX 4060Ti GPU 硬件平台开展, 软件环境配置为 PyTorch 2.0.0、CUDA 11.8。所有模型均使用统一训练参数, 不加载预训练权重, 在完全一致的实验条件下完成训练与测试。

训练阶段输入图像分辨率设置为 640×640 , 最大训练轮次为 300 个 Epoch, Batch Size 取 16; 优化器选用 SGD, 初始学习率 0.01, 动量 Momentum 为 0.937, 权重衰减 Weight Decay 设置为 0.0005, 早停耐心值 Patience 为 50, Num Workers 设置为 8。前 3 个 Epoch 启用 Warmup 预热策略, 对应 Warmup Momentum 为 0.8, Warmup Bias Learning Rate 为 0.1。同时, 所有模型均采用完全相同的数据增强方案, 进一步确保各组对比实验之间的公平可比。

3.3. 性能评价指标

为评价模型的检测性能和计算复杂度, 本文采用精确率(Precision, P)、召回率(Recall, R)、平均精度均值(mean Average Precision, mAP)、参数量(Parameters, Params)、浮点运算次数(Floating Point Operations, FLOPs)和每秒处理帧数(Frames Per Second, FPS)作为指标。

3.3.1. 精确率与召回率(Precision and Recall)

根据交并比(IoU)阈值判断预测框与真实框的匹配结果, 检测结果分为真正例(TP)、假正例(FP)和假负例(FN)。

精确率表示模型给出的目标预测中正确预测所占的比例:

$$P = \frac{TP}{TP + FP} \quad (6)$$

召回率表示真实目标中被模型成功检出的比例：

$$R = \frac{TP}{TP + FN} \quad (7)$$

3.3.2. 平均精度与平均精度均值(Average Precision and Mean Average Precision)

平均精度(Average Precision, AP)可表示给定 IoU 阈值 θ 下精确率 - 召回率曲线的面积。平均精度均值(mAP)为各类别 AP 的平均值：

$$mAP@ \theta = \frac{1}{N} \sum_{i=1}^N \int_0^1 P_i(R_i) dR_i \quad (8)$$

其中， N 表示目标类别数量。

本文采用 $mAP@50$ (IoU 阈值 0.50)和 $mAP@50:95$ (IoU 从 0.50 至 0.95、步长 0.05 的 10 个阈值下 mAP 的平均)两个指标：

$$mAP@50:95 = \frac{1}{10} \sum_{\theta \in \{0.50, 0.55, \dots, 0.95\}} mAP@ \theta \quad (9)$$

3.3.3. 模型复杂度与推理速度

本文使用参数量、FLOPs 和 FPS 分别描述模型的可学习参数规模、单次前向理论运算量和单位时间处理图像数。

$$Params = (C_{in} \cdot k^2 + 1) \cdot C_{out} \quad (10)$$

$$FLOPs = 2HW (C_{in} \cdot k^2 + 1) C_{out} \quad (11)$$

4. 实验

4.1. 消融实验

为观察 GD-RiceNet 各组成部分加入后的累计变化，本文以 YOLO11s 为基础模型，依次采用 GhostHGNetV2、Slim-neck、DySample 和 DyHead 构建不同配置。该实验属于逐步累积消融，主要反映各模块在前一配置基础上的增量变化，而不是各模块相互独立的单因素比较。结果如表 1 所示。

Table 1. Quantitative results of the cumulative ablation configurations in GD-RiceNet

表 1. 消融实验中 GD-RiceNet 各组件的定量结果

Model	GhostHGNetV2	Slim-neck	DySample	DyHead	Params (M)	GFLOPs	mAP@50 (%)	mAP@50:95 (%)	FPS
YOLO11s	-	-	-	-	9.43	21.5	56.7	33.9	185
YOLO11s-A	□	-	-	-	7.38	17.4	56.5	33.5	172
YOLO11s-B	□	□	-	-	7.02	16.9	56.6	33.7	172
YOLO11s-C	□	□	□	-	7.54	17.2	57.0	34.7	185
Ours	□	□	□	□	6.56	12.7	57.3	35.3	169

首先，以 GhostHGNetV2 替换 YOLO11s 的原始特征提取结构，得到 YOLO11s-A。与基线(FPS = 185)相比，参数量由 9.43 M 降至 7.38 M，GFLOPs 由 21.5 降至 17.4，分别减少 21.7%和 19.1%；mAP@50 由 56.7%变为 56.5%，mAP@50:95 由 33.9%变为 33.5%。该结果表明骨干替换显著降低了模型复杂度，检测指标出现了小幅下降。

在 YOLO11s-A 基础上采用 Slim-neck, 得到 YOLO11s-B。参数量进一步降至 7.02 M, GFLOPs 降至 16.9; mAP@50 和 mAP@50:95 分别为 56.6%和 33.7%, 较 YOLO11s-A 略有回升。该结果说明 Slim-neck 进一步降低了模型复杂度并基本保持检测指标。

随后, 在 YOLO11s-B 基础上采用 DySample, 得到 YOLO11s-C。mAP@50 由 56.6%提高至 57.0%, mAP@50:95 由 33.7%提高至 34.7%; 参数量和 GFLOPs 分别变为 7.54 M 和 17.2, FPS 为 185 (DySample 的采样偏移学习不增加显著推理延迟, FPS 与基线持平)。上述结果提示动态上采样与更高的检测指标相关, 但带来一定参数和计算变化。尤为值得注意的是, mAP@50:95 取得 1.0 个百分点的增益, 这也印证了 DySample 自适应采样机制可以更好保留病斑边界与空间细节, 进而提升高 IoU 阈值条件下的目标定位性能。

最后, 在 YOLO11s-C 基础上将原始检测头替换为 DyHead, 构建最终 GD-RiceNet。由于 DyHead 采用统一的注意力机制替代原有多个独立尺度的卷积检测分支, 模型参数量降至 6.56 M, GFLOPs 降至 12.7; mAP@50 和 mAP@50:95 分别达到 57.3%和 35.3%, FPS 为 169。FPS 的下降主要源于 DyHead 中跨尺度注意力操作在 GPU 上的执行效率低于理论 FLOPs 所反映的水平, 但其实际推理速度仍满足实时检测需求。相较 YOLO11s 基线, mAP@50 和 mAP@50:95 分别提高 0.6 和 1.4 个百分点。从模块增量消融结果进一步分析, 引入 DyHead 后 mAP@50:95 实现 0.6 个百分点的提升, 说明该模块在尺度、空间及任务维度上的动态特征筛选能力, 能够有效强化多尺度水稻病斑的特征表达。

总体而言, GhostHGNetV2 与 Slim-neck 负责轻量化特征压缩, DySample 与 DyHead 负责动态特征补偿, 共同形成“特征压缩-动态补偿”的协同机制。

4.2. 对比实验

为评价 GD-RiceNet 在相同数据划分下的综合表现, 本文将其与 RT-DETR-r18 [23]、RetinaNet [24]、YOLOv5s [25]、YOLOv8s [26]、YOLOv10s [27]和 YOLO11s 进行比较, 结果如表 2 所示。

Table 2. Comprehensive performance comparison between GD-RiceNet and representative detection algorithms
表 2. GD-RiceNet 与主流算法的综合性能对比

Model	Params (M)	GFLOPs	mAP@50 (%)	mAP@50:95 (%)	FPS
RT-DETR-r18	19.88	57.0	48.8	28.9	71
RetinaNet	36.48	147.5	52.9	28.0	14
YOLOv5s	9.12	24.0	56.1	32.8	204
YOLOv8s	11.14	28.6	57.0	33.0	188
YOLOv10s	7.23	21.6	53.6	32.4	164
YOLO11s	9.43	21.5	56.7	33.9	185
Ours	6.56	12.7	57.3	35.3	169

由表 2 可知, GD-RiceNet 在参数量(6.56 M)、计算量(12.7 GFLOPs)、检测精度(mAP@50 为 57.3%, mAP@50:95 为 35.3%)和推理速度(169 FPS)四个维度上均表现出较强的综合性能。在表中模型范围内, GD-RiceNet 的参数量和 GFLOPs 均处于较低水平, 而 mAP@50:95 为最高值。

具体而言, 与轻量级 YOLO 系列模型的对比呈现出一致的规律。与 YOLO11s 相比, GD-RiceNet 参数量减少 30.4%, GFLOPs 降低 40.9%, mAP@50 和 mAP@50:95 分别提高 0.6 和 1.4 个百分点, FPS 由 185 降至 169。与 YOLOv10s 相比, GD-RiceNet 参数量减少 9.3%, GFLOPs 降低 41.2%, mAP@50 和

mAP@50:95 分别提高 3.7 和 2.9 个百分点。上述结果表明, 在本文数据集和训练设置下, GD-RiceNet 以可控的推理速度下降换取了模型复杂度的显著降低和检测精度的全面提升, 在各项指标之间取得了较好的折中。与 YOLOv8s 和 YOLOv5s 相比, GD-RiceNet 同样在更低复杂度下实现了更高的 mAP@50:95, 进一步验证了这一趋势。与 RT-DETR-r18 和 RetinaNet 相比, GD-RiceNet 的优势更为突出: 参数量分别降低约 67.0%和 82.0%, mAP@50:95 分别提高 6.4 和 7.3 个百分点, 差距主要源于两者较高的模型复杂度与计算开销。

综合来看, GD-RiceNet 在推理速度方面虽略低于 YOLOv5s (204 FPS)、YOLOv8s (188 FPS)和 YOLO11s (185 FPS), 但结合其仅为 6.56 M 的参数量和 12.7 GFLOPs 的计算量, 该模型在显著降低复杂度的前提下仍保持了 169 FPS 的较快推理能力, 能够在检测精度与推理速度之间实现有效平衡。

4.3. 实验结果可视化对比

除整体指标外, 本文进一步采用精确率 - 召回率(P-R)曲线观察各模型在不同病害类别上的检测表现, 如图 7 所示(蓝色粗线为综合结果, 其余颜色对应各类别)。

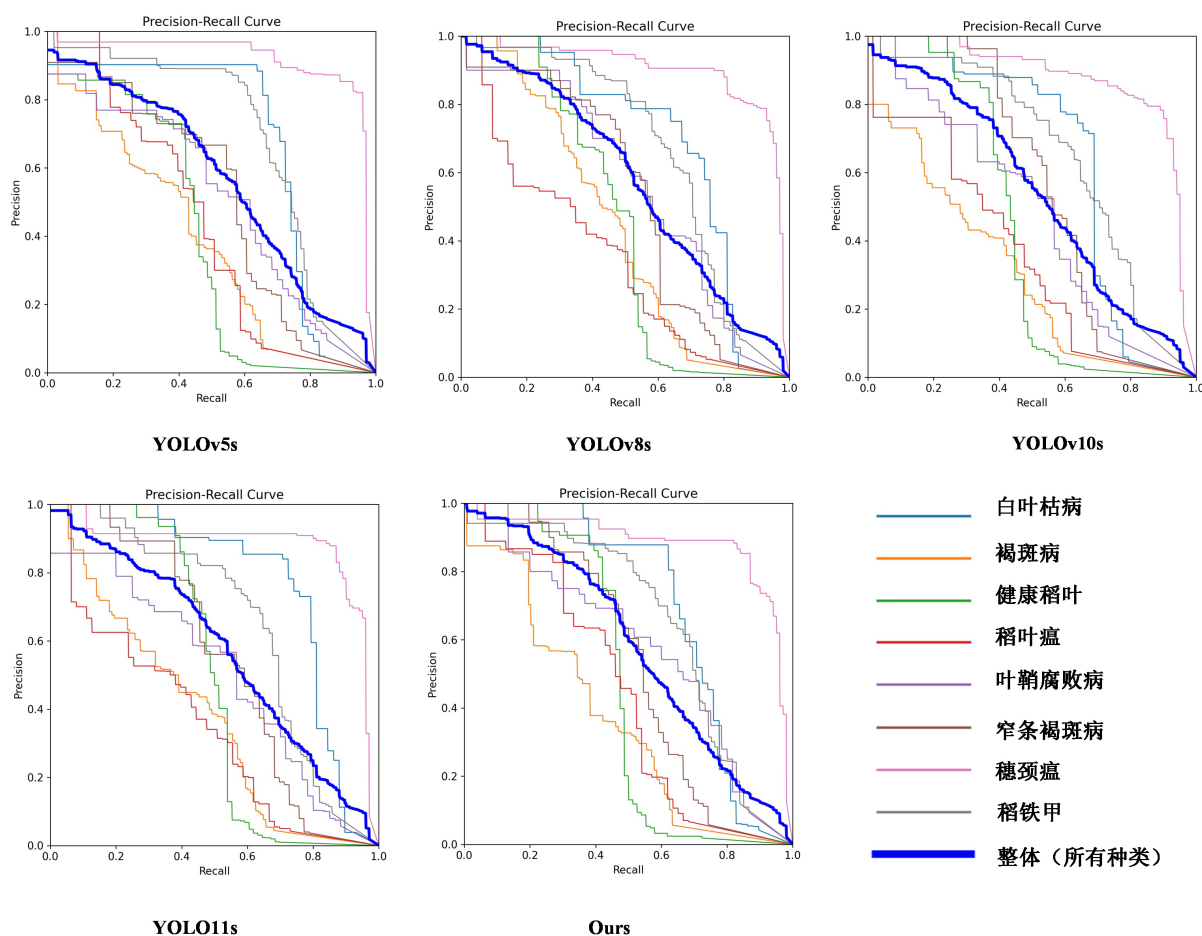


Figure 7. PR curves of YOLO baselines across rice leaf diseases

图 7. YOLO 基线模型在水稻叶部病害上的 PR 曲线

从综合结果看, GD-RiceNet 的 P-R 曲线覆盖范围较大(mAP@50 为 57.3%), 且在多个召回率区间内保持了较高的精确率, 与表 2 中整体指标趋势一致。在类别层面, 穗颈瘟在多个模型中均取得了较高的

AP, 而叶瘟和褐斑病等类别检测难度较大。其中, GD-RiceNet 在叶瘟上的 AP 达到 0.436, 高于 YOLO11s 的 0.346, 表明模型对该小病斑类别的检出能力有所提升。在细菌性条斑病和窄褐斑病等类别中, GD-RiceNet 同样保持了较完整的 P-R 曲线, 说明模型对病害形态和尺度变化具有一定适应能力。综上, 图 7 从类别粒度补充了不同病害的精确率与召回率变化特征, 与表 2 整体结果形成互补印证。

4.4. 检测结果可视化

为直观分析模型在田间图像中的响应和预测情况, 本文从 Grad-CAM 特征响应与检测结果两个方面进行可视化。

4.4.1. Grad-CAM 特征响应分析

为观察网络改进过程中模型关注区域的变化, 本文采用 Grad-CAM 方法[28]对 YOLO11s 基线、累计改进配置 YOLO11s-A、YOLO11s-B、YOLO11s-C 以及最终 GD-RiceNet 进行特征响应可视化。从图 8 可以看出, 不同网络配置对病害区域和背景纹理的响应分布存在一定差异。YOLO11s 基线在部分样本中的高响应区域较为分散, 除病斑位置外, 叶脉、叶片边缘及复杂背景区域也出现了一定程度的激活。引入 GhostHGNv2 和 Slim-neck 后, 部分样本中的响应范围逐渐向叶片病害相关区域集中; 加入 DySample 后, 小面积病斑及边界附近的响应得到一定增强。最终 GD-RiceNet 在多数代表性样本中能够保持对病斑区域的较连续响应, 高响应区域与病害位置具有较好的空间对应关系, 同时部分背景区域的非任务相关响应有所减弱。需要指出的是, 不同类别和样本之间仍存在响应差异, GD-RiceNet 在个别复杂背景样本中仍表现出一定程度的响应扩散。

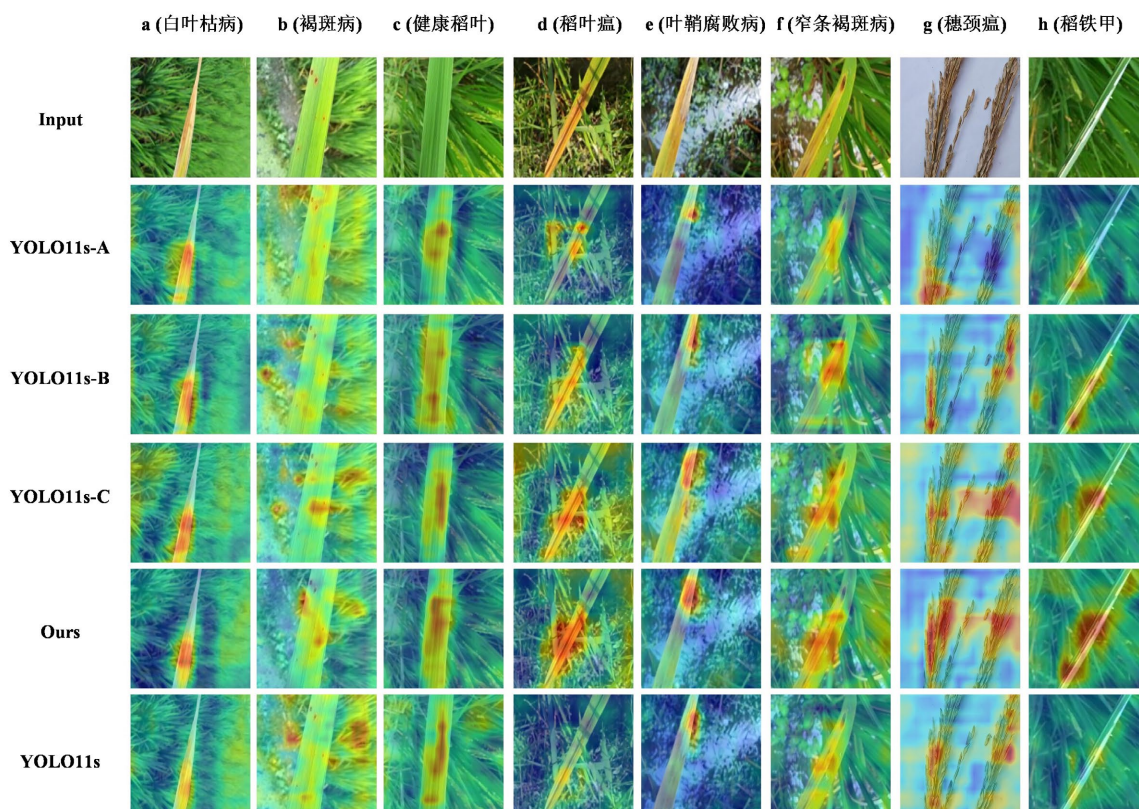


Figure 8. Comparative Grad-CAM visualizations of the baseline model and GD-RiceNet

图 8. 基线模型与 GD-RiceNet 的 Grad-CAM 可视化对比

4.4.2. 检测结果定性分析

为进一步比较不同模型在复杂田间环境下的实际检测表现，选取 5 组具有叶片遮挡、背景干扰、低对比度病斑及多目标共存等特征的代表性测试图像，对 YOLOv8s、YOLOv10s、YOLO11s 和 GD-RiceNet 的检测结果进行可视化，结果如图 9 所示。其中，第一行为真实标注，其余各行分别为不同模型在相同测试条件下的预测结果。

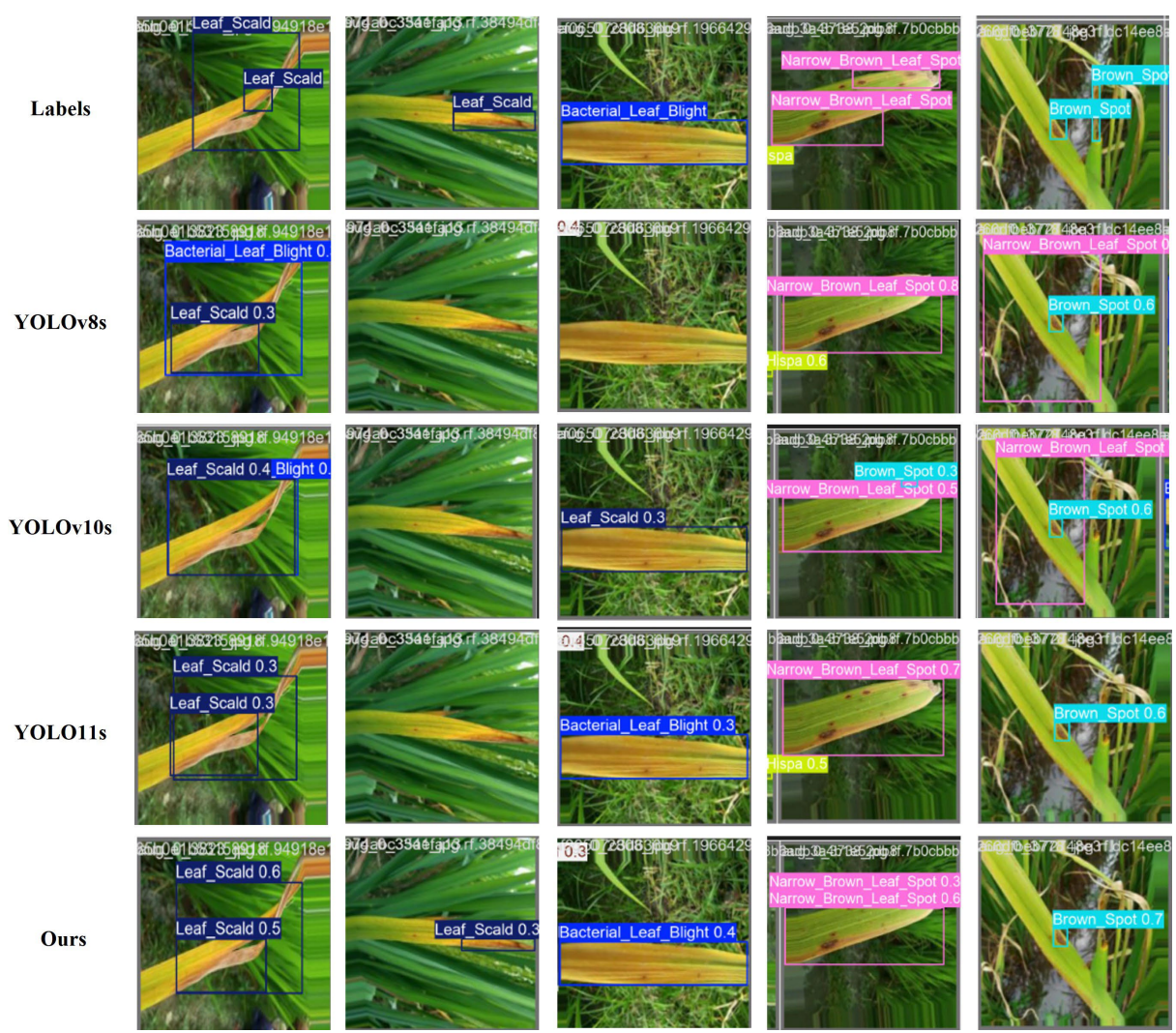


Figure 9. Qualitative comparison of detection results under complex field conditions
图 9. 复杂田间环境下检测结果的定性对比

从图 9 可以看出，不同模型在病害类别判断、目标检出完整性和边界定位方面存在一定差异。在部分多病斑或低对比度样本中，YOLOv8s 和 YOLOv10s 出现了漏检或类别混淆；YOLO11s 能够检出部分主要病害区域，但在个别小面积病斑和复杂背景样本中仍存在漏检。相比之下，GD-RiceNet 在图中前 3 组样本上表现出一定优势：对于同一叶片上的多个叶鞘腐败病区域，模型能够给出与真实标注较为一致的类别预测；对于背景干扰下较不明显的叶鞘腐败病和白叶枯病区域，GD-RiceNet 能够检出部分被其他模型遗漏或误分类的目标，检测框与主要病害区域具有较好的空间对应关系。

与此同时，GD-RiceNet 并未在所有样本上均取得完全正确的检测结果。在后 2 组包含窄褐斑病、稻

铁甲和褐斑病的复杂样本中,模型仍存在个别小目标漏检现象。这说明尽管 GD-RiceNet 改善了部分复杂场景下的病害检出和类别判断能力,但对尺度较小、外观相似或相互邻近的病害目标仍有进一步优化空间。

总体而言,图 9 所展示的代表性样本表明, GD-RiceNet 在部分低对比度病斑、多目标共存及复杂背景场景中具有较好的检测表现,与表 2 中 mAP 指标的整体变化趋势基本一致。

5. 结论

面向复杂田间环境下水稻叶部病害检测任务,本文受水稻叶脉重复纹理特性的启发,为轻量化网络设计提供了明确的任务动机。以 YOLO11s 为基线,构建轻量级检测网络 GD-RiceNet: 骨干网络采用 GhostHGNv2 以降低标准卷积的计算开销,颈部引入 Slim-neck 实现低代价的多尺度特征融合,同时搭配 DySample 动态上采样与 DyHead 动态检测头,以缓解轻量化过程中空间细节信息丢失与特征选择能力不足的问题。

实验结果表明, GD-RiceNet 参数量为 6.56 M, 计算量为 12.7 GFLOPs, mAP@50 和 mAP@50:95 分别达到 57.3%和 35.3%, 在当前测试条件下推理速度达 169 FPS。与 YOLO11s 相比,参数量和计算量分别下降 30.4%和 40.9%, mAP@50 与 mAP@50:95 分别提升 0.6 和 1.4 个百分点。消融实验、对比实验及可视化结果一致表明,该模型在本文数据集与实验设定下实现了检测精度与模型复杂度的良好折中。

基金项目

重庆科技大学硕士研究生创新计划项目“融合动态卷积与全局注意力机制的植物病害检测”(YKJCX2521224)。

参考文献

- [1] Sarfaraz, H. (2023) Rice Production, Quality, and Nutrition: A Comprehensive Review on Challenges, Opportunities, and Global Perspectives. *International Journal of Agriculture and Sustainable Development*, **5**, 90-103.
- [2] Conde, S., Catarino, S., Ferreira, S., Temudo, M.P. and Monteiro, F. (2025) Rice Pests and Diseases around the World: Literature-Based Assessment with Emphasis on Africa and Asia. *Agriculture*, **15**, Article 667. <https://doi.org/10.3390/agriculture15070667>
- [3] Mukherjee, R., Ghosh, A., Chakraborty, C., De, J.N. and Mishra, D.P. (2025) Rice Leaf Disease Identification and Classification Using Machine Learning Techniques: A Comprehensive Review. *Engineering Applications of Artificial Intelligence*, **139**, Article ID: 109639. <https://doi.org/10.1016/j.engappai.2024.109639>
- [4] Zahoor, I. and Mushtaq, A. (2023) Water Pollution from Agricultural Activities: A Critical Global Review. *International Journal of Chemical and Biochemical Sciences*, **23**, 164-176.
- [5] Saleem, M.H., Potgieter, J. and Arif, K.M. (2019) Plant Disease Detection and Classification by Deep Learning. *Plants*, **8**, Article 468. <https://doi.org/10.3390/plants8110468>
- [6] Yao, Q., Guan, Z., Zhou, Y., Tang, J., Hu, Y. and Yang, B. (2009) Application of Support Vector Machine for Detecting Rice Diseases Using Shape and Color Texture Features. 2009 *International Conference on Engineering Computation*, Hong Kong SAR, 2-3 May 2009, 79-83. <https://doi.org/10.1109/icec.2009.73>
- [7] Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S. and Batra, N. (2020) PlantDoc: A Dataset for Visual Plant Disease Detection. *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*, Hyderabad, 5-7 January 2020, 249-253. <https://doi.org/10.1145/3371158.3371196>
- [8] Ren, S., He, K., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/tpami.2016.2577031>
- [9] Liu, J. and Wang, X. (2021) Plant Diseases and Pests Detection Based on Deep Learning: A Review. *Plant Methods*, **17**, Article No. 22. <https://doi.org/10.1186/s13007-021-00722-9>
- [10] Cao, Q., Zhao, D., Li, J., Li, J., Li, G., Feng, S., et al. (2024) Pyramid-YOLOv8: A Detection Algorithm for Precise Detection of Rice Leaf Blast. *Plant Methods*, **20**, Article No. 149. <https://doi.org/10.1186/s13007-024-01275-3>

- [11] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., *et al.* (2021) Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, 10-17 October 2021, 9992-10002. <https://doi.org/10.1109/iccv48922.2021.00986>
- [12] Nugroho, H., Chew, J.X., Eswaran, S. and Tay, F.S. (2024) Resource-Optimized CNNs for Real-Time Rice Disease Detection with ARM Cortex-M Microprocessors. *Plant Methods*, **20**, Article No. 159. <https://doi.org/10.1186/s13007-024-01280-6>
- [13] Li, H., Li, J., Wei, H., Liu, Z., Zhan, Z. and Ren, Q. (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>
- [14] Lin, T., Dollar, P., Girshick, R., He, K., Hariharan, B. and S. Belongie. (2017) Feature Pyramid Networks for Object Detection. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 936-944. <https://doi.org/10.1109/cvpr.2017.106>
- [15] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>
- [16] Kiratiratanapruk, K., Temniranrat, P., Kitvimonrat, A., Sinthupinyo, W. and Patarapuwadol, S. (2020) Using Deep Learning Techniques to Detect Rice Diseases from Images of Rice Fields. In: Fujita, H., Fournier-Viger, P., Ali, M. and Sasaki, J., Eds., *Trends in Artificial Intelligence Theory and Applications*, Springer, 225-237. https://doi.org/10.1007/978-3-030-55789-8_20
- [17] Liu, W., Lu, H., Fu, H. and Cao, Z. (2023) Learning to Upsample by Learning to Sample. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, 1-6 October 2023, 6004-6014. <https://doi.org/10.1109/iccv51070.2023.00554>
- [18] Dai, X., Chen, Y., Xiao, B., Chen, D., Liu, M., Yuan, L., *et al.* (2021) Dynamic Head: Unifying Object Detection Heads with Attentions. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 7369-7378. <https://doi.org/10.1109/cvpr46437.2021.00729>
- [19] Han, K., Wang, Y., Tian, Q., Guo, J., Xu, C. and Xu, C. (2020) GhostNet: More Features from Cheap Operations. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 13-19 June 2020, 1577-1586. <https://doi.org/10.1109/cvpr42600.2020.00165>
- [20] Ma, N., Zhang, X., Zheng, H. and Sun, J. (2018) ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 122-138. https://doi.org/10.1007/978-3-030-01264-9_8
- [21] Baidu PaddlePaddle Vision Team (2023) PP-HGNetV2 (High Performance GPU Network V2). PaddleClas Documentation. https://github.com/PaddlePaddle/PaddleClas/blob/release/2.6/docs/en/models/PP-HGNetV2_en.md
- [22] Nabil, H.R. (2026) Rice Leaf Spot Disease Annotated Dataset. Kaggle. <https://www.kaggle.com/datasets/hadiurrahmannabil/rice-leaf-spot-disease-annotated-dataset>
- [23] Zhao, Y., Lv, W., Xu, S., Wei, J., Wang, G., Dang, Q., *et al.* (2024) DETRS Beat Yolox on Real-Time Object Detection. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 16-22 June 2024, 16965-16974. <https://doi.org/10.1109/cvpr52733.2024.01605>
- [24] Lin, T., Goyal, P., Girshick, R., He, K. and Dollar, P. (2020) Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **42**, 318-327. <https://doi.org/10.1109/tpami.2018.2858826>
- [25] Jocher, G. (2020) YOLOv5 by Ultralytics, Version 7.0. <https://github.com/ultralytics/yolov5>
- [26] Jocher, G., Chaurasia, A. and Qiu, J. (2023) Ultralytics YOLOv8, Version 8.0.0. <https://github.com/ultralytics/ultralytics>
- [27] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., *et al.* (2024) YOLOv10: Real-Time End-to-End Object Detection. *Advances in Neural Information Processing Systems* 37, Vancouver, 10-15 December 2024, 107984-108011. <https://doi.org/10.52202/079017-3429>
- [28] Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D. (2017) Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 618-626. <https://doi.org/10.1109/iccv.2017.74>