

doi: 10.16104/j.issn.1673-1891.2026.01.012

YOLO-SEG: 一种用于水稻病害 实时检测的高性能轻量化模型

王 智, 冯依虎

(亳州学院电子与信息工程系, 安徽 亳州 236800)

摘要:为实现复杂田间环境下水稻病害的快速、精准识别,提出一种基于YOLOv11n改进的高性能轻量化模型YOLO-SEG。该模型通过协同优化,在网络中集成了空间到深度卷积(spatial to depth convolution, SPD-Conv)无损下采样模块以增强微小病斑特征提取,引入高效多尺度注意力(efficient multi-scale attention)模块以聚焦关键特征,并采用轻量级分组混洗卷积(group-shuffle convolution, GSConv)以平衡精度与效率。在包含6种常见病害的自建数据集上,YOLO-SEG模型的平均精度均值(mAP@0.5)达到了95.6%,较基线模型提升了4.4%。同时,模型参数量仅为2.7 MB,推理速度高达120 FPS。以上表明,YOLO-SEG模型在检测精度、模型复杂度和推理速度之间取得了平衡,为智能农业领域的病害实时检测提供了高效、可靠的技术支持。

关键词:水稻病害检测;深度学习;目标检测;YOLOv11;轻量化模型

中图分类号:TP183;TP391.4;S435.111 文献标志码:A 文章编号:1673-1891(2026)01-0103-12

YOLO-SEG: A High-Performance Lightweight Model for Real-Time Detection of Rice Diseases

WANG Zhi, FENG Yihu

(Department of Electronic and Information Engineering, Bozhou University, Bozhou 236800, Anhui, China)

Abstract: To achieve rapid and accurate identification of rice diseases in complex field environments, a high-performance lightweight model named YOLO-SEG is proposed, which is improved based on YOLOv11n. Through synergistic optimization, the model integrates a lossless downsampling module (SPDConv) to enhance the feature extraction of tiny disease spots, introduces an efficient multi-scale attention (EMA) module to focus on key features, and adopts lightweight convolution (GSConv) to balance detection accuracy and inference efficiency. On a self-constructed dataset containing six common rice diseases, the mean average precision (mAP@0.5) of the YOLO-SEG model reaches 95.6%, which is 4.4 percentage points higher than that of the baseline model. Meanwhile, the model has only 2.7 MB of parameters and an ultra-high inference speed of 120 FPS. The results show that the YOLO-SEG model achieves a superior balance among detection accuracy, model complexity and inference speed, and provides an efficient and reliable technical support for the real-time disease detection in the field of smart agriculture.

Keywords: rice disease detection; deep learning; object detection; YOLOv11; lightweight model

收稿日期:2025-10-03

基金项目:安徽省教育厅自然科学重点基金项目(2024AH051298);安徽省高等学校省级质量工程项目(2024xsxx085);亳州学院教育教学改革研究一般项目(2025XJXM056)。

第一作者简介:王智(1997—),男,安徽亳州人,助教,硕士,主要研究方向为应用数学、深度学习。E-mail:zhiwangbzy@163.com。

0 引言

水稻作为全球最重要的粮食作物之一,是维系世界粮食安全与社会稳定的基石^[1]。然而,水稻的稳产与持续增产面临着各类病害的严重威胁。特别是叶部病害,因其具有传播速度快、破坏性强的特点,一旦爆发,常导致作物产量与品质的显著下降,造成巨大的经济损失^[2]。以细菌性枯萎病为例,该病害在流行期间可造成 20%~30% 的产量损失,极端情况下减产幅度可高达 50%,甚至颗粒无收。目前,化学防治虽是控制水稻叶部病害的主要手段,但化学农药的过量施用已引发了对生态环境安全的广泛担忧。因此,发展快速、精准的水稻叶部病害识别技术,以实现病害的早期预警与靶向干预,对于保障水稻丰产、降低经济损失、推动农业绿色可持续发展,具有至关重要的理论意义与应用价值。

当前,依赖植保专家经验的水稻病害诊断方法主观性强、效率低下,难以满足现代农业大规模、实时监测的需求,常因诊断延迟而错失最佳防治窗口^[3]。为克服此局限性,研究者们尝试应用传统计算机视觉与机器学习技术进行自动识别^[4]。这类方法的核心思路是:通过颜色、纹理、形态等人工设计的特征来区分病斑与健康组织,再利用支持向量机、决策树等分类器进行识别^[5]。然而,其性能瓶颈在于对人工设计特征的严重依赖。这些特征在面向真实、复杂的田间环境时,其判别力与泛化能力存在固有局限。面对光照波动、植株摇曳、叶片遮挡、多病并发等动态多变的场景,手工特征难以稳定、准确地提取出高区分度的病害信息,导致较高的漏检与误报率,使该技术路线在实际应用中面临巨大挑战。与传统方法相比,深度学习凭借其卓越的特征自学习与泛化能力,已在农业病害识别领域展现出显著优势并得到广泛应用^[6-9]。Pan 等^[10]构建了一个 2 阶段识别框架,先通过 YOLOX 模型提取

病斑特征,再利用孪生网络进行分类,有效解决了小样本及复杂背景下的识别难题,平均精度均值(mean average precision, mAP)@0.5 达到了 95.58%。Jia 等^[11]则改进了 YOLOv7 模型,采用轻量化的 MobileNetV3 压缩模型参数,并融合坐标注意力机制与 SIoU 损失函数,将 mAP@0.5 提升至 93.7%。为解决深层网络中的梯度消失问题,Stephen 等^[12]在预训练的 ResNet 架构中集成了自注意力机制,显著强化了模型的特征选择能力,其中 ResNet34 模型的识别准确率高达 98.54%,展现了优越的多类别分类性能。这些研究为基于深度学习的病害智能检测奠定了坚实基础。

基于深度学习的植株病害识别性能在很大程度上依赖于模型对关键特征的高效学习。然而,水稻病害识别任务面临诸多挑战:叶片细长的形态增加了小尺寸病斑特征的提取难度;病害类型繁多且病征高度相似,使得模型的特征辨识能力受到严峻考验。尽管通过加深网络层数可以提升识别精度,但这往往伴随计算量和参数规模的大幅增加,从而限制了模型在移动端的实时应用。针对这些问题,研究者们提出了多种结构改进与优化方案。首先,在微小目标特征提取方面,传统下采样会丢失细粒度信息,对检测微小病斑极为不利。为解决此问题,Wu 等^[13]将空间到深度卷积(spatial to depth convolution, SPDCConv)^[14]引入 YOLOv7,它通过像素重排保留所有原始特征信息,显著提升了模型对低分辨率图像与小目标的检测性能;其次,在注意力机制方面,为了让模型聚焦于关键病害特征、抑制复杂背景噪声,多种注意力模块被集成到检测网络中。Jiang 等^[15]将挤压与激励(squeeze-and-excitation, SE)^[16]模块嵌入到 DenseNet 网络中,提出了 3 种不同的嵌入方式。通过在特征通道维度上进行权重自适应,SE 模块显著增强了 DenseNet 对水稻病害特征的提取能力。Li 等^[17]将卷积块注意力模块(convolutional block attention module, CBAM)^[18]集

成于 YOLOv5s 的颈部网络,通过在通道与空间维度上进行特征的自适应聚焦,有效提高了对百香果害虫的检测与识别率。Lan 等^[19]则将高效多尺度注意力(efficient multi-scale attention, EMA)^[20]模块嵌入 YOLOv5 主干网络的第一个 C3 模块中形成 C3-EMA 结构,旨在强化模型对水稻穗边缘等细节特征的提取能力。然而,部分注意力模块在提升性能的同时也带来了计算开销。Wang 等^[21]在 YOLOv5 的颈部网络中引入了分组混洗卷积(group-shuffle convolution, GSConv),在保证不同特征层之间的融合能力的同时,可以最大限度地减少模型参数和计算的数量。

综上所述,本文为应对水稻病害检测中普遍存在的背景复杂、小目标特征提取困难及模型轻量化与高精度难以兼顾的挑战,提出一种基于 YOLOv11n 框架的改进型高性能检测模型,命名为 YOLO-SEG。该模型通过一系列协同的结构优化实现了综合性能的提升。具体而言,在主干与颈部网络中引入 SPDCConv 模块,以无损的方式增强模型对低分辨率图像和小尺寸目标的特征提取与融合能

力。同时,采用 EMA 模块替换原有 C2PSA 模块,以更优的效率聚焦于关键特征信息。最后,为平衡精度与效率,颈部网络集成了轻量级卷积 GSConv,在保证性能的同时显著降低了计算与参数复杂度,从而全面地优化了检测性能。

1 数据集的构建

1.1 图像采集

本研究以 6 种常见的水稻叶、穗部病害为研究对象,包括白叶枯病(bacterial blight)、条斑病(leaf streak)、稻瘟病(blast)、稻曲病(false smut)、纹枯病(sheath blight)及东格鲁病(tungro)。研究所用图像数据集于 2024 年 3—9 月,在位于安徽省阜阳市颍上县的国家现代农业产业园基地完成采集。采集设备为智能手机,共计获取病害图像样本 2 966 幅,原始分辨率为 4 624×3 472 像素。图 1 展示了自建数据集中的部分病害样本。为降低后续网络训练过程中的内存与显存开销,所有图像尺寸都被统一预处理至 640×640 像素。



图 1 自制水稻病害数据集部分样本

1.2 数据集预处理

为提升模型的鲁棒性与泛化能力,并缓解原始数据集中可能存在的类别不均衡问题,本文应用了多种数据增强技术以扩充训练数据集。这些技术

主要包括几何变换、光学变换、椒盐噪声及随机擦除。部分增强效果的示例如图 2 所示。

为构建高质量的病害图像数据集,对数据增强后的图像进行筛选,在剔除信息严重丢失或无效的

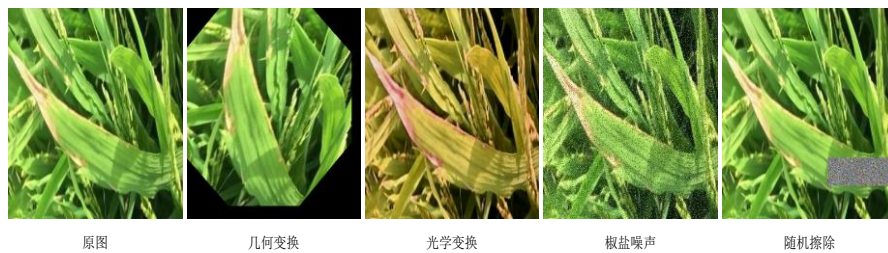


图2 数据增强效果

样本后,获得了一个包含 20 318 幅图像的集合。本文采用 X-AnyLabeling 工具为所有图像标注了目标边界框,并按照 8:1:1 的比例将该数据集划分为训练集、验证集和测试集,其详细的样本数量分布如表 1 所示。

表 1 水稻病害数据集构成

水稻病害	样本总数	训练集 样本数	验证集 样本数	测试集 样本数
白叶枯病	2 547	2 037	255	255
条斑病	3 101	2 487	310	310
稻瘟病	3 000	2 400	300	300
稻曲病	5 239	4 191	524	524
纹枯病	2 142	1 714	214	214
东格鲁病	4 289	3 431	429	429

2 水稻病害检测模型设计

2.1 模型结构概述

单次检测目标检测算法(you only look once, YOLO)系列是先进的单阶段目标检测框架,而 YOLOv11n 作为其极致轻量化的代表,专为在移动或嵌入式设备上实时检测而设计。其网络架构由骨干网络、颈部和检测头 3 部分构成,能高效地完成从特征提取、融合到目标解析的全过程。尽管 YOLOv11n 速度优势显著,但其面对复杂的田间水稻病虫害图像时也暴露出局限性。这些图像普遍存在背景干扰严重、不同病害视觉特征相似,以及同一病害因时期、光照不同而形态尺寸多变等问题。YOLOv11n 的轻量化网络在下采样时会损失部分细节纹理,这对于准确识别微小病斑和早期病害等小

目标构成了严峻挑战。

为弥补上述不足,本文以 YOLOv11n 为基准,提出了一种改进的高性能轻量化水稻病害检测模型 YOLO-SEG,结构如图 3 所示。该模型的核心目标是在维持轻量化特性的同时,对网络关键组件进行针对性强化,以提升对微小目标的特征捕获能力。首先,在主干网络层中,引入 SPDCConv 模块替换标准下采样层,以从源头保留细粒度特征,并嵌入 EMA 注意力机制来聚焦关键病征、抑制背景噪声。其次,在颈部网络中,使用 SPDCConv 模块与 GSCConv 模块,旨在增强多尺度特征融合的同时,有效平衡性能与计算开销。

2.2 模型结构优化

2.2.1 SPDCConv 模块

针对自然环境下水稻早期病害呈微小病斑状,其细粒度特征在传统下采样中易丢失的检测难题,引入 SPDCConv 模块作为下采样工具。其核心在于空间到深度(space-to-depth, SPD)的无损转换机制,整个过程如图 4 所示。对于一个形状为 $S \times S \times C_1$ 的特征图 X ,SPD 首先根据 N_{scale} 因子对其进行切片重组。这一切片操作表达式如式(1)所示。

$$f_{x,y} = X[x:S:N_{scale}, y:S:N_{scale}] \quad (1)$$

式中: x 和 y 的取值范围为 0 到 $N_{scale} - 1$, N_{scale} 为重排因子。

该切片操作生成了一系列子特征图,这些子图随后在通道维度上进行拼接,从而实现了从原始特征图到新特征图的转换。这一转换的本质,是将特征图的空间维度与通道深度进行置换:其形状从

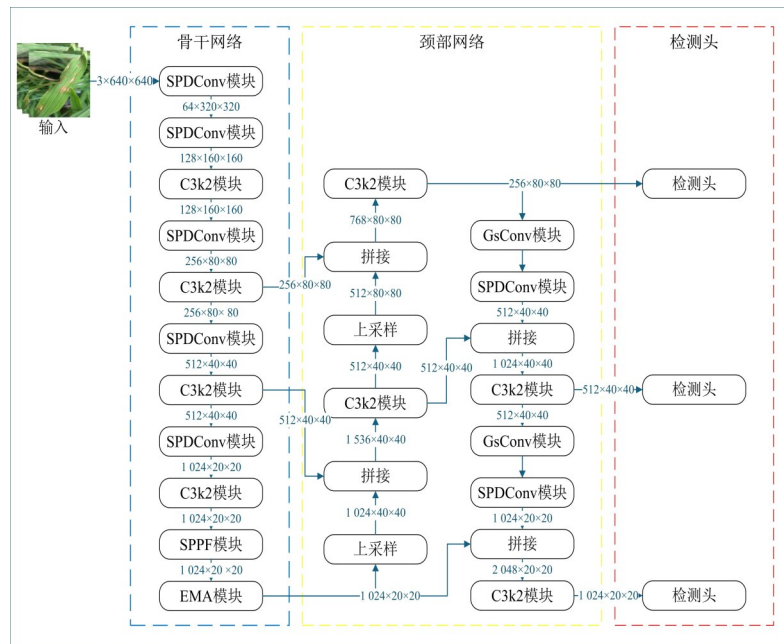


图 3 YOLO-SEG 结构

$S \times S \times C_1$ 变为 $(\frac{S}{N_{scale}}, \frac{S}{N_{scale}}, N_{scale}^2 \times C_1)$, 在缩小空间分辨率的同时, 确保了原始像素信息的完整性。

无步长卷积作为 SPD 模块的后续处理单元, 旨在对 SPD 转换后产生的高维特征图进行学习及降维。具体而言, 该卷积层采用 1 的步长和 C_2 ($C_2 < N_{scale}^2 \times C_1$) 个滤波器, 旨在对 SPD 层输出的高维特

征图 $(\frac{S}{N_{scale}}, \frac{S}{N_{scale}}, N_{scale}^2 \times C_1)$ 进行处理。它将该特征图映射为一个更为紧凑的新特征表示 $(\frac{S}{N_{scale}}, \frac{S}{N_{scale}}, C_2)$, 并使其蕴含更丰富的语义信息。这一操作既能从重排的通道中有效学习特征, 又能显著降低特征图的通道数, 为后续网络层提供计算成本更低的输入。

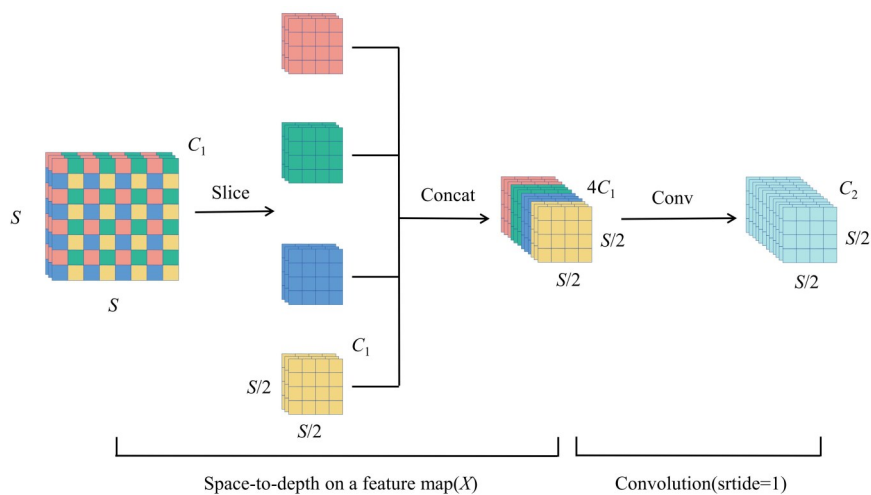


图 4 $N_{scale} = 2$ 时 SPDCConv 结构

2.2.2 EMA 模块

EMA 模块的设计原则, 是在不进行通道降维的前提下, 高效捕获多尺度与跨维度的特征依赖关

系, 这一特性可以有效解决水稻病害识别中细粒度特征易丢失与微小病斑难检测的问题。该模块的详细结构如图 5 所示。

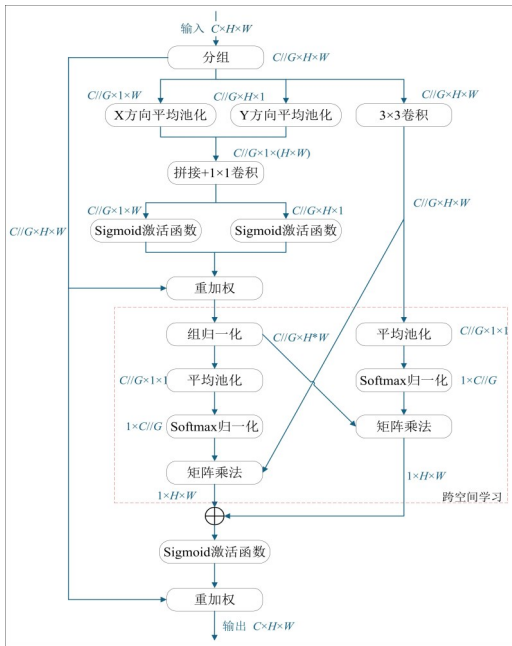


图5 EMA 结构

EMA 模块首先对输入特征图进行分组处理。具体而言,对于一个给定的输入特征 $X \in \mathbb{R}^{C \times H \times W}$,该模块会沿着通道维度,将其均匀地划分为 $G(G \ll C)$ 个子特征组,旨在从不同角度学习和捕获多样的语义特征。该分组过程如式(2)所示。

$$X = [X_0, X_1, \dots, X_{C/G-1}] \quad (2)$$

式中:每个子特征组 $X_i \in \mathbb{R}^{C/G \times H \times W}$ 。

接着,通过一个包含 3 条并行路径的子网络来提取分组特征的注意力权重。其中 2 条并行路径位于 1×1 卷积分支中,第 3 条路径则是 3×3 卷积分支。为了在捕获跨通道依赖的同时降低计算开销, 1×1 分支首先沿空间的高度和宽度维度分别进行一维全局平均池化,将空间信息编码为 2 个独立的特征向量。随后,这 2 个向量沿宽度方向拼接,并共同送入一个共享的 1×1 卷积层。该卷积层的输出被分解并经过 2 个并行的 Sigmoid 函数激活,以拟合出一个二维的注意力分布。最后,通过逐元素相乘来聚合这 2 条路径的注意力图,从而实现高效且复杂的跨通道特征交互。另一方面, 3×3 分支通过一个 3×3 卷积来捕获局部的跨通道交互,以扩展特征空间。

接下来是跨空间学习阶段,该阶段旨在融合并行子网络中 1×1 分支与 3×3 分支的输出信息,以生

成 2 种互补的空间注意力图。首先,为了捕获全局上下文和长距离依赖关系,对 1×1 分支的输出应用二维全局平均池化。该操作将整个空间信息编码为一个特征向量,其计算公式如式(3)所示。

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (3)$$

式中: z_c 为对第 c 个通道进行全局平均池化得到的输出, $x_c(i, j)$ 为特征图在第 c 个通道、空间位置 (i, j) 处的取值, H 和 W 分别表示特征图的高度和宽度。

池化后的特征经过 Softmax 函数处理,以高效地生成一个近似二维高斯分布的全局注意力图。通过将该图与 3×3 分支的输出特征图进行点积运算,便得到了第 1 个空间注意力图。与此同时,同样利用二维全局平均池化来编码 3×3 分支的全局空间信息,并与 1×1 分支的输出进行点积运算,便得到了第 2 个空间注意力图,它完整地保留了特征的精细空间结构。最终,每组输出特征图通过聚合 2 个生成的空间注意力权重值并经 Sigmoid 函数处理后得出。该方法能捕捉像素级成对关系,并突出所有像素的全局上下文信息。EMA 通过建模像素间复杂空间依赖,生成高精度注意力图,实现对不规则微小目标的精准聚焦,从而显著提升检测性能。

2.2.3 GSConv 模块

在深度学习网络中,尤其是轻量化模型设计中,计算效率与精度始终难以兼顾。标准卷积具备强大的特征表达能力,却因计算量庞大影响推理速度;深度可分离卷积虽能显著降低计算成本,但削弱了通道间信息交互,导致精度下降。为平衡两者,GSConv 模块被提出,其在保持高效率的同时,尽可能保留标准卷积的特征提取能力,其结构示意图如图 6 所示。

GSConv 模块的整体流程旨在高效地融合不同卷积操作所提取的特征。首先,一个具有 C_1 个通道的输入特征图经过一个标准的卷积模块,该模块由一个二维卷积层、一个二维批量归一化层和一个激活函数构成,生成具有 $C_2/2$ 个通道的中间特征。随

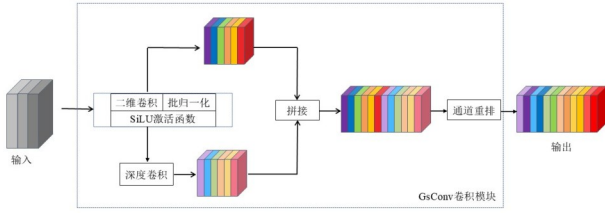


图6 GSConv 结构

后,该中间特征被送入2个并行的分支进行处理:其中一个分支直接将该特征保留,作为主干信息;另一个分支则通过一个深度可分离卷积操作,以更低的计算成本提取另一组 $C_2/2$ 通道的特征。接着,来自2个分支的特征图在通道维度上进行拼接形成一个具有 C_2 个通道的融合特征图。最后,为了促进2个分支特征之间的信息交互与流通,对拼接后的特征图执行一个通道混洗操作,最终得到该模块的输出结果。

3 实验验证

3.1 实验配置

本文所有实验均在统一配置的计算机平台上进行。硬件配置为 Intel (R) Core (TM) i5-1135G7 CPU (主频 2.40 GHz)、NVIDIA GeForce RTX3090 GPU,操作系统为 Windows10(64位)。软件环境方面,深度学习框架采用 PyTorch2.2.2,开发环境为 PyCharm2024.2.3,编程语言为 Python3.10,并使用 CUDA12.1 进行 GPU 加速。在模型训练阶段,所有网络的输入图像尺寸统一调整为 640×640 像素。训练总轮次(epoch)设置为 300,批处理大小(batch size)为 32。模型优化方面,采用随机梯度下降(SGD)优化器进行训练,初始学习率设置为 0.01。

3.2 评价指标

为对本文提出的目标检测模型的性能进行全面、客观的量化评估,采用精确率(precision, P)、召回率(recall, R)及 mAP 为评价指标,其计算公式如式(4)~(6)所示。

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (4)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (5)$$

$$N_{mAP} = \frac{1}{N} N_{AP_i} \quad (6)$$

式(4)~(6)中: N_{TP} 表示真实为正类且被模型预测为正类的样本数, N_{FP} 表示真实为负类但被模型预测为正类的样本数, N_{FN} 为真实为正类但被模型预测为负类的样本数, N 为类别总数, N_{AP_i} 为第 i 个类的平均精度。

mAP 通常与一个或多个交并比(IoU)阈值一同呈现。例如,mAP@0.5 指在 IoU 阈值为 0.5 时计算的 mAP。 N_{mAP} 为 mAP 的数值。参数量(parameters, P)是指模型中所有需要通过训练学习的权重和偏置的总和。帧率(frames per second, FPS)是指模型每秒可以处理的图像帧数。这是评估实时性能最直观、最常用的指标。高 FPS 值意味着更流畅、更及时的诊断能力。

3.3 实验结果分析

3.3.1 不同模型检测结果对比

为了对本文提出的 YOLO-SEG 模型的性能进行客观的量化分析,采用严格一致的实验环境与参数配置,将其与当前具有代表性的 Faster R-CNN、YOLOv10n、YOLOv11n、YOLOv12n 以及 YOLOv13n 模型进行了深入的对比实验。图 7 展示了所有模型在训练过程中的 mAP@0.5 收敛曲线。从图中可以观察到,本研究的 YOLO-SEG 模型在训练初期便展现出极为迅猛的精度爬升速率;更为重要的是,在训练 60 轮后,YOLO-SEG 的 mAP@0.5 值开始超越其他对比模型,最终收敛于一个远超其他模型的精度点,这充分证明了本文提出的模型在精度性能上的优越性。

在最终的检测性能上,改进后的 YOLO-SEG 模型表现优异(表 2)。其精确率、召回率及 mAP@0.5 分别达到了 90.1%、90.3% 和 95.6%。与作为其改进基线的 YOLOv11n 模型相比,其核心指标全面提升;同时,与其他基线模型相比,其性能优势也同样突

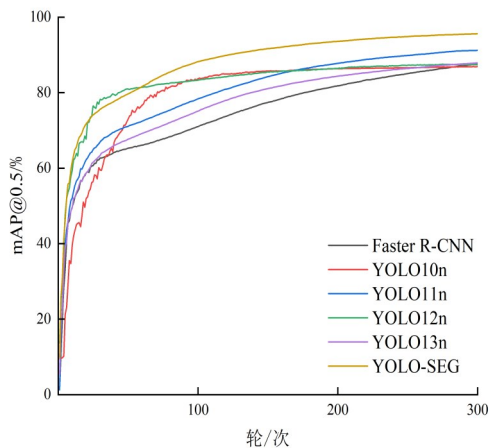


图 7 不同模型检 mAP@0.5 变化曲线

出。例如 mAP@0.5 值较 Faster R-CNN、YOLOv10n、YOLOv12n 和 YOLOv13n 分别高出 8.0%、8.7%、8.2% 和 7.7%，充分证明了改进策略的优越性。此外，在模型轻量化方面，YOLO-SEG 在实现性能跃升的同时，也保持了与基线模型相当的复杂度。其参数量仅从 2.6 MB 小幅增加至 2.7 MB，帧率也维持在 120 FPS 的高水平，与 YOLOv11n 几乎持平。综合来看，上述定量比较结果有力地证明了 YOLO-SEG 模型在检测精度、模型参数量和推理速度之间取得了平衡，其综合性能不仅超越了基线模型，也优于当前其他主流模型。

表 2 不同模型的性能比较

模型	精确率/%	召回率/%	mAP@0.5/%	参数量/MB	帧率/FPS
Faster R-CNN	89.2	88.7	87.6	41.3	13
YOLOv10n	87.3	86.6	86.9	2.3	125
YOLOv11n	88.9	87.8	91.2	2.6	122
YOLOv12n	88.3	87.6	87.4	2.6	121
YOLOv13n	86.3	85.8	87.9	2.5	124
YOLO-SEG	90.1	90.3	95.6	2.7	120

由于 Faster R-CNN 参数量较大、推理帧率较低，难以满足模型轻量化与实时监测的需求，因此本文选取 YOLOv10n、YOLOv11n、YOLOv12n、YOLOv13n 及本文提出的 YOLO-SEG 模型，对稻瘟病、白叶枯病、稻曲病、纹枯病、条斑病和东格鲁病 6 种

水稻病害进行了实际检测对比，其结果如图 8 所示。

从图 8 中可以看出，在面对形态细长、与叶片背景高度相似的病害(如白叶枯病、条斑病)时，YOLOv10n、YOLOv11n、YOLOv12n、YOLOv13n 等对比模型普遍表现出定位不精确和识别断裂的问题。它们的检测框往往是碎片化的，无法形成一个连续、完整的矩形框覆盖整个条状病变区域，这导致了对病害范围的严重低估和大量的冗余检测。同样，在检测稻瘟病和稻曲病这类尺寸小、分布密集的病斑时，这些对比模型常常只能识别出其中几个特征较为明显的病斑，而对大量处于初期阶段或被叶片轻微遮挡的小病斑则完全忽略，造成了较高的漏检率。

相比之下，本文改进的 YOLO-SEG 模型在所有 6 种病害的检测任务中，均展现出了一致的优越性能和强大的泛化能力。对于白叶枯病和条斑病，YOLO-SEG 能够生成更宏观、更贴合病害实际蔓延形态的检测框，显著提升了检测的准确性和完整性。在处理稻瘟病和稻曲病这类小而密集的病斑时，YOLO-SEG 依然能够实现精准的捕获，几乎没有遗漏，大幅提升了对早期微小病灶的检出能力。综上所述，无论是检测框的质量还是对复杂场景下各类目标的召回率，YOLO-SEG 的整体表现均超过其他对比模型，其检测结果与病害的真实情况最为吻合。

3.3.2 消融实验

为系统验证本文针对水稻叶部与穗部病害识别所设计的 YOLO-SEG 模型中 SPDCConv、EMA 及 GSConv 3 个组件的有效性，基于自制的水稻病害数据集样本数据分别对 YOLOv11n、YOLO-S (YOLOv11n+SPDCConv)、YOLO-E (YOLOv11n+EMA)、YOLO-G (YOLOv11n+GSConv)、YOLO-SE (YOLOv11n+SPDCConv+EMA)、YOLO-SG (YOLOv11n+SPDCConv+GSConv)、YOLO-EG (YOLOv11n+EMA+GSConv) 和 YOLO-SEG 这 8 种模型进行性能测试，结果如表 3 所示。

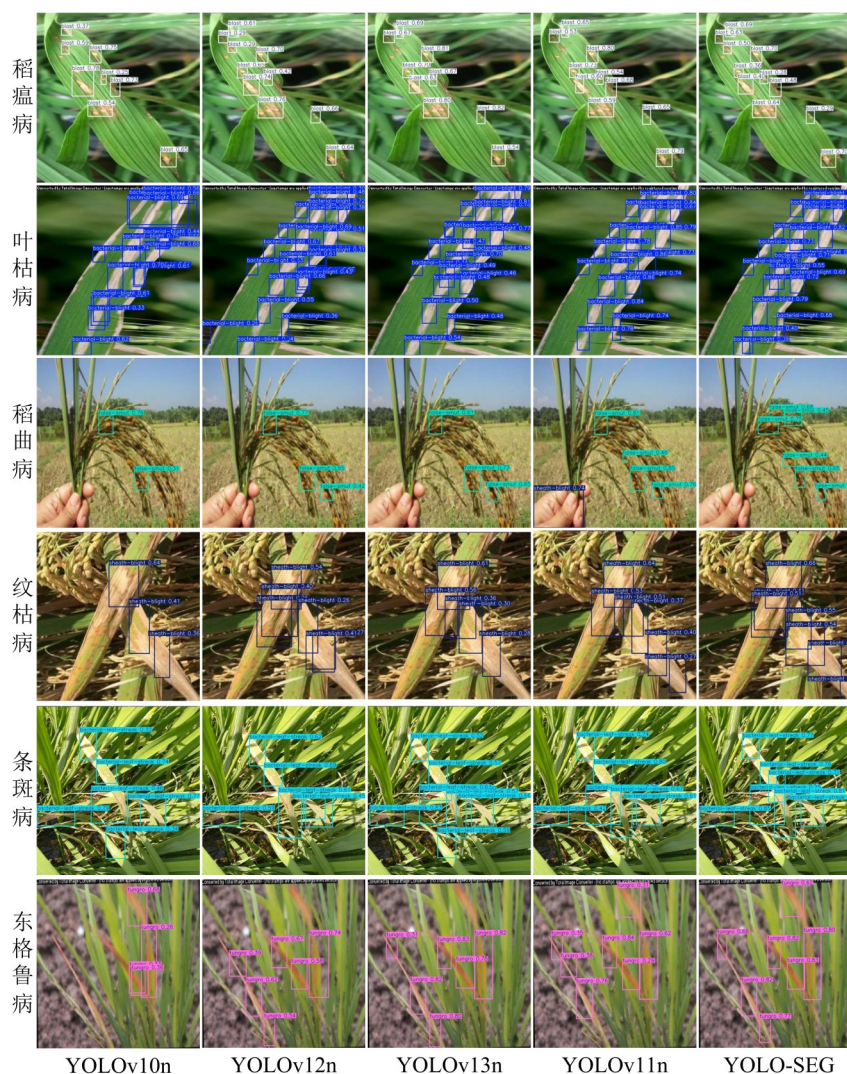


图 8 不同模型检测结果对比结果

表 3 不同模块组合的模型性能参数

模型	精确率/%	召回率/%	mAP@0.5/%	参数量/MB	帧率/FPS
YOLOv11n	88.9	87.8	91.2	2.6	122
YOLO-S	88.5	86.7	93.1	3.0	115
YOLO-E	89.4	88.9	94.6	2.1	128
YOLO-G	85.8	85.2	90.8	2.1	130
YOLO-SE	89.0	87.8	93.9	2.8	119
YOLO-SG	87.2	86.0	91.9	2.6	121
YOLO-EG	87.6	87.1	92.7	2.2	126
YOLO-SEG	90.1	90.3	95.6	2.7	120

由表 3 可知,本文所提出的不同模块在集成至基线模型后所带来的独特效果与协同潜力。具体

而言,SPDConv 模块主要贡献了 1.9% 的 mAP 显著增长,但随着参数量的增加与帧率下降,表明其专注于提升检测精度,但牺牲了一定的模型轻量化与推理速度。引入 EMA 注意力模块后,模型性能获得了最为显著的提升。得益于其高效的特征增强能力,与基准 YOLOv11n 相比,YOLO-E 的精确率、召回率和 mAP 分别提升了 0.5%、1.1% 和 3.4%,同时将参数量降低了 0.5 MB 并实现了 6 FPS 的帧率提升,展现了该模块在提升精度的同时优化模型复杂度和运行效率的卓越能力。GsConv 模块作为一种轻量化组件,其集成使模型参数量成功减少 0.5 MB,并将推理速度提升了 8 FPS,是所有模型中效率最高的。然而,此优化的代价是精确率、召回率和

mAP的显著下降,这明确了其在效率与精度之间进行权衡的设计定位。

在探索模块间的协同作用时,双模块组合模型的结果提供了更深刻的协同潜力。YOLO-SE模型结合了SPDConv与EMA,在mAP上取得了2.7%的显著增益。然而,这是以参数量增加0.2 MB及帧率下降3 FPS为代价的,表明该组合策略侧重于最大化检测精度,但在模型效率上有所妥协。而YOLO-EG则展现了最为均衡的协同效应,该模型在EMA模块的驱动下,实现了1.5%的mAP提升,并且得益于GsConv,其参数量和推理速度均优于基线模型,成功地在提升核心性能的同时,强化了模型的效率优势。相比之下,YOLO-SG的组合揭示了复杂权衡,GsConv的引入虽维持了与基线持平的参数量且帧率仅有微小下降,但SPDConv带来的mAP增益被削弱,说明单纯叠加可能导致特征辨别力的损失。

最终,集成了全部3种模块的YOLO-SEG模型,展现了高效的设计协同效应。该设计的精妙之处在于,GsConv的轻量化特性有效对冲了其他增强模块带来的参数增长与速度损耗。与基线模型YOLOv11n相比,YOLO-SEG在参数量仅增加0.1 MB、帧率基本持平的情况下,实现了精确率、召回率及mAP分别为1.2%、2.5%与4.4%的全面提升,成功验证了多模块集成策略在达成模型高精度、高效率与轻量化三者间平衡的有效性。

为深入评估与量化YOLO-SEG模型对不同病害的检测能力,本文具体分析了各病害类别的精确率-召回率(P-R)曲线(图9)及关键性能指标(表4)。P-R曲线下方的面积即为平均精度(AP),是衡量模型对单一类别检测性能的核心指标,而所有类别的AP均值(mAP)则反映了模型的整体稳健性。本次实验的AP值均在交并比阈值为0.5的条件下计算。

由图9可知,模型展现了良好的综合性能,整体平均精度(mAP@0.5)达到0.956。各病害的P-R曲线形态饱满且均趋向于图的右上角,这有力地证明

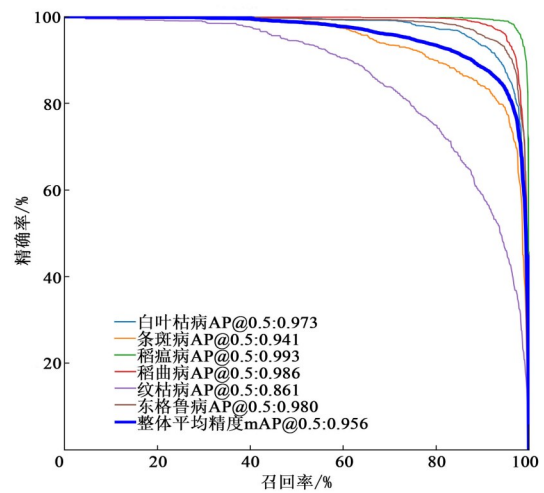


图9 YOLO-SEG模型P-R曲线

了模型在保持高精确率的同时也能实现高召回率。其中,稻瘟病、稻曲病和东格鲁病的曲线最为突出,其AP值均超过0.98,表明模型对这些病害的识别达到了极高的水准。

表4 不同病害检测性能

病害类型	精确率/%	召回率/%	AP@0.5/%
白叶枯病	91.6	92.2	97.3
条斑病	84.1	89.6	94.1
稻瘟病	96.2	97.9	99.3
稻曲病	94.1	95.6	98.6
纹枯病	82.6	71.4	86.1
东格鲁病	92.0	95.2	98.0

由表4可知,YOLO-SEG模型在各类病害检测中整体表现优异,与P-R曲线所呈现的趋势一致,稻瘟病的精确率和召回率均位居前列,体现了模型对该病害最强的检测能力。然而,在纹枯病和条斑病2类病害的检测中,模型性能相对较弱。特别是纹枯病,其召回率仅为71.4%,精确率为82.6%,均为所有类别中的最低值,导致平均精度(AP)仅为86.1%,明显低于其他类别。这一差距与P-R曲线中纹枯病曲线靠内、未能紧贴右上角的分布特征完全一致。

为深入剖析这一现象,本文借助类激活图技术对模型在决策过程中的特征关注区域进行了可视化

分析,结果如图 10 所示。由图 10 可知,YOLO-SEG 模型能够准确捕捉多数病害的核心视觉特征,但在识别形态不规则、纹理模糊的病害时,其注意力聚焦存在不足。具体而言,对于稻瘟病、白叶枯病和稻曲病等形态清晰、对比度明显的病害,模型的激活区域(高亮区域)能够完整覆盖主要病斑区域,表明模型已有效学习到这些类别的判别性特征。相反,在纹枯病和条斑病样本中,热力图的激活区域分布较为弥散,且未能紧密贴合叶片表面的实际病斑边界。造成该问题的主要原因包括:数据层面上,纹枯病和条斑病样本数量相对较少,且在光照条件、拍摄角度和病斑形态上存在较大差异,导致模型学习到的特征分布不均衡;模型层面上,当前的 EMA 注意力机

制在复杂纹理区域的特征聚焦能力仍不足,对边界模糊的病斑表现欠佳;训练层面上,现有的数据增强策略对这两类病害的特征变换覆盖不充分,未能充分模拟其真实的田间变异情况。

针对上述不足,未来可从以下方向改进:优化数据增强策略,在纹理扰动和仿生模糊模拟等方面进行针对性增强,以提升模型对复杂环境的适应性;改进注意力机制,可引入具有自适应特征加权能力的混合注意力模块,强化模型对不规则边界与细粒度纹理的响应能力;采用类别平衡与特征重采样策略,在训练过程中提升弱类别样本的权重,减少模型对主导类别的过拟合,从而提升整体检测均衡性与鲁棒性。

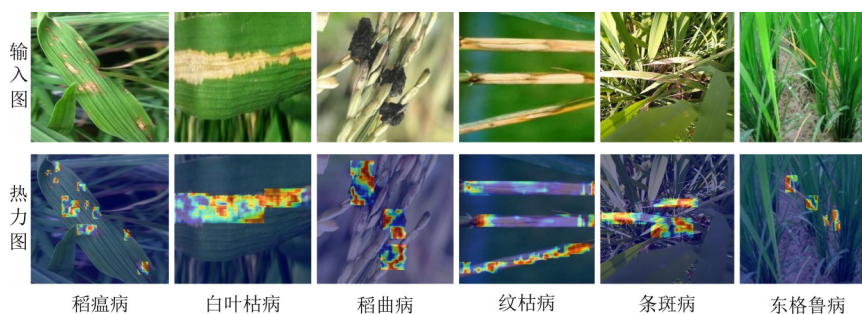


图 10 YOLO-SEG 水稻病害检测

4 结束语

本文为解决复杂田间环境下水稻病害检测的难题,提出了一种改进的轻量化模型 YOLO-SEG。该模型通过协同集成 SPDConv、EMA 及 GSConv 模块,显著提升了对微小病害的检测能力,并成功地在检测精度、模型大小与推理速度之间取得了平衡。实验证明,YOLO-SEG 在 mAP@0.5 达到 95.6% 的同时,参数量仅为 2.7 MB,帧率高达 120 fps。在 6 类主要病害的检测中,模型均取得出色表现,其中稻瘟病、稻曲病与东格鲁病的平均精度超过 98%,

体现出良好的精度与泛化能力。模型结构紧凑、推理高效,具备在移动端与无人机等边缘设备上实时部署的潜力。然而,YOLO-SEG 在识别纹枯病等形态不规则、纹理模糊的病害时仍存在局限,主要受边界模糊与特征聚焦不足影响。未来研究将重点引入 Transformer 结构与多模态特征融合机制,以提升模型对复杂背景和不规则病斑的表征能力;同时结合半监督学习与增量学习扩充数据规模,并开展田间实地部署与系统集成测试,构建可嵌入农业监测终端的智能病害诊断系统,实现病害识别与防治决策的智能一体化应用。

参考文献:

- [1] 朱梦远,杨红兵,李志伟.高光谱图像和叶绿素含量的水稻纹枯病早期检测识别[J].光谱学与光谱分析,2019,39(6):1898-1904.
- [2] 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.

- [3] ZHANG Z, YANG Y, XU X, et al. GVC-YOLO: a lightweight real-time detection method for cotton aphid-damaged leaves based on edge computing[J]. *Remote Sensing*, 2024, 16(16): 3046.
- [4] PANTAZI X E, MOSHOUD, TAMOURIDOU A A. Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers[J]. *Computers and Electronics in Agriculture*, 2019, 156: 96-104.
- [5] CHOUHAN S S, SINGH U P, JAIN S. Automated plant leaf disease detection and classification using fuzzy based function network[J]. *Wireless Personal Communications*, 2021, 121(3): 1757-1779.
- [6] FERENTINOS K P. Deep learning models for plant disease detection and diagnosis[J]. *Computers and electronics in agriculture*, 2018, 145: 311-318.
- [7] NAGARAJU M, CHAWLA P, UPADHYAY S, et al. Convolution network model based leaf disease detection using augmentation techniques[J]. *Expert Systems*, 2022, 39(4): e12885.
- [8] ARUN R A, UMAMAHESWARI S. Effective multi-crop disease detection using pruned complete concatenated deep learning model[J]. *Expert Systems with Applications*, 2023, 213: 118905.
- [9] LI B, TANG J, ZHANG Y, et al. Ensemble of the deep convolutional network for multiclass of plant disease classification using leaf images[J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2022, 36(4): 2250016.
- [10] PAN J, WANG T, WU Q. RiceNet: A two stage machine learning method for rice disease identification[J]. *Biosystems Engineering*, 2023, 225: 25-40.
- [11] JIA L, WANG T, CHEN Y, et al. MobileNet-CA-YOLO: an improved YOLOv7 based on the MobileNetV3 and attention mechanism for rice pests and diseases detection[J]. *Agriculture*, 2023, 13(7): 1285.
- [12] STEPHEN A, PUNITHA A, CHANDRASEKAR A. Designing self attention-based ResNet architecture for rice leaf disease classification[J]. *Neural Computing and Applications*, 2023, 35(9): 6737-6751.
- [13] WU W, ZHANG J, ZHOU G, et al. Esg-yolo: a method for detecting male tassels and assessing density of maize in the field [J]. *Agronomy*, 2024, 14(2): 241.
- [14] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects[C]//Joint European conference on machine learning and knowledge discovery in databases. Cham: Springer Nature Switzerland, 2022: 443-459.
- [15] JIANG M, FENG C, FANG X, et al. Rice disease identification method based on attention mechanism and deep dense network [J]. *Electronics*, 2023, 12(3): 508.
- [16] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Salt Lake City, UT: IEEE, 2018: 7132-7141.
- [17] LI K, WANG J, JALIL H, et al. A fast and lightweight detection algorithm for passion fruit pests based on improved YOLOv5 [J]. *Computers and Electronics in Agriculture*, 2023, 204: 107534.
- [18] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [C]//Proceedings of the European conference on computer vision (ECCV). 2018: 3-19.
- [19] LAN M, LIU C, ZHENG H, et al. Rice-yolo: in-field rice spike detection based on improved yolov5 and drone images [J]. *Agronomy*, 2024, 14(4): 836.
- [20] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning [C]//2023 IEEE international conference on acoustics, speech and signal processing (ICASSP), 2023: 1-5.
- [21] WANG J, QIN C, HOU B, et al. LCGSC-YOLO: A lightweight apple leaf diseases detection method based on LCNet and GSConv module under YOLO framework [J]. *Frontiers in Plant Science*, 2024, 15: 1398277.