

基于改进 YOLO11n 的葡萄果叶病害检测方法

吴世琛¹, 毛玉明^{1*}, 胡慧中¹, 张文康¹, 邢思瑶¹, 段会清¹, 丁青艳²

(1. 山东交通学院信息科学与电气工程学院, 济南 250000; 2. 山东省计算中心(国家超级计算济南中心), 济南 250000)

摘要: 为提高葡萄果叶病害的检测精度和速度, 解决葡萄因生长环境和病害种类复杂造成的误检、漏检等问题, 该研究提出了一种基于改进 YOLO11n 的葡萄病害检测模型 RLSL-YOLO11。首先, 为提高葡萄病害特征的提取能力, 将 YOLO11n 主干网络替换为融合 RepConv 的 HGNetv2, 构建改进型主干网络 RHGNetv2。其次, 引入基于 LSKA (large separable kernel attention) 注意力机制设计的改进型池化模块 LSPF, 用于强化模型的多尺度特征建模能力。再次, 采用 Slim-neck 架构优化颈部特征融合网络, 以降低模型的整体计算复杂度。最后, 设计了一种轻量级共享卷积分离器批量归一化检测头 LSCSBN (lightweight shared convolutional separator batch normalized detection head), 以进一步降低参数数量和计算复杂度, 并提高模型对多尺度病害特征的定位和提取能力。在该研究构建的数据集检测中, RLSL-YOLO11 模型的准确率、召回率、mAP_{0.5} 和 mAP_{0.5-0.95} 分别为 82.8%、76.1%、83.1% 和 52.0%, 与基准模型 YOLO11n 相比, mAP_{0.5} 和 mAP_{0.5-0.95} 分别提升了 1.9 和 6.0 个百分点; 模型参数量、计算复杂度和模型权重分别降低 12.0%、14.2% 和 8.9%。研究表明, RLSL-YOLO11 模型在有效提升综合检测能力的同时也降低了计算资源的需求, 为葡萄果叶病害检测的轻量化和实际应用提供了参考。

关键词: 病害; 葡萄; 深度学习; YOLO; 目标检测; 人工智能

doi: 10.11975/j.issn.1002-6819.202503032

中图分类号: S24

文献标志码: A

文章编号: 1002-6819(2025)-14-0134-08

吴世琛, 毛玉明, 胡慧中, 等. 基于改进 YOLO11n 的葡萄果叶病害检测方法[J]. 农业工程学报, 2025, 41(14): 134-141.

doi: 10.11975/j.issn.1002-6819.202503032 <http://www.tcsae.org>

WU Shichen, MAO Yuming, HU Huizhong, et al. Detecting grape leaf diseases using improved YOLO11n[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(14): 134-141. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202503032 <http://www.tcsae.org>

0 引言

葡萄是一种营养丰富的重要经济作物, 在农业生产中具有较高的经济价值。然而, 作为浆果类植物, 葡萄在生长过程中极易受到病害侵袭, 这些病害常导致其产量和品质下降, 进而造成显著的经济损失^[1]。当前, 在葡萄种植过程中, 病害识别主要依赖人工目测和经验判断, 但该方法存在识别成本高、效率低、误判率高以及易受种植环境干扰等问题^[2]。常见的葡萄病害包括灰霉病^[3]、霜霉病^[4]、白粉病^[5]和裂果病^[6], 这些病害可对葡萄的花序、果实及茎叶等部位造成严重损伤^[7]。由于葡萄不同生长周期对病害的敏感程度不同, 病害识别的及时性和准确性对葡萄的最终产量和品质具有重要影响^[8]。

鉴于传统人工病害检测存在效率低、准确率不高等问题, 研究人员开始深入研究基于计算机视觉的自动化病害识别方法, 以提高识别效率和精度。例如, ZHANG 等^[9]提出了一种基于多特征融合的大豆叶片病害检测模型 MF3R-CNN, 该模型通过跳跃连接的方式融合特征提

取网络中不同层的特征, 实现了多尺度信息的有效集成, 最终获得了 83.34% 的 mAP 指标。WANG 等^[10]基于迁移学习提出了一种轻量化的 SSD 模型, 通过优化特征提取网络结构与预测卷积核, 实现了柑橘害虫的高效检测与计数。毛锐等^[11]基于 Faster R-CNN 构建了一种用于识别小麦条锈病和黄矮病的模型, 利用卷积核拆解与下采样延迟策略对深度残差主干网络进行优化, mAP 为 91.06%。

在农业病害检测领域, 时效性是衡量检测系统性能的核心指标之一, 因此, 选用计算资源占用低、检测效率高的目标检测算法对于实际部署具有重要意义。YOLO^[12] (You Only Look Once) 作为一种兼具检测精度高、推理速度快、部署灵活等优点的高效目标检测算法, 已在农业病虫害场景中得到了广泛应用。例如, 孙俊等^[13]提出了一种基于改进 YOLOv7 的棉田虫害检测模型 ECSF-YOLOv7, 该模型融合 EfficientFormerV2 主干网络、CBAM 注意力机制、Slim-neck 架构以及 Focal-IOU 损失函数, 有效提升了模型在复杂棉田环境中的检测精度。马超伟等^[14]针对小麦病害识别任务, 提出了一种基于改进 YOLOv8 的轻量化检测方法, 该方法引入 PP-LCNet 主干网络、全局注意力机制 (GAM)、CARAFE 上采样模块及 Wise-IOU 损失函数, 在降低模型参数数量的同时显著提升了检测精度。孙丰刚等^[15]基于改进 YOLOv5s 和迁移学习提出了一种用于苹果果实病害识别的模型, 通过引入幻影结构、注意力机制、双向特征金字塔网络 (BiFPN) 以及 Transformer 模块, 实现了模型的轻量化

收稿日期: 2024-11-30 修订日期: 2025-06-01

基金项目: 国家重点研发计划项目 (2023YFC3304900), 中央引导地方科技发展资金项目 (YDZX2023050)

作者简介: 吴世琛, 研究方向为人工智能技术, 智慧农业。

Email: 1667807462@qq.com

*通信作者: 毛玉明, 副教授, 研究方向为人工智能技术、智慧物联技术。Email: 34293593@qq.com

与精度提升，为果实病害的快速、准确识别提供了有效方案。张惠莉等^[16]针对葡萄果叶病害识别任务，提出了一种结合 GhostNetV2 主干特征提取网络的改进 YOLOv8 模型，在检测精度和推理速度方面均表现出良好性能。尽管上述研究在特定作物病虫害检测中借助 YOLO 系列目标检测算法取得了良好效果，但在实际应用中仍面临诸多挑战。例如，检测模型对复杂种植环境（如光照变化、叶片遮挡、果实与背景色差小等）适应性不足，导致检测精度下降；部分性能提升模块的引入导致模型计算资源消耗较高，限制了其在资源受限场景下的应用和推广。为解决这些问题，本研究构建了包含葡萄灰霉病、霜霉病、白粉病以及裂果病的 4 类病害数据集，并提出了一种基于改进 YOLO11n 的轻量级葡萄病害检测模型 RLSL-YOLO11，旨在提升葡萄病害的检测精度，同时降低部署成本，实现病害实时监测。

1 材料和方法

1.1 数据集制作

本研究所用葡萄果叶病害图像采集自山东省临沂市郯城县葡萄种植基地，该基地采用大棚栽培模式。本研究主要关注葡萄果实与叶片上的病害，拍摄对象仅包括果实与叶片，未涉及枝干等其他部位。原始数据集图像共 2 449 张，其中 326 张为公开网络收集，2 123 张为园区拍摄，拍摄设备为尼康 D3400，拍摄时间为 2023 年 12 月—2024 年 5 月，拍摄图片原始像素为 2 992×2000。使用 X-AnyLabeling-GPU 软件手动标记病害部位和名称，数据集的病害图片数量和标签分布结果如表 1 所示，数据集样本数据如图 1 所示。

表 1 葡萄病害名称和数量分布

Table 1 Names and quantitative distribution of grapevine diseases

病害 Diseases	图片数 Number of images	标签数 Number of labels
白粉病 Powdery mildew	529	2 324
霜霉病 Downy mildew	720	1 740
灰霉病 Botrytis cinerea	800	1 558
裂果病 Fruit cracking disease	400	1 548
总的 All	2 449	7 170

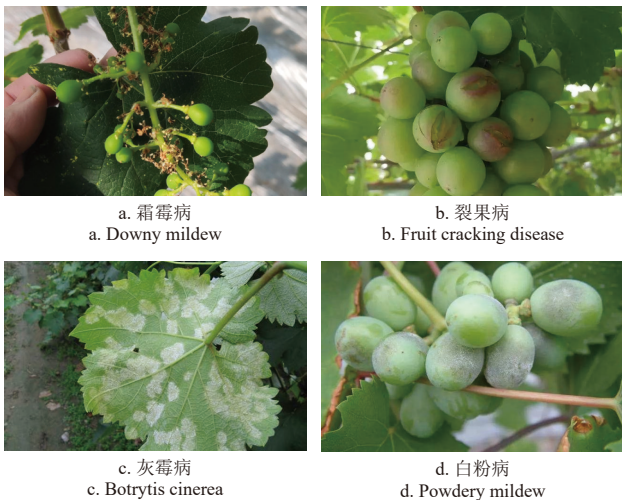


图 1 葡萄病害样本示例

Fig.1 Examples of grape disease samples

为提升模型的泛化能力并有效缓解过拟合问题，本研究引入数据增强技术对原始图像数据进行扩充。具体增强方法包括：图像随机旋转、上下左右翻转、亮度调整以及静态裁剪等。上述技术在一定程度上模拟了葡萄在实际生长过程中果叶的空间姿态及复杂环境下的干扰情况，从而增强了模型对真实场景的适应能力，通过数据增强技术，原始图像数量由 2 449 张扩展至 5 000 张。将数据集按照 8:1:1 的比例划分为训练集、验证集和测试集，训练集包括 4 000 张图片，验证集和测试集各 500 张图片。

1.2 方法

1.2.1 RLSL-YOLO11

YOLO11 由 Ultralytics 团队开发，其网络结构主要包括主干网络 (Backbone)、颈部特征融合网络 (Neck) 以及检测头 (Head)。YOLO11 在模型训练、推理速度、基础参数量以及计算复杂度等方面均实现了显著优化，同时引入了 C3k2 以及 C2PSA 等全新的特征提取模块。根据不同的计算资源需求，YOLO11 提供了五个版本，其中 YOLO11n 具有最少的参数量与计算复杂度，适用于计算资源受限的边缘设备。综合考虑模型的检测速度、精度表现与部署成本等因素，本研究选用 YOLO11n 作为基准模型。针对 YOLO11n 模型，本研究主要进行了以下结构改进：

1) 将 YOLO11n 原主干网络替换为基于 RepConv 改进后的 HGNetv2，即 RHGNetv2，以增强模型的特征提取能力；

2) 引入大核可分离卷积注意力机制 LSKA (large separable kernel attention)，改进池化结构，构建 LSPF 模块以替代原有的 SPPF 模块，提高模型对关键区域的响应能力；

3) 采用 Slim-neck 架构对原 YOLO11n 的颈部特征融合网络进行优化，以降低模型整体计算复杂度；

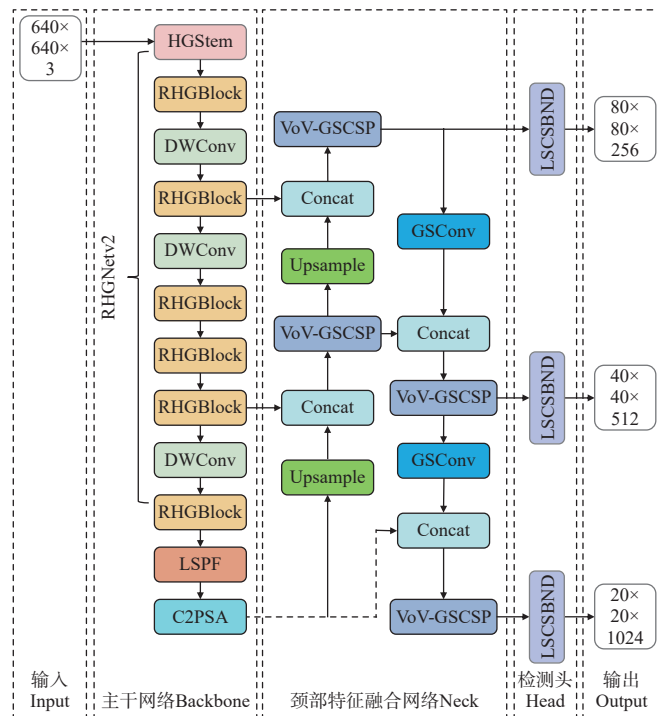
4) 提出一种轻量化检测头，命名为共享卷积分离器批量归一化检测头 LSCSBN (lightweight shared convolutional separator batch normalized detection head)，用于替换 YOLO11n 原检测头模块，以进一步提高检测效率与精度。

经改进后得到 RLSL-YOLO11 模型，其网络结构如图 2 所示。

1.2.2 RHGNetv2 主干网络

YOLO11n 原主干网络在追求轻量化的同时，特征提取能力相对不足，难以有效识别葡萄果实与叶片病害中复杂的纹理和边缘信息，易导致误检和漏检。HGNetv2 是百度飞桨视觉团队提出的 RT-DETR^[17] 目标检测算法的主干网络，兼顾轻量化和检测精度，且适用于计算资源受限场景。HGNetv2 主干网络由 HGStem，HGBlock，DWConv 组成，其核心模块 HGBlock 负责提取多样化特征，是特征表达的关键组件，同时也是占计算开销比例最大的部分。为进一步提升模型对细粒度病害特征的提取能力，并降低网络计算资源的开销，本研究借鉴 RepConv^[18] 的结构重参数化思想，对 HGBlock 模块进行改进，并将改进后的模块命名为 RHGBlock，其结构如图 3 所示。本研究特别考虑了与 YOLO11n 中其他关键

模块(如 C2PSA 注意力模块)的协同作用。C2PSA 模块融合了通道与空间注意力机制,能够引导模型更聚焦于关键区域,尤其适用于遮挡严重或背景复杂的病害检测场景。RHGBlock 在与 C2PSA 模块协同作用下,进一步强化多尺度特征表达,保持轻量化计算优势的同时,有效提升网络对葡萄果实和叶片病害细粒度特征的提取能力。



注: RHGNetv2 为改进后得到主干网络, DWConv 为深度可分离卷积模块; LSPF 为改进后得空间金字塔池化模块; C2PSA 为具有注意力机制的特征提取和处理模块; Upsample 为上采样操作; LSCSBND 为改进后的检测头。Note: RHGNetv2 is the backbone network obtained after improvement, and DWConv is the depth-separable convolutional module. LSPF is the improved spatial pyramid pooling module. C2PSA is a feature extraction and processing module with an attention mechanism; Upsample is an upsampling operation; LSCSBND is an improved detection head.

图2 RLSL-YOLO11 的网络结构

Fig.2 Network structure of RLSL-YOLO11

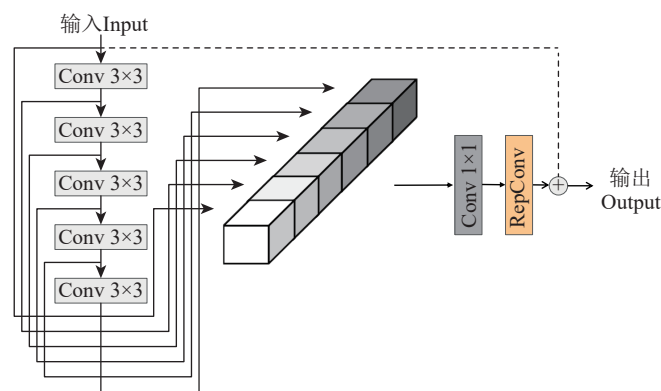


图3 RHGBlock 结构

Fig.3 Structure of RHGBlock

RepConv (re-parameterized convolution, 重参数化卷积)是一种结构重参数化技术,旨在保持模型精度的前提下提升推理效率。其核心思想在于训练与推理阶段结构解耦:在训练阶段引入多分支结构增强特征表达;而

在推理阶段则通过参数融合将多分支结构等效转换为单一卷积分支,从而实现结构简化。

本研究将 RepConv 引入 HGNetv2,对其核心模块 HGBlock 进行改进,旨在利用 RepConv 的多分支结构提升对小目标特征的提取能力,从而增强小目标检测精度。改进中首先对 3×3 卷积层与 BN^[19] (batch normalization) 层进行融合,其计算过程如式(1)和式(2)所示:

$$\text{Conv}(x) = W(x) + b \quad (1)$$

$$\text{BN}(x) = \gamma \times \frac{x - M}{\sqrt{v}} + \beta \quad (2)$$

其中,式(1)为卷积层计算公式,式(2)为 BN 层计算公式。Conv(x) 表示卷积操作的最终输出结果,是加了偏置之后的特征图,BN(x) 为归一化并线性变换后的输出。

将卷积结果带入到 BN 公式中进行化简如式(3)所示:

$$\left\{ \begin{aligned} \text{BN}(\text{Conv}(x)) &= \gamma \times \frac{\text{Conv}(x) - M}{\sqrt{v}} + \beta \\ &= \gamma \times \frac{W(x) + b - M}{\sqrt{v}} + \beta \\ &= \frac{\gamma \times W(x)}{\sqrt{v}} + \left(\frac{\gamma \times (b - M)}{\sqrt{v}} + \beta \right) \end{aligned} \right. \quad (3)$$

式中, $\text{BN}(\text{Conv}(x))$ 表示对卷积后的输出进行归一化处理, $W(x)$ 表示输入 x 与卷积核权重 W 进行卷积操作后的输出结果, M 指在结构重参数化过程中对多个卷积分支的权重取平均, γ 代表可学习参数权重, b 代表卷积层的偏置项, β 代表权衡参数, $\frac{\gamma \times W(x)}{\sqrt{v}}$ 为融合后卷积的参数权重值, $\frac{\gamma \times (b - M)}{\sqrt{v}} + \beta$ 为融合后卷积的权衡参数值。随后,将 1×1 卷积核的维度调换,并与 3×3 卷积核权重进行卷积操作,如式(4)所示:

$$P_{\text{out}} = \text{Conv}_{3\times 3}(\text{Conv}_{1\times 1}(x)) \quad (4)$$

式中 P_{out} 代表经过两次卷积处理后的最终输出特征图。

最后,将所有卷积权重通过填充对齐后相加,得到 RHGBlock,该结构在训练时采用多分支卷积,推理时将分支参数重参数化合并为单一主分支。

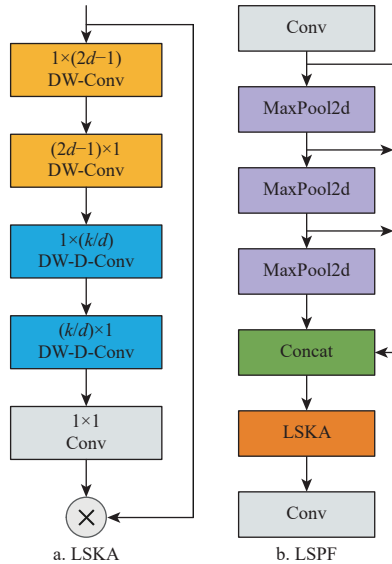
1.2.3 LSPF 模块

SPPF (spatial pyramid pooling - fast) 模块是空间金字塔池化 SPP^[20] (spatial pyramid pooling) 的一种轻量化改进版本。它的主要作用是通过多尺度池化操作融合不同感受野的特征,有效提升模型对目标尺度变化的鲁棒性。尽管该设计通过参数共享降低了计算复杂度,但堆叠式池化操作可能导致局部细节信息丢失,特别是在检测密集且细小目标时表现不足,而特征拼接操作会进一步加剧特征信息的衰减。针对葡萄果叶病害图像中密集小病斑的检测需求,本研究通过融合注意力机制重构 SPPF 模块,以强化关键区域的特征响应。

LSKA^[21]是在大核注意力 LKA^[22] (large kernel attention) 的基础上改进而来,融合了“大核”感受野与“可分离”卷积的优势。传统注意力机制通过特征重加权提升检测精度,但大核卷积通常带来较高的参数数量和

计算开销。大核卷积具有较大感受野，能够有效捕捉图像中的全局信息和大尺度特征，如物体轮廓，但对小目标细节的敏感性较弱，可能导致局部特征被全局信息淹没，影响检测性能。LSKA 通过将标准深度卷积（DW-Conv）和深度膨胀卷积（DW-D-Conv）分解为水平和垂直的 1D 卷积内核级联，实现了局部细节与全局上下文的有效融合。该设计既缓解了大感受野带来的局部信息丢失问题，又避免了额外计算模块的引入，在保持等效感受野的同时显著降低了计算复杂度，提升了模型性能与计算效率。

基于 LSKA 的思想，本研究对 SPPF 模块进行改进，提出了一种融合注意力机制的模块，命名为 LSPF，结构如图 4 所示。该模块在保持 SPPF 多尺度特征融合优势的基础上，利用注意力机制动态调节不同区域的权重，增强对细节信息的关注能力，从而提升模型在多尺度特征提取与融合过程中的表现，进而提高检测性能。



注：MaxPool2d 为最大池化层；LSKA 为大可分离核注意力；DW-D-Conv 为深度扩张卷积； d 为膨胀率； k 为卷积核边长。
 Note: MaxPool2d is the maximum pooling layer; LSKA is the Large Separable Kernel Attention; DW-D-Conv is the Depth Expansion Convolution; d is the expansion rate; k is the convolution kernel edge length.

图 4 SKA 与 LSPF 结构
 Fig.4 Structure of LSKA and LSPF

1.2.4 Slim-neck 架构

为兼顾葡萄果叶病害检测的精度和实时性，降低计算资源需求，本研究引入 Slim-neck 架构^[23]改进 YOLO11n 的颈部特征融合网络（Neck）。Neck 连接主干网络（Backbone）与检测头（Head），主要负责多层特征的融合与传递。Slim-neck 通过将传统卷积 Conv 替换为轻量级卷积 GSConv 实现轻量化，GSConv 结合了标准卷积（standard convolution, SC）、深度可分离卷积（depthwise separable convolution, DSC）与通道混洗（Shuffle）技术，其中架构核心为分组卷积（grouped convolution, GConv），可有效加速卷积计算，结构如图 5 所示。GSConv 在提升推理速度的同时，保持了较强的特征学习与融合能力。

为进一步降低 Neck 部分的计算开销，在基于 GSConv 的基础上，将 Neck 部分的 C3k2 模块替换为 VoV-GSCSP

模块。VoV-GSCSP 采用单次跨阶段聚合策略设计，其结构如图 6 所示，用于实现不同阶段特征图之间的信息高效融合，兼顾深层特征的语义表达与浅层特征细节信息，从而提升特征利用效率，平衡计算成本与检测精度。通过结合 GSConv 与 VoV-GSCSP，构建出 Slim-neck 架构。

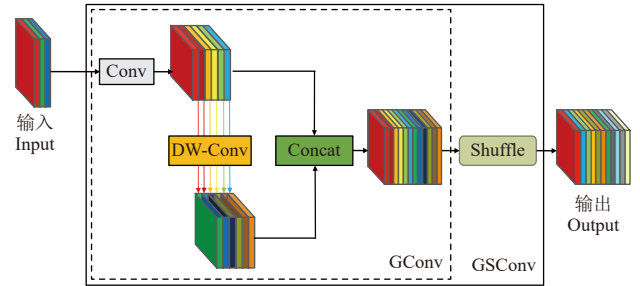


图 5 GSConv 模块结构
 Fig.5 Structure of GSConv module

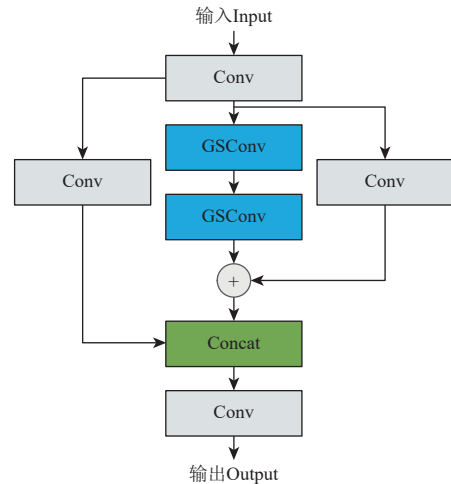


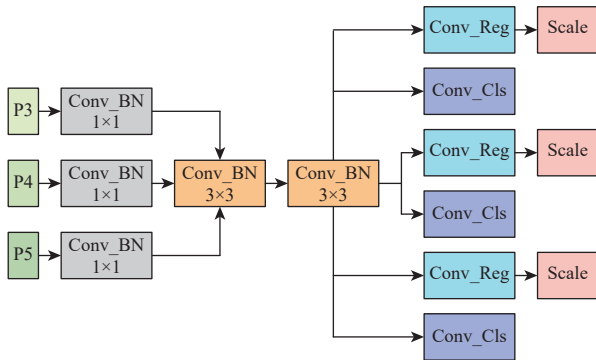
图 6 VoV-GSCSP 模块结构
 Fig.6 Structure of VoV-GSCSP module

1.2.5 LSCSBND 检测头

YOLO11 沿用 YOLOv8 的解耦头（decoupled head）设计，即在不同的网络层级使用独立的检测头。这种任务解耦的处理方法能够减少多类别检测中的相互干扰，提升检测精度。该设计通过两次 3×3 卷积和一次 1×1 卷积串联来实现，但是会导致参数量和计算复杂度显著增加。为此，提出一种轻量级共享卷积分离器批量归一化检测头 LSCSBND。鉴于不同层级特征的统计差异，归一化层（normalization layer）不可缺少，但批量归一化（batch normalization, BN）对批量大小敏感，直接引入 BN 易导致滑动平均值出现偏差。为解决此问题，本研究借鉴 NASFP 的权重共享策略^[24]，结合 C3k2 模块中动态残差连接的思想，实现卷积层权重共享而 BN 独立计算，解决 BN 对批量大小敏感的问题。LSCSBND 检测头结构如图 7 所示。

P3、P4、P5 分别表示不同尺度的特征图层，用于检测不同大小的目标。LSCSBND 检测头接收来自 P3、P4、P5 的特征图后，先通过带 BN 的 1×1 共享卷积调整通道维度，再利用两组带 BN 的 3×3 共享卷积实现跨层级特征融合，降低参数量。输出阶段采用独立权重的 1×1 卷积分支分别处理定位与分类任务，实现任务解耦。同时，

引入 Scale 层进行特征缩放, 提升模型对多尺度目标的检测能力。



注: P3 代表第 3 层特征图, 分辨率较高、信息更细腻, 用于检测小目标; P4 代表第 4 层特征图, 中等分辨率, 用于检测中等大小目标; P5 代表第 5 层特征图, 分辨率最低、语义最强, 用于检测大目标。

Note: P3 represents Layer 3 feature map with higher resolution and more delicate information for detecting small targets; P4 represents Layer 4 feature map with medium resolution for detecting medium-sized targets; P5 represents Layer 5 feature map with the lowest resolution and strongest semantics for detecting large targets.

图 7 LSCSBND 检测头结构

Fig.7 Structure of LSCSBND detection head

表 2 RLSL-YOLO11 与其他模型的对比结果

Table 2 Comparison results of RLSL-YOLO11 with other models

模型 Models	准确率 Precision P/%	召回率 Recall R/%	mAP _{0.5} /%	mAP _{0.5-0.95} /%	参数量 Params/M	计算复杂度 Floating point operations FLOPs/G	模型权重 Model size/MB	帧率 Frames per second/ (帧·s ⁻¹)
Faster R-CNN	76.4	73.4	77.7	41.0	41.50	207.0	167.01	28.0
RT-DETR-R18	85.4	78.8	82.8	52.7	19.88	57.0	38.60	88.4
RT-DETR-R34	86.6	80.3	82.7	54.0	31.11	88.0	60.11	66.2
YOLOv3-tiny	79.6	74.3	79.6	41.1	8.67	12.9	23.20	394.2
YOLOv5n	78.5	76.6	78.9	44.0	2.50	7.1	5.05	316
YOLOv6n	78.7	75.3	80.1	45.7	4.23	11.8	8.31	388.6
YOLOv7-tiny	77.1	73.4	78.5	38.5	6.02	13.0	11.70	219.5
YOLOv8n	79.2	77.0	81.1	47.6	3.01	8.1	5.98	338.9
YOLOv9-t	78.3	76.7	80.7	45.3	2.62	10.7	5.84	227.7
YOLOv10n	77.6	74.0	78.5	44.4	2.27	6.5	5.53	255.8
YOLO11n	81.2	76.5	81.2	46.0	2.58	6.3	5.25	243.2
YOLOv12n	80.0	78.3	81.3	47.3	2.56	6.3	5.30	149.4
RLSL-YOLO11	82.8	76.1	83.1	52.0	2.27	5.4	4.78	192.3

注: mAP₅₀ 为 IoU 阈值为 0.5 时的平均精度均值, mAP₅₀₋₉₅ 为 IoU 阈值为 0.5 到 0.95, 步长为 0.05 时 mAP 的平均值。

Note: mAP₅₀ is the average precision when the intersection over union threshold is 0.5, mAP₅₀₋₉₅ is the average value of mAP for IoU thresholds of 0.5 to 0.95 in steps of 0.05.

如表 2 所示, RLSL-YOLO11 在准确率、召回率、mAP_{0.5} 和 mAP_{0.5-0.95} 分别达 82.8%、76.1%、83.1% 和 52.0%, 较基准模型 YOLO11n 在 mAP_{0.5} 和 mAP_{0.5-0.95} 上分别提升了 1.9 和 6.0 个百分点。在模型参数量、计算复杂度和模型权重分别减少了 12.0%、14.2% 和 8.9%。RT-DETR-R18 在参数量是 RLSL-YOLO11 模型的 8.7 倍, 计算复杂度 10.5 倍的情况下, 在 mAP_{0.5} 和 mAP_{0.5-0.95} 的指标上并没有明显的领先。RLSL-YOLO11 模型在参数量仅为 RT-DETR-R34 的 7.3%, 且计算复杂度仅为 6.1% 的情况下, 实现了相近的 mAP_{0.5} 和 mAP_{0.5-0.95} 的指标。

在本研究所涉及的葡萄病害中, 白粉病因其病斑细小且密集, 属典型的密集小目标, 同时易受环境光照、灰尘及果粉等白色干扰物影响, 增加了检测难度, 对模型检测能力提出了更高要求。为避免单一病害精度提升对整体检测精度造成误导, 表 3 对比列出了 YOLO11n 与 RLSL-YOLO11 在各类病害上的 mAP_{0.5} 指标。结果表明, RLSL-YOLO11 在所有病害类别上均实现了精度提升, 未出现某一类别指标异常波动的情况, 从而有效验证了

1.2.6 试验环境及评价指标

试验平台所用操作系统为 Window10 (22H2), 深度学习框架为 Pytorch (2.2.2), CPU 为 Intel(R) Core(TM) i5-12490F, GPU 为 NVIDIA GeForce RTX 4060Ti 16GB。初始学习率 (Learning rate) 为 0.01, 动量 (momentum) 为 0.937, 优化器 (optimizer) 选用 SGD, 批量设置 (batch) 为 32, 其他超参数沿用默认设置, 统一训练 500 轮 (epochs), 并最终基于测试集数据统计模型性能指标。

试验采用准确率 (precision, P)、召回率 (recall, R)、平均精度均值 (mAP_{0.5} 和 mAP_{0.5-0.95}) 作为模型性能的主要评价指标。模型参数量 (params)、计算复杂度 (floating point operations, FLOPs)、模型权重 (model size) 和每秒帧数 (frames per second, FPS) 作为模型的次要评价指标。

2 结果与分析

2.1 不同模型结果对比

为了验证改进模型的性能, 在相同试验环境下使用自建葡萄果叶病害数据集对选取模型进行训练并收集测试集数据, 对比结果如表 2 所示。

改进模型在各类病害检测中的整体性能提升。

表 3 模型病害 mAP_{0.5} 指标对比

Table 3 Comparison of model disease mAP_{0.5} metrics

病害 Disease	YOLO11n	RLSL-YOLO11
灰霉病 Botrytis cinerea	82.0	84.8
霜霉病 Downy mildew	87.0	89.4
白粉病 Powdery mildew	61.9	63.5
裂果病 Fruit cracking disease	93.8	94.6
总的 All	81.2	83.1

2.2 主干网络对比试验

为验证 RHGNetv2 主干网络对模型性能的提升效果, 本研究选取了多个主流的主干网络进行比较, 如表 4 所示。

从表 4 可以得出, RHGNetv2 主干网络的 mAP_{0.5} 比 ConvNeXtV2、InceptionNeXt、MobileNetV4、StarNet、EfficientViT、HGNetv2、GhostHGNetv2 分别高 1.8、2.5、4.7、2.9、0.1、2.3、1.4 个百分点。在参数量方面, RHGNetv2 仅为 2.13M, 远低于如 ConvNeXtV2 (5.39M) 和 InceptionNeXt (26.17M) 等主干网络, 同时计算复杂度

度仅为 5.7G，在所列主干网络中也处于较低水平，仅略高于参数最少的 StarNet，但精度明显优于 StarNet。

表 4 不同主干网络的性能对比

Table 4 Performance comparison of different backbone networks

主干网络 Backbone	mAP _{0.5} /%	mAP _{0.5-0.95} /%	参数量 Params/M	FLOPs/G
ConvNeXtV2 ^[25]	80.3	44.9	5.39	12.6
InceptionNeXt ^[26]	79.6	44.1	26.17	72.7
MobileNetV4 ^[27]	77.4	43.4	5.43	21.0
StarNet ^[28]	79.2	45.0	1.94	5.0
EfficientViT ^[29]	82.0	47.2	3.74	7.9
HGNetv2	79.8	44.9	2.14	5.7
GhostHGNetv2 ^[30]	80.7	47.6	3.14	7.7
RHGNetv2	82.1	49.3	2.13	5.7

2.3 消融试验

在本研究中，为验证各改进模块对模型性能的贡献，利用消融试验来验证改进模块对算法性能的影响。共设计了 8 组试验，依次引入 RHGNetv2、LSPF、Slim-neck

和 LSCSBND 四个模块，评估其对检测精度和模型效率的影响，表 5 为 RLSL-YOLO11 模型消融试验数据。

从表 5 可以得出，本研究所提出的 RHGNetv2、LSPF、Slim-neck 和 LSCSBND 均在不同程度上提升了模型的检测性能。其中，RHGNetv2 在减少参数量和模型体积的同时，将 mAP_{0.5} 提升至 82.1%，mAP_{0.5-0.95} 提升至 49.3%，表现出良好的特征提取能力；LSPF 提升了检测精度，但略微增加了模型复杂度；Slim-neck 和 LSCSBND 则在保持精度的同时有效压缩了参数和模型大小，体现出良好的轻量化效果。最终，将 4 个模块全部引入后，构建的 RLSL-YOLO11 模型在精度（mAP_{0.5} 为 83.1%，mAP_{0.5-0.95} 为 52%）和效率（参数量 2.27M，FLOPs 为 5.4G，模型权重 4.78MB）之间达到了最佳平衡，验证了各模块在提升检测性能和降低资源消耗方面的有效性和协同优势。

表 5 RLSL-YOLO11 模型消融试验

Table 5 RLSL-YOLO11 model ablation experiments

YOLO11n	RHGNetv2	LSPF	Slim-Neck	LSCSBND	mAP _{0.5} /%	mAP _{0.5-0.95} /%	参数量 Params/M	FLOPs/G	模型权重 Model size/MB
√	-	-	-	-	81.2	46.0	2.58	6.3	5.25
√	√	-	-	-	82.1	49.3	2.13	5.7	4.5
√	-	√	-	-	82.1	48.2	2.86	6.5	5.77
√	-	-	√	-	82.0	47.5	2.57	5.9	5.26
√	-	-	-	√	82.7	48.8	2.46	6.2	5.0
√	√	√	-	-	82.8	51.3	2.41	5.9	5.03
√	√	√	√	-	82.5	50.5	2.39	5.5	5.04
√	√	√	√	√	83.1	52.0	2.27	5.4	4.78

注：“√”表示采用此项；“-”表示未采用此项。

Note: "√" represents that this item is adopted, and "-" represents that this item is not adopted.

2.4 测试结果可视化

本章节将数据集的真实标注和 YOLO11n 模型检测结果以及 RLSL-YOLO11 模型检测结果进行可视化，如图 8 所示。

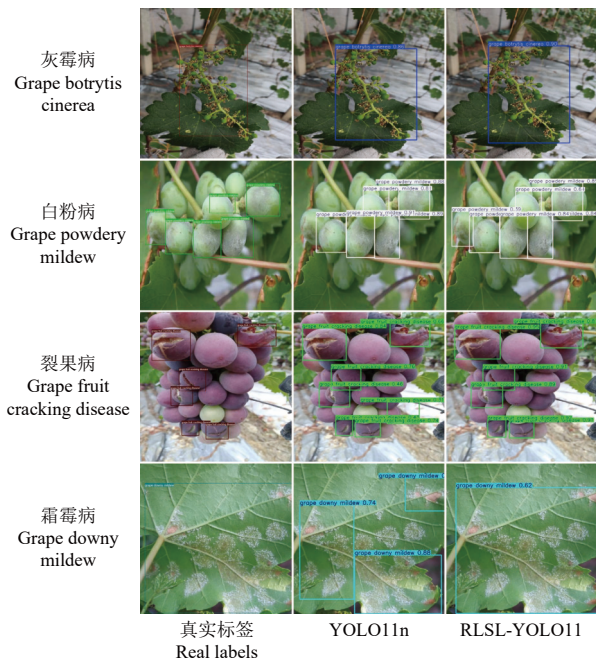


图 8 YOLO11n 与 RLSL-YOLO11 检测结果对比

Fig.8 Comparison of YOLO11n and RLSL-YOLO11 detection results

图 8 可见，YOLO11n 在检测葡萄白粉病与裂果病时存在明显的漏检与误检现象，在灰霉病与霜霉病的检测中亦表现出检测置信度较低、检测框密集的问题。相比之下，改进后的 RLSL-YOLO11 模型在各类病害目标的检测精度与稳定性方面均有显著提升，表现出更优的目标定位能力和更高的置信度分布。

2.5 泛化试验

为进一步验证 RLSL-YOLO11 模型的泛化性能，本节选取了 IDADP^[31] (an image dataset for IDADP-grape disease identification) 数据集进行测试，该数据集包含了白粉病、霜霉病、灰霉病等 7 种高质量病害图像，选取其中 2023 张图像进行测试，测试结果如表 6 所示。

表 6 RLSL-YOLO11 模型泛化结果

Table 6 RLSL-YOLO11 model generalization results

模型 Models	P/%	R/%	mAP _{0.5} /%	mAP _{0.5-0.95} /%
YOLOv5n	77.8	72.0	78.1	44.2
YOLOv8n	77.3	72.0	77.8	45.3
YOLOv10n	75.3	69.4	75.5	42.3
YOLO11n	77.6	72.8	78.2	44.2
RLSL-YOLO11	79.6	72.5	79.3	47.9

试验结果表明，改进后的 RLSL-YOLO11 模型在目标检测关键指标上优于其它基准模型。相较于 YOLO11n，其 mAP_{0.5} 和 mAP_{0.5-0.95} 分别提升了 1.1 和 3.7 个百分点。该结果不仅证实了模型鲁棒性的提升，还通过更严苛的 mAP_{0.5-0.95} 指标凸显了其泛化性能的增强。

3 结论

本研究为提高葡萄果叶病害的检测精度,降低模型计算资源需求,解决葡萄因生长环境和病害种类复杂造成的误检、漏检以及模型计算开销大等问题,提出了一种融合轻量化结构与高效注意力机制的 RLSL-YOLO11 检测模型,主要研究结论如下:

1) 本研究基于 YOLO11n 模型构建改进结构,在主干网络中引入轻量高效的 RHGNetv2 网络,增强特征表达能力的同时控制参数规模;在 SPPF 模块中引入 LSKA 注意力机制构建 LSPF 模块,提升多尺度特征提取能力;在颈部特征融合网络中采用 GSCov 与 VoV-GSCSP 形成 Slim-neck 架构,有效压缩模型计算复杂度;在检测头部分设计了轻量级共享卷积分离器批量归一化检测头 LSCSBND,提升检测速度。与原始 YOLO11n 相比, $mAP_{0.5}$ 和 $mAP_{0.5-0.95}$ 分别提升了 1.9 和 6.0 个百分点,模型参数量、计算复杂度和模型权重分别降低 12.0、14.2 和 8.9 个百分点。

2) 与当前主流检测模型对比,RLSL-YOLO11 在精度和轻量化方面具有显著优势,平均精度均值 $mAP_{0.5}$ 为 83.1%, $mAP_{0.5-0.95}$ 为 52.0%,模型体积仅为 4.78MB,帧数达 192.3FPS,RLSL-YOLO11 在精度、参数量、模型体积与推理速度等方面均实现较优平衡。

本研究提出的 RLSL-YOLO11 模型在自建葡萄果叶病害数据集上取得了较为理想的检测效果,但仍存在不足之处。受限于当前数据采集条件,数据集在样本规模、病害种类、生长期及环境复杂性等方面尚显不足。后续研究将持续扩充数据规模,增强数据多样性,以提高模型在不同实际场景中的适应性。为提升模型在实际应用中的部署效率与可移植性,还将引入模型剪枝和知识蒸馏等轻量化优化策略,在保障检测精度的基础上进一步压缩模型体积,降低计算资源消耗与硬件依赖。后续工作将聚焦于模型在边缘设备或移动终端上的部署试验,评估其在田间环境中的实时性能与稳定性。

[参 考 文 献]

- 于海森, 石琳, 柯忻, 等. 葡萄均衡营养肥对鲜食葡萄品质的影响[J]. 肥料与健康, 2020, 47(1): 21-26.
YU Haisen, SHI Lin, KE Xin, et al. Effect of grape balanced nutrition fertilizer on quality of table grape[J]. Fertilizer & Health, 2020, 47(1): 21-26. (in Chinese with English abstract)
- 何雨霜, 王琢, 王湘平, 等. 深度学习在农作物病害图像识别中的研究进展[J]. 中国农机化学报, 2023, 44(2): 148-155.
HE Yushuang, WANG Zhuo, WANG Xiangping, et al. Research progress of deep learning in crop disease image recognition[J]. Journal of Chinese Agricultural Mechanization, 2023, 44(2): 148-155. (in Chinese with English abstract)
- 李博雅, 鲁妍璇, 谢家贝, 等. 武夷菌素防治葡萄灰霉病的作用及机理[J]. 中国生物防治学报, 2023, 39(3): 676.
LI Boya, LU Yanxuan, XIE Jiabei, et al. Suppression and mechanism of wuyiencin on grape gray mold[J]. Chinese Journal of Biological Control, 2023, 39(3): 676. (in Chinese with English abstract)
- 何涛, 陈吉越, 孙家秀, 等. 葡萄霜霉病对葡萄叶片代谢的影响[J]. 植物病理学报, 2022, 52(5): 821-828.
HE Tao, CHEN Jiyue, SUN Jiayou, et al. Impacts of grape downy mildew on metabolism in grape leaves[J]. Acta Phytopathologica Sinica, 2022, 52(5): 821-828. (in Chinese with English abstract)
- 贾静, 怡张玮, 燕继晔, 等. 葡萄白粉病抗性鉴定方法优化及品种抗性评价[J]. 植物保护, 2021, 47(1): 160-164.
JIA Jingyi, ZHANG Wei, YAN Jiye, et al. Optimization of grape powdery mildew resistance identification methods and evaluation of variety resistance[J]. Plant Protection, 2021, 47(1): 160-164. (in Chinese with English abstract)
- 冯学梅, 杨玲, 梁玉文. 大青葡萄裂果发生原因与预防措施[J]. 农业科学, 2022, 12: 154.
FENG Xuemei, YANG Ling, LIANG Yuweng. Causes and preventive measures of fruit cracking of "Daqing" grape[J]. Hans Journal of Agricultural Sciences, 2022, 12: 154. (in Chinese with English abstract)
- 韦慧玲, 梁成斌, 王金海, 等. 绳驱柔性机械手研究进展及其在农业机器人中的应用[J]. 智能化农业装备学报 (中英文), 2024, 5(4): 95-106.
WEI Huiling, LIANG Chengbin, WANG Jinhai, et al. Research progress of cable-driven flexible manipulator and its application in agricultural robots[J]. Journal of Intelligent Agricultural Mechanization, 2024, 5(4): 95-106. (in Chinese with English abstract)
- 兰玉彬, 王天伟, 陈盛德, 等. 农业人工智能技术: 现代农业科技的翅膀[J]. 华南农业大学学报, 2020, 41(6): 1-13.
LAN Yubin, WANG Tianwei, CHEN Shengde, et al. Agricultural artificial intelligence technology: Wings of modern agricultural science and technology[J]. Journal of South China Agricultural University, 2020, 41(6): 1-13. (in Chinese with English abstract)
- ZHANG K, WU Q, CHEN Y. Detecting soybean leaf disease from synthetic image using multi-feature fusion faster R-CNN[J]. Computers and Electronics in Agriculture, 2021, 183: 106064.
- WANG L, SHI W, TANG Y, et al. Transfer learning-based lightweight SSD model for detection of pests in citrus[J]. Agronomy, 2023, 13(7): 1710.
- 毛锐, 张宇晨, 王泽玺, 等. 利用改进 Faster-RCNN 识别小麦条锈病和黄矮病[J]. 农业工程学报, 2022, 38(17): 176-185.
MAO Rui, ZHANG Yuchen, WANG Zexi, et al. Recognizing stripe rust and yellow dwarf of wheat using improved Faster-RCNN[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(17): 176-185. (in Chinese with English abstract)
- HAN X, CHANG J, WANG K. You only look once: Unified, real-time object detection[J]. Procedia Computer Science, 2021, 183(1): 61-72.
- 孙俊, 贾忆琳, 吴兆祺, 等. 基于改进 YOLOv7 的棉田虫害检测[J]. 农业工程学报, 2024, 40(10): 176-184.
SUN Jun, JIA Yilin, WU Zhaoqi, et al. Detecting pests in cotton fields using improved YOLOv7[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(10): 176-184. (in Chinese with English abstract)
- 马超伟, 张浩, 马新明, 等. 基于改进 YOLOv8 的轻量化小麦病害检测方法[J]. 农业工程学报, 2024, 40(5): 187-195.
MA Chaowei, ZHANG Hao, MA Xinming, et al. Method for the lightweight detection of wheat disease using improved YOLOv8[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(5): 187-195. (in Chinese with English abstract)
- 孙丰刚, 王云露, 兰鹏, 等. 基于改进 YOLOv5s 和迁移学习的苹果果实病害识别方法[J]. 农业工程学报, 2022, 38(11): 171-179.
SUN Fenggang, WANG Yunlu, LAN Peng, et al. Identification of apple fruit diseases using improved YOLOv5s and transfer learning[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(11): 171-179. (in Chinese with English abstract)
- 张惠莉, 代晨龙, 任景龙, 等. 基于 YOLOv8-GSGF 模型的葡萄病害识别方法研究[J]. 农业机械学报, 2024, 55(11): 75-83.
ZHANG Huili, DAI Chenlong, REN Jinglong, et al. Grape disease identification method based on YOLOv8-GSGF model[J]. Transactions of the Chinese Society for Agricultural

- Machinery*, 2024, 55(11): 75-83. (in Chinese with English abstract)
- [17] ZHAO Y, LV W, XU S, et al. Detsr beat yolos on real-time object detection[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA, 2024: 16965-16974.
- [18] DING X, ZHANG X, MA N, et al. Repvgg: Making vgg-style convnets great again[C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Virtual Event, Canada, 2021: 13733-13742.
- [19] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning (ICML). Lille, France, PMLR, 2015: 448-456.
- [20] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [21] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: Rethinking the large kernel attention design in cnn[J]. *Expert Systems with Applications*, 2024, 236: 121352.
- [22] GUO M H, LU C Z, LIU Z N, et al. Visual attention network[J]. *Computational Visual Media*, 2023, 9(4): 733-752.
- [23] LI H, LI J, WEI H, et al. Slim-neck by GSConv: A lightweight-design for real-time detector architectures[J]. *Journal of Real-Time Image Processing*, 2024, 21(3): 62.
- [24] GHIASI G, LIN T Y, LE Q V. Nas-fpn: Learning scalable feature pyramid architecture for object detection[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA, 2019: 7036-7045.
- [25] WOO S, DEBNATH S, HU R, et al. Convnext v2: Co-designing and scaling convnets with masked autoencoders[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada, 2023: 16133-16142.
- [26] YU W, ZHOU P, YAN S, et al. Inceptionnext: When inception meets convnext[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA, 2024: 5672-5683.
- [27] QIN D, LEICHNER C, DELAKIS M, et al. MobileNetV4: Universal models for the mobile Ecosystem[C]//European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 78-96.
- [28] MA X, DAI X, BAI Y, et al. Rewrite the Stars[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA, 2024: 5694-5703.
- [29] LIU X, PENG H, ZHENG N, et al. Efficientvit: Memory efficient vision transformer with cascaded group attention[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada, 2023: 14420-14430.
- [30] DAI D, WU H, WANG Y, et al. LHSDNet: A Lightweight and High-Accuracy SAR Ship Object Detection Algorithm[J]. *Remote Sensing*, 2024, 16(23): 4527.
- [31] 袁媛, 陈雷. IDADP-葡萄病害识别研究图像数据集[J]. *中国科学数据*, 2022, 7(1): 86-90.
YUAN Yuan, CHEN Lei. An image dataset for IDADP-grape disease identification[J]. *China scientific data*, 2022, 7(1): 86-90.

Detecting grape leaf diseases using improved YOLO11n

WU Shichen¹, MAO Yuming^{1*}, HU Huizhong¹, ZHANG Wenkang¹, XING Siyao¹,
DUAN Huiqing¹, DING Qingyan²

(1. School of Information Science and Electrical Engineering, Shandong Jiaotong University, Jinan 250000, China;

2. Shandong Computer Science Center (National Supercomputer Center in Jinan), Jinan 250000, China)

Abstract: Precise and rapid detection of the grape fruit and leaf diseases is of vital importance to the stable development of the grape industry. The growth environment of the grapes is also dominated by various highly complex factors, such as the climate and soil. Meanwhile, there are the numerous types of the diseases, with the complex and interwoven symptoms. However, traditional detection cannot fully meet the requirements to distinguish different diseases, due mainly to the frequent misdetections and missed detections, thus leading to the low yield and quality of grapes. In this study, a novel model was proposed to detect the grape fruit and leaf disease using the improved YOLO11n - RLSL-YOLO11n. Four common and highly harmful diseases were taken as the grape botrytis cinerea, downy mildew, powdery mildew and fruit cracking. According to the efficient detection of the YOLO series algorithms, the RLSL-YOLO11 model was achieved in the better performance after a series of the improvements. Firstly, the RepConv convolution module was introduced into the backbone network. A multi-scale fusion backbone network, RHGNetv2 was established using the HGNetv2 architecture. The performance of the improved model was enhanced to capture the disease features at different scales. After that, the disease features were more comprehensively perceived to effectively reduce the number of parameters and computational complexity of the lightweight model after optimization on the network structure. Secondly, an SPPF (Spatial Pyramid Pooling - Fast) module was obtained to further enhance the feature extraction of the improved model using the LSKA (Large Separable Kernel Attention) attention mechanism. The LSPF module was then introduced the large-scale separable convolution and attention mechanism, in order to focus more on the feature extraction of the disease area. At the same time, the features of the diseases were recognized and distinguished in the complex backgrounds, in order to effectively reduce the background interference. Furthermore, the Slim Neck architecture was adopted to optimize the Neck feature fusion network neck of the YOLO11n model. The feature fusion path was simplified to reduce the redundant calculations. A high recognition accuracy rate was maintained to further reduce the computational complexity for the operational efficiency of the improved model. Finally, a lightweight shared convolution separator batch normalized detection head (LSCSBND) was designed to further enhance the lightweight degree of the model. The detection head was effectively reduced the number of parameters and computational complexity. The shared convolution kernels and batch normalization were simultaneously improved to locate and then extract multi-scale disease features. A dataset was constructed to verify the performance of the RLSL-YOLO11 model. Specifically, 2,449 original images were contained in the four kinds of grape fruit and leaf diseases. The RLSL-YOLO11 model was achieved in the accuracy rate, recall rate, mAP0.5 and mAP0.5-0.95 of 82.8%, 76.1%, 83.1% and 52.0%, respectively. The mAP0.5 and mAP0.5-0.95 of YOLO11n were improved by 1.9 and 6.0 percentage points, respectively, compared with the baseline model. At the same time, the number of model parameters, computational complexity and model weights were reduced by 12.0%, 14.2% and 8.9% respectively. This finding can provide a new solution to the precise detection of the grape fruit and leaf diseases. Strong support can also offer for the lightweight deployment and practical application of the disease detection in modern agriculture.

Keywords: disease; grapes; deep learning; YOLO; target detection; artificial intelligence