doi:10.6041/j.issn.1000-1298.2025.02.035

基于改进 YOLO v8n 的辣椒穴盘育苗播种质量轻量级检测方法

孔德航[1,2]　刘云强[1,2]　崔　巍[1,2]　吴海华[1,2]　张学东[1,2]　宁义超[1,2]

(1. 中国农业机械化科学研究院集团有限公司, 北京 100083;
2. 农业装备技术全国重点实验室, 北京 100083)

摘要: 针对辣椒穴盘育苗播种质量实时、准确检测难和边缘设备算力有限等问题,基于 YOLO v8n 设计了一种轻量级检测算法 YOLO v8n-SCS(YOLO v8n improved with StarNet, CAM and SCConv)。采用"单格训练+整盘检测"的技术策略以降低训练成本,提高训练效率。采用 StarNet 轻量级网络和上下文增强模块(Context augmentation module,CAM)作为主干网络,在降低模型复杂程度同时,实现深层特征多感受野信息融合;采用空间与通道重建卷积(Spatial and channel reconstruction convolution,SCConv)优化中间层 C2f(Cross stage partial network fusion)模块的瓶颈结构,增强模块特征提取能力和提高模型推理速度;融合 P2 检测层,将基线 3 个检测头减至 1 个,增强模型对小目标的检测性能。结果表明,YOLO v8n-SCS 模型参数量为 1.2×10^6、内存占用量为 2.7 MB、浮点数运算量为 7.6×10^9,在穴盘单格数据集上,其 mAP_{50} 为 98.3%、mAP_{50-95} 为 83.8%、帧率为 112 f/s,相比基线模型 YOLO v8n,参数量降低 62.5%、mAP_{50} 提升 2.5 个百分点、mAP_{50-95} 提升 2.1 个百分点、浮点数运算量降低 14.6%、帧率提升 23.1%;在整盘检测任务中,其检测帧率为 21 f/s,检测准确率为 98.2%,相比基线模型检测帧率提升 8.2%、准确率提升 1.1 个百分点,对于播种速度 800 盘/h 以内的 72 穴育苗盘和 600 盘/h 以内的 128 穴育苗盘,其平均检测准确率大于 96%,且单粒率、重播率和漏播率检测误差小于 1.8%。本文模型在性能和计算成本之间取得了很好的平衡,降低了部署边缘设备计算要求,满足辣椒穴盘育苗播种质量在线检测需求,为育苗播种生产线智能化升级提供了技术支持。

关键词: 辣椒种子; 穴盘育苗; 播种质量检测; 改进 YOLO v8n; 轻量级模型

中图分类号: TP391; S223　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)02-0381-12　　**OSID**:

Lightweight Detection Method for Seeding Quality of Chili Seedling Trays Based on Improved YOLO v8n

KONG Dehang[1,2]　LIU Yunqiang[1,2]　CUI Wei[1,2]　WU Haihua[1,2]　ZHANG Xuedong[1,2]　NING Yichao[1,2]

(1. *Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing* 100083, *China*
2. *State Key Laboratory of Agricultural Equipment Technology*, *Beijing* 100083, *China*)

Abstract: Aiming to address the challenges of real-time and accurate detection of the seeding quality of chili seedling trays, considering the computing power limitations of edge devices, a lightweight detection algorithm YOLO v8n-SCS (YOLO v8n improved with StarNet, CAM, and SCConv) was proposed based on YOLO v8n. Meanwhile, the technical strategy of "single-cell training + whole-tray detection" was adopted to reduce training costs and improve training efficiency. Firstly, the StarNet lightweight network and the CAM (Context augmentation module) were used as the backbone network to achieve multi-receptive field information fusion of deep features while reducing the complexity of the model. Secondly, the spatial and channel reconstruction convolution (SCConv) was employed to optimize the bottleneck structure of the intermediate layer cross stage partial network fusion (C2f) module to enhance the feature extraction ability of the module and improve the model inference speed. Finally, the P2 detection layer was fused and the detection heads were reduced to one to enhance the model's detection performance for

收稿日期: 2024-10-11　修回日期: 2024-10-30

基金项目: 国家重点研发计划项目(2023YFD2001200)

作者简介: 孔德航(1997—),男,硕士生,主要从事农业机械装备与技术研究,E-mail: k1955055125@163.com

通信作者: 吴海华(1981—),男,研究员,博士生导师,主要从事农业机械装备与技术研究,E-mail: caamswhh@163.com

small targets. The results showed that the YOLO v8n - SCS model had a parameter quantity of 1.2×10^6, a memory occupation of 2.7 MB, and a computation amount of 7.6×10^9 FLOPs. On the single-cell dataset of the seedling trays, its mAP_{50} was 98.3%, mAP_{50-95} was 83.8%, and the frame rate was 112 f/s. Compared with the baseline model YOLO v8n, the parameter quantity was reduced by 62.5%, mAP_{50} was increased by 2.5 percentage points, mAP_{50-95} was increased by 2.1 percentage points, the floating-point operations were reduced by 1.3×10^9, and the frame rate was increased by 23.1%. In the whole-tray detection task, its detection frame rate was 21 f/s and the detection accuracy rate was 98.2%. Compared with the baseline model, the detection frame rate was increased by 8.2% and the accuracy rate was increased by 1.1 percentage points. For 72-cell seedling trays with a seeding speed within 800 trays/h and 128-cell seedling trays with a seeding speed within 600 trays/h, its average detection accuracy was above 96%, and the detection errors of single-seed rate, reseeding rate, and miss-seeding rate were less than 1.8%. This study achieved a good balance between performance and computational cost, reduced the computing power requirements for deploying edge devices, met the online detection needs for the seeding quality of chili seedling trays, and provided key technical support for the intelligent upgrade of the seedling production line.

Key words: chili seeds; plug seedling; seeding quality detection; improved YOLO v8n; lightweight model

0 引言

辣椒是一种重要的经济作物,其种植面积和产量逐年增加,且育苗移栽率超过 80%。穴盘育苗技术通过育苗精量播种生产线进行穴盘苗批量化生产,因其高效、节约和标准化等优势,已成为工厂化育苗重要组成部分[1]。辣椒种子小且扁平,在自动化播种过程中易出现漏播、重播等问题,影响育苗质量和其后续生长。因此,实现对播种质量实时、准确检测是保障育苗播种生产线高效运行的关键。

传统的播种质量检测主要依赖人工观察和抽样统计,存在效率低、误差大、无法实时反馈等问题。相关学者应用传感器技术和图像处理技术实现了白菜[2-3]、水稻[4]、干豆[5]和玉米[6]等种子检测与分类,但此类方法依赖种子表型和尺寸特征,对小籽粒种子检测精度不高。

随着计算机视觉和深度学习技术发展,基于图像识别的检测方法逐渐成为研究热点[7-10]。以上研究在模型架构和检测方式上进行了探索,但未充分考虑边缘设备算力限制。

近年来,诸多学者在模型轻量化改进上进行了研究[11-14]。以上研究对一些主流的单阶段目标检测算法进行了轻量化改进,均衡了模型检测精度和计算量。

综上所述,目前穴盘播种质量检测研究主要集中在水稻、玉米等大籽粒种子上,而对辣椒种子研究较少,且聚焦于提升检测效果,对于检测效率提升和边缘设备算力限制考虑相对较少。辣椒种子尺寸较小,语义信息有限,且呈不规则扁平状,在穴盘中易出现重叠或倾斜,增加了识别难度;在不同光照条件下种子对比度变化明显,其颜色与育苗基质背景相近;育苗基质中含有蛭石、珍珠岩等干扰物,易对检测精度造成影响。

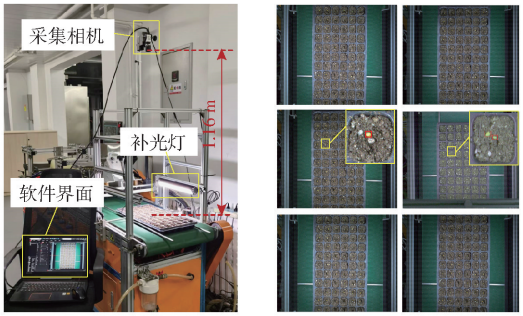
针对以上问题,本文基于改进 YOLO v8n 提出一种辣椒穴盘育苗播种质量轻量级检测方法 YOLO v8n - SCS(YOLO v8n improved with StarNet, CAM and SCConv)。同时,根据穴盘结构特点(单格图像之间背景近似,整盘图像可以视为单格图像拼接而成),采用“单格训练 + 整盘检测”的技术策略以降低训练成本,提高训练效率。在自制的数据集上进行模型检测性能评价和播种单粒率、重播率、漏播率评估,以期为辣椒穴盘育苗播种质量实时、准确检测提供关键技术支持。

1 试验材料

1.1 数据采集

图像数据采集于 2023 年 12 月 13 日,在农业装备技术全国重点实验室进行。试验选用寿禾牛角椒种子作为检测对象,采用 72 穴标准育苗盘,并使用北京中农富通园艺有限公司生产的专用育苗基质土。试验时,将配有 1 200 万像素高清 12 mm 定焦镜头的大华 A3A20CG8 型相机固定在育苗精量播种生产线上方,使镜头前端与传送带保持 1.16 m 的垂直距离,并调节补光灯亮度确保图像清晰。以帧率 30 f/s 进行图像采集,如图 1a 所示。为提高模型泛化能力并确保数据集多样性,图像采集在育苗播种生产线不同播种速度下进行。试验共采集 80 幅图像,每幅图像分辨率为 4 000 像素 × 3 000 像素,其中处于静止和 600、800、1 000、1 200 盘/h 状态下图像数量比为 3∶2∶1∶1∶1。图 1b 为部分采集图像,

其中辣椒种子平均尺寸约为 30 像素 ×30 像素。



(a) 图像采集场景　　(b) 部分采集图像

图 1　数据采集情况

Fig. 1　Data collection diagrams

1.2　穴盘单格数据集构建

采用穴盘单格图像构建模型训练集。为实现单格图像快速提取，降低前期数据处理工作量，首先对采集的穴盘图像进行裁剪，去除穴盘周围背景；再采用模板匹配对裁剪后穴盘图像进行单格图像提取。

鉴于相机固定后采集图像尺寸一致，且穴盘在传送带两侧限位杆作用下横向位置相对居中，因此采用固定区域裁剪方法，依据图像中穴盘的左上角和右下角坐标点进行裁剪，去除无效背景区域，利于后续处理，如图 2 所示。

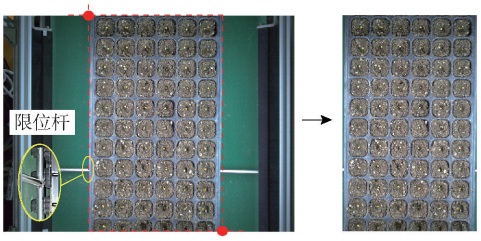


图 2　固定区域裁剪

Fig. 2　Fixed area cropping

鉴于不同穴格背景和亮度存在一定差异，因此采用对光照、噪声有一定鲁棒性的归一化相关模板匹配法进行穴盘单格图像获取[15-16]，如图 3 所示。

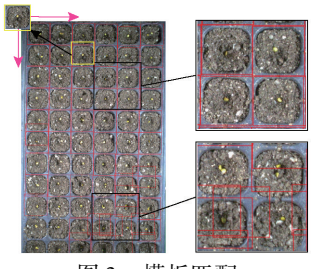


图 3　模板匹配

Fig. 3　Template matching

匹配时，从原图像中选取一穴格作为模板，将模板在原图像上进行像素级滑动，并计算原图像 S 与模板图像 M 之间的归一化相关系数 R，计算式为

$$R(x,y)=\left[\sum_{u=0}^{U-1}\sum_{v=0}^{V-1}(M(u,v)-E(M))\cdot(S(x+u,y+v)-E(S_{xy}))\right]\Big/\left(\sqrt{\sum_{u=0}^{U-1}\sum_{v=0}^{V-1}(M(u,v)-E(M))^2}\cdot\sqrt{\sum_{u=0}^{U-1}\sum_{v=0}^{V-1}(S(x+u,y+v)-E(S_{xy}))^2}\right) \quad (1)$$

式中　x、y——坐标匹配起点，像素

U、V——模板图像 M 尺寸，像素

$M(u,v)$——模板图像 M 在位置 (u, v) 的灰度

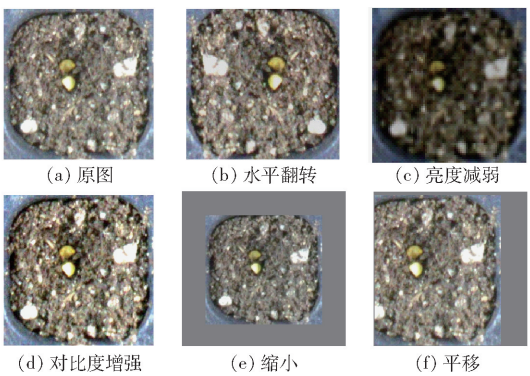
$E(M)$——模板图像 M 灰度均值

$S(x+u, y+v)$——原图像 S 在位置 $(x+u, y+v)$ 的灰度

$E(S_{xy})$——原图像 S 中以 (x, y) 为起点的大小为 $U \times V$ 区域灰度均值

将 R 同设定的匹配置信度 Φ 相比较，若 $R \geqslant \Phi$，则认为在该位置存在匹配，否则认为不存在匹配。试验时，设定 $\Phi = 0.14$，匹配成功率为 84%，共获取到穴盘单格图像 3 472 幅。图中匹配结果显示少数区域存在匹配重叠的问题，因此进一步对获取的图像进行筛选清洗，去除不合格图像，以提高训练数据质量。

为有效提升模型性能和泛化能力，采取随机翻转、颜色变换、随机缩放和平移方式对清洗后 2 876 幅穴盘单格图像进行数据增强，增强效果如图 4 所示。数据增强后数据集扩充为 4 620 幅，使用 LabelMe 软件进行标注，按照比例 8∶1∶1对标注完成的穴盘单格图像进行划分，得到训练集图像 3 696 幅，验证集和测试集图像各 462 幅。



(a) 原图　(b) 水平翻转　(c) 亮度减弱

(d) 对比度增强　(e) 缩小　(f) 平移

图 4　数据增强

Fig. 4　Data augmentation

2　检测方法

2.1　技术策略

本研究技术策略可概括为“单格训练 + 整盘检

测”,如图 5 所示。在单格训练阶段,使用构建的数据集对 YOLO v8n－SCS 网络进行训练,迭代优化模型权重。在整盘检测阶段,首先对采集的整盘图像按照图 2 方法进行固定区域裁剪,减少模型推理数据量;再根据模型下采样深度,确定输入穴盘图像调节目标尺寸 L;最后输入到加载训练最优权重的 YOLO v8n－SCS 网络中进行检测,获取穴盘内辣椒种子播种情况。

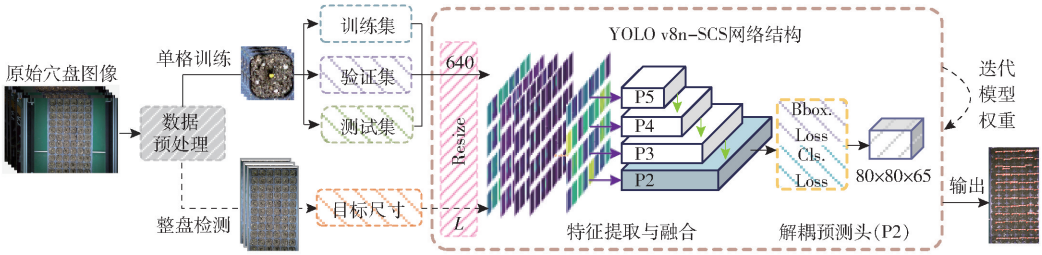


图 5 技术路线图

Fig. 5 Technology roadmap

2.2 YOLO v8n－SCS 网络架构

YOLO 系列算法是单阶段目标检测算法的代表,在目标检测任务中取得了多次 SOTA 表现,相比于 Faster R－CNN[17]、Mask R－CNN[18]等两阶段目标检测算法具有更快的推理速度和更多的版本选择性。YOLO v8[19]是 Ultralytics 公司推出的最新 YOLO 系列算法,相比之前的 YOLO v3[20]、YOLO v5[21]、YOLO v7[22]等版本,在识别精度和速度上都有提升,有 YOLO v8n、YOLO v8s、YOLO v8m、YOLO v8l 和 YOLO v8x 5 个版本。其中,YOLO v8n 作为轻量级版本,在小样本数据集上不容易出现过拟合,且更适合于在资源有限的移动设备端运行。为实现辣椒穴盘播种质量在线检测,精确获取穴盘内辣椒种子位置和数量,以 YOLO v8n 为基线模型进行设计,网络整体依然采用特征金字塔网络(Feature pyramid network,FPN)与路径聚合网络(Path aggregation network,PAN)架构,改进后模型网络结构如图 6 所示。使用轻量级网络 StarNet 对原有的主干网络(Backbone)进行了替换,并引入上下文增强模块(Context augmentation module,CAM)实现深层特征的多感受野信息融合;在中间层(Neck)部分,融合了 P2 检测层,并使用空间与通道重建卷积(Spatial and channel reconstruction convolution,SCConv)替换基线 C2f 模块的瓶颈结构,以实现在增强模块特征提取能力的同时,提高模型推理速度;将基线的 3 个检测头(Head)减至 1 个,集中关注小尺度特征,以

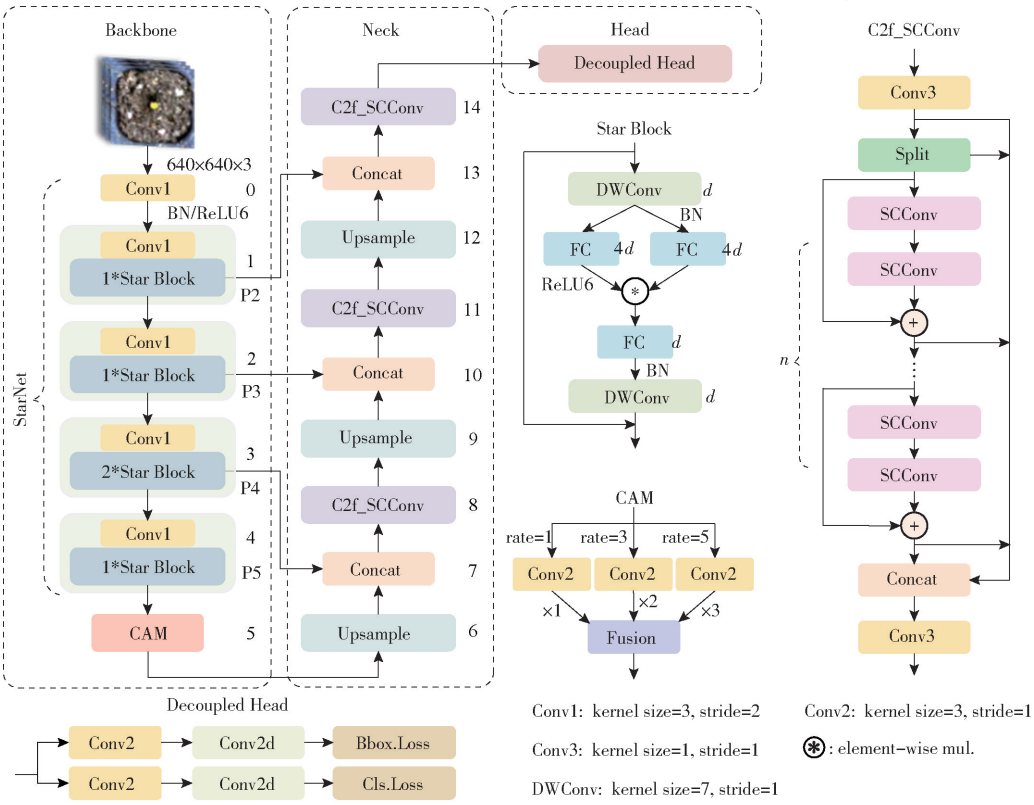


图 6 YOLO v8n－SCS 网络结构图

Fig. 6 Architecture of YOLO v8n－SCS

增强模型对小目标的检测性能。

2.2.1　StarNet 特征提取网络

主干网络选取会显著影响模型整体性能和计算复杂性，因此需兼顾两者间平衡。GhostNet[23]、ShuffleNetV2[24]和 FasterNet[25]等高效网络，在实现轻量化同时，通过特征重用和特征重组来增强对高维非线性特征表征能力，提高网络效率。

StarNet[26]网络采用元素乘法操作显著地提升了自身隐式维度，其网络结构如图 6 所示，对应 YOLO v8n－SCS 主干网络前 5 层，完成 5 次下采样操作。StarNet 遵循传统分层网络结构，首先使用 1 个 3×2 的卷积层降低分辨率，其次通过 5 个由 3×2 的卷积和 Star Block 组成的功能块实现特征提取。而每个功能块 Star Block 的深度（即重复次数）可单独进行设置，大大提升了网络灵活性。

如图 7 所示，Star Block 模块采用深度 7×1 卷积 DWConv（Depth-wise convolution）完成输入特征提取，有效减少了计算量；然后经过批归一化 BN（Batch normalization）高效实现输出的标准化处理；由两路扩展因子为 4 的 1×1 卷积进行特征升维，并选择其中一路使用 ReLU6 进行激活；之后对两路输出执行元素乘法（即 Star），实现全局交互同时，指数式放大特征隐式维数；依次通过 1×1 的卷积和 7×1 的深度卷积实现特征降维和提取。

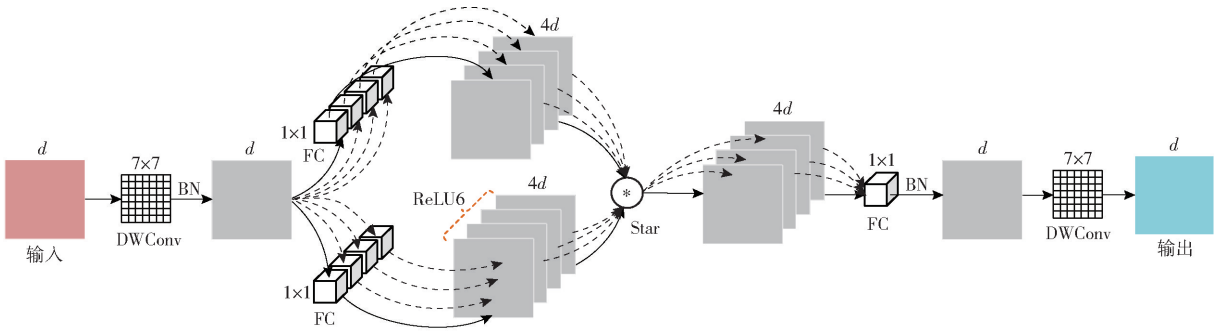


图 7　Star Bolck 模块结构图

Fig. 7　Architecture of Star Bolck module

记网络输入特征为 z，初始特征维数为 d，进行第 1 个星型操作的第 i、j 两路通道权重为 $w^i_{(1,1)}$、$w^j_{(1,2)}$，引入偏置后整体权重为 $W_{(1,1)}$、$W_{(1,2)}$。记“*”表示元素乘法，则第 l 个星型操作输出 O_l 为

$$O_l = \boldsymbol{W}^{\mathrm{T}}_{(l,1)} O_{l-1} \boldsymbol{W}^{\mathrm{T}}_{(l,2)} O_{l-1} \quad (O_{l-1} \in R^{\left(\frac{d}{\sqrt{2}}\right)^{2^l}}) \tag{2}$$

向量空间维数为 $(d/\sqrt{2})^{2^l}$，推导见文献[26]，模型的星型操作次数为 5，当输入通道数为 128 时，网络隐式特征维数约为 90^{32}。可见 StarNet 网络在小巧紧凑的架构上，拥有庞大的高维空间属性。同时，网络采用 ReLU6 作为激活函数，使其在移动端设备运行时也能有很好的分辨率。

2.2.2　CAM 特征增强模块

数据集图像中含有蛭石、珍珠岩等与辣椒种子形态相近，但像素幅值存在差异的干扰物，因此可采取特征增强的方式将干扰物与待检测的辣椒种子加以区分。自注意力机制将特征映射到不同的空间[27]，通过点积相乘构建注意力矩阵。然而，这种实现并不高效，并且导致注意力复杂度随着数量增加呈二次比例增长。基线模型采用快速空间金字塔池化的方式（Spatial pyramid pooling-fast，SPPF）融合不同尺度的上下文信息，但池化层不可学习，并且对于空间分辨率降低的小目标在经历多次池化后会造成该感受野内的目标信息丢失，而导致其信息无法重建[28]。

CAM[29]有效避免了以上问题，该模块采用空洞卷积（Atrous convolution）[30]替代池化（Pooling）操作，在不提高计算量的情况下，结合不同扩张率（Dilation rate）获取图像多尺度信息，其结构见图 6，输入信号在经过扩张率为 1、3、5 的 3×1 空洞卷积后进行特征融合。融合方式共分为 3 种，如图 8 所示。图 8a 和图 8c 分别为加权融合（Weight）和拼接融合（Concat）操作，即将特征图直接注入到空间和通道维度；图 8b 为自适应融合方法（Adaptive），假设将输入尺寸表示为 (b_s, C, H, W)，则可以通过执行卷积操作获得 $(b_s, 3, H, W)$ 的空间自适应权重。

2.2.3　SCConv 卷积模块

在网络中间层部分，影响模型性能的主要是 C2f（Cross stage partial network fusion）跨阶段部分网络融合模块。C2f 模块瓶颈结构由 n 个特征提取模块组成，其中每个模块由两个 3×1 卷积构成，如图 9a 所示。

普通卷积在训练过程中会提取到冗余特征，造成计算资源浪费，采用 SCConv[31]模块替换基线 C2f 模块的瓶颈结构以减少冗余特征提取，提升模型目标检测准确性，如图 9b 所示。

SCConv 利用特征之间的空间和通道冗余来进行卷积压缩，其模块结构如图 10 所示，其中空间重

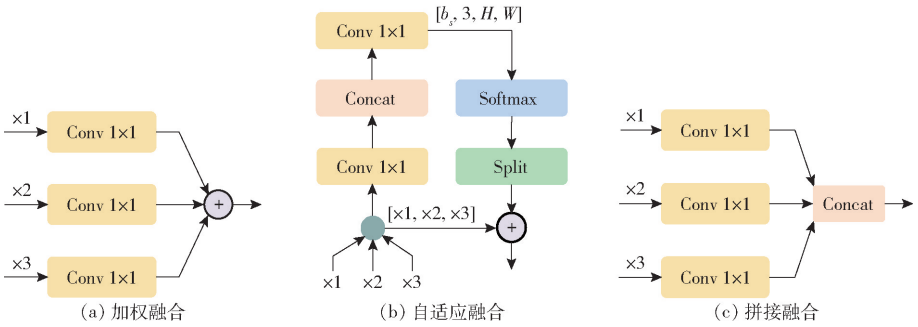


图 8 CAM 模块 3 种融合方式

Fig. 8 Three fusion modes of CAM module

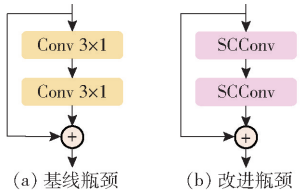


图 9 C2f 改进前后瓶颈结构对比

Fig. 9 Comparison of two bottleneck structures in C2f module

建单元 SRU(Spatial reconstruction unit)和信道重建单元 CRU(Channel reconstruction unit)是 SCConv 中的两个核心模块,两者旨在最大限度地减少特征图中冗余。SRU 主要关注特征图的空间维度,通过设定阈值 T 将归一化后输入特征分成两部分,然后对这两部分进行交叉重构操作,结合加权后两种不同信息量特征,加强彼此间信息流,既精确地捕获了每个特征块特征信息,又显著降低了整体空间冗余。CRU 主要关注特征图通道维度,通过设定分割比 α 将通道划分为两部分,利用逐点卷积 PWC(Point-wise convolution)提取的浅层特征丰富分组卷积 GWC(Group-wise convolution)提取的代表性特征和直接输入特征,并与全局平均池化 GAP(Global average pooling)提取的全局空间信息相融合,通过特征重用处理了冗余特征,提高了模型推理速度。改进 C2f_SCConv 模块如图 6 所示。

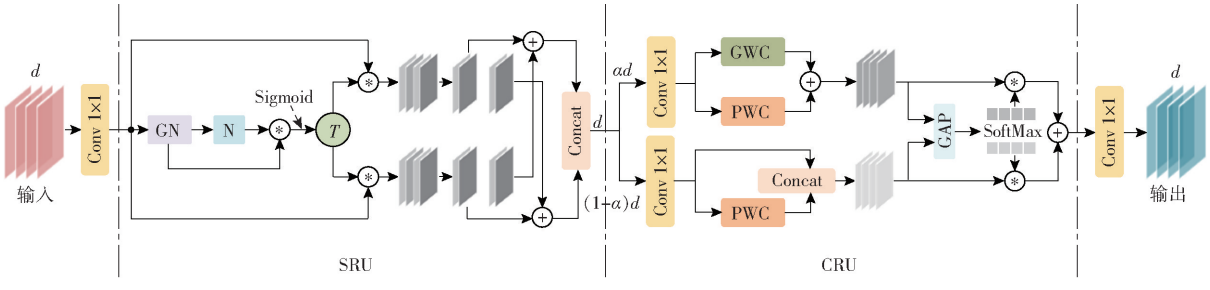


图 10 SCConv 模块结构图

Fig. 10 Architecture of SCConv module

2.3 模型训练与评价指标

模型训练所用 CPU 型号为 Intel Core i7 - 10700 CPU 2.90 GHz, 32 GB 运行内存; GPU 型号为 NVIDIA GeForce RTX 4060 Ti, 16 GB 显存;运行环境为 Windows 10 操作系统。

训练时的输入图像尺寸设置为 640 像素 × 640 像素,批处理大小(Batch size)设置为 18。为加快训练速度,初始权重设置为在 COCO 数据集上预训练网络权重。为防止过拟合和适当加快收敛速度,初始学习率取 0.03,迭代次数为 100。如图 11 所示,在迭代次数为 85 左右时,模型损失值下降到最低。

采用精确率(Precision)、召回率(Recall)、平均精度均值(mAP_{50}、mAP_{50-95})、参数量(Parameters)和浮点数运算量(FLOPs)等作为指标对模型性能进行评价。

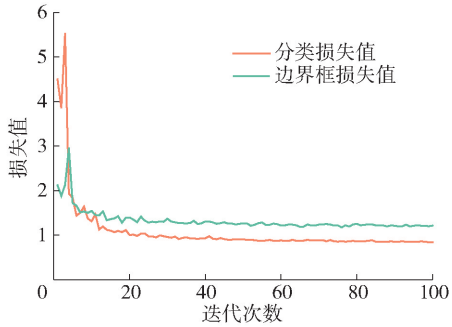


图 11 YOLO v8n - SCS 训练损失值变化曲线

Fig. 11 YOLO v8n - SCS training loss curves

3 试验结果与分析

3.1 单格训练试验

3.1.1 消融试验

为直观地展示本研究方法对检测效果的影响,

以基线模型 YOLO v8n 作对比进行消融试验。试验结果如表 1 所示，采用轻量级网络 StarNet 与特征增强模块 CAM 对主干网络进行改进后，模型参数量减少 12.4%、浮点数运算量降低 1.4×10^9、mAP_{50} 提升 0.7 个百分点、mAP_{50-95} 提升 0.9 个百分点，在不降低精度的同时，实现模型轻量化；引入 P2 小目标特征层、采用 SCConv 替换 C2f 模块瓶颈结构完成对中间层的改进后，模型 YOLO v8n－SCConv－P2 对小目标关注增强，其 mAP_{50} 提升 1.7 个百分点、mAP_{50-95} 提升 1.5 个百分点，但参数量增加 15.3%，浮点数运算量增加 4.9×10^9；检测头削减方案 YOLO v8n－OneHead 在略微降低精确率的同时，显著降低了模型参数量和计算量，弥补了中间层改进带来的问题。集成上述改进就是本研究最终算法 YOLO v8n－SCS，如图 12 所示，该算法能有效地减少因目标倾斜、微弱和重叠导致的漏检和误检，相比基线模型 mAP_{50} 提升 2.5 个百分点、mAP_{50-95} 提升 2.1 个百分点、参数量降低 62.5%、浮点数运算量降低 1.3×10^9。

表 1　消融试验结果

Tab. 1　Ablation experiment results

模型	精确率/%	召回率/%	mAP_{50}/%	mAP_{50-95}/%	参数量	浮点数运算量
YOLO v8n	86.7	88.5	95.8	81.7	3.2×10^6	8.9×10^9
YOLO v8n－StarNet	87.3	86.4	95.4	81.4	2.2×10^6	6.5×10^9
YOLO v8n－CAM	91.1	90.3	96.3	82.8	5.0×10^6	9.8×10^9
YOLO v8n－StarCAM	87.9	89.5	96.5	82.5	2.6×10^6	6.8×10^9
YOLO v8n－P2	87.2	88.6	96.3	82.2	2.9×10^6	1.2×10^{10}
YOLO v8n－SCConv	90.3	91.2	96.8	82.6	3.5×10^6	8.7×10^9
YOLO v8n－SCConv－P2	91.7	92.4	97.5	83.2	3.4×10^6	1.3×10^{10}
YOLO v8n－OneHead	86.8	88.7	95.2	81.3	1.6×10^6	6.1×10^9
YOLO v8n－SCS	88.6	91.3	98.3	83.8	1.2×10^6	7.6×10^9

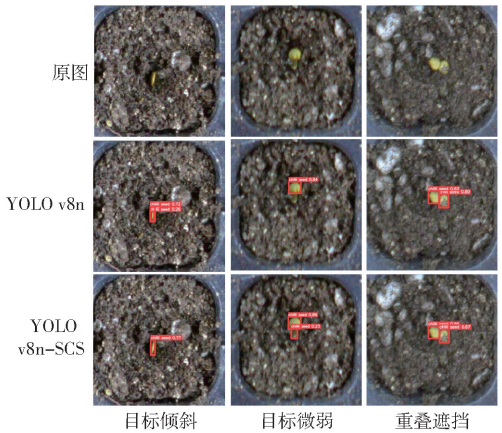


图 12　基线与改进算法检测效果对比

Fig. 12　Comparison of baseline and improved algorithm

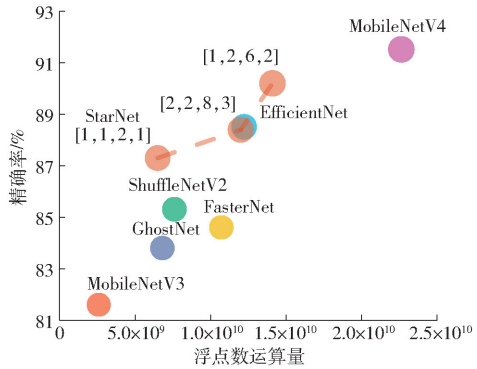


图 13　不同轻量化网络改进效果对比

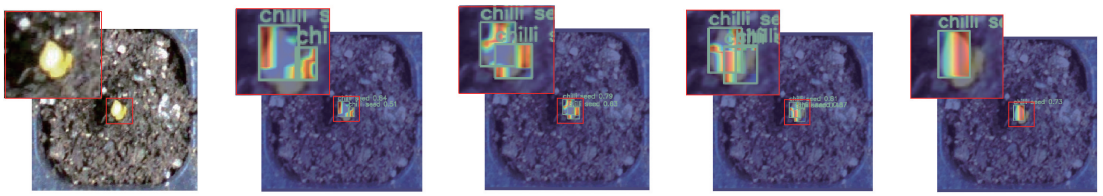
Fig. 13　Comparison of improvement effects of different lightweight networks

3.1.2　对比试验

图 13 预试验选取不同深度的 StarNet 与一些主流轻量化网络 GhostNet、ShuffleNetV2、FasterNet、EfficientNet[32]、MobileNetV3[33] 和 MobileNetV4[34] 对基线模型 YOLO v8n 主干网络改进效果对比情况，可见基于 StarNet 的改进模型在均衡计算量与检测精确率上更有优势。考虑常用边缘设备算力限制，如 Respberry Pi－4B 算力为 1.35×10^{10}、Pentium D 算力为 1.18×10^{10}，即 StarNet 深度取[1，1，2，1]时，其改进模型更能适应较多边缘设备运行需求。

图 14 为 CAM 模块 3 种特征融合方式与原模块 SPPF 产生的热力图对比，检测结果分别为 2、2、3、1。由热力图可知，Concat 方式效果最佳，对重叠且颜色相近种子也能产生较强的特征响应；Weight 和 Adaptive 方式对相交处特征关注过多，导致种子自身区域特征微弱，易造成错检；SPPF 由于池化会造成种子部分细节特征丢失，导致对下方种子感知较弱，造成漏检。热力图差异说明 CAM 更加关注边缘区域特征，同时能够降低特征提取过程中的信息损失，较暗的基质背景增强了模块对种子微弱特征捕捉。

为保证试验公平性，选取模型体量相当且代码开源的目标检测算法（YOLO v5、YOLO v6[35]、YOLO v7、YOLO v9[36]、YOLO v10[37]）轻量级版本在构建的穴盘单格数据集上进行训练，训练次数为 100，结果如表 2 所示。



(a) 原图 (b) 拼接方式 (c) 加权方式 (d) 自适应方式 (e) SPPF

图 14 CAM 与 SPPF 模块产生的热力图对比

Fig. 14 Comparison of heat maps generated by CAM and SPPF

表 2 对比试验结果

Tab. 2 Results of comparative trial

模型	mAP_{50}/%	mAP_{50-95}/%	参数量	浮点数运算量	内存占用量/MB	帧率/(f·s^{-1})
YOLO v8n	95.8	81.7	3.2×10^6	8.2×10^9	5.9	91
YOLO v8n-SCS	98.3	83.8	1.2×10^6	7.6×10^9	2.7	112
YOLO v5n	94.6	81.0	2.5×10^6	7.2×10^9	5.3	77
YOLO v6n	93.5	80.5	4.2×10^6	1.2×10^{10}	8.7	83
YOLO v7-tiny	92.2	78.6	6.0×10^6	1.3×10^{10}	45.0	79
YOLO v9-t	96.1	81.2	2.0×10^7	7.6×10^9	4.6	63
YOLO v10n	96.5	81.5	2.7×10^6	8.2×10^9	5.8	90

由表 2 可知,改进模型 YOLO v8n-SCS 拥有最高的检测平均精度均值 mAP_{50} 和 mAP_{50-95},模型内存占用量为 2.7 MB,浮点数运算量仅为 7.6×10^9,所需计算量比 YOLO v5n 之外的其他检测算法都少,且最高检测帧率为 112 f/s,其综合表现不弱于后续的 SOTA 模型 YOLO v9-t、YOLO v10n。分析其原因,主要是因为 YOLO 后续版本的改进主要集中于优化模型架构、提升模型性能和适应性方面,而非针对提升小目标检测能力,因此在面对小目标检测时表现欠佳。可见,本文算法 YOLO v8n-SCS 在提升检测性能的同时,减少了模型整体参数量、降低了算力需求、提高了检测效率,证明了本改进模型的有效性与先进性。

3.2 整盘检测试验

3.2.1 输入图像尺寸对整盘检测效果影响

相关研究表明,较大输入图像检测结果更接近实际[38],但缺少量化标准。本文模型训练与预测任务所用穴盘单格图像与整盘图像长边尺寸比为 1∶14,使得两种任务下提取的输入特征差异很大,即需探究输入图像尺寸变化对模型检测效果的影响规律,以保证模型良好的识别效果。鉴于模型主干网络对输入图像共进行 5 次下采样(最小特征图为原图的 $1/2^5$),为保证每个像素点有效,应将图像输入尺寸调整为 32 整数倍。其输入图像尺寸调节原则为

$$\begin{cases}32(k-1)\leqslant\max(w,h)\leqslant32k\\L=32(k+n)\end{cases}\tag{3}$$

式中 w、h——原图像宽、高,像素

其中,基系数 k 根据原图像尺寸确定,可变系数 n 预先设定,两者均为整数,后者负责约束调节目标尺寸 L,调节后图像宽、高变为原来的 $L/\max(w,h)$。

从采集图像中随机抽取 5 幅作为原图样本,将其按照图 2 方式裁剪后作为裁剪样本,2 组尺寸分别为 4 000 像素×3 000 像素、1 550 像素×3 000 像素。经测试,采用 OpenCV 进行单幅图像裁剪耗时远小于 0.01 ms,因此可忽略该预处理操作的影响。以每组 5 幅图像检测平均耗时和平均准确率为评价指标,分别采用基线模型 YOLO v8n 与改进模型 YOLO v8n-SCS 对裁剪前后图像按照不同输入尺寸进行预测,试验结果如图 15 所示。

由图 15 可知,平均耗时与 n 变化呈正相关,平均准确率与 n 变化呈二次相关,固定区域裁剪能够显著缩减检测耗时近 75%、程度性尺寸放大操作能够平均提高检测准确率约 0.86 个百分点,且 YOLO v8n-SCS 相比 YOLO v8n 具有更好的预测表现。这是因为大尺寸图像所需计算量大,从而增加了检测耗时,而裁剪操作缩减了输入图像尺寸,有助于提高检测效率。同时,尺寸缩减会导致图像部分细节特征丢失,而尺寸增加到一定范围时,插值带来的噪声对小目标影响不可忽略,导致准确率下降。图 15b 呈现出个别小尺寸输入相比大尺寸输入拥有更高的检测平均准确率,这归因于此时特征图尺寸更接近网络最佳特征提取尺寸。

当 $n\in[10,20]$ 时,两模型检测平均准确率表现最佳,其中 YOLO v8n-SCS 检测平均准确率大于等于 96%,结合检测平均耗时,取 $n=11$,此时模型检测帧率为 21 f/s,检测准确率为 98.2%,相比基线

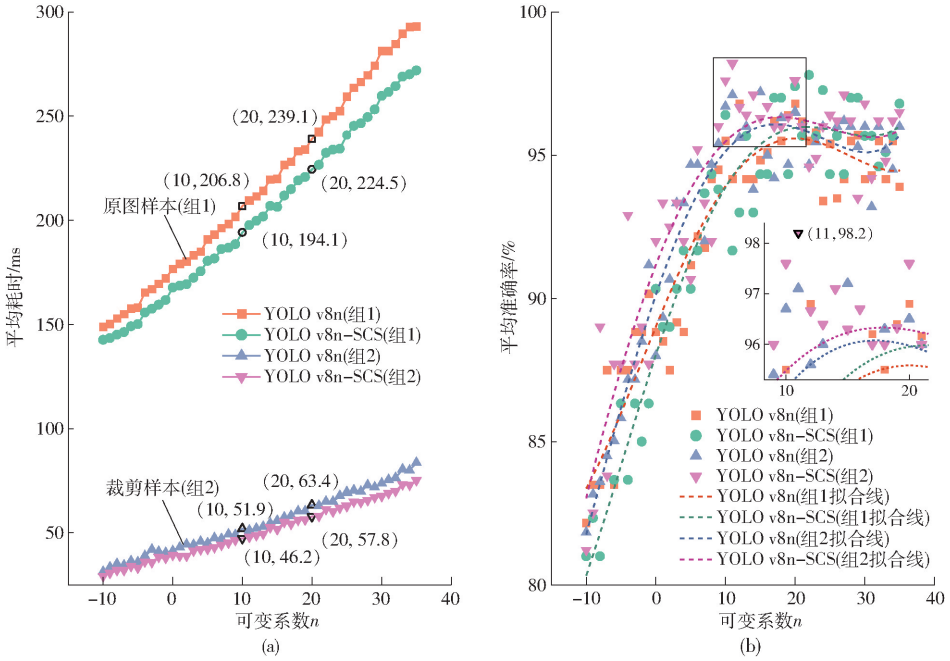


图 15　输入图像尺寸对 2 种模型预测结果的影响

Fig. 15　Effects of different input image sizes on prediction results of two models

模型检测帧率提升 8.2%、准确率提升 1.1 个百分点。

3.2.2　整盘检测与分块检测效果对比

由上述分析可知，图像尺寸越小，检测耗时越短，为综合比较整盘检测效果，以 5 幅裁剪后图像作为试验对象，依照图 16 方法对其分块。选取每种分块方式下 5 幅图像进行预测，使可变系数 n 为 11，按照式(3)设定图像输入尺寸，得到检测耗时和平均准确率随划分块数变化情况如图 17 所示。其中，预估整盘耗时近似为单块检测耗时与划分块数的乘积，忽略单块图像预加载时长。

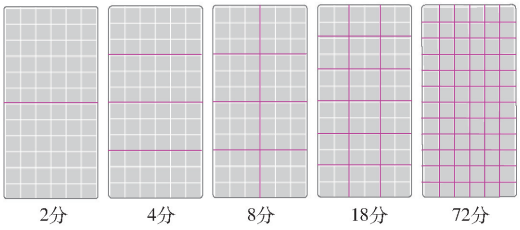


图 16　穴盘分块检测划分方法

Fig. 16　Dividing methods for regional detection of seedling trays

由图 17a 可知，对整盘进行划分后，随着分块数量增加，单块图像检测耗时降低，但整盘检测耗时会逐渐增加；由图 17b 可知，相比分块检测，整盘检测对模型检测准确率损失低于 0.31 个百分点。综合分析，整盘检测更有优势，在实现高效率和准确率的同时，其后处理也更方便，同时 YOLO v8n－SCS 相比基线模型拥有更佳的性能表现。

3.2.3　整盘检测效果评价

为综合评价播种质量检测效果，将训练好的模型集成到试验平台嵌入式系统(英伟达 Jetson Nano B401 开发板)，并按照前文结论对输入图像尺寸进行设置，如图 18 所示。试验准备 72 穴和 128 穴两种规格育苗盘各 5 盘，分别在不同播种速度下进行测试。其中，图像获取通过穴盘到位触发激光红外

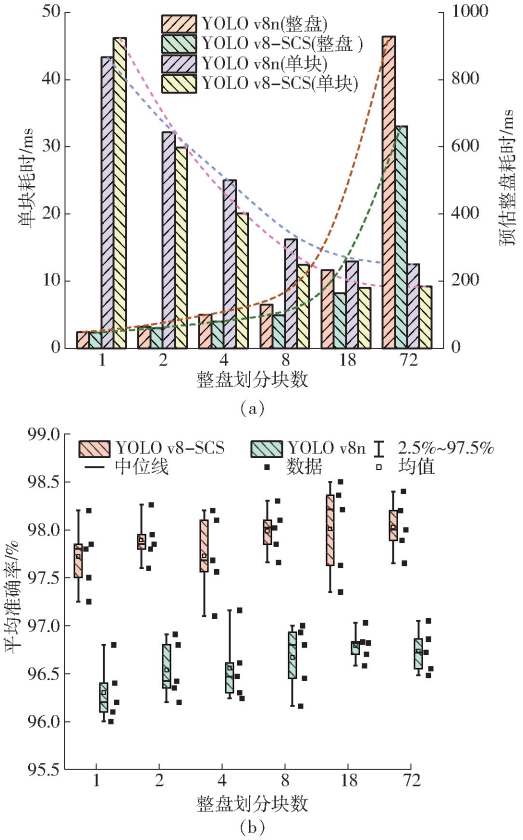


图 17　不同分块方法检测效果对比

Fig. 17　Comparison of detection results of different dividing methods

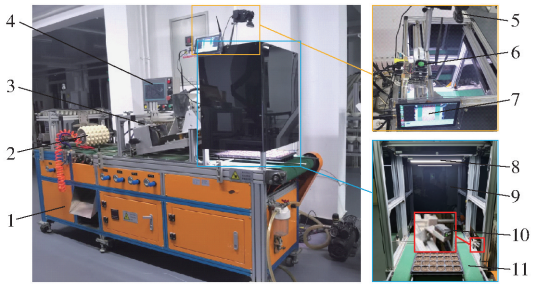


图 18 试验平台

Fig. 18 Test platform

1. 装置箱体 2. 压穴装置 3. 播种装置 4. 控制系统 5. 摄像头 6. Jetson Nano 嵌入式系统 7. 触摸屏 8. 补光灯 9. 遮光板 10. 激光红外传感器 11. 传送带

传感器来进行关键帧采集。

以检测准确率为评价指标,得到部分整盘检测效果图如图 19a 所示(黄色框中为漏检种子),不同播种速度下整盘检测准确率如图 19b 所示,3 个指标随播种速度变化曲线如图 19c、19d 所示。

可见,随着播种速度的增加,整盘检测准确率随之下降,同时检测单粒率、重播率随之下降,而检测漏播率随之上升。这主要是由于相机帧率限制,在高播种速度下采集图像会出现程度性混合模糊[39],这种像素级变化会致使特征微弱、重叠和遮挡状态下的辣椒种子丢失部分特征,影响其特征提取效果,导致模型整盘检测准确率下降,其中

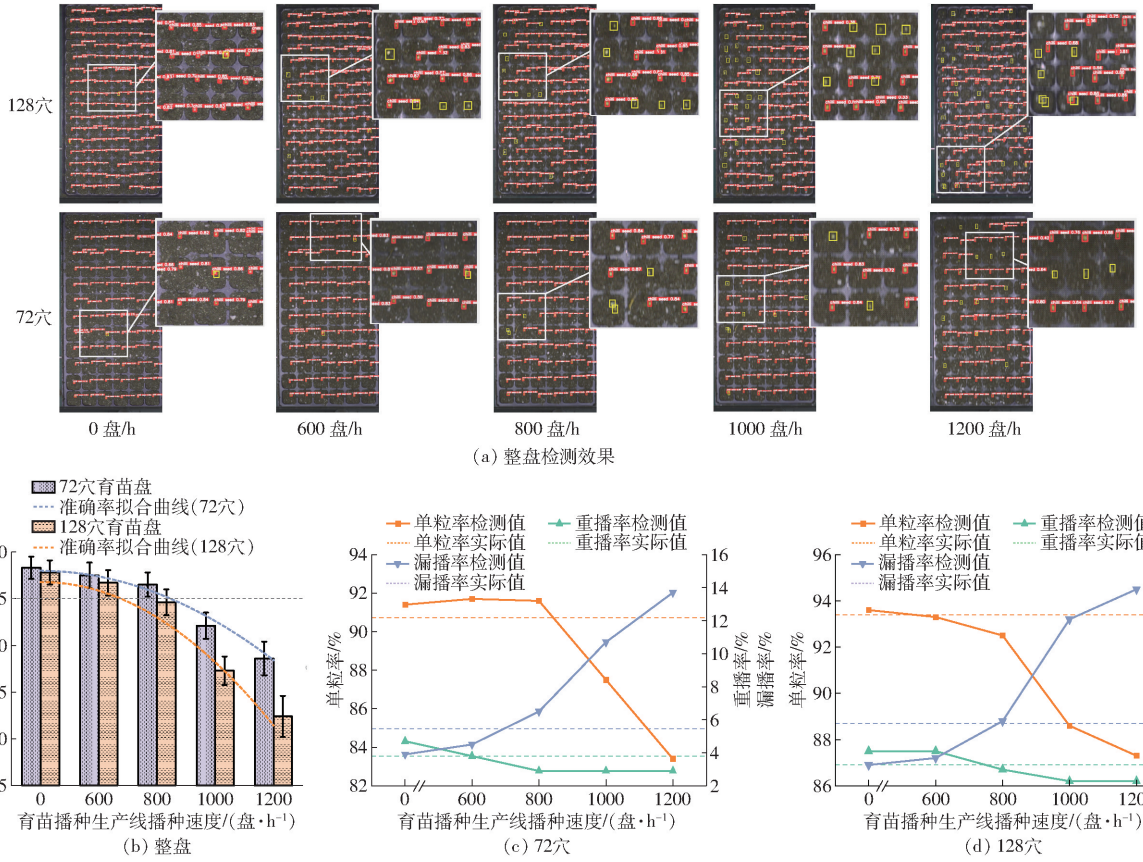


图 19 YOLO v8n - SCS 整盘检测效果

Fig. 19 Whole-tray detection results of YOLO v8n - SCS

128 穴育苗盘播种种子更多,准确率下降也更加明显。在本试验条件下,改进模型 YOLO v8n - SCS 对 72 穴育苗盘静态检测准确率为 98.3%,在播种速度 600 盘/h 以内检测准确率不小于 97.5%,在播种速度 800 盘/h 以内检测准确率不小于 96.5%;对 128 穴育苗盘静态检测准确率为 97.8%,在播种速度 600 盘/h 以内检测准确率不小于 96.7%,即在以上工况下,改进模型检测精度不小于 95%,同时检测单粒率、重播率和漏播率波动维持在 1.8% 以内。因此,本文模型可用于辣椒穴盘育苗精量播种质量检测。

4 结论

(1)以 YOLO v8n 为基线模型提出了一种改进模型 YOLO v8n - SCS:采用 StarNet 网络轻量化 Backbone,并引入 CAM 模块进行特征增强;融合 P2 小目标特征层,采用 SCConv 模块替换基线 C2f 模块的瓶颈结构完成对 Neck 部分改进;将基线检测头 Head 削减到 1 个,得到改进模型 YOLO v8n - SCS。

(2)在穴盘单格数据集训练表现上,相比基线模型 YOLO v8n,YOLO v8n - SCS 参数量降低 62.5%、mAP_{50} 提升 2.5 个百分点、mAP_{50-95} 提升

2.1 个百分点、浮点数运算量降低 1.3×10^9；相比主流检测算法，YOLO v8n－SCS 模型内存占用量仅为 2.7 MB，浮点数运算量仅为 7.6×10^9，检测帧率达 112 f/s，对穴盘辣椒种子检测效果更加显著。

(3)在整盘检测表现上，首先分析了输入图像尺寸对整盘检测效果影响，得出该任务下通过固定区域裁剪去除穴盘图像周围背景可缩减检测耗时近 75%。当可变系数 n 为 11 时，按照式(3)调节输入图像尺寸检测效果最佳，此时模型检测帧率为 21 f/s，检测准确率为 98.2%，相比基线模型检测帧率提升 8.2%、准确率提升 1.1 个百分点。其次，由分块检‖测效果对比试验得出，分块方式检测帧率低于 16 f/s，且对模型检测准确率提升效果有限。综合来看，整盘检测更具优势。

(4)验证了 YOLO v8n－SCS 在育苗播种生产线不同播种速度下的整盘检测效果。结果显示，改进模型 YOLO v8n－SCS 对 72 穴育苗盘静态检测准确率为 98.3%，在播种速度 800 盘/h 以内检测准确率不小于 96.5%；对 128 穴育苗盘静态检测准确率为 97.8%，在播种速度 600 盘/h 以内检测准确率不小于 96.7%。在以上条件下，改进模型能够取得不小于 95% 的检测精度，同时单粒率、重播率和漏播率检测误差小于 1.8%。因此，本文模型可通过结合关键帧检测的方式，实现辣椒穴盘育苗精量播种质量在线检测。

参 考 文 献

[1] 杨昌敏，易文裕，邱云桥，等. 育苗精量播种机研究现状及发展分析[J]. 中国农机化学报，2022，43(4)：183－188.
YANG Changmin，YI Wenyu，QIU Yunqiao，et al. Research status and development analysis of seed-raising precision seeder [J]. Journal of Chinese Agriculture Mechanization，2022，43(4)：183－188.（in Chinese）

[2] 胡建平，陆黎. 磁吸式穴盘播种器图像监控系统设计[J]. 农业机械学报，2006，37(11)：88－91.
HU Jianping，LU Li. Design of the image monitored control system for the magnetic type precision seeder[J]. Transactions of the Chinese Society for Agricultural Machinery，2006，37(11)：88－91.（in Chinese）

[3] XIA H，ZHEN W，LIU Y，et al. Optoelectronic measurement system for a pneumatic roller-type seeder used to sow vegetable plug-trays[J]. Measurement，2021，170：108741.

[4] 周海波，李晨铭，张镔，等. 水稻钵体育秧智能补种决策方法与试验[J]. 农业机械学报，2020，51(12)：122－129.
ZHOU Haibo，LI Chenming，ZHANG Bin，et al. Decision and experiment on intelligent reseeding system of rice bowl nursing seedling[J]. Transactions of the Chinese Society for Agricultural Machinery，2020，51(12)：122－129.（in Chinese）

[5] KOKLU M，OZKAN I A. Multiclass classification of dry beans using computer vision and machine learning techniques[J]. Computers and Electronics in Agriculture，2020，174：105507.

[6] BAI J，HAO F，CHENG G，et al. Machine vision-based supplemental seeding device for plug seedling of sweet corn[J]. Computers and Electronics in Agriculture，2021，188：106345.

[7] 董文浩. 基于嵌入式系统的机器视觉杂交稻智能恒量播种装置研究[D]. 广州：华南农业大学，2020.
DONG Wenhao. Research on machine vision hybrid rice intelligent constant seeding device based on embedded system[D]. Guangzhou：South China Agricultural University，2020.（in Chinese）

[8] DENG X，LIANG S. Detection of rice seed hole seeding amount based on deep learning[C]//2023 IEEE 6th Information Technology，Networking，Electronic and Automation Control Conference（ITNEC）. IEEE，2023：1110－1114.

[9] 闫泽禹. 番茄穴盘自动播种装置与漏播检测系统设计及关键技术研究[D]. 杨凌：西北农林科技大学，2024.
YAN Zeyu. Design and key technology research of tomato plug tray automatic planting device and missing seeding detection system[D]. Yangling：Northwest A&F University，2024.（in Chinese）

[10] SABANCI K，ASLAN M F，ROPELEWSKA E，et al. A convolutional neural network-based comparative study for pepper seed classification：analysis of selected deep features with support vector machine[J]. Journal of Food Process Engineering，2022，45(6)：e13955.

[11] LI J，LI J，ZHAO X，et al. Lightweight detection networks for tea bud on complex agricultural environment via improved YOLO v4[J]. Computers and Electronics in Agriculture，2023，211：107955.

[12] THAKURIA A，ERKINBAEV C. Real-time canola damage detection：an end-to-end framework with semi-automatic crusher and lightweight ShuffleNetV2_YOLOv5s[J]. Smart Agricultural Technology，2024，7：100399.

[13] 苗荣慧，李志伟，武锦龙. 基于改进 YOLO v7 的轻量化樱桃番茄成熟度检测方法[J]. 农业机械学报，2023，54(10)：225－233.
MIAO Ronghui，LI Zhiwei，WU Jinlong. Lightweight maturity detection of cherry tomato based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(10)：225－233.（in Chinese）

[14] 罗斌，李家超，周亚男，等. 基于改进 YOLO v8 的轻量化稻瘟病孢子检测方法[J]. 农业机械学报，2024，55(11)：32－38.
LUO Bin，LI Jiachao，ZHOU Ya'nan，et al. Lightweight rice blast spores detection method based on improved YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(11)：32－38.（in Chinese）

[15] SARVAIYA J N，PATNAIK S，BOMBAYWALA S. Image registration by template matching using normalized cross-correlation

[C]//2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies. IEEE, 2009: 819 - 822.
[16] BRUNELLI R. Template matching techniques in computer vision: theory and practice[M]. John Wiley & Sons, 2009.
[17] GIRSHICK R. Fast R - CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2015:1440 - 1448.
[18] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R - CNN[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961 - 2969.
[19] JIANG P, ERGU D, LIU F, et al. A review of Yolo algorithm developments[J]. Procedia Computer Science, 2022, 199: 1066 - 1073.
[20] REDMON J, FARHADI A. Yolov3: an incremental improvement[J]. arXiv Preprint, arXiv:1804.02767, 2018.
[21] HORVAT M, JELEČEVIĆ L, GLEDEC G. A comparative study of YOLOv5 models performance for image localization and classification[C]//Central European Conference on Information and Intelligent Systems. Faculty of Organization and Informatics Varazdin, 2022:349 - 356.
[22] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:7464 - 7475.
[23] HAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020:1580 - 1589.
[24] MA N, ZHANG X, ZHENG H T, et al. Shufflenet v2: practical guidelines for efficient cnn architecture design[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018:116 - 131.
[25] CHEN J, KAO S hong, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:12021 - 12031.
[26] MA X, DAI X, BAI Y, et al. Rewrite the stars[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024:5694 - 5703.
[27] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale[J]. arXiv Preprint, arXiv:2010.11929, 2020.
[28] ZAFAR A, AAMIR M, MOHD N N, et al. A comparison of pooling methods for convolutional neural networks[J]. Applied Sciences, 2022, 12(17):8643.
[29] KISANTAL M. Augmentation for small object detection[J]. arXiv Preprint, arXiv:1902.07296, 2019.
[30] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J]. arXiv Preprint, arXiv:1706.05587, 2017.
[31] LI J, WEN Y, HE L. Scconv: spatial and channel reconstruction convolution for feature redundancy[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023:6153 - 6162.
[32] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[C]//International Conference on Machine Learning, PMLR, 2019:6105 - 6114.
[33] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea: IEEE, 2019:1314 - 1324.
[34] QIN D, LEICHNER C, DELAKIS M, et al. MobileNetV4—universal models for the mobile ecosystem[J]. arXiv Preprint, arXiv:2404.10518, 2024.
[35] LI C, LI L, JIANG H, et al. YOLOv6: a single-stage object detection framework for industrial applications[J]. arXiv Preprint, arXiv:2209.02976, 2022.
[36] WANG C Y, YEH I H, LIAO H Y M. Yolov9: learning what you want to learn using programmable gradient information[J]. arXiv Preprint, arXiv:2402.13616, 2024.
[37] WANG A, CHEN H, LIU L, et al. Yolov10: real-time end-to-end object detection[J]. arXiv Preprint, arXiv:2405.14458, 2024.
[38] RUKUNDO O. Effects of image size on deep learning[J]. Electronics, 2023, 12(4):985.
[39] ZHANG K, REN W, LUO W, et al. Deep image deblurring: a survey[J]. International Journal of Computer Vision, 2022, 130(9):2103 - 2130.