

CottonBud-YOLOv5s 轻量型棉花顶芽检测算法

赵露强¹, 彭强吉^{2,3}, 兰玉彬¹, 康建明^{2,3}, 张敬文¹, 代建龙⁴, 陈玉龙^{1*}

(1. 山东理工大学农业工程与食品科学学院, 淄博 255000; 2. 山东省农业机械科学研究院, 济南 250100; 3. 山东省丘陵山区智慧农业装备重点实验室, 济南 250100; 4. 山东省农业科学院, 济南, 250100)

摘要: 针对棉花机械打顶作业过程中, 边缘移动设备算力受限实时性差, 运动模糊、小目标遮挡导致难以检测的问题, 该研究基于 YOLOv5s 模型提出 CottonBud-YOLOv5s 轻量型棉花顶芽检测模型, 该模型采用 ShuffleNetv2 主干网络和 DySample 动态上采样模块替换原始模块降低计算量, 提高模型检测速度; 头部 (head) 和颈部 (neck) 分别引入 ASFFHead 检测头和 GC (global context) 全局注意力模块增强模型尺度不变性和上下文特征提取能力, 提高小目标遮挡和运动模糊图像的检测性能。通过消融试验和模型对比试验, 验证 CottonBud-YOLOv5s 棉花顶芽检测模型的可行性。试验结果表明: 引入 ASFFHead 检测头和 GC 全局注意力机制后, 小目标平均精度 $AP_{0.5:0.95}$ 和平均召回率 $AR_{0.5:0.95}$ 值比引入前分别提升 3.6、2.1 个百分点, 中目标平均精度 $AP_{0.5:0.95}$ 和平均召回率 $AR_{0.5:0.95}$ 值分别提升 4.1、3.5 个百分点, 大目标平均精度 $AP_{0.5:0.95}$ 和平均召回率 $AR_{0.5:0.95}$ 值分别提升 6.5、5.9 个百分点; 与 Faster-RCNN、TOOD、RTDETR、YOLOv3s、YOLOv5s、YOLOv9s 和 YOLOv10s 检测模型相比检测速度分别提升 26.4、26.7、24.2、24.8、11.5、18.6、15.6 帧/s, 平均精度均值分别提升 14.0、13.3、5.5、0.9、0.8、0.2、1.5 个百分点, 召回率分别提升 16.8、16.0、3.2、2.0、0.8、0.5、1.2 个百分点, CottonBud-YOLOv5s 模型平均精度均值达到 97.9%, 召回率达到 97.2%, CPU 检测速度达到 27.9 帧/s。由模型可视化分析可知 CottonBud-YOLOv5s 模型在单株、多株、运动模糊、小目标遮挡的整体检测性能优于其他检测模型。该模型具有较高的检测精度、鲁棒性和检测速度, 适用于密植环境下棉花顶芽的精准检测, 可为棉花机械化打顶提供视觉检测基础。

关键词: 目标检测; 遮挡; 运动模糊; 小目标; 棉花顶芽; 卷积神经网络; YOLOv5s

doi: 10.11975/j.issn.1002-6819.202407140

中图分类号: TP181; TP183; S24

文献标志码: A

文章编号: 1002-6819(2025)-04-0175-10

赵露强, 彭强吉, 兰玉彬, 等. CottonBud-YOLOv5s 轻量型棉花顶芽检测算法[J]. 农业工程学报, 2025, 41(4): 175-184.

doi: 10.11975/j.issn.1002-6819.202407140 <http://www.tcsae.org>

ZHAO Luqiang, PENG Qiangji, LAN Yubin, et al. CottonBud-YOLOv5s lightweight cotton bud detection algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(4): 175-184. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202407140 <http://www.tcsae.org>

0 引言

棉花是中国的重要经济作物和战略资源^[1]。打顶作业是棉花种植过程中的重要环节, 通过去除其主茎顶芽来抑制顶端生长优势, 促使棉花多结铃, 提高种植效益^[2]。目前, 棉花打顶主要采用人工打顶、化控打顶和机械打顶 3 种方式^[3]。人工打顶劳动强度大、效率低; 化控打顶技术要求高, 需要多次喷施, 危害劳动人员健康, 造成环境污染^[4]; 机械打顶具有节约劳动力、操作简便、效率高等优点, 是棉花打顶的必然选择^[5]。随着农业科技的发展, 机械打顶正朝着智能化和精准化的方向迅速发展^[6]。如何提高田间密植环境下棉花顶芽的识别精度, 是实现棉花智能化精准打顶作业的关键前提。

早期研究者主要通过激光、红外、超声波等技术检测棉花顶芽位置。史增录等^[7]利用激光对射传感器检测植株高度来确定棉花顶芽位置; 王玲等^[8]运用红外测距实验装置和计算机图像处理技术结合棉花的农学生长特性测量实现了单朵棉花的测距; 闫毅敏等^[9]设计了一种基于红外测距传感器和超声波测距传感器的棉株顶尖高度测量系统, 提高了对棉株顶尖的识别精度。该类方法主要通过检测棉花植株高度来控制机械机构进行打顶, 误差较大。

卷积神经网络的出现, 推动目标检测进入了全新的阶段, 研究人员针对不同的农业应用场景, 开展改进算法研究, 以进一步提高检测精度和效率^[10-13]。范迎迎等^[14]提出一种 BP 神经网络算法的棉花识别方法, 能够有效提高棉花识别精度并简化识别过程; KOIRALA 等^[15]应用 YOLOv3 算法来自动识别芒果生长过程中的圆锥花序数量, 但检测时实时性较差; 张日红等^[16]基于 YOLOv4 检测算法提出轻量化菠萝苗心检测模型, 实时性良好但检测准确率有待进一步提高; 彭明霞等^[17]在 RPN 网络中引入特征金字塔网络生成目标候选框对 Faster R-CNN 结构进行优化, 在棉间杂草检测上取得较好的效果, 但

收稿日期: 2024-07-16 修订日期: 2025-01-06

基金项目: 山东省农机研发制造推广应用一体化试点项目 (NJYTHSD-202317)

作者简介: 赵露强, 研究方向为视觉检测与智能控制。

Email: luqiang.zhao@qq.com

※通信作者: 陈玉龙, 博士, 副教授, 研究方向为智能农机装备。

Email: cyl06471@sdu.edu.cn

计算量增加使检测速度变慢; WU 等^[18] 基于 YOLOv4 引入残差块增强小目标杂草的检测能力, 但模型计算量增大; 刘海涛等^[19] 提出了一种基于 YOLOv4 算法的识别方法, 该方法未能有效识别顶芯, 实际打顶效果有待进一步提高。XU 等^[20] 基于 YOLOv5 检测算法增加 P2 特征层, 增强害虫密集目标检测性能的同时, 也造成计算量增大、实时性变差; 张新月等^[21] 提出改进型 YOLOv8n 轻量化红花识别模型, 该模型计算量小、识别速度快, 但检测准确率低于基础模型; 陈柯屹等^[22] 提出改进型 Faster-RCNN 检测算法, 解决密植环境中棉花顶芽的精准识别问题, 但是 Fast R-CNN 算法需要采集大量的数据训练模型, 在样本不足时检测准确度明显降低; 彭炫等^[23] 基于 YOLOv5s 算法增加小目标检测层和 CPP-CBAM 注意力机制, 提升了棉花顶芽检测的精度, 同时增大了模型计算量。

上述学者针对小目标检测困难、实时性能差等问题开展研究, 取得了一些有益效果, 但在运动模糊、小目标遮挡检测方面存在巨大挑战。本文以密植环境下棉花顶芽为检测对象, 针对棉花打顶作业过程中边缘移动设备算力受限难以满足实时要求, 运动模糊、小目标遮挡难以检测的问题, 提出一种轻量型棉花顶芽检测模型为 CottonBud-YOLOv5s, 该模型以 YOLOv5s 模型为基础, 通过替换主干网络为 ShuffleNetv2 轻量级主干网络和 DySample 轻量化动态上采样模块, 显著减少计算量并提升检测速度; 引入自适应空间特征融合检测头 ASFFHead 和 GC (global context) 全局注意力机制, 增强特征的尺度不变性和上下文特征的提取能力, 以期提高对运动模糊图像、小目标遮挡的检测性能, 为棉花机械化打顶视觉检测提供基础。

1 材料与方法

1.1 棉花顶芽图像采集

2023 年 7 月, 在山东省滨州市无棣县景国合作社机采棉基地进行棉花顶芽图像的采集工作。采集时间段为每天 08:00—19:00。拍摄设备为索尼 HDR-CX350E, 拍摄方式为垂直向下拍摄棉株顶部。选取 1 200 张棉花顶芽图像, 建立棉花顶芽数据集。数据集包含不同时间段下棉花顶芽单株、多株、运动模糊、遮挡等情况的图像, 如图 1 所示。

1.2 构建数据集

使用 LabelImg 工具对选取的 1 200 张原始图片进行单类别标注, 标注内容为棉花顶芽, 标记为 bud。为提升模型泛化能力, 通过 Python 脚本对每张标注好的图片进行数据增强处理, 包含上下镜像、左右镜像各一次, 随机旋转两次, 随机亮度调整一次, 合计生成 7 200 张带有标记的棉花顶芽图像。按照 8:1:1 的比例将数据增强后的图像随机分为训练集 (5 760 张)、验证集 (720 张) 和测试集 (720 张), 供后续模型训练和测试任务使用。

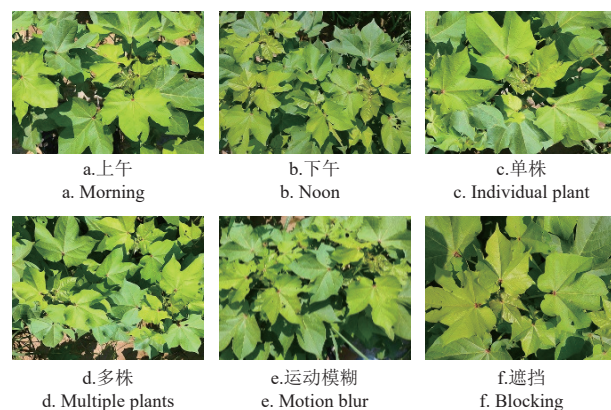


图 1 棉花顶芽样本图像

Fig.1 Cotton bud sample image

2 棉花顶芽检测算法

2.1 CottonBud-YOLOv5s 网络模型

YOLOv5 为一阶段目标检测算法, 其性能良好, 是 YOLO 算法系列中训练时间和推理速度较优的模型, 其中, YOLOv5s 网络深度和特征图宽度较小, 相较于其他版本推理速度快, 更适合部署于边缘移动设备^[24-25], 本文基于 YOLOv5s 模型对棉花顶芽检测模型进行改进。

将 ShuffleNetv2 轻量级主干网络替代原始 CSPDarknet53 主干网络, 同时 DySample 轻量化动态上采样模块替换最邻近上采样模块, 动态调整上采样策略, 减少计算量并提高模型检测速度; 为提高运动模糊图像、小目标遮挡的检测性能, 基于数据驱动金字塔特征融合策略 ASFF 提出自适应空间特征融合检测头 ASFFHead, 引入 GC 全局注意力机制, 增强模型尺度不变性和上下文特征提取能力。具体改进方法如下: 第 1 步, 在 yolov5s.yaml 中更改主干网络为 ShuffleNetv2 主干网络、更改上采样为 DySample 动态上采样、更改检测头 Head 为 ASFFHead, 添加 GC 全局注意力机制; 第 2 步, 将 ShuffleNetv2、DySample、ASFFHead、GC 模块增加至 common.py 文件中; 第 3 步, 通过试验进一步调整网络层结构和训练参数, 实现改进模块与 YOLOv5s 目标检测网络的集成, 并将改进后模型命名为 CottonBud-YOLOv5s, 模型网络结构如图 2 所示。

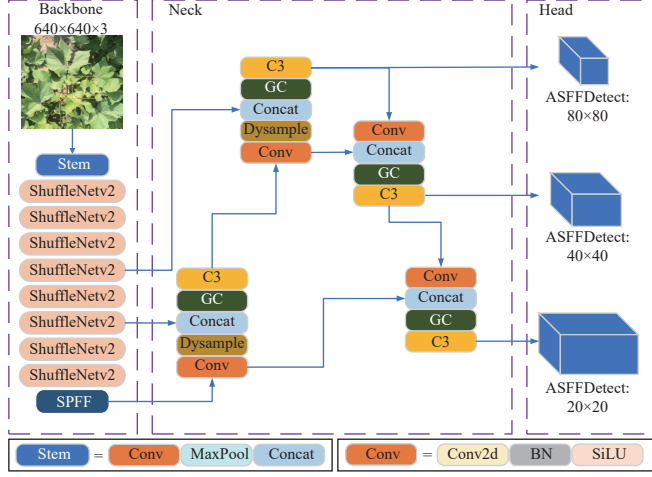
2.2 模型轻量化策略

为解决实际应用中边缘移动设备算力受限, 模型运行缓慢的问题, 基于 YOLOv5s 模型进行改进优化。采用 ShuffleNetv2 轻量级主干网络和 DySample 轻量化动态上采样模块替换原始模块, 减少计算量并提高模型检测速度。

2.2.1 ShuffleNetv2 轻量级主干网络

ShuffleNetv2^[26] 是 2018 年 7 月由旷视科技提出的轻量级网络, 在速度和精度方面进行了均衡。ShuffleNetv2 的基本模块如图 3a 所示, 采用通道切分 (channel split) 将输入特征图在通道维度分成 2 个通道数相同的分支, 左分支为恒等映射减少网络碎片, 右分支包含 3 个通道数一样的卷积使输入输出通道数相同; 2 个分支进行

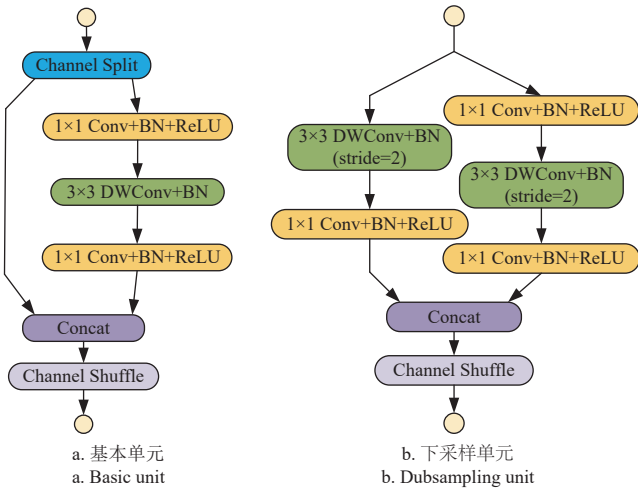
Concat 操作得到特征图的输出通道与原特征图的输入通道相同, 其中使用通道混洗 (channel shuffle) 操作来保证 2 个分支的信息进行交互。下采样模块如图 3b 所示, 移除通道切分, 且 2 个分支均采用原特征图, 并修改卷积单元步长 (stride) 为 2, Concat 后特征图大小减半通道数翻倍。



注: Backbone 为主干网络层; Neck 为颈部网络层; Head 为头部网络层; Stem 模块为 Conv+MaxPool+Concat; ShuffleNetv2 为轻量级主干网络; SPPF 为空间金字塔池化结构; Conv 模块为 Conv2d+BN+SiLU, Conv2d 为二维卷积层, BN 表示批量归一化, SiLU 为激活函数; C3 由 Conv 模块集成的卷积结构; Dysample 为动态上采样模块; Concat 为张量拼接操作; GC 为 Global Context 全局注意力模块; MaxPool 为最大池化操作; ASFFDetect 为自适应空间特征融合检测头。
Note: Backbone is the backbone network layer; Neck is the neck network layer; Head is the head network layer; The Stem module is Conv+MaxPool+Concat; ShuffleNetv2 is a lightweight backbone network; SPPF is a spatial pyramid pooling structure; the Conv module is Conv2d+BN+SiLU, Conv2d is a two-dimensional convolutional layer, BN represents batch normalization, and SiLU is an activation function; C3 is a convolution structure integrated by the Conv module; Dysample is a dynamic upsampling module; Concat is a tensor concatenation operation; GC is a Global Context global attention module; MaxPool is a maximum pooling operation; ASFFDetect is an adaptive spatial feature fusion detection head.

图 2 CottonBud-YOLOv5s 模型结构

Fig.2 The CottonBud-YOLOv5s model architecture



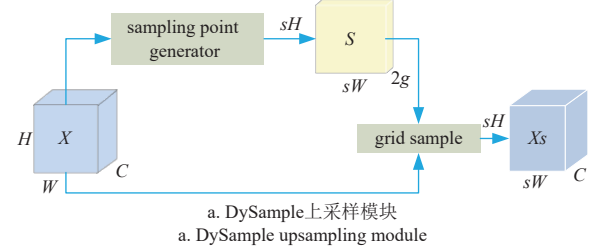
注: ReLU 为非线性激活函数; Channel Split 为通道切分操作; Channel Shuffle 为通道混洗操作; DWConv 为深度可分离卷积; Stride 为步长。
Note: ReLU is a nonlinear activation function; Channel Split is a channel splitting operation; Channel Shuffle is a channel shuffling operation; DWConv is a depthwise separable convolution; Stride is the step size.

图 3 ShuffleNetv2 网络结构

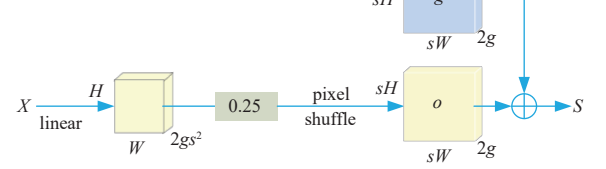
Fig.3 ShuffleNetv2 network architecture

2.2.2 DySample 轻量级动态上采样

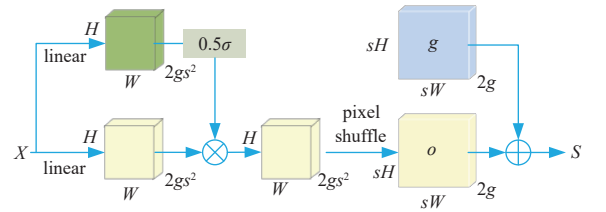
DySample^[27] 基于动态采样引入采样点生成器的概念, 避免动态卷积的方法, 并采用点采样的方式重新定义上采样过程。DySample 上采样模块如图 4a 所示, 给定大小为 $C \times H \times W$ 的特征图 X 和大小为 $2g \times H \times W$ 的采样集 S , 其中第一个维度的 2 代表 x 和 y 的坐标值; C 为特征图的通道数; H 、 W 为特征图的高度和宽度; g 为分组数。



Static Scope Factor



Dynamic Scope Factor

b. 采样点发生器
b. Sampling point generator

注: S 、 X 、 X_s 分别为采样点集合、输入特征图、上采样后的特征图; H 、 W 、 C 分别为输入特征图的高度、宽度、通道数; sH 、 sW 分别为采样点高度、宽度; g 为分组数; o 为输出特征图; σ 为动态范围因子。
Note: S , X , and X_s represent the sampling point set, input feature map, and upsampled feature map, respectively; H , W , and C represent the height, width, and number of channels of the input feature map, respectively; sH and sW represent the height and width of the sampling points, respectively; g is the number of groups; o is the output feature map; σ is the dynamic range factor.

图 4 DySample 上采样网络结构

Fig.4 DySample upsampling network architecture

采样函数 (grid sample) 使用 S 中的位置将给定的双线性插值 X 重新采样为大小 $C \times sH \times sW$ 的特征图 X_s 。此过程由式 (1) 定义:

$$X_s = \text{grid sample}(X, S) \quad (1)$$

采样集 S 由采样点发生器 (sampling point generator) 生成, 如图 4b 所示。DySample 上采样引入了“静态偏移范围”和“动态偏移范围”的概念, 通过乘以静态因子 (0.25) 或动态因子 (0.5σ) 来局部限制采样位置的移动范围, 从而减少重叠和预测误差。大小为 $C \times H \times W$ 的特征图 X 通过双线性插值上采样 (linear)、静态或动态因子、像素混洗 (pixel shuffle) 操作将其重塑为大小 $2g \times sH \times sW$ 的偏移 o , 那么采样集 S 是偏移 o 和原始采样 g 的总和, 即:

$$o = \text{linear}(X) \quad (2)$$

$$S = g + o \quad (3)$$

2.3 检测性能增强策略

密植环境下棉花顶芽特征较小易被棉叶遮挡, 且棉花打顶作业过程中的运动易导致棉花顶芽图像模糊, 给棉花顶芽检测带来困难, 为使网络模型能够更准确检测棉花顶芽位置和特征, 对检测头 Head 改进, 并在颈部引入 GC 全局注意力机制, 增强模型尺度不变性和上下文特征提取能力, 进而捕捉棉花顶芽的细节特征和增强模型对整体场景的感知能力。

2.3.1 ASFFHead 检测头

ASFF^[28] 自适应特征融合方法主要用于解决 Level 特征图尺度之间不一致问题, 将 ASFF 嵌入到 Head 中以增强对不同尺度对象的处理能力。ASFFHead 模块结构如图 5 所示, Level-1、Level-2 和 Level-3 指的是特征金字塔中不同层级的特征, 每个层级都有不同的空间分辨率。ASFF-1、ASFF-2 和 ASFF-3 表示应用了 ASFF 机制的不同层级的特征融合。

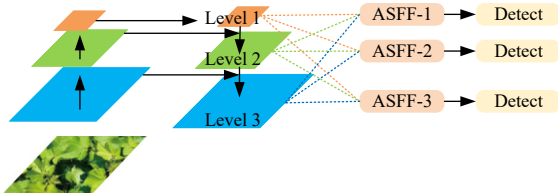


图 5 ASFFHead 检测头网络结构

Fig.5 ASFFHead detection head network architecture

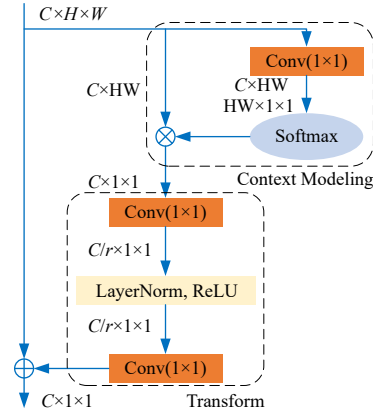
ASFF-1 是对 Level-3 的特征图做 3×3 MaxPool 和 3×3 Conv, 对 Level-2 的特征图做 3×3 Conv; ASFF-2 是对 Level-3 的特征图做 3×3 Conv, 对 Level-1 的特征图做 1×1 Conv, 并重置为原图分辨率的 2 倍; ASFF-3 是对 Level-2 的特征图做 1×1 Conv, 并重置为原图分辨率 2 倍, 对 Level-1 的特征图做 1×1 Conv, 并重置为原图分辨率 4 倍。通过 ASFF 机制的不同层级的特征融合, 所有特征便具有相同的空间维度。

2.3.2 GC 注意力机制

视觉领域常用通道注意力机制和空间注意力机制, 通道注意力机制使模型能学习不同通道的重要程度, 并对其进行加权; 其中代表性方法 SENet^[29] 网络, 首先对输入特征图进行全局平均池化, 获取每个通道的全局信息, 再通过两个全连接层和 ReLU、Sigmoid 激活函数的输出作为通道重要性的权重; 将权重与原始输入特征图的各个通道进行加权处理, 有效提升重要通道特征的表达能力。空间注意力机制是对空间位置的加权, 让模型更加关注目标检测任务相关区域; NLNet (non-local neural networks) 网络^[30] 通过对长距离依赖关系建模, 并对远距离信息进行加权融合, 增强全局空间信息的关联性。

GC 全局注意力机制^[31] 融合了简化后的 Non-local 模块和 SENet 模块的优点, 既有 NLNet 全局上下文建模能力, 又像 SENet 模块一样轻量化。GC 全局注意力模块

网络结构图如图 6 所示, 首先采用 1×1 卷积 Conv 和 Softmax 函数来获取注意力权重, 然后与原始输入特征相乘获得全局上下文特征, 上下文信息建模 (context modeling) 使用 NLNet 中的机制; 特征转换 (transform) 借鉴 SENet 模块, GC 模块中将 $C \times 1 \times 1$ 卷积用瓶颈转换模块代替, 该模块将参数量从 $C \times C$ 减少到 $2 \times C \times C/r$, 其中 r 是瓶颈比率。Layer Normalization 层的作用是在特征维度上进行归一化, 提高模型泛化性。



注: \oplus 表示元素加法; \otimes 表示矩阵乘法; HW 为张量总空间位置数; r 为瓶颈比率。

Note: \oplus represents element-wise addition; \otimes represents matrix multiplication; HW represents the total number of spatial positions in a tensor; r represents the bottleneck ratio.

图 6 GC 注意力模块网络结构

Fig.6 GC attention module network architecture

2.4 试验平台及评价指标

试验训练环境: 使用中科视拓科技有限公司的线上服务器 (AutoDL), 操作系统为 Ubuntu 20.04, 内存为 120 G, 显卡为 RTX4090, 显存为 24 G, CPU 为 Intel Xeon Gold 6430 16 核, CUDA 11.3 并行计算框架, Python 3.8.10, PyTorch 1.11.0。

试验推理环境: 操作系统为 Windows 11, 内存为 32 G, CPU 为 Intel i9-12900H 14 核, Python 3.8.19, PyTorch 1.13.0。

模型训练参数设置: 单机单卡, 输入尺寸为 $3 \times 640 \times 640$, 每批次样本数量为 16, 多线程为 16, 优化器为 SGD (stochastic gradient descent), 训练次数 (epoch) 为 200 轮, 初始学习率为 0.001, 权重衰减系数为 0.000 5, 使用 Mosaic、Mixup 数据增强方式。

为了准确评估模型的性能, 采用以下指标^[32]: 精度 P (precision, %), 平均精度均值 mAP (mean average precision, %), 平均精度 AP (average precision, %), 召回率 R (recall, %), 参数量 (params, M), 计算量 (flops/G), 检测速度 FPS (frames per second, 帧/s)。其中, 精度 P 表示预测正确的样本数占预测为正样本数的比例, 如式 (4) 所示:

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (4)$$

召回率 R 表示预测正确的样本数占实际为正样本数的比例, 如式 (5) 所示:

$$R = \frac{T_p}{T_p + F_N} \times 100\% \quad (5)$$

平均精度 AP 即为 P-R 曲线的面积, 平均精度均值 mAP 为平均精确度 AP 的均值, 如式 (6) 所示:

$$\text{mAP} = \frac{\sum_{i=1}^N \int_0^1 P(R) dR}{N} \times 100\% \quad (6)$$

式中 T_p 表示正确检测棉花顶芽的样本数; F_p 表示被误测为棉花顶芽的样本数; F_N 表示未被正确检测的棉花顶芽样本数; N 表示类别数量, 本文为棉花顶芽, 故 $N=1$ 。

3 试验结果与分析

3.1 模型改进试验结果

3.1.1 ShuffleNetv2 主干网络改进试验

本试验以 YOLOv5s 目标检测模型为基础, 遵循单一变量原则, 仅替换不同轻量级主干网络 MobileOne、MobileNetV3、GhostNet、ShuffleNetv2 进行比较。由表 1 可知, 使用 ShuffleNetv2 轻量级主干网络作为 YOLOv5s 主干网络时, 计算量为 2.20 G, 精度 P 为 95.4%, 平均精度均值 mAP 为 96.2%, 召回率 R 为 94.8%, 检测速度为 40.7 帧/s。由于 ShuffleNetv2 引入的通道切分、通道混洗操作避免分组卷积, 增强通道间信息交互, 减少计算量并提高了特征提取能力, 与改进前 CSPDarkNet53 主干网络相比, 计算量约为原来的 1/8, 检测速度约为原来的 2.5 倍, 检测精度 P 、平均精度均值 mAP、召回率 R 分别减小 0.7、0.9、1.6 个百分点, 略有降低。

表 1 不同主干网络棉花顶芽试验

Table 1 Test of different backbone networks on cotton buds

主干网络 Backbone network	计算量 Flops/G	检测速度 FPS/(帧·s ⁻¹)	$P/\%$	mAP/%	$R/\%$
CSPDarkNet53	15.80	16.4	96.1	97.1	96.4
MobileOne	8.00	23.5	88.7	91.9	91.4
MobileNetV3	1.30	34.0	81.8	84.0	81.3
GhostNet	13.40	16.7	95.8	96.7	96.1
ShuffleNetv2	2.20	40.7	95.4	96.2	94.8

注: P 为检测精度; mAP 为平均精度均值; R 为召回率。下同。

Note: P is the detection precision; mAP is the mean average precision; R is the recall rate. The same below.

GhostNet 主干网络的检测速度不满足实时性要求, 与 MobileOne、MobileNetV3 主干网络相比, ShuffleNetv2 主干网络检测精度 P 分别提升 6.7、13.6 个百分点, 平均精度均值 mAP 分别提升 4.3、12.2 个百分点, 召回率 R 分别提升 3.4、13.5 个百分点。因此, 本文选择提取特征充分且检测速度都相对良好的 ShuffleNetv2 主干网络对 YOLOv5s 模型进行设计与改进。

3.1.2 Dysample 上采样模块改进试验

为了弥补替换轻量级主干网络 ShuffleNetv2 带来的精度损失, 同时保持轻量化。使用 Dysample 上采样模块替换 Nearest 原始上采样模块, 并对比不同上采样模块 Nearest、Bilinear、Carafe、Dysample 对模型性能指标的影响, 试验结果如表 2 所示。

由表 2 可知, 4 种上采样模块的检测速度均满足实时性要求, 其中, Dysample 动态上采样模块与 Nearest、

Bilinear、Carafe 上采样模块相比, 检测精度 P 分别提升 0.1、0.4、0.8 个百分点, 平均精度均值 mAP 分别提升 0.6、0.8、0.9 个百分点, 与 Nearest、Carafe 上采样模块相比召回率 R 分别提升 1.1、0.9 个百分点, 与 Bilinear 上采样模块召回率持平。替换 ShuffleNetv2 轻量级主干网络和 Dysample 动态上采样模块对 YOLOv5s 进行轻量化改进时, 改进后模型在参数量、计算量、检测速度上明显优于改进前的 YOLOv5s 模型, 在检测精度 P 、平均精度均值 mAP、召回率 R 上也明显优于其他轻量化组合。故选用 ShuffleNetv2 轻量级主干网络和 Dysample 动态上采样模块进行轻量化改进是较优组合。

表 2 不同上采样模块棉花顶芽对比试验

Table 2 Test of different upsampling modules on cotton buds

上采样模块 Upsampling module	检测速度 FPS/(帧·s ⁻¹)	$P/\%$	mAP/%	$R/\%$
Nearest	40.7	95.4	96.2	94.8
Bilinear	41.6	95.1	96.0	95.9
Carafe	38.3	94.7	95.9	95.0
Dysample	40.0	95.5	96.8	95.9

3.1.3 ASFFHead 检测头改进试验

替换 ShuffleNetv2 轻量级主干网络和 Dysample 动态上采样模块对 YOLOv5s 进行轻量化改进后, 为提高模型对运动模糊图像、小目标遮挡的检测性能, 提出 ASFFHead 改进型检测头。为了测试改进型检测头 ASFFHead 的性能, 与 DecoupledHead、CLLAHead、原始 Head 检测头进行对比试验, 试验结果如表 3 所示。

表 3 不同检测头棉花顶芽对比试验

Table 3 Test of different detection heads on cotton buds

检测头 Detection head	计算量 Flops/G	$P/\%$	mAP/%	$R/\%$
Head	2.30	95.4	96.2	94.8
CLLAHead	2.90	94.9	96.5	94.6
DecoupledHead	12.30	96.3	97.4	97.0
ASFFHead	4.40	95.5	96.8	95.9

由表 3 可知, DecoupledHead 在精度 P 、平均精度均值 mAP、召回率 R 上略高于其他检测头, 但 DecoupledHead 与 Head、CLLAHead、ASFFHead 相比, 计算量分别增加 10、9.4、7.9 G, 计算量大量增加导致推理时间过长。ASFFHead 与 Head、CLLAHead 相比, 在精度 P 、平均精度均值 mAP、召回率 R 增加的同时计算量增加较小。ASFFHead 检测头通过自适应地调整不同尺度特征的权重, ASFF 通过特征金字塔网络 (FPN) 生成多个尺度的特征图, 不同尺度的特征图与对应的权重图逐像素相乘, 再进行加权求和, 生成融合后的特征图, 该特征图在空间和语义上具有高度一致性, 从而提高模型对小目标的感知能力, 更准确地检测小目标。

3.1.4 GC 注意力机制改进试验

在使用 ShuffleNetv2 轻量级主干网络、Dysample 动态上采样模块、ASFFHead 检测头改进 YOLOv5s 的基础上, 为测试 GC 注意力机制对模型的影响, 颈部引入 CBAM、GE、MLCA、GC 不同注意力模块进行对比试验, 试验结果如表 4 所示。

由表 4 可知, GC 注意力机制与 CBAM、GE、MLCA 注意力机制相比, 精度 P 分别提升 2.7、1.6、1.8 个百分

点, 平均精度均值 mAP 分别提升 2.2、1.0、2.6 个百分点, 召回率 R 分别提升 2.5、1.8、4.8 个百分点。GC 全局注意力机制融合了 Non-local 模块的优势, 继承了 NLNet 的全局上下文建模能力, 保留了更多棉花顶芽的细节特征, Layer Normalization 层的引入在特征维度上进行归一化处理, 进一步提升模型的泛化能力。

表 4 不同注意力机制对比

Table 4 Comparison of different attention mechanisms

注意力机制 Attention mechanisms	$P/\%$	mAP/%	$R/\%$
CBAM	94.2	95.7	94.7
GE	95.2	96.9	95.4
MLCA	95.1	95.3	92.4
GC	96.9	97.9	97.2

为了更直观地展示 GC 注意力机制对特征细节的增强作用, 采用 Grad-CAM 技术进行可视化解释和分析, GC 注意力机制引入前后, 模型对棉花顶芽识别的热力图如图 7 所示。

由图 7 可知, 引入 GC 注意力模块前的模型在识别棉花顶芽时, 关注部位杂乱且分散, 且聚焦部位包含较多背景信息。而引入 GC 注意力模块后的模型, 在单株、多株、运动模糊和遮挡多种情况下, 所提模型都将注意力聚焦在了棉花顶芽部位。棉花顶芽识别的热力图可视化对比表明, 引入 GC 注意力模块后的模型具备更加高效准确的特征提取能力, 且所提模型在识别棉花顶

芽时所依据的图像特征信息可信。

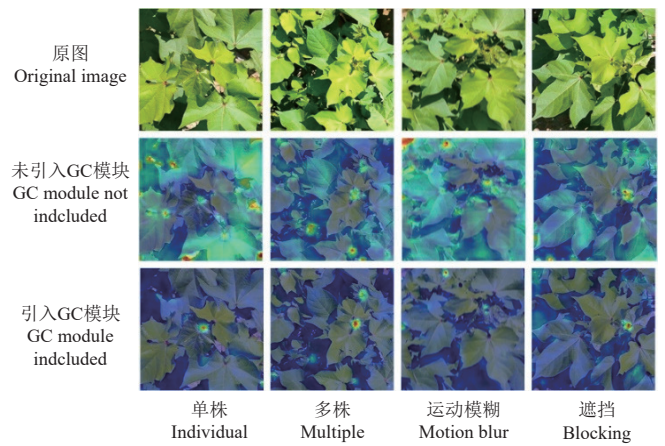


图 7 GC 注意力模块对棉花顶芽部位关注能力的可视化

Fig.7 Visualization of the GC attention module's focus on cotton bud locations

3.1.5 CottonBud-YOLOv5s 模型消融试验

为了验证模型轻量化后引入 ASFFHead 检测头和 GC 全局注意力模块对大、中、小目标检测性能的影响。首先, 定义目标框与原图面积的比值小于 0.002 5 为小目标, 大于 0.022 5 为大目标, 其余为中目标^[33]。随后, 通过消融试验评估各模块对各类目标检测性能的影响, 结果如表 5 所示。

表 5 CottonBud-YOLOv5s 消融试验

Table 5 CottonBud-YOLOv5s ablation experiments

ASFFHead 模块 ASFFHead module	GC 注意力模块 GC attention module	参数量 Params /M	计算量 Flops/G	Small		Medium		Large	
				AP _{0.5:0.95} /%	AR _{0.5:0.95} /%	AP _{0.5:0.95} /%	AR _{0.5:0.95} /%	AP _{0.5:0.95} /%	AR _{0.5:0.95} /%
—	—	4.05	2.30	31.2	41.4	48.2	56.3	61.1	67.0
√	—	9.49	4.40	30.9	37.6	51.4	59.3	64.7	70.5
—	√	4.21	2.30	24.1	31.1	43.1	53.0	57.1	63.6
√	√	9.65	4.40	34.8	43.5	52.3	59.8	67.6	72.9

注: “—”表示不使用此模块, “√”表示使用此模块; Small、Medium、Large 分别代表小目标、中目标、大目标; AP_{0.5:0.95} 代表 IOU 阈值 0.5–0.95 的平均精度值, AR_{0.5:0.95} 代表 IOU 阈值 0.5–0.95 的平均召回率。

Note: “—” indicates not using this module, “√” indicates using this module; Small, Medium, Large represent small, medium, and large objects respectively; AP_{0.5:0.95} represents the average precision value over the IOU threshold range from 0.5 to 0.95, and AR_{0.5:0.95} represents the average recall value over the IOU threshold range from 0.5 to 0.95.

由表 5 可知, 与不使用 ASFFHead 检测头和 GC 注意力机制模块时相比, 仅使用 GC 全局注意力模块时, 大中小目标 AP_{0.5:0.95} 和 AR_{0.5:0.95} 值均有所下降; 仅使用 ASFFHead 检测头时, 中目标和大目标 AP_{0.5:0.95} 和 AR_{0.5:0.95} 值均有所提高, 小目标 AP_{0.5:0.95} 和 AR_{0.5:0.95} 值均略有下降; 同时使用 ASFFHead 检测头和 GC 全局注意力模块时, 小目标 AP_{0.5:0.95} 和 AR_{0.5:0.95} 值分别提升 3.6、2.1 个百分点, 中目标 AP_{0.5:0.95} 和 AR_{0.5:0.95} 值分别提升 4.1、3.5 个百分点, 大目标 AP_{0.5:0.95} 和 AR_{0.5:0.95} 值分别提升 6.5、5.9 个百分点, 参数量、计算量略有增加。

通过消融试验可知, 在棉花顶芽数据集中, 轻量化改进后的模型引入 ASFFHead 检测头和 GC 注意力模块对大中小目标检测性能均有提升, 可满足模型轻量化的同时提高棉花顶芽的检测精度。

3.2 不同目标检测算法对比试验

基于 YOLOv5s 改进的 CottonBud-YOLOv5s 模型与目前主流的目标检测模型 Faster-RCNN、TOOD、RTDETR、YOLOv3s、YOLOv5s、YOLOv9s、YOLOv10s 进行对比试验, 结果如表 6 所示。

表 6 不同目标检测模型对比

Table 6 Comparison of different object detection models

模型 Model	主干网络 Backbone network	计算量 Flops/G	检测速度 FPS /(帧·s ⁻¹)	mAP/%	$R/\%$
Faster-RCNN	Resnet50	178.00	1.5	83.9	80.4
TOOD	Resnet101	168.00	1.2	84.6	81.2
RTDETR	Resnet50	130.50	3.7	92.4	94.0
YOLOv3s	Darknet53	283.00	3.1	97.0	95.2
YOLOv5s	CSPDarknet53	15.80	16.4	97.1	96.4
YOLOv9s	GELAN	27.40	9.3	97.7	96.7
YOLOv10s	CSPNet	24.80	12.3	96.4	96.0
CottonBud-YOLOv5s	ShuffleNetv2	4.40	27.9	97.9	97.2

CottonBud-YOLOv5s 模型与 Faster-RCNN、TOOD、RTDETR、YOLOv3s、YOLOv5s、YOLOv9s、YOLOv10s 相比, 计算量分别减小 173.6、163.6、126.1、278.6、11.4、23.0、20.4G, 检测速度分别提升 26.4、26.7、24.2、24.8、11.5、18.6、15.6 帧, 平均精度均值 mAP 分别提升 14.0、13.3、5.5、0.9、0.8、0.2、1.5 个百分点, 召回率 R 分别提升 16.8、16.0、3.2、2.0、0.8、0.5、1.2 个百分点。在棉花顶芽数据集上, CottonBud-YOLOv5s 模型在计算量、检测速度、平均精度均值 mAP、召回率 R 上

都优于其他 7 种模型, 满足在边缘移动设备上棉花顶芽检测对精度和实时性要求。

3.3 模型检测效果可视化

为了更加清楚地观察改进效果, 针对单目标、多目

标、运动模糊、遮挡这 4 种情况, 各随机选取一张照片进行目标检测。对比 Faster-RCNN、TOOD、RTDETR、YOLOv3s、YOLOv5s、YOLOv9s、YOLOv10s、CottonBud-YOLOv5s 不同模型检测效果, 如图 8 所示。

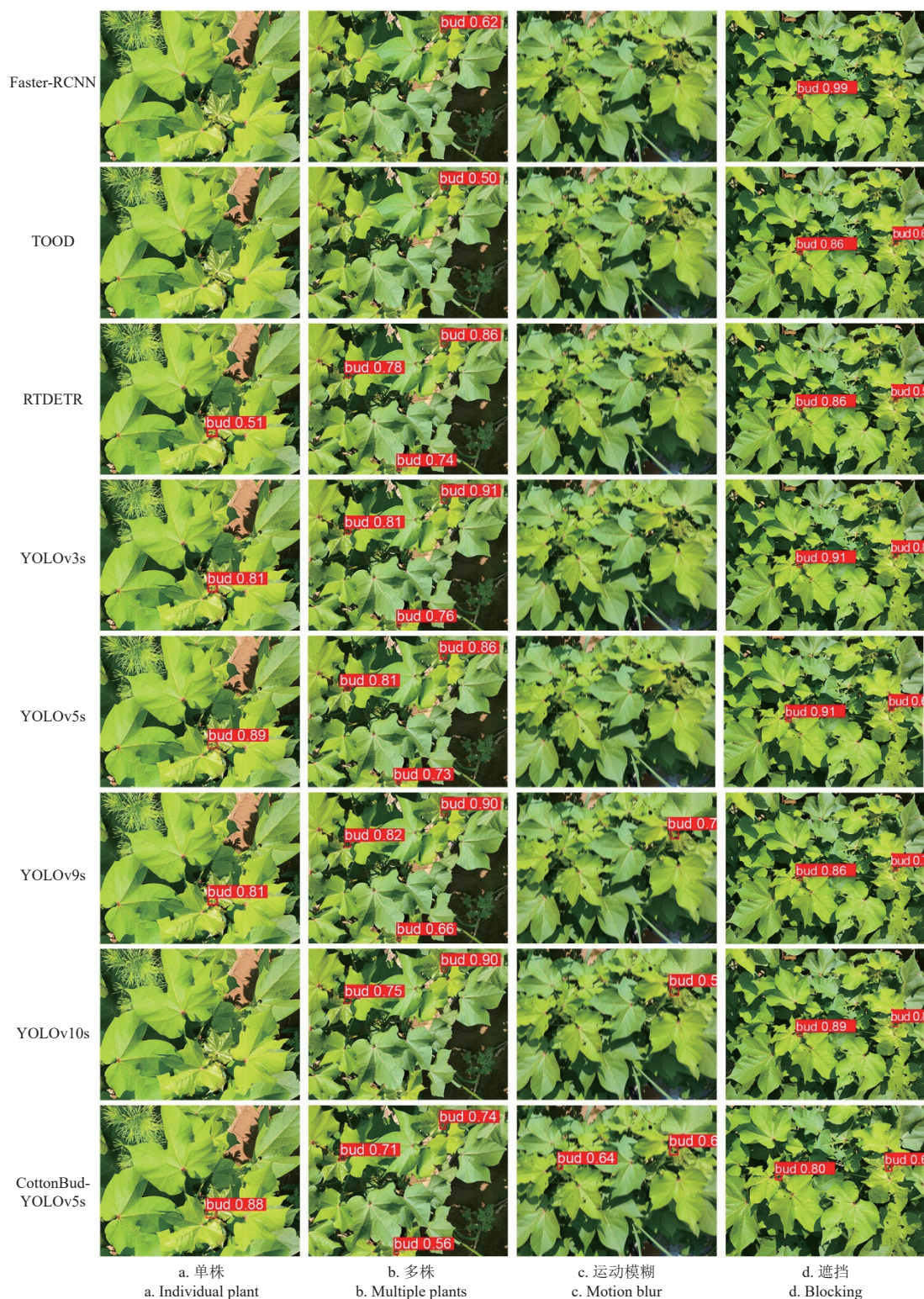


图 8 不同模型检测效果图

Fig.8 Detection effect diagram of different models

单株、多株棉花顶芽检测时, Faster-RCNN、TOOD、YOLOv10s 模型存在漏检问题; 运动模糊图像检测时, Faster-RCNN、TOOD、RTDETR、YOLOv3s、YOLOv5s、

YOLOv9s、YOLOv10s 模型均存在漏检, 仅 CottonBud-YOLOv5s 正确检测到棉花顶芽; 遮挡的棉花顶芽检测时, Faster-RCNN 模型存在漏检现象, TOOD、RTDETR、

YOLOv3s、YOLOv5s、YOLOv9s、YOLOv10s、CottonBud-YOLOv5s 模型均可以正确检测棉花顶芽。CottonBud-YOLOv5s 模型在单目标、多目标、运动模糊、遮挡场景下棉花顶芽检测的综合性能最好。

基于以上试验分析, CottonBud-YOLOv5s 模型轻量化的同时, 提升了运动模糊图像、小目标遮挡的检测性能, 可满足棉花顶芽检测实时性与精准性的要求。

4 结 论

针对边缘移动设备算力受限难以满足实时性要求, 棉花打顶作业过程中的运动模糊图像、小目标遮挡难以检测的问题, 基于 YOLOv5s 算法模型, 提出了一种轻量化 CottonBud-YOLOv5s 棉花顶芽检测模型, 主要结论如下:

1) 通过将 YOLOv5s 主干网络和上采样模块分别替换为 ShuffleNetv2 轻量级主干和 Dysample 轻量级上采样模块, 引入全新自适应空间特征融合检测头 ASFFHead 和 GC 全局注意力模块, 有效减少了模型的计算量, 提升了检测速度, 增强了模型对运动模糊和小目标遮挡的检测性能。

2) CottonBud-YOLOv5s 棉花顶芽检测模型平均精度均值 mAP 达到 97.9%, 召回率 R 达到 97.2%, CPU 检测速度达到 27.9 帧/s, 与主流目标检测模型 Faster-RCNN、TOOD、RTDETR、YOLOv3s、YOLOv5s、YOLOv9s 和 YOLOv10s 进行试验比较, 检测速度分别提升 26.4、26.7、24.2、24.8、11.5、18.6、15.6 帧/s, 平均精度均值 mAP 分别提升 14.0、13.3、5.5、0.9、0.8、0.2、1.5 个百分点, 召回率 R 分别提升 16.8、16.0、3.2、2.0、0.8、0.5、1.2 个百分点。

3) 可视化结果表明, 在单株、多株、运动模糊、小目标遮挡情况下 CottonBud-YOLOv5s 的检测性能优于其他检测模型, 具有较高的检测精度、鲁棒性和检测速度, 适用于密植环境下棉花顶芽的精准检测, 为棉花机械化打顶提供视觉检测基础。

未来将继续优化模型的泛化能力、扩展其在多样农业场景中的应用、提升在低端硬件上的实时性能、增强模型的鲁棒性, 并通过融合多模态数据进一步提升检测精度, 发展成为智慧农业领域的关键技术, 推动农业现代化的进程。

[参 考 文 献]

- [1] 李继福, 何俊峰, 陈佛文, 等. 中国棉花生产格局与施肥研究现状——基于 CNKI 数据计量分析[J]. *中国棉花*, 2019, 46(4): 17-24.
LI Jifu, HE Junfeng, CHEN Fowen, et al. Research on the production pattern and fertilization of cotton in china: Based on CNKI data metric analysis[J]. *Chinese Cotton*, 2019, 46(4): 17-24. (in Chinese with English abstract)
- [2] 韩长杰, 燕长辉, 仇世龙, 等. 基于机器视觉的双圆盘式棉花打顶装置设计与试验[J]. *农业机械学报*, 2023, 54(5): 36-46.
- [3] 李涛, 杨德松. 棉花不同打顶方式应用分析[J]. *新疆农垦科技*, 2021, 44(2): 14-15.
- [4] 韩鑫, 韩金鸽, 陈允琳, 等. 基于顶芽智能识别的棉花化学打顶系统研究[J]. *农业机械学报*, 2024, 55(3): 145-152.
HAN Xin, HAN Jing, CHEN Yunlin, et al. Research on the chemical topping system of cotton based on apical bud intelligent recognition[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(3): 145-152. (in Chinese with English abstract)
- [5] 彭强吉, 康建明, 宋裕民, 等. 3MDZ-4 型自走式棉花打顶喷药联合作业机设计[J]. *农业工程学报*, 2019, 35(14): 30-38.
PENG Qiangji, KANG Jianming, SONG Yumin, et al. Design of 3MDZ-4 self-propelled cotton topping and spraying joint operation machine[J]. *Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE)*, 2019, 35(14): 30-38. (in Chinese with English abstract)
- [6] 卢秀茹, 贾肖月, 牛佳慧. 中国棉花产业发展现状及展望[J]. *中国农业科学*, 2018, 51(1): 26-36.
LU Xiuru, JIA Xiaoyue, NIU Jiahui. Current situation and prospect of cotton industry development in China[J]. *Chinese Agricultural Science*, 2018, 51(1): 26-36. (in Chinese with English abstract)
- [7] 史增录, 张学军, 丁永前, 等. 基于自动检测与控制的棉花打顶装置的设计[J]. *中国农机化学报*, 2017, 38(12): 58-62.
SHI Zenglu, ZHANG Xuejun, DING Yongqian, et al. Design of cotton topping device based on automatic detection and control[J]. *Journal of Chinese Agricultural Mechanization*, 2017, 38(12): 58-62. (in Chinese with English abstract)
- [8] 王玲, 邹小昱, 刘思瑶, 等. 棉花采摘机器人红外测距技术研究[J]. *农业机械学报*, 2014, 45(7): 61-66.
WANG Ling, ZOU Xiaoyu, LIU Siyao, et al. Infrared distance measurement used for cotton picker robot[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2014, 45 (7): 61-66. (in Chinese with English abstract)
- [9] 闫毅敏, 王维新, 李霞, 等. 基于自动控制的棉株顶尖高度测量系统设计与试验[J]. *甘肃农业大学学报*, 2018, 53(5): 176-184.
YAN Yimin, WANG Weixin, LI Xia. Design and test of cotton top height measurement system based on automatic control[J]. *Journal of Gansu Agricultural University*, 2018, 53(5): 176-184. (in Chinese with English abstract)
- [10] GANESAN G, CHINNAPPAN J. Hybridization of ResNet with YOLO classifier for automated paddy leaf disease recognition: An optimized model[J]. *Journal of Field Robotics*, 2022, 39(7): 1085-1109.
- [11] LI D, AHMED F, WU N, et al. YOLO-JD: A deep learning network for jute diseases and pests detection from images[J]. *Plants*, 2022, 11(7): 937.
- [12] BAKHSHAYESHI I, ERFANI E, TAGHIKHAH F R, et al. An

- intelligence cattle re-identification system over transport by siamese neural networks and YOLO[J]. IEEE Internet of Things Journal, 2023, 11(2): 2351-2363.
- [13] PEREZ-PORRAS F J, TORRES-SANCHEZ J, LOPEZ-GRANADOS F, et al. Early and on-ground image-based detection of poppy (*Papaver rhoeas*) in wheat using YOLO architectures[J]. Weed Science, 2023, 71(1): 50-58.
- [14] 范迎迎, 钱育蓉, 杨柳, 等. 基于 BP 神经网络的遥感影像棉花识别方法[J]. 计算机工程与设计, 2017, 38(5): 1356-1360.
- FANG Yingying, QIAN Yurong, YANG Liu, et al. Cotton recognition method for remote sensing image based on BP neural network[J]. Computer Engineering and Design, 2017, 38(5): 1356-1360. (in Chinese with English abstract)
- [15] KOIRALA A, WALSH K B, WANG Z, et al. Deep learning for mango (*Mangifera indica*) panicle stage classification[J]. Agronomy, 2020, 10(1): 2073-4395.
- [16] 张日红, 区建爽, 李小敏, 等. 基于改进 YOLOv4 的轻量化菠萝苗心检测算法[J]. 农业工程学报, 2023, 39(4): 135-143.
- ZHANG Rihong, QU Jianshuang, LI Xiaomin, et al. Lightweight pineapple seedling heart detection algorithm based on improved YOLOv4[J]. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE), 2023, 39(4): 135-143. (in Chinese with English abstract)
- [17] 彭明霞, 夏俊芳, 彭辉. 融合 FPN 的 Faster R-CNN 复杂背景下棉田杂草高效识别方法[J]. 农业工程学报, 2019, 35(20): 202-209.
- PENG Mingxia, XIA Junfang, PENG Hui. Efficient weed identification method in cotton field with complex background based on Faster R-CNN fused with FPN[J]. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE), 2019, 35(20): 202-209. (in Chinese with English abstract)
- [18] WU H, WANG Y, ZHAO P, et al. Small-target weed-detection model based on YOLO-V4 with improved backbone and neck structures[J]. Precision Agriculture, 2023, 24(6): 2149-2170.
- [19] 刘海涛, 韩鑫, 兰玉彬, 等. 基于 YOLOv4 网络的棉花顶芽精准识别方法[J]. 中国农业科技导报, 2022, 24(8): 99-108.
- LIU Haitao, HAN Xin, LAN Yubin, et al. Precision identification of cotton terminal bud based on YOLOv4 network[J]. Journal of Agricultural Science and Technology, 2022, 24(8): 99-108. (in Chinese with English abstract)
- [20] XU W, XU T, THOMASSON J A, et al. A lightweight SSV2-YOLO based model for detection of sugarcane aphids in unstructured natural environments[J]. Computers and Electronics in Agriculture, 2023, 211: 107961.
- [21] 张新月, 胡广锐, 李浦航, 等. 基于改进 YOLOv8n 的轻量化红花识别方法[J]. 农业工程学报, 2024, 40(13): 163-170.
- ZHANG Xinyue, HU Guangrui, LI Puhang, et al. Recognizing safflower using improved lightweight YOLOv8n[J]. Journal of Agricultural Engineering, 2024, 40(13): 163-170. (in Chinese with English abstract)
- [22] 陈柯屹, 朱龙付, 宋鹏, 等. 融合动态机制的改进型 Faster R-CNN 识别田间棉花顶芽[J]. 农业工程学报, 2021, 37(16): 161-168.
- CHEN Keyi, ZHU Longfu, SONG Peng, et al. Improved faster R-CNN with fusion dynamic mechanism for identifying cotton apical buds in the field[J]. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE), 2021, 37(16): 161-168. (in Chinese with English abstract)
- [23] 彭炫, 周建平, 许燕, 等. 改进 YOLOv5 识别复杂环境下棉花顶芽[J]. 农业工程学报, 2023, 39(16): 191-197.
- PENG Xuan, ZHOU Jianping, XU Yan, et al. Identifying cotton apical buds in complex environments using improved YOLOv5[J]. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE), 2023, 39(16): 191-197. (in Chinese with English abstract)
- [24] 白强, 高荣华, 赵春江, 等. 基于改进 YOLOv5s 网络的奶牛多尺度行为识别方法[J]. 农业工程学报, 2022, 38(12): 163-172.
- BAI Qiang, GAO Ronghua, ZHAO Chunjiang, et al. Multi-scale behavior recognition of dairy cows based on improved YOLOv5s network[J]. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE), 2022, 38(12): 163-172. (in Chinese with English abstract)
- [25] 尚钰莹, 张倩如, 宋怀波. 基于 YOLOv5s 的深度学习在自然场景苹果花朵检测中的应用[J]. 农业工程学报, 2022, 38(9): 222-229.
- SHANG Yuying, ZHANG Qianru, SONG Huaibo. Application of deep learning based on YOLOv5s in detecting apple flowers in natural scenes[J]. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE), 2022, 38(9): 222-229. (in Chinese with English abstract)
- [26] MA N, ZHANG X, ZHENG H T, et al. Shufflenetv2: Practical guidelines for efficient cnn architecture design[C]//Proceedings of the European Conference on Computer Vision (ECCV). Cham, Springer, 2018: 122-138.
- [27] LIU W, LU H, FU H, et al. Learning to upsample by learning to sample[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE, 2023: 6004-6014.
- [28] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection [EB/OL]. (2019-11-25) [2025-01-06]. <https://doi.org/10.48550/arXiv.1911.09516>.
- [29] XU G, GAO Y, LI J, et al. InSAR phase denoising: A review of current technologies and future directions[J]. IEEE Geoscience and Remote Sensing Magazine, 2020, 8(2): 64-82.
- [30] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 7794-7803.
- [31] CAO Y, XU J, LIN S, et al. GCNet: Non-Local Networks meet squeeze-excitation networks and beyond[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul, USA: IEEE, 2020: 1971-1980.
- [32] 兰玉彬, 孙斌书, 张乐春, 等. 基于改进 YOLOv5s 的自然场景下生姜叶片病虫害识别[J]. 农业工程学报, 2024, 40(1): 210-216.
- LAN Yubin, SUN Binshu, ZHANG Lechun, et al.

- Identification of ginger leaf diseases and pests in natural scenes based on improved YOLOv5s[J]. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE), 2024, 40(1): 210-216. (in Chinese with English abstract)
- [33] 顾洪宇, 李志合, 李涛, 等. 基于 YOLOv5 的马铃薯种薯芽眼轻量化检测算法[J]. 农业工程学报, 2024, 40(14):

126-136.

GU Hongyu, LI Zhihe LI Tao, et al. Lightweight detection algorithm for potato seed tubers' bud eyes based on YOLOv5[J]. Transactions of the Chinese Society of Agricultural Engineering(Transactions of the CSAE), 2024, 40(14): 126-136. (in Chinese with English abstract)

CottonBud-YOLOv5s lightweight cotton bud detection algorithm

ZHAO Luqiang¹, PENG Qiangji^{2,3}, LAN Yubin¹, KANG Jianming^{2,3}, ZHANG Jingwen¹,
DAI Jianlong⁴, CHEN Yulong^{1*}

(1. College of Agricultural Engineering and Food Science, Shandong University of Technology, Zibo 255000, China; 2. Shandong Academy of Agricultural Machinery Sciences, Jinan 250100, China; 3. Shandong Key Laboratory of Intelligent Agricultural Equipment in Hilly and Mountainous Areas, Jinan 250100, China; 4. Shandong Academy of Agricultural Sciences, Jinan 250100, China)

Abstract: Cotton mechanical topping is one of the most important cultural practices to improve crop yield during production. The shoots of cotton topping can be cut at about 10–20 cm from the top of plants. However, the performance of mechanical topping has been limited to computing power and real-time transport in several edge-moving devices at present. The detection can also be confined to the motion blur and small target occlusion. In this study, a lightweight detection model of a cotton bud (named CottonBud-YOLOv5s) was proposed using the well-known YOLOv5s architecture. Both performance and efficiency were optimized to detect the cotton buds in complex field environments. The ShuffleNetv2 backbone network was utilized to enhance the overall performance of the CottonBud-YOLOv5s model. The computational complexity was reduced to maintain the high accuracy of detection. In addition, the DySample dynamic upsampling module was integrated to replace the original ones. The computational costs were further reduced to improve the speed of detection. As such, the improved model was run more efficiently on edge devices with limited computing power. Real-time performance was also achieved during cotton mechanical topping. Moreover, the ASFFHead detection head and GC (global context) attention mechanism were also introduced into the head and neck components, in order to handle the varying object scales and complex contextual information. The scale invariance was significantly improved to extract the context-based features, which was crucial to detect the small targets that occluded or blurred due to the various motions in fields. Ultimately, the robustness of the model was improved to perform the best in real-world conditions. A series of ablation and comparison tests were conducted to validate the efficacy of the CottonBud-YOLOv5s model. The experimental results demonstrated that the introduction of the ASFFHead detection head and the GC global attention mechanism led to notable improvements in detection accuracy. Specifically, the average precision (AP) at 0.5:0.95 for small targets increased by 3.6 percentage points, while the average recall rate (AR) at the same threshold was improved by 2.1 percentage points. In the medium-sized targets, the AP and AR increased by 4.1 and 3.5 percentage points, respectively. In the large targets, the AP and AR increased by 6.5 and 5.9 percentage points, respectively. The improved model performed the best to detect the targets across a range of sizes. Furthermore, the CottonBud-YOLOv5s model shared significant improvements in the detection speed, compared with the state-of-the-art detection models, including Faster-RCNN, TOOD, RTDETR, YOLOv3s, YOLOv5s, YOLOv9s, and YOLOv10s. Specifically, the speed outperformed with the increases of 26.4, 26.7, 24.2, 24.8, 11.5, 18.6, and 15.6 frames per second, respectively. Additionally, the mean average precision (mAP) was improved by 14.0, 13.3, 5.5, 0.9, 0.8, 0.2, and 1.5 percentage points. The recall rate substantially increased by 16.8, 16.0, 3.2, 2.0, 0.8, 0.5, and 1.2 percentage points, respectively. Overall, the CottonBud-YOLOv5s model achieved a remarkable mean average precision (mAP) of 97.9%, a recall rate of 97.2%, and a CPU detection speed of 27.9 frames per second, indicating exceptional performance in both accuracy and speed. Visual analysis confirmed that the CottonBud-YOLOv5s model excelled in various detection scenarios, including the single-plant, multi-plant, motion blur, and small target occlusion conditions. Its superior performance in these areas highlighted its robustness and effectiveness in real-world agricultural environments, where such challenges were commonly encountered. In conclusion, the CottonBud-YOLOv5s model can offer a promising solution to the precise, real-time detection of cotton buds in densely planted environments, indicating high detection accuracy, enhanced robustness, and efficient computational performance. The finding can provide a solid visual detection for cotton mechanized topping in automated agricultural practices.

Keywords: object detection; blocking; motion blur; small target; cotton bud; convolutional neural network; YOLOv5s