

doi:10.3969/j.issn.1003-3114.2025.05.016

引用格式:曹丽英,刘洋,王喜林,等.基于同色背景下的黄瓜果实识别方法研究[J].无线电通信技术,2025,51(5):1036-1045. [CAO Liying, LIU Yang, WANG Xilin, et al. A Method for Cucumber Fruit Recognition Under Uniform-color Backgrounds[J]. Radio Communications Technology, 2025, 51(5): 1036-1045.]

基于同色背景下的黄瓜果实识别方法研究

曹丽英,刘洋,王喜林*,周恒宇,姜冬辉
(吉林农业大学信息技术学院,吉林长春130118)

摘要:与异色背景相比,同色背景下目标的识别特征较少,更易受到遮挡和复杂背景的干扰导致同色背景下对黄瓜果实的识别检测一直是研究领域的重难点之一。针对该问题,提出了一种基于YOLOv11n的目标检测网络YOLO-ACG。引入自适应动态下采样(Adaptive Dynamic Downsample, ADown)模块,融合可变形卷积和通道注意力机制,实现跨尺度特征自适应采样;构建Ghost_HGNetV2网络结构,其中高分辨率组卷积(High-resolution Group Stem, HGStem)将输入图像的通道数压缩,生成固有的特征映射,实现高效特征提取,Ghost_HGBlock模块采用知识蒸馏技术增强特征表达能力;引入上下文与空间特征校准网络结构(Context and Spatial Feature Calibration Network, CSFCN),该结构包含上下文特征校准(Context Feature Calibration, CFC)和空间特征校准(Spatial Feature Calibration, SFC),通过聚合每个像素相关的上下文信息,利用校准空间特征,确保网络正确理解图像的空间布局,进而更加精确地区分具有相似颜色的黄瓜果实和背景。经实验验证,改进后模型精确率提高4.64个百分点,召回率提高5.07个百分点,F1提高4.89个百分点,mAP值提高4.48个百分点。消融、对比实验表明,YOLO-ACG在同色背景下黄瓜果实识别中明显减少了误检和漏检的问题,且具有更高的检测精度。

关键词:特征提取;YOLO;果实识别;目标检测

中图分类号:TP391

文献标志码:A

开放科学(资源服务)标识码(OSID):

文章编号:1003-3114(2025)05-1036-10



A Method for Cucumber Fruit Recognition Under Uniform-color Backgrounds

CAO Liying, LIU Yang, WANG Xilin*, ZHOU Hengyu, JIANG Donghui
(College of Information Technology, Jilin Agricultural University, Changchun 130118, China)

Abstract: Compared with a different-color backgrounds, recognizing and detecting cucumber fruits under uniform-color backgrounds remains a key challenge due to limited distinguishing features and increased susceptibility to occlusion and background interference. To address this, we propose YOLO-ACG, a detection network based on YOLOv11n. An Adaptive Dynamic Downsample (A-Down) module is introduced, combining deformable convolution and channel attention to achieve adaptive cross-scale feature sampling. A Ghost_HGNetV2 architecture is designed, where the High-resolution Group Stem (HGStem) reduces input channels to extract efficient intrinsic features, and the Ghost_HGBlock applies knowledge distillation to enhance feature representation. A Context and Spatial Feature Calibration Network (CSFCN) network structure is introduced, which includes Context Feature Calibration (CFC) and Spatial Feature Calibration (SFC). The CFC module aggregates context information relevant to each pixel, while the SFC module leverages calibrated spatial features to ensure accurate understanding of spatial layout the image. Together, they enable the network to more precisely distinguish cucumber fruits from backgrounds with similar colors. Experimental results show that the improved model achieves 4.64 percentage points increase in precision, recall by 5.07 percentage points, F1 by 4.89 percentage points, and mAP by 4.48 percentage

收稿日期:2025-06-06

基金项目:吉林省科技发展计划项目(20250601061RC);吉林省教育厅科学研究项目(JJKH20250571KJ)

Foundation Item: Jilin Provincial Science and Technology Development Project (20250601061RC); Scientific Research Project of Jilin Provincial Department of Education(JJKH20250571KJ)

points. Ablation and comparative experiments confirm that YOLO-ACG significantly reduces false positives and missed detections, offering effective technical support for cucumber fruits recognition in complex, uniform-color environments.

Keywords: feature extraction; YOLO; fruit identification; object detection

0 引言

黄瓜是全球普遍种植的温室蔬菜,且营养价值丰富。然而,黄瓜的采摘过程繁重,且极度依赖人工,直接导致劳动力成本增加。黄瓜的绿色果实与周围环境不易区分,且采摘时往往受叶片遮挡,导致目标检测复杂度提高,在复杂环境条件下提升黄瓜果实检测的准确性,成为当前研究的焦点。研制智能且自动化的采摘系统^[1],降低人力消耗,提高生产效率,成为急需攻克的难题。

同色背景下的作物识别检测是研究的重难点之一。由于检测目标特征少、定位精度要求高、易受同色背景干扰等问题,建立准确可靠的同色环境下目标检测算法具有重要的研究价值。近年来,许多研究学者通过不同方式提取特征以改进这一问题。如 Elena 等^[2]运用基于颜色的分割算法和边缘检测算子对橄榄的边缘进行检测,并借助对图像的分析,客观评估不同橄榄样品的成熟度指数;周文静等^[3]通过改进红绿色差和大津算法(Otsu's Method, Otsu),提取 R、G 特征图直方图后展开分析,对所得到的点乘特征图进行 Otsu 运算,然后做形态学处理,进而将葡萄果穗从复杂背景中分离出来;Chuquimarca 等^[4]认为不同颜色空间在不同应用场景中可能表现更优,故而采用 RGB 和轮廓图像的多输入架构来提升水果识别精度;Mahendran 等^[5]通过灰度共生矩阵提取香蕉纹理、能量、对比度、相关性等特征,并结合卷积深度神经网络进行类别划分,从而辨别健康与患病香蕉。这些传统方法均存在人工成本高、识别速度慢等缺点。

作为深度学习重要组成的卷积神经网络(Convolutional Neural Network, CNN),在图像分类、目标检测和语义分割等任务中取得了不错的成果^[6-8]。在农业图像识别领域, CNN 相对于传统方法,可显著提升检测精度与效率^[9]。常添春等^[10]提出将改进的 CNN 应用于水果类别分类,先构建 8 层 CNN,再用参数整流线性单元替代普通整流线性单元,为防止过拟合,在各全连接层引入随机失活技术。该方法显著提高了水果分类精确率,随着深度学习框架的发展,像 TensorFlow 这类工具的诞生,使 CNN 构建更便捷,基于 TensorFlow 搭建的 CNN,如 VGG 等结构,通过不断叠加卷积层与池化层,经全连接层后,最终通过 soft-

max 分类,已成为农作物识别的标准方法^[11]。实际应用中,为进一步增强检测精度,提高检测速度,研究者们采用了若干先进技术。例如,陈锋军等^[12]提出的基于改进 YOLOv7 的油茶果实成熟度检测法,将卷积注意力机制引入主干和头层,增强了树干特征提取与目标聚焦能力,准确性较高;李莹等^[13]针对柑橘果实成熟度检测提出基于改进 YOLOv8s 的方法,通过添加混合注意力变换器模块提升果实成熟度特征的代表能力;Yang 等^[14]采用深度可分离卷积技术降低计算复杂度,通过双路径注意力门模块提升网络对西红柿与背景的区分能力。上述研究凸显出深度学习检测算法在农作物识别中的强大潜力,显著提升了成熟度检测的准确性、稳定性及智能水平。

上述文献多采用轻量化和引入注意力机制的方式提升农作物果实目标检测性能,与传统算法相比,轻量级算法在参数、计算量等方面均得到一定优化,但算法精度仍受复杂环境影响较大。针对上述问题,设计同色背景下黄瓜果实识别检测方法 YOLO-ACG,以 YOLOv11n 为基础框架,结合 YOLO 算法、ADown 模块、Ghost_HGNetV2 和 CSFCN 结构。首先,引入 HGStem 模块作为网络入口,基于 HGStem 与动态通道分配完成多粒度特征提取;其次,在主干网络中用 ADown 模块替换普通卷积 Conv 模块,实现空间与上下文分离,解决同色背景干扰问题;同时建立时域特征蒸馏与 Ghost 卷积融合架构,得到高效递归特征生成路径,实现轻量化;最后,引入 CFC 和 SFC,防止像素与上下文不匹配及特征图像与输入图像在空间上错位引发的空间误差,提高模型在复杂环境下对小目标的识别效果。

1 材料与方 法

1.1 实验数据集

数据集通过采集黄瓜图像构建。图像涵盖黄瓜在实际生长中的多种形态,如重叠、遮阴及带附着果实的图像等,共获得 450 张原始黄瓜有效图片。

通过数据增强方法处理图像,包括垂直翻转、水平旋转、随机亮度调整,以及添加噪声、马赛克等,得到增强后图片 3 150 张,如图 1 所示。黄瓜数据集总计 3 600 张图片。训练时,训练集、验证集和测试集比例为 7 : 2 : 1。



图 1 图像数据增强

Fig. 1 Image data enhancement

1.2 技术架构

与前几代 YOLO 算法相比, YOLOv11n 算法在主干网络上使用 C3k2 替代之前的 C2f 模块提升计算效率;引入 C2PSA 模块,结合通道和注意力机制增强特征提取能力;引入空间金字塔池化-快速模块,通过连续最大池化操作提升计算速度并保留多尺度特征提取能力。原始 YOLOv11n 虽在速度、检测精确度和计算效率上较前几代有部分提升,但在复杂环境和同色背景下的图像特征提取及检测精确度方面仍有改进空间。基于上述分析,主要从精确获取黄瓜果实特征信息、避免同色背景导致的黄瓜果实识别误检和错检、提升目标检测精确度入手。

本文基于 YOLOv11n 提出改进模型。首先,在主干部分引入 HGStem 模块,优化特征提取效能,同时降低计算复杂度;其次,选取 Ghost_HGBlock 模块,通过深度残差结构与 Ghost 卷积整合,降低模型参数量,提升特征表达能力;同时在多尺度下采样阶段融入 ADown 模块,提升多尺度特征表达水平;最后,在 Neck 模块中置入 CFC_CRB 和 SFC_G2 模块,增强多尺度特征融合水平,同时提升模型在复杂背景下对目标区域的关注能力,具体结构如图 2 所示。

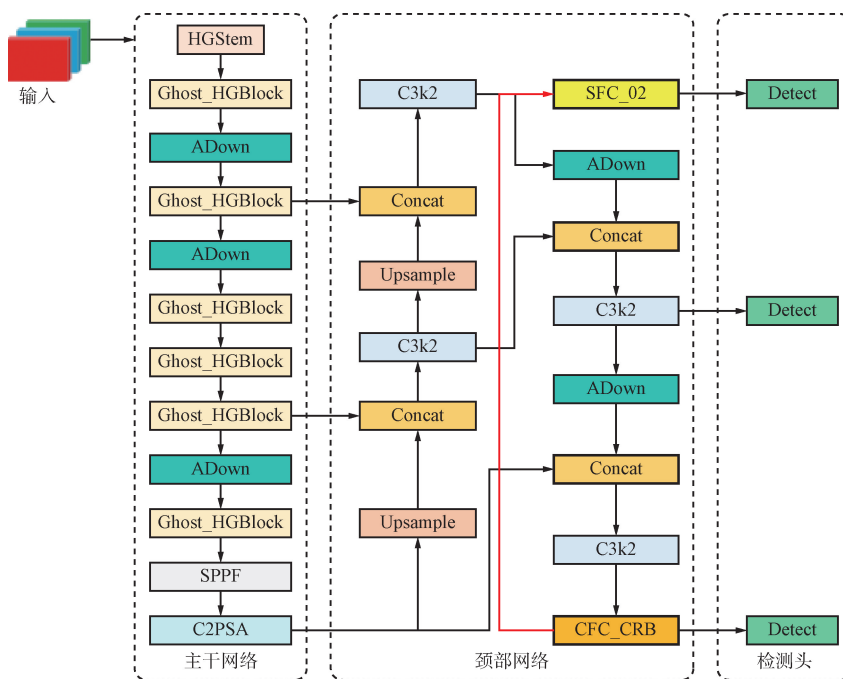


图 2 改进 YOLO-ACG 网络模型

Fig. 2 Improved YOLO-ACG network model

1.3 ADown 模块

在同色背景下的黄瓜果实识别任务中,传统下

采样方法依赖标准卷积实现特征图空间压缩,但在多尺度特征提取中易导致上下文信息损失,难以有

效抑制背景噪声。为提升模型对细节的感知能力,在主干网络中引入 ADown^[15] 模块替代常规卷积采样结构 ADown 模块融合平均池化与最大池化的优势^[16-17],兼顾全局背景信息与显著局部特征提取。其与普通卷积的结构如图 3 所示。输入特征图先通过平均池化层,在通道维度上分为子特征图 X1 和 X2。通过并行路径, X1 直接进行卷积操作, X2 通过最大池化层提取池化窗口内的最大值以减少特征图空间尺寸,再经卷积层下采样提取局部特征。最后将 2 条路径输出的特征图拼接合并,合并后的特征图作为最终输出。

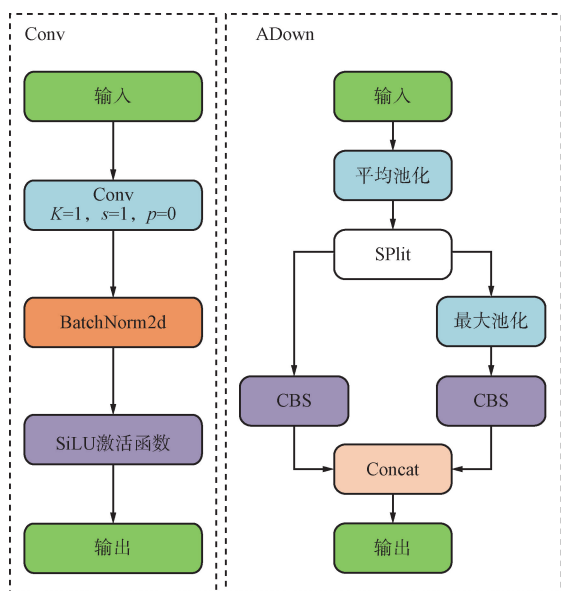


图 3 Conv 与 ADown 网络模型结构

Fig. 3 Conv and ADown network model structure

具体的实现步骤如下。

通道分裂:输入 X 通过 2×2 平均池化,步长为 1 进行预处理,以最小化局部信息损失。列式如下:

$$X' = AvgPool2d(X, 2, 1, 0) \quad (1)$$

然后, X' 沿着通道维度 (dim=1) 分为 X1 和 X2:

$$X1, X2 = split(X', dim = 1) \quad (2)$$

不同的下采样方式:卷积下采样 X1 部分,通过步长为 2 和填充为 1 的 3×3 卷积对 X1 特征进行提取和下采样:

$$y1 = Conv(x1, 3, 2, 1) \quad (3)$$

最大池化下采样 X2 部分:

通过步长为 2 和填充为 1 的 3×3 最大池化对 X2 进行下采样:

$$X2 = MaxPool2d(X2, 3, 2, 1) \quad (4)$$

然后, X2 通过 1×1 卷积进行通道变换:

$$y2 = Conv(X2, 1, 1, 0) \quad (5)$$

最后进行特征整合。y1 和 y2 在通道维度上拼接形成最终输出:

$$Y = Concat(y1, y2, dim = 1) \quad (6)$$

1.4 Ghost_HGNetV2 架构

在黄瓜果实识别中,原始 YOLOv11n 算法识别检测黄瓜果实时,往往会出现误检和漏检现象,因同色背景下黄瓜叶片与果实颜色相近,特征提取过程中难以获得有效特征信息。针对这一问题,引入轻量化特征提取结构 Ghost_HGNetV2^[18]。该架构由 HGStem 模块与 Ghost_HGBlock 模块协同构成,具备较强特征表达能力与结构灵活性,可有效缓解目标特征不明显、边界模糊等问题,提升检测精度。图 4 展示了 HGStem 网络模型结构。

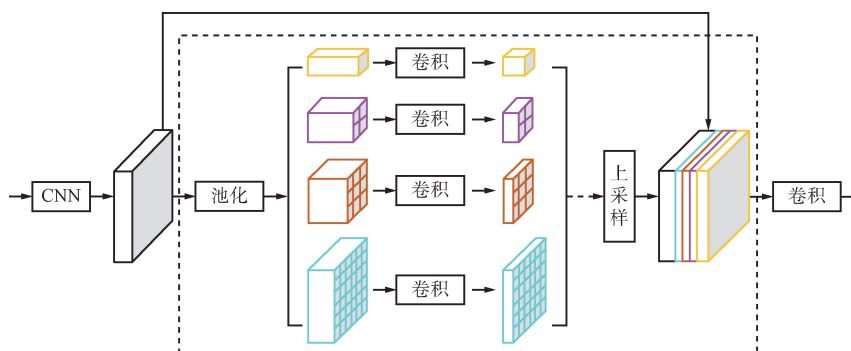


图 4 HGStem 网络模型结构

Fig. 4 HGStem network model structure

HGStem 模块作为网络的输入端,替代传统固定结构的 Stem 层。其设计融合了 HGStem 与动态通道分配机制,具备多尺度特征提取能力。

首先,进行可变形卷积特征池化。在 3×3 卷积

层中引入偏移量参数 $\delta \in R^2$, 通过空间注意力机制动态调整感受野:

$$D_{conv}(x) = \sum_{i=1}^{k^2} \sum_{j=1}^c w_i \cdot \sigma(x * \delta_i) \cdot \alpha_j \quad (7)$$

式中: σ 为 Sigmoid 激活函数, $\alpha \in [0,1]^C$ 为通道注意力权重。

其次,进行多尺度特征金字塔融合。通过 1×1 Ghost 卷积生成辅助特征图:

$$\mathcal{G}_{\text{conv}}(x) = \mathcal{D}_{\text{conv}}(x) \oplus \text{Ghost}(x) = \mathcal{D}_{\text{conv}}(x) * \mathcal{W}_g, \quad (8)$$

式中:Ghost 卷积通过特征蒸馏生成 1/4 参数量的冗余特征。

最后,进行动态通道分配。采用门控机制动态选择特征通道:

$$\mathcal{G}_{\text{gate}}(x) = \sigma(\mathcal{W}_g x) \odot \mathcal{G}_{\text{conv}}(x) + \sigma(\mathcal{W}_d x) \odot \mathcal{D}_{\text{conv}}(x). \quad (9)$$

Ghost_HGBlock 作为主干网络的基本构件,在保持结构轻量化的同时,实现特征增强与信息动态传递。模块包含 3 个部分:时域特征蒸馏层通过跨阶段特征迁移机制提升深层特征稳定性;轻量化残

差连接采用深度可分离卷积结构压缩模型参数,降低运算开销;动态计算调度机制基于特征复杂度指数 (Feature Complexity Index, FCI) 分配计算资源,提升整体推理效率。列式如下:

$$\mathcal{J}_{\text{distill}}(x) = \mathcal{W}_t \cdot \text{ReLU}(\mathcal{W}_d \cdot x + \mathcal{W}_s \cdot \mathcal{G}_{\text{prev}}), \quad (10)$$

$$\mathcal{R}_{\text{block}}(x) = \mathcal{W}_d \cdot \text{DepthwiseConv}(x) + \mathcal{W}_p \cdot \text{PointwiseConv}(x) + \beta \cdot x, \quad (11)$$

$$\text{FCI}(x) = \frac{\sum_{c=1}^C \left| \frac{\partial \mathcal{L}}{\partial x_c} \right|}{\|x\|_2}. \quad (12)$$

因此 Ghost_HGNetV2 架构兼具信息保留能力与运算高效性;其高分辨率特征入口增强模型对纹理细节与目标边缘的捕捉能力,Ghost 卷积机制显著减少冗余计算量,动态通道与计算调度策略进一步提升模型对复杂目标结构的适应能力。图 5 展示了 Ghost_HGBlock 网络模型结构。

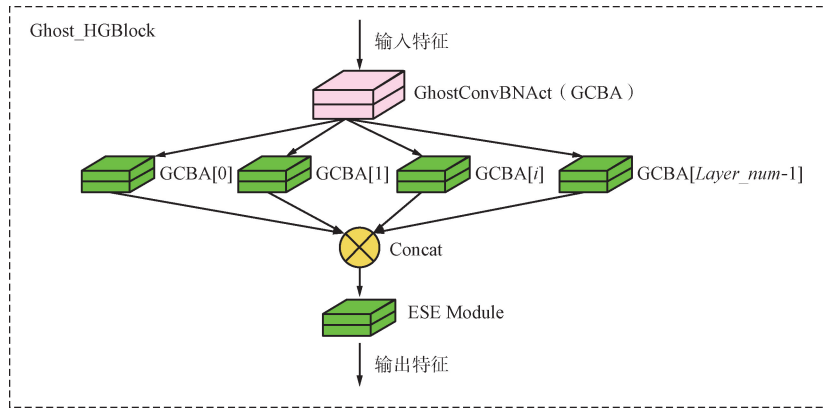


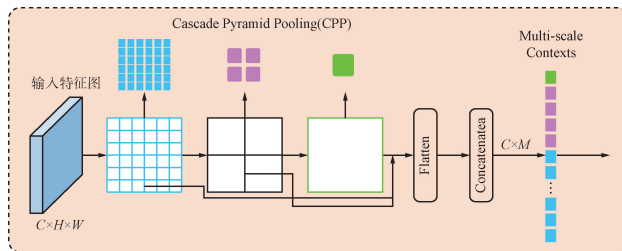
图 5 Ghost_HGBlock 网络模型结构

Fig. 5 Ghost_HGBlock network model structure

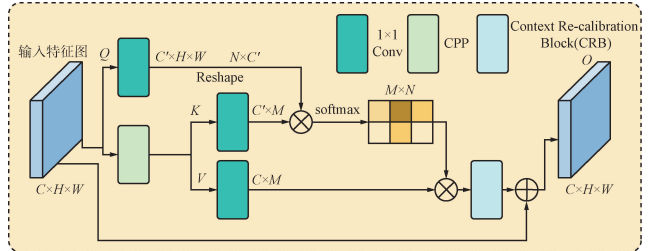
1.5 上下文与空间特征校准网络结构

在黄瓜果实识别检测中,原 YOLOv11n 模型因骨干网络无法有效捕获小目标特征,检测小目标时表现较弱,且易受背景噪声和复杂场景影响,检测精度下降。

为解决此问题,在 Neck 层引入 CSFCN 网络架构,通过引入多尺度特征融合机制,CFC^[19]可捕获不同尺度目标信息,SFC^[20]通过空间注意力机制增强模型对目标位置的感知能力,进而提高小目标检测能力,具体如图 6 所示。



(a) CFC 网络模型结构



(b) SFC 网络模型结构

图 6 CFC 与 SFC 网络模型结构

Fig. 6 CFC and SFC network model structure

CFC 模块是 CSFCN 的核心创新之一,该模块旨在解决像素与上下文不匹配问题。传统上下文建模方式通常聚合固定区域上下文信息,未顾及不同像素对上下文需求的差异,导致像素与上下文不匹配,影响分割精度。CFC 模块借助级联金字塔池化模块捕获多尺度上下文内容,基于像素-上下文相似度聚合各像素私有上下文,校正上下文特征。SFC 模块解决网络因多次下采样操作导致输出特征图与输入图像空间方位错位问题,这种错位进一步加剧分割误差,尤其在边界区域。SFC 模块依据通道维度将特征分割为多组子特征,通过可学习采样传播子特征,实现 SFC 功能。

CSFCN 的设计使网络在各阶段均可动态适应输入特征的复杂状况。CFC 提升语义层面的上下文感知能力,SFC 增强空间定位准确性,二者相互补充,共同优化目标特征表达质量。这使 CSFCN 在保持较低计算复杂度的同时,显著提升小目标检测精确率,尤其适用于实时性要求高、目标特征较少的检测场景。综上,CSFCN 通过引入上下文和空间双重校准机制,强化模型对小目标和复杂场景的鲁棒性。

2 结果与分析

2.1 实验环境

实验在 Windows 11 系统上运行,采用 PyTorch 开发深度学习模型。实验环境参数如表 1 所示。训练过程采用随机梯度下降优化,初始学习率 0.001,动量因子 0.937,权重衰减 0.000 5。输入图像归一化为 640 pixel×640 pixel,批次大小 32,训练 300 轮次。

表 1 实验参数配置

Tab. 1 Experimental parameter configuration

类别	配置
CPU	12th Gen Intel Core i7-12700KF@ 3.60 GHz
GPU	NVIDIA GeForce RTX 4060ti 16 GB
系统环境	Windows 11
框架	PyTorch 2.1.0
编程语言	Python 3.8

2.2 评价指标

本文以精确率(Precision, P)、召回率(Recall, R)、平均精度(mean Average Precision, mAP)、F1、模型参数量(Parameters)、每秒浮点运算次数(Floating Point Operations Per Second, FLOPs)和推理时间(Inference)作为算法性能的主要评价指标。其中 mAP 是评估算法性能的重要指标,F1 作为 P 和 R 的调和平均数,能更全面地评估模型综合表现。

上述主要指标通过以下公式计算:

$$AP_i = \frac{TP}{TP + FP}, \quad (13)$$

$$mAP = \frac{\sum_{i=1}^Q AP_i}{Q}, \quad (14)$$

$$P = \frac{TP}{TP + FP}, \quad (15)$$

$$R = \frac{TP}{P}, \quad (16)$$

$$F1 = \frac{2 \times R \times P}{R + P}. \quad (17)$$

2.3 模型改进实验分析

为验证本文所提出 YOLO-ACG 的有效性,首先进行与原 YOLOv11 模型的对比实验,实验结果如表 2 所示。

表 2 模型改进实验
Tab. 2 Model improvement experiments

单位:%

类别	YOLOv11				YOLO-ACG			
	P	R	F1	mAP	P	R	F1	mAP
成熟	83.67	79.32	81.43	86.51	87.96	83.58	85.71	90.94
未成熟	80.43	66.95	73.07	77.04	85.41	72.82	78.61	81.57
整体	82.05	73.13	77.33	81.77	86.69	78.20	82.22	86.25

实验结果显示,在成熟、未成熟和整体分类中,改进后模型的 P、R、F1 和 mAP 值均优于原模型。首先,未成熟类别提升最为明显,P、R、F1 和 mAP 值分别提高 4.98、5.87、5.54 和 4.53 个百分点;其次,整体类别也较为出色,上述指标分别提

高 4.64、5.07、4.89 和 4.48 个百分点;最后,成熟类别也有提升,分别提高 4.29、4.26、4.28 和 4.43 个百分点。综上,本文提出的 YOLO-ACG 算法相较于原 YOLOv11 算法,更适用于黄瓜果实的识别检测。

2.4 消融实验分析

为评估改进算法有效性,在相同设备、数据集及

相同实验条件下进行 8 次消融实验,以验证引入模块的有效性。实验对比结果如表 3 所示。

表 3 消融实验

Tab. 3 Ablation experiments

模块				P	R	F1	mAP
YOLOv11n	ADown	CSFCN	Ghost_HGNetV2				
√				82.05	73.13	77.33	81.77
√	√			83.44	74.54	78.73	83.22
√		√		83.21	76.25	79.57	84.03
√			√	85.39	74.60	79.63	83.86
√	√	√		85.56	77.50	81.33	85.73
√		√	√	85.09	77.93	81.35	85.76
√	√		√	86.78	77.22	81.72	86.15
√	√	√	√	86.69	78.20	82.22	86.25

单位:%

第 1 组实验选用 YOLOv11n 作为基准模型进行实验。第 2~4 组实验是在基准模型的基础上分别添加 ADown、CSFCN、Ghost_HGNetV2 模块或网络,其模型的 F1 和 mAP 与第 1 组实验数据相比分别提升了 1.40 与 1.45、2.24 与 2.26 个百分点,以及 2.30 与 2.09 个百分点。第 5 组实验是在第 2 组上添加了 CSFCN 结构,其模型的 F1 和 mAP 与第 1 组实验数据相比分别提升了 4.00 与 3.96 个百分点。第 6 组实验是在第 3 组上添加 Ghost_HGNetV2 结构,其模型的 F1 和 mAP 与第 1 组实验数据相比分别提升了 4.02 与 3.99 个百分点。第 7 组实验是在第 2 组上添加 Ghost_HGNetV2 结构,其模型的 F1 和 mAP 与第

1 组实验数据相比分别提升了 4.39 与 4.38 个百分点。第 8 组实验是将 YOLO-ACG 算法进行实现,最终模型的 F1 和 mAP 与第 1 组实验数据相比分别提升了 4.89 与 4.48 个百分点,进一步说明了改进后的模型具有很好的泛化能力,验证了方案的可行性。

2.5 对比实验分析

为了全面评估 YOLO-ACG 模型的性能,系统地将其与各种先进的目标检测模型进行了比较,除了对 YOLOv8n、YOLOv9c、YOLOv10b、YOLOv10l 和 YOLOv11n 不同版本的 YOLO 模型进行评估外,还引入了 Faster R-CNN 和 SSD 算法进行对比实验。实验结果如表 4 所示。

表 4 对比实验

Tab. 4 Comparative experiments

算法	P/%	R/%	F1/%	mAP/%	Parameters/M	FLOPs/G	Inference/ms
Faster R-CNN	82.58	74.42	78.27	80.94	42.80	180.7	92.42
SSD	79.42	70.83	74.88	77.24	24.10	45.6	30.25
YOLOv8n	78.01	67.99	72.65	76.53	2.68	6.8	15.79
YOLOv9c	83.05	73.30	77.87	81.57	21.35	84.1	31.66
YOLOv10b	77.74	72.58	75.07	80.51	19.00	91.6	27.48
YOLOv10l	80.73	72.86	76.59	81.50	24.31	120.0	33.47
YOLOv11n	82.05	73.13	77.33	81.77	2.58	6.0	18.12
YOLO-ACG	86.69	78.20	82.22	86.25	3.59	8.8	22.00

对比实验结果显示, YOLO-ACG 模型的 F1 和 mAP 均高于其他模型。F1 值相比 YOLOv8n、YOLOv9c、YOLOv10b、YOLOv10l、YOLOv11n 分别提高了 9.57、4.35、7.15、5.63 和 4.89 个百分点; mAP

值相比 YOLOv8n、YOLOv9c、YOLOv10b、YOLOv10l、YOLOv11n 分别提高了 9.72、4.68、5.74、4.75 和 4.48 个百分点。同时,相较于 SSD 和 Faster R-CNN, YOLO-ACG 的 F1 分别高出 7.34 和 3.95 个百

分点, mAP 分别提升了 9.01 和 5.31 个百分点, 在精度与召回率方面也均有明显优势, 充分验证了模型在复杂环境下的鲁棒性与稳定性。

尽管 YOLO-ACG 在参数量和计算复杂度上略高于 YOLOv11n, 但相比于 YOLOv9c、YOLOv10l 和 Faster R-CNN 等模型, 其整体依旧保持了轻量化优势。同时, YOLO-ACG 的推理时间为 22.00 ms, 显著优于 Faster R-CNN, 具备较强的实时性与部署灵活性。综上所述, YOLO-ACG 在保持计算资源友好的前提下, 实现了更优的检测性能, 尤其适用于对精度和速度要求并重的边缘设备和实际农业场景果实检测任务。

2.6 检测结果分析

在本研究设置的同色背景实验场景中, 黄瓜果实与叶片颜色接近, 存在遮挡、纹理重叠等因素, 构成典型的背景噪声干扰环境。图 7 展示了不同模型在该干扰背景下的可视化检测结果及检测度的对比。YOLOv8n、YOLOv10b、YOLOv10l 和 YOLOv11n 分别在未成熟果实识别中出现漏检情况, 检测度分别为 0.45、0.72、0.65 和 0.70, 在成熟果实检测中亦存在检测误差, 检测度分别为 0.62、0.73、0.70 和 0.54。上述结果表明, 同色背景对传统模型识别精度具有干扰作用, 影响识别准确性。

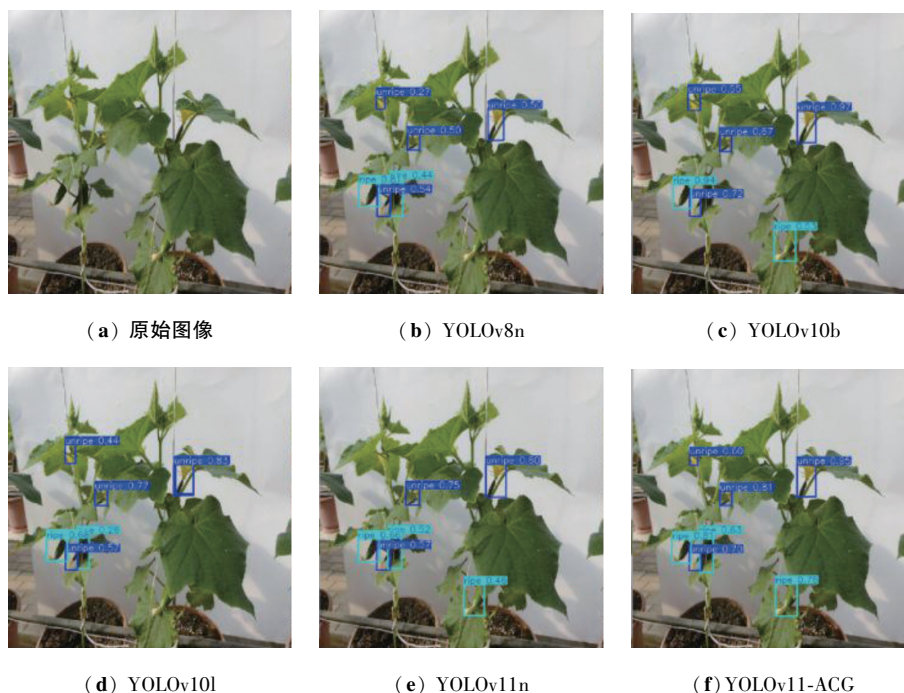


图 7 黄瓜检测前后对比

Fig. 7 Comparison of cucumber before and after testing

相比之下, YOLO-ACG 模型在同样的背景噪声条件下表现稳定, 未出现漏检与错检现象, 未成熟果实与成熟果实的检测度分别达到 0.76 和 0.73, 显著高于其他模型。结果充分说明引入的 ADown 采样模块、Ghost_HGNetV2 结构与 CSFCN 校准机制在复杂同色背景下有效增强了模型对细粒度目标特征的捕捉能力, 并提升了模型对背景噪声干扰的鲁棒性。

2.7 异色背景下检测对比分析

为进一步验证改进模型的鲁棒性和实用性, 使用 Labeled-cucumber-dataset 数据集作为异色背景数据集, 其特点是黄瓜根、茎、叶覆盖较少, 其他背景为

白色, 使得环境对黄瓜果实检测影响较小, 能够较好地突出黄瓜果实在异色背景下的检测效果。

如表 5 所示, 改进后的 YOLO-ACG 模型在 P、R、F1 和 mAP 四项关键性能指标上均优于原始 YOLOv11n 模型, 分别提升了 4.77、5.56、5.17 和 3.22 个百分点。此外, YOLO-ACG 的 Inference 为 11.73 ms, 虽然相比 YOLOv11n 略有增加, 但仍保持较高的实时性, 表明本文所提出的模型在异色背景下仍具备良好的泛化能力和检测稳定性。尽管在参数量和计算复杂度方面略有增加, 但相较于性能的整体提升而言, 这一资源消耗仍处于合理可控范围内。

表 5 异色背景对比实验

Tab. 5 Heterochromatic background contrast experiment

模型	P/%	R/%	F1/%	mAP/%	Parameters/M	FLOPs/G	Inference/ms
YOLOv11n	80.58	78.61	79.58	85.05	2.58	6.0	8.41
YOLO-ACG	85.35	84.17	84.75	88.27	3.59	8.8	11.73

综上所述, YOLO-ACG 模型在同色背景与异色背景 2 类场景中均展现出优异的检测性能, 尤其在颜色接近、背景复杂的同色背景条件下表现出更强的特征区分与目标定位能力, 凸显其在复杂环境下的鲁棒性与实用性, 为后续温室、采摘机器人等环境中的果实识别应用提供了有力支撑。

3 结束语

为解决黄瓜果实在种植环境中因叶片遮挡、颜色相近等问题导致的检测困难, 本文聚焦于同色背景下果实识别的关键挑战, 提出了一种改进的 YOLO-ACG 算法。引入 HGStem 模块作为网络入口, 基于 HG-Stem 与动态通道分配完成多粒度特征提取; 建立时域特征蒸馏与 Ghost 卷积融合架构, 得到高效的递归特征生成路径; 引入 CFC 和 SFC, 防止像素与上下文之间的不匹配和特征图像与输入图像在空间上错位引发的空间误差, 明显提高了模型在多尺度目标检测和小物体识别中的水平, 同时减少了计算资源的占用。YOLO-ACG 算法在黄瓜果实数据集上对比原算法, 在 P、R 和 mAP 值上均有提升。mAP 值提升了 4.64 个百分点, 且算法实时性能强、鲁棒性好。未来工作中, 将尝试将 YOLO-ACG 算法嵌入式植入, 为黄瓜果实的自动化和智能化采摘提供检测支持。

参考文献

[1] XIAO X, WANG Y N, JIANG Y M. Review of Research Advances in Fruit and Vegetable Harvesting Robots[J]. Journal of Electrical Engineering & Technology, 2024, 19(1): 773-789.

[2] ELENA G, VINCENT B, FERNÁNDEZ A J P, et al. Determination of the Olive Maturity Index of Intact Fruits Using Image Analysis[J]. Journal of Food Science and Technology, 2015, 52(3): 1462-1470.

[3] 周文静, 赵康, 马晓晓, 等. 基于改进红绿色差和 Otsu 的葡萄果穗图像分割[J]. 中国农机化学报, 2023, 44(1): 172-177.

[4] CHUQUIMARCA L, VINTIMILLA B, VELASTIN S. Classifying Healthy and Defective Fruits with a Multi-input Architecture and CNN Models[C]//Proceedings of the 14th International Conference on Pattern Recognition Systems (ICPRS). London: IEEE, 2024: 1-7.

[5] MAHENDRAN T, SEETHARAMAN K. Feature Extraction and Classification Based on Pixel in Banana Fruit for Disease Detection Using Neural Networks[C]//2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT). Ahmedabad: IEEE, 2023: 1-7.

[6] PURWONO P, MA'ARIF A, RAHMANIAR W, et al. Understanding of Convolutional Neural Network (CNN): A Review[J]. International Journal of Robotics and Control Systems, 2022, 2(4): 739-748.

[7] 周耀威, 孔令军, 戴琪, 等. 基于多头自注意力机制的轻量级包裹破损检测算法[J]. 无线电通信技术, 2025, 51(1): 161-170.

[8] 郑涵今, 薛健, 吕科, 等. 基于 X-Ray 图像的锂电池电极检测方法[J]. 无线电通信技术, 2023, 49(3): 532-540.

[9] 曾平平, 李林升. 基于卷积神经网络的水果图像分类识别研究[J]. 机械设计与研究, 2019, 35(1): 23-26.

[10] 常添春, 王联国. 基于 CBAM-ResNet50 模型的水果图像识别[J]. 软件工程与应用, 2024, 13: 61-78.

[11] PANG B, NIJKAMP E, WU Y N. Deep Learning with TensorFlow: A Review[J]. Journal of Educational and Behavioral Statistics, 2020, 45(2): 227-248.

[12] 陈锋军, 陈闯, 朱学岩, 等. 基于改进 YOLOv7 的油茶果实成熟度检测[J]. 农业工程学报, 2024, 40(5): 177-186.

[13] 李莹, 刘梦莲, 何自芬, 等. 基于改进 YOLOv8s 的柑橘果实成熟度检测[J]. 农业工程学报, 2024, 40(24): 157-164.

[14] YANG G L, WANG J X, NIE Z L, et al. A Lightweight YOLOv8 Tomato Detection Algorithm Combining Feature Enhancement and Attention[J]. Agronomy, 2023, 13(7): 1824.

[15] LIANG X M, JIA H J, WANG H, et al. ASE-YOLOv8n: A Method For Cherry Tomato Ripening Detection[J]. Agronomy, 2025, 15(5): 1088.

[16] GAO Y J, GU X Y, ZHANG H Y, et al. Runtime Performance Prediction for Deep Learning Models with Graph Neural Network[C]//2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP). Melbourne: IEEE, 2023: 368-380.

[17] HASAN M M, NISHI J S, HABIB M T, et al. A Deep Learning Approach to Recognize Bangladeshi Shrimp Species[C]//2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). Lucknow: IEEE, 2023: 1-5.

- [18] YU J Q, GUAN Y L, LEI Z, et al. GLD-YOLO: A Lightweight Real-time Object Detection Algorithm Based on Improved YOLOv8[C]//2025 IEEE International Conference on Industrial Technology (ICIT). Wuhan: IEEE, 2025:1-6.
- [19] 罗浚锐,张宝华,张艳月,等. 基于上下文自校准双重注意力的目标检测算法[J]. 激光与光电子学进展, 2021, 58(12):182-188.
- [20] LI K G, GENG Q C, WAN M X, et al. Context and Spatial Feature Calibration for Real-time Semantic Segmentation [J]. IEEE Transactions on Image Processing, 2023, 32: 5465-5477.

作者简介:

曹丽英 女,(1978—),博士,教授。主要研究方向:智慧农业、农业遥感。

刘洋 男,(1999—),硕士研究生。主要研究方向:农业信息化。

(*通信作者)王喜林 男,(1973—),硕士,讲师。主要研究方向:数理统计。

周恒宇 男,(2006—)。主要研究方向:农业信息化。

姜冬辉 男,(2000—),硕士研究生。主要研究方向:智慧农业。