DOI: 10.12398/j.issn.2096-7217.2025.04.004

# 基于改进YOLOv8n的樱桃成熟度检测研究

李瑞鑫[1]，曹悦[2]，司秀丽[1,3]，马丽[2*]

（1. 吉林农业大学信息技术学院，吉林长春，130118；2. 无锡学院，江苏无锡，214105；
3. 吉林省农业人工智能研究基地，吉林长春，130118）

**摘要：**精确检测樱桃果实情况对提升果实品质、优化采摘时机及实现果园自动化管理具有重要意义。然而，由于樱桃果实体积小、颜色多样、背景复杂及重叠遮挡等因素，检测仍面临诸多挑战。为此，提出一种基于YOLOv8n架构的轻量化改进模型ACS-YOLOv8n，以提升检测精度与效率。该模型在主干网络中引入自适应下采样模块（ADown），融合平均池化与最大池化优势，结合多路径卷积操作，有效保留全局信息与局部显著特征，增强小目标检测能力。同时，采用Context Guided Block（CGB）替代传统C2f模块，通过局部、上下文与全局特征的联合提取，提升模型对目标与环境关系的感知能力与多尺度特征提取效果。在Neck部分，引入基于GSConv与VoV-GSCSP的Slim-Neck结构，优化特征融合方式，显著降低计算冗余，兼顾速度与精度。试验结果显示，ACS-YOLOv8n在复杂果园环境下的精度达92.94%，召回率90.66%，*mAP*50为96.57%；相较YOLOv8n，精度提升2.62个百分点，召回率提升1.57个百分点，平均精度提升2.08个百分点，同时模型参数量与计算量分别减少31.90%和33.33%，模型体积为原始模型的71.43%。该模型在复杂环境中表现出良好的检测效果与鲁棒性，为樱桃自动采摘与果园智能化管理提供了有效技术支撑。

**关键词：**樱桃；成熟度检测；YOLOv8n；多尺度特征提取；目标检测；深度学习

**中图分类号：**S24 **文献标识码：**A **文章编号：**2096-7217（2025）04-0040-10

李瑞鑫，曹悦，司秀丽，马丽．基于改进YOLOv8n的樱桃成熟度检测研究[J]．智能化农业装备学报（中英文），2025，6（4）：40－49

LI Ruixin，CAO Yue，SI Xiuli，MA Li. Research on cherry ripeness detection based on improved YOLOv8n [J]. Journal of Intelligent Agricultural Mechanization，2025，6(4)：40－49

## 0 引言

樱桃果实富含活性物质[1]，兼具抗氧化等功能[2]，有较高的经济价值[3]。目前，美国、土耳其、西班牙、伊朗和乌克兰是全球主要产区。中国为重要生产国，山东、陕西、河南等地为主产区，2022年产量达680 kt，位居世界第15位。全球总产量于2022年达4 687 kt[4]。当前，樱桃采摘主要依赖人工，受成熟期短、劳动力紧缺和劳动强度大等限制，影响种植规模；而成熟度则直接关系到果品品质、采摘管理及市场效益[5]。

为推动果实采摘的智能化与精细化，近年来，国内外学者在苹果[6]、香蕉[7]、芒果[8]、柿子[9]等果实的成熟度检测方面，积极探索机器视觉与光谱分析等方法，取得了诸多进展。例如，薛建新等[10]基于高光谱成像实现杏果成熟阶段识别；苏斐等[11]将MobileNetv1引入YOLO-v3，实现绿熟期番茄检测；张志远等[12]通过优化特征融合提出改进YOLO v5模型用于樱桃识别；苗荣慧等[13]则结合MobileNetv3与注意力机制，构建轻量YOLO v7模型检测樱桃番茄成熟度。这些研究为果实检测提供了重要参考。

YOLO系列算法以其结构轻量、检测高效的优势，已成为目标检测领域的主流方法，并在复杂果园环境中具有较高适用性。然而，基于YOLO的樱桃检测研究仍相对不足。针对这一问题，本研究在YOLOv8n模型的基础上，结合樱桃果实在形态尺寸与颜色分布方面的特征，对网络结构进行有针对性的优化设计。研究从多尺度特征提取、上下文信息融合与模型轻量化3个方面展开，以实现对不同成熟度樱桃的精准检测。

## 1 材料与方法

### 1.1 试验材料

本研究所用樱桃成熟度数据集采集于2024年5月

收稿日期：2025-04-22 修回日期：2025-06-11

基金项目：国家自然科学基金项目（U19A2061）；无锡学院科研启动项目（2024r007）

第一作者：李瑞鑫，男，2000年生，山东滨州人，硕士研究生；研究方向为智慧农业。E-mail：liruixin0116@163.com

*通信作者：马丽，女，1979年生，吉林长春人，博士，教授；研究方向为智慧农业。E-mail：mali@cwxu.edu.cn

1—20日，地点为中国山东省潍坊市临朐县五井镇某樱桃园。图像由iPhone11手机于每日09:00—18:00拍摄，拍摄距离为10～50 cm，涵盖樱桃的左、右及正面角度，以及重叠、粘附等典型生长状态，以增强数据多样性并降低模型过拟合风险。如图1所示，共采集1 681张JPEG格式图像（3 024像素×4 034像素），经剔除模糊、曝光严重、遮挡明显及与研究无关图像后，最终筛选出1 201张高质量图像构建成熟度检测数据集。




**图1　不同角度和光线拍摄的樱桃果实**

Figure 1　Cherry fruits photographed from different angles and lighting conditions

### 1.2　数据标注与构建

本研究使用Make Sense图像标注软件对樱桃图像进行手动矩形标注，标注框紧贴目标，以减少背景干扰。完成标注后，软件生成与图像名称对应的txt文件，每行记录1个目标的标注信息。目标樱桃分为3类：成熟樱桃、未成熟樱桃和半成熟樱桃[14]。如图2所示。




**图2　不同种类的樱桃果实**

Figure 2　Different varieties of cherry fruits

深度学习网络在特征提取过程中依赖大量标注数据，但噪声和标注错误可能影响模型性能，降低准确性和泛化能力。当数据中存在大量错误标注时，模型难以学习有效模式，导致表现下降。在数据量受限的情况下，过拟合问题尤为突出，网络可能会过度依赖噪声或干扰信息，降低鲁棒性并显著影响测试精度。为解决这一问题，本研究提出离线数据增强策略，通过随机旋转、亮度调整和马赛克等技术，提高模型的泛化能力和鲁棒性，减轻过拟合影响[15]。数据增强后，原有1 201张标注图像扩增至6 416张，数据集划分为训练集（4 997张）、验证集（714张）和测试集（705张）。图3展示了增强后的图像，表1列出了标签分布。




**图3　樱桃果实数据增强**

Figure 3　Cherry fruits data augmentation

**表1　数据集的标签分布表**

Table 1　Label distribution of the dataset

| 数据集分类 | 未成熟/个 | 半成熟/个 | 成熟/个 |
|---|---|---|---|
| 训练集 | 22 870 | 13 405 | 13 063 |
| 验证集 | 3 197 | 1 890 | 1 704 |
| 测试集 | 3 109 | 1 815 | 1 873 |
| 总计 | 29 176 | 17 110 | 16 640 |

## 2　改进YOLOv8n的樱桃检测方法

### 2.1　YOLOv8算法

YOLO（you only look once）算法[16]是一种单阶段目标检测方法，以高效性和精确性著称。该模型将输入图像划分为多个网格，同时，模型对各网格内目标的置信度及其对应的边界框进行预测，并最终通过非极大值抑制算法筛选最优检测结果。凭借速度与精度的良好平衡，YOLO系列算法在目标检测领域得到了广泛应用。

YOLOv8[17]是YOLO系列中的较轻量化版本，在检测精度和速度方面表现突出，特别是在主流目标检测模型中展现了显著优势。相比YOLOv5[18]，YOLOv8在模型设计上做了不少优化。它在特征金字

塔网络(FPN[19])和路径聚合网络(PANet[20])的上采样阶段省略了一些卷积结构,并用C2f模块取代了C3模块。在Head部分,YOLOv8采用了解耦头设计,损失函数由分类损失和回归损失组成,其中分类损失采用VFL Loss[21],回归损失结合了CIoU Loss和分布式焦点损失(DFL-distribution focal loss)[21]。此外,它还引入了无锚框(anchor-free)预测框设计,使检测流程更加简洁高效。

在实际应用中,轻量化模型不仅能降低计算成本,还能提升检测效率。因此,本研究在YOLOv8n的基础上进行了优化,使其更适用于樱桃成熟度检测。改进后的模型命名为ACS-YOLOv8n,主要优化点包括以下几个方面:首先,在主干网络中,使用ADown模块替代传统的下采样操作,并引入C2f_ContextGuided模块以增强特征提取能力。其次,在Neck部分,用GSConv模块替代了常规卷积操作,以进一步提高计算效率。此外,C2f模块被改进为VoV-GSCSP模块。这些优化措施在保持模型轻量化的同时,显著提升了检测性能。图4展示了YOLOv8的原始网络结构,而图6则为改进后YOLOv8的网络结构,其中红色标记部分突出了关键的改进点。

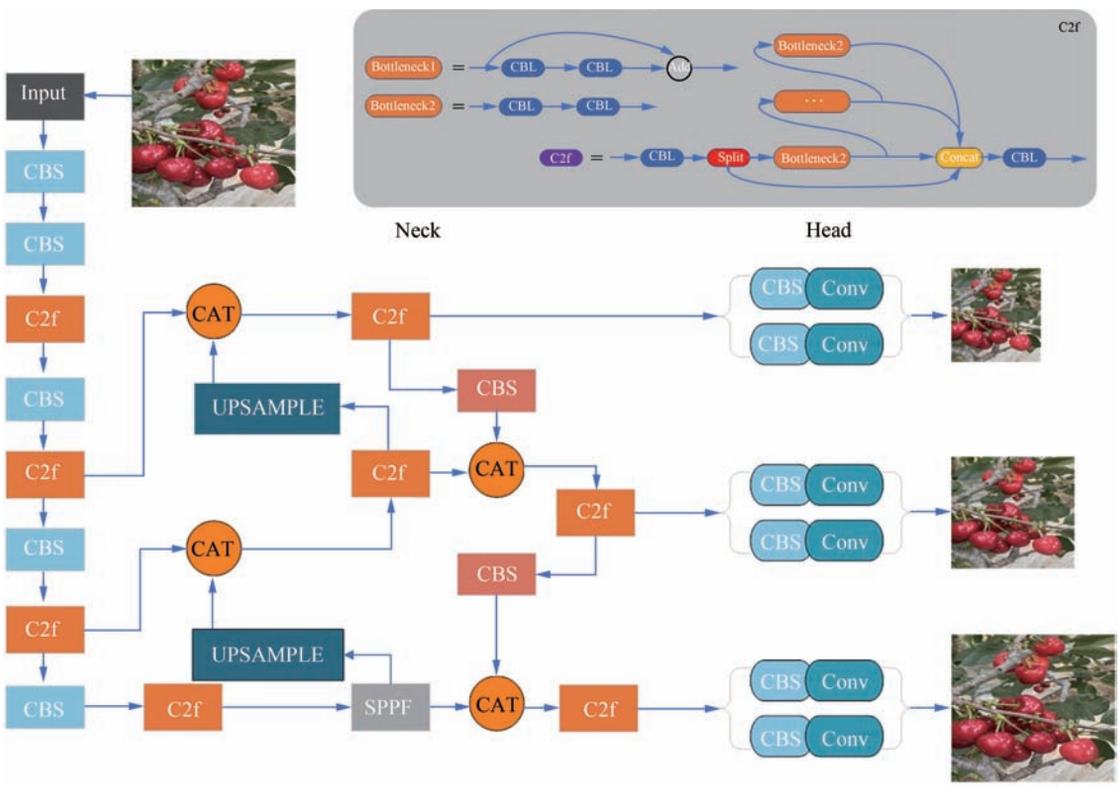



**图4 YOLOv8模型**

Figure 4 YOLOv8 baseline model

## 2.2 引入Adown自适应下采样模块

标准卷积下采样通过增大步长缩小特征图尺寸,但这可能导致小目标樱桃的关键信息丢失,影响后续层对小目标的识别能力。为提高小目标检测精度,本研究引入基于ADown[22]的下采样方法。ADown结合最大池化和平均池化的优势,最大池化突出显著特征,平均池化保留全局信息,通过卷积运算减少信息损失并保留多层次特征,增强对细节的感知能力。如图5所示。图片左侧展示了传统卷积下采样的流程,而右侧则呈现了ADown下采样的实现方法。传统方法中,输入特征图依次经过卷积层(Conv)、批量归一化层和非线性激活函数SiLU处理,最终得到下采样后的输出特征图。

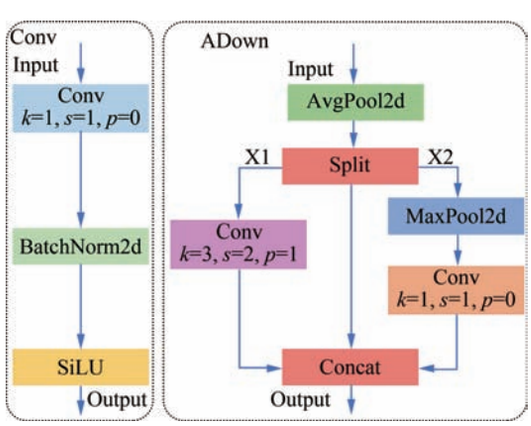



**图5 Adown结构图**

Figure 5 Adown module architecture

与传统方法不同,ADown下采样采用了更精细的策略:首先对输入特征图进行平均池化,然后在通道维度上将

结果分为两个部分，分别命名为X1和X2。X1直接送入卷积层进行处理，而X2则先通过最大池化提取局部最大值，再经过卷积层实现下采样。最后，两条路径的特征图经过拼接(Concatenation)整合成最终输出。通过这种改进设计，ADown不仅在保持全局信息的同时加强对局部显著特征的提取，还显著提升了模型对小目标的检测性能。

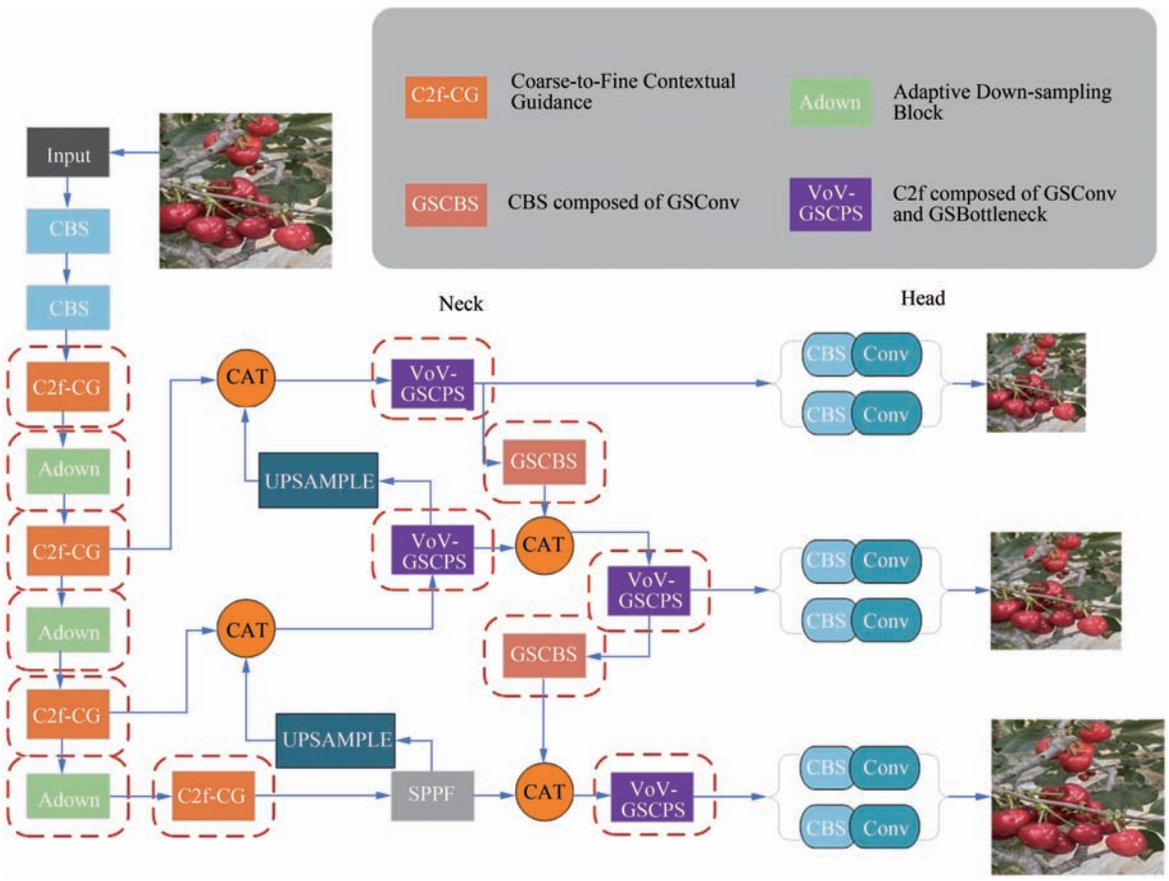



**图 6　ACS-YOLOv8n网络结构图**

Figure 6　ACS-YOLOv8n network architecture

## 2.3　Context Guided Block模块

在YOLOv8结构中，C2f模块由多个Bottleneck单元组成，每个单元包含两层卷积操作，用于提取深层特征信息。尽管这种设计有助于学习多尺度特征，但也增加了参数量，影响了模型的轻量化性能。为此，本研究提出了Context Guided Block(CGB)结构[23]，替代C2f模块中的Bottleneck，旨在减少模型规模并提升对上下文信息的理解能力。CGB结构包括4个核心部分：局部特征提取单元、周围上下文捕获单元、联合特征融合单元和全局上下文建模单元，如图7所示。

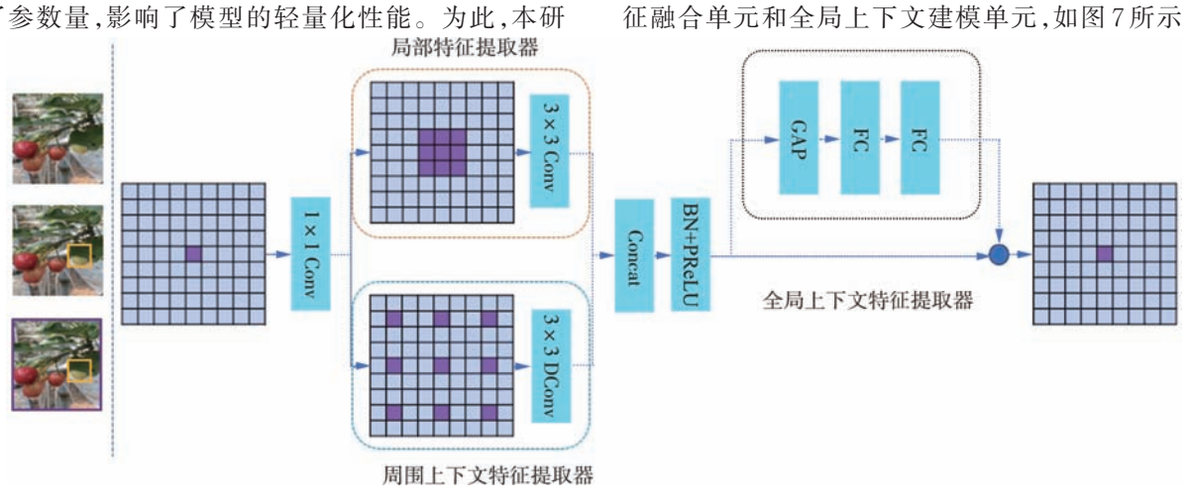



**图 7　Context Guided Block模块结构图**

Figure 7　Context Guided Block module architecture

在传统下采样方法中,模型常面临信息丢失、分辨率降低和感受野受限等问题,影响精细识别任务的精度。CGB通过融合局部与全局上下文信息来增强模型对目标和背景的关联理解,从而提高识别准确度。CGB的工作原理可分为两步:(1)局部特征提取通过等效3×3卷积从周围8个相邻特征提取局部信息;同时,周围上下文提取模块采用3×3空洞卷积扩大感受野,捕获更丰富的上下文特征。局部和上下文特征通过联合特征提取单元融合,经过批量归一化和参数化ReLU激活函数优化特征表示。(2)全局上下文整合通过全局平均池化层提取全局信息,再利用多层感知机(MLP)对全局特征进行深度处理,增强模型的整体上下文理解能力。最终,局部和全局特征通过加权操作结合,生成更强输出特征图。引入CGB结构替代Bottleneck后,试验结果表明该设计显著提升了樱桃成熟度检测任务的性能和准确性。

### 2.4 Slim-Neck模块

本改进模型采用了Slim-Neck结构[24],该结构由GSConv和VoV-GSCSP两部分组成。图8展示了该模块的网络结构。GSConv的引入旨在加速预测计算。在卷积神经网络(CNN)中,图像通常需要经过从空间信息逐步向通道信息转化的过程。密集卷积能尽可能保留隐藏连接,但增加了计算成本,而稀疏卷积则通过切断部分连接来降低计算复杂度[25]。然而,GSConv在保留必要连接的同时,有效减少了模型深度所带来的计算阻力。在Neck阶段,由于特征图已经较为细长且不需要进一步压缩,因此GSConv的应用显得特别合适。它能够在保留特征连接的同时减少冗余信息,从而避免了传统卷积操作在这一阶段可能带来的额外计算负担。因此,GSConv的引入不仅优化了特征提取的效率,还提升了模型的整体性能。

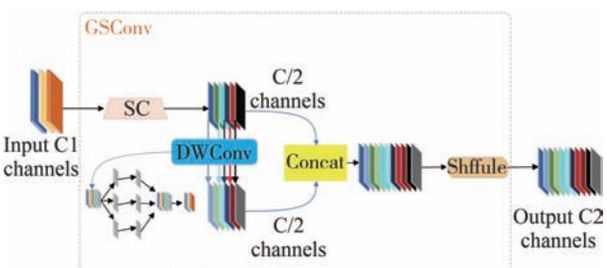



**图8 GSConv模块结构图**

Figure 8 GSConv module architecture

GSConv提供了一种有效的方式,既能降低模型复杂度,又能保持较高精度,从而在模型的准确性和速度之间实现了更好的平衡[26]。此外GSConv引入了Slim-Neck设计范式,这种设计显著提升了检测器的计算成本效益。基于此,本改进模型在Neck部分使用VoV-GSCSP模块替代了原有的C2f模块。这一设计不仅提高了模型的效率,还进一步提升了其性能。图9展示了VoV-GSCSP的结构。在GSbottleneck结构中,包含两个GSConv和一个标准卷积操作,这些操作分别应用于输入特征。标准卷积操作在该结构中扮演着促进梯度流动的角色,能够有效避免梯度消失问题。此外,该设计类似于残差连接,通过促使网络学习输入与输出之间的残差关系,从而提升了训练效率和模型性能。GSbottleneck的最终输出不是通过简单的特征拼接,而是采用特征相加的方式。这种设计强化了残差学习机制,使得模型能够更加高效地利用输入特征,从而形成一个紧凑而高效的瓶颈结构[27-28]。

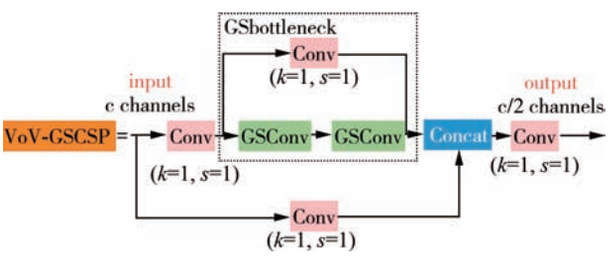



**图9 VoV-GSCSP and GSbottleneck 模块结构**

Figure 9 VoV-GSCSP and GSbottleneck module architectures

### 2.5 模型训练与算法性能指标

#### 2.5.1 模型训练

本研究在统一试验环境下完成模型训练与测试。硬件包括Intel i5-12490F CPU、NVIDIA RTX 4060 GPU(8 GB显存)、32 GB内存,操作系统为Windows 10(64位),软件环境为Python 3.8、PyTorch 2.1.0、CUDA 12.1与Pycharm 2022.1.1。训练设置为:学习率0.01,动量0.937,权重衰减0.000 5,批量大小16,训练轮次300,输入图像尺寸为640×640。该配置确保了试验结果的稳定性与可比性。

#### 2.5.2 评价指标

在本研究中,我们采用多种评价指标来全面评估所提出模型的性能,包括*mAP*50、精度(*Precision*)、召回率(*Recall*)、参数量以及*GFLOPs*。具体而言,*mAP*(mean average precision)通过计算不同IoU阈值下的平均精度来衡量模型的性能,计算*mAP*50,如公式(1),其中$N$为类别数,$AP(i)$为类别$i$的平均精度。精度(*Precision*)定义为正确预测的正样本与所有预测为正样本的比例,计算如公式(2),其中*TP*为真阳性(true positives),*FP*为假阳性(false positives)。召回率(*Recall*)则定义为正确预测的正样本与所有真实正样本的比例,公式如式(3),其中*FN*为假阴性(false negatives)。此外,参数量表示模型的可训练参数总数,通常以百万兆为单位。最后,*GFLOPs*(giga

floating point operations per second)用于衡量模型推理时的计算复杂度，计算公式见式(4)，其中 *Total FLOPs* 表示推理过程中所需的总浮点运算次数。该指标有助于全面评估模型的准确性和计算效率。

$$mAP=\frac{1}{N}\sum_{i=1}^{N}AP(i) \tag{1}$$

$$Precision=\frac{TP}{TP+FP} \tag{2}$$

$$Recall=\frac{TP}{TP+FN} \tag{3}$$

$$GFLOPs=\frac{Total\ FLOPs}{10^{9}\times Inference\ Time} \tag{4}$$

式中：*TP*——正确检测出樱桃目标的数量；

*FP*——错误检测樱桃目标的数量；

*FN*——未被检出樱桃目标的数量。

## 3　试验结果与分析

### 3.1　消融实验

首先，本研究通过将YOLOv8n骨干网络中的3个卷积层替换为ADown模块(一种下采样卷积)，实现了模型准确率提升1.67个百分点，同时将参数数量减少了9.30%，从而使模型更加轻量化。接着，在骨干网络中3个C2f模块的基础上，增加了带有上下文模块的C2f，从而减少了30.24%的参数量，并通过有效的特征提取和上下文信息融合，提升了图像检测任务的准确性与效率。其次，在Neck部分，4个C2f模块被替换为VoV-GSCSP模块，同时将向上连接的两个卷积替换为GSConv，进一步减少了11.30%的模型参数数量。使用GSConv和VoV-GSCSP保留了通道的隐式连接性，对于精简后的特征图，GSConv的应用减轻了深层对数据流的阻力，并显著减少了推理时间。经过这些改进后，模型参数量减少了31.90%，模型大小减少了1.7 MB，*GFLOPs*计算量减少了33.33%，准确率提升了2.62个百分点，*mAP*50提升2.08个百分点，同时保持了相同的检测速度。整体来看，改进后的模型在检测性能上得到了显著提升，同时更加轻量化(表2是消融实验效果对比)，使得YOLOv8n网络更适合用于樱桃成熟度数据集目标检测任务。

**表2　消融实验结果图**

Table 2　Results of ablation experiments

| Model | ADown | C2f-CG | Slim -Neck | *Precision*/% | *Recall* /% | *mAP*50 /% | *Parameter*/M | *Size* /M | *GFLOPs* |
|---|---|---|---|---|---|---|---|---|---|
| YOLOv8n | | | | 90.32 | 89.09 | 94.49 | 3.01 | 5.95 | 8.1 |
| | √ | | | 91.99 | 88.84 | 95.63 | 2.73 | 5.46 | 7.6 |
| | | √ | | 89.07 | 88.10 | 94.32 | 2.10 | 5.12 | 6.7 |
| | | | √ | 91.61 | 88.35 | 95.14 | 2.80 | 5.64 | 7.3 |
| | √ | | √ | 92.46 | 88.44 | 95.35 | 2.67 | 5.11 | 6.8 |
| | √ | √ | √ | 92.94 | 90.66 | 96.57 | 2.05 | 4.25 | 5.4 |

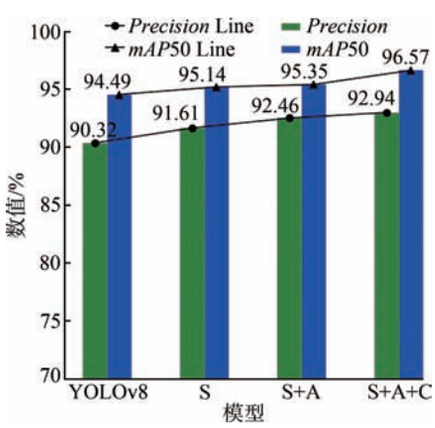



**图10　消融实验的*mAP*和*Precision*结果图**

Figure 10　*mAP* and *Precision* results of ablation experiments

图10逐个说明了每个模块的*mAP*50和精度值的增加。

在图11中，S表示使用GSConv替换Neck部分中的常规卷积，以及使用VoV-GSCSP模块替换Neck部分中的部分常规C2f模块；A表示使用ADown替换Backbone部分中的部分常规卷积；C表示替换主干网络中的C2f模块。

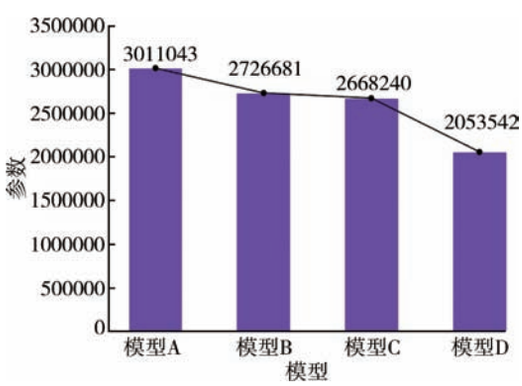



**图11　消融实验的参数量结果图**

Figure 11　Parameter results of ablation experiments

图 11 中,模型 A 是 YOLOv8n 原模型的参数量,模型 B 是加入 Adown 模块的参数量,模型 C 是使用 Adown 和 Slim-Neck 模块的参数量,模型 D 是我们改进之后的 ACG-YOLOv8n 模型的参数量,由图 11 可见,改进后的 ACG-YOLOv8n 模型在所有模型中具有最少的参数量。相比于 YOLOv8n 原模型,参数量减少了 31.90%。

### 3.2 与其他算法对比分析实验

通过比较 YOLOv3-tiny(Darknet-19)、YOLOv5(CSPDarknet53)、YOLOv6(EfficientRep)和 YOLOv8n(Darknet53)等主流模型,验证了本改进模型的检测效果优于其他模型。此外,改进模型的参数量、模型大小和 *GFLOPs* 计算量均明显低于其他模型。选择 YOLOv3-tiny 作为对比模型,是因为 YOLOv8n 本身是轻量级模型,YOLOv3 的原始版本尺寸较大,因此我们选择了其轻量级版本与 YOLOv8n 进行比较。表 3 展示了不同模型的试验结果对比。

**表 3 对比试验结果表**

Table 3 Comparison experiment results

| Model | *Precision*/% | *Recall*/% | *mAP*50/% | *Parameter*/M | *Size*/M | *GFLOPs* |
|---|---|---|---|---|---|---|
| YOLOv3-tiny | 87.02 | 85.45 | 91.98 | 12.13 | 23.2 | 18.9 |
| YOLOv5n | 83.19 | 82.74 | 88.88 | 2.50 | 5.02 | 7.1 |
| YOLOv6 | 85.12 | 81.62 | 89.20 | 4.23 | 8.28 | 11.8 |
| YOLOv8n | 90.32 | 89.09 | 94.49 | 3.01 | 5.95 | 8.1 |
| ACS-YOLOv8n | 92.94 | 90.66 | 96.57 | 2.05 | 4.25 | 5.4 |

试验结果显示,ACS-YOLOv8n 模型在验证集上的 *mAP*50 达到了 96.57%,相比于 YOLOv8n 模型的 94.49% 提高了 2.08 个百分点;同时,ACS-YOLOv8n 的参数量和计算量分别为 YOLOv8n 模型的 68.10% 和 66.67%。相比于 YOLOv3-tiny,ACS-YOLOv8n 在准确率 *Precision* 上提高了 5.92 个百分点,而在参数量和计算量上分别减少了 81.69% 和 71.43%。与 YOLOv5n 和 YOLOv6 等模型相比,ACS-YOLOv8n 在精度上有显著优势,且保持较低的计算量和参数量。通过改进,ACS-YOLOv8n 实现了良好的精度和效率平衡,适合在边缘计算环境中进行部署。图 12 显示了 YOLO 系列模型训练过程的 *mAP*50 值的比较,这表明本改进的模型更有效。

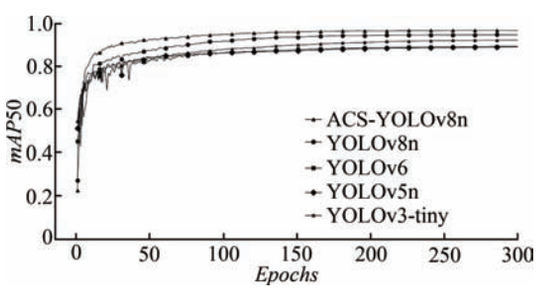



**图 12 训练过程中 *mAP*50 对比图**

Figure 12 Comparison of *mAP*50 during the training process

消融实验与对比实验结果表明,改进模型在检测性能上优于 YOLOv8n 及其他主流模型。为防止过拟合,数据集按 80% 训练、10% 验证、10% 测试的比例进行划分,确保模型在不同数据集上的泛化能力。图 13 显示了 ACS-YOLOv8n 模型与 YOLOv8n 模型在多个测试集上的检测性能比较,表明该模型可以很好地泛化到新数据。

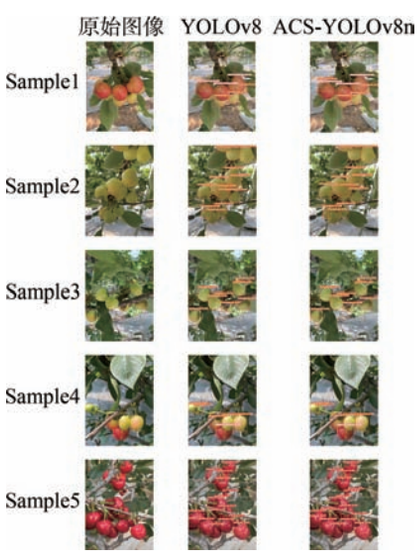



**图 13 不同模型的检测效果图**

Figure 13 Detection results of different models

如表 4 所示,经过 300 轮训练后,改进后的 ACS-YOLOv8n 模型在多个评估指标上均优于 YOLOv8n 模型。具体而言,ACS-YOLOv8n 在整体准确率、召回率、*mAP*50 和 *mAP*50-95 分别提升了 2.62 个百分点、1.63 个百分点、2.08 个百分点和 7.56 个百分点。在成熟类别的检测中,ACS-YOLOv8n 表现出更加显著的改进,其准确率、召回率、*mAP*50 和 *mAP*50-95 分别提高了 3.24 个百分点、3.14 个百分点、3.04 个百分点和 11.19 个百分点。对于半成熟类别,模型在上述 4 项指标上分别提升了 3.30 个百分点、5.85 个百分点、3.53

个百分点和8.73个百分点，同样展现出良好的检测性能。在未成熟类别中，模型的准确率和*mAP*50-95也分别提高了1.32个百分点和2.77个百分点。这些结果充分表明，ACS-YOLOv8n在整体性能以及各成熟度类别上的检测精度均实现了显著提升，验证了所提出改进策略的有效性。

**表4　各类别图像对比试验结果表**

Table4　Comparison of experimental results of each category of images

| Level | Model | *Precision*/% | *Recall*/% | *mA*P50/% | *mAP*50-90/% |
|---|---|---|---|---|---|
| Immature | YOLOv8n | 91.29 | 91.89 | 95.59 | 82.93 |
| | ACS-YOLOv8n | 92.61 | 87.59 | 95.28 | 85.70 |
| Semi-mature | YOLOv8n | 88.78 | 87.05 | 93.40 | 81.33 |
| | ACS-YOLOv8n | 92.08 | 92.90 | 96.93 | 90.06 |
| Mature | YOLOv8n | 90.88 | 88.34 | 94.47 | 77.69 |
| | ACS-YOLOv8n | 94.12 | 91.48 | 97.51 | 88.88 |
| All | YOLOv8n | 90.32 | 89.09 | 94.49 | 80.65 |
| | ACS-YOLOv8n | 92.94 | 90.66 | 96.57 | 88.21 |

## 4　讨论与结论

1)本研究提出的ACS-YOLOv8n模型在樱桃成熟度检测任务中表现出色。通过引入自适应下采样模块(ADown)、上下文引导模块(CGB)和Slim-Neck结构，模型在复杂果园环境下实现了检测精度与推理效率的显著提升。试验结果表明，改进模型在保持轻量化的同时，检测精度达到92.94%，*mAP*50达到96.57%，参数量减少31.90%，模型大小仅为YOLOv8n的71.43%，在各项性能指标上均优于YOLOv3-tiny、YOLOv5n等主流轻量化模型。

2)本研究的主要创新在于模块设计与结构优化。ADown模块通过多路径卷积融合全局与局部信息，缓解小目标检测中的特征损失；CGB模块增强了模型对果实与环境上下文关系的理解能力，提升了遮挡和重叠场景下的检测表现；Slim-Neck结构在降低模型复杂度的同时，保持了多尺度特征的有效融合，为轻量化目标检测提供了可行路径。

3)尽管改进模型在静态图像条件下取得了良好效果，但本研究的数据主要来源于单一区域，模型的泛化能力尚需进一步验证。同时，当前实验仍基于静态图像，未在动态采摘或视频流场景中评估模型的实时性与鲁棒性，存在一定的局限性。

4)未来工作可从多个方向展开：一是采集多地区、多品种樱桃图像数据，以增强模型的跨环境适应能力；二是引入多光谱或其他辅助信息提升特征表达；三是结合迁移学习进一步优化成熟度分级；四是探索模型在移动采摘平台中的实际部署与应用可行性，推动樱桃智能采摘系统的发展与落地。

## 参　考　文　献


[1] 伊贤贵，李蒙，王贤荣．樱亚属分类学研究进展[J]．南京林业大学学报(自然科学版)，2024，48(3)：46－57.
YI Xiangui，LI Meng，WANG Xianrong. A review on the taxonomy study of *Prunus* subgen. *Cerasus* (Mill) A. Gray [J]. Journal of Nanjing Forestry University (Natural Science Edition)，2024，48(3)：46－57.

[2] 付陈龙，李蒙，田昌芬，等．四种气候变化情景下中国樱桃的潜在适生区预测[J]．南京林业大学学报(自然科学版)，2024，48(4)：235－242.
FU Chenlong，LI Meng，TIAN Changfen，et al. Prediction of potential suitable regions of *Prunus pseudocerasus* based on MaxEnt model under climate change [J]. Journal of Nanjing Forestry University (Natural Science Edition)，2024，48(4)：235－242.

[3] 董渭雪．樱桃黄酮组分及降尿酸作用研究[D]．汉中：陕西理工大学，2020.
DONG Weixue. Study on the flavonoids in cherries and the effect of reducing uric acid [D]. Hanzhong：Shaanxi University of Technology，2020.

[4] ERTEK TOSUN N，UZUNDUMLU A S. Cherry production forecasting of leading countries in 2024－2028 [J]. Applied Fruit Science，2025，67(2)：67.

[5] 麦昌青，梁秋南．浅析三亚台农芒果可溶性固形物含量Brix值测定[J]．农业科学，2018，8(5)：486－495.
MAI Changqing，LIANG Qiunan. A brief analysis of brix value determination of the soluble solids tainnong mango in Sanya [J]. Agricultural Science，2018，8(5)：486－495.

[6] PALUMBO M，CEFOLA M，PACE B，et al. Computer vision system based on conventional imaging for non-destructively evaluating quality attributes in fresh and

packaged fruit and vegetables [J]. Postharvest Biology and Technology, 2023, 200: 112332.

[7] CHU X, MIAO P, ZHANG K, et al. Green banana maturity classification and quality evaluation using hyperspectral imaging [J]. Agriculture, 2022, 12(4): 530.

[8] WENDEL A, UNDERWOOD J, WALSH K. Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform [J]. Computers and Electronics in Agriculture, 2018, 155: 298−313.

[9] MUNERA S, BESADA C, BLASCO J, et al. Astringency assessment of persimmon by hyperspectral imaging [J]. Postharvest Biology and Technology, 2017, 125: 35−41.

[10] 薛建新, 张淑娟, 张晶晶. 基于高光谱成像技术的沙金杏成熟度判别[J]. 农业工程学报, 2015, 31(11): 300−307.
XUE Jianxin, ZHANG Shujuan, ZHANG Jingjing. Ripeness classification of Shajin apricot using hyperspectral imaging technique [J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(11): 300−307.

[11] 苏斐, 张泽旭, 赵妍平, 等. 基于轻量化YOLO-v3的绿熟期番茄检测方法[J]. 中国农机化学报, 2022, 43(3): 132−137.
SU Fei, ZHANG Zexu, ZHAO Yanping, et al. Detection of mature green tomato based on lightweight YOLO-v3 [J]. Journal of Chinese Agricultural Mechanization, 2022, 43(3): 132−137.

[12] 张志远, 罗铭毅, 郭树欣, 等. 基于改进YOLO v5的自然环境下樱桃果实识别方法[J]. 农业机械学报, 2022, 53(S1): 232−240.
ZHANG Zhiyuan, LUO Mingyi, GUO Shuxin, et al. Cherry fruit detection method in natural scene based on improved YOLO v5 [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(S1): 232−240.

[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.

[14] WEI Y, YAO S, WU F, et al. Maturity classification and quality determination of cherry using VNIR hyperspectral images and comprehensive chemometrics [J]. International Journal of Fruit Science, 2024, 24(1): 32−48.

[15] ZHANG C, BENGIO S, HARDT M, et al. Understanding deep learning (still) requires rethinking generalization [J]. Communications of the ACM, 2021, 64(3): 107−115.

[16] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]// Proceedings of the IEEE Conference on Computer Vision And Pattern Recognition. 2017: 7263−7271.

[17] HUSSAIN M. YOLOv5, YOLOv8 and YOLOv10: The go-to detectors for real-time vision [J]. arXiv preprint arXiv: 2407.02988, 2024.

[18] LOU H, DUAN X, GUO J, et al. DC-YOLOv8: small-size object detection algorithm based on camera sensor [J]. Electronics, 2023, 12(10): 2323.

[19] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 2117−2125.

[20] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8759−8768.

[21] CASAS E, RAMOS L, ROMERO C, et al. A comparative study of YOLOv5 and YOLOv8 for corrosion segmentation tasks in metal surfaces [J]. Array, 2024: 100351.

[22] 王小桥, 黄友锐. 一种轻量化的甜菜与杂草实例分割方法[J/OL]. 重庆工商大学学报(自然科学版): 1−10[2025−05−10]. https://link.cnki.net/urlid/50.1155.n.20241015.1118.002.
WANG Xiaoqiao, HUANG Yourui. A lightweight instance segmentation method for sugar beets and weeds [J]. Journal of Chongqing Technology and Business University (Natural Science Edition): 1−10[2025−05−10]. https://link.cnki.net/urlid/50.1155.n.20241015.1118.002.

[23] 桂余鹏,胡蓉华,崔艳荣,等. 基于轻量化YOLO v8-Rice的水稻虫害检测方法[J]. 江苏农业科学, 2024, 52(20): 277−284.
GUI Yupeng, HU Ronghua, CUI Yanrong, et al. A rice pest detection method based on lightweight YOLO v8-Rice [J]. Jiangsu Agricultural Sciences, 2024, 52(20): 277−284.

[24] 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.

[25] 张哲鼎, 王超柱, 陈骏, 等. 基于YOLOX和DeepSORT的淡水鱼行为检测算法研究[J]. 智能化农业装备学报(中英文), 2025, 6(3): 88−97.
ZHANG Zheding, WANG Chaozhu, CHEN Jun, et al. Research on freshwater fish behavior detection algorithm based on YOLOX and DeepSORT [J]. Journal of

Intelligent Agricultural Mechanization, 2025, 6(3): 88－97.

[26] 陈家政，付根平，黄伟锋，等. 基于改进YOLOv5s的云机器人视觉导航方法[J]. 智能化农业装备学报(中英文)，2025，6(3)：98－110.
CHEN Jiazheng, FU Genping, HUANG Weifeng, et al. Visual navigation method based on improved YOLOv5s and cloud robot [J]. Journal of Intelligent Agricultural Mechanization, 2025, 6(3): 98－110.

[27] 吴擎，韦润轩，周乐，等. 基于改进YOLOv5s的轻量化鲜茶叶识别方法[J]. 智能化农业装备学报(中英文)，2025，6(1)：1－14.
WU Qing, WEI Runxuan, ZHOU Le, et al. Lightweight fresh tea leaf recognition method based on improved YOLOv5s [J]. Journal of Intelligent Agricultural Mechanization, 2025, (1): 1－14.

[28] 刘畅，孙雨，杨晶，等. 基于3C-YOLOv8n和深度相机的葡萄识别与定位方法[J]. 智慧农业(中英文)，2024，6(6)：121－131.
LIU Chang, SUN Yu, YANG Jing, et al. Grape recognition and localization method based on 3C-YOLOv8n and depth camera [J]. Smart Agriculture, 2024, 6(6): 121－131.

# Research on cherry ripeness detection based on improved YOLOv8n

LI Ruixin[1], CAO Yue[2], SI Xiuli[1, 3], MA Li[2*]

(1. *School of Information Technology, Jilin Agricultural University, Changchun* 130118, *China*;
2. *Wuxi University, Wuxi* 214105, *China*;
3. *Jilin Agricultural Artificial Intelligence Research Base, Changchun* 130118, *China*)

**Abstract:** Accurate detection of cherry fruits is crucial for improving fruit quality, optimizing harvesting timing, and automating orchard management. However, due to factors such as small fruit size, diverse colors, complex backgrounds, and overlapping occlusions, detection still faced many challenges. To this end, this paper proposed a lightweight and improved model, ACS-YOLOv8n, based on the YOLOv8n architecture to enhance detection accuracy and efficiency. The model introduced an adaptive downsampling module (ADown) in the backbone network, which integrates the advantages of average pooling and max pooling, and combines multi-path convolution operations to effectively retain both global information and local salient features, thereby enhancing the detection capability of small targets. Meanwhile, the Context Guided Block (CGB) was used to replace the traditional C2f module, improving the model's ability to perceive the relationship between the target and the environment and enhancing the effect of multi-scale feature extraction through the joint extraction of local, contextual, and global features. In the neck part, the Slim-Neck structure based on GSConv and VoV-GSCSP was introduced to optimize the feature fusion method, significantly reduce computational redundancy, and balance speed and accuracy. The experimental results showed that ACS-YOLOv8n achieved 92.94% precision, 90.66% recall, and 96.57% mAP50 in complex orchard environments. Compared with YOLOv8n, the precision increased by 2.62 precentage point, the recall by 1.57 precentage point, and the average precision by 2.08 precentage point, while the number of model parameters and the computational volume were reduced by 31.90% and 33.33%, respectively, and the model size was 71.43% of the original model. The proposed model demonstrated good detection performance and robustness in complex environments, providing effective technical support for automatic cherry harvesting and intelligent orchard management.

**Keywords:** cherry; ripeness detection; YOLOv8n; multi-scale feature extraction; object detection; deep learning

**Foundation items:** National Natural Science Foundation of China(U19A2061); Wuxi University Research Start-up Project (2024r007)

**First Author:** LI Ruixin, E-mail: liruixin0116@163.com

*** Correspondence Author:** MA Li, E-mail: mali@cwxu.edu.cn