doi:10.6041/j.issn.1000-1298.2026.04.029

基于 MSCS－YOLO 的非结构化环境中草莓成熟度识别

王永胜[1]　丁　宇[1]　张若晨[2,3]　许洪光[1]　范　玥[4]

（1.河北东方学院人工智能学院，廊坊 065000；2.北京市农林科学院智能装备技术研究中心，北京 100097；
3.国家农业智能装备工程技术研究中心，北京 100097；4.河北东方学院经济管理学院，廊坊 065000）

摘要：针对草莓个体较小、个体间遮挡严重的问题，提出了一种基于 MSCS－YOLO 的非结构化环境中草莓成熟度识别方法。在 YOLO v8n 模型的 Neck 部分引入多尺度扩展注意力机制（Multi-scale dilated attention，MSDA），扩大模型的感受野，解决草莓果实较小、特征易被忽略的问题。同时，利用改进的 C2f－Triplet attention 结构替换 Neck 部分的 C2f 结构，从 3 个维度更全面地捕捉草莓图像的信息，增强模型在果实遮挡情况下的目标识别能力。将改进的 SAHead 检测头嵌入到 YOLO v8n 模型，提升模型在非结构化环境中对不同成熟度草莓的识别精度。试验结果表明，MSCS－YOLO 模型在成熟、中等成熟及未成熟 3 类草莓识别的任务中，平均精度均值达到 94.22%，较 YOLO v8n 和 RTDETR－L 模型分别提高 1.38、5.42 个百分点；对成熟和中等成熟草莓识别平均精度分别达到 96.35% 和 92.00%，较 YOLO v8n 模型分别高 0.82、3.66 个百分点。MSCS－YOLO 模型在各种光照条件下（包括夜晚、晴天、直射阳光和光线照射）都展现了更佳的识别表现和更高的准确性。此外，改进模型内存占用量为 6.42 MB，相较于 YOLO v7－tiny 和 YOLO v9c 模型分别减少 45.22% 和 86.93%，在保持与 YOLO v8n 相近模型体积的同时，实现了精度与效率的协同优化。因此，MSCS－YOLO 模型在资源有限的环境中更具部署和应用的优势，为后期对草莓成熟度的实际应用提供了可靠的技术支持。

关键词：非结构化环境；草莓成熟度识别；SAHead；YOLO v8n

中图分类号：S641.2　　**文献标识码**：A　　**文章编号**：1000-1298(2026)04-0296-13　　**OSID**：

Strawberry Maturity Recognition in Unstructured Environments Based on MSCS－YOLO

WANG Yongsheng[1]　DING Yu[1]　ZHANG Ruochen[2,3]　XU Hongguang[1]　FAN Yue[4]

（1. *Artificial Intelligence College*, *Hebei Oriental College*, *Langfang* 065000, *China*
2. *Research Center for Intelligent Equipment Technology*, *Beijing Academy of Agriculture and Forestry Sciences*, *Beijing* 100097, *China*
3. *National Engineering Research Center for Intelligent Equipment in Agriculture*, *Beijing* 100097, *China*
4. *Economics and Management College*, *Hebei Oriental College*, *Langfang* 065000, *China*）

Abstract: Aiming to address the problems of small strawberry individuals and serious inter-individual occlusion, a strawberry ripeness detection method was proposed based on MSCS－YOLO in an unstructured environment. The method introduced the multi-scale dilated attention (MSDA) mechanism in the Neck part of the YOLO v8n model, which enlarged the sensory field of the model and solved the problem of small strawberry fruits and easy to ignore features. Meanwhile, the improved C2f－Triplet attention structure was utilized to replace the C2f structure in the Neck part to capture the information of the strawberry image more comprehensively from the three dimensions, which enhanced the model's target recognition ability in the case of fruit occlusion. Embedding the improved SAHead detection head into the YOLO v8n model enhanced the model's recognition accuracy for strawberries with different ripeness levels in unstructured environments. The experimental results showed that the MSCS－YOLO model achieved an average accuracy of 94.22% in the task of recognizing three types of strawberries: ripe, moderately ripe

收稿日期：2025－08－20　修回日期：2025－09－28

基金项目：国家重点研发计划项目（2024YFD2001005－04）、河北省级一流本科专业-物联网工程项目（冀教高函[2024]27 号）、河北省级一流本科立项建设课程-嵌入式系统与设计项目（冀教高函[2020]25 号）和河北东方学院校级科研项目（XJYB2025069）

作者简介：王永胜（1998—），男，助教，主要从事机器视觉技术研究，E-mail：878074375@qq.com

通信作者：许洪光（1986—），男，副教授，主要从事软件工程和人工智能研究，E-mail：690133307@qq.com

and unripe, which was 1.38 percentage points and 5.42 percentage points higher than that of the YOLO v8n and RTDETR－L models, respectively; among them, the accuracy of recognizing ripe and moderately ripe strawberries achieved 96.35% and 92.00%, which was 0.82 percentage points and 3.66 percentage points higher than that of the YOLO v8n model, respectively. The MSCS－YOLO model demonstrated better recognition performance and higher accuracy regardless of evening, sunny day, direct sunlight or light irradiation conditions. In addition, the model size of the improved model was 6.42 MB, which was 45.22% and 86.93% smaller than that of the YOLO v7－tiny and YOLO v9c models, respectively, and achieved the synergistic optimization of accuracy and efficiency while maintaining a similar model size with YOLO v8n. Therefore, the MSCS－YOLO model was more advantageous for deployment and application in resource-limited environments, and it can provide reliable technical support for later practical applications on strawberry maturity.

Key words: unstructured environments; strawberry maturity recognition; SAHead; YOLO v8n

0　引言

草莓作为一种兼具高经济价值与丰富营养的特色浆果，其产业规模正不断扩大，对智能化采摘装备的需求日益迫切。根据国家统计局数据显示，截至 2024 年 9 月，我国草莓种植面积已达 9.2×10^4 hm^2。在非结构化环境中，由于草莓果实体积较小、易受树枝和叶片遮挡，以及易受光照变化的影响，传统的人工采摘方式存在效率低、劳动强度大、易产生主观误判、果实损伤率高等问题，难以满足产业现代农业的生产需求，严重制约了采摘自动化的发展[1－2]。

在传统机器视觉领域，研究者们主要依赖人工特征提取与分类器构建实现果实检测。例如，谢志勇等[3]通过图像分割对其进行 Hough 变换，实现了成熟草莓的识别。侯贵洋等[4]采用颜色特征与形态特征相结合的方法，实现了对不同成熟度草莓分类。ZHANG 等[5]通过分析二维颜色直方图，并利用颜色映射与反投影技术，对草莓果实的成熟度和品质进行了分级。WANG 等[6]提出基于几何中心匹配的检测方法，并结合像素阈值技术对荔枝果实进行分类，平均识别率达到 92.00% ~94.17%。尽管这些方法在一定程度上提升了目标识别精度，但仍高度依赖人工图像分割与手动特征提取，识别效果易受操作人员经验影响，难以满足大规模、复杂场景下自动化需求。

近年来，深度学习技术快速发展为农业检测领域带来了革命性的变革。针对草莓果实体积小、种植环境复杂等挑战，研究者提出了多种基于深度学习的优化模型。WANG 等[7]提出 DSE－YOLO 模型，通过双维度特征增强策略，将草莓检测精度提升至 86.58%。DU 等[8]基于 YOLO v7 构建的 DCNv3－SA－WIoU－YOLO 模型，在遮挡场景下对草莓检测精度达到 86.7%。而 YANG 等[9]融合 LW－Swin Transformer 模块与 YOLO v8s，构建了 LS－YOLO v8s 模型，在草莓成熟度分级任务上达到 94.4% 的检测精度。同时，LIU 等[10]基于任务对齐机制开发的草莓检测算法，在密集小目标场景下平均精度达 93.9%。这些研究表明，深度学习方法能够通过自动特征学习与上下文信息捕捉，有效提升复杂农业场景下的目标检测性能。

尽管现有方法在草莓检测任务中取得了较高精度，但由于草莓个体小、果实密集以及叶片遮挡等因素，小目标和遮挡仍然是影响检测性能的主要挑战。针对小目标检测，研究者提出了多尺度特征融合、膨胀卷积等方法，以增强模型对小尺寸目标的感知能力[11－12]。针对遮挡目标，研究者常利用注意力机制，减轻背景干扰，增强对目标区域的关注[13－14]，提高模型在遮挡情况下的识别精度。

基于上述研究进展与现有技术挑战，本文提出一种基于 MSCS－YOLO 的草莓成熟度检测模型，以解决非结构化环境下草莓果实体积小、特征易被忽略的问题。旨在增强模型对遮挡场景中目标识别能力，提升不同成熟度草莓检测精度，为智能采摘系统构建提供技术支持。

1　试验材料

1.1　数据采集

为了确保数据集能够充分反映非结构化环境下草莓生长特征，在江苏省某草莓种植园采集了不同环境下的草莓图像。采集设备为 iPhone 15 Pro Max，光圈值为 F2.2，焦距为 2 mm。数据采集时间为 2023 年 12 月 12—17 日，共采集 1 645 幅 RGB 图像，涵盖傍晚、晴天、阳光照射和灯光照射 4 种光照条件。

傍晚图像 404 幅，晴天 468 幅，阳光照射 492 幅，灯光照射 281 幅，所有图像分辨率均为 4 032 像素 ×3 024 像素，图 1 为 4 种不同环境下草莓图像。从图 1 可知，草莓果实存在较为严重的遮挡现象，由图 1d 可知，由于灯光直射，草莓轮廓较为模糊，特征不够明显，这些因素均会增加草莓识别的难度。

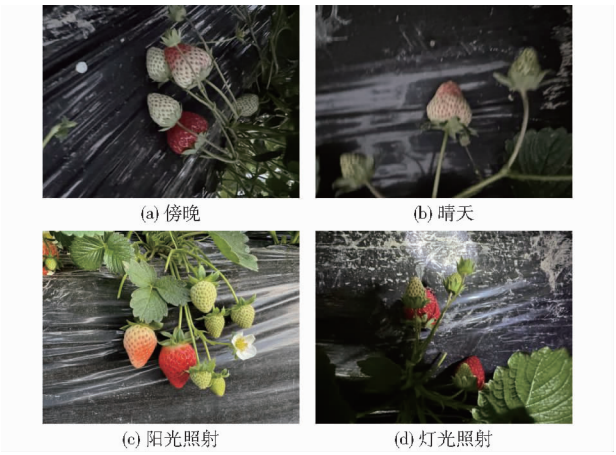

图 1 不同环境下草莓图像

Fig. 1 Images of strawberry in different environments

1.2 数据集制作

在模型训练前,采用 Labellmg 软件对所采集草莓图像进行人工标注。文献[12]基于果皮着色面积与市场需求的成熟度划分标准,将草莓成熟度分为 3 个等级:成熟(果皮基本全部变红)、中等成熟(果皮红绿相间)和未成熟(果皮主要为绿色)。标注过程中每幅图像均生成对应的 XML 标注文件,并统一转换为 Pascal VOC 格式。为提高模型泛化能力[15-17],将数据集按比例 7∶2∶1 划分为训练集、验证集与测试集,分别为 1 151 幅、329 幅和 165 幅。

为明确草莓目标的尺度分布特征,依据相对尺度定义标准[18-19],对数据集中草莓实例进行了划分。依据标注边界框的像素面积占图像像素总面积的比重,将草莓实例划分为 3 类:小目标(比重 <10%)、中目标(10% ≤比重 <25%)和大目标(比重≥25%)。经统计,本数据集中共包含草莓实例 12 410 个,其中小目标 12 388 个,中目标 22 个,大目标 0 个。各类别大、中、小目标实例统计结果如表 1 所示。

表 1 各类别大中小目标实例数量

Tab. 1 Number of large, medium and small target instances in each category

类别	小目标数量	中目标数量	大目标数量
成熟	2 516	15	0
中等成熟	1 729	5	0
未成熟	8 143	2	0

由表 1 可知,小目标占比较高,表明草莓识别任务面临显著的尺度多样性挑战,尤其需要模型具备强大的小目标检测能力。

2 草莓成熟度识别方法

2.1 YOLO v8 模型结构

YOLO v8 模型是目前主流的目标检测算法,模型兼顾了推理速度和检测精度[20-22],模型包括输入、骨干、Neck 和输出部分。

YOLO v8 的输入部分处理预处理后的草莓图像并将其送入模型。骨干网络部分借鉴了 CSP 模块的思想,将 YOLO v5 中 C3 模块替换为 C2f 模块。其中,C2f 模块的首个卷积层采用 3 × 3 卷积核,实现模型进一步轻量化;同时,该模块引入更多跳层连接和额外的 Split 操作,有效缓解因深度增加而导致的梯度消失问题,从而可显著提升特征提取能力。

在 Neck 部分,YOLO v8 延续了 PAN[23](Path aggregation network)的设计思路,通过融合不同尺度的特征信息,进一步增强特征表达能力。

在输出端采用解耦头结构(Decoupled - Head),将分类和检测任务分离,使得每个任务能够独立优化,从而减少相互干扰,提升目标分类和定位的准确性。此外,YOLO v8 由 Anchor - Based 方法转变为 Anchor - Free 方法,可摆脱对预定义锚点的依赖,使得模型能够更灵活地适应各种尺寸和形状的目标,提高模型的鲁棒性。

2.2 模型选择

YOLO v8 模型包括 5 个版本[24],这些版本在网络结构的深度和宽度上存在着明显差异。图 2 为各版本模型参数量(Parmeters)和浮点运算量(GFLOPS),较为复杂的模型(除 YOLO v8n 外)拥有更多参数,从而导致更高的计算成本与内存占用量,在实际部署时面临资源瓶颈。相比之下,YOLO v8n 参数量最少、运算量最低,更加适合资源有限的应用场景。因此,为了满足本文实际需求,选择在 YOLO v8n 模型的基础上进行改进。

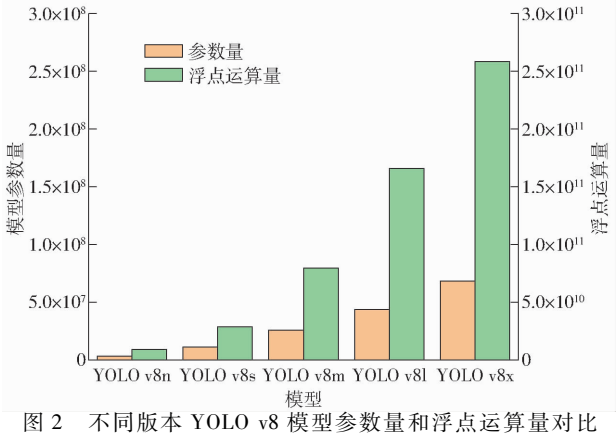

图 2 不同版本 YOLO v8 模型参数量和浮点运算量对比

Fig. 2 Comparison of parameter quantity and calculation amount of different YOLO v8 versions

2.3 模型改进

2.3.1 多尺度扩展注意力模块

YOLO v8n 模型 Neck 部分融合特征来自 Backbone 中不同尺度的 3 个特征图,当输入尺寸为

640 像素 ×640 像素时，这些特征图的尺寸分别为 20 像素 ×20 像素、40 像素 ×40 像素和 80 像素 ×80 像素，分别融合小、中、大目标特征信息，且其感受野随着特征图分辨率降低而逐渐增大。在对不同成熟度草莓特征进行融合时，如果草莓尺度较大，感受野最大的特征图可能并不能提取到有效的高级语义信息，而当草莓尺度过小时，感受野最小的特征图也可能无法保留完整的轮廓与位置信息。

在实际种植环境中，草莓果实通常尺寸较小，有效特征也较少。为了获取更多尺寸下的草莓特征并增强特征融合能力，在 YOLO v8n 的 Neck 部分进行改进，在 3 幅不同尺寸特征图后加入多尺度空洞注意力[25]（Multi-scale dilated attention，MSDA）模块，进一步增大模型感受野，MSDA 工作原理如图 3 所示。

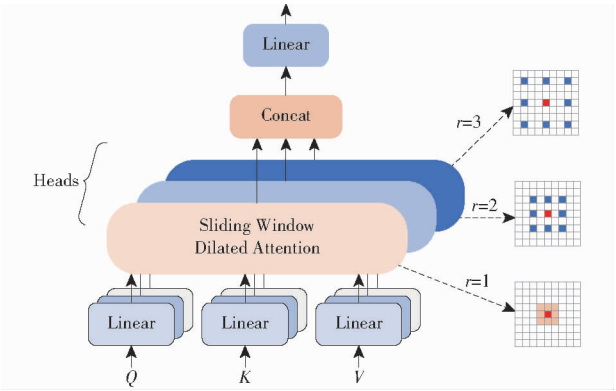

图 3　MSDA 工作原理

Fig. 3　Working principle of MSDA

如图 3 所示，MSDA 模块首先通过线性投影将输入特征图 X 转换为查询、键和值。然后，将特征图的通道分成 n 个不同的部分，在每部分上分别应用不同扩张率的多尺度滑动窗口扩张注意力操作（Sliding window dilated attention，SWDA），这些操作在以红色查询块为中心的窗口内，对周围的彩色块进行信息交互。

图 3 中 3 种扩张率（$r=1,2,3$）分别对应 3 ×3、5 ×5、7 ×7 的感受野，每部分自注意力操作均与其特定扩张率和感受野相匹配，从而使模型能够在不同的尺度下捕捉图像特征。然后，各部分产生的特征被融合，输入到一个线性层进行特征聚合。引入 MSDA 模块后，模型不仅能够提取草莓的局部细节，还能捕捉更广泛的上下文信息，从而提升模型对草莓的整体特征学习能力，解决草莓目标较小且特征易被忽略的问题。

2.3.2　C2f－Triplet attention 改进

在实际种植环境中，草莓叶片与果实之间的遮挡常常导致目标特征部分丢失，从而影响检测精度。为了解决这一问题，引入三重注意力机制[26]（Triplet attention），以更全面地捕捉草莓图像中的信息，从而有效提升在小果实遮挡情况下的识别能力。

如图 4 所示，三重注意力共包括 3 个分支。上分支主要计算通道维度 C 和空间维度 W 的注意力权重。该分支对输入张量执行池化操作，以提取全局特征信息，然后通过一个卷积层进行特征提取，并利用 Sigmoid 函数生成注意力权重。中分支进一步扩展特征捕获能力，负责提取通道维度 C 与空间维度 H 和 W 之间的相互依赖关系。其计算过程与上部分支相同，既先进行 Z 池化和卷积操作，再通过 Sigmoid 函数生成注意力权重。下分支主要用于捕获空间维度之间的依赖性，保持输入的身份，但仍执行 Z 池化、卷积操作和 Sigmoid 激活，计算空间维度的注意力权重。最终，3 个分支的输出经过排列变换后，通过平均聚合操作进行融合，形成最终的三重注意力输出。这种结构通过不同维度的信息交互，能够更全面地捕捉草莓图像的特征，提高模型对复杂环境下草莓目标的感知能力。

在 YOLO v8n 结构中，C2f 模块作为一个残差块，通过残差连接优化网络的训练过程，使模型能够更有效地学习输入与输出之间的直接映射关系；其结构相对简单、计算资源消耗较低，在计算资源受限的情况下能有效降低模型复杂度并提高运行效率。

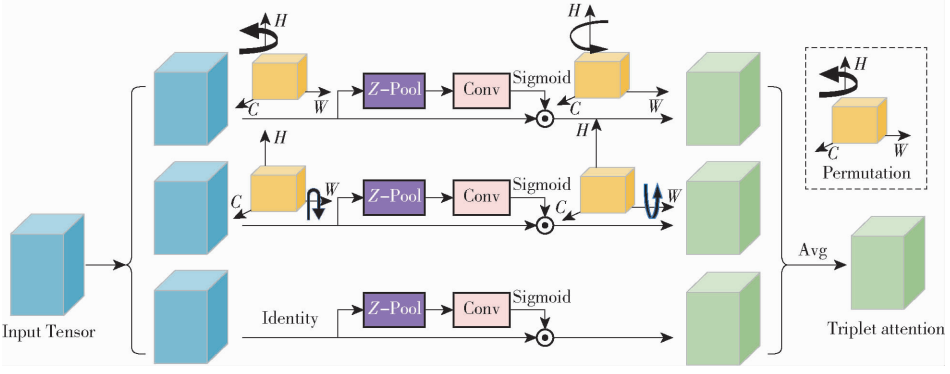

图 4　三重注意力实现流程图

Fig. 4　Flowcharts for specific realization of triple attention

然而,该结构在特征学习方面存在明显不足,尤其难以充分捕获草莓目标的多尺度特征和遮挡情况下的细节信息。为此,在 YOLO v8n 的 Neck 部分对 C2f 结构进行改进,采用三重注意力机制替换其中的 Bottleneck 模块,构建 C2f – Triplet attention 模块,结构如图 5 所示。

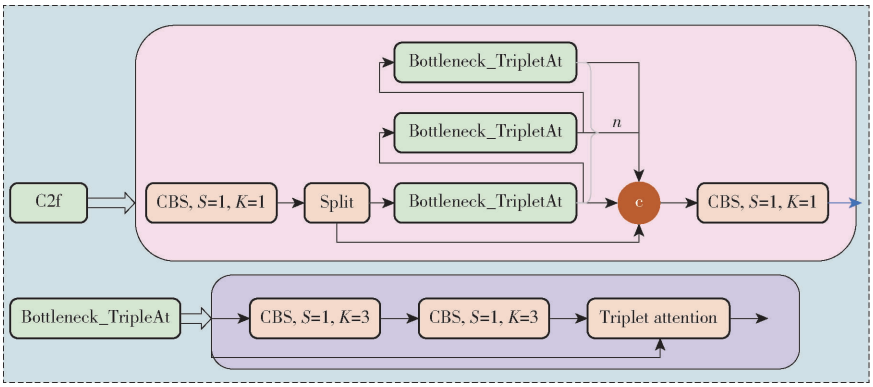

图 5 C2f – Triplet attention 结构

Fig. 5 C2f – Triplet attention structure

模型在学习草莓特征后,经过 YOLO v8n 结构中 Neck 部分时,C2f – Triplet attention 模块能够进一步从通道、高度和宽度 3 个维度深度挖掘草莓目标的特征信息。这种多维度特征学习方式使得模型在面对草莓叶片与果实之间的相互遮挡时,仍能有效提取关键特征,避免因遮挡导致的特征丢失问题,从而提升草莓目标的检测精度和鲁棒性。

2.3.3 自注意力检测头

YOLO v8n 模型的检测头负责将特征提取网络(Backbone)和特征融合网络(Neck)所生成的多尺度特征图转换为具体的检测结果。检测结果包括目标类别、目标位置(边界框)和目标置信度 3 部分,其结构如图 6 所示。

作为目前最先进的目标检测算法之一,YOLO v8n 模型的检测头采用了解耦结构,将检测任务拆分为 Cls 分类和 Box 回归 2 个独立分支,并去除 Obj loss 计算分支,在检测性能和精度方面均取得了显著提升。但是,由于草莓通常生长在非结构化环境

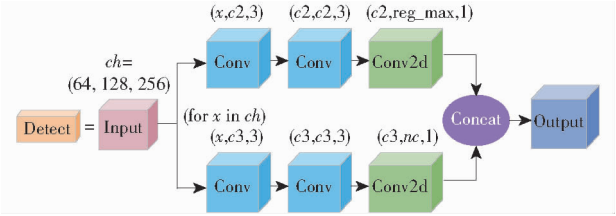

图 6 YOLO v8 模型检测头

Fig. 6 YOLO v8 model detection head

中,易受光照条件(如灯光、阳光照射)的影响,检测过程存在特征丢失、误检漏检等问题,影响检测的稳定性和准确性。

为提升模型在非结构化环境下的鲁棒性,在 YOLO v8n 模型检测头的基础上进行改进,提出了一种新型检测头(SAHead),其结构如图 7 所示。

在 YOLO v8n 的检测头中引入自注意力模块[27](SAttention),构建了 SAHead 模块,通过计算特征间的全局相关性,使检测头能够有效捕捉长距离依赖关系,并动态调整特征重要性。这不仅增强了模型的特征表达能力,使其能更全面地提取草莓

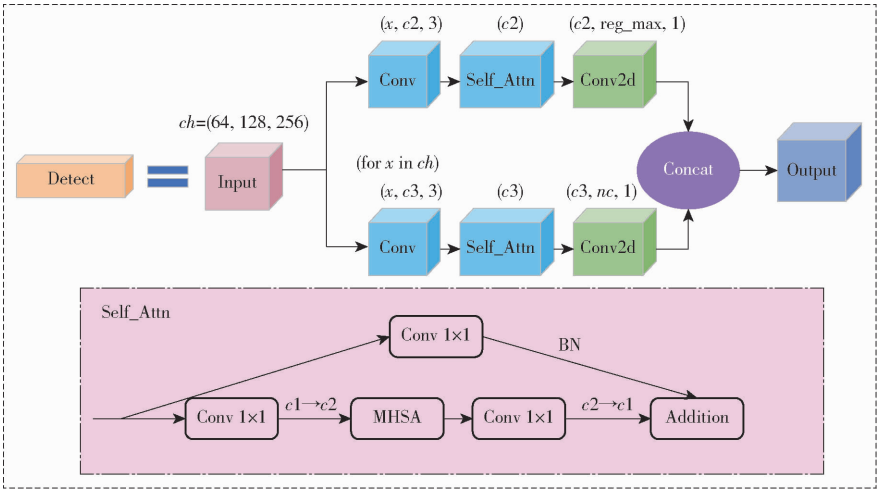

图 7 SAHead 结构图

Fig. 7 SAHead structure

目标的全局特征，还可有效缓解因光照变化导致的特征丢失问题，从而提升在非结构化环境下对草莓成熟度的检测精度与鲁棒性。

2.4　MSCS－YOLO 模型

基于 YOLO v8n 构建 MSCS－YOLO 模型，进行了 3 项关键改进，以提升草莓成熟度在非结构化环境下的检测精度。

首先，对 Neck 部分进行了优化，在 3 个不同尺寸的特征图后引入多尺度空洞注意力，扩大模型的感受野，增强模型对草莓有效特征的融合能力，以解决草莓目标较小且特征易被忽略的问题。其次，将 Neck 部分的 C2f 结构中 Bottleneck 模块改进为 C2f－Triplet attention 模块，以解决草莓叶片以及果实间导致草莓出现遮挡，致使特征丢失的问题。最后，利用提出的 SAHead 检测头作为 MSCS－YOLO 的检测头，以避免灯光、阳光照射的影响，提升模型在非结构化环境下对草莓成熟度识别能力。MSCS－YOLO 模型结构图如图 8 所示。

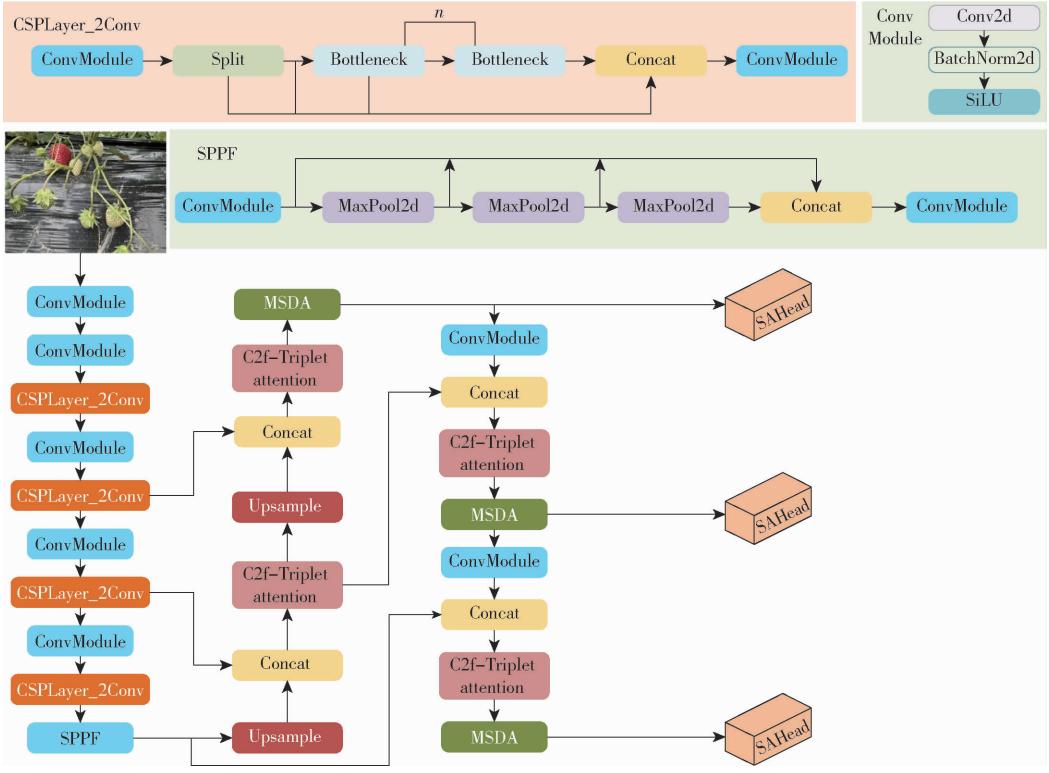

图 8　MSCS－YOLO 模型结构

Fig. 8　MSCS－YOLO model structure

2.5　模型训练与评估

2.5.1　试验环境

所有试验均在表 2 描述的试验设备及试验环境下进行训练。试验超参数设置为输入图像尺寸 640 像素 × 640 像素，初始学习率（Initial learning rate）0.01，动量因子（Momentum）0.937，迭代次数（Epoch）300，线程数（Num workers）8，训练批次大小（Batch size）8，优化器设置为 SGD。在模型评估过程，使用训练得到的最优权重在测试集上进行推理验证。在预测过程，非极大值抑制（NMS）的 IoU 阈值设置为 0.7，并保留置信度分数高于 0.25 的预测框，作为最终性能评估的依据。

2.5.2　评价指标

利用精确率（Precision，P）、召回率（Recall，R）、平均精度（Average precision，AP）、F1 值、平均精度均值（mean Average Precision，mAP）、模型内存占用量（Model size）、检测速度（Frames per second，FPS）对模型性能进行评估。

表 2　硬件配置和试验环境

Tab. 2　Hardware configuration and experimental environment

硬件	配置	环境	版本
操作系统	Windows10	Python	3.7.9
CPU	Intel Core i5－12600KF	PyTorch	1.13
GPU	Nvidia GeForce RTX3060	CUDA	11.6
RAM	16GB	CUDNN	8.4

3　试验结果与分析

3.1　消融试验

为了验证改进方法对草莓成熟度识别的有效性，在本文草莓成熟度识别数据集上分别对加入不同改进项的模型进行测试，试验结果如表 3 所示。

表 3 消融试验结果

Tab. 3 Ablation test results %

模型	改进模块			精确率	召回率	mAP
	MSDA	C2f - Triplet attention	SAHead			
A	×	×	×	90.14	87.16	92.84
B	√	×	×	87.13	89.70	93.27
C	×	√	×	89.30	88.51	93.62
D	×	×	√	89.42	87.55	93.55
E	√	√	×	89.74	88.67	93.50
F	√	√	√	90.89	87.57	94.22

注:√表示使用此模块,×表示未使用此模块。

由表 3 可以看出,在单独引入 MSDA 改进项时,尽管精确率由 90.14% 降至 87.13%,但召回率从 87.16% 提升至 89.70%,平均精度均值由 92.84% 提高至 93.27%。表明此改进模块能够扩大模型感受野,增强对小尺度草莓目标的特征提取能力,从而提升模型的召回性能。

单独引用 C2f - Triplet attention 改进模块时,模型精确率、召回率和平均精度均值均获得提升,其中平均精度均值达到 93.62%。表明此改进模块能够增强模型在果实遮挡场景下的判别能力,提升复杂情况下草莓识别的召回率与整体识别精度。

单独使用 SAHead 检测头改进模块时,模型平均精度均值提升至 93.55%,在复杂光照条件下展现出良好的适应性。表明此改进模块能够增强模型对光照变化(如强光、阴影及不均匀光照)的鲁棒性,提高草莓成熟度识别的准确性。

综合而言,3 种改进模块分别针对小目标漏检、目标间遮挡及复杂光照条件等问题提供了有效解决方案。当 3 种改进模块协同融合时,模型在精确率(90.89%)与平均精度均值(94.22%)均达到最高值,表明其在不同挑战性场景下具备更强的综合检测能力与稳定性,能够提升在非结构化环境中草莓成熟度识别的整体性能。

3.2 模型对比试验

3.2.1 不同模型识别结果

为了检验本文模型(MSCS - YOLO)在草莓成熟度识别方面的性能,在本文草莓成熟度识别数据集上,对比分析 MSCS - YOLO 与当前主流目标检测模型(YOLO v5n、YOLO v6n[28]、YOLO v7 - tiny[29]、YOLO v8n、YOLO v9c[30]、YOLO v10n[31]、RTDETR - L[32])的性能表现,同时纳入文献[12]小目标检测模型及文献[14]抗遮挡模型作为对比。不同模型试验结果如表 4 所示,不同模型对不同成熟度识别结果如表 5 所示。

表 4 不同模型试验结果

Tab. 4 Different model test results

模型	精确率/%	召回率/%	F1 值/%	mAP/%	模型内存占用量/MB	检测速度/(f·s^{-1})
YOLO v5n	88.67	86.10	87.37	93.01	5.01	217.39
YOLO v6n	89.23	84.86	86.99	91.53	8.28	232.56
YOLO v7 - tiny	88.78	88.05	88.41	93.31	11.70	156.25
YOLO v8n	90.14	87.16	88.62	92.84	5.95	196.08
YOLO v9c	87.85	89.64	88.74	93.04	49.10	56.18
YOLO v10n	88.32	86.19	87.24	92.39	5.48	227.27
RTDETR - L	86.10	87.32	86.71	88.80	56.30	47.17
文献[12]	90.07	87.17	88.60	93.32	6.88	140.85
文献[14]	89.36	86.94	88.13	92.33	5.62	151.52
MSCS - YOLO	90.89	87.57	89.20	94.22	6.42	156.25

表 5 不同模型对不同成熟度的识别结果

Tab. 5 Recognition results of different models for different levels of maturity %

模型	成熟			中等成熟			未成熟		
	精确率	召回率	平均精度	精确率	召回率	平均精度	精确率	召回率	平均精度
YOLO v5n	91.42	90.51	95.21	82.14	80.95	88.88	92.44	86.84	94.93
YOLO v6n	92.08	89.80	94.78	83.40	77.77	86.10	92.19	87.02	93.71
YOLO v7 - tiny	92.55	91.09	95.55	82.25	82.72	89.77	91.54	90.34	94.61
YOLO v8n	92.49	90.35	95.53	85.26	82.74	88.34	92.69	88.41	94.66
YOLO v9c	91.86	91.51	95.84	82.44	83.93	88.62	89.25	93.48	94.67
YOLO v10n	89.86	88.03	94.91	83.69	82.74	88.63	91.42	87.80	93.62
RTDETR - L	90.68	90.35	92.75	81.20	82.25	81.28	86.43	89.37	92.36
文献[12]	92.40	90.73	95.40	85.53	82.74	89.59	92.28	88.03	94.96
文献[14]	90.86	92.28	95.23	85.93	80.00	87.80	91.29	88.56	93.94
MSCS - YOLO	92.25	91.92	96.35	88.60	81.55	92.00	91.83	89.25	94.32

由表 4 可以看出,在草莓成熟度识别的精确率方面,MSCS - YOLO 模型表现最为突出,达到 90.89%。尽管召回率略低于 YOLO v7 - tiny 和 YOLO v9c 模型,但仍优于 YOLO v8n 模型,并且与文献[12]和文献[14]模型相比也具有明显优势。相较于 YOLO v7 - tiny 和 YOLO v9c 模型,MSCS -

YOLO 模型内存占用量（6.42 MB）分别减少 45.22% 和 86.93%，在资源有限的环境中 MSCS－YOLO 模型更具部署和应用的便利性。

在综合性能指标方面，模型 MSCS－YOLO F1 值达到 89.20%，在所有对比模型中处于最高水平，显著优于文献[12]和文献[14]的模型；mAP 达到 94.22%，也居于最高水平，相较 YOLO v8n、RTDETR－L 和 YOLO v9c 模型分别高 1.38、5.42、1.18 个百分点，并高于文献[12]和文献[14]的模型。表明 MSCS－YOLO 不仅在通用检测性能上优于主流模型，同时在小目标检测和遮挡处理方面也表现更为优越。

在检测速度方面，MSCS－YOLO 模型达到 156.25 f/s，展现出良好的实时检测能力。虽然低于部分轻量化模型（YOLO v5n：217.39 f/s、YOLO v6n：232.56 f/s、YOLO v8n：196.08 f/s、YOLO v10n：227.27 f/s），但远高于参数量更大的模型（YOLO v9c：56.18 f/s、RTDETR－L：47.17 f/s）。

由表 5 可以看出，由于成熟度较高的草莓特征较为明显，现有主流目标检测模型能够有效学习并识别这些特征，因此在识别精度、召回率和平均精度方面表现出较高的性能。相比之下，对于中等成熟度的草莓，由于其特征不够突出，当前主流模型的识别效果普遍较差；而在未成熟草莓的识别中，不同模型的表现存在一定差异。

在识别成熟及中等成熟草莓方面，MSCS－YOLO 模型平均精度分别达到 96.35%、92.00%，明显高于其他主流模型并优于文献[12]和文献[14]模型，相较 YOLO v8n 模型分别高 0.82、3.66 个百分点，说明其在特征不够突出的场景下也具备更强的识别能力。对于未成熟草莓，MSCS－YOLO 模型的表现与大多数主流模型基本持平，但仍略高于文献[12]和文献[14]。

3.2.2　不同环境下不同模型识别结果

为了验证本文模型在不同环境下对草莓成熟度识别的性能，对测试集中傍晚图像 36 幅、晴天图像 46 幅、阳光照射图像 56 幅以及灯光照射图像 27 幅，分别采用 YOLO v5n、YOLO v6n、YOLO v7－tiny、YOLO v8n、YOLO v9c、YOLO v10n、RTDETR－L、文献[12]模型、文献[14]模型以及本文 MSCS－YOLO 模型进行识别对比，识别结果如表 6～9 所示。

表 6　傍晚图像识别结果

Tab.6　Recognition results in evening　　%

模型	精确率			召回率			平均精度			平均精度均值
	成熟	中等成熟	未成熟	成熟	中等成熟	未成熟	成熟	中等成熟	未成熟	
YOLO v5n	94.17	90.28	93.56	91.67	92.91	82.67	96.84	95.64	94.52	95.67
YOLO v6n	97.02	75.83	92.41	91.67	83.33	82.67	95.94	86.19	93.05	91.73
YOLO v7－tiny	96.44	87.43	90.66	97.22	92.77	90.67	97.65	95.72	94.60	95.99
YOLO v8n	97.10	85.28	94.07	92.89	96.57	86.67	96.48	92.38	95.79	94.89
YOLO v9c	97.11	86.02	93.17	97.22	90.00	90.93	97.87	91.89	95.28	95.02
YOLO v10n	89.89	87.72	90.45	88.89	96.67	88.44	95.59	92.15	93.04	93.59
RTDETR－L	90.48	79.11	86.80	88.89	83.33	87.69	91.15	84.69	90.42	88.75
文献[12]	94.43	84.56	91.93	94.11	96.67	91.15	96.57	91.53	95.29	94.46
文献[14]	97.11	86.87	92.40	93.38	90.00	86.00	96.70	92.13	92.98	93.94
MSCS－YOLO	91.51	79.16	83.89	100	90.00	93.73	98.06	92.35	93.29	94.57

表 7　晴天图像识别结果

Tab.7　Recognition results on sunny days　　%

模型	精确率			召回率			平均精度			平均精度均值
	成熟	中等成熟	未成熟	成熟	中等成熟	未成熟	成熟	中等成熟	未成熟	
YOLO v5n	95.93	88.44	96.05	89.90	88.30	91.59	97.22	93.92	97.75	96.30
YOLO v6n	93.07	83.36	91.96	89.52	86.71	93.20	96.21	91.54	96.41	94.72
YOLO v7－tiny	95.81	80.55	95.44	87.16	90.38	95.15	96.00	91.53	97.71	95.08
YOLO v8n	96.94	90.07	95.44	90.52	88.46	93.85	97.32	94.62	97.58	96.51
YOLO v9c	94.98	81.78	92.92	90.19	86.31	97.09	96.68	89.74	97.59	94.67
YOLO v10n	93.25	86.52	92.97	92.03	86.54	94.13	96.92	92.63	96.24	95.26
RTDETR－L	96.46	80.82	93.82	89.52	82.69	87.38	94.05	81.56	95.20	90.27
文献[12]	96.17	90.19	95.98	91.43	88.40	92.82	97.91	95.41	97.87	97.07
文献[14]	95.08	88.32	94.12	92.05	86.54	93.19	96.95	91.71	96.78	95.14
MSCS－YOLO	94.39	93.45	95.01	91.43	88.46	95.47	97.07	94.99	97.45	96.50

表 8 阳光照射图像识别结果

Tab. 8 Recognition results under sunlight %

模型	精确率			召回率			平均精度			平均精度均值
	成熟	中等成熟	未成熟	成熟	中等成熟	未成熟	成熟	中等成熟	未成熟	
YOLO v5n	87.69	74.86	92.10	92.49	73.58	85.16	95.00	93.24	93.72	90.65
YOLO v6n	89.12	86.06	94.45	87.01	69.87	84.11	93.13	79.64	93.25	88.68
YOLO v7 - tiny	88.57	81.25	94.61	90.59	73.58	88.34	95.42	83.08	93.90	90.80
YOLO v8n	85.43	76.26	89.02	93.51	78.79	88.69	94.95	83.50	93.32	90.59
YOLO v9c	88.62	76.44	87.75	90.98	77.36	92.23	94.35	85.74	93.93	91.34
YOLO v10n	86.06	77.77	94.46	80.19	73.58	84.39	91.66	82.39	93.69	89.25
RTDETR - L	88.59	85.88	87.54	90.71	81.13	86.93	93.33	80.19	92.31	88.61
文献[12]	90.27	80.88	90.68	92.21	73.58	96.57	94.33	82.26	93.53	90.04
文献[14]	84.42	78.52	89.45	94.81	75.85	87.63	93.71	81.44	93.29	89.48
MSCS - YOLO	89.23	84.33	89.61	93.51	75.47	88.36	95.29	88.29	93.14	92.24

表 9 灯光照射图像识别结果

Tab. 9 Recognition results of light exposure %

模型	精确率			召回率			平均精度			平均精度均值
	成熟	中等成熟	未成熟	成熟	中等成熟	未成熟	成熟	中等成熟	未成熟	
YOLO v5n	77.35	77.57	73.60	90.24	83.87	84.88	91.72	84.89	87.33	87.98
YOLO v6n	89.03	92.74	85.92	92.68	77.47	78.02	95.72	93.17	87.74	92.21
YOLO v7 - tiny	88.57	81.25	94.61	90.59	73.58	88.34	95.42	83.08	93.90	90.80
YOLO v8n	84.10	82.32	82.63	90.28	78.79	77.91	92.18	84.95	85.43	87.52
YOLO v9c	89.55	90.28	85.81	92.68	84.43	84.40	94.56	92.23	86.64	91.14
YOLO v10n	92.50	86.87	83.20	90.21	84.85	74.42	95.14	90.29	83.41	89.61
RTDETR - L	92.15	80.35	77.96	90.24	78.79	84.88	91.18	80.48	86.01	85.89
文献[12]	85.47	86.62	88.42	86.10	78.50	71.04	91.10	89.34	87.58	89.34
文献[14]	84.35	87.29	77.38	92.68	90.91	86.05	95.36	90.54	86.39	90.76
MSCS - YOLO	93.08	93.25	88.15	92.68	83.72	77.86	95.67	94.09	87.96	92.57

由表6可知,在傍晚条件下,MSCS - YOLO 模型对草莓成熟度识别平均精度均值为 94.57%,虽然略低于 YOLO v8n 模型(94.89%)和 YOLO v9c 模型(95.02%),但依然处于较高水平,比识别精度最低的 RTDETR - L 模型高 5.82 个百分点。特别是在成熟和未成熟类别召回率分别为 100% 和 93.73%,比 YOLO v8n 模型分别高 7.11、7.06 个百分点,也明显优于文献[12]和文献[14],表明该模型在傍晚低光照环境下能够有效减少漏检现象。

由表7可知,在晴天条件下,MSCS - YOLO 模型对中等成熟度草莓识别精确率达到 93.45%,较 YOLO v8n 模型提升 3.38 个百分点,且显著高于文献[12]和文献[14]。本文模型对未成熟草莓的召回率达到 95.47%,优于 YOLO v8n、文献[12]和文献[14]。综合平均精度指标,MSCS - YOLO 在 3 种成熟度草莓的识别上均优于其他对比模型,证明其在晴天条件下具有显著优势。

由表8可知,在阳光照射条件下,MSCS - YOLO 模型对中等成熟草莓的识别精确度高于 YOLO v5n、YOLO v7 - tiny、YOLO v8n、YOLO v9c 和 YOLO v10n 模型,比 YOLO v8n 模型高 3.8 个百分点,仅略低于 YOLO v6n 和 RTDETR - L 模型,对于未成熟草莓的识别精确度也高于 YOLO v8n 原模型。在召回率方面,MSCS - YOLO 模型对于成熟草莓的表现尤为出色,除与 YOLO v8n 持平外,均高于其他模型。在平均精度均值方面,MSCS - YOLO 模型达到 92.24%,均高于其他所有模型,比 YOLO v8n 高 1.65 个百分点,比平均精度均值最低的 RTDETR - L 模型高 3.63 个百分点。表明在阳光照射条件下,MSCS - YOLO 模型在识别成熟和中等成熟草莓时更具优势。

由表9可知,在灯光照射条件下,MSCS - YOLO 模型对于成熟和中等成熟草莓识别的精确率均高于其他模型,对于未成熟草莓的识别精确率仅低于 YOLO v7 - tiny 和文献[12]模型。在召回率方面,MSCS - YOLO 模型对成熟草莓的表现尤为出色,相比于 YOLO v5n、YOLO v7 - tiny、YOLO v8n、YOLO v10n、RTDETR - L 和文献[12]模型更具优势,并与 YOLO v6n 和 YOLO v9c 模型持平。特别是对于成熟草莓识别的召回率,MSCS - YOLO 模型明显优于

YOLO v8n 模型，高出 2.40 个百分点。此外，MSCS－YOLO 模型在 3 种成熟度草莓的平均精度均值上也均高于其他所有模型，比 YOLO v8n 模型高 5.05 个百分点，比平均精度均值最低的 RTDETR－L 模型高 6.68 个百分点。

3.2.3　不同模型识别效果

为了验证本文模型（MSCS－YOLO）对不同成熟度草莓的识别效果，从测试集中随机选取不同环境下图像各 1 幅，并使用训练好的 10 个模型分别对这些图像进行识别，结果如图 9 所示。

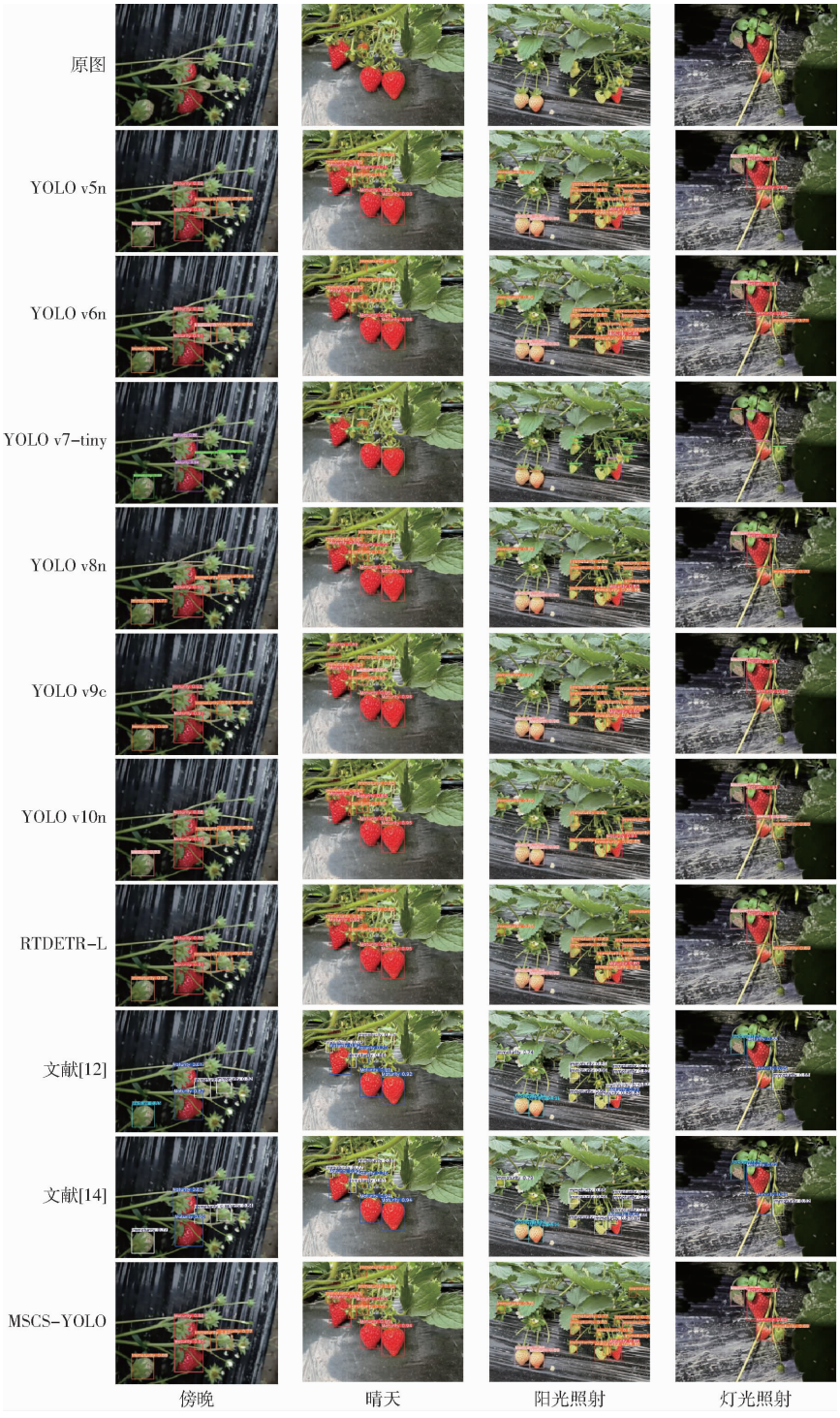

图 9　不同模型检测结果对比

Fig. 9　Comparison of different models detection results

由图 9 可知，傍晚时由于光线较暗，草莓特征模糊，YOLO v5n、YOLO v6n、YOLO v10n 和文献[12]模型均出现了将未成熟的草莓误识别为中等成熟的现象，而 YOLO v8n 模型漏检 1 个成熟草莓。相比之下，YOLO v7－tiny、YOLO v9c、RTDETR－L、文献[14]模型以及 MSCS－YOLO 模型准确识别出了每个草莓的成熟度。

晴天情况下，虽然视觉效果较优，但草莓枝叶与

果实之间的相互遮挡导致所有模型均存在不同程度的漏检现象。

阳光照射情况下,由于部分草莓个体较小,YOLO v7 - tiny、YOLO v8n 模型对未成熟较小果实的草莓出现了漏检现象,YOLO v10n 模型由于 2 个未成熟草莓形成了粘连,出现了将 2 个未成熟草莓误识别为 1 个未成熟草莓的现象,其余模型准确识别出了草莓果实。

灯光照射情况下,由于灯光干扰,草莓视觉上显示不够明显,YOLO v5n 和 YOLO v9c 模型均漏检 1 个未成熟的草莓,YOLO v7 - tiny 模型误将成熟草莓同时识别为中等成熟和成熟 2 种状态,YOLO v10n 模型误将成熟的草莓识别为中等成熟状态,RTDETR - L 和 YOLO v8n 模型漏检了 1 个成熟草莓,而 YOLO v6n、文献[12]模型、文献[14]模型、MSCS - YOLO 模型准确识别出了所有草莓的成熟度。

综上,MSCS - YOLO 模型在傍晚光照不足、晴天遮挡、阳光直射以及灯光照射等复杂场景下均展现出更优的鲁棒性与准确性,相较于 YOLO 系列模型及文献[12]模型,误检和漏检率显著降低,检测结果更加稳定。而与文献[14]模型相比,MSCS - YOLO 在不同光照条件下表现出相当甚至更优的检测性能。试验结果表明,MSCS - YOLO 在草莓成熟度智能检测中具备更强的应用潜力,为智慧农业的产业化推广提供了可靠的技术支撑。

由于 Grad - CAM[33-34]可以通过热力图显示模型对草莓成熟度特征的学习情况,为进一步验证本文模型(MSCS - YOLO)在草莓成熟度识别方面的性能,分别利用 MSCS - YOLO、YOLO v8n 模型对选取的 4 幅不同环境下图像进行了特征图可视化,结果如图 10 所示。

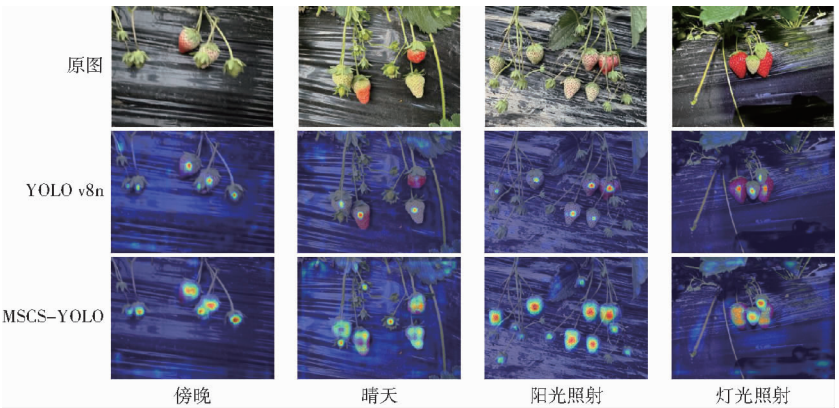

图 10 特征可视化结果

Fig. 10 Heat map of test results

从图 10 可以看出,在晴天和阳光照射条件下,YOLO v8n 模型未能充分学习到被遮挡草莓的特征,导致漏检现象的发生。而 MSCS - YOLO 模型则能够准确地捕捉每个草莓的特征,无漏检现象。在傍晚和灯光照射的条件下,YOLO v8n 模型仅能够学习每个草莓的部分特征,影响了对草莓成熟度的识别精度,相比之下,MSCS - YOLO 模型能够学习草每个莓的所有特征。综上所述,MSCS - YOLO 模型在草莓特征学习的完整性方面表现更为优异,因此在草莓成熟度识别上,相较于 YOLO v8n 模型具有更明显优势。

4 结论

(1)提出了一种基于 MSCS - YOLO 的草莓成熟度识别方法,通过在 YOLO v8n 模型的 Neck 部分引入多尺度扩展注意力机制,有效扩大了模型的感受野,解决了草莓目标较小且特征易被忽略的问题。其次,将 Neck 部分的 C2f 结构改进为 C2f - Triplet attention 结构,从 3 个维度更全面地捕捉草莓图像的信息,增强了模型在果实遮挡情况下目标识别能力。同时,采用改进的 SAHead 检测头代替 YOLO v8n 原始检测头,提升了模型在非结构环境下对不同成熟度草莓的识别精度。

(2)MSCS - YOLO 模型在 3 种成熟度草莓识别任务上的平均精度达到 94.22%,比 YOLO v8n 模型高 1.38 个百分点,比 RTDETR - L 模型高 5.42 个百分点。特别是在成熟和中等成熟度草莓识别方面,平均精度分别达到 96.35% 和 92.00%,较 YOLO v8n 模型分别高 0.82、3.66 个百分点,显著优于其他主流检测模型。

(3)无论是在傍晚、晴天、阳光直射还是灯光照射环境下,MSCS - YOLO 模型均展现出更强的识别能力和更高的精度。改进模型内存占用量仅为 6.42 MB,与 YOLO v8n 模型基本持平,但相比 YOLO v7 - tiny 和 YOLO v9c 模型分别减少 45.22% 和

86.93%，大幅降低了计算资源需求，使其更适用于资源受限环境，具备高效部署与应用推广的潜力，为农业领域的自动化果实采摘和智能农业管理提供了可靠的技术支持。

参 考 文 献

[1] 刘莫尘，褚镇源，崔明诗，等. 基于改进 YOLO v8－Pose 的红熟期草莓识别和果柄检测[J]. 农业机械学报，2023，54(增刊 2)：244－251.
LIU Mochen, CHU Zhenyuan, CUI Mingshi, et al. Red ripe strawberry recognition and stem detection based on improved YOLO v8－Pose[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 2): 244－251. (in Chinese)

[2] 刘华，赵祎平，王琳. 中国草莓生产空间布局变迁及影响因素研究[J]. 中国农业资源与区划，2024，45(12)：49－59.

[3] 谢志勇，张铁中，赵金英. 基于 Hough 变换的成熟草莓识别技术[J]. 农业机械学报，2007，38(3)：106－109.
XIE Zhiyong, ZHANG Tiezhong, ZHAO Jinying. Ripened strawberry recognition based on Hough transform[J]. Transactions of the Chinese Society for Agricultural Machinery, 2007, 38(3): 106－109. (in Chinese)

[4] 侯贵洋，赵桂杰，王璐瑶. 草莓采摘机器人图像识别系统研究[J]. 软件，2018，39(6)：184－188.

[5] ZHANG D, LEE D J, TIPPETTS B J, et al. Date maturity and quality evaluation using color distribution analysis and back projection[J]. Journal of Food Engineering, 2014, 131: 161－169.

[6] WANG C, TANG Y, ZOU X, et al. Recognition and matching of clustered mature litchi fruits using binocular charge-coupled device (CCD) color cameras[J]. Sensors, 2017, 17(11): 2564.

[7] WANG Y, YAN G, MENG Q, et al. DSE－YOLO: detail semantics enhancement YOLO for multi-stage strawberry detection [J]. Computers and Electronics in Agriculture, 2022, 198: 107057.

[8] DU X, CHENG H, MA Z, et al. DSW－YOLO: a detection method for ground-planted strawberry fruits under different occlusion levels[J]. Computers and Electronics in Agriculture, 2023, 214: 108304.

[9] YANG S, WANG W, GAO S, et al. Strawberry ripeness detection based on YOLO v8 algorithm fused with LW-Swin Transformer[J]. Computers and Electronics in Agriculture, 2023, 215: 108360.

[10] LIU H, WANG X, ZHAO F, et al. Upgrading swin-B transformer-based model for accurately identifying ripe strawberries by coupling task-aligned one-stage object detection mechanism[J]. Computers and Electronics in Agriculture, 2024, 218: 108674.

[11] 杨震宇，汪小昱，祁子涵，等. 基于改进 YOLO v8 的草莓识别与果梗采摘关键点检测[J]. 农业工程学报，2024，40(18)：167－175.
YANG Zhenyu, WANG Xiaochan, QI Zihan, et al. Recognizing strawberry to detect the key points for peduncle picking using improved YOLO v8 model[J]. Transactions of the CSAE, 2024, 40(18): 167－175. (in Chinese)

[12] 张晖耀，黄力湘，陈继清，等. 基于改进 YOLO v5s 的草莓成熟度检测算法[J]. 南京农业大学学报，2025，48(4)：990－999.
ZHANG Huiyao, HUANG Lixiang, CHEN Jiqing, et al. Strawberry ripeness detection algorithm based on improved YOLO v5s [J]. Journal of Nanjing Agricultural University, 2025, 48(4): 990－999. (in Chinese)

[13] 汤泽政，伍奕桦，徐新明，等. 基于改进 YOLO v7－Tiny 的成熟草莓识别模型研究[J]. 江西农业大学学报，2023，45(6)：1528－1542.
TANG Zezheng, WU Yihua, XU Xinming, et al. Research on ripe-strawberry recognition model based on improved YOLO v7－Tiny[J]. Acta Agriculturae Universitis Jiangxiensis, 2023, 45(6): 1528－1542. (in Chinese)

[14] 沈桂芳，张平. 基于改进 YOLO v8 的温室草莓成熟度智能实时识别[J]. 江苏农业科学，2025，53(5)：62－72.
SHEN Guifang, ZHANG Ping. Intelligent real-time recognition of strawberry maturity in greenhouses based on improved YOLO v8[J]. Jiangsu Agricultural Sciences, 2025, 53(5): 62－72. (in Chinese)

[15] 杨断利，王永胜，陈辉，等. 基于改进 YOLO v6－tiny 的蛋鸡啄羽行为识别与个体分类[J]. 农业机械学报，2023，54(5)：268－277.
YANG Duanli, WANG Yongsheng, CHEN Hui, et al. Feather pecking abnormal behavior identification and individual classification method of laying hens based on improved YOLO v6－tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(5): 268－277. (in Chinese)

[16] 杨断利，王永胜，陈辉，等. 基于 SEEC－YOLO v5s 的散养蛋鸡日常行为识别与统计系统[J]. 农业机械学报，2023，54(9)：316－328.
YANG Duanli, WANG Yongsheng, CHEN Hui, et al. Daily behavior recognition and real-time statistics system of free-range laying hens based on SEEC－YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 316－328. (in Chinese)

[17] 杨断利，王永胜，陈辉，等. 复杂环境下蛋鸡个体识别与自动计数系统研究[J]. 农业机械学报，2023，54(6)：297－306.
YANG Duanli, WANG Yongsheng, CHEN Hui, et al. Individual identification and automatic counting system of laying hens under complex environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 297－306. (in Chinese)

[18] CHEN C,LIU M Y, TUZEL O, et al. RCNN for small objectdetection[C]// Proceeding of Asian Conference on Computer Vision. Cham: Springer, 2016: 214230.

[19] 公徐路,张淑娟.基于改进 YOLO v5s 的苹果叶片小目标病害轻量化检测方法[J].农业工程学报,2023,39(19):175 - 184.
GONG Xulu, ZHANG Shujuan. Lightweight detection of small target diseases in apple leaf using improved YOLO v5s[J]. Transactions of the CSAE, 2023, 39(19): 175 - 184. (in Chinese)

[20] 张宏建,孙智霖,亓新春,等.基于改进 YOLO v8 的苹果树树干精准识别方法[J].农业机械学报,2024,55(增刊 1):246 - 255,262.
ZHANG Hongjian, SUN Zhilin, QI Xinchun, et al. Accurate apple tree trunk recognition method based on improved YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(Supp. 1): 246 - 255, 262. (in Chinese)

[21] 孔德航,刘云强,崔巍,等.基于改进 YOLO v8n 的辣椒穴盘育苗播种质量轻量级检测方法[J].农业机械学报,2025,56(2):381 - 392.
KONG Dehang, LIU Yunqiang, CUI Wei, et al. Lightweight detection method for seeding quality of chili seedling trays based on improved YOLO v8n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2): 381 - 392. (in Chinese)

[22] 杨断利,齐俊林,陈辉,等.基于改进 YOLO v8n 模型的散养蛋鸡个体行为识别方法与差异分析[J].农业机械学报,2024,55(11):112 - 123.
YANG Duanli, QI Junlin, CHEN Hui, et al. Individual behavioral identification and differential analysis of free-range laying hens based on improved YOLO v8n model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 112 - 123. (in Chinese)

[23] 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.

[24] 罗柳茗,李岩舟,石美琦,等.基于 YOLO v8n 的甘蔗杂草识别方法[J].中国农机化学报,2025,46(2):237 - 244.
LUO Liuming, LI Yanzhou, SHI Meiqi, et al. A identification method for sugarcane weed based on YOLO v8n[J]. Journal of Chinese Agricultural Mechanization, 2025, 46(2): 237 - 244. (in Chinese)

[25] JIAO J, TANG Y M, LIN K Y, et al. Dilateformer: multi-scale dilated transformer for visual recognition[J]. IEEE Transactions on Multimedia, 2023, 25: 8906 - 8919.

[26] MISRA D, NALAMADA T, ARASANIPALAI A U, et al. Rotate to attend: convolutional triplet attention module[C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 3139 - 3148.

[27] LIN Z, FENG M, SANTOS C N, et al. A structured self-attentive sentence embedding[J]. arXiv Preprint, arXiv:1703.03130, 2017

[28] LI C, LI L, JIANG H, et al. YOLO v6: a single-stage object detection framework for industrial applications[J]. arXiv Preprint, arXiv: 2209.02976,2022.

[29] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLO v7: 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.

[30] WANG C Y, YEH I H, MARK L H Y. Yolov9: learning what you want to learn using programmable gradient information[C]// European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 1 - 21.

[31] WANG Ao, CHEN Hui, LIU Lihao, et al. YOLO v10: real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984 - 108011.

[32] ZHAO Yi'an, LÜ Wenyu, XU Shangliang, et al. Detrs beat YOLOs on real-time object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965 - 16974.

[33] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]// Proceedings of the IEEE International Conference on Computer Vision, 2017: 618 - 626.

[34] 韩源涛,张聪,詹晓芸,等. 基于改进 ConvNeXt 模型的农作物害虫细粒度识别[J]. 农业工程学报,2025,41(4):185 - 192.
HAN Yuantao, ZHANG Cong, ZHAN Xiaoyun, et al. Fine-grained identification of crop pests using an enhanced ConvNeXt model[J]. Transactions of the CSAE, 2025, 41(4): 185 - 192. (in Chinese)