doi:10.6041/j.issn.1000-1298.2026.13.003

面向种植农业复杂场景的 YOLO 算法应用综述

李振波[1,2]　刘　欣[1,3]　竺明昊[1,4]　韩米娜[2,3]

(1.中国农业大学信息与电气工程学院,北京 100083;2.中国农业大学国家数字渔业创新中心,北京 100083;
3.农业农村部农业信息化标准化重点实验室,北京 100083;4.农业农村部国家数字农产品流通创新分中心,北京 100083)

摘要:YOLO 系列算法具有高推理速度、高检测精度和高度模块化的特点,已经成为智慧农业应用中目标检测的首选方法。本文系统梳理了 YOLO 算法在种植农业领域的研究进展与应用现状。首先,根据视觉任务的实现方式将种植农业领域视觉任务划分为果实检测类任务、农情分析类任务和采摘机器人视觉三大类,构建模型适配性评估框架,明确 YOLO 模型在各类任务中的适配程度及其应用现状。其次,针对复杂农业种植场景中普遍存在的小目标、遮挡、光照变化多样及算力限制等共性挑战,归纳主流的 YOLO 模型的改进策略,涵盖模型结构与模块优化、注意力机制引入等主流改进方法。最后,从视觉任务性能主导因素、高质量数据需求、模型版本选择和优化策略等方面综合分析 YOLO 系列算法在种植农业中的应用规律,并从感知能力突破与泛化能力提升两个方向展望未来发展趋势。

关键词:YOLO;目标检测;实例分割;关键点检测;姿态估计;轻量化

中图分类号:TP391　　**文献标识码**:A　　**文章编号**:1000-1298(2026)13-0045-23　　**OSID**:

Survey of YOLO Algorithm Applications for Complex Scenarios in Crop Agriculture

Li Zhenbo[1,2]　Liu Xin[1,3]　Zhu Minghao[1,4]　Han Mi'na[2,3]

(1. *College of Information and Electrical Engineering*, *China Agricultural University*, *Beijing* 100083, *China*
2. *National Innovation Center for Digital Fishery*, *China Agricultural University*, *Beijing* 100083, *China*
3. *Key Laboratory of Agricultural Informatization Standardization*, *Ministry of Agriculture and Rural Affairs*, *Beijing* 100083, *China*
4. *National Digital Agricultural Products Circulation Innovation Sub-center*, *Ministry of Agriculture and Rural Affairs*, *Beijing* 100083, *China*)

Abstract: YOLO series algorithms, characterized by high inference speed, high detection accuracy, and a highly modular architecture, have become the preferred methods for object detection in smart agriculture. The research progress and application status of YOLO algorithms in the field of crop planting agriculture were systematically reviewsed. Firstly, according to the implementation mode of visual tasks, visual tasks in planting agriculture were classified into three categories: fruit detection tasks, crop condition analysis tasks, and robotic harvesting vision tasks. Based on this classification, a model suitability evaluation framework was constructed to clarify the adaptability and application status of YOLO models in different tasks. Secondly, in response to common challenges in complex agricultural planting scenarios, such as small targets, occlusion, diverse illumination conditions, and limited computing resources, the mainstream improvement strategies for YOLO models were summarized, including model structure and module optimization, as well as the introduction of attention mechanisms. Finally, from the perspectives of key factors dominating visual task performance, the demand for high-quality data, model version selection, and optimization strategies, the application patterns of YOLO series algorithms in planting agriculture were comprehensively analyzed, and the future development trends were discussed from two directions: breakthroughs in perception capability and improvements in generalization ability.

Key words: YOLO; object detection; instance segmentation; keypoint detection; pose estimation; lightweight

收稿日期:2026-04-01　修回日期:2026-04-30

基金项目:财政部和农业农村部:国家现代农业产业技术体系(岗位专家)项目(BAIC10-2025-E14)

作者简介:李振波(1978—),男,教授,博士,主要从事计算机视觉研究,E-mail:lizb@cau.edu.cn

0 引言

种植业是农业生产的基础，国家统计局数据显示，2024 年我国种植业产值达到 9.0 万亿元，占农业总产值的 55.05%[1]，是农业智能化转型的关键领域。目前，智慧农业已成为现代农业的主要发展方向之一，2024 年，农业农村部相继发布《关于大力发展智慧农业的指导意见》[2]系统阐明了智慧农业发展的总体方向，提出需着力破解信息感知、智能决策、精准作业各环节的瓶颈问题。《全国智慧农业行动计划（2024—2028 年）》[3]进一步明确了到 2030 年农业生产信息化率达到 35% 的发展规划。2026 年中央一号文件强调需促进人工智能与农业发展相结合，加快农业生物制造关键技术创新[4]。同年农业农村部关于落实《中共中央国务院关于锚定农业农村现代化扎实推进乡村全面振兴的意见》的实施意见[5]将加强人工智能在农业农村领域应用列为重点任务，标志着智慧农业从政策倡导迈向全面实施。其中，机器视觉作为智慧农业的核心感知技术，具有低成本、非接触、无损检测及远程监控等优点，能够代替人工为田间管理决策、产量预测及农机作业提供关键的视觉信息支撑，适用于大规模智慧农业应用场景。

机器视觉技术通过视觉传感器获取目标图像，并利用计算机对图像进行处理、分析与理解，从而实现对目标的检测、识别、定位与测量，已广泛应用于种植业中，其技术发展主要经历传统机器学习和深度学习两个阶段。传统机器学习方法依赖于人工设计特征进行识别，面对作物个体差异及复杂环境背景时，存在泛化能力弱、识别精度不稳定及计算效率低等问题。随着卷积神经网络的发展，深度学习算法已被广泛应用于种植领域中。R-CNN 通过引入“候选区域生成-分类回归”的串行机制，显著提升了目标检测精度和鲁棒性，但其计算开销较大。Faster R-CNN 将候选区域生成与分类回归统一到同一网络中，实现了端到端训练，显著提升了检测速度，但仍难以满足智慧农业实时作业的需求。YOLO（You only look once）[6-14]算法创造性地将目标检测任务重构为单一回归问题，具备推理速度快、检测精度高及模型轻量化等优势，其推理速度为 Faster R-CNN 的 6 倍，已成为当前智慧农业视觉研究中应用最广泛的目标检测方法之一。如图 1 所示，自 YOLO v4[7]发布以来，YOLO 系列算法确立了“Backbone-Neck-Head”三段式架构，研究者可依据任务需求及边缘设备的算力限制，对特定模块进行重构和优化。YOLO v5 凭借轻量化设计和完善的部署生态，推动了 YOLO 方法在实际场景中的落地。随后，YOLO v8 构建了统一的多任务框架，可在同一体系下支持目标检测、实例分割与姿态估计，拓展了模型的视觉理解能力。YOLO 11 则进一步提高了模型效率和对细粒度特征的提取能力。

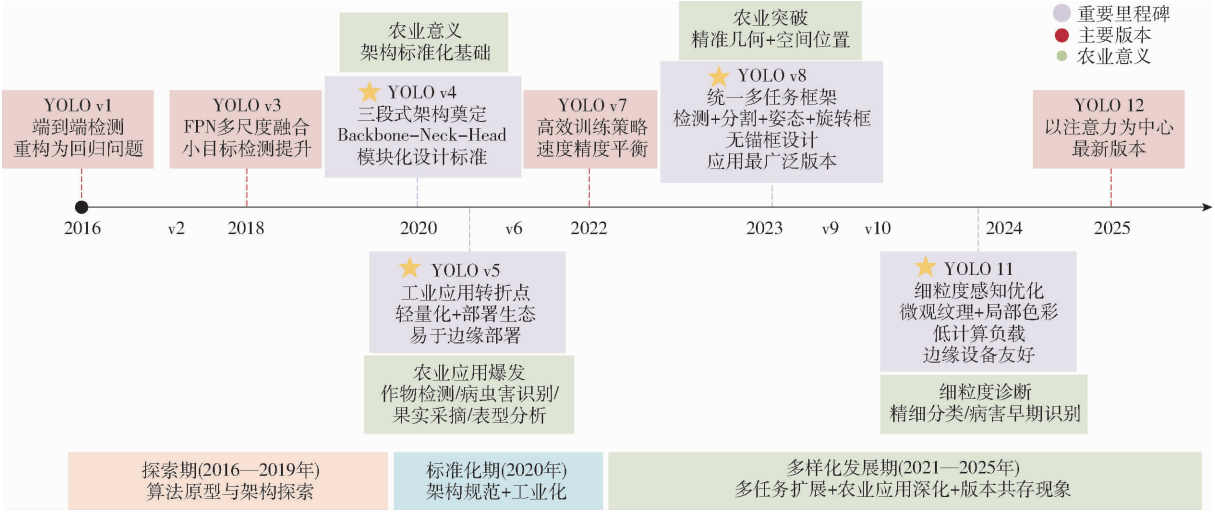


图 1 面向农业场景的 YOLO 系列算法研究进展与应用图

Fig. 1 Research progress and applications of YOLO series algorithms in agricultural scenarios

种植业是农业的核心产业，种植业视觉任务具有目标尺度差异大、遮挡严重、光照变化频繁和背景纹理复杂等特点，同时部署环境通常受限于算力与存储资源，因此面向实际生产场景的视觉算法研究面临较大挑战。根据应用需求，本文将种植业视觉任务归纳为果实检测类任务、农情分析任务和采摘机器人视觉任务 3 类。针对上述任务，研究者围绕 YOLO 系列模型的骨干网络、特征融合结构和检测头等关键模块开展了大量改进研究，并推动其在种植业中的应用与落地。

目前，已有综述对视觉算法在农业场景中的应用进行了系统调研，Chakrabarty 等[15]详述了不同监测场景下害虫识别的技术进展与差异化挑战，李伟强等[16]梳理了果实识别技术的研究进展，翟肇裕等[17]调研了 YOLO 在动植物表型研究中的应用现状。然而，现有文献仍缺乏从种植业应用需求出发，系统比较 YOLO 在种植领域视觉任务中的适用边界与优化策略的综合性综述。基于此，本文对 2023—2026 年种植业文献进行了全面调研，研究对象涵盖当前种植业的核心作物，包括蔬菜、水果、粮食及经济作物。聚焦从萌发出苗期到成熟采收期的全生育周期，围绕果实检测、产量预测、表型估计、病虫害识别、品质分级及机械化采收等视觉感知任务展开分析，这些任务是智慧种植领域研究的核心内容。在此基础上，梳理 YOLO 模型结构与模块优化、注意力机制引入等主流改进方法，为研究者提供从 YOLO 模型适配度分析到基线选型再到算法优化的系统性参考。本文的整体研究框架如图 2 所示。

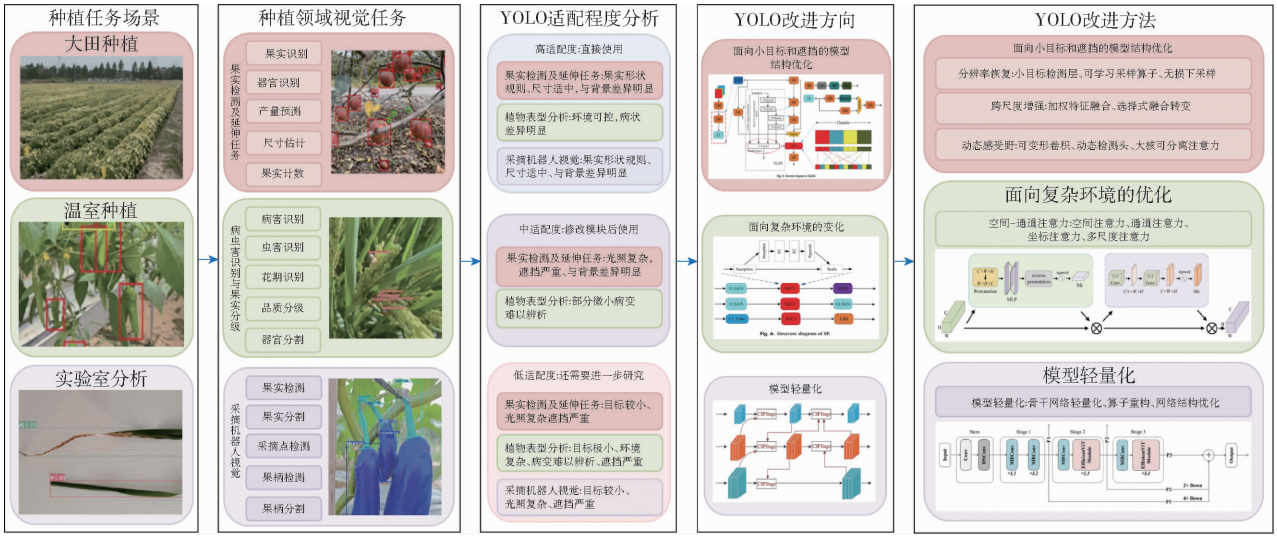


图 2　种植农业领域 YOLO 模型适配性分析与改进方法

Fig. 2　Applicability analysis and improvement of YOLO models in crop farming

1　种植业视觉感知任务

农作物生长周期包括作物萌发出苗期、营养生长期、生殖生长期和成熟采收期 4 个阶段，对应不同阶段的生产管理需求，具体如表 1 所示。

表 1　不同生长阶段对应的种植领域视觉感知任务及应用目标

Tab. 1　Visual perception tasks and application objectives in crop production across different growth stages

生长阶段	主要视觉任务	应用目标
萌发出苗期	幼苗检测与计数	出苗率检测、缺苗补种
营养生长期	病虫害识别	为田间管理提供决策依据
生殖生长期	花序检测、挂果量统计	为产量预估提供数据支持
成熟采收期	果实检测、成熟度判断、品质分级、采摘点定位	采收决策、品质评估与自动采摘

从视觉任务的实现方式看，上述视觉感知任务可进一步归纳为果实检测类、农情分析类和采摘机器人视觉 3 类。其中，果实检测类任务主要覆盖生殖生长期至采收期，以器官目标检测为基础，并延伸至产量预估、动态计数与尺寸估计等应用；农情分析类任务侧重于营养生长期的病虫害监测以及采收期的品质分级；采摘机器人视觉任务则主要集中于成熟采收期，结合目标检测、关键点定位与姿态估计，为机械臂提供精确的空间位置与几何形状信息。如图 3 所示（图中括号内数字分别表示该类研究的文献数量及其在所有场景中的占比，如农情分析为 35 篇，占比 38.0%），对近年文献进行统计分析，发现多数研究集中于目标检测这一基础任务，而面向实例分割、姿态估计等精细化作业研究则相对较少，表明当前农业模型的环境理解能力仍处于初级阶段。

在评价指标方面，目标检测任务主要采用平均精度均值（mean Average Precision，mAP）、精确率（Precision）和召回率（Recall）等指标。其中，Precision 衡量模型检测结果的准确性；Recall 衡量模型对目标的检出能力；mAP 综合考虑不同置信度阈值下的 Precision 和 Recall，衡量模型的整体检测性能。产量预测任务采用决定系数（R^2）衡量预测值与真实值的拟合程度，其值越接近 1 表示预测越准确；相对误差则直接反映预测偏差的大小。定位任务采用平均绝对误差（Mean absolute error，MAE）和均方根误差（Root mean square error，RMSE）衡量预测位置与真实位置的偏差程度，数值越小表示定位精度越高。为了系统比较 YOLO 系列模型在不同

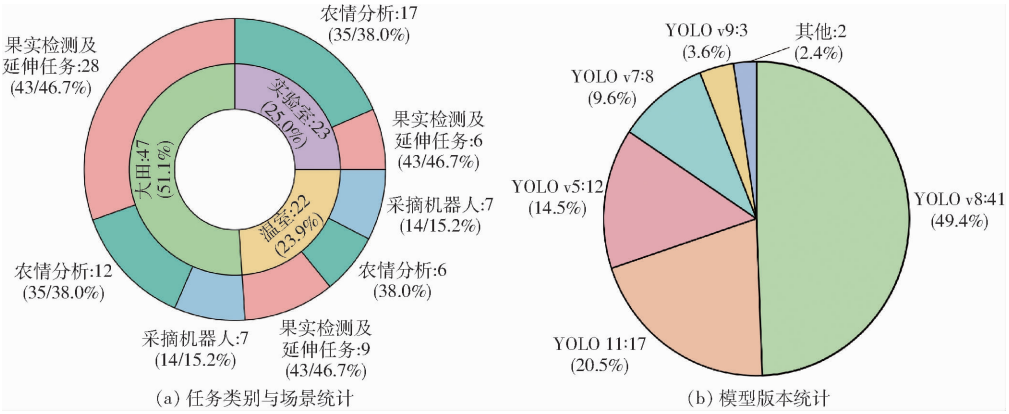


图 3 文献统计图

Fig. 3 Statistics of reviewed literature

农业视觉任务中的适配性,本文结合任务特点、文献结果分布和实际应用需求,构建基于关键性能指标的适配性评价体系。已有研究表明,农业视觉任务的性能评价具有显著的应用导向特征[18-19],如 Caballero-Ramirez 等[20]在对比实验表明,人工综合准确率为 92.4%,YOLO v5 为 93.8%,说明当检测精度达到 90% 左右时,模型已接近人工判别水平。结合本文所统计文献的结果分布,多数能够有效支撑识别、预警和作业执行的研究,其检测性能通常集中在 90% 及以上。基于此,本文以任务常用评价指标为基础,并结合性能稳定性、场景复杂度和作业需求,对 YOLO 模型适配水平进行分级。

1.1 果实检测类任务

果实检测是种植业智能化生产中应用最广泛、最基础的视觉感知任务,旨在通过机器视觉算法自动识别田间或温室中的果实、花序等器官,通过对检测结果进一步分析可实现挂果量统计、尺寸估计与产量预测等高级任务,如图 4 所示,为采收决策与供应链管理提供数据支撑。本文以 mAP50 作为果实检测类任务适配性分析的重要参考指标。多数目标尺度适中、背景差异明显的果实检测研究,其 mAP50 结果稳定分布在 90% 以上;而性能处于 90% 附近的研究通常对遮挡、光照和拍摄条件较为敏感;明显低于 90% 的研究则往往仍存在较突出的漏检与误检问题。基于这一文献分布特征,本文以 mAP50 等于 90% 作为果实检测类任务适配性分级的主分界线,构建高适配、中适配和低适配 3 级评价体系,用于不同研究之间的横向比较。



图 4 果实检测及延伸任务[21-25]

Fig. 4 Fruit detection and its extended tasks[21-25]

1.1.1 YOLO 在果实检测中的应用

种植场景中的果实检测需在遮挡、复杂光照及目标与背景色相近等条件下,实现对不同尺寸和形态果实的准确识别。其中,果实尺寸是影响检测精度的核心因素:较大的果实尺寸能够为模型提供更丰富的纹理、边缘等细节特征与空间信息,更有利于特征提取。对于苹果、草莓和柑橘等尺寸较大的果实,YOLO 基础模型即可达到较高的检测精度[26-28],研究重点应从精度提升转向模型轻量化[22,26,29-31]与泛化能力增强[32-33],以提升模型在实际生产场景中的可部署性。通过模型剪枝或主干、模块替换,轻量化改进可提升模型在边缘设备上的实时性,如吕强等[28]针对柑橘检测任务,通过将模型主干网络替换为幻影网络(GhostNet)大幅降低模型计算量,并进一步部署至安卓手机平台。通过模拟算法或风格迁移算法构建多场景数据集,泛化性改进可提升模型在不同环境下的鲁棒性,Wu 等[21]通过雾模拟算法构建包含多种天气条件的苹果图像数据集,并集成卷积注意力模块(Convolutional block attention module, CBAM)和 CARAFE 上采样算子,提高模型在复杂环境条件下的检测精度。

当果实尺度较小时,模型可提取的有效特征有限,遮挡与复杂光照条件易引入大量噪声,干扰特征

表达。此时，果实颜色与背景之间的对比度是影响检测精度的核心因素，对比度越高，目标与背景的可分性越强。因此，根据环境特点针对性引入高分辨率检测头[22,34-35]或跨尺度特征融合[36-38]等方法，可有效提高特征分辨率并抑制环境噪声干扰。如 Bai 等[39]在 YOLO v7 中引入 Swin Transformer 高分辨率检测头以提高模型对小目标的特征提取能力，mAP50 提升至 92.1%。刘祥等[40]在 YOLO v8 中引入 CAHSFPN 特征融合模块（Context-augmented hierarchical scale feature pyramid network，CAHSFPN）和 Focaler-MPDIoU 损失函数（Focaler minimum point distance intersection over union loss）提高对不同尺度棉花的识别精度，检测精度提升至 96.8%。

在极端场景下[41-43]，果实尺寸极小且颜色与背景相似，遮挡和复杂光照条件引入的噪声难以与有效特征区分，YOLO 模型易出现严重漏检或误检。高分辨率检测头或多尺度融合结构主要通过对已提取特征进行加权提高模型性能，难以从根本上弥补信息缺失，因此即使经过优化，检测精度也难以达到实际应用需求[23,44]。未来研究可通过引入多模态方法提高图像的输入信息密度以寻求突破[45-46]。此外，表 2 统计了不同农作物的特点、相关研究的效果和 YOLO 模型的适配程度。结果表明，即使针对同一检测任务和模型，不同研究间的精度差异也较为显著。这种差异一定程度上源于拍摄距离、相机角度等数据采集条件的不同。如同样采用 YOLO v7 进行温室草莓检测，在不同的采集条件下检测精度差异可达 10.2%[33,39,47]。由此可见，优化相机部署位置、统一拍摄距离与角度，可从数据源头提高 YOLO 模型的检测性能。

表 2　果实检测类任务统计

Tab. 2　Statistics of fruit detection tasks

作物	造成准确度差异的因素	适配性	场景	任务	模型	性能参数/%	文献序号
苹果	检测目标尺寸；与背景差异程度（检测部位不同）	高	大田	苹果检测	YOLO v8	mAP50：93.0～94.3	[21]
		低	大田	苹果花检测	YOLO v7	mAP：83.4～86.8	[47]
草莓	检测目标尺寸；遮挡与光照条件（摄像头距离与拍摄角度）	高	温室	草莓检测	YOLO v8	mAP50：94.7～97.2	[33]
		中	温室	草莓检测	YOLO v7	mAP50：87.5～92.1	[39]
		低	温室	草莓检测	YOLO v7	mAP50：84.5～86.7	[48]
柑橘	果实形状规则；尺寸适中	高	大田	柑橘检测	YOLO v5	mAP50：96.6～97.8	[28]
番茄	检测部位尺寸；与背景差异和遮挡（检测部位差异与相机距离）	高	温室	番茄果柄检测	YOLO v8	mAP50：90.3～91.7	[25]
		低	大田	番茄主茎检测	YOLO 11	mAP50：55.8～57.4	[43]
油茶果	检测目标尺寸；颜色；光照（果实成熟度与拍摄角度）	高	大田	油茶果检测	YOLO v5	AP：98.3～98.4	[49]
		低	大田	油茶果检测	YOLO 11	mAP50：85.1～87.6	[50]
棉花	检测目标尺寸；遮挡程度（相机距离与种植密度）	高	大田	棉花检测	YOLO v8	mAP50：96.0	[45]
	目标极小；遮挡严重；易与背景混淆	低	大田	棉花花蕾/花朵/棉铃	YOLO 11	mAP50：77.6～79.8	[51]
大豆	与背景差异（实验环境）	高	实验室	大豆根瘤	YOLO v8	mAP 95.4～98.2	[52]
		低	大田	大豆幼苗	YOLO v9	mAP：43.1～47.5	[42]
辣椒	与背景差异明显（生长阶段）	高	实验室	辣椒种子	YOLO v8	mAP50：95.8～98.3	[53]
花朵	遮挡与光照条件（拍摄角度，开花密度）	高	大田	金银花检测	YOLO v5	mAP：91.2～91.8	[31]
		高	温室	大马士革玫瑰检测	YOLO v9	mAP50：95.9～96.2	[54]
		低	大田	杭白菊检测	YOLO v8	mAP：83.2～84.9	[55]
幼苗	检测目标尺寸；与背景差异和遮挡（作物种类、相机距离、背景颜色）	高	大田	生菜幼苗	YOLO v5	Accuracy：98.7	[56]
		低	大田	小麦幼苗	YOLO 11	mAP50：82.3～84.5	[57]
		低	大田	玉米幼苗	YOLO v8	mAP：78.4～79.5	[46]
黄瓜	目标尺寸适中；与背景差异明显	高	温室	黄瓜分割	YOLO v8	mAP50：95.8～96.2	[58]
	目标尺度差异大；遮挡严重	低	温室	黄瓜叶/花	YOLO 11	mAP50：76.6～85.4	[59]

注：使用 mAP50 作为主要性能评价指标，若文献未给出 mAP50 则依次使用 mAP、平均精度（Average precision，AP）、准确率（Accuracy）、召回率（Recall）和精准率（Precision）作为评价指标。

1.1.2　基于 YOLO 的多模型融合应用

YOLO 检测器仅能完成果实的检测，若结合时序或回归模型对检测结果进行进一步分析，则可实现挂果量统计、产量预测和尺度估计等高级任务。在此类任务中，YOLO 模型作为上游感知模块为时序关联、统计回归和几何估计提供输入。因此，该类任务的性能不仅取决于 YOLO 本身的检测精度，还取决于检测结果与下游模型之间的误差传递关系。相关任务统计如表 3 所示。

挂果量统计任务通常采用“检测—跟踪—计

表 3 基于 YOLO 的多模型融合应用任务统计

Tab. 3 Statistics of YOLO-based multi-model fusion application tasks

适配性	因素	场景	任务	模型	性能参数	文献序号
高	果实形状规则;尺寸适中	大田	柑橘产量预测	YOLO v8	AP:91.9% ~92.5%,R^2: 0.98	[22]
低	目标较小;光照复杂;遮挡严重	大田	柑橘尺寸估计	YOLO v8	mAP:87% ~88%,R^2: 0.76	[23]
高	目标较小;与背景差异明显	大田	棉花计数	YOLO v8	mAP: 96.0%,R^2: 0.92	[45]
中	目标较小且密集	大田	棉铃产量预测	YOLO v8	mAP: 86.7% ~96.8% 平均相对误差:7.22%	[40]

注:使用 mAP50 作为主要性能评价指标,若文献未给出 mAP50 则使用 mAP 作为评价指标。对于产量预测、尺寸估计和计数任务,统一以 R^2 作为下游任务性能指标。若文献未给出 R^2,则使用文献中给出的结果评价指标。

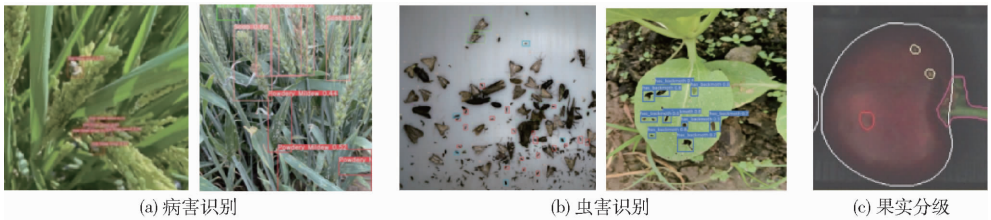
数”的级联框架,即先利用 YOLO 完成单帧果实检测,再通过 ByteTrack[60]、DeepSORT[61] 或光流法对检测结果进行跨帧关联,最后结合撞线或区域统计规则获得果树或田块尺度的挂果总量。在该类任务中,检测精度决定轨迹建立的上限,而跟踪算法决定跨帧去重与关联的稳定性。YOLO 模型需要解决果实在运动中因视角变化导致的检测不稳定问题,同一果实在视角变化、遮挡或光照波动下频繁失检与重现,会导致跟踪轨迹被重复创建,进而使挂果量被高估[45]。此外,在检测-跟踪级联框架下,跟踪器自身的误差同样构成瓶颈。因此,该类任务的优化重点不应局限于提高单帧检测精度,更应从检测与跟踪协同的角度提高时序一致性和目标关联稳定性。

产量预测任务通常以 YOLO 检测结果为基础,进一步通过统计或回归模型建立果实数量与实际产量之间的映射关系。YOLO 模型的检测数量与最终产量之间并非简单的线性关系,果实大小、成熟度差异以及漏检、误检等都会引入误差。刘祥等[40]将 YOLO 模型检测结果作为棉花产量,检测精度虽高达 96.8%,但产量估计精度仅达到 92.8%。回归模型具有一定的误差平滑能力,因此,相较于直接将检测结果等同于产量,采用回归模型对检测结果进行再建模,更有利于减少检测误差对最终预测结果的影响。Zeng 等[22]针对柑橘产量预测任务,通过简单线性回归模型将检测计数转化为产量预测值,当检测精度(AP)为 91% 时,产量预测精度可达到 98%。

尺寸估计任务首先通过 YOLO 模型对图像中的果实进行定位与分割,获取目标边界框或掩码区域。由于果实在二维图像中的像素尺寸受拍摄距离、镜头畸变和视角变化的影响,难以直接换算为真实物理尺寸,因此需要结合深度图像、自然参照物或几何先验信息估算果实直径,如 Qu 等[23]以成熟健康的叶片作为自然参照物,结合离线叶片尺寸分布统计,将图像中的像素尺寸换算为果实实际直径,其果径估计决定系数 R^2 为 0.76。该方法精度受限于参照物尺寸分布的稳定性。相比之下,深度信息提供的几何约束能更有效地控制误差传播。Zhou 等[62]使用深度相机采集多色苹果图像,通过融合分割掩码与深度信息进行点云重建与球体拟合,误差仅为 0.768 mm。

1.2 农情分析任务

农情分析旨在监控作物的生长状态与外观形态信息,为田间管理、采摘后处理和市场流通提供客观量化依据。其中,病虫害监测与品质分级是研究热点,对视觉算法有着较高的要求。如图 5 所示,农情分析任务关注作物个体或局部区域的视觉属性,包括颜色、纹理、形态异常及外观等级等,这对 YOLO 模型的细粒度识别能力提出了更高要求。农情分析任务同样以 mAP50 为主评价指标。表 4 和表 5 显示,当病斑特征明显、虫害与背景差异较大或品质等级差异清晰时,多数研究结果稳定分布在 90% 以上,此时 YOLO 模型通常能够较好完成病虫害识别与品质分级任务。当结果分布在 90% 附近时,模型性能往往对病斑细微差异、背景干扰和采集条件较为敏感;当结果整体低于 90% 时,则通常表明类间差异弱、目标极小或背景干扰强等因素仍对识别性能形成显著制约。基于此,本文以 mAP50 等于 90% 作为农情分析任务适配性分级的主分界线,对农情分析任务进行高



(a) 病害识别 (b) 虫害识别 (c) 果实分级

图 5 植物农情分析[63-67]

Fig. 5 Crop condition analysis[63-67]

表 4　病虫害检测任务统计

Tab. 4　Statistics of pest and disease detection tasks

任务类型	因素	适配性	场景	任务	模型	性能参数/%	文献序号
虫害识别	目标与背景差异明显	高	实验室	水稻害虫	YOLO v8	Precision：93.3～95.5 Recall：90.1～90.5	[65]
	目标极小，易与背景混淆	中	温室	温室昆虫	YOLO v8	mAP50：88.3～92.2	[68]
	目标极小，易与背景混淆，背景存在干扰	低	实验室	豇豆	YOLO 11	mAP50：84.4～88.1	[37]
		低	大田/实验室	番茄害虫	YOLO v5	mAP：79.6～84.3	[69]
		低	大田	大田害虫	YOLO v8	mAP50：75～80.1	[35]
病害识别	环境可控，病状明显	高	实验室	葡萄叶病	YOLO v8	mAP50：99.5	[70]
		高	实验室/大田	苹果叶病	YOLO v8	mAP50：92.4～94.6	[71]
	部分病害种类和微小病变难以辨析	高	大田	水稻病害	YOLO v5	mAP：95.4～94.2	[63]
		中	大田	番茄叶病	YOLO 11	mAP50：89.8～91.7	[72]
		中	大田	小麦病害	YOLO v8	mAP50：89.5～91.4	[64]
		低	实验室	燕麦病害	YOLO v7	mAP50：80.9～88.5	[36]
		低	大田	西瓜叶病	YOLO v9	mAP50：84～89	[73]
		低	实验室	黄瓜病害	YOLO v7	mAP50：82.6～86.5	[74]

注：使用 mAP50 作为主要性能评价指标，若文献未给出 mAP50，则依次使用 mAP、Recall 和 Precision 作为评价指标。

表 5　果实分级任务统计

Tab. 5　Statistics of fruit grading tasks

作物	因素	适配性	场景	任务	模型	性能参数/%	文献序号
马铃薯	不同分级差异明显	高	实验室	缺陷检测	YOLO v8	mAP：94.2～95.8	[75]
樱桃		高	实验室	缺陷检测	YOLO v8	果体 mAP50～95：99.5～99.5 果梗 IoU：92.15，瑕疵 mIoU：65.97	[67]
蓝靛果	目标较小，品质差异难以辨析	中	实验室	缺陷检测	YOLO v8	mAP50：89.9～94.3	[76]
可可豆	不同类别特征不明显 难以辨析	低	实验室	发酵程度分级	YOLO v8	mAP50：64.4～70.4	[77]
菖头		低	大田	关键点识别与分级	YOLO 11	尺寸分类精度：92.87 弯曲度分类精度：84.72	[78]

注：使用 mAP50 作为主要性能评价指标，若文献未给出 mAP50 则使用 mAP 作为评价指标。对于实例分割任务，若文献未给出 mAP50 或 mAP，则使用 mAP50－95、IoU 或 mIoU 作为评价指标。对于关键点识别与分类任务，若文献未给出 mAP50 或 mAP，则使用分类精确率或文献中给出的结果评价指标。

适配、中适配和低适配划分。

1.2.1　YOLO 在病虫害检测中的应用

病虫害是威胁作物产量与品质安全的关键因素，研究表明害虫导致的作物产量损失可达 26%[15]，及时有效的防治可有效降低种植过程中的经济损失。病虫害检测算法需在作物枝叶遮挡和复杂光照条件下实现对植物病虫害的准确定位与识别。

相较于植物果实，病斑与害虫尺寸更小。因此对于定位任务，病虫害特征与背景的区分度是影响检测精度的主要因素。由表 4 可知，当区分度较高时，YOLO 基础模型即可达到较高的检测精度[68－69]。然而，由于病斑和害虫的特征细微，复杂光照和遮挡条件易引入大量噪声，显著影响 YOLO 模型对病虫害的识别能力。通过引入注意力机制、多尺度特征融合等改进策略可有效提高检测精度，如马超伟等[64]在 YOLO v8 中引入全局注意力机制（Global attention mechanism，GAM）模块与 CARAFE 模块，提高模型对小麦病害特征的提取能力，检测精度提升 1.9 个百分点。郑俊键等[66]在 YOLO v5 中引入 CBAM 模块实现对小白菜害虫特征的高效筛选，检测精度提升 3.3 个百分点。当病虫害特征与背景的区分度过低时，单一 RGB 传感器存在信息密度上限，导致可用特征匮乏，此时仅依靠常规的模块改进难以进一步提升检测精度[35,69]。未来，通过多光谱与高光谱成像可弥补可见光波段下目标与背景较为相似的问题，为低区分度病虫害检测提供更丰富的光谱维度信息。

对于病虫害分类任务，其准确率取决于不同类别之间的视觉差异。当病虫害类间特征相似时，模型的分类性能将显著下降。针对这一问题，可通过引入注意力机制、细粒度特征提取等策略增强模型对局部细微差异的感知能力，提升 YOLO 模型的类间区分能力。如 Zhang 等[36]引入秩扩张网络（Rank expansion networks，ReXNet）提高模型对燕麦病害细微病变纹理的提取能力，识别精度达到 88.5%。此外，病虫害图像的标注需要专业植保知识，且病斑边界模糊、不

同病害性状相似导致高质量病虫害数据集的标注耗时长、成本高。由于 YOLO 模型普遍采用数据驱动的训练方式,因此高质量病虫害数据集的缺失已成为限制害虫识别模型泛化能力的首要瓶颈[15]。

1.2.2 YOLO 在果实分级中的应用

品质分级是农产品产后处理与流通环节的核心技术,直接关系产品的市场价值与消费者体验。果实品质的评定涉及颜色、形状、尺寸、表面瑕疵等多个维度,不同等级间的外观差异往往细微且具有主观性,因此等级间的视觉差异程度直接决定了分级精度。如表 5 所示,果实分级通常在实验室环境下进行,其核心挑战在于细粒度特征的提取与判别。当不同等级间特征区分度较高时,如识别樱桃碰伤、开裂等瑕疵或马铃薯缺陷,YOLO 算法检测精度可达 90% 以上[67,75]。

当分级涉及细粒度外观差异时,模型对特征提取能力的要求显著提高。若差异仍可通过外观表征区分,通过引入特征增强策略可有效提升分类精度。如 Xue 等[76]在 YOLO v8 中引入大型可分离核注意力模块与动态检测头,提升模型对蓝靛果早期损伤特征的提取能力,检测精度达到 94.3%。此外,当环境可控时通过调整采集条件,如镜头距离或背景光,亦可有效提升分级精度。

若分级标准依赖于内部品质仅靠视觉信息难以区分,如发酵程度,检测难度将显著增加。例如,可可豆不同发酵阶段差异极小,YOLO v8 检测精度仅为 64.4%[77],即便使用大选择性核模块(Large selective kernel block, LSK)等多种改进策略,检测精度仍仅有 70.4%。这表明,若分级所依据的特征在可见光波段下缺乏,改进策略的提升空间将受到根本性制约。此时可结合高光谱成像、近红外光谱等技术引入内部品质的直接表征信息。此外,工业化分拣场景对算法的实时性提出了较高要求,分拣线上农产品流量大、速度快,模型需在保证精度的同时具备高吞吐能力,因此轻量化部署成为该类任务的重要改进方向[75]。

1.3 采摘机器人视觉

机器人视觉是实现农业自动化采收的核心技术,如图 6 所示,果实检测与采摘点检测是两项基础任务。机器人采摘任务遵循感知—决策—执行的闭环架构,首先通过 YOLO 模型获取果实掩码、抓取点等位置信息;然后,结合深度图像或立体视觉将二维像素坐标转换为三维空间位置;最后,由路径规划模块引导机械臂末端执行器完成果实的精准抓取或剪切操作。采摘机器人视觉任务同时涉及目标检测、实例分割、关键点定位和三维位姿估计,其适配性不能仅由单一视觉指标判定。关键点检测和实例分割等任务中普遍采用 mAP50 和 mIoU 作为核心指标,结合文献[19]及本文统计结果,当 mAP50 大于 90% 或 mIoU 大于 80% 时,YOLO 模型通常能够较好地满足机械臂无碰撞抓取的精度要求。反之,采摘成功率往往低于 70%,难以满足番茄等作物的商业化采摘需求。



(a) 采摘点检测 (b) 果实检测

图 6 采摘机器人视觉检测结果[79-83]

Fig. 6 Vision tasks for harvesting robots[79-83]

机器人采摘过程中相机与作物距离较近,果实尺寸对 YOLO 模型性能的影响相对有限。其核心挑战在于,YOLO 模型需在机械臂运动过程中适应相机视角,遮挡关系与光照条件的持续变化,实时准确地区分成熟与未成熟果实,同时对果梗等细小目标进行精确的分割与关键点定位。由表 6 可知,即使静态场景下检测精度较高,实际采摘成功率仍可能因动态遮挡与光照突变等因素显著下降。Ma 等[33]针对温室草莓采收任务改进 YOLO v8 模型,关键点定位精度(mAP50)达到 96%。然而,其抓取成功率在遮挡条件下从 92.6% 下降至 83.3%。刘芹等[83]针对茄子采摘任务提出 YOLO-CRC,检测精度达 94.1%,但因光照不足难以准确识别采摘点,抓取成功率仅为 70%。此外,在极端小目标与密集遮挡场景下,YOLO 模型的性能仍面临较大挑战。由表 6 可知,番茄茎秆分割任务[80-81]的分割与检测准确率相对较低,其茎秆直径仅 3~5 mm,在图像中仅占数个像素,且温室番茄通常采用密植模式,果实与枝叶相互遮挡,使分割难度进一步增加。未来针对遮挡问题,可结合几何先验知识训练补全模型,推断被遮挡部分的关键点位置;针对光照问题,根据光照条件自动对采摘目标进行补光,可有效降低模型对光照条件的依赖。除检测精度外,实时性同样是采摘任务落地部署的关键约束。采摘机器人需在动态环境

表 6　机器人采摘任务典型研究成果

Tab. 6　Typical research results of robotic harvesting tasks

作物/目标部位	因素	适配性	场景	任务	模型	性能参数/%	成功率/%	文献序号
草莓果实	果实与背景差异明显	高	温室	检测	YOLO v8	mAP50：93.7～96.0	88	[33]
芒果果柄		高	大田	分割与定位	YOLO v8	mIoU：81.9	92	[79]
大马士革玫瑰花朵		高	温室	检测	YOLO v9	mAP50：95.9～96.2		[54]
南瓜果实		高	大田	检测与采摘定位	YOLO v8	mAP50：91.8*	95	[82]
茄子采摘点		高	温室	定位	YOLO v8	mAP50：92.1～94.1		[83]
西兰花花球		高	大田	分割与姿态估计	YOLO v8	$mAP50_{box}$：96.1～97.3 $mAP50_{mask}$：96.1～97.3		[84]
葡萄采摘点	光照复杂	中	大田	定位	YOLO 11	mAP：87.4～93.6	91.3	[38]
樱桃番茄成熟果	目标较小、尺度变化大、遮挡严重	低	温室	检测与采摘点定位	YOLO v8	AP：82.5～85.3 mIoU：76.4～76.8		[85]
番茄植株/茎秆		低	温室	姿态估计与定位	YOLO	F1：87.8 mIoU：74.8		[80]
番茄茎秆		低	温室	分割与定位	YOLO v5	mAP：87.6～90.7* IoU：70.6～72.0		[81]

注：使用 mAP50 作为主要性能评价指标，若文献未给出 mAP50，则依次使用 mAP、AP 作为评价指标。对于分割任务，若文献未给出上述指标，则使用 mIoU 或 IoU 作为评价指标；对于关键点定位与姿态估计任务，若文献未给出上述指标，则使用 F1 或文献中给出的结果评价指标。成功率表示机器人在实际采摘或定位试验中的作业成功率，若文献未给出则留空。* 表示该结果为多场景或多条件试验下的平均值。

中持续进行目标检测与位姿估计，检测延迟将直接影响机械臂的响应速度与采摘效率。然而，机器人平台通常搭载算力有限的边缘计算设备，如 Jetson 系列开发板、树莓派或工控机等，难以支撑大参数量模型的实时推理。因此，模型轻量化成为采摘任务落地部署的关键环节，研究者需在压缩率与模型精度之间寻求平衡。

2　面向农业场景的 YOLO 模型优化策略

不同于传统目标检测任务，农业场景中的果实检测主要侧重于多尺度目标的精准检测，植物表型分析则要求对微小病斑和缺陷实现高灵敏度识别，而采摘机器人视觉则对关键点定位精度与实时性提出了更高要求。如图 7 所示，YOLO 系列模型采用高度模块化的架构设计，其骨干网络、特征金字塔、检测头等核心组件均可根据具体任务需求进行针对性调整。本节将相关优化策略归纳为结构优化、注意力机制和轻量化设计 3 类，并结合典型应用案例详细阐述各类方法的技术原理与适用场景。

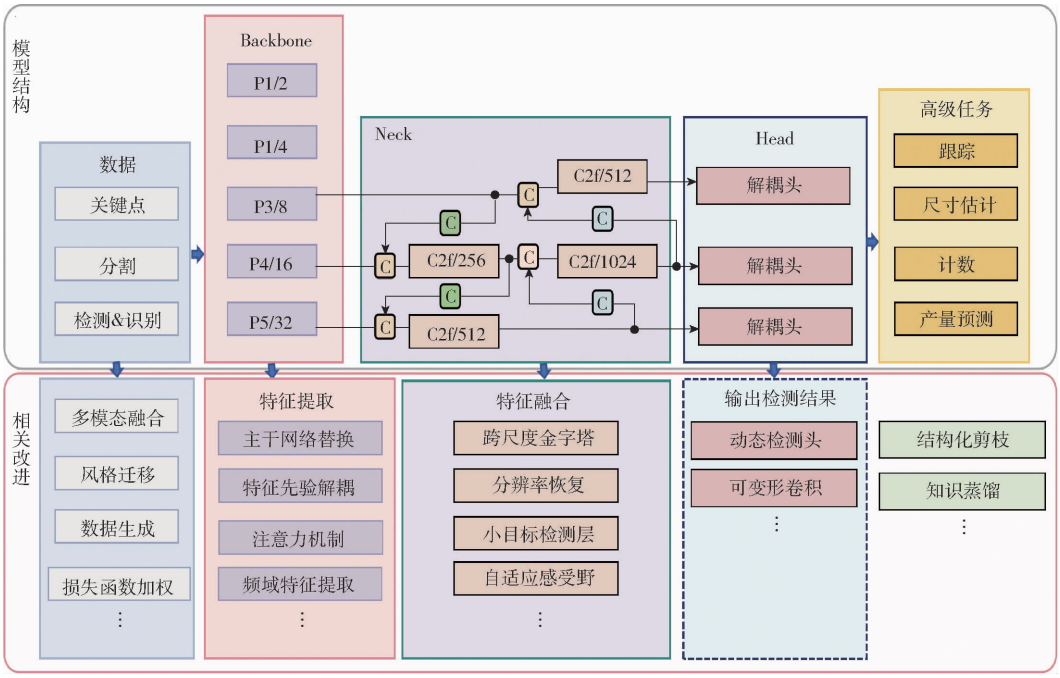
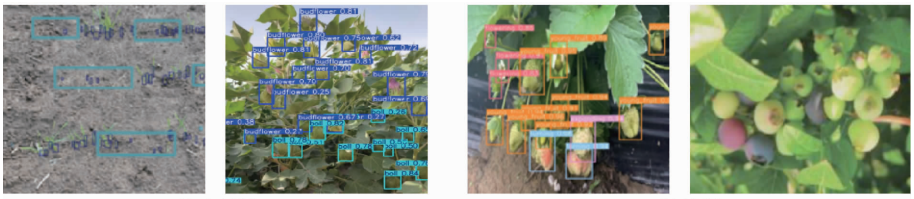


图 7　YOLO 结构与改进方式

Fig. 7　YOLO architecture and improvement strategies

2.1 面向小目标和遮挡的模型结构优化

如图 8 所示,遮挡和小目标问题是果实检测和病虫害识别的核心挑战,YOLO 模型采用多级下采样策略提取语义特征,浅层特征图中的边缘、纹理等空间细节信息随分辨率降低而逐步丢失,进而导致漏检与误检。以虫害识别为例,原始图像中 32 像素×32 像素的害虫在经过 P5 层 32 倍下采样后仅保留单一像素点,造成关键空间信息的丢失。类似地,被遮挡害虫仅有的局部可见区域也难以在低分辨率特征图中得到充分保留。由表 7 可知,研究者主要从细粒度特征提取、多尺度特征融合和形态自适应感知 3 个方向进行优化。



(a) 小目标 (b) 遮挡

图 8 面向小目标和遮挡的模型结构优化结果[22,27,51,57]

Fig. 8 Results of model architecture optimization for small object detection and occlusion handling[22,27,51,57]

表 7 细粒度特征提取优化典型研究成果

Tab. 7 Typical research results of fine-grained feature extraction optimization

改进类型	分辨率恢复	跨尺度增强	感受野/动态	对象	模型	性能参数/百分点	文献序号
小目标检测层	P2			柑橘果实	YOLO v8	AP: 2.2	[22]
	P2		DyHead	大田农作物害虫	YOLO v8	mAP50:5.1*	[35]
	P2		动态 C2f(C2f-Dynamic)	杭白菊	YOLO v8	mAP: 1.7*	[55]
	P2		多尺度空洞卷积(CAM)	辣椒种子	YOLO v8	mAP50:2.5*	[53]
	P2	BiFPN		小麦真菌	YOLO v8	mAP50:4.4	[85]
	Q1		SA(空间注意力)	樱桃番茄	YOLO v8	AP:2.8	[86]
特征重构与下采样	CARAFE	混合跨阶段卷积网络(VoV-GSCSP)		苹果果实	YOLO v8	mAP: 1.3*	[21]
	CGAFPN	多尺度特征聚合(MKFA)		板栗果实	YOLO 11	mAP: 5.5*	[34]
	CARAFE			小麦病害	YOLO v8	mAP50: 1.9*	[64]
	CARAFE	CDCP(DBB+PSA 极化注意力)		茄子采摘点	YOLO v8	mAP50: 1.5*	[83]
	DySample			油茶果实	YOLO 11	mAP50: 2.5*	[50]
	DySample			番茄叶病	YOLO 11	mAP50: 1.9*	[71]
	SPD-Conv	BiFPN		芒果果实	YOLO v8	mAP: 0.2	[79]
	SPD-Conv			番茄害虫	YOLO v5	mAP: 2.2	[70]

注:表中性能参数表示相较于基线模型的提升值。使用 mAP50 作为主要性能评价指标,若文献未给出 mAP50,则依次使用 mAP、AP 作为评价指标。* 表示该数值为文献中最终组合模型相较于基线模型的整体提升值,并非表中所列单一模块或模块组合的独立增益(下同)。

2.1.1 细粒度特征提取

针对小目标信息丢失问题,研究者主要提出构建小目标检测层和特征重构与无损下采样两种重构策略。构建小目标检测层的核心思路是在网络浅层(如 P2 层,4 倍下采样)引出额外分支或针对性设计超大尺度输出头,使 YOLO 模型在特征降维前完成检测,直接避免下采样造成的空间分辨率损失。然而,额外的预测分支会增加计算开销,特征重构利用可学习的算子替代传统插值或卷积,在采样过程中动态地保存或恢复局部细节信息,实现了从固定采样向动态特征重构的跨越。

构建小目标检测层是最基础且行之有效的策略,对极小目标的改善效果显著。文献[22]在 YOLO v8 模型的 P2 层构建小目标检测层,保留了小目标的边缘与纹理细节,在柑橘果实检测任务中 AP 提升 2.2 个百分点。文献[85]通过将 P2 小目标检测层与双向特征金字塔(Bidirectional feature pyramid network, BiFPN)组合,在小麦真菌检测任务中 mAP50 提升 4.4 个百分点。对于采摘点等更微小的目标,P2 层的分辨率仍显不足,可进一步将预测头构建于接近原图分辨率的 Q1 层。如 Qin 等[86]针对樱桃番茄采摘点定位任务,在 YOLO v8 模型的 Q1 构建小目标检测层,AP 提升 1.0 个百分点。然而,小目标检测层虽然恢复了特征图细节,但浅层特征往往缺乏语义信息容易引入背景噪声,需配合注意力机制或多尺度特征聚合进行抑制。如李

茂等[34]提出将内容引导的小目标检测层(Context-guided attention feature pyramid network, CGAFPN)与多尺度特征聚合模块(Multi-scale kernel feature aggregation, MKFA)相结合,通过融合浅层几何特征与深层语义特征,有效缓解了板栗果实因尺度差异导致的漏检问题,mAP 提升 5.3 个百分点。

在特征重构方面,现有研究主要从优化上采样和减少下采样信息损失两方面展开。优化的上采样方法主要解决传统的双线性插值上采样导致边缘特征模糊的问题,常见方法包括内容感知上采样算子 CARAFE 和轻量化动态上采样模块 DySample。CARAFE 能够根据局部语义信息动态生成采样核,避免了传统插值造成的特征退化,适用于病虫害检测、果实分级等对边缘细节敏感的任务。文献[31,64]分别在 YOLO v5 与 YOLO v8 中引入 CARAFE 模块,在金银花与小麦病害检测任务中,分别实现了 mAP 提升 0.9、2.2 个百分点。然而 CARAFE 计算开销较大,动态上采样模块(Dynamic upsampler, DySample),通过点采样策略实现高效特征重构,更适用于对实时性要求较高的场景。文献[50,72]在 YOLO 11 中引入了 DySample 模块,在油茶果实与番茄叶病检测中 mAP50 分别提升 0.1、0.7 个百分点。此外,将上采样算子与跨阶段特征融合网络深度结合可以进一步提升检测效果,高质量的上采样特征能够有效引导多尺度特征的精准对齐,减少多尺度融合过程中的特征错位。如文献[21,83]在 YOLO v8 中将 CARAFE 跨阶段特征融合网络深度结合,在苹果检测与茄子采摘点定位中 mAP50 分别提升 0.3、1.5 个百分点。

除优化上采样方法恢复特征外,从源头减少下采样过程中的信息流失同样关键。空间深度转换卷积(Space-to-depth convolution, SPD-Conv),通过将空间信息转换至通道维度避免下采样过程中小目标信息的丢失,适用于小目标检测、害虫识别等目标像素占比极低的任务。Tang 等[69]在 YOLO v5 中引入 SPD-Conv 模块,在番茄害虫检测任务中 mAP 提升 2.2 个百分点。

综上所述,细粒度特征提取策略旨在解决深层语义模糊与浅层细节匮乏之间的矛盾。在实际应用中,研究者在模块选型时应充分考虑目标尺度分布与场景复杂度。对于目标尺度相对均匀、背景较为简洁的场景,特征重构策略即可满足需求。而对于目标极小、存在密集遮挡或易与背景混淆的复杂场景,构建小目标检测层仍是更为直接有效的选择,但需要通过结合注意力机制抑制背景噪声。

2.1.2　多尺度特征融合

YOLO 采用特征金字塔结构(Feature pyramid networks, FPN),通过逐层相加与拼接实现多尺度特征融合。然而,这种融合方式易导致深层语义特征覆盖浅层的几何细节,在大小目标同框的农业场景中产生明显的尺度偏向问题。多尺度特征融合策略通过加权特征融合或选择式特征融合,确保大、小目标都能在特征金字塔中找到匹配的响应层级。

加权特征融合是解决多尺度问题的首选方案,双向特征金字塔通过双向跨尺度连接与可学习权重,使不同层级特征能够充分交互,成为农业场景中最主流的改进方案。庄煜等[38]在 YOLO 11 中引入 BiFPN,显著增强了模型对葡萄穗梗等细长目标的多尺度表征能力,mAP 提升 6.2 个百分点。针对高密度种植或严重遮挡的场景,研究者在 BiFPN 基础上通过增加层级间的连接密度或引入注意力机制,进一步优化融合结构以提升特征对齐能力[59]。

全连接式的 BiFPN 虽然通过密集的跨层连接实现充分的特征交互,但随着融合层级增加计算开销显著上升。因此,以 SFF、Slim-Neck 为代表的选择式融合通过按需激活 YOLO 模型的不同层次减少冗余计算,更契合边缘部署的应用需求,成为另一重要的优化方向。刘祥等[40]引入选择性特征融合机制(Selective feature fusion, SFF),利用高级特征作为注意力权重筛选低级特征中的关键信息,实现了多尺度特征的高效融合,在棉铃检测任务中 mAP 提升 7.1 个百分点。Zhang 等[36]引入轻量化颈部结构(Slim-Neck)并结合解耦检测头(Decoupled head),在保持较低计算开销的同时提升了燕麦叶病的多尺度检测性能,mAP50 提升 5.8 个百分点。但这类方法性能高度依赖尺度分布的先验假设,一旦实际场景与训练数据的尺度分布发生偏移,选择机制的有效性将大幅下降。

此外,由表 8 可知,优化融合拓扑结构的性能提升相对有限,EMA-HSFPN 进一步引入了可变形卷积和注意力机制,性能提升显著。这表明单纯优化特征融合拓扑结构效果存在上限,将跨尺度融合与可变形卷积、动态检测头等自适应机制相结合[25,37,81],能够在特征对齐的基础上进一步增强模型对目标形变与遮挡的适应能力。此外,在与 SPD-Conv 联合使用时,BiFPN 在 mAP50 的增益仅为 0.2 个百分点[79],表明在已具备无损下采样能力的网络中,加权特征融合的提升效果呈现递减趋势。因此,研究者在模块组合时应审慎评估各组件之间的功能重叠程度,避免冗余堆叠。

表 8 多尺度特征融合优化典型研究成果

Tab. 8 Typical research results of multi-scale feature fusion optimization

融合方式	跨尺度增强	感受野/动态	对象	模型	性能参数/百分点	文献序号
全连接特征融合	BiFPN		葡萄穗梗	YOLO 11	mAP: 6.2*	[38]
	BiFPN		柑橘果实	YOLO v5	AP: 1.2*	[28]
	EMA - HSFPN		黄瓜器官	YOLO 11	mAP50: 10.3	[59]
选择式特征融合	Slim - Neck	Decoupled Head	燕麦叶病	YOLO v7	mAP50: 7.6*	[36]
	上下文引导残差 FPN(CGRFPN)	大核可分离注意力(LSKM - CGRFPN)	豇豆害虫	YOLO 11	mAP50: 0.5*	[37]
	SFF		棉铃	YOLO v8	mAP: 10.1*	[40]

注:性能参数列所示数值均为相较于基线模型的性能提升值,而非改进模型的最终绝对性能。使用 mAP50 作为主要性能评价指标,若文献未给出 mAP50,则依次使用 mAP、AP 作为评价指标。

2.1.3 形态自适应感知

农业生物目标普遍存在非刚体形态特征,如黄瓜的自然弯曲、病斑的不规则扩散等,标准卷积固定的采样网格易引入背景噪声。形态自适应感知机制通过可变形卷积、动态检测头和大核注意力机制,帮助模型根据目标的形状、尺度与语义信息动态调整特征提取过程,替代固定的几何先验。如表 9 所示,形态自适应感知机制呈现出从不规则轮廓精准对齐到目标特征显著性增强再到全局上下文感知的递进发展方式。

表 9 形态自适应感知优化典型研究成果

Tab. 9 Typical research results of morphology-adaptive perception optimization

改进类型	感受野/动态	对象	模型	性能参数/百分点	文献序号
可变形卷积	DCNv2	柑橘果实	YOLO v8	mAP50:0.5	[23]
	DCNv4,TDADH	小麦幼苗	YOLO 11	mAP50:2.2	[57]
动态检测头	DCN - ELAN	草莓果实	YOLO v7	mAP50:2.2*	[47]
	DyHead	水稻病害	YOLO v5	mAP50: 0.5	[63]
	DyHead	黄瓜病害	YOLO v7	mAP50: 3.9*	[74]
	TADDH	棉花器官	YOLO 11	mAP: 0.81	[51]
大核分离注意力	LSKA,DyHead	蓝靛果果实	YOLO v8	mAP50: 4.4*	[76]
	CP_LSK	水稻分蘖	YOLO 11	mAP50box: 0.8* mAP50mask: 2.3*	[87]

注:表中性能参数均为相较于基线模型的性能提升值,而非改进模型的最终绝对性能。使用 mAP50 作为主要性能评价指标,若文献未给出 mAP50,则依次使用 mAP 作为评价指标。对于实例分割任务,若文献未给出上述指标,则使用 mAP50box、mAP50mask 作为评价指标。

可变形卷积(Deformable convolutional networks, DCN)是解决不规则轮廓精准对齐的首选策略,DCN 通过引入可学习的二维空间偏移量,允许采样点自适应对齐目标轮廓,适用于果实形态不规则但尺度相对单一的检测任务。Qu 等[23]在 YOLO v8 中引入 DCNv2,增强了对柑橘果实不规则轮廓的感知能力,mAP50 提升 0.5 个百分点。然而,可变形卷积的偏移量学习高度依赖训练数据的形态多样性,当目标形态超出训练分布时,对齐效果将显著下降。

可变形卷积仅调整空间采样位置,忽略了特征在通道和尺度维度上的相关性。动态检测头通过多维特征重校准策略弥补了这一不足,现有研究主要围绕特征重校准和任务级特征对齐两个方向展开。动态检测头(Dynamic head, DyHead)将多尺度特征金字塔建模为包含层级、空间与通道的三维张量,通过串联的注意力模块进行多维特征重校准,适用于目标尺度跨度大且形态多样的复杂农业场景。文献[63,74]验证了动态检测头在抑制不规则病斑背景噪声上的有效性,在水稻病斑与黄瓜霜霉病病斑检测任务中 mAP50 分别提升 0.5、2.3 个百分点。然而,DyHead 中多级注意力的堆叠会显著增加计算开销,在边缘部署时需兼顾动态自适应能力与轻量化设计。任务对齐动态检测头(Task-aligned dynamic detection head, TADDH)进一步引入任务对齐机制,使分类与回归分支能够聚焦各自最优的特征区域,适用于目标姿态变化显著且对定位精度要求较高的场景。Shao 等[51]在 YOLO 11 中引入任务对齐动态检测头 TADDH,实现棉花分类与回归特征的精准匹配,棉花器官识别 mAP 提升 0.81 个百分点。

在品质分级任务中,细微缺陷在像素层面上往往与背景非常相似。此时,动态机制无法有效处理这一问题。LSKA 在品质分级与实例分割任务中展现出独特优势[37,76,87],通过扩大感受野捕获更大范围的上下文信息,使模型能够区分目标与背景在纹理、颜色上的细微差异,适用于缺陷检测、品质分级

等需要精细纹理判别的任务。此外，由于大核卷积的参数量与感受野呈平方增长，因此在边缘端部署时需采用可分离或稀疏化策略加以缓解。

值得注意的是，上述形态自适应感知模块并非孤立使用，由表 9 可知，动态检测头适合与其他改进策略灵活搭配，实现从特征提取到检测输出的协同优化。例如，DCN 与 TDADH 的联合使用可同时增强模型的空间感知与任务判别能力[57]，与跨尺度融合模块联合使用可增强模型对不同尺度目标的同步能力，与 LSKA 结合可在扩大有效感受野的同时增强模型对目标细小缺陷的感知能力[76]。

2.2　环境噪声抑制

如图 9 所示，在非结构化农业环境中，光照波动、恶劣天气以及复杂的背景纹理均会引入噪声，导致目标信息在深层特征金字塔中被背景所稀释。针对这一问题，研究者通过引入空间-通道注意力机制重构清晰的特征表示，确保模型在复杂田间条件下的鲁棒性。如表 10 所示，根据农业场景的不同特点，注意力机制可分为空间注意力、通道注意力、坐标注意力和多尺度注意力 4 类。



(a) 背景与果实易混淆　　(b) 复杂光照条件

图 9　复杂环境示例[34,47,88]

Fig. 9　Examples of complex environmental scenarios[34,47,88]

表 10　典型注意力机制优化方法

Tab. 10　Statistics of attention mechanism optimization methods

注意力机制类型	对象	模型	性能参数/百分点	文献序号
空间注意力机制(CBAM)	小白菜虫害	YOLO v5	mAP50：12.9*	[66]
	大豆根瘤	YOLO v8	mAP：2.8*	[52]
	番茄害虫	YOLO v5	mAP：0.8	[69]
	水稻分蘖	YOLO 11	mAP50box：0.8*，mAP50mask：2.3*	[87]
通道注意力机制(SE)	黄瓜病害	YOLO v7	mAP50：3.9*	[74]
	双孢蘑菇	YOLO v8	mAP50:95：2.3	[89]
坐标注意力(CA)	棉铃	YOLO v8	mAP：10.1*	[40]
	葡萄穗梗	YOLO 11	mAP：2.6*	[38]
多尺度注意力(EMA)	苹果花序	YOLO v7	mAP：3.4*	[47]
	马铃薯	YOLO v8	mAP50：2.7	[90]
	黄瓜果实	YOLO 11	mAP50：10.3*	[59]
全局注意力(GAM)	玉米幼苗	YOLO v8	mAP：3.8*	[46]
	番茄果柄	YOLO v8	mAP50：1.4*	[25]
	苹果果实	YOLO v8	mAP50：0.1	[29]
洗牌注意力机制(SA)	水稻病害	YOLO v5	mAP：0.5	[63]
	草莓果实	YOLO v7	mAP：0.7	[48]
多通道空间注意力(C2f_MSCA)，四维动态注意力(C2f_MDAC)	欧李果实	YOLO v8	mAP50:13.9*	[41]
矩形自校准注意力(RCM+RCMV)	草莓	YOLO 11	mAP50：3.8	[91]
多路径坐标注意力(MPCA)	苹果分割	YOLO v8	mAP：0.8*	[62]
三重注意力(Triplet Attention)	苹果叶片	YOLO v8	mAP50：2.2*	[71]
双层路由注意力(BRA)	菌头	YOLO 11	Pose-P：1.9*，Box-P：0.9*	[78]
上下文聚合注意力(CA*)	樱桃果实	YOLO v8	IoU_{steam}：0.2*，$mIoU_{steam}$：2.7*	[67]

注：性能参数均为相较于基线模型的性能提升值，而非改进模型的最终绝对性能。使用 mAP50 作为主要性能评价指标，若文献未给出 mAP50，则依次使用 mAP 作为评价指标；对于实例分割、姿态估计或关键点任务，若文献未给出上述指标，则使用 mAP50box、mAP50mask、mIoU、IoU、Pose-P、Box-P 或文献中主要结果指标作为评价指标。

空间注意力主要用于目标空间分布复杂的场景，如茂密植株中的果实定位、细长茎秆的精确分割。当背景复杂时，应优先使用 CBAM 等空间注意力，其通过串联通道注意力与空间注意力模块，依次筛选关键特征与核心区域，成为农业场景中应用最广泛的注意力机制。文献[56,63,69,87]广泛验证

了其在抑制枝叶干扰方面的有效性,在不同作物检测中 mAP 取得 0.2 ~ 3.2 个百分点的提升。然而,在目标与背景纹理高度相似时,如绿色未成熟果实与绿叶、土壤中的根瘤与土块,空间注意力难以有效区分,需要通过特征通道加权增强细粒度判别能力。SE 通过全局平均池化学习通道间的依赖关系,并经过全连接层对通道进行加权。文献[74,89]在 YOLO 模型中引入 SE,提升了模型对病斑颜色与菌盖纹理等细粒度特征的提取能力,mAP 分别提升 0.2、0.5、2.3 个百分点。由表 10 可知,SE 在双孢菇检测中的精度提升最为显著,表明其性能增益主要取决于目标本身的纹理复杂度而非环境因素,与空间注意力的场景依赖特性形成互补。

对于机器人采摘等需精确空间坐标的任务,如采摘点定位、果柄抓取,传统的空间注意力往往会因下采样而丢失位置信息。坐标注意力(Coordinate attention, CA)在通道注意力基础上引入位置编码,通过将空间信息分解为水平与垂直两个方向,在抑制背景的同时保留了精确的空间坐标。文献[38,40]在葡萄穗梗定位与棉铃检测中引入 CA,增强了模型对细长目标的空间感知能力,mAP 分别提升 7.5、2.6 个百分点。这种方法能够提供精确的位置信息,但其效果高度依赖目标的空间分布规律,对形态不规则或姿态多变的目标表现受限。

针对目标尺度差异显著的场景,如同一视野内大小果实共存、不同生长阶段的花序检测,多尺度注意力能够在不同感受野范围内并行提取特征并自适应融合。高效多尺度注意力(Efficient multi-scale attention, EMA)通过并行多分支结构处理尺度差异。文献[47,59,90]在苹果花序、马铃薯芽眼及黄瓜果实检测任务中验证了其能有效应对从花蕾到成熟果实的巨大形态差异,检测精度提升 2.7 ~ 11.8 个百分点。针对光照条件复杂多变的场景,如逆光、阴影遮挡、晨昏光照不均,全局注意力模块(Global attention mechanism, GAM)通过三维注意力机制同时建模通道、空间与跨通道交互,捕获全局上下文信息,补偿局部光照畸变。文献[20,29,63]在苹果花序、黄瓜果实及马铃薯芽眼检测任务中引入 GAM,增强了模型对全局场景的理解能力与抗背景干扰能力,检测精度分别提升 0.1、1.4、0.2 个百分点。然而,并行结构显著增加了计算开销,在边缘设备部署时需注意平衡精度与效率。

综上所述,各类注意力机制在农业场景中的适用性呈现明显的任务依赖特性。当基础模型性能已接近上限时,引入注意力机制难以带来显著提升,反而可能因模型复杂度增加而引发过拟合问题。因此,研究者在选择模块时应充分评估影响模型性能的核心因素,避免引入复杂注意力机制导致计算冗余。同时,注意力模块的组合设计展现出超越单一模块的潜力[41,43,78],表明不同类型注意力机制在特征空间中的作用具有互补性。

2.3 模型轻量化

农业边缘设备通常受算力和存储资源的制约,而深度神经网络中并非所有参数均对最终性能产生实质性贡献。模型轻量化的核心目标是在确保关键特征及语义信息完整的前提下,降低模型参数量和计算开销。如表 11 所示,目前的优化策略主要通过骨干网络替换、算子重构和网络结构优化剔除冗余参数并优化计算单元,在尽量保持精度的前提下提高模型部署效率。

骨干网络替换是最直接的轻量化策略,常用的轻量化骨干包括 Ghost 系列、ShuffleNet 系列、EfficientNet 及 MobileNet 系列等,各骨干在压缩效率与精度保持方面各有侧重。Ghost 系列主干通过廉价线性变换生成特征图,具有参数压缩比高、精度损失小的特点,是应用最广泛的轻量化方法。文献[28,63]利用 GhostNet 替换 YOLO 主干,在柑橘果实与水稻病害检测任务中将参数量压缩至原模型的 50%。在此基础上,刘祥等[40]进一步引入重参数化思想,在 YOLO v8 中引入重参数化跨阶段特征融合结构(RepGhostCSPELAN),通过训练多分支、推理单分支的策略,在棉铃检测中将模型计算量从 8.1×10^9 降至 5.5×10^9。

在边缘部署场景下,不同任务对模型的压缩需求各异,需针对性地选择骨干网络替换方案。ShuffleNet 系列通过通道混洗与分组卷积大幅减少参数量,适用于存储资源严格受限的场景。Ma 等[29]采用 ShuffleNetv2 替换 YOLO v8 的骨干网络并结合幽灵 C2f 模块(C2f - Ghost),在苹果果实检测任务中将参数量压缩 60% 以上。EfficientNet 通过同步缩放网络深度、宽度与输入分辨率,实现计算量与参数量的均衡压缩,适用于算力与存储均需同步削减的场景。范天浩等[31]引入 EfficientNet 的复合缩放策略,通过缩减网络深度、宽度与分辨率,在金银花识别任务中将模型计算量从 1.58×10^{10} 降至 7.3×10^9。MobileNet 系列基于深度可分离卷积降低计算量,可有效提升推理速度。文献[38,88]分别在葡萄穗梗与蓝莓检测中引入 MobileNetV3 和 MobileNetV4,参数压缩比均低于 15%,表明 MobileNet 系列相较于 Ghost 系列在参数压缩方面优势有限,更适用于对推理延迟敏感而非参数量受限的应用场景。

表 11　模型轻量化典型方法

Tab. 11　Typical model lightweighting methods

重构	对象	模型	浮点计算量		参数量		文献序号
			改进前	改进后	改进前	改进后	
GhostNet	柑橘果实	YOLO v5	1.58×10^{10}	8.30×10^{9}	7.97×10^{6}	3.57×10^{6}	[28]
ShuffleNetv2，C2f－Ghost	苹果果实	YOLO v8	8.1×10^{9}	3.9×10^{9}	3.00×10^{6}	1.18×10^{6}	[29]
MobileNetV3	蓝莓果实	YOLO v8			7.9×10^{6}	7.00×10^{6}	[88]
EfficientNet	金银花	YOLO v5	1.58×10^{10}	7.30×10^{9}	7.01×10^{6}	3.76×10^{6}	[31]
ReXNet	燕麦叶病	YOLO v7	1.32×10^{10}	1.13×10^{10}			[36]
MobileNetV4	葡萄穗梗	YOLO 11			2.29×10^{6}	2.22×10^{6}	[38]
GhostNet	水稻病害	YOLO v5	1.6×10^{10}	8.2×10^{9}	5.73×10^{6}	3.01×10^{6}	[63]
轻量级 CPU 卷积神经网络（PP－LCNet）	小麦病害	YOLO v8	8.1×10^{9}	5.2×10^{9}	3×10^{6}	1.8×10^{6}	[64]
FasterNet	双孢蘑菇	YOLO v8	1.2×10^{9}	9×10^{8}	3.3×10^{6}	2.0×10^{6}	[89]
网络缩放，PDWFasterNet，DWSConv	苹果果实	YOLO v8	8.2×10^{9}	2.3×10^{9}	3×10^{6}	7×10^{5}	[26]
高效多尺度注意力融合 FasterNet 模块（C2f－FE），DWConv	大田农作物害虫	YOLO v8	8.1×10^{9}	6×10^{9}	3.0×10^{6}	2.1×10^{6}	[35]
分组洗牌高效层聚合网络（GS－ELAN）	草莓幼苗	YOLO v7	1.19×10^{11}	1.10×10^{11}	3.78×10^{7}	3.45×10^{7}	[39]
GhostConv，RepGhostCSPELAN	棉铃	YOLO v8	8.1×10^{9}	5.5×10^{9}	3.2×10^{6}	1.9×10^{6}	[40]
DCN－ELAN	草莓果实	YOLO v7	1.03×10^{11}	9.95×10^{10}	3.65×10^{7}	3.24×10^{7}	[48]
C2f－Faster	苹果叶片	YOLO v8	8.1×10^{9}	6.4×10^{9}	3.0×10^{6}	2.3×10^{6}	[71]
RepViT 模块（C2f_RVB）	茄子采摘点	YOLO v8	1.20×10^{10}	1.13×10^{10}			[83]
ADown	番茄果实	YOLO 11	6.5×10^{9}	5.6×10^{9}	2.60×10^{6}	2.05×10^{6}	[30]
ADown	番茄叶部病害	YOLO 11	6.3×10^{9}	5.1×10^{9}	2.58×10^{6}	2.09×10^{6}	[72]
ADown	黄瓜果实	YOLO 11	1.02×10^{10}	9.20×10^{9}	3.4×10^{6}	2.7×10^{6}	[59]
VoV－GSCSP	苹果果实	YOLO v8	2.84×10^{10}	2.54×10^{10}	1.11×10^{7}	1.05×10^{7}	[21]
CCFM	番茄果柄	YOLO v8			3.0×10^{6}	2.3×10^{6}	[25]
组洗牌轻量颈部（GSConv－Slim－neck）	苹果花序	YOLO v7			6.0×10^{6}	5.6×10^{6}	[47]
Ghost 液态动态卷积（LGC）	油茶果实	YOLO 11	5.8×10^{9}	5.1×10^{9}			[50]
双重卷积双分支轻量化（EDLAN）	棉花器官	YOLO 11	6.9×10^{9}	8.4×10^{9}	2.6×10^{6}	2.1×10^{6}	[51]
DCNv4，TDADH	小麦幼苗	YOLO 11	6.3×10^{9}	7.8×10^{9}	2.5×10^{6}	2.1×10^{6}	[57]

算子重构是通过高效计算单元替代标准卷积或原有模块，在尽量保持检测精度的前提下降低计算量与内存访问开销，是模型轻量化的重要途径。常见方法包括部分卷积（Partial conv，PConv）和深度可分离卷积（Depthwise separable convolution，DSConv）。PConv 仅对部分通道进行空间卷积，适用于兼顾精度与速度的场景。Chen 等[44]在 YOLO v5 中引入 PConv 用于葡萄果实检测，推理速度从136 ms 优化至 90 ms。DSConv 将标准卷积分解为深度卷积和逐点卷积，更适用于对实时性要求较高的任务。此外，上述算子可与其他模块组合使用以获得更高的压缩比。文献[71，89]将 PConv 与 FasterNet 模块结合构建 C2f－Faster 模块，应用于苹果叶片病害与双孢菇表型检测，参数量压缩至原有的 88% 以下。Liu 等[26]进一步结合 PConv 与 DSConv，构建 PDWFasterNet 模块替换 YOLO v8 的 C2f 模块，兼顾部分卷积的精度保持优势与深度可分离卷积的高压缩能力，在苹果果实检测中将参数量从 3.0×10^{6} 压缩至 7.9×10^{5}，实现约 73% 的参数压缩。总体来看，单一算子的压缩能力存在一定上限，不同轻量化算子的组合设计更有利于在检测精度与推理效率之间取得平衡。

网络结构优化则从模型的拓扑连接层面进行轻量化，通过重新设计颈部特征融合路径或下采样方式精简冗余计算，常见方法包括 Slim-neck 和自适应下采样（Adaptive downsampling，ADown）。Slim-neck 通过混合卷积（Grouped shuffle convolution，GSConv）与基于 VoVNet 的分组洗牌跨阶段部分结构（VoVNet-based Group－Shuffle Cross Stage Partial，VoV－GSCSP）减少网络颈部特征融合层的计算冗余，适用于对模型整体规模有大幅压缩需求的场景。Wu 等[21]将其应用于苹果果实检测，将模型计算量从 2.85×10^{10} 压缩至 3.4×10^{9}，压缩比超过 88%。ADown 通过可学习的下采样策略替代传统池化，在降低分辨率的同时保留更多有效特征，适合作为基础轻量化组件与其他策略叠加使用。文献[30，72]分别在番茄果实与番茄叶病检测中引入 ADown 模块，参数量降低近 20%。

综上所述，不同轻量化策略在压缩机制、适用场

景与性能边界上存在显著差异。骨干网络替换通过整体架构重构实现最大幅度的参数与计算量压缩，适用于模型规模与目标平台算力差距悬殊的场景。算子重构在模块级别以高效计算单元替代标准卷积，在有限的精度损失下实现推理加速，且不同算子间的互补组合可进一步突破单一算子的压缩上限。网络结构优化从拓扑连接层面精简冗余路径，对模型整体架构改动最小，适用于对检测精度敏感、不宜大幅调整网络结构的任务。在实际部署中，上述策略可根据平台约束组合使用，以实现精度与效率的最优权衡。

2.4 边缘设备部署

边缘部署能够将目标检测与识别功能直接下沉至田间终端，在降低通信延迟的同时提高系统实时性，并在一定程度上增强数据处理的本地化能力。目前，低功耗边缘设备，如嵌入式开发板和物联网节点，已广泛用于田间自动化监测。这些设备计算资源有限，因此需要对模型进行轻量化改进并根据具体硬件进行针对性优化[12]。

边缘部署平台分为 GPU、ARM 开发板、x86 平台和移动端 4 类，不同平台在算力、功耗和适用场景上存在显著差异。GPU 加速板以 NVIDIA Jetson 系列为代表，具备高算力、低功耗的特点，适用于对实时性和算力要求较高的场景，如机器人采摘任务，可在运动过程中同步完成果实检测、成熟度判别与采摘点定位[25,34,48,53,83]。ARM 开发板包括 Raspberry Pi、ROCK 和 OrangePi，功耗较低且成本低廉，适用于田间物联网节点的长时无人值守监测，当检测到病斑或害虫时即可触发本地警告[30,68]。x86 工控平台以工控机和 Intel NUC 为代表，具备工业级稳定性与成熟的软硬件生态，且可配备 GPU 加速卡实现高帧率推理，适用于对吞吐量与可靠性要求极高的农产品分选流水线场景[30,85]。移动终端以 Android 手机为代表，具备便携性强、普及率高的特点，为农户现场即时诊断提供低成本解决方案[28,65,70]。

此外，如表 11、12 所示，当前关于轻量化改进的研究明显多于真实部署验证研究，说明农业视觉模型从实验室精度提升到生产环境落地之间仍存在显著工程化鸿沟。现有研究通常以浮点运算量和参数量评价轻量化程度，但理论指标与实际推理速度并非严格线性对应，内存带宽、算子调度和缓存命中率等硬件因素同样会显著影响部署性能。因此，后续研究应更加重视模型在真实边缘设备上的系统性验证，重点关注推理速度、功耗、稳定性和连续运行表现，建立从算法设计到硬件落地的完整评估体系。

表 12 模型部署典型成果

Tab. 12 Typical results of model deployment

部署设备	对象	模型	文献序号
Jetson Nano B01	辣椒种子	YOLO v8	[53]
Jetson Nano	板栗果实	YOLO 11	[34]
Jetson Nano	茄子采摘点	YOLO v8	[83]
Jetson Xavier NX	草莓果实	YOLO v7	[48]
Jetson Orin Nano	番茄果柄	YOLO v8	[25]
Raspberry Pi 4B	温室昆虫	YOLO v8	[68]
工控机	小麦真菌	YOLO v8	[85]
Intel NUC	番茄果实	YOLO 11	[30]
OrangePi Plus5	番茄果实	YOLO 11	[30]
Android 手机	水稻害虫	YOLO v8	[65]
Android 手机	柑橘果实	YOLO v5	[28]
Android 手机	葡萄叶病害	YOLO v8	[70]

3 YOLO 算法在种植农业中应用分析

对影响 YOLO 模型性能的关键因素、数据需求、模型版本适用性及优化策略选择进行综合分析。如图 10 所示，农业场景中的 YOLO 模型设计不应局限于单一模块改进或版本迭代，而应围绕任务瓶颈、场景特征与部署需求进行系统性决策，从而为后续模型选型、方法组合与工程落地提供理论参考。

3.1 种植领域视觉任务性能主导因素

本研究将种植业视觉任务分为果实检测任务、农情分析任务和采摘机器人视觉任务 3 类，基于第 1 节中对 3 类任务的文献梳理与适配性评估可知，YOLO 模型性能上限的关键因素并非模型版本或模块堆叠方式，而是目标在场景中可被稳定利用的特征密度。特征密度是指模型在给定图像或视频中能够提取并用于判别的有效视觉信息总量，其本质上由目标尺寸、目标与背景的对比度以及环境干扰强度共同决定。目标尺寸越大，模型能够获取的有效像素与语义信息越丰富；目标与背景对比度越高，特征可分性越强；而遮挡、复杂光照和背景纹理干扰则会稀释有效信息、引入噪声，从而降低特征密度。

在果实检测类任务中，成熟果实通常尺寸较大且与背景色差明显，特征密度高，基础 YOLO 模型即可达到较高的检测精度。随着目标尺寸缩小或背景对比度降低，特征密度下降，模型性能随之显著退化。在农情分析任务中，病斑与害虫目标尺寸更小，且类间视觉差异细微，特征密度受限于类间可判别特征的匮乏，对模型的细粒度感知能力提出了更高要求。在采摘机器人视觉任务中，虽然相机与目标距离较近使得尺寸对模型的影响相对有限，但动态遮挡、光照变化等因素导致特征密度在时序上持续

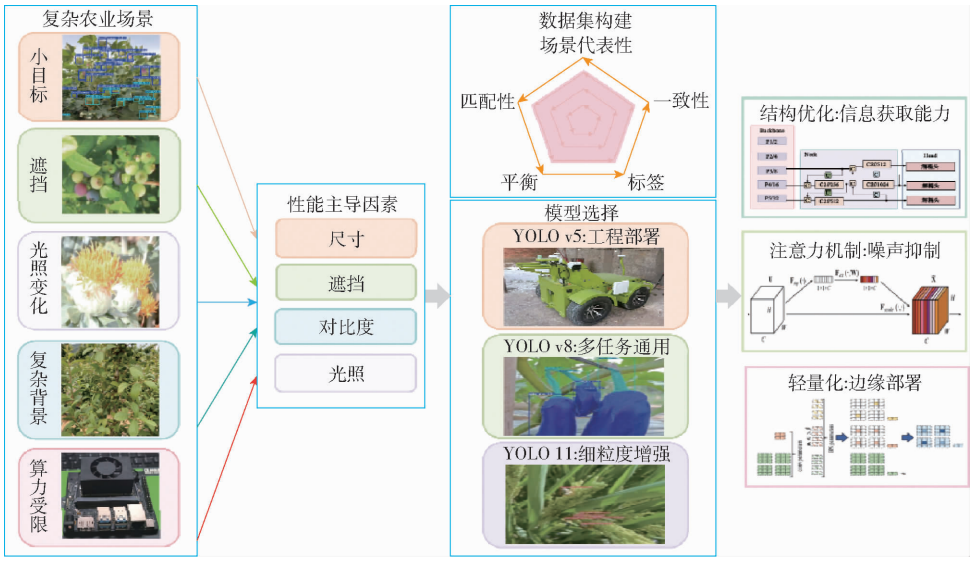


图 10　农业视觉任务分析与优化流程图

Fig. 10　Lowchart of agricultural visual task analysis and optimization

波动,模型需要在动态环境中维持稳定的高置信度输出。

因此,在任务规划与模型设计阶段,应首先评估目标的特征密度水平。当特征密度充足时,研究重点应由单纯追求检测精度转向轻量化、泛化性和部署效率优化;当特征密度不足时,则应优先从数据采集、数据增强和结构优化 3 个层面提升可用特征的数量与质量。为了量化不同任务中可利用视觉信息的差异,图 11 给出了适配等级与有效视觉线索密度得分的对应关系(图中散点表示各文献样本,散点的颜色与形状用于区分任务类别)。该得分依据文献中对目标尺寸、背景对比度、环境可控性、遮挡程度、光照复杂度和背景混淆程度等因素的描述进行规则化编码获得。图 11 显示,高适配案例整体分布于较高得分区间,而低适配案例集中于低分区间,表明 YOLO 模型的检测性能主要受限于目标可分性和场景干扰强度。

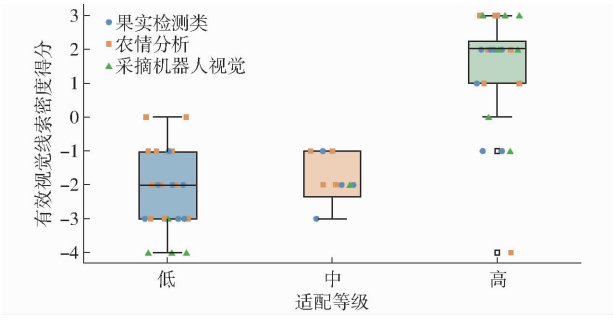


图 11　适配等级与有效视觉线索密度得分的关系

Fig. 11　Relationship between adaptation level and effective visual cue density score

3.2　高质量数据需求

YOLO 模型属于典型的数据驱动方法,其性能上限不仅取决于样本规模,更取决于数据质量与数据多样性。对于种植领域任务而言,高质量数据不应仅理解为样本数量充足,而应同时具备场景代表性、采集一致性、标签可靠性、类别平衡性和部署匹配性。相较于单纯的模型结构改进,满足上述要求的数据集往往更能从根本上提升模型的检测精度与跨场景泛化能力。其中,场景代表性是指训练数据应覆盖实际生产中可能出现的天气、光照、生长阶段、种植密度和背景条件;采集一致性强调相机位置、拍摄距离、视角、分辨率和曝光条件的规范化;标签可靠性要求标注边界清晰、类别定义明确,并尽量降低主观误差;类别平衡性用于避免模型对高频类别过拟合而忽视少样本类别;部署匹配性则要求训练数据分布尽可能接近最终应用环境,以减小训练与部署之间的域偏移。

在数据采集阶段,拍摄条件的一致性与规范性会直接影响模型学习到的特征分布。前文不同研究对同类任务的比较表明,即使针对同一检测对象和相近模型版本,不同研究之间的精度差异仍可能较为显著,这在很大程度上与拍摄距离、相机角度、背景复杂度以及光照条件等数据源差异有关。因此,研究者应尽可能规范相机布设方式,控制图像采集距离和视角范围,并提高照明条件的稳定性,从源头减少噪声输入。

在数据标注阶段,标签质量往往比样本规模更容易成为瓶颈。尤其是在病虫害识别、品质分级等任务中,早期病斑边界模糊、不同病害外观相似、轻微缺陷等级区分主观性强,这些因素都会导致标注一致性下降,进而引入标签噪声并削弱模型判别能力。今后的研究应更加重视高质量数据集构建机制,例如采用专家复核、分层标注规范、多轮一致性

校验以及面向难样本的精细标注策略,以提高训练数据的可信度和泛化价值。总体来看,高质量数据并非模型优化之外的附属条件,而是提升农业视觉模型性能最基础、最有效的前提。

3.3 模型版本选择策略

YOLO 系列模型经过持续迭代,不同版本在检测精度、推理效率、任务支持能力与工程生态方面呈现出不同特点。对于种植领域研究而言,模型选型不宜简单理解为新版本对旧版本的线性替代,而应结合任务复杂度、部署平台和工程成熟度进行综合判断。

YOLO v5 具有部署工具链完善、边缘平台兼容性强等优势,在移动端、嵌入式平台和田间物联网节点等场景中仍具有较高实用价值。对于以目标检测为主、对系统稳定性要求较高的应用,如田间监测、移动诊断和轻量化实时识别,YOLO v5 仍是较优选择。YOLO v8 通过解耦检测头、Anchor-free 设计以及统一的多任务框架,在目标检测、实例分割和关键点检测等任务中均表现出较强的适应能力,因此特别适合采摘机器人视觉等需要在同一体系下同时完成目标识别、掩码提取和关键点定位的复杂任务场景。YOLO 11 则在轻量化结构设计和细粒度特征提取能力方面进一步增强,更适用于小目标、复杂背景和高密度遮挡等对特征表达要求更高的任务。

从种植领域近年的研究趋势来看,YOLO v8 在农业视觉研究中使用较为广泛,主要原因在于其在精度、多任务统一性和社区生态之间取得了较好的平衡;YOLO v5 则凭借成熟的部署生态,在实际工程落地中仍保持重要地位;YOLO 11 虽然出现时间较晚,但在轻量化和复杂场景感知方面展现出较强潜力。由此可见,种植领域中不同 YOLO 版本正在形成“工程部署型”“多任务通用型”和“细粒度增强型”并存的格局。未来随着 YOLO 11 相关部署生态的逐步完善,其在农业边缘设备和复杂场景中的应用比例有望进一步提升,但短期内不同版本仍将长期共存,并服务于不同层次的应用需求。

3.4 模型优化策略

本文将 YOLO 模型的优化策略归纳为结构优化、注意力机制与轻量化设计 3 类,三者分别对应“信息获取—噪声抑制—资源约束”3 个层面的问题,存在明确的协同与互斥关系。研究者需根据具体任务场景针对性地选择优化策略。

对于果实检测类任务,当目标尺寸较大、轮廓清晰且与背景差异明显时,基础模型通常已能获得较高检测精度,此时继续追求复杂结构改进的边际收益有限,研究重点应转向模型轻量化、跨场景泛化和工程部署优化。当目标尺寸较小或存在密集遮挡时,模型的主要瓶颈在于浅层细节信息不足与多尺度特征表达能力不强,因此应优先采用小目标检测层、特征重构和多尺度融合等结构优化策略,以增强对弱特征目标的识别能力。对于农情分析任务,其核心难点更多体现在病斑、害虫和缺陷之间的细微差异识别,因此应优先引入通道注意力、大核注意力或多尺度注意力等方法,以增强对纹理、颜色和局部异常特征的筛选能力;但当目标在可见光波段下本身缺乏足够判别信息时,单纯依赖模块改进的提升空间往往有限,更需要借助多模态数据或高质量数据增强策略。对于采摘机器人视觉任务,模型需要同时满足动态场景下的稳定感知、高精度定位以及边缘端实时运行等多重要求,因此坐标注意力、动态检测头、轻量化骨干与高效颈部结构的组合往往更具应用价值。图 12 显示(图中散点表示各文献相对基线模型的性能增益样本,每个点对应一项研究成果;不同箱体颜色仅用于区分优化策略类别,散点颜色反映样本点的重叠密度,颜色越深表示该取值附近聚集的样本越多),不同优化策略的性能增益分布存在明显差异,其中多尺度特征融合的中位收益相对更高,注意力机制的波动范围更大,说明不同方法对应的性能瓶颈并不相同。

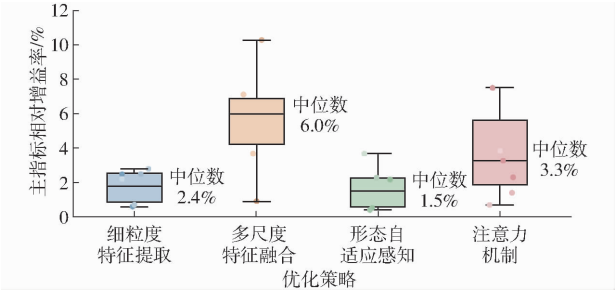


图 12 不同优化策略的性能增益分布

Fig. 12 Distribution of performance gains across different optimization strategies

需要指出的是,不同模块之间既存在功能互补,也存在边际收益递减。若多个模块作用于不同维度,其联合使用往往能够实现协同增强,例如“小目标检测层 + 空间注意力”可形成“细节恢复—噪声抑制”的互补,“动态检测头 + 可变形卷积”可形成“多维重校准—空间对齐”的联动效果。相反,若多个模块功能相近或作用路径高度重叠,则继续叠加往往只能带来有限增益,甚至引入额外计算负担和训练不稳定问题。因此,农业场景中的模型改进不应追求模块数量的增加,而应建立在对性能瓶颈准确诊断的基础上,优先选择作用维度互补、能够回应任务核心矛盾的改进方案。

4　展望

基于上述分析，农业场景中 YOLO 模型的未来发展应重点从感知能力突破与泛化能力提升两个方向推进。在感知能力方面，单一 RGB 图像所提供的信息维度有限，这已成为制约复杂农业场景检测性能进一步提升的重要瓶颈。对于密集遮挡场景，深度传感器能够利用几何距离信息区分前景与背景，从而提高目标分离与空间定位能力；对于低对比度病虫害识别和品质分级任务，多光谱、高光谱和近红外成像能够提供可见光之外的反射率信息，从而增强对病变组织、成熟度差异和内部品质属性的表征能力；对于挂果量统计和机器人采摘等场景，多视角成像能够通过跨视角几何映射恢复被遮挡目标的完整形态，降低漏检率并提升空间定位精度。未来研究的关键，不仅在于引入多模态信息，更在于如何在保持 YOLO 实时性优势的前提下实现多模态信息的高效融合。

在泛化能力方面，农业场景天然具有跨地区、跨季节、跨品种和跨管理方式的显著差异，模型在实际部署中不可避免地面临严重域偏移问题。仅依赖单一场景、单一季节采集的数据进行训练，往往难以支持跨环境稳定应用。未来可进一步加强域适应、域泛化与迁移学习在农业视觉中的应用，通过风格迁移和特征对齐减少训练域与目标域之间的表观差异；结合教师—学生架构和知识蒸馏，在提升轻量化模型泛化能力的同时保留大模型的鲁棒性；利用扩散模型、基础分割模型和自动数据生成方法构建更丰富的合成样本，以缓解稀有类别样本不足、类别不平衡和困难样本覆盖不足等问题。

总体而言，YOLO 在种植领域的研究已由“单一检测精度提升”逐步转向“多任务感知、跨场景泛化与边缘部署协同优化”的综合发展阶段。未来，只有将多模态感知、高质量数据构建、场景化模型设计与真实部署验证深度结合，才能推动 YOLO 模型从实验室研究迈向复杂农业生产环境的规模化应用。

参　考　文　献

[1] 国家统计局. 农林牧渔业总产值[EB/OL]. [2026－04－26]. https://data.stats.gov.cn/easyquery.htm?cn=C01&zb=A0D04&sj=2024.

[2] 农业农村部. 农业农村部关于大力发展智慧农业的指导意见[EB/OL]. (2024－10－23)[2026－04－26]. https://www.moa.gov.cn/govpublic/SCYJJXXS/202410/t20241025_6465040.htm.

[3] 农业农村部. 农业农村部关于印发《全国智慧农业行动计划(2024—2028 年)》的通知[EB/OL]. (2024－10－23)[2026－04－26]. https://www.moa.gov.cn/govpublic/SCYJJXXS/202410/t20241025_6465041.htm.

[4] 中共中央，国务院. 中共中央 国务院关于锚定农业农村现代化 扎实推进乡村全面振兴的意见[EB/OL]. (2026－01－03)[2026－04－26]. https://www.drc.gov.cn/DocView.aspx?chnid=378&docid=2909616&leafid=1333.

[5] 农业农村部. 关于落实《中共中央国务院关于锚定农业农村现代化扎实推进乡村全面振兴的意见》的实施意见[EB/OL]. (2026－01－19)[2026－04－26]. https://nyt.hubei.gov.cn/bmdt/ztzl/zxzt/fzxxtjjj/202602/t20260210_5874644.shtml.

[6] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779－788.

[7] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[PP/OL]. Vl. arXiv(2020－04－23)[2026－04－26]. https://doi.org/10.48550/arXiv.2004.10934.

[8] Khanam R, Hussain M. What is YOLOv5: a deep look into the internal features of the popular object detector[PP/OL]. Vl. arXiv(2024－07－30)[2026－04－26]. https://doi.org/10.48550/arXiv.2407.20892.

[9] Varghese R, Sambath M. YOLOv8: a novel object detection algorithm with enhanced performance and robustness[C]// 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). Piscataway: IEEE, 2024: 1－6.

[10] Khanam R, Hussain M. YOLO11: an overview of the key architectural enhancements[PP/OL]. Vl. arXiv(2024－10－23)[2026－04－26]. https://doi.org/10.48550/arXiv.2410.17725.

[11] Tian Y, Ye Q, Doermann D. YOLOv12: attention-centric real-time object detectors[J]. Advances in Neural Information Processing Systems, 2026,38:78433－78457.

[12] Li C, Li L, Jiang H, et al. YOLOv6: a single-stage object detection framework for industrial applications[PP/OL]. Vl. arXiv(2022－09－07)[2026－04－26]. https://doi.org/10.48550/arXiv.2209.02976.

[13] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464－7475.

[14] Wang A, Chen H, Liu L, et al. YOLOv10: real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984－108011.

[15] Chakrabarty S, Deb C K, Marwaha S, et al. Application of artificial intelligence in insect pest identification: a review[J]. Artificial Intelligence in Agriculture, 2026, 16(1): 44－61.

[16] 李伟强，王东，宁政通，等. 计算机视觉下的果实目标检测算法综述[J]. 计算机与现代化，2022(6)：87－95.
Li Weiqiang, Wang Dong, Ning Zhengtong, et al. Survey of fruit object detection algorithms based on computer vision[J]. Computer and Modernization, 2022(6): 87－95. (in Chinese)

[17] 翟肇裕，张梓涵，徐焕良，等. YOLO 算法在动植物表型研究中应用综述[J]. 农业机械学报，2024，55(11)：1－20.
Zhai Zhaoyu, Zhang Zihan, Xu Huanliang, et al. A review of applications of YOLO algorithms in animal and plant phenotyping research[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 1－20. (in Chinese)

[18] Wosner O, Farjon G, Bar-Hillel A. Object detection in agricultural contexts: a multiple resolution benchmark and comparison to human[J]. Computers and Electronics in Agriculture, 2021, 189: 106404.

[19] Tang Y, Chen M, Wang C, et al. Recognition and localization methods for vision-based fruit picking robots: a review[J]. Frontiers in Plant Science, 2020, 11: 510.

[20] Caballero-Ramirez D, Baez-Lopez Y, Limon-Romero J, et al. An assessment of human inspection and deep learning for defect identification in floral wreaths[J]. Horticulturae, 2023, 9(11): 1213.

[21] Wu H, Mo X, Wen S, et al. DNE－YOLO: a method for apple fruit detection in diverse natural environments[J]. Journal of King Saud University-Computer and Information Sciences, 2024, 36(9): 102220.

[22] Zeng T, Tong J, Sun X, et al. YOLO－FCAP: an improved lightweight object detection model based on YOLOv8n for citrus yield prediction in complex environments[J]. Smart Agricultural Technology, 2025, 12: 101590.

[23] Qu H, Du H, Tang X, et al. Citrus fruit diameter estimation in the field using monocular camera[J]. Biosystems Engineering, 2025, 252: 47－60.

[24] Zhang Z, Tanimoto Y, Iwata M, et al. Data generation using Pix2Pix to improve YOLO v8 performance in UAV-based Yuzu detection[J]. Smart Agricultural Technology, 2025, 10: 100777.

[25] Weng W, Lai Z, Cui Z, et al. GCD－YOLO: a deep learning network for accurate tomato fruit stalks identification in unstructured environments[J]. Smart Agricultural Technology, 2025, 12: 101465.

[26] Liu Z, Abeyrathna R M R D, Sampurno R M, et al. Faster－YOLO－AP: a lightweight apple detection algorithm based on improved YOLOv8 with a new efficient PDWConv in orchard[J]. Computers and Electronics in Agriculture, 2024, 223: 109118.

[27] Yang H, Yang L, Wu T, et al. MFD－YOLO: a fast and lightweight model for strawberry growth state detection[J]. Computers and Electronics in Agriculture, 2025, 234: 110177.

[28] 吕强，林刚，蒋杰，等. 基于改进 YOLOv5s 模型的自然场景中绿色柑橘果实检测[J]. 农业工程学报，2024，40(18)：147－154.
Lyu Qiang, Lin Gang, Jiang Jie, et al. Detecting green citrus fruit in natural scenes using improved YOLOv5s model[J]. Transactions of the CSAE, 2024, 40(18): 147－154. (in Chinese)

[29] Ma B, Hua Z, Wen Y, et al. Using an improved lightweight YOLOv8 model for real-time detection of multi-stage apple fruit in complex orchard environments[J]. Artificial Intelligence in Agriculture, 2024, 11: 70－82.

[30] Deng X, Huang T, Wang W, et al. SE－YOLO: a sobel-enhanced framework for high-accuracy, lightweight real-time tomato detection with edge deployment capability[J]. Computers and Electronics in Agriculture, 2025, 239: 110973.

[31] 范天浩，顾寄南，王文波，等. 基于改进 YOLOv5s 的轻量化金银花识别方法[J]. 农业工程学报，2023，39(11)：192－200.
Fan Tianhao, Gu Ji'nan, Wang Wenbo, et al. Lightweight honeysuckle recognition method based on improved YOLOv5s[J]. Transactions of the CSAE, 2023, 39(11): 192－200. (in Chinese)

[32] Zhang Y, Chen L, Li X, et al. Multi-arm robotic system and strategy for the automatic packaging of apples[J]. Artificial Intelligence in Agriculture, 2026, 16(1): 578－591.

[33] Ma Z, Dong N, Gu J, et al. STRAW－YOLO: a detection method for strawberry fruits targets and key points[J]. Computers and Electronics in Agriculture, 2025, 230: 109853.

[34] 李茂，肖洋轶，宗望远. 基于改进 YOLO 11 的多尺度板栗果实识别方法[J]. 农业机械学报，2025，56(5)：443－454.
Li Mao, Xiao Yangyi, Zong Wangyuan. Multi-scale chestnut fruit recognition method based on improved YOLO 11[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 443－454. (in Chinese)

[35] 张小艾，乔红波，席磊，等. 基于改进 YOLOv8 的大田农作物害虫小目标检测方法[J]. 农业工程学报，2025，41(13)：225－233.
Zhang Xiaoai, Qiao Hongbo, Xi Lei, et al. Detecting small objects for field crop pests using improved YOLOv8[J]. Transactions of the CSAE, 2025, 41(13): 225－233. (in Chinese)

[36] Zhang Y, Xu Y, Xu T, et al. RSD－YOLO: an improved YOLOv7－tiny framework for oat disease severity identification with integration of ReXNet and decoupled head[J]. Smart Agricultural Technology, 2025, 12: 101433.

[37] Wang C, Dai Y, Sun P P, et al. MRLCM－YOLO:a lightweight and multi scale enhanced model for detecting cowpea pests in complex field environments[J]. Smart Agricultural Technology, 2026, 13: 101676.

[38] 庄煜，徐琨林，李美琪，等. 基于改进 YOLO v11n 的轻量化葡萄穗梗采摘点定位方法与试验[J]. 农业机械学报，2025，56(9)：618－627.

Zhuang Yu, Xu Kunlin, Li Meiqi, et al. Lightweight grape rachis picking-point localization method and experiment based on improved YOLO v11n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(9): 618 – 627. (in Chinese)

[39] Bai Y, Yu J, Yang S, et al. An improved YOLO algorithm for detecting flowers and fruits on strawberry seedlings[J]. Biosystems Engineering, 2024, 237: 1 – 12.

[40] 刘祥, 项若雪, 班成龙, 等. 基于改进 YOLO v8 的轻量化棉铃识别模型与产量预测方法研究[J]. 农业机械学报, 2025, 56(5): 130 – 140.
Liu Xiang, Xiang Ruoxue, Ban Chenglong, et al. Lightweight cotton boll recognition model and yield prediction method based on improved YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 130 – 140. (in Chinese)

[41] Lv L, Li J, Zhao Y. DMP – YOLO: dense multi-scale perception for complex scenes YOLO algorithm prunus humilis small target detection[J]. Smart Agricultural Technology, 2025, 12: 101461.

[42] Liang B, Hu L, Liu G, et al. YOLOv9 – GSSA model for efficient soybean seedlings and weeds detection[J]. Smart Agricultural Technology, 2025, 12: 101134.

[43] Cardellicchio A, Renò V, Cellini F, et al. Incremental learning with domain adaption for tomato plant phenotyping[J]. Smart Agricultural Technology, 2025, 12: 101324.

[44] Chen J, Chen H, Xu F, et al. Real-time detection of mature table grapes using ESP – YOLO network on embedded platforms [J]. Biosystems Engineering, 2024, 246: 122 – 134.

[45] Tan C, Sun J, Paterson A H, et al. Three-view cotton flower counting through multi-object tracking and RGB – D imagery[J]. Biosystems Engineering, 2024, 246: 233 – 247.

[46] Tang B, Zhou J, Zhao C, et al. Using UAV-based multispectral images and CGS – YOLO algorithm to distinguish maize seeding from weed[J]. Artificial Intelligence in Agriculture, 2025, 15(2): 162 – 181.

[47] 陈军, 曹晓明, 刘光耀, 等. 基于 ESW – YOLO 的化学疏花苹果花序状态检测方法研究[J]. 农业机械学报, 2025, 56(11): 433 – 440.
Chen Jun, Cao Xiaoming, Liu Guangyao, et al. Detection method for apple inflorescence status in chemical flower thinning based on ESW – YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(11): 433 – 440. (in Chinese)

[48] 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.

[49] 周宏平, 金寿祥, 周磊, 等. 基于多模态图像的自然环境下油茶果识别[J]. 农业工程学报, 2023, 39(10): 175 – 182.
Zhou Hongping, Jin Shouxiang, Zhou Lei, et al. Recognition of *Camellia oleifera* fruits in natural environment using multi-modal images[J]. Transactions of the CSAE, 2023, 39(10): 175 – 182. (in Chinese)

[50] 伍德林, 刘英豪, 黄可玥, 等. 基于 YOLO – LGC 的油茶果计数及空间定位分布[J]. 农业机械学报, 2025, 56(12): 568 – 580.
Wu Delin, Liu Yinghao, Huang Keyue, et al. *Camellia oleifera* fruit counting and spatial localization distribution based on YOLO – LGC[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(12): 568 – 580. (in Chinese)

[51] Shao J, Zhao Q, Gong Z, et al. ENT – YOLO: an improved lightweight YOLO for cotton organ detection in mulched drip irrigation systems in southern Xinjiang[J]. Agricultural Water Management, 2026, 323: 110054.

[52] Zhao Y, Zhang W, Zhang L, et al. Development of an enhanced hybrid attention YOLO v8s small object detection method for phenotypic analysis of root nodules[J]. Artificial Intelligence in Agriculture, 2026, 16(1): 12 – 43.

[53] 孔德航, 刘云强, 崔巍, 等. 基于改进 YOLOv8n 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 pepper plug seedling sowing quality based on improved YOLO v8n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2): 381 – 392. (in Chinese)

[54] Fatehi F, Bagherpour H, Parian J A. Enhancing the performance of YOLOv9t through a knowledge distillation approach for real-time detection of bloomed damask roses in the field[J]. Smart Agricultural Technology, 2025, 10: 100794.

[55] 喻陈楠, 伍永红, 周杰, 等. 基于改进 YOLO v8n 的非结构环境下杭白菊检测方法[J]. 农业机械学报, 2025, 56(5): 405 – 414.
Yu Chennan, Wu Yonghong, Zhou Jie, et al. Detection method for Hangzhou white chrysanthemum in unstructured environments based on improved YOLO v8n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 405 – 414. (in Chinese)

[56] Xiang M, Gao X, Wang G, et al. An application oriented all-round intelligent weeding machine with enhanced YOLOv5[J]. Biosystems Engineering, 2024, 248: 269 – 282.

[57] 宋绪斌, 王春颖, 李明, 等. 基于表型机器人的小麦关键生育期表型检测方法[J]. 农业机械学报, 2025, 56(3): 67 – 79.

Song Xubin, Wang Chunying, Li Ming, et al. Phenotypic detection method for wheat at key growth stages based on a phenotyping robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 67 - 79. (in Chinese)

[58] 夏天, 谢纯, 李琳一, 等. 基于改进 YOLO v8 的复杂温室环境黄瓜果实分割方法[J]. 农业机械学报, 2025, 56(5): 433 - 442.
Xia Tian, Xie Chun, Li Linyi, et al. Cucumber fruit segmentation method in complex greenhouse environments based on improved YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(5): 433 - 442. (in Chinese)

[59] 张宇轩, 柳平增, 李杰, 等. 基于改进 YOLOv11n - seg 的设施黄瓜植株表型自动化提取方法[J]. 农业工程学报, 2025, 41(24): 191 - 200.
Zhang Yuxuan, Liu Pingzeng, Li Jie, et al. Extracting facility cucumber phenotypes using improved YOLOv11n - seg[J]. Transactions of the CSAE, 2025, 41(24): 191 - 200. (in Chinese)

[60] Zhang Y, Sun P, Jiang Y, et al. Bytetrack: multi-object tracking by associating every detection box[C]// European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 1 - 21.

[61] Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric[C]// 2017 IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE, 2017: 3645 - 3649.

[62] Zhou J, Chen M, Zhang M, et al. Improved YOLOv8 for multi-colored apple fruit instance segmentation and 3D localization[J]. Artificial Intelligence in Agriculture, 2026, 16(1): 381 - 396.

[63] 周成, 陈章彬, 杜雅刚, 等. 面向边缘计算的水稻病害检测方法与装置研究[J]. 农业机械学报, 2025, 56(4): 353 - 362.
Zhou Cheng, Chen Zhangbin, Du Yagang, et al. Rice disease detection method and device for edge computing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(4): 353 - 362. (in Chinese)

[64] 马超伟, 张浩, 马新明, 等. 基于改进 YOLOv8 的轻量化小麦病害检测方法[J]. 农业工程学报, 2024, 40(5): 187 - 195.
Ma Chaowei, Zhang Hao, Ma Xinming, et al. Method for the lightweight detection of wheat disease using improved YOLOv8[J]. Transactions of the CSAE, 2024, 40(5): 187 - 195. (in Chinese)

[65] 王兴旺, 查海涅, 卢浩男, 等. 基于 YOLO v8 STSF 的多类别害虫识别算法与监测系统研究[J]. 农业机械学报, 2025, 56(6): 228 - 236.
Wang Xingwang, Zha Hainie, Lu Haonan, et al. Multi-category pest recognition algorithm and monitoring system based on YOLO v8 STSF[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6): 228 - 236. (in Chinese)

[66] 郑俊键, 兰玉彬, 熊万杰, 等. 基于 YOLOv5s 改进模型的小白菜虫害识别方法[J]. 农业工程学报, 2024, 40(13): 124 - 133.
Zheng Junjian, Lan Yubin, Xiong Wanjie, et al. Method for identification of Pak choi pests and diseases based on improved YOLOv5s[J]. Transactions of the CSAE, 2024, 40(13): 124 - 133. (in Chinese)

[67] 吴晓瑶, 张佳琪, 李功燕, 等. 基于改进 YOLOv8n - seg 的樱桃叠果与瑕疵实时分割[J]. 农业工程学报, 2025, 41(8): 310 - 317.
Wu Xiaoyao, Zhang Jiaqi, Li Gongyan, et al. Real-time segmentation of overlapped cherry fruits and defects using improved YOLOv8n - seg model[J]. Transactions of the CSAE, 2025, 41(8): 310 - 317. (in Chinese)

[68] Zhang Zongqi, Chutichaimaytar P, Ahamed T. YOLO - GIR: a new insect detection algorithm using inception-based variable-size convolutions[J]. Smart Agricultural Technology, 2025, 12: 101651.

[69] Tang Y, Luo F, Wu P, et al. An improved YOLO network for small target insects detection in tomato fields[J]. Computers and Electronics in Agriculture, 2025, 239: 110915.

[70] Mamun S B, Payel I J, Ahad M T, et al. Grape Guard: a YOLO-based mobile application for detecting grape leaf diseases[J]. Journal of Electronic Science and Technology, 2025, 23(1): 100300.

[71] 徐洪丽, 安凤, 史宇轩. 基于改进 YOLO v8 的苹果叶片病害检测研究[J]. 农业机械学报, 2025, 56(10): 530 - 538.
Xu Hongli, An Feng, Shi Yuxuan. Apple leaf disease detection based on improved YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(10): 530 - 538. (in Chinese)

[72] Yao J, Li Y, Xia Z, et al. WTAD - YOLO: a lightweight tomato leaf disease detection model based on YOLO11[J]. Smart Agricultural Technology, 2025, 12: 101349.

[73] Madadum H, Nasir F E, Haruehansapong K. Optimizing watermelon leaf disease detection using SAM-based augmentation with YOLO for practical agricultural solutions[J]. Smart Agricultural Technology, 2025, 12: 101326.

[74] 乔琛, 韩宗桓, 张一丁, 等. 基于改进 YOLO v7 的黄瓜霜霉病菌侵染结构识别方法[J]. 农业机械学报, 2025, 56(11): 528 - 537.
Qiao Chen, Han Zonghuan, Zhang Yiding, et al. Recognition method for cucumber downy mildew infection structures based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(11): 528 - 537. (in Chinese)

[75] Liao H, Wang G, Jin S, et al. HCRP - YOLO: a lightweight algorithm for potato defect detection[J]. Smart Agricultural Technology, 2025, 10: 100849.

[76] Xue L, Cong H, Xu D, et al. DBH - YOLO: a novel algorithm for surface defects detection of postharvest blue honeysuckle

berry fruit[J]. Postharvest Biology and Technology, 2026, 233: 114045.

[77] Al Riza D F, Tulsi A A, Momin A. Assessing cacao beans fermentation degree with improved YOLOv8 instance segmentation [J]. Computers and Electronics in Agriculture, 2024, 227: 109507.

[78] 刘浩蓬, 杨云潇, 康启新, 等. 基于 BMR－YOLO 11n－Pose 的藠头关键点识别与分类方法[J]. 农业机械学报, 2025, 56(11): 490－498.
Liu Haopeng, Yang Yunxiao, Kang Qixin, et al. Keypoint recognition and classification method for *Allium chinense* based on BMR－YOLO 11n－Pose[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(11): 490－498. (in Chinese)

[79] Li H, Huang J, Gu Z, et al. Positioning of mango picking point using an improved YOLOv8 architecture with object detection and instance segmentation[J]. Biosystems Engineering, 2024, 247: 202－220.

[80] Du X, Meng Z, Ma Z, et al. Comprehensive visual information acquisition for tomato picking robot based on multitask convolutional neural network[J]. Biosystems Engineering, 2024, 238: 51－61.

[81] Liang X, Wei Z, Chen K. A method for segmentation and localization of tomato lateral pruning points in complex environments based on improved YOLOV5[J]. Computers and Electronics in Agriculture, 2025, 229: 109731.

[82] Yang L, Noguchi T, Hoshino Y. Development of a pumpkin fruits pick-and-place robot using an RGB－D camera and a YOLO based object detection AI model[J]. Computers and Electronics in Agriculture, 2024, 227: 109625.

[83] 刘芹, 欧阳婧玟, 周智彬, 等. 基于 YOLO－CRC 的茄子及茄子梗分割与定位方法[J]. 农业工程学报, 2025, 41(19): 196－205.
LIU Qin, Ouyang Jingwen, Zhou Zhibin, et al. Segmentation and localization method for eggplants and stems based on YOLO－CRC [J]. Transactions of the CSAE, 2025, 41(19): 196－205. (in Chinese)

[84] He Z, Yuan F, Zhou Y, et al. Stereo vision based broccoli recognition and attitude estimation method for field harvesting[J]. Artificial Intelligence in Agriculture, 2025, 15(3): 526－536.

[85] Lei Y, Yang C, Zhao J, et al. PPBM－YOLO: a lightweight and accurate detection model for multi-species airborne spores in wheat disease cross-infection scenarios[J]. Computers and Electronics in Agriculture, 2026, 241: 111325.

[86] Qin J, Chen Z, Zhang Y, et al. YOLO－CT: a method based on improved YOLOv8n－Pose for detecting multi-species mature cherry tomatoes and locating picking points in complex environments[J]. Measurement, 2025, 254: 117954.

[87] 孙燕, 汪苗, 苏宇辰, 等. 基于改进 YOLO11－seg 的田间水稻分蘖数原位检测方法[J]. 农业机械学报, 2025, 56(9): 596－606, 656.
Sun Yan, Wang Miao, Su Yuchen, et al. In-situ detection method for rice tiller number in fields based on improved YOLO 11－seg[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(9): 596－606, 656. (in Chinese)

[88] 田有文, 覃上声, 闫玉博, 等. 基于改进 YOLOv8 的田间复杂环境下蓝莓成熟度检测[J]. 农业工程学报, 2024, 40(16): 153－162.
Tian Youwen, Qin Shangsheng, Yan Yubo, et al. Detecting blueberry maturity under complex field conditions using improved YOLOv8[J]. Transactions of the CSAE, 2024, 40(16): 153－162. (in Chinese)

[89] 苗全龙, 周扬, 李建涛, 等. 基于 YOLO v8－ABSeg 的双孢蘑菇表型参数提取方法[J]. 农业机械学报, 2025, 56(3): 158－168.
Miao Quanlong, Zhou Yang, Li Jiantao, et al. Phenotypic parameter extraction method for *Agaricus bisporus* based on YOLO v8－ABSeg[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 158－168. (in Chinese)

[90] 黄华, 张存东, 张晖旺, 等. 基于 YOLO v8_EGW 的马铃薯种薯芽眼检测方法[J]. 农业机械学报, 2025, 56(8): 458－466, 506.
Huang Hua, Zhang Cundong, Zhang Huiwang, et al. Potato seed tuber bud-eye detection method based on YOLO v8_EGW [J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(8): 458－466, 506. (in Chinese)

[91] Zhen J, Wang Y, Liu X, et al. Detection of strawberry bloom phenology based on YOLO－RCMC and flower opening scale [J]. Smart Agricultural Technology, 2025, 12: 101275.