doi:10.6041/j.issn.1000-1298.2025.12.052

基于 YOLO－LGC 的油茶果计数及空间定位分布

伍德林[1,2]　刘英豪[1]　黄可玥[1]　李功磊[1]　蒋　杰[1]　王荣炎[1,2]
（1.安徽农业大学工学院，合肥 230036；2.安徽省智能农机装备工程研究中心，合肥 230036）

摘要：深度学习技术在农业视觉任务中应用广泛，确定采摘参数是油茶机械化采摘的首要任务，而基于深度学习的果实自动检测与目标计数对于精准设定这些参数至关重要。然而，果园复杂环境中遮挡、光照变化等因素，直接影响检测方法的效率与准确性。本研究基于 YOLO 11n 提出一种高效轻量级模型 YOLO－LGC，旨在提升在复杂果园环境下目标检测精度和空间定位能力。在主干网络中引入 LGC（LNN Ghost Dynamic Convolution）模块，通过自适应内核选择优化特征提取，提升复杂环境下检测精度；采用可编程梯度信息（Programmable Gradient Information，PGI）机制，动态调节梯度传播，解决遮挡和光照变化导致的特征冲突；融合动态上采样模块（Dynamic UpSample），在 PGI 引导下增强特征保留能力，提高空间定位精度。本研究结合轴对齐边界框（AABB）和惯性测量单元（IMU）以及卡尔曼滤波将油茶果实在高度 0～1.5 m、1.5～1.8 m、1.8 m 以上空间分布进行可视化，通过分析果实空间关系提供更准确的果实位置信息。试验结果表明，与 YOLO 11n 相比，YOLO－LGC 检测精确率从 82.0% 提升至 83.0%，召回率从 76.4% 提升至 78.0%，mAP@0.5 从 85.1% 提升至 87.6%，检测速度从 229.8 f/s 提升至 299.7 f/s，浮点运算量小幅提升至 6.7×10^9，在保持高推理速度优势的同时增强检测性能，优于现有主流检测方法，空间分布误差率小于 5%，能够满足实际检测需求，该研究结果可为油茶果机械化采摘参数匹配提供依据。

关键词：深度学习；目标计数；油茶果；YOLO－LGC；空间定位；果园环境
中图分类号：S24；TP391.4　　**文献标识码**：A　　**文章编号**：1000-1298(2025)12-0568-13　　**OSID**：

Counting and Spatial Distribution of *Camellia oleifera* Fruits Based on YOLO－LGC

WU Delin[1,2]　LIU Yinghao[1]　HUANG Keyue[1]　LI Gonglei[1]　JIANG Jie[1]　WANG Rongyan[1,2]
(1. *College of Engineering*, *Anhui Agricultural University*, *Hefei* 230036, *China*
2. *Anhui Provincial Engineering Research Center for Intelligent Agricultural Machinery Equipment*, *Hefei* 230036, *China*)

Abstract: Deep learning methods are widely applied in agricultural vision tasks. Determining harvesting parameters is a primary task for the mechanized harvesting of *Camellia oleifera* fruits, and automated fruit detection and object counting based on deep learning are crucial for accurately setting these parameters. However, factors such as occlusion and illumination variations in complex orchard environments directly affect the efficiency and accuracy of detection methods. An efficient, lightweight model named YOLO－LGC was proposed based on YOLO 11n, aiming to enhance target detection accuracy and spatial localization capability in complex orchard environments. Firstly, the LNN ghost dynamic convolution (LGC) module was introduced into the backbone network to optimize feature extraction through adaptive kernel selection, thereby improving detection accuracy in complex environments. Secondly, the programmable gradient information (PGI) mechanism was adopted to dynamically adjust gradient propagation, addressing feature conflicts caused by occlusion and illumination changes. Finally, the Dynamic UpSample module was integrated, guided by PGI, to enhance feature retention and improve spatial localization accuracy. Combined axis-aligned bounding boxes (AABB), an inertial measurement unit (IMU), and Kalman filtering to visualize the spatial distribution of *Camellia oleifera* fruits at heights of 0～1.5 m, 1.5～1.8 m, and above 1.8 m. By analyzing the spatial relationships of the fruits, more

收稿日期：2025－07－01　修回日期：2025－07－29
基金项目：安徽省高校协同创新项目（GXXT－2023－111）和安徽省自然科学基金项目（2208085ME132）
作者简介：伍德林（1970—），男，教授，主要从事现代农业装备与技术研究，E-mail：wudelin@126.com
通信作者：王荣炎（1992—），女，讲师，博士，主要从事中药材生产机械化与装备研究，E-mail：wry@ahau.edu.cn

accurate positional information was provided. Experimental results showed that compared with YOLO 11n, YOLO－LGC improved detection precision from 82.0% to 83.0%, recall from 76.4% to 78.0%, mAP@0.5 from 85.1% to 87.6%, and FPS from 229.8 f/s to 299.7 f/s, with computational complexity slightly increased to 6.7×10^9. While maintaining the advantage of high inference speed, YOLO－LGC enhances detection performance, outperforming existing mainstream detection methods. The spatial distribution error rate was less than 5%, meeting practical detection requirements. The findings can provide a basis for matching mechanized harvesting parameters for *Camellia oleifera* fruits.

Key words: deep learning; object counting; *Camellia oleifera*; YOLO－LGC; spatial localization; orchard environment

0　引言

油茶是中国一种重要的油料作物[1]，2020 年我国共种植 4.53×10^7 hm^2 的油茶，总产量逾 6.2×10^5 t，油茶产业达到 1 160 亿元的总产值，尽管油茶的产量与种植规模逐年攀升，但当前仍以手工采摘为主导。在此背景下，实现油茶果的机械化采摘，以降低生产成本，已成为行业迫切需求，而在复杂果园环境中实现油茶果实的精准实时检测，更是支撑智能果园管理、目标计数、产量精准估算及自动化采摘环节的核心技术要点[2]。尽管其意义重大，但在自然果园环境中，果实检测面临着诸多严峻挑战，例如果实相互重叠、被树叶和树枝遮挡，以及光照条件变化无常等[3]。

传统的图像处理方法，如基于特征方法或颜色分割方法，成效有限。LV 等[4]提出一种结合动态阈值分割和特征分析的苹果检测方法，该方法对于未被遮挡的果实效果良好，但在处理相互重叠或被树叶遮挡的目标时则鲁棒性较差。同样，DORJ 等[5]开发一种利用 HSV 颜色空间和分水岭分割的柑橘检测算法，与人工计数的相关系数达到 0.93，但在光照条件变化的情况下性能会下降。这些方法严重依赖人工提取的特征，并且对环境因素十分敏感，这使得它们在实时和动态的果园环境中效果欠佳。随着深度学习的发展，数据驱动的目标检测方法在非结构化环境中展现出卓越的鲁棒性和准确性。以目标检测领域流行的 YOLO 模型为例，由于其能够自动提取高维特征，已成为果实检测任务的基础。李茂等[6]提出一种基于改进 YOLO 11 模型的多尺度板栗果实识别方法 YOLO 11－MCS，通过引入多尺度关键特征聚合模块、CGAFPN 网络和共享卷积分离批量归一化检测头，有效提升板栗识别的精确率与召回率，试验结果表明，该方法在检测性能和部署效率上均满足实际应用需求。李丽等[7]提出一种多通道信息融合网络 YOLO v5－citrus，旨在解决柑橘果实识别精确率低、果实分类模糊和定位不精准等问题，通过引入多通道信息融合模块和聚类算法，提高柑橘采摘状态识别精度，并通过三维坐标定位系统优化机器人采摘路径，试验结果表明该模型具有较强的鲁棒性，适应复杂环境下柑橘果实的精准识别与定位。WU 等[8]进一步证明，将 YOLO v7 与数据增强相结合可以在复杂的田间环境中提高检测精度和泛化能力。

在油茶果自然场景检测中，尤其是在复杂场景、小目标识别以及实时检测的条件下，主流模型常常难以在检测精度、实时性能和模型复杂度之间取得平衡。本研究提出 YOLO－LGC 模型，在复杂果园环境中针对油茶果实进行实时目标检测、计数和空间分布识别而设计的轻量级检测框架。模型集成可编程梯度信息（Programmable Gradient Information，PGI）[9]，动态调节梯度传播，解决遮挡和光照变化导致的特征冲突；采用动态上采样模块（Dynamic UpSampling）以最小化特征损失，通过向 C3K2 模块嵌入 LGC（LNN Ghost dynamic Conv）提高计算效率，通过 PGI 模块实现高效的多尺度特征融合。为进一步提高动态场景的检测精度，将卡尔曼滤波与惯性测量单元（IMU）数据相结合用于姿态估计，同时采用轴对齐边界框（AABB）碰撞检测算法避免对果实重复计数，以保证目标检测鲁棒性和可靠性。

1　材料与方法

1.1　试验环境

油茶果实图像采集于中国安徽省安庆市桐城市孔城镇的安徽旺飞农业发展有限公司的油茶果园。数据采集于 2023 年至 2024 年油茶果成熟季节（9—11 月），通过随机抽样选择 150 棵长势良好的油茶果树，对油茶果进行图像采集，采集时间为 07:00—11:00 以及 13:00—19:00。此外，采集过程考虑各种光照条件，包括过曝、逆光、直射光和侧光。图 1 为数据采集设备以及油茶果实样本图像。

1.2　数据采集

图像采集自树龄为 8 年的油茶树，包括长林 1 号、中科大果和湘林 1 号等品种。种植区域位于安



(a) 油茶树体 (b) 图像采集设备 (c) 油茶果实

图 1 油茶果图像采集

Fig. 1 Image acquisition of *Camellia oleifera* fruit

徽省西南部,属于亚热带季风气候区,气候温和,四季分明,雨量充沛,日照充足。果实图像使用 OAK－D－Pro 相机拍摄,相机分辨率为 4 032 像素×3 040 像素,焦距为 4.81 mm,采用卷帘快门,对角线视野(FOV)为 81°,水平视野为 69°,垂直视野为 55°。相机最大帧率为 60 f/s。

1.3 数据集构建

使用 RobotFlow 工具,由一名标注人员对采集的油茶果实图像进行统一标注。为确保数据质量,对数据集进行了仔细筛选。最终,共获得 1 676 幅图像(含 8 637 个边界框),每幅图像均配有包含目标类别与坐标信息的 TXT 文件。该数据集被划分为训练集(1 464 幅)、验证集(141 幅)和测试集(71 幅),具体分布见表 1,分别用于模型训练、参数调整与性能评估。

YOLO 11n 具有内置的图像增强功能,在训练过程中会自动应用数据增强技术,例如随机亮度调整、旋转、翻转、裁剪和缩放等。这些数据增强技术进一步提高训练集和验证集的多样性,增强模型的鲁棒性和泛化能力,为油茶果实检测模型开发提供坚实的数据基础。

表 1 数据集构建

Tab. 1 Dataset construction 幅

数据集	图像数量	标注框数量
训练集	1 464	5 799
验证集	141	1 938
测试集	71	900
总数据集	1 676	8 637

1.4 改进 YOLO 11n 检测模型

通过协同改进对 YOLO 11n 框架进行架构创新,将设计的 LGC 幽灵液态卷积模块融入主干网络 C3K2 模块中,并采用一个带有空间自适应内核的动态上采样模块,与 PGI 多分支拓扑结构协同工作,解决分辨率缩放过程中出现的特征退化问题,以保留对于被遮挡果实识别至关重要的纹理细节。此外,经 PGI 增强的特征层级结构在多分辨率表示之间建立双向连接,提高对视觉相似叶片图案以及部分被遮挡目标的判别能力。所提出的检测模型,如图 2 所示。其中,CBS 为卷积模块;SPPF 为高效空间金字塔池化模块;Conact 为特征拼接层;C2PSA 为 C2 特征提取与 PSA 注意力机制结合的模块;CBLiner 为线性卷积模块;Silence 为对齐网络结构;

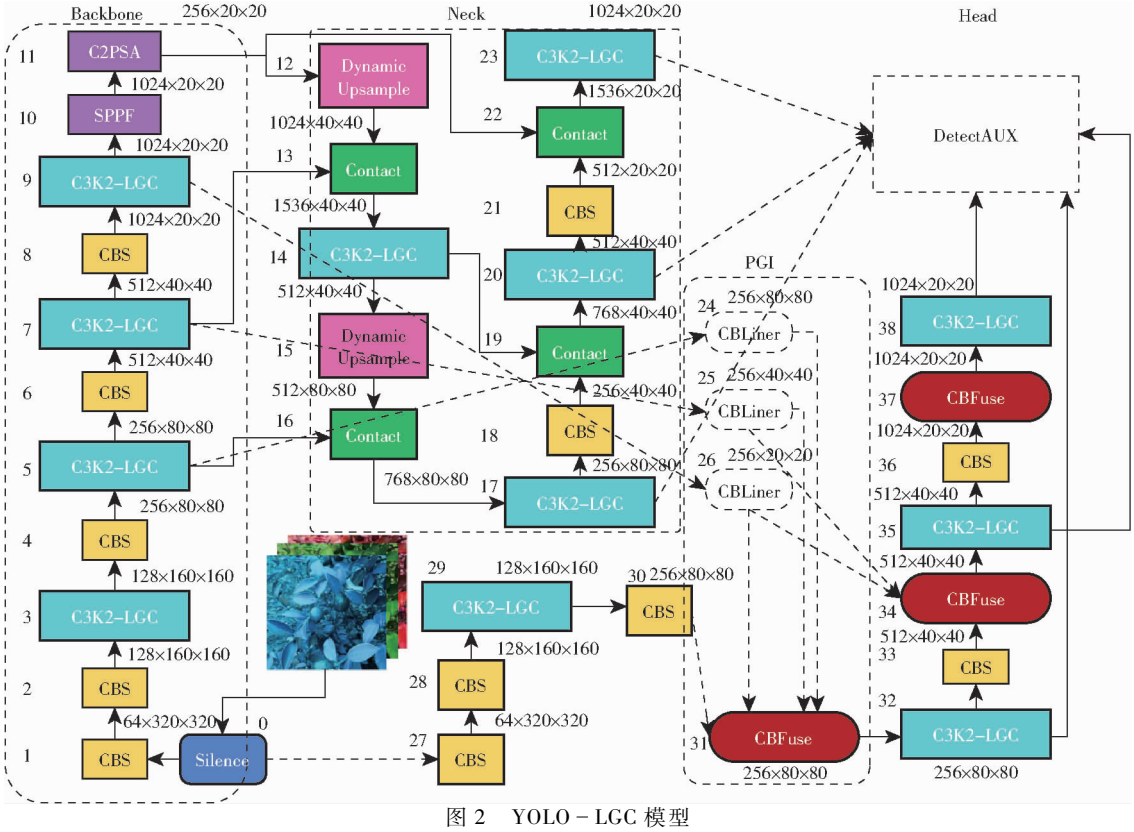


图 2 YOLO－LGC 模型

Fig. 2 YOLO－LGC model

Dynamic Upsample 为动态上采样模块；CBFuse 为特征融合模块；DetectAUX 为辅助检测头。

1.4.1　C3K2－LGC 幽灵液态卷积模块

C3K2 结构的灵感来源于跨阶段局部网络（CSPNet）[9]，通过在不同阶段对特征图进行拆分以优化梯度流动，并有效减少计算冗余，C3K2－LGC 模块通过将原 C3K2 模块中多个 Bottleneck 结构中 Conv2D 层替换为 GhostDynamicConv 层，形成多个 Bottleneck _ GhostDynamicConv 子模块，GhostDynamicConv 单元通过堆叠作用增强特征提取能力，提升计算效率和特征表达能力；同时，在每个 Bottleneck_GhostDynamicConv 子模块后引入液态神经网络（LNN）层，通过微分方程动态建模时序信息，以增强特征在遮挡或光线变化场景下的鲁棒性，如图 3 所示。C3K2－LGC 模块核心是动态卷积，通过多个动态加权共享卷积核替代传统静态卷积核[10]，LNN GhostDynamicConv（LGC）模块结构如图 4 所示。

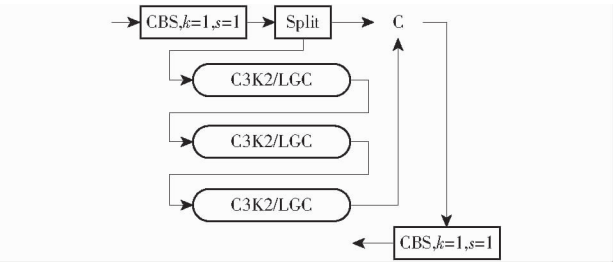


图 3　C3K2－LGC 模块

Fig. 3　C3K2－LGC module

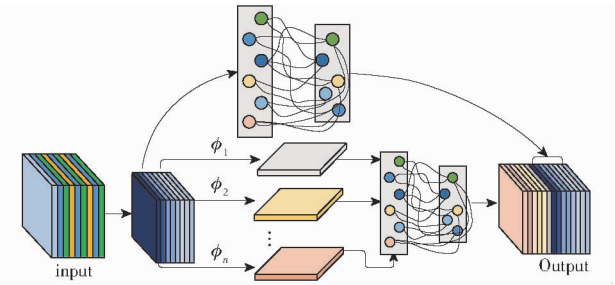


图 4　LNN GhostDynamicConv（LGC）模块

Fig. 4　LNN GhostDynamicConv（LGC） module

C3K2－LGC 模块处理过程：对输入图像特征图进行分组卷积操作，提取空间特征后生成主特征图；通过动态权重生成器对主特征图进行语义分布提取，生成卷积核权重，其中动态权重生成器由一个1×1 卷积层和 Sigmoid 激活函数组成，用于捕捉所述特征图中特征的上下文信息并生成权重系数，以增强 YOLO－LGC 模型对油茶果目标在不同光照和背景下的适应性；卷积核权重对主特征图进行卷积操作，生成 Ghost 特征图，其中生成的 Ghost 特征图通过线性变换从主特征图中生成，进一步降低计算复杂度，同时保留丰富的特征信息；Ghost 特征图和主特征图进行通道级别拼接，生成初始融合特征图；初始融合特征图输入液态神经网络（LNN）层，通过微分方程动态建模时序信息，生成动态增强特征图。

LNN GhostDynamic 表达式为

$$X_m = \mathrm{ReLU}(\mathrm{BatchNorm}(\mathrm{GroupConv}(X))) \quad (1)$$

式中　X_m——生成的特征图

X——输入图像特征图

GroupConv——分组卷积操作，用于提取特征

BatchNorm——批归一化，用于加速训练并稳定模型

ReLU——激活函数，用于引入非线性

对主特征图进行全局平均池化与 1×1 卷积，结合 Sigmoid 函数生成通道注意力权重，表达式为

$$W_d = \mathrm{Sigmoid}(\mathrm{Conv1\times1}(\mathrm{AvgPool}(X_m))) \quad (2)$$

式中　W_d——通道注意力权重

Conv1×1——1×1 卷积操作，用于生成通道权重

AvgPool——平均池化操作，用于下采样并提取全局特征

Sigmoid——激活函数，用于生成 0～1 之间的权重

利用上述权重对主特征图进行线性变换，得到增强后的特征图，表达式为

$$X_g = \mathrm{LinearTransform}(X_m, W_d) \quad (3)$$

式中　X_g——线性变换后的特征图

LinearTransform——线性变换，用于根据权重调整特征图

将原始主特征图与增强后特征图沿通道维度拼接，形成融合特征图 X_f，表达式为

$$X_f = \mathrm{Concat}(X_m, X_g) \quad (4)$$

式中　Concat——拼接操作，用于将主特征图和调整后的特征图按通道维度进行合并

将融合特征图输入液态神经网络（LNN），其状态随时间演化的微分方程表达式为

$$\frac{\mathrm{d}X_{\mathrm{LNN}}(t)}{\mathrm{d}t} = -\alpha X_{\mathrm{LNN}}(t) + \sigma(\boldsymbol{W}_{\mathrm{LNN}} X_f(t) + b_{\mathrm{LNN}})$$

$$X_{\mathrm{LNN}}(0) = X_f \quad (5)$$

式中　$X_{\mathrm{LNN}}(t)$——液态神经网络状态，表示随时间变化的特征

α——衰减因子，控制特征在时间上的变化速率

$\boldsymbol{W}_{\mathrm{LNN}}$——液态神经网络权重矩阵，用于调整特征图

$X_f(t)$——当前时刻特征图

b_{LNN}——偏置项,用于调整网络的输出

σ——激活函数,对计算结果进行归一化处理

在时刻 T 提取液态神经网络的输出特征,经过批归一化、ReLU 激活和 1×1 卷积操作,得到最终输出特征图 Y,表达式为

$$Y = \text{Conv1} \times 1(\text{ReLU}(\text{BatchNorm}(X_{LNN}(T)))) \tag{6}$$

式中 $X_{LNN}(T)$——液态神经网络在最终时刻 T 输出的特征图

BatchNorm——批归一化,用于稳定输出特征

每个卷积核权重是通过全局平均池化和全连接层计算的,并结合挤压激励(Squeeze-and-excitation)方法生成归一化的注意力系数。这些注意力系数会根据输入特征动态调整卷积核的权重,从而增强模型的特征表示能力。权重聚合过程表达式为

$$W = \sum_{k=1}^{k} \pi_k(x) W_d \tag{7}$$

式中 W——最终卷积核权重(加权后的卷积核)

$\pi_k(x)$——归一化注意力系数

注意力系数通过挤压激励机制计算得到,为每个卷积核分配一个动态权重,以便根据输入特征的不同进行调整,注意力系数和为 1,即

$$\sum_{k=1}^{k} \pi_k(x) = 1 \tag{8}$$

式中 $\sum_{k=1}^{k} \pi_k(x)$——对所有注意力系数进行求和

式(8)中 $\pi_k(x)$ 是基于输入 x 动态生成的注意力系数,满足归一化约束。Ghost 模块通过降低成本的线性变换生成额外特征图,从而有效降低计算复杂度。Ghost 模块通过公式(9)生成冗余特征图 $\hat{Y}$,其中 φ 表示线性变换函数,$W_{primary}$ 用于生成主特征图的卷积核权重,W_{ghost} 用于生成冗余特征图的权重。

$$\hat{Y} = \text{Conv}(x, W_{primary}) + \varphi(\hat{y}, W_{ghost}) \tag{9}$$

1.4.2 Dynamic UpSample 动态上采样模块

在油茶果实检测模型中,传统上采样方法通常采用固定的插值策略,这种方法在处理复杂场景时存在一定局限性,对低分辨率特征图的重构能力不足,容易导致特征信息丢失或细节模糊,难以适应多尺度目标的检测需求。为此,本研究提出用动态上采样模块取代传统上采样方法,并将其应用于网络颈部,以高效地重构低分辨率特征图。图 5 为基于采样的动态上采样过程。其中 X 表示特征图;C 表示特征图通道数;H 与 W 表示特征图宽度和厚度;s 代表线性操作中的采样比例因子;H_S 与 W_S 表示像素重组后特征图的高度与宽度;$2g$ 表示像素重组后的通道数;grid sample 表示采样函数;X' 表示采样后特征图。

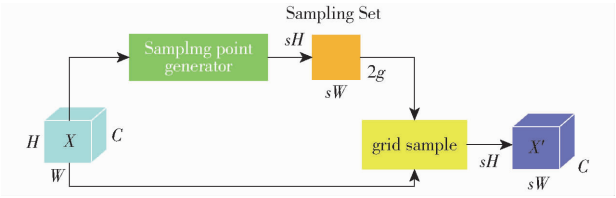


图 5 动态上采样过程

Fig. 5 Dynamic upsampling process

动态上采样模块解决了传统上采样方法在高层语义信息传输方面的局限性[11]。传统插值方法采用固定的采样策略,难以保留深层特征的空间信息,尤其是在复杂背景和多目标检测场景中,目标细节常常会丢失。根据上下文特征动态调整采样位置,动态上采样模块能够精确地重构低分辨率特征图。生成高分辨率特征图为多尺度特征融合提供高质量的输入,提高处理精细细节的能力。

Dynamic UpSample 采样首先输入特征图 X 与采样比例因子 s,特征图维度为 $C \times H \times W$。其次采样点生成器生成采样集 P,并利用残差链接以及 grid sample 函数生成最终 $C \times sH \times sW$ 特征图 X',表达式为

$$X' = \text{grid sample}(X, s) \tag{10}$$

主干网络中幽灵液态卷积(LGC)生成稀疏特征图,通过稀疏计算有效地降低计算需求。然而,仅靠稀疏特征不足以支持高分辨率目标细节的检测,特别是在多尺度特征融合过程中,不同尺度特征的有效传输和精确对齐至关重要。

1.4.3 PGI(可编程梯度信息)机制

为解决上述挑战,在模型中引入可编程梯度信息(PGI)的新型辅助监督框架,如图 6 所示,其中 Pooling 表示池化模块;Unpooling 表示反池化模块;Prediction Head 表示预测头;Auxiliary Branch 表示发辅助分支;Main Branch 表示主分支。

PGI 架构包括主分支与辅助可逆分支 2 个部分,主分支在推理阶段独立运行,确保部署效率,并将计算复杂度严格控制在线性范围内,其计算复杂度范围为

$$C_{inf} = O(n) \quad (n \in N^+) \tag{11}$$

式中 n——网络深度,即神经网络中的层数

N^+——正整数集的标准符号

O——表示计算复杂度与问题规模 n 呈线性关系

C_{inf}——模型在推理阶段计算复杂度

式(11)表明模型计算复杂度与网络层数呈线性关系,有助于保证系统可扩展性与实时性。在

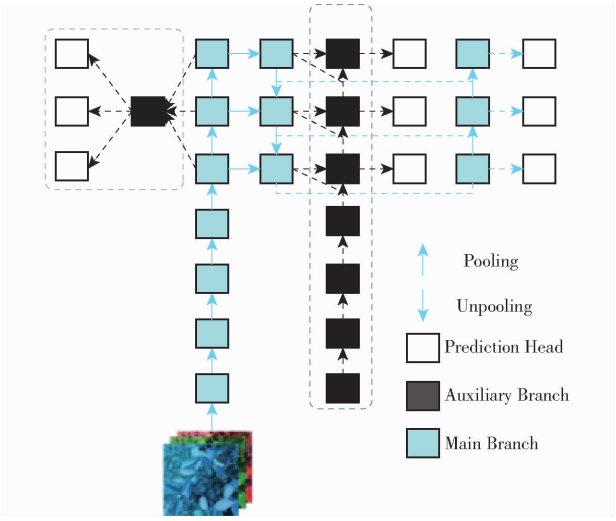


图 6　PGI 结构图

Fig. 6　PGI structure diagram

训练阶段，辅助分支通过一系列可逆映射函数提升信息保真度。通过可逆函数 $\mathscr{F}_{rev}$ 和参数 $\boldsymbol{\Theta}$ 对第 i 层特征图 X_i 进行更新，生成下一层特征图 X_{i+1}，公式为

$$X_{i+1}=v_{\epsilon}(r_{\gamma}(X_i))=X_i+\mathscr{F}_{rev}(X_i;\boldsymbol{\Theta}) \tag{12}$$

式中　v_{ϵ}——特征变换函数

r_{γ}——特征重标定函数

为保证深层特征表达中语义信息不丢失，PGI 设计了信息损失阈值约束，表达式为

$$\Delta I(Y,X)=I(Y,X_i)-I(Y,X_{i+1})<0.05I(Y,X_0) \tag{13}$$

其中 $\Delta I(Y,X)$ 表示标签 Y 与特征图 X_{i+1} 的交互信息，用以平衡在传递过程中信息的保持程度。为进一步优化误差反向传播路径，PGI 机制引入动态掩码控制机制。掩码 M 由特征图 X 通过非线性函数 σ 和映射函数 $\mathscr{F}_{\text{mask}}$ 计算得到，公式为

$$M=\sigma(\mathscr{F}_{\text{mask}}(X;\theta_m)) \tag{14}$$

式中　θ_m——掩码生成函数 $\mathscr{F}_{\text{mask}}$ 的可学习参数集合，在训练过程中通过梯度下降优化

利用掩码 M 对原始特征图进行主元素逐元素调整，实现关键特征的选择性增强，表达式为

$$\hat{X}=v_{\epsilon}(r_{\gamma}(X)\odot M) \tag{15}$$

式中　$\hat{X}$——经过增强后的输出特征图

X——输入特征图，来自网络前一层的特征张量

为评估 PGI 机制对误差控制的作用，定义 PGI 误差 ε_{PGI} 与基础误差 $\varepsilon_{\text{base}}$ 的比值 ζ_{error} 作为评价指标，表达式为

$$\zeta_{\text{error}}=\frac{\|\varepsilon_{\text{PGI}}\|_2}{\|\varepsilon_{\text{base}}\|_2}\leqslant 0.368 \tag{16}$$

PGI 架构通过在推理阶段仅使用主分支，消除额外计算开销。其辅助可逆分支针对深度网络中的信息瓶颈问题，通过可逆变换动态重构梯度流，保留关键语义特征完整性。同时，多级辅助信息系统解决深度监督范式中的误差累积问题，优化不同语义层次的梯度传播路径。这种设计不仅确保任务特征的完整性，还对各种网络规模和复杂度表现出灵活适应性，优于传统深度监督方法。

通过建立可微的梯度路径，PGI 在特征编码阶段可动态调节 LGC 梯度传播方向。核心组件——梯度路由器采用门控机制，有选择地增强有效特征图的梯度流强度，同时抑制冗余特征的梯度更新。在特征解码阶段，PGI 与动态上采样的自适应网格采样形成一个闭环反馈系统，通过可编程梯度场引导高分辨率特征的空间重构过程。这种轻量级设计确保在资源受限设备上的实时性能，使其适合油茶果实检测和空间分布分析的实际应用中。

1.4.4　多源信息融合的油茶果三维定位

为实现油茶果实的精确定位和三维空间定位，本研究首先运用 YOLO－LGC 模型对彩色图像中的果实进行检测，获取每个果实的边界框。将每个检测到的果实中心计算为边界框中心，记（x_{center}，y_{center}）[12]。

基于相机内参矩阵和深度相机生成的深度，将图像坐标系中心点映射到相机坐标系的深度。相机内参矩阵为

$$\text{Camera_matrix}=\begin{bmatrix}f_x & 0 & c_x\\ 0 & f_y & c_y\\ 0 & 0 & 1\end{bmatrix} \tag{17}$$

$$X=\frac{(x_{\text{center}}-c_x)Z}{f_x} \tag{18}$$

$$Y=\frac{(y_{\text{center}}-c_y)Z}{f_y} \tag{19}$$

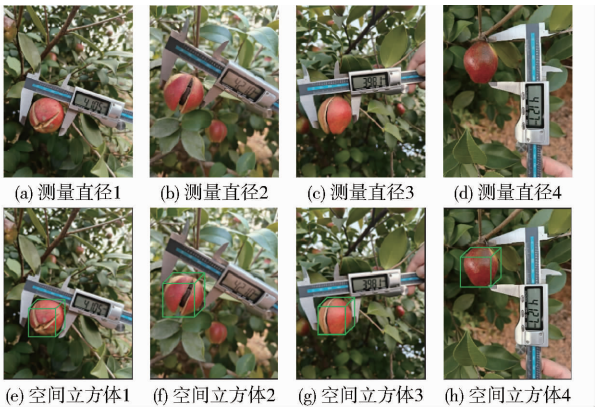
式中　c_x、c_y——图像主点（光学中心）在图像坐标系中的坐标像素

f_x、f_y——图像在 x 和 y 方向上的焦距，像素

将图像平面像素坐标转换为三维空间的相机坐标（X,Y,Z），进而为果实的三维定位提供精确的空间信息。在获取三维坐标后，本研究为每个油茶果实在三维空间中生成一个立方体区域。该立方体以三维深度中心点为中心，其边长由果实测量平均半径（20 mm）决定，如图 7 所示，最终得到一个边长为 40 mm 的立方体。为避免冗余的空间记录，所生成的立方体中心所有三维坐标都存储在一个列表中，并在后续处理中进行碰撞检测。

1.4.5　AABB 碰撞检测与果实计数

为解决实时空间建模过程中检测抖动而产生重



(a) 测量直径1 (b) 测量直径2 (c) 测量直径3 (d) 测量直径4

(e) 空间立方体1 (f) 空间立方体2 (g) 空间立方体3 (h) 空间立方体4

图 7 油茶果测量的直径和相应算法生成的空间立方体

Fig. 7 Measured diameters of *Camellia oleifera* fruits and corresponding cubic spaces generated by algorithm

复计数问题,并确保油茶果实定位的唯一性,本研究采用轴对齐边界框(AABB)碰撞检测算法[13]。检测框的抖动可能会导致将同一个果实错误识别为多个不同的物体,这会影响定位精度。在实时应用中也观察到类似的挑战,AABB 算法为检测到的果实构建最小边界框,并在连续的帧之间对空间立方体进行碰撞检测[14]。两个立方体区域 A 和 B 之间不发生碰撞的条件定义为

$$X^{A}_{max} < X^{B}_{min} \text{ or } X^{A}_{min} > X^{B}_{max} \tag{20}$$

$$Y^{A}_{max} < Y^{B}_{min} \text{ or } Y^{A}_{min} > Y^{B}_{max} \tag{21}$$

$$Z^{A}_{max} < Z^{B}_{min} \text{ or } Z^{A}_{min} > Z^{B}_{max} \tag{22}$$

如果这些条件中的任何一个不成立,那这些立方体就被认为是发生碰撞。当检测到碰撞时,新立方体会被赋予现有立方体的 ID[15],从而确保同一个果实不会被多次计数。否则,就会分配一个新 ID,并记录该果实空间坐标。这种方法整合了 YOLO-LGC 目标检测和深度映射技术,以果实深度坐标为中心生成半径为 20 mm 的空间立方体。轴对齐边界框(AABB)算法能消除由检测框抖动所导致的重复计数情况,可提高空间定位准确性和计算效率。

1.4.6 IMU 和卡尔曼滤波高度估计及降噪

在油茶果实空间定位与分布识别中,精确的高度估计至关重要。然而,油茶种植园地形崎岖,相机移动时的震动会影响定位精度。为此,结合卡尔曼滤波和惯性测量单元(IMU)数据[16],优化高度估计并增强定位系统鲁棒性,更新公式为

$$\hat{x}_k \mid k = \hat{x}_{k \mid k-1} + K_k(z_k - H_k \hat{x}_{k \mid k-1}) \tag{23}$$

$$P_k \mid k = (I - K_k H_k) P_{k \mid k-1} \tag{24}$$

为提高空间定位准确性,使用卡尔曼滤波平滑深度图像数据,因深度图像常因光照变化、相机移动和地形不规则性等环境因素引入噪声。在深度图像中引入高斯噪声和随机抖动以模拟现实噪声,再通过卡尔曼滤波抑制噪声,获得更平滑的深度图[17]。该方法将每个像素视为独立状态,利用卡尔曼滤波器的预测-观测框架迭代更新像素值,有效减少随机波动,保留关键结构提升空间定位的可靠性。

将卡尔曼滤波与惯性测量单元(IMU)的加速度和陀螺仪数据相结合,可实时优化高度估计。IMU 作为一种惯性传感器采集的加速度和角速度可用于计算相机位姿变化,设初始时刻深度为 Z_0,速度为 v_{z_0},位姿为 $Z[0]$,通过加速度 a_z 进行积分可得到相机在 t 时刻深度 $Z(t)$、速度 $v_z[i]$ 以及位姿 $Z[i]$,计算公式为

$$Z(t) = Z_0 + \int_{t_0}^{t}\int_{t_0}^{t} a_z(\tau)\,\mathrm{d}\tau^2 \tag{25}$$

$$v_z[i] = v_z[i-1] + a_z[i]\Delta t \tag{26}$$

$$Z[i] = Z[i-1] + v_z[i]\Delta t \tag{27}$$

为验证卡尔曼滤波结合 IMU 数据的优化效果,本研究设计 1 个动态轨迹试验,包含 1 个高度范围在[0.5 m,1.5 m]之间的波动轨迹。在每个时间步长,实际高度、基于 IMU 的高度估计值以及经过卡尔曼滤波后高度估计值都被记录,如表 2 所示。

表 2 Kalman 滤波处理

Tab. 2 Kalman filtering

时间/s	实际高度/m	IMU 高度估计/m	Kalman 处理高度/m	IMU 估计误差/m	Kalman 估计误差/m
0	0.50	0.62	0.54	0.12	0.04
2	1.00	1.12	1.03	0.12	0.03
4	1.50	1.37	1.47	0.13	0.03
6	1.00	0.92	0.97	0.08	0.03
8	0.50	0.55	0.52	0.05	0.02
10	0.50	0.63	0.53	0.13	0.03
12	1.00	0.98	1.07	0.02	0.07
14	1.50	1.33	1.42	0.17	0.08
16	1.00	1.12	1.04	0.12	0.04
18	0.50	0.42	0.46	0.08	0.04
20	1.00	0.83	1.02	0.17	0.02
平均误差				0.11	0.04

同样地,将卡尔曼滤波应用于带有噪声的深度图像,并对处理后的结果进行可视化如图 8 所示。

由表 2 可知,基于 IMU 高度估计由加速度积分导致的漂移积累,经计算平均误差为 0.11 m。相比之下,经过卡尔曼滤波后高度估计将平均误差降低到 0.04 m,明显地提升了精度。由图 8 可知,在深度图像处理中噪声抑制效果良好,进一步表明卡尔曼滤波在提升数据质量方面有效。

通过将 IMU 数据与卡尔曼滤波相结合,增强了

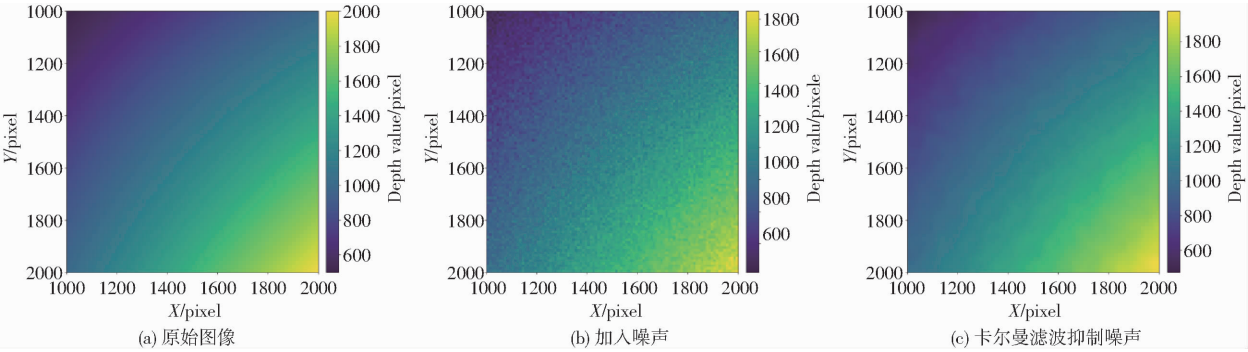


图 8　深度图像对比

Fig. 8　Comparison of depth images

系统在地形不平的动态环境中鲁棒性，为油茶果实实时空间定位和分布识别奠定坚实基础。

2　结果与分析

2.1　试验平台

在一个标准化的计算平台上对模型进行训练和测试。硬件环境包括 1 台运行 Windows 10 系统的台式计算机，Python 版本为 3.12。中央处理器（CPU）是英特尔（R）酷睿（TM）i7－8750H @ 2.20 GHz（2.21 GHz），拥有 8 个内核，图形处理器（GPU）为 NVIDIA GeForce RTX 3080Ti，所有代码均在 PyCharm 中使用 PyTorch 2.3.0 版本和 CUDA 12.1 版本实现。

在油茶果实检测的训练过程中，输入图像统一设置为 640 像素 ×640 像素，批量大小设置为 32。模型训练 110 个周期，使用 4 个工作线程。采用随机梯度下降（SGD）优化器[18]，初始学习率为 0.01，权重衰减参数为 0.000 5，动量参数为 0.937。

2.2　模型评价指标

本研究从多个方面综合评估模型的性能，包括检测精确率 P（precision）、召回率 R（recall）、平均精度均值（mAP@0.5）、检测速度 FPS、浮点运算量 FLOPs。这些指标不仅反映模型的检测性能，还决定其部署性能与资源需求[19]。

2.3　YOLO 不同模型试验结果

为全面评估 YOLO－LGC 性能，将其与其他主流轻量级目标检测模型进行对比，与模型 YOLO v5s[20]、YOLO v7－tiny[21]、YOLO v8n[22]、YOLO v10n[23]、YOLO 11n[24] 和 YOLO 12n[25] 对比试验结果见表 3。不同模型雷达图如图 9 所示。

表 3　不同模型试验结果

Tab. 3　Experimental results from different models

模型	精确率/%	召回率/%	mAP@0.5/%	检测速度/(f·s^{-1})	浮点运算量
YOLO v5n	82.0	76.2	86.8	119.7	4.1×10^{9}
YOLO v8n	78.0	74.0	84.8	189.7	8.1×10^{9}
YOLO v7－tiny	76.0	74.4	83.6	135.2	1.3×10^{10}
YOLO v10n	79.0	74.0	82.3	226.5	6.5×10^{9}
YOLO 11n	82.0	76.4	85.1	229.8	5.8×10^{9}
YOLO v12n	82.0	77.1	85.6	222.6	7.6×10^{9}
YOLO－LGC	83.0	78.0	87.6	299.7	6.7×10^{9}

由表 3 可知，YOLO－LGC 达到了更高的检测精确率（83.0%）、召回率（78.0%）和 mAP@0.5（87.6%），同时保持较低的浮点运算量（6.7×10^{9}）和较高的检测速度（299.7 f/s）。与 YOLO v5n 和 YOLO v7－tiny 模型相比，YOLO－LGC 的 mAP@0.5 分别提高 0.8、4.0 个百分点，充分体现其在检测精度方面的优势。此外，与 YOLO v8n 和 YOLO v10n

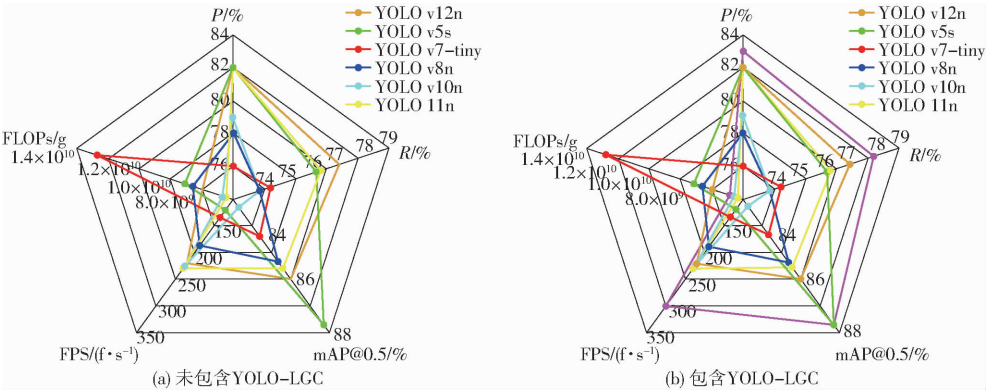


图 9　不同模型雷达图

Fig. 9　Radar charts of different models

相比,YOLO－LGC 在保持相对较低模型计算复杂度的同时,在检测精确率、召回率和平均精度均值方面取得了提升,对比 YOLO12,YOLO－LGC 在检测精确率提高 1 个百分点,召回率提高 0.9 个百分点,mAP@0.5 提高 2 个百分点,且计算复杂度降低 9.0×10^8。图 10 为不同模型在相同图像上的检测结果,图中每列 4 幅图像内容相同,每列代表不同模型对该组图像的检测效果,图 10b(YOLO v7－tiny)、图 10e(YOLO 11n)和图 10k(YOLO v10n)存在误检,而图 10d(YOLO v10n)则存在漏检。此外,在图 10i～10n 中,只有 YOLO－LGC 成功避免漏检,而其他模型则不同程度地存在漏检问题,红圈代表误检,绿框代表漏检。这些结果充分表明 YOLO－LGC 在处理遮挡和检测小目标方面的卓越能力,进一步凸显其在复杂果园环境中的技术优势。



图 10 可视化对比不同模型油茶果实检测结果

Fig. 10 Visualization of oil tea fruit detection across different models

图 11 为 YOLO 11n 与 YOLO－LGC 在油茶果检测任务中的可视化结果。图 11 第 1 行为 YOLO 11n 的检测结果,第 2 行为对应图像的热力图分析;第 3 行为 YOLO－LGC 对相同图像的检测结果,第 4 行为其对应的热力图。第 1 行的检测结果反映出,YOLO 11n 出现误检(第 1 幅)和漏检(第 2 幅)现象,而 YOLO－LGC 在第 3 行中有效避免此类问题。热力图显示,YOLO－LGC 在有遮挡和重叠目标的区域表现出更集中和显著的注意力响应,说明模型能更准确地捕捉关键特征。这一改进源于可编程梯度信息(PGI)机制引导的特征重建策略,增强了模型在复杂背景下对上下文信息的利用能力,进一步验证 YOLO－LGC 在油茶果实时检测中的技术优势与抗干扰能力。

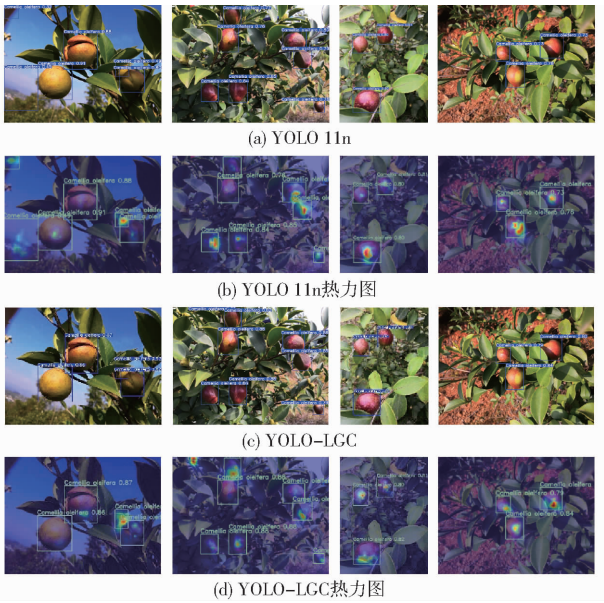


图 11 YOLO 11n 与 YOLO－LGC 油茶果实检测热力图可视化

Fig. 11 Visualization of *Camellia oleifera* fruit detection heatmaps of YOLO 11n and YOLO－LGC

2.4 消融试验

通过消融试验系统地分析幽灵液态卷积模块(LGC)、动态上采样模块(DU)以及可编程梯度信息(PGI)模块对 YOLO－LGC 模型性能的贡献,验证各

模块的必要性和有效性。为确保试验的独立性，诸如 AABB 碰撞检测、卡尔曼滤波以及 IMU 数据融合等外部算法仅在实时检测阶段使用，且不会直接影响消融试验的结果。所有试验均在相同的硬件环境和超参数设置下进行，以确保试验结果的可靠性和可比性。

消融试验结果如表 4 所示，当同时集成 LGC、DU 和 PGI 模块时，模型达到最佳性能，精确率、召回率和平均精度均值 mAP@0.5 分别达到 83.0%、78.0% 和 87.6%。

表 4　YOLO 11n 模型消融试验结果

Tab.4　Ablation results of YOLO 11n model

Base	LGC	DU	PGI	精确率/%	召回率/%	mAP@0.5/%	检测速度/(f·s^{-1})	浮点运算量
√	×	×	×	82.0	76.4	85.1	229.8	5.8×10^9
√	√	×	×	83.0	73.6	85.4	234.1	5.1×10^9
√	×	√	×	84.0	75.5	85.2	315.9	6.3×10^9
√	×	×	√	80.0	76.2	86.0	226.4	8.6×10^9
√	√	√	×	83.0	76.4	85.6	328.5	4.3×10^9
√	√	×	√	81.0	72.2	84.9	247.5	8.1×10^9
√	×	√	√	83.0	78.0	85.6	208.4	8.7×10^9
√	√	√	√	83.0	78.0	87.6	299.7	6.7×10^9

注：√表示使用该模块，×表示不使用该模块。

消融试验结果验证了 LGC、DU 和 PGI 模块在提升 YOLO－LGC 模型性能方面的关键作用，同时也证明多模块集成策略在目标检测任务中的优势。在模块协同方面，LGC 与 DU 的前置协同起到重要作用。LGC 率先实施通道轻量化操作，通过剔除冗余特征有效降低计算负载，这一过程如同为后续操作进行一次“特征清理”，为 DU 的动态上采样扫除“无效特征噪声”。而 DU 则针对 LGC 精炼后留下的核心通道，精准地恢复高分辨率信息。两者形成一种“效率优化-精细特征重建”的正向闭环，使得模型在浮点运算量降低至 4.3×10^9 的情况下，能够保持 mAP@0.5 微增至 85.6%，确保基础特征处理过程的高效性和精准性。在油茶果面临不同光照、背景干扰等复杂场景时，PGI 能够对难样本学习进行定向强化。得益于前置模块已完成的特征提纯，PGI 无需再处理冗余特征干扰，其梯度调节能够更专注地聚焦于真实目标，最终实现“复杂场景适应力与模型效率的双重突破”mAP@0.5 提升至 87.6%，检测速度 299.7 f/s 和浮点运算量为 6.7×10^9，都在可控范围内，充分验证多模块协同所产生的强大效能，对比其他组合，PGI＋DU＋LGC 能够成功融合的关键在于“模块功能的阶梯式互补”：LGC 通过做“减法”优化效率，DU 做“加法”以重建精细特征，PGI 做“乘法”突破复杂场景，这 3 个模块在“特征处理路径和设计逻辑”上深度适配，从而协同释放出模型的性能潜力。

2.5　油茶果实地检测结果评价

在非结构化的油茶果园环境中，果实定位的准确性和稳定性直接影响着采摘机械的作业性能以及产量估算的精度[26－28]。

YOLO－LGC 与 YOLO 11n 相比主要优势在于提升了实际场景的适用性。此外，即便在存在遮挡和果实重叠等具有挑战性的条件下，YOLO－LGC 仍能保持较高的检测精度，实际场景下对油茶果的实时检测结果如图 12 所示。通过对比检测果实数量与实际数量，验证了空间定位准确性，系统评估 YOLO－LGC 性能。在遮挡和果实重叠情况下，深度信息与 AABB 碰撞检测算法结合，使果实检测总数误差率均低于 5%，如表 5 所示。

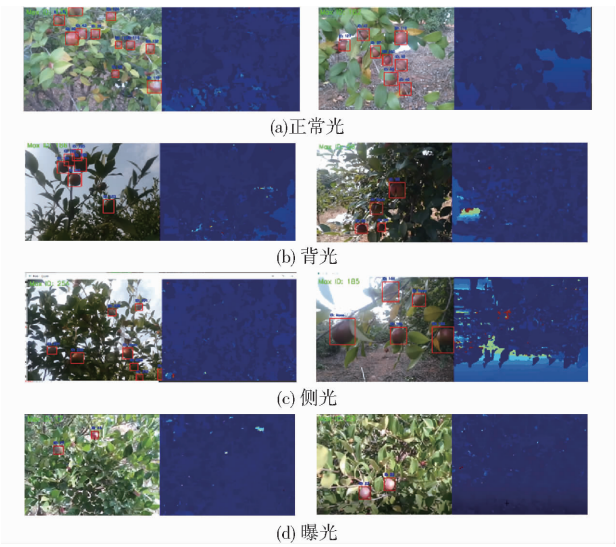


图 12　实际场景下对油茶果的实时检测

Fig.12　Real-time detection of *Camellia oleifera* fruits in actual scenarios

图 13 为 YOLO－LGC 在自然果园环境中的实时检测后的空间位置分布，通过 OAK 相机部署获取的空间坐标信息进行收集，并借助 Matplotlib[29] 可视化工具生成三维分布图。其中 0～1.5 m 为蓝色，1.5～1.8 m 为绿色，1.8 m 以上为红色，完成不同高

表 5 油茶树检测果实数量与实际数量对比

Tab. 5 Comparison of detected and actual fruit counts for *Camellia oleifera* trees

树体样本	区域	检测数量/个	实际数量/个	误差/个	误差率/%
树体 1	高	108	110	-2	-1.82
	中	79	82	-3	-3.66
	低	6	8	-2	-25.00
	总数	193	200	-7	-3.5
树体 2	高	182	180	2	1.11
	中	54	56	-2	-3.57
	低	25	27	-2	-7.41
	总数	261	263	-2	-0.67
树体 3	高	163	162	1	0.62
	中	38	39	-1	-2.56
	低	11	12	-1	-8.33
	总数	212	213	-1	-0.47
树体 4	高	24	26	-2	-7.69
	中	100	105	-5	-4.76
	低	150	155	-5	-3.23
	总数	274	286	-12	-4.19

度区域果实数量的可视化统计分析。

3 结论

(1)基于 YOLO 11n 提出 YOLO-LGC 油茶果实时检测与空间定位分布模型,通过向 C3K2 模块引入轻量化幽灵液态卷积(LGC)模块,高效提取稀疏特征,减少冗余;通过将动态上采样替换网络中的上采样增强特征图分辨率,解决特征丢失问题;引入可编程梯度信息(PGI)模块优化梯度传播,提升特征提取效率。与原始 YOLO 11n 相比,YOLO-LGC 模型平均精度均值从 85.1% 提升至 87.6%,召回率从 76.4% 提升至 78.0%,检测速度从 229.8 f/s 提升至 299.7 f/s;浮点运算量从 5.8×10^9 小幅增至 6.7×10^9,性能优于当前主流未改进的 YOLO 系列网络模型。

(2)引入卡尔曼滤波平滑深度数据并融合 IMU 数据,结合 AABB 轴对齐碰撞算法增强动态果园环境中的鲁棒性,减少果实的重复计数,实现油茶果的空间准确定位,提升了模型在果园动态环境中的目标跟踪鲁棒性。

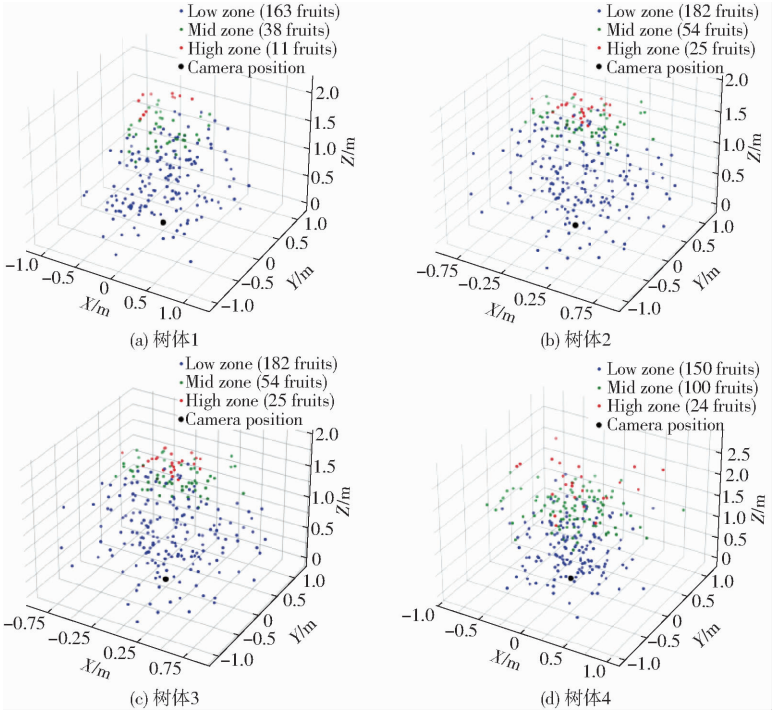


图 13 不同树体检测后油茶果实在空间状态中的数量层次分布

Fig. 13 Hierarchical distributions of number of *Camellia oleifera* fruits in spatial state after detection on different tree bodies

(3)将 YOLO-LGC 进行双目相机部署,实现实际场景下油茶果的的实时识别计数与定位。试验结果表明,在各种环境光照条件下均能实现良好的油茶果检测识别能力。

(4)通过 Matplotlib 可视化技术,将油茶果的空间坐标信息按照高度划分为 0~1.5 m、1.5~1.8 m 和 1.8 m 以上 3 个层次,以直观呈现其空间分布特征,为油茶果的高效采摘与产量评估提供有力的技术支持。

参 考 文 献

[1] 高自成, 李立君, 王勇桦, 等. 基于 Inventor 的油茶果采摘机器人机械臂力学分析[J]. 中南林业科技大学学报,2015,

35(7):121－124,130.
GAO Zicheng, LI Lijun, WANG Yongye, et al. Mechanical arm strength analysis of *Camellia* fruit picking robot based on Inventor[J]. Journal of Central South University of Forestry and Technology, 2015,35(7):121－124,130. (in Chinese)
[2] 迟诚, 康勇军. 扎实推动油茶产业实现高质量发展[J]. 中国林业产业,2019(11):10－11.
[3] 王昱, 易振峰, 谭文超, 等. 遮挡条件下多视角甜椒果实点云三维重构方法[J]. 农业机械学报,2024,55(5):218－225.
WANG Yu, YI Zhenfeng, TAN Wenchao, et al. Multi perspective point cloud reconstruction method for sweet pepper fruit under occlusion conditions[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(5):218－225. (in Chinese)
[4] LV J D, ZHAO D A, JI W, et al. Recognition of apple fruit in natural environment[J]. Optik,2016,127:1354－1362.
[5] DORJ U O, LEE M, YUN S S, et al. An yield estimation in citrus orchards via fruit detection and counting using image processing[J]. Computers and Electronics in Agriculture, 2017,140: 103－112.
[6] 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 detection method based on improved YOLO 11 model[J]. Transactions of the Chinese Society for Agricultural Machinery,2025,56(5):443－454. (in Chinese)
[7] 李丽, 梁继元, 张云峰, 等. 基于改进 YOLO v5 的复杂环境下柑橘目标精准检测与定位方法[J]. 农业机械学报,2024,55(8):280－290.
LI Li, LIANG Jiyuan, ZHANG Yunfeng, et al. Accurate detection and localization method of citrus targets in complex environments based on improved YOLO v5[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(8):280－290. (in Chinese)
[8] WU D, JIANG S, ZHAO E, et al. Detection of *Camellia oleifera* fruit in complex scenes by using YOLO v7 and data augmentation[J]. Applied Sciences, 2022(12): 11318.
[9] WANG C, YEH I, MARK L. YOLO v9: learning what you want to learn using programmable gradient information[C]// European Conference on Computer Vision. Springer, Cham 2025.
[10] CHEN I, DAI X, LIU M, et al. Dynamic convolution: attention over convolution kernels[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2020:11030－11039.
[11] WANG J, CHEN K, XU R, et al. Carafe: content-aware reassembly of features[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision,2019:3007－3016.
[12] ZHENG Z, WANG P, LIU W, et al. Distance-iou loss: faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020(34):12993－13000.
[13] HOR K, SUNG N J, MA J, et al. A fast parallel processing algorithm for triangle collision detection based on AABB and octree space slicing in Unity3D[J]. IEEE Access, 2025,13:4759－4773.
[14] 朱晓林, 刘晓敏, 洪玫, 等. X3D 环境下混合式自适应包围盒生成算法研究[J]. 实验技术与管理,2023,40(12):131－136.
ZHU Xiaolin, LIU Xiaomin, HONG Mei, et al. Research on hybrid adaptive bounding box generation algorithm in X3D environment[J]. Experimental Technology and Management, 2023, 40(12): 131－136. (in Chinese)
[15] 李景顺, 刘美, 孟亚男, 等. 应用轻量化 FEB－YOLO 模型的荔枝果实动态识别计数方法[J]. 计算机测量与控制,2025,33(2):229－237,261.
LI Jingshun, LIU Mei, MENG Ya'nan, et al. Dynamic recognition and counting method for litchi fruit using lightweight FEB－YOLO model[J]. Computer Measurement and Control, 2025, 33(2): 229－237,261. (in Chinese)
[16] 林桂花. 基于无人机技术的精准农业机械导航定位方法[J]. 南方农机,2025,56(8):54－56.
[17] 林俊杰, 李笑, 刘晓明, 等. 基于 YOLO v4 的挖掘机器人土坑检测与三维重建[J]. 机械设计与制造,2025,10:216－225.
LIN Junjie, LI Xiao, LIU Xiaoming, et al. Detection and 3D reconstruction of soil pit for excavating robot based on YOLO v4[J]. Mechanical Design and Manufacturing, 2025, 10:216－225. (in Chinese)
[18] GOWER R M, LOIZOU N, QIAN X, et al. SGD: general analysis and improved rates[C]// Proceedings of International Conference on Machine Learning Research(PMLR), New York, USA,2019:5200－5209.
[19] PSIROUKIS V, ESPEJO-GARCIA B, CHITOS A, et al. Cspnet: a new backbone that can enhance learning capability of CNN[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020:390－391.
[20] JOCHER G, NISHIMURA K, MINEEVA T. YOLO v5 release v6.0[EB/OL]. 2021－10－12[2022－08－10]. https://github.com/ultralytics/yolov5/releases/tag/v6.0.
[21] 王会征, 孙良晨, 李新龙, 等. 基于改进 YOLOv7－tiny 的番茄叶片病虫害检测方法[J]. 农业工程学报,2024,40(10):194－202.
WANG Huizheng, SUN Liangchen, LI Xinlong, et al. Detecting tomato leaf pests and diseases using improved YOLOv7－tiny[J]. Transactions of the CSAE,2024, 40(10): 194－202. (in Chinese)
[22] 岳凯, 张鹏超, 王磊, 等. 基于改进 YOLOv8n 的复杂环境下柑橘识别[J]. 农业工程学报,2024,40(8):152－158.
YUE Kai, ZHANG Pengchao, WANG Lei, et al. Recognizing citrus in complex environment using improved YOLOv8n[J]. Transactions of the CSAE, 2024, 40(8):152－158. (in Chinese)
[23] 韩誉, 刘浩然, 林文树. 基于改进 YOLOv10 和无人机影像的树种识别[J]. 森林工程,2025,41(5):922－935.

HAN Yu, LIU Haoran, LIN Wenshu. Tree species identification based on improved YOLOv10 and UAV RGB imagery[J]. Forest Engineering, 2025,41(5): 922 - 935. (in Chinese)

[24] 王展, 吴浩东, 陈复扬. 基于改进 YOLO11 的太阳电池微缺陷检测方法[J]. 激光与光电子学进展,2025,62(20):1 - 16.
WANG Zhan, WU Haodong, CHEN Fuyang. Solar cell microdefect detection method based on improved YOLO11[J]. Laser and Optoelectronics Progress, 2025,62(20): 1 - 16. (in Chinese)

[25] TIAN Y, YE Q, DOERMANN D. YOLO v12: attention-centric real-time object detectors[J]. arXiv preprint arXiv: 2502.12524,2025.

[26] 邹荣, 李金炎, 王权, 等. 基于相机与激光雷达融合的番茄果实三维定位研究[J]. 农机化研究,2025,47(10):1 - 10,18.
ZOU Rong, LI Jinyan, WANG Quan, et al. Research on three-dimensional positioning of tomato fruit based on fusion of camera and LiDAR[J]. Journal of Agricultural Mechanization Research, 2025, 47(10): 1 - 10,18. (in Chinese)

[27] 闫英. 基于深度学习的樱桃番茄识别与定位技术研究[D]. 北京:北京信息科技大学,2024.
YAN Ying. Research on recognition and positioning technology of cherry tomatoes based on deep learning[D]. Beijing: Beijing Information Science & Technology University,2024. (in Chinese)

[28] 王元昊. 柑橘采摘机器人目标识别定位与末端执行器设计[D]. 重庆:西南大学,2024.
WANG Yuanhao. Target recognition and positioning for citrus harvesting robot and design of end-effector[D]. Chongqing: Southwest University,2024. (in Chinese)

[29] BISONG E. Matplotlib and seaborn[M]// Building machine learning and deep learning models on google cloud platform: a comprehensive guide for beginners. Berkeley, CA: Apress,2019:151 - 165.

(上接第 567 页)

[20] YANG H, YUAN H, WANG Y, et al. Rape seedling density estimation in-field conditions based on improved multi-column convolutional neural network[J]. Agronomy Journal, 2024, 116(3): 810 - 825.

[21] CHEN Y, JIANG H, ZHANG X, et al. Refined feature fusion for in-field high-density and multi-scale rice panicle counting in UAV images[J]. Computers and Electronics in Agriculture, 2023, 211: 108032.

[22] LAN S, REN Z, WU Y, et al. SaccadeNet: a fast and accurate object detector[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020: 10397 - 10406.

[23] WOO S, PARK J, LEE J Y, KWEON I S, Cbam: convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 - 19.

[24] GAO J, WANG Q. SCAR: spatial-/channel-wise attention regression networks for crowd counting[J]. Neurocomputing, 2019, 363: 1 - 8.