

基于深度学习热力图回归的樱桃分级检测方法

宋雪珺 高 磊 郭晓霞

(山西工程技术学院, 山西 阳泉 045000)

摘要: [目的] 解决樱桃筛选效率低、成本高的问题。[方法] 提出一种基于热力图回归方法 HRNet-YT, 用于自动识别樱桃大小和果梗有无, 实现高效筛选。HRNet-YT 通过构建多个平行子网络实现多尺度信息融合, 保持高分辨率表达, 确保果梗和果萼关键点热力图的空间准确性。结合热力图技术捕捉丰富的上下文信息, 并优化损失函数以提升模型的鲁棒性和精度。[结果] HRNet-YT-W48 (384×288) 在数据集上的检测准确率为 87.3%, 关键点平均精度 (AP, OKS=0.5) 为 0.22。[结论] 试验提出的方法在樱桃关键点检测任务中具有较高的准确性和适应性。

关键词: 深度学习; 热力图回归; 樱桃; 分级检测; 关键点检测; 多尺度特征融合

Cherry grading screening based on deep learning-driven heatmap regression

SONG Xuejun GAO Lei GUO Xiaoxia

(Shanxi Institute of Technology, Yangquan, Shanxi 045000, China)

Abstract: [Objective] To solve the problems of low efficiency and high cost in cherry screening. [Methods] A heatmap regression method HRNet-YT was proposed to automatically identify cherry size and the presence of fruit stems, thereby realizing efficient screening. HRNet-YT utilized multiple parallel subnetworks to achieve multi-scale information fusion while maintaining high-resolution representations, ensuring the spatial accuracy of heatmaps for stem and calyx keypoints. By leveraging heatmap techniques to capture rich contextual information and optimizing the loss function, the model's robustness and precision were enhanced. [Results] HRNet-YT-W48 (384×288) achieves a detection precision of 87.3% and an keypoint average precision (AP, OKS=0.5) of 0.22 on the dataset. [Conclusion] The proposed method demonstrates high precision and adaptability in the cherry keypoint detection task.

Keywords: deep learning; heatmap regression; cherry; grading screening; keypoint detection; multi-scale feature fusion

随着农业现代化进程的加速,智能化和自动化技术正逐步渗透到农业生产的各个环节。樱桃作为一种高经济价值的水果,其市场需求和商业价值受到果实大小、外观质量及贮藏时间等因素的影响。在樱桃的分级过程中,果实的大小直接关系到其市场售价,有无果梗则对果实的贮藏时间及运输过程中的损耗有着重要影响。当前,樱桃筛选主要依赖人工操作,通过观察大小、颜色及果梗有无进行分类,存在效率低、成本高、标准化不足等问题。

目前,针对樱桃大小和果梗判别的研究主要聚焦于传统图像处理和基于关键点坐标回归的关键点检测方法。传统图像方法通过提取颜色^[1]、纹理^[2-3]和形状^[4-5]

等视觉特征实现分级,且在光照条件稳定、背景简单的环境中能够获得较好的分类效果。例如,使用果实的形状和边缘特征,通过检测圆度、长宽比等参数来实现自动化分级。然而,在复杂的背景环境下(如果园场景或光照条件多变的室内环境),图像特征容易受到干扰,导致分级精度显著下降^[6-8]。此外,传统图像方法对果实细节的捕捉能力有限,难以应对樱桃果梗等细小部件的判别需求,导致鲁棒性和检测精度不足。

近年来,基于深度学习的方法逐渐成为农业智能化的重要技术路径,为农业自动化提供了新的解决方案^[9-11]。针对樱桃大小分级和果梗检测问题,基于关键点检测的深度学习方法因其优越的性能备受关[12-15]。

通信作者:宋雪珺(1986—),女,山西阳泉技术学院中级实验师,学士。E-mail: 373600965@qq.com

收稿日期:2024-12-23 改回日期:2025-03-03

引用格式:宋雪珺,高磊,郭晓霞. 基于深度学习热力图回归的樱桃分级检测方法[J]. 食品与机械, 2026, 42(1): 72-78.

Citation: SONG Xuejun, GAO Lei, GUO Xiaoxia. Cherry grading screening based on deep learning-driven heatmap regression[J]. Food & Machinery, 2026, 42(1): 72-78.

现有关键点检测算法通常采用二维坐标回归方法,直接预测每个关键点的位置,通过定位果梗和果萼等关键部位,为樱桃的大小分级和果梗判别提供了技术支持^[16-18]。实际应用中,樱桃果梗由于尺寸小且细节模糊,难以被准确识别,定位精度较低,特别是对小尺寸果实的检测效果不佳^[19]。

热力图回归具有更强的空间表达能力,能够在复杂场景中提供更加精确的定位,尤其在处理小尺寸果实和细节模糊时,表现出显著优势。基于热力图回归方法的樱桃关键点检测网络(HRNet-YT)可通过多个平行子网络实现多尺度信息融合,保持高分辨率特征表达,并优化损失函数,提升模型精度和鲁棒性。试验拟利用基于热力图回归方法 HRNet-YT 对樱桃进行分级,通过为每个关键点生成热力图(高斯分布),预测其在图像中的概率分布,并与坐标点直接回归方法进行对比,旨在为樱桃的标准化生产和商品化增值提供依据。

1 HRNet-YT 模型结构

HRNet-YT 的整体架构采用多分支的设计理念,将深度分布到每个阶段,通道数分布到每个分辨率分支,其中每个分支专注于不同尺度特征的提取,并通过平行网络进行信息融合,如图 1 所示。该模型由 4 个阶段组成,每个阶段都由多个平行子网络构成,负责处理不同尺度的信息。这些阶段通过逐步的特征提取与融合,最终得到对樱桃果萼和果梗关键点的精确预测。

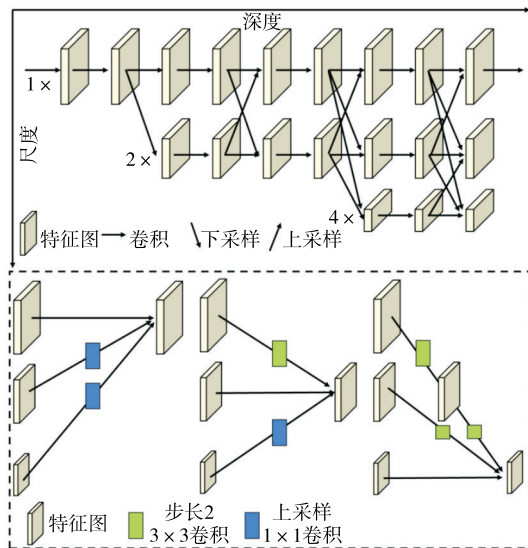


图 1 HRNet-YT 整体架构

Figure 1 Overall architecture of HRNet-YT

第 1 阶段,输入图像经过一个卷积层,用于提取图像的低级纹理和边缘信息。第 2 阶段,通过两个并行分支提取多尺度信息,其中一个分支执行 2 倍下采样以捕捉低分

辨率的特征,另一个分支保持输入特征图的高分辨率,并进一步进行卷积操作,以提取更细致的特征,为后续阶段提供多尺度特征表示。第 3 阶段,主要通过加深网络的深度来增强特征表示的能力,并对两种分辨率分支特征图进行融合,为后续阶段的精细特征提取和关键点检测提供更丰富的特征表示。第 4 阶段,添加一个 4 倍下采样分支,以进一步增强模型对更大范围特征的感知。首先,进行两分支与三分支的特征融合,之后进行 2 倍下采样分支与 4 倍下采样分支的特征融合,最后进行 3 种分辨率分支的特征融合。通过这样层次化特征融合,模型能够在不同尺度上充分整合信息。最终,网络输出一个预测热力图预测关键点在图像中的位置与概率分布,从而定位出樱桃的果梗、果萼。融合模块的目的是每个分支的输出都能从其他并行分支的输出接收信息。以 3 个表征的融合模块为例,输入 3 个分支表征 $\{R_r^i, r=1, 2, 3\}$, 输出也是 3 个表征 $\{R_r^o, r=1, 2, 3\}$, 其中 r 表示分辨率分支, i 表示输入, o 表示输出。输出表征为:

$$R_r^o = f_{1r}(R_1^i) + f_{2r}(R_2^i) + f_{3r}(R_3^i). \quad (1)$$

其中, $f_{xr}(\cdot)$ 为转换函数,第 1 个下标 x 对应第 x 个输入,第 2 个下标 r 对应第 r 个输出。

如果 $x=r$, $f_{xr}(R)=R$; 如果 $x < r$, $f_{xr}(R)$ 对应的表征 R 通过 $r-x$ 个步长为 2 的 3×3 卷积进行下采样; 如果 $x > r$, $f_{xr}(R)$ 通过双线性插值进行上采样,并连接 1 个 1×1 卷积对通道数进行对齐。

热力图回归的关键点检测方法本质是输出一个以每个真实关键点为中心的高斯分布,将 IoU (intersection over union) 损失从矩形框扩展到关键点上(在关键点检测任务上 OKS 指标被视为 IoU),针对每个单独的关键点计算 OKS (object keypoint similarity),并累加到最终的 OKS 损失^[20]:

$$\text{OKS} = \frac{\sum_i \exp\left(-\frac{d_i^2}{2s_i^2\sigma_i^2}\right) \delta(v_i > 0)}{\sum_i \delta(v_i > 0)}, \quad (2)$$

$$\mathcal{L}_{\text{oks}} = 1 - \sum_{n=1}^N \text{OKS}, \quad (3)$$

式中:

d_i ——第 i 个关键点的预测和真实位置之间的欧几里得距离;

s_k ——目标尺度大小(一般选取边界框的面积);

σ_i ——每个关键点的标准误差(通常是一个事先设定的常数,用于反映不同关键点的重要性和尺度不确定性);

$\sum_i \delta(v_i > 0)$ ——归一化因子;

$\delta(v_i > 0)$ ——可见性标记(用于忽略不可见的关键点)。

2 模型训练

试验使用包含 501 幅图像的樱桃数据集,共标注了 801 个樱桃样本。将数据集随机划分为训练集(301 幅)、验证集(100 幅)和测试集(100 幅)。为了统一模型输入尺寸,所有图像在预处理阶段被缩放至 256×192 或 384×288 。随后,图像像素值被归一化到区间 $[0, 1]$,以加速网络的训练收敛并减轻数据分布对模型的影响。将标注樱桃果萼大小和有无果梗的关键点转换为二维高斯分布热力图,如图 2 所示。高斯分布在关键点附近具有明显的中心特征,同时引入一定的模糊性,以缓解数据噪声对模型学习的干扰。

$$H_{ij} = \exp\left[-\frac{(x_i - x_c)^2 + (y_i - y_c)^2}{2\sigma^2}\right], \quad (4)$$

式中:

(x_c, y_c) ——关键点的真实坐标;

(x_i, y_i) ——热力图像素点坐标;

σ ——高斯分布的标准差,通常为 3~5 像素。

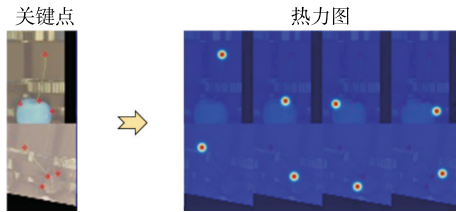


图 2 标签处理结果

Figure 2 Label processing results

为了提升模型的泛化,训练过程中引入多种数据增强策略,包括旋转、水平翻转及亮度调整等操作。数据增强能够模拟不同场景下的图像变化,使模型在复杂背景与光照条件下表现更加鲁棒。模型训练基于 Pytorch 框架实现,参数优化采用 Adam 优化器,其具有自适应学习率调整机制,能够在初始阶段快速收敛。学习率初始值设置为 0.001,并结合余弦退火学习策略逐步降低学习率。

$$\eta_t = \eta_{\min} + 0.5(\eta_{\max} - \eta_{\min}) \left[1 + \cos\left(\frac{t}{T_{\max}}\pi\right) \right], \quad (5)$$

式中:

η_t ——第 t 次迭代的学习率;

η_{\max}, η_{\min} ——最小和初始学习率;

T_{\max} ——模型训练总迭代轮次。

为了全面评估模型在樱桃关键点任务中的性能,采用 PCK (probability of correct keypoint) 和 AP (average precision) 作为评价指标。PCK 强调单点的位置精度,AP 注重目标内所有关键点整体一致性和定位质量。PCK 定义为正确关键点的比例,其计算式为

$$PCK_i^k = \frac{\sum_p \left(\frac{d_{pi}}{d_p^{\text{def}}} \leq T_k \right)}{\sum_p 1}, \quad (6)$$

式中:

T_k ——阈值;

d_{pi} ——第 p 个樱桃第 i 个关键点预测值与真实值之间的欧式距离;

d_p^{def} ——归一化系数。

AP 的计算步骤为:

(1) 预测值与真实值的关键点相似性 (object keypoint similarity, OKS)。

(2) 设定阈值 (0.50, 0.55, ..., 0.95) 并根据这些阈值判断预测是否正确。

(3) 对于每一个阈值,计算对应的精度和召回率曲线,计算 AP,最终对所有阈值的 AP 进行平均,得到最终的平均 AP。

3 结果与分析

试验表明,从果萼关键点偏离角度上看,HRNet-YT 热力图回归检测模型相较于坐标回归方法在准确性上表现更为优越,热力图回归模型的检测点和目标的实际位置更加接近。热力图回归方法通过对每个关键点位置的热力图预测,形成一组具有高置信度的热力分布,能够有效缓解坐标回归方法可能出现的偏移或不准确的情况。

HRNet-YT 模型检测结果为了全面评估模型的性能和关键组件的性能,分别对骨干网络和损失函数进行消融试验,探究其对模型性能影响 (PCK 用于反映检测结果的整体匹配程度,关注定位精度;OKS 结合物体尺度和关键点的重要性权重衡量预测关键点与真实关键点的距离关系)。ResNet 凭借其残差结构和深度可扩展性,在 ImageNet 等基准数据集上取得了广泛验证。因此,为了聚焦热力图回归与坐标回归在相同网络下的性能差异,专注于回归策略的影响,采用 ResNet 作为对照组,结果见表 1。

由表 1 可知,模型结构方面,HRNet-YT 在 PCK 指标上表现出强劲优势,特别是 w48_384x288 模型, PCK 达到 87.3%, 优于其他所有模型。表明 HRNet-YT 能够精准地定位关键点,特别是在高分辨率下,多尺度特征融合与高分辨率表示的优势更加显著。ResNet 的 PCK 波动较大, res101_256x192 模型的 PCK 为 58.5%, 远低于 res50 和 res152。这可能是由于网络深度增加但特征融合不足,导致其性能下降。ResNet-res152 在 256×192 分辨率下的 PCK 达到 88.2%, 优于其他分辨率和模型,说明较深的网络层数能提升特征提取能力。在更高分辨率 (384×288) 下, PCK 反而下降到 80.7%, 可能是由于过深网络难以有

表 1 不同骨干网络与输入分辨率对模型性能的影响

Table 1 Impact of different backbone networks and input resolutions on model performance

检测方法	模型	模型参数	参数量/M	FLOPs/G	PCK/%	AP(OKS=0.5)
热力图回归	HRNet-YT	w32_256×192	28.50	7.12	72.7	0.21
		w32_384×288	28.50	16.02	75.8	0.23
		w48_256×192	63.60	14.61	82.3	0.07
		w48_384×288	63.60	32.88	87.3	0.22
坐标回归	HRNet-YT	w48_384×288	15.92	8.22	78.5	0.05
热力图回归	ResNet	res50_256×192	34.00	8.98	74.3	0.33
		res50_384×288	34.00	20.21	83.5	0.24
		res101_256×192	52.99	12.37	58.5	0.06
		res101_384×288	52.99	27.83	84.3	0.26
		res152_256×192	68.63	15.76	88.2	0.13
		res152_384×288	68.63	35.44	80.7	0.27
坐标回归	ResNet	res101_384×288	16.44	8.78	81.5	0.22

效利用更高分辨率特征。在分辨率方面,高分辨率(384×288)通常能显著提升PCK,HRNet-w48的PCK从82.3%提升至87.30%,OKS的从0.07提升至0.22,原因在于高分辨率输入可以捕获更细致的局部特征,提高关键点定位的精度。

HRNet-YT通过并行多尺度网络进行特征融合,对关键点局部细节特征捕捉更加敏感,对关键点分布缺乏全局约束。而ResNet的残差结构和深度网络层次更偏向全局语义表达,因此,在宽松匹配条件下(OKS=0.5),其预测结果能更好地覆盖目标区域的整体分布。HRNet-YT

与ResNet-101在使用热力图回归方法时,其精度显著优于坐标回归方法。热力图回归方法通过预测每个像素点的概率分布,更好地保留了空间信息,从而捕捉到更加细粒度的特征,有效提升了关键点定位的性能。例如,热力图回归方法的PCK比坐标回归方法的高8.8%,体现了其显著优势。相比之下,坐标回归方法通过直接预测关键点坐标,未能有效利用像素级的空间特征,导致其精度明显下降。

由图3和图4可知,热力图能够通过生成概率分布更好地应对背景干扰,准确地定位关键点;而坐标回归方法

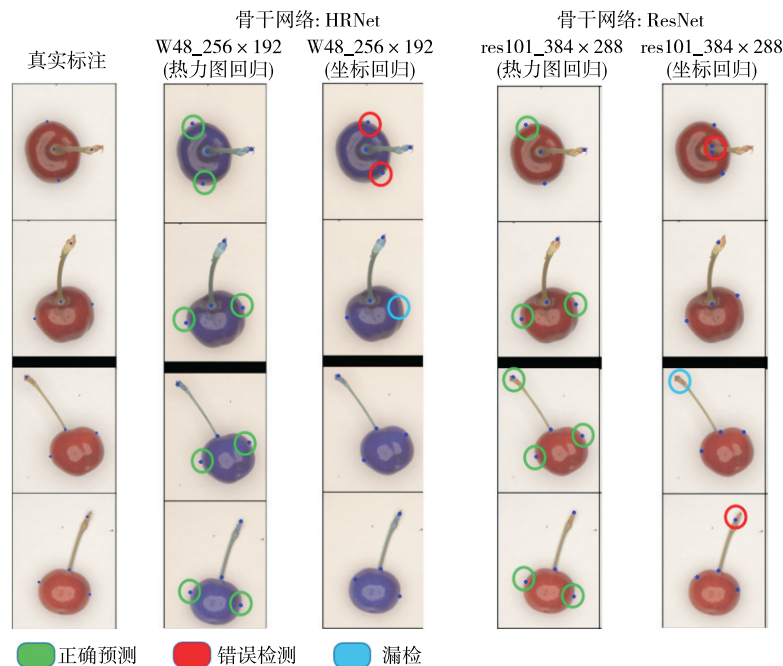


图 3 热力图回归与坐标回归方法的对比结果

Figure 3 Comparison results of heatmap regression and coordinate regression methods

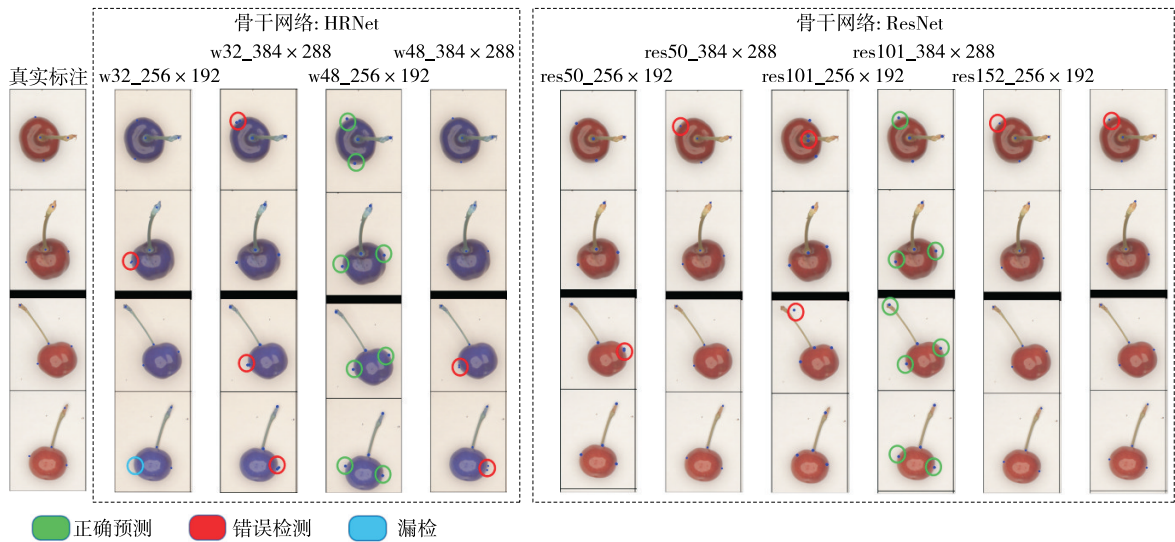


图 4 模型在不同分辨率和结构下的检测结果

Figure 4 Detection results of model under different resolutions and architectures

缺乏空间分布的支持,对关键点的定位更加依赖全局特征,这使其在复杂任务中容易受到干扰,表现出明显的定位偏差和漏检问题。

由表 2 可知,OXS 损失在 PCK 和 AP 两个指标上表现最佳,PCK 达到 87.3%,AP 为 0.33%。OXS 损失的优势在于能够利用关键点空间相似性优化热力图分布,精确地提升关键点定位精度,从而有效地获取目标的空间布局 and 结构。相比之下,MSE 损失虽然表现良好(PCK 为 84.3%,AP 为 0.26%),但其逐像素计算误差的方式忽略了关键点与周围像素的空间关系,仅关注数值差异,导致无法有效学习到热力图的全局信息,精度较 OXS 稍逊。L1 损失和 KL 损失的表现明显较差,其中 L1 损失对较大误差敏感,优化过程中难以平滑过度,影响整体表现(PCK 为 69.5%,AP 为 0.07%)。KL 损失过度关注概率分布之间的差异性,导致梯度信息不足,无法有效优化热力图。Focal Loss 的表现最差,原因在于其主要解决类别不平衡问题,并未针对连续分布优化热力图,因此未能提升检测性能。综上,OXS 损失在樱桃关键点检测任务中能够有效优化热力图分布,提升模型精度和稳定性。

在樱桃果梗和果萼的关键点检测任务中,基于 HRNet-YT-w48 模型的 5 种损失函数(MSE 损失、OXS 损失、L1 损失、KL 损失和 Focal Loss)表现如图 5 所示。OXS 损失因其能够根据目标关键点的空间关系加权优化,表现出色,整体精度高于其他损失函数。MSE 损失虽然能够正确检测关键点,但在果梗和果萼检测中存在轻微的偏移,表明其对空间部分的优化不足,导致精度轻微下降。L1 损失在细节特征的捕捉上较弱,漏检和误检率

表 2 不同损失函数对模型性能的影响

Table 2 Impact of different loss functions on model performance

损失函数	PCK/%	AP(OXS=0.5)/%
MSE	84.3	0.26
OXS	87.3	0.33
L1	69.5	0.07
KL	59.5	0.03
Focal Loss	48.0	0.01

较高。KL 损失在果梗关键点的定位上表现稳定,但在果萼的检测中误差较大,可能是由于 KL 损失过于关注概率分布之间的差异,无法有效处理果萼的空间复杂性。Focal Loss 虽然在处理类别不平衡问题中表现优异,但在连续分布优化中未能发挥其预期效果,导致在樱桃关键点检测任务中的性能较差。综上,OXS 损失凭借其在空间关系优化方面的优势,能够提供更高的检测精度和更稳定的性能。

4 结论

针对樱桃的大小分级与有无果梗检测问题,提出了一种基于改进的高分辨率网络(HRNet-YT)的关键点热力图回归方法。结果表明,热力图回归方法相较于坐标回归方法具有明显优势,能够通过优化关键点的空间关系,显著减少偏移、漏检和误检问题。在骨干网络的选择上,HRNet-YT 相较于 ResNet 系列展现出更强的特征提取能力,尤其是在高分辨率条件下,检测准确率达 87.3%,关键点平均精度(AP,OXS=0.5)为 0.22%,展现了在关键点空间精确性和上下文信息捕捉上的优势。在损失函数的

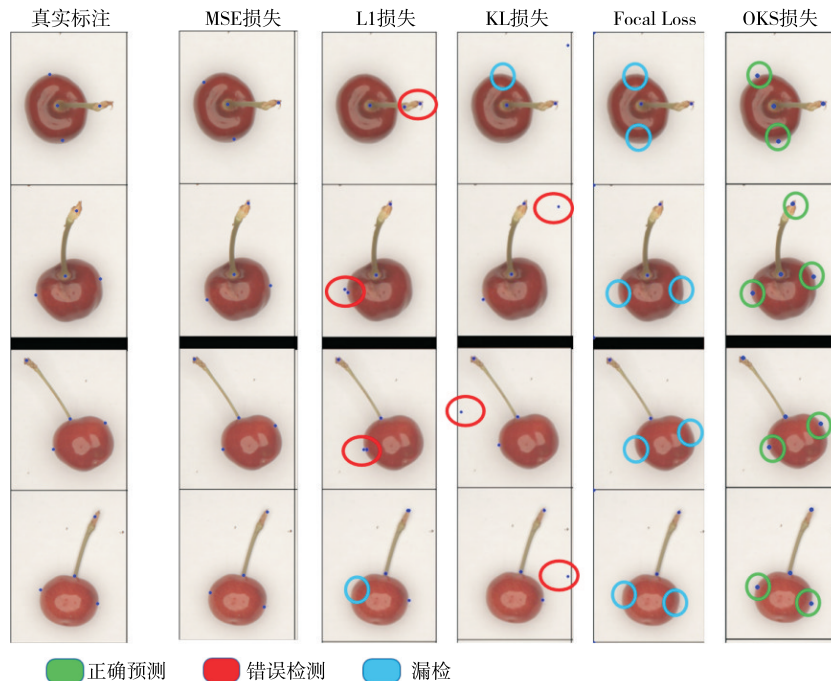


图5 不同损失函数的模型性能对比

Figure 5 Comparison of model performance of different loss functions

选择上,OKS损失能够充分利用目标关键点空间分布信息,避免传统损失函数对离群点的敏感性,采用OKS损失函数的HRNet-YT在PCK和AP(OKS=0.5)指标上均表现最佳,分别为87.3%和0.33%,显著优于其他损失函数。综上,基于高分辨率表示和多尺度融合机制的HRNet-YT热力图回归方法在樱桃关键点检测任务中具有较高的适应性和鲁棒性,系统能够显著提高分级效率,推动樱桃的标准化生产和商品化增值。后续可围绕标注效率提升、模型轻量化及复杂场景泛化进行研究,同时需要在工业环境中验证系统的长期稳定性与成本效益,推动从实验室到生产线的落地应用。

参考文献

- [1] BALAN V, IVANOV I, ȘARBAN V, et al. Changes in cherries size and quality according to color[J]. *Lucrări Științifice Seria Horticultură*, 2017, 60(2): 139-144.
- [2] ROPELEWSKA E, POPIŃSKA W, SABANCI K, et al. Cultivar identification of sweet cherries based on texture parameters determined using image analysis[J]. *Journal of Food Process Engineering*, 2021, 44(7): e13724.
- [3] 王昭. 基于机器视觉的樱桃缺陷检测[D]. 北京: 北京林业大学, 2019: 18-27.
WANG Z. Cherries defect detection based on machine vision [D]. Beijing: Beijing Forestry University, 2019: 18-27.
- [4] 裴悦琨, 叶家敏, 姜艳超, 等. 基于机器视觉的樱桃形状及大小检测技术[J]. *食品工业*, 2020, 41(8): 199-202.
- PEI Y K, YE J M, JIANG Y C, et al. The cherry shape and size detection technology based on machine vision[J]. *The Food Industry*, 2020, 41(8): 199-202.
- [5] 杨立伟, 赖文聪, 刘刚, 等. 基于改进Harris角点检测的虚拟樱桃叶片重建方法[J]. *农业机械学报*, 2022, 53(S1): 213-217.
YANG L W, LAI W C, LIU G, et al. Virtual leaf reconstruction method based on improved Harris corner detection[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(S1): 213-217.
- [6] 魏冉, 裴悦琨, 姜艳超, 等. 基于改进Faster R-CNN模型的樱桃缺陷检测[J]. *食品与机械*, 2021, 37(10): 98-105, 201.
WEI R, PEI Y K, JIANG Y C, et al. Detection of cherry defects based on improved Faster R-CNN model[J]. *Food & Machinery*, 2021, 37(10): 98-105, 201.
- [7] 李兴旭, 陈雯柏, 王一群, 等. 基于级联视觉检测的樱桃番茄自动采收系统设计与试验[J]. *农业工程学报*, 2023, 39(1): 136-145.
LI X X, CHEN W B, WANG Y Q, et al. Design and experiment of an automatic cherry tomato harvesting system based on cascade vision detection[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2023, 39(1): 136-145.
- [8] 代东南, 马睿, 刘起, 等. 基于特征融合注意力机制的樱桃缺陷检测识别研究[J]. *山东农业科学*, 2024, 56(3): 154-162.
DAI D N, MA R, LIU Q, et al. Research on cherry defect detection and recognition based on feature fusion attention mechanism[J]. *Shandong Agricultural Sciences*, 2024, 56(3): 154-162.

- [9] 赵博, 柳苏纯, 张巍朋, 等. 面向樱桃番茄采摘识别的轻量化 Transformer 架构优化研究[J]. 农业机械学报, 2024, 55(10): 62-71, 105.
- ZHAO B, LIU S C, ZHANG W P, et al. Performance optimization of lightweight transformer architecture for cherry tomato picking[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 62-71, 105.
- [10] 沈超, 夏秀波, 杨玮, 等. 温室樱桃番茄果实深度学习识别模型研究[J]. 北方农业学报, 2023, 51(5): 114-122.
- SHEN C, XIA X B, YANG W, et al. Research on deep learning recognition model of greenhouse cherry tomato fruit[J]. Journal of Northern Agriculture, 2023, 51(5): 114-122.
- [11] QIN X Y, CAO J X, ZHANG Y H, et al. Development of an optimized YOLO-PP-based cherry tomato detection system for autonomous precision harvesting[J]. Processes, 2025, 13(2): 353.
- [12] DU W S, JIA Z H, SUI S S, et al. Table grape inflorescence detection and clamping point localisation based on channel pruned YOLOV7-TP[J]. Biosystems Engineering, 2023, 235: 100-115.
- [13] AMIRI N, WANG G H, JANABI-SHARIFI F. Keypoint detection technique for image-based visual servoing of manipulators[C]// 2024 IEEE 20th International Conference on Automation Science and Engineering (CASE). Bari, Italy: IEEE, 2024: 3 557-3 562.
- [14] LI Y J, FENG Q C, LIU C, et al. MTA-YOLACT: multitask-aware network on fruit bunch identification for cherry tomato robotic harvesting[J]. European Journal of Agronomy, 2023, 146: 126812.
- [15] 杨震宇, 汪小昆, 祁子涵, 等. 基于改进 YOLOv8 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 Z Y, WANG X C, QI Z H, et al. Recognition strawberry to detect the key points for peduncle picking using improved YOLOv8 model[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(18): 167-175.
- [16] 刘敬宇, 裴悦琨, 常志远, 等. 基于改进 YOLOX 模型的樱桃缺陷及分级检测[J]. 食品与机械, 2023, 39(1): 139-145.
- LIU J Y, PEI Y K, CHANG Z Y, et al. Cherry defect and classification detection based on improved YOLOX model[J]. Food & Machinery, 2023, 39(1): 139-145.
- [17] 缙新科, 常英. 基于机器视觉的樱桃番茄在线分级系统设计[J]. 计算机与数字工程, 2020, 48(7): 1 798-1 803.
- GOU X K, CHANG Y. Design of cherry tomato online grading system based on machine vision[J]. Computer and Digital Engineering, 2020, 48(7): 1 798-1 803.
- [18] 张永飞, 裴悦琨, 姜艳超, 等. 基于深度卷积神经网络的樱桃分级检测[J]. 食品研究与开发, 2021, 42(14): 138-144.
- ZHANG Y F, PEI Y K, JIANG Y C, et al. Cherry grading using a deep convolutional neural network[J]. Food Research and Development, 2021, 42(14): 138-144.
- [19] ISLAM M A, JIA S, BRUCE N D B. How much position information do convolutional neural networks encode? [EB/OL]. arXiv preprint arXiv. (2020-01-22) [2024-06-11]. <https://arxiv.org/abs/2001.08248v1>.
- [20] MAJI D, NAGORI S, MATHEW M, et al. Yolo-pose: enhancing yolo for multi person pose estimation using object keypoint similarity loss[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022: 2 637-2 646.