doi:10.6041/j.issn.1000-1298.2026.05.014

基于改进 YOLO 13n 模型与边缘计算的芒果识别方法研究

李宏伟[1,2]　高际涛[1]　陈家升[1]　姚　烨[1]　韦　锦[1]　靳震震[1,2]　贺德强[1,2]
(1. 广西大学机械工程学院, 南宁 530004; 2. 广西制造系统与先进制造技术重点实验室, 南宁 530004)

摘要: 在劳动力成本不断上升、人口老龄化日益严重的背景下,探究基于轻量化深度学习模型的芒果识别技术和边缘计算模型部署方法,对于推进芒果采摘机器人设备轻质化发展至关重要。本研究提出了一种基于改进 YOLO 13n 模型的芒果识别方法,并将其部署于边缘计算设备上。首先,将 SE 模块引入到 YOLO 13n 模型的骨干网络中,以提升模型的特征表达能力。其次,在颈部网络中添加 CBAM 模块,以融合更多特征信息,强化芒果识别的通道特征,凸显图像中芒果区域。然后,使用新引入的 SE 模块和 CBAM 模块重塑了 YOLO 13n 架构,得到改进的 YOLO 13n 模型,最后,通过比较模型部署方式,实现了改进 YOLO 13n 模型在边缘计算设备上的部署。试验结果表明,改进 YOLO 13n 模型的芒果检测性能为:精确率 94.5%,召回率 91.2%,$AP_{0.5}$ 95.6%,$AP_{0.75}$ 94.9%,超过了包括 YOLO v8n、YOLO v9s、YOLO v10n、YOLO 11n、YOLO 12n 和 YOLO 13n 在内的众多轻量级先进模型。此外,与原始 YOLO 13n 模型相比,改进 YOLO 13n 模型的精确率、召回率、$AP_{0.5}$ 和 $AP_{0.75}$ 分别提高了 0.6、0.6、1.1、1.5 个百分点,表明改进模型具有整体提升芒果识别性能的效果。为验证改进 YOLO 13n 模型在边缘计算设备上的有效性,将该模型部署在 NVIDIA Jetson Orin Nano 上,单幅图像最高推理速度达到 31.71 f/s,与原始 YOLO 13n 模型相比,推理速度仅下降约 0.4 f/s,仍然高于 30 f/s,达到实时性芒果识别要求。测定 50 幅随机样本图像推理时间,其平均推理时间为 36.88 ms,表明模型对芒果实时识别的稳定性和可靠性。本研究可为研发轻质化芒果采摘机器人提供技术支撑。

关键词: 芒果; 识别; 采摘机器人; 边缘计算; YOLO 13n
中图分类号: S223.2　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)05-0159-08　　**OSID**:

Mango Recognition Method Using Improved YOLO 13n Model and Edge Computing

LI Hongwei[1,2]　GAO Jitao[1]　CHEN Jiasheng[1]　YAO Ye[1]　WEI Jin[1]　JIN Zhenzhen[1,2]　HE Deqiang[1,2]
(1. *School of Mechanical Engineering*, *Guangxi University*, *Nanning* 530004, *China*
2. *Guangxi Key Laboratory of Manufacturing System and Advanced Manufacturing Technology*, *Nanning* 530004, *China*)

Abstract: Aiming at the problems of rising labor costs and an aging population, it is crucial to explore mango recognition technology and edge computing model deployment methods based on lightweight deep learning models to promote the development of lightweight and low-cost mango picking robots. A mango recognition method based on the improved YOLO 13n model was proposed by using YOLO 13n as the baseline model, and it was deployed on edge computing devices. Firstly, the SE module was introduced into the backbone network of YOLO 13n to improve the feature expression ability of the model. Secondly, a CBAM module was added to the neck network to fuse more feature information and strengthen the channel features of mango recognition, and highlight the mango region in the image. Then, the introduced SE module and CBAM module were used to reshape the YOLO 13n architecture to obtain the improved YOLO 13n model, and finally, the deployment of the improved YOLO 13n model on edge computing devices was realized by comparing the model deployment methods. The experimental results showed that the mango detection performance of the improved YOLO 13n model was P was 94.5%, R

收稿日期: 2025-11-07　修回日期: 2025-12-02
基金项目: 国家自然科学基金项目(32501773)、广西自然科学基金项目(2025GXNSFBA069589)和北海市科技计划项目(北科合 2023174001)
作者简介: 李宏伟(1992—),男,副教授,博士,主要从事机器视觉与智能农业机器人研究,E-mail: lihongwei@gxu.edu.cn
通信作者: 贺德强(1973—),男,教授,博士,主要从事智能农林装备、智能检测和故障诊断技术研究,E-mail: hdqlqy@gxu.edu.cn

was 91.2%, $AP_{0.5}$ was 95.6% and $AP_{0.75}$ was 94.9%, which exceeded many lightweight models, including YOLO v8n, YOLO v9s, YOLO v10n, YOLO 11n, YOLO 12n and YOLO 13n. In addition, compared with the original YOLO 13n model, the P, R, $AP_{0.5}$ and $AP_{0.75}$ of the improved YOLO 13n model were improved by 0.6, 0.6, 1.1 and 1.5 percentage points, respectively, and the lightweight characteristics were maintained. To verify the effectiveness of the improved YOLO 13n model on edge computing devices, the improved model was deployed on the NVIDIA Jetson Orin Nano. The maximum inference speed for a single image reached 31.71 f/s. Compared with the original YOLO 13n model, the inference speed only decreased by about 0.4 f/s, still exceeding 30 f/s, meeting the real-time demand for mango recognition. By measuring the inference time on 50 randomly selected sample images, the average inference time was 36.88 ms, indicating the stability and reliability in real-time mango recognition. The research result can provide technical support for the development of lightweight, low-cost mango picking robots.

Key words: mango; recognition; picking robot; edge computing; YOLO 13n

0 引言

随着人们生活水平提高,消费者对健康安全、环境友好的植物性食品需求日益增长。被誉为“热带水果之王”的芒果因其诱人的香气、细腻的口感和丰富的营养成分,深受人们的喜爱[1]。联合国粮农组织数据显示:我国是世界第二大芒果生产国,2021年种植面积达 3.7×10^5 hm^2[2],芒果产业已成为农民增收的重要经济产业。然而,目前芒果收获主要依靠人工采摘,面临劳动力成本不断上升、人口老龄化日益严重等问题,开发智能芒果采摘机器人技术的紧迫性日益凸显,而轻量化的芒果识别模型以及芒果识别深度学习模型边缘计算部署是推进芒果采摘机器人设备轻质化发展的关键技术基础。

在农业机器人领域,视觉系统作为采摘机器人的“眼睛”,其目标识别能力直接决定了机器人作业效率与精度,但在复杂自然环境下,光照变化、目标遮挡、小目标检测等共性挑战制约着技术应用。在光照变化影响方面,果园环境中光照强度存在显著的时空差异,上午强光易产生目标阴影过暗、下午斜射光导致高光反射,以及阴天弱光环境下图像对比度不足,这些变化会导致芒果颜色特征失真、纹理信息模糊,造成识别精度不稳定。

目标遮挡是影响目标识别的另一大挑战性难题。芒果常被叶片、枝条部分交叉遮挡,现有深度学习模型若缺乏针对性的特征增强机制,易将遮挡区域误判为背景或漏检目标,尤其在两阶段检测模型中,候选区域生成阶段可能直接排除部分遮挡的芒果目标,导致召回率降低;而单阶段模型虽速度优势明显,但对遮挡区域的特征区分能力不足,易出现假阳性结果。

小目标检测则进一步加剧了目标识别难度,也是当前影响目标识别的另一挑战性难题。采摘机器人作业时,相机与芒果的距离通常在 0.3 ~ 1.0 m 之间,部分远处或处于树冠内层的芒果在图像中像素占比仅为 2% ~5%,属于典型小目标[3]。这类目标的特征信息稀缺,易与背景中的细小枝条、落叶等干扰物混淆,且现有模型在多尺度特征融合过程中,小目标的高层语义特征易被大目标特征覆盖,导致模型对小目标的识别能力弱于大目标。

近年来,野外环境下的芒果识别技术研究已取得一定的进展。在早期,研究者主要采用图像处理技术对芒果进行识别。研究表明[4-6],基于图像处理的芒果识别技术尽管取得了一定的进展,但对变异光照等不可控因素较为敏感,导致目标图像阈值设置、椭圆拟合的参数难以在不同场景应用中保持一致性,从而限制了相关算法的场景适应性。

随着人工智能技术的快速发展,一些研究者尝试将深度学习算法应用于芒果识别[7-8]。然而,R-CNN 是一种典型的两阶段目标检测方法,意味着在追求识别精度的同时难以兼顾推理速度,不利于机器人采摘作业的实时控制。随着具有快速推理速度的单阶段目标检测方法(如 YOLO 系列, SSD)兴起,研究者不断进行单阶段目标检测方法的改进与应用,先后产生了基于 YOLO v2[9]、改进 YOLO v2[10-11]、Light-YOLO v3[12]、YOLOMS[13]、YOLO v8[14-15]的芒果识别模型。此外,还有一些研究者采用语义分割方法,突破了芒果识别技术[16]。上述基于深度学习的研究尽管在芒果识别效果上取得了新突破,但特定数据集易导致出现局部最优解问题,而这些特定数据集较多采用具有高分辨率成像相机拍摄获得,这与集成在采摘机器人上的低分辨率成像工业相机大不相同。现有研究表明,高分辨率成像与低分辨率成像的目标识别结果存在差异[17],因而在开放果园环境机器人采摘作业中采用低分辨率成像的轻量化深度学习模型感知机理仍未清晰。

为解决机器人采摘应用场景低分辨率成像芒果识别问题,本研究以深度学习模型 YOLO 13n 为基

线模型，探索引入 SE 模块和 CBAM 模块重塑 YOLO 13n 架构的改进算法，并将其部署于边缘计算设备，以期为研发轻质化芒果采摘机器人提供技术支撑。

1　材料与方法

1.1　图像采集

本研究构建了一个具有代表性和多样性的芒果图像数据集。该数据集涵盖不同品种、成熟度及外观特征的芒果图像，原始图像共计 699 幅，旨在为后续模型训练与性能评估提供可靠的数据支撑。图像拍摄于广西大学芒果种质资源圃（北纬 22.84°，东经 108.29°）。拍摄时间为 2025 年 6—7 月，图像采集在自然光条件下进行，拍摄设备为 Intel RealSense D455 型深度相机，拍摄距离范围为 0.3 ~ 1.0 m，图像分辨率为 640 像素 ×480 像素。采集的图像采用多方位的拍摄角度，同时涵盖了上午、下午、晴天、阴天不同环境下的样本，从而提升数据集的代表性与泛化能力。

1.2　数据集构建

共采集了 699 幅原始图像，对其进行裁剪与尺寸归一化预处理，并在此基础上开展数据增强操作，包括旋转、翻转、亮度调整及颜色扰动等，以模拟复杂场景下的视觉差异并增强模型的鲁棒性。增强后的图像数量为 2 095 幅。

数据标注采用 LabelImg 工具完成。采用矩形边框对芒果目标进行标注，在芒果图像标注过程中，图像中较小且较远的芒果并非采摘过程中的作业目标，因此未进行标注。

经过预处理与标注后，最终构建的数据集涵盖了多种场景，包括上午、下午、晴天、阴天等条件下的芒果图像样本，从而形成了一个多元化且具有代表性的数据集。随后，将数据集按照比例约 8∶1∶1 划分为训练集（1 676 幅）、验证集（209 幅）与测试集（210 幅）。

1.3　YOLO 13n 模型

YOLO 13n 模型是一种面向轻量化应用场景设计的目标检测模型，旨在保证较高检测精度的同时大幅减少参数量与计算开销，使其能够在嵌入式设备和移动终端等资源受限平台上实现部署。该模型在保持 YOLO 系列一贯的实时性与鲁棒性基础上，进一步优化了网络结构与训练策略，从而兼顾速度与精度的平衡。基于其轻量化设计、优越的检测性能以及良好的部署适应性，本文根据预试验结果，选择 YOLO 13n 作为基线模型，用于芒果识别。

1.3.1　SE 模块

在芒果目标识别过程中，芒果与背景叶片、枝条之间常存在颜色与纹理的相似性，容易导致模型在特征提取阶段出现通道特征利用不足的问题。为此，引入 SE（Squeeze-Excitation）模块。其中，Squeeze（全局信息聚合）利用全局平均池化对每个特征通道进行统计，将空间信息压缩为通道描述向量，以反映该通道在全局特征表达中的重要性；而 Excitation（通道权重学习）通过两层全连接网络与 Sigmoid 激活函数生成通道权重，进而对原特征进行加权调整，使模型能够突出关键特征通道并抑制背景冗余通道。SE 模块通过“压缩-激励”的方式自适应地为不同通道分配权重，从而突出与芒果目标相关的关键信息，抑制冗余背景特征，为后续操作提供更纯净的特征输入，有效提升了模型的特征表达能力。SE 模块的网络结构如图 1 所示。

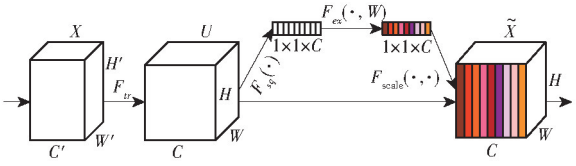


图 1　SE 模块网络结构图

Fig. 1　Network structure of SE module

1.3.2　CBAM 模块

在复杂采摘场景中，芒果的尺度差异显著且容易被叶片部分遮挡，单一通道注意力难以充分捕获目标的空间显著性特征。为进一步提升检测性能，引入空间注意力分支 CBAM（Convolutional block attention module）模块，形成“通道 + 空间”双维度的注意力机制，其中通道注意力分支综合全局平均池化与最大池化的信息，通过共享多层感知机制学习通道权重；而空间注意力分支在通道注意力优化后的特征图上，通过通道维度的平均与最大池化生成双通道特征图，再经卷积压缩为单通道空间权重图，该空间注意力图能够在目标区域生成高响应值，在背景区域生成低响应值，实现空间维度上的特征聚焦。CBAM 模块的双维度注意力机制使模型能够在特征融合阶段同时考虑“通道重要性”与“空间显著性”，即能够在通道维度和空间维度上同时建模，既可以强化芒果相关的关键通道特征，又可以突出图像中芒果所在的显著区域，从而实现更精确的目标定位与识别。CBAM 模块的网络结构图如图 2 所示。

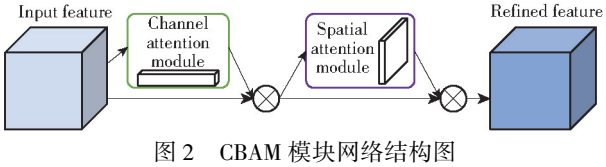


图 2　CBAM 模块网络结构图

Fig. 2　Network structure of CBAM module

1.4 YOLO 13n 模型改进

以 YOLO 13n 为基线模型，并使用 1.3.1 节与 1.3.2 节中提及的 SE 模块和 CBAM 模块对基线模型进行改进。采用“分阶段嵌入 + 串联融合”策略，将 SE 与 CBAM 模块集成至基线模型 YOLO 13n 中。一方面，在作用机理上，将 SE 模块添加到骨干网络 Backbone A2C2f 模块后面，增强芒果与背景相似特征的识别能力，抑制冗余背景特征，提升模型的特征表达能力和检测精度；此外，在颈部网络 Neck 中，紧随最后两个 DS - C3k2 模块后方，各分别添加一个 CBAM 模块，以期通过结合通道注意力与空间注意力，融合更多特征信息，强化芒果识别的通道特征，以及凸显图像中芒果区域。

另一方面，在融合策略上，采用模块串联融合方式，顺序为 SE→CBAM，即：先利用 SE 模块完成通道筛选，保留显著特征后，再通过 CBAM 的空间注意力聚焦关键目标区域，避免空间注意力在冗余通道上浪费计算资源，实现注意力的层次化协同。这种融合策略不仅保证了模块间的信息互补，也使得通道与空间注意力在不同特征层次上发挥最优作用，从而实现检测精度与模型轻量化的平衡。改进的 YOLO 13n 模型结构图如图 3 所示。

1.5 试验环境与参数设置

为保证试验结果的可比性，本文所有研究均在 Windows 10 操作系统环境下完成，采用 PyTorch 深度学习框架实现。试验平台的硬件配置包括：CPU 为 13thGen Intel Core i7 - 13700F 2.10 GHz，GPU 为 NVIDIA RTX 4090，内存容量为 16 GB。训练过程中，系统超参数设置如下：输入图像尺寸为 640 像素 × 640 像素，训练轮次设为 300，批大小为 16，并开启 8 个线程以加速训练。

1.6 评价指标

评估指标主要分为两方面：性能指标和复杂度指标。性能指标包括 4 个关键指标：精确率 P、召回率 R、平均精度 $AP_{0.5}$ 和平均精度 $AP_{0.75}$，精确率用于评估检测框中被正确识别为芒果的比例，召回率用于评估真实芒果目标中被成功检测出的比例，而 $AP_{0.5}$ 和 $AP_{0.75}$ 则代表了精确率和召回率的综合度量。此外，使用浮点运算次数、模型内存占用量和帧率 3 个指标来评估模型的复杂度。

2 结果与分析

2.1 不同模型检测定量对比结果与可视化分析

2.1.1 改进 YOLO 13n 模型与不同模型定量对比结果

为验证改进 YOLO 13n 模型的芒果识别性能，将改进 YOLO 13n 模型与 YOLO v8n[18]、YOLO v9s[19]、YOLO v10n[20]、YOLO 11n[21]、YOLO 12n[22]、YOLO 13n[23] 模型进行对比试验。由于 YOLO v9 未推出“n”(nano，轻量化模型系列)版本，故采用 YOLO v9 模型中网络深度较浅、参数量最少、模型内存占用量最小的 YOLO v9s 版本，作为本文的一种对比模型。

为保障结果可信度，所有模型均在相同训练环

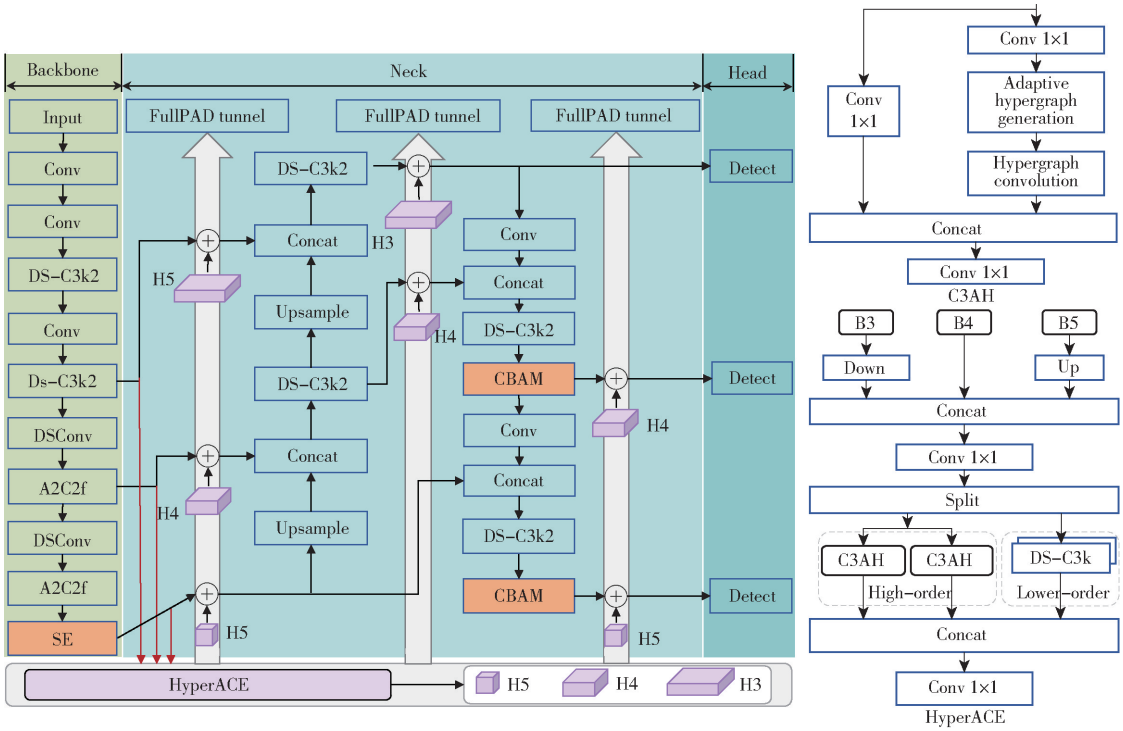


图 3 改进的 YOLO 13n 模型网络结构图

Fig. 3 Network structure of improved YOLO 13n model

境和参数设置下进行。从试验结果可知(见表 1)，改进 YOLO 13n 模型的 P、R、$AP_{0.5}$、$AP_{0.75}$分别达到 94.5%、91.2%、96.7%、94.9%，相比基线模型分别提升 0.6、0.6、1.1、1.5 个百分点，而且保持了相同的计算量，尽管模型内存占用量产生轻微的提高(提高了 0.3%)。根据性能指标观测可见，基于基线模型 YOLO 13n 添加 SE 模块和 CBAM 模块，能够全面提升芒果识别能力；与 YOLO 13n 相比，保持了近乎相同的浮点运算次数。不容忽视的是，改进 YOLO 13n 模型赋能了其本身更多的注意力，使得改进模型的帧率相对基线模型明显下降，这为后续开展模型的部署研究奠定了基础。

表 1　不同模型对比试验结果

Tab. 1　Comparison experiment results of different models

模型	P/%	R/%	$AP_{0.5}$/%	$AP_{0.75}$/%	浮点运算次数	帧率/(f·s^{-1})	模型内存占用量
YOLO v8n	91.5	88.1	93.2	90.1	8.1×10^{9}	168	6.3×10^{6}
YOLO v9s	91.8	85.2	93.7	91.2	2.67×10^{10}	106	1.53×10^{7}
YOLO v10n	93.8	88.9	93.4	90.9	8.2×10^{9}	132	5.8×10^{6}
YOLO 11n	94.8	89.1	93.9	92.3	6.3×10^{9}	155	5.5×10^{6}
YOLO 12n	93.3	87.2	94.2	92.0	5.8×10^{9}	107	5.5×10^{6}
YOLO 13n	93.9	90.6	95.6	93.4	6.2×10^{9}	90	5.4×10^{6}
改进 YOLO 13n	94.5	91.2	96.7	94.9	6.2×10^{9}	82	5.7×10^{6}

2.1.2　不同模型检测结果可视化分析

为更直观地比较芒果识别性能，图 4 可视化了改进 YOLO 13n 模型与其他对比模型对芒果在不同天气条件下的识别效果。在不同天气条件

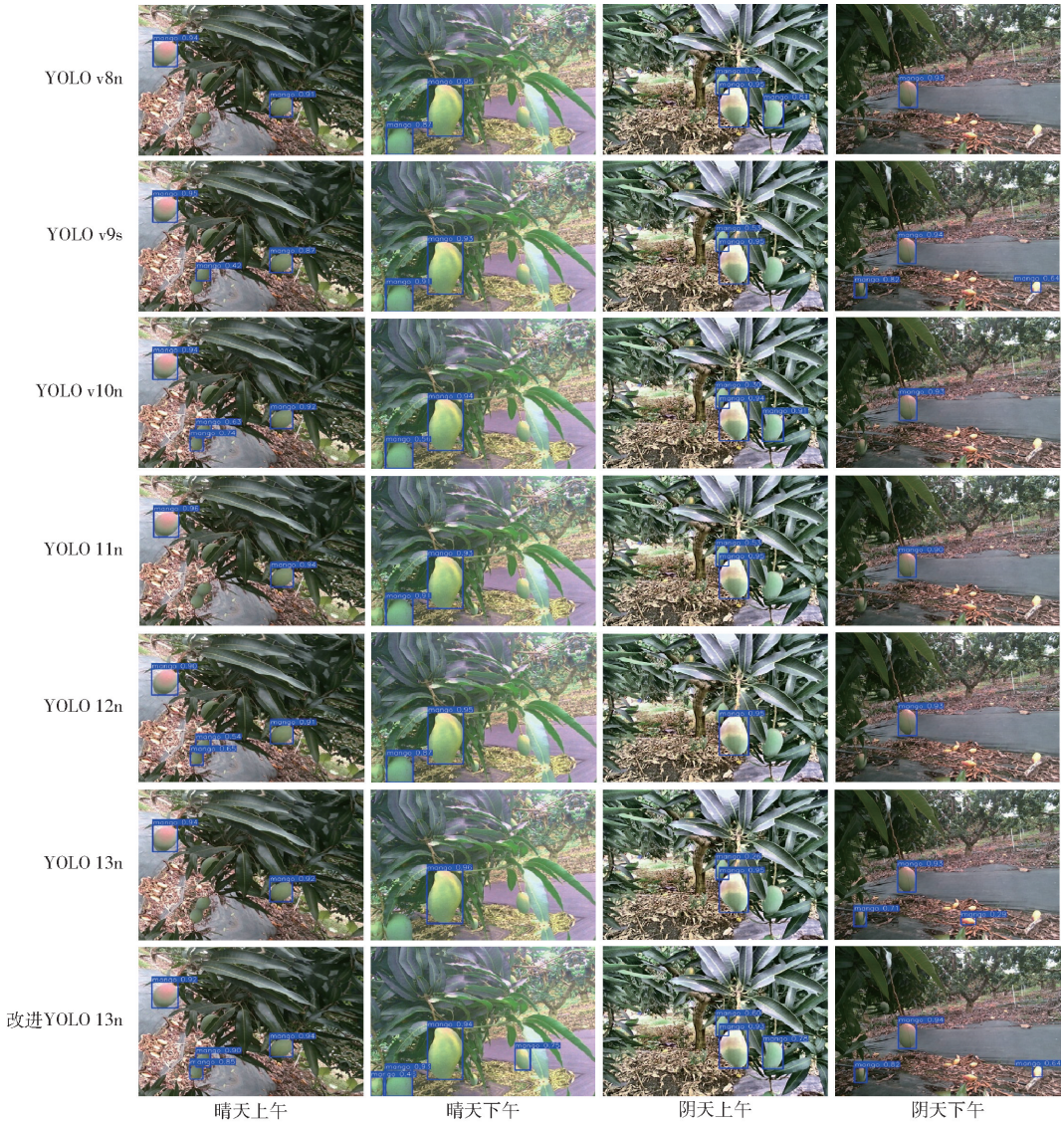


图 4　不同模型检测结果可视化图像

Fig. 4　Visualization diagrams of detection results of different models

下,大多数模型都成功检测出了未被遮挡且清晰可见的芒果。但对于部分遮挡芒果,改进 YOLO 13n 模型具有更强的识别能力,表明添加 SE 模块和 CBAM 模块对提升特征提取能力的有效性。在晴天下午条件下,仅改进 YOLO 13n 模型能够识别距离较远且目标较小的芒果,表明添加 SE 模块和 CBAM 模块对捕获目标空间信息的优越性。

2.2 消融试验结果

本研究对基线模型 YOLO 13n 进行了改进,为进一步验证各添加模块对芒果识别的作用,在芒果数据集上进行了消融试验,结果如表 2 所示。

表 2 消融试验结果

Tab. 2 Ablation experiment results

序号	SE	CBAM	P/%	R/%	$AP_{0.5}$/%	$AP_{0.75}$/%	浮点运算次数	帧率/(f·s^{-1})	模型内存占用量
1	×	×	93.9	90.6	95.6	93.4	6.2×10^9	90	5.4×10^6
2	√	×	94.6	88.2	95.8	94.1	6.2×10^9	89	5.4×10^6
3	×	√	94.4	89.9	96.3	94.3	6.2×10^9	88	5.4×10^6
4	√	√	94.5	91.2	96.7	94.9	6.2×10^9	82	5.7×10^6

注:“×”表示不包含该模块,“√”表示包含该模块。

从表 2 可以看出,在 YOLO 13n 模型主干网络中,当仅添加 SE 模块到 A2C2f 模块之后,可以提升芒果识别精确率和平均精度,但导致召回率有所下降。当在 YOLO 13n 颈部网络中,在最后两个 DS-C3k2 模块后方各分别添加一个 CBAM 模块,同样能够提升芒果识别精确率和平均精度,但召回率相对仅添加 SE 模块时的下降幅度较小。当同时在骨干网络中添加 SE 模块、在颈部网络中添加 CBAM 模块,能够弥补单方面添加 SE 模块或 CBAM 模块的不足,全面提升芒果识别性能。相对基线模型 YOLO 13n,改进 YOLO 13n 模型的性能指标 P、R、$AP_{0.5}$、$AP_{0.75}$分别提升 0.6、0.6、1.1、1.5 个百分点,表明改进 YOLO 13n 模型的芒果识别性能相对基线模型均有提高。在复杂度方面,改进模型相对基线模型具有相近的浮点运算次数,模型内存占用量略微增加,适于边缘计算设备部署。

此外,为直观地观测消融试验中添加各模块的芒果识别效果,生成不同模型的热力图(见图 5)。可以看出,仅添加 SE 模块的 YOLO 13n + SE 模型能够增强相似特征的表达能力,仅添加 CBAM 模块的 YOLO 13n + CBAM 模型能够凸显芒果空间特征信息,而融合 SE 模块和 CBAM 模块的 YOLO 13n + SE + CBAM 模型能够较全面提升芒果识别效果(见图 5 红色区域)。

2.3 改进 YOLO 13n 模型的边缘计算部署

要实现设备轻质化,在边缘计算设备上兼顾识别性能和识别速度至关重要。根据现有文献[16]可知,TensorRT 推理优化器、C ++ 编译和模型 API

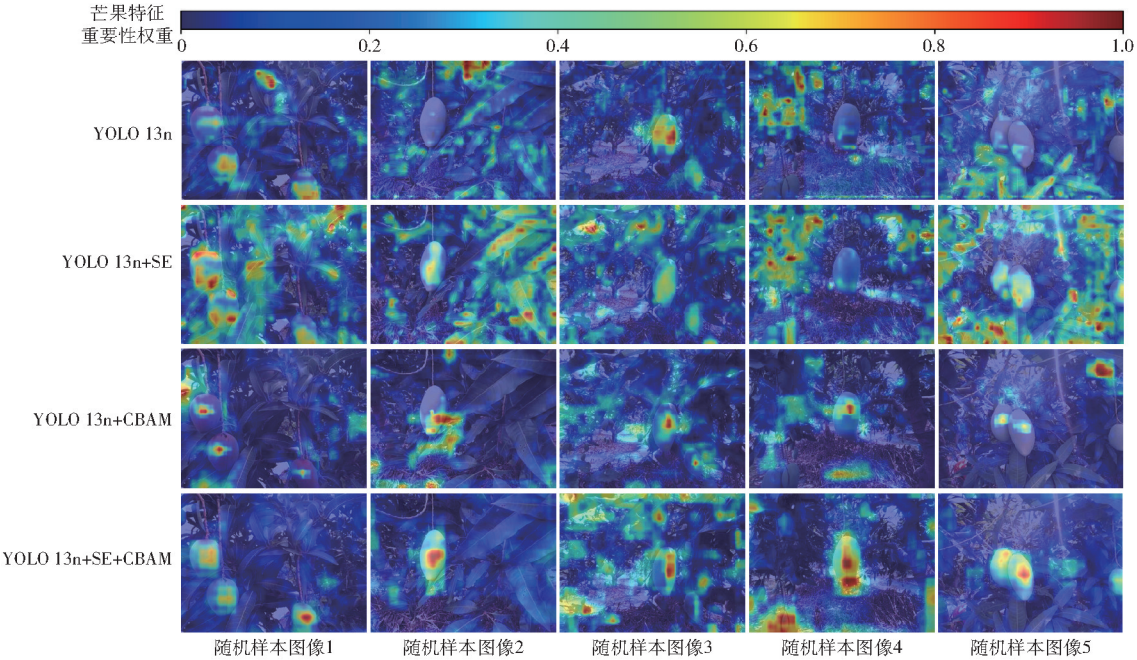
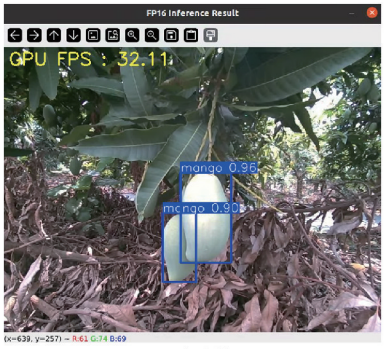


图 5 消融试验热力图

Fig. 5 Heat map of ablation experiment

相结合的模型部署方式能够获得较佳推理性能。为此，采用该种方式对本文提出的改进 YOLO 13n 模型进行部署（见图 6），以帧率作为改进 YOLO 13n 模型在 NVIDIA Jetson Orin Nano 边缘计算设备上的性能评价指标。经试验可得，在 NVIDIA Jetson Orin Nano 环境下，模型的帧率达到 31.71 f/s，相较于模型改进前的边缘计算部署，推理速度仅下降约 0.4 f/s，但仍高于 30 f/s，达到实时性芒果识别要求。



(a) 改进前

(b) 改进后

图 6　YOLO 13n 模型改进前后的边缘计算部署

Fig. 6　Edge computing deployment before and after improvement of YOLO 13n model

此外，随机选择 50 幅样本图像，分别基于部署的原始 YOLO 13n、改进 YOLO 13n 模型，测定其推理时间，结果如图 7 所示。由图 7 得知，50 幅随机样本图像在边缘计算部署下，改进前 YOLO 13n 模型的推理时间为 34.35 ~ 39.01 ms，其平均推理时间为 36.82 ms；而改进后 YOLO 13n 模型的推理时间为 34.43 ~ 38.85 ms，其平均推理时间为 36.88 ms，说明尽管改进 YOLO 13n 模型经历了架构重塑，总体平均推理时间略微增加（0.06 ms），但差异不明显，验证了芒果在边缘计算下实时识别的稳定性和可靠性。

为进一步验证改进 YOLO 13n 模型在 NVIDIA Jetson Orin Nano 环境下的综合性能，以 P、R、$AP_{0.5}$、$AP_{0.75}$、浮点运算次数和单幅图像推理时间作为模型在 NVIDIA Jetson Orin Nano 边缘计算设备上的性能评价指标。基于 PyTorch 深度学习框架，对改进前 YOLO 13n 模型、改进后 YOLO 13n 模型进行

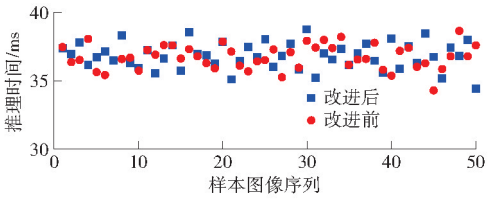


图 7　改进 YOLO 13n 模型在边缘计算上的推理时间

Fig. 7　Inference time of improved YOLO 13n model on edge computing

NVIDIA Jetson Orin Nano 部署，获得这两种模型在相同硬件与数据条件下的平均精度对比结果（见表 3）。对比结果可得，在边缘计算环境下，YOLO 13n 模型改进前后的精度结果与 1.5 节试验环境下的平均精度结果较为一致，表明尽管将模型部署到更为紧凑的边缘计算平台，模型仍然具有较高精度。

表 3　YOLO 13n 模型改进前后 NVIDIA Jetson Orin Nano 环境下性能对比结果

Tab. 3　Comparison of detection performance before and after YOLO 13n improvement on NVIDIA Jetson Orin Nano environment

模型	P/%	R/%	$AP_{0.5}$/%	$AP_{0.75}$/%	浮点运算次数
改进前 YOLO 13n	93.2	89.4	96.3	93.6	6.2×10^9
改进后 YOLO 13n	94.4	91.6	96.5	94.3	6.2×10^9

综上，通过上述平均精度、推理速度等方面数据对比，充分论证了本文提出方法在边缘计算设备上的综合性能优势。

3　讨论

针对机器人采摘应用场景低分辨率成像芒果识别精度不足问题，通过在 YOLO 13n 基线模型的骨干网络引入 SE 模块、在颈部网络添加 CBAM 模块，构建了改进 YOLO 13n 模型，其芒果识别性能（P 为 94.5%、R 为 91.2%、$AP_{0.5}$ 为 96.7%、$AP_{0.75}$ 为 94.9%）不仅优于 YOLO v8n、YOLO v9s、YOLO v10n、YOLO 11n 和 YOLO 12n 等主流轻量级模型，还较改进前 YOLO 13n 模型在 P、R、$AP_{0.5}$ 和 $AP_{0.75}$ 上分别提升 0.6、0.6、1.1、1.5 个百分点，且保持轻量化特征（浮点运算次数为 6.2×10^9，模型内存占用量为 5.7×10^6）。研究说明：采用“分阶段嵌入-串联融合”策略，将 SE 和 CBAM 这两种注意力机制重塑 YOLO 13n 模型，能够起到强化特征表达和突出芒果空间区域的作用，层次化协同提升总体目标检测能力，较好地实现了检测精度与模型轻量化的平衡。

动态移动采摘机器人平台受限于空间大小、能源供应等条件。近年来，一些研究者进行了深度学

习模型与边缘计算相结合的研究[15],以发展更为紧凑的采摘作业机器人,但是,上述研究对检测实时性方面略欠考虑。基于此,本文采用 TensorRT 推理优化器、C++编译、模型 API 相结合的部署方式,将构建的改进型 YOLO 13n 模型部署于 NVIDIA Jetson Orin Nano 边缘计算设备,单幅图像最高推理速度可达 31.71 f/s,50 幅随机样本平均推理时间仅 36.88 ms,相较于模型改进前,模型推理速度仅下降 0.4 f/s,仍然高于 30 f/s,说明本研究的部署方式能够有效提升模型推理速度,使其达到目标实时识别要求,为轻质化芒果采摘机器人的研发提供关键技术支撑,同时也为农业机器人在目标识别轻量化设计与边缘端高效部署的结合应用提供了可借鉴的技术方案。

尽管研究取得了一定的进展,但仍有一些不足之处。例如,本文只考虑了模型在 NVIDIA Jetson Orin Nano 环境下的部署,缺乏在其他边缘计算设备上的模型部署验证;此外,改进的 YOLO 13n 模型尚未进行模型在边缘计算上的能耗评估,这些都是未来研究需要探究的问题。

4 结论

(1)基于 YOLO 13n 基线模型,引入 SE 模块和 CBAM 模块,构建改进的 YOLO 13n 模型,能够有效提升芒果整体识别性能,体现了 SE 模块优异的特征表达能力,以及 CBAM 模块突出的空间信息提取能力。

(2)改进 YOLO 13n 模型在 NVIDIA Jetson Orin Nano 上的部署试验表明,使用 TensorRT 推理优化器、C++编译和模型 API 相结合的模型部署方式,模型推理速度高于30 f/s,满足芒果实时识别要求。

(3)鉴于采摘机器人受制于有限空间大小、能源供应等条件,以及边缘计算部署要求轻量级模型,通过构建改进的 YOLO 13n 模型,基于低分辨率成像的芒果图像,实现边缘计算下的实时芒果识别,为研发轻质化芒果采摘机器人提供技术支撑。

参 考 文 献

[1] MALDONADO-CELIS M E, YAHIA E M, BEDOYA R, et al. Chemical composition of mango (*Mangifera indica* L.) fruit: nutritional and phytochemical compounds[J]. Frontiers in Plant Science, 2019, 10: 1073.

[2] FAOSTAT[EB/OL]. [2025-09-08]. https://www.fao.org/faostat/zh/#data.

[3] CHEN C, LIU M Y, TUZEL O, et al. R-CNN for small object detection[C]// Proceeding of Asian Conference on Computer Vision. Cham: Springer, 2016: 214-230.

[4] PAYNE A B, WALSH K B, SUBEDI P P, et al. Estimation of mango crop yield using image analysis-segmentation method[J]. Computers and Electronics in Agriculture, 2013, 91: 57-64.

[5] PAYNE A, WALSH K, SUBEDI P, et al. Estimating mango crop yield using image analysis using fruit at 'stone hardening' stage and night time imaging[J]. Computers and Electronics in Agriculture, 2014, 100: 160-167.

[6] NANAA K, RIZON M, ABD RAHMAN M N, et al. Detecting mango fruits by using randomized hough transform and backpropagation neural network[C]// 2014 18th International Conference on Information Visualisation. IEEE, 2014: 388-391.

[7] STEIN M, BARGOTI S, UNDERWOOD J. Image based mango fruit detection, localisation and yield estimation using multiple view geometry[J]. Sensors, 2016, 16(11): 1915.

[8] CHANDA A, VORASEYANONT P, SIRICHAROEN P. Deep learning-based mango fruit detection and counting[C]// 2024 21st International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). IEEE, 2024: 1-5.

[9] 薛月菊,黄宁,涂淑琴,等. 未成熟芒果的改进 YOLOv2 识别方法[J]. 农业工程学报,2018,34(7): 173-179.
XUE Yueju, HUANG Ning, TU Shuqin, et al. Immature mango detection based on improved YOLOv2[J]. Transactions of the CSAE, 2018, 34(7): 173-179. (in Chinese)

[10] 熊俊涛,刘振,林睿,等. 自然环境下树上绿色芒果的无人机视觉检测技术[J]. 农业机械学报,2018, 49(11): 23-29.
XIONG Juntao, LIU Zhen, LIN Rui, et al. Unmanned aerial vehicle vision detection technology of green mango on tree in natural environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(11): 23-29. (in Chinese)

[11] XIONG J, LIU Z, CHEN S, et al. Visual detection of green mangoes by an unmanned aerial vehicle in orchards based on a deep learning method[J]. Biosystems Engineering, 2020, 194: 261-272.

[12] XU Z F, JIA R S, SUN H M, et al. Light-YOLO v3: fast method for detecting green mangoes in complex scenes using picking robots[J]. Applied Intelligence, 2020, 50(12): 4670-4687.

[13] ZHANG B, XIA Y, WANG R, et al. Recognition of mango and location of picking point on stem based on a multi-task CNN model named YOLOMS[J]. Precision Agriculture, 2024, 25(3): 1454-1476.

[14] LI H, HUANG J, GU Z, et al. Positioning of mango picking point using an improved YOLO v8 architecture with object detection and instance segmentation[J]. Biosystems Engineering, 2024, 247: 202-220.

(下转第 176 页)

[17] SHEN L, SU J, HE R, et al. Real-time tracking and counting of grape clusters in the field based on channel pruning with YOLO v5s[J]. Computers and Electronics in Agriculture, 2023, 206: 107662.
[18] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[J]. arXiv Preprint, arXiv:2410.17725, 2024.
[19] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021 - 12031.
[20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132 - 7141.
[21] WANG Q, WU B, ZHU P, et al. ECA - Net: efficient channel attention for deep convolutional neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 11534 - 11542.
[22] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 - 19.
[23] ZHANG Q L, YANG Y B. SA - Net: shuffle attention for deep convolutional neural networks[C]// ICASSP 2021—2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021: 2235 - 2239.
[24] 张宸嘉, 朱磊, 俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20):64 - 72.
ZHANG Chenjia, ZHU Lei, YU Lu. Review of attention mechanism in convolutional neural networks[J]. Computer Engineering and Applications, 2021, 57(20):64 - 72. (in Chinese)
[25] YANG L, ZHANG R Y, LI L, et al. Simam: a simple, parameter-free attention module for convolutional neural networks[C]// International Conference on Machine Learning. PMLR, 2021: 11863 - 11874.
[26] TAN M, LE Q. EfficientNetv2: smaller models and faster training[C]// International Conference on Machine Learning. PMLR, 2021: 10096 - 10106.
[27] TANG Y, HAN K, GUO J, et al. GhostNetv2: enhance cheap operation with long-range attention[J]. Advances in Neural Information Processing Systems, 2022, 35: 9969 - 9982.
[28] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314 - 1324.
[29] MA N, ZHANG X, ZHENG H T, et al. ShuffleNet v2: practical guidelines for efficient CNN architecture design[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 116 - 131.
[30] CHATTOPADHAY A, SARKAR A, HOWLADER P, et al. Grad-CAM + +: generalized gradient-based visual explanations for deep convolutional networks[C]// 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018: 839 - 847.

(上接第 **166** 页)

[15] GU Z, HE D, HUANG J, et al. Simultaneous detection of fruits and fruiting stems in mango using improved YOLO v8 model deployed by edge device[J]. Computers and Electronics in Agriculture, 2024, 227: 109512.
[16] ZHENG C, CHEN P, PANG J, et al. A mango picking vision algorithm on instance segmentation and key point detection from RGB images in an open orchard[J]. Biosystems Engineering, 2021, 206: 32 - 54.
[17] LI H, CHEN J, GU Z, et al. Optimizing edge-enabled system for detecting green passion fruits in complex natural orchards using lightweight deep learning model[J]. Computers and Electronics in Agriculture, 2025, 234: 110269.
[18] Ultralytics. YOLO v8[CP/OL]. (2023 - 01 - 10)[2025 - 10 - 08]. https://github.com/ultralytics/ultralytics.
[19] WANG C Y, YEH I H, LIAO H Y M. YOLO v9: learning what you want to learn using programmable gradient information[C]// European Conference on Computer Vision. Springer Nature Switzerland, 2024: 1 - 21.
[20] WANG A, CHEN H, LIU L, et al. YOLO v10: real-time end-to-end object detection[C]// arXiv Preprint, arXiv:2405.14458, 2024.
[21] KHANAM R, HUSSAIN M. YOLO v11: an overview of the key architectural enhancements[C]// arXiv Preprint, arXiv:2410.17725, 2024.
[22] TIAN Y, YE Q, DOERMANN D. YOLO v12: attention-centric real-time object detectors[C]// arXiv Preprint, arXiv:2502.12524, 2025.
[23] LEI M, LI S, WU Y, et al. YOLO v13: real-time object detection with hypergraph-enhanced adaptive visual perception[C]// arXiv Preprint, arXiv:2506.17733, 2025.