

基于 YOLOv13-HyperACE 的甘蔗整杆收获识别方法

林舒哈¹, 肖诗正¹, 侯明鑫^{1*}, 李星晨¹, 杨桢毅¹, 牛钊君², 孟庆河²

1. 广东海洋大学机械工程学院, 广东 湛江 524088

2. 中国热带农业科学院农业机械研究所, 广东 湛江 524013

DOI: 10.61369/TACS.2026050039

摘 要 : 针对甘蔗收获过程中整杆甘蔗的精准识别问题, 提出一种基于改进 YOLOv13 目标检测模型的识别方法。通过构建甘蔗整杆数据集, 引入超图的自适应关联增强注意力 HyperACE(Hypergraph-based Adaptive Correlation Enhancement), 提升模型对密集堆叠甘蔗的检测性能。实验结果表明, 该方法在复杂场景下对单根及堆叠甘蔗的平均识别精度 (mAP) 达到 92.3%, 较原始 YOLOv13 提升 6.7%, 满足机械臂抓取对甘蔗整杆收获识别的实时性与准确性需求。

关 键 词 : 甘蔗整杆识别; 收获; YOLO; 深度学习; 抓取

Sugarcane Whole Stalk Harvest Recognition Method Based on YOLOv13-HyperACE

Lin Shuhan¹, Xiao Shizheng¹, Hou Mingxin^{1*}, Li Xingchen¹, Yang Zhenyi¹, Niu Zhaojun², Meng Qinghe²

1. School of Mechanical Engineering, Guangdong Ocean University, Zhanjiang, Guangdong 524088

2. Agricultural Machinery Research Institute, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang, Guangdong 524013

Abstract : Addressing the issue of precise identification of whole sugarcane stalks during the harvesting process, a recognition method based on an improved YOLOv13 object detection model is proposed. By constructing a dataset of whole sugarcane stalks and introducing HyperACE (Hypergraph-based Adaptive Correlation Enhancement), an adaptive correlation enhancement method based on hypergraph, the detection performance of the model for densely stacked sugarcane is improved. Experimental results show that the average precision (mAP) of this method for single and stacked sugarcane in complex scenes reaches 92.3%, which is 6.7% higher than that of the original YOLOv13, meeting the real-time and accuracy requirements for mechanical arm grasping and harvesting identification of whole sugarcane stalks.

Keywords : sugarcane whole stalk recognition; harvesting; YOLO; deep learning; capture

引言

甘蔗作为广西省重要的糖料作物, 其机械化收获水平直接影响农业生产效率^[1-3]。在整杆式收获作业中, 机械臂需精准识别田间单根及小堆堆叠的甘蔗目标, 以实现无损抓取^[4-6]。然而, 田间环境复杂多变, 甘蔗常呈现细长形态且易发生密集堆叠与相互遮挡, 传统视觉识别方法难以在保证实时性的同时兼顾检测精度, 导致机械臂抓取成功率受限^[7-9]。

针对上述难点, 本文提出一种基于改进 YOLOv13 (You Only Look Once version 13) 的甘蔗整杆识别方法。通过构建专用甘蔗数据集, 并引入超图自适应关联增强注意力机制 (HyperACE) 优化策略, 旨在提升模型对密集粘连目标的特征提取能力与定位精度^[10]。实验表明, 该方法能有效适应复杂田间工况, 为甘蔗收获机械臂的精准作业提供了可靠的技术支撑。

一、材料与方法

(一) 数据集构建

为适配田间甘蔗整杆收获的复杂识别场景, 本文以广西甘蔗主产区的实际收获工况为依据, 构建甘蔗整杆收获目标检测专用数据集。数据采集于自然光照、多云、局部阴影、田间杂草与残

膜干扰等多种环境, 采集设备为分辨率 1280×720 的工业相机, 共获取田间收获状态下甘蔗图像 1580 张, 覆盖单根整杆、2-6 根小堆堆叠、局部重叠遮挡三类典型目标。使用 Label Studio 标注工具对样本进行人工标注, 统一类别标签为 sugarcane, 以最小外接矩形完整包裹甘蔗杆体或堆叠整体, 确保目标边界与实际抓取区域一致。按 7:2:1 的比例将数据集随机划分为训练集、验证集与

项目信息:

广西重点研发计划项目 (桂科 AB23026069;

桂农科 AB241484034);

广东省自然科学基金 (2025A1515012901);

广东海洋大学创新训练项目 (202410566012)。

通讯作者: 侯明鑫, 教授 / 博士, houmx@gdou.edu.cn

测试集，其中训练集1106张、验证集316张、测试集158张，各子集样本分布保持均衡。



图1 基于 LabelStudio 的甘蔗整杆数据集标注

(二) 模型改进方法

针对甘蔗整杆在田间自然生长状态下呈现的细长杆状几何特征，以及在机械化收获过程中极易出现的密集堆叠、相互遮挡、光照反射不均等复杂视觉干扰问题，本文选取当前性能卓越的 YOLOv13 作为基础检测框架。然而，标准的 YOLOv13 模型虽然在通用目标检测任务中表现出极高的实时性与准确性，但在处理甘蔗这种长宽比悬殊 (Aspect Ratio Extreme)、且堆叠区域特征极易混淆的特定农业场景时，其骨干网络 (Backbone) 与特征金字塔 (Neck) 往往难以建立长距离的语义关联，导致对堆叠甘蔗边界框的回归精度不足，易产生漏检或误检。

为解决上述痛点，本研究提出了一种基于超图自适应关联增强注意力机制 (Hypergraph-based Adaptive Correlation Enhancement, HyperACE) 的改进策略。该策略的核心思想在于打破传统卷积神经网络 (CNN) 仅通过局部感受野进行特征提取的局限，利用超图 (Hypergraph) 结构强大的非欧几里得数据建模能力，显式地捕捉甘蔗目标之间复杂的高阶空间关联。具体而言，本研究将 HyperACE 模块深度嵌入至 YOLOv13 的 Neck 网络关键特征融合层中，旨在通过自适应增强密集区域的特征响应，抑制背景噪声干扰，从而提升模型对整杆及小堆甘蔗的精细化识别能力。

1. HyperACE 模块的数学建模与构建逻辑

HyperACE 模块的设计初衷是为了解决密集目标检测中的“特征混淆”问题。在田间场景中，多根甘蔗往往紧密贴合，传统注意力机制 (如 SE、CBAM) 仅关注通道或空间的二阶统计量，难以区分紧密相邻的目标实例。为此，本研究构建了一个基于超图的注意力增强单元，其数学表达过程如下：

首先，假设输入特征图 $F_m \in \mathbb{R}^{C \times H \times W}$ 经过降维处理后得到节点特征矩阵 $X \in \mathbb{R}^{N \times d}$ ，其中 $N=H \times W$ 为特征节点数， d 为特征维度。HyperACE 模块的核心在于构建一个自适应的超图结构 $H=(V, E)$ ，其中 V 代表特征节点集合， E 代表超边集合。

与传统图结构中一条边仅连接两个节点不同，超图允许一条超边 $e_k \in E$ 连接任意数量的节点，这使得模型能够将属于同一根甘蔗 (即具有高度相关性的特征点) 的像素点聚合在一起。超边的构建并非预定义，而是通过可学习的参数动态生成：

$$S = \sigma(X\Phi W_s) \quad (1)$$

式 (1) 中， $S \in \mathbb{R}^{N \times K}$ 为关联矩阵 (Incidence Matrix)， K 为超边数量， Φ 为映射函数， W_s 为可学习权重， σ 为 Sigmoid 激活函数。该过程实现了对甘蔗特征的自适应聚类。

随后，引入超图卷积层进行信息传播。为了增强模型对甘蔗细长形态的适应性，本研究设计了一种加权的超图拉普拉斯算子：

$$F_{out} = D_v^{-1} H P D_e^{-1} W_{att} X \quad (2)$$

式 (2) 中， H 为关联矩阵， P 为超边权重矩阵， D_v 和 D_e 分别为节点度矩阵和超边度矩阵。 W_{att} 是本模块引入的自适应注意力权重，用于在聚合邻居信息时，放大属于甘蔗茎秆的特征响应，抑制土壤、杂草等背景噪声的干扰。通过这种高阶关联建模，模型能够学习到甘蔗之间“谁与谁相邻”、“谁属于同一簇”的深层语义信息，从而有效区分堆叠在一起的多根甘蔗。

2. 基于 YOLOv13 的网络架构融合策略

在具体的网络架构实现上，本文并未盲目堆叠模块，而是经过消融实验验证，选择将 HyperACE 模块嵌入在 YOLOv13 的 Neck 部分的特征金字塔 (FPN) 关键节点。具体位置位于 P3/P4/P5 特征层的上采样与下采样融合路径之后。

选择该位置的逻辑在于：在 Backbone 提取的基础特征中，甘蔗的细长纹理信息尚处于初级阶段，而经过 Neck 的特征融合后，多尺度特征已经初步聚合。此时引入 HyperACE，可以利用超图结构对融合后的特征进行“语义精修”。它不仅能够利用深层特征的语义信息来指导浅层特征的定位 (Top-down refinement)，还能通过自适应关联增强，将分散在不同尺度上的甘蔗特征 (如根部、中部、梢部) 重新关联，形成一个完整的实例表达。

为了保证模型的实时性以满足机械臂抓取的控制周期要求，本研究对 HyperACE 模块的计算复杂度进行了优化。通过引入分组卷积 (Group Convolution) 和降维策略，将超图构建的节点数控制在合理范围内，避免了因图计算带来的过大的计算开销。同时，在激活函数的选择上，采用了最新的 SiLU (Sigmoid Linear Unit) 变体，以提升梯度的流动效率。

3. 针对甘蔗收获场景的特异性优化

甘蔗收获作业通常在户外强光、阴影交替的复杂光照条件下进行，本研究在 HyperACE 模块中额外增加了一个光照鲁棒性分支。该分支通过并行的全局平均池化 (Global Average Pooling) 和最大池化 (Global Max Pooling)，捕捉图像的全局光照分布特征，并将该特征作为先验知识注入到超图的节点特征中。这一设计使得模型在面对甘蔗表面因光照产生的高光或阴影遮挡时，依然能够保持稳定的特征提取能力，有效防止因局部光照变化导致的特征断裂。

综上所述，本文提出的 YOLOv13-HyperACE 模型，通过在特征融合阶段引入超图自适应关联增强机制，不仅保留了 YOLOv13 原有的高速推理优势，更显著增强了其对密集堆叠甘蔗的高阶语义理解能力。这种改进为后续机械臂手的精准抓取提供了更为精确的二维定位信息与实例分割掩码，是实现甘蔗整杆智能化收获的关键技术环节。

二、实验与结果分析



图2 YOLOv13-HyperACE 整杆甘蔗识别结果

为全面评估改进后的 YOLOv13-HyperACE 模型在甘蔗整杆收获场景下的检测性能，本研究在自建的甘蔗整杆数据集上进行了详尽的对比实验与消融实验。实验环境配置如下：操作系统为 Ubuntu 24.04 LTS，深度学习框架采用 PyTorch 2.4.1，GPU 加速使用 NVIDIA RTX5090D（24GB 显存），CPU 为 Intel Core i9-14900K。训练过程中，批量大小（Batch size）设置为 32，初始学习率设为 0.01，动量（Momentum）为 0.937，并采用余弦退火策略进行学习率衰减。为了验证模型的泛化能力，所有实验均进行了 300 个 Epoch 的训练，并选取验证集上 mAP 最高的权重作为最终模型。

（一）模型整体性能对比分析

为了验证本文提出的 YOLOv13n-HyperACE 算法的有效性，本研究选取了原始 YOLOv13 模型、以及目前主流的 YOLOv8 和 YOLOv10 作为基准对比模型。评价指标涵盖了平均精度均值（mAP@0.5）、精确率（Precision）、召回率（Recall）以及单张图片推理时间（FPS）。

实验结果如表 1 所示。从数据中可以看出，原始 YOLOv13 模型在测试集上的 mAP@0.5 为 85.6%，虽然在通用目标检测中表现优异，但在面对田间密集堆叠的甘蔗场景时，其精确率和召回率均出现了明显的瓶颈。相比之下，引入 HyperACE 机制后的改进模型表现出了显著的性能提升。实验数据显示，改进后的模型 mAP@0.5 达到了 92.3%，较原始模型提升了 6.7 个百分点。这一数据的提升得益于 HyperACE 模块对甘蔗细长杆状特征的高阶关联建模能力。同时，改进模型的精确率（Precision）达到了 91.8%，较原始模型提升了 6.2%。这意味着在基于本文方法的机械臂抓取作业中，视觉系统产生误报（如将杂草误认为甘蔗）的概率大幅降低，从而保证了甘蔗整杆抓取动作的可靠性。同时，召回率（Recall）提升至 90.5%，提升了 7.1%，这表明模型在面对甘蔗相互遮挡或光照不均的情况时，漏检率显著下降，能够更完整地识别出田间所有的收获目标。

表 1 不同模型在甘蔗数据集上的性能对比

模型名称	参数量 (M)	mAP@0.5 (%)	精确率 (%)	召回率 (%)	推理速度 (ms)
YOLOv8s	3.0	84.2	85.1	83.3	25.4
YOLOv10s	2.3	86.5	87.9	85.2	22.1
YOLOv13n (原始)	3.1	85.6	85.6	83.4	18.9
YOLOv13n-HyperACE (本文)	3.4	92.3	91.8	90.5	19.2

表 1 表明，虽然引入了 HyperACE 模块导致模型参数量略有增加（从 .1M 增加到 3.4M），但推理时间仅增加了 0.3ms，维持在 19.2ms（约 52FPS），依然满足机械臂实时控制的帧率要求（通常 >30FPS）。这有力地证明了本文提出的算法在精度与速度之间取得了良好的平衡，具备在嵌入式农业机器人平台上部署的潜力。

（二）定性分析与可视化结果讨论

除了定量数据，本研究还选取了多组具有代表性的田间复杂

场景图像进行可视化分析，以直观展示模型的改进效果。

在甘蔗处于强光照射下，且背景中存在大量反光的残膜。原始 YOLOv13 模型由于受到高光干扰，将部分残膜误检为甘蔗（False Positive），同时漏检了阴影中的一根甘蔗。而本文提出的模型，得益于 HyperACE 中的光照鲁棒性分支，能够有效抑制背景噪声，准确地识别出了所有目标，且边界框回归非常精准。

在密集堆叠场景中，多根甘蔗相互交叉重叠。原始模型在重叠区域产生了严重的粘连现象，仅输出了一个巨大的边界框，无法精准识别。相比之下，改进后的模型利用超图的自适应关联能力，成功将重叠的甘蔗分离识别为独立的个体，清晰地勾勒出了每一根甘蔗的轮廓。这种像素级的精准定位，为后续机械臂手的无损分离抓取提供了可靠的视觉依据。

三、结束语

本文针对甘蔗整杆收获过程中密集堆叠、局部遮挡导致识别精度不足的问题，提出一种基于 YOLOv13HyperACE 的整杆甘蔗识别方法。通过构建田间甘蔗专用数据集，在 YOLOv13 网络中嵌入 HyperACE 超图自适应关联增强注意力，有效提升了模型对细长杆状目标与密集堆叠场景的检测性能。实验结果表明，所提方法 mAP@0.5 达 92.3%，较原始模型提升 6.7 个百分点，在复杂田间环境下具备良好的精度与鲁棒性，可稳定输出目标定位信息，满足机械臂精准抓取的工程需求。该研究可为甘蔗智能收获装备的视觉系统提供有效技术方案，对推动南方糖料作物收获机械化、智能化发展具有实际应用价值。

参考文献

[1] 刘金雨, 王冠琳, 廖柳生等. 新发展格局下广西甘蔗产业转型升级研究 [J]. 大众科技, 2026, 28(01): 105-109.

[2] 黄晓玲. 广西糖用甘蔗高产栽培管理技术及其应用实践的深入讨论 [J]. 种子世界, 2026(01), 07-209.

[3] 欧剑荣, 匡昭敏, 李莉等. 广西甘蔗机收作业气候适宜度评价 [J]. 甘蔗糖业, 2025, 54(02): 53-61.

[4] 文飞. 甘蔗收获机械化技术与收获机械推广分析 [J]. 广西农业机械化, 2022(03): 16-18.

[5] 朱昌炜. 基于双目视觉的田间甘蔗空间定位及抓取研究 [D]. 南宁: 广西大学, 2022.

[6] 吴鹏坤. 甘蔗收获机械推广应用存在的问题与措施研究 [J]. 河北农机, 2021(10): 91-92.

[7] 张邦. 基于机器视觉的堆叠蔗种识别与自动取投技术研究 [D]. 桂林: 桂林理工大学, 2025.

[8] 李隆钦. 基于双目 DSO 算法的蔗根定位与抓取研究 [D]. 南宁: 广西大学, 2024.

[9] 胡秋生. 机器人柔性触觉末端设计及触觉协同抓取研究 [D]. 合肥: 合肥工业大学, 2024.

[10] 廖伟辉, 赖景暖, 屠博等. 基于 DSA-YOLOv13 的钢结构表面缺陷检测方法 [J]. 环境技术, 2025, 43(10): 86-93.