

DOI: 10.13652/j.spjx.1003.5788.2025.80514

基于改进 YOLOv11n 模型的香蕉成熟度识别方法

胡俐蕊¹ 王佳星^{2,3} 胡泽坤⁴

(1. 北部湾大学电子与信息工程学院, 广西 钦州 535000; 2. 桂林理工大学计算机科学与工程学院, 广西 桂林 541006; 3. 桂林理工大学广西嵌入式技术与智能系统重点实验室, 广西 桂林 541004; 4. 北部湾大学机械与船舶海洋工程学院, 广西 钦州 535000)

摘要: [目的] 提高香蕉成熟度的识别效率。[方法] 提出基于改进 YOLOv11n 模型的香蕉成熟度识别方法。在 YOLOv11n 模型中, 引入改进后的极化自注意力机制, 提升骨干网络在多个香蕉果实分布场景下的特征提取能力; 使用内容感知的特征重组模块替换原始上采样, 增大感受野, 更有效地聚合上下文信息; 将斯库拉交并比损失函数作为新的边界框损失函数, 计算真实框与预测框之间的向量角度, 更好地解决真实框与预测框的匹配问题, 降低漏检和错检机率。[结果] 改进后的方法在平均精度均值 0.50 和平均精度均值 0.50~0.95 指标上分别提高了 1.4% 和 3.0%, 识别精度高于现有方法。[结论] 改进方法有效提升了香蕉成熟度识别的准确性和效率, 具有较高的实用价值。

关键词: 香蕉成熟度; YOLOv11n 模型; 极化自注意力; 内容感知的特征重组; 斯库拉交并比

Recognition for banana ripeness based on improved YOLOv11n

HU Lirui¹ WANG Jiaying^{2,3} HU Zekun⁴

(1. College of Electronics and Information Engineering, Beibu Gulf University, Qinzhou, Guangxi 535000, China; 2. College of Computer Science and Engineering, Guilin University of Technology, Guilin, Guangxi 541006, China; 3. Guangxi Key Laboratory of Embedded Technology and Intelligent System, Guilin University of Technology, Guilin, Guangxi 541004, China; 4. College of Mechanical and Marine Engineering, Beibu Gulf University, Qinzhou, Guangxi 535000, China)

Abstract: [Objective] To improve the efficiency of recognition for banana ripeness. [Methods] A method for recognizing banana ripeness is developed based on improved YOLOv11n. A modified polarized self-attention mechanism is introduced into YOLOv11n to enhance the feature extraction capability of the backbone network across various banana distribution scenarios. The original upsampling is replaced with a module of content-aware reassembly of features, which enlarges the receptive field to more effectively aggregate contextual information. Scylla intersection over union (SIoU) is adopted as the new bounding box loss, which calculates the vector angle between ground truth and predicted boxes to better address the matching problem between them and reduce instances of missed and false detection. [Results] The improved method achieves increases of 1.4% and 3.0% in mean Average Precision 0.50 (mAP_{0.50}) and mean Average Precision 0.50~0.95 (mAP_{0.50-0.95}), respectively, with the recognition accuracy surpassing other existing methods. [Conclusion] The proposed method effectively enhances the accuracy and efficiency of recognition for banana ripeness, demonstrating high practical value.

Keywords: banana ripeness; YOLOv11n; polarized self-attention; content-aware reassembly of features; scylla intersection over union

作为全球重要水果之一, 香蕉在不同成熟阶段呈现 色, 再到过熟的褐色^[1]。在贮藏、分拣、运输和销售过程中, 成熟度是香蕉分级的关键指标, 对其品质有显著影

基金项目: 广西重点研发计划项目(编号: 桂科 AB25069378); 钦州市科技计划项目(编号: 202116602)

通信作者: 胡泽坤(1995—), 男, 北部湾大学助理研究员, 硕士。E-mail: 843537018@qq.com

收稿日期: 2025-06-04 改回日期: 2025-08-20

引用格式: 胡俐蕊, 王佳星, 胡泽坤. 基于改进 YOLOv11n 模型的香蕉成熟度识别方法[J]. 食品与机械, 2026, 42(2): 126-132.

Citation: HU Lirui, WANG Jiaying, HU Zekun. Recognition for banana ripeness based on improved YOLOv11n[J]. Food & Machinery, 2026, 42(2): 126-132.

响。传统上,成熟度依赖于人员的主观判断,但这种方法易受个体差异影响,导致结果不一致^[2]。此外,人工检测耗时费力,尤其在大规模生产和供应链中,可能导致效率低下并增加成本。

目前,香蕉成熟度的自动识别方法备受关注,在多种技术路线和算法模型上取得了显著进展。Da Conceição Luiz等^[3]通过分析香蕉不同成熟阶段化学成分引起的红外光谱特征变化,实现了成熟度识别。莫松涛等^[4]截取了香蕉果实上、中、下3个部位的横切面,提取棱角数值作为特征,结合遗传算法优化的支持向量机模型,成熟度判别准确率达到86.2%。王灵敏等^[5]提出了一种基于AlexNet^[6]和Softmax分类器的香蕉图像分类算法,准确率达95.56%。针对AlexNet在移动设备应用中的局限性,蒋瑜等^[7]进一步设计了轻量化的Mini-AlexNet模型,通过调整网络结构并采用全局平均池化,大幅降低了参数量和内存开销。Kahfi等^[8]利用灰度共生矩阵与K近邻算法构建了香蕉成熟度判别模型,该模型在成熟度识别方面具有较高的准确性。韩雪等^[9]将视觉几何组网络(VGGNet^[10])特征提取与极限梯度提升算法相结合,并通过贝叶斯优化调整超参数,成功建立了高精度的香蕉成熟度判别方法,准确率达到91.25%。

尽管上述方法在单个香蕉果实的成熟度识别中取得了较好效果,但在实际应用中仍存在诸多挑战。光谱技术易受环境因素影响,导致测量结果不稳定且设备成本较高;图像分类方法虽取得一定进展,但难以实现对果实的精确定位,且多集中于单果识别,难以满足多果实复杂场景的需求。相比之下,目标检测技术不仅能够精确定位香蕉图像中的多个果实,还能实现果实成熟度的有效识别,因此更适用于复杂的实际应用场景。

在目标检测算法中,YOLO模型因其高效且实时识别的能力被广泛应用。试验拟基于YOLOv11n模型进行香蕉成熟度识别,并改进YOLOv11n模型在提取颜色渐变特征方面的不足。将经过改进的极化自注意力(PSA^[11])机制替换跨阶段局部金字塔挤压注意力(C2PSA)模块中的多头注意力机制,以增强骨干网络的特征提取能力。将原始上采样模块替换为内容感知的特征重组(CARAFE^[12])模块,动态生成上采样核,更有效地聚合上下文信息。使用斯库拉交并比(SIoU^[13])损失函数代替原来的完整交并比(CIoU^[14])损失函数,设计预测框与真实框的角度匹配机制,使得预测框在训练时能够有方向地与真实框进行匹配,提高识别准确率,并降低漏检和错检机率。该试验旨在为香蕉成熟度智能识别提供技术支撑。

1 材料与方法

1.1 图像收集

采用Roboflow平台上广受认可且具有较高影响力的

香蕉成熟度公开数据集(<https://app.roboflow.com/wob/banana-ripeness-skmyq>),该数据集共有5 234张图像,包含6个类别:新鲜成熟、新鲜未成熟、过熟、成熟、腐烂和未成熟。将数据集按7:2:1划分为训练集、测试集和验证集,部分香蕉图像如1所示。

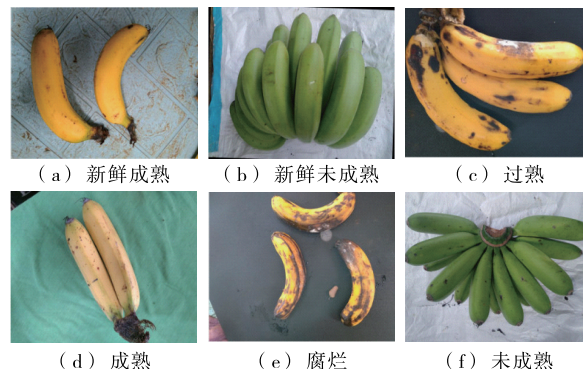


图1 部分不同成熟度的香蕉图像

Figure 1 Bananas at varying stages of ripeness

1.2 改进后的模型

YOLOv11n模型是一款全新一阶段目标检测算法,旨在提升检测的准确性和速度。该模型的主干网络负责特征提取与信息优化两项关键任务,可有效增强对多尺度及复杂场景特征的捕捉能力。模型颈部采用单向特征金字塔结构,通过融合浅层与深层特征,显著提升语义表达能力,增强目标的识别与定位效果。试验对YOLOv11n模型结构进行了优化调整,改进后的模型结构如图2所示。首先,将C2PSA中的多头注意力机制替换为改进的极化自注意力机制,进一步融合空间与通道维度信息,使模型能够在不同尺度下精准捕捉与香蕉成熟相关的关键特征,改进后的模块命名为C2PSA-imp;其次,优化模型颈部,将传统的上采样操作替换为内容感知的特征重组CARAFE结构,扩大感受野并改善上下文信息的聚合,更好地利用成熟香蕉的特征,提升识别精度;最后,采用形状感知交并比SIoU损失函数取代完整交并比CIoU损失函数,通过计算真实框与预测框的向量角度,增强边界框在形状和方向上的一致性,显著提升回归的稳定性与精确度。改进后的模型在复杂场景下表现更加优异。改进后的模型命名为YOLOv11n-PCS。

1.2.1 改进C2PSA 在C2PSA模块中,原有的多头注意力机制模块虽然能够捕捉局部空间细节,但在建模通道间的全局关系时存在局限性,难以充分挖掘整体颜色分布等全局语义信息,从而限制了对香蕉成熟度特征的全面理解。为了解决这个问题,引入了改进的PSA注意力机制,以替代原有的多头注意力机制。

PSA为像素级回归任务设计的融合空间与通道注意

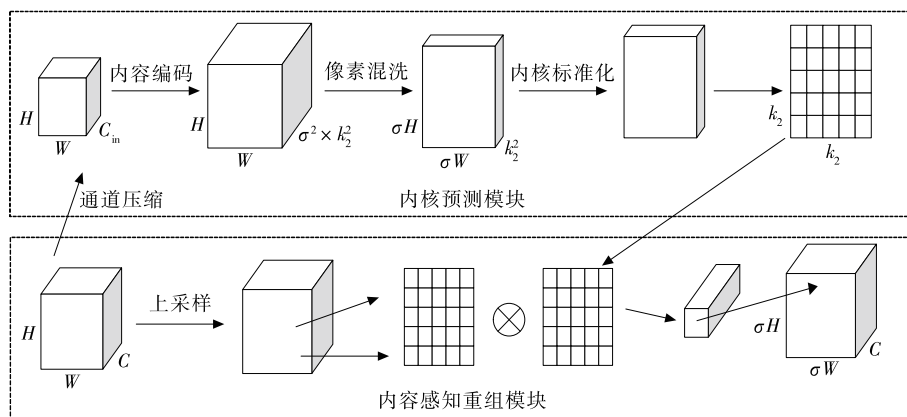


图4 CARAFE 模块图

Figure 4 CARAFE module

本,提升CARAFE的效率。②内容编码将压缩后的特征图作为输入,通过使用卷积核大小为 $k_2 \times k_2$ 的卷积层对内容进行编码生成重组核。③使用像素混洗方法将通道维在空间维展开,得到形状为 $\sigma H \times \sigma W \times k_2^2$ 的上采样重组核,每个重组核在空间上使Softmax函数进行归一化。归一化强制将特征值的总和为1,从而使得CARAFE算法能够更灵活地处理不同区域的特征,而不引起整体缩放或平均值的改变。

特征重组模块将预测出的上采样核与特征图的每一层特征进行逐层乘积,生成最终的上采样输出,这使得模型能够动态地调整上采样操作,提高对不同特征的适应性。

1.2.3 改进损失函数 YOLOv11n模型在计算边界框损失时采用了CIoU损失函数,忽视了预测框与真实框之间的方向不匹配问题。当预测框与真实框的方向不一致时,可能导致预测框在训练中随机移动,进而减缓收敛速度,降低训练效率,影响模型性能。

为解决此问题,引入SIoU损失函数,通过计算真实框与预测框之间的向量角度,有效克服了方向不匹配带来的影响。其计算式为:

$$L_{\text{SIoU}} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2}, \quad (1)$$

$$\Delta = (1 - e^{(\Lambda - 2)\rho_x})^\theta + (1 - e^{(\Lambda - 2)\rho_y})^\theta, \quad (2)$$

$$\Lambda = 1 - 2 \sin^2 \left(\arcsin(x) - \frac{\pi}{4} \right), \quad (3)$$

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega t}), \quad (4)$$

式中:

Δ ——距离损失;

Λ ——角度损失;

Ω ——形状损失;

ρ_x, ρ_y ——真实框的中心点与预测框的中心点XY轴

的差值;

t ——当前计算的维度变量;

w, h ——宽、高;

θ ——形状损失的非线性调整参数;

ω ——形状损失的缩放参数。

2 结果与分析

2.1 实现环境

该模型训练在Linux操作系统的环境中完成,所用硬件设备包括NVIDIA RTX 3060 GPU(显存容量为8GB)以及Intel Xeon E5-2683 CPU。训练过程采用Python 3.8编程语言,基于PyTorch 2.0.1深度学习框架实现,并配合CUDA 11.8版本支持GPU加速计算。训练参数方面,为了促进梯度更新的稳定性和收敛速度,动量设置为0.937;为了抑制模型过拟合,权重衰减设置为0.0005;为了充分学习数据中的特征信息,模型训练200轮。

2.2 消融试验

为确保改进后模型的有效性与合理性,以YOLOv11n为基础模型进行消融试验,结果见表1。组合1~组合5分别引入不同的模块。由表1可知,前3个组合在参数量和计算量未显著增加的情况下,平均精度均值0.50(mAP_{0.50})分别提升了1.1,0.2,0.4个百分点,平均精度均值0.50~0.95(mAP_{0.50-0.95})分别提升了2.2,0.4,0.8个百分点,验证了各模块的有效性。组合4相较于组合1在两个指标上略有提升,而组合5相较于组合4在mAP_{0.50-0.95}上提升了0.6个百分点,且未增加额外参数。最终,改进模型相较于YOLOv11n在mAP_{0.50}和mAP_{0.50-0.95}上分别提升了1.4和3.0个百分点,表明该改进的有效性。

为了验证改进后的模型在香蕉成熟度识别方面的良好表现,随机选取一组图像进行测试,结果如图5所示。由图5可知,对于正常光的第一幅图像、第二幅图像和昏暗光图像,原始模型YOLOv11n分别漏检了1个腐烂

表 1 消融试验结果

Table 1 Ablation study results

模型	C2PSA-imp	CARAFE	SIoU	mAP _{0.50} /%	mAP _{0.50-0.95} /%	参数量/M	计算量/G
YOLOv11n				95.1	74.0	2.6	6.3
组合 1	✓			96.2	76.2	2.6	6.3
组合 2		✓		95.3	74.4	2.7	6.5
组合 3			✓	95.5	74.8	2.6	6.3
组合 4	✓	✓		96.4	76.4	2.7	6.5
组合 5	✓	✓	✓	96.5	77.0	2.7	6.5

(rotten)香蕉、将 2 个香蕉的新鲜成熟 (fresh ripe) 错误识别为成熟 (ripe)、将 1 个香蕉的成熟 (ripe) 错误识别为过熟 (overripe)。改进后的模型则在检测香蕉个体时表现优异,准确识别了所有香蕉及其成熟度,说明改进模型在香蕉成熟度识别任务中的更强能力和更高可靠性。

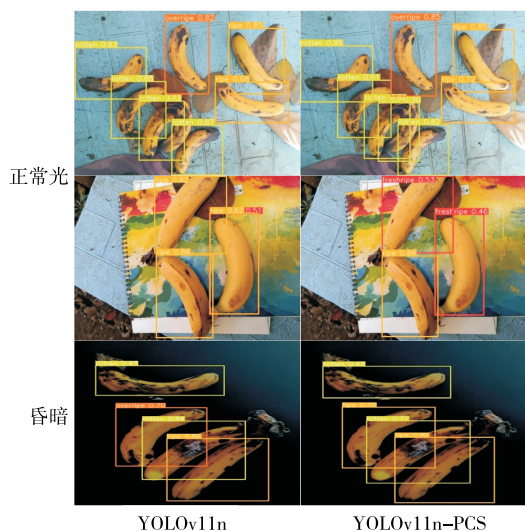


图 5 模型改进前后推理对比

Figure 5 Inference comparison before and after model improvement

2.3 注意力机制对比试验

为了验证 PSA-imp 注意力机制在特征提取阶段的优越性,将其与 PSA^[11]、Effective Squeeze and Extraction Attention^[18]、Coordinate Attention^[19]、Shuffle Attention^[20]和 Normalization-based Attention^[21]进行对比,结果见表 2。由表 2 可知,PSA-imp 在 mAP_{0.50} 和 mAP_{0.50-0.95} 两个指标上均表现出优异性能,相较于其他注意力机制均有显著提升,表明 PSA-imp 能更有助于骨干网络提取特征信息,抑制背景干扰。

2.4 损失函数对比试验

为了验证 SIoU 损失函数在衡量预测框与真实框之间误差方面的优越性,将 SIoU^[13]与 CIoU^[14]、高效交并比

表 2 不同注意力机制对比结果

Table 2 Comparison of different attention mechanisms

注意力机制	mAP _{0.50} /%	mAP _{0.50-0.95} /%
PSA	95.5	75.3
PSA-imp	96.2	76.2
Effective Squeeze and Extraction Attention	94.9	74.1
Coordinate Attention	94.8	73.9
Shuffle Attention	95.4	74.2
Normalization-based Attention	94.7	73.7

(EIoU^[22])、优化型交并比 (WIoU^[23]) 损失函数进行对比,结果见表 3。由表 3 可知,SIoU 损失函数在 mAP_{0.50} 和 mAP_{0.50-0.95} 两个评价指标上均取得了最佳表现,分别达到了 96.5% 和 77.0%,优于 CIoU、EIoU 和 WIoU 损失函数。因此,选取 SIoU 作为最终模型的损失函数,凸显了其在香蕉成熟度识别任务中的更佳性能和适用性。

表 3 不同损失函数对比结果

Table 3 Comparison of different loss functions

损失函数	mAP _{0.50} /%	mAP _{0.50-0.95} /%
CIoU	96.4	76.4
EIoU	96.2	76.4
WIoU	96.3	76.7
SIoU	96.5	77.0

2.5 不同目标检测模型对比试验

为了深入验证 YOLOv11n-PCS 模型的优越性,对多种目标检测算法进行比较,结果见表 4。由表 4 可知,相较于二阶段目标检测算法快速区域卷积神经网络 (Faster R-CNN^[24]),YOLOv11n-PCS 在 mAP_{0.50} 和 mAP_{0.50-0.95} 上分别高出了 7.4 和 17.0 个百分点,同时在参数量和计算量方面远远小于 Faster R-CNN。在与一些经典的一阶段算法如单次多框检测器 (SSD^[25])、Retina Net^[26]、全卷积单阶段目标检测器 (FCOS^[27])、实时检测转换器 (RT-DETR^[28]) 的对比中,YOLOv11n-PCS 的各项指标也展现出显著优势。最后,与 YOLO 系列算法的比较中,原始的 YOLOv11n 识

别精度指标上超越了之前的 YOLO 系列算法,而 YOLOv11n-PCS 在性能提升方面进一步突破,在 $mAP_{0.50}$ 和 $mAP_{0.50-0.95}$ 上分别高出了 1.4 和 3.0 个百分点,同时参数数量和计算量也只有略微增加。综上,试验提出的 YOLOv11n-PCS 能够在保持参数数量和计算量相当的同时进一步提高香蕉成熟度的识别精度,减少漏检和错检问题。

表 4 不同目标检测模型对比结果

Table 4 Comparison of different object detection models

模型	$mAP_{0.50}/\%$	$mAP_{0.50-0.95}/\%$	参数量/ M	计算量/ G
Faster R-CNN	89.1	60.0	28.3	470.5
SSD	94.6	66.8	24.4	30.7
Retina Net	90.9	63.0	36.4	73.5
FCOS	96.5	72.7	32.1	80.7
RT-DETR	91.8	70.0	32.0	103.5
YOLOv5s	92.8	67.7	7.0	15.8
YOLOv7-tiny	94.2	65.9	6.0	13.1
YOLOv8n	94.3	73.0	3.0	8.1
YOLOv9t	94.8	73.8	2.1	10.7
YOLOv10n	94.0	72.7	2.7	8.2
YOLOv11n	95.1	74.0	2.6	6.3
YOLOv11n-PCS	96.5	77.0	2.7	6.5

2.6 已有方法对比

为了验证所提方法的优越性,对所使用的数据集进行筛选,选取仅包含单个香蕉果实的图像组成新的子集。筛选后的图像数量为 1 845 张,随后按照 7:2:1 比例重新划分为训练集、验证集和测试集。基于该数据集开展试验,并将试验方法与 VGGNet+ 极限梯度提升^[9]、AlexNet^[5]、Mini-AlexNet^[7]以及灰度共生矩阵+K 近邻^[8]方法进行对比,结果见表 5。由表 5 可知,YOLOv11n-PCS 模型在香蕉成熟度识别任务中表现最为优异。相比于结合深度学习的 VGGNet+ 极限梯度提升模型,YOLOv11n-PCS 在准确率和召回率方面均表现更佳,体现出其在特征提取和分类效果上的优势。尽管 AlexNet 和 Mini-AlexNet 在深度学习模型中表现良好,但其指标仍未达到 YOLOv11n-PCS 的水平,尤其在召回率方面存在明显差距。灰度共生矩阵+K 近邻方法的各项指标均低于 YOLOv11n-PCS,进一步验证了该模型在香蕉成熟度检测上的综合优势。

总体来看,YOLOv11n-PCS 不仅在准确率上领先,而且在精确率和召回率方面表现出较高的稳定性和鲁棒性,充分彰显了其在香蕉成熟度自动识别领域的应用潜力和实用价值。

表 5 已有方法对比

Table 5 Comparison of existing methods %

方法	准确率	精确率	召回率
VGGNet+ 极限梯度提升	91.5	90.1	88.7
AlexNet	88.7	87.6	88.1
Mini-Alexnet	86.2	84.1	85.9
灰度共生矩阵+K 近邻	80.2	82.2	78.5
YOLOv11n-PCS	92.1	91.2	90.8

3 结论

为了提升实际环境中香蕉成熟度的识别效率,试验基于改进的 YOLOv11n 目标检测模型完成了香蕉成熟度识别任务。通过引入改进的 PSA 注意力机制、CARAFE 模块和 Siou 损失函数,显著提高了模型在香蕉成熟度识别中的性能。改进后的模型在 $mAP_{0.50}$ 和 $mAP_{0.50-0.95}$ 指标上分别提升了 1.4% 和 3.0%,同时保持了较低的参数量。尽管试验使用的数据集包含了 6 种不同成熟度的香蕉图像,数据集的规模和质量依然对模型性能产生了重要影响。未来可以着重于更大规模且更加多样化的数据集进一步验证模型的泛化能力。

参考文献

[1] 陈铭中, 钟旭美, 林海生, 等. UV-C 处理对采后香蕉贮藏防御性成分和品质的影响[J]. 江苏农业学报, 2022, 38(2): 528-538.
CHEN M Z, ZHONG X M, LIN H S, et al. Effects of UV-C treatment on defensive components and quality of postharvest banana during storage[J]. Jiangsu Journal of Agricultural Sciences, 2022, 38(2): 528-538.

[2] KOOTSTRA G, WANG X, BLOK P M, et al. Selective harvesting robotics: current research, trends, and future directions[J]. Current Robotics Reports, 2021, 2(1): 95-104.

[3] DA CONCEIÇÃO LUIZ L, NASCIMENTO C A, BELL M J V, et al. Use of mid infrared spectroscopy to analyze the ripening of Brazilian bananas[J]. Food Science and Technology, 2022, 42: e74221.

[4] 莫松涛, 董涛, 赵汐璇, 等. 基于遗传算法与 SVM 的香蕉果实成熟度判别模型[J]. 果树学报, 2022, 39(12): 2 418-2 427.
MO S T, DONG T, ZHAO X X, et al. Discriminant model of banana fruit maturity based on genetic algorithm and SVM[J]. Journal of Fruit Science, 2022, 39(12): 2 418-2 427.

[5] 王灵敏, 蒋瑜. 基于深度学习的香蕉成熟度自动分级[J]. 食品与机械, 2022, 38(11): 149-154.
WANG L M, JIANG Y. Automatic classification of banana ripeness based on deep learning[J]. Food & Machinery, 2022, 38 (11): 149-154.

[6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J].

- Communications of the ACM, 2017, 60(6): 84-90.
- [7] 蒋瑜, 王灵敏. 基于改进 Alexnet 的轻量化香蕉成熟度检测[J]. 食品与机械, 2024, 40(5): 128-136.
JIANG Y, WANG L M. Lightweight banana ripeness detection based on improved Alexnet[J]. Food & Machinery, 2024, 40(5): 128-136.
- [8] KAHFI A H, HASAN M, HASANAH R L. Classification of banana ripeness based on color and texture characteristics[J]. Journal of Computer Networks, Architecture and High Performance Computing, 2022, 5(1): 10-17.
- [9] 韩雪, 张磊, 赵雅菲, 等. 基于 CNN 和 XgBoost 的香蕉成熟度判别[J]. 食品与机械, 2024, 40(4): 127-135, 178.
HAN X, ZHANG L, ZHAO Y F, et al. Banana ripeness determination based on CNN and XgBoost[J]. Food & Machinery, 2024, 40(4): 127-135, 178.
- [10] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]// Proceedings of the International Conference on Learning Representations (ICLR). San Diego, USA: [s.n.], 2015: 1-14.
- [11] LIU H J, LIU F Q, FAN X Y, et al. Polarized self-attention: Towards high-quality pixel-wise mapping[J]. Neurocomputing, 2022, 506: 158-167.
- [12] WANG J Q, CHEN K, XU R, et al. CARAFE: content-aware ReAssembly of FEatures[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2019: 3 007-3 016.
- [13] GEVORGYAN Z. SIoU loss: more powerful learning for bounding box regression[DB/OL]. (2022-05-25) [2025-08-20]. <https://arxiv.org/pdf/2205.12740>.
- [14] ZHENG Z H, 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(7): 12 993-13 000.
- [15] LIU S J, HE N, WANG C, et al. Lightweight human pose estimation algorithm based on polarized self-attention[J]. Multimedia Systems, 2023, 29(1): 197-210.
- [16] ELFWING S, UCHIBE E, DOYA K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning[J]. Neural Networks, 2018, 107: 3-11.
- [17] JIANG Y W. Surface defect detection of steel based on improved YOLOv5 algorithm[J]. Mathematical Biosciences and Engineering, 2023, 20(11): 19 858-19 870.
- [18] LEE Y, PARK J. CenterMask: real-time anchor-free instance segmentation[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA: IEEE, 2020: 13 903-13 912.
- [19] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN: IEEE, 2021: 13 708-13 717.
- [20] 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). Toronto, ON: IEEE, 2021: 2 235-2 239.
- [21] LIU Y, SHAO Z, TENG Y, et al. NAM: normalization-based attention module[DB/OL]. (2021-11-24) [2025-08-20]. <https://arxiv.org/pdf/2111.12419>.
- [22] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [23] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[DB/OL]. (2023-01-24) [2025-08-20]. <https://arxiv.org/pdf/2301.10051>.
- [24] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1 137-1 149.
- [25] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]// Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [26] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]// 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 2 999-3 007.
- [27] TIAN Z, SHEN C H, CHEN H, et al. FCOS: fully convolutional one-stage object detection[C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 9 626-9 635.
- [28] ZHAO Y A, LV W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection[C]// 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA: IEEE, 2024: 16 965-16 974.