

DOI: 10.13733/j.jcam.issn.2095-5553.2025.02.038

李涛, 买买提明·艾尼, 古丽巴哈尔·托乎提, 等. 基于改进YOLOv8的小棚架下无核白葡萄果梗识别[J]. 中国农机化学报, 2025, 46(2): 259-263, 278

Li Tao, Mamtimin·Geni, Gulbahar·Tohti, et al. Seedless white grape stem recognition under small trellises based on improved YOLOv8 [J]. Journal of Chinese Agricultural Mechanization, 2025, 46(2): 259-263, 278

基于改进 YOLOv8 的小棚架下无核白葡萄果梗识别 *

李涛¹, 买买提明·艾尼^{1, 2}, 古丽巴哈尔·托乎提¹, 杨佳雨¹

(1. 新疆大学智能制造现代产业学院, 乌鲁木齐市, 830047;

2. 乌鲁木齐佰博机电科技有限公司, 乌鲁木齐市, 830002)

摘要:小棚架下准确识别无核白葡萄果梗是葡萄采摘机器人自动采摘任务的关键。针对新疆吐鲁番市小棚架下鲜食无核白葡萄果梗识别效果差的问题,提出一种基于 YOLOv8 的无核白葡萄果梗识别模型 Small—YOLO,实现无核白葡萄果梗的自动识别。在原有的模型结构上改进目标检测头,提高浅层与深层网络的特征融合,增加对无核白葡萄果梗信息提取的能力。在浅层网络中采用可变形卷积 DCN 以增强卷积操作对形态变化的建模能力,使得卷积核可以更好地适应输入特征图中的不规则变形,有助于提高视觉模型在处理不同尺度、形态和变换目标时的性能。增加坐标注意力机制 CA,优化无核白葡萄果梗识别的准确率。结果表明,改进后的识别模型对无核白葡萄果梗平均精度 mAP 值达到 76.2%。与 YOLOv3—tiny、YOLOv5n、YOLOv6、YOLOv7、YOLOv8n 等算法相比, mAP 值分别提升 23.9%、8%、7.6%、9.2%、7%,同时保持较快的检测速度,实现在小棚架下无核白葡萄机械采摘可能性。

关键词:无核白葡萄;采摘机器人;果梗识别;坐标注意力机制;可变形卷积;视觉模型

中图分类号:S663.1; TP391.4 **文献标识码:**A **文章编号:**2095-5553 (2025) 02-0259-06

Seedless white grape stem recognition under small trellises based on improved YOLOv8

Li Tao¹, Mamtimin·Geni^{1, 2}, Gulbahar·Tohti¹, Yang Jiayu¹

(1. School of Intelligent Manufacturing Modern Industry, Xinjiang University, Urumqi, 830047, China;

2. Urumqi Baybol Metronocs Technology Co. Ltd., Urumqi, 830002, China)

Abstract: Accurate identification of seedless white grape stem under small trellises is the key to the automatic picking task of grape picking robot. Aiming at the problem of poor recognition effect of fresh seedless white grape stem under Small trellises in Turfan City, Xinjiang, a seedless white grape stems recognition model Small—YOLO based on YOLOv8 was proposed to realize automatic recognition of seedless white grape stem, improve the target detection head on the original model structure, and improve the feature fusion of shallow and deep networks. The ability of extracting information from the seedless white grape stem was increased. Deformable convolutional DCN is used in the shallow network to enhance the modeling ability of convolution operation on morphological changes, so that the convolutional kernel can better adapt to the irregular deformation in the input feature map, which is helpful to improve the performance of the visual model when dealing with targets of different scales, shapes and transformations, increase the coordinate attention mechanism CA, optimize the accuracy rate of the seedless white grape peduncle recognition. The test results showed that the mAP value of the improved recognition model for the seedless white grape stem reached 76.2%. Compared with YOLOv3—tiny, YOLOv5n, YOLOv6, YOLOv7, YOLOv8n and other algorithms, the mAP value was improved by 23.9%, 8%, 7.6%, 9.2%, 7%, respectively, while maintaining a fast detection speed, realizing the possibility of the seedless white grape mechanical picking under small tallows.

Keywords: seedless white grapes; harvesting robot; stem recognition; coordinate attention mechanism; deformable convolution; visual model

收稿日期:2023 年 10 月 24 日 修回日期:2023 年 12 月 18 日

* 基金项目:国家自然科学基金资助项目(12162031)

第一作者:李涛,男,1995 年生,四川南充人,硕士;研究方向为机械设计和图像识别。E-mail: 1378800938@qq.com

通讯作者:买买提明·艾尼,男,1958 年生,乌鲁木齐人,博士,教授;研究方向为农牧机械设计理论与方法、图像识别等。E-mail: mgjeni@263.com

0 引言

2022 年,中国葡萄产量 15 378 kt,新疆葡萄产量完成 3 165 kt,占全国葡萄总产量的 20.58%^[1],新疆吐鲁番市葡萄种植面积达 36 km²,占新疆葡萄种植总面积的 39.6%,占全国葡萄种植总面积的 8.1%^[2],其中无核白葡萄产量占比 90%^[3]。吐鲁番市葡萄种植采用小棚架模式,小棚架高 1.6 m,宽 2 m,长 10~50 m,葡萄架低矮,采摘高度较为低矮,人工采摘效率低,用工季节葡萄采摘费昂贵。在图像处理领域,对于小目标的定义有多种方式,一种是将大小低于图像高度 10% 的检测目标或大小低于图像高度的 20% 的识别目标定义为小目标^[4]。另一种方式,如 MSCOCO 数据集^[5]则将像素绝对值小于 32×32 的目标定义为小目标。

在现有葡萄识别研究中,识别对象主要是葡萄果实,且通常针对篱架和大棚架种植模式下的葡萄,这些场景相对开阔。Liu 等^[6]在研究葡萄采摘机器人中提出一种葡萄串分割算法,检测准确率为 88.0%。Reis 等^[7]在夜间利用闪光灯拍摄白葡萄与红葡萄在进行图像处理得到葡萄串,正确率为 97%。Wu 等^[8]使用 YOLOv5 目标检测算法来识别葡萄串,并应用关键点检测模型来定位葡萄果梗,平均准确率约为 92%。Qiu 等^[9]用一种基于改进 YOLOv4 的葡萄成熟度检测和视觉预定定位算法,平均准确率达到 93.52%,平均检测时间为 10.82 ms。罗陆锋等^[10]采用基于深度学习的葡萄果梗识别与最优采摘定位方法,利用 Mask R-CNN 模型在不同天气光照下该方法对葡萄果梗的检测精确率平均值为 88%。李慧鹏等^[11]使用 PSPNet(MobileNetv2)语义分割模型分割葡萄图像,采用霍夫变换检测果梗边缘上的直线段并进行直线拟合,采摘点定位准确率为 91.94%,定位时间为 187.47 ms。然而,在吐鲁番地区的小棚架环境中,针对吐鲁番特有的无核白葡萄的果梗识别研究尚属空白。

本文提出一种基于 YOLOv8n 的无核白葡萄果梗识别模型(Small-YOLO),为评估 Small-YOLO 模型的性能,分析吐鲁番小棚架下无核白葡萄果梗数据集,在无核白葡萄果梗数据集中,进行 YOLOv3-tiny、YOLOv5n、YOLOv6、YOLOv7 和 YOLOv8n 模型试验对比。此外,还分析基于 YOLOv8n 改进 Small 模型、Small+DCN 模型、Small+CA 模型和 Small-YOLO 模型的性能,得到各个模型改进的有效性。

1 数据获取

数据来源于新疆维吾尔自治区吐鲁番市高昌区(北纬 41°12'~43°40',东经 87°16'~91°55')。数据采集

时间是葡萄采收时期,采集设备是 OnePlus Ace 2V, vivo 手机。为确保数据的多样性与可靠性,分别采集经过农艺处理、近景、强光和遮光等场景中无核白葡萄果梗图像,共 3 994 张(分辨率为 3 468 像素×4 624 像素的 RGB 图像)。图 1 为小棚架下无核白葡萄果梗样本。



(a) 农艺处理 (b) 近景 (c) 强光

图 1 无核白葡萄果梗样本

Fig. 1 Sample of seedless white grape stem

2 改进 YOLOv8n 模型

2.1 YOLOv8n 模型

YOLOv8n 模型主要由输入端、主干网络、颈部网络和输出端 4 部分组成。YOLOv8n^[12]在输入端会对输入图片大小进行归一化处理。主干网络主要由 Conv^[13]、C2f、SPPF 等模块组成。Conv 包括 Conv2d、BatchNorm2d 和 SiLU 激活函数。C2f 模块是对 C3 模块的改进,主要用来增强特征融合能力,提高检测速度,实现模型整体的轻量化。SPPF 是 YOLOv8n 对 SPP^[14]模块改进的一个空间金字塔池化层,可以扩展接受域,实现局部和全局特征融合,丰富特征信息。颈部网络通过 FPN^[15]+PAN^[16]的结构来解决特征信息的多尺度融合问题。输出端采用解耦头结构将分类和检测头分离,使用 Anchor-Free^[17]的思想。YOLOv8n 模型结构如图 2 所示。

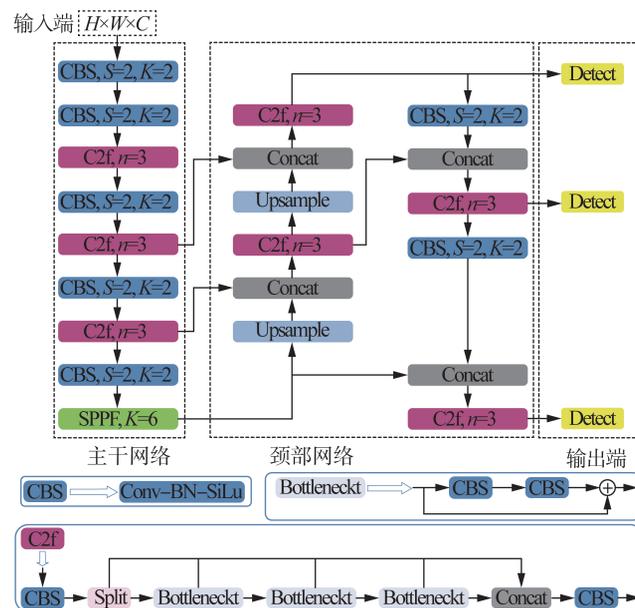


图 2 YOLOv8n 模型结构

Fig. 2 YOLOv8n model structure

2.2 Small—YOLO 模型

2.2.1 YOLOv8n 模型的改进

随着网络模型层数增多,深层特征图的感受野更大,但分辨率低,特征图信息更少;浅层的特征图含有小的感受野,特征细节信息丰富。采用改进的 Small 模型将 YOLOv8n 模型主干网络第 2 层与颈部网络 14 层融合引出一个目标检测头,提取浅层的特征图中无核白葡萄果梗的特征信息,解决无核白葡萄果梗特征信息随着网络模型层数增多而特征信息丢失问题,由于无核白葡萄果梗数据集中没有大目标,将深层检测头去掉,减少模型参数量,在此基础上使用 160×160 、 80×80 、 40×40 网络的检测头,得到的整体信息更加全面。Small—YOLO 模型结构是在 Small 模型基础上,为解决浅层的特征图中无核白葡萄果梗特征信息提取不充分问题,将浅层的标准卷积替换为可变形卷积 DCN,在主干网络和颈部网络添加 CA 坐标注意力机制^[18],分别加在第 3 层和 21 层,充分提取无核白葡萄果梗的特征。Small—YOLO 模型结构如图 3 所示。

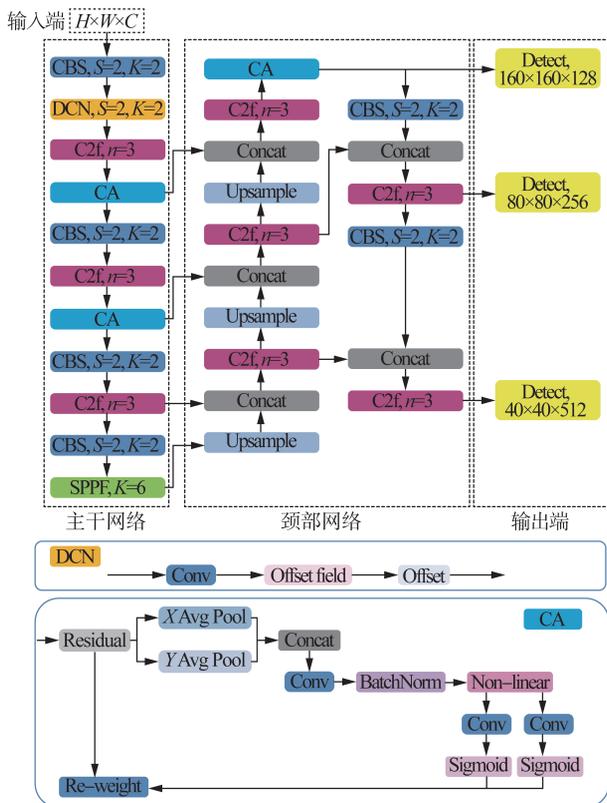


图 3 Small—YOLO 模型结构

Fig. 3 Small—YOLO model structure

2.2.2 坐标注意力机制 CA

受小棚架种植模式影响,光线、遮挡因素严重,无核白葡萄果梗识别效果不好,采用坐标注意力机制 CA 增强网络模型对无核白葡萄果梗特征提取能力。坐标注意力机制 CA 的核心思想是利用水平和垂直方

向上的平均池化计算坐标信息,将坐标信息与原始特征图相结合,从而提高特征的提取能力。坐标注意力机制 CA 的工作流程包括以下步骤:首先,输入特征图经过自适应平均池化,得到水平和垂直方向上的均值信息;然后,这两个均值信息被连接在一起,经过卷积和归一化操作,降低通道数;接下来,特征图被分割为水平和垂直部分,分别通过卷积层计算注意力权重;最后,特征图根据注意力权重进行加权,得到增强后的表示。捕获特征图中的细节信息和上下文信息,且具有较小计算开销。

2.2.3 可变形卷积 DCN

标准卷积核通常采用规则的采样点对输入特征图进行卷积操作。在同一卷积层中,所有卷积核的感受野都相同。由于输入特征图中不同位置可能对应不同尺度、形态、视角和几何变化的目标,如何适应这些目标的多样性是视觉识别的一项重要挑战。为解决该问题,采用可变形卷积模块,以增强卷积操作对形态变化的建模能力。如图 3 所示,可变形卷积 DCN 对输入特征图进行卷积生成位置偏移量,根据位置偏移量对卷积和进行像素位置变换,输出特征图。可变形卷积引入对卷积核中每个采样点位置的自适应偏移,从而实现灵活的采样,不再受限于规则的格点采样,使得卷积核可以更好地适应输入特征图中的不规则变形,有助于提高视觉模型在处理不同尺度、形态和变换目标时的性能。

2.3 平台配置与训练策略

试验平台在 Windows10 操作系统和 cudnn11.0 环境下进行, GPU 配置 NVIDIA GeForce RTX3070, 显存 8 GB, 调用 GPU 进行训练, 中央处理器 Intel (R) Core (TM) i7—10875 H CPU @2.30 GHz, 运行内存为 16 GB, 固态硬盘 500 GB, 所有程序在 pytorch1.12 深度学习框架下用 Python 语言编写, 并使用 NVIDIA CUDA11.6 并行运算驱动加速训练。

在所有试验训练参数设置中,输入图像大小为 640 像素 \times 640 像素。优化器采用 SGD 驱动量优化器。初始学习率设置为 0.001,衰减系数为 0.01,动量为 0.9,批次大小为 16 个,共进行 300 轮训练。在训练过程中使用了热身训练,为提高模型的鲁棒性,在模型训练过程中使用马赛克增强,通过随机剪裁、缩放、翻转、色彩变化的图像增强操作后,再随机拼接成一张图片进行训练,丰富数据的多样性。

2.4 评价指标

采用召回率 R 、精确率 P 、平均精度均值 mAP 和每秒检测图像帧数 FPS 作为评价指标。

$$R = \frac{TP}{TP + FN} \times 100\% \quad (1)$$

$$P = \frac{TP}{TP + FP} \times 100\% \quad (2)$$

$$mAP = \int_0^1 P(R) dR \quad (3)$$

式中： TP ——模型将正样本识别为正样本，即正确识别出目标的情况；

FP ——模型将正样本识别为负样本，即未被识别到目标的情况；

FN ——负样本被识别为正样本，即背景被错误认为是目标的情况。

3 结果与分析

3.1 无核白葡萄果梗数据集分析

从最初的 3 994 张 RGB 图像中剔除没有果梗、虚影、拖影、重复的图像后，剩余 3 435 张 RGB 图像作为原始数据集。使用 YOLO 格式的数据集，采用 LabelImg 图像标注工具在 RGB 图像上进行葡萄果梗标注。为精确识别无核白葡萄果梗，其余未标注部分由 LabelImg 默认为背景。将标记好的图像按照 8:1:1 的比例划分成训练集、验证集与测试集，其中训练集图像 2 747 张，验证集图像 344 张，测试集图像 344 张。如图 4 所示。无核白葡萄果梗框的宽相对于整幅图比例集中在 0~0.025，无核白葡萄果梗框高相对于整幅图比例集中在 0~0.15，可以看出无核白葡萄果梗更细。如图 5 所示，果梗框的面积相对于整幅图面积占比 0~0.005 为主要集中区域，因此无核白葡萄果梗属于小目标的检测范围。

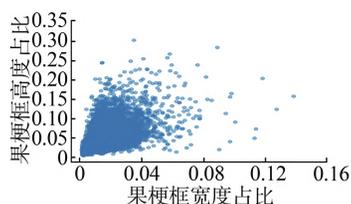


图 4 无核白葡萄果梗框的宽高相对于整图比例

Fig. 4 Width and height of the seedless white grape stem frame relative to the whole map

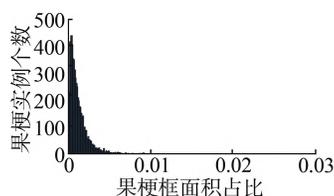


图 5 无核白葡萄果梗框面积相对于整图占比分布

Fig. 5 Area of stem frame of the seedless white grape stem relative to the whole map accounted for the proportion of layout map

3.2 YOLOv8n 模型与其他模型对比

为验证 YOLOv8n 在无核白葡萄果梗数据集上检测效果，分别将 YOLOv3—tiny、YOLOv5n、YOLOv6、YOLOv7 和 YOLOv8n 五种网络模型在无核白葡萄数据集上进行训练，在相同迭代次数内，YOLOv8n 的 mAP 值比 YOLOv3—tiny、YOLOv5n、YOLOv6 网络模型高 16.9%、1%、0.6%、2.2%，说明 YOLOv8n 模型对无核白葡萄果梗特征提取效果更好，试验结果如表 1 所示。

表 1 5 种模型在无核白葡萄果梗数据集检测效果对比

Tab. 1 Comparison of detection performance on the seedless white grape stem dataset for five models

模型	$mAP/\%$	$P/\%$	$R/\%$	$FPS/(\text{帧} \cdot \text{s}^{-1})$
YOLOv3—tiny	52.3	66.9	46.4	21.6
YOLOv5n	68.2	73.7	62.9	17.2
YOLOv6	68.6	76.2	62.6	16.9
YOLOv7	67.0	73.2	64.5	14.0
YOLOv8n	69.2	76.2	66.6	21.5

3.3 消融试验

将 YOLOv8n 模型作为基准模型，以平均精度均值 mAP 、精确率 P 、召回率 R 和 FPS 作为评价指标，在无核白葡萄果梗数据集上进行 4 组消融试验，验证不同改进点的有效性。通过对 YOLOv8n 模型改进进行消融试验分别取模型在无核白葡萄果梗验证集进行验证，试验结果如表 2 所示。

表 2 消融试验性能指标对比

Tab. 2 Comparison of performance of ablation experiment

模型	$mAP/\%$	$P/\%$	$R/\%$	$FPS/(\text{帧} \cdot \text{s}^{-1})$
基准模型	69.2	76.2	66.6	21.5
Small	74.0	75.0	65.0	13.0
Small+DCN	74.8	76.3	68.3	14.5
Small+CA	75.2	77.8	71.6	20.5
Small—YOLO	76.2	78.0	70.4	25.0

Small 模型表示对 YOLOv8n 模型检测头进行改进，Small+DCN 模型表示在 Small 模型的基础上将主干网络第 2 层的标准卷积替换成可变形卷积 DCN，Small+CA 模型表示在 Small 模型的基础上，主干网络第 3 层和第 6 层以及颈部网络第 20 层添加坐标注意力机制 CA，Small—YOLO 模型是融合以上 3 种模型。改进 Small—YOLO 相对 YOLOv8n 基准模型在 mAP 、 P 、 R 上分别高 7%、2%、3.8%， FPS 基本保持不变。采用 Small 模型相对基准模型在 mAP 、 P 、 R 上均有提升，但是 FPS 下降了 40%，采用 Small+DCN 模型使得召回率 R 提升 1.7%，精确率 P 提升 0.1%，平均精度均值 mAP 值提升 5.6%，采用 Small+CA 模型使得平均精度均值 mAP 值提升 6%，精确率 P 提升 1.6%，召回率 R 提升 5%。可见所用方法都能有效提升网络性能。

3.4 小棚架下无核白葡萄果梗识别效果

利用改进后的 Small—YOLO 模型对无核白葡萄果梗数据集进行训练与验证,并在验证集中进行测试,测试效果如图 6 所示。将无核白葡萄果梗识别效果图中果梗局部放大,发现无核白葡萄在小棚架下农艺处理后、近景、遮光远景及强光场景中,无核白葡萄果梗都能被识别出来,说明 Small—YOLO 模型在复杂情景中鲁棒性较好。枝叶与果梗十分相近的情况下也会出现错误识别情况,在错误识别果梗的情况下,其置信度较低,因此可以通过设置置信度阈值来减少错误识别,在后续的实际部署上可以规避此类问题。

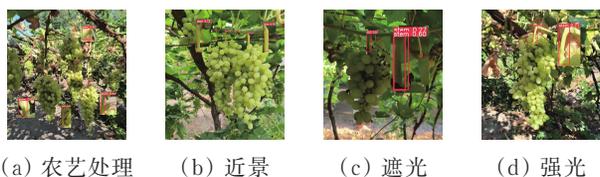


图 6 小棚架下无核白葡萄果梗识别效果

Fig. 6 Effect of identification of the seedless white grape stem under small trellises

4 结论

提出 Small—YOLO 模型实现无核白葡萄果梗识别,完成鲜食无核白葡萄无损采摘机器人的关键环节,建立无核白葡萄果梗数据集,改进模型检测头,采用 DCN 动态卷积和 CA 坐标注意力机制提升模型识别效果,实现对小目标识别,并与其他算法相比。

1) 在无核白葡萄果梗数据集上,相比 YOLOv3—tiny、YOLOv5 n、YOLOv6、YOLOv7、YOLOv8 模型,Small—YOLO 模型在平均精度均值上分别提高 23.9%、8%、7.6%、9.2%、7%,在准确率上提高 11.1%、4.3%、1.8%、4.8%、1.8%,在召回率上提升 24%、7.5%、7.8%、5.9%、3.8%。

2) 无核白葡萄果梗数据集中,果梗面积占比在 0.005 以下,属于小目标,采用改进 Small—YOLO 模型的 FPS 由 21.5 帧/s 提高至 25 帧/s,有效提升模型识别果梗性能。

3) 针对目标特征提取不充分问题,Small—YOLO 模型采用坐标注意力机制 CA 和将浅层标准卷积替换为可变形卷积 DCN,有效提升模型的识别性能。该模型特别针对吐鲁番市小棚架下无核白葡萄的果梗进行识别,为采摘机器人在无核白葡萄果梗采摘作业中的应用提供研究基础。

参 考 文 献

[1] 国家统计局. 中国统计年鉴[J]. 北京: 中国统计出版社,

2022.

- [2] 吾尔尼沙·卡得尔, 刘凤之, 刘丽媛. 新疆吐鲁番葡萄产业发展及转型升级建议[J]. 中国果树, 2021(11): 94—97.
- [3] 满保德. 新疆吐鲁番地区葡萄产业发展对策研究[D]. 石河子: 石河子大学, 2017.
- [4] Akyon F C, Altinuc S O, Temizel A. Slicing aided hyper inference and fine-tuning for small object detection [C]. 2022 IEEE International Conference on Image Processing (ICIP). IEEE, 2022: 966—970.
- [5] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context [C]. Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6—12, 2014, Proceedings, Part V 13. Springer International Publishing, 2014: 740—755.
- [6] Liu S, Whitty M. Automatic grape bunch detection in vineyards with an SVM classifier [J]. Journal of Applied Logic, 2015, 13(4): 643—653.
- [7] Reis M J C S, Morais R, Peres E, et al. Automatic detection of bunches of grapes in natural environment from color images [J]. Journal of Applied Logic, 2012, 10(4): 285—290.
- [8] Wu Z, Xia F, Zhou S, et al. A method for identifying grape stems using keypoints [J]. Computers and Electronics in Agriculture, 2023, 209: 107825.
- [9] Qiu C, Tian G, Zhao J, et al. Grape maturity detection and visual pre-positioning based on improved Yolov4 [J]. Electronics, 2022, 11(17): 2677.
- [10] 罗陆锋, 邹湘军, 熊俊涛, 等. 自然环境下葡萄采摘机器人采摘点的自动定位[J]. 农业工程学报, 2015, 31(2): 14—21.
- Luo Lufeng, Zou Xiangjun, Xiong Juntao, et al. Automatic positioning for picking point of grape picking robot in natural environment [J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(2): 14—21.
- [11] 李惠鹏, 李长勇, 李贵宾, 等. 基于深度学习的多品种鲜食葡萄采摘点定位[J]. 中国农机化学报, 2022, 43(12): 155—161.
- Li Huipeng, Li Changyong, Li Guibing, et al. Picking point positioning of multi-variety table grapes based on deep-learning [J]. Journal of Chinese Agricultural Mechanization, 2022, 43(12): 155—161.
- [12] Terven J, Córdova-Esparza D M, Romero-González J A. A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO—NAS [J]. Machine Learning and Knowledge Extraction, 2023, 5(4): 1680—1716.

(下转第 278 页)

林业大学学报(自然科学版), 2024, 48(4): 93-103.

Zhao Yugang, Liu Wenping, Zhou Yan, et al. UAV forestry land-cover image segmentation method based on attention mechanism and improved DeepLabV3+ [J]. Journal of Nanjing Forestry University (Natural Sciences Edition), 2024, 48(4): 93-103.

- [9] Gao C, Zhao F, Zhang Y, et al. Research on multitask model of object detection and road segmentation in unstructured road scenes [J]. Measurement Science and Technology, 2024, 35(6): 065113.
- [10] Sotelo M A, Rodriguez F J, Magdalena L, et al. A color vision-based lane tracking system for autonomous driving on unmarked roads [J]. Autonomous Robots, 2004, 16(1): 95-116.
- [11] 石金进. 基于视觉的智能车辆道路识别与障碍物检测方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2018.
- [12] Yang G, Wang Y, Yang J, et al. Fast and robust vanishing point detection using contourlet texture detector for unstructured road [J]. IEEE Access, 2019, 7: 139358-139367.
- [13] 王谦, 何朗, 王展青, 等. 基于改进DeepLabv3+的遥感影像道路提取算法[J]. 计算机科学, 2024, 51(8): 168-175.
- Wang Qian, He Lang, Wang Zhanqing, et al. Road extraction algorithm for remote sensing images based on improved DeepLabv3+ [J]. Computer Science, 2024, 51(8): 168-175.
- [14] Tung N X, Son T G. Hyperspectral image classification using an encoder-decoder model with depthwise separable convolution, squeeze and excitation blocks [J]. Earth Science Informatics, 2024, 17(1): 527-538.
- [15] Chen K, Wang H, Zhai Y. A lightweight model for real-time detection of vehicle black smoke [J]. Sensors, 2023, 23(23): 124-131.
- [16] 杨萍, 张汐. 改进DeepLabv3+的道路表面裂缝检测方法[J/OL]. 计算机工程, 1-10 [2024-09-14].

<https://doi.org/10.19678/j.issn.1000-3428.0069114>.

Yang Ping, Zhang Xi. Improved DeepLabv3+ road surface crack detection method [J/OL]. Computer Engineering, 1-10 [2024-09-14]. <https://doi.org/10.19678/j.issn.1000-3428.0069114>.

- [17] He C, Liu Y, Wang D, et al. Automatic extraction of bare soil land from high-resolution remote sensing images based on semantic segmentation with deep learning [J]. Remote Sensing, 2023, 15(6): 1646.
- [18] Chen Z, Chen J, Yue Y, et al. Tradeoffs among multi-source remote sensing images, spatial resolution, and accuracy for the classification of wetland plant species and surface objects based on the MRS_DeepLabV3+ model [J]. Ecological Informatics, 2024, 81: 102594.
- [19] 郝夏楠, 庞亚军, 麻月欣, 等. 基于MobileNetV3网络的轻量级激光雷达点云图像语义分割[J/OL]. 激光杂志, 1-9 [2024-09-14]. <http://kns.cnki.net/kcms/detail/50.1085.TN.20240619.1515.002.html>.
- Hao Xiannan, Pang Yajun, Ma Yuexin, et al. Semantic segmentation of lightweight LiDAR point cloud images based on MobileNetV3 network [J/OL]. Laser Journal, 1-9 [2024-09-14]. <http://kns.cnki.net/kcms/detail/50.1085.TN.20240619.1515.002.html>.
- [20] Yin Y, Guo Y, Deng L, et al. Improved PSPNet-based water shoreline detection in complex inland river scenarios [J]. Complex & Intelligent Systems, 2023, 9(1): 233-245.
- [21] Islam M S, Ibrahim A M, Hoque K E, et al. Advancement in the automation of paved roadways performance patrolling: A review [J]. Measurement, 2024: 114734.
- [22] Saeedizadeh N, Jalali S M J, Khan B, et al. A new optimization approach based on neural architecture search to enhance deep U-Net for efficient road segmentation [J]. Knowledge-Based Systems, 2024, 296: 111966.

(上接第 263 页)

- [13] Liu Z, Mao H, Wu C Y, et al. A convnet for the 2020s [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11976-11986.
- [14] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [15] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern

Recognition, 2017: 2117-2125.

- [16] Wang C, Xu C, Wang C, et al. Perceptual adversarial networks for image-to-image transformation [J]. IEEE Transactions on Image Processing, 2018, 27(8): 4066-4079.
- [17] Tian Z, Shen C, Chen H, et al. FCOS: Fully convolutional one-stage object detection [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 9627-9636.
- [18] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.