

doi:10.6041/j.issn.1000-1298.2026.08.010

基于改进 DeepLabV3 + 的青贮饲料收获机跟随料车车斗识别方法

余永峰^{1,2} 李雷霞^{1,2} 王振坤^{1,2} 杨泽宇^{1,2} 赫志飞^{1,2} 谢胜建^{1,2} 王名举^{1,2}

(1. 中国农业机械化科学研究院集团有限公司, 北京 100083; 2. 中机美诺科技股份有限公司, 北京 100083)

摘要: 青贮饲料收获机跟随料车车斗的准确识别是实现自动抛料装车的首要前提,针对青贮收获田间作业环境复杂、料车种类繁多,现有料车车斗识别方法难以适用各类装载料车识别的问题,利用图像语义分割技术提出一种改进 DeepLabV3 + 的青贮饲料收获机跟随料车车斗识别方法。该方法将轻量化的 MobileNetV3 网络作为主干网络进行特征提取,降低模型复杂度;在 ASPP 模块中以条形池化替换原有的平均池化加强对车斗的识别能力,并将原有的 3 个空洞卷积分支扩充为 4 个空洞深度可分离卷积分支,提升对大目标的识别能力;在解码器部分融合了多尺度特征,最大程度地弥补损失的车斗边界细节特征,并在拼接融合后添加 C2f-GS 模块进一步增强模型的特征融合能力。在构建的料车车斗数据集上设计消融试验和对比试验对模型改进的有效性进行验证,试验结果表明,改进 DeepLabV3 + 模型 mIoU 和 mPA 达到 94.25% 和 96.62%,分割效果优于其他分割模型,且模型参数量和浮点运算量显著低于其他模型。与原 DeepLabV3 + 模型相比,改进模型的 mIoU 和 mPA 分别提高 4.77、3.71 个百分点,识别准确性得到有效提升。利用基于 RANSAC 算法的料车车斗拟合方法对分割后的结果进行车斗边界拟合,对不同情况下的分割结果进行拟合试验,结果表明拟合的车斗与原始标注车斗的 mIoU 平均值达到 94.49%,二者质点位置偏差仅为 25.47 像素,拟合结果基本能够覆盖车斗区域,满足青贮饲料收获机田间作业情况下的需求。

关键词: 青贮饲料收获机; 图像语义分割; DeepLabV3 +; 车斗识别; 边界拟合

中图分类号: S817.1 文献标识码: A 文章编号: 1000-1298(2026)08-0100-11 OSID: 

Recognition Method of Silage Harvester Follow-up Trailer Hopper Based on Improved DeepLabV3 +

YU Yongfeng^{1,2} LI Leixia^{1,2} WANG Zhenkun^{1,2} YANG Zeyu^{1,2} HE Zhifei^{1,2} XIE Shengjian^{1,2}
WANG Mingju^{1,2}

(1. Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing 100083, China

2. Menoble Co., Ltd., Beijing 100083, China)

Abstract: Accurate recognition of the silage harvester follow-up trailer hopper is the first prerequisite for realizing automatic throwing and loading, for the problem that the complex operational environment of silage harvesting, the variety of trailer types, and existing methods for trailer hopper recognition struggle to apply to all kinds of loading hopper identification, using the technology of image semantic segmentation, a silage harvester follow-up trailer hopper recognition method was proposed with improved DeepLabV3 +. The method took the lightweight MobileNetV3 network as the backbone network for feature extraction to reduce the model complexity. In the ASPP module, the original average pooling was replaced by strip pooling to enhance the recognition ability of the hopper, and the original three atrous convolution branches were expanded into four atrous depthwise separable convolution branches to improve the recognition of large targets. Moreover, the multi-scale features in the decoder part were integrated to compensate for the lost hopper boundary detail features. After concatenation and fusion, the C2f-GS module was added to further enhance model's feature fusion capabilities. The ablation tests and comparison tests were conducted on the constructed hopper dataset to verify the effectiveness of the improved model, and the test results showed that the values of the mean intersection over union (mIoU)

收稿日期: 2025-01-07 修回日期: 2025-02-25

基金项目: 国家重点研发计划项目(2022YFD2001905)

作者简介: 余永峰(2001—),男,硕士生,主要从事农业装备智能测控技术研究,E-mail: yuyf0603@163.com

通信作者: 李雷霞(1982—),女,研究员,主要从事智能农机收获装备电控系统研究,E-mail: lileixia1016@163.com

and mean pixel accuracy (mPA) of the proposed improved DeepLabV3 + model reached 94.25% and 96.62%, which was better than that of the other segmentation models, and the model parameters and FLOPs were significantly lower than that of the other models. Compared with the original DeepLabV3 + model, the mIoU and mPA of the improved model were increased by 4.77 and 3.71 percentage point, model recognition accuracy was effectively enhanced. Using the hopper fitting method based on the RANSAC algorithm, the boundary of the hopper was fitted from the segmented results, fitting experiments conducted under various conditions revealed that the mIoU of the fitted hopper and the original labeled image reached 94.49%, and the deviation of the center-of-mass point position between the two was only 25.47 pixels, which indicated that the fitting result basically could cover the area of the hopper, meeting the operational requirements for silage harvesting in the field.

Key words: silage harvester; image semantic segmentation; DeepLabV3 +; hopper recognition; boundary fitting

0 引言

青贮饲料是由玉米、苜蓿、黑麦草等青绿饲料通过厌氧微生物发酵制成的优质饲料,因其具有营养价值高、贮存时间长、适口性好等优点被广泛用来喂养牛、羊等反刍动物^[1-2]。目前,我国青贮饲料收获主要采用青贮饲料收获机进行机械化收获作业,该收获机在田间作业时可一次性完成作物的收割、喂入、切碎和籽粒破碎等操作,再通过抛送装置抛入收获机后方或侧方同向行驶的跟随料车车斗内,以完成整个收获作业过程^[3-4]。在抛料装车作业环节,国产收获机目前仍以肉眼识别、手动控制为主,操作人员需要长时间专注观察跟随料车车斗的位置以调节抛臂左右及高低方向,使得饲料最大程度地抛送到车斗内,其工作强度较大,易造成饲料抛送损失,因此研究青贮饲料收获机自动抛料装车技术对于提升收获效率具有重要意义。实现自动抛料装车的首要前提是完成对跟随料车车斗的准确识别,苗中华等^[5]利用深度相机获取跟随料车车斗的点云数据,利用车斗高度位于世界坐标系中最高处的原则实现了对料车车斗边缘的识别,但由于田间作业环境复杂、料车种类繁多等原因,该方法难以用于各类装载料车的识别,因此迫切需要寻找一种识别精度高、适应性强的识别方法用于青贮饲料收获机跟随料车车斗的识别,开发自动抛料装车系统,以提高青贮饲料收获机的作业效率。

近年来,随着深度学习算法和机器视觉技术的不断发展,不少学者将基于深度学习的图像分割技术应用到农业领域^[6-8],DeepLabV3 +^[9]作为图像分割领域的经典网络之一,已在农机作业导航线提取、农作物类型识别、田块分割等方面得到了广泛应用。为了实现火龙果园视觉导航路径识别,周学成等^[10]利用 MobileNetV2 替换 DeepLabV3 + 的主干特征提取网络,并引入深度可分离卷积及坐标注意力

机制实现了对果园道路区域的分割。武锦龙等^[11]提出一种 SE—RepVGG 与 MobileViT 融合的主干网络结构用于荞麦苗期无人机遥感图像的分割。顾文娟等^[12]将 GhostNet 网络和多分支解码结构引入到 DeepLabV3 + 网络中,并结合 CA 注意力机制提高了对番茄图像中小目标类别识别的准确性。曹英丽等^[13]针对稻田中杂草和水稻彼此遮挡、难以区分的问题,提出了一种多分支感受野级联融合的 ASPP 模块,从而实现对杂草特征的共享,最终使得杂草识别平均像素准确率达到 95.67%。

上述研究将 DeepLabV3 + 模型用于农业领域不同场景识别当中,均取得了不错的效果,可为青贮饲料收获机跟随料车车斗识别提供参考。但青贮饲料的收获过程不同于其他常规作物,其收获作业过程面临干扰因素多、作业环境复杂、光照多变,导致识别难度较大,因此,为了实现对青贮饲料收获机田间作业环境下跟随料车车斗的准确识别,本文提出一种基于改进 DeepLabV3 + 的料车车斗识别方法。采用 MobileNetV3 轻量级网络作为主干特征提取网络以降低模型复杂度,引入条形池化替换空洞空间金字塔模块的平均池化以实现对料车车斗更好的识别效果,并扩展原有的空洞卷积分支,在解码器部分融合多尺度特征,并在拼接融合后添加 C2f—GS 模块进一步增强模型的特征融合能力,提高对料车车斗识别准确性,最后利用基于 RANSAC 算法的料车车斗拟合方法实现对车斗边界准确拟合,以期青贮饲料收获机自动抛料装车技术提供支持。

1 材料与方法

1.1 图像采集与数据集构建

本研究所需的青贮饲料收获机跟随料车车斗图像数据采集自河北省石家庄市元氏县马村镇,拍摄设备为青岛维感科技有限公司生产的 NYX660 型深度相机,将相机搭载在中机美诺 9580 型青贮饲料收

获机上采集图像,由 24 V 车载电源对其进行供电,图像数据通过以太网传输并保存在计算机内;采用与相机配套的 ScepterGUITool 工具通过手动触发进行图像采集,分辨率设置为 1 600 像素 × 1 200 像素,保存为 JPG 格式文件。根据抛臂机械结构设计相机连接固定支架,在数据采集时通过相机支架将 NYX660 型深度相机安装在抛臂的中上部,调节相机角度使其向下倾斜,确保当抛臂升至最高位置时能使跟随料车图像出现在相机视角范围内,图 1 为青贮饲料收获机田间作业时的图像采集系统。



图 1 图像采集系统

Fig. 1 Image acquisition system

1. 青贮饲料收获机 2. 跟随料车 3. 深度相机 4. 计算机采集平台

考虑青贮饲料收获机在田间作业过程能够采集的跟随料车种类较少,所采集的料车车斗图像不够丰富,为进一步增强本文研究的通用性和适应性,采取手动拍摄方式进行数据集的扩充。常规青贮饲料收获机抛臂升至最高处时抛头与地面垂直距离为 5 m 多,将相机安装在抛臂的中上部,在田间作业过程相机与地面垂直距离为 4 ~ 5 m,因此采用特制的 5 m 长碳纤维拍摄杆完全模拟青贮饲料收获机田间作业过程相机的拍摄视角,在各乡村、郊区采集农用厢式货车图像共 1 264 幅,筛选得到图像 740 幅,共 19 种农用货车类型,部分货车车斗图像样例如图 2a、2b 所示;在青贮饲料收获机田间作业环境采集跟随料车车斗图像 1 920 幅,剔除掉部分曝光严

重及重复图像得到图像 800 幅,如图 2c 所示。

1. 2 数据增强

利用 Labelme 软件对采集到的料车车斗总数据集进行语义标注,将除料车车斗外的其他像素标注为背景类,标注后的图像以 .json 格式存储,并将其转为 8 位灰度图作为标签图,如图 2d 所示。为进一步增强模型的泛化能力,通过对料车车斗原图像和标签同时进行随机旋转、随机明亮度、镜像、随机模糊以及添加高斯噪声和椒盐噪声等操作丰富数据集,最终获得数据集共 2 000 幅,其数据集构建如图 2 所示,随后按照比例 9:1 划分训练集和验证集对模型进行训练和测试。

2 料车车斗分割模型设计与边缘拟合

2. 1 原始 DeepLabV3 + 网络模型

DeepLab 系列模型^[14-16]因其出色的性能和准确性,在多个基准数据集上取得了领先成绩,被广泛应用于各个领域。DeepLabV3 + 是 DeepLab 系列的最新版本,在之前模型的基础上增加 1 个解码器模块,专门用于恢复对象边界,从而改善边界细节的分割质量。在主干网络部分,DeepLabV3 + 采用 Xception^[17]网络进行特征提取,得到多个有效语义特征层,其中将 1 个深层特征输入到空洞空间金字塔模块(Atrous spatial pyramid pooling, ASPP)中,通过多个空洞卷积分支和池化操作捕获多尺度的上下文信息;将 1 个浅层特征直接输入到解码器中,并与处理后的深层特征进行融合,最后通过 3 × 3 卷积和 4 倍上采样获得与原图分辨率大小相同的预测结果。

2. 2 改进 DeepLabV3 + 网络模型

青贮饲料收获机在田间作业时,需要快速识别跟随料车车斗位置,从而形成决策并控制抛臂做出反应,使得物料能够精准地抛送到特定位置,这对部



图 2 数据集构建

Fig. 2 Data set construction

署的网络模型识别性能提出了更高的要求,为保证在有限的车载硬件条件下能够部署网络模型时料车车斗的准确识别,对原 DeepLabV3 + 网络模型进行改进:①将更为轻量级的 MobileNetV3 网络替换原模型中 Xception 网络进行特征提取,以减小参数计算量并降低模型的复杂度。②根据料车车斗图像的特点,将 ASPP 模块中平均池化替换为条形池化,并扩展原有 3 个空洞卷积分支。③在解码器部分融合多尺度特征,最大程度地弥补损失的车斗边界细节特征。④在 Concat 融合操作后加入改进的 C2f 模块增强对多尺度特征的融合,改进后整体网络结构如图 3 所示。

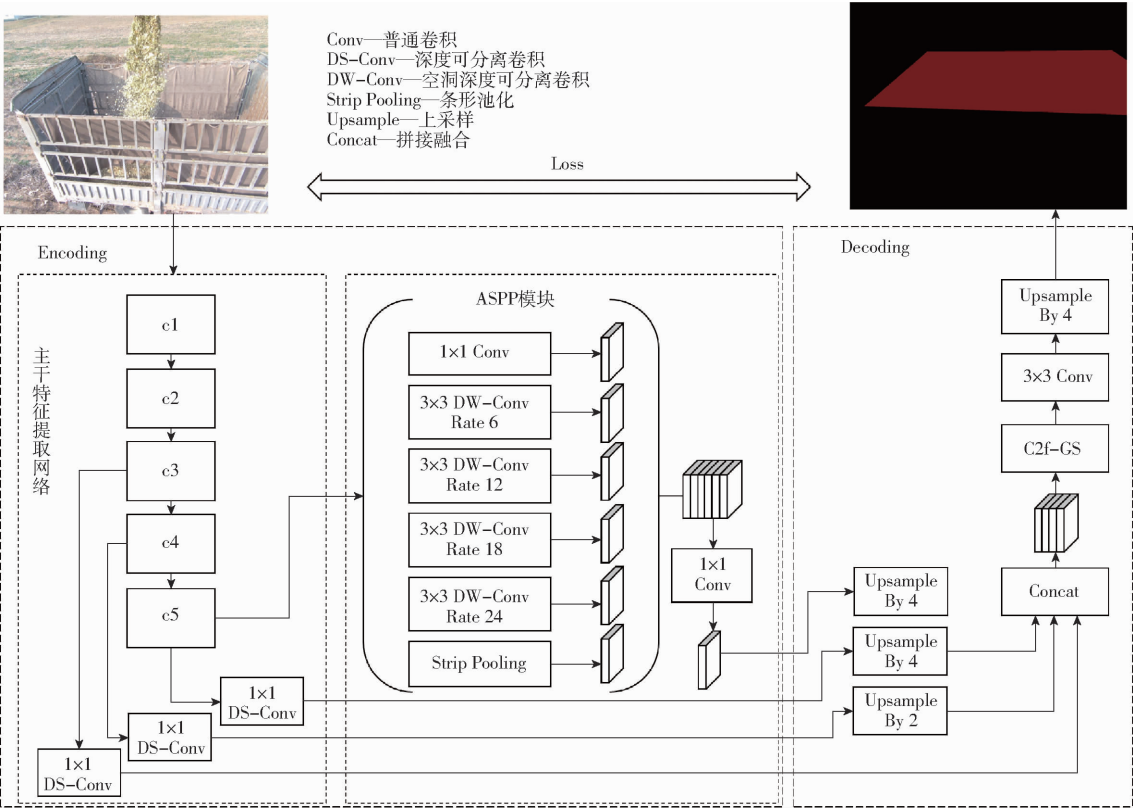


图 3 改进 DeepLabV3 + 模型的网络结构

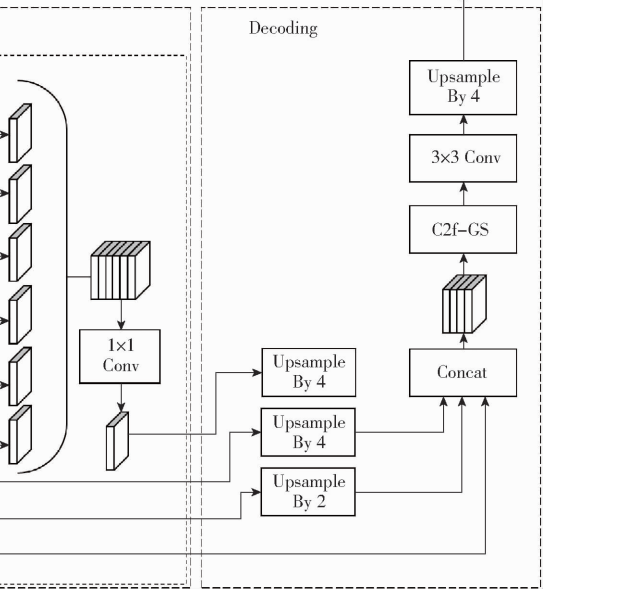
Fig. 3 Improving network structure of DeepLabV3 + models

2.2.1 轻量级主干特征提取网络

原 DeepLabV3 + 模型采用 Xception 网络作为其主干特征提取网络,但 Xception 网络由于参数规模和运行速度控制不佳,并不适用于青贮饲料收获机这样的移动式设备,因此需要对主干特征提取网络进行轻量化改进。

MobileNetV3^[18]是 Google 在 2019 年提出的一种轻量级卷积神经网络结构,专为移动和嵌入式设备优化,以提高速度和准确性。该版本模型结合 V1 中的深度可分离卷积和 V2 中线性瓶颈的反向残差结构,能够显著减少计算量和参数量,对计算过程更加友好,然后引入 1 个优化的激活函数 Hard-swish 函数,在保证模型非线性表达能力的同时通过简化计算步骤使得计算密集度明显降低,使其更加适合移动设备上的高效推理任务。此外,MobileNetV3 还增加了挤压和激励轻量级注意力模块(Squeeze and excit,SE),在训练过程,SE 模块为每个通道赋予权重,以便模型全局用于强调具有较高权重的特征并

抑制不重要的特征。如图 4 为 MobileNetV3 特有的基本网络单元,在料车车斗图像输入后,通过 1 个 1 × 1 卷积提升其通道维数,然后利用 3 × 3 大小的深度可分离卷积(Dwise)对提升后的高维数据进行处理,提取料车车斗特征,之后通过 SE 注意力机制模块中的池化(Pool)、全连接(Fc)和激活函数处理进一步强化重要特征,最后再利用 1 个 1 × 1 卷积减少通道数得到最终的输出结果^[19],其中 + 表示残差连接,只有在输入通道数与输出通道数相等且步距为 1 时出现^[20]。



抑制不重要的特征。如图 4 为 MobileNetV3 特有的基本网络单元,在料车车斗图像输入后,通过 1 个 1 × 1 卷积提升其通道维数,然后利用 3 × 3 大小的深度可分离卷积(Dwise)对提升后的高维数据进行处理,提取料车车斗特征,之后通过 SE 注意力机制模块中的池化(Pool)、全连接(Fc)和激活函数处理进一步强化重要特征,最后再利用 1 个 1 × 1 卷积减少通道数得到最终的输出结果^[19],其中 + 表示残差连接,只有在输入通道数与输出通道数相等且步距为 1 时出现^[20]。

为适应不同场景和需求,MobileNetV3 分为 Large 和 Small 2 个版本,Small 版本虽然比 Large 版本参数量和计算量更低,但同时网络模型精度也有所下降,综合考虑,本文采用 MobileNetV3_large 网络替换 DeepLabV3 + 原模型 Xception 网络,并通过其他改进措施进一步降低模型参数,提升运行速度。

2.2.2 ASPP 模块改进

原 DeepLabV3 + 网络的 ASPP 模块包含 5 个分

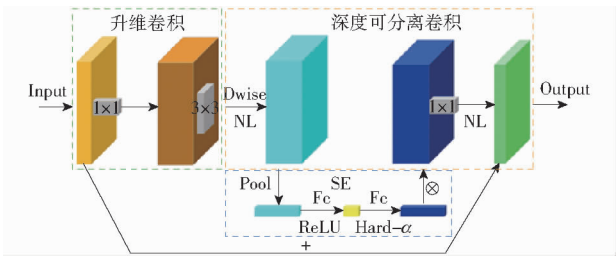


图 4 MobileNetV3 基本网络单元
Fig. 4 MobileNetV3 basic network unit

支,分别是 1 个 1×1 的标准卷积、3 个空洞率递增的 3×3 空洞卷积以及 1 个平均池化层^[21],根据料车车斗图像的特点,本研究将 ASPP 模块中平均池化层采用条形池化层进行替换,并将 3 个空洞卷积分支改为空洞率为 6、12、18、24 的 4 个空洞深度可分离卷积分支。

传统池化操作通常采用方形的池化核,并以固定步长滑动覆盖整个特征图,这种设计主要用于捕获局部邻域信息,而无法有效地捕获全局或长距离的依赖关系。考虑当跟随料车在与青贮饲料收获机并排行驶或是后跟车行驶时,大中型料车车斗图像呈现为长条状的四边形,传统池化由于卷积核覆盖范围限制可能无法有效地捕获料车车斗的全局依赖关系,针对此问题,本研究在 DeepLabV3 + 的 ASPP 模块中采用条形池化模块 (Strip Pooling) 替换平均池化模块,有效地捕获料车车斗的上下文信息。条形池化模块网络结构图如图 5 所示,它通过在某一空间维度配置一条狭长的池化窗口来捕获各个区域之间的远程关系,并沿其他空间维度保持较窄的池化窗口使其只捕获局部上下文信息,从而避免无关区域的污染信息对预测结果形成干扰^[22]。在条形池化模块中,对于 1 个输入尺寸为 $H \times W$ 特征图,采用纵向和横向长条形池化核转换得到 $H \times 1$ 和 $1 \times W$ 2 个低维特征图,利用一维卷积对 2 个特征图进行扩充,使其尺寸匹配,并通过逐像素相加的方式合并信息,恢复为原尺寸大小的特征图。在此基础上,采用 1×1 卷积和 Sigmoid 激活函数生成一个权重矩阵,与原输入特征图进行逐元素相乘得到最终的输出特征图^[23],从而实现对原特征图关键信息的有效增强,同时能抑制不重要的背景噪声。

空洞深度可分离卷积将传统空洞卷积过程分为逐通道空洞卷积 (Depthwise atrous convolution, DW) 和逐点卷积 (Pointwise convolution, PW) 2 个相对简单的步骤进行,其结构图如图 6 所示^[19],逐通道空洞卷积允许网络在每个通道上独立提取更大范围的特征,而逐点卷积则利用 1×1 卷积核整合不同通道之间的信息,这样分离操作能有效地减少卷积过程计算量。

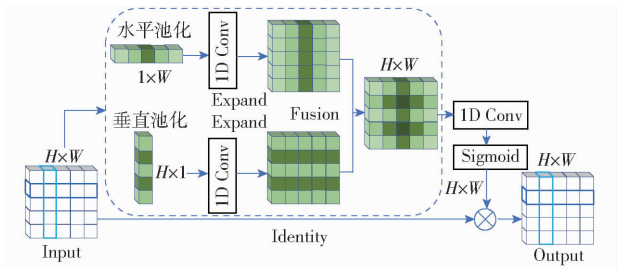


图 5 条形池化模块
Fig. 5 Strip Pooling module

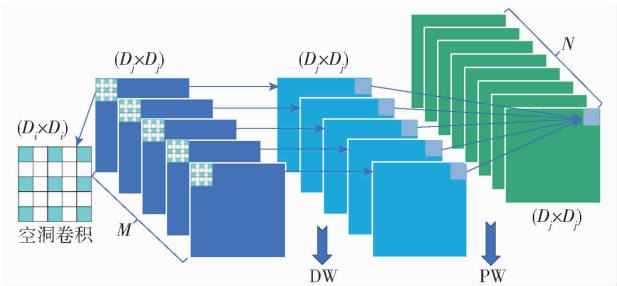


图 6 空洞深度可分离卷积
Fig. 6 Atrous depth-separable convolution

假设输入特征图尺寸为 $D_i \times D_i$, 通道数为 M , 卷积核大小为 $D_j \times D_j$, 输出特征图通道数为 N , 则传统空洞卷积对应的计算量 Q_1 为

$$Q_1 = D_i D_i M N D_j D_j \quad (1)$$

而空洞深度可分离卷积中的 DW 卷积计算量 Q_2 为

$$Q_2 = D_i D_i M D_j D_j \quad (2)$$

PW 卷积计算量 Q_3 为

$$Q_3 = M N D_i D_i \quad (3)$$

将 DW 卷积和 PW 卷积计算量相加,即可得空洞深度可分离卷积计算量,与传统空洞卷积计算量比值为

$$\frac{Q_2 + Q_3}{Q_1} = \frac{D_i D_i M D_j D_j + M N D_i D_i}{D_i D_i M N D_j D_j} = \frac{1}{N} + \frac{1}{D_j^2} \quad (4)$$

本文 ASPP 模块使用的空洞卷积核大小均为 3×3 , 即可计算出原普通空洞卷积的计算量为空洞深度可分离卷积计算量的 8 ~ 9 倍^[24], 因此将原 ASPP 模块中 3 个空洞卷积采用深度分离的方式进行卷积,能够在保持相似性能的同时显著降低模型参数数量和计算复杂度,减少计算资源的消耗,提升计算效率。在此基础上,考虑青贮饲料收获机跟随料车车斗识别为大目标分割问题,原空洞率 6、12、18 的卷积核对像素占比大的目标特征捕获效果不够强,因此在 ASPP 模块中增加 1 个空洞率为 24 的空洞深度可分离卷积分支^[25],从而捕获更大范围的料车车斗上下文信息,提高分割精度。

2.2.3 多尺度特征融合

DeepLabV3 + 网络在解码器部分会将得到的浅

层特征层和深层特征层进行融合,这样能够利用到不同层次的语义信息,但是仅采用主干网络部分单一的 1/4 层进行融合会导致损失较多重要的车斗边界细节特征,为了更好地利用主干网络下采样过程中各个层级的语义信息并最大程度地弥补损失的边界细节特征,借鉴 FPN^[26]的思想设计多尺度特征融合模块^[12],如图 7 所示。将 MobileNetV3 网络输出的 1/4、1/8、1/16 分支的特征层均通过 1×1 卷积将通道数调整一致,然后对 1/8 分支和 1/16 分支的特征层进行上采样操作,将 2 个分支的特征层尺寸调整为跟 1/4 分支的特征层相同。同时 ASPP 模块输出的深层特征层也通过 1×1 卷积和上采样操作将尺寸和通道数调整为与 1/4 分支相同的特征层,将得到的 4 个特征层进行 Concat 融合,最后通过 1 个 3×3 卷积和 2 倍上采样操作即可得到与输入图像尺寸相同的料车车斗分割图像。

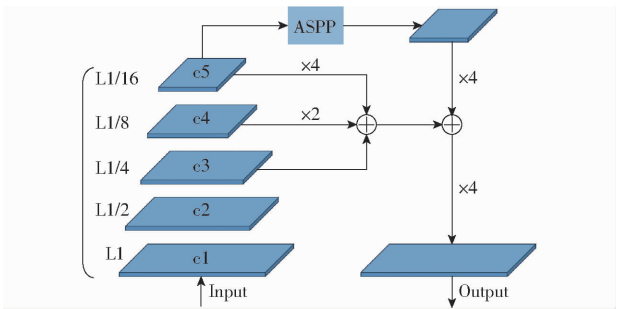


图 7 多尺度特征融合
Fig. 7 Multi-scale feature fusion

2.2.4 C2f-GS 模块

DeepLabV3 + 模型在解码器部分通过 Concat 结合来自浅层和深层的特征,本文引入多个浅层语义特征层,为更好地融合这些来自不同层次的语义特征,在解码器的 Concat 融合操作后添加 1 个 C2f 模块增强模型的特征融合能力。C2f 是 YOLO v8 中引进的创新模块,巧妙地设计了特征融合过程,能够有效地整合不同尺度和层次的特征。特征图在输入后被拆分成 2 部分,一部分直接传递,另一部分经过多个 Bottleneck 块处理后再与直接传递的部分进行拼接,从而充分利用不同尺度的特征信息,提高模型对不同尺度物体的检测能力。同时通过拼接融合和加法融合等多种特征融合方式,能够保留各层特征的共同信息,实现不同层次特征的有效整合^[27]。

考虑到青贮饲料收获作业环境复杂,料车车斗的大小、形状等都存在一定的重复性,C2f 模块采用普通卷积进行特征提取往往会保留更多的冗余信息,无法充分利用不同通道之间的相关性,导致计算资源的浪费^[28],为解决上述问题,采用 GSConv (Group-shuffle convolution) 重构原始 C2f 模块中的

Bottleneck 结构,以减少冗余特征的提取,有效降低计算量,改进后的结构称为 C2f-GS,如图 8 所示。GSConv 主要由分组卷积和通道混洗 2 部分组成,分组卷积能够将输入特征进行分组,让其中部分通道执行标准卷积,以保留详细的特征表达,部分剩余通道执行深度卷积,然后通过通道混洗将 2 种卷积的信息进行混合,从而提升特征提取的效率和能力^[29]。通过在 DeepLabV3 + 中解码器的 Concat 融合后添加 C2f-GS,不仅能够有效地聚合多尺度的特征信息,并且能够在不牺牲性能的前提下减少模型的参数量,提高模型的计算效率。

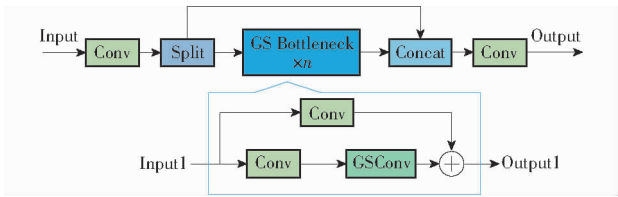


图 8 C2f-GS 结构
Fig. 8 C2f-GS structure

2.3 车斗边界拟合

根据跟随料车车斗分割图,一定的直线拟合操作得到车斗的 4 条边界。目前在农业领域常用最小二乘法、霍夫变换和随机抽样一致性算法 (RANSAC) 等对农机作业导航线提取^[8,10,30]。考虑车斗形状为四边形,采用常规的最小二乘法难以准确地拟合出料车车斗,而霍夫变换在受到噪声干扰、部分遮挡等情况下易造成误检和漏检的情况,因此综合来看选择随机抽样一致性算法进行料车车斗的边界拟合。

青贮饲料收获机在田间收获作业过程,不同类型的跟随料车在深度相机视野范围内成像情况不同,小型料车车斗能够完整呈现在相机视野范围内,而大型料车车斗仅能呈现部分,此时边缘检测算法无法将图像边线作为车斗轮廓的一部分将其检测出来,因此,为保证能够准确拟合出相机视野范围内完整的车斗边界,本文基于 RANSAC 算法提出适用于青贮饲料田间收获场景的料车车斗拟合方法,其流程图如图 9 所示。

(1) 为保证边缘检测时能检测出深度相机视野范围内完整的封闭车斗轮廓,对网络模型识别得到的分割图像四周进行一定像素黑色背景的填充。

(2) 采用 Canny 算法进行边缘检测,利用高斯平滑处理能够在一定程度上减小噪声对边缘的影响。

(3) 对边缘检测后的结果筛选出一定数量边界点,减少随机抽样一致性算法的迭代时间,同时求出所有边界点的质心坐标。

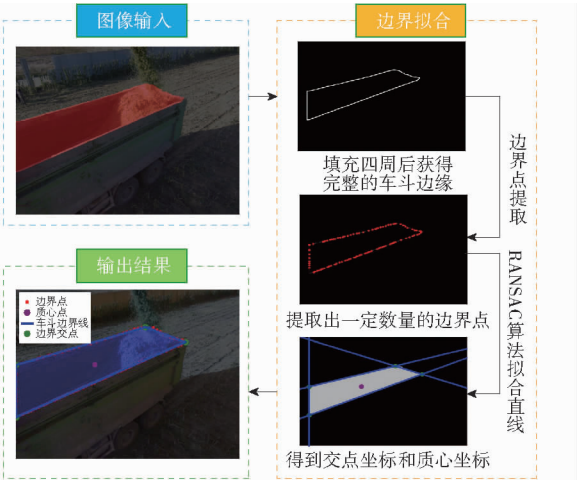


图 9 车斗边界拟合流程图

Fig. 9 Bucket boundary fitting flow chart

(4) 利用随机抽样一致性算法依次拟合出车斗的 4 条边界,每拟合完 1 条边界后删除边界一定范围内的所有边界点,防止对后续的边界拟合产生干扰。

(5) 求出 4 条边界直线的交点,由于车斗在相机中的成像不一定为规则的四边形,4 条拟合直线可能存在 4~6 个交点。

(6) 根据各交点与质心之间的距离,筛选出最近的 4 个交点,并通过各交点与质心之间的夹角依次连接 4 个交点形成封闭轮廓,即可得到拟合的车斗边界。

3 试验结果与分析

3.1 试验环境及参数设置

本试验所用硬件平台为 11th Gen Intel(R) Core (TM) i7 - 11700 处理器、NVIDIA GeForce RTX 3080Ti 显卡、12 GB 内存,软件环境为 Windows 10、Cuda、Anaconda3、Python 3.7。

通过多次试验,确定网络模型训练时可以较快达到局部收敛的网络参数,学习率设置为 1×10^{-4} ,下降方式为余弦退火,批量处理样本数量 (Batch size) 为 8,选择 ADAM 优化器,模型迭代次数 (epoch) 为 100。

3.2 评价指标

为客观评价模型对青贮饲料收获机跟随料车车斗识别效果,选取平均交并比 (mIoU)、平均精度均值 (mPA)、模型参数量 (Parameters) 和浮点运算量 (FLOPs) 作为模型性能评价指标。

为验证青贮饲料收获机跟随料车车斗边界拟合方法的准确性,从料车车斗边界拟合准确程度和车斗区域偏移程度 2 方面对拟合结果进行衡量。

(1) 车斗边界拟合准确程度采用交并比 (IoU)

进行量化,与分割结果交并比类似,该处交并比表示拟合出车斗 4 条边界包围的车斗面积与原图通过 Labelme 软件标注的标签图之间的重叠程度,计算式为

$$IoU = \frac{P_N \cap P_L}{P_N \cup P_L} \times 100\% \tag{5}$$

式中 P_N ——拟合出车斗区域像素

P_L ——标签图中车斗区域像素

(2) 车斗区域偏移程度利用拟合出车斗 4 个交点的质心坐标与标签图中车斗边缘质心坐标之间的位置偏差表示,计算式为

$$D_\theta = \sqrt{(x_N - x_L)^2 + (y_N - y_L)^2} \tag{6}$$

式中 D_θ ——2 质心坐标之间位置偏差

(x_N, y_N) ——拟合出车斗边界 4 个交点质心坐标

(x_L, y_L) ——标签图中车斗边缘质心坐标

3.3 不同模型分割效果

3.3.1 模型训练结果

为验证本研究改进 DeepLabV3+ 模型用于料车车斗识别相较于原 DeepLabV3+ 模型的提升,将 2 模型以相同环境和参数设置在构建的跟随料车车斗数据集上进行训练,其损失值和 mIoU 变化曲线如图 10 所示,总体上二者损失值变化趋势相同,均随迭代次数增加而递减并最终趋于稳定,mIoU 则随迭代次数增加而递增,当达到 80 时,二者的 mIoU 均达到平稳不再增加。但是与原 DeepLabV3+ 网络模型相比,改进 DeepLabV3+ 网络模型的损失值和 mIoU 曲线收敛速度更快,且总体损失值更低、mIoU 更高,后期波动更小。

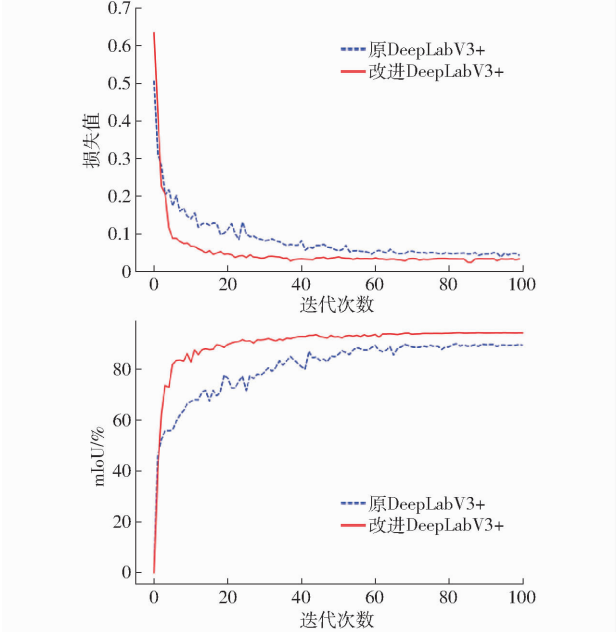


图 10 改进前后网络模型训练结果

Fig. 10 Model training result before and after improvement

3.3.2 消融试验

为了说明 DeepLabV3 + 模型改进效果,分析每

一个改进举措后模型在本文构建的数据集上识别效果,结果如表 1 所示。

表 1 不同改进模型车斗分割性能试验结果

Tab.1 Test results on dump body partitioning performance of different improved models								
模型编号	主干网络	ASPP 模块	多尺度特征融合模块	C2f-GS 模块	mIoU/%	平均精度均值/%	模型参数量	浮点运算量
1	Xception	×	×	×	89.48	92.91	3.71×10^7	5.545×10^{10}
2	MobileNetV3	×	×	×	91.51	94.70	1.11×10^7	2.086×10^{10}
3	MobileNetV3	✓	×	×	93.26	95.95	6.84×10^6	1.638×10^{10}
4	MobileNetV3	✓	✓	×	93.72	96.78	6.34×10^6	1.666×10^{10}
5	MobileNetV3	✓	✓	✓	94.25	96.62	6.80×10^6	1.567×10^{10}

注:✓表示引入此模块,×表示未引入此模块。

由表 1 可知,原 DeepLabV3 + 模型对料车车斗识别的 mIoU 和平均精度均值分别为 89.48% 和 92.91%,模型参数量较大,复杂度较高;在原模型基础上将轻量级的 MobileNetV3 网络作为主干网络进行特征提取,使得模型 mIoU 和平均精度均值分别提升 2.03、1.79 个百分点,同时模型参数量仅为原模型 30%,有效地降低了模型复杂度;模型 3 在模型 2 基础上进行了 ASPP 模块的调整,mIoU 和平均精度均值分别提升 1.75、1.25 个百分点,同时模型参数量也有所降低;模型 4 对解码器模块进行改进,在浮点运算量小幅增加的情况下使得模型 mIoU 和平均精度均值进一步提升,同时模型参数量仅为原模型的 17.11%。模型 5 在 Concat 融合操作后添加了改进的 C2f 模块,使得模型 mIoU 和平均精度均值达到 94.25% 和 96.62%。与原模型相比,改进模型的 mIoU 和平均精度均值分别提高 4.77、3.71 个百分点,识别准确性得到有效提升,模型参数量仅为 6.80×10^6 ,在保持低参数量的同时有效地提升了分割精度,保证了在有限环境配置下的部署和应用。

3.3.3 不同模型性能试验

为进一步评价改进 DeepLabV3 + 模型对料车车斗的分割性能,选择 U-Net、PSNet、HRNet 语义分割模型在相同的训练参数和数据集上进行训练,对训练结果进行对比,结果如表 2 所示。由表 2 可知,U-Net 模型的 mIoU 和平均精度均值虽然比原 DeepLabV3 + 模型更高,达到 91.83% 和 94.83%,但浮点运算量显著高于其他网络模型;PSNet 和 HRNet 模型 mIoU 和平均精度均值分别为 88.00%、89.16% 和 94.36%、93.26%,在跟随料车车斗数据集上表现一般,总体模型参数量较大;本文改进 DeepLabV3 + 模型在平均交并比方面比 U-Net、PSPNet、HRNet 和原 DeepLabV3 + 分别提高 2.42、6.25、5.09、4.77 个百分点,在平均精度均值方面分别提高 1.79、2.26、3.36、3.71 个百分点。此外,改进模型模型参数量和浮点运算量分别为 $6.80 \times$

表 2 不同模型性能对比

Tab.2 Performance comparison of different models				
模型	mIoU/%	平均精度 均值/%	模型参 数量	浮点运 算量
U-Net	91.83	94.83	2.489×7	2.258×10^{11}
PSPNet	88.00	94.36	4.671×7	5.921×10^{10}
HRNet	89.16	93.26	6.585×7	9.384×10^{10}
DeepLabV3 +	89.48	92.91	3.705×7	5.545×10^{10}
本文	94.25	96.62	6.800×6	1.567×10^{10}

10^6 、 1.567×10^{10} ,显著低于其他模型。

青贮饲料收获机在田间收获时,需要快速并准确地识别到跟随料车车斗位置,从而形成决策以控制抛臂做出反应,因此需要综合考虑网络分割准确性和推理速度。改进 DeepLabV3 + 模型在保证较高的分割准确性的同时有效降低了模型参数量及模型复杂度,能极大地满足青贮饲料收获机移动端的部署要求。

3.4 车斗边界拟合效果试验

为验证青贮饲料收获机跟随料车车斗边界拟合方法的准确性,对跟随料车侧面跟车与后跟车 2 种不同情况分割结果进行拟合,如图 11 所示,其中侧面跟车时车斗在深度相机中可能存在完整图像、单边截断图像和双边截断图像 3 种不同的情况,如图 11a~11c 所示。由图 11 可知,拟合算法对不同情况料车车斗均能达到较好的拟合结果。

选取侧面跟车 3 种不同图像与后跟车场景的图像分割结果进行车斗边界拟合,测试其与原标签图的 IoU 以及质心坐标位置偏差,测试结果如表 3 所示。由表 3 可知,不同图像情况拟合车斗与原标签图 mIoU 平均值为 94.49%,其中侧面跟车时 3 种不同图像 mIoU 达到 94.41%、96.38%、95.71%,表明拟合方法可以比较准确地拟合出跟随料车车斗边界。后跟车情况下 mIoU 值略低于侧面跟车,其原因是后跟车情况车斗在相机中成像面积相对较小,拟合结果可能产生较大相对偏差,且在后跟车情况

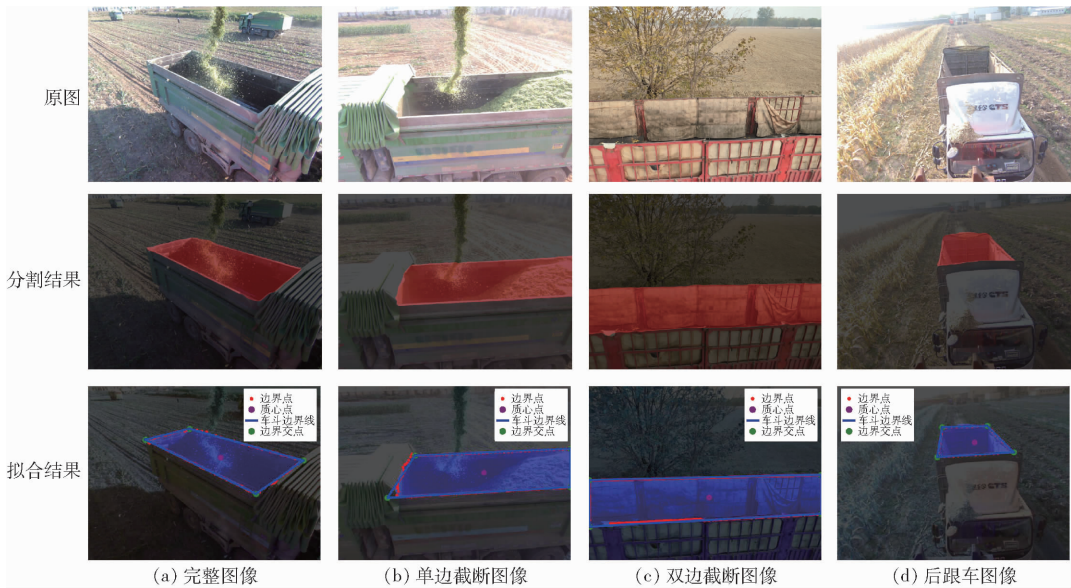


图 11 不同情况下的车斗拟合结果

Fig. 11 Bucket fitting results for different scenarios

表 3 不同图像拟合车斗与标签图交并比及质心位置偏差

Tab. 3 Intersection ratio and centre-of-mass position deviation of fitted buckets to labeled diagrams in different images									
图像	图序号	拟合车斗与标签图交并比			mIoU/%	质心点位置偏差			平均位置 偏差/像素
		交集/像素	并集/像素	IoU/%		x/像素	y/像素	D_{θ} /像素	
完整图像	1	328 620	342 286	96. 01	94. 41	- 16. 64	21. 48	27. 18	23. 19
	2	613 057	650 710	94. 21		6. 86	9. 25	11. 52	
	3	254 420	270 590	94. 02		3. 41	3. 96	5. 23	
	4	81 186	88 230	92. 02		- 21. 94	- 0. 11	21. 95	
	5	507 402	529 736	95. 78		56. 32	- 3. 34	56. 41	
单边截断图像	1	313 680	327 141	95. 89	96. 38	28. 3	25. 46	38. 07	32. 07
	2	273 050	280 391	97. 38		3. 68	12. 65	13. 17	
	3	390 765	399 859	97. 73		41. 71	37. 92	56. 37	
	4	493 462	517 477	95. 36		18. 48	32. 14	37. 07	
	5	294 315	308 048	95. 54		8. 51	13. 17	15. 68	
双边截断图像	1	687 045	708 154	97. 02	95. 71	18. 58	- 8. 90	20. 60	17. 79
	2	763 924	805 716	94. 81		- 17. 68	30. 99	35. 69	
	3	398 089	415 657	95. 77		- 5. 79	6. 19	8. 47	
	4	529 451	540 823	97. 90		- 1. 09	17. 19	17. 22	
	5	433 559	465 926	93. 05		5. 81	3. 85	6. 96	
后跟车图像	1	83 931	93 966	89. 32	91. 44	- 4. 16	40. 66	40. 87	28. 82
	2	70 974	76 612	92. 64		36. 84	16. 59	40. 40	
	3	202 633	222 062	91. 25		- 21. 11	1. 47	21. 16	
	4	177 066	192 764	91. 86		- 28. 75	14. 87	32. 37	
	5	188 282	204 372	92. 13		7. 79	- 5. 04	9. 29	
总平均值					94. 49				25. 47

车头部分可能产生一定的遮挡,导致车斗分割不完全从而导致拟合车斗产生偏差。在质心点位置偏差方面,平均位置偏差均相对较低,总的平均值为 25.47 像素,远低于车斗区域的像素值,因此拟合车斗的平均位置偏差在可接受范围之内,表明拟合车斗区域基本能够完全覆盖实际车斗范围。

因此本文提出的的青贮饲料收获机跟随料车车斗识别方法识别准确性高、识别效果稳定、适应性强,能够满足青贮饲料收获机田间实际作业情况下的需求。

4 结论

(1)提出一种改进 DeepLabV3 + 模型用于青贮

饲料收获机跟随料车车斗分割,以轻量化的 MobileNetV3 网络作为主干网络进行特征提取,降低模型复杂度;在 ASPP 模块中以条形池化替换原有的平均池化加强对车斗的识别能力,并将原有的 3 个空洞卷积分支扩充为 4 个空洞深度可分离卷积分支,提升对大目标的识别能力;在解码器部分融合了多尺度特征,最大程度地弥补了损失的车斗边界细节特征,并在拼接融合后添加 C2f-GS 模块进一步增强模型的特征融合能力。

(2) 在构建的料车车斗数据集上,改进 DeepLabV3 + 模型 mIoU 和平均精度均值达到

94.25% 和 96.62%,且模型参数量和浮点运算量为 6.80×10^6 、 1.567×10^{10} ,整体性能优于其他分割模型,与原 DeepLabV3 + 模型相比,改进模型的 mIoU 和平均精度均值分别提高 4.77、3.71 个百分点,识别准确性得到有效提升。

(3) 利用基于 RANSAC 算法的料车车斗拟合方法对不同情况的分割结果进行车斗边界拟合,结果显示拟合的车斗区域与原标签图的 mIoU 平均值达到 94.49%,二者质心点位置偏差仅为 25.47 像素,表明拟合结果基本能够完全覆盖车斗区域,满足青贮饲料收获机田间实际作业情况需求。

参 考 文 献

[1] 闫威明,陈雅坤,杨鹏标,等.青贮饲料质量评定方法研究进展[J].中国畜牧兽医,2024,51(1):135-144.
YAN Weiming, CHEN Yakun, YANG Pengbiao, et al. Research progress on quality assessment methods of silage [J]. China Animal Husbandry & Veterinary Medicine, 2024, 51(1): 135-144. (in Chinese)

[2] 刘杰,杨蜜,金孟达,等.青贮饲料营养价值、影响因素及其在家畜生产中的应用[J].中国畜牧杂志,2024,60(2):58-64.
LIU Jie, YANG Mi, JIN Mengda, et al. Nutritional value and affecting factors of silage and its application in livestock production [J]. Chinese Journal of Animal Science, 2024, 60(2): 58-64. (in Chinese)

[3] 王泽群,华荣江,赫志飞,等.自走式青饲料收获机主要机械部件技术现状[J].农业工程,2022,12(3):20-25.
WANG Zequn, HUA Rongjiang, HE Zhifei, et al. Technical status of main mechanical parts of self-propelled forage harvester [J]. Agricultural Engineering, 2022, 12(3): 20-25. (in Chinese)

[4] TIAN F Y, XIA K L, WANG J, et al. Design and experiment of self-propelled straw forage crop harvester[J]. Advances in Mechanical Engineering, 2021, 13(7):1-12.

[5] 苗中华,陈苏跃,何创新,等.基于3D视觉的青饲机拖车车斗自动识别与定位方法[J].农业机械学报,2019,50(5):43-49.
MIAO Zhonghua, CHEN Suyue, HE Chuangxin, et al. Automatic identification and location method of forage harvester trailer hopper based on 3D vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 43-49. (in Chinese)

[6] 韩振浩,李佳,苑严伟,等.基于U-Net网络的果园视觉导航路径识别方法[J].农业机械学报,2021,52(1):30-39.
HAN Zhenhao, LI Jia, YUAN Yanwei, et al. Path recognition of orchard visual navigation based on U-Net [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 30-39. (in Chinese)

[7] 苏童,王琳,班超,等.基于语义分割的中后期玉米行间路径导航线检测[J].农业机械学报,2024,55(10):275-285.
SU Tong, WANG Lin, BAN Chao, et al. Interrow path navigation line detection of maize in the middle and late period based on semantic segmentation [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10):275-285. (in Chinese)

[8] 肖珂,夏伟光,梁聪哲.复杂背景下果园视觉导航路径提取算法[J].农业机械学报,2023,54(6):197-204,252.
XIAO Ke, XIA Weiguang, LIANG Congzhe. Visual navigation path extraction algorithm in orchards under complex background [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 197-204, 252. (in Chinese)

[9] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Proceedings of the European Conference on Computer Vision, 2018: 801-818.

[10] 周学成,肖明玮,梁英凯,等.基于改进DeepLabv3+的火龙果园视觉导航路径识别方法[J].农业机械学报,2023,54(9):35-43.
ZHOU Xuecheng, XIAO Mingwei, LIANG Yingkai, et al. Navigation path recognition between dragon orchard using improved DeepLabv3+ network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 35-43. (in Chinese)

[11] 武锦龙,吴虹麒,李浩,等.基于改进DeepLabV3+的养麦苗期无人机遥感图像分割识别方法研究[J].农业机械学报,2024,55(5):186-195.
WU Jinlong, WU Hongqi, LI Hao, et al. Segmentation and recognition of buckwheat seedling remote sensing images from unmanned aircraft based on improved DeepLabV3+ [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(5): 186-195. (in Chinese)

[12] 顾文娟,魏金,阴艳超,等.基于改进DeepLabv3+的番茄图像多类别分割方法[J].农业机械学报,2023,54(12):261-271.

- GU Wenjuan, WEI Jin, YIN Yanchao, et al. Multi-category segmentation method of tomato images based on improved DeepLabv3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(12): 261 – 271. (in Chinese)
- [13] 曹英丽, 赵雨薇, 杨璐璐, 等. 基于改进 DeepLabv3 + 的水稻田杂草识别方法[J]. 农业机械学报, 2023, 54(12): 242 – 252.
- CAO Yingli, ZHAO Yuwei, YANG Lulu, et al. Weed identification method in rice field based on improved DeepLabv3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(12): 242 – 252. (in Chinese)
- [14] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. Computer Science, 2014(4): 357 – 361.
- [15] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834 – 848.
- [16] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[J]. arXiv Preprint arXiv:1706.05587, 2017.
- [17] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251 – 1258.
- [18] 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.
- [19] 马宏兴, 董凯兵, 王英菲, 等. 基于改进 YOLO v5s 的轻量化植物识别模型研究[J]. 农业机械学报, 2023, 54(8): 267 – 276.
- MA Hongxing, DONG Kaibing, WANG Yingfei, et al. Lightweight plant recognition model based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(8): 267 – 276. (in Chinese)
- [20] 陈法法, 董海飞, 何向阳, 等. 改进 U-Net 网络的钢结构表面锈蚀图像分割方法[J]. 电子测量与仪器学报, 2024, 38(2): 49 – 57.
- CHEN Fafa, DONG Haifei, HE Xiangyang, et al. Segmentation method of steel structure surface corrosion image with improved U-Net network [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(2): 49 – 57. (in Chinese)
- [21] 乌兰, 苏力德, 贾立国, 等. 基于改进 DeepLabv3 + 网络的马铃薯根系图像分割方法[J]. 农业工程学报, 2023, 39(3): 134 – 144.
- WU Lan, SU Lide, JIA Ligu, et al. Image segmentation method of potato root based on improved DeepLabv3 + network[J]. Transactions of the CSAE, 2023, 39(3): 134 – 144. (in Chinese)
- [22] HOU Q B, ZHANG L, CHENG M M, et al. Strip pooling: rethinking spatial pooling for scene parsing[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 4003 – 4012.
- [23] 夏晓华, 苏建功, 王耀耀, 等. 基于 DeepLabv3 + 的轻量化路面裂缝检测模型[J]. 激光与光电子学进展, 2024, 61(8): 0812001.
- XIA Xiaohua, SU Jiangong, WANG Yaoyao, et al. Lightweight pavement crack detection model based on DeepLabv3 + [J]. Laser & Optoelectronics Progress, 2024, 61(8): 0812001. (in Chinese)
- [24] 李翠明, 王华, 徐龙儿, 等. 基于改进 DeepLabv3 + 的光伏电站道路识别方法[J]. 上海交通大学学报, 2024, 58(5): 776 – 782.
- LI Cuiming, WANG Hua, XU Long'er, et al. Road recognition method of photovoltaic plant based on improved DeepLabv3 + [J]. Journal of Shanghai Jiaotong University, 2024, 58(5): 776 – 782. (in Chinese)
- [25] 刘慧, 姜建滨, 沈跃, 等. 基于改进 DeepLab V3 + 的果园场景多类别分割方法[J]. 农业机械学报, 2022, 53(11): 255 – 261.
- LIU Hui, JIANG Jianbin, SHEN Yue, et al. Multi-category segmentation of orchard scenes based on improved DeepLab V3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(11): 255 – 261. (in Chinese)
- [26] LIN T Y, DOLLÜR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117 – 2125.
- [27] LI X, LI G, ZHANG Z. A study on the performance improvement of a conical bucket detection algorithm based on YOLO v8s [J]. World Electric Vehicle Journal, 2024, 15(6): 238.
- [28] 许德刚, 王双臣, 尹柯栋, 等. 改进 YOLO v8 的城市车辆目标检测算法[J]. 计算机工程, 2025, 51(11): 377 – 391.
- XU Degang, WANG Shuangchen, YIN Kedong, et al. Improved YOLO v8 urban vehicle target detection algorithm [J]. Computer Engineering and Applications, 2025, 51(11): 377 – 391. (in Chinese)
- [29] 刘忠, 卢安舸, 崔浩, 等. 基于改进 YOLO v8 的轻量化荷叶病虫害检测模型[J]. 农业工程学报, 2024, 40(19): 168 – 176.
- LIU Zhong, LU Ange, CUI Hao, et al. Lightweight model for detecting lotus leaf diseases and pests using improved YOLO v8 [J]. Transactions of the CSAE, 2024, 40(19): 168 – 176. (in Chinese)
- [30] 张彦斐, 魏鹏, 宫金良, 等. 复杂环境中苹果树识别与导航线提取方法[J]. 农业机械学报, 2022, 53(10): 220 – 227.
- ZHANG Yanfei, WEI Peng, GONG Jinliang, et al. Apple tree recognition and navigation line extraction in complex environment[J]. Transactions of the Chinese Society for Agricultural Engineering, 2022, 53(10): 220 – 227. (in Chinese)