

基于改进 YOLO v8n-seg 的羊只实例分割方法

王福顺^{1,2} 王旺¹ 孙小华³ 王超^{1,2} 袁万哲⁴

(1. 河北农业大学信息科学与技术学院, 保定 071001; 2. 河北省农业大数据重点实验室, 保定 071000;

3. 河北软件职业技术学院数字传媒系, 保定 071000; 4. 河北农业大学动物医学院/中兽医学院, 保定 071000)

摘要: 羊只实例分割是实现羊只识别和跟踪、行为分析和健康管理、疾病监测等任务的重要前提。针对规模化羊场复杂养殖环境中,羊只个体存在遮挡、光线昏暗、个体颜色与背景相似等情况所导致的羊只实例错检、漏检问题,提出了一种基于改进 YOLO v8n-seg 的羊只实例分割方法。以 YOLO v8n-seg 网络作为基础模型进行羊只个体分割任务,首先,引入 Large separable kernel attention 模块以增强模型对实例重要特征信息的捕捉能力,提高特征的代表性及模型的鲁棒性;其次,采用超实时语义分割模型 DWR-Seg 中的 Dilation-wise residual 模块替换 C2f 中的 Bottleneck 模块,以优化模型对网络高层特征的提取能力,扩展模型感受野,增强上下文语义之间的联系,生成带有丰富特征信息的新特征图;最后,引用 Dilated reparam block 模块对 C2f 进行二次改进,多次融合从网络高层提取到的特征信息,增强模型对特征的理解能力。试验结果表明,改进后的 YOLO v8n-LDD-seg 对羊只实例的平均分割精度 mAP_{50} 达到 92.08%, $mAP_{50,90}$ 达到 66.54%, 相较于 YOLO v8n-seg, 分别提升 3.06、3.96 个百分点。YOLO v8n-LDD-seg 有效提高了羊只个体检测精度,提升了羊只实例分割效果,为复杂养殖环境下羊只实例检测和分割提供了技术支持。

关键词: 羊只; 个体检测; 实例分割; 改进 YOLO v8n-LDD-seg 网络

中图分类号: TP391

文献标识码: A

文章编号: 1000-1298(2024)08-0322-11

OSID:



Sheep Instance Segmentation Method Based on Improved YOLO v8n-seg

WANG Fushun^{1,2} WANG Wang¹ SUN Xiaohua³ WANG Chao^{1,2} YUAN Wanzhe⁴

(1. College of Information Science and Technology, Hebei Agricultural University, Baoding 071001, China

2. Hebei Key Laboratory of Agricultural Big Data, Baoding 071000, China

3. Department of Digital Transmit, Hebei Software Institute, Baoding 071000, China

4. College of Veterinary Medicine/Traditional Chinese Veterinary Medicine,

Hebei Agricultural University, Baoding 071000, China)

Abstract: Sheep instance segmentation is an important prerequisite for sheep identification and tracking, behavior analysis and management, and disease monitoring. Aiming at the problem of false detection and missed detection of sheep instance detection caused by the occlusion of sheep individuals, dim light, and the similarity of individual color and background in the complex breeding environment of large-scale sheep farms, a sheep instance segmentation method based on improved YOLO v8n-seg was proposed. The YOLO v8n-seg network was used as the basic model for the individual sheep segmentation task. Firstly, the large separable kernel attention module was introduced to enhance the ability of the model to capture important feature information of the instance, which improved the representativeness of the features and the robustness of the model. Secondly, the bottleneck module in C2f was replaced by the expansion-wise residual module in DWR-Seg, a hyperreal-time semantic segmentation model, to optimize the ability of the model to extract high-level network features, expanding the receptive field of the model, and enhanced the relationship between context semantics. Generate new feature maps with rich feature information. Finally, the dilated reparam block module was used to further improve C2f, and the feature information extracted from the high level of the network was fused several times to enhance the

收稿日期: 2024-04-15 修回日期: 2024-05-11

基金项目: 河北省重点研发计划项目(22327403D)和河北省现代农业产业技术体系羊产业创新团队专项资金项目(HBCT2024250204)

作者简介: 王福顺(1981—),男,副教授,博士,主要从事大数据分析挖掘和信息智能化处理研究,E-mail: xxwfs@hebau.edu.cn

通信作者: 袁万哲(1978—),男,教授,博士,主要从事兽用生物制品研究,E-mail: yuanwanzhe@126.com

understanding ability of the model. The experimental results showed that the average segmentation accuracy of the improved YOLO v8n-LDD-seg for sheep cases reached 92.08% at mAP_{50} and 66.54% at $mAP_{50:90}$. Compared with YOLO v8n-seg, mAP_{50} and $mAP_{50:95}$ were improved by 3.06 percentage points and 3.96 percentage points, respectively. YOLO v8n-LDD-seg effectively improved the detection accuracy of individual sheep, improved the segmentation effect of sheep instances, and provided technical support for the detection and segmentation of sheep instances in complex breeding environments.

Key words: sheep; individual detection; instance segmentation; improved YOLO v8n-LDD-seg network

0 引言

信息技术的快速发展大大加快了养殖业的数字化和智能化水平,其中计算机视觉在智慧养殖中发挥着重要作用^[1-3]。在计算机视觉应用于畜禽规模化养殖和管理的过程中,实例分割是一项重要任务,其目的是将图像中的畜禽个体检测并分割出来,是实现个体的精准定位、目标跟踪、行为分析、健康监测等应用的重要前提和基础。

目前,基于视觉的实例检测和分割大致分为传统图像处理^[4-7]和深度学习^[8-12]2种方法。传统图像处理没有自动化训练过程,缺乏灵活性,且不能自动提取图像中的高层次特征,对复杂环境下的实例处理效果不佳,不具备很好的鲁棒性。而深度学习方法可以通过训练自动优化模型,并且适用于多种图像处理任务,具有较强的通用性。因此,深度学习在畜牧检测方面得到广泛应用^[13-15]。QIAO等^[16]为解决真实饲养场环境下的牛只实例分割和轮廓提取问题,通过改进 Mask R-CNN 对牛图像数据集进行了训练和测试,试验结果表明,该方法对牛只的平均分割精度达到了 92%,进一步提高了牛只分割性能。XIAO等^[17]通过改进 Mask R-CNN 分割奶牛的俯视图图像,提取奶牛背部的形状特征,随后利用支持向量机(SVM)实现了对自由栏牛的识别,试验结果表明,改进后的 Mask R-CNN 在奶牛背部平均分割精度达到 97.39%,同时基于所获得的形状特征训练的 SVM 分类器,对奶牛的识别准确率达到 98.67%。CHU等^[18]通过 YOLO v7 和 CenterNet 分别构建奶牛乳房温度特征向量、乳房大小特征向量,将二者特征融合后,利用基于包裹体的支持向量机算法进行自动检测奶牛乳房炎,试验结果表明,该方法准确率达到 88.61%。HU等^[19]采用帧差和分割跨度分析的部分分割算法,分割奶牛侧视图像中的头部、躯干和腿部,利用融合特征训练的支持向量机分类器实现奶牛个体的识别,试验结果表明,该方法在包含 93 头奶牛侧视图像的数据集上取得 98.36% 的奶牛识别准确率,优于现有的研

究成果。梁秀英等^[20]通过 YOLO v8n-seg 实现对小鼠个体的实例分割,试验结果表明,该模型 mAP_{50} 达到 99.2%,为后续小鼠多目标跟踪研究提供了基础。陈佳等^[21]通过 Mask R-CNN 和 YOLACT 2 种实例分割算法实现对肉鸡个体质量估测,试验结果表明,以 Mask R-CNN 和 YOLACT 进行单鸡体质量估测平均准确率分别为 97.23% 和 97.49%,群鸡场景中体质量估测最低准确率为 90.50%。方鹏等^[22]通过调整主干网络、引入注意力机制以及可变形卷积层的方法来改进 Mask R-CNN 实例分割方法实现对鸡只个体轮廓的提取,试验结果表明,优化后网络模型精度均值达到 90.37%。耿艳利等^[23]基于 PointNet 网络,结合注意力模块构建语义分割模型,针对不同分割部位设计猪只体尺测量方法,试验结果表明,改进的语义分割模型在自制数据集上的准确率达到 86.3%。张姝瑾等^[24]在 YOLO v8n-seg 的基础上,通过加入多尺度融合模块与双向跨尺度加权特征金字塔结构,实现目标奶牛前肢、后肢、乳房、尾部、腹部、头部、颈部和躯干 8 个部位的分割,试验结果表明,平均精度均值达到 97.1%,可以准确、快速地实现奶牛身体部位精准分割。王树才等^[25]通过表面凹凸性的点云分割方法实现对羊胴体的精准分割,试验结果表明,羊胴体点云分割平均总体准确率达到 92.2%。覃学标等^[26]使用 ResNet-101 作为主干网络,首先对鱼、鱼头和鱼尾进行目标检测,对图像中具有完整轮廓的鱼类目标进行提取,然后在此区域内运用边缘支持的方法对图像中的鱼进行分割,试验结果表明,该方法对鱼类分割精度达到 85.66%。除此之外,实例检测和分割也被广泛应用于畜禽自动计数研究^[27-29]。

综合分析,目标检测与实例分割技术在牛、猪等畜禽养殖中已经有所应用并取得了一定的成果,但在羊养殖方面的研究还相对匮乏。而且,对于规模化棚舍羊场,其养殖环境较复杂,由于存在羊只个体相互遮挡、棚舍光线强弱不均、个体颜色与背景相似等问题,严重影响羊只个体的识别精度和实例分割效果。因此本文提出一种基于改进 YOLO v8n-seg

的羊只实例分割方法,以期实现对规模化羊场复杂养殖环境下羊只个体精准检测和实例分割。

1 数据获取与预处理

1.1 数据获取

本试验所用羊只视频和图像均于 2023 年 2 月 4 日至 7 月 15 日在河北衡水志豪畜牧科技有限公

司羊只养殖基地采集。采集设备为海康威视摄像头,分辨率为 1 902 像素 × 1 080 像素。摄像头分布于养殖基地多个位置,从多种视角捕捉羊只在不同背景下的形态。考虑到夜间监控视频画质模糊,为确保图像清晰度,选择每天 08:00—18:00 时间段内进行视频片段截取,并运用关键帧提取技术得到羊只图像,如图 1 所示。

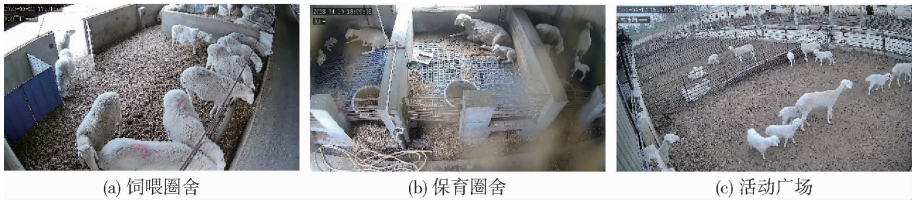


图 1 关键帧提取的羊只图像

Fig.1 Key frame extraction of sheep image

图 1 为饲喂圈舍、保育圈舍和活动广场的图像示例。场景多样性丰富了数据集,让模型能够学习并识别羊只在各种环境中的表现,使模型具备更好的适应性。

1.2 数据预处理

在截取的视频片段中,存在监控画面在一段时

间内并无明显变化的情况。因此,为了确保图像的清晰度和质量,通过数据清洗,去除采集到的无羊和高度相似的图像。数据清洗后,将图像进行对比度和亮度增强处理,图像增强前后对比如图 2 所示。

经过数据清洗和图像增强处理后,选出 1 150 幅图像作为羊只个体实例分割数据集。使用开源图

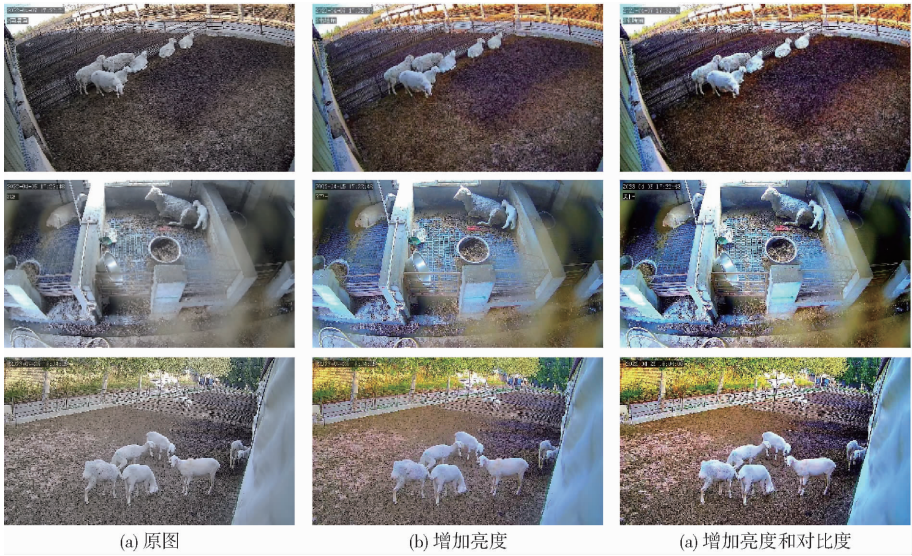


图 2 图像增强前后对比

Fig.2 Comparison of images before and after enhancement

像标注工具 LabelMe^[30]对数据集中的羊只个体实例进行标注,标签命名为“sheep”,其余部分视为“背景”。标注信息保存至与图像名称对应的 json 文件中,通过 txt 文件转换,将 json 文件转换成模型训练需要的 txt 文件格式。将标注好的数据集按照比例 7:3 划分为训练集和验证集,其中训练集包含 805 幅图像,验证集包含 345 幅图像。经过检测,数据集中共标注 8 491 个羊只实例。如图 3 所示,羊只实例数量在不同图像之间呈现出显著波动,有些图像中仅出现 1 个羊只实例,而有些图像中羊只实例数量

高达 48 个。这种差异展现了数据集多样性,对于模型学习具有重要意义,不仅为模型学习提供了丰富的特征信息,而且使模型在复杂环境中能够更全面地学习和适应各种情况,从而增强了模型泛化性和鲁棒性。

2 羊只实例分割模型

2.1 YOLO v8n-seg 分割网络

YOLO v8n-seg 网络结构主要由输入层 (Input)、主干网络 (Backbone)、颈部网络 (Neck)、分

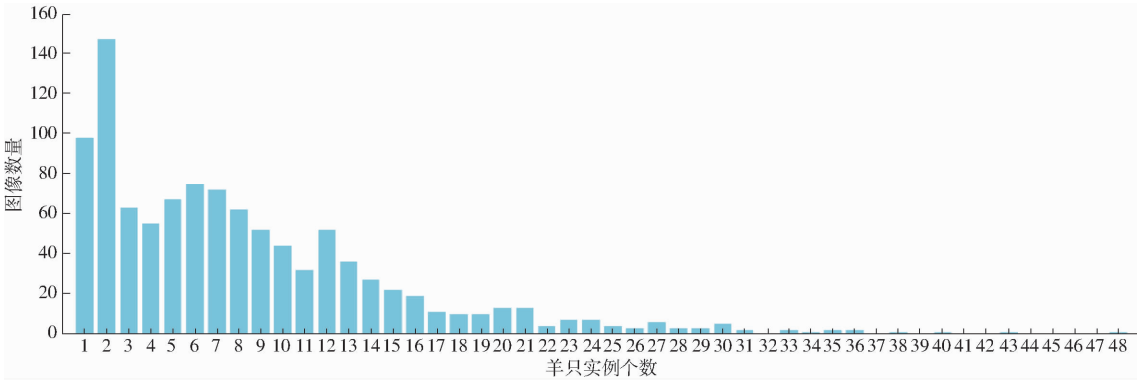


图 3 相同羊只实例个数的图像数量

Fig. 3 Number of images with the same number of sheep instances

割层 (Segment) 组成。输入层主要用来接收需要分割的图像样本,为模型的训练提供必需数据;主干网络主要是从图像中提取各种尺度和类别的实例特征,常见的主干网络结构包含 Darknet、ResNet 等,通过串联的卷积和池化层逐层深化需要提取的特征;颈部网络的作用在于整合主干网络所提取到的特征信息,增强模型对特征的理解;分割层主要通过分析各层特征之间的相互关联来完成类别预测和像素级的实例分割等精细化任务。将待分割的图像输入到 YOLO v8n-seg 模型,模型将执行目标检测和实例分割 2 个阶段。在目标检测阶段,模型依据算法逻辑识别出图像中的各个目标实例,并标注出实例的边界框以及类别标签。在实例分割阶段,模型对图像中的每个像素点进行精准分类,完成像素级别的分割。YOLO v8n-seg 同时实现了目标检测和实例分割,提高了模型对目标实例的精准检测。

2.2 YOLO v8n-seg 模型优化

2.2.1 基于 LSKA 的 SPPF-LSKA 模块

羊只实例分割数据集在经过数据预处理后,因光照、拍摄角度等外界因素影响,部分羊只图像仍存在曝光过度以及羊只实例与背景高度相似等问题,导致模型在训练过程中,提取到的特征缺乏代表性和区分度,进而影响模型准确性和泛化能力。为了解决上述问题,提高模型对羊只实例具有代表性特征信息的提取能力,在 SPPF 模块中引入大型可分离核注意力 (Large separable kernel attention, LSKA)^[31] 模块。将其命名为 SPPF-LSKA, SPPF-LSKA 和 LSKA 结构如图 4 所示。

当特征图进入 SPPF-LSKA 模块后,通过池化层对不同尺度的特征信息进行压缩、降维操作,避免较多噪声和冗余信息介入,扩大模型感受野。之后通过连接层,将局部区域内的特征值合并传入 LSKA 模块中。LSKA 采用核分解(将深度卷积层的 2D 卷积核分解为水平和垂直 1D 卷积核)和串联卷

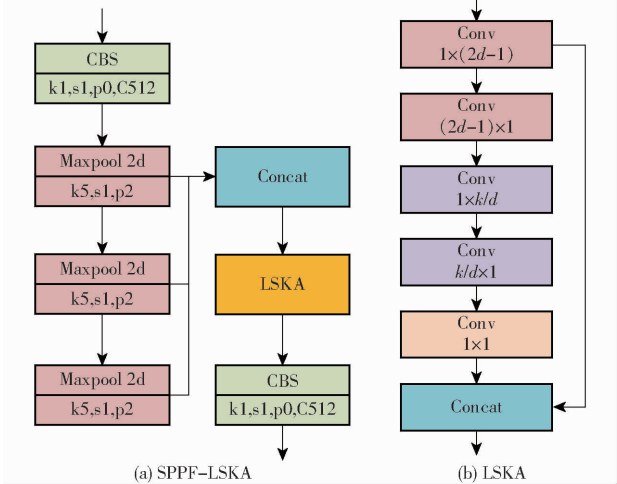


图 4 改进 SPPF-LSKA 和 LSKA 结构

Fig. 4 Improved SPPF-LSKA and LSKA structures

积的策略,保留实例中的重要特征信息,提取出更加鲁棒和具有代表性的特征,提高模型对羊只实例图像理解能力。其输出公式为

$$Z_1^C = \sum_{H,W} W_{(2d-1) \times 1}^C * \left(\sum_{H,W} W_{1 \times (2d-1)}^C * F_1^C \right) \tag{1}$$

$$Z_2^C = \sum_{H,W} W_{\frac{k}{d} \times 1}^C * \left(\sum_{H,W} W_{1 \times \frac{k}{d}}^C * Z_1^C \right) \tag{2}$$

$$A^C = W_{1 \times 1} * Z_2^C \tag{3}$$

$$F^C = A^C \otimes F_1^C \tag{4}$$

式中 *、 $\otimes$ ——卷积、哈达玛乘积
d——扩张率 k——内核大小
 F_1^C ——输入特征图
 Z_1^C ——深度卷积输出,由输入特征图 F_1^C 和大小为 $1 \times (2d-1)$ 、 $(2d-1) \times 1$ 的卷积得到
 Z_2^C ——全局空间信息,由 Z_1^C 和大小为 $\frac{k}{d} \times 1$ 、 $1 \times \frac{k}{d}$ 的扩张深度卷积得到
 A^C ——注意力图谱,由 Z_2^C 与 1×1 内核卷积得到

F^C ——注意力图谱 A^C 和特征图谱 F_1^C 的哈达玛乘积

2.2.2 引入上下文特征融合 C2f-DWR-DRB

2.2.2.1 基于 DWR 的 C2f-DWR 模块

实例分割数据集为不同位置、不同角度下拍摄得到的羊群图像组合而成。其中,由于“保育圈舍”位于室内,因光线不足,羊只实例的可见度降低,导致在该位置采集的图像中部分羊只实例成像昏暗,影响了模型实例分割效果,如图 5 所示。

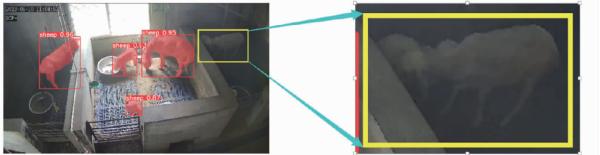


图 5 成像昏暗实例分割效果图

Fig. 5 Segmentation effect of imaging dim instance

为了提高特征表示,增强模型对网络高层细节特征的提取能力,将超实时语义分割模型 DWR-Seg 中的 Dilation-wise residual (DWR)^[32] 模块替换 C2f 模块中的 Bottleneck 模块,用以引入膨胀卷积技术以扩大感受野并捕捉更多上下文信息,从而提升细节特征的表达能力。DWR 模块采用残差方式设计,首先将特征输入到一个 3×3 的卷积,得到具有不同大小的区域残差特征,然后使用不同空洞率的空洞深度卷积通过分支结构对不同尺寸的区域特征进行形态学过滤,之后通过 1×1 的逐点卷积进行特征融合,有效获取不同网络层之间的特征信息,新特征图加强了上下文特征的联系,形成一个更为丰富的特征表示。将改进后的 C2f 模块命名为 C2f-DWR。DWR 和 C2f-DWR 模块结构如图 6 所示。其公式表示为

$$C_1(x) = \text{SiLU}(\text{BN}(\text{Conv}(x))) \tag{5}$$

$$C_2(x, d) = D_d \text{DConv}(C_1(x)) \tag{6}$$

$$\text{DWR}(x) = \text{PConv}(\text{BN}(\Gamma_d\{C_2(x, d)\})) + x \tag{7}$$

式中 x ——输入特征映射

$\text{Conv}()$ —— 3×3 卷积

$\text{BN}()$ ——Batch Normalization 层

$\text{SiLU}()$ ——激活函数

$D_d \text{DConv}()$ ——具有空洞率 d 的空洞深度方向 3×3 的卷积

$\text{PConv}()$ ——逐点卷积

Γ_d ——所有 d 的级联运算

$\text{DWR}()$ ——DWR 模块

2.2.2.2 基于 DRB 的 C2f-DWR-DRB 模块

羊群生活环境复杂,导致数据集图像中普遍存在羊只实例被遮挡的现象。遮挡因素分为两方面,一方面是外界环境如圈舍栏杆和栏舍围墙带来的遮

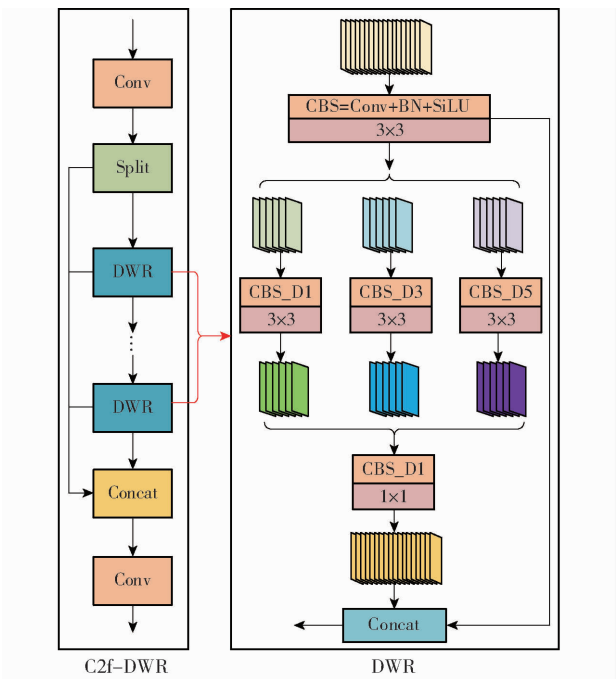


图 6 C2f-DWR 和 DWR 模块结构图

Fig. 6 Structure diagram of C2f-DWR and DWR module

挡,另一方面是羊群中羊只个体间的相互遮挡。这些遮挡导致被遮挡区域的特征信息缺失,使得实例特征不完整,进而影响模型在训练过程中对羊只完整特征信息的提取,影响模型的实例分割效果,如图 7 所示。

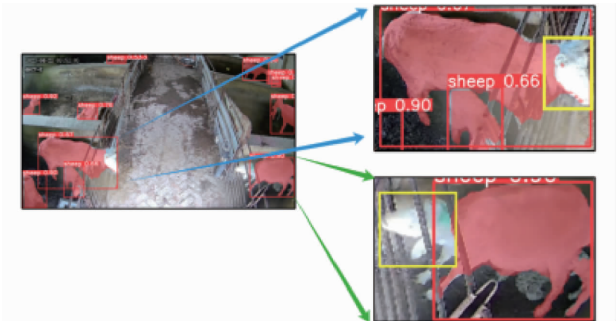


图 7 特征信息缺失分割效果图

Fig. 7 Feature information missing segmentation effect

为了解决因特征信息缺失导致羊只实例分割不完整的问题,对 C2f 进行二次改进,在 C2f-DWR 模块的基础上,通过引入 Dilated reparam block (DRB) 模块改进 DWR 模块,并将其命名为 C2f-DWR-DRB 模块。改进 DWR-DRB 和 DRB 结构如图 8 所示。DRB 模块首先将输入的特征图通过多个卷积层进行多层次特征提取和转换,使每个卷积层学习到不同层次的特征表示,然后通过多个连接层将来自不同卷积层和连接层的特征信息进行多次融合,提高特征多样性和丰富性,从而增强模型对输入数据的理解能力,提高模型对实例的分割效果。

2.2.3 实例分割增强网络 YOLO v8n-LDD-seg

在 YOLO v8n-seg 分割网络的基础上,将改进

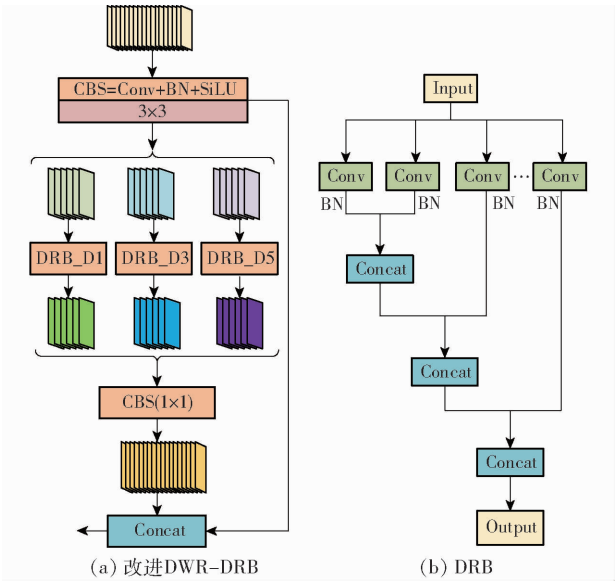


图 8 改进 DWR-DRB 和 DRB 结构图

Fig.8 Structure diagram of improved DWR-DRB

SPPF-LASK 模块和 C2f-DWR-DRB 模块整合到模型中,提出一种全新的实例分割增强网络 YOLO v8n-LDD-seg。具体改进分为 3 点:①为了有效区分图像中的羊只实例和背景信息,提高模型对羊只

实例的理解能力和分割效果,在 SPPF 模块中引入大型可分离核注意力 (LSKA) 模块,旨在增强模型对实例重要特征的提取能力,使其更好地适应复杂场景。②为了提高模型对多角度、曝光度不足的羊只实例分割效果,引用超实时语义分割模型 DWR-Seg 中的 DWR 模块替换 C2f 模块中的 Bottleneck 模块,旨在扩展模型感受野,增强特征表示和不同网络层之间的特征联系,使其更有效地检测和分割昏暗环境下的羊只实例。③为了提高模型对被遮挡羊只的实例分割效果,改进 C2f-DWR 模块,新增 DRB 细节特征融合模块替换 DWR 中的卷积模块,旨在增强模型对特征信息的融合能力,使其更易理解提取到的丰富特征,有效检测羊只实例的边缘信息,提高被遮挡羊只实例的分割效果。改进后的 YOLO v8n-LDD-seg 网络结构如图 9 所示。

2.4 参数设置

试验中,模型的改进和运行均在表 1 所示配置的服务器上进行。尽管各模型之间存在结构上的差异,但在参数设置上统一。本试验设定模型训练轮次总计为 400 轮;为了避免模型出现过拟合现象,设

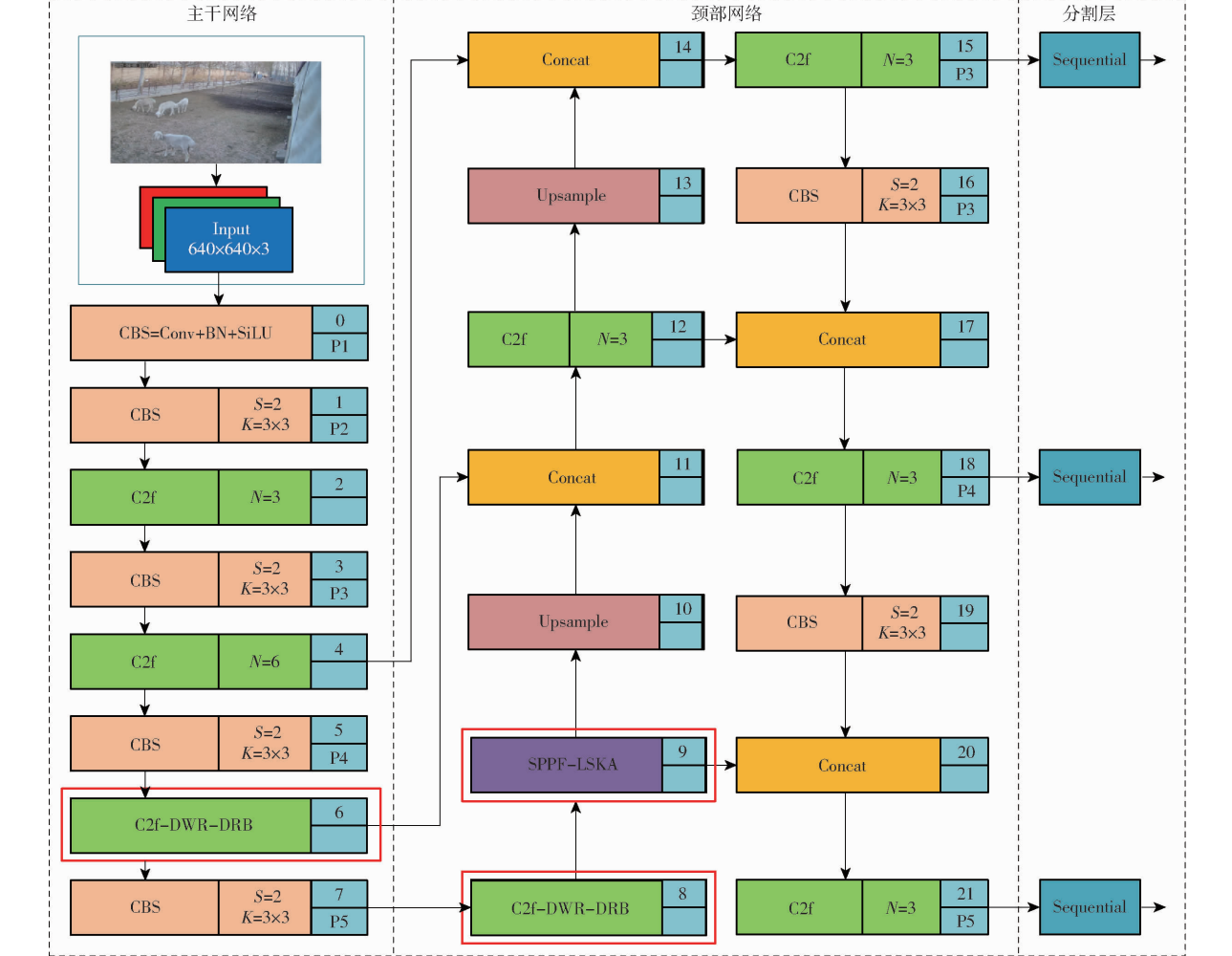


图 9 改进后的 YOLO v8n-LDD-seg 网络结构图

Fig.9 Structure diagram of improved YOLO v8n-LDD-seg

表 1 服务器配置与运行环境

Tab.1 Server configuration and operating environment

软硬件	配置	运行包	版本
操作系统	Windows 10	Python	3. 7. 16
CPU	Intel(R) Core(TM) i9 – 13900K	PyTorch	1. 13. 0
GPU	NVIDIA GeForce RTX 3090	CUDA	11. 6

置早停机机制的等待轮数为 50 轮;训练时每个批次所处理的图像数量设定为 16 幅;输入图像分辨率为 640 像素 × 640 像素;为获得更佳的计算效率,运用 GPU 设备进行模型训练;在数据加载阶段,为确保数据能够高效地输送至模型中,工作线程数设为 8;为了优化模型参数并最小化损失函数,试验选用动量参数为 0.937 的 SGD 优化器;为了保证试验可重复性,设置随机种子为 0,从而便于复现试验结果。

2.5 评价指标

本研究采用了多个评价指标来衡量不同模型对羊只实例分割的效果,包括 mAP_{50} 、 $mAP_{50:95}$ 、精确率 (Precision, P)、召回率 (Recall, R) 以及 F1 值。

在模型评估方面,试验选用平均分割准确率、参

数量、计算量、内存占用量以及帧速率作为评估各个模型在羊只实例分割任务中的性能。平均分割准确率是性能评估中的关键指标,反映模型在整体上对所有羊只分类和分割准确度。参数量表示模型复杂性,参数越少意味着模型结构越简洁。内存占用量反映模型对计算资源的要求,特别是在资源有限的设备上,低内存占用量尤显重要。而帧速率则体现了模型实时性能,即每秒能够处理的图像帧数,高帧速率对于实时分割任务至关重要。因此,综合考虑各个指标是决定模型性能的关键。

3 结果分析

3.1 模型精度对比

将改进后的 YOLO v8n – LDD – seg 分别与 YOLO v8n – seg 和 YOLO v9 – seg 进行模型分割精度对比,模型分割精度对比如表 2 所示。通过表 2 可以看出,相较于 YOLO v9 – seg, YOLO v8n – seg 对羊只实例的分割精度略有降低,随着对 YOLO v8n – seg 的改进, YOLO v8n – LDD – seg 在精确率、召回率、 mAP_{50} 和 $mAP_{50:95}$ 指标上均取得了提升。

表 2 模型分割精度对比

Tab.2 Comparison of model segmentation accuracy

模型	LSKA	DWR	DRB	P	R	mAP_{50}	$mAP_{50:95}$	F1 值
YOLO v9 – seg	×	×	×	88. 50	85. 15	91. 30	64. 91	86. 79
YOLO v8n – seg	×	×	×	88. 38	83. 32	89. 02	62. 58	85. 65
	✓	×	×	88. 81	83. 62	90. 20	65. 16	86. 31
YOLO v8n – LDD – seg	✓	✓	×	90. 50	83. 99	91. 15	65. 48	86. 98
	✓	✓	✓	89. 56	85. 14	92. 08	66. 54	87. 20

注:“ × ”表示不使用该模块,“ ✓ ”表示使用该模块。下同。

相较于 YOLO v9 – seg, YOLO v8n – seg mAP_{50} 和 $mAP_{50:95}$ 分别降低 2.28、2.33 个百分点,而 YOLO v8n – LDD – seg 则分别提高 0.78、1.63 个百分点。相较于 YOLO v8n – seg,改进后的 YOLO v8n – LDD – seg 模型引入 LSKA 使得模型 mAP_{50} 提升 1.18 个百分点,而 $mAP_{50:95}$ 提升了 2.58 个百分点。进一步采用 LSKA + C2f – DWR 改进方案,精确率提升至 90.50%,召回率提升至 83.99%, mAP_{50} 提升 2.13 个百分点, $mAP_{50:95}$ 提升 2.90 个百分点。最后,采用 LSKA + C2f – DWR – DRB 改进方案,召回率提升至 85.14%,但精确率略微降低至 89.56%,原因是改进后的模型更加适应数据集的特性,使得在更多情况下能够成功地检测并分割羊只实例目标,从而提升了召回率。但在一些情况下,由于模糊边界或者复杂背景,可能会导致一些误检,从而降低了精确率。但模型 mAP_{50} 提升 3.06 个百分点, $mAP_{50:95}$ 提升 3.96 个百分点。

综上所述,改进后的 YOLO v8n – LDD – seg 在检测和分割羊只实例任务上表现优异,相较于 YOLO v8n – seg 有着显著的性能提升,并且在分割效果上也优于 YOLO v9 – seg 模型。

3.2 模型参数对比

将改进后的 YOLO v8n – LDD – seg 分别与 YOLO v8n – seg 和 YOLO v9 – seg 进行模型参数对比,模型参数对比如表 3 所示。

通过表 3 可知, YOLO v9 – seg 拥有最大的参数量和计算量,相较于 YOLO v8n – LDD – seg,分别提高 24 104 640 和 1.32×10^{11} ,但是高参数量和计算量并不意味着更好的分割效果。对于羊只实例分割任务, YOLO v9 – seg 的 mAP_{50} 、 $mAP_{50:95}$ 较 YOLO v8n – LDD – seg 分别低 0.78、1.63 个百分点。综合来看, YOLO v8n – LDD – seg 不仅分割精度优于 YOLO v9 – seg,而且参数量、计算量和内存占用量也具有明显优势。相较于 YOLO v8n – seg,尽管改进后的

表 3 模型参数对比

Tab.3 Comparison of model parameters

模型	LSKA	DWR	DRB	mAP ₅₀ /%	mAP _{50,95} /%	参数量	计算量	内存占用量/MB	帧速率/(f·s ⁻¹)
YOLO v9-seg	×	×	×	91.30	64.91	27 362 899	1.44 × 10 ¹¹	53.06	121.95
YOLO v8n-seg	×	×	×	89.02	62.58	3 258 259	1.20 × 10 ¹⁰	6.49	303.03
	√	×	×	90.20	65.16	3 531 155	1.22 × 10 ¹⁰	7.02	263.16
YOLO v8n-LDD-seg	√	√	×	91.15	65.48	3 470 099	1.21 × 10 ¹⁰	6.92	277.78
	√	√	√	92.08	66.54	3 368 979	1.19 × 10 ¹⁰	6.77	312.50

YOLO v8n-LDD-seg 参数量和内存占用量有一定程度的增加,但在羊只实例分割任务中,其性能却得到了明显提升。具体而言,改进后的 YOLO v8n-LDD-seg 模型引入 LASK 模块后,虽然模型参数量、计算量和内存占有量增加较为明显,分别增加 272 896、 2.0×10^8 、0.53 MB,但是模型 mAP_{50,95} 提升 2.58 个百分点,这表明 LASK 的加入使模型在更高要求的实例分割精度方面取得显著提高,增加了模型对更高实例分割要求的关注度。C2f-DWR-DRB 在 LSKA 的基础上,进一步提升了模型分割精度,mAP₅₀ 和 mAP_{50,95} 分别达到 92.08% 和 66.54%,且一定程度上降低了模型参数量和计算量,原因是 DWR 模块设计了更高效特征提取机制,扩张卷积在不同网络层上捕获特征,有助于提高网络对上下文信息的感知和表征能力。通过这种方式,模型可以在不增加参数量的情况下获得更丰富的特征表示,从而减少需要学习的参数量。DRB 模块对网络参数进行重新组织和精简,以提高参数效率和利用率。通过对参数进行重新调整,更有效地利用参数空间,从而减少不必要的参数冗余和计算量,降低整体参数量和计算量。

综上所述,尽管改进的 YOLO v8n-LDD-seg 模型在参数量上有所增长,但其计算复杂度相较于原始的 YOLO v8n-seg 模型实现了有效降低。这种改进确保了模型在执行羊只实例分割任务时性能得到显著提升。在保持效率的同时,模型精确度和鲁棒性也得到了兼顾与优化,实现更为准确的羊只实例检测和分割。

3.3 模型效果对比

针对 YOLO v8n-seg 在规模化羊场复杂养殖环境中,对实例特征提取不充分,导致分割效果不理想的问题,在 YOLO v8n-seg 基础上进一步改进。改进后的 YOLO v8n-LDD-seg 模型整合大型可分离卷积核注意力模块 LSKA、增强特征语义模块 DWR 和细节特征融合模块 DRB,旨在增强模型对羊只实例特征的捕捉能力,尤其是对昏暗、与背景相似以及遮挡条件下的羊只实例分割性能,进一步提升分割精度。

对比 YOLO v8n-seg 和改进后的 YOLO v8n-LDD-seg 模型在不同场景下羊只实例上的分割和检测效果,如图 10 所示。通过观察图像可以发现,在羊只实例检测和分割过程中,改进后的 YOLO v8n-LDD-seg 相对于 YOLO v8n-seg 表现了更好的分割效果。



图 10 不同场景下 YOLO v8n-LDD-seg 和 YOLO v8n-seg 效果对比

Fig. 10 Comparison of effects of YOLO v8n-LDD-seg and YOLO v8n-seg in different scenarios

图 10b 相对于图 10a,改进后的模型能够识别并分割出图像改进点 1 位置出现的昏暗实例,这表明改进后的模型对于昏暗情况下的羊只实例目标能够更好地进行检测和分割。图 10d 相对于图 10c,在图像改进点 2、改进点 3 因围栏遮挡导致羊只实例头部丢失的情况下,改进后的模型展现了更高的稳定性和精确度,能够更好地处理困难情况,避免因遮挡而导致分割区域缺失的问题。此外,改进点 4 位置的两只羊因相互遮挡被 YOLO v8n-seg 检测成一个实例。在改进后的模型中得到了有效解决,模

型将它们准确地区分开,避免了将多个实例错误检测为单一实例的情况。图 10f 相对于图 10e,改进后的模型不仅成功地分割出图像边界处改进点 5 位置上的不完整羊只实例,而且在整体羊只实例的目标置信度上也得到显著提高,这表明改进后的模型更有效地理解了羊只实例的特征和边缘信息,从而在分割任务中表现出更高的置信度和更好的分割效果。图 10h 相对于图 10g,图像改进点 6 位置上的趴卧羊只由于存在部分遮挡而未被 YOLO v8n-seg 正确地分割出来。改进后的模型对于遮挡羊只和被遮挡羊只都得到更好的目标边缘分割效果,进一步证明了改进后的模型在理解场景和羊只实例方面的能力得到提升。

综上所述,改进后的 YOLO v8n-LDD-seg 模型在规模化羊只养殖的不同场景下,羊只实例的检测与分割任务上的表现显著优于原始模型。该模型有效地提升了在复杂场景下羊只实例的处理效果,从而在实际应用场景中实现更高的精确性和稳定性。

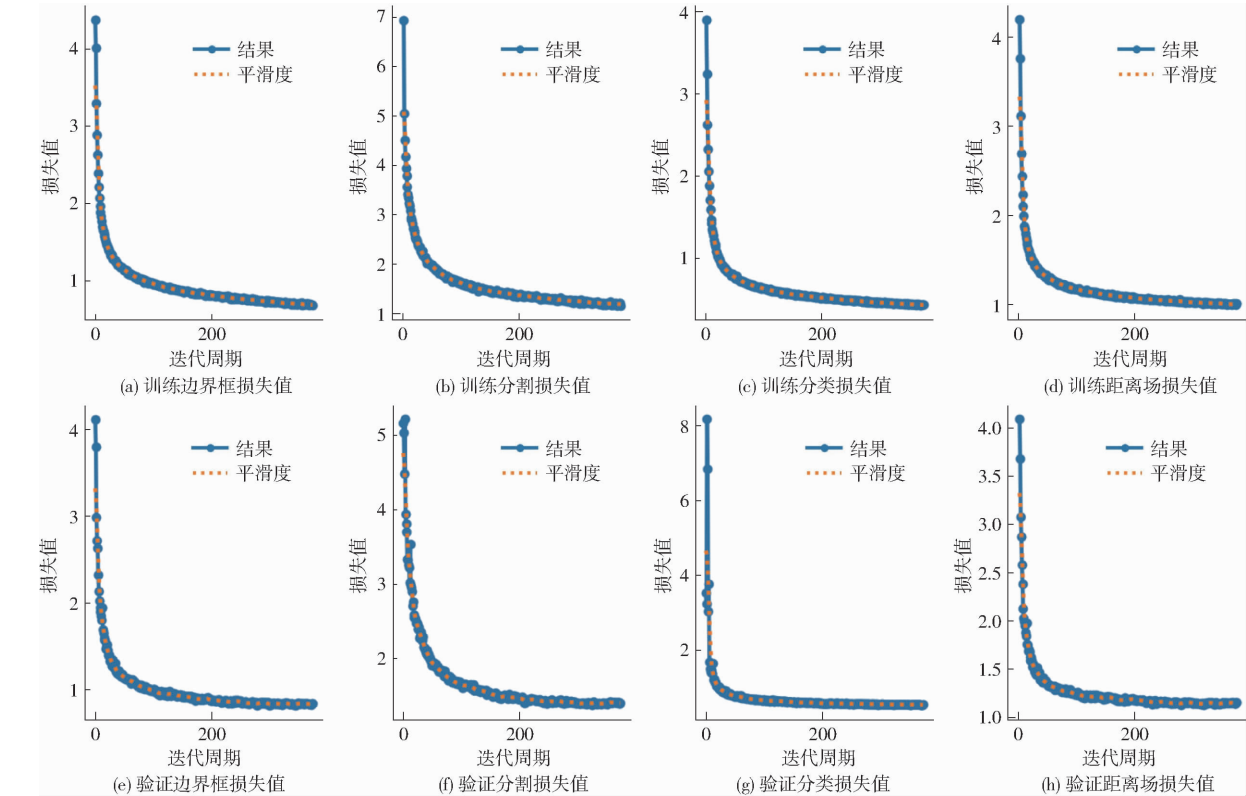


图 11 训练过程损失函数变化曲线

Fig. 11 Change curves of each loss function during training

综上所述,平滑的损失值变化曲线表明模型能够在训练和验证过程中对于不同场景下羊只实例进行准确检测和分割,反映了模型在进行羊只实例分割任务时的稳定性,以及对于复杂羊只养殖场景的良好适应性,同时意味着模型具有更好的泛化能力,在面对新场景时也能够表现出色。

3.4 模型稳定性

为进一步验证模型稳定性,图 11 展示了 YOLO v8n-LDD-seg 模型分割羊只实例训练过程中各个损失函数变化曲线。边界框损失函数代表预测边界框和真实标注框之间的差异;分割损失函数代表像素级别上模型预测的分割掩膜和真实分割掩膜之间的差异;分类损失函数代表模型对每个目标物体类别预测准确性;距离场损失函数代表模型预测的距离场与真实距离场之间的差异。从图 11 可以看出,无论训练还是验证,各个损失函数变化曲线最终都稳定趋于平滑状态。平滑的边界框损失变化曲线表明模型在定位羊只实例的表现上更加精确,边界框位置和尺寸与真实情况更加吻合;平滑的分割损失值变化曲线表示模型能够更准确地捕捉羊只实例中不同区域的语义信息,从而分割出羊只的形状和轮廓;平滑的分类损失值变化曲线和距离场损失变化曲线分别说明模型对于不同类别特征的分类学习和边界位置的预测具有稳定性。

4 结论

(1) 针对规模化羊场复杂养殖环境下,拍摄到的羊只实例图像受光照、拍摄角度、外界环境等复杂场景以及羊只互相遮挡等因素影响,造成模型对羊只分割精度低和分割效果不佳的问题,本文提出了

一种改进 YOLO v8n-seg 的羊只实例分割方法。

(2)LSKA 采用核分解和串联卷积的策略,使得模型能够在多个卷积层之间传递和整合信息,更全面地理解图像的结构,进一步增强特征的鲁棒性和代表性;DWR 模块通过提取网络高层的特征信息,将其融合生成新的特征图,增强上下文特征之间的联系,形成一个更为丰富的特征表示;DRB 模块通过多层次的特征提取和融合,增强模型对数据的理解,从而提高模型对羊只边缘信息的学习和认识。

(3)改进后的 YOLO v8n-LDD-seg 在羊只实例的平均分割精度达到 92.08%,mAP_{50:90} 达到 66.54%。相较于 YOLO v8n-seg,mAP₅₀ 和 mAP_{50:95}

分别提升 3.06、3.96 个百分点。虽然改进后的模型在参数量上略有提升,但是考虑到模型对实例分割精度的显著提升,这些参数量的增加是值得的,同时计算量的降低也提高了模型计算效率。

(4)由于距离监控位置较远的羊只实例在监控画面中往往呈现出模糊和较小的外观特征,造成小目标羊只实例边缘信息不明显。在对数据集图像进行分割和检测过程中,模型不仅存在将多个小目标实例误认为是一个实例的情况,而且存在部分小目标羊只实例被识别成背景的情况,对小目标羊只实例的分割效果欠佳。未来的研究可以重点关注提升模型对小目标羊只实例分割效果的研究,以进一步提高模型对整体分割精度的表现。

参 考 文 献

[1] LELIVELD L M C, PROVOLO G. A review of welfare indicators of indoor-housed dairy cow as a basis for integrated automatic welfare assessment systems[J]. *Animals*, 2020, 10(8): 1430-1447.

[2] KANG J, ULLAH Z, GWAK J. MRI-based brain tumor classification using ensemble of deep features and machine learning classifiers[J]. *Sensors*, 2021, 21(6): 2222-2242.

[3] 张宏鸣,孙扬,赵春平,等.反刍家畜典型行为监测与生理状况识别方法研究综述[J]. *农业机械学报*,2023,54(3):1-21.
ZHANG Hongming, SUN Yang, ZHAO Chunping, et al. Review on typical behavior monitoring and physiological condition identification methods for ruminant livestock[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023,54(3): 1-21. (in Chinese)

[4] ALBUQUERQUE P L F, GARCIA V, JUNIOR A S O, et al. Automatic live fingerlings counting using computer vision[J]. *Computers and Electronics in Agriculture*, 2019, 167: 105015.

[5] AWALLUDIN E A, MUHAMMAD W N A W, ARSAD T N T, et al. Fish larvae counting system using image processing techniques[C]//*Journal of Physics: Conference Series*. IOP Publishing, 2020, 1529(5): 052040.

[6] YEH C T, LIND M S. Portable device for ornamental shrimp counting using unsupervised machine learning[J]. *Sensors & Materials*, 2021, 33(9): 3027-3036.

[7] NGO T N, WU K C, YANG E C, et al. A real-time imaging system for multiple honey bee tracking and activity monitoring[J]. *Computers and Electronics in Agriculture*, 2019, 163: 104841.

[8] WANG Y, CHU M, KANG X, et al. A deep learning approach combining DeepLabV3+ and improved YOLOv5 to detect dairy cow mastitis[J]. *Computers and Electronics in Agriculture*, 2024, 216: 108507.

[9] ZHENG Z, LI J, QIN L. YOLO-BYTE: an efficient multi-object tracking algorithm for automatic monitoring of dairy cows[J]. *Computers and Electronics in Agriculture*, 2023, 209: 107857.

[10] MA S, ZHANG Q, LI T, et al. Basic motion behavior recognition of single dairy cow based on improved Rexnet 3D network [J]. *Computers and Electronics in Agriculture*, 2022, 194: 106772.

[11] YU Z, LIU Y, YU S, et al. Teat detection of dairy cows based on deep learning neural network FS-YOLOv4 model[J]. *Computers and Electronics in Agriculture*, 2022, 200: 107224.

[12] WANG D, TANG J L, ZHU W, et al. Dairy goat detection based on Faster R-CNN from surveillance video[J]. *Computers and Electronics in Agriculture*, 2018, 154: 443-449.

[13] 景庄伟,管海燕,臧玉府,等.基于深度学习的点云语义分割研究综述[J]. *计算机科学与探索*, 2021, 15(1): 1-26.
JING Zhuangwei, GUAN Haiyan, ZANG Yufu, et al. Survey of point cloud semantic segmentation based on deep learning[J]. *Journal of Frontiers of Computer Science and Technology*, 2021, 15(1): 1-26. (in Chinese)

[14] QIAO Y, CAPPELLE C, RUICHEK Y, et al. ConvNet and LSH-based visual localization using localized sequence matching [J]. *Sensors*, 2019, 19(11): 2439.

[15] KUMAR S, PANDEY A, SATWIK K S R, et al. Deep learning framework for recognition of cattle using muzzle point image pattern[J]. *Measurement*, 2018, 116: 1-17.

[16] QIAO Y, TRUMAN M, SUKKARIEH S. Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming[J]. *Computers and Electronics in Agriculture*, 2019, 165: 104958.

[17] XIAO J, LIU G, WANG K, et al. Cow identification in free-stall barns based on an improved Mask R-CNN and an SVM[J]. *Computers and Electronics in Agriculture*, 2022, 194: 106738.

[18] CHU M, LI Q, WANG Y, et al. Fusion of udder temperature and size features for the automatic detection of dairy cow mastitis

using deep learning[J]. Computers and Electronics in Agriculture, 2023, 212: 108131.

[19] HU H, DAI B, SHEN W, et al. Cow identification based on fusion of deep parts features[J]. Biosystems Engineering, 2020, 192: 245 – 256.

[20] 梁秀英, 贾学镇, 何磊, 等. 基于 YOLO v8n – seg 和改进 Strongsort 的多目标小鼠跟踪方法[J]. 农业机械学报, 2024, 55(2): 295 – 305, 345.

LIANG Xiuying, JIA Xuezhen, HE Lei, et al. Multi-object mice tracking based on YOLO v8n – seg and improved Strongsort [J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 295 – 305, 345. (in Chinese)

[21] 陈佳, 刘龙申, 沈明霞, 等. 基于实例分割的白羽肉鸡体质量估测方法[J]. 农业机械学报, 2021, 52(4): 266 – 275.

CHEN Jia, LIU Longshen, SHEN Mingxia, et al. Breeding white feather broiler weight estimation method based on instance segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 266 – 275. (in Chinese)

[22] 方鹏, 郝宏运, 李腾飞, 等. 基于注意力机制和可变形卷积的鸡只图像实例分割提取[J]. 农业机械学报, 2021, 52(4): 257 – 265.

FANG Peng, HAO Hongyun, LI Tengfei, et al. Instance segmentation of broiler image based on attention mechanism and deformable convolution[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 257 – 265. (in Chinese)

[23] 耿艳利, 季燕凯, 岳晓东, 等. 基于点云语义分割的猪只体尺测量方法研究[J]. 农业机械学报, 2023, 54(7): 332 – 338, 380.

GENG Yanli, JI Yankai, YUE Xiaodong, et al. Pigs body size measurement based on point cloud semantic segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(7): 332 – 338, 380. (in Chinese)

[24] 张姝瑾, 许兴时, 邓洪兴, 等. 基于 YOLO v8n – seg – FCA – BiFPN 的奶牛身体分割方法[J]. 农业机械学报, 2024, 55(3): 282 – 289, 391.

ZHANG Shujin, XU Xingshi, DENG Hongxing, et al. Segmentation model of cow body parts based on YOLO v8n – seg – FCA – BiFPN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3): 282 – 289, 391. (in Chinese)

[25] 王树才, 白宇, 赵世达, 等. 基于表面凹凸性的羊胴体点云分割方法[J]. 农业机械学报, 2022, 53(7): 387 – 394.

WANG Shucai, BAI Yu, ZHAO Shida, et al. Point cloud segmentation of sheep carcass based on surface convexity[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 387 – 394. (in Chinese)

[26] 覃学标, 黄冬梅, 宋巍, 等. 基于目标检测及边缘支持的鱼类图像分割方法[J]. 农业机械学报, 2023, 54(1): 280 – 286.

QIN Xuebiao, HUANG Dongmei, SONG Wei, et al. Fish image segmentation method based on object detection and edge support[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 280 – 286. (in Chinese)

[27] XU B, WANG W, FALZON G, et al. Automated cattle counting using Mask R – CNN in quadcopter vision system[J]. Computers and Electronics in Agriculture, 2020, 171: 105300.

[28] HUANG E, MAO A, GAN H, et al. Center clustering network improves piglet counting under occlusion[J]. Computers and Electronics in Agriculture, 2021, 189: 106417.

[29] CAO Y, CHEN J, ZHANG Z. A sheep dynamic counting scheme based on the fusion between an improved-sparrow-search YOLOv5x – ECA model and few-shot deepsort algorithm[J]. Computers and Electronics in Agriculture, 2023, 206: 107696.

[30] RUSSELL B C, TORRALBA A, MURPHY K P, et al. LabelMe: a database and web-based tool for image annotation[J]. International Journal of Computer Vision, 2008, 77: 157 – 173.

[31] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN [J]. Expert Systems with Applications, 2024, 236: 121352.

[32] WEI H, LIU X, XU S, et al. DWRSeg: dilation – wise residual network for real-time semantic segmentation[J]. arXiv Preprint, arXiv:2212.01173, 2022.