

基于 YOLO v8n-seg-FCA-BiFPN 的奶牛身体分割方法

张姝瑾^{1,2} 许兴时^{1,2} 邓洪兴^{1,2} 温毓晨^{1,2} 宋怀波^{1,2}

(1. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100;

2. 农业农村部农业物联网重点实验室, 陕西杨凌 712100)

摘要: 奶牛身体部位的精准分割广泛应用于奶牛体况评分、姿态检测、行为分析及体尺测量等领域。受奶牛表面污渍和遮挡等因素的影响, 现有奶牛部位精准分割方法实用性较差。本研究在 YOLO v8n-seg 模型的基础上, 加入多尺度融合模块与双向跨尺度加权特征金字塔结构, 提出了 YOLO v8n-seg-FCA-BiFPN 奶牛身体部位分割模型。其中, 多尺度融合模块使模型更好地提取小目标几何特征信息, 双向跨尺度加权特征金字塔结构实现了更高层次的特征融合。首先在奶牛运动通道处采集奶牛侧面图像作为数据集, 为保证数据集质量, 采用结构相似性算法剔除相似图像, 共得到 1 452 幅图像。然后对目标奶牛的前肢、后肢、乳房、尾部、腹部、头部、颈部和躯干 8 个部位进行标注并输入模型训练。测试结果表明, 模型精确率为 96.6%, 召回率为 94.6%, 平均精度均值为 97.1%, 参数量为 3.3×10^6 , 检测速度为 6.2 f/s。各部位精确率在 90.3% ~ 98.2% 之间, 平均精度均值为 96.3%。与原始 YOLO v8n-seg 相比, YOLO v8n-seg-FCA-BiFPN 的精确率提高 3.2 个百分点, 召回率提高 2.6 个百分点, 平均精度均值提高 3.1 个百分点, 改进后的模型在参数量基本保持不变的情况下具有更强的鲁棒性。遮挡情况下该模型检测结果表明, 精确率为 93.8%, 召回率为 91.67%, 平均精度均值为 93.15%。结果表明, YOLO v8n-seg-FCA-BiFPN 网络可以准确、快速地实现奶牛身体部位精准分割。

关键词: 奶牛; 身体部位分割; 语义分割; FCABasicBlock; BiFPN; YOLO v8n

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298 (2024) 03-0282-08

OSID:



Segmentation Model of Cow Body Parts Based on YOLO v8n-seg-FCA-BiFPN

ZHANG Shujin^{1,2} XU Xingshi^{1,2} DENG Hongxing^{1,2} WEN Yuchen^{1,2} SONG Huaibo^{1,2}

(1. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

2. Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China)

Abstract: The fine segmentation of cow body parts has significant applications in research fields such as cow body condition scoring, posture estimation, behavior recognition, and body measurement. Due to the limited practicality of existing segmentation methods for different cow body parts, an improved YOLO v8n-seg model named YOLO v8n-seg-FCA-BiFPN was proposed for cow body part segmentation tasks. The improved model added FCA channel attention mechanism to the YOLO v8n backbone feature extraction network to better extract the geometric feature information of small targets, and used repeated weighted bidirectional features in the network feature fusion layer. The BiFPN was used to achieve the purpose of increasing the coupling of features at each scale. In order to validate the model performance, side-view images of cows at the channel were collected for network training. To ensure the quality of the dataset, the structural similarity algorithm was used to remove similar redundant images, resulting in a total of 1 452 images. LabelMe software was used to label the target cows, which were divided into eight parts, forelimbs, hindlimbs, udders, tails, belly, head, neck, and trunk, and was sent to the training model. The test results showed that the precision was 96.6%, the recall was 94.6% and the mean

收稿日期: 2023-07-17 修回日期: 2023-08-18

基金项目: 国家重点研发计划项目(2023YFD1301800)和国家自然科学基金项目(32272931)

作者简介: 张姝瑾(1999—), 女, 硕士生, 主要从事图像处理研究, E-mail: zhangshujin@nwfau.edu.cn

通信作者: 宋怀波(1980—), 男, 教授, 博士生导师, 主要从事图像处理和模型识别研究, E-mail: songyangfeifei@163.com

average precision was 97.1% , the parameters number was 3.3×10^6 , and the detection speed was 6.2 f/s. The precision of each part was from 90.3% to 98.2% , and the mean average precision was 96.3% . The YOLO v8n-seg-FCA-BiFPN network could realize accurate segmentation of various parts of dairy cows. Compared with the original YOLO v8n, the precision, recall and mean average precision of YOLO v8n-seg-FCA-BiFPN were 3.2 percentages points, 2.6 percentages points and 3.1 percentages points higher than that of YOLO v8n-seg, respectively. The precision under occlusion was 93.8% , the recall value was 91.67% , and the mean average precision was 93.15% . The volume of the improved model remained unchanged and had strong robustness. Under occlusion, the precision was 93.8% , the recall was 91.67% , and the mean average precision was 93.15% . The overall results showed that the research can provide necessary technical support for precise segmentation of dairy cows' body parts.

Key words: dairy cows; body part segmentation; semantic segmentation; FCABasicBlock; BiFPN; YOLO v8n

0 引言

随着我国智能化养殖程度不断提高,奶牛疫病监测、姿态检测、行为分析、体尺测量及体况评估等受到研究人员的重视,奶牛部位精准分割可为捕捉奶牛运动细节、分析奶牛行为奠定基础^[1-6]。受养殖场牛只遮挡、围栏遮挡及泥污干扰等因素影响,牛只部位精准分割困难,精度有待提高^[7-8]。实现复杂环境下奶牛部位精准分割对智能化奶牛养殖具有重要意义^[9]。

基于图像构建骨架图,利用骨架图实现身体各部位分割是研究精准分割问题的常用方法。宋怀波等^[10]基于部分亲和场的卷积神经网络(Convolutional neural networks, CNN),选用 VGG-19 作为前馈网络提取图像特征并通过最优匹配组合关键点和部分亲和场连接出奶牛骨架结构。不同关键点相似性为 0.75 时,骨架提取精确率为 93.40%,召回率为 94.20%。JIA 等^[11]提出了一种基于模板匹配的奶牛身体部位分割方法,将 RGB-D 图像应用到基于模板的端点分类模型中,提取奶牛骨架信息构建骨架图,实现了骨架相似性匹配从而区分每个骨架的所属类别,侧视图和后视图奶牛部位分割精确率与召回率分别为 96.00% 和 91.00%。

近年来,诸多学者开始利用深度学习方法解决目标各部位精准分割问题。CHEN 等^[12]构建了用于提取鸭只形状的 CNN 模型,实现了鸭只体质量的无接触测量,均方误差为 2.15%。CHEN 等^[13]提出了猪体多部位检测跟踪算法,通过 YOLACT 和 Zhangsu 算法计算关键点间的位置关系,实现了猪只不同部位的有效区分,精确率为 90.00%。HU 等^[14]提出了一种改进的 PointNet++ 点云分割模型,该模型可将猪只整体点云分割为头部、耳朵、躯干、四肢和尾巴,从而实现猪只的体型测量,与人工

测量结果相比,实际平均相对误差为 3.39%。赵凯旋等^[15]提出了一种基于帧差法联合调整 RGB 通道系数用决策树森林机器学习方法实现奶牛头部、颈部、躯干、前肢、后肢以及尾部的分割,该算法精确率为 86.29%。但该方法在光照突变等环境下的分割效果尚需改进。JIANG 等^[16]提出了精细分割奶牛部位的 FLYOLO v3 模型,该模型利用均值滤波算法将 LeakyReLU 函数应用到过滤层以解决图像在卷积后特征图产生粒子噪声的问题。此操作有效区分奶牛头部、背部和腿部,精确率为 99.18%,平均初检率为 94.00%。

综上所述,在奶牛身体部位精确分割方面主要存在以下问题亟待解决:①奶牛养殖场环境复杂,牛体之间互相遮挡、污垢干扰、围栏墙壁的遮挡均对后续精准提取奶牛各部位造成较大影响。②相比于躯干,奶牛乳房部位和尾部目标面积更小且易被遮挡,识别难度更高。为实现自然场景下奶牛部位精准分割,本研究利用 YOLO v8n(You only look once v8 nano)^[17-18]模型检测精度高、多尺度目标检测效果好和检测速度快的优点,提出一种 YOLO v8n-FCA-BiFPN 网络模型,以期实现复杂环境下奶牛身体部位的精准分割。

1 数据集制备与实验平台搭建

1.1 数据集制备

1.1.1 数据来源

本研究所用数据采集于 2022 年 12 月 21 日至 2023 年 3 月 10 日,采集地点为陕西杨凌科元克隆股份有限公司的规模奶牛养殖场。实验对象为泌乳期荷斯坦奶牛。数据采集过程如图 1 所示,图像获取设备为索尼 HDR-CX290E 型相机,分辨率为 3 984 像素(水平)×2 240 像素(垂直),图像保存为 JPEG 格式。设备采用支架固定在奶牛通道一侧,拍摄方向垂直于奶牛行进方向,摄像头距离地面 1 m,

距离栏杆 1.5 ~ 2 m,采集时间为 09:00—15:00,采集包含 1~4 头奶牛的侧面图像,比例约为 4:3:2:1,每幅图像中至少包含一头完整奶牛,各类数据均有不同程度遮挡。人工筛选 56 段视频片段,每段时长 5 ~ 20 s,帧率 25 f/s。

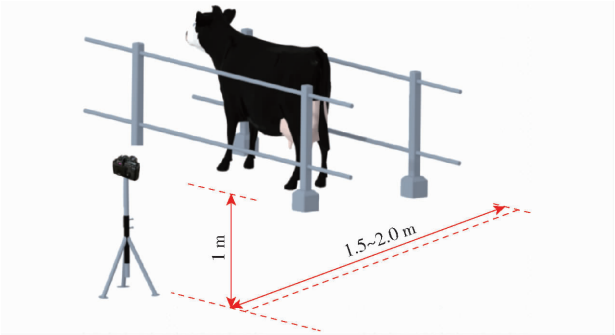


图 1 数据采集示意图
Fig. 1 Data acquisition diagram

1.1.2 数据集构建

图像分辨率统一调整为 1 920 像素(水平) × 1 080 像素(垂直)。为避免相邻帧图像相似度过高导致过拟合,采用结构相似性(Structural similarity,

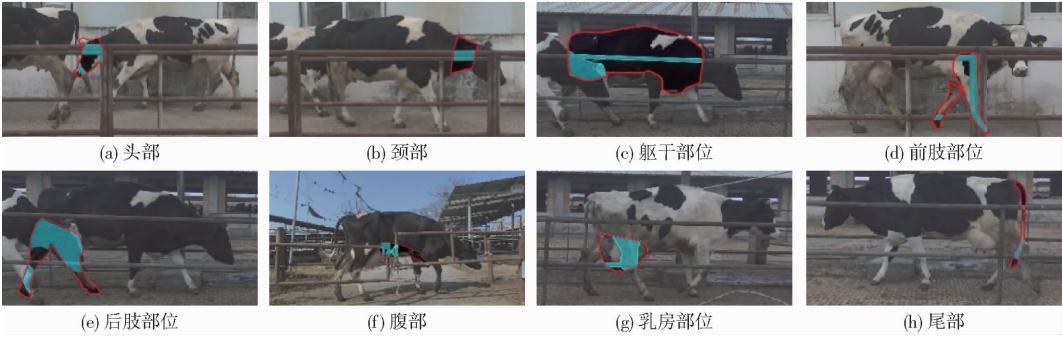


图 2 奶牛各部位遮挡示意图
Fig. 2 Shelter diagrams of cows' each part

表 1 数据集详细信息				
Tab.1 Details of data set				
奶牛部位	遮挡		未遮挡	
	训练集	测试集	训练集	测试集
前肢	293	98	796	265
后肢	328	110	760	253
乳房	506	169	583	194
尾部	331	110	759	252
腹部	421	140	669	222
头部	384	128	706	235
颈部	163	55	926	308
躯干	91	30	999	332

由表 1 可知,训练集和测试集中均包括各部位遮挡与未遮挡奶牛图像。且测试集与训练集在各部位遮挡与未遮挡图像数量比例基本保持在 3:1 左右。

1.1.3 奶牛身体区域标注

标注信息主要包括奶牛位置信息边界框和各分

SSIM) 算法^[19],剔除冗余图像,计算公式为

$$SSIM_{x,y} = \frac{(2\mu_x\mu_y + C_1)(2\delta_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\delta_x^2 + \delta_y^2 + C_2)} \quad (1)$$

式中 $SSIM_{x,y}$ ——图像 x,y 的相似度
 δ_x^2, δ_y^2 ——图像 x,y 方差
 μ_x, μ_y ——图像 x,y 均值
 δ_{xy} ——图像 x,y 协方差
 C_1, C_2 ——任意常数

剔除重复度过高冗余图像后最终获得 1 452 幅图像用于训练和测试,其中训练集图像 1 089 幅,测试集图像 363 幅,训练集与测试集比例为 3:1。

由于奶牛尾部具有平衡身体、驱赶蚊虫等作用,奶牛在运动过程中左右摆尾导致尾部更易被遮挡。与此同时,尾部形状变换较大,静止时呈细长条状占比面积较小,甩动时尾部四散且易被污泥污染,颜色与背景接近识别难度较大。如图 2 所示,按照遮挡部位可分为头部遮挡、颈部遮挡、躯干部位遮挡、前肢部位遮挡、后肢部位遮挡、腹部遮挡、乳房部位遮挡和尾部遮挡。数据集详细信息如表 1 所示。

割部位,分割部位包括头部、颈部、躯干、前肢、后肢、腹部、乳房和尾部。如图 3 所示,本研究使用 LabelMe 标注工具对原始图像进行多边形标注,标注时拐点尽可能接近奶牛实际轮廓,图像中出现的所有奶牛部位均进行标注。

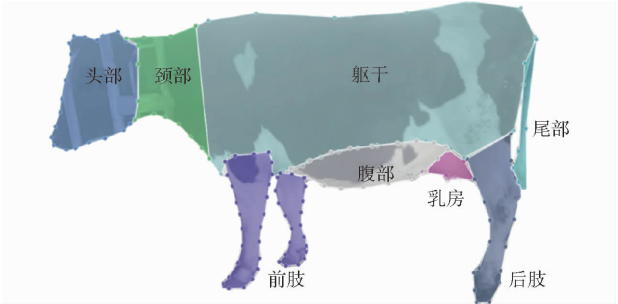


图 3 各部位轮廓示意图
Fig. 3 Outline diagram of each part

1.2 实验平台搭建

所有实验均在 Windows 11 操作系统下进行,处

理器型号为 AMD R7-5800H with Radeon Graphics 16CPU@3.20 GHz, 显卡型号为 NVIDIA GeForce RTX 3060。深度学习框架为 PyTorch 1.9, 编程环境为 PyCharm, 编程语言为 Python 3.7。

2 研究方法

2.1 YOLO v8 算法网络结构

YOLO v8 系列模型兼顾识别精度与速度, 能够高效处理现阶段目标检测算法。相较 YOLO v8 系列其他版本, Nano 版本可在保证精度的同时减少模型参数量与计算量, 如图 4 所示。

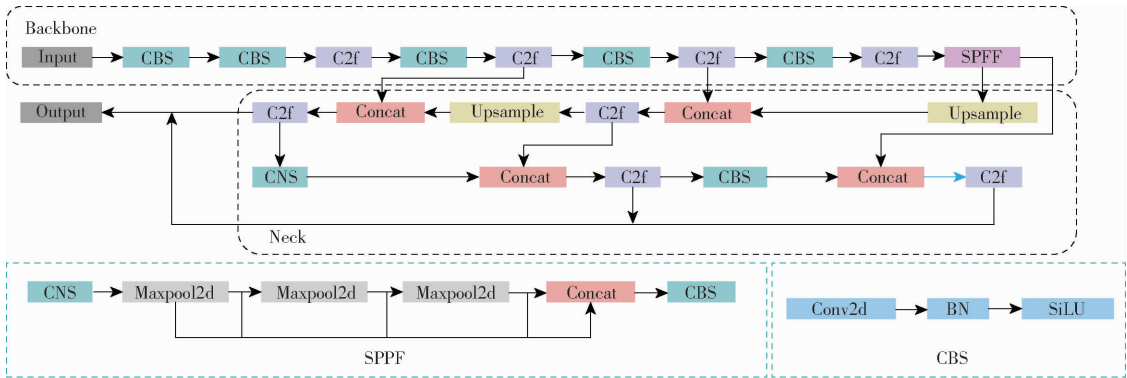


图 4 YOLO v8 网络结构图

Fig.4 Network structure of YOLO v8

(2)特征融合层(Neck):采用特征金字塔网络 (Feature pyramid network, FPN)^[22]与路径聚合网络 (Path aggregation network, PAN)^[23]相融合的结构, 采集更多的语义信息以及位置信息, 保证不同尺寸目标均能准确预测。YOLO v8n 采用 2 个上采样模块、多个 C2f 模块和解耦合头。YOLO 计算真实情况与预测回归框间的交并比 (Intersection over union, IoU), 非极大值抑制 (Non-maximum suppression, NMS) 计算每个回归框的置信度得分, 选择保留得分最高的边界框以确保精准度。

YOLO v8n-seg 主要解决目标分割问题, 其损失值函数 L 由多标签分类的交叉熵损失值 L_{cls} 、交并比 (Complete intersection over union, CIoU) 损失值 L_{box} 以及 Generalized Focal 损失值 L_{gfl} 组成。

2.2 YOLO v8n-seg 算法结构优化

在特征提取层部分首个 C2f 模块后增设频域通道注意力机制 FCABasicBlock, 同时输出特征层由原来 3 层增加至 4 层, 有助于特征融合层读取更丰富的语义信息; 对特征融合层进行优化, 采用双向跨尺度加权特征融合的方式读取特征信息以实现各尺度特征融合。

2.2.1 特征提取层优化

采用增加通道注意力模块以更好地解决深层网络特征图几何信息较少、浅层网络语义特征不

本研究使用的 YOLO v8n 模型主要由以下模块构成:

(1)特征提取层(Backbone):主干特征提取网络采用跨阶段局部网络^[20]方式, 通过分割梯度流使梯度流通过不同网络路径传播, 提高 CNN 的学习能力, 有效解决梯度消失等问题; 使用 C2f 模块在确保算法模型轻量级的同时获得更丰富的梯度流信息; 对空间金字塔池化结构 (Spatial pyramid pooling, SPP)^[21]进行了改进, 使用快速空间金字塔池化 (Spatial pyramid pooling-fast, SPFF) 结构提高模型推理速度。

足的问题, 具体改进如图 5 所示。在 Backbone 的第 1 个 C2f 模块后添加注意力机制 FCABasicBlock, 同时增加 1 层输出特征层。FCABasicBlock 以频域通道注意力网络 (Frequency channel attention, FCA)^[24]为基础, 融合了 Linear、ReLU 和 Sigmoid 模块。FCA 模块则基于压缩和激励网络 (Squeeze-and-excitation networks, SE)^[25]模块进行了改进, 其主要差异体现在通道压缩方法上。全局平均池化 (Global average pooling, GAP) 属于 2D 离散余弦变换的特殊情况, 通道注意力机制单独使用 GAP, 只能保留最低频率特征信息, 造成其他频率信息大量丢失。FCA 采用多频二维离散余弦变换对特征图进行处理, 输入被分为 n 份, 每份指定相应的 2D 离散余弦变换频率分量, 其结果采用串联方式进行拼接, 最后得到多频注意力机制框架。将该模块放入网络中, 即使出现冗余通道也能从中提取不同信息。在未增加计算量且没有使用额外参数的同时获得较为明显的性能增益。

2.2.2 特征融合层优化

网络 Neck 层对网络性能具有较大影响。如图 6a 所示, FPN 建立了具有横向连接自上而下的架构, 用于构建所有尺度的高级语义特征图。如图 6b 所示, PAN 在 FPN 的基础上增加了自下而上的融合

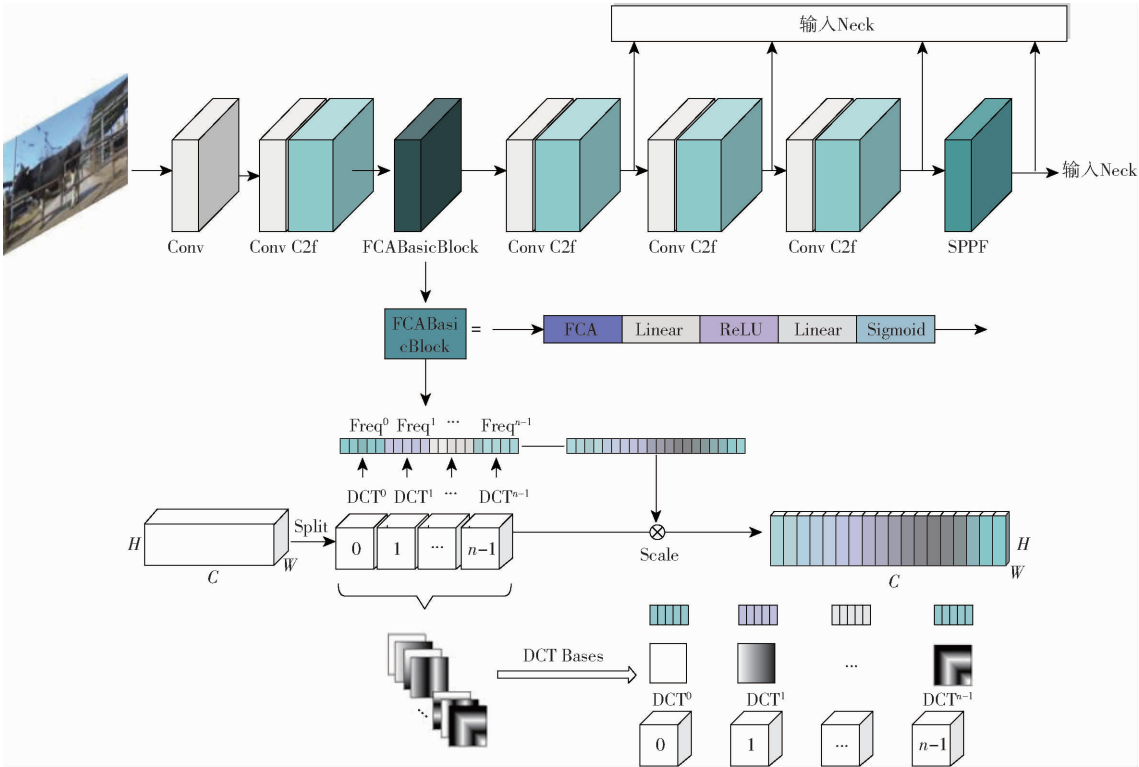


图 5 融合 FCA 的特征提取层优化原理图

Fig. 5 Backbone optimization schematic of FCA

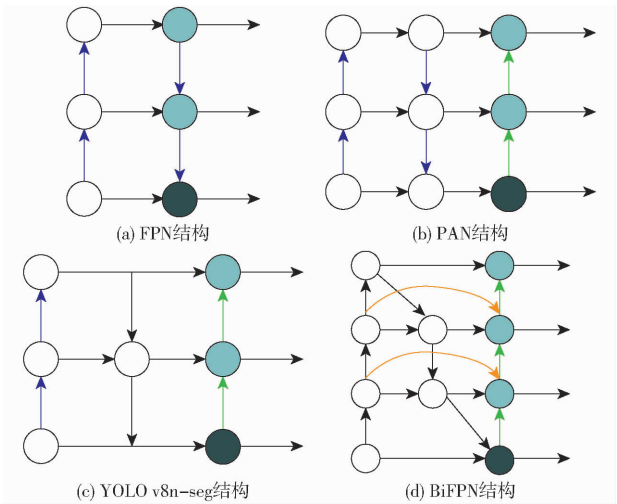


图 6 不同 Neck 网络结构对比

Fig. 6 Comparison diagrams of different Neck network structures

框架,使得特征层不受单侧信息流限制,同时拥有顶层语义信息和底层位置信息双层含义。如图 6c 所示,YOLO v8n-seg 网络结构融合 FPN 与 PAN 结构,但是目标信息主要集中在顶层语义信息,小尺寸目标位置信息被删除,导致小尺寸目标检测效果不佳。为了进一步提高目标检测的准确性,本研究使用 BiFPN^[26] 结构作为特征融合的方法。与以上 3 种基于临近特征融合的方式不同,图 6d 所示 BiFPN 采用双向跨尺度和加权特征融合的方式以实现更高层次的特征融合。

2.3 YOLO v8n-seg 参数设置与评价指标

设定模型输入尺寸 40 像素(水平)×640 像素(垂直),为加快训练速度,模型采用 COCO 数据集预训练权重,初始学习率设为 0.1,权重衰减系数为 0.000 5,共训练 250 轮次。

采用精确率 (Precision, P)、召回率 (Recall, R)、平均精度均值 (Mean average precision, mAP)、参数量 (Params)、浮点运算量 (Floating point operations, FLOPs) 共 5 个指标对模型进行评价,考虑实际应用需要,对检测速度进行评价。

3 结果与分析

3.1 训练结果

经过 250 轮次训练,YOLO v8n-seg-FCA-BiFPN 模型在测试集中 P 为 96.6%, R 为 94.6%,mAP 为 97.1%,参数量为 3.3×10^6 ,浮点运算量为 1.21×10^7 ,检测速度为 6.2 f/s。各部位检测具体结果如表 2 所示。结果表明,不同部位按照精确率由高到低排序依次是头部、前肢、腹部、躯干、颈部、乳房、后肢和尾部,头部精确率为 98.1%,尾部精确率为 90.3%;按照召回率由高到低排序依次是前肢、腹部、后肢、头部、躯干、颈部、乳房和尾部,前肢召回率为 98.7%,尾部召回率为 82.0%;mAP 最高为 99.4%,检测部位为前肢,最低为 87.9%,检测部位为尾部。

表 2 YOLO v8n-seg-FCA-BiFPN 模型训练结果

Tab.2 YOLO v8n-seg-FCA-BiFPN model training results

部位	精确率	召回率	平均精度均值	%
前肢	98.1	98.7	99.4	
后肢	95.3	96.5	98.0	
乳房	96.0	92.6	97.1	
尾部	90.3	82.0	87.9	
腹部	97.9	98.4	98.8	
头部	98.2	96.2	97.8	
颈部	97.2	95.9	98.6	
躯干	97.6	96.2	98.4	

3.2 检测效果

本研究对不同部位遮挡图像进行检测,测试集共包含不同场景下不同部位被遮挡的图像,检测框左上方显示检测类别与检测结果置信度,检测效果如图 7 所示,可以看出,该模型对牛只遮挡和栏杆遮挡等各种复杂场景下的奶牛部位分割均有较好的识别效果。

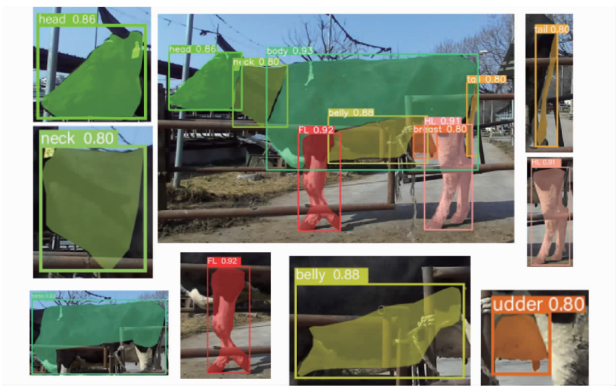


图 7 YOLO v8n-seg-FCA-BiFPN 检测效果
Fig.7 YOLO v8n-seg-FCA-BiFPN detection effect

3.3 不同分割算法性能比较

为评估模型对奶牛部位分割的检测效果,本研究在与 YOLO v8n-seg-FAC-BiFPN 模型相同的训练集以及测试集上对 YOLACT++^[27]、Mask R-CNN^[28]以及 YOLO v8n-seg 3 种模型进行了性能评估,各模型均训练 250 轮次,训练结果如表 3 所示。由表 3 可知,与改进前模型 YOLO v8n-seg

表 3 4 种实例分割模型的试验结果

Tab.3 Results of four instance segmentation models

模型	精确率/%	召回率/%	平均精度均值/%	参数量	浮点运算量
YOLACT++	79.8	81.1	83.5		
Mask R-CNN	83.1	85.6	86.4		
YOLO v8n-seg	93.4	92.0	94.0	3.1×10 ⁶	1.20×10 ⁷
YOLO v8n-seg-FAC-BiFPN	96.6	94.6	97.1	3.3×10 ⁶	1.21×10 ⁷

相比,YOLO v8n-seg-FAC-BiFPN 在参数量和浮点运算量基本保持不变的情况下,精确率提高 3.2 个百分点,召回率提高 2.6 个百分点,mAP 提升 3.1 个百分点。与其余两种模型相比,YOLO v8n-seg-FAC-BiFPN 的精确率分别高于 YOLACT++和 Mask R-CNN 模型 16.8、13.5 个百分点,召回率高于 YOLACT++和 Mask R-CNN 模型 13.5、9.0 个百分点,mAP 高于 YOLACT++和 Mask R-CNN 模型 13.6、10.7 个百分点。本模型在召回率与其他模型相近的情况下,表现出在精确率、mAP 等方面的优势,证明本研究模型切实可行,为研究遮挡情况下奶牛部位精准分割提供了技术支撑。

3.4 消融实验

3.4.1 不同注意力机制应用于 YOLO v8n-seg 性能比较

为评估注意力机制对 YOLO v8n-seg 网络性能提升效果,实验分别基于卷积注意力机制模块(Convolutional block attention module, CBAM)^[29]、SE 和 FCA 注意力机制对网络进行改进。各改进模型训练次数均为 250 轮次,训练结果如表 4 所示。结果表明,与 YOLO v8n-seg-CBAM 相比,YOLO v8n-seg-FCA 模型在参数量、浮点运算量与检测速度保持不变的前提下,精确率提升 0.1 个百分点,召回率提升 2.0 个百分点,mAP 提升 0.6 个百分点;与 YOLO v8n-seg-SE 相比,YOLO v8n-seg-FCA

表 4 3 种注意力机制性能

Tab.4 Performance of three attention mechanisms

模型	精确率/%	召回率/%	平均精度均值/%	参数量	浮点运算量	检测速度/(f·s ⁻¹)
YOLO v8n-seg-CBAM	95.4	92.7	96.5	3.2×10 ⁶	1.20×10 ⁷	6.7
YOLO v8n-seg-SE	95.4	95.0	97.0	3.2×10 ⁶	1.20×10 ⁷	6.7
YOLO v8n-seg-FCA	95.5	94.7	97.1	3.2×10 ⁶	1.20×10 ⁷	6.7

模型精确率和 mAP 均提高 0.1 个百分点,召回率降低 0.3 个百分点,参数量、浮点运算量与检测速度保持不变。由上述分析可知,FCA 注意力机制综合优势较为明显。

3.4.2 BiFPN 应用于 YOLO v8n-seg 性能比较

为更好地验证 YOLO v8n-FCA-BiFPN 网络性能,本文对比单独改动特征提取层的 YOLO v8n-seg-FCA 模型与单独改动特征融合层的

YOLO v8n-seg-BiFPN 模型,结果如表 5 所示。由表 5 可知,本研究模型相比 YOLO v8n-seg-FCA,精确率提升 1.1 个百分点,相比于 YOLO v8n-seg-BiFPN 模型,精确率、召回率和平均精度均值分别提高 0.8、0.5、0.3 个百分点,参数量、浮点运算量和检测速度基本不变。由上述分析可知,YOLO v8n-seg-FCA-BiFPN 模型改进方式更有助于部分遮挡情况下奶牛部位的精准分割。

表 5 YOLO v8n-FCA-BiFPN 网络性能
Tab.5 Performance of YOLO v8n-FCA-BiFPN

模型	精确率/%	召回率/%	平均精度均值/%	参数量	浮点运算量	检测速度/(f·s ⁻¹)
YOLO v8n-seg-FCA	95.5	94.7	97.1	3.2×10 ⁶	1.20×10 ⁷	6.7
YOLO v8n-seg-BiFPN	95.8	94.1	96.8	3.3×10 ⁶	1.21×10 ⁷	6.5
YOLO v8n-seg-FCA-BiFPN	96.6	94.6	97.1	3.3×10 ⁶	1.21×10 ⁷	6.2

4 讨论

4.1 不同部位遮挡情况检测效果

针对不同部位遮挡情况,本研究验证 YOLO v8n-seg-FCA-BiFPN 网络的识别效果,本节主要针对奶牛各部位进行遮挡测试。测试前根据遮挡位置不同,制作了不同部位遮挡的奶牛图像测试集,将测试集图像分别输入到网络中得到结果如表 6 所示。由表 6 可知,YOLO v8n-seg-

FCA-BiFPN 网络可以准确识别遮挡状况下各部位标签及轮廓信息,但各项指标略低于未遮挡情况。以前肢为例,无遮挡情况下精确率、召回率和平均精度均值分别为 98.1%、98.7% 和 99.4%,遮挡情况下分别为 96.6%、94.1% 和 98.9%,分别下降 1.5、4.6、0.5 个百分点。YOLO v8n-seg-FCA-BiFPN 网络对不同部位遮挡具有一定鲁棒性,该网络满足养殖场实际检测需求。

表 6 奶牛各部位遮挡情况识别效果
Tab.6 Impact of different parts of a dairy cow being obscured on recognition effectiveness %

评价指标	前肢	后肢	乳房	尾部	腹部	头部	颈部	躯干
精确率(无遮挡)	98.1	95.3	96.0	90.3	97.9	98.2	97.2	97.6
精确率(部分遮挡)	96.6	93.1	94.3	81.2	96.2	95.9	96.1	97.0
召回率(无遮挡)	98.7	96.5	92.6	82.0	98.4	96.2	95.9	96.2
召回率(部分遮挡)	94.1	92.1	91.3	71.3	97.3	95.2	96.0	96.0
平均精度均值(无遮挡)	99.4	98.4	97.1	87.9	98.8	97.8	98.6	98.4
平均精度均值(部分遮挡)	98.9	97.9	96.3	75.9	92.1	97.8	93.1	93.2

4.2 奶牛部位误检漏检分析

图 8 给出了部分奶牛部位误检漏检示例。图 8a 为模型存在无法正确分离牛体与背景的问题,图 8b 实验结果存在个别部位漏检现象。

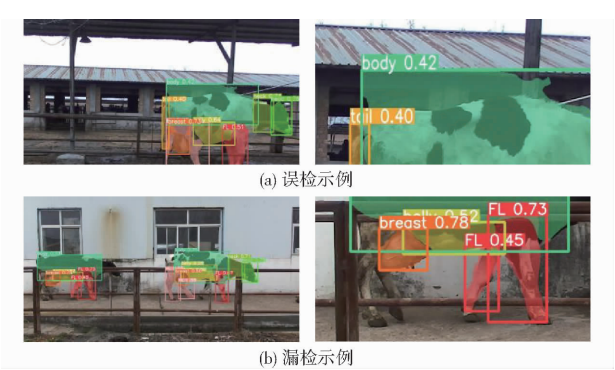


图 8 误检漏检示例

Fig.8 False detection and missed detection example display

出现误检漏检的原因可能是:①由 4.1 节可知,遮挡会对检测精度造成一定影响,随着遮挡面积的增加,精度也随之下降,发生误检漏检。②光照不足导致背景与牛体颜色相近,模型无法正确识别边缘信息,造成漏检误检现象。③由于清洁不到位导致奶牛体表附着泥污,检测过程中难以提取关键点信息,造成个别部位误检漏检。

4.3 不同部位检测效果分析

奶牛运动过程中的姿态变化同样可能导致漏检。如图 9 所示,奶牛尾巴细长且运动灵活,检测过程中常出现遮挡,易发生漏检,影响检测效果;奶牛趴卧过程中前肢、后肢部位常与地面接触,表面易被泥污污染从而影响检测效果;走动过程中后腿遮挡等问题同样会对奶牛乳房部位精准定位造成一定阻力,发生乳房漏检。

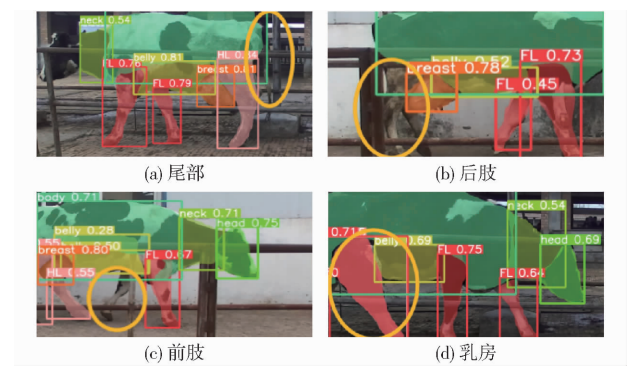


图9 各部位漏检示例

Fig.9 Missing examples of each part

5 结论

(1)将 YOLO v8-seg 网络与 FCABasic-Block

注意力机制、BiFPN 结构相融合,提出 YOLO v8n-seg-FCA-BiFPN 网络改进模型。改进后网络模型精确率为 96.6%,召回率为 94.6%,平均精度均值为 97.1%,参数量为 3.3×10^6 ,浮点运算量为 1.21×10^7 ,检测速度为 6.2 f/s。精确率、召回率和平均精度均值提升明显,证明该操作可有效提升模型性能。

(2)针对荷斯坦奶牛不同部位的遮挡进行了讨论。实验结果表明,在遮挡状况下该模型仍可较好地识别各部位轮廓信息与标签,具有一定的鲁棒性。

(3)本模型能够在不改变模型参数量和浮点运算量的前提下明显增加目标检测精度,检测速度略有提升,更有利于模型实现在遮挡情况下奶牛身体部位的精准分割。

参 考 文 献

[1] ZUCALI M,BAVA L,TAMBURINI A,et al. Effects of season, milking routine and cow cleanliness on bacterial and somatic cell counts of bulk tank milk[J]. Journal of Dairy Research, 2011,78(4):436-441.

[2] ELMOSLEMANY A M, KEEFE G P, DOHOO I R, et al. Risk factors for bacteriological quality of bulk tank milk in Prince Edward Island dairy herds. Part 1: overall risk factors[J]. Journal of Dairy Science, 2009, 92(6):2644-2652.

[3] YUAN K,LIU Z P,WANG Z Y. Effective application of dairy cow feeding behavior in nutritional management[J]. Journal of Dairy Science and Technology,2008,5:237-242.

[4] 米法英,王翠芳,敖长金,等. 动物福利对奶牛生产性能的影响[J]. 中国奶牛,2016(9):4-8.
MI Faying, WANG Cuifang, AO Changjin, et al. Effect of animal welfare on the performance in dairy cows[J]. China Dairy Cattle,2016(9):4-8. (in Chinese)

[5] ANA B, RAMIRO F. Associations between on-farm welfare, milk production, and reproductive performance in dairy herds in northwestern Spain[J]. Journal of Applied Animal Welfare Science,2021,1(24):29-38.

[6] VALENTIN S, LANCELOT R, ROCHE M. Identifying associations between epidemiological entities in news data for animal disease surveillance[J]. Artificial Intelligence in Agriculture,2021,5:163-174.

[7] HOSSAIN M E, KABIR A, ZHENG L,et al. A systematic review of machine learning techniques for cattle identification: datasets, methods and future directions[J]. Artificial Intelligence in Agriculture,2022,6:138-155.

[8] 吴宇峰,李一鸣,赵远洋,等. 基于计算机视觉的奶牛体况评分研究综述[J]. 农业机械学报,2021,52(增刊):268-275.
WU Yufeng, LI Yiming, ZHAO Yuanyang, et al. Review of research on body condition score for dairy cows based on computer vision[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(Supp.):268-275. (in Chinese)

[9] 司永胜,王朝阳,张艳,等. 基于深度与传统特征融合的非限制条件下奶牛个体识[J]. 农业机械学报,2023,54(6):272-279.
SI Yongsheng, WANG Zhaoyang, ZHANG Yan,et al. Individual identification of dairy cows under unrestricted conditions based on fusion of deep and traditional features[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 272-279. (in Chinese)

[10] 宋怀波,李振宇,吕帅朝,等. 基于部分亲和场的行走奶牛骨架提取模型[J]. 农业机械学报, 2020, 51(8): 203-213.
SONG Huaibo, LI Zhenyu, LÜ Shuaichao, et al. Skeleton extraction model of walking daily cows based on partial affinity field [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 203-213. (in Chinese)

[11] JIA N, KOOTSTRA G, KOERKAMP P G, et al. Segmentation of body parts of cows in RGB-depth images based on template matching[J]. Computers and Electronics in Agriculture, 2021, 180(4): 105897.

[12] CHEN R, ZHAO Y, YANG Y, et al. Online estimating weight of white Pekin duck carcass by computer vision[J]. Poult Science, 2023, 102(2): 102348.

[13] CHEN F, LIANG X, CHEN L, et al. Novel method for real-time detection and tracking of pig body and its different parts[J]. International Journal of Agricultural and Biological Engineering, 2020, 13(6): 144-149.

[14] HU H, YU J, YIN L, et al. An improved PointNet++ point cloud segmentation model applied to automatic measurement method of pig body size[J]. Computers and Electronics in Agriculture, 2023, 205: 107560.

[15] 赵凯旋,李国强,何东健. 基于机器学习的奶牛深度图像身体区域精细分割方法[J]. 农业机械学报,2017,48(4): 173-179.
ZHAO Kaixuan, LI Guoqiang, HE Dongjian. Fine segment method of cows' body parts in depth images based on machine learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(4): 173-179. (in Chinese)

imaging for detection of bruises in apples[J]. Computers and Electronics in Agriculture, 2016, 127: 652 – 658.

[29] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 3431 – 3440.

[30] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C] // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Proceedings, Part III 18. Springer International Publishing, 2015: 234 – 241.

[31] WANG J, SUN K, CHENG T, et al. Deep high-resolution representation learning for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(10): 3349 – 3364.

[32] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2881 – 2890.

[33] 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 (ECCV), 2018: 801 – 818.

[34] 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.

[35] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481 – 2495.

~~~~~

(上接第 289 页)

[16] JIANG B, WU Q, YIN X, et al. FLYOLO v3 deep learning for key parts of dairy cow body detection[J]. Computers and Electronics in Agriculture, 2019, 166: 104982.

[17] JOCHER G. YOLO v5 release v6[EB/OL]. 2021 – 10 – 12[2023 – 07 – 06]. <https://github.com/ultralytics/YOLOv5/releases>.

[18] JOCHER G. Ultralytics YOLOv8[EB/OL]. 2023 – 01 – 11[2023 – 07 – 06]. <https://github.com/ultralytics/ultralytics>.

[19] WANG Z, PAN W, CUPPENS-BOULAHIA N, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600 – 612.

[20] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

[21] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9): 1904 – 1916.

[22] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[23] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[24] QIN Z, ZHANG P, WU F, et al. FcaNet: frequency channel attention networks[C] // International Conference on Computer Vision, 2020: 783 – 792.

[25] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation networks[C] // 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

[26] TAN M, PANG R, LE V Q, et al. EfficientDet: scalable and efficient object detection[C] // 2020 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

[27] BOLYA D, ZHOU C, XIAO F, et al. YOLACT + + better real-time instance segmentation[C] // 2022 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2022.

[28] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R – CNN [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 42(2): 386 – 397.

[29] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[C] // 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.