

基于改进 YOLO v8n - pose 的奶山羊体尺关键点检测方法

李书琴 王井超

(西北农林科技大学信息工程学院, 陕西杨凌 712100)

摘要: 针对奶山羊姿态多变导致边缘轮廓特征提取困难、体尺关键点误检漏检等问题, 本文提出一种基于改进 YOLO v8n - pose 的奶山羊体尺关键点检测方法。在主干网络构建 C2f_Ghost 模块, 保持模型轻量化的同时增强特征提取能力; 引入双域选择机制 (DSM) 提高模型对关键点所在边缘信息的关注, 增强对边缘轮廓特征提取能力; 引入全局注意力机制 (GAM) 获取全局特征, 增强对关键点之间特征的提取。试验结果表明, 改进模型 mAP_{0.5} 为 94.5%, 相较 YOLO v8n - pose 提升 3.3 个百分点, 并降低参数量和浮点运算量。与主流模型 HRNet、Lite - HRNet、YOLO v5 - pose、YOLO v7 - W6 - pose、YOLO v8s - pose、YOLO - NAS - pose - N 和 YOLO 11n - pose 相比, mAP_{0.5} 分别提升 4.3、5.0、5.9、3.7、3.2、3.0、3.5 个百分点。奶山羊体高、十字部高、体长、胸宽、腹宽和胸围平均绝对误差分别为 3.43、3.21、3.27、1.33、1.34、5.64 cm; 平均绝对百分比误差分别为 4.72%、4.36%、4.26%、6.21%、4.84% 和 5.83%, 经坐标系矫正后胸围误差可降低至 4.63%。本文方法能有效提升体尺关键点检测精度, 为奶山羊体尺关键点检测和体尺获取提供技术支持。

关键词: 奶山羊; 关键点检测; YOLO v8n - pose; 羊只体尺测量

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2026)14-0068-12

OSID:



Body Size Key Point Detection Method of Dairy Goats Based on Improved YOLO v8n - pose

Li Shuqin Wang Jingchao

(College of Information Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: Aiming at the challenges posed by the highly variable postures of dairy goats, which result in difficulty in extracting edge contour features and frequent misdetection or omission of body size keypoints, an enhanced keypoint detection method was proposed based on an improved YOLO v8n - pose architecture. Firstly, a C2f_Ghost module was integrated into the backbone network to enhance the model's feature extraction capability while maintaining its lightweight nature, making it more suitable for real-time applications. Secondly, a dual-domain selection mechanism (DSM) was designed to emphasize edge-related information in both spatial and frequency domains, thereby enhancing the model's sensitivity to boundary features and improving the localization of keypoints near complex contours. Thirdly, a global attention mechanism (GAM) was employed to effectively capture long-range dependencies and global context, which strengthened the interaction and structural consistency among keypoints. Experimental results showed that the improved model achieved a mAP_{0.5} of 94.5%, outperforming YOLO v8n - pose by 3.3 percentage points while reducing parameters and computation. Compared with mainstream models, HRNet, Lite - HRNet, YOLO v5 - pose, YOLO v7 - W6 - pose, YOLO v8s - pose, YOLO - NAS - pose - N, and YOLO 11n - pose, mAP_{0.5} was improved by 4.3, 5.0, 5.9, 3.7, 3.2, 3.0, and 3.5 percentage points, respectively. The mean absolute errors for body height, withers height, body length, chest width, abdominal width, and chest girth were 3.43 cm, 3.21 cm, 3.27 cm, 1.33 cm, 1.34 cm, and 5.64 cm, with mean absolute percentage errors of 4.72%, 4.36%, 4.26%, 6.21%, 4.84%, and 5.83%, respectively. After coordinate correction, chest girth error was further reduced to 4.63%. These results demonstrated the effectiveness of the proposed method in improving keypoint detection accuracy which can provide technical support for dairy goat body measurement.

Key words: dairy goats; key point detection; YOLO v8n - pose; sheep body size measurement

0 引言

羊养殖业是中国畜牧业的重要组成部分,奶山羊作为关键分支,智能化养殖至关重要。其中,体尺参数能反映家畜生长状态和健康水平^[1]。传统方法依赖人工接触式测量,耗费大量人力物力,还因应激导致测量准确性较低。因此,研发无接触式体尺测量技术成为奶山羊智能养殖关注任务之一^[2]。

目前,家畜体尺关键点检测方法主要分为基于二维图像和基于三维点云 2 类方法^[3]。基于二维图像主要依靠单视角或多视角图像获取家畜边缘轮廓信息,利用图像分割方法将家畜从背景中分离,使用边缘检测算法提取边缘轮廓,再通过计算曲率定位关键点位置^[4]。赵建敏等^[5]基于 YOLO v5 检测图像中牛只体尺特征部位,结合边缘检测算法获取体尺测量关键点。关键点位于局部特征中,采用图像局部方法可以增强关键点定位准确性^[6]。陆明洲等^[7]基于双视角图像,获取羊体二值图并定位山羊体尺关键点。

基于三维点云方法主要通过多视角深度相机获取点云数据,重建家畜三维点云模型^[8],然后在点云中提取家畜表型特征并定位关键点^[9],但这种方法对点云模型有较高要求,通常需要点云补全和去噪处理^[10],家畜姿态也会影响三维点云重建结果^[11]。为提高关键点检测精度,常通过点云分割将家畜点云划分为头部、胸部、腹部等区域,并在局部区域内提取特征^[12]。Jin 等^[13]基于 PointStack 将羊只点云分割并定位关键点,实现了 5% 以内的体尺误差。赵健等^[14]基于 SGPointNet + 改进模型对奶牛点云分割,实现体尺计算平均误差均低于 4%。

然而,与 RGB 图像相比,家畜点云上的关键点检测过程复杂,且技术不如图像关键点检测成熟。Nguyen 等^[15]通过检测 RGB 图像关键点,利用相机内部参数计算空间坐标,结合点云实现关键点定位与体尺测量。张淑瑾等^[16]基于 YOLO v8n - seg 改进图像分割算法,将奶牛分割为 8 个部位,平均精度均值为 97.1%。翁智等^[17]基于肉牛图像的分割,获取关键点图像坐标并结合深度值转换为三维坐标,体尺平均相对误差低于 7%。Yang 等^[18]基于卷积神经网络跨尺度融合优化关键点检测,关键点平均精度均值为 92.8%,并将关键点像素坐标转换到三维空间,体尺平均相对误差低于 7%。Singh 等^[19]通过获取奶牛的 RGB 图像和深度信息,分割奶牛身体部位并提取关键点像素位置,获取三维坐标数据计算体尺,平均绝对百分比误差低于 6%。Li 等^[20]基于改进 YOLO - pose 实

现蒙古马的体尺关键点检测,并结合点云测量体尺,实现关键点平均精度均值为 92.5%,体尺平均绝对百分比误差降低到 3% 以下。

目前,虽然基于机器视觉的家畜关键点检测和体尺获取取得一定进展,但还存在边缘轮廓特征提取困难、体尺关键点定位偏移和误检、非线性体尺参数获取困难等问题^[21]。针对以上问题,本文提出一种基于 YOLO v8n - pose 改进的关键点检测模型,通过增强边缘信息选择及关键点特征提取,实现羊只体尺关键点快速准确检测,并结合相机参数和深度图转换为三维坐标映射到点云,以期实现高效精准的羊只体尺参数获取。

1 材料与数据

1.1 数据采集

奶山羊视频于西北农林科技大学畜牧教学试验基地采集,羊只月龄范围为 5 ~ 27 月,采集设备为 Azure Kinect 型深度相机与便携式计算机(惠普, Windows 11),每只奶山羊采集视频时间约为 10 s。使用 3 台相机从奶山羊两侧和顶部视角采集,两侧相机距离奶山羊 90 ~ 100 cm,顶部相机距离为 80 ~ 90 cm。提取视频中的图像帧并进行筛选,选取 2 000 幅羊只侧视图像,1 000 幅俯视图像,训练集、测试集和验证集按照 7:1:2 划分。图 1 为 3 个视角下采集的奶山羊图像。



图 1 羊只图像数据集

Fig. 1 Sheep images dataset

1.2 关键点选取及标注

使用标注软件 Labelme 对图像进行标注,标注羊只目标框和关键点位置信息,侧视视角标注 6 个体尺关键点,俯视视角标注 4 个体尺关键点。图 2 为奶山羊关键点标注及体尺测量示意图。

2 奶山羊体尺关键点检测方法

YOLO - pose^[22]是基于 YOLO v5 框架提出的关键点检测算法,YOLO v8n - pose 是基于 YOLO v8 结合目标检测和关键点位置估计的模型,因其较小的计算资源和内存占用量需求更适用于养殖场景^[23]。但由于家畜关键点位置分布和特征与人体关键点存在差异导致出现大量错检漏检。

针对奶山羊体尺关键点检测任务,对 YOLO v8n - pose 模型改进,提出了一种关键点检测模型 YOLO

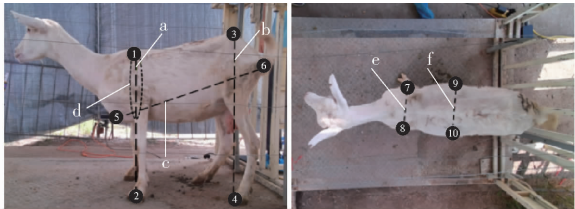


图2 奶山羊体尺关键点图示

Fig.2 Diagram of key points of dairy goat body size
a. 体高 b. 十字部高 c. 体斜长 d. 胸围 e. 胸宽 f. 腹宽
1. 鬃甲顶点 2、4. 前蹄后蹄与地面交点 3. 臀部最高点 5. 胸骨节点 6. 坐骨结节 7、8. 肩胛后角左右点 9、10. 腹部最宽点

v8n-SK,采用 GhostNet 中的 Ghost 模块改进原有的 C2f 模块,减少通道信息的冗余并增强模型特征提取能力。其次,融合双域选择机制来使模型重点关注关键点所在边缘轮廓区域,增强对边缘轮廓特征提取能力。最后,在颈部网络中引入全局注意力机制,通过放大全局交互式表示来提高网络的性能,更好关注全局特征,提取关键点对之间的特征关系。改进的网络模型如图 3 所示。整体信息流动可表示为以下过程,对于输入图像 $X \in \mathbf{R}^{H \times W \times 3}$,首先经过一系列卷积与 C2f_Ghost 模块,获得中间特征 F ,随后送入到 DSM,其中 SSM 模块提取退化区域的空间分布,生成空间选择后的特征 F_s ,FSM 则对 F_s 进行高频增强,剔除冗余低频信息,获得增强后的特征图 $\hat{F}$,随后送入 SPPF 模块融合不同尺度的上下文信息,

在 Neck 阶段,通过上采样并与主干网络中提取的浅层特征进行拼接融合,通过 C2f_Ghost 模块进一步调整提炼,得到 3 个不同尺度的融合特征,即

$$F_n^{(i)} = \text{C2f_Ghost}(\text{Concat}(\text{Upsample}(F_{\text{sppf}}), F^{(i)})) \tag{1}$$

式中 $F_n^{(i)}$ ——融合特征
 F_{sppf} ——快速空间金字塔特征
 $F^{(i)}$ ——浅层特征

融合后的不同尺度特征图被送入 GAM,通过联合通道注意与空间注意实现全局建模,最后通过检测头输出相应的目标框及关键点检测结果。

2.1 C2f_Ghost 模块

羊只实际养殖环境不适合部署高算力设备,对模型轻量化有一定要求,为快速检测羊只体尺关键点,需要降低模型复杂度。YOLO v8 使用的 C2f 模块存在额外的瓶颈结构从而带来多余的卷积,导致通道信息的冗余和额外的特征提取。GhostNet^[24]是一种轻量级深度学习模型,通过 Ghost Module 和 Ghost BottleNeck 实现高效特征提取。

Ghost 模块流程分为 3 步,如图 4 所示,首先,使用较少的卷积核生成一组内在特征图,并舍弃了普通卷积中的偏置项;之后通过廉价的线性操作将所得内在特征图扩展为更多的 Ghost 特征图,计算式为

$$y_{i,j} = \varphi_{i,j}(y_i) \tag{2}$$

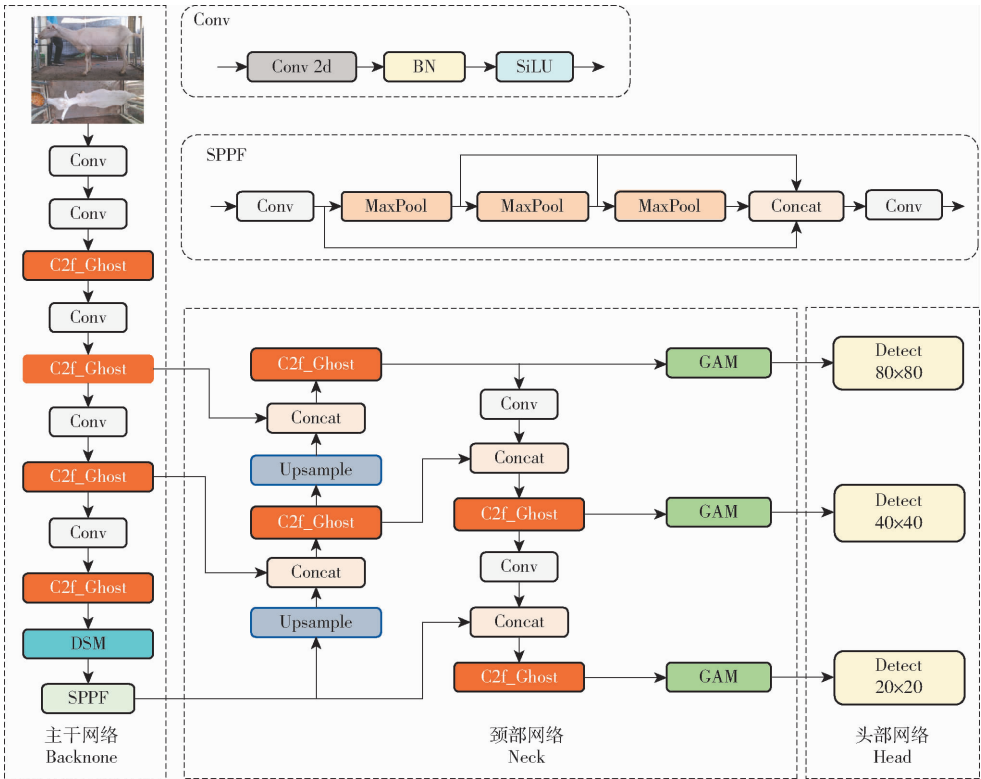


图3 YOLO v8n-SK 网络结构

Fig.3 YOLO v8n-SK network architecture

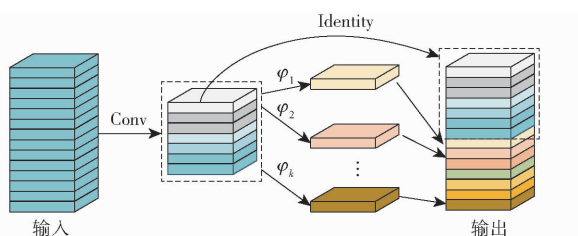


图 4 Ghost 模块示意图

Fig. 4 Ghost module diagram

式中 y_i ——第 i 个内在特征图

$\varphi_{i,j}$ ——线性运算

$y_{i,j}$ ——Ghost 特征图

最后通过将内在特征图和 Ghost 特征图进行拼接,生成和标准卷积相同数量的特征图。

将原 C2f 模块中所有瓶颈结构替换为 Ghost BottleNeck,瓶颈结构中的卷积模块使用 Ghost 模块,得到 C2f_Ghost 模块,可以有效减少特征冗余,避免低质量重复特征。C2f_Ghost 结构如图 5a 所示。

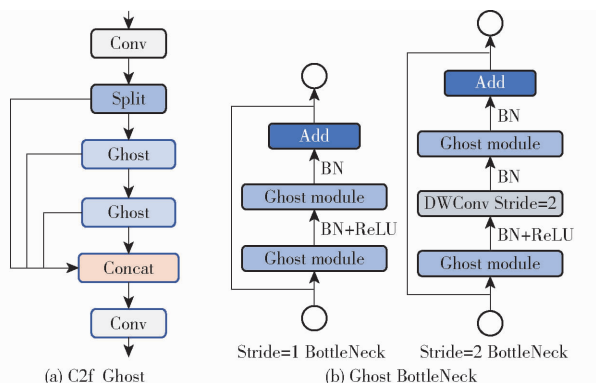


图 5 C2f_Ghost 模块及 Ghost BottleNeck 结构

Fig. 5 Structure of C2f_Ghost module and BottleNeck

Ghost BottleNeck 有 2 个种类,如图 5b 所示,第 1 个瓶颈结构步长为 1,不对输入特征图进行高和宽的压缩,仅调整通道维度。首先用 1 个 Ghost module 进行特征提取,再用 1 个 Ghost module 进一步提炼特征,从而加深网络深度并增强表达能力。当需要对特征层进行空间维度下采样时,采用第 2 个 Ghost BottleNeck,使用步长为 2 的深度卷积对输入特征图的高和宽进行压缩,实现尺度调整。在残差分支中,将输入特征与主干路径输出进行元素级相加,保证信息流的传递与梯度的稳定更新。

2.2 双域选择机制

奶山羊体尺关键点通常位于羊只的边缘轮廓区域,因此边缘轮廓成为体尺关键点检测任务中需要重点关注的特征区域,但由于姿态的变化边缘特征易受到干扰导致提取困难。然而,YOLO v8n-pose 模型缺少对边缘和复杂区域的有效关注。FocalNet^[25]旨在高效地捕获图像中不同区域的多尺度信息,更加关注边缘位置和困难区域等关键信息,

提升对重点关注区域特征提取能力;双域选择机制 (Dual-domain selection mechanism, DSM) 是 FocalNet 中的信息选择与增强机制,通过在空间域和通道域双域内态筛选和聚焦信息,实现对高价值特征的保留与强化。

DSM 包括 2 个组件:空间选择模块 (Spatial selection module, SSM) 和频率选择模块 (Frequency selection module, FSM)。DSM 模型结构如图 6 所示,其中 F 表示特征图, h 表示不同网络层级, s 表示不同分辨率, SSM 将 F 作为输入,通过深度卷积层学习得到特征退化可能发生的位置分布。然后 FSM 通过从特征中去除低频信号放大高频信号或难以处理的区域,实现对复杂边缘信息的增强。

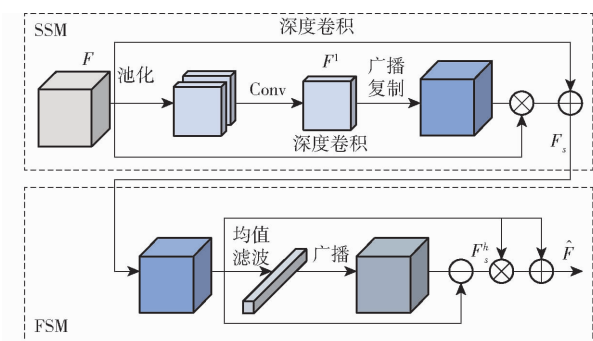


图 6 DSM 模块结构

Fig. 6 Structure of DSM module

SSM 有助于网络在空间域中关注重要区域,为 FSM 提供信息退化的初始位置。SSM 由 3 个分支组成,其中主分支基于 CBAM 设计,通过聚焦于特征退化的空间位置生成全局特征表示。对输入特征图 F ,首先在通道维度上分别进行最大池化和平均池化,通过卷积层得到通用空间特征图,即

$$F^1 = \text{Conv}_3([\text{AvgPool}(F), \text{MaxPool}(F)]) \quad (3)$$

式中 Conv_3 —— 3×3 卷积

考虑到不同通道的退化模式差异,SSM 通过深度卷积对输入特征图 F 进行通道分离变换,并使用结果 F^1 广播复制得到原有通道维度特征图,并通过元素相乘和相加生成特征 F_s ,即

$$F_s = \text{DConv}_{5,7}(F) \otimes T(F^1, C) + \text{DConv}_3(F) \quad (4)$$

式中 $\text{DConv}_{5,7}$ —— 5×5 和 7×7 卷积核尺寸级联深度卷积层

DConv_3 —— 3×3 卷积核尺寸深度卷积

$\otimes$ ——逐元素乘法

$T(F^1, C)$ ——沿通道复制 F^1 特征 C 次函数

将空间选择的特征 F_s 输入到 FSM 中进行频率选择,FSM 通过去除低频成分,进一步强化羊只边缘与背景信息之间高频上的关键差异。首先,应用

均值滤波器生成低频特征,然后从输入中减去生成的低频信号,得到互补的高频特征,即

$$F_s^h = F_s - \text{Mean}(F_s) \tag{5}$$

此过程通过通道维度的全局平均池化实现,并通过广播操作恢复到原有通道维度。最后,输出通道特征 F_s^h 和 F_s 之间的乘法和残差连接生成增强后特征 $\hat{F}$,即

$$\hat{F} = F_s^h \otimes F_s + F_s \tag{6}$$

经过 DSM 模块处理后,羊只边缘重要区域得到强化,进一步提升了模型对边缘特征的提取能力。

2.3 全局注意力机制

由于羊只体尺关键点分布在边缘轮廓上,检测过程易出现误检和偏移,并且体尺信息是根据关键点来测量的,全局注意力机制^[26] (Global attention mechanism, GAM)可以更好地关注关键点特征信息之间的联系。因此,在 YOLO v8n - pose 上集成了

GAM,通过在 3 个尺度检测头前引入 GAM 增强对边缘关键点特征信息提取,捕获各关键点之间的依赖关系。

卷积注意力机制模块 (Convolutional block attention module, CBAM)^[27] 是一种结合了空间和通道的注意力机制模块。基于 CBAM 架构, GAM 通过整合序列化的通道与空间注意力策略,对原有结构进行优化。结构如图 7 所示,分为通道注意力子模块和空间注意力子模块。GAM 模块计算过程为

$$F_2 = M_c(F_1) \otimes F_1 \tag{7}$$

$$F_3 = M_s(F_2) \otimes F_2 \tag{8}$$

式中 M_c ——通道注意力加权结果

M_s ——空间注意力输出结果

如式(7)、(8)所示,对于输入特征 F_1 ,经过通道注意力模块获取中间加权特征 F_2 ,再经过空间注意力模块获得输出特征 F_3 。

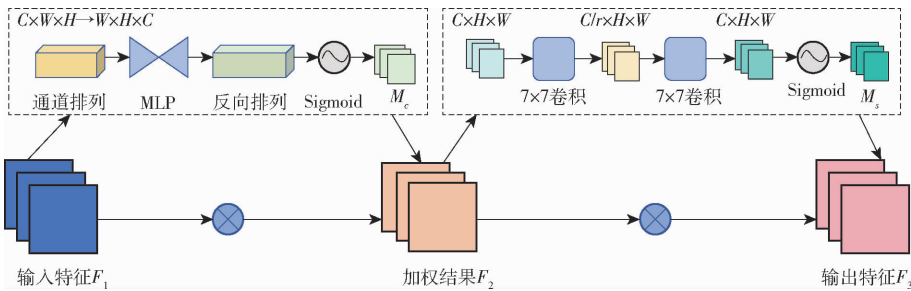


图 7 GAM 模块结构图

Fig. 7 GAM module structure diagram

通道注意力子模块首先对输入特征图进行维度转换,采用三维排列方式以保留空间、通道和批次维度的信息,转换后的特征图被输入到多层感知器 (MLP),来增强跨维度的通道-空间依赖性。随后,特征图被转换回原始维度,通过 Sigmoid 激活函数进行输出,生成通道注意力图,捕捉通道间的全局关联性,有效提取关键特征。空间注意力子模块主要通过卷积操作进行特征处理,首先,使用 1 个 7×7 卷积核对输入特征图进行卷积操作,以减少通道数提取重要的空间特征。然后使用另 1 个 7×7 卷积核对减少后的特征图进行卷积操作,恢复并增强空间信息。最后,经过 Sigmoid 激活函数处理,输出空间注意力图,表示每个空间位置的重要性,使模型自适应地关注图像中的关键区域。

3 试验与结果

3.1 试验环境与参数设置

试验环境及参数设置如表 1 所示。在训练过程中,所有输入图像分辨率统一调整为 640 像素 $\times$ 640 像素,批次大小为 16,共进行 200 轮迭代。

表 1 试验环境与参数设置

Tab. 1 Experimental environment

参数	数值
GPU	NVIDIA GeForce A40
CPU	Intel(R) Xeon(R) Gold 5318Y
操作系统	Ubuntu 16. 04
深度学习框架	Pytorch 1. 12
编程语言	Python 3. 8
运算结构	Cuda 11. 3
初始学习率	0. 01
动量系数	0. 937
权重衰减系数	0. 000 5

3.2 评价指标

3.2.1 关键点检测评价指标

关键点相似度是评估关键点检测的常用指标,采用平均精度均值 (mean Average Precision, mAP_{0.5})、参数量、浮点运算量和推理速度作为评价指标。其中 mAP_{0.5} 表示 OKS 阈值为 0.5 时的检测精度。

3.2.2 体尺参数测量评价指标

体尺参数是衡量关键点检测效果的重要信息,采用平均绝对误差 (Mean absolute error, MAE) 和平

均绝对百分比误差 (Mean absolute percentage error, MAPE) 评估羊只体尺测量结果。其中 MAE 表示测量值和真实值之间的绝对误差, MAPE 表示测量值和真实值之间差值占真实值的百分比。

3.3 结果与分析

3.3.1 模型性能试验

为验证改进的模型在奶山羊体尺关键点检测中的有效性, 针对奶山羊数据集进行试验与对比, 分别测试奶山羊侧视和俯视 2 个视角。选取当前主流的关键点检测模型进行试验, 包括 HRNet^[28]、Lite - HRNet^[29]、YOLO v5n - pose、YOLO v7 - W6 - pose^[30]、YOLO v8s - pose、YOLO v8n - pose、YOLO - NAS - pose - N 和 YOLO 11n - pose^[31], 试验结果如表 2 所示。由表 2 可知, YOLO v8n - SK 相较于 HRNet、Lite - HRNet、YOLO v5n - pose、YOLO v7 - W6 - pose、YOLO v8s - pose、YOLO v8n - pose、YOLO - NAS - pose - N 和 YOLO 11n - pose, 在 mAP_{0.5} 上分别提升 4.3、5.0、5.9、3.7、3.2、3.3、3.0、3.5 个百分点; 其中 Lite - HRNet、YOLO v5n - pose 两者参数量低, 推理速度快, 但针对边缘轮廓关键点检测的精度表现不佳; 相较于 HRNet、YOLO v7 - W6 - pose 和 YOLO v8s - pose 等复杂度更高的模型, YOLO v8n - pose 在参数量上分别减少 2.49×10^7 、 6.22×10^7 和 7.9×10^6 , 并且精度上表现优秀。和 YOLO - NAS - pose - N、YOLO 11n - pose 等最新轻量级关键点检测模型相比, YOLO v8n - pose 展现了更为平衡的精度与速度表现, 适用于体尺关键点检测任务。YOLO v8n - SK 达到更最高的检测精度, 更适用于奶山羊体尺关键点检测任务。试验表明 YOLO v8n - SK 展现了更强的关键点检测能力, 在处理具有边缘轮廓特征的体尺关键点检测任务中具有显著优势。

表 2 关键点检测模型性能

Tab.2 Performance of key point detection models

模型	平均精度 均值/%	参数量	浮点运算 量	推理速度/ (f·s ⁻¹)
HRNet	90.2	2.84×10^7	1.44×10^{10}	94.2
Lite - HRNet	89.5	1.80×10^6	2.60×10^9	168.1
YOLO v5n - pose	88.6	1.90×10^6	5.10×10^9	159.5
YOLO v7 - W6 - pose	90.8	6.67×10^7	8.26×10^{10}	90.4
YOLO v8s - pose	91.3	1.14×10^7	3.15×10^{10}	120.6
YOLO v8n - pose	91.2	3.50×10^6	8.90×10^9	132.5
YOLO - NAS - pose - N	91.5	9.70×10^6	1.45×10^{10}	125.4
YOLO 11n - pose	91.0	3.00×10^6	7.90×10^9	145.3
YOLO v8n - SK	94.5	3.30×10^6	8.60×10^9	150.9

为更直观体现改进模型的有效性, 针对奶山羊

侧视图和俯视图进行关键点检测结果分析, 图 8 为各项主流模型与本文方法的检测结果。由图 8a 可知, HRNet、YOLO v5n - pose 和 YOLO v7 - W6 - pose 在关键点检测上存在显著问题, 主要表现为蹄部关键点的偏移以及背部关键点偏离轮廓, 导致部分关键点被误检为背景区域或落入羊体内部。YOLO v8n - pose 在处理姿态倾斜时, 背部和蹄部关键点的检测精度也较差, 未能有效捕捉这些区域的特征变化。由图 8b 可知, HRNet、YOLO v5n - pose、YOLO v7 - W6 - pose 和 YOLO v8n - pose 都出现了关键点误检情况, 尤其是部分关键点被错误地定位到地面上。相比之下, YOLO v8n - SK 能够显著减少关键点误检和漏检现象, 并有效缓解姿态变化带来的关键点位置偏移的问题, 展现了更强的对关键点的检测能力。

3.3.2 消融试验

为验证所引入的 C2f_Ghost 模块、双域选择机制 (DSM) 和全局注意力机制 (GAM) 对奶山羊体尺关键点检测模型性能的影响, 进行消融试验, 结果如表 3 所示。由表 3 可知, 与初始模型 YOLO v8n - pose 相比, 引入 C2f_Ghost 模块参数量减少 1.1×10^6 , 浮点运算量下降 1.3×10^9 , 推理速度提高 25.9 f/s, 同时 mAP_{0.5} 提升 0.2 个百分点, 说明该模块可以在保证精确率的同时降低网络计算复杂度, 提升检测效率; 单独引入 DSM 和 GAM 时, mAP_{0.5} 分别提升 1.4、1.7 个百分点, 说明这 2 个模块能使模型更关注羊只边缘轮廓区域及关键点特征, 有效增强对关键点的感知能力, 提升检测精度。当 C2f_Ghost 模块与 DSM 模块和 GAM 注意力机制分别结合使用时, mAP_{0.5} 分别提升 1.8、2.0 个百分点, 同时推理速度分别提高 25.1、20.7 f/s, 进一步验证了在保持参数量适中的前提下, 能够兼顾检测精度与推理效率。最终, 当同时引入 3 个模块后, mAP_{0.5} 提升 3.3 个百分点, 参数量相比原始模型减少 2×10^5 , 推理速度提升 18.4 f/s, 浮点运算量降低 3×10^8 , 充分证明了本文改进模型在性能和效率上的综合优势。试验表明所引入的模块能够有效改善模型的特征提取能力、增强对边缘轮廓区域和关键点特征关注, 证明关键点检测模型设计的合理性和有效性。

3.3.3 奶山羊体尺测量结果

为进一步验证关键点检测的有效性, 通过计算体尺参数评判检测效果。由于胸围等体尺信息需要结合空间坐标和点云模型, 但图像中关键点为二维坐标, 缺少深度信息, 无法直接映射为空间坐标。因此根据像素坐标在深度图中获取深度, 并利用相机内部参数矩阵计算空间坐标值, 实现像素坐标系向

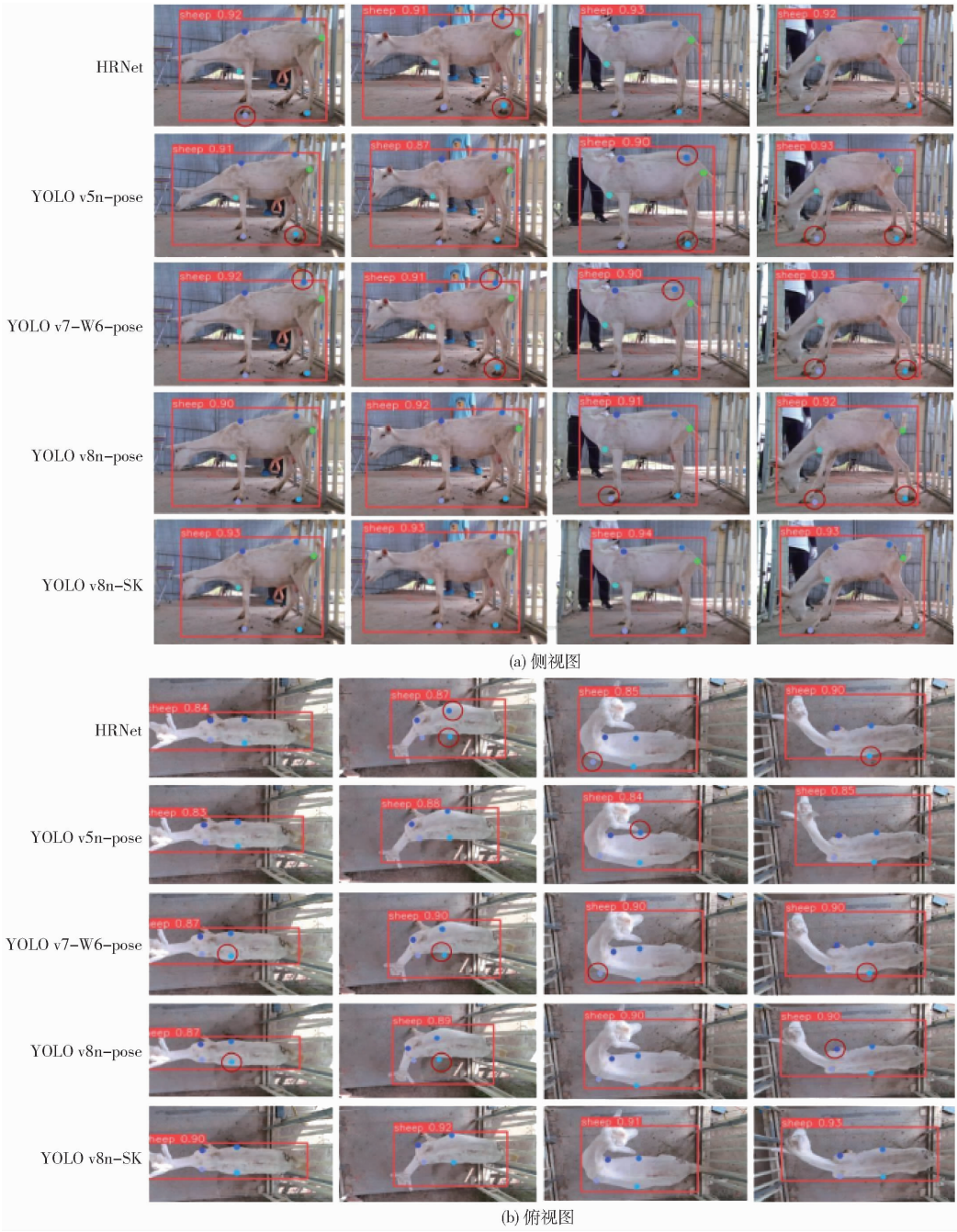


图 8 关键点检测结果

Fig. 8 Results of key point detection

表 3 消融试验结果

Tab. 3 Ablation experiment results

C2f_Ghost	DSM	GAM	平均精度 均值/%	参数量	浮点运 算量	推理速度/ (f·s ⁻¹)
√	×	×	91.4	2.4×10 ⁶	7.60×10 ⁹	168.4
×	√	×	92.6	3.8×10 ⁶	8.90×10 ⁹	118.4
×	×	√	92.9	3.9×10 ⁶	9.60×10 ⁹	104.7
√	√	×	93.0	3.0×10 ⁶	8.50×10 ⁹	157.6
√	×	√	93.2	3.1×10 ⁶	8.90×10 ⁹	153.2
×	√	√	94.3	4.2×10 ⁶	1.02×10 ¹⁰	103.8
√	√	√	94.5	3.3×10 ⁶	8.60×10 ⁹	150.9

注:√表示采用此模块,×表示未采用此模块。

空间坐标的转换,并通过点云映射实现奶山羊点云中关键点的定位,转换流程^[32]如图 9 所示。

体高等线性体尺参数由关键点三维坐标之间的欧氏距离获取;胸围提取基于胸宽关键点,胸围曲线提取和计算过程如图 10 所示。首先,根据胸宽点坐标在 X 轴即躯体方向上获取点云切片,包含 150 ~ 300 个点,能较为全面地反映胸围轮廓结构信息。之后将切片点云投影到 YOZ 平面并去除重复点,采用 KDTree 算法对投影后的点集进行排序,能有效避免点云分布中可能出现的跳跃、重复或轨迹错乱等问题,确保点集在空间上的连贯性。排序完成后,使用 3 次 B 样条曲线对点云轮廓进行拟合获得胸围

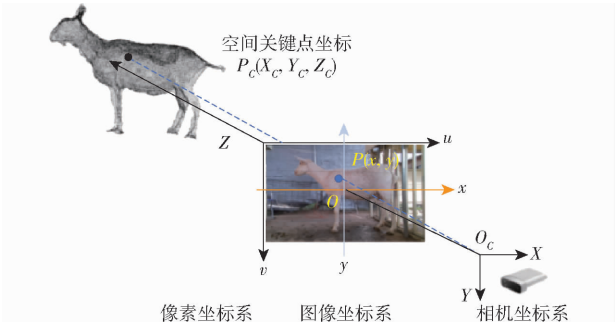


图 9 坐标转换示意图

Fig. 9 Coordinate conversion diagram

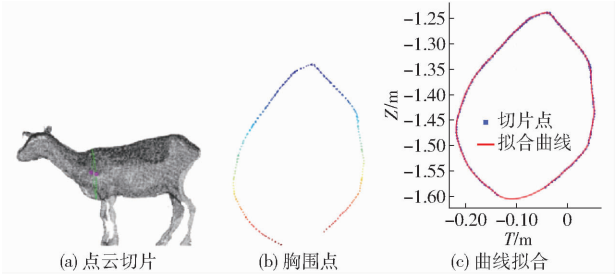


图 10 胸围曲线拟合

Fig. 10 Chest curve fitting

曲线,根据拟合后的轮廓曲线,计算胸围。

针对 50 只奶山羊的 6 项体尺参数进行测量,分别为体高、十字部高、体长,胸宽,腹宽和胸围。对比预测值与真实值,各项体尺 MAE 和 MAPE 如表 4 所示,采用箱线图分析 MAPE,如图 11 所示。由图 11 可知,对于体高、十字部高、体长和腹宽测量,MAPE 分别为 4.72%、4.36%、4.26% 和 4.84%,均低于 5%,验证了关键点检测的准确性,能有效计算体尺信息。对于胸宽和胸围测量,MAPE 达到 6.21% 和 5.83%,相较其他体尺参数,易受羊只姿态影响,其中胸宽测量误差主要由于姿态倾斜导致胸宽测点选取不精确,胸围测量也因姿态倾斜难以获取标准点云切片。

表 4 奶山羊体尺测量结果

Tab. 4 Body size measurement results of dairy goats		
体尺信息	平均绝对误差/cm	平均绝对百分比误差/%
体高	3.43	4.72
十字部高	3.21	4.36
体长	3.27	4.26
胸宽	1.33	6.21
腹宽	1.34	4.84
胸围	5.64	5.83

胸围的测量由于姿态的影响,直接使用主成分分析(Principal component analysis,PCA)校准坐标系时,所得坐标不符合测量需求,如图 12a 所示。主要由于羊只脖子及头部点云占据整体点云比重较大,抬头、低头和扭头都会对 PCA 造成影响,因此在计

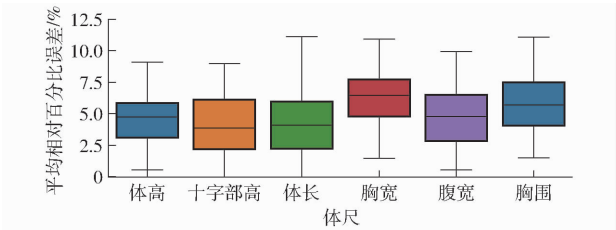


图 11 体尺信息测量误差

Fig. 11 Body size information measurement error

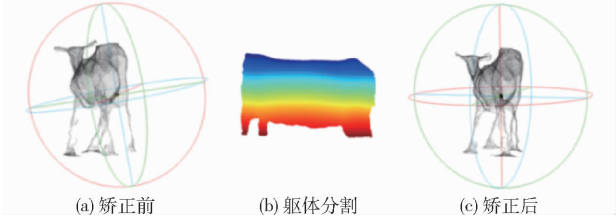


图 12 羊只姿态矫正

Fig. 12 Sheep posture correction

算点云质心后,通过直通滤波只保留躯体点云,对羊只点云分割,从而减少脖子点云影响。只针对躯体使用 PCA 方法,记录变换矩阵并应用到脖子等剩余点云,最后进行拼接,可实现羊只点云坐标系的矫正,从而获取更准确的胸围点云切片,有效缓解姿态对于胸围数据的影响,经姿态矫正后胸围误差可减少 1.2 个百分点。

为进一步量化姿态倾斜对于胸围体尺误差的影响,按照不同的倾斜角度对羊只点云进行分类验证,分别为标准、倾斜 0° ~ 45°、倾斜 45° ~ 90° 和大于 90° 共 4 个组别,结果如表 5 所示。由表 5 可知,随着倾斜角度增加,胸围误差呈明显增大趋势。这是由于头部倾斜角度越大,胸部区域变形越严重。结果表明,姿态倾斜角度对胸围体尺的准确度有显著影响,尤其在大角度倾斜情况下,测量误差较大。

表 5 不同倾斜角度胸围测量结果

Tab. 5 Chest measurement results at different tilt angles

参数	标准	0° ~ 45°	45° ~ 90°	大于 90°
MAE/cm	4.16	4.28	4.89	5.95
MAPE/%	4.42	4.53	5.11	6.32

由于点云重建依赖于相机标定和配准,因此对点云配准误差进行分析。羊只点云重建是由 3 个视角点云进行配准,采用基于相机标定的粗配准方法和基于 ICP 的精配准方法。以顶部视角为目标点云,两侧点云为源点云,由于顶部点云和两侧点云只是存在部分重叠,因此分析不同重叠率下配准均方根误差(Root mean square error, RMSE),当重叠率分别为 40%、30%、20% 和 10% 时, RMSE 分别为 6.86、5.71、3.87、4.95 mm。当重叠率为 20% 时,

RMSE 达到最低,可以获取更为精准的羊只点云模型,为后续体尺测量奠定基础。

3.3.4 倾斜姿态下关键点检测结果分析

羊只相较猪只、牛只等家畜更易动,倾斜姿态较

多,对背部和蹄部关键点检测造成显著影响。为进一步验证倾斜姿态下 YOLO v8n-SK 关键点检测的准确性,针对奶山羊侧视图中倾斜姿态数据集进行测试,体尺关键点检测结果如图 13 所示。

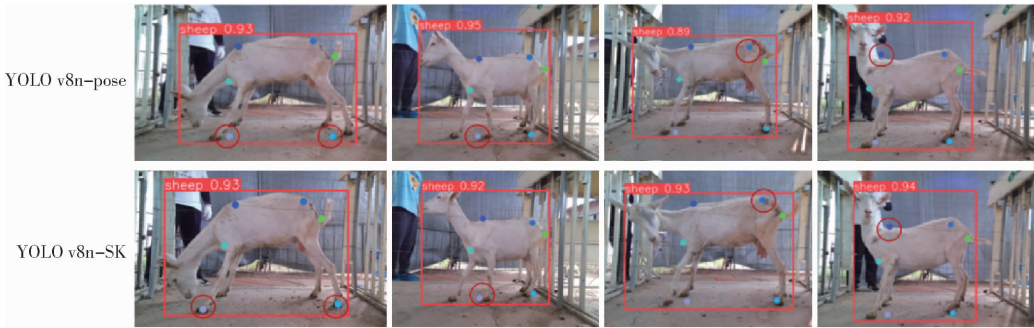


图 13 倾斜姿态检测结果

Fig. 13 Detection results of non-standard posture

由图 13 可知, YOLO v8n-pose 在处理倾斜姿态时,背部关键点中部分关键点落在背景区域或羊体内部,且由于姿态的影响,蹄部位置表现出较大的不稳定性,从而导致明显的误检和偏移现象,蹄部关键点被定位到地面上。而改进的 YOLO v8n-SK 模型能够显著缓解倾斜姿态带来的干扰,更准确地识别羊只边缘轮廓关键点和蹄部位置,表现出更强的鲁棒性。

3.3.5 体尺关键点检测模型泛化性分析

为了进一步验证奶山羊体尺关键点检测模型的泛化性,针对各种养殖场景进行试验分析,包括不同光照、不同羊只品种、自由活动状态、室内室外以及多羊只遮挡场景。

针对光照影响,对自然光照下上午和下午的光照动态变化以及极端光照条件下的关键点进行检测。首先,对自然光照下数据集中各视角下 100 幅羊只图像进行 Gamma 变换处理,模拟暗光和曝光 2 种极端光照场景。试验结果如图 14 所示,在 4 种光照条件下,关键点检测模型表现出良好稳定性。自然光照下,关键点检测模型对于上午和下午光照动态变化具有稳定的检测结果。在曝光条件下,关键点也能够被准确定位。在暗光条件下,由于羊只边界与背景之间的差异趋于模糊,导致部分背部关键点定位精度受到一定影响。试验结果表明本文模型能够应对各种光照变化对检测精度的挑战。

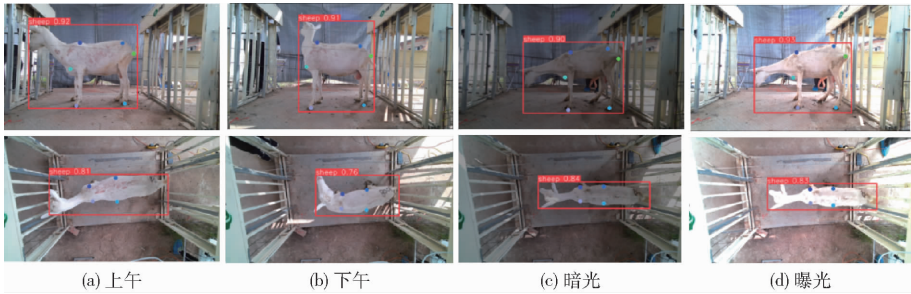


图 14 不同光照条件下检测结果

Fig. 14 Detection results under different light conditions

针对不同羊只品种泛化性验证,在湖羊数据集上开展了试验验证。该数据集中羊只月龄范围为 6~48 月。试验同样针对侧视图和俯视图 2 个视角下的关键点检测任务进行评估。检测结果如图 15 所示。由图 15 可知,模型在整体上能够较为准确地识别关键点位置,关键点分布与湖羊体态结构一致,表现出良好的跨品种适应性。然而,由于湖羊毛发相对浓密,尤其在背部和胸部区域,边缘轮廓更为平滑模糊,导致部分关键点的定位精度受到一定影响。

在上述区域,关键点存在向内部偏移的现象,被定位至羊只躯体内部。结果表明在面临毛发遮挡、边界模糊等复杂感知条件下,模型在边缘关键点检测方面仍存在进一步优化空间。

为进一步验证模型的泛化性,结合养殖场景下的多样化条件,开展了多角度试验分析。考虑到奶山羊在自由活动状态下的图像采集可能受到不同拍摄距离及光照环境的影响,分别针对室内与室外场景进行了试验,结果如图 16 所示。由

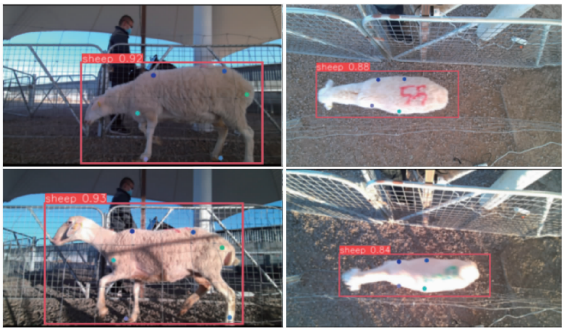


图 15 不同羊只品种检测结果

Fig. 15 Test results of different sheep breeds



图 16 不同养殖场景检测结果

Fig. 16 Test results of various farming conditions

图 16 可知,在室外条件下,模型能够准确识别大多数关键点。然而,在室内环境中,由于光线较暗、存在明显阴影等因素,对模型的检测性能产生一定干扰,尤其是蹄部关键点识别出现定位偏差。此外针对羊只重叠遮挡条件进行验证分析,结果显示,在羊只出现遮挡或重叠时,模型关键点检测的准确性受到较大影响,尤其在背部区域。当羊只重叠较为严重时,边缘关键点难以被有效区分,遮挡羊只的关键点更易出现漏检现象,难以通过体尺关键点获取有效体尺参数。

3.3.6 模块检测效果可视化分析

为更直观验证消融试验各模块贡献,针对 C2f_Ghost 模块的减少冗余特征功能,采用特征图进行分析,针对 DSM 和 GAM 增强边缘特征区域和关键点位置的感知能力,采用热力图进行分析。

首先,对网络中部分层级的 C2f 与 C2f_Ghost 模块的输出特征图进行可视化,结果如图 17 所示,左侧红色边框内为通过部分普通卷积提取得到的本体特征图,右侧绿色边框中为基于本体特征图生成的

Ghost 特征图。图 17a 中,左上角为输入原始图像,右侧为 C2f 模块提取的特征图。可观察到特征图中存在较多相似的冗余信息,不同通道间存在结构重复、信息重叠现象,表明传统卷积在特征提取过程中存在冗余。分析可得,Ghost 特征图在保持关键信息的同时,与本体特征图也存在较大差异,呈现出更加多样化的表达形式,有效提升了特征的利用效率,减少了无效或重复特征的产生,验证了 C2f_Ghost 模块在特征提取与冗余压缩方面的优势。

为进一步分析模型对关键区域的感知能力,采用热力图的方式对 DSM 和 GAM 模块进行可视化,结果如图 18 所示。由图 18 可知,除羊只目标的中心区域具有明显响应外,其边缘区域同样呈现出较强的关注度。这表明 DSM 能显著增强对羊只边缘等区域的感知能力,引导模型更加关注姿态变化带来的边缘轮廓变化,具备良好的边界感知能力,从而增强对姿态变化的适应性和检测的稳定性。图 18c 为经 GAM 增强后关键点检测结果热力图,可观察到在各关键点位置呈现明显的高响应区域,表明 GAM 可以更好聚焦于关键点位置,提升了对关键结构信息的表达能力,从而显著提高关键点检测的准确性与鲁棒性。

3.3.7 低资源场景下推理速度测试

为验证 YOLO v8n - SK 在实际养殖场景下的部署可行性,在低资源设备上推理速度测试。Jetson Nano P3450 开发板是 NVIDIA 推出的一款面向边缘 AI 应用的开发平台,搭载四核 ARM Cortex - A57 处理器和基于 Maxwell 架构的 128 核心 CUDA GPU,并配备 4 GB LPDDR4 内存。测试过程中,输入图像尺寸设为 640 像素 × 640 像素,关键点数量为侧视图 6 个、俯视图 4 个。YOLO v8n - SK 未加速状态下的推理速度为 3.5 f/s,通过部署 TensorRT 并采用 FP16 精度优化后,推理速度可提升至 10.2 f/s。试验结果表明本模型在 Jetson Nano 等边缘设备上的部署具有可行性,可以满足养殖环境中推理速度需求。

4 结论

(1) 基于 YOLO v8n - pose 模型提出了 YOLO v8n - SK 模型。该模型在主干网络和颈部网络中集成了 C2f_Ghost 模块,以提高关键点检测效率并降低计算复杂度;引入 DSM 模块,增强了网络对边缘轮廓区域特征提取能力;在颈部网络不同层级引入 GAM 注意力机制模块,使得模型在预测过程中更加关注目标关键点的位置信息,显著提高了关键点检测效果。通过结合深度相机参数,将模型检测得到

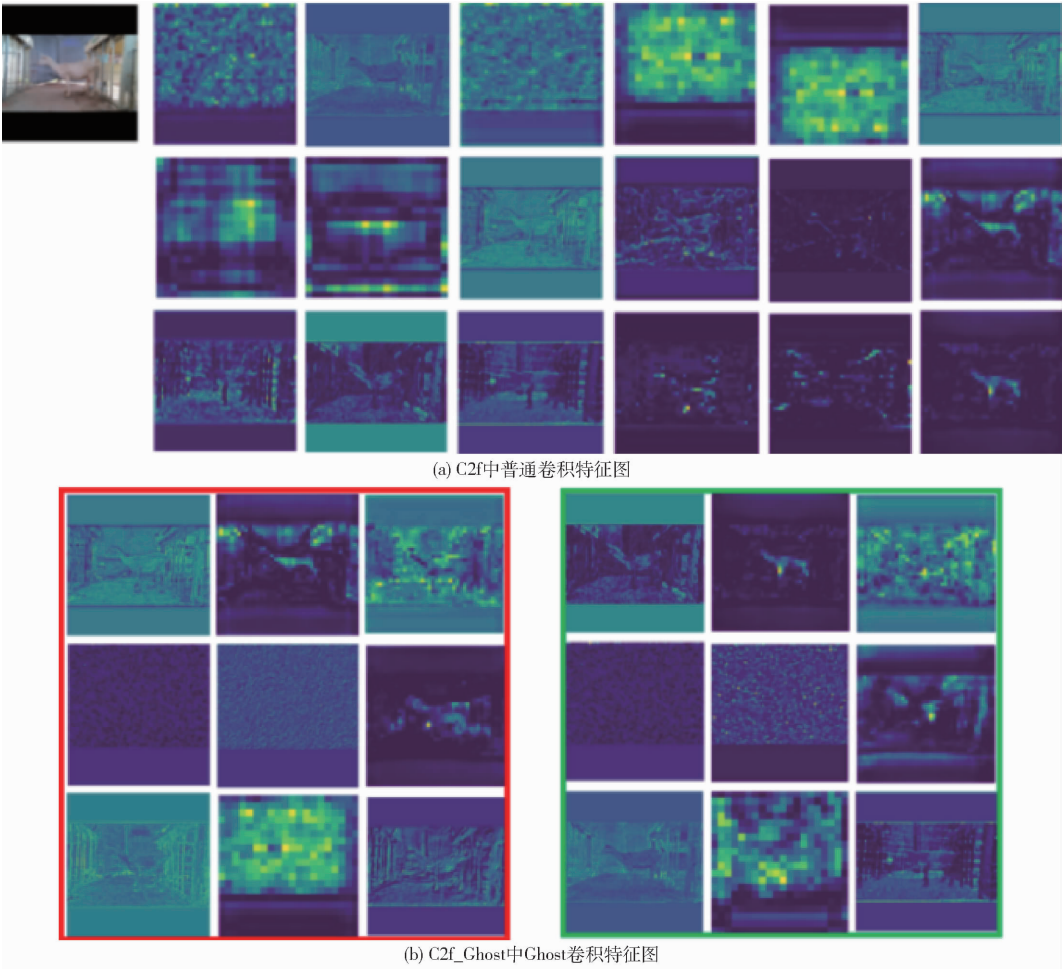


图 17 特征图可视化结果

Fig. 17 Results of feature map visualization

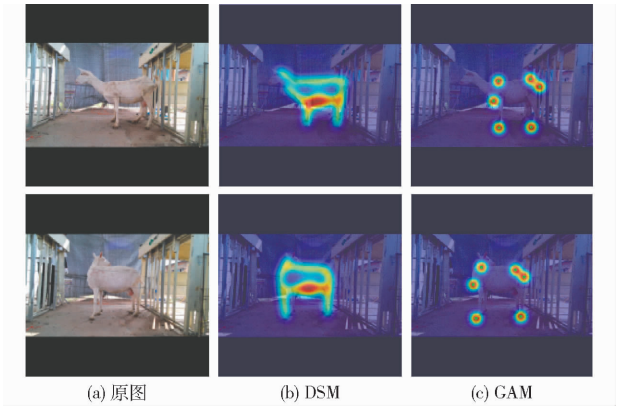


图 18 热力图可视化结果

Fig. 18 Results of heatmap visualization

的关键点二维坐标转换为三维坐标并映射到点云模型中,并实现了奶山羊体高、十字部高、体斜长、胸宽、腹宽和胸围等 6 项体尺参数测量。

(2)YOLO v8n-SK 模型相比 YOLO v8n-pose 模型在 $mAP_{0.5}$ 提升 3.3 个百分点,参数量降低 2×10^5 。测得奶山羊体高、十字部高、体长、胸宽和腹宽平均绝对误差分别为 3.43、3.21、3.27、1.33、1.34 cm;平均相对误差分别为 4.72%、4.36%、4.26%、6.21% 和 4.84%。胸围参数通过进一步坐标系矫正可以实现相对误差为 4.63%;表明所提方法可满足实际养殖过程中奶山羊的关键点检测和体尺测量需求。

参 考 文 献

[1] 初梦苑,司永胜,李前,等. 家畜体尺自动测量技术研究进展[J]. 农业工程学报, 2022,38(13):228-240.
Chu Mengyuan,Si Yongsheng,Li Qian, et al. Research progress on automatic measurement technology of livestock body size [J]. Transactions of the CSAE, 2022,38(13):228-240. (in Chinese)

[2] 张飞宇,王美丽,王正超. 引入 Transformer 和尺度融合的动物骨骼关键点检测模型构建[J]. 农业工程学报, 2021, 37(23):179-185.
Zhang Feiyu,Wang Meili,Wang Zhengchao. Construction of animal skeleton key point detection model by introducing Transformer and scale fusion[J]. Transactions of the CSAE,2021,37(23):179-185. (in Chinese)

[3] 赵建敏,赵成,夏海光. 基于 Kinect v4 的牛体尺测量方法[J]. 计算机应用, 2022,42(5):1598-1606.
Zhao Jianmin,Zhao Cheng,Xia Haiguang. Cattle body size measurement method based on Kinect v4[J]. Computer Applications, 2022,42(5):1598-1606. (in Chinese)

[4] Zheng Z., Gao J B., Weng Z. Measurement of body size parameters and body weight prediction in beef cattle based on image

- analysis[J]. *Journal of Intelligent & Fuzzy Systems*, 2024,47(1-2):155-167.
- [5] 赵建敏, 关晓鹏. 基于双目深度估计的牛体尺测量方法设计[J]. *光电子·激光*, 2022,33(4):429-435.
Zhao Jianmin, Guan Xiaopeng. Design of cattle body size measurement method based on binocular depth estimation[J]. *Optoelectronics and Laser*, 2022,33(4):429-435. (in Chinese)
- [6] 张婧婧, 程芸涛. 基于图像分割的马体尺测量系统设计[J]. *计算机技术与发展*, 2020,30(3):177-180.
Zhang Jingjing, Cheng Yuntao. Design of horse body size measurement system based on image segmentation[J]. *Computer Technology and Development*, 2020,30(3):177-180. (in Chinese)
- [7] 陆明洲, 光二颖, 陈子康, 等. 基于双视角图像的山羊体尺自动测量方法[J]. *农业机械学报*, 2023,54(8):286-295.
Lu Mingzhou, Guang Erying, Chen Zikang, et al. Automatic measurement method of goat body size based on dual-view images[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023,54(8):286-295. (in Chinese)
- [8] Li J, Ma W, Li Q, et al. Multi-view real-time acquisition and 3D reconstruction of point clouds for beef cattle[J]. *Computers and Electronics in Agriculture*, 2022,197:106987.
- [9] Ruchay A, Kober V, Dorofeev K, et al. Accurate body measurement of live cattle using three depth cameras and non-rigid 3-D shape recovery[J]. *Computers and Electronics in Agriculture*, 2020,179:105821.
- [10] Yang G, Xu X, Song L, et al. Automated measurement of dairy cows body size via 3D point cloud data analysis[J]. *Computers and Electronics in Agriculture*, 2022,200:107218.
- [11] Li J, Ma W, Bai Q, et al. A posture-based measurement adjustment method for improving the accuracy of beef cattle body size measurement based on point cloud data[J]. *Biosystems Engineering*, 2023,230:171-190.
- [12] 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.
- [13] Jin B, Wang G, Feng J, et al. PointStack based 3D automatic body measurement for goat phenotypic information acquisition[J]. *Biosystems Engineering*, 2024,248:32-46.
- [14] 赵健, 周国源, 王智文, 等. 基于 SGPointNet++ 模型的奶牛点云分割与表型自动测定系统设计[J]. *农业机械学报*, 2025,56(3):180-187.
Zhao Jian, Zhou Guoyuan, Wang Zhiwen, et al. Design of automatic determination system for point cloud segmentation and morphology of dairy cows based on SGPointNet++ model[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025,56(3):180-187. (in Chinese)
- [15] Nguyen A H, Holt J P, Knauer M T, et al. Towards rapid weight assessment of finishing pigs using a handheld mobile RGB-D camera[J]. *Biosystems Engineering*, 2023,226:155-168.
- [16] 张淑瑾, 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*, 2025,56(3):282-289.
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*, 2025,56(3):180-187. (in Chinese)
- [17] 翁智, 范琦, 郑志强. 基于多模态图像信息及改进实例分割网络的肉牛体尺自动测量方法[J]. *智慧农业(中英文)*, 2024,6(4):64-75.
Weng Zhi, Fan Qi, Zheng Zhiqiang. Automatic measurement method of beef cattle body size based on multimodal image information and improved instance segmentation network[J]. *Smart Agriculture*, 2024,6(4):64-75. (in Chinese)
- [18] Yang G, Qiao Y, Deng H, et al. One-stage keypoint detection network for end-to-end cow body measurement[J]. *Engineering Applications of Artificial Intelligence*, 2025,146:110333.
- [19] Singh N, Devi I, Dudi K, et al. Development of attention-enabled multi-scale pyramid network-based models for body part segmentation of dairy cows[J]. *Journal of Biosystems Engineering*, 2024,49(2):186-201.
- [20] Li M, Su L, Zhang Y, et al. Automatic measurement of mongolian horse body based on improved YOLO v8n-pose and 3D point cloud analysis[J]. *Smart Agriculture*, 2024,6(4):91-102.
- [21] Guo H, Li Z, Ma Q, et al. A bilateral symmetry based pose normalization framework applied to livestock body measurement in point clouds[J]. *Computers and Electronics in Agriculture*, 2019,160:59-70.
- [22] Maji D, Nagori S, Mathew M, et al. YOLO-Pose: enhancing YOLO for multi person pose estimation using object keypoint similarity loss[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022:2637-2646.
- [23] Kadam P, Fang G, Amirabdollahian F, et al. Hand pose detection using YOLO v8-pose[C]//*2024 IEEE Conference on Engineering Informatics (ICEI)*. IEEE, 2024:1-6.
- [24] Han K, Wang Y, Tian Q, et al. GhostNet: more features from cheap operations[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020:1580-1589.
- [25] Cui Y, Ren W, Cao X, et al. Focal network for image restoration[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023:13001-13011.
- [26] Liu Y, Shao Z, Hoffmann N, et al. Global attention mechanism: retain information to enhance channel-spatial interactions [PP/OL]. V1.arXiv (2021-12-10) [2025-03-20]. <https://doi.org/10.48550/arXiv.2112.05561>.
- [27] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[C]//*Proceedings of the European Conference on Computer Vision (ECCV)*, 2018:3-19.
- [28] 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*, 2021,43(10):3349-3364.
- [29] Yu C, Xiao B, Gao C, et al. Lite-HRNet: a lightweight high-resolution network[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021:10440-10450.
- [30] Wang C, Bochkovskiy A, Liao H, et al YOLO v7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023:7464-7475.
- [31] Khanam R, Hussain M. YOLO v11: an overview of the key architectural enhancements[PP/OL]. V1.arXiv (2024-10-23) [2025-03-20]. <https://doi.org/10.48550/arXiv.2410.17725>.
- [32] Du A, Guo H, Lu J, et al. Automatic livestock body measurement based on keypoint detection with multiple depth cameras[J]. *Computers and Electronics in Agriculture*, 2022,198:107059.