doi:10.6041/j.issn.1000-1298.2026.02.025

基于机器视觉的舍饲羊只行为识别方法

黄小平[1,2]　豆子豪[1]　郭阳阳[1,2]　廖振慧[1]　韩　聪[1]　魏诗磊[1]

(1. 安徽大学互联网学院, 合肥 230039; 2. 农业生态大数据分析与应用技术国家地方联合工程研究中心, 合肥 230039)

摘要: 在规模化绵羊养殖场中,畜禽的行为特征能够有效反映其健康状况及环境适应能力。针对传统舍饲羊只行为识别方法在羊群密度变化条件下存在的监测效率低、识别精度不足等问题,提出了一种基于改进 YOLO 11n 模型的舍饲羊只行为识别方法。在羊圈斜上方安装 2D 摄像机,采集羊群的视频数据,并构建包含站立、进食、饮水和休息 4 种行为的舍饲羊行为数据集。在 YOLO 11n 模型基础上,结合 CARAFE(Content-aware reassembly of features)上采样结构,并引入高效多尺度注意力机制 EMA(Efficient multi-scale attention)与动态检测头 DyHead(Dynamic feature learning for head detection)构成 YOLO-CFED 模型,提升羊只行为检测的特征提取与识别能力。结果表明,相较原 YOLO 11n 模型,改进 YOLO-CFED 模型在自建数据集上的性能提升显著:识别精确率(Precision)为 95.6%(提升 1.2 个百分点)、召回率(Recall)为 93%(提升 0.4 个百分点)、mAP@0.5 为 94%(提升 0.3 个百分点)、mAP@0.5:0.95 为 82.4%(提升 1.3 个百分点)、F1 值为 93.4%(提升 0.9 个百分点)。该方法能够有效识别羊只 4 种主要行为,为实现羊只行为智能化监测与健康管理提供了有力技术支持。

关键词: 羊只; 行为识别; YOLO 11; CARAFE; 目标检测

中图分类号: TP391　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)02-0256-09　　**OSID**:

Housed Sheep Behavior Recognition Method Based on Machine Vision

HUANG Xiaoping[1,2]　DOU Zihao[1]　GUO Yangyang[1,2]　LIAO Zhenhui[1]　HAN Cong[1]　WEI Shilei[1]

(1. *School of Internet*, *Anhui University*, *Hefei* 230039, *China*

2. *National Engineering Research Center for Agro-Ecological Big Data Analysis and Application*, *Hefei* 230039, *China*)

Abstract: In large-scale sheep farming operations, the behavioral characteristics of livestock serve as effective indicators of their health status and environmental adaptability. Aiming to address the challenges of low monitoring efficiency and insufficient recognition accuracy inherent in traditional methods for housed sheep behavior recognition, particularly under varying flock densities, a novel behavior recognition method was proposed based on an improved YOLO 11n model. Initially, 2D cameras were installed diagonally above the sheep enclosures to capture video data of the flock, from which a housed sheep behavior dataset was constructed, comprising four distinct behaviors: standing, feeding, drinking, and resting. Subsequently, the YOLO 11n model was enhanced by integrating the content-aware reassembly of features (CARAFE) upsampling structure, the efficient multi-scale attention (EMA) mechanism, and the dynamic head (DyHead) framework, resulting in the proposed YOLO-CFED model. This was designed to improve the feature extraction and recognition capabilities for sheep behavior detection. The experimental results showed that the improved YOLO-CFED model significantly improved its performance on self built datasets: the recognition precision reached 95.6%, the recall reached 93%, an mAP@0.5 reached 94%, an mAP@0.5:0.95 reached 82.4% and an F1 score of 93.4%, all indicators were superior to that of the original YOLO 11n model. The proposed method effectively identified the four primary behaviors of sheep, thereby offering robust technical support for the implementation of intelligent behavioral monitoring and health management in sheep farming.

Key words: sheep; behavior recognition; YOLO 11; CARAFE; object detection

收稿日期: 2025-06-20　修回日期: 2025-08-06

基金项目: 安徽省科技厅自然科学基金项目(2308085MC103)、安徽省高等学校自然科学基金项目(2023AH050082)和安徽省自然科学基金项目(2408085QC104)

作者简介: 黄小平(1984—),男,副教授,博士,主要从事畜牧养殖信息研究,E-mail: hxping@mail.ustc.edu.cn

0　引言

畜牧业是农业的重要组成部分，而羊只养殖作为畜牧业的重要分支，其智能化管理对提高生产效率、保障动物福利具有重要意义[1]。畜禽行为能够有效反映其健康状态和环境适应能力[2-4]，不同的行为模式反映了动物本身的不同状态，这是评估动物健康状况和福利的重要指标之一[5-6]。长时间监测羊只行为，可判断羊只是否出现异常状态，如饮水时间短可能是腹痛的表现[7-8]，躺卧时间长可能是疾病和痛苦的表征[9]，进食时间短可能患有消化系统疾病。随着机器视觉技术的发展，研究人员采用深度学习和机器学习的评估方法评估羊只健康状况[10-12]。

KLEANTHOUS 等[13]通过机器学习和传感器相结合研究绵羊的采食、活跃和静止行为。GIOVANETTI 等[14]采用三轴加速度计对放牧绵羊监测采食、反刍和休息行为。MAGNO 等[15]结合三轴加速度计和机器学习技术开发了用于判别绵羊采食、反刍等行为的电子设备。BARWICK 等[16]通过不同加速度计部署方法对羊只的放牧、站立和行走行为预测。翟亚红等[17]提出基于改进的 YOLO v5n 模型对羊圈斜上方的视频序列识别绵羊的进食、躺卧以及站立行为。王旺等[4]提出基于改进的 YOLO v8s 对羊只站立、行走、采食、饮水、趴卧行为进行检测识别。HUANG 等[18]提出 YOLO v5 - CBAM 的山羊和绵羊区分的检测方法。宋怀波等[19-21]在奶牛的反刍行为、跛行和呼吸检测等方面开展了许多工作，也取得了较好的实验效果。

本研究基于 YOLO 11n 模型，采用 CARAFE 模块替代基准模型中的上采样结构，提升遮挡场景下特征提取能力；使用 DyHead 替换基准检测头，解决不同羊群密度带来的多尺度识别难题；在检测头前引入 EMA 注意力机制，增强模型对复杂行为特征的判别能力。以期更好地满足现代化羊只养殖牧场对智能行为监测的实际需求。

1　材料与方法

1.1　数据采集

羊只行为数据集于 2024 年 9 月在安徽省合肥市长丰县福满园生态有限公司标准化养殖场采集。数据采集设备为安装在羊圈斜上方的 2D 海康威视网络摄像机，图像分辨率为 1 920 像素 × 1 080 像素，帧率为 25 f/s。为提升数据集多样性并更贴近实际养殖环境，从 2 个不同角度对羊圈进行拍摄，并分别构建包含 6、9、12 只羊只的不同密度场景，如图 1a 所示。此外，为确保数据集的代表性和完整性，进行了 24 h 不间断数据采集。采集过程全面覆盖了正常光照、弱光照、强光照、遮挡、模糊及夜晚等多种环境条件，其中场景 2 因包含 3 个相邻羊圈，通过后期处理去除两侧羊圈干扰数据，数据集样例如图 1b 所示。

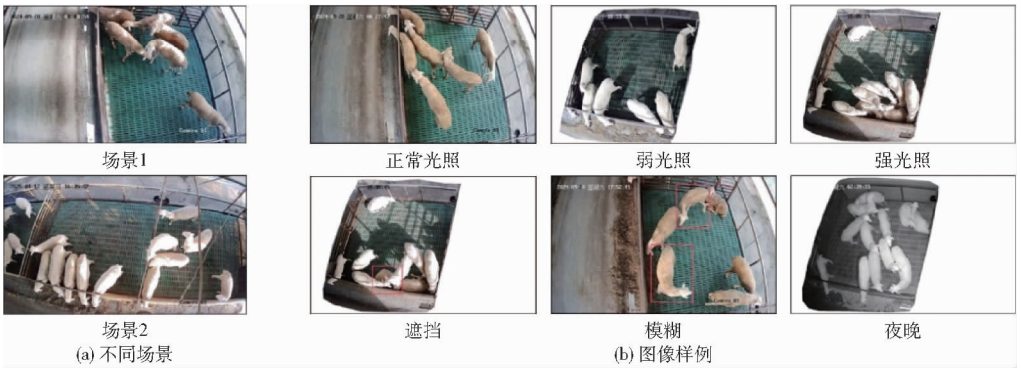


图 1　数据集场景和图像样例

Fig. 1　Dataset scenarios and image samples

羊只日常主要由站立、进食、休息和饮水 4 种核心行为构成，占据其大部分时间。由于这些行为与健康状况紧密相关，分析其模式变化是自动化健康监测的关键一环。例如，进食减少、休息过多、饮水异常等现象，常预示着羊只可能患有口腔、胃肠道或发热性等疾病，图 2a 为 4 种状态的羊只示例图。羊只大部分时间处于站立和休息状态，由于休息状态特征较为单一，所以样本保留数量较少，站立行为标注框数量最多。相比之下，饮水行为持续时间短且无明显规律，标注框数量最少。统计显示，在 4 263 幅图像中，站立状态羊只共 19 105 只、进食状态共 7 031 只、休息状态共 3 431 只、饮水状态共 1 635 只，各行为分布情况如图 2b 所示。

1.2　数据预处理

从 24 h 监控视频中截取关键片段，采用间隔帧采样技术提取原始图像。在对场景 2 中两侧羊圈填充白色背景消除干扰后，使用 Python 库的 LabelImg 标注工具对每幅图像中的羊只目标进行图像标注，

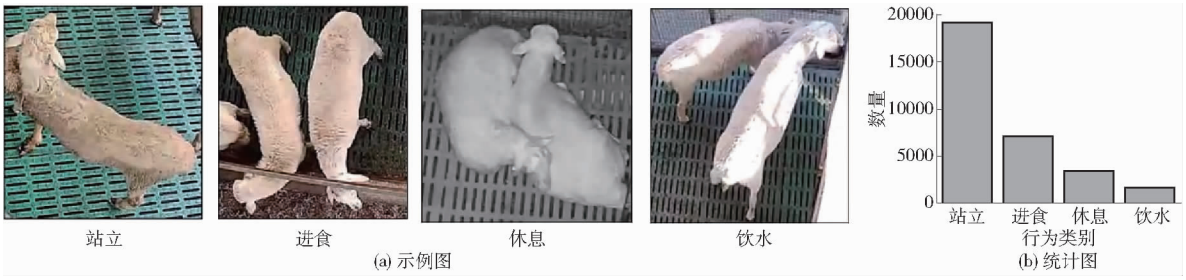


图 2 各状态示例图和统计图

Fig. 2 Example diagrams and statistical charts of each state

并生成 txt 格式的标注文件。为提高数据的多样性和模型的泛化能力,对图像进行了随机变化处理,包括旋转、缩放、上下翻转、镜像及添加噪声等操作,最终制作 4 263 幅有效图像。为了方便模型训练与评估,把图像按照训练集、测试集和验证集比例 8∶1∶1 划分,得到 3 410 幅训练集样本、427 幅测试集样本和 426 幅验证集样本。

2 羊群行为分析方法

2.1 YOLO 11n 目标检测模型

在羊只多行为识别分类任务中,YOLO 11 网络相较于其前代版本(如 YOLO v8、YOLO v10 等)在精度和检测速度方面均表现出显著优势。YOLO 11 系列包含 n、s、m、l、x 5 种不同规模模型,其中 YOLO 11n 作为最轻量化版本,具有参数量少、计算效率高的特点,特别适用于羊圈内羊只的实时多行为检测。

YOLO 11n 模型主要由 Input、Backbone、Neck、Head 4 部分组成。Input 用于接收输入图像并进行预处理,确保输入图像的一致性;Backbone 作为骨干网络通过多层卷积、池化和特殊模块(C3k2)从输入图像中逐步提取图像特征;Neck 作为 Backbone 和 Head 之间的桥梁,用于对骨干网络得到的特征进行特征融合和增强,提高模型对不同大小目标的检测能力;Head 是模型的输出部分,包含多个并行的检测分支,用于执行边界框回归、目标分类和置信度预测等任务。

2.2 基于改进 YOLO 11n 的羊只行为识别模型

2.2.1 模型 YOLO-CFED 改进

在羊只行为识别任务中,通过对 YOLO 11n 进行系统性改进,提出 YOLO-CFED 模型。该模型在保持实时性的同时,可提升羊只行为识别精度,尤其在处理羊群密集、目标遮挡等复杂场景时表现出较强的鲁棒性。具体而言,引入 3 个关键改进模块:①采用内容感知的 CARAFE 模块替换传统 Upsample 上采样操作,以保留更细致的行为特征。②引入动态特征提取的 DyHead 替换原有 Detect 检测头,增强模型对多尺度目标的适应能力。③在检测头的 3 个检测分支前分别引入 EMA 注意力模块,提升关键行为特征提取能力。YOLO-CFED 模型整体架构如图 3 所示。

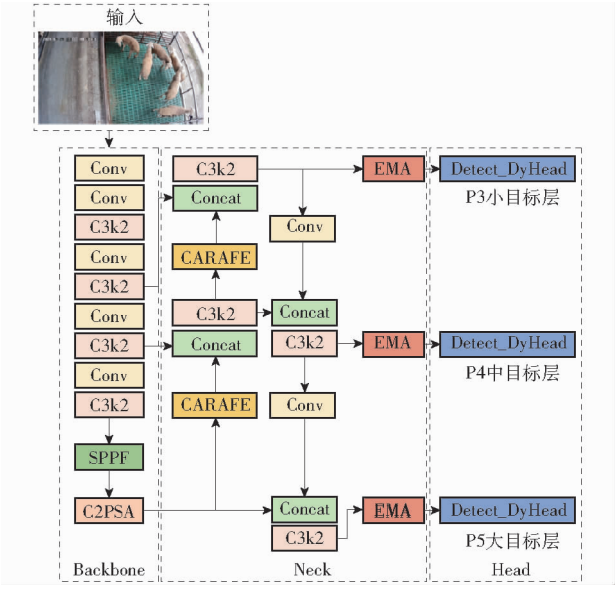


图 3 YOLO-CFED 模型网络结构图

Fig. 3 YOLO-CFED model network structure diagram

2.2.2 CARAFE 上采样

针对圈养羊只的识别场景,当羊群数量增加到 6、9、12 只时,羊群密度的提升使得个体间的相互遮挡现象更加普遍。在 12 只羊只场景中,甚至可能出现完全遮挡情况。这种遮挡现象导致输入图像中羊只的有效视觉区域显著减少,从而使得模型在通过骨干网络提取特征并进行上采样恢复分辨率时,易丢失关键细节信息,尤其是羊只轮廓边缘相关特征的提取。而 YOLO 11 采用的最近邻插值上采样方法是内容无关的,无法根据特征的重要性进行自适应调整,限制了模型对被遮挡或细节模糊的羊只行为进行准确判断的能力。

为了解决上采样过程中特征信息丢失和细节恢复不足的问题,引入 CARAFE 模块,并用其替换 YOLO 11n 的 Upsample 上采样方法,使模型能够更有效地聚焦于关键特征,从而在遮挡现象导致的有效视觉区域减少的情况下,优化特征的上采样效果。CARAFE 结构如图 4 所示,主要由核预测模块和内

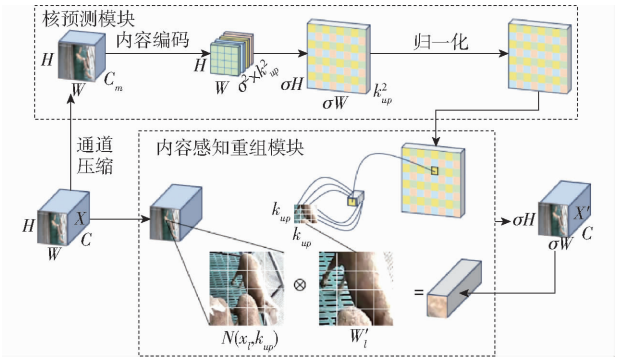


图 4　CARAFE 结构图

Fig. 4　CARAFE structure diagram

容感知重组模块 2 部分构成，其中 H 表示特征图高度，W 表示特征图宽度，C 表示特征图通道数，C_m 表示通道压缩后的通道数，σ 表示上采样的缩放因子，K_{up} 表示上采样核大小，σH 表示上采样后特征图高度、σW 表示上采样后特征图宽度。

具体而言，首先在核预测模块中，将输入特征图 $H \times W \times C$ 通过通道压缩转换为 $H \times W \times C_m$。随后，对压缩后的特征图进行内容编码，生成 $H \times W \times \sigma^2 \times k_{up}^2$ 的中间特征，并通过归一化操作使卷积核权重之和为 1，得到上采样核 $\sigma H \times \sigma W \times k_{up}^2$。最后，在内容感知重组模块中，将归一化后的特征图每个位置映射回原始输入特征图中，提取其中心的 $k_{up} \times k_{up}$ 区域，并与上采样核进行点积运算，生成最终的输出特征图 $\sigma H \times \sigma W \times C$。由于 CARAFE 在同一位置的不同通道间共享相同的上采样核，能够根据局部区域的内容动态调整上采样过程，从而更有效地保留图像的语义信息和细节特征。

2.2.3　动态检测头 DyHead

羊只行为识别不仅需要准确定位每只羊，还需要准确分类其行为。在包含不同数量羊只的场景中，羊只的尺度变化（远近距离）、外观相似性给检测头带来了挑战。YOLO 11 的检测头难以同时优化定位精度和行为分类精度，尤其是在处理不同尺度目标时，其特征表示的适应性和有效性可能不足。此外，密集场景下的个体相互干扰也对检测头的区分能力提出了更高要求。为提升模型对羊群多尺度行为的动态感知能力，采用 DyHead 替换 YOLO 11n 原始检测头。DyHead 模块通过协同整合尺度感知注意力、空间感知注意力及任务感知注意力 3 种机制，实现了跨尺度特征的自适应融合，其结构如图 5 所示，其中，L 表示层级、S 表示空间。

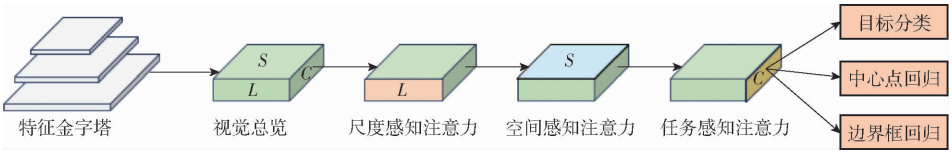


图 5　DyHead 结构图

Fig. 5　DyHead structure diagram

DyHead 的张量 $\boldsymbol{F} \in \boldsymbol{R}^{L \times S \times C}$ 在层级、空间和通道维度上应用 3 种注意力机制

$$W(\boldsymbol{F}) = \pi(\boldsymbol{F})\boldsymbol{F} \tag{1}$$

式中　W——经注意力加权后的特征图

$\pi(\cdot)$——注意力函数

由于直接在所有维度学习注意力函数的计算成本很高，将其分解为 3 个顺序注意力机制：尺度感知注意力 π_L、空间感知注意力 π_S 和任务感知注意力 π_C。由 $\pi_L(\cdot)$、$\pi_S(\cdot)$ 和 $\pi_C(\cdot)$ 三重注意力机制的协同作用构成，使 DyHead 在复杂多变的羊群行为识别任务中展现出优异的特征提取和行为判别能力，其中 $\pi_L(\cdot)$、$\pi_S(\cdot)$ 和 $\pi_C(\cdot)$ 分别为应用于层级 L、空间 S 和通道 C 的注意力函数

$$W(\boldsymbol{F}) = \pi_C(\pi_S(\pi_L(\boldsymbol{F})\boldsymbol{F})\boldsymbol{F})\boldsymbol{F} \tag{2}$$

2.2.4　EMA 注意力机制

在 YOLO 11n 模型中，动态检测头 DyHead 对不同尺度的特征图（P3、P4 和 P5）进行目标检测和分类时，其性能未达到预期效果。为此，在 P3、P4 和 P5 特征图之前引入了 EMA，以增强特征提取和跨维度交互能力。实验结果表明，EMA 的引入显著提升了模型对羊群复杂行为识别能力，尤其是在区分站立、进食、饮水和休息行为时表现突出。EMA 通过分组通道并构建并行分支（1 × 1 和 3 × 3 卷积核），能够同时捕捉不同尺度的空间信息，从而有效提取局部细节和全局轮廓特征。例如，在识别羊只站立和休息行为时，由于姿态幅度差异较大，EMA 的多尺度处理能力可同时关注腿部弯曲状态等局部细节和整体身体轮廓等全局特征，显著提升分类精度。

在检测头部分引入 EMA 后，模型能够更高效地聚焦于关键特征。图 6 为 EMA 模块结构设计，其中通过分组操作（Groups）将通道数 c 划分为 g 个跨通道维度的子特征图，h 和 w 分别表示特征图高度和宽度。EMA 对每个分组的特征图采用坐标信息嵌入分支和 3 × 3 卷积分支提取注意力权重。首先，坐标信息嵌入分支通过 X Avg Pool（沿水平方向平均池化）和 Y Avg Pool（沿垂直方向平均池化）操作，分别对特征图的 2 个空间方向进行一维全局平均池

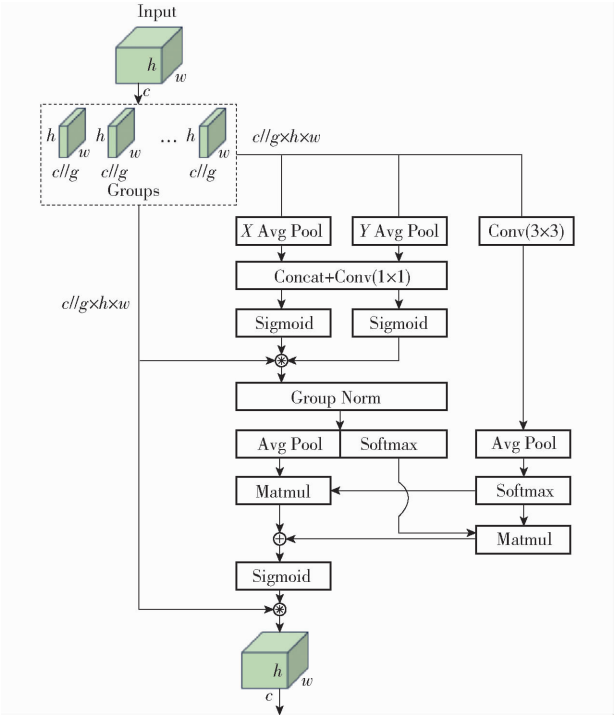


图 6 EMA 结构图

Fig. 6 EMA structure diagram

化。随后,通过 1×1 卷积和特征融合进行降维和特征提取,再通过 Sigmoid 激活函数拟合线性卷积后与原始输入特征相乘;接着,经归一化、平均池化、Softmax 操作和矩阵乘法后,与 3×3 卷积分支进行交互学习。另一分支通过堆叠 3×3 卷积核捕获多尺度特征,之后经过平均池化和 Softmax 操作,与 2 个 1×1 坐标分支进行交互学习,并在像素相加后利用 Sigmoid 函数拟合线性卷积的二维分布,最终与原始输入特征相乘,生成 EMA 的输出模块。通过上述设计,EMA 能够在空间和通道维度上高效聚合多尺度特征,增强模型对复杂背景中羊只行为的建模能力。结合 DyHead 的动态特性,EMA 可显著提升检测头的性能,在处理羊群行为识别任务时表现出色。

3 实验平台及模型评价指标

3.1 实验平台

在 Ubuntu 24.04 操作系统环境下搭建实验平台,硬件配置采用 Intel i5 - 12490F 处理器、32 GB DDR4 内存和 NVIDIA GeForce RTX 4070Ti 显卡。软件开发环境基于 Visual Studio Code 集成开发环境,使用 Python 3.9 编程语言和 PyTorch 2.2.2 深度学习框架。实验设置方面,模型训练批次大小为 16,输入图像分辨率统一调整为 640 像素×640 像素,初始学习率设为 0.01,经过 300 训练轮次确保模型充分收敛。为全面评估 YOLO - CFED 模型的性能优势,选取当前主流的 Faster R - CNN、YOLO v5s、YOLO v8n、YOLO 11n、YOLO 12n 以及 RT - DETR 等目标检测模型作为对比基准。该实验环境确保了所有对比模型在相同条件下进行公平的性能评估。

3.2 评价指标

构建包含检测精度和模型复杂度 2 个维度的综合评价体系,以全面评估 YOLO - CFED 模型在羊只行为检测任务中的性能表现。检测精度方面,采用精确率(Precision)、召回率(Recall)、平均精度均值(mAP)和 F1 值衡量模型对羊只行为的分类能力。其中平均精度均值采用 mAP@0.5 和 mAP@0.5:0.95 2 个层级的指标:mAP@0.5 反映模型在 IoU 阈值为 0.5 时的检测性能,mAP@0.5:0.95 则评估模型在不同 IoU 阈值(0.5 ~ 0.95)范围内的综合表现,二者结合可全面考察模型的检测准确性和鲁棒性。选取模型内存占用量、参数量和浮点运算量评估模型复杂度。

4 结果与分析

4.1 行为识别实验

YOLO 11n 与 YOLO - CFED 模型进行对比,结果如表 1 所示。在站立、进食、休息、饮水羊只行为的检测任务中,相较于基准模型 YOLO 11n,YOLO - CFED 召回率、mAP@0.5 和 mAP@0.5:0.95 均取得提升(精确率见表 2)。表明 YOLO - CFED 模型具有更优越的检测能力。

4.2 不同模型下检测结果对比

为全面评估 YOLO - CFED 模型性能表现,在羊只行为数据集上将其与当前主流检测模型进行对比实验。如表 2 所示,YOLO - CFED 精确率为 95.6%、mAP@0.5:0.95 为 82.4%,优于其他对比模型。具体而言,相比于 Faster R - CNN、YOLO

表 1 YOLO 11n 与 YOLO - CFED 实验对比

Tab. 1 Comparison of YOLO 11n and YOLO - CFED experiments %

模型	召回率				mAP@0.5				mAP@0.5:0.95			
	站立	进食	休息	饮水	站立	进食	休息	饮水	站立	进食	休息	饮水
YOLO 11n	96.4	95.6	89.1	89.4	97.1	95.4	92	90.1	83.8	81.8	79.3	79.5
YOLO - CFED	97.8	96.2	90.1	88.1	97.2	95.5	93	90.3	84.7	83.2	81.6	80.2

表 2　不同模型检测结果

Tab. 2　Different model experiment results　%

模型	精确率				精确率	召回率	mAP@0.5	mAP@0.5:0.95
	站立	进食	休息	饮水				
Faster R－CNN	90.2	86.2	87.2	84.9	87.2	92.0	93.9	56.6
YOLO v5s	96.1	95.2	85.6	93.3	92.6	90.2	92.5	80.9
YOLO v8n	96.3	94.2	93.3	91.7	93.9	92.8	93.7	81.4
YOLO 11n	96.7	94.7	96.3	90.1	94.4	92.6	93.7	81.1
YOLO 12n	96.4	95.6	95.0	92.6	94.9	92.2	93.1	80.9
RT－DETR	96.2	93.9	96.9	94.9	95.5	92.4	91.4	80.2
YOLO－CFED	96.5	95.6	96.1	94.2	95.6	93.0	94.0	82.4

v5s、YOLO v8n、YOLO 11n、YOLO 12n 和 RT－DETR 模型，其精确率分别提升 8.4、3、1.7、1.2、0.7、0.1 个百分点；其 mAP@0.5:0.95 提升 25.8、1.5、1、1.3、1.5、2.2 个百分点。结果证明，YOLO－CFED 在羊只行为识别任务中整体性能最优，为羊只健康管理提供了有效技术手段。

在羊只行为识别场景中，模型复杂度低并不能决定检测性能。如表 3 所示，虽 YOLO－CFED 参数量、浮点运算量和内存占用量高于 YOLO v8n、YOLO 11n 和 YOLO 12n 等轻量级模型，但其在检测精度上展现出显著优势。这种以适度增加模型复杂度换取检测性能提升的设计思路是合理的。YOLO－CFED 相比其他模型 F1 值更优，较 YOLO 11n 提高 0.9 个百分点，其在羊只识别场景中能保持稳定的检测质量，符合羊只行为识别这类对检测精度要求严苛的应用需求。

表 3　不同模型复杂度对比

Tab. 3　Comparison of different model complexities

模型	mAP@0.5/%	F1 值/%	参数量	浮点运算量	内存占用量/MB
YOLO v5s	92.5	91.4	7.030×10^{6}	1.6×10^{10}	14.4
YOLO v8n	93.7	93.3	3.006×10^{6}	8.1×10^{9}	6.3
YOLO 11n	93.7	93.4	2.582×10^{6}	6.3×10^{9}	5.5
YOLO 12n	93.1	93.5	2.557×10^{6}	6.3×10^{9}	5.5
RT－DETR	91.4	93.9	1.988×10^{7}	5.7×10^{10}	40.5
YOLO－CFED	94.0	94.3	3.253×10^{6}	7.9×10^{9}	6.9

4.3　消融实验

为验证 YOLO－CFED 模型的有效性，在保持数据集和训练参数一致前提下，对模型各组件进行消融实验。以 YOLO 11n 模型为基础，依次引入 CARAFE、EMA 和 DyHead 模块进行对比。结果如表 4 所示，其中“√”表示采用该模块，“×”表示不采用该模块。

由表 4 可知，基准模型 YOLO 11n（模型 1）精确率、召回率、mAP@0.5 和 mAP@0.5:0.95 分别为 94.4%、92.6%、93.7% 和 81.1%。在引入 CARAFE 模块后（模型 2），精确率提升 1.2 个百分点，mAP@0.5 和 mAP@0.5:0.95 分别提高 0.2、0.6 个百分点，但召回率下降 0.3 个百分点，表明该模块虽能通过扩大感受野增强特征聚焦能力，但召回率仍需改进。EMA 注意力机制的加入（模型 3）使精确率提升 0.6 个百分点，召回率提升 0.5 个百分点，mAP@0.5 和

表 4　消融实验结果

Tab. 4　Results of ablation experiment　%

模型	CARAFE	EMA	DyHead	精确率	召回率	mAP@0.5	mAP@0.5:0.95
模型 1	×	×	×	94.4	92.6	93.7	81.1
模型 2	√	×	×	95.6	92.3	93.5	81.7
模型 3	×	√	×	95.0	93.1	94.0	81.8
模型 4	×	×	√	94.5	93.4	93.8	83.2
模型 5	√	√	×	94.7	92.4	93.9	82.3
模型 6	√	×	√	95.2	92.3	94.1	82.8
模型 7	×	√	√	94.7	93.0	94.0	82.7
模型 8	√	√	√	95.6	93.0	94.0	82.4

mAP@0.5:0.95 分别提高 0.3、0.7 个百分点，证实了该机制的有效性。DyHead 模块（模型 4）所有指标均取得提升，其中 mAP@0.5:0.95 提升 2.1 个百分点，表现最优。

在模块组合实验中，CARAFE 与 EMA 的组合（模型 5）mAP@0.5:0.95 较优，但召回率下降 0.2 个百分点；CARAFE 与 DyHead 组合（模型 6）使 mAP@0.5:0.95 提升 1.7 个百分点，展现出良好的协同效应；EMA 与 DyHead 组合（模型 7）精确率、召回率和 mAP@0.5 提升幅度较小。整合 CARAFE、EMA 和 DyHead 3 个模块（模型 8）后，精确率为 95.6%（提升 1.2 个百分点），召回率为 93%（提升 0.4 个百分点），mAP@0.5 为 94%（提升 0.3 个百分点），mAP@0.5:0.95 为 82.4%（提升 1.3 个百分点），在所有方案中表现最优。如图 7 所示，消融实验证实了各模块的改进效果，模型 YOLO－CFED 在遮挡场景下的识别表现尤为突出，各项指标均有提升，充分验证了其在羊只行为识别任务中的优越性能。



图 7 消融实验效果

Fig. 7 Effect diagram of ablation experiment

4.4 羊只不同行为下检测结果分析

为验证模型检测性能，选取 5 幅包含不同数量羊只（6、9、12 只）的白天和夜间场景图像进行测试，测试结果如图 8 所示。由图 8 可知，YOLO－CFED 模型检测性能显著优于 YOLO 11n 模型。具体而言，在第 1 幅和第 4 幅图像测试中，YOLO 11n 模型出现了明显漏检现象，而 YOLO－CFED 模型则实现了所有羊只的准确检测，且检测精度更高。此外，在行为识别方面，YOLO 11n 模型存在明显误判现象：在第 1 幅图像中将进食行为误判为站立，在第 3 幅图像中将站立行为误判为饮水；相比之下，YOLO－CFED 模型能够准确区分羊只不同行为状态。综合所有测试场景结果，YOLO－CFED 模型在检测效果上全面优于 YOLO 11n 模型。该模型不仅有效解决了漏检和误检问题，而且在各种羊群规模和光照条件下都展现出优异的检测能力和泛化性能。值得注意的是，即使在夜间场景和羊群密集等具有挑战性复杂条件下，YOLO－CFED 模型仍能保持稳定的检测效果，充分证明了其卓越的鲁棒性。

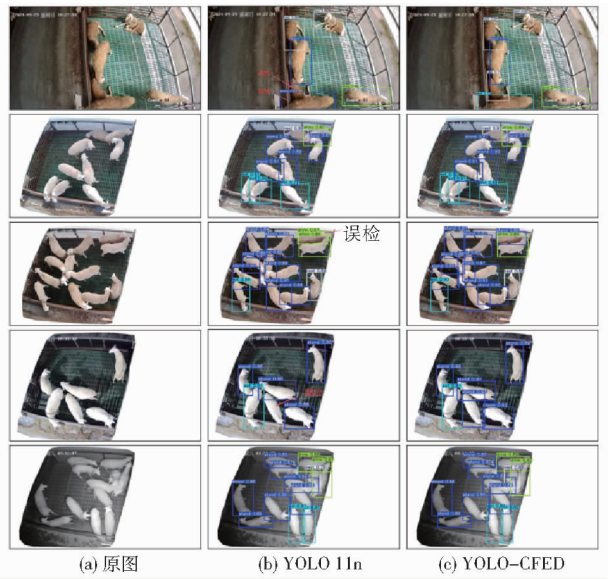


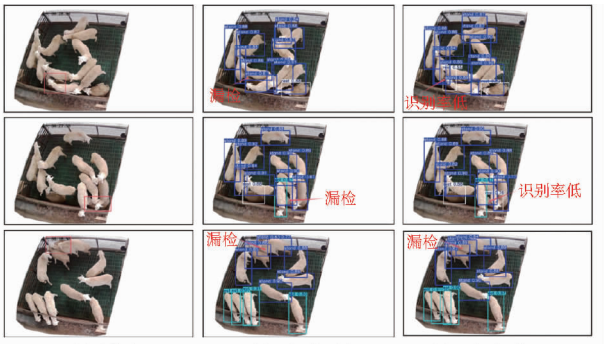
图 8 YOLO 11n 与 YOLO－CFED 检测效果对比

Fig. 8 Comparison of detection performance between YOLO 11n and YOLO－CFED

尽管当前模型通过多模块优化已在一定程度提升了遮挡场景的检测能力，但在极端遮挡情况下仍存在明显局限。当羊只相互遮挡面积超过 60% 时，模型检测率会出现显著下降；而当遮挡程度达到 80% 以上或完全遮挡时，系统不仅难以准确识别具体行为类别，甚至会出现完全漏检情况。另外，在特定行为场景如饮水过程中，由于羊头部位被遮挡导致的特征缺失，会造成饮水行为的误判（如将饮水行为误判为站立行为）。

当前模型存在的技术瓶颈，如图 9 所示。未来

可通过引入时序建模方法，利用视频帧间的运动信息和历史轨迹预测，结合遮挡区域的上下文特征推理，提升模型对短暂完全遮挡目标的持续跟踪能力，从而改善遮挡场景下检测和行为识别准确率。



(a) 原图　(b) YOLO v11n　(c) YOLO-CFED

图 9　遮挡场景下检测效果对比

Fig. 9　Comparison of detection effects in occluded scenes

5　结论

(1)针对羊只 4 种行为(站立、进食、休息和饮水)，提出了一种羊只行为识别模型 YOLO－CFED。针对羊只遮挡导致的特征损失、有效视觉区域减少的问题，引入 CARAFE 上采样模块以提升特征还原能力；针对不同密度场景下羊只尺度变化，引入 DyHead 模块增强多尺度适应性；同时，在 DyHead 的 P3、P4、P5 层级前引入 EMA 注意力机制，进一步加强特征提取能力和跨维度信息交互。结果表明，YOLO－CFED 模型在羊只站立、进食、休息和饮水 4 种行为精确率分别达到 96.5%、95.6%、96.1% 和 94.2%，表现出优异的检测性能。

(2)羊只行为数据集测试结果表明，改进 YOLO－CFED 模型优于对比模型性能。消融实验进一步验证了 CARAFE、DyHead 和 EMA 模块对整体识别效果提升。相比基线模型，改进 YOLO－CFED 在精确率、召回率、mAP@0.5、mAP@0.5:0.95、F1 值分别提高 1.2、0.4、0.3、1.3、0.9 个百分点。与 YOLO v5s、YOLO v8n、YOLO 11n、YOLO 12n 和 RT－DETR 等主流检测模型相比，精确率分别提升 3.0、1.7、1.2、0.7、0.1 个百分点，其余各项指标同样优于对比模型。本文模型能够有效实现对不同密度及遮挡的舍饲羊多种行为的精准监测，具有良好的应用前景。

参　考　文　献

[1] 黄小平，冯涛，郭阳阳，等. 基于改进 YOLO v5s 的轻量级奶牛体况评分方法[J]. 农业机械学报，2023，54(6)：287－296.
HUANG Xiaoping, FENG Tao, GUO Yangyang, et al. Lightweight dairy cow body condition scoring method based on improved YOLO v5s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 287－296. (in Chinese)

[2] 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)

[3] 张曦宇，宣传忠，武佩，等. 基于声信号的畜禽行为信息监测研究进展[J]. 黑龙江畜牧兽医，2017(11)：63－68.
ZHANG Xiyu, XUAN Chuanzhong, WU Pei, et al. Review on monitoring technology for behaviors of livestock and poultry based on acoustic signals[J]. Heilongjiang Animal Husbandry and Veterinary Medicine, 2017(11): 63－68. (in Chinese)

[4] 王旺，王福顺，张伟，等. 基于改进 YOLO v8s 的羊只行为识别方法[J]. 农业机械学报，2024，55(7)：325－335，344.
WANG Wang, WANG Fushun, ZHANG Wei, et al. Sheep behavior recognition method based on improved YOLO v8s[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(7): 325－335, 344. (in Chinese)

[5] PEDEN R S E, TURNER S PBOYLE L A, et al. The translation of animal welfare research into practice: the case of mixing aggression between pigs[J]. Applied Animal Behavior Science, 2018, 204: 1－9.

[6] 夏元天，寇旭鹏，薛洪成，等. 基于改进 YOLO－MAO 检测框架的笼养白羽肉鸡行为检测方法[J]. 农业机械学报，2024，55(11)：103－111.
XIA Yuantian, KOU Xupeng, XUE Hongcheng, et al. Behavior detection algorithm for caged white-feather broilers based on improved YOLO detection framework[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 103－111. (in Chinese)

[7] RICHESON J T, LAWRENCE T E, WHITE B J. Using advanced technologies to quantify beef cattle behavior[J]. Translational Animal Science, 2018, 2(2): 223－229.

[8] SCHWARTZKOPF-GENSWEIN K S, HUISMA C. Validation of a radio frequency identification system for monitoring the feeding patters of feedlot cattle[J]. Livestock Production Science, 1999, 60(1): 27－31.

[9] CHARLTON G, GAULD C, VERONESI F, et al. Assessing the accuracy of leg mounted sensors for recording dairy cow behavioral activity at pasture, in cubicle housing and a straw yard[J]. Animals, 2022, 12(5): 638.

[10] NOOR A, ZHAO Y, KOUBAA A, et al. Automated sheep facial expression classification using deep transfer learning[J]. Computers and Electronics in Agriculture, 2020, 175: 105528.

[11] XU Y, NIE J, CEN H, et al. Spatio-temporal-based identification of aggressive behavior in group sheep[J]. Animals, 2023, 13(16): 2636.

[12] HU T, YAN R, JJANG C, et al. Grazing sheep behaviour recognition based on improved YOLOV5[J]. Sensors, 2023, 23(10): 4752.
[13] KLEANTHOUS N, HUSSAIN A, KHAN W, et al. Deep transfer learning in sheep activity recognition using accelerometer data[J]. Expert Systems with Applications, 2022, 207: 117925.
[14] GIOVANETTI V, DECANDIA M, MOLLE G, et al. Automatic classification system for grazing, ruminating and resting behaviour of dairy sheep using a tri-axial accelerometer[J]. Livestock Science, 2017, 196: 42 - 48.
[15] MAGNO N A, SILVIA H N, DANIEL D S C, et al. Discrimination of ingestive behavior in sheep using an electronic device based on a triaxial accelerometer and machine learning[J]. Computers and Electronics in Agriculture, 2024, 218: 108657.
[16] BARWICK J, LAMB D W, DOBOS R, et al. Categorising sheep activity using a tri-axial accelerometer[J]. Computers and Electronics in Agriculture, 2018, 145: 289 - 297.
[17] 翟亚红,王杰,徐龙艳,等.基于改进 YOLO v5n 的舍养绵羊行为识别方法[J].农业机械学报,2024,55(4):231 - 240.
ZHAI Yahong, WANG Jie, XU Longyan, et al. Behavior recognition of domesticated sheep based on improved YOLO v5n[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 231 - 240. (in Chinese)
[18] HUANG X, HUANG F, HU J, et al. Automatic face detection of farm images based on an enhanced lightweight deep learning model[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2024, 38(12): 2456009.
[19] 宋怀波,牛满堂,姬存慧,等.基于视频分析的多目标奶牛反刍行为监测[J].农业工程学报,2018,34(18):211 - 218.
SONG Huaibo, NIU Mantang, JI Cunhui, et al. Monitoring of multi-target cow ruminant behavior based on video analysis technology[J]. Transactions of the CSAE, 2018, 34(18): 211 - 218. (in Chinese)
[20] 宋怀波,姜波,吴倩,等. 基于头颈部轮廓拟合直线斜率特征的奶牛跛行检测方法[J].农业工程学报, 2018, 34(15): 190 - 199.
SONG Huaibo, JIANG Bo, WU Qian, et al. Detection of dairy cow lameness based on fitting line slope feature of head and neck outline[J]. Transactions of the CSAE, 2018, 34(15): 190 - 199. (in Chinese)
[21] 宋怀波,吴禧华,阴旭强,等. 基于 Lucas - Kanade 稀疏光流算法的奶牛呼吸行为检测[J]. 农业工程学报, 2019, 35(17): 215 - 224.
SONG Huaibo, WU Dihua, YIN Xuqiang, et al. Respiratory behavior detection of cow based on Lucas - Kanade sparse optical flow algorithm[J]. Transactions of the CSAE, 2019, 35(17): 215 - 224. (in Chinese)

(上接第 160 页)

[24] ESTER M, KRIEGEL H P, SANDER J, et al. A density-based algorithm for discovering clusters in large spatial databases with noise[C]// Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD - 96), 1996: 226 - 231.
[25] LI Q, XUE Y. Total leaf area estimation based on the total grid area measured using mobile laser scanning[J]. Computers and Electronics in Agriculture, 2023, 204: 107503.
[26] DEMANTKÉ J, MALLET C, DAVID N, et al. Dimensionality-based scale selection in 3D LiDAR point clouds[J]. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2012, 38: 97 - 102.
[27] WEINMANN M, JUTZI B, MALLET C. Semantic 3D scene interpretation: a framework combining optimal neighborhood size selection with relevant features[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2014, 2: 181 - 188.
[28] KASHANI A G, OLSEN M J, PARRISH C E, et al. A review of LiDAR radiometric processing: from ad hoc intensity correction to rigorous radiometric calibration[J]. Sensors, 2015, 15(11): 28099 - 28128.
[29] 刘旭,李秋洁,徐幼林. LiDAR 强度校正方法研究综述[J].遥感技术与应用,2023,38(1):15 - 25.
LIU Xu, LI Qiujie, XU Youlin. Review on intensity correction method for LiDAR[J]. Remote Sensing Technology and Application, 2023, 38(1): 15 - 25. (in Chinese)
[30] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45: 5 - 32.
[31] LOH W Y. Classification and regression trees[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2011, 1(1): 14 - 23.
[32] LI Qiujie, YAN Yu, LI Weizheng. Coarse-to-fine segmentation of individual street trees from side-view point clouds[J]. Urban Forestry & Urban Greening, 2023, 89: 1 - 12.
[33] WEINMANN M, MALLET C, BR EDIF M. A classification-segmentation framework for the detection of individual trees in dense MMS point cloud data acquired in urban areas[J/OL]. Remote Sens., 2017. https://doi.org/10.3390/rs903277.
[34] LI Qiujie, XUE Yuxi. Real-time detection of street tree crowns using mobile laser scanning based on pointwise classification [J]. Biosystems Engineering, 2023, 231: 20 - 35.
[35] ZHANG Y, SUN P, JIANG Y, et al. Bytetrack: multi-object tracking by associating every detection box[C]// European Conference on Computer Vision, 2022: 1 - 21.