doi:10.6041/j.issn.1000-1298.2026.09.032

基于改进 YOLO v8s + DeepSort 的奶山羊发情行为识别与跟踪方法

廖　娟[1]　张子睿[1]　申　琦[2]　饶　元[3]　何豪旭[3]　张顺龙[1]　王　奎[1]　董　萧[1]

（1. 安徽农业大学工学院，合肥 230036；2. 安徽农业大学动物科技学院，合肥 230036；
3. 安徽农业大学信息与人工智能学院，合肥 230036）

摘要：奶山羊发情行为的准确识别和跟踪对提高养殖繁殖效率和管理水平至关重要。针对传统养殖对奶山羊发情期缺乏准确、及时的行为分析与监控方法问题，提出了一种基于改进 YOLO v8s + DeepSort 的奶山羊发情行为识别与跟踪方法。采用轻量化、无参数的 SimAM 注意力机制增强 YOLO v8s 的特征提取能力，并将其旋转边界框模型（YOLO v8s_obb）、常规矩形框模型（YOLO v8s）分别标记并检测奶山羊的爬跨行为和非爬跨行为；检测得到的行为类别信息和边界框参数信息与 DeepSort 跟踪算法结合，在其 IoU 匹配过程中引入匹配帧数计数机制，增强爬跨与被爬跨奶山羊之间的关联性，精准确定处于发情状态的奶山羊，获取对应的 ID，确保发情行为在连续图像帧中能够准确跟踪。研究结果表明，改进模型对奶山羊爬跨行为和非爬跨行为的平均精度较 YOLO v8s_obb 和 YOLO v8s 分别提升了 0.90、2.64 个百分点；在发情山羊跟踪性能方面，与 DeepSort、BotSort、StrongSort 和 ByteTrack 相比，本文方法的 HOTA 和 IDF1 分数最高，分别达到 72.4%、80.3%，可为奶山羊的繁殖管理提供有力支撑。

关键词：奶山羊；发情行为；爬跨行为；识别与跟踪；YOLO v8s；DeepSort

中图分类号：S818.9　　**文献标识码**：A　　**文章编号**：1000-1298(2026)09-0339-11　　**OSID**：

Estrus Behavior Recognition and Tracking Method of Dairy Goats Based on Improved YOLO v8s and DeepSort Algorithm

LIAO Juan[1]　ZHANG Zirui[1]　SHEN Qi[2]　RAO Yuan[3]　HE Haoxu[3]　ZHANG Shunlong[1]
WANG Kui[1]　DONG Xiao[1]

(1. *School of Engineering*, *Anhui Agricultural University*, *Hefei* 230036, *China*
2. *College of Animal Science and Technology*, *Anhui Agricultural University*, *Hefei* 230036, *China*
3. *School of Information and Artificial Intelligence*, *Anhui Agricultural University*, *Hefei* 230036, *China*)

Abstract: In modern dairy goat farming, accurate recognition and tracking of estrus behavior is crucial for improving breeding efficiency and management standards. Aiming at the lack of accurate and timely behavior analysis and monitoring methods for the estrus of dairy goats in traditional breeding, an estrus behavior recognition and tracking method of dairy goats was proposed. Firstly, the lightweight SimAM attention mechanism was used to enhance the feature extraction capability of YOLO v8s, and a rotating bounding box model (YOLO v8s_obb) and a conventional rectangular box model (YOLO v8s) were used to detect the mounting and the no mounting dairy goats, respectively. Then, combining the detected behavior category information and bounding box parameter information with the DeepSort tracking algorithm, the matching frame count mechanism was introduced into IoU matching process to enhance the correlation between the mounting and the no mounting dairy goats, and accurately determined the dairy goats in estrus, and obtained stable ID, which can ensure the accurate tracking of estrus behavior in successive image frames. The results showed that the AP of the improved model for detecting mounting and non-mounting behaviors in dairy goats was improved by 0.90 percentage points and 2.64 percentage

收稿日期：2025-01-04　修回日期：2025-02-24

基金项目：国家自然科学基金项目（32371993）、安徽省自然科学基金青年基金项目（2308085QC106）和安徽省教育厅高校科研项目（2023AH050982）

作者简介：廖娟（1986—），女，副教授，博士，主要从事机器视觉和农业智能信息处理研究，E-mail：liaojuan@ahau.edu.cn

通信作者：董萧（1992—），男，讲师，博士，主要从事计算机视觉和智慧养殖相关研究，E-mail：dx012210@ahau.edu.cn

points, respectively, compared with that of YOLO v8s_obb and YOLO v8s. In terms of tracking performance of estrus goats, compared with DeepSort, BotSort, StrongSort and ByteTrack, the HOTA and IDF1 scores of this method were the highest, which reached 72.4% and 80.3%, respectively, which could provide strong support for breeding management of dairy goats.

Key words: dairy goats; estrus behavior; mounting behavior; recognition and tracking; YOLO v8s; DeepSort

0 引言

奶山羊是一种以产奶为主的小型反刍动物,近年来随着羊奶制品消费需求的增加,我国奶山羊养殖规模化水平持续增高,智能化是奶山羊产业发展的必然趋势[1]。奶山羊养殖业中,及时、准确地评估羊只的发情状况是羊繁殖周期中的重要环节,直接影响到羊群的繁殖效率。传统的奶山羊发情评估多依赖人工监测的方法观察奶山羊发情行为,需要消耗大量的时间和人力,且准确率难以保证。奶山羊发情期短,发情持续时间仅为 24 ~ 48 h[2],在此期间,发情母羊会接受其他山羊的爬跨,也可能会爬跨其他山羊[3]。爬跨行为是发情行为重要的外在表现形式,但由于其频率低、持续时间短、发生位置随机[4],人工往往难以及时发现,导致发情奶山羊无法及时配种和奶山羊场经济损失。因此,及时、准确地监测识别奶山羊发情状况对现代化奶山羊养殖具有重要意义。

奶山羊发情时生理状况和行为特征与非发情时不同,根据动物在发情期间的生理变化和行为特征,国内外研究人员采用多类别传感器和大数据挖掘技术,研究和分析动物的体征数据,如步履数、声音、体温等,以此判断动物是否处于发情阶段[5]。WANG 等[6]在奶牛脖子上安装电子标签获取加速度和位置数据,监测步数和躺卧时间等 7 项行为指标,设计了反向传播神经网络预测奶牛发情期,准确率为 95.46%。WANG 等[7]将双通道声音采集装置固定在奶牛颈部采集声音数据,提出了基于最优特征组合和最优时间窗的奶牛发情检测方法,平均准确率为 97.62%。这些方法在动物身上植入或佩戴相关传感器,其准确率高,但没有完全解决成本问题和动物福利保护问题[8],且对动物产生一定的应激不适,如佩戴传感器的羊只会出现频繁躁动、体温升高、进食减少、停止反刍等一系列的表现,给羊群的管理和健康保障带来了一定的困难。

随着计算机视觉技术和智慧畜牧业的兴起,研究人员逐步探索基于视频图像分析监测动物发情行为的方法。YU 等[9]基于 YOLO v3,构建了一种轻量级的母羊发情行为检测模型,准确率为 99.44%,模型内存占用量为 40.6 MB。WANG 等[4]提出了一种改进的 YOLO v8n 用于识别发情的奶牛,与 YOLO v8 相比,平均识别精度提升了 5.40%。王少华等[10]提出了一种基于改进 YOLO v3 模型的奶牛发情行为识别方法,识别准确率为 99.15%,召回率为 97.62%。这些研究在动物发情行为检测领域取得了一定进展,但主要局限于对发情行为的初步检测,难以有效地在群体中区分个体。此外,动物发情行为通常是一种动态且连续的行为过程,仅依赖静态检测难以全面获取其行为特征。为获取动物的行为类别、持续时间以及位置等关键信息,大部分学者采用目标跟踪算法,通过给动物分配身份编号(Identity,ID)[11-12],以跟踪动物的行为轨迹。但奶山羊的发情行为具有复杂的交互特性,如非发情山羊可能因应激或群体竞争而表现出类似爬跨行为。此外,发情母羊可能因遮挡导致检测目标丢失,从而误将发情山羊判定为非发情个体。为准确识别和跟踪发情奶山羊,需要在相邻帧之间关联多个目标信息,解决因羊只运动引起的目标位置变化以及羊只之间的遮挡问题。

为此,本文提出一种基于改进 YOLO v8s 与 DeepSort 的奶山羊发情行为的检测和跟踪方法。在 YOLO v8s 的主干网络中引入无参数的 SimAM 注意力模块,不额外增加网络参数的同时,增强重要特征的有效提取,强化奶山羊特征,抑制背景特征,并结合爬跨奶山羊的形态特征,分别以旋转边界框和标准矩形框表示爬跨羊只和非爬跨羊只;将检测到的羊只行为信息和边界框参数引入 DeepSort 跟踪算法,基于 IoU 计算和连续匹配帧数计数机制,寻找与爬跨羊只关联度最大的非爬跨奶山羊,并给两只山羊分配统一的发情 ID,同时显示发情行为,实现在真实群养环境下奶山羊的发情行为跟踪,为无接触式的动物发情行为的连续监测和动态分析提供技术支持。

1 数据集制作与模型建立

1.1 数据采集和制作

本研究视频采集于安徽省滁州市安徽农业大学江淮流域试验站的羊圈内,羊圈长 45 m、宽 36 m、高

1.7 m，羊圈内安装 2 个 2.5 英寸定焦双光 4G 圆顶相机（DS－2DE3Q140MY－T 型，HIKVISION），分辨率为 2 560 像素×1 440 像素，帧率为 25 f/s，格式为 MP4，摄像头安装高度为 2.65 m，安装的摄像机和场地如图 1 所示。



图 1　摄像机安装位置及拍摄图像

Fig. 1　Camera installation position and shooting images

以多个不同品种的奶山羊群体作为研究对象，以确保数据的多样性和代表性，视频选自 2024 年 2—3 月收集的 45 只奶山羊的活动视频。在站内专家的指导下，将奶山羊后肢置于地面，前肢搭在另一只奶山羊身上，且身体与地面呈较大倾斜角度的动作定义为爬跨行为；将无明显身体倾斜角度变化，而是正常行走、站立或休息等状态定义为非爬跨行为。通过观看奶山羊活动视频，人工选取包含上述行为的视频片段 20 段，并利用视频分帧技术分帧，保存格式为 JPG，共获取 1 200 帧有效爬跨图像和非爬跨图像。为扩大数据集的样本规模，提高模型的泛化能力，对数据集进行数据增强操作。首先，针对发情期间奶山羊之间相互遮挡导致位置的不确定性，对数据集图像进行水平翻转、水平移动和垂直移动；其次，针对白天不同时间的光线复杂多变的特点，进行亮度增强和亮度降低，最终得到 3 600 幅数据集图像。使用开源标注工具 RoLabelImg 对爬跨行为标注，使用 LabelImg 对非爬跨行为进行标注。本文的标签数量为 2，旋转框标签命名为“mount”，矩形框标签命名为“normal”。数据集按照比例 7∶2∶1 随机划分为训练、验证和测试集，得到训练集 2 520 幅图像，验证集 720 幅图像，测试集 360 幅图像。

1.2　奶山羊发情行为的识别与跟踪方法

本文提出了一种自然场景下奶山羊发情行为的检测和跟踪方法，包括爬跨行为和非爬跨行为识别以及发情行为跟踪两部分，其技术路线如图 2 所示。首先，将山羊行为视频帧分别输入到 YOLO v8s_obb 网络和 YOLO v8s 网络，得到旋转框处理结果（爬跨行为）和标准矩形框的处理结果（非爬跨行为）；其次，根据两个检测器提供的置信度、检测框参数和行为类别等信息判断当前是否为有效帧，如果不是，则跳过当前帧；如果为有效帧，将检测结果输入到 DeepSort 跟踪器中，通过 IoU 匹配和匹配帧数计数机制，寻找爬跨山羊和被爬跨山羊间的关联，并判定当前时刻爬跨和被爬跨对象的匹配帧数（TTL）是否符合发情行为标准，如果是，则对两只山羊分配统一的发情 ID，同时显示发情行为。

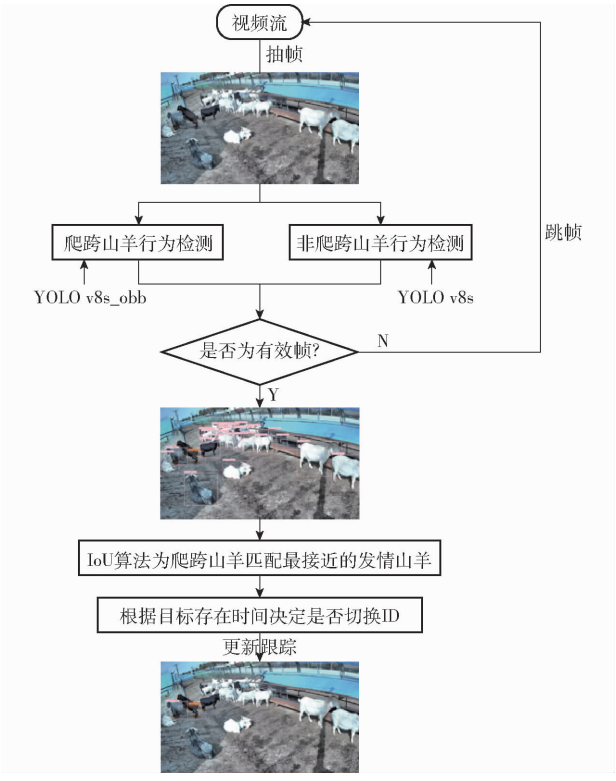


图 2　奶山羊发情行为识别和跟踪算法流程图

Fig. 2　Flowchart of estrus behavior detection and tracking algorithm for dairy goats

1.3　基于 YOLO v8 的奶山羊行为识别

YOLO 系列模型因其检测准确性高、实时性好在农业工程领域得到广泛应用，而 YOLO v8 汲取了先前版本的优点，具有检测精度高、推理速度快等优点[13]。YOLO v8 模型结构主要由主干网络、颈部网络和检测头 3 部分组成，其中，主干网络使用 C2f、CBS 和 SPPF 模块实现图像特征提取，颈部网络采用特征金字塔网络＋路径聚合网络（FPN＋PAN）结构进行多尺度特征融合，而检测头负责处理颈部网络输出的特征图，在特征图上定位和识别目标，生成目标边界框的位置和类别预测。YOLO v8 分为 s、m、l、x 等模型，主要的区别在于网络中特定位置的特征提取模块和卷积核的数量，综合考虑模型精度和模型内存占用量，以 YOLO v8s 作为基础构建奶山羊行为识别模型。但羊圈内背景杂乱，羊只之间会存在遮挡，且奶山羊具有个体间相似度高和差异

小的特点,易导致误检、漏检。为提高模型对奶山羊个体识别任务所需要重要信息的关注,在 YOLO v8s 的特征提取网络中引入 SimAM 无参注意力机制。此外,由于具有爬跨行为的奶山羊通常是后肢置于地面,前肢则放在另一只奶山羊身上,导致整体身体与地面呈较大的倾斜角度,为能更好定位和识别爬跨的羊只,在保证骨干网络结构一致的前提下,采用 YOLO v8s_obb 和 YOLO v8s 分别检测具有爬跨行为的羊只和四肢完全接触地面的非爬跨羊只。图 3 所示为奶山羊行为识别网络模型结构。

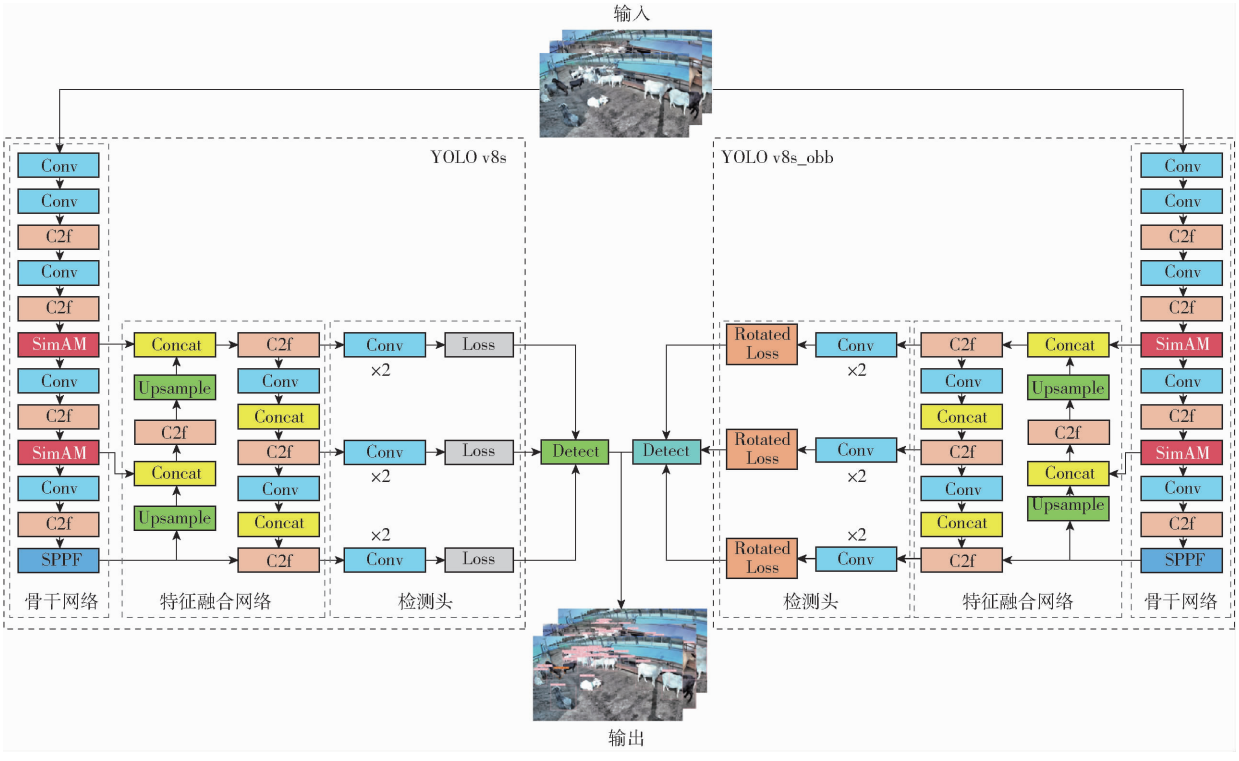


图 3 奶山羊行为识别网络模型结构图

Fig. 3 Network model structure diagram of dairy goat mounting behavior recognition

1.3.1 骨干网络中引入注意力机制

YOLO v8 骨干网络采用 CSPDarkNet53 网络结构,网络中的多层卷积结构使模型更好地学习到图像的深层语义特征,但多层堆叠卷积层在提取图像特征时通常会引入一些冗余特征,干扰模型对羊圈内奶山羊个体的正确识别,同时奶山羊个体之间的外貌特征相似度高,增加了准确识别复杂背景下羊圈内羊只的难度。为此,在不增加参数量的基础上,在 YOLO v8s 和 YOLO v8s_obb 的主干网络中引入无参数的 SimAM 注意力模块[14],抑制背景等非重要信息,提高模型对检测目标关键特征的关注能力。改进后的主干网络结构组成参数如表 1 所示。

SimAM 是一种无需增加额外参数便可从输入的特征图中生成一维通道和二维空间特征的权重,利用局部区域的均值和方差作为输入特征的衡量依据,构建能量函数,通过最小化能量函数,对输入特征图计算注意力权重,对重要特征信息赋予更大的权重,可有效过滤对空间和通道非重要特征信息,以增强模型特征学习能力。以特征图 F 为 SimAM 的输入,定义能量函数 E_{ijk} 为

表 1 加入 SimAM 的骨干网络结构参数

Tab. 1 Backbone network structure parameters with SimAM module

输入尺寸	操作	输出通道数	SimAM	步长
640×640×3	Conv 3×3	32	–	2
320×320×32	Conv 3×3	64	–	2
160×160×64	C2fBottleneck1	64	–	1
160×160×64	Conv 3×3	128	–	2
80×80×128	C2fBottleneck2	128	–	1
80×80×128	C2fBottleneck3	128	+	1
80×80×128	Conv 3×3	256	–	2
40×40×256	C2fBottleneck4	256	–	1
40×40×256	C2fBottleneck4	256	+	1
40×40×256	Conv 3×3	512	–	2
20×20×512	C2fBottleneck4	512	+	1
20×20×512	SPPF	512	–	1

注:SimAM 代表卷积注意力模块;“–”表示当前的网络层未使用该结构或策略;“+”表示当前网络层使用该结构或策略。

$$E_{ijk}=\frac{1}{2}(F_{ijk}-\mu)^2-\lambda F_{ijk}^2 \tag{1}$$

其中,F_{ijk}表示输入特征图 F 的第 i 通道、第 j 行、第 k

列的元素，μ 为特征图局部区域的均值，$\lambda=\frac{1}{\sigma^2+\varepsilon}$，$\lambda$ 为超参数，σ^2为特征图局部区域的方差，ε 为防止分母为零的小常数。通过对能量函数 E_{ijk}求最小值，得到优化后的注意力特征 A_{ijk}，即

$$A_{ijk}=\frac{\mu}{1-2\lambda} \tag{2}$$

将注意力特征 A_{ijk}与原始特征图逐元素相乘生成细化后的特征图 F'表达式为

$$F'=F\odot A_{ijk} \tag{3}$$

式中　⊙——逐元素乘法

将轻量级 SimAM 注意力模块集成到 YOLO v8s 和 YOLO v8s_obb 的骨干网络，获取通道与空间位置的三维权重，提升模型对奶山羊特征的感知能力并抑制噪声干扰，提高模型在自然场景下对奶山羊的识别能力。

1.3.2　旋转边界框表示

传统的动物爬跨行为检测常使用标准矩形边界框来表示目标对象，但矩形框无法准确刻画爬跨行为的方向性和形状特征，尤其是在爬跨行为表现出明显角度变化时，基于标准矩形框表示来计算交并比（IoU）会引入多余面积，导致检测精度下降。为此，引入了旋转边界框（Oriented bounding box，obb）表示，旋转边界框能够根据爬跨奶山羊的动作方向动态调整边界框角度，从而更精确地贴合目标实际形状。

如图 4 所示，基于旋转框的 IoU 计算方式能更准确地反映爬跨行为的真实边界范围，可有效减少无关区域的干扰，与传统矩形框相比，在复杂行为场景下表现出更高的可靠性，可提升模型对爬跨行为的检测精度。同时，旋转边界框表示通过精确标定爬跨行为的边界和角度信息，能够对奶山羊识别后的跟踪匹配提供更可靠的对象，在匹配过程中，跟踪

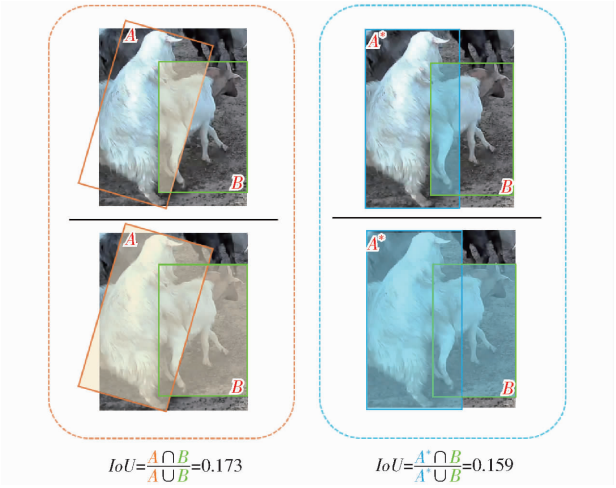


图 4　IoU 计算方法的对比

Fig. 4　Comparison of computing methods for IoU

算法能够利用旋转框提供的方向性信息，快速确定爬跨奶山羊与其潜在交互对象（即非爬跨奶山羊）之间的关联度，以提升匹配的准确性和计算效率。

1.3.3　奶山羊行为识别模型评价指标

使用精确率（P）、召回率（R）、平均精度均值（mAP）、平均精度（AP）、F1 分数、加权精确率（P'）、加权召回率（R'）、加权平均精度均值（$m_{\text{weighting}}$）、加权 F1 分数（$F_{1\text{weighting}}$）来评估奶山羊行为识别模型的性能。设置 $IoU\geqslant 0.5$ 的阈值，以保证正确检测。指标的计算式为

$$P'=w_{\text{mount}}P_{\text{mount}}+w_{\text{normal}}P_{\text{normal}} \tag{4}$$

$$R'=w_{\text{mount}}R_{\text{mount}}+w_{\text{normal}}R_{\text{normal}} \tag{5}$$

$$F_{1\text{weighting}}=w_{\text{mount}}F_{1\text{mount}}+w_{\text{normal}}F_{1\text{normal}} \tag{6}$$

$$m_{\text{weighting}}=w_{\text{mount}}A_{\text{mount}}+w_{\text{normal}}A_{\text{normal}} \tag{7}$$

其中，w_{mount}和 w_{normal}分别表示爬跨和非爬跨行为对应的加权系数；P_{mount}和 P_{normal}分别表示检测奶山羊爬跨和非爬跨行为的精确率；R_{mount}和 R_{normal}分别表示检测奶山羊爬跨和非爬跨行为的召回率；$F_{1\text{mount}}$和 $F_{1\text{normal}}$分别表示检测奶山羊爬跨和非爬跨行为的 F1 分数；A_{mount}和 A_{normal}分别表示检测奶山羊爬跨和非爬跨行为的平均精度。

1.4　基于 DeepSort 奶山羊发情行为跟踪

在奶山羊发情行为监测中，除识别发情行为的静态特征，还需在动态视频中对发情奶山羊进行轨迹跟踪，以准确定位爬跨和被爬跨的羊只。本研究以 DeepSort 算法为基础实现羊圈内羊只多目标跟踪。对于输入的视频帧图像，首先通过改进的 YOLO v8s 和 YOLO v8s_obb 网络识别出爬跨羊只和非爬跨羊只，获取行为类别、位置及边界框信息，并赋予每个目标唯一的 ID，如图 5a、5b 所示；接着，通过级联匹配获取关联度最大的发情组（即 IoU 最大的爬跨山羊和非爬跨山羊），卡尔曼滤波器基于上一帧的状态信息预测目标在下一时刻的位置，将当前检测到的目标与之前的运动轨迹进行匹配，实现发情羊只运动轨迹的跟踪。

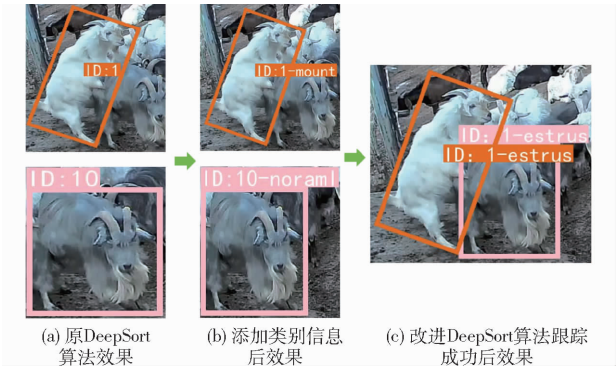


图 5　羊只目标的 ID 与状态信息

Fig. 5　Dairy goat target ID and status information

为提升匹配效果,增加第 2 轮交并比匹配对匹配失败检测框进行处理。改进后的匹配过程为:第 1 轮交并比匹配完成初步匹配,第 2 轮通过引入匹配帧数计数,延长失败检测框的存活帧数,当成功匹配的帧数存活数大于等于 5 帧,则确认这两只山羊产生了发情行为,跟踪算法将重新为这对发情山羊分配统一的发情 ID 和行为类别,并显示在视频中,如图 5c 所示。改进后算法流程如图 6 所示,其中虚线框部分为算法改进内容。轨迹初始化为未确认态,满足连续 5 帧都成功匹配,将未确认态转化为确认态。

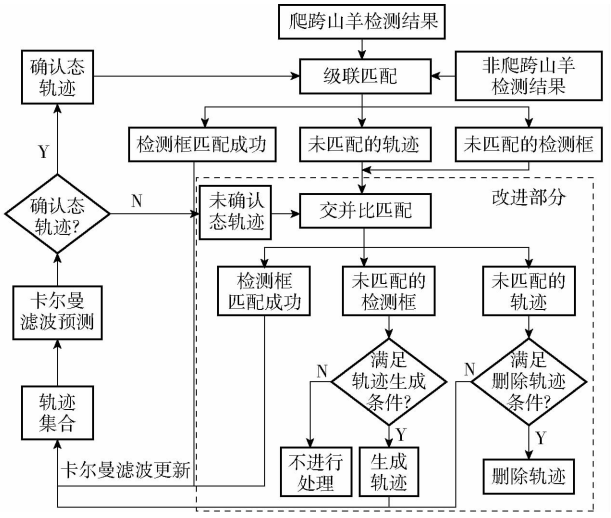


图 6 改进后奶山羊发情行为跟踪算法流程图

Fig. 6 Improved algorithm flowchart of tracking algorithm for estrus behavior of dairy goats

1.4.1 引入匹配帧数 TTL(Time to live)计数

进行羊圈的发情奶山羊跟踪时,由于羊圈的羊群经常产生运动或遮挡现象,随着视频帧的增长,检测结果难以与原轨迹匹配,导致未匹配的检测结果生成新轨迹,同一目标易被分配不同的 ID,造成羊只身上的 ID 发生偏移现象。针对此问题,在 DeepSort 的匹配算法中引入匹配帧数计数 TTL(Time to live),用于统计爬跨奶山羊与可能产生发情行为的奶山羊被跟踪的帧数。在匹配跟踪时,每只识别到的爬跨奶山羊都被分配了唯一的 ID,基于 IoU 计算爬跨奶山羊与非爬跨奶山羊的边界框重叠度,以最大 IoU 重叠对应的爬跨奶山羊与非爬跨奶山羊为候选奶山羊发情组,在后续帧中候选奶山羊发情组的匹配上则实现 TTL 加 1,当 TTL 大于等于 5 时,则认定此对奶山羊产生了发情行为,并为这两只奶山羊统一分配相同的发情 ID,如图 7 所示;如果后续帧中有新的非爬跨奶山羊被候选为同一爬跨奶山羊的发情对象,统计爬跨奶山羊与该新的非爬跨奶山羊被跟踪的帧数 TTL,比较统计的帧数可确定爬跨奶山羊的发情对象,若 TTL 大于等于 5 则表示爬跨奶山羊重新找到一个新的发情对象,并为该爬跨奶山羊和被爬跨的奶山羊重新分配一个新的 ID,否则,发情对像保持原始匹配结果,如图 8 所示。由图 8 示例可知,在第 5 帧中,算法为检测到的发情山羊统一分配新的发情 ID,在第 10 帧中,算法通过计算最大 IoU 找到了新的候选非爬跨山羊(ID 为 2),在最后一帧(第 14 帧)中,由于 ID 为 2 的非爬跨山羊并没有保持 5 帧的最大 IoU,因此算法认为不是该爬跨山羊的发情对象。

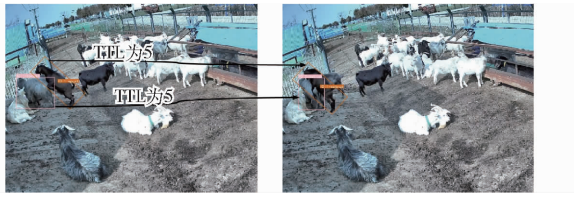
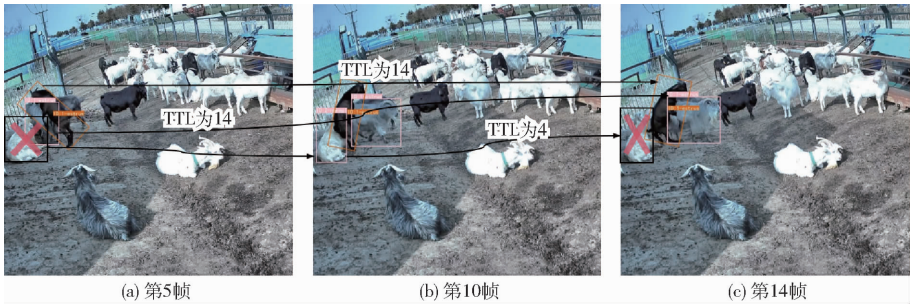


图 7 爬跨山羊与非爬跨山羊匹配过程示例

Fig. 7 Example of matching process between mounting and non-mounting dairy goats



(a) 第5帧 (b) 第10帧 (c) 第14帧

图 8 爬跨山羊发情对象匹配过程示例

Fig. 8 Example of estrus object matching process for climbing dairy goat

1.4.2 奶山羊发情行为跟踪算法评价指标

奶山羊发情跟踪性能评估选取高阶跟踪准确性 HOTA、跟踪准确度 MOTA、目标 ID 切换次数 IDS、跟踪精度 MOTP 和正确目标跟踪比例 IDF1 作为评价指标。HOTA 可以综合衡量检测和关联精度,能全面评估跟踪系统的整体性能;MOTA 可以衡量跟踪错误,包括误检、漏检和 ID 切换对系统整体精度的影响;MOTP 可衡量跟踪目标与其匹配

位置之间的平均距离，反映空间定位精度；IDF1 用于衡量跟踪过程中 ID 一致性的准确性，即目标的正确关联性；其中 HOTA、MOTA、IDF1 的数值越高反映算法越好，MOTP、IDS 的数值越低反映算法越好。

2　结果与分析

2.1　试验平台

针对自然场景下活动区域中奶山羊进行发情行为的检测和跟踪试验。试验在搭载 2.30 GHz Intel Xeon(R) W－2123 处理器、128 GB 内存、4 TB 硬盘的 Ubuntu 18.04 平台上进行。使用的 GPU 为 NVIDIA GeForce RTX 2080Ti。使用的编程语言为 Python 3.8，深度学习框架为 PyTorch 1.10.0。在模型训练过程中，采用预训练模型的迁移学习和参数微调来加速收敛。

2.2　奶山羊目标检测结果分析

2.2.1　基准模型的选择

为验证基准模型 YOLO v8 选取的有效性，选取了 SSD[15]、Faster R－CNN[16]、EfficientDet－D1[17]、YOLO v5[18]、YOLO v7[19]、YOLO v10[20] 及 YOLO 11[21] 等 7 种主流目标检测模型，对羊只的爬跨和非爬跨行为识别进行了性能对比试验，试验中均以标准矩形边界框表示目标对象，对比结果如表 2 所示。由表 2 可知，YOLO v8 的 mAP 值达 88.12%，明显优于 SSD、Faster R－CNN、EfficientDet－D1 等模型。相比于 YOLO 早期版本，较 YOLO v5 和 YOLO v7 分别提高了 2.38、5.48 个百分点。YOLO v8 的 F1 分数和 mAP 略低于 YOLO v10 和 YOLO 11，但 YOLO v8 计算复杂度显著低于 YOLO v10 和 YOLO 11。综合考虑上述各模型的识别性能和计算复杂度，本研究以 YOLO v8 作为基准模型进行后续的奶山羊发情行为识别研究。

表 2　不同目标识别模型的性能对比

Tab.2　Performance comparison of different object detection models

模型	*P*/%	*R*/%	F1 分数/%	mAP/%	模型内存占用量/MB	浮点运算量
SSD	85.00	73.40	78.10	80.64	92.20	2.74×10^{11}
Faster R－CNN	82.20	61.80	71.40	71.74	139.10	1.70×10^{11}
EfficientDet－D1	83.30	68.90	76.57	75.54	17.70	1.15×10^{10}
YOLO v7	85.20	84.50	80.17	82.64	71.40	1.03×10^{11}
YOLO v5	87.90	76.00	82.77	85.74	40.20	4.79×10^{10}
YOLO v8	90.52	84.30	86.10	88.12	23.20	3.20×10^{10}
YOLO v10	90.33	82.90	87.13	88.71	24.40	5.91×10^{10}
YOLO 11	91.00	84.25	87.56	89.24	20.10	6.80×10^{10}

为验证采用 YOLO v8_obb 识别爬跨行为的有效性，采用 YOLO v8 和 YOLO v8_obb 对奶山羊爬跨行为进行识别，结果如表 3 所示，支持旋转框检测的 YOLO v8_obb 在爬跨行为检测上的各项指标均优于标准 YOLO v8，可以对奶山羊爬跨行为进行较好识别。此外，实际场景图像中，爬跨行为的数量要远小于非爬跨行为，如果基于 YOLO v8_obb 同时识别奶山羊爬跨和非爬跨行为，由于 YOLO v8_obb 训练过程中需边界框角度参数，此时大量角度为 0 的非爬跨行为边界框参数会严重干扰爬跨行为的边界框训练效果，导致最后混合在一起训练结果为非爬跨行为的边界框达不到非旋转效果，同时爬跨行为的边界框达不到精准定位奶山羊的效果，如图 9 所示。为解决这一问题，可在训练集填充大量具有爬跨行为的奶山羊数据，保证行为类别的均衡化，但这将显著增加数据集制作和计算成本。然而，YOLO v8_obb 和 YOLO v8 的骨干网络设计一致，即训练后特征提取部分的权重参数基本一致，其差异主要在检测头部分，即旋转框和非旋转框的损失函数设计不同。在实际部署中，将预训练得到的主干网络权重保存并加载，然后根据不同的任务需求，分别加载 YOLO v8_obb 和 YOLO v8 各自的检测头权重，实现对奶山羊爬跨行为和非爬跨行为的识别。这

表 3　YOLO v8 和 YOLO v8_obb 的爬跨行为识别性能

Tab.3　Performance of YOLO v8 and YOLO v8_obb for mounting behavior detection　%

模型	*P*	*R*	F1 分数	A_{mount}
YOLO v8_obb	95.10	89.70	93.50	94.38
YOLO v8	91.73	85.80	86.30	88.92

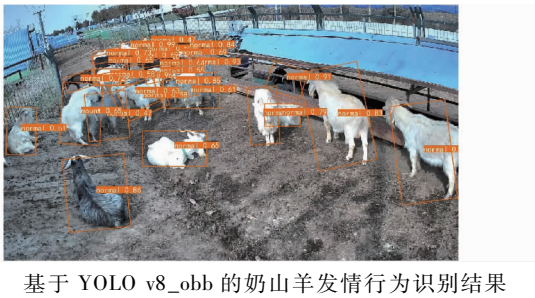


图 9　基于 YOLO v8_obb 的奶山羊发情行为识别结果

Fig.9　Detection results of estrus behavior of dairy goats based on YOLO v8_obb

种资源共享的策略能够显著降低系统对硬件资源的依赖,使得模型在边缘设备、移动设备或其他资源受限的环境中可运行。综上,本研究采用 YOLO v8_obb 检测爬跨行为奶山羊,YOLO v8 检测非爬跨行为奶山羊的方法,以实现对奶山羊发情行为的有效识别。

2.2.2 改进前后 YOLO v8s 模型性能评估

为了验证 SimAM 注意力机制的有效性,开展了 SimAM 集成前后的 YOLO v8s 和 YOLO v8s_obb 模型对比试验,改进前后的 YOLO v8s 和 YOLO v8s_obb 模型对奶山羊发情行为的识别效果如图 10(图中红色虚线框代表原模型的漏检部分,蓝色虚线框代表原模型误检部分)所示。由图 10 可知,对于稀疏、颜色区别明显的奶山羊,改进前后均有较好的检测效果,但 YOLO v8s 对遮挡羊只存在一些漏检现象,YOLO v8s_obb 对爬跨行为的羊只有误检现象。而改进后的模型增强了对羊只特征的感知能力,提升了模型对遮挡羊只的识别效果。



图 10 改进前后的 YOLO v8s 和 YOLO v8s_obb 模型对奶山羊的检测效果

Fig. 10 Detection effects of dairy goats before and after improvement of YOLO v8s and YOLO v8s_obb

改进前后的 YOLO v8s 和 YOLO v8s_obb 模型对奶山羊识别的精确率、召回率、F1 分数以及平均精度如表 4 所示。由表 4 可知,相比于原 YOLO v8s_obb 模型,改进的 YOLO v8s_obb 对具有爬跨行为的奶山羊的平均精度和精确率分别提升了 0.90、2.10 个百分点;相比于原 YOLO v8s 模型,改进的 YOLO v8s 对具有非爬跨行为的奶山羊的平均精度和精确率分别提升了 2.64、1.70 个百分点。进一步证明了加入了 SimAM 注意力模块后的模型对于本研究中的奶山羊行为的识别有着更好的效果。

表 4 改进前后 YOLO v8s 和 YOLO v8s_obb 模型的识别性能对比

Tab. 4 Comparison of recognition performance of YOLO v8s and YOLO v8s_obb models before and after improvement %

模型	类别	*P*	*R*	F1 分数	AP
YOLO v8s_obb	爬跨	95.10	89.70	93.50	94.38
YOLO v8s	非爬跨	89.30	82.80	85.90	87.32
改进 YOLO v8s_obb	爬跨	97.20	90.20	93.16	95.28
改进 YOLO v8s	非爬跨	91.00	84.33	87.47	89.96

2.2.3 不同检测模型的性能对比

为验证本文奶山羊发情行为识别模型:改进 YOLO v8s + 改进 YOLO v8s_obb 的有效性,在同一数据集上,将其与 SSD、Faster R - CNN、EfficientDet - D1、YOLO v5m + YOLO v5m_obb、YOLO v7 + YOLO v7_obb 及 YOLO v8m + YOLO v8m_obb 这 6 类目标检测算法进行对比试验,对比试验结果如表 5 所示。由表 5 可知,基于标准矩形框和旋转矩形框的 YOLO 算法识别奶山羊的爬跨行为和非爬跨行为的精确率、召回率、F1 分数及 AP 均高于基于标准矩形框的 SSD、Faster R - CNN、EfficientDet - D1。为确保结果的公平性与可比性,设计了一种加权评价方法,根据数据集中爬跨行为和非爬跨行为的类别比例(1:9)对不支持旋转框模型的各项指标(精确率、召回率、F1 分数及 AP)进行加权计算;对支持旋转框的模型,分别采用其旋转框版本和标准矩形框版本的各项指标进行加权计算;最后,各类算法在统一的加权评价指标体系下进行试验对比和分析,对比试验结果如表 6 所示。

表 5 不同目标检测模型对爬跨行为和非爬跨行为的识别性能

Tab. 5 Detection performance of different object detection models for mounting behavior and no mounting behavior %

模型	爬跨行为				非爬跨行为			
	P	*R*	F1 分数	AP	*P*	*R*	F1 分数	AP
SSD	84.93	71.50	77.25	79.52	85.07	75.30	78.95	81.76
Faster R - CNN	81.33	55.12	70.80	68.30	83.07	68.48	72.00	75.18
EfficientDet - D1	81.20	63.52	74.71	72.37	85.40	74.28	78.43	78.71
YOLO v7	–	–	–	–	87.28	88.57	83.20	86.23
YOLO v7_obb	93.30	87.39	88.39	89.51	–	–	–	–
YOLO v5m	–	–	–	–	89.62	77.50	85.64	87.13
YOLO v5m_obb	92.74	80.33	87.45	90.70	–	–	–	–
YOLO v8m	–	–	–	–	91.16	85.63	88.13	89.14
YOLO v8m_obb	95.12	88.24	92.61	93.91	–	–	–	–
改进 YOLO v8s	–	–	–	–	91.00	84.33	87.47	89.96
改进 YOLO v8s_obb	97.20	90.20	93.16	95.28	–	–	–	–

注:“ - ”表示该算法不支持该类别的检测,未计算相关指标。

表 6　不同模型加权指标后的性能对比

Tab. 6　Performance comparison of different methods after weighting metrics　%

模型	P'	R'	$F_{1\text{weighting}}$	$m_{\text{weighting}}$
SSD	85.06	74.92	78.78	81.54
Faster R－CNN	82.90	67.14	71.88	74.49
EfficientDet－D1	84.98	73.20	78.06	78.08
YOLO v7 + YOLO v7_obb	87.88	88.45	83.72	86.56
YOLO v5m + YOLO v5m_obb	89.93	77.78	85.82	87.49
YOLO v8m + YOLO v8m_obb	91.56	85.89	88.58	89.62
改进 YOLO v8s + 改进 YOLO v8s_obb	91.62	84.92	88.04	90.49

表 6 的试验结果表明，本研究提出的方法具有最高的加权精确率和加权 mAP，且具有较小模型内存占用量，相较于 SSD、Faster R－CNN、YOLO v5m + YOLO v5m_obb、EfficientDet－D1、YOLO v7 + YOLO v7_obb 及 YOLO v8m + YOLO v8m_obb，加权精确率分别提高了 6.56、8.72、6.64、3.74、1.69、0.06 个百分点，加权 mAP 分别上升了 8.95、16.00、12.41、3.93、3.00、0.87 个百分点；与 EfficientDet－D1 相比，单个模型内存占用量虽大了 5.5MB，但所提出的方法在加权精确率、加权召回率、加权 F1 分数和加权 mAP 方面分别高出 6.64、11.72、9.98、12.41 个百分点。YOLO v8m + YOLO v8m_obb 虽具有较高的加权召回率和加权 F1 分数，但其加权精确率、加权 mAP 均低于所提出的模型，且模型尺寸较大，不利于设备端的部署和应用。与 YOLO v8m 相比，YOLO v8s 模型内存占用量缩小了 53.2%，浮点运算量降低了 59.3%，且在加权精确率和加权 mAP 指标上表现出良好的性能。综上，本文改进的奶山羊发情行为识别方法效果优于其他 6 种算法，能够准确地识别自然场景中奶山羊的爬跨行为和非爬跨行为，为后续的发情行为跟踪提供更准确的信息。

2.3　奶山羊发情行为跟踪结果分析

为评估本文所提出奶山羊发情行为跟踪方法的性能，随机抽取了 10 段奶山羊发情视频进行跟踪试验，跟踪结果如图 11 所示。在图 11 中，视频 03 和视频 04 分别拍摄于晴天的上午和阴天的下午，且发情的奶山羊在视频中存在一定的遮挡。从图 11 可以看出引入匹配帧数统计的多目标跟踪方法可以准确地跟踪奶山羊之间的发情行为，如视频中当两只奶山羊正处于发情状态时，羊群内运动的羊只经过发情奶山羊前面会引起遮挡，本文方法仍能准确匹配跟踪两只正处于发情态的羊只，没有出现 ID 错误切换和检测框错位的情况。



(a) 视频03的跟踪结果

(b) 视频04的跟踪结果

图 11　发情奶山羊的视频跟踪结果

Fig. 11　Video tracking results of estrus dairy goats

为了进一步验证改进算法的有效性，使用 DarkLabel 对上述 10 段视频中爬跨行为进行逐帧标注，并与其他 4 种多目标跟踪算法 DeepSort[22]、BotSort[23]、StrongSort[24] 和 ByteTrack[25] 进行比较试验，试验结果如表 7 所示。通过对比，本文算法的 HOTA 和 IDF1 在 5 种算法中最高，分别为 72.4% 和 80.3%，表明本文算法具有较好的连续目标跟踪精度和稳定性。在目标身份错误切换次数方面，该方法切换次数为 46 次，比 ByteTrack 算法减少了 58.6%，

表 7　不同算法跟踪结果对比

Tab. 7　Comparison of tracking results of different algorithms

算法	HOTA/%	MOTA/%	MOTP/%	IDF1/%	IDS/次
DeepSort	63.4	57.5	82.9	69.1	79
BotSort	71.7	79.2	86.1	80.2	91
StrongSort	70.6	79.7	85.3	77.8	122
ByteTrack	68.3	74.9	82.9	80.0	111
改进算法	72.4	79.1	85.9	80.3	46

这对于实现自然场景下奶山羊发情行为的跟踪具有重要意义。

为进一步验证本研究跟踪方法的有效性，在一个视频片段中抽取了前后两帧的跟踪结果，并与 DeepSort、BotSort、StrongSort 和 ByteTrack 的结果进行比较，如图 12（红色虚线框内的目标表示没有被跟踪或错误跟踪，蓝色虚线框表示由于目标 ID 切换导致的错误跟踪）所示。在图 12a 中，DeepSort 采用级联匹配并结合外观特征关联虽成功跟踪到了左边 ID 为 1 的发情奶山羊，但右边羊群场景较为复杂，导致误跟踪了 ID 为 2 的奶山羊，实际上这两只山羊并未产生发情行为；图 12b 中 BotSort 通过在跟踪过程中结合重识别和相机运动补偿模块来减少错误跟踪的情况，虽在前一帧成功识别到了 ID 为 1 的发情奶山羊，但在后一帧却丢失了目标；StrongSort 在两帧中都成功跟踪到了唯一的发情奶山羊，但在第 2 帧却发生了 ID 的切换，由#1 切换到了#2，表明其未能保证连续帧中身份的一致性，如图 12c 所示；图 12d 中的 ByteTrack 通过关联每个检测框并保留低置信度的检测框来处理重叠或遮挡的对象，有助于减少漏检和轨迹碎片化，仍存在前一帧中没有成功跟踪到发情行为的不足。本文通过旋转框 IoU 的计算方法和引入匹配帧数统计机制，增强了持续跟踪发情羊只的能力，能更好地适应奶山羊的位置移动和遮挡变化，获得了良好的跟踪性能。

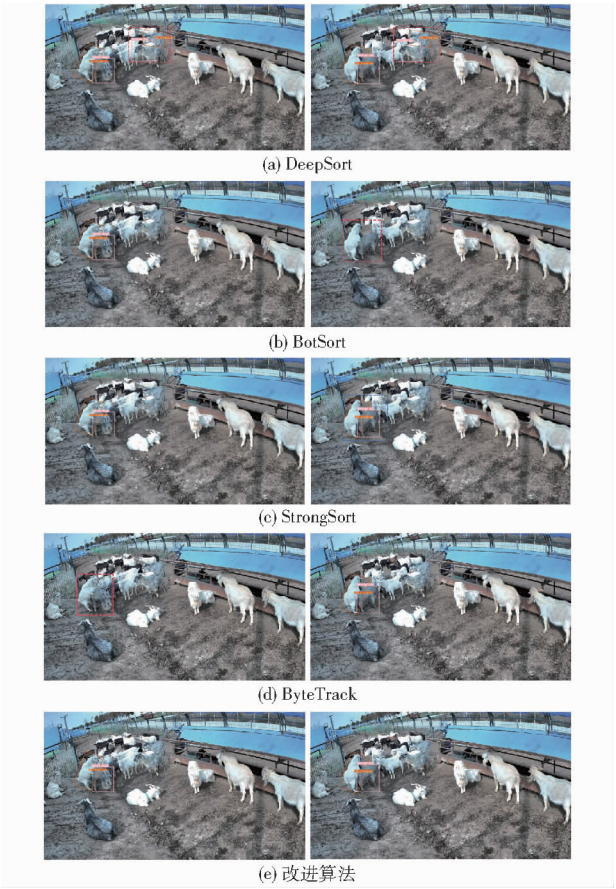


图 12 不同跟踪算法的跟踪结果对比示例

Fig. 12 Comparative example of tracking results for different tracking algorithms

3 结论

（1）及时准确地识别跟踪发情的奶山羊对于实施有效的繁殖管理策略至关重要，本研究采用机器视觉方法提供可视化的跟踪结果，可以直观地观察到奶山羊的发情行为变化，从而帮助畜牧从业者在最佳时间安排交配，为奶山羊的精准繁殖计划和管理决策提供技术支持。

（2）在 YOLO v8s 和 YOLO v8s_obb 的骨干网络中融合了 SimAM 注意力模块，并采用标准目标框和旋转框分别表示非爬跨奶山羊和爬跨奶山羊，提高了奶山羊行为检测精度。改进后的 YOLO v8s 和 YOLO v8s_obb 对非爬跨行为和爬跨行为的平均精度分别提高了 2.64、0.90 个百分点。

（3）将检测的羊只行为信息引入跟踪算法，分配爬跨行为羊只的 ID，基于 IoU 和匹配帧数计数机制，寻找爬跨羊只与被爬跨羊只检测框之间的关联性，实现了奶山羊发情行为的准确跟踪。本文跟踪方法在所有测试算法中取得了最高的 HOTA（72.4%）和 IDF1（80.3%），并有着最低的 IDS（46 次）。

参 考 文 献

[1] 宁纪锋，张静，杨蜀秦，等. 基于 DiMP 的类特定养殖奶山羊跟踪方法[J]. 农业机械学报，2023，54(6)：280－286，400.
NING Jifeng, ZHANG Jing, YANG Shuqin, et al. Tracking of class-specific dairy goats based on DiMP[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 280－286, 400. (in Chinese)

[2] SILPER B F, ROBLES I, MADUREIRA A M L, et al. Automated and visual measurements of estrous behavior and their sources of variation in Holstein heifers. I: walking activity and behavior frequency[J]. Theriogenology, 2015, 84(2): 312－320.

[3] 王旭贞. 母羊的发情鉴定技术[J]. 养殖与饲料，2018(12)：24－25.
WANG Xuzhen. Estrus identification techniques for ewes[J]. Culture and Feed, 2018(12): 24－25. (in Chinese)

[4] WANG Z, HUA Z, WEN Y, et al. E－YOLO: recognition of estrus cow based on improved YOLOv8n model[J]. Expert Systems with Applications, 2024, 238: 122212.

[5] 刘忠超，何东健. 基于卷积神经网络的奶牛发情行为识别方法[J]. 农业机械学报，2019，50(7)：186－193.

LIU Zhongchao, HE Dongjian. Cow estrus behavior recognition method based on convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019,50(7):186 – 193. (in Chinese)

[6] WANG J, ZHANG Y, WANG J, et al. Using machine-learning technique for estrus onset detection in dairy cows from acceleration and location data acquired by a neck-tag[J]. Biosystems Engineering, 2022, 214: 193 – 206.

[7] WANG J, CHEN H, WANG J, et al. Identification of oestrus cows based on vocalisation characteristics and machine learning technique using a dual-channel-equipped acoustic tag[J]. Animal, 2023, 17(6): 100811.

[8] 汪开英, 赵晓洋, 何勇. 畜禽行为及生理信息的无损监测技术研究进展[J]. 农业工程学报, 2017, 33(20): 197 – 209.
WANG Kaiying, ZHAO Xiaoyang, HE Yong. Research progress on nondestructive monitoring technology of animal behavior and physiological information[J]. Transactions of the CSAE, 2017, 33(20): 197 – 209. (in Chinese)

[9] YU L, PU Y, CEN H, et al. A lightweight neural network-based method for detecting estrus behavior in ewes[J]. Agriculture, 2022, 12(8): 1207.

[10] 王少华,何东健.基于改进 YOLO v3 模型的奶牛发情行为识别研究[J].农业机械学报,2021,52(7): 141 – 150.
WANG Shaohua, HE Dongjian. Research on cow estrus behavior recognition based on improved YOLO v3 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021,52(7): 141 – 150. (in Chinese)

[11] MAR C C, ZIN T T, TIN P, et al. Cow detection and tracking system utilizing multi-feature tracking algorithm[J]. Scientific Reports, 2023, 13(1): 17423.

[12] HUANG X, HU Z, QIAO Y, et al. Deep learning-based cow tail detection and tracking for precision livestock farming[J]. IEEE/ASME Transactions on Mechatronics, 2023, 28:1213 – 1221.

[13] GLENN J, AYUSH C, JING Q. Ultralytics YOLO(Version 8.0.0)[CP/OL]. https://github.com/ultralytics/ultralytics.

[14] YANG L, ZHANG R Y, LI L, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks[J]. International Conference on Machine Learning,2021,139:11863 – 11874.

[15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multiBox detector[M]// Cham: Springer International Publishing, 2016: 21 – 37.

[16] REN S, HE K, GIRSHICK R, et al. Faster R – CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137 – 1149.

[17] TAN M, PANG R, LE Q V. EfficientDet: scalable and efficient object detection[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020: 10778 – 1078.

[18] GLENN J. Ultralytics/YOLOv5: v6.0[CP/OL]. https://github.com/ultralytics/yolov5.

[19] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2023: 7464 – 7475.

[20] WANG A, CHEN H, LIU L, et al. YOLOv10: real-time end-to-end object detection[C]// arXiv:2405.14458,2024.

[21] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[J]. arXiv:2410.17725,2024.

[22] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with deep association metric[C]// IEEE International Conference on Image Processing, 2017:3645 – 3649.

[23] DU Y, ZHAO Z, SONG Y, et al. StrongSORT: make DeepSORT great again[J]. IEEE Transactions on Multimedia, 2023, 25: 8725 – 8737.

[24] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT – SORT: robust associations multi-pedestrian tracking[EB/OL]. arXiv: 2206.14651, 2022.

[25] ZHANG Y, SUN P, JIANG Y, et al. ByteTrack: multi-object tracking by associating every detection box[M]// AVIDAN S, BROSTOW G, CISSÉ M, et al. Computer Vision-ECCV 2022. Cham: Springer Nature Switzerland, 2022: 1 – 21.