doi:10.6041/j.issn.1000-1298.2024.05.028

基于改进 YOLO v8 的牛只行为识别与跟踪方法

付辰伏[1,2]　任力生[1,2]　王　芳[1,2]

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

摘要:随着我国畜牧业的快速发展,牛只养殖由分散性养殖逐渐向精准化养殖转变。针对分散养殖中农户无法对每头牛只健康状况给予足够关注的问题,通过分析牛只行为模式结合视觉方向特征,设计了综合管理方法来准确识别和跟踪牛只行为。首先,采用改进 YOLO v8 算法对牛只进行目标监测,其中,在 Backbone 和 Neck 端使用 C2f - faster 结构,增强模型特征提取能力;引入上采样算子 CARAFE,拓宽感受视野进行数据特征融合;针对牛只幼仔检测加入 BiFormer 注意力机制,以识别牛只小面积特征;更换动态目标检测头 DyHead,融合尺度、空间和任务感知;然后,使用 Focal SIoU 函数,解决正负样本分配不均衡和 CIoU 局限性的问题。最后,将 YOLO v8 检测到的行为类别信息引入 BoTSORT 算法中,实现在复杂场景下牛只多目标行为识别跟踪。实验结果表明,提出的 FBCD - YOLO v8n(FasterNet、BiFormer、CARAFE、DyHead)模型在牛只行为数据集上,相比较 YOLO v5n、YOLO v7tiny 和原 YOLO v8n 模型的 mAP@0.5 分别提升 3.4、3.1、2.4 个百分点,尤其牛只回舔行为识别平均精度提高 7.4 个百分点。跟踪方面,BoTSORT 算法的 MOTA 为 96.1%,MOTP 为 78.6%,IDF1 为 98.0%,HOTA 为 78.9%;与 ByteTrack、StrongSORT 算法比,MOTA 和 IDF1 显著提升,跟踪效果良好。研究表明,在牛舍养殖环境下,本研究构建的多目标牛只行为识别跟踪系统,可有效帮助农户监测牛只行为,为牛只的自动化精准养殖提供技术支持。

关键词:牛只;目标监测;行为识别;多目标跟踪;YOLO v8;BoTSORT

中图分类号:TP391.4　　**文献标识码**:A　　**文章编号**:1000-1298(2024)05-0290-12　　**OSID**:

Method for Cattle Behavior Recognition and Tracking Based on Improved YOLO v8

FU Chenfu[1,2]　REN Lisheng[1,2]　WANG Fang[1,2]

(1. *College of Information Science and Technology*, *Hebei Agricultural University*, *Baoding* 071001, *China*
2. *Hebei Key Laboratory of Agricultural Big Data*, *Baoding* 071001, *China*)

Abstract: With the rapid development of animal husbandry in China, the transition from farmers' dispersed cattle breeding to precision husbandry has become increasingly important. Efficient management of breeding, behavior monitoring, disease prevention, and health assurance pose significant challenges. Traditionally, farmers have struggled to provide adequate attention to each cow. To address these challenges, a comprehensive approach was developed that accurately identified and tracked cattle behavior by analyzing behavior patterns and visual characteristics. Firstly, the improved YOLO v8 algorithm was employed for cattle target detection. The model's feature extraction capabilities were enhanced by incorporating the C2f - faster structure into the Backbone and Neck. The upsampling operator CARAFE was introduced to expand the perception field for data feature fusion. To identify small area characteristics of young cattle, the BiFormer attention mechanism was integrated into the detection process, replacing the dynamic target detection head DyHead. This allowed to effectively integrate scale, space, and task perception. Furthermore, the issue of the uneven distribution of positive and negative samples and the limitations of CIoU was addressed by utilizing the Focal SIoU function. Finally, the behavior category information detected by YOLO v8 was incorporated into the BoTSORT algorithm to

收稿日期:2023 - 09 - 18　修回日期:2023 - 11 - 08

基金项目:河北省省级科技计划项目(19220119D)

作者简介:付辰伏(1999—),男,硕士生,主要从事人工智能和图像处理研究,E-mail:2972212942@qq.com

通信作者:王芳(1971—),女,教授,博士,主要从事计算机网络、智能信息处理和农业信息化研究,E-mail:8335901@qq.com

enable multi-target behavior recognition and tracking in complicated situations. The experiments demonstrated significant performance improvements. The proposed FBCD - YOLO v8n model outperformed both the YOLO v5n, YOLO v7tiny, and the original YOLO v8n models, with an increase of 3.4 percentage points, 3.1 percentage points, and 2.4 percentage points in mAP@0.5, respectively, on the bovine behavior dataset. Notably, the accuracy of bovine back licking behavior recognition was increased by 7.4 percentage points. Regarding tracking, the BoTSORT algorithm achieved an MOTA of 96.1%, MOTP of 78.6%, HOTA of 78.9%, and IDF1 of 98.0%. Compared with ByteTrack and StrongSORT algorithms, the proposed method of MOTA and IDF1 scores demonstrated significant tracking improvements. This research demonstrated that the multi-objective cattle behavior recognition and tracking system developed can provide effective assistance to farmers in monitoring cattle behavior within the cattle barn environment. It offered crucial technical support for automated and precise cattle breeding.

Key words: cattle; object detection; behavior recognition; multi-object tracking; YOLO v8; BoTSORT

0　引言

牛养殖业中,牛只的行为与其健康状态之间存在密切关联[1]。牛只的行为可以作为反映牛只健康状况的重要指标,通过观察和分析牛只的行为,可以及时发现潜在的健康问题或异常情况,如牛只高频次的回舔身体可能原因有母牛产犊、调节体温、牛只身体表皮有蜱虫等,由此来警惕牛只产犊、牛舍温度过高、患有皮肤病、牛舍环境脏乱等状况[2]。目前,通过人工监测的方法观察牛只行为,需要消耗大量的时间、金钱和人力,无法实现个体化和精细化管理。为了解决目前存在的问题,计算机视觉作为智慧农业领域的热门技术,可以长时间对牛只进行监控,代替人类视觉对牛只图像进行分析,以此判断牛只是否健康[3]。

目前国内外学者针对牛只行为识别跟踪展开了大量研究,主要分为两方面:针对牛只行为的识别,利用 YOLO 系列模型目标检测识别牛只的日常行为;关于牛只的跟踪,大部分学者采用多目标跟踪(MOT)算法,通过给牛只分配身份编号(Identity, ID),来跟踪牛只的运动轨迹。

目标检测方面,因为牛的花纹特征较为明显,个体较大,目标检测较为容易。白强等[4]提出了一种改进的 YOLO v5s 奶牛多尺度行为识别方法,通过改进 YOLO v5s 的网络结构,添加 SE 注意力机制优化检测器,提取奶牛特征,识别奶牛站立、喝水、行走、躺卧等行为。王政等[5]融合 YOLO v5n 与通道剪枝算法,对主干的特征提取网络进行修剪,构建轻量化网络模型结构,提高检测速度。实验结果表明其模型在复杂环境下仍能对奶牛的发情行为保持较高的识别率。WANG 等[6]在 YOLO v5 模型上引入基于空洞空间金字塔池化和 C3GC - 3 特征增强模块,实现了高层语义信息的多尺度融合和在密集奶牛图像中的特征提取能力,并结合 CIoU 损失函数,提高奶牛爬跨行为检测精度。

跟踪方面,牛只移动速度较慢,相对静态目标更容易被识别和跟踪。李琦等[7]设计了一种跟踪草原牛的方法,其融合 YOLO v3[8]与 DeepSORT 跟踪算法,结合 PID 算法控制云台,用摄像头控制跟踪草原牛,实现了对草原牛的跟踪。CHO 等[9]采用 YOLO v5 和 DeepSORT 算法[10-11]将其应用于黑牛的识别与跟踪上,并发现黑牛跟踪时都具有相同或相似的纹理特征,即使人眼也不容易识别。因此,所提出的系统仍然存在一些 ID 切换和虚假检测问题需要解决。涂淑琴等[12]为改善 DeepSORT 算法引发的猪只重叠与遮挡造成的猪只身份编号频繁跳变,改进了 DeepSORT 算法的轨迹匹配过程,降低了 ID 错误频率。涂淑琴等[13]还提出了改进 ByteTrack 算法,利用 YOLOX[14]对猪进行识别,设计并实现 BYTE 数据关联的轨迹插值后处理策略,相比 DeepSORT 算法性能显著提升。SU 等[15]比较了一些先进 MOT 算法,如 DeepSORT、StrongSORT 和定制的轻量级跟踪算法,提出了将改进的 StrongSORT 应用于黑牛身上,发现 YOLO v7 和 StrongSORT 适用于黑牛发情行为的跟踪,同时也可以用于检测牛只其他行为,如跛行和爬跨等行为。

本文采用 YOLO 系列中最先进的 YOLO v8 算法和前沿的 BoTSORT 算法。在 YOLO v8 基础上通过替换特征提取模块、融合注意力机制、更换上采样、添加动态目标检测器、更改损失函数等对 YOLO v8 进行改进,提高牛只行为的识别率。然后,对比多个 MOT 算法如 ByteTrack[16]、StrongSORT[17]和 BoTSORT[18]等选取最高跟踪率的 BoTSORT 作为牛只的跟踪算法。最后将检测到的牛只行为信息引入跟踪算法,输出每头牛只的 ID,同时显示相应行为,在白天、夜晚、雨天、风沙等复杂天气下实现在牛舍养殖环境中对多头牛只行为的识别和跟踪。

1 实验数据与方法

1.1 数据采集

本研究视频采集于河北省邢台市新河县某小型养殖场牛舍,摄像头安装于牛舍西南角电线杆上,图像分辨率为 2 560 像素 ×1 440 像素,帧率为 20 f/s,视频保存格式为 MP4。共选取 2023 年 4 月和 10 月不同时段视频 59 段,牛只 17 头,包含白天、夜晚、雨天、风沙等自然天气场景。在识别牛只站立、躺卧、采食、饮水等日常行为的同时,加入牛只回头舔舐自己身体(回舔)行为。牛舍养殖环境如图 1 所示。



图 1 牛舍环境实例图

Fig. 1 Example of cowshed environment

1.2 数据预处理

将 2023 年 4 月采集的 49 段视频利用视频分帧技术分帧,保存格式为 jpg,去除冗余图像后,筛选 3 259 幅图像通过 LabelImg 软件进行牛只行为标注,将数据集按 7∶2∶1 划分为训练集、验证集、测试集,得到训练集图像 2 281 幅,验证集图像 652 幅,测试集图像 326 幅。多目标跟踪数据集通过选择不同场景和不同时间段的 2023 年 4 月和 10 月的 16 段视频,每段视频时长 60 s,利用 DarkLabel 软件标注牛只跟踪轨迹,制作 MOT17 数据集[19],记录检测目标 ID、边框大小、坐标,保存格式为 txt,其中 8 段用于训练,8 段用于测试。重识别数据集是通过将 17 头牛只的不同视觉角度、行为姿态等图像构建 ReID 数据集,共 10 744 幅图像,每头牛只平均有 632 幅图像。

1.3 多目标牛只跟踪方法

本文所提出的基于牛舍养殖环境下多目标牛只行为跟踪技术路线如图 2 所示。首先,将牛只行为视频中第 i 帧图像输入 YOLO v8 目标检测器,检测出每头牛只置信度、检测框和行为类别,得到第 i 帧的检测结果;然后,将检测结果分为高分检测框和低分检测框,并与相机运动补偿后的第 $i-1$ 帧跟踪轨迹预测算法得到的预测框进行目标关联匹配,得到第 i 帧的匹配结果;最后,依据当前匹配结果更新轨迹状态,循环重复处理连续视频帧图像,输出牛只多目标跟踪图像结果。

2 牛只行为识别方法

2.1 YOLO v8 目标检测

针对牛只的目标检测,选取 YOLO 系列中最新的模型 YOLO v8,相比 YOLO v5 和 YOLO v7,在牛只的多行为识别分类上,YOLO v8 比前者在精度和速度上都有提升。YOLO v8 主要分为 n、s、m、l、x 等模型,其中 YOLO v8n 模型内存占用量较小,在保持较高速度的同时,能够准确地检测常见大小的目标物体,更适合用于含有成年牛只和幼仔牛只的牛舍,因此选择在 YOLO v8n 模型基础上进行改进。

YOLO v8n 模型结构主要由 Backbone、Neck、Head 构成。主干网络 Backbone,使用 CSP 思想,通过 Conv 模块使用 3 ×3 的卷积核进行卷积操作提取特征。为了避免梯度消失和减轻训练难度,使用残差连接,C2f 模块将输入的特征图转换为全连接层输入。最后 SPPF 模块接收不同尺度池化窗口,对特征图池化拼接。Neck 端主要采用自下而上的特征提取和自上而下的特征融合。上采样模块

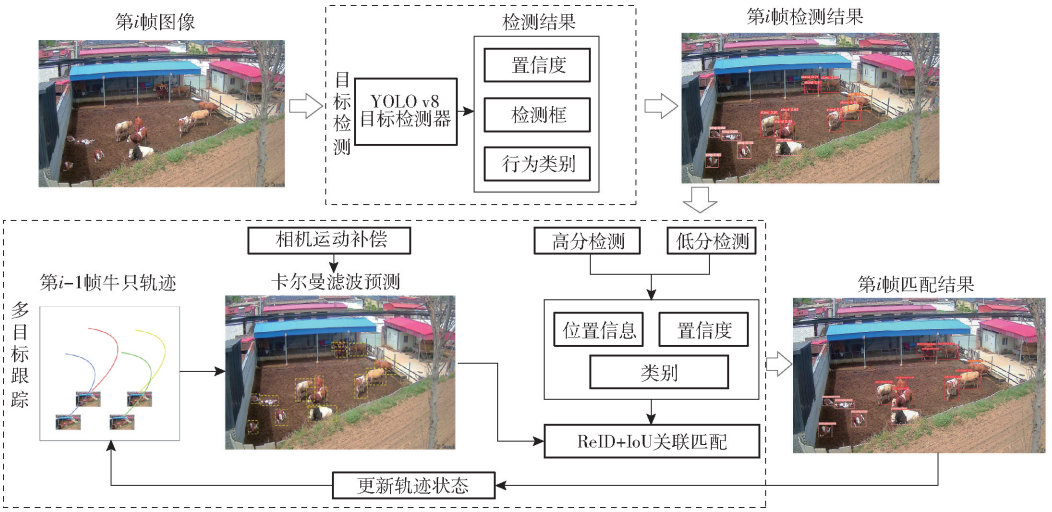


图 2 多目标牛只行为跟踪技术路线图

Fig. 2 Technical route diagram for multi target bull behavior tracking

Upsample 使用邻近插值算法，经过上采样后将特征图尺寸扩大 2 倍。Concat 模块将来自主干网络 Backbone 的特征图与上采样操作后得到的特征图进行连接，以融合低级和高级的语义信息。特征金字塔（FPN）的 Neck 部分通过卷积操作进行降维和通道数调整。Head 端检测器 Detect 由一系列卷积层和全连接层组成，负责在特征图上定位和识别目标，生成边界框的位置和类别预测。

2.2　基于改进 YOLO v8n 的牛只行为识别模型

2.2.1　特征提取 C2f－faster

牛舍养殖环境下，牛只活动场地较为拥挤，背景杂乱，牛只遮挡严重。当出现雨天、风沙天气时，电子监控摄像头表面会有水滴、尘土，摄像头会剧烈晃动等都会导致画面模糊。以上情况下，卷积神经网络在提取图像特征时通常会引入一些冗余特征和噪声，这会干扰模型对牛只行为的正确识别。卷积操作在每个位置上都进行一次滤波，同时卷积神经网络的多层堆叠也会增加模型的复杂度，使得模型需要更多的计算资源。

消除干扰特征冗余，减少模型复杂度，有利于提高目标检测精度。MobileNets、ShuffleNets 和 GhostNet 等利用深度卷积（DWConv）[20] 和分组卷积（GConv）[21] 来提取空间特征。然而，在减少浮点运算量的过程中，操作符通常会遭受内存访问增加的副作用。FasterNet[22] 提出了一个新的部分卷积 PConv，利用特征映射冗余性，系统地在部分输入通道上应用规则卷积（Conv），而不影响其他通道的变化。并在 PConv 基础上附加一个点向卷积（PWConv），以充分有效地利用来自所有通道的信息。每个 FasterNet Block 都有一个 PConv 层，后面连接 2 个 PWConv，或 1 ×1Conv 层。它们共同形成反向残差块的结构，其中中间层具有扩展的通道数，并且还添加了快捷连接以重复使用输入特征。基于 FasterNet Block 构造的 FasterNet 在分类、检测等方面有很高的性能，其结构如图 3a 所示。本研究将 YOLO v8 网络中的 C2f 模块进行改进，把 FasterNet Block 嵌入 C2f 中，形成新的结构 C2f－faster，用于在牛只目标检测中进行特征提取和处理，在减少参数量和计算复杂度的同时，保持一定的感受野范围和非线性表示能力，其结构如图 3b 所示。

2.2.2　CARAFE 上采样

YOLO v8 的 Upsample 上采样层以 nearest 最近邻插值方法为主，依赖于最近邻像素值，无法观测获取到图像的细微变化和密集语义信息，使图像失真。当图像放大时，会使图像模糊化，感受视野较小。所以本文采用大感受视野的轻量级上采样算子

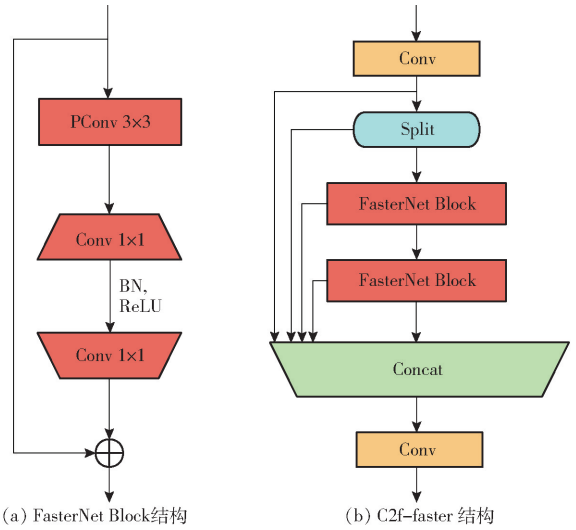


图 3　FasterNet Block 结构和 C2f－faster 结构

Fig. 3　FasterNet Block structure and C2f－faster structure

CARAFE，它可以在大范围内接收聚合语义信息。CARAFE[23] 采用特征感知重组对特征进行上采样，利用底层信息对每个位置进行预测重组核、定义重组特征。CARAFE 流程如图 4 所示，包括上采样核预测模块和特征重组模块两部分。首先将 $H \times W \times C$ 特征图通过 1 ×1 卷积压缩为 $H \times W \times C_m$，再经过内容编码和上采样核预测，得到 $\sigma H \times \sigma W \times K_{up}^2$ 的上采样核，对其进行归一化操作后，使其卷积核权重和为 1。然后特征重组模块，将输出特征图中的每个位置映射回输入特征图，取出其中心 $K_{up} \times K_{up}$ 的区域，并和预测后该点的上采样核作点积得到输出值。相同位置的不同通道共享同一个上采样核，最后输出 $\sigma H \times \sigma W \times C$ 的输出特征图。

2.2.3　BiFormer 注意力机制

成年牛只目标个体较大，动作幅度较明显，易于检测；幼仔牛只个体目标较小，识别精度低，容易漏检。针对幼仔牛只的小目标检测，在 YOLO v8n 模型基础上，在 Backbone、Neck 端之间引入动态稀疏注意力 BiFormer，以查询自适应的方式关注小面积牛只外观特征标记，而不会分散其他不相关标记的注意力，因此它具有良好的性能和计算效率。

在 SPFF 后添加 BiFormer 注意力机制模块[24]，可以去除冗余背景特征，聚焦牛只目标检测区域，并与 CARAFE 上采样相连接，帮助模型在低分辨率的特征图上进行关注度的加权，减少在高分辨率上的计算复杂性，准确定位牛只目标，从而提高模型的运行效率。BiFormer 以 BRA 为核心构建块，采用 4 层金字塔结构。具体步骤：阶段 1 使用重叠图像块嵌入，阶段 2 ~4 图像合并模块，降低分辨率，增加通道数。然后使用连续的 BiFormer Block 来转换输入数据的特征。BiFormer 结构图如图 5a 所示。每个

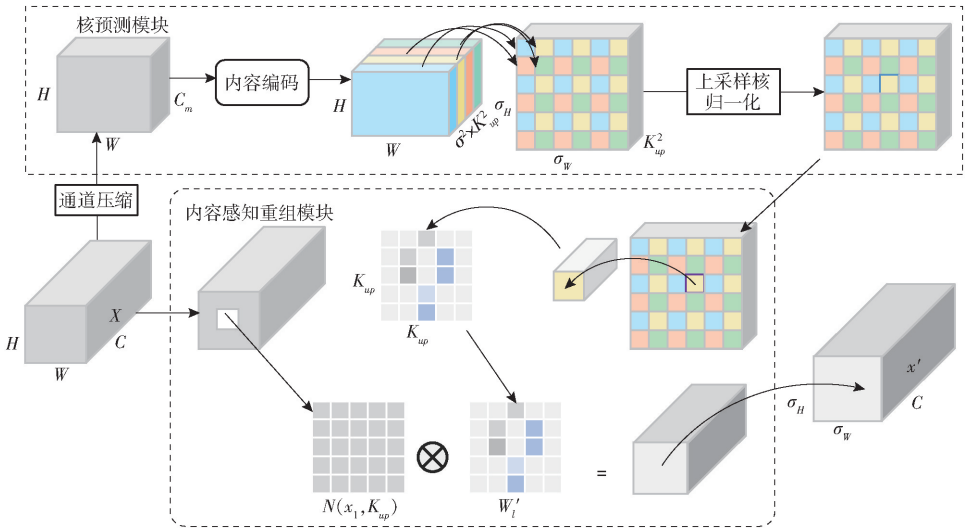
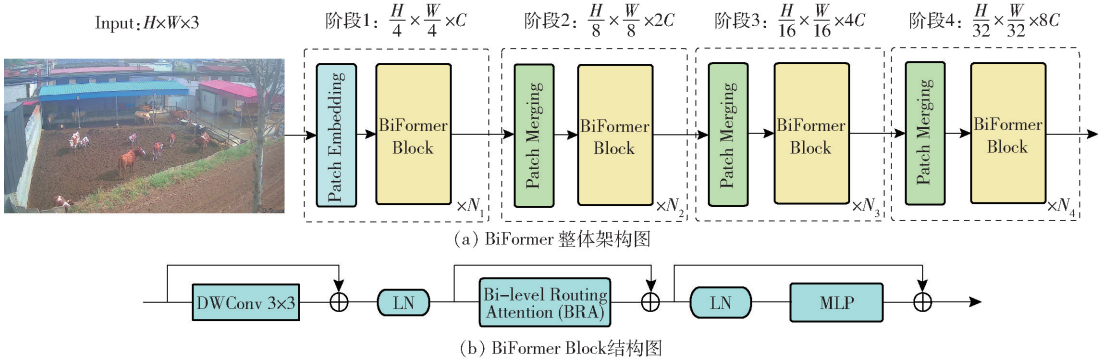


图 4 CARAFE 流程图

Fig. 4 CARAFE process diagram

BiFormer Block 开始都先应用 3×3 深度可分离卷积隐式编码相对位置信息,然后 BRA 模块捕捉双向相对注意力,最后使用双层感知器(MLP)模块进行关系建模和位置嵌入。其结构如图 5b 所示。



(a) BiFormer 整体架构图

(b) BiFormer Block结构图

图 5 BiFormer 结构图

Fig. 5 BiFormer structure diagram

2.2.4 动态检测头 DyHead

YOLO v8 模型是将输入图像通过 Backbone 提取特征,得到不同尺度的特征图(P3、P4 和 P5),特征图再经过 Detection Head 进行目标检测、分类,得到识别结果。在牛舍环境下检测牛只行为,因为多类型行为较为相似,不便于区分,需要高精度检测,研究面临许多问题挑战。首先,牛只处于生长发育阶段,大小不一,会使多尺度不同的牛只出现在同幅图像上,需要检测器有较高的多尺度感知能力。其次,牛只活动范围大,行为姿势多变,检测位置易发生变化,检测器需要空间感知牛只在不同空间位置之间的关系。最后,检测器需要具有任务感知能力,因为牛只检测有不同表征(例如:边界框、中心和角点),这些表征具有完全不同的目标和约束。为了提高牛只检测精度,引入 DyHead[25] 提高目标检测模型头部的表示能力,DyHead 检测范式如图 6 所示。

DyHead 使用新的动态头部框架来统一关注目标检测头,分别从尺度感知的特征层次、空间感知的空间位置和任务感知的输出通道内连贯地组合多个自注意机制这 3 个方面提高检测器能力。DyHead 根据牛只尺度大小引入尺度感知注意力(Scale-aware attention),根据不同尺度的语义重要性动态融合特征。利用空间感知注意力(Spatial-aware

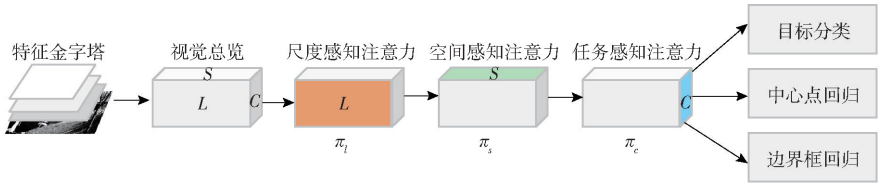


图 6 DyHead 检测范式

Fig. 6 DyHead detection paradigm

attention)在空间维度上处理牛只特征,并关注空间位置和特征层次之间的判别区域。最后部署任务感知注意力(Task-aware attention),根据牛只行为的不同表征,进行目标分类、中心点回归和边框回归。其中 π_l、π_s、π_c 分别是作用于维度 L、S、C 的 3 种不同的注意力函数。

2.2.5　Focal SIoU 损失函数

YOLO v8 默认采用完全交并比(CIoU)做边框回归损失函数,通过预测框与目标框的纵横比,来评估预测框与真实框位置差异。但 CIoU 没有考虑到预测框与目标框之间方向的匹配性,因此本文引入了 SIoU[26] 作为检测框回归损失函数。SIoU 在 CIoU 基础上添加所需回归之间的角度因素,重新定义惩罚指标。牛舍环境下牛只行为数据集相对不平衡,站立行为较多,回舔行为较少,两种行为相似,为解决数据集中正负样本分布不均衡的问题,采用 FocalLoss 函数[27],其通过降低简单负样本权重进行训练的方式来使检测模块区分站立、回舔行为。基于以上情况,FocalLoss 结合 SIoU 构建 Focal SIoU 损失函数。FocalLoss 函数公式为

$$FocalLoss = -\alpha(1-p_t)^{\gamma}\lg p_t \tag{1}$$

式中　p_t——模型预测概率

α——平衡参数,用来平衡正、负样本的比例

γ——可调节的聚焦参数,用来减少易分类样本损失

SIoU 表达式为

$$L_{SIoU} = L_{IoU} - (\Delta + \Omega)/2 \tag{2}$$

式中　L_{SIoU}——SIoU 损失值

Δ——距离损失值

Ω——形状损失值

L_{IoU}——IoU 损失值

基于以上模块对 YOLO v8n 进行改进得到的 FBCD－YOLO v8n(FasterNet、BiFormer、CARAFE 和 DyHead)结构如图 7 所示。

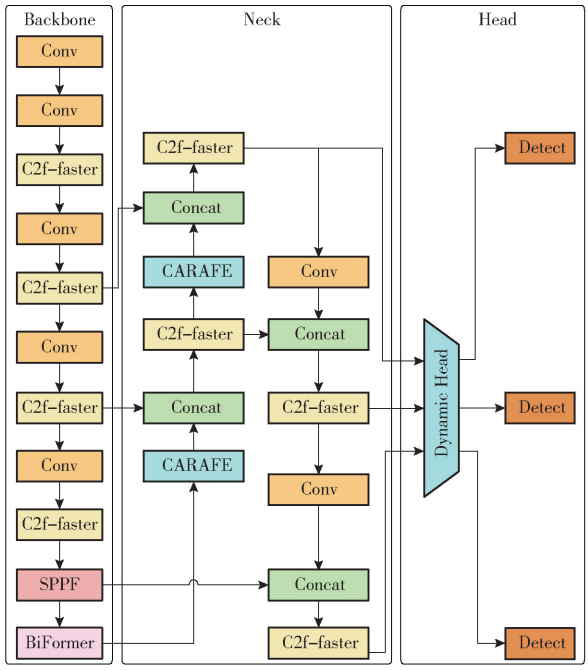


图 7　FBCD－YOLO v8n 网络模型

Fig. 7　FBCD－YOLO v8n network model

3　牛只跟踪方法

3.1　MOT 跟踪算法

多目标跟踪算法通过目标检测算法来跟踪场景内的目标物体,从而估计视频中多个目标的运动轨迹,给每个目标分配一个唯一 ID 标识符。目前跟踪步骤分为:运动模型和外观估计,通过卡尔曼滤波(Kalman filter,KF)[28] 预测后面序列帧的运动轨迹(Tracklets)边界框;目标定位和外观模型,通过预测轨道边界框和检测边界框之间的交并比(IoU)和重识别(ReID),将视频序列的最新帧检测结果与当前轨迹集合相关联。

目前主流的跟踪算法有 DeepSORT、ByteTrack、StrongSORT,本文选用 BoTSORT 算法,它是在 ByteTrack 的基础上进行改进,整体采用 SORT 框架,结合了运动特征和外观信息的优势,加入摄像机运动补偿和一个更精确的卡尔曼滤波状态向量,实现了在复杂场景下对多目标牛只行为的稳定跟踪。

3.1.1　卡尔曼滤波器

为了模拟图像上运动轨迹,通常使用具有恒定速度模型的离散卡尔曼滤波器对运动中的检测目标进行运动建模。经典跟踪器 DeepSORT 中提出的卡尔曼滤波状态表征试图估计框的纵横比而不是宽度,这导致宽度估计不准确。SORT 的卡尔曼滤波状态向量为 7 元组,DeepSORT 的卡尔曼滤波状态向量为 8 元组,其将 SORT 元组中第 4 维负责边界框纵横比的预测改为边界框的高度预测。而 BoTSORT 为获得更好的性能,重新选择定义卡尔曼滤波的状态向量,BoTSORT 将纵横比改成了预测宽度,直接估计边界框的宽度和高度。

3.1.2　摄像机运动补偿

跟踪器依赖于预测目标运动轨迹边框和检测目标轨迹边框的重叠,动态化的摄像机会使视频画面晃动,边框位置发生变化,造成预测匹配不准确。牛舍的摄像头在面对风沙天气时,由于风的影响,摄像头会轻微晃动,Box 的像素位置会发生剧烈变化导致牛只跟丢。由于缺乏摄像机运动数据、固有矩阵,可将相邻帧之间的图像配准,看成近似采用摄像机的刚性运动在图像平面上的投影,即利用 OpenCV[29] 视觉库中带有仿射变换的算法的视频防

抖模块,这种方法称为全局运动补偿(Global motion compensation,GMC)。这项技术可理解为相机中牛只运动等同于坐标系的变换,牛只的位置、运动方向等需要重新投影到新坐标系中,因此需要求解坐标系的变换矩阵。相机对牛只运动补偿示意图如图 8 所示。

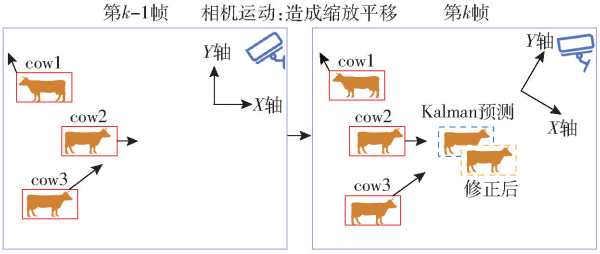


图 8 牛只运动补偿示意图

Fig. 8 Cattle motion compensation schematic

首先,通过 OpenCV 提取图像特征点,并且通过特征点进行稀疏光流跟踪,使用 RANSAC 计算求解获取运动背景下目标的仿射变换矩阵,进而预测从第 $k-1$ 帧的坐标系转换为第 k 帧坐标系的牛只边界框,称为运动补偿。仿射变换矩阵平移部分仅影响边界框的中心位置,而其他部分则影响所有状态向量和噪声矩阵,需要通过相机运动校正方程解决。相机运动校正方程为

$$\boldsymbol{K}_k=\boldsymbol{P}'_{k|k-1}\boldsymbol{H}_k^{\mathrm{T}}(\boldsymbol{H}_k\boldsymbol{P}'_{k|k-1}\boldsymbol{H}_k^{\mathrm{T}}+\boldsymbol{R}_k)^{-1} \tag{3}$$

$$\hat{\boldsymbol{x}}_{k|k}=\hat{\boldsymbol{x}}'_{k|k-1}+\boldsymbol{K}_k(\boldsymbol{z}_k-\boldsymbol{H}_k\hat{\boldsymbol{x}}'_{k|k-1}) \tag{4}$$

$$\boldsymbol{P}_{k|k}=(\boldsymbol{I}-\boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}'_{k|k-1} \tag{5}$$

式中 $\hat{\boldsymbol{x}}_{k|k-1}$、$\hat{\boldsymbol{x}}'_{k|k-1}$——时刻 k 时,相机运动补偿前、后的 KF 预测状态向量

$\boldsymbol{P}_{k|k-1}$、$\boldsymbol{P}'_{k|k-1}$——校正前、后 KF 的预测协方差矩阵

$\boldsymbol{R}_k$——测量误差的协方差矩阵

$\boldsymbol{z}_k$——测量向量 $\boldsymbol{I}$——单位矩阵

$\boldsymbol{H}_k$——测量矩阵,用于将状态向量映射到测量空间,以便在卡尔曼滤波的测量更新步骤中进行状态估计的修正

$\boldsymbol{K}_k$——增益矩阵,用于将预测的状态估计与实际测量结果进行融合

3.1.3 IoU + ReID 的融合机制

为了更准确地跟踪目标位置,BoTSORT 将检测目标的外观特征与跟踪器相融合。采用 FastReID 库中的 BoT(SBS)[30] 之上的更强基线,以 ResNeSt50 作为主干,采用指数移动平均(EMA)机制来更新轨迹外观状态。由于外观状态容易受到物体遮挡影响,所以只考虑高置信度检测。为了匹配平均轨迹外观状态和新检测的嵌入向量,测量余弦相似度,使用 IoU + ReID 融合机制的策略,即 IoU 距离矩阵和余弦距离矩阵来进行匹配。首先,对 IoU 分值,低余弦相似度或距离较远的侯选者会被拒绝。然后,使用矩阵每个元素中的最小值作为成本矩阵 $\boldsymbol{C}$ 的最终值。IoU - ReID 融合管道公式为

$$C_{i,j}=\min(d_{i,j}^{\mathrm{iou}},\hat{d}_{i,j}^{\mathrm{cos}}) \tag{6}$$

其中

$$\hat{d}_{i,j}^{\mathrm{cos}}=\begin{cases}0.5d_{i,j}^{\mathrm{cos}} & (d_{i,j}^{\mathrm{cos}}<\theta_{\mathrm{emb}}\text{且}\ d_{i,j}^{\mathrm{iou}}<\theta_{\mathrm{iou}})\\ 1 & (\text{其他})\end{cases} \tag{7}$$

式中 $C_{i,j}$——成本矩阵 $\boldsymbol{C}$ 的(i,j)元素

$d_{i,j}^{\mathrm{iou}}$——第 i 个轨迹预测边界框与第 j 个检测边界框间的 IoU 距离,表示运动成本

min——最小值函数

$d_{i,j}^{\mathrm{cos}}$——平均轨迹外观描述符 i 和新检测描述符 j 之间的余弦距离

$\hat{d}_{i,j}^{\mathrm{cos}}$——新外观成本

θ_{iou}——接近阈值,用于舍弃不可能的轨迹和检测对

θ_{emb}——外观阈值,用于将轨迹外观状态和检测嵌入向量的正关联与负关联分开

3.2 BoTSORT 跟踪算法流程

基于多目标跟踪的 BoTSORT 算法利用 YOLO v8 对牛只进行目标检测,将卡尔曼滤波器(KF)、摄像机运动补偿(CMC)、IoU + ReID 的融合机制进行修改和改进,将它们集成到跟踪算法 ByteTrack 中。BoTSORT 跟踪算法主要步骤如下:①YOLO 目标检测。②预处理和阈值选择。③外观特征提取,ReID 模型提取外观特征。④估计相机运动,CMC 对卡尔曼滤波器和预测边界框修正。⑤数据关联,涉及数据与之前检测到帧中的数据关联起来。⑥轨迹管理,包括更新卡尔曼滤波器和外观信息,以及创建轨迹和删除轨迹,经过后处理,得到跟踪结果。图 9 为使用 BoTSORT 进行牛只跟踪的过程。

3.3 实验平台与评价指标

本实验在 Windows 10 系统环境下进行,编程平台为 pycharm,编程语言为 python 3.8,深度学习框架为 pytorch 1.12.1。硬件环境为 AMD 锐龙 R5 5600 处理器、64 GB 内存、NVIDIA GeForce RTX 4070 显卡。

实验通过 YOLO v5n、YOLO v7tiny 和 YOLO v8n 共 3 个目标检测模型训练牛只行为标签数据集,输入图像尺寸为 640 像素 ×640 像素,图像批量大小为 16,学习率为 0.01,模型训练 300 轮,在保证训练收敛的前提下,再对数据集进行评价。跟踪方面,对于 SORT 跟踪算法,IoU 阈值为 0.7,conf 阈值为 0.5,其余参数取默认值。

为了验证牛只目标检测性能,采用准确率(Precision,P)、查全率(Recall,R)、帧速率(Frames per second,FPS)和平均精度均值(mean average precision,mAP)作为评价指标。实验中,mAP 取 IoU

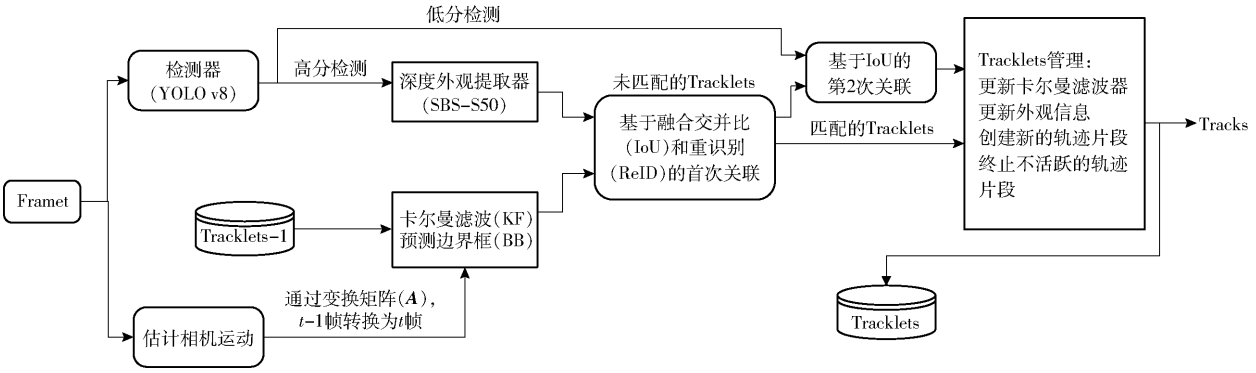


图 9　BoTSORT 跟踪算法流程图

Fig. 9　BoTSORT tracking algorithm flowchart

阈值为 0.5(mAP@0.5)以及不同 IoU 阈值 0.5 ~ 0.95(mAP@0.5:0.95),综合评价模型的准确性。对于牛只跟踪的评价,本研究选取多目标跟踪准确率(Multiple object tracking accuracy,MOTA)、多目标跟踪精确度(Multiple object tracking precision,MOTP)、识别平均数比率(Identification F1,IDF1)、高阶跟踪精度(Higher order tracking accuracy,HOTA)、大多数跟踪成功率(Mostly tracked ratio,MTR)、主要跟踪丢失率(Mostly lost ratio,MLR)6 个评价指标,衡量跟踪模型性能,其中 P、R、mAP 越高,表示目标检测识别越准确;FPS 越高,表示目标检测速度越快;MOTA、MOTP、IDF1、HOTA、MTR 越高,而 MTR 越低,反映跟踪器性能越好。

4　实验与结果分析

4.1　目标检测实验与结果分析

针对牛只行为数据集,分别训练了 YOLO v5n、YOLO v7tiny、YOLO v8n 和 FBCD - YOLO v8n 模型,实验对比了牛只行为识别准确率、查全率、平均精度均值和帧速率来评估模型性能,实验结果如表 1 所示。

表 1　不同模型实验结果

Tab. 1　Different models experiment results

模型	P/%	R/%	mAP@0.5/%	帧速率/(f·s^{-1})
YOLO v5n	91.8	87.9	91.4	24.4
YOLO v7tiny	90.5	86.2	91.7	14.2
YOLO v8n	91.1	89.3	92.4	32.5
FBCD - YOLO v8n	93.8	90.0	94.8	24.2

相比较 YOLO v5n 和 YOLO v7tiny,YOLO v8n 查全率提高 1.4、3.1 个百分点,平均精度均值提高 1、0.7 个百分点,帧速率提高 8.1、18.3 f/s,准确率比 YOLO v5n 略低。说明 YOLO v8n 比 YOLO v5n 和 YOLO v7tiny 模型性能更强,在 YOLO v8n 模型上改进更适合牛只的行为分类。在 YOLO v8n 基础上改进的 FBCD - YOLO v8n 准确率为 93.8%,查全率为 90%,平均精度均值为 94.8%,帧速率为 24.2 f/s,比较其他 3 个模型,综合性能更优。

对比常用损失函数 CIoU、EIoU、GIoU、DIoU 和 SIoU 在 YOLO v8n 上收敛后的平均精度均值,如表 2 所示。实验结果表明,使用 SIoU 损失函数相比较其他 IoU 精度更高,YOLO v8n 精度提高到 93.6%;SIoU 损失函数引入分类 FocalLoss 后 YOLO v8n(Focal SIoU)模型精度提高 0.2 个百分点,比 YOLO v8n(CIoU)提高 1.4 个百分点。说明对于牛只行为数据集,在处理类别不平衡和难分类样本的情况下,结合 FocalLoss 和边界框损失函数的回归角度可以在目标检测任务中提升模型性能。

表 2　不同 IoU 损失模型平均精度均值对比

Tab. 2　Performance comparison of different IoU loss

%

模型	IoU	Focal + IoU
YOLO v8n + CIoU	92.4	93.2
YOLO v8n + EIoU	92.9	92.9
YOLO v8n + GIoU	92.9	93.5
YOLO v8n + DIoU	93.0	93.6
YOLO v8n + SIoU	93.6	93.8

为验证 FBCD - YOLO v8n 模型的有效性,从牛只总行为的 mAP 和牛只不同类别行为的平均精度(AP)进行分析,设计 C2f - faster、BiFormer、CARAFE、DyHead 等改进方法的消融实验,如表 3 所示。从实验结果来看,采用 YOLO v8n 模型对牛只行为识别的分类表现良好,尤其站立和躺卧行为精度较高,平均精度分别为 97% 和 99.1%,而在回舔行为上表现稍显不足,仅为 81.3%。分析原因为:牛只躺卧行为时,四肢身体全部卧于地面,特征明显,容易识别;而其他 4 种行为都是四肢立于地面,行为特征较为相似,其中站立和回舔区别在于牛只的头部扭动姿势,不容易区分,容易混淆,所以识别精度较低。后续实验随着不同模块的引入,模型的性能逐渐提升,特别是在回舔和饮水行为上,使用

特定方法后性能提升较为显著。

表 3 消融实验对比结果

Tab. 3 Comparative results of ablation experiment

%

模型	mAP@0.5	平均精度				
		站立	躺卧	饮食	回舔	饮水
YOLO v8n	92.4	97.0	99.1	95.6	81.3	89.0
YOLO v8n + C2f - faster	92.7	97.0	99.2	95.7	83.8	88.1
YOLO v8n + BiFormer	93.0	96.8	99.2	95.3	83.7	89.8
YOLO v8n + CARAFE + DyHead	93.3	97.2	99.2	95.7	84.0	90.5
YOLO v8n + Focal SIoU	93.8	97.4	99.2	96.1	85.7	90.6
FBCD - YOLO v8n	94.8	97.2	99.2	96.1	88.7	92.6

YOLO v8n + C2f - faster 较 YOLO v8n 对牛只回舔行为识别平均精度提高了 2.5 个百分点。在 YOLO v8n 中加入 BiFormer 注意力机制，回舔、饮水行为识别的平均精度提高 2.4、0.8 个百分点。分析原因是验证集中有幼仔牛只回舔和饮水，BiFormer 检测到小目标特征行为信息，识别正确。同时将 CARAFE + DyHead 两个模块加入 YOLO v8n 中，回舔、饮水行为提升效果明显，分别提升 2.7、1.5 个百分点。分析原因是在大感受视野下，CARAFE 和 DyHead 对处理动态和姿态变化较大的行为时尤为有效。结合之前实验模块，引入 Focal SIoU 损失函数，构建 FBCD - YOLO v8n，牛只行为 mAP@0.5 为 94.8%，站立为 97.2%，躺卧为 99.2%，饮食为 96.1%，回舔为 88.7%，饮水为 92.6%。综合来看，FBCD - YOLO v8n 模型在处理各种牛只行为分类任务中都表现出色。FBCD - YOLO v8n 模型牛只行为训练过程曲线如图 10 所示。

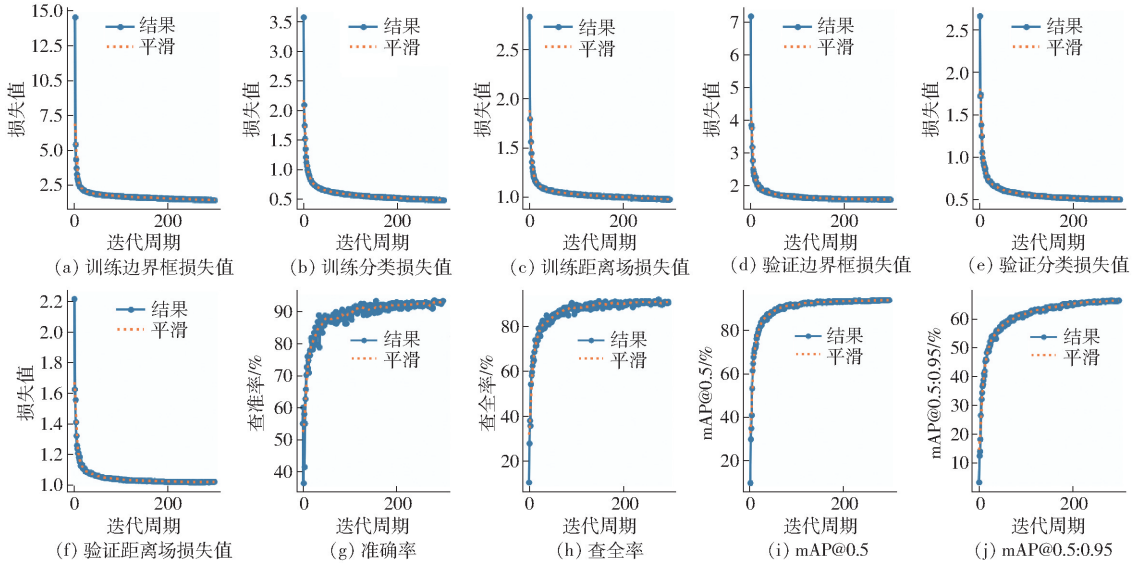


图 10 训练过程曲线

Fig. 10 Training process curves

4.2 重识别实验与结果分析

牛只的重识别模型实验旨在解决在不同帧中，可能出现因角度、光照变化等造成的同一头牛外观特征不一样，从而导致目标识别和跟踪方法在这种情况下效果不佳的问题。通过重识别模型，可以在不同时间和场景中准确地将同一头牛的不同外观关联起来，从而提高跨时间段、跨场景的牛只识别能力。

实验使用 Market 1501 数据集格式，制作牛只重识别数据集，采用 ResNet50 模型训练牛只重识别，并获得其权重文件。在牛只重识别数据集上训练重识别模型，批量大小为 32，迭代 100 轮，其余参数取默认值。

重识别实验结果如图 11 所示。训练后使用 Tensorboard 观察重识别准确率，其纵坐标表示重识别模型预测准确率，值越接近 1，说明模型提取特征能力越强，即重识别效果越好；横坐标表示训练步数，指模型参数根据训练数据进行一次更新的次数。当训练步数为 2.34×10^4 时，准确率曲线趋于平稳，此时模型已收敛，准确率为 100%。实验表明，重识

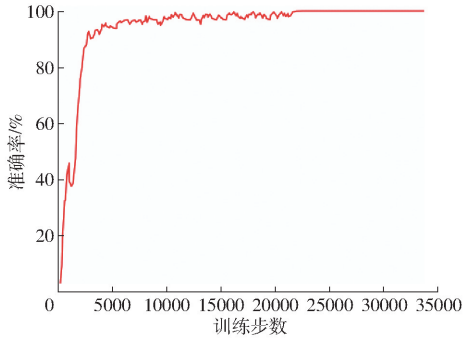


图 11 重识别模型准确率变化曲线

Fig. 11 Re-identification model accuracy change curve

别模型可以较好地提取不同牛只外观特征。

4.3　牛只行为跟踪实验与结果分析

关于牛只的跟踪实验，选取几种较为先进的 MOT 跟踪算法做对比实验。实验结果如表 4 所示，从 MOTA 来看，BoTSORT 算法表现最佳，准确率达到 96.1%，而 ByteTrack 和 StrongSORT 分别为 95.2% 和 94.6%。在 MOTP 方面，BoTSORT 算法为 78.6%，稍低于其他两个算法。在 IDF1 和 HOTA 方面，BoTSORT 表现最好，为 98.0% 和 78.9%，相比 StrongSORT 和 ByteTrack，IDF1 提高 1.7、3.2 个百分点，HOTA 提高 1.1、1.3 个百分点。在跟踪覆盖率上，BoTSORT 和 StrongSORT 都为 93.7%，这意味着 BoTSORT 和 StrongSORT 算法都能够很好地捕捉到大多数真实目标，几乎没有漏报。在跟踪错误率上，所有算法都为 0，表明跟踪目标没有误报。整体来看，MOTP 和 HOTA 精度略低。分析 MOTP 略低的原因是牛只在春季和秋季活动较为频繁，跟踪数据集采集于 4 月和 10 月，牛舍内牛只运动繁多、移动速度较快且行为姿势复杂多变，牛只跟踪时跟踪器预测的位置与实际位置之间的误差会增加，导致位置估计不准确，从而降低了 MOTP。分析 HOTA 精度略低的原因是牛舍场景复杂，活动范围较小，牛只活动较为拥挤，牛只运动时有围墙、树木和其他牛只遮挡，这可能会导致 ID 切换，因为牛只被遮挡或与其他牛只发生交互，导致跟踪系统难以正确识别牛只目标间的关联关系，进而降低 HOTA 得分。综合评价，BoTSORT 算法在多个指标下表现最佳，特别是在 MOTA、IDF1、MTR 和 MLR 方面表现出色，该算法在保持较高检测率的同时，较少漏检，并且没有误报，非常适用于牛只跟踪应用领域。

表 4　不同 MOT 算法跟踪性能对比

Tab.4　Performance comparison of different MOT

%

MOT 算法	MOTA	MOTP	IDF1	HOTA	MTR	MLR
ByteTrack	95.2	78.8	94.8	77.6	92.8	0
StrongSORT	94.6	78.9	96.3	77.8	93.7	0
BoTSORT	96.1	78.6	98.0	78.9	93.7	0

为了测试 BoTSORT 算法的性能及鲁棒性，在设计跟踪实验时增加了一些具有挑战性、更接近实际应用场景的元素。首先，实验引入天气场景、牛只活动程度等因素，作为实验的控制变量。其次，牛只跟踪数据集选取了 4 月和 10 月两个时间段的视频，数据采集间隔较长，同时考虑到牛只在这段时间内生长并且身躯发生变化，牛只数量由 4 月的 15 头增加到 10 月的 17 头。实验共选取 8 段不同场景条件下的视频作为多目标牛只跟踪测试集，其中视频 01、02、05、06、07、08 采集于 4 月，视频 03、04 采集于 10 月。实验结果如表 5 所示。

表 5　不同场景下 BoTSORT 算法跟踪性能对比

Tab.5　Performance comparison of BoTSORT algorithm in different scenes

%

视频编号	天气场景	牛只数量	活动程度	MOTA	MOTP	IDF1	HOTA	MTR	MLR
01	白天	14	稀疏	100	77.8	100	78.9	100	0
02	白天	14	拥挤	97.2	82.5	98.6	82.4	92.9	0
03	白天	17	稀疏	94.5	80.9	97.2	81.6	86.7	0
04	白天	17	拥挤	92.1	80.1	96.0	80.1	86.7	0
05	雨天	15	拥挤	93.5	79.2	96.8	76.4	93.3	0
06	风沙	12	拥挤	99.7	73.5	99.8	74.0	100	0
07	夜晚(19:30)	13	稀疏	98.7	76.7	99.3	77.9	100	0
08	夜晚(20:00)	13	拥挤	90.5	79.8	95.1	76.8	92.3	0

在图 12a 白天的稀疏活动场景中(视频 01)，BoTSORT 的 MOTA 和 IDF1 能够达到 100%，显示出较高的跟踪能力。相比其他天气场景，白天对牛只跟踪(视频 01、02、03、04)的 HOTA 都高，分别为 78.9%、82.4%、81.6%、80.1%。可以看出，BoTSORT 算法对牛只目标检测、位置估计和目标关联在白天时跟踪更为合适。对 4 月和 10 月白天场景下跟踪对比发现，10 月期间稀疏和拥挤场景下(视频 03、04)，牛只跟踪 MOTA、IDF1 和 MTR 稍微降低。分析原因可能是：10 月跟踪的牛只数量增加，而 YOLO v8 目标检测牛只数据集采集于 4 月，新加入的牛只的行为特征在训练数据中未被充分覆盖，并且之前训练的牛只在 10 月时身体尺寸和行为姿态发生了变化，从而影响跟踪性能；10 月的实验中，如图 12c 和图 12d 所示，摄像头视野出现树木枝叶和草丛遮挡情况。这种遮挡因素增加了跟踪难度，新的环境因素引入了更多不确定性，可能导致 MOTA、IDF1 和 MTR 等指标略微降低。在各个拥挤的场景中(视频 02、04、05、06、08)，如图 12e 雨天中牛只交叉重叠，BoTSORT 同样表现出色，MOTA 和 MOTP 都在较高水平上，表明算法在拥挤场景中牛只位置估计准确，能够有效跟踪物体。尤其图 12f



(a) 视频01,第250帧牛只跟踪结果 (b) 视频02,第250帧牛只跟踪结果
(c) 视频03,第250帧牛只跟踪结果 (d) 视频04,第250帧牛只跟踪结果
(e) 视频05,第82帧牛只跟踪结果 (f) 视频06,第135帧牛只跟踪结果
(g) 视频07,第120帧牛只跟踪结果 (h) 视频08,第120帧牛只跟踪结果

图 12 BoTSORT 跟踪示例

Fig. 12 BoTSORT tracking examples

风沙天气下,MOTA 和 IDF1 达到 99.7% 和 99.8%,MTR 和 MLR 分别为 100% 和 0,代表风沙天气下多目标跟踪牛只没有漏检和误报。分析原因是 BoTSORT 算法中引入了相机运动补偿,降低了摄像头晃动情况下牛只跟踪预测偏差,提高了风沙天气下对牛只跟踪的准确率。夜晚场景下(图 12g 和图 12h),分别为 19:30 和 20:00 两个时段(视频 07、08),并且在不同的活动程度下(稀疏和拥挤)使用了 BoTSORT 算法进行目标跟踪,BoTSORT 仍然保持高水平的性能,但牛只拥挤情况相对于稀疏场景跟踪准确率稍微下降。分析原因是夜晚的光线较暗,目标间相互干扰增加,造成了一些轻微的跟踪误差,尽管 MOTA 和 MOTP 稍有下降,但跟踪效果仍然很好。综合来看,BoTSORT 算法在不同天气场景和活动程度下都展现了优秀的跟踪性能,特别是在拥挤场景中也能够保持高水平的识别和跟踪准确性。

5 结论

(1)提出了基于改进 YOLO v8 的牛只行为识别与跟踪方法。采用 FBCD - YOLO v8n 算法对牛只进行目标检测,获得牛只在图像中的姿态信息。利用 BoTSORT 算法实现对牛只的精确跟踪。将改进后的 YOLO v8 检测到的行为类别信息输入到 BoTSORT 算法中,实现在牛舍养殖环境下对牛只的多目标行为识别跟踪。

(2)实验结果表明,改进后的 YOLO v8n 模型目标检测的 mAP 为 94.8%,对比 YOLO v5n、YOLO v7tiny 和 YOLO v8n 提高 3.4、3.1、2.4 个百分点。牛只的重识别实验准确率为 100%。BoTSORT 跟踪算法实验结果的 MOTA 为 96.1%,MOTP 为 78.6%,IDF1 为 98.0%,HOTA 为 78.9%,MTR 为 93.7%,MLR 为 0。对比 ByteTrack 和 StrongSORT 算法,本文方法跟踪性能更优。

参 考 文 献

[1] YAZDANBAKHSH O, ZHOU Y, DICK S. An intelligent system for livestock disease surveillance[J]. Information Sciences, 2017, 100(378): 26 - 47.

[2] JENSEN M B. Behaviour around the time of calving in dairy cows[J]. Applied Animal Behaviour Science, 2012, 139(3 - 4): 195 - 202.

[3] CHEN C, ZHU W, NORTON T. Behaviour recognition of pigs and cattle: journey from computer vision to deep learning[J]. Computers and Electronics in Agriculture, 2021, 187: 106255.

[4] 白强,高荣华,赵春江,等. 基于改进 YOLO v5s 网络的奶牛多尺度行为识别方法[J]. 农业工程学报,2022,38(12):163 - 172.
BAI Qiang, GAO Ronghua, ZHAO Chunjiang, et al. Multi-scale behavior recognition method for dairy cows based on improved YOLO v5s network[J]. Transactions of the CSAE, 2022, 38(12): 163 - 172. (in Chinese)

[5] 王政,许兴时,华志新,等. 融合 YOLO v5n 与通道剪枝算法的轻量化奶牛发情行为识别[J]. 农业工程学报,2022,38(23):130 - 140.
WANG Zheng, XU Xingshi, HUA Zhixin, et al. Lightweight recognition for the oestrus behavior of dairy cows combining YOLO v5n and channel pruning[J]. Transactions of the CSAE, 2022, 38(23): 130 - 140. (in Chinese)

[6] WANG R, GAO Z, LI Q, et al. Detection method of cow estrus behavior in natural scenes based on improved YOLO v5[J]. Agriculture, 2022, 12(9): 1339.

[7] 李琦,尚绛岚,李宝山. 基于 YOLO v3 和 Deep SORT 的草原牛跟踪系统[J]. 传感器与微系统,2021,40(6):83 - 85,88.
LI Qi, SHANG Jianglan, LI Baoshan. Grassland cattle tracking system based on YOLO v3 and Deep SORT[J]. Transducer and Microsystem Technologies, 2021, 40(6): 83 - 85,88. (in Chinese)

[8] REDMON J, FARHADI A. YOLO v3: an incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[9] CHO C A, THI T Z, IKUO K. Black cow tracking by using deep learning-based algorithms[J]. 革新的コンピューティング·情

報·制御に関する速報—B:応用, 2022, 13(12): 1313.
[10] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017: 3645 - 3649.
[11] 曹天一,李宝山,王月明,等. 基于起卧频率异常的牛产犊自动预警[J]. 农业工程学报,2022,38(增刊 1):201 - 208.
CAO Tianyi, LI Baoshan, WANG Yueming, et al. Automatic warning of calving based on the abnormal frequency of standing up and lying down[J]. Transactions of the CSAE, 2022, 38(Supp. 1): 201 - 208. (in Chinese)
[12] 涂淑琴,刘晓龙,梁云,等. 基于改进 DeepSORT 的群养生猪行为识别与跟踪方法[J]. 农业机械学报,2022,53(8):345 - 352.
TU Shuqin, LIU Xiaolong, LIANG Yun, et al. Behavior recognition and tracking method of group-housed pigs based on improved DeepSORT algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8): 345 - 352. (in Chinese)
[13] 涂淑琴,汤寅杰,李承桀,等. 基于改进 ByteTrack 算法的群养生猪行为识别与跟踪技术[J]. 农业机械学报,2022,53(12):264 - 272.
TU Shuqin, TANG Yinjie, LI Chengjie, et al. Behavior recognition and tracking of group-housed pigs based on improved ByteTrack algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(12): 264 - 272. (in Chinese)
[14] GE Z, LIU S, WANG F, et al. Yolox: exceeding yolo series in 2021[J]. arXiv preprint arXiv:2107.08430, 2021.
[15] SU N M,THI T Z,PYKE T, et al. Comparing state-of-the-art deep learning algorithms for the automated detection and tracking of black cattle[J]. Sensors, 2023, 23(1): 532.
[16] 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.
[17] DU Y, ZHAO Z, SONG Y, et al. StrongSORT: make DeepSORT great again[J]. arXiv preprint arXiv:2202.13514, 2022.
[18] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT - SORT: robust associations multi-pedestrian tracking[J]. arXiv preprint arXiv:2206.14651, 2022.
[19] MILAN A, LEAL-TAIXÉ L, REID I, et al. MOT16: a benchmark for multi-object tracking[J]. arXiv preprint arXiv:1603.00831, 2016.
[20] SIFRE L, MALLAT S. Rigid-motion scattering for texture classification[J]. arXiv preprint arXiv:1403.1687, 2014.
[21] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25(6): 84 - 90.
[22] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021 - 12031.
[23] WANG J, CHEN K, XU R, et al. Carafe: content-aware reassembly of features[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision,2019: 3007 - 3016.
[24] ZHU L, WANG X, KE Z, et al. BiFormer: vision transformer with bi-level routing attention[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 10323 - 10333.
[25] DAI X, CHEN Y, XIAO B, et al. Dynamic head: unifying object detection heads with attentions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 7373 - 7382.
[26] GEVORGYAN Z. SIoU loss: more powerful learning for bounding box regression[J]. arXiv preprint arXiv:2205.12740, 2022.
[27] TSUNG-YI L,PRIYA G, ROSS G, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(2):318 - 327.
[28] SIMON D. Kalman filtering[J]. Embedded Systems Programming, 2001, 14(6): 72 - 79.
[29] BRADSKI G. The OpenCV library[J]. Software Tools for the Professional Programmer, 2000, 25(11): 120 - 123.
[30] LUO H, GU Y, LIAO X, et al. Bag of tricks and a strong baseline for deep person re-identification[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2019: 1487 - 1495.