doi:10.6041/j.issn.1000-1298.2025.10.053

基于 YOLO v9c 和改进 ByteTrack 的群养羊只多目标跟踪方法

郑　芳[1,2]　夏传宇[1]　杜小勇[1,2]　周　勇[3]　田　芳[1,2]　李国亮[1,2]

(1. 华中农业大学信息学院, 武汉 430070; 2. 农业农村部智慧养殖技术重点实验室, 武汉 430070;
3. 甘肃省金昌市畜牧兽医总站, 金昌 737100)

摘要: 基于视频的群养羊只跟踪对于规模化、智能化、无人化养殖具有重要作用。然而,由于羊群存在严重的遮挡、重叠及移动速度过快等问题,在复杂场景下对多只羊只进行精确跟踪较为困难。针对上述问题,为提高跟踪技术对群养羊只的适应性,本文提出一种基于 YOLO v9c 与改进 ByteTrack 相结合的羊只多目标跟踪方法。在目标检测方面,将羊只的行为划分为站立、躺卧和进食 3 种状态。在多目标跟踪方面,对 ByteTrack 做了两点改进:引入时间距离匹配模块(Time and distance matching module,TDMM),将未匹配成功的高分框与未匹配轨迹按照丢失轨迹的丢失时长与欧氏距离相结合形成身份关联系数矩阵,再次进行匹配;引入 ID 延时分配机制,除第 1 帧外,将 ID 分配模块移至第 3 次匹配并加入条件,防止 ID 过早分配。试验结果表明,HOTA 为 72.051%,MOTA 为 88.326%,IDF1 为 88.237%,IDSW 为 8。与 ByteTrack 相比 MOTA 提高 0.242 个百分点,HOTA 提高 2.21 个百分点,IDF1 提高 5.734 个百分点,ID 跳变次数降低了约 46.67%。与算法 Bot-SORT 和 OC-SORT 相比,HOTA 和 IDF1 有明显上升,并且 ID 跳变次数大幅降低。多羊只复杂场景情况下测试结果表明,基于改进 ByteTrack 算法具有良好的多目标跟踪性能,可以有效提高对群养羊只跟踪的准确性和可靠性。该算法在与 YOLO v9c 目标检测算法结合对群养羊只进行多目标跟踪并保存跟踪结果时,平均帧率为 47.1 f/s 相较于 Bot-SORT 算法(34.2 f/s)提高约 37.7%。该算法能够实时可靠地监测羊只,为羊只养殖场管理者及时发现羊只行为异常以及监测羊只健康状况提供了有效的技术手段。

关键词: 群养羊只; 目标跟踪; 时间距离匹配模块; ID 延时分配; YOLO v9c; 改进 ByteTrack

中图分类号: TP391.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)10-0585-11　　**OSID**:

Multi-objective Tracking Method for Sheep in Flock Breeding Based on YOLO v9c and Improved ByteTrack

ZHENG Fang[1,2]　XIA Chuanyu[1]　DU Xiaoyong[1,2]　ZHOU Yong[3]　TIAN Fang[1,2]　LI Guoliang[1,2]

(1. *Collage of Informatics*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*
2. *Key Laboratory of Smart Breeding Technology*, *Ministry of Agriculture and Rural Affairs*, *Wuhan* 430070, *China*
3. *Jinchang Animal Husbandry and Veterinary Station*, *Gansu Province*, *Jinchang* 737100, *China*)

Abstract: Video-based group sheep tracking plays an important role in large-scale, intelligent and unmanned breeding. However, due to the severe occlusion, overlap and excessive movement speed of the flock, it is difficult to accurately track multiple sheep in complex scenes. In order to solve the above problems and improve the adaptability of tracking technology to group sheep, a multi-objective sheep tracking method was proposed based on the combination of YOLO v9c and improved ByteTrack. In terms of object detection, the behavior of sheep was divided into three states: standing, grovel and eating. In terms of multi-target tracking, two improvements were made to ByteTrack: the time and distance matching module (TDMM) was introduced, and the unmatched high-score frame and unmatched trajectory were combined according to the loss time of the lost trajectory and the Euclidean distance to form an identity association coefficient matrix, and the matching was carried out again. The ID delay allocation

收稿日期: 2024-06-11　修回日期: 2024-07-08

基金项目: 国家自然科学基金项目(31872978)和国家重点研发计划青年科学家项目(2021YFD1300800)

作者简介: 郑芳(1980—),女,副教授,博士,主要从事机器学习和智慧农业研究,E-mail: zhengfang@mail.hzau.edu.cn

通信作者: 李国亮(1972—),男,教授,博士,主要从事智慧农业和生物信息研究,E-mail: guoliang.li@mail.hzau.edu.cn

mechanism was introduced, and in addition to the first frame, the ID allocation module was moved to the third match and conditions were added to prevent premature ID allocation. The results showed that the HOTA was 72.051%, the MOTA was 88.326%, the IDF1 was 88.237%, and the IDSW was 8. Compared with the original ByteTrack, MOTA was increased by 0.242 percentage points, HOTA was increased by 2.21 percentage points, IDF1 was increased by 5.734 percentage points, and the number of ID hops was decreased by about 46.67%. Compared with the representative algorithms Bot - SORT and OC - SORT, the number of ID hops was significantly increased in HOTA and IDF1, and the number of ID hops was greatly reduced. The test results in the complex scenario of multiple sheep showed that the improved ByteTrack algorithm had good multi-target tracking performance, which can effectively improve the accuracy and reliability of group sheep tracking. When the algorithm was combined with the YOLO v9c object detection algorithm to track the sheep in groups and save the tracking results, the average video processing speed was 47.1 f/s, which was about 37.7% higher than that of the Bot - SORT algorithm of 34.2 f/s. The algorithm can reliably monitor sheep in real time, which can provide an effective technical means for sheep farm managers to detect abnormal sheep behavior and monitor the health status of sheep in time.

Key words: sheep in flock breeding; target tracking; TDMM; delayed ID allocation; YOLO v9c; improved ByteTrack

0 引言

作为世界上最大的畜牧国之一,羊养殖业是我国畜牧业的重要组成部分[1]。2023 年上半年中国羊出栏量达 $1.453\ 6\times10^8$ 只,同比增加 6.62×10^6 只。我国羊养殖业正由个体化养殖朝着规模化、集约化方向发展。但随着养殖规模与养殖密度逐渐增大,羊群患病概率呈上升趋势,羊只精细化管理也面对重大挑战。羊只行为反映其生理与心理健康状况[2-3],而传统人工监测方式存在难以持续监测、工作强度大、效率低、误诊率高等问题[4-5]。

近年来,人工智能技术在各个领域中快速融合发展,羊只养殖也逐步向智能化方向转变。基于射频识别技术的可穿戴设备在羊只监测方面取得了一定成功[6]。但也存在缺点,例如可能需要在牲畜耳朵上打孔,会对牲畜造成伤害,穿戴设备会因牲畜运动而容易脱落以及在大型养殖环境下成本高[7-8]。计算机视觉技术在动物监测方面得到广泛应用,可提供一种无接触式、自动化的方式来实时监测羊只行为[9]。基于深度学习的非接触式计算机视觉技术对群养羊只的监测已经成为一种有效的解决方案[10-12]。然而,在羊只监测方向,目前研究多为针对羊只行为以及羊只计数统计,在实际较大规模羊只养殖场景,不仅需要对羊只行为进行监测还需要确定到具体的羊只,以实现真正的精细化、无人化、智能化管理。由于羊只生活习性,羊群会经常性地拥挤在一起,从而出现严重的长时间遮挡问题,且羊只外观特征极为相似的特点又为目标跟踪增加了难度。因此,迫切需要一种高效的方法来实现对群养羊只的准确实时多目标跟踪。

基于检测的跟踪(Tracking by detection, TBD)算法[13]相比于联合检测跟踪(Joint detection and embedding, JDE)算法[14-15],TBD 算法检测精度更高。目前有两类主流的目标检测算法,第 1 类是一阶段检测算法,可直接在网络中同时预测类别和边界框位置,YOLO 系列[16-18]为其代表性的算法。第 2 类是二阶段检测算法,第 1 阶段提取感兴趣的区域,第 2 阶段进行分类和定位,例如 Faster R - CNN 算法[19]。其中 YOLO 系列具有更快的检测速度,能够满足对多目标跟踪实时性具有较高要求的场景。在跟踪方面,以观察为中心的 OC - SORT[20]算法是在 SORT 基础上,通过修复 SORT 对状态估计噪声的敏感性,随时间的误差累积和以估计为中心的问题这 3 个限制进行遮挡和非线性运动下的跟踪,实现了较为先进的性能表现。但在实际羊只养殖场景下,群养羊只具有非线性、非匀速的运动特点且经常存在相互遮挡的场景。Bot - SORT[21]是在 ByteTrack 基础上改进了卡尔曼滤波,引入相机运动补偿算法,使得轨迹预测更加精准;其在一定程度上可以缓解由于群养羊只运动不规则性和遮挡导致的 ID 频繁跳动的问题,但在复杂场景下,如羊群运动量增大,该算法抑制 ID 频繁跳动效果会显著降低。

针对上述问题,本研究提出一种基于 YOLO v9c 和改进 ByteTrack 的群养羊只跟踪算法。YOLO v9c 作为算法检测器,负责提供预测边界框位置,在 ByteTrack 的基础上引入时间距离匹配模块和 ID 延时分配机制,时间距离匹配可以有效缓解群养羊只长时间遮挡以及运动等问题导致的目标丢失,ID 延时分配机制可以解决检测端误检导致 ID 频繁跳动的问题,以期实现对群养羊只的高效、准确、实时跟踪。

1　数据采集与处理

1.1　视频数据采集

试验数据采集自甘肃省金昌市某羊场，视频采集使用海康威视 4K 超高清臻真彩 POE、焦距 2.8 mm 的枪形摄像机，通过手机 APP 远程连接监控录像机获取视频数据。采集目标为 3 个羊栏内的临产前群养羊只，视频拍摄的画面中每个羊栏 7 ~ 9 只羊，于 2023 年 2 月 14—28 日采集全天 24 h 的视频，每 24 min 保存为一段。视频分辨率为 3 840 像素 × 2 160 像素，格式为 mp4，帧率为 20 f/s。为确保数据的多样性与可靠性，分别选取白天、夜晚、密集场景、活动频繁等场景下图像。视频采集摄像机与羊栏位置及形状如图 1 所示。

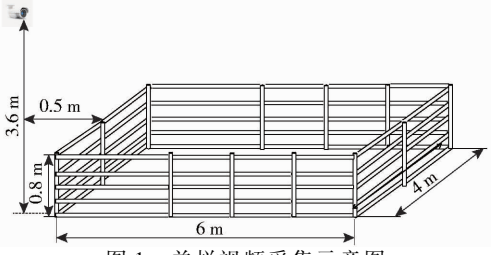


图 1　羊栏视频采集示意图

Fig. 1　Schematic of video collection of sheep pen

1.2　数据集构建

群养羊只目标检测数据集与跟踪测试数据集均是基于上述采集的多段 24 min 的视频片段所构建。

1.2.1　群养羊只目标检测数据集

首先，人工挑选不同羊栏、不同时间段的视频；然后使用 Python 编码对筛选出的视频进行关键帧截取。本数据集每 240 帧（即 12 s）截取 1 帧图像，既避免相邻两帧图像相似性过高，又保证了数据集的完备性，所得图像可以涵盖到羊只在羊圈内位置与羊只行为状态；最后使用通过设定阈值相似度高的图像进行筛除，保留相似度较低的图像作为数据集。使用 Labelimg 工具对图像进行标注，主要标注羊只 3 种状态，即站立（stand）、进食（eat）、趴卧（grovel）。标注数据以 YOLO 格式保存，标注文件与图像文件一一对应，生成的标注 TXT 文件中包括每一只羊的位置坐标和类别信息，最终共标注图像 1 500 幅，训练集、验证集和测试集按比例8∶1∶1划分。目标检测数据集羊只数据如图 2 所示。

1.2.2　群养羊只跟踪测试数据集

为了验证改进 ByteTrack 算法在各个场景下的有效性，选取白天羊只活动频繁、夜晚羊只活动频繁、白天羊只活动闲散、夜晚羊只活动闲散 4 种视频共 6 段，且这 6 段视频不包含在训练集中，总时长约 9 min。使用 Python 脚本对测试视频进行处理，通过间隔抽帧的方法将视频帧率降为 10 f/s，并将视频图像尺寸调整为 1 920 像素 ×1 080 像素。



(a) 羊栏(白天)　　(b) 羊栏(夜晚)

图 2　目标检测数据集

Fig. 2　Object detection dataset

2　方法

2.1　YOLO v9 目标检测算法

传统深度学习中，深度网络输入数据在逐层进行特征提取和空间变换时，会丢失大量信息。目前，可以缓解这一问题的方法主要有可逆架构、掩模建模以及引入深度监督概念。然而，这些方法存在不同的缺点，例如，使用可逆架构需要额外的层来结合重复输入数据，这会显著增加推理成本，并且在训练过程中难以对高阶语义信息进行建模；掩模建模，其重构损失有时会与目标损失冲突，此外，大多数掩码机制也会与数据产生不正确的关联；对于深度监督机制来说，其会产生误差累积。为了解决上述问题，YOLO v9 团队提出可编程梯度信息（Programmable gradient information，PGI）[22]。不同于一般的路径聚合网络（Path aggregation network，PAN）[23]，PGI 思想是通过辅助可逆分支（Auxiliary reversible branch）与多级辅助信息（Multi-level auxiliary information）生成可靠的梯度，使深层特征仍能保持关键特征，如图 3 所示。辅助可逆分支可在不同的语义层次上对梯度信息传播进行编程，从而获得最佳的训练结果，避免了传统的多路径特征融合深度监督过程可能造成的语义损失。并且，PGI 的可逆架构是建立在辅助分支上的，只在训练时参与权重调节，因此在推理时无额外成本。此外，YOLO v9 还在高效聚合网络（Efficient layer aggregation network，ELAN）的基础上设计了广义高效聚合网络（Generalized efficient layer aggregation network，GELAN），GELAN 考虑了参数量、计算复杂性、准确性和推理速度，允许用户随意选择适合不同推理设备的计算块。

本研究将 YOLO v9c 网络结构分为 5 个部分，分别为输入（Input）、主干（Backbone）、头部（Head）、多级辅助信息和辅助可逆分支，如图 4 所示。在 YOLO v9c 中，多级辅助信息与辅助可逆分支两部分仅在训练时使用，在推理时仅使用输入、主干部分和头部网络。

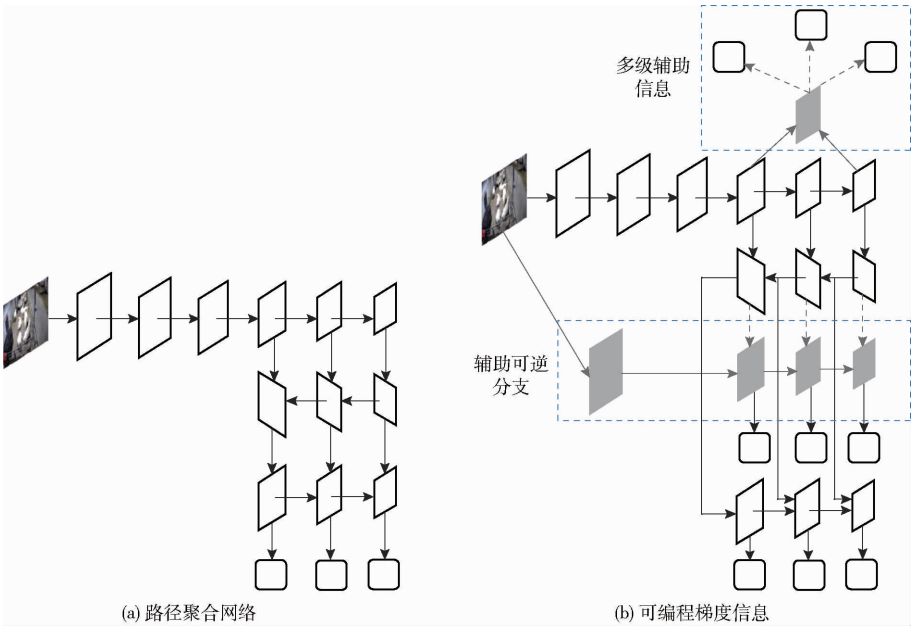


图 3 PAN 与 PGI 结构对比

Fig. 3 Comparison of PAN and PGI structures

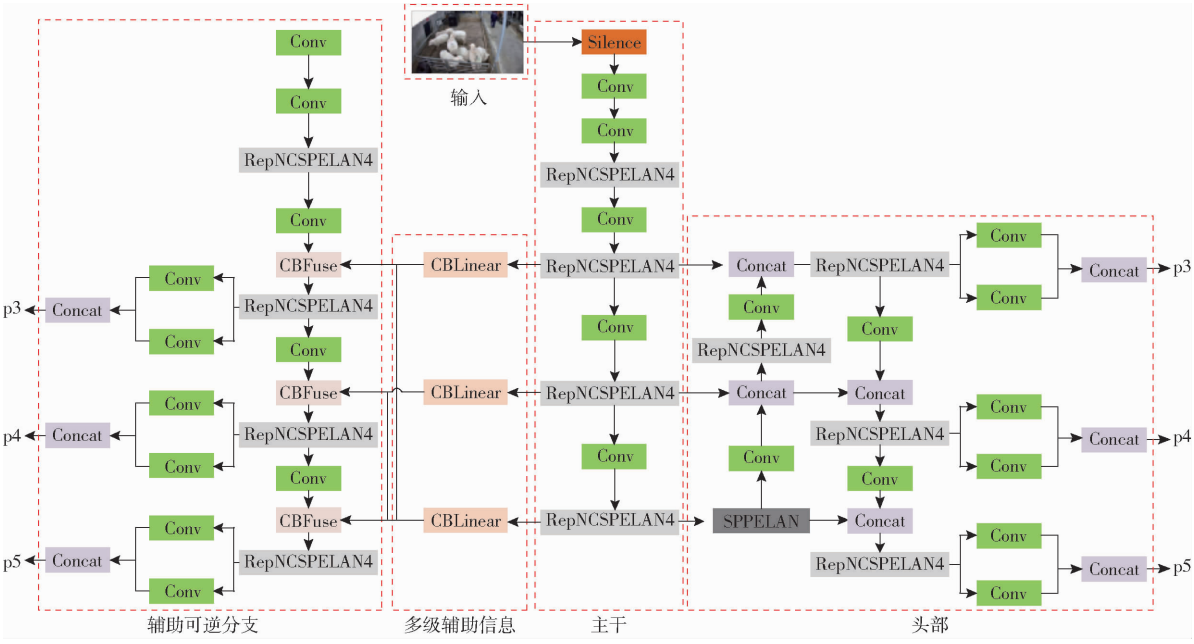


图 4 YOLO v9c 网络结构图

Fig. 4 Network structure diagram of YOLO v9c

2.2 多目标跟踪算法 ByteTrack

2.2.1 ByteTrack 算法原理

ByteTrack 是基于目标检测的跟踪算法,不采用 ReID 策略,没有深度方法,无需额外训练,仅使用目标跟踪所得到的检测框进行跟踪。当目标轨迹被遮挡严重时,目标检测算法检测出的被严重遮挡目标的检测框置信度会降低,ByteTrack 的最大创新点就是对低分检测框的使用,可以有效缓解被严重遮挡目标丢失跟踪的问题。ByteTrack[24-25]将检测器预测的检测框按照置信度一定的阈值分为高分检测框和低分检测框;首先,进行第 1 次关联,即高分检测框与轨迹之间的关联匹配。使用卡尔曼滤波器预测边界框位置,并与当前帧的高分框计算交并比(Intersection over union,IoU),获取两两之间 IoU 的关系损失矩阵。根据 IoU 关系损失矩阵,使用匈牙利算法对轨迹和检测框进行匹配。其次,进行第 2 次关联,即低分检测框与轨迹之间的关联匹配。使用低分检测框与未匹配成功的轨迹之间计算两两的 IoU 矩阵,再使用匈牙利算法进行匹配,若有未成功匹配的低分检测框,则直接丢弃。然后,进行后处理,对未激活轨迹再次与匹配失败高分检测框匹配,若匹配成功标记为已激活轨迹,否则丢弃轨迹。最

后，还未成功匹配的高分框若高于一定阈值，则生成新的未激活轨迹。

ByteTrack 算法虽在一定程度上缓解了被严重遮挡物体丢失跟踪的问题，但也可能带来跟踪目标 ID 频繁跳动问题，并且当跟踪目标长时间被遮挡或者移动速度过快时，ByteTrack 仍会频繁出现对跟踪目标丢失跟踪的问题。

2.2.2　ByteTrack 算法改进

为了进一步提高 ByteTrack 算法对群养羊只跟踪的准确性与可靠性，对 ByteTrack 做出两点改进：在匹配流程中加入时间距离匹配模块；采用 ID 延时分配机制，对成功匹配的新轨迹 ID 分配进行一定限制。如图 5 所示，其中，改进 1 为时间距离匹配模块，改进 2 为 ID 延时分配机制。

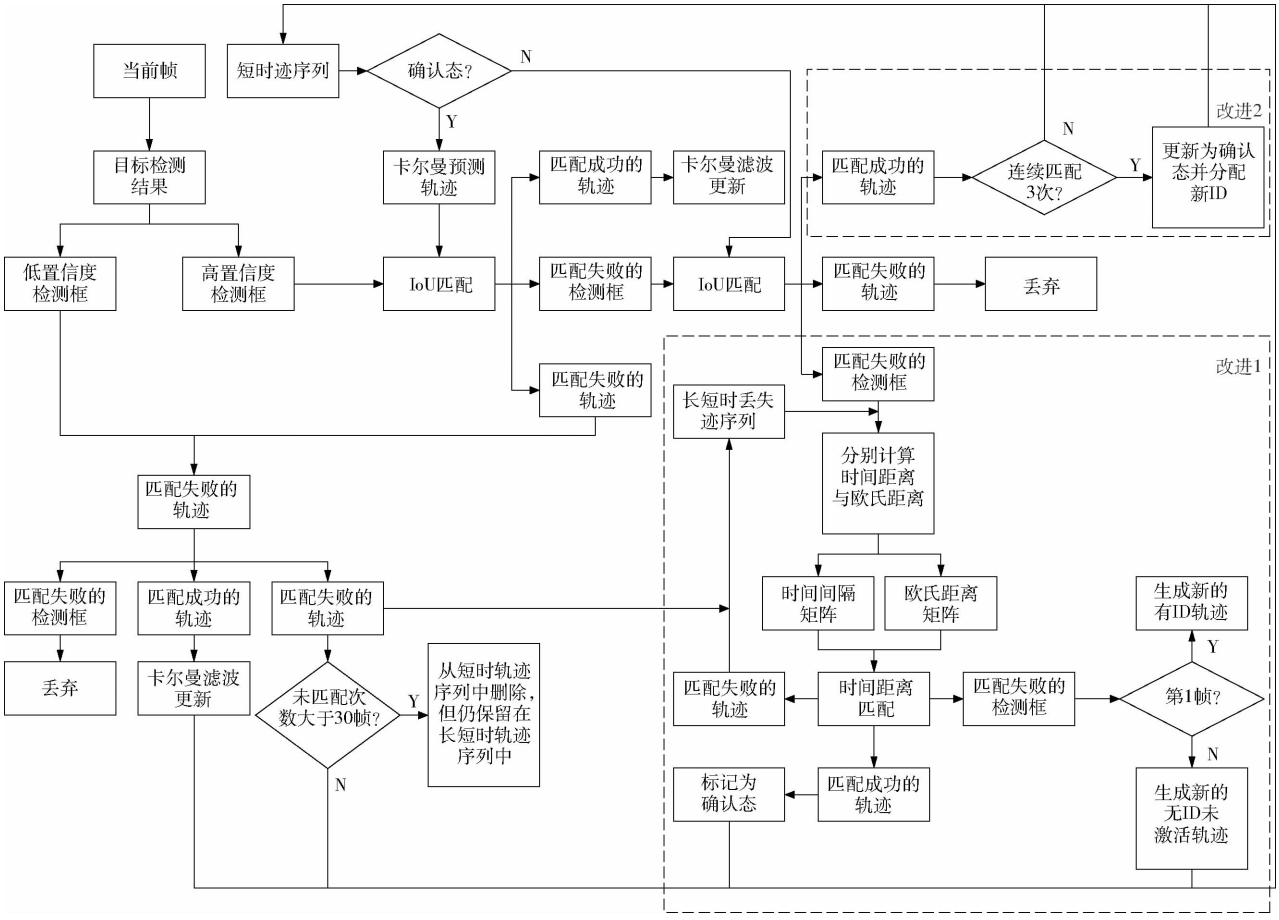


图 5　改进 ByteTrack 匹配流程图

Fig. 5　Improved ByteTrack matching flowchart

2.2.2.1　时间距离匹配模块

ByteTrack 算法在 3 次匹配结束后会获得 3 种结果：已匹配的轨迹框（Matched tracks）和边界框（Matched detections）、未成功匹配的轨迹（Lost tracks）、未成功匹配的高分边界框（Unmatched detections）。未成功匹配有两种原因：①正常状态，即出现新目标。②异常状态，即可能是旧目标被长时间遮挡后又恢复为未遮挡状态或移动速度过快。在多羊只养殖场景下，由于羊只生活习性，其会频繁性地挤在一起，这就有可能导致羊只长时间处在严重遮挡的状态下，以及由于羊只较为敏感，易受到惊吓而奔跑，导致跟踪器难以准确跟踪。

针对长时间遮挡后又恢复新状态这一场景，设计一种长短时丢失轨迹队列（Long and short time sequence of lost tracks，LSLT），可以保存短时间与长时间的丢失轨迹。针对目标长时间遮挡与移动速度过快易导致算法丢失跟踪，引入时间距离匹配模块（TDMM），即综合考虑丢失轨迹丢失时长以及丢失轨迹与每个未匹配预测框之间的欧氏距离（中心坐标之间距离）并再次进行匹配，如图 6 所示。TDMM 首先将未匹配成功的高分检测框（Hight score

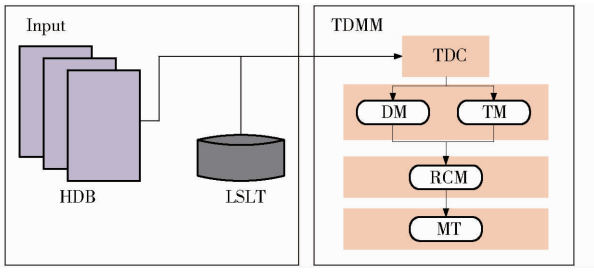


图 6　TDMM 结构图

Fig. 6　TDMM structure diagram

detection boxes, HDB)和长短时丢失轨迹送入时间距离矩阵计算器(Time matrix and distance matrix calculator, TDC),时间距离矩阵计算器首先计算两两之间欧氏距离 d_{ij}(负相关于轨迹与检测框之间的匹配程度)生成距离矩阵(Distance matrix, DM),计算公式为

$$d_{ij} = \sqrt{(t_{xi} - b_{xj})^2 + (t_{yi} - b_{yj})^2} \tag{1}$$

式中 d_{ij}——轨迹 i 与检测框 j 距离

t_{xi}、t_{yi}——轨迹 i 横、纵坐标

b_{xj}、b_{yj}——检测框 j 横、纵坐标

然后计算每个失跟踪轨迹丢失跟踪的时长并归一化 T_i(正相关于轨迹与检测框之间的匹配程度),得到时间矩阵(Time matrix, TM),计算公式为

$$T_i = \frac{l_i}{c} \tag{2}$$

式中 T_i——轨迹 i 丢失时长并归一化结果

l_i——轨迹 i 丢失时的帧

c——当前帧

最后结合两矩阵,计算轨迹与检测框两两之间的关系系数 f_{ij},得到关系系数矩阵(Relation coefficient matrix, RCM),计算公式为

$$f_{ij} = \alpha T_i + \beta \frac{1}{d_{ij} + \Delta} \tag{3}$$

式中 f_{ij}——轨迹 i 与检测框 j 的关系系数

α——调节时间权重

β——调节距离倒数权重

Δ——极小数,避免分母为 0

关系系数越大,轨迹与检测框相关性越大。利用关系系数矩阵,再进行轨迹与检测框之间的匹配,得到成功匹配的轨迹(Matched tracks, MT)并更新已匹配轨迹状态,未成功匹配的检测框进入下一个模块,未成功匹配的轨迹若丢失时长超过最大时长限制则丢弃。

2.2.2.2 ID 延时分配

ByteTrack 在 3 次匹配结束后,即高分框与已激活轨迹匹配、低分框与已激活轨迹匹配以及高分框与未激活轨迹匹配,如果仍未成功匹配的高分框置信度高于一定阈值则认为出现了新的目标,就会为该目标分配新的未激活轨迹以及新的 ID,在下一帧高分框与未激活轨迹匹配时,若匹配成功则将轨迹标记为已激活轨迹。在标准化的羊只养殖场景中,每个羊栏中羊只数量相对固定,很少会出现忽然多出一只羊的场景,且由于羊只外观特征极为相似,检测器难以准确分辨,对于新的高分目标直接分配新 ID,需要检测器确保测出来的确实为新目标,这对目标检测算法提出了很大挑战。为了解决这一问题,在 ByteTrack 跟踪器里加入 ID 延时分配机制。当出现新目标时(除第 1 帧外),ID 延时分配机制为新目标分配一个新的未激活轨迹,但不立即分配新 ID。在此后若高分框连续多次与未激活轨迹成功匹配后则认为出现新目标,将轨迹标记为已激活状态且分配新 ID,若有 1 次未匹配成功则认为是检测器误判,删除此未激活轨迹,不分配新 ID。

3 结果与分析

3.1 实验环境

硬件配置为 Intel(R) Core(TM) i5 12400F;16 GB 内存;1 个 NVIDIA GeForce RTX 3060Ti 8 GB 显存版本。软件环境为 Python 3.11;CUDA 12.0。

在目标检测部分,以 COCO 预训练模型作为初始权重,利用随机梯度下降法(Stochastic gradient, SGD)进行优化,初始学习率为 0.01,动量为 0.937,权重衰减为 5×10^{-4},迭代次数为 150,类别数为 3,图像尺寸统一调整为 640 像素 ×640 像素。

在多目标跟踪部分,均采用 YOLO v9c 作为目标检测部分,低分检测框置信度为 0.1 ~0.5,置信度得分大于 0.5 为高分检测框,高分匹配失败的检测框若置信度大于 0.6 则创建新轨迹。

改进 BytTrack 与其他算法进行对照实验时,改进 ByteTrack 的 ID 延时分配中连续匹配次数限制 N 取 3,时间系数 α 取 0.3,距离系数 β 取 600。

3.2 评价指标

为了验证 YOLO v9c 行为识别性能,采用精确率(Precision)、召回率(Recall)、平均精度均值(mAP@0.5、mAP@0.5:0.95)及参数量作为评价指标[26],其中 mAP@0.5 即固定 IoU 阈值为 0.5 时的平均精度的均值;mAP@0.5:0.95 即在 IoU 阈值从 0.5 到 0.95(步长 0.05)区间内,共 10 个不同阈值点上所测得的平均精度的均值。

跟踪部分采用多目标跟踪精度(Multiple object tracking accuracy, MOTA)、高阶跟踪精度(Higher order tracking accuracy, HOTA)、F1 得分(Identification F1 score, IDF1)、ID 跳变(ID switch, IDSW)作为多目标跟踪性能评价指标[19,27-28]。MOTA 是最广泛使用的指标,其综合考虑漏检个数(FN)、误检个数(FP)以及 ID 切换(IDSW)次数,值越大越好;HOTA 将执行精确检测、关联和定位的效果平衡到一个统一用于比较跟踪器的度量中;IDF1 是 F1 的改进版,用来衡量算法在识别并跟踪多个目标时的准确度,其值越大越好;IDSW 指同一个物体在视频中 ID 的变化次数,ID 每变化 1 次,ID 跳变就

增加 1 次，反映跟踪算法的稳定性和准确性，其值越小越好。为验证改进 ByteTrack 算法实时性，采用平均帧率作为评价指标。

3.3　YOLO v9c 目标检测模型实验结果

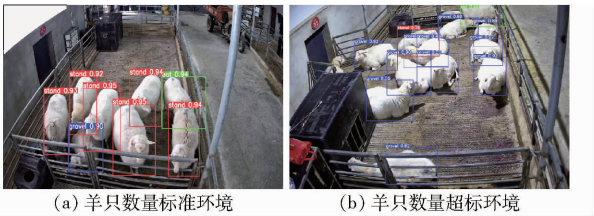
为了验证 YOLO v9c 目标检测效果，将其与 YOLO v9e 模型以及 YOLO v8 中 m、l、x 3 种不同大小的模型进行比较。由表 1 可知，YOLO v9c 模型在目标检测方面的精确率仅低于 YOLO v8l 模型，召回率高于 YOLO v8m 模型，mAP@0.5 仅低于 YOLO v8x 模型，mAP@0.5:0.95 都略低于其他 4 种模型。总体来看 YOLO v9c 和其他 4 种模型在目标检测的性能表现基本一致，但 YOLO v9c 模型参数量最低，略低于 YOLO v8m 模型，远低于 YOLO v9e、YOLO v8l、YOLO v8x 模型。YOLO v9c 模型相比于 YOLO v8m 模型，在参数量略低于其的情况下，精确率、召回率、mAP@0.5 分别较 YOLO v8m 模型高 0.2、2.9、1.6 个百分点，mAP@0.5:0.95 较 YOLO v8m 模型仅低 0.3 个百分点。

表 1　不同目标检测模型性能

Tab.1　Performance of different object detection models

模型	精确率/%	召回率/%	mAP@0.5/%	mAP@0.5:0.95/%	参数量
YOLO v9c	94.0	93.5	97.1	77.3	2.55×10^7
YOLO v9e	93.0	93.7	96.5	79.3	5.81×10^7
YOLO v8m	93.8	90.6	95.5	77.6	2.62×10^7
YOLO v8l	94.9	93.6	95.4	77.5	4.37×10^7
YOLO v8x	92.4	95.9	97.3	78.0	6.82×10^7

图 7 为训练后的 YOLO v9c 网络模型在不同场景下群养羊只目标检测效果。如图 7b 所示，在羊只数量超过标准养殖环境下，仍可提供较为精确的行为识别与目标检测效果，表明 YOLO v9c 可以在复杂场景下提供可靠的目标检测结果。故采用检测精度高、检测速度快的 YOLO v9c 模型作为群养羊只的目标检测部分，为后续的 ByteTrack 多目标跟踪算法提供优质的预测框。



(a) 羊只数量标准环境　　(b) 羊只数量超标环境

图 7　YOLO v9c 行为识别与目标检测效果

Fig.7　Effect of YOLO v9c behavior recognition and object detection

3.4　基于改进的 ByteTrack 多目标跟踪结果

3.4.1　消融实验

为了验证基于 ByteTrack，引入时间距离模块与 ID 延时分配机制的改进算法的有效性，进行消融实验，结果如表 2 所示。消融实验结果表明，单独采用本研究所设计的 TDMM 模块相比于 ByteTrack 算法 ID 跳变次数减少 5 次，MOTA、HOTA、IDF1 分别增加 0.191、2.134、5.61 个百分点。HOTA 达到 71.975%。IDF1 达到 88.113%。在此基础上对算法加入 ID 延时分配机制后，ID 跳变次数降低 7 次（降低约 46.7%），MOTA、HOTA 和 IDF1 得到进一步提升，分别为 88.326%、72.051% 和 88.237%。ByteTrack 算法经过加入 TDMM 和 ID 延时机制的双重改进后，相比于原来的算法，ID 跳变次数下降为 8 次，MOTA、HOTA 和 IDF1 分别提高 0.242、2.21、5.734 个百分点。

表 2　改进 ByteTrack 消融实验结果

Tab.2　Improved ByteTrack ablation assay

多目标跟踪算法	时间距离匹配模块	ID 延时分配机制	ID 跳变/次	多目标跟踪精度/%	高阶跟踪精度/%	F1 得分/%
ByteTrack	×	×	15	88.084	69.841	82.503
改进	√	×	10	88.275	71.975	88.113
	×	√	9	87.368	69.868	83.092
	√	√	8	88.326	72.051	88.237

注：√表示加入此改进模块，×表示不加入此改进模块。

为了更直观地说明本文算法的有效性，图 8 为不同环境下视频跟踪效果。图 8a 为夜晚场景，原 ByteTrack 算法中在 ID 为 7 的羊只由于运动速度快致使 ByteTrack 跟丢后，再次得到目标后赋予其新的 ID，其 ID 由 7 变为 12，后又变为 16；而改进后算法的 TDMM 则会在目标轨迹丢失后根据轨迹丢失的时间以及与预测框的位置信息再次进行匹配，防止本来 ID 为 7 的羊只被赋予新的 ID。图 8b 为白天场景，由于羊只拥挤在一起，在羊只未改变情形下原 ByteTrack 出现新的轨迹，后导致羊只 5 的 ID 发生

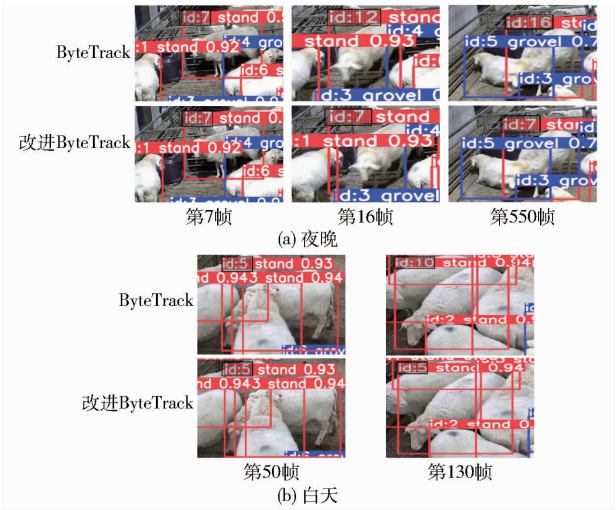


图 8　ByteTrack 与改进 ByteTrack 跟踪效果对比

Fig.8　Original ByteTrack and improved ByteTrack tracking effect

跳变,变为羊只 10;在改进 ByteTrack 中的 ID 延时分配机制发现当出现新的高置信度预测框时,并不会立即赋予新的 ID,而是会首先认为出现检测失误,当多次出现该新轨迹时,才会为其分配 ID,有效抑制 ID 频繁跳变的问题。

ID 延时分配中连续匹配次数限制:如图 9 所示,当连续匹配次数 N(创建新轨迹也算一次匹配)处于 1 ~ 3 时再分配新 ID,MOTA、IDF1 和 HOTA 呈明显上升趋势,ID 切换频率呈明显下降趋势;并且在 $N=3$ 时,各项指标达到最优。

时间系数 α 取值:如图 10 所示,可以看出当 α 从 0 到 0.3 时 MOTA、IDF1 和 HOTA 呈上升趋势,ID 切换频率呈下降趋势;在 α 在 0.3 之后各项性能指标趋于平缓,故 α 取 0.3。

距离系数 β 取值:如图 11 所示,在 β 处于 0 ~ 600 时,MOTA 和 IDF1 呈上升趋势,HOTA 先上升后稍微有所下降,ID 切换次数呈明显下降趋势。β 在 600 之后各项指标趋于平缓,故 β 取 600。

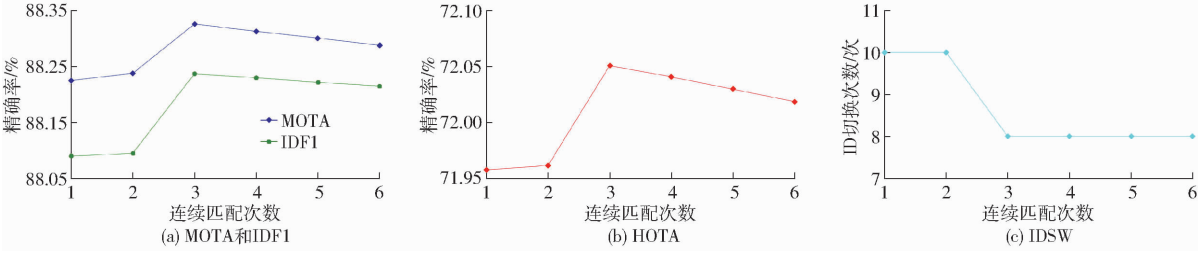


图 9 连续匹配次数限制与性能关系

Fig. 9 Relationship between limit of consecutive matches and performance

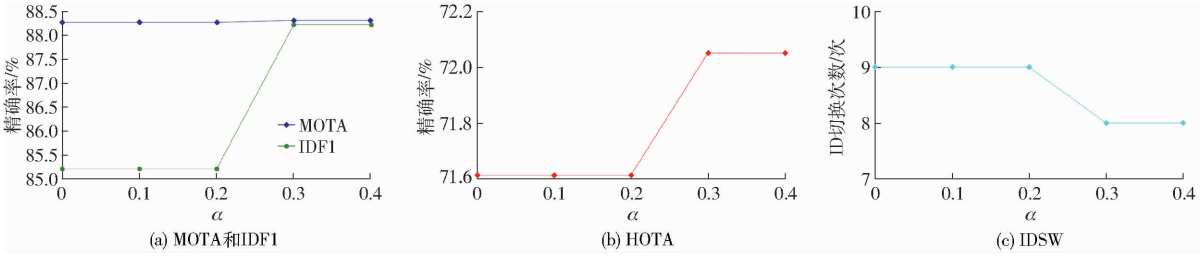


图 10 时间系数 α 与性能关系

Fig. 10 Relationship between time coefficient α and performance

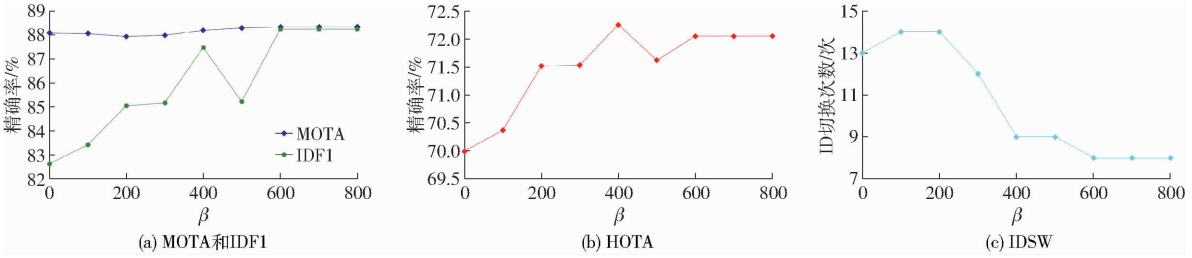


图 11 距离系数 β 与性能关系图

Fig. 11 Relationship between distance coefficient β and performance

3.4.2 多目标跟踪算法对比实验

为了进一步验证本文算法的有效性,与类似于 ByteTrack 的基于目标框运动与位置信息进行匹配的两种多目标跟踪算法(OC - SORT 和 Bot - SORT)进行对比,结果如表 3 所示。由表 3 可知,改进算法比 OC - SORT 算法 ID 跳变次数减少 29 次,MOTA 增加 2.623 个百分点,HOTA 增加 1.058 个百分点以及 IDF1 提高 3.572 个百分点。相比于 Bot - SORT,ID 跳变次数减少了 6 次,HOTA 增加 4.67 个百分点,IDF1 增加 8.83 个百分点。Bot - SORT 和 OC - SORT 均为基于卡尔曼滤波(Kalman filter,KF)运动模型的多目标跟踪器,在复杂场景下,检测器难以提供高质量的预测框时,卡尔曼滤波便会难以预测目标轨迹未来运动位置,可能导致目标轨迹 ID 频繁切换等问题,本研究通过在 ByteTrack 的基础改进可以在复杂场景难以预测目标轨迹运动位置时,再通过 TDMM 和 ID 延时分配机制有效匹配正确轨迹。综上可以验证算法的有效性和优势。

表 3 多算法对比结果

Tab. 3 Comparison results of multiple algorithms

多目标跟踪算法	ID 跳变次数/次	多目标跟踪精度/%	高阶跟踪精度/%	F1 得分/%
OC - SORT	37	85.703	70.993	84.665
Bot - SORT	14	88.805	67.360	79.392
本文	8	88.326	72.051	88.237

为验证改进 ByteTrack 算法的实时性，选择一段 100 s 的视频并用不同的跟踪算法对视频中的群养羊只进行行为识别与跟踪以验证各算法的实时性。为确保实验的严谨性，每个算法测试该视频段 3 次，再取平均值。表 4 为仅显示跟踪效果不保存跟踪结果的多算法效率对比，表 5 为既显示跟踪效果又保存跟踪结果的多算法效率对比。可见本文改进 ByteTrack 算法效率与 ByteTrack 算法和 OC－SORT 算法效率接近，且都远超 Bot－SORT 算法运行效率。

表 4　不保存结果的多算法效率对比结果

Tab. 4　Comparison results of multiple algorithms without saving results

多目标跟踪算法	平均预处理时长/ms	平均推理时长/ms	平均后处理时长/ms	平均总时长/ms	平均帧率/(f·s^{-1})
ByteTrack	1.4	12.3	1.2	14.9	67.1
OC－SORT	1.4	11.9	1.0	14.3	69.9
Bot－SORT	1.5	25.5	1.6	28.6	34.9
本文	1.5	12.9	1.1	15.5	64.5

表 5　保存结果的多算法效率对比结果

Tab. 5　Comparison of efficiency of multiple algorithms for saving results

多目标跟踪算法	平均预处理时长/ms	平均推理时长/ms	平均后处理时长/ms	平均总时长/ms	平均帧率/(f·s^{-1})
ByteTrack	1.6	17.1	1.2	19.9	50.2
OC－SORT	1.6	17.6	1.1	20.3	49.3
Bot－SORT	1.6	26.1	1.5	29.2	34.2
本文	1.6	18.4	1.2	21.2	47.1

3.4.3　羊只轨迹可视化

为了更直观地比较本改进算法相比于其他非 ReID 算法在实际养殖环境下的性能差距，图 12 为两段视频中的 4 只不同羊只的真实轨迹图与采用不同跟踪算法所得到的轨迹图的对比。图中羊只的 ID 每变化 1 次，其轨迹颜色也会随之改变。在对运动较为复杂频繁的轨迹 1 进行跟踪时，采用本算法跟踪时，ID 仅变化 1 次；而采用 ByteTrack 与 Bot－SORT 算法时，ID 变化 6 次；采用 OC－SORT 算法时，ID 变化 4 次。在对另外 3 个轨迹跟踪时，本算法的 ID 切换次数、跟踪稳定性都明显优于 ByteTrack、Bot－SORT 与 OC－SORT 3 种算法。

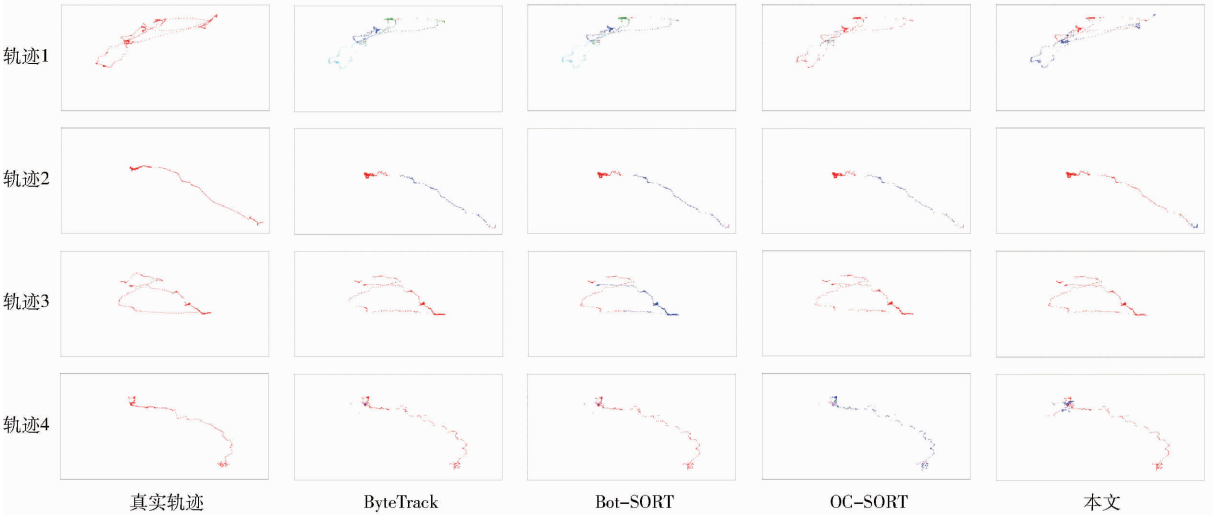


图 12　羊只活动轨迹

Fig. 12　Sheep activity trajectory

3.4.4　改进 ByteTrack 行为识别与跟踪

采用 YOLO v9c 与本文改进 ByteTrack 相结合，可以有效对群养羊只进行行为识别与多目标跟踪。如图 13 所示，YOLO v9c 对不同羊只的不同行为做到了精确识别；本文改进 ByteTrack 多目标跟踪算法可以长时间对群养羊只进行精准跟踪，显著缓解了由于遮挡导致的跟踪丢失问题。

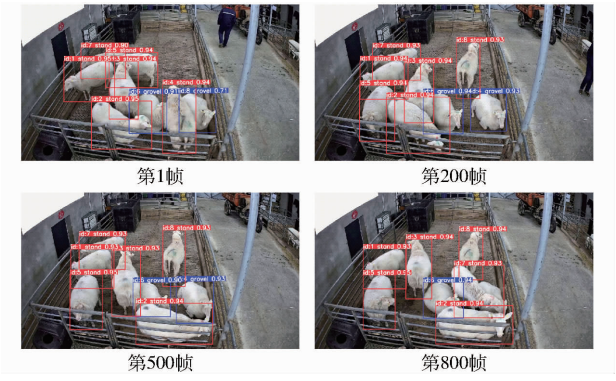


图 13　改进 ByteTrack 行为识别与跟踪效果

Fig. 13　Improved ByteTrack behavior recognition and tracking renderings

4　结论

(1)针对实际羊只饲养环境下，对羊只多目标跟踪时由于存在长时间遮挡或目标移动速度过快现

象导致羊只难以被持续稳定跟踪等问题,提出了将 YOLO v9 与改进 ByteTrack 相结合的 TBD 算法,实现了群养羊只的稳定跟踪。

(2)将 YOLO v9 作为 TBD 算法的目标检测模型,通过 YOLO v9 的 PGI 概念在不增加推理成本的基础上,使目标检测模型具有更好的性能表现,为后续改进 ByteTrack 算法提供了高质量的目标检测框。

(3)设计了时间距离匹配模块(TDMM),该模块的核心概念是丢失轨迹再匹配,主要步骤是建立一个长短时丢失轨迹队列,将队列中轨迹与高置信度检测框计算欧氏距离矩阵再结合各个轨迹丢失的时长得到关系系数矩阵。并将 TDMM 与 ID 延时分配机制一起运用在 ByteTrack 多目标跟踪算法上,在实际相对封闭的羊只养殖场景下,得到了较为可靠的跟踪结果。

(4)YOLO v9c 与改进 ByteTrack 算法相结合的 TBD 算法在符合实际羊只养殖的测试视频下获得了优于 ByteTrack、Bot - SORT 及 OC - SORT 的性能,ID 跳变次数大幅下降,MOTA 达到 88.326%,HOTA 达到 72.051%,IDF1 达到 88.237%。证明了改进 ByteTrack 多目标跟踪算法的有效性,为羊只养殖领域的研究与应用提供了一种有效的技术支持。

参 考 文 献

[1] 赵亮, 宋强. 肉羊养殖技术问题及创新[J]. 畜牧兽医科技信息,2023(6): 110 - 112.

[2] PLOMIN R,DEFRIES J C,LOEHLIN J C. Genotype-environment interaction and correlation in the analysis of human behavior [J]. Psychological Bulletin,1977,84(2):309.

[3] 金玫蕾. 动物行为学在现代生命科学研究中的应用[J]. 实验动物与比较医学,2008,28(1):1 - 3.

[4] CHENG M, YUAN H B,WANG Q F, et al. Application of deep learning in sheep behaviors recognition and influence analysis of training data characteristics on the recognition effect[J]. Computers and Electronics in Agriculture,2022,198:107010.

[5] AIVARENGA F A P,BORGES I,PALKOVIC L, et al. Using a three-axis accelerometer to identify and classify sheep behaviour at pasture[J]. Applied Animal Behaviour Science,2016,181:91 - 99.

[6] XU J,YU L,ZHANG J,et al. Automatic sheep counting by multi-object tracking[C]// IEEE International Conference on Visual Communications and Image Processing,2022:257.

[7] RANI R B,WAHAB D,DUNG G B D,et al. Cattle health monitoring and tracking system[J]. Recent Trends in Electronics and Communication,2022,777:789 - 795.

[8] ZHENG Z,QIN L. PrunedYOLO-Tracker: an efficient multi-cows basic behavior recognition and tracking technique[J]. Computers and Electronics in Agriculture,2022,213:108172.

[9] SU Q,TANG J,ZHAI M,et al. An intelligent method for dairy goat tracking based on Siamese network[J]. Computers and Electronics in Agriculture,2022,193:106636.

[10] 苏影超. 基于机器视觉的羊只行为识别研究[D]. 呼和浩特:内蒙古大学,2023.
SU Yingchao. Study on sheep behavior recognition based on machine vision[D]. Hohhot:Inner Mongolia University,2023. (in Chinese)

[11] 袁洪波,曹润柳,程曼. 融合 Res3D、BiLSTM 和注意力机制的羊只行为识别方法[J]. 农业机械学报,2024,55(4):221 - 230.
YUAN Hongbo, CAO Runliu, CHENG Man. Fusion of Res3D, BiLSTM and attention mechanism for sheep behavior recognition method[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(4):221 - 230. (in Chinese)

[12] CAO Y,CHEN J,ZHANG Z. A sheep dynamic counting scheme based on the fusion between an improved-sparrow-search YOLOv5x - ECA model and few-shot deepsort algorithm[J]. Computers and Electronics in Agriculture,2023,206:107696.

[13] 刘军, 后士浩, 张凯, 等. 基于增强 Tiny YOLOV3 算法的车辆实时检测与跟踪[J]. 农业工程学报,2019,35(8): 118 - 125.
LIU Jun,HOU Shihao,ZHANG Kai,et al. Real-time vehicle detection and tracking based on enhanced Tiny YOLOV3 algorithm [J]. Transactions of the CSAE,2019,35(8):118 - 125. (in Chinese)

[14] 涂淑琴, 黄磊, 梁云, 等. 基于 JDE 模型的群养生猪多目标跟踪[J]. 农业工程学报, 2022, 38(17): 186 - 195.
TU Shuqin, HUANG Lei, LIANG Yun, et al. Multiple object tracking of group-housed pigs based on JDE model[J]. Transactions of the CSAE,2022,38(17):186 - 195. (in Chinese)

[15] GUO Q,SUN Y, MIN L,et al . Video-based detection and tracking with improved re-identification association for pigs and laying hens in farms[C]// Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications,2022: 69 - 78.

[16] 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)

[17] 杨断利,齐俊林,陈辉,等. 基于改进 YOLO v8n 模型的散养蛋鸡个体行为识别方法与差异分析[J]. 农业机械学报,2024,55(11):112 - 123.
YANG Duanli,QI Junlin,CHEN Hui,et al. Individual behavioral identification and differential analysis of free-range laying hens based on improved YOLO v8n model[J]. Transactions of the Chinese Society for Agricultural Machinery,2024, 55(11):112 - 123. (in Chinese)

[18] 刘成,岳训. 基于 YOLO v5 + DeepSORT 算法的羊群游走同步群体决策行为研究[J]. 农业机械学报,2024,55(6):229 - 236.
LIU Cheng, YUE Xun. Experiment of synchronized group decision-making behavior under herding walk model based on YOLO v5 + DeepSORT algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(6): 229 - 236. (in Chinese)

[19] 黄成龙,张忠福,华向东,等. 基于改进 Faster R - CNN 和 Deep Sort 的棉铃跟踪计数[J]. 农业机械学报,2023,54(6):205 - 213.
HUANG Chenglong, ZHANG Zhongfu, HUA Xiangdong, et al. Cotton boll tracking and counting based on improved Faster R - CNN and Deep Sort[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 205 - 213. (in Chinese)

[20] CAO J, WENG X, PANG J, et al. Observation-Centric SORT: rethinking sort for robust multi-object tracking[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 9686 - 9696.

[21] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT - SORT: robust associations multi-pedestrian tracking[EB/OL]. (2022 - 06 - 29)[2024 - 04 - 28]. https://arxiv.org/abs/2206.14651.

[22] WANG C, YEH I, LIAO H. YOLOv9: learning what you want to learn using programmable gradient information[EB/OL]. (2024 - 02 - 21)[2024 - 04 - 28]. https://arxiv.org/abs/2402.13616.

[23] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 8759 - 8768.

[24] ZHANG Y, SUN P, JIANG Y, et al. ByteTrack: multi-object tracking by associating every detection box[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2022: 1 - 21.

[25] 王雒,李飚,傅瑞罡. 基于改进 ByteTrack 算法的红外地面多目标跟踪方法[J]. 计算机科学,2023,50(9):176 - 183.
WANG Luo, LI Biao, FU Ruigang. Infrared ground multi-object tracking method based on improved ByteTrack algorithm[J]. Computer Science, 2023, 50(9): 176 - 183. (in Chinese)

[26] PADILLA R, NETTO S L, DA SILVA E A B. A survey on performance metrics for object-detection algorithms[C]// 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). IEEE, 2020: 237 - 242.

[27] 韩锟,彭晶莹. 基于改进 YOLOX 与多级数据关联的行人多目标跟踪算法研究[J]. 铁道科学与工程学报,2024, 21(1): 94 - 105.
HAN Kun, PENG Jingying. Pedestrian multi-object tracking algorithm based on improved YOLOX and multi-level data association[J]. Journal of Railway Science and Engineering, 2024, 21(1): 94 - 105. (in Chinese)

[28] 梁秀英,贾学镇,何磊,等. 基于 YOLO v8n-seg 和改进 Strongsort 的多目标小鼠跟踪方法[J]. 农业机械学报,2024, 55(2):295 - 305,345.
LIANG Xiuying, JIA Xuezhen, HE Lei, et al. Multi-target mice tracking based on YOLO v8n-seg and improved Strongsort[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 295 - 305, 345. (in Chinese)

(上接第 557 页)

[24] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad - CAM: visual explanations from deep networks via gradient-based localization[J]. arXiv Preprint, arXiv: 1610.02391, 2016.

[25] ZHU X K, LYU S C, WANG X, et al. TPH - YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2778 - 2788.

[26] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLO v7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. arXiv Preprint, arXiv: 2207.02696, 2022.

[27] LOU H T, DUAN X H, GUO J M, et al. DC - YOLO v8: small-size object detection algorithm based on camera sensor[J]. Electronics, 2023, 12(10): 2323.

[28] WANG C Y, YEH I H, LIAO H Y M. YOLO v9: learning what you want to learn using programmable gradient information[J]. arXiv Preprint, arXiv: 2402.13616, 2024.