

基于 DeepSORT 算法的肉牛多目标跟踪方法

张宏鸣^{1,2} 汪 润¹ 董佩杰¹ 孙红光¹ 李书琴^{1,2} 王红艳^{2,3}

(1. 西北农林科技大学信息工程学院, 陕西杨凌 712100; 2. 宁夏智慧农业产业技术协同创新中心, 银川 750004; 3. 西部电子商务股份有限公司, 银川 750004)

摘要: 肉牛的运动行为反映其健康状况,在实际养殖环境下如何识别肉牛并对其进行跟踪,对感知肉牛的运动行为至关重要。基于 YOLO v3 改进算法 (LSRCCEM - YOLO),利用视频监控实现了实际养殖环境下的肉牛实时跟踪。该方法采用 MobileNet v2 作为目标检测骨干网络,根据肉牛分布不均、目标尺度变化较大的特点,提出通过添加长短距离语义增强模块 (LSRCCEM) 进行多尺度融合,结合 Mudeep 重识别模型实现了肉牛多目标跟踪。结果表明:在目标检测方面,LSRCCEM - YOLO 的 mAP 值达到了 92.3%,模型参数量仅为 YOLO v3 的 10%,相比 YOLO v3 - tiny 也降低了 31.34%;在肉牛重识别方面,采用基于调整感受野的 Mudeep 模型,获得了更多的多尺度特征,其 Rank - 1 指标达到了 96.5%;多目标跟踪的多目标跟踪准确率相对于 DeepSORT 算法从 32.3% 提高到了 45.2%,ID switch 次数降低了 69.2%。本文方法可为实际环境下的肉牛行为实时跟踪、行为感知提供技术参考。

关键词: 肉牛; 多目标跟踪; 目标检测; 重识别; 注意力机制; 长短距离语义增强模块

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298 (2021)04-0248-09 OSID: 

Beef Cattle Multi-target Tracking Based on DeepSORT Algorithm

ZHANG Hongming^{1,2} WANG Run¹ DONG Peijie¹ SUN Hongguang¹ LI Shuqin^{1,2} WANG Hongyan^{2,3}

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

2. Ningxia Intelligent Agricultural Industry Technology Collaborative Innovation Center, Yinchuan 750004, China

3. West China Electronic Bussiness Co., Ltd., Yinchuan 750004, China)

Abstract: The behavior of beef cattle reflects its health status. How to recognize and track beef cattle in a real breeding environment is very important to perceive the behavior of beef cattle. Wearable devices have limited accuracy in sensing motion behavior and are easily damaged, while monitoring devices are widely used in farms and have a long lifespan. Based on the improved YOLO v3 algorithm (LSRCCEM - YOLO), surveillance video was used to achieve real-time tracking of beef cattle in a real breeding environment. MobileNet v2 was used as the object detection backbone network. According to the uneven distribution of beef cattle and the large change of target scale, long-short range context enhancement module (LSRCCEM) was proposed for multi-scale fusion, combined with the Mudeep ReID model to achieve multiple targets for beef cattle track. The experimental results showed that in beef cattle object detection, the mAP index of LSRCCEM - YOLO reached 92.3%, and the model parameter amount was only 10% of YOLO v3, which was also reduced by 31.34% compared with YOLO v3 - tiny; in terms of beef cattle re-identification (ReID), based on adjusting the Mudeep model of the receptive field obtained more multi-scale features, and its Rank - 1 index reached 96.5%. Compared with the original DeepSORT algorithm, the MOTA index of multi-target tracking was increased from 32.3% to 45.2%, and the number of ID switch was decreased by 69.2%. This method can provide technical reference for real-time tracking and behavior perception of beef cattle in real environment.

Key words: beef cattle; multi-target tracking; object detection; re-identification; attention mechanism; LSRCEM

收稿日期: 2020 - 07 - 22 修回日期: 2020 - 08 - 16

基金项目: 国家重点研发计划项目 (2020YFD1100601)、宁夏智慧农业产业技术协同创新中心项目 (2017DC53) 和国家自然科学基金项目 (41771315)

作者简介: 张宏鸣 (1979—),男,教授,博士,主要从事空间大数据管理与智慧农业研究, E-mail: zhbm@nwsuaf.edu.cn

0 引言

牛肉是人类主要肉食之一^[1]。肉牛的行为可以反映其健康状况^[2-3],进而反映牛肉的品质。近年来,随着人们对牛肉需求的不断增加,肉牛的养殖规模不断扩大,精细化管理难度增加。实时感知肉牛的运动行为、掌握其运动状态有助于及时发现肉牛异常,从而保证牛肉质量和产量。多目标实时跟踪是肉牛行为实时感知的基础,对于肉牛养殖具有重要意义。

目前,众多学者在动物行为的实时感知方面进行了诸多研究。部分学者研制了可穿戴设备,利用安装在动物身上的追踪设备收集动物的运动信息,获取其行为数据^[4-6],再借助机器学习^[7-8]、深度学习^[9-10]方法分析数据,从而实现对动物行为的监测。由于动物行为的不确定性,可穿戴设备的损坏率较高^[11],导致制作、维护成本增加,不能满足大型养殖场的使用需求。相比之下,利用养殖场已有的监控视频对牲畜进行管理具有使用便利、普及度高、成本低的优点,非常适合大型养殖场对牲畜的精细化管理。

利用监控视频对牲畜进行跟踪需满足实时性要求,而多目标跟踪的速度在很大程度上取决于其检测到目标的速度。传统的目标检测算法如背景减法^[12]、帧间差分法^[13]和光流法^[14-15]无法满足实时性的要求。目前,基于深度学习的目标检测算法不断成熟,该算法可划分为 One-stage 和 Two-stage 检测算法两类。One-stage 算法主要有 YOLO 系列算法^[16-18]、SSD^[19-21]、RetinaNet^[22]等。何东健等^[23]采用 YOLO v3 算法,通过优化锚点框和改进网络结构实现了对挤奶奶牛的个体识别,识别准确率达 95.91%,mAP 为 95.16%。Two-stage 算法主要有 R-CNN^[24]、Fast R-CNN^[25]、Faster R-CNN^[26]等。SUN 等^[27]利用 Faster R-CNN 算法对猪进行多目标跟踪,该算法具有良好的抗遮挡能力,能较为精准地跟踪多个目标。相对于 Two-stage 算法,One-stage 算法速度更快、实时性更高,而 YOLO v3 算法精度高于大部分 One-stage 算法中的其他检测算法。

在多目标跟踪算法方面,基于匈牙利算法(KM 匹配)的后端追踪优化算法,如 SORT^[28]、DeepSORT^[29]算法,能够满足实时跟踪的要求,其中 DeepSORT 算法在 SORT 算法的基础上,通过提取深度表观特征明显提高了多目标的跟踪效果。然而其跟踪效果依赖目标检测器的精确度和特征区分程度,其跟踪速度与目标检测速度密切相关。

基于此,本文提出一种改进 YOLO v3 算法和 Mudeep 目标重识别算法相结合的 DeepSORT 肉牛多目标跟踪方法,针对 YOLO v3 算法在多尺度环境下识别效果较差的问题,通过添加长短距离语义增强模块(Long-short range context enhancement module,LSRCEM)进行多尺度融合,建立 LSRCEM-YOLO 模型,提高 YOLO v3 算法的检测精度,通过修改 YOLO v3 模型的网络结构,简化模型复杂度,增强实时性,通过 Mudeep 目标重识别算法提取肉牛具有鉴别性的特征,减少标号变化现象,以期实现肉牛多目标实时跟踪。

1 材料与方法

1.1 视频采集

实验视频采集于某肉牛养殖场,为其养殖场内部监控设备所拍摄视频,拍摄对象为成年安格斯牛与西门塔尔牛,摄像机位于矩形肉牛养殖棚短边的天棚上,斜向下俯视拍摄,能够拍摄到牛棚内的大部分肉牛活动区域。

供试视频拍摄时间为 2019 年 7 月 21—29 日,共 1 328 段视频,平均时长为 35 min,经筛选,剔除夜间、无目标、镜头污染等无效片段,保留有效视频 66 段,MP4 格式,分辨率为 1 920 像素(水平)×1 080 像素(垂直),视频帧率为 24 f/s,视频场景如图 1 所示。



图 1 养殖场监控视频画面
Fig.1 Monitoring video of the farm

1.2 技术路线

本研究旨在利用深度学习方法通过视频监控对肉牛进行实时跟踪,本文方法技术路线如图 2 所示,包括 5 部分:

- (1)数据集构建:筛选可用视频,对视频进行剪切,提取视频关键帧,通过对关键帧进行标注、裁剪,分别构建肉牛目标检测、目标重识别、目标跟踪数据集。
- (2)肉牛目标检测模块训练:利用肉牛目标检测数据集训练 LSRCEM-YOLO 模型用于肉牛目标检测,作为多目标跟踪的检测器。
- (3)肉牛重识别网络构建:利用肉牛重识别数

据集训练 Mudeep^[30]用于肉牛目标重识别。

(4) 肉牛多目标实时跟踪的实现: 结合 LSRCEM - YOLO 模型、Mudeep 模型与 DeepSORT 模型, 实现肉牛的多目标实时跟踪。

(5) 结果分析: 通过对目标检测效果、重识别效果、多目标跟踪效果进行评价, 全面评估肉牛多目标跟踪模型的效果。

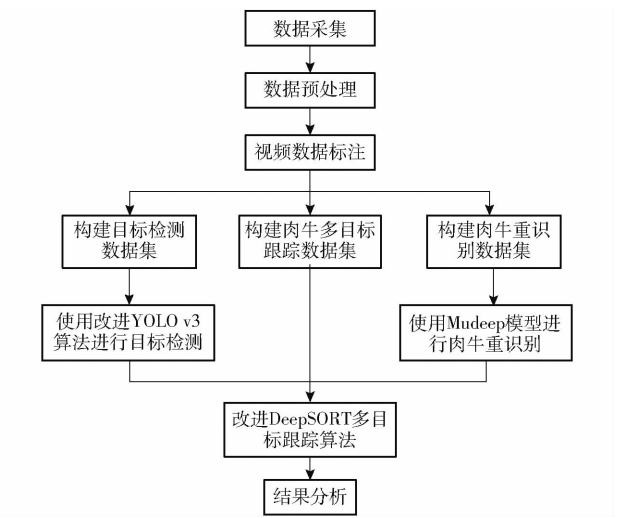


图2 本文方法技术路线图
Fig.2 Technology roadmap of the proposed method

1.3 数据集构建

构建3个数据集分别用于训练目标检测模型、重识别模型和验证多目标跟踪效果。视频中的肉牛个体之间花色差别较大、姿态复杂, 为了增强数据集的多样性, 将有效视频分为两部分处理, 第1部分使用 DarkLabel 进行视频标注, 每10帧提取1帧加入数据集。第2部分首先使用 ffmpeg 进行关键帧提取, 使用 LabelImg 进行人工标注。两部分数据集构成了肉牛目标检测数据集, 最终确定了7456幅图像作为训练集, 1491幅图像作为测试集。肉牛目标检测部分数据如图3所示。



图3 目标检测数据集样例
Fig.3 Samples of object detection dataset

为了提高重识别结果的准确性, 首先需要通过 对视频数据进行人工筛选, 保证所有视频中的肉牛个体都是唯一存在的, 不同视频中不能出现同一头

牛, 再使用 DarkLabel 对视频进行标注, 根据标注过程中的不同标号来区分不同个体, 最后按照 Market - 1501^[31]数据集格式构建肉牛重识别数据集。肉牛重识别数据集包含99头不同肉牛的共6195幅图像, 平均每头牛约30幅图像为训练集, 30幅图像为测试集。肉牛重识别数据如图4所示。



图4 重识别数据集样例
Fig.4 Samples of re-identification dataset

为验证本文提出的肉牛多目标跟踪方法的有效性, 选取实际场景下的监控视频进行测试, 视频长度为54s, 来验证本文方法追踪效果。

1.4 多目标跟踪方法

本文结合 DeepSORT 算法实现肉牛多目标跟踪, DeepSORT 算法流程如图5所示, 主要包含4个步骤: ①对视频进行预处理, 使用 OpenCV 将视频处理成视频帧, 然后使用目标检测算法提取深度特征得到候选框, 使用非极大值抑制 (NMS) 算法去除重叠框, 得到最终检测结果。②使用卡尔曼滤波算法对目标在视频中下一帧的位置和状态进行预测, 对比检测器提供的检测框, 将置信度更高的检测框作为预测结果。③使用匈牙利算法对前后两帧之间若干目标进行最优匹配, 得到目标在视频中的轨迹, 其中利用重识别算法计算余弦距离来度量表观特征的距离, 减少标号变化 (ID switch) 的发生。④输出结果, 同时对跟踪器进行参数更新, 重新开始目标检测。

DeepSORT 算法是一个两阶段的算法, 目标检测算法关系到其跟踪的准确度与速度, 重识别算法可以减少跟踪过程中 ID switch 的发生, 因此可以将其中的目标检测模型和重识别模型分别进行优化, 最终得到优化后算法。

1.4.1 目标检测

YOLO v3 是一个比较成熟的单阶段目标检测算法, 其在匹配过程中将目标图像分为 $S \times S$ 个格子,

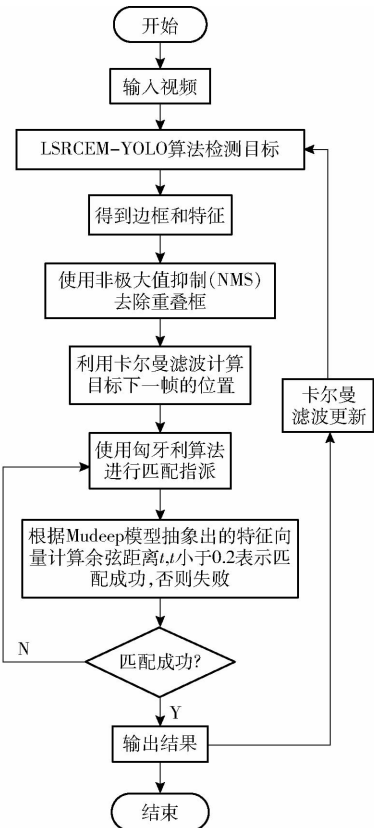


图 5 多目标跟踪算法流程图

Fig. 5 Multi-target tracking algorithm flow chart

每个单元格负责检测中心点落于此单元格的目标,每个单元格需要预测出 B 个边界框,每个边界框包含 5 个预测输出量 (x, y, w, h, conf), 其中 x, y 为目标框中心坐标值, w, h 表示边界框的长和宽, conf 表示框中含有目标的可信度。若检测类别总数为 C , 则每个单元格需预测 $B \times (5 + C)$ 个值。YOLO v3 借鉴残差网络结构,在 YOLO v2 中 Darknet19 的基础上,通过不断堆叠 $3 \times 3, 1 \times 1$ 的卷积和残差连接构成了含 53 个卷积层的深度神经网络 Darknet53^[32]。考虑到本文采集的视频中肉牛经常处于远离摄像头位置以及 YOLO v3 对小目标检测的良好效果,本文采用 YOLO v3 作为肉牛多目标跟踪模型的目标检测模块,同时,为了提高模型在边缘设备上的运行速度,本文对 YOLO v3 模型进行了改进:

(1) 使用 MobileNet v2 替换 Darknet53 作为 YOLO v3 的骨干网络。MobileNet v1 利用深度可分离结构在牺牲较小检测精度的情况下,大幅降低了计算量与网络层参数,降低了模型内存。MobileNet v2 在 MobileNet v1 的基础上,使用残差结构(图 6),进行先升维、再卷积、再降维的策略,在保持计算量变化不大的情况下获得了更高的检测精度。

(2) 利用长短距离语义增强模块 (LSRCCEM) (图 7) 融合 YOLO v3 的 3 个检测分支结果的同时自

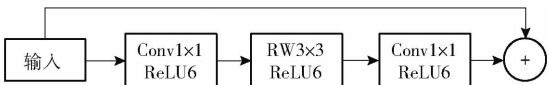


图 6 MobileNet v2 的残差结构

Fig. 6 Residual structure in MobileNet v2

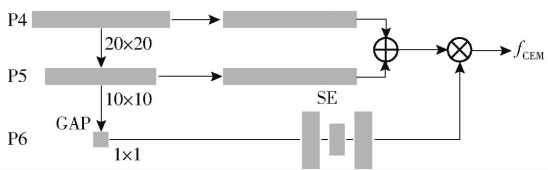


图 7 长短距离语义增强模块结构图

Fig. 7 Long-short range context enhancement module structure diagram

适应学习目标的多尺度特征。LSRCCEM 参考 ThunderNet^[33] 中的 CEM 模块,CEM 通过扩大模型感受野,融合多尺度局部信息和全局信息来获取更有鉴别力的特征。但是骨干网络中不同层所包含的语义信息并不相同,浅层网络包含物体的边缘、颜色、角点等信息;深层网络包含抽象的语义信息,然而 CEM 使用的特征融合方式是将各个模块的特征相加,相当于将所有层视为平等,忽略了各个层的不同,故引入长距离注意力机制和短距离注意力机制,通过两者配合来扩大模型感受野,并通过动态调整高层信息与浅层信息的分配来完成特征融合。

短距离注意力机制启发自 YOLO v3 - SPP 中的 SPP 模块,在空间最大池化模块 (Spatial MaxPooling Module) 中,3 个分支各自使用了 3 个 13×13 的池化核来进行特征提取;长距离注意力机制使用压缩激活模块 (SE),将特征金字塔高层提取到的信息以相乘的方式融合到浅层网络,以深层网络的语义信息引导低层网络的语义提取。YOLO v3 模型与 LSRCCEM - YOLO 结构如图 8 所示。

1.4.2 肉牛重识别

Mudeep 模型主要由多尺度层、降维层和显著性融合层组成 (图 9)。多尺度层由 Multi-scale-A 和 Multi-scale-B 组成,使用多分支结构实现多尺度特征的提取,每个分支的感受野不同,提升了特征的鉴别力和多样性。降维层 (Reduction) 位于两个多尺度层之间,用来降低特征的空间分辨率,同时使用了多分支的方式,综合最大池化层和步长卷积以弥补由于减少层数带来的空间信息损失。显著融合层 (Saliency-based learning fusion) 用来融合多尺度层的输出,使用显著性的学习策略筛选多尺度层输出的冗余信息,将多尺度层的输出通过相加的方式进行融合,然后将特征图通过全局平均池化得到一个归一化后的特征向量。Mudeep 通过引入空间可分离卷积,将普通的 3×3 卷积分解为 1×3 和 3×1 卷积,降低了模型所需的运算量,提高了模型的运算速

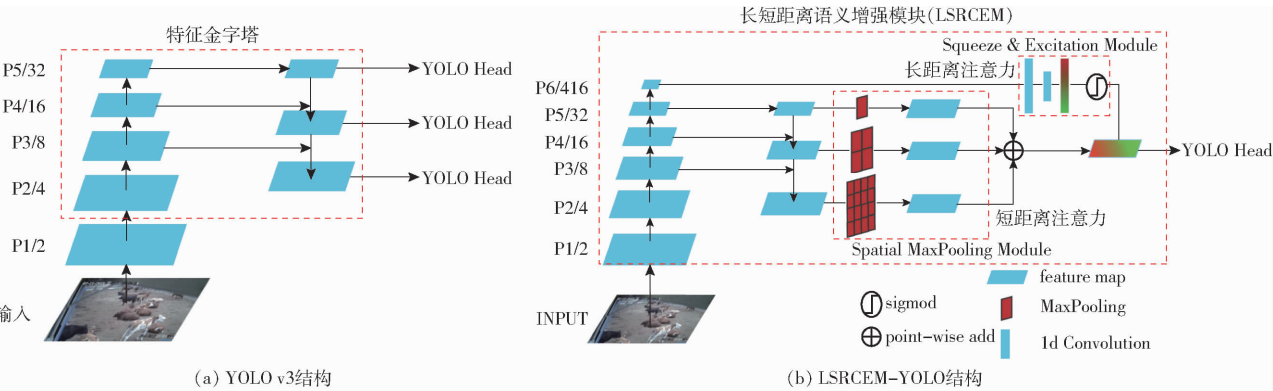


图 8 YOLO v3 改进前后的结构图

Fig. 8 Structure diagram before and after improvement of YOLO v3

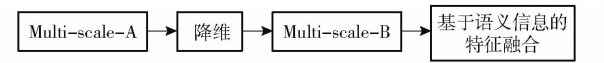


图 9 Mudeep 重识别模型总体框架

Fig. 9 General framework of Mudeep

度,且在 CUHK01、CUHK03、VIPeR 数据集上都有优秀的表现。

由于 Mudeep 相对于 Wide ResNet 具有更小的运算量,能够更好实现肉牛的实时跟踪,故本文选用 Mudeep 取代 Wide ResNet 作为 DeepSORT 中的重识别模块,通过肉牛的重识别,获取不同肉牛之间更有区分性的特征,降低跟踪过程中 ID switch 发生的次数。

1.5 评价指标

1.5.1 目标检测评价指标

目标检测与分类任务不同,检测模型输出的结果是非结构化的,无法事先得知检测得到的物体数量、位置、大小等,因此检测任务需要引入交并比 (IoU) 来量化预测框 A 和真实框 B 的贴合程度,设置阈值为 0.5,若 IoU 大于 0.5,则认为正确检测,否则认为是错误检测,交并比计算公式为

I_{oU} = \frac{S_{A \cap B}}{S_{A \cup B}} \tag{1}

式中 S_{A \cap B}——预测框与真实框重叠区域面积
S_{A \cup B}——预测框与真实框覆盖区域面积
I_{oU}——预测框与真实框的交并比

一般目标图中存在背景和目标,预测框也分为正确和错误,因此产生 4 种样本,真正例 (True positive, TP) 表示预测框与真实框正确匹配,假正例 (False positive, FP) 表示模型对正例进行错误分类,一般指 IoU 小于阈值的检测目标,假反例 (False negative, FN) 表示模型未检测到的真实框,真反例 (True negative, TN) 表示本身是背景,也未被检测出来。

根据以上 4 种样本,引入准确率 (Precision, P) 和召回率 (Recall, R) 来评价分类效果,准确率表示

被分为正例的所有实例中真正例占比,值越大说明分类越精确;召回率表示被分为正例的实例占有真实框的比例,值越大表示检测覆盖程度越好。然而准确率与召回率是对立的关系,若准确率提高,召回率会有一定的下降,为了平衡两者,引入准确率和召回率的调和平均数 F1 (F1-score) 和平均检测精度 (mAP) 来综合衡量。以召回率为横轴,准确率为纵轴,绘制准确率-召回率 (PR) 曲线,则 PR 曲线与坐标轴围成的面积为 mAP。mAP 是目标检测中重要的评价标准,能够较好地评价目标检测模型。

此外,引入模型的参数量 (Params) 来评价模型的复杂程度,参数量越小,模型所需计算量越少,越适用于边缘设备。

1.5.2 肉牛重识别评价指标

DeepSORT 模型中使用的重识别模型的主要功能是用于提取具有鉴别力的特征,能够让前后两帧属于相同目标的对象相似度更高,让不同目标间的相似度降低,以降低多目标跟踪中 ID switch 现象的发生。Rank-n 表示检测结果中置信度最高的 n 幅图像中有正确结果的概率,因此本文使用 Rank-n 作为重识别模型的评价指标。

1.5.3 多目标跟踪评价指标

本文选用 4 个指标评价多目标跟踪效果:
(1) ID switch 表示一条跟踪轨迹中改变目标标号的次数,值越小越好。
(2) 多目标跟踪准确率 (Multiple object tracking accuracy, MOTA),主要考虑跟踪过程中所有对象的匹配错误,主要是 FP、FN、ID switch。MOTA 给出的是非常直观的衡量跟踪算法在检测物体和保持轨迹时的性能,与目标检测进度无关,MOTA 值越大表示模型的性能越好。MOTA 计算公式为

M_{OTA} = 1 - \frac{\sum (A_{FP} + A_{FN} + A_{ID})}{\sum A_{GT}} \tag{2}

式中 A_{FP}——假正例出现次数

A_{FN} ——假反例出现次数
 M_{OTA} ——多目标跟踪准确率
 A_{ID} ——ID switch 次数
 A_{GT} ——标注的目标个数

(3)多目标跟踪精确度 (Multiple object tracking precision, MOTP), 用来量化检测器的定位精度, MOTP 的值越大表示检测器的精度越大,计算公式为

$$M_{OTP} = \frac{\sum_{t,i} d_{i,t}}{\sum_t c_t} \times 100\%$$
 (3)

式中 c_t ——当前帧匹配成功的数目
 $d_{i,t}$ ——检测框和真实框的交并比
 i ——当前检测目标
 t ——帧序号
 M_{OTP} ——多目标跟踪精确度

(4)FPS,模型每秒处理的图像帧数,值越大处理效果越好。

2 结果与分析

2.1 目标检测结果分析与精度评价

对比 YOLO v3 算法、YOLO v3 - tiny 算法以及 LSRCEM - YOLO 算法在肉牛目标检测数据集上的实验结果如表 1 所示。

表 1 不同算法的目标检测结果对比
Tab.1 Comparison of target detection results of different algorithm

模型	准确率/%	召回率/%	mAP/%	F1/%	参数量
YOLO v3	89.1	93.3	91.8	91.1	4.7×10^8
YOLO v3 - tiny	87.0	91.9	89.9	89.4	6.7×10^7
LSRCEM - YOLO	89.7	94.0	92.3	92.7	4.6×10^7

LSRCEM - YOLO 的 mAP 为 92.3%, 比 YOLO v3 高 0.5 个百分点,比 YOLO v3 - tiny 高 2.4 个百分点,其模型参数量仅为 YOLO v3 的 10%,为 YOLO v3 - tiny 的 68%。LSRCEM - YOLO 算法不仅在性能上相较 YOLO v3 和 YOLO v3 - tiny 有整体的提高,也降低了参数处理规模。

YOLO v3 - tiny 与 LSRCEM - YOLO 的 mAP 和损失值随迭代次数的变化情况如图 10、11 所示。

由 mAP 变化曲线可知,LSRCEM - YOLO 的精度高于 YOLO v3 - tiny,收敛速度更快,且其 mAP 稳定性高于 YOLO v3 - tiny。

通过损失值的变化曲线可以看出,LSRCEM - YOLO 的收敛速度快于 YOLO v3 - tiny,且收敛后的损失值小于 YOLO v3 - tiny。

YOLO v3 - tiny 和 LSRCEM - YOLO 在相同置信

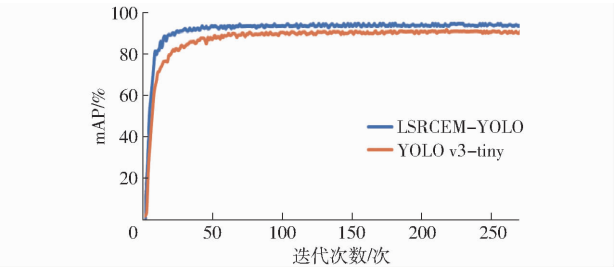


图 10 mAP 变化曲线

Fig. 10 Change curves of mAP

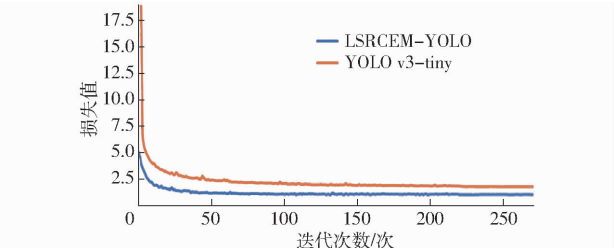


图 11 损失值变化曲线

Fig. 11 Change curves of loss

度(0.5)下进行结果对比,黄色虚线表示漏检目标(图 12),漏检对象多为目标过小或者存在遮挡的情况,当目标数量变多,YOLO v3 - tiny 相较于 LSRCEM - YOLO 的检测效果变差,由于 LSRCEM - YOLO 选择了 3 个不同尺度的特征进行融合,因此其更能兼顾浅层网络关注的小目标与深层网络关注的大目标。

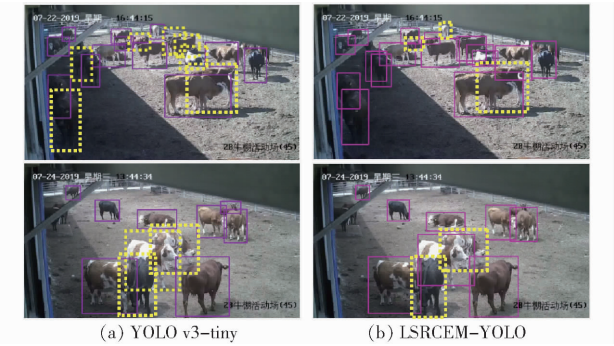


图 12 YOLO v3 - tiny 与 LSRCEM - YOLO 测试结果对比图

Fig. 12 Comparison of test results between YOLO v3 - tiny and LSRCEM - YOLO

实验结果表明,LSRCEM - YOLO 算法兼顾速度与精度,适用于肉牛目标检测。

2.2 肉牛重识别结果分析与精度评价

Mudeep 模型对相同或不同个体肉牛的识别结果进行可视化(图 13),若是同一目标,热图集中于需要关注的区别于其他肉牛的特征信息。若是不同目标,模型通过关注不同肉牛之间的表观特征如花纹等,从而区别出特征信息属于不同个体。说明 Mudeep 模型可以动态学习不同尺度、不同位置的有区分度的特征,自动聚焦于不同目标之间最有区分度的区域。

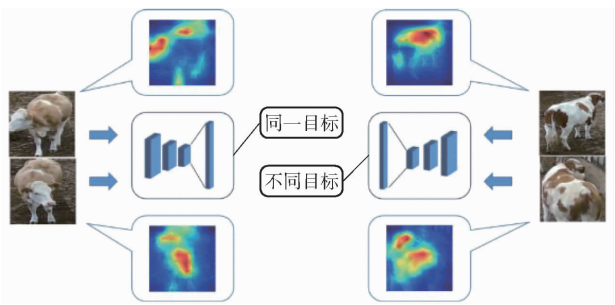


图 13 Mudeep 识别结果

Fig. 13 Identification results of Mudeep

Mudeep 重识别模型的 Rank - 1、Rank - 5、Rank - 10 达到 96.5%、98.6%、99.1%，具备提取到具有高鉴别力特征的能力。

训练后的 Mudeep 模型参数量为 3.5×10^6 ，远小于 Wide ResNet 模型的参数量 4.4×10^7 。

2.3 多目标跟踪结果分析与精度评价

为验证本文方法在肉牛多目标跟踪方面的表现，在肉牛多目标跟踪数据集上进行了测试，并将结果与 YOLO v3 - DeepSORT 算法进行了对比，结果如表 2 所示。

表 2 本文方法与 YOLO v3 - DeepSORT 实验结果对比

Tab. 2 Experimental results of this algorithm compared with YOLO v3 - DeepSORT

算法	多目标跟踪准确	多目标跟踪精确	目标标号改变次数	帧速率/ (f·s ⁻¹)
	率/%	度/%		
YOLO v3 - DeepSORT	32.3	22.4	39	9
本文方法	45.2	34.0	12	21

由表 2 可知，本文方法的多目标跟踪准确率为 45.2%，较 YOLO v3 - DeepSORT 算法提高了 12.9 个百分点，多目标跟踪精确度提高了 11.6 个百分点，目标标号改变次数降低了 69.2%，视频处理速度也由 9 f/s 提高到了 21 f/s。

将同一视频片段用两种不同算法进行跟踪，结

果如图 14 所示。黄色虚线框为 YOLO v3 - DeepSORT 漏检目标，可以看出，相较于 YOLO v3 - DeepSORT 算法，本文方法的漏检目标明显较少，检测框更加贴合肉牛目标，对于遮挡度高的肉牛有更好的检测跟踪能力。

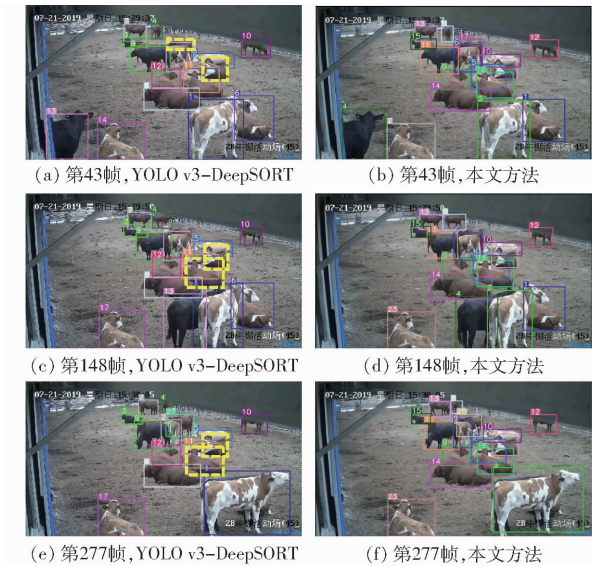


图 14 跟踪结果

Fig. 14 Tracking results

3 结论

(1)在 YOLO v3 模型的基础上，将模型的骨干网络 Darknet53 替换为 MobileNet v2，并增加长短距离语义增强模块(LSRCCEM)，提出了基于 LSRCEM - YOLO 的肉牛多目标跟踪方法。

(2)本文方法目标检测的 mAP 值达 92.3%，采用 Mudeep 模型的肉牛重识别 Rank - 1 指标为 96.5%，多目标跟踪准确率提高到了 45.2%，目标标号改变次数降低了 69.2%。

(3)构建的融合了 LSRCEM - YOLO 和 Mudeep 模型的 DeepSORT 肉牛多目标跟踪模型能够实现实际养殖环境下的肉牛多目标实时跟踪，可为实现肉牛大规模养殖的精准管理提供良好的技术支持。

参 考 文 献

[1] 任继周,李发弟,曹建民,等.我国牛羊肉产业的发展现状、挑战与出路[J].中国工程科学,2019,21(5):67-73.
REN Jizhou, LI Fadi, CAO Jianmin, et al. Development status, challenges, and solutions of China's beef and mutton industry [J]. Strategic Study of CSAE, 2019, 21(5): 67-73. (in Chinese)

[2] 赵圣波,王淞,林佳琳.肉牛食道阻塞的临床症状和综合疗法[J].现代畜牧科技,2017(8):126.
ZHAO Shengbo, WANG Song, LIN Jialin. Clinical symptoms and comprehensive treatment of esophageal obstruction in beef cattle [J]. Modern Animal Husbandry Technology, 2017(8): 126. (in Chinese)

[3] MATTACHINI G, RIVA E, PROVOLO G. The lying and standing activity indices of dairy cows in free-stall housing[J]. Applied Animal Behaviour Science, 2010, 129(1):18-27.

[4] 马志鹏,李啸.野生动物跟踪与定位技术的设计与研究[J].南方农机,2017,48(6):72-73.
MA Zhipeng, LI Xiao. Design and research of wildlife tracking and positioning technology [J]. China Southern Agricultural Machinery, 2017, 48(6): 72-73. (in Chinese)

[5] 王学诚,张沕琳.基于图像传感的快速被试动物追踪系统设计与实现[J].微纳电子与智能制造,2019,1(3):119-125.

- WANG Xuecheng, ZHANG Milin. Design of an image sensor based fast tracking algorithm for animal experiments [J]. Micro Nano Electronics and Intelligent Manufacturing, 2019, 1(3): 119–125. (in Chinese)
- [6] MA Congguo, WANG Yeqin, YING Genwang. The pig breeding management system based on RFID and WSN [C] // International Conference on Information & Computing, 2011: 30–33.
- [7] 任晓惠, 刘刚, 张森, 等. 基于支持向量机分类模型的奶牛行为识别方法[J/OL]. 农业机械学报, 2019, 50(增刊): 290–296. REN Xiaohui, LIU Gang, ZHANG Miao, et al. Dairy cattle's behavior recognition method based on support vector machine classification model[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(Supp.): 290–296. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=2019s045&flag=1. DOI:10.6041/j.issn.1000-1298.2019.S0.045. (in Chinese)
- [8] 王俊, 张海洋, 赵凯旋, 等. 基于最优二叉决策树分类模型的奶牛运动行为识别[J]. 农业工程学报, 2018, 34(18): 202–210. WANG Jun, ZHANG Haiyang, ZHAO Kaixuan, et al. Cow movement behavior classification based on optimal binary decision-tree classification model[J]. Transactions of the CSAE, 2018, 34(18): 202–210. (in Chinese)
- [9] 田富洋, 王冉冉, 刘莫尘, 等. 基于神经网络的奶牛发情行为辨识与预测研究[J/OL]. 农业机械学报, 2013, 44(增刊1): 277–281. TIAN Fuyang, WANG Ranran, LIU Mochen, et al. Oestrus detection and prediction in dairy cows based on neural networks[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(Supp.1): 277–281. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=2013s150&flag=1. DOI:10.6041/j.issn.1000-1298.2013.S1.050. (in Chinese)
- [10] 周雅婷, 许童羽, 陈春玲, 等. 基于神经网络算法的肉牛采食行为检测方法[J]. 沈阳农业大学学报, 2016, 47(6): 752–757. ZHOU Yating, XU Tongyu, CHEN Chunling, et al. Detection method of beef cattle feeding behavior based on neural network algorithm [J]. Journal of Shenyang Agricultural University, 2016, 47(6): 752–757. (in Chinese)
- [11] MILAN H F M, MAIA A S C, GEBREMEDHIN K G. Technical note: device for measuring respiration rate of cattle under field conditions[J]. Journal of Animal Science, 2016, 94(12): 5434–5438.
- [12] ZHAO K X, HE D J. Target detection method for moving cows based on background subtraction[J]. International Journal of Agricultural and Biological Engineering, 2015, 8(1): 42–49.
- [13] 何东健, 孟凡昌, 赵凯旋, 等. 基于视频分析的犊牛基本行为识别[J/OL]. 农业机械学报, 2016, 47(9): 294–300. HE Dongjian, MENG Fanchang, ZHAO Kaixuan, et al. Recognition of calf basic behaviors based on video analysis [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(9): 294–300. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20160940&flag=1. DOI:10.6041/j.issn.1000-1298.2016.9.40. (in Chinese)
- [14] 吴进, 董国豪, 李乔深. 基于区域卷积神经网络和光流法的目标跟踪[J]. 电讯技术, 2018, 58(1): 6–12. WU Jin, DONG Guohao, LI Qiaoshen. Object tracking based on region convolution neural network and optical flow method [J]. Telecommunication Technology, 2018, 58(1): 6–12. (in Chinese)
- [15] 赵凯旋, 何东健, 王恩泽. 基于视频分析的奶牛呼吸频率与异常检测[J/OL]. 农业机械学报, 2014, 45(10): 258–263. ZHAO Kaixuan, HE Dongjian, WANG Enze. Respiratory rate and abnormal detection of dairy cows based on video analysis [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(10): 258–263. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20141040&flag=1. DOI:10.6041/j.issn.1000-1298.2014.10.040. (in Chinese)
- [16] REDMON J, FARHADI A. YOLOv3: an incremental improvement [J]. arXiv e-prints, 2018: 1804.02767.
- [17] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263–7271.
- [18] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779–788.
- [19] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector [C] // European Conference on Computer Vision. Springer, 2016: 21–37.
- [20] LI Z, ZHOU F. FSSD: feature fusion single shot multibox detector[J]. arXiv e-prints, 2017: 1712.00960.
- [21] FU C Y, LIU W, RANGA A, et al. DSSD: deconvolutional single shot detector[J]. arXiv e-prints, 2017: 1701.06659.
- [22] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980–2988.
- [23] 何东健, 刘建敏, 熊虹婷, 等. 基于改进 YOLO v3 模型的挤奶奶牛个体识别方法[J/OL]. 农业机械学报, 2020, 51(4): 250–260. HE Dongjian, LIU Jianmin, XIONG Hongting, et al. Individual identification of dairy cows based on improved YOLO v3 [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(4): 250–260. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20200429&flag=1. DOI:10.6041/j.issn.1000-1298.2020.04.029. (in Chinese)
- [24] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580–587.

- [25] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision,2015:1440-1448.
- [26] 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 & Machine Intelligence, 2017, 39(6): 1137-1149.
- [27] SUN L, ZOU Y, LI Y, et al. Multi target pigs tracking loss correction algorithm based on Faster R-CNN[J]. International Journal of Agricultural and Biological Engineering, 2018, 11(5): 192-197.
- [28] BEWLEY A, GE Z, OTT L, et al. Simple online and realtime tracking[C]//2016 IEEE International Conference on Image Processing (ICIP),2016:3464-3468.
- [29] 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), 2017:3645-3649.
- [30] QIAN X, FU Y, JIANG Y G, et al. Multi-scale deep learning architectures for person re-identification[C]//Proceedings of the IEEE International Conference on Computer Vision,2017:5399-5408.
- [31] ZHENG L, SHEN L, TIAN L, et al. Scalable person re-identification: a benchmark[C]//IEEE International Conference on Computer Vision,2015:1116-1124.
- [32] WANG H, YU Y, CAI Y, et al. A comparative study of state of the art deep learning algorithms for vehicle detection[J]. IEEE Intelligent Transportation Systems Magazine, 2019, 11(2): 82-95.
- [33] QIN Z, LI Z, ZHANG Z, et al. ThunderNet: towards real-time generic object detection[J]. arXiv e-prints, 2019:1903.11752.

(上接第 196 页)

- [21] DASH J, CURRAN P J. Evaluation of the MERIS terrestrial chlorophyll index (MTCI) [J]. Advances in Space Research, 2007, 39(1): 100-104.
- [22] HE L, SONG X, FENG W, et al. Improved remote sensing of leaf nitrogen concentration in winter wheat using multi-angular hyperspectral data[J]. Remote Sensing of Environment, 2016, 174: 122-133.
- [23] WU C, NIU Z, TANG Q, et al. Estimating chlorophyll content from hyperspectral vegetation indices: modeling and validation [J]. Agricultural and Forest Meteorology, 2008, 148(8-9): 1230-1241.
- [24] BARNES E M, CLARKE T R, RICHARDS S E, et al. Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data [M]. Arizona: American Society of Agronomy, 2000.
- [25] GITELSON A A, KEYDAN G P, MERZLYAK M N. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves[J]. Geophysical Research Letters, 2006, 33(11): 431-433.
- [26] 刘潭, 许童羽, 于丰华, 等. 基于 PROSAIL 模型偏差补偿的水稻叶绿素含量遥感估测[J/OL]. 农业机械学报, 2020, 51(5): 156-164.
LIU Tan, XU Tongyu, YU Fenghua, et al. Remote sensing estimation of rice chlorophyll content based on PROSAIL model deviation compensation[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5): 156-164. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20200517&flag=1. DOI: 10.6041/j.issn.1000-1298.2020.05.017. (in Chinese)
- [27] 张喆, 丁建丽, 李鑫, 等. TVDI 用于干旱区农业旱情监测的适宜性[J]. 中国沙漠, 2015, 35(1): 220-227.
ZHANG Zhe, DING Jianli, LI Xin, et al. Suitability of TVDI used to monitor agricultural drought in arid area[J]. Journal of Desert Research, 2015, 35(1): 220-227. (in Chinese)
- [28] 姚雄, 余坤勇, 杨玉洁, 等. 基于随机森林模型的林地叶面积指数遥感估算[J/OL]. 农业机械学报, 2017, 48(5): 159-166.
YAO Xiong, YU Kunyong, YANG Yujie, et al. Estimation of forest leaf area index based on random forest model and remote sensing data[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(5): 159-166. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170519&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2017.05.019. (in Chinese)
- [29] 郑兴明, 丁艳玲, 赵凯, 等. 基于 Landsat 8 OLI 数据的玉米冠层含水量反演研究[J]. 光谱学与光谱分析, 2014, 34(12): 3385-3390.
ZHENG Xingming, DING Yanling, ZHAO Kai, et al. Estimation of vegetation water content from Landsat 8 OLI data[J]. Spectroscopy and Spectral Analysis, 2014, 34(12): 3385-3390. (in Chinese)
- [30] MUTANGA O, ADAM E, CHO M A. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm[J]. International Journal of Applied Earth Observation and Geoinformation, 2012, 18: 399-406.
- [31] 刘贞强. 基于机器学习的北方水稻叶面积指数反演[D]. 哈尔滨: 黑龙江大学, 2018.
LIU Zhenqiang. Inversion of rice leaf area index in North China based on machine learning[D]. Harbin: Heilongjiang University, 2018. (in Chinese)