

基于 DiMP 的类特定养殖奶山羊跟踪方法

宁纪锋^{1,2} 张 静¹ 杨蜀秦^{2,3} 胡沈荣⁴ 蓝贤勇⁴ 王勇胜⁵

(1. 西北农林科技大学信息工程学院, 陕西杨凌 712100; 2. 农业农村部农业物联网重点实验室, 陕西杨凌 712100;
3. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100; 4. 西北农林科技大学动物科技学院, 陕西杨凌 712100;
5. 西北农林科技大学动物医学院, 陕西杨凌 712100)

摘要: 为准确实时跟踪羊只目标,进行疾病异常预警,实现奶山羊精细化养殖,本文基于 DiMP 跟踪模型,利用奶山羊跟踪对象单一且图像样本丰富的特点,结合迁移学习和类特定融合方法,设计了一种类特定的奶山羊目标跟踪模型,能够有效克服 DiMP 算法在跟踪类特定目标时定位精度不足的缺点。利用构建的奶山羊视频跟踪数据训练集对跟踪算法进行迁移训练,加快模型收敛速度,使评估网络预测出的边界框更贴合奶山羊真实框的位置和尺寸。在线跟踪阶段,针对目标模板仅采用第 1 帧特征制作整个序列的调制向量,导致该调制向量相对整个跟踪阶段特征不具代表性,与后续帧差异大的缺点,使用训练集制作包含奶山羊各种姿态的类调制向量,以指数消融方式更新奶山羊类调制向量与第 1 帧调制向量间的比重,增强边界框回归任务中的奶山羊特征与背景的判别性。提出的算法在测试集上的 AUC(Area under curve)和精准度(Precision)分别为 76.20% 和 60.19%,比 DiMP 方法分别提升 6.17、14.18 个百分点,跟踪速度为 30 f/s,满足实时跟踪的要求。实验结果表明,提出的类特定奶山羊目标跟踪方法可用于监测复杂场景下奶山羊的运动,为奶山羊精细化管理提供了技术支持。

关键词: 奶山羊; 目标跟踪; 类特定; 迁移学习; DiMP

中图分类号: TP399 文献标识码: A 文章编号: 1000-1298(2023)06-0280-07 OSID: 

Class-special Real-time Dairy Goat Tracking Method Based on DiMP

NING Jifeng^{1,2} ZHANG Jing¹ YANG Shuqin^{2,3} HU Shenrong⁴ LAN Xianyong⁴ WANG Yongsheng⁵

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

2. Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs,
Yangling, Shaanxi 712100, China

3. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

4. College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China

5. College of Animal Medical, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: In the process of fine breeding of dairy goats, the accurate and real-time tracking of goat targets is an important basis for their behavior recognition and disease abnormality warning. Based on the DiMP tracking model, a kind specific target tracking model was designed for dairy goats, which can effectively overcome the disadvantage of insufficient positioning accuracy of DiMP algorithm in tracking specific targets. The migration training of the tracking algorithm was carried out by using the constructed dairy goat video tracking data training set to accelerate the convergence speed of the model and make the boundary box predicted by the evaluation network more fit the position and size of the real frame of the dairy goat. In the online tracking stage, aiming at the disadvantage that the target template only used the first frame features to produce the modulation vector of the whole sequence, which led to unrepresentative characteristics of the modulation vector relative to the whole tracking stage, the training set was used to produce the class modulation vector containing various poses of the dairy goat, and the proportion between the class modulation vector of the dairy goat and the modulation vector of the first frame was updated by exponential ablation to enhance the discrimination between characteristics and background of dairy goats

in the boundary box regression task. The AUC and accuracy of the proposed algorithm on the test set were 76.20% and 60.19%, respectively, which were 6.17 and 14.18 percentage points higher than that of the DiMP method. The tracking speed was 30 frames per second (f/s), which met the requirements of real-time tracking. The experimental results showed that the proposed target tracking method can be used to monitor the movement of milk goats in complex scenes, and it can provide technical support for fine management of dairy goats.

Key words: dairy goat; object tracking; class-specific; transfer learning; DiMP

0 引言

奶山羊是一种以产奶为主的小型反刍动物,其乳制品脂肪球径小,钙镁含量高于牛奶,是现代乳业的重要原料之一^[1]。随着智慧畜牧和精准饲养理念的提出^[2-3],研究者们通过部署传感器采集实时数据,获取家畜身体状况信息和生理指标,从而对养殖过程进行疾病监测、行为识别和异常预警^[4-6]。近年来,人工智能模型驱动畜牧业高质量发展的相关研究显著增加,主要集中于动物行为检测和识别领域^[7-9],其中,计算机视觉技术在动物福利、牲畜生产、监测动物方面表现出巨大潜力^[10-11],为畜牧管理者生产决策提供评估和分析的依据。因此,设计一种鲁棒、实时的奶山羊跟踪算法对奶山羊智能化管理中的行为分析和精准饲喂具有重要意义^[12]。

相较牛和猪等其他畜牧动物,奶山羊通体雪白,躯体无明显花纹,群聚时相似目标干扰现象严重^[13]。同时,奶山羊运动矫健、敏捷,对跟踪算法的速度和鲁棒性有更高需求^[14]。传统的监测方法通过在家畜躯体穿绑可穿戴设备监测运动轨迹,但是其会限制家畜的活动。ZHANG 等^[15]设计了一种基于视觉图像分析的测量方法,利用固定位置装置捕获视觉图像,无需人工接触测量绵羊的体尺参数。VAYSSADE 等^[16]用 GPS 设备记录奶山羊位置间隔,结合阈值和监督分类跟踪其活动。WANG 等^[17]基于 Faster R-CNN^[18]在监控录像中提取关键帧,改进区域前景的分割,使算法适应视频监控目标检测的时空连续性。SU 等^[19]在 SiamRPN^[20]骨干网络中加入多卷积残差块和降采样多卷积残差块,提取奶山羊高维深度语义信息,建立实时和低成本的奶山羊跟踪算法。随后,其又提出 AMTracker^[21]方法,使用 EfficientNet^[22]为骨干网络,融合多层特征并引入注意力机制提升模板分支和类分支的相关度,使用无锚框的跟踪网络定位奶山羊位置也取得了良好的结果。这些研究表明深度学习方法识别和跟踪奶山羊的可行性,但算法的改进仅局限于增强骨干网络、增加功能性模块等传统思路^[23],未充分挖掘奶山羊独有的外观细节并增强网络对奶山羊类

的针对性学习,跟踪时对相似目标干扰抵抗性差,易发生跟踪漂移^[24]。

本文提出类特定的实时奶山羊跟踪算法,区别于通用跟踪算法使用的基准数据集 COCO^[25]、LaSOT^[26]、TrackingNet^[27]、Got-10K^[28]等预训练网络参数,而是利用跟踪目标类的唯一性,用奶山羊数据集训练区分度更大且判别力更强的奶山羊识别定位模型。其次,在线跟踪时,对边界框回归网络调制向量以指数平滑方式融合奶山羊类特征进行在线更新,适应跟踪目标的表观变化。在奶山羊跟踪数据集上,与代表性跟踪算法的比较结果验证提出方法的有效性。

1 数据获取与处理

1.1 数据采集与标注

实验地点位于陕西省杨凌农业高新技术示范区西北农林科技大学奶山羊基地,于 2021 年 6—7 月采用索尼 FDR-AX100E 型和索尼 A5000 型两种摄像机拍摄。视频拍摄时长为 30~60 s,索尼 FDR-AX100E 型摄像机拍摄的视频帧速率为 29.97 f/s,分辨率为 1 440 像素×1 880 像素,索尼 A5000 型摄像机拍摄的视频帧速率为 25 f/s,分辨率为 1 920 像素×1 880 像素。使用 Labelme 图像标注工具对采集到的视频进行标注。

1.2 数据集构建与分析

奶山羊数据集由 52 个视频序列组成,其中训练集包括 18 个视频,测试集包括 34 个视频。视频平均长度为 701 帧,共计 36 430 幅奶山羊图像,图像中包含模糊、严重遮挡、光线不足、离开视野等不同情况,综合羊只体尺、光线明暗、羊只数量、运动状态、模糊程度和远近景情况,将拍摄视频分为单只羊类、小目标类、快速运动类和群羊类。如图 1 所示,单只羊类图像清晰易识别(图 1a);小目标类多远景,羊静立、久卧或慢速运动,部分存在遮挡(图 1b);快速运动类羊剧烈运动,多模糊、形变较大(图 1c);群羊类指视野中多只羊,目标混入羊群出现遮挡,易发生相似目标干扰,对算法判别能力要求高(图 1d)。奶山羊分类数据集数量划分情况如表 1 所示。

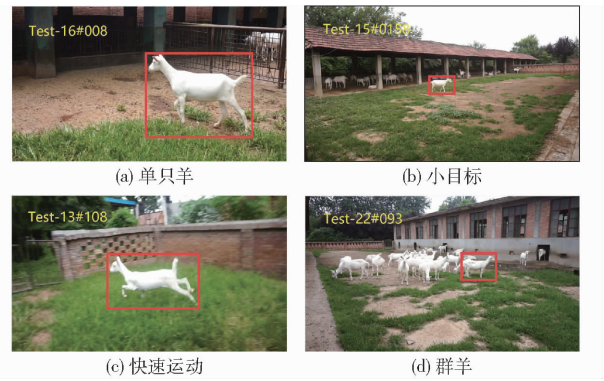


图 1 奶山羊跟踪数据集示例

Fig. 1 Examples of dairy goat tracking data set

表 1 奶山羊数据集不同类视频平均帧和视频数
Tab.1 Average frames and number of different types of videos in dairy goat dataset

数据集	单只羊平均 帧数/视频数	小目标平均 帧数/视频数	快速运动平均 帧数/视频数	群羊平均帧 数/视频数
训练集	692/7	684/2	687/6	546/3
测试集	680/13	675/6	805/5	754/10

2 研究方法

2.1 DiMP 跟踪方法

DiMP^[29]是一种判别式目标跟踪模型,包含分类分支和边界框回归分支:前者可区分目标和背景,预测模块主要用于训练判别式目标模型 f 、初始化模块和循环优化器模块。初始化器根据训练集得到模型的初始解,然后将其送入循环优化器获得判别式目标模型 f ,其参数以数据驱动的方式学习,如标签置信分数 y_c ,空间权重函数 v_c ,正则化参数 λ 以及目标掩码 m_c ,图像 x 与滤波器 f 经过卷积操作得到预测的目标中心,计算与真实目标中心 c 的残差,得到判别前后背景的分类模型;后者继承自 ATOM 模型^[30],基于目标模型的调制向量预测当前跟踪目标的位置与尺寸,参考 IoUNet^[31],提出了目标特定的 IoU 预测,适用于不同种类目标对象,以训练学到统一的经验权重,进行通用目标类的识别和定位。

2.2 基于类特定实时奶山羊跟踪算法

DiMP 网络采用目标特定策略预测跟踪目标边界框的尺寸与位置,该策略虽然提高了算法对通用目标跟踪时的普适性和泛化能力,但在特定类目标跟踪方面,将极大抑制异类目标独有特征的表达和传递,降低算法对该既定目标类的敏感性和适应性。在线跟踪时,边界框回归分支使用第 1 帧作为整个跟踪序列的调制向量,由于跟踪初期后续帧与第 1 帧相似度高,调制向量可为边界框精准回归提供支撑,但随着帧数增加,后续帧目标变化,调制向量仅包含第 1 帧特征,相较整个跟踪序列的形态变化,

特征单一且不具备代表性,难以应对表观变化较大的跟踪目标^[32]。

本文提出的类特定跟踪算法包括 2 方面:基于奶山羊训练集的类适应性训练和基于类特定融合的 IoU - Refine 分支改进。类适应性训练方法使跟踪网络在识别特定跟踪类时,借鉴目标检测中采用迁移学习训练网络识别特定类的思想,对执行特定类任务的跟踪网络应用迁移学习方法,解决 DiMP 原始模型和具体应用域之间不一致问题,有效过滤冗余信息并加速模型收敛。同理,由于跟踪目标仅有奶山羊类,因此可以有效增强 IoU - Refine 分支的融合效果,类调制向量提取奶山羊类丰富的语义特征和先验知识,与跟踪序列第 1 帧调制向量以指数消融方式进行融合,自适应地优化特征,弥补仅采用第 1 帧调制向量造成的特征不具代表性问题。

2.2.1 基于奶山羊训练集的 DiMP 类特定跟踪模型

由于通用跟踪算法对跟踪目标的不可知性,一般采用目标特定的策略设计跟踪网络,本文利用跟踪目标的类特定性,将奶山羊数据集在 DiMP 网络原始训练权重基础上微调,建立对奶山羊类敏感适应的判别式分类器模型 f 。具体过程为,首先选择一个奶山羊视频序列,随机挑出若干帧,一部分作为训练集 M_{train} ,另一部分作为测试集 M_{test} 。然后,骨干网络 ResNet^[33]使用 MoCov2^[34]在 ImageNet^[35]上初始化训练出的模型,将 M_{train} 和 M_{test} 分别输入到特征提取网络得到参考分支和测试分支的特征信息 S_{train} 和 S_{test} 。基于参考分支特征 S_{train} 提取的目标特征得到初始化模型,输入循环优化器得到最终的判别模型 f ,并与测试集提取的特征进行卷积得到打分,再结合测试标签 S_{test} 进行评价,反向传播以更新骨干网络参数。通过该方法可为后续正负样本分类和边界框评估提供更精准的特征映射,以此增加跟踪奶山羊过程中对目标和背景的判别能力。

2.2.2 基于 IoU - Refine 分支的类特定调制向量

由于 DiMP 所使用的边界框回归网络的参考模板仅用视频序列的第 1 帧制作调制向量,难以完整反映奶山羊跟踪过程中出现的不同形态变化且后续帧和第 1 帧形态存在较大差异。因此,本文在线跟踪阶段,随着跟踪的进行,以指数平滑方式在第 1 帧调制向量中逐通道融合奶山羊类调制向量,构造随时域更新的调制向量,以适应奶山羊表面模型的变化,改进的向量融合方式如图 2 所示。

首先,对于在线跟踪阶段, x 表示跟踪帧, B 表示跟踪帧对应的边界框。从训练集每个序列中随机取 2~3 幅图像,共取出 40 幅图像组成奶山羊类数据集 $S = \{(\hat{x}_j, \hat{B}_j)\}_{j=1}^{\hat{n}}$,其中 $\hat{n}$ 表示类数据集帧数,

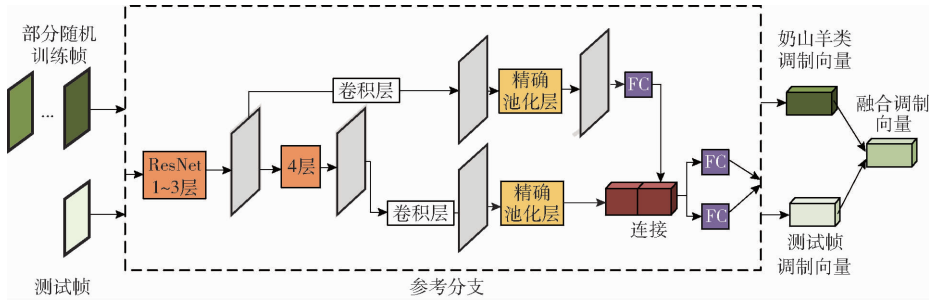


图 2 类调制向量的融合网络

Fig. 2 Fusion of class modulation vector and the first frame modulation vector

类数据集的奶山羊图像来自 18 个训练集视频序列,与在线跟踪的测试集无交集。其中每对 $(\hat{x}_j, \hat{B}_j)$ 经过边界框评估网络的参考分支生成对应的调制向量 $\hat{c}(\hat{x}_j, \hat{B}_j)$, S 生成的所有调制向量逐通道加权融合并取平均值,得到融合奶山羊各种跟踪姿态的类调制向量 $\hat{c}(\hat{x}_0, \hat{B}_0)$,尺寸为 $1 \times 1 \times D$, D 为类调制向量的维度, $\hat{c}(\hat{x}_0, \hat{B}_0)$ 包含奶山羊类丰富的形态与姿势。

其次,对奶山羊测试集 $S_{\text{test}} = \{x_i, B_i\}_{i=1}^n$ 中的帧在线跟踪时,首先将 (x_i, B_i) 经过边界评估网络的参考分支生成调制向量 $\bar{c}(x_i, B_i)$,尺寸为 $1 \times 1 \times D$,计算与类调制向量 $\hat{c}(\hat{x}_0, \hat{B}_0)$ 的融合权值,逐通道融合成调制向量 $c(x_i, B_i)$,尺寸为 $1 \times 1 \times D$,作为评估分支当前帧的外观模板。在跟踪过程中,随着跟踪帧数的变化,采用指数平滑方式更新两者融合权值,公式为

$$\alpha = ae^{-T} + b \tag{1}$$

$$c(x_i, B_i) = \alpha \bar{c}(x_i, B_i) + (1 - \alpha) \hat{c}(\hat{x}_0, \hat{B}_0) \tag{2}$$

式中 a, b ——自适应权重

T ——跟踪帧数

α ——融合权重比

跟踪序列测试帧产生调制向量为 $z(x_i, B_i)$,尺寸为 $K \times K \times D$, K 为池化层空间输出,通过通道乘法与 $c(x_i, B_i)$ 融合,送入模块 g 预测当前帧的边界框,计算当前目标预测边界框与真实框的 IoU。

3 实验与结果分析

3.1 模型训练

实验硬件环境为 GeForce RTX 2080Ti GPU,显存 12 GB,操作系统为 Ubuntu 1604LTS,使用 PyTorch 1.7.1 深度学习架构,编程语言及版本为 Python。改进后 DiMP 模型采用 MoCov2 在 ImageNet 训练得到的权重对骨干网络进行初始化,其余部分使用 DiMP 预训练模型权重进行初始化。使用奶山羊训练集训练 50 个迭代周期(epoch),批处理量设置为 4,输入图像的分辨率为 288 像素 $\times$ 288 像素,使用

Adam 优化器优化模型,初始学习率设置为 2×10^{-4} ,动量参数为 0.9,调制向量更新频率为 150。

3.2 评价指标

采用 5 项指标评价算法的跟踪效果。为了对模型的性能进行恰当的排序,需要明确性能参数的优先级。AUC 是 ROC (Receiver operating characteristic) 曲线下的面积,描述分类器对正、负例的分类能力,值越高表示算法判别能力越强。精确率(Precision)为目标中心点坐标与真实值的误差距离在给定阈值范围内的视频帧数占总视频帧数的百分比,值越高表示算法定位能力越强。OP50 (Overlap precision of 50%) 和 OP75 (Overlap precision of 75%) 分别为真实框和预测框重叠率超过 50% 和 75% 的视频帧数占总视频帧数的百分比,重叠率为重叠区域面积与预测矩形面积、真实矩形面积、重叠区域面积和的比值。归一化精确率(Norm Precision)使用目标中心点坐标与真实值的距离归一化距离替换 Precision 的绝对距离进行评估。

3.3 实验结果分析

3.3.1 不同算法实验结果对比

为了验证改进模型的有效性,将 ATOM 模型、DiMP 模型、SiamRPN 模型在测试集上的结果与本文改进模型进行了对比实验,结果如表 2 所示,(+) 表示将该模型在奶山羊训练集上参数微调后在测试集上进行对比实验。

由表 2 可知,改进 DiMP 模型各项指标均优于其他 6 种目标跟踪模型。ATOM、SiamRPN、DiMP 和改进 DiMP 模型的速度分别为 26、27、33、30 f/s。在 AUC 方面,改进模型为 76.20%,分别比 ATOM、DiMP 和 SiamRPN 模型高 10.40、6.17、6.09 个百分点。改进模型的识别准确率为 60.19%,分别比 ATOM、DiMP 和 SiamRPN 模型高 18.93、14.18、8.50 个百分点。改进模型的 OP50、OP75 和 Norm Precision 分别为 89.55%、78.60%、87.55%,与 DiMP 模型相比,提升 5.21、15.68、6.90 个百分点。其中,ATOM、DiMP 模型在奶山羊训练集上进

表 2 奶山羊测试数据集实验结果

Tab.2 Experimental results of dairy goat test dataset

%

跟踪模型	AUC	Precision	OP50	OP75	Norm Precision
ATOM	65.80	41.26	82.10	50.38	79.97
ATOM(+)	68.33(+2.53)	44.72(+3.46)	81.93(+0.17)	60.48(+10.10)	80.48(+0.51)
SiamRPN	70.11	51.69	86.36	62.71	83.83
SiamRPN(+)	55.84(-14.27)	28.13(-23.56)	68.26(-18.10)	28.46(-34.25)	77.71(-6.12)
DiMP	70.03	46.01	84.34	62.92	80.65
DiMP(+)	75.50(+5.47)	60.01(+14.00)	88.64(+4.30)	78.15(+15.23)	86.75(+6.10)
改进 DiMP	76.20	60.19	89.55	78.60	87.55

行参数微调后,识别性能和定位精度有效提高,但 SiamRPN 使用奶山羊训练集微调参数后结果下降,分析原因是由于 SiamRPN 作为孪生系列算法,目标定位依靠跟踪帧与模板帧间的交互获取目标特征信息,且网络结构多为浅层,参数量少,难以在微调训练中学习复杂特征信息,导致鲁棒性差和泛化能力下降。

改进 DiMP 模型在 OP75 和 Precision 上均有显著提升,表明在 DiMP 采用类特定训练和融合类特征调制向量的奶山羊跟踪模型,识别性能和定位精度有效提高,验证了改进 DiMP 模型的有效性。图 3 为改进 DiMP 模型在奶山羊数据集上训练时的损失函数变化曲线,可见,该方法收敛速度较快,当迭代次数为 20 时,训练损失值由 0.23 降至 0.09。

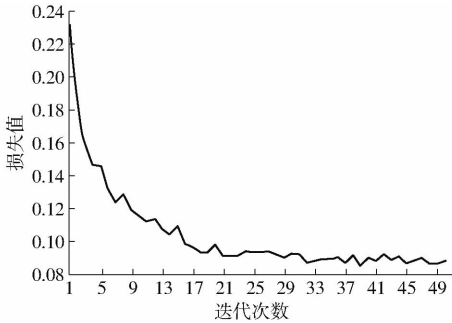


图 3 损失值收敛曲线

Fig.3 Convergence curves of loss value

3.3.2 不同类别数据集上实验结果对比

采用 ATOM、DiMP、SiamRPN 和改进 DiMP 模型分别对单只羊测试集、小目标测试集、快速运动测试集和群羊测试集进行跟踪,以比较不同算法对不同场景的跟踪效果。场景中包括了一些困难场景,例如,小目标测试集特征不明显、群羊测试集多出现遮挡和相似目标干扰以及快速运动测试集多出现形变和模糊等现象。表 3 通过改进 DiMP 模型与其它 3 种跟踪算法对 4 类奶山羊跟踪场景 AUC 结果加以比较。

可以看出,在单只羊测试集、快速运动测试集和群羊测试集上,改进 DiMP 模型相较其他算法各项指标都有明显提高。改进 DiMP 模型比 DiMP 模型

表 3 不同算法在分类测试集上的 AUC

Tab.3 AUC results of different algorithms on

classified test set

%

跟踪模型	测试集			
	单只羊	小目标	快速运动	群羊
ATOM	71.41	67.27	63.99	57.39
DiMP	76.09	71.47	68.26	61.70
SiamRPN	70.91	76.14	67.27	65.62
改进 DiMP	81.49	71.60	74.76	69.84

在单只羊、快速运动和群羊上的 AUC 分别高出 5.4、6.5、8.14 个百分点,说明改进 DiMP 模型对于形变、模糊、相似目标干扰问题的鲁棒性显著增强,进一步验证了改进算法良好的判别能力。

在小目标类别中,SiamRPN 的跟踪效果优于改进 DiMP、DiMP、ATOM 3 种算法,改进 DiMP 的 AUC 无明显提升,推测原因是由于小目标类奶山羊多为远景,目标小且特征不明显,SiamRPN 用模板信息与搜索区域信息进行相似性度量预测边界框的方式,优于 DiMP 算法中 IoU 预测模块的目标模板为待搜索图像提供调制向量预测边界框的方式,且跟踪过程无明显模糊和形变,DiMP 和改进 DiMP 在线训练时循环优化器模块判别正、负样本的过程对该数据集提升效果有限。

3.3.3 消融实验

表 4 为消融实验用于验证改进 DiMP 模型各个模块对跟踪性能的影响。可以看出,采用迁移学习方法使 AUC 和精确率分别提升 5.47 个百分点和 14.00 个百分点;而最后通过添加类特定融合方法,使得改进模型最终的 AUC 和精确率达到 76.20% 和 60.19%,进一步提升了精度。实验结果表明,改进 DiMP 模型在各项指标上都比原始模型有显著提升,AUC 和精确率分别提升 6.17、14.18 个百分点,验证了提出方法的有效性。

3.3.4 定性评价

图 4 为 3 个不同类型奶山羊视频序列的跟踪结果。可以看出,出现相似目标干扰和模糊的跟踪困难场景时,原始 DiMP 模型预测框过大,未贴合目标,且存在跟踪漂移现象,而改进 DiMP 模型预测的

表 4 改进 DiMP 模型消融实验结果							
Tab.4 Ablation analysis of experimental results of improved DiMP model							
模型	类特定训练	类特定融合	AUC	Precision	OP50	OP75	Norm Precision
DiMP	×	×	70. 03	46. 01	84. 34	62. 92	80. 92
改进 DiMP	×	√	70. 61(+0. 58)	46. 76(+0. 75)	84. 84(+0. 09)	63. 89(+0. 89)	81. 02(+0. 10)
	√	×	75. 50(+4. 89)	60. 01(+13. 25)	88. 64(+3. 80)	78. 15(+14. 26)	86. 75(+5. 73)
	√	√	76. 20(+0. 70)	60. 19(+0. 18)	89. 55(+0. 91)	78. 60(+0. 45)	87. 55(+0. 72)

注：×表示不采用此方法；√表示采用此方法。

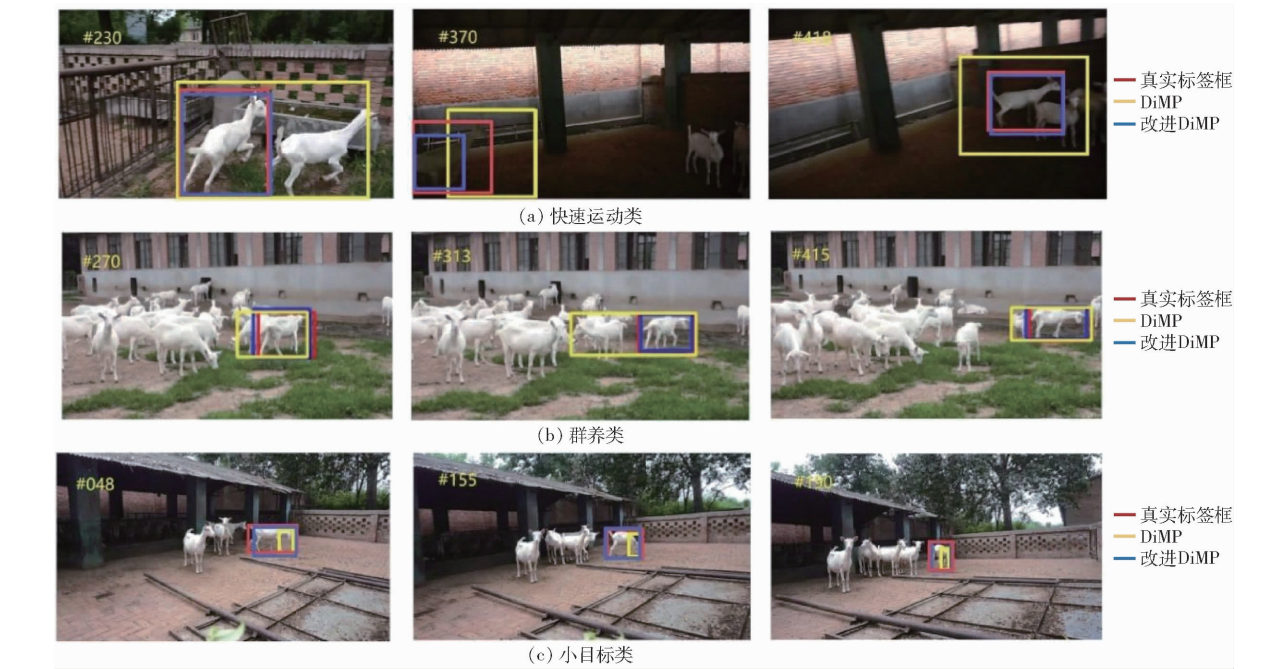


图 4 改进 DiMP 跟踪方法跟踪结果示例

Fig.4 Example of improved DiMP tracking method

目标框能够准确框住目标。表明提出的跟踪方法能够适应奶山羊目标表观的变化,从而得到准确的跟踪结果。

4 结论

(1)为了实现复杂情况下精确、鲁棒的奶山羊目标跟踪,本文提出了一种类特定的 DiMP 奶山羊目标跟踪方法,充分利用奶山羊跟踪对象单一固定的优势,结合迁移学习和指数融合类调制向量,构造适应奶山羊类外观变化的判别性网络和高质量的外观融合模板,从而获得准确的预测边界框和可靠的位置信息,提高了算法在面对奶山羊相似目标干

扰、模糊、遮挡和外观变化时的跟踪精度和成功率。在 4 种场景类别奶山羊测试集上的实验结果验证了改进方法的有效性。

(2)基于奶山羊目标跟踪的类特定先验性条件,采用迁移学习策略,使 DiMP 模型向奶山羊跟踪任务收敛,AUC 从 70. 61% 提升到 75. 50% ,显著增强了目标跟踪性能。

(3)使用类特定方法代替目标特定方法进行跟踪时,采用该类数据集训练的模型,并在参考模板中融合该类的特征信息,可得到更强的奶山羊表观特征,AUC 从 75. 50% 提升到 76. 20% ,进一步提升了跟踪性能。

参 考 文 献

[1] 武学峰. 咸阳奶山羊产业概况[J]. 中国畜牧业, 2017(7) : 78.

[2] BERCKMANS D, LEUVEN K U, MEASURE D O, et al. Automatic on-line monitoring of animals by precision livestock farming [J]. Livestock Production and Society, 2006, 287: 27 – 30.

[3] 宣传忠, 武佩, 张丽娜, 等. 羊咳嗽声的特征参数提取与识别方法[J]. 农业机械学报, 2016, 47(3) : 342 – 348. XUAN Chuazhong, WU Pei, ZHANG Li'na, et al. Feature parameters extraction and recognition method of sheep cough sound [J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(3) : 342 – 348. (in Chinese)

[4] BANHAZI T M, LEHR H, BLACK J, et al. Precision livestock farming: an international review of scientific and commercial aspects[J]. International Journal of Agricultural Biological Engineering, 2012, 5(3) : 1 – 9.

- [5] RUSHEN J, CHAPINAL N, DE P, et al. Automated monitoring of behavioural-based animal welfare indicators[J]. *Animal Welfare*, 2012, 21(3): 339.
- [6] 张宏鸣, 周利香, 李永恒, 等. 基于改进 MobileFaceNet 的羊脸识别方法[J]. *农业机械学报*, 2022, 53(5): 267–274.
ZHANG Hongming, ZHOU Lixiang, LI Yongheng, et al. Sheep face recognition method based on improved MobileFaceNet[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(5): 267–274. (in Chinese)
- [7] 赵世达, 王树才, 白宇, 等. 基于生成对抗网络与 ICNet 的羊骨架图像实时语义分割[J]. *农业机械学报*, 2021, 52(2): 329–339.
ZHAO Shida, WANG Shucai, BAI Yu, et al. Real-time semantic segmentation of sheep skeleton image based on generative adversarial network and ICNet[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(2): 329–339. (in Chinese)
- [8] QIAO Y, KONG H, CLARK C, et al. Intelligent perception for cattle monitoring: a review for cattle identification, body condition score evaluation, and weight estimation[J]. *Computers Electronics in Agriculture*, 2021, 185: 106143.
- [9] BAO J, XIE Q J. Artificial intelligence in animal farming: a systematic literature review[J]. *Journal of Cleaner Production*, 2022, 331: 129956.
- [10] LIAKOS K G, BUSATO P, MOSHOU D, et al. Machine learning in agriculture: a review[J]. *Sensors*, 2018, 18(8): 2674.
- [11] VAN HERTEM T, ROOIJAKKERS L, BERCKMANS D, et al. Appropriate data visualisation is key to precision livestock farming acceptance[J]. *Computers Electronics in Agriculture*, 2017, 138: 1–10.
- [12] JIANG M, RAO Y, ZHANG J Y, et al. Automatic behavior recognition of group-housed goats using deep learning[J]. *Computers Electronics in Agriculture*, 2020, 177: 105706.
- [13] 尚诚, 王美丽, 宁纪锋, 等. 联合损失优化下的高相似度奶山羊身份识别[J]. *中国图象图形学报*, 2022, 27(4): 1137–1147.
SHANG Cheng, WANG Meili, NING Jifeng, et al. Joint loss optimization based high similarity identification for milch goats [J]. *Journal of Image and Graphics*, 2022, 27(4): 1137–1147. (in Chinese)
- [14] 韩丁, 王斌, 王亮, 等. 基于改进 VGGNet 的羊个体疼痛识别方法[J]. *农业机械学报*, 2022, 53(6): 311–317.
HAN Ding, WANG Bin, WANG Liang, et al. Individual pain recognition method of sheep based on improved VGGNet[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(6): 311–317. (in Chinese)
- [15] ZHANG A L, WU B P, WUYUN C T, et al. Algorithm of sheep body dimension measurement and its applications based on image analysis[J]. *Computers Electronics in Agriculture*, 2018, 153: 33–45.
- [16] VAYSSADE J A, ARQUET R, BONNEAU M. Automatic activity tracking of goats using drone camera[J]. *Computers Electronics in Agriculture*, 2019, 162: 767–772.
- [17] WANG D, TANG J L, ZHU W J, et al. Dairy goat detection based on Faster R-CNN from surveillance video[J]. *Computers Electronics in Agriculture*, 2018, 154: 443–449.
- [18] SUN X, WU P, HOI S. Face detection using deep learning: an improved Faster R-CNN approach[J]. *Neurocomputing*, 2018, 299: 42–50.
- [19] SU Q, TANG J L, ZHAI J, et al. Automatic tracking of the dairy goat in the surveillance video[J]. *Computers Electronics in Agriculture*, 2021, 187: 106254.
- [20] LI B, YAN J J, WU W, et al. High performance visual tracking with Siamese region proposal network[C] // *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: 8971–8980.
- [21] SU Q, TANG J, ZHAI M, et al. An intelligent method for dairy goat tracking based on Siamese network[J]. *Computers Electronics in Agriculture*, 2022, 193: 106636.
- [22] TAN M X, LE Q. EfficientNet: rethinking model scaling for convolutional neural networks[C] // *International Conference on Machine Learning (ICML)*, 2019: 6105–6114.
- [23] 陈志良, 石繁槐. 结合双模板融合与孪生网络的鲁棒视觉目标跟踪[J]. *中国图象图形学报*, 2022, 27(4): 1191–1203.
CHEN Zhiliang, SHI Fanhuai. Double template fusion based siamese network for robust visual object tracking[J]. *Journal of Image and Graphics*, 2022, 27(4): 1191–1203. (in Chinese)
- [24] 连丽容, 罗文婷, 秦勇, 等. 双目机器视觉及 RetinaNet 模型的路侧行人感知定位[J]. *中国图象图形学报*, 2021, 26(12): 2941–2952.
LIAN Lirong, LUO Wenting, QIN Yong, et al. Roadside pedestrian detection and location based on binocular machine vision and RetinaNet[J]. *Journal of Image and Graphics*, 2021, 26(12): 2941–2952. (in Chinese)
- [25] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[C] // *Proceedings of the European Conference on Computer Vision (ECCV)*, 2014: 740–755.
- [26] FAN H, LIN L, YANG F, et al. Lasot: a high-quality benchmark for large-scale single object tracking[C] // *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 5374–5383.
- [27] MULLER M, BIBI A, GIANCOLA S, et al. TrackingNet: a large-scale dataset and benchmark for object tracking in the wild [C] // *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 300–317.
- [28] HUANG L H, ZHAO X, HUANG K Q, et al. Got-10K: a large high-diversity benchmark for generic object tracking in the wild[J]. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 2019, 43(5): 1562–1577.

on the organosolv pulping of beech wood[J]. Green Chemistry, 2014, 16(3): 1569 – 1578.

[19] KOBAYASHI T, SAKAI Y. Hydrolysis rate of pentosan of hardwood in dilute sulfuric acid[J]. Bulletin of the Agricultural Chemical Society of Japan, 1956, 20(1): 1 – 7.

[20] TIZAZU B Z, MOHOLKAR V S. Kinetic and thermodynamic analysis of dilute acid hydrolysis of sugarcane bagasse[J]. Bioresour. Technol., 2018, 250(1): 197 – 203.

[21] LU Y, MOSIER N S. Kinetic modeling analysis of maleic acid-catalyzed hemicellulose hydrolysis in corn stover[J]. Biotechnol. Bioeng., 2008, 101(6): 1170 – 1181.

[22] 谭丽红, 马乐凡, 刘定军, 等. 楠竹水预水解半纤维素溶出动力学研究[J]. 中国造纸, 2012, 27(3): 4.
TAN Lihong, MA Lefan, LIU Dingjun, et al. Kinetic model of hemicellulose dissolution during water prehydrolysis of *Phyllostachys heterocyclava*[J]. China Pulp & Paper, 2012, 27(3): 4. (in Chinese)

[23] SLUITER A, HAMES B, RUIZ R, et al. Determination of structural carbohydrates and lignin in biomass. Lab:NREL/TP – 510 – 42618[S]. Anal. Proced, 2008.

[24] YAN Y, ZHANG C, LIN Q, et al. Microwave-assisted oxalic acid pretreatment for the enhancing of enzyme hydrolysis in the production of xylose and arabinose from bagasse[J]. Molecules, 2018, 23(4): 862.

[25] JACOBSEN S E, WYMAN C E. Cellulose and hemicellulose hydrolysis models for application to current and novel pretreatment processes[J]. Appl. Biochem. Biotechnol., 2000, 53:81 – 96.

[26] ESTEGLALIAN A, HASHIMOTO A G, FENSKE J J, et al. Modeling and optimization of the dilute-sulfuric-acid pretreatment of corn stover, poplar and switchgrass[J]. Bioresour. Technol., 1997, 59(2 – 3): 129 – 136.

[27] CONNER A H, LORENZ L F. Kinetic modeling of hardwood prehydrolysis. Part III. Water and dilute acetic acid prehydrolysis of southern red oak[J]. Wood Fiber Sci., 1986(2), 248 – 263.

[28] TAN H, YANG R, SUN W, et al. Peroxide-acetic acid pretreatment to remove bagasse lignin prior to enzymatic hydrolysis[J]. Ind. Eng. Chem. Res., 2010, 49(4): 1473 – 1479.

[29] 王琼, 庄新姝, 余强, 等. 超低浓度马来酸催化水解纤维素的机理研究[J]. 可再生能源, 2011, 29(2): 49 – 54.
WANG Qiong, ZHUANG Xinzhu, YU Qiang, et al. Research on mechanism of cellulose hydrolysis catalyzed by extremely low maleic acid[J]. Renewable Energy Resources, 2011, 29(2): 49 – 54. (in Chinese)

[30] MOSIER N S, SARIKAYA A, LADISCH C M, et al. Characterization of dicarboxylic acids for cellulose hydrolysis[J]. Biotechnol. Prog., 2001, 17(3): 474 – 480.

[31] LU Y, MOSIER N S. Biomimetic catalysis for hemicellulose hydrolysis in corn stover[J]. Biotechnol. Prog., 2007, 23(1): 116 – 123.

[32] SHI S, GUAN W, KANG L, et al. Reaction kinetic model of dilute acid-catalyzed hemicellulose hydrolysis of corn stover under high-solid conditions[J]. Ind. Eng. Chem. Res., 2017, 56(39): 10990 – 10997.

[33] SCHMORAK J. The chemistry of cellulose and wood[M]. Jerusalem: Israel Program for Scientific Translation, 1966.

[34] MALONEY M T, CHAPMAN T W, BAKER A J. Dilute acid hydrolysis of paper birch: kinetics studies of xylan and acetyl-group hydrolysis[J]. Biotechnology and Bioengineering, 1985, 27(3): 355 – 361.

[35] YUAN Q, LIU S, MA M G, et al. The kinetics studies on hydrolysis of hemicellulose[J]. Front. Chem., 2021, 9: 781291.

(上接第 286 页)

[29] BHAT G, DANELLJAN M, GOOL L V, et al. Learning discriminative model prediction for tracking[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 6182 – 6191.

[30] DANELLJAN M, BHAT G, KHAN F S, et al. Atom: accurate tracking by overlap maximization[C] // IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 4660 – 4669.

[31] JIANG B, LUO R, MAO J, et al. Acquisition of localization confidence for accurate object detection[C] // Proceedings of the European Conference on Computer Vision (ECCV), 2018: 784 – 799.

[32] 尹宽, 李均利, 胡凯, 等. 融入时序和速度信息的自适应更新目标跟踪[J]. 中国图象图形学报, 2021, 26(4): 883 – 897.
YIN Kuan, LI Junli, HU Kai, et al. Adaptive update object tracking algorithm incorporating timing and speed information[J]. Journal of Image and Graphics, 2021, 26(4): 883 – 897. (in Chinese)

[33] HE K, FAN H, WU Y, et al. Momentum contrast for unsupervised visual representation learning[C] // Proceedings of the European Conference on Computer Vision (ECCV), 2020: 9729 – 9738.

[34] CHEN X, FAN H, GIRSHICK R, et al. Improved baselines with momentum contrastive learning[J]. arXiv preprint, arXiv: 2003.04297, 2020.

[35] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211 – 252.