

DOI: 10.16652/j.issn.1004-373x.2023.07.008

引用格式: 韩晓冰, 王雨田, 黄综浏, 等. 改进 YOLOv5+DeepSort 的行人跟踪算法[J]. 现代电子技术, 2023, 46(7): 33-38.

# 改进 YOLOv5+DeepSort 的行人跟踪算法

韩晓冰, 王雨田, 黄综浏, 张玮良

(西安科技大学 通信与信息工程学院, 陕西 西安 710000)

**摘要:** 针对复杂环境道路行人跟踪易发生身份丢失、切换的问题, 提出一种改进的 YOLOv5 检测并结合 DeepSort 跟踪算法。检测阶段, 融合注意力模块 CBAM 与 YOLOv5 颈部网络增强对行人特征的提取; 用 SIoU 边界框损失函数代替 CIoU 边界框损失函数, 加速边界框回归的同时提高准确定位度。跟踪阶段, 改进 DeepSort 利用拓展卡尔曼滤波器对非线性环境行人位置进行预测, 通过匈牙利算法匹配预测和检测轨迹, 优化复杂环境下行人身份切换频繁的问题。最后连接改进后的 YOLOv5 与 DeepSort 算法, 对 MOT-16 数据集进行检测跟踪。实验结果表明: 改进 YOLOv5 算法较原算法平均精准度提高 4%, 结合 DeepSort 跟踪, 平均跟踪精确度为 63.5%, 比原始算法提升了 3.4%; 行人身份切换次数减少 52 次, 比原始算法减少了 6.5%。

**关键词:** 行人跟踪; YOLOv5; DeepSort; 特征提取; 注意力机制; 拓展卡尔曼滤波

**中图分类号:** TN911.73-34; TP391

**文献标识码:** A

**文章编号:** 1004-373X(2023)07-0033-06

## Improved YOLOv5+DeepSort algorithm for pedestrian tracking

HANG Xiaobing, WANG Yutian, HUANG Zongliu, ZHANG Weiliang

(College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710000, China)

**Abstract:** An improved YOLOv5 detection combined with DeepSort tracking algorithm is proposed to solve the problem that the pedestrian identity is prone to loss and switch when tracking the pedestrian on roads in complex environments. In the detection stage, CBAM (convolutional block attention module) and the YOLOv5 neck network are integrated to enhance the extraction capacity of pedestrian features, and the CIoU bounding box loss function is replaced by the SIoU bounding box loss function, which accelerates the bounding box regression and improves the positioning accuracy. In the tracking stage, the extended Kalman filter (EKF) is used in the improved DeepSort to predict the position of pedestrians in a nonlinear environment, and the Hungarian algorithm is adopted to match the prediction and detection trajectories to optimize the frequent pedestrian identity switching in complex environments. Finally, the improved YOLOv5 and DeepSort algorithms are connected to detect and track the MOT-16 dataset. The experimental results show that the average accuracy of the improved YOLOv5 algorithm is increased by 4% compared with the original algorithm; in combination with Deepsort tracking algorithm, the average tracking accuracy is 63.5%, which is 3.4% higher than the original algorithm, and the number of pedestrian identity switching is reduced by 52 times, which is 6.5% less than the original algorithm.

**Keywords:** pedestrian tracking; YOLOv5; DeepSort; feature extraction; CBAM; EKF

## 0 引言

近年来, 计算机视觉在行人多目标跟踪<sup>[1]</sup>领域得到广泛的应用, 该技术的发展有助于视频监控、辅助驾驶、人机交互等<sup>[2-4]</sup>。行人多目标跟踪算法大多基于 Trace-By-Detection<sup>[5-7]</sup>, 即检测和跟踪, 通过检测器将结果传输到跟踪器中, 进行卡尔曼预测和匈牙利算法匹配, 得到匹配轨迹, 输出跟踪结果。

文献[8]提出了卷积块注意模块(Convolutional Block

Attention Module, CBAM), 它可以嵌入任何卷积网络架构中, 更加关注重要特征、抑制次要特征, 有效细化了中间特征。文献[9]介绍了一种全新的损失函数 SIoU, 其中考虑到回归之间的向量角度, 重新定义了惩罚指标, 能有效提高训练速度和推理准确性。文献[10]利用简单在线和实时跟踪 Sort 算法, 使用卡尔曼滤波器与匈牙利算法相结合的思想, 在线性环境下取得了不错的效果, 但在非线性环境下, 跟踪效果较差。文献[11]在 Sort 的基础上提出了 DeepSort 算法, 较之前算法引入了外观信息,

收稿日期: 2022-09-25

修回日期: 2022-10-17

基金项目: 国家自然科学基金项目(51774234); 陕西省科技厅项目(2020GY-029); 陕西省科技厅项目(2018GY-151)

预训练时使用关联度量的同时保留了卡尔曼滤波与匈牙利算法结合的方式,在跟踪效果上进一步提升。文献[12]针对传统的JDE算法对小目标跟踪较差问题,引入注意力机制以及特征融合的方法,利用特征金字塔(FPN)与金字塔池化(SPP)有效提升了原算法对小目标的跟踪能力。文献[13]使用YOLOv4检测行人,再利用拓展卡尔曼滤波预测行人在下一帧的位置,将遮挡目标送入轨迹异常修正算法,缓解了因遮挡带来的身份切换问题,具有很好的参考价值。

上述行人目标检测算法,一方面检测精确度与实时性很难兼顾;另一方面在面对复杂路况、行人遮挡等环境,身份切换问题仍有待优化。基于上述原因,本文提出了一种改进的YOLOv5+DeepSort算法,主要工作如下:

- 1) 添加CBAM注意力模块,增强模型特征提取能力,更加关注重要特征,快捷高效地分析复杂场景信息。
- 2) 通过SiLU Loss 损失函数改善定位精度,优化训练阶段的收敛速度,提高推理准确性。
- 3) 通过拓展卡尔曼滤波器对跟踪算法流程进行优化,减少因遮挡带来的行人身份切换频繁问题。

## 1 YOLOv5+DeepSort 算法

### 1.1 YOLOv5 算法

YOLOv5是由作者Glenn Jocher提出,一共有五种大小模型,分别为YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。其中:YOLOv5n的卷积层数最少、检测速率最快、检测精度最低;另外的4个模型,卷积层数从小到大依次排列,随着模型复杂度增大检测速率逐渐降低,检测精度逐渐提高。YOLOv5结构包括四部分:Input、Backbones、Neck、Prediction,如图1所示。

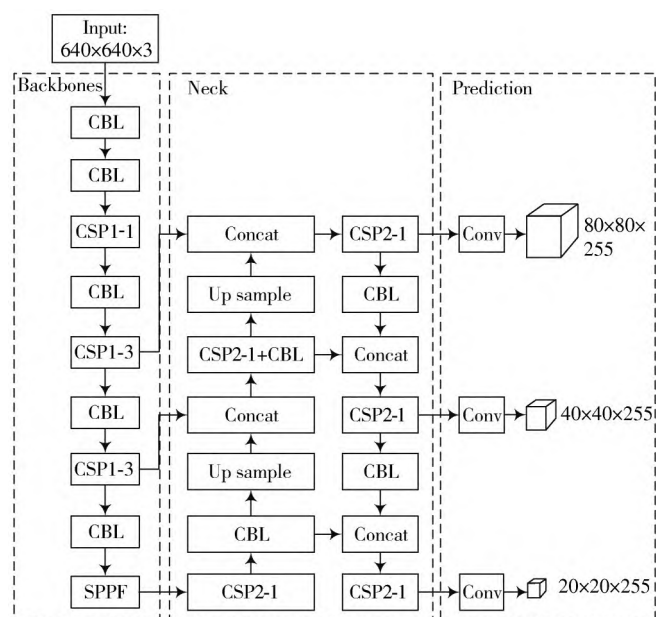


图1 YOLOv5网络总结构图

Input端包含Mosaic增强、动态锚框、自适应图片处理等;Backbones包括跨阶段局部网络CSP、空间金字塔池化SPPF,前者有助于减少计算量,后者提高检测精确度;Neck采用了FPN+PAN的结构,下采样增强语义信息,上采样增强定位信息;Prediction是最后检测输出端。原YOLOv5算法在检测远处小目标行人以及遮挡行人时,特征提取不充分出现部分漏检现象,而行人检测器的精确度会直接影响后续的跟踪效果,因此训练出一个高精度的检测器至关重要。

### 1.2 DeepSort 算法

DeepSort是对Sort追踪算法的改进,跟踪过程中,利用卡尔曼滤波对目标下一帧位置进行预测,再进行级联匹配,最后利用匈牙利算法进行数据关联,提升跟踪效果,减少身份切换问题,整体流程如图2所示,输入视频到检测网络得到行人位置信息,再传输到跟踪网络中进行数据关联并对前后帧的人匹配,得到跟踪结果。

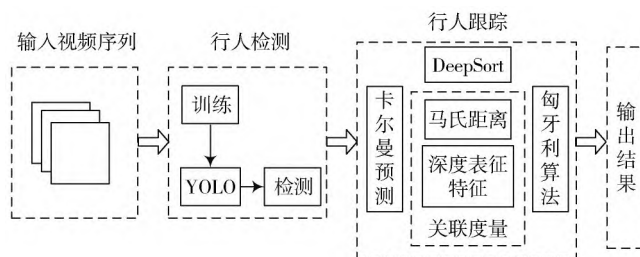


图2 行人检测总流程图

DeepSort计算相似度利用了目标的运动信息和外观信息。对于运动信息,用马氏距离判断预测目标与检测目标的关联度。马氏距离表达式为:

$$d^{(1)}(i, j) = (d_j - y_i)^T S_i^{-1} (d_j - y_i) \quad (1)$$

式中: $d_j$ 是检测框 $j$ 的位置; $y_i$ 是跟踪器 $i$ 预测的位置; $S_i$ 是检测与预测位置的协方差矩阵。

当目标长久遮挡或视角抖动,则要引入外观信息,通过余弦距离来解决因遮挡带来身份切换的问题。

余弦距离表达式为:

$$d^{(2)}(i, j) = \min \{ 1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in R_i \} \quad (2)$$

式中: $r_j$ 是检测框 $d_j$ 的本征向量; $r_k^{(i)}$ 是跟踪器 $i$ 对应最近100帧本征向量的集合; $R_i$ 为外观特征向量库。

为了充分利用两种信息,采用线性加权的方式求和:

$$c_{i,j} = \lambda d^{(1)}(i, j) + (1 - \lambda) d^{(2)}(i, j) \quad (3)$$

式中: $\lambda$ 为超参数,当且仅当度量值 $c_{i,j}$ 存在于 $d^{(1)}(i, j)$ 和 $d^{(2)}(i, j)$ 之间,才认为目标关联。

DeepSort算法卡尔曼滤波只能对线性环境下的目标进行状态预测。在复杂环境如行人遮挡时,卡尔曼滤波不能很好地进行预测,因此改进卡尔曼滤波很有必要。

## 2 改进YOLOv5+DeepSort算法

### 2.1 YOLOv5 算法改进

#### 2.1.1 添加CBAM

注意力机制源于人的直觉,人脑在处理大量信息时,会对重要信息相对关注,对次要信息相对忽略。联合注意力机制(CBAM)结构如图3所示。将输入特征依次经过通道注意力模块与空间注意力模块,并分别进行元素相乘得到与输入时相同维度的输出特征。

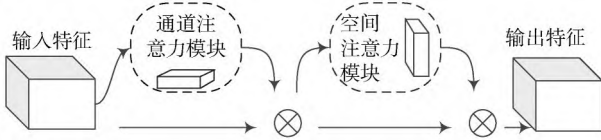


图3 联合注意力机制总结结构图

通道注意力模块如图4所示。首先对提取到的特征进行池化操作,得到特征分别为  $F_{avg}^c$  和  $F_{max}^c$ ,再通过共享全连接层进行按元素求和以及 sigmoid 激活操作,最后得到通道特征图  $M_c(F)$ ,其表达式为:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad (4)$$

式中:  $\sigma$  为 logistid 函数;  $W_0 \in R^{C/r \times C}$  和  $W_1 \in R^{C \times C/r}$  为权重,  $r$  是缩减率。

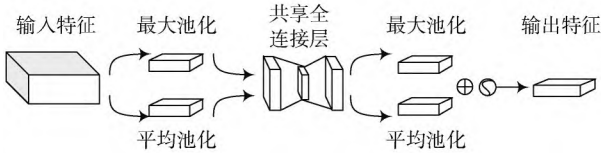


图4 通道注意力机制模块

空间注意力模块如图5所示,输入特征为上面操作得到的  $M_c(F)$ 。首先分别进行最大和平均池化,并将两个结果连接起来利用  $3 \times 3$  卷积降低维度,最后通过 logistid 激活函数得到特征图  $M_s(F)$ 。其表达式为:

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \\ = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (5)$$

式中  $f^{7 \times 7}$  表示卷积核大小。

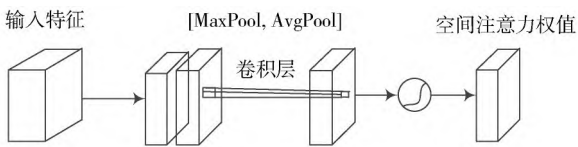


图5 空间注意力机制模块

注意力机制是对特征图重构,其显著效果是突出重要特征、抑制次要特征。Backbones 是特征提取的关键

部分, Neck 层是对特征进行融合送往头部预测,故将 CBAM 添加在 Backbones 与 Neck 之间,结构如图6所示。

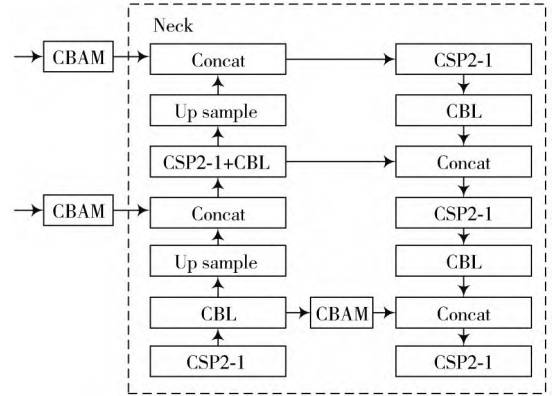


图6 CBAM添加位置

#### 2.1.2 改进损失函数

YOLOv5 使用 CIoU Loss<sup>[14]</sup> 作为边界框回归损失函数,计算预测框与真实框的距离。公式如下:

$$\text{CIoU} = \text{IoU} - \frac{p^2(b, b^{\text{gt}})}{c^2} - av \quad (6)$$

$$L_{\text{CIoU}} = 1 - \text{CIoU}$$

式中:  $\text{IoU}$  表示预测框与真实框的交并比;  $p^2(\cdot)$  为欧氏距离;  $b$  表示预测框中心点坐标;  $b^{\text{gt}}$  表示真实框中心点坐标;  $c$  表示预测框与真实框最小区域的对角线长度;  $a$  表示一个平衡参数;  $v$  表示两框长宽比的一致性。  $a$  与  $v$  的公式如下:

$$a = \frac{v}{1 - \text{IoU} + v} \\ v = \frac{4}{\pi^2} \left( \arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w}{h} \right)^2 \quad (7)$$

式中:  $w^{\text{gt}}, h^{\text{gt}}$  和  $w, h$  分别表示真实框、预测框的宽度及高度。

CIoU 虽然考虑了预测框与真实框的重叠面积、中心点距离以及长宽比,但公式中  $v$  反映的是两者长宽比的差异,而不是长宽与其置信度的真实差异,有时会阻碍模型的优化;而 SIoU 作为一种新的损失函数,重新定义了惩罚度量,考虑了期望回归的向量夹角,能有效提高回归速度和推理的准确性。

SIoU 由四部分组成:角度函数、距离函数、形状函数和 IoU。

角度函数公式如下:

$$\Lambda = 1 - 2\sin^2 \left( \arctan(x) - \frac{\pi}{4} \right) \quad (8)$$

式中  $x$  为预测框和真实框中心点的高度差与其欧氏距离的比值。

距离函数公式如下:



$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma p_t}) = 2 - e^{-\gamma p_x} - e^{-\gamma p_y} \quad (9)$$

式中: $\gamma$ 为 $2 - \Delta$ ;  $p_x$ 为预测框和真实框中心点宽度差与最小外接框宽度比值的平方;  $p_y$ 为预测框和真实框中心点高度差与最小外接框高度比值的平方。

形状函数公式如下:

$$\Omega = \sum_{t=w,h} (1 - e^{-w_t})^\theta = (1 - e^{-w_w})^\theta + (1 - e^{-w_h})^\theta \quad (10)$$

式中: $w_w$ 表示预测框与真实框的宽度差绝对值与两者最大宽度比值; $w_h$ 表示预测框与真实框的高度差绝对值与两者最大高度比值; $\theta$ 为常量,取值范围为[2,6]。

综上,最终的SIoU损失函数表达式如下:

$$L_{\text{SIoU}} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2} \quad (11)$$

式中:IoU为预测框与真实框交并比; $\Delta$ 为距离函数; $\Omega$ 为形状函数。本文选用SIoU来改进边界框损失函数。

## 2.2 DeepSort算法改进

### 2.2.1 拓展卡尔曼滤波

卡尔曼滤波只适用于线性系统中,但在行人跟踪领域存在大量非线性因素,扩展卡尔曼滤波(EKF)通过对非线性函数进行泰勒展开、线性化截断,忽略高阶项,完成函数的近似线性化。

扩展卡尔曼滤波算法在迭代过程中,非线性系统可以用如下公式表示:

$$\begin{aligned} x_k &= f(x_{k-1}) + w_k \\ y_k &= g(x_k) + v_k \end{aligned} \quad (12)$$

式中: $w_k$ 和 $v_k$ 是高斯白噪声; $f(x_{k-1})$ 是状态转移函数; $g(x_k)$ 是量测的非线性函数。迭代公式如下:

$$\begin{aligned} x_k &\approx f(\hat{x}_{k-1} + F_{k-1}(x_{k-1} - \hat{x}_{k-1})) + w_k \\ y_k &\approx g(x_k + Y_k(x_k - \hat{x}_k)) + v_k \end{aligned} \quad (13)$$

式中: $\hat{x}_{k-1}$ 是目标 $k-1$ 帧的估计值; $\hat{x}_k$ 为目标第 $k$ 帧的预测值; $F_{k-1}$ 与 $Y_k$ 分别为函数 $x_k, y_k$ 的雅克比矩阵。

### 2.2.2 改进DeepSort算法

优化DeepSort算法流程图见图7。

优化步骤如下:

- 1) 用YOLOv5检测视频第一帧图像多目标检测框。
- 2) 用卷积神经网络提取各目标特征将其输出,初始化扩展卡尔曼滤波器。
- 3) 根据上一帧目标检测框信息,由扩展卡尔曼预测目标当前帧的位置,输出目标预测框。
- 4) 计算目标检测框与预测框的面积交并比(IoU),进行级联匹配。
- 5) 用匈牙利算法进行最优匹配,成功匹配则输出目标坐标。匹配失败,为目标生成暂时跟踪器,若后面

连续3帧均匹配成功,则视为一个新目标;若后面连续3帧均匹配失败,则视为目标丢失并保留目标信息,接下来60帧没有匹配到该目标,则删除该目标跟踪器。

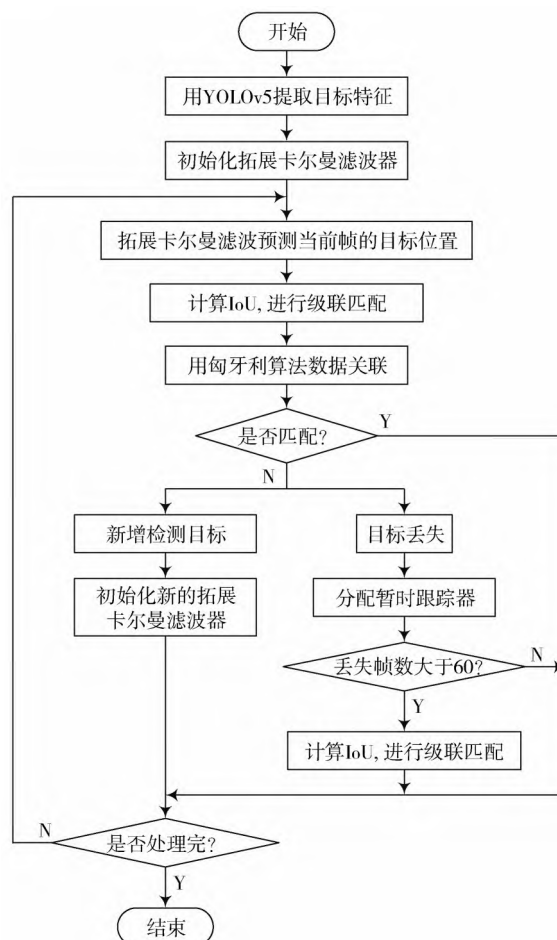


图7 改进DeepSort算法流程图

6) 迭代进行步骤3)~步骤5),直到所有图像处理完成。

## 3 实验结果与分析

### 3.1 YOLOv5消融实验

在实验阶段,采用公共数据集 WiderPerson 拥挤场景行人检测数据集,共包含 13 382 张图片。数据集划分为训练集 8 000 张、验证集 1 000 张、测试集 4 382 张。实验的硬件环境是 NVIDIA RTX2080Ti GPU,软件环境是 Python 3.6, PyTorch 1.7。batch\_size 为 16, epochs 为 1 000,采用 YOLOv5s 预训练模型加速训练。为验证第 1 节的改动进行消融实验,判断每个改进点的有效性,依次在 YOLOv5 的基础上加入 CBAM、SIoU Loss。采用精确度(Precision)、召回率(Recall)和平均精确度(AP)作为评价指标,实验结果见表 1。

由表 1 可知,总体改进 YOLOv5 相比于原算法 Precision 提升了 5%, Recall 提升了 7%, AP 提升了 4%,验证了改

进模型的有效性。

表1 不同方法评估结果

Models	Precision	Recall	AP
baseline	0.85	0.78	0.85
YOLOv5+CBAM	0.89	0.83	0.88
YOLOv5+SIoU	0.86	0.83	0.86
YOLOv5+CBAM+SIoU	0.90	0.85	0.89

为了更加直观地表现不同方法改进YOLOv5与baseline在行人检测时的优势,图8将展示两者在WiderPerson测试数据集的具体图片。图8左侧均为baseline检测结果,右侧均为改进YOLOv5后检测结果。由图8a)可看出,baseline对远处行人存在2个漏检,而添加CBAM后,成功检测到这2人;由图8b)可看出,baseline对部分遮挡的人存在1处漏检,引入Siou后,成功检测到这1人;由图8c)可看出,baseline对遮挡、远处行人共存在5处漏检,同时引入CBAM+SioU后,成功检测到这5人。



图8 不同方法改进与baseline结果对比

3.2 改进YOLOv5+DeepSort实验

本阶段采用行人公共数据集MOT-16进行实验,该数据集是用于评价多目标跟踪算法性能的常用数据集,有多种多样行人场景,包括14个视频序列,有7个视频序列详细标注了行人身份和边框位置,用于训练多目标行人跟踪算法;另外7个视频序列作为测试集。

实验采用评价标准为:跟踪准确度(MOTA)、跟踪精度(MOTP)、目标轨迹正确跟踪80%以上的个数

(MT)、目标轨迹正确跟踪20%以下的个数(ML)、目标ID变换次数(IDSW),跟踪准确度与跟踪精度计算公式如下:

$$MOTA = 1 - \frac{\sum_t (m_t + n_t + s_t)}{\sum_t g_t}$$
$$MOTP = \frac{\sum_{i,t} d_i}{\sum_t c_t}$$

(14)

式中: $t$ 表示第 $t$ 帧; $m_t$ 表示 $t$ 帧漏检目标数目; $n_t$ 表示 $t$ 帧误检目标数目; $s_t$ 表示 $t$ 帧身份切换数目; $g_t$ 表示 $t$ 帧出现目标总数目; $d_i$ 表示第 $t$ 帧目标 $i$ 预测位置与真实位置之间的距离; $c_t$ 表示第 $t$ 帧中成功匹配的目标数。

该次实验均在MOT-16的测试集视频序列上进行,为了更好地凸显本文算法的性能,设计了两组不同实验进行对照比较。

实验1:将改进YOLOv5分别与原DeepSort以及改进DeepSort算法结合,放在测试集视频测试,分析改进DeepSort算法后优劣性,后续将改进YOLOv5+改进DeepSort称为改进算法。

实验2:分析在多样场景下改进算法的鲁棒性。将改进算法与几种主流算法对比,分析改进算法的优劣性。

实验1结果见表2,改进DeepSort算法后,各项指标都略有提升,跟踪准确度(MOTA)提升了1.2%,跟踪精度(MOTP)提升了0.2%,目标轨迹正确跟踪80%以上的个数(MT)提升了1.7%,目标轨迹正确跟踪20%以下的个数(ML)降低了0.1%,目标身份切换总数(IDSW)减少了31次,可见改进DeepSort能减少行人切换次数。

表2 DeepSort算法改进前后在MOT-16数据集的实验结果

算法	MOTA ↑	MOTP ↑	MT ↑	ML ↓	IDSW ↓
改进YOLOv5+原DeepSort	62.3	79.3	32.5	18.7	772
改进算法	63.5	79.5	34.2	18.6	741

图9将展示DeepSort算法改进前后检测MOT-16数据集的具体测试图,此检测图片来自MOT16-07序列的第450帧、第470帧、第498帧,分别对应行人遮挡前、行人遮挡中、行人遮挡后。从图9可看出,改进DeepSort前,行人1号遮挡后变成15号,而改进后,行人3号遮挡后仍为3号。

实验2结果见表3。将改进算法与Sort、DeepSort、JDE、原始YOLOv5+DeepSort进行对比,改进算法的跟踪准确度(MOTA)为63.5%,比DeepSort高出2.1%;跟踪精度(MOTP)为79.5%,比Sort算法低0.1%;目标轨迹正确跟踪80%以上的个数(MT)为34.2%,也是五种算法中最好的;目标轨迹正确跟踪20%以下的个数(ML)仅比DeepSort算法高0.3%;目标身份切换总数(IDSW)为741次,在五种算法中排第一;与原始YOLOv5+DeepSort

算法对比每项指标均有所提升,其中跟踪准确度(MOTA)提升3.4%,目标身份切换总数(IDSW)减少52次。



a) 改进YOLOv5+原DeepSort跟踪结果



b) 改进算法跟踪结果

图9 算法改进前后跟踪对比

表3 本文算法与主流算法在MOT-16数据集的实验结果

算法	MOTA ↑	MOTP ↑	MT/% ↑	ML/% ↓	IDSW ↓
Sort	59.8	79.6	25.4	22.7	1 423
DeepSort	61.4	79.1	32.8	18.2	781
JDE	62.1	75.4	35.4	20.0	1 544
原始YOLOv5+ DeepSort	60.1	78.5	30.9	18.8	793
Ours	63.5	79.5	34.2	18.5	741

综上所述,改进算法在整体效果上是五种算法中最优的算法。此外,在实验的硬件条件下,改进算法检测速率达到每秒25帧,能满足实时检测要求。

## 4 结 语

对YOLOv5+DeepSort算法进行改进,在检测算法颈部融合多个CBAM模块,增强了模型的特征提取能力;再用Siou改进边界回归损失函数,提高了模型的定位精度,改善了行人密集带来的漏检问题;跟踪算法用拓展卡尔曼滤波器(EKF)替换卡尔曼滤波器,增强非线性环境下算法的鲁棒性。由实验结果可知,改进算法在数据集MOT-16上与改进前的算法或Sort、DeepSort、JDE算法相比,整体效果是最优的,行人跟踪准确度得到提

升,且减少了因遮挡带来的行人身份切换次数,但是改进算法在较暗的环境下跟踪效果较差,后续的研究将增强此类数据,进一步提升算法的鲁棒性。

注:本文通讯作者为王雨田。

## 参 考 文 献

- [1] 樊璐,张轶.一种新的行人多目标跟踪算法[J].计算机应用与软件,2021,38(4):190-196.
- [2] 温宇.多特征融合和改进神经网络的运动视频镜头转换检测[J].现代电子技术,2021,44(1):74-77.
- [3] GUPTA A, ANPALAGAN A, GUAN L, et al. Deep learning for object detection and scene perception in self-driving cars: survey, challenges, and open issues [J]. Array, 2021,10(10): 57-77.
- [4] BAUMGARTL H, SAUTER D, SCHENK C, et al. Vision-based hand gesture recognition for human-computer interaction using MobileNetV2 [C]// 2021 IEEE 45th Annual Computers Software and Applications Conference. Madrid: IEEE, 2021: 1667-1674.
- [5] LUO Wenhan, XING Junliang, ZHANG Xiaoqin, et al. Multiple object tracking: a literature review [J]. Artificial intelligence, 2017, 293(2): 3-26.
- [6] 金立生,华强,郭柏苍,等.基于优化DeepSort的前方车辆多目标跟踪[J].浙江大学学报(工学版),2021,55(6):1056-1064.
- [7] 黄凯文,凌六一,王成军,等.基于改进YOLO和DeepSort的实时多目标跟踪算法[J].电子测量技术,2022,45(6):7-13.
- [8] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]// Proceedings of the European Conference on Computer Vision. Munich: ECCV, 2018: 3-19.
- [9] GEVORGYAN Z. Siou loss: more powerful learning for bounding box regression [J]. Computer science, 2022, 31(5): 127-139.
- [10] BEWLEY A, GE Z, OTT L, et al. Simple online and real-time tracking [C]// 2016 IEEE International Conference on Image Processing. Arizona: IEEE, 2016: 3464-3468.
- [11] WOJKE N, BEWLEY A, PAULUS D. Simple online and real-time tracking with a deep association metric [C]// 2017 IEEE International Conference on Image Processing. Beijing: IEEE, 2017: 3645-3649.
- [12] 晏康,曾凤彩,何宁,等.引入注意力机制的JDE多目标跟踪方法[J].计算机工程与应用,2022,58(21):189-196.
- [13] 朱新丽,才华,寇婷婷,等.行人多目标跟踪算法[J].吉林大学学报(理学版),2021,59(5):1161-1170.
- [14] 李永上,马荣贵,张美月.改进YOLOv5s+DeepSort的监控视频车流量统计[J].计算机工程与应用,2022,58(5):271-279.

作者简介:韩晓冰(1965—),男,教授,研究方向为数字化矿山技术和视频图像处理。

王雨田(1996—),男,硕士研究生,研究方向为目标检测与跟踪。

黄综浏(1998—),男,硕士研究生,研究方向为目标检测与跟踪。

张玮良(1997—),男,硕士研究生,研究方向为目标检测。