

文章编号: 1671-251X(2025)02-0065-12

DOI: 10.13272/j.issn.1671-251x.2024110022

基于自适应链接优化的井下行人抗遮挡跟踪方法研究

路洋, 董立红, 叶鸥

(西安科技大学 通信与信息工程学院, 陕西 西安 710054)

摘要: 针对煤矿井下行人因遮挡频繁和外观混淆导致轨迹匹配不准确的问题, 提出了一种基于自适应链接优化的井下行人抗遮挡跟踪方法。首先, 根据目标置信度变化率和交并比计算, 对目标进行遮挡判定, 筛选出潜在遮挡目标。然后, 在匹配级联阶段, 引入潜在遮挡目标的非线性动态特征, 并结合历史轨迹信息扩展轨迹链接优化模块的轨迹对输入, 同时在轨迹对输入进行时域块处理后添加通道先验卷积注意力机制, 增强时域表征能力。轨迹对输入向量经压缩与融合处理后, 由多层感知器输出轨迹相似性得分, 与原有匹配级联阶段中卡尔曼滤波器的总成本函数相结合, 优化匹配决策, 有效缓解轨迹匹配过程中的错误匹配问题。最后, 在交并比匹配阶段, 通过计算断裂率和 ID 切换率的变化量, 引入自适应 RB 因子, 构建反馈机制, 动态调整匹配决策中的交并比阈值, 以适应因长时间遮挡导致的轨迹断裂问题。采用所提方法与 DeepSORT, YOLOv7-SAM, OSNet, FuCoLoT 对煤矿井下典型视频序列进行对比实验, 结果表明, 所提方法的跟踪准确度 (MOTA), 跟踪精度 (MOTP) 和身份 F1 (IDF1) 分别为 76.17%, 84.13%, 74.9%, 较 DeepSORT 分别提升了 14.9%, 1.83% 和 10.93%, 较 YOLOv7-SAM 分别提升了 1.57%, 0.4% 和 0.37%, 较 OSNet 分别提升了 2.83%, 0.77% 和 1.27%, 较 FuCoLoT 分别提升了 2.5%, 0.08% 和 1.8%, 说明所提方法能够有效解决煤矿井下目标在遮挡情形下的跟踪误匹配问题。

关键词: 井下人员跟踪; 多目标跟踪; 目标遮挡; 跟踪误匹配; 轨迹链接优化; 轨迹断裂

中图分类号: TD76

文献标志码: A

Research on anti-occlusion tracking method for underground mine personnel based on adaptive link optimization

LU Yang, DONG Lihong, YE Ou

(College of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an 710054, China)

Abstract: To address the issue of inaccurate trajectory matching caused by frequent occlusions and appearance confusion of underground mine personnel in coal mines, an anti-occlusion tracking method for underground mine personnel based on adaptive link optimization was proposed. Firstly, occlusion detection of the targets was performed based on the target confidence change rate and intersection-over-union (IoU) calculation to identify potential occluded targets. Secondly, in the matching cascade stage, nonlinear dynamic features of potential occluded targets were introduced, and historical trajectory information was incorporated to expand the trajectory pair input for the trajectory link optimization module. Additionally, after performing time-domain block processing on the trajectory pair input, a channel prior convolutional attention mechanism was added to enhance the time-domain representation capability. After compression and fusion processing of the trajectory pair input vectors, a trajectory similarity score was output by the multilayer perceptron. This score was combined with the

收稿日期: 2024-11-18; 修回日期: 2025-02-01; 责任编辑: 王晖, 郑海霞。

基金项目: 国家自然科学基金青年项目 (62303375)。

作者简介: 路洋 (2000—), 男, 安徽合肥人, 硕士研究生, 主要研究方向为计算机视觉和图像处理、人工智能, E-mail: luyang20000520@163.com。

引用格式: 路洋, 董立红, 叶鸥. 基于自适应链接优化的井下行人抗遮挡跟踪方法研究[J]. 工矿自动化, 2025, 51(2): 65-75, 137.

LU Yang, DONG Lihong, YE Ou. Research on anti-occlusion tracking method for underground mine personnel based on adaptive link optimization[J]. Journal of Mine Automation, 2025, 51(2): 65-75, 137.



扫码移动阅读

total cost function of the Kalman filter in the original matching cascade stage to optimize matching decisions, effectively alleviating the issue of incorrect matching during the trajectory matching process. Finally, in the IoU matching stage, an adaptive RB factor was introduced by calculating the variations in fracture rate and ID switch rate to construct a feedback mechanism. This mechanism dynamically adjusted the IoU threshold in the matching decision to address trajectory fragmentation caused by long-term occlusion. Comparative experiments were conducted on typical video sequences from underground coal mines using the proposed method, DeepSORT, YOLOv7-SAM, OSNet, and FuCoLoT. The results showed that the proposed method achieved the multiple object tracking accuracy (MOTA) of 76.17%, the multiple object tracking precision (MOTP) of 84.13%, and the identity F1 (IDF1) of 74.9%. Compared to DeepSORT, these values improved by 14.9%, 1.83%, and 10.93%, respectively. Compared to YOLOv7-SAM, they improved by 1.57%, 0.4%, and 0.37%, respectively. Compared to OSNet, they improved by 2.83%, 0.77%, and 1.27%, respectively. Compared to FuCoLoT, they improved by 2.5%, 0.08%, and 1.8%, respectively. This demonstrates that the proposed method can effectively address the issue of tracking mismatches in occlusion scenarios in underground coal mine targets.

Key words: underground mine personnel tracking; multiple target tracking; target occlusion; tracking mismatch; trajectory link optimization; trajectory fragmentation

0 引言

目标跟踪是计算机视觉领域中一个重要研究方向,在煤矿井下智能监控系统中,借助目标跟踪技术对井下人员、车辆、运输设备等运动目标进行实时自动跟踪,是行人计数、异常识别等高级视觉任务的基础^[1]。煤矿井下环境复杂,行人目标经常受到被设备、支架等遮挡和短暂消失(如进入盲区)的影响,传统目标跟踪算法在应对外观特征丢失和轨迹断裂问题上表现出准确性和鲁棒性不足,易导致错误匹配和目标丢失。因此,在遮挡频繁且目标密集的复杂场景中,传统目标跟踪算法难以满足多目标跟踪的精度和连续性需求^[2]。

目标遮挡是目标跟踪中的常见问题,1个或多个目标在视场中被其他物体部分或完全遮挡,在跟踪过程中将面临目标丢失的问题,从而导致对目标数量的估计产生偏差^[3]。在煤矿领域,目标遮挡呈现更多特有的问题:视频质量存在差异,有的较为模糊;背景信息较少,整体呈现暗灰色调。许多学者针对目标遮挡的问题进行了研究。文献^[4]提出了一种基于深度估计的尺度自适应目标跟踪方法,利用目标深度信息估计目标尺度,对目标尺度变化及形变具有很好的适应性。然而,该方法对深度信息的依赖较高,在实际应用中,深度信息的获取可能会受到环境因素(如低光照、遮挡等)影响,从而降低精度。文献^[5]提出的行人检测方法在双阶段检测器 Faster R-CNN 的基础架构上引入注意力机制网络,提高了模型在遮挡时的检测能力。但该方法计算复杂度较高,在实际应用中可能面临性能瓶颈。文献^[6]基于跟踪学习检测(Track Learning Detection, TLD)

框架,提出了一种融合多特征的复杂场景动态目标长时间视觉跟踪算法,使用融合多特征的 MeanShift 算法^[7]替代 TLD 框架中的光流法,对目标形变有较强的鲁棒性。然而,MeanShift 算法对目标初始位置的选择较为敏感,当目标发生大幅度移动时,算法可能无法准确收敛。文献^[8]提出了一种基于注意力机制的无监督矿井人员跟踪算法,结合相关滤波和孪生网络在跟踪任务上的优势,构建轻量化目标跟踪模型。虽然该算法可有效减少计算开销,但无监督算法的跟踪性能常常无法与有监督算法相比,尤其是在复杂背景和多目标情况下。文献^[9]提出了一种改进的异常抑制相关滤波跟踪方法,专门用于在红外视频中追踪地下矿山中的人体。通过分析响应图的峰型形态和峰值旁瓣比变化特征,设计了自适应外观模型更新策略,提高了井下人员目标的跟踪精准性和鲁棒性。然而,该方法主要依赖红外视频的特征提取,在低光环境下可能表现较好,但在有较强背景干扰的环境中,其鲁棒性会受到影响,导致跟踪精度降低。文献^[10]提出了基于深度学习的矿井巷道人员计数算法,通过引入全尺度特征提取模型 OSNet 优化 DeepSORT,替换原有卷积神经网络模块。OSNet 采用深度可分离卷积和多尺度特征流,增强了目标细粒度特征的提取能力。但该算法在训练时需要较大规模的数据集,并且深度可分离卷积的计算开销较高,可能影响算法在实时检测中的应用效率。文献^[11]提出了基于深度学习的井下运动目标跟踪算法,采用一种局部-全局匹配网络来提高算法在不同场景下的跟踪精度,并设计了一个无锚框的分类回归网络,降低模型计算复杂度。尽管该

算法有效减少了计算量,但无锚框设计可能导致在密集场景中目标检测精度下降。文献[12]提出了基于改进YOLOv5s和DeepSORT的井下人员检测及跟踪算法,通过改进YOLOv5s网络的自注意力模块,提高井下不同尺度人员目标的检测精度,采用更深卷积层替换DeepSORT算法中小型残差网络,提升对人员外观信息的提取能力。但该算法的主要问题是自注意力模块可能增加模型的计算复杂度,且在高遮挡或低光照环境下,性能提升可能有限。文献[13]提出了基于改进YOLOv7和DeepSORT的井下人员检测与跟踪算法,通过在YOLOv7的Neck模块中融入SimAM注意力机制,以及在DeepSORT特征提取网络中引入ShuffleNetV2轻量化模块,显著提升了井下复杂环境中人员目标检测和跟踪的实时性能。但面对多发遮挡场景,该算法可能无法充分发挥其优势。

针对遮挡频繁和外观混淆导致轨迹匹配不准确的问题,本文提出了一种基于自适应链接优化的井下行人抗遮挡跟踪方法。该方法在DeepSORT框架的基础上,通过遮挡判定模块、轨迹链接优化模块(CoLinker)及自适应RB因子反馈机制实现优化。首先,在遮挡判定模块中,结合目标置信度变化率与交并比(IoU),筛选潜在遮挡目标,为后续处理提供可靠输入;然后,在匹配级联阶段,通过CoLinker模块引入非线性动态特征与通道先验卷积注意力机制(Channel Prior Convolutional Attention, CPCA)^[14],增强时序特征的代表能力,优化轨迹匹配决策;最后,在IoU匹配阶段,通过自适应RB因子构建反馈机制动态调整匹配阈值,缓解因长期遮挡导致的轨迹断裂问题,从而在复杂环境中保持高鲁棒性,最终实现对遮挡目标的跟踪优化。

1 井下行人目标遮挡分析

目标遮挡在计算机视觉和视频监控领域中指的是目标物体在视场中被其他物体部分或完全覆盖的情况。根据目标的遮挡程度,可以分为无遮挡(0)、轻度遮挡(1%~10%)、部分遮挡(10%~35%)、严重遮挡(35%~80%)和完全遮挡($\geq 80\%$)^[15]。为了对目标遮挡过程进行描述,建立目标遮挡模型。当遮挡发生时总会出现被遮挡目标ID1(绿框)和障碍目标ID2(红框),如图1所示。

当2个目标发生遮挡时,跟踪失败往往发生在被遮挡目标上。为了及时判定目标发生遮挡的时刻,通过记录一段视频序列中行人目标的置信度(图2),探究置信度是否可以作为遮挡判定依据。



图1 遮挡发生时被遮挡目标和障碍目标

Fig. 1 Occluded and obstacle targets during occlusion

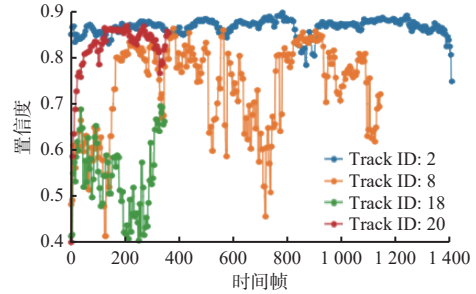


图2 一段视频序列井下行人目标置信度变化

Fig. 2 Variation in underground mine personnel target in a video sequence

由图2可看出,Track ID: 8和Track ID: 18在特定时刻波动较大,目标之间的遮挡和相互影响导致置信度下降,说明将置信度作为遮挡判定因素较合理。

考虑到视频跟踪的时序连续性,跟踪任务中的连续帧特性使置信度波动具有时序意义。通过记录行人目标的置信度,进一步分析其在连续帧中的变化情况。在 k 时刻,置信度变化率为

$$\Delta O_{\text{track_id},k} = \begin{cases} \frac{O_{\text{track_id},k} - O_{\text{track_id},k-1}}{O_{\text{track_id},k-1}} & O_{\text{track_id},k} - O_{\text{track_id},k-1} > 0 \\ 0 & \text{其他} \end{cases} \quad (1)$$

式中 $O_{\text{track_id},k}$, $O_{\text{track_id},k-1}$ 分别为第 k , $k-1$ 时刻跟踪对象的置信度分数。

通常目标检测概率取决于目标状态,将置信度变化率作为遮挡发生时对目标状态的影响因素有所欠缺。为了更精确地评估目标在复杂环境中的遮挡问题,优化跟踪算法,提高模型的鲁棒性和适应性,本文采用ID切换率 R_k 和跟踪断裂率 B_k 来评估第 k 时刻的目标遮挡概率,实现对遮挡目标的更高效跟踪。ID切换率 R_k 表示在第 k 时刻的跟踪切换次数与累计跟踪次数的比值,反映了目标在被遮挡时的ID切换频率。跟踪断裂率 B_k 表示在第 k 时刻跟踪目标中未被跟踪到的目标数占前一时刻总目标数的比例,反映了目标跟踪过程中的断裂情况。

$$R_k = S_k / N_k \quad (2)$$

$$N_k = \sum_{b=1}^k |D_b| \quad (3)$$

$$S_k = \sum_{b=1}^{k-1} S_w(b, b+1) \quad (4)$$

$$B_k = \frac{|D_{k-1} \cap (-D_k)|}{|D_{k-1}|} \quad (5)$$

式中： N_k 为累计到第 k 时刻的所有有效跟踪实例的数量； S_k 为累计到第 k 时刻的ID切换数量； $|D_b|$ 为在时刻

b 检测到的目标； $S_w(b, b+1)$ 为在第 b 时刻到第 $b+1$ 时刻之间发生ID切换的数量； $|D_{k-1} \cap (-D_k)|$ 为跟踪断裂的目标集合， D_{k-1} 为在第 $k-1$ 时刻被检测且跟踪的目标， $-D_k$ 为在第 k 时刻没有被跟踪到的目标， $-$ 为逻辑非。

2 自适应链接优化井下行人抗遮挡跟踪方法

基于自适应链接优化的井下行人抗遮挡跟踪方法架构如图 3 所示。

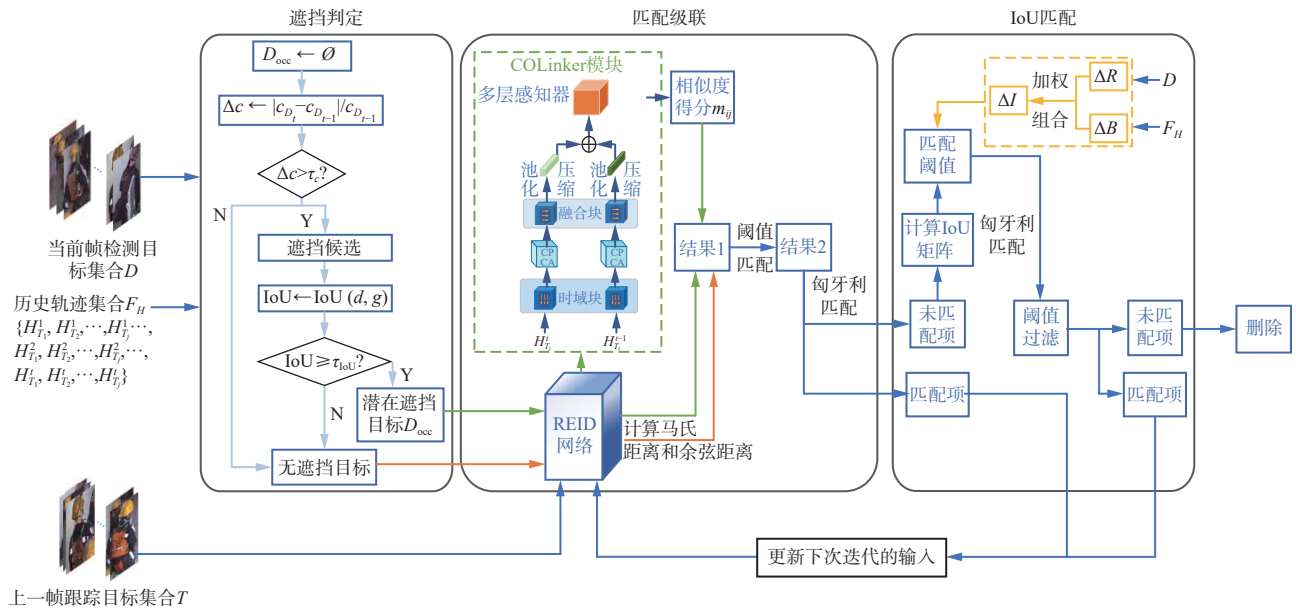


图 3 基于自适应链接优化的井下行人抗遮挡跟踪方法架构

Fig. 3 Architecture of anti-occlusion tracking method for underground mine personnel based on adaptive link optimization

传统 DeepSORT 算法缺乏对遮挡状态的动态判定机制，难以区分正常目标与潜在遮挡目标，误匹配问题突出。本文提出遮挡判定模块，并将其引入 DeepSORT 算法的检测与轨迹初始化阶段。通过结合目标置信度变化率和IoU，对目标的遮挡状态进行精准判定。通过检测目标在相邻帧间的置信度变化，快速判定可能被遮挡的目标，结合目标与轨迹框的IoU，进一步过滤误判的干扰目标。

传统 DeepSORT 算法依赖静态外观特征与固定运动模型，难以适应行人目标因遮挡导致的运动突变。本文在匹配级联模块中，引入基于非线性动态特征的 COLinker 模块。接收来自 REID 网络的输入，并在计算马氏距离和余弦距离的基础上添加相似度得分。在 COLinker 模块中，通过扩展轨迹对的输入向量并引入 CPCA 注意力机制，增强了轨迹链接过程对时间序列特征的捕捉能力。轨迹对输入经过多层特征融合和压缩处理后，通过多层感知器 (Multilayer Perceptron, MLP) 输出轨迹相似性得分，并将该得分与卡尔曼滤波器的总成本函数相结合，

从而优化匹配决策。

传统 DeepSORT 算法在 IoU 匹配阶段采用固定阈值，无法根据遮挡程度动态调整，导致轨迹断裂与 ID 切换率升高。本文在 IoU 匹配阶段提出了基于自适应 RB 因子的反馈机制。通过计算轨迹的断裂率和 ID 切换率的变化量，动态调整 IoU 匹配阶段的阈值，使模型能够在复杂遮挡场景中更灵活地进行匹配优化。

2.1 遮挡判定

遮挡判定的基本原理是通过检测目标置信度的剧烈变化并结合目标检测框与轨迹框的IoU来识别潜在的遮挡目标。当目标被部分遮挡或完全遮挡时，检测目标的置信度在相邻帧间发生显著下降。本文通过计算置信度的变化率来量化该变化，并设置阈值来判定是否发生遮挡。

遮挡判定算法的输入包括当前帧的检测集合和上一帧完成关联的轨迹集合，以及置信度变化阈值和IoU阈值。

首先，初始化潜在遮挡目标集合 D_{occ} ，用于存储

符合遮挡条件的检测目标。其次,遍历检测目标集合 D ,计算每个目标的置信度变化率 Δc ,即该目标在当前帧和上一帧的置信度差值的归一化值。如果目标的置信度变化率 Δc 超过预设阈值 τ_c ,则判定该目标处于遮挡状态,需进一步验证。然后,对每个潜在遮挡目标,与上一帧轨迹集合 F_H 中对应轨迹框进行IoU计算。如果目标检测框 d 与轨迹框 g 间的IoU大于设定阈值 τ_{IoU} ,则进一步确认目标处于遮挡状态,并将其加入潜在遮挡目标集合。通过置信度变化率与IoU的双重筛选,以确保对目标遮挡状态的判定更加可靠。最后,遮挡判定模块输出潜在遮挡目标集合,为后续的COLinker模块提供输入。在轨迹链接优化模块中,这些潜在遮挡目标将作为候选目标,与轨迹片段进一步匹配优化。通过遮挡判定模块的前置筛选,能够有效减少后续轨迹链接过程中的干扰目标,从而提升遮挡目标的跟踪精度。

2.2 COLinker 模块

现有算法利用状态估计和协方差矩阵来预测每个跟踪目标的新位置,通过匹配策略关联跟踪目标^[16]。但是煤矿井下环境恶劣,矿工着装统一,传统DeepSORT算法中仅依赖欧氏距离和余弦相似度的匹配策略,由于过度侧重空间位置特征,难以适应目标运动模式多样化的井下跟踪场景^[17]。为了解决这一问题,本文引入了轨迹链接模块AFLink^[18],以优化DeepSORT算法中目标匹配过程。但AFLink缺乏对目标运动特征的考虑,在目标运动模式多样化或遮挡频繁的场景下,匹配精度和鲁棒性较差。针对这一不足,本文首先扩展了AFLink的输入,加入目标的非线性动态特征,构建OLinker模块,以提高目标匹配的精度和鲁棒性。对扩展输入向量进行时域块处理后,为了加强对连续帧之间的关注度,聚焦关键特征,引入CPCA注意力机制,构建轨迹链接优化COLinker模块,改善轨迹之间的相似性评估和匹配精度,从而优化轨迹链接过程。具体优化过程如图4所示。

对于每个被判定为潜在遮挡状态的跟踪目标 T_j 和检测目标 D_j ,通过YOLOv9模型^[19]获取外观特征。结合跟踪历史中位置序列 $Q_{T_j}^t$,为非线性动态特征的计算提供基础。

$$Q_{T_j}^t = \{C_{T_j}^1, C_{T_j}^2, \dots, C_{T_j}^t\} \quad (6)$$

$$C_{T_j}^t = \left(x_{T_j}^t + \frac{w_{T_j}^t}{2}, y_{T_j}^t + \frac{h_{T_j}^t}{2} \right) \quad (7)$$

式中: $C_{T_j}^t$ 为第 j 个跟踪目标 T_j 在 t 时刻的中心点位置; $(x_{T_j}^t, y_{T_j}^t, w_{T_j}^t, h_{T_j}^t)$ 为边界框四元组, $x_{T_j}^t, y_{T_j}^t, w_{T_j}^t, h_{T_j}^t$ 分

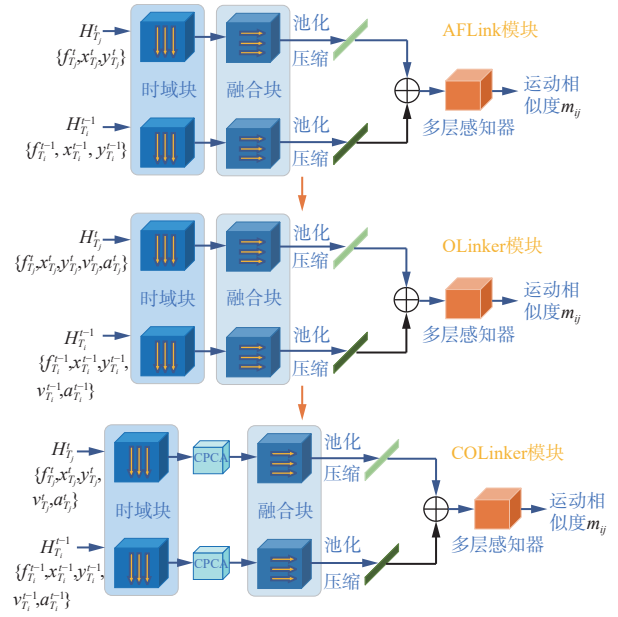


图4 轨迹链接优化过程

Fig. 4 Trajectory link optimization process

别为跟踪目标 T_j 在 t 时刻的坐标。

目标 T_j 在 t 时刻的速度 $v_{T_j}^t$ 、加速度 $a_{T_j}^t$ 可表示为

$$v_{T_j}^t = Q_{T_j}^t - Q_{T_j}^{t-1} \quad (8)$$

$$a_{T_j}^t = v_{T_j}^t - v_{T_j}^{t-1} \quad (9)$$

式中: $v_{T_j}^{t-1}$ 为目标 T_j 在 $t-1$ 时刻的速度; $Q_{T_j}^t, Q_{T_j}^{t-1}$ 分别为目标 T_j 在 $t, t-1$ 时刻的位置。

CPCA利用通道注意力机制突出特征图中关键的通道信息,其结构如图5所示。首先通过多尺度卷积核提取输入特征图中的多层次信息。然后利用通道注意力图,动态分配各个通道的权重,从而突出有用的通道特征。最后空间注意力机制结合通道级别特征,进一步通过生成空间注意力图调整不同区域的关注程度。

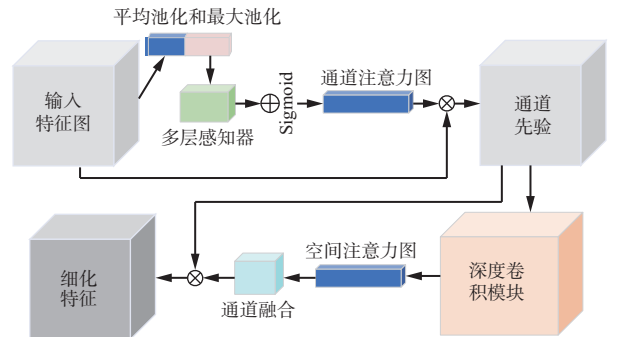


图5 CPCA结构

Fig. 5 CPCA structure

COLinker的匹配级联模块在匹配过程不仅考虑外观相似度 s_{ij} 和空间位置的马氏距离 $M(D_i, T_j)$,还引入了运动相似度 m_{ij} ,以提高目标匹配的有效性。 m_{ij} 由COLinker模块输出的概率 $P(S|D_i, T_j)$ 转换而

来,表示 2 个轨迹片段之间属于同一目标的概率。

$$m_{ij} = L(H_{T_j}^t, H_{T_i}^{t-1}) \quad (10)$$

$$H_{T_j}^t = \{f_{T_j}^t, x_{T_j}^t, y_{T_j}^t, v_{T_j}^t, a_{T_j}^t\} \quad (11)$$

式中 $L(\cdot)$ 为使用马氏距离计算轨迹片段的相似度函数,它将输入的轨迹片段 $H_{T_j}^t$ 和 $H_{T_i}^{t-1}$ 转换为张量 X_j 和 X_i ,以计算第 i 个和第 j 个跟踪目标是否为同一目标; $H_{T_j}^t$ 为 t 时刻的第 j 个跟踪目标轨迹段; $H_{T_i}^{t-1}$ 为 $t-1$ 时刻的第 i 个跟踪目标轨迹段; $f_{T_j}^t$ 为目标 T_j 在 t 时刻的外观特征。

因此匹配级联模块的输出总成本函数可表示为

$$G_{ij} = \lambda_1 M(D_i, T_j) + \lambda_2 (1 - s_{ij}) + \lambda_3 (1 - m_{ij}) \quad (12)$$

式中 λ_1 , λ_2 和 λ_3 为权重参数,用来平衡马氏距离、外观相似度和运动相似度的影响。

2.3 自适应 RB 因子

当长时间发生遮挡且设定匹配阈值过高时,可能因为重叠区域未达到高阈值要求而不被视为同一目标^[20];当遮挡发生较少时,适当提高匹配阈值,可以使模型更关注遮挡区域。为此,本文提出了自适应 RB 因子,依据 k 时刻的 R_k 和 B_k 及 $k-1$ 时刻的 R_{k-1} 和 B_{k-1} ,动态调整匹配阈值,以确保在遮挡区域内的正确跟踪。

1) 计算跟踪断裂率和 ID 切换率的变化量:

$$\Delta B = \alpha(B_k - B_{k-1}) \quad (13)$$

$$\Delta R = \alpha(R_k - R_{k-1}) \quad (14)$$

式中 α 为转换系数,将变化率转换为百分比, $\alpha=100$ 。

2) 定义 R_k 和 B_k 的权重 ω_R 和 ω_B ,计算 IoU 阈值的总调整量 ΔI :

$$\Delta I = \omega_R \Delta R + \omega_B \Delta B \quad (15)$$

3) 为了确保自适应调整策略在一个合理的范围内进行,对 IoU 阈值进行更新。

$$I_k = \max(\min(I_{k-1} - \Delta I, I_{\max}), I_{\min}) \quad (16)$$

式中: I_{k-1} 为 $k-1$ 阈值; I_{\max} 和 I_{\min} 分别为 IoU 阈值的上下限值。

3 实验验证

3.1 煤矿井下人员多目标跟踪数据集

为验证基于自适应链接优化的井下行人抗遮挡跟踪方法的性能,选取井下人员视频数据来构建煤矿井下人员多目标跟踪数据集 (Mine Personnel Detection Dataset, MPDD),包括井下闸机出入口、井底硐室和井下工作面 3 种场景的图像数据(图 6),共计 14 个视频片段,共 64 770 帧。使用 DarkLabel 软件进行数据标注,并将其转换为标注的 MOT (Multiple Object Tracking, 多目标跟踪)数据集格式。数据集详细情况与数据特点见表 1。

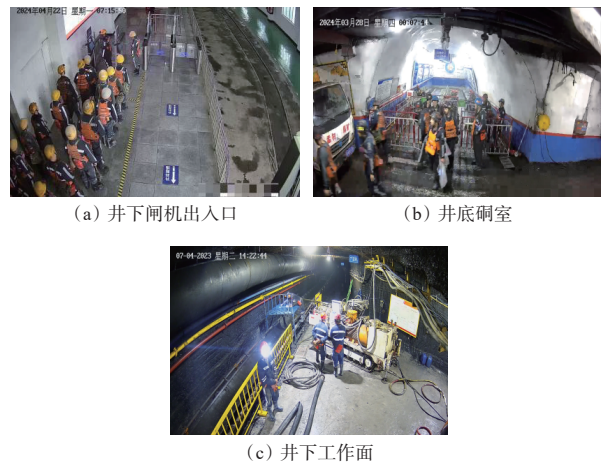


图 6 3 种场景的图像数据

Fig. 6 Image data of 3 scenarios

表 1 数据集详细情况与特征描述

Table 1 Dataset details and feature descriptions

场景	视频时长/s	分段数	抽取图像数/张	数据特点
井下闸机出入口	670	7	7 884	视频分辨率高;拍摄区域合适;遮挡情况多;入口处目标运动模糊程度较低
井底硐室	82	3	1 634	视频分辨率低;拍摄区域狭长;遮挡情况较少;入口处目标运动模糊程度较高
井下工作面	1 407	4	3 346	视频分辨率低;拍摄区域合适;遮挡情况很少;目标模糊程度较低;图像质量差,低光照和局部曝光情况严重

3.2 实验准备

实验在 64 位 Win11 系统下进行,使用 PyCharm 平台编程,深度学习框架选择 Pytorch1.1.1,并搭配 CUDA11.6 计算框架。图像处理相关任务使用 OpenCV-4.9.0.80 库。实验硬件环境:CPU 为 AMD Ryzen 7 7735H@.20 GHz,内存为 32 GiB,显卡为 NVIDIA

GeForce GTX 4060,硬盘为 1 TiB NVMe SSD。所有训练和测试均在该环境下完成。

为了更加客观地评估本文方法针对井下人员多目标的跟踪性能,分别采用不同评价指标对链接优化模块 COLinker 和井下人员跟踪性能进行分析。

在验证 COLinker 模块性能时, 选择 MOT20 数据集, 并使用精确率、召回率和 F_1 分数作为评价指标^[21]。精确率用于衡量模型预测的正样本中有多少是真正的正样本, 从而评估模型的准确性。召回率用于衡量所有正样本中有多少被模型正确预测, 反映模型的敏感性。 F_1 分数是精确率和召回率的调和平均数, 综合评价模型的性能。

在验证自适应链接优化的井下行人抗遮挡跟踪方法性能方面, 本文在 MPTD 数据集上选取 MOTA (Multiple Object Tracking Accuracy)、MOTP (Multiple Object Tracking Precision) 和 IDF1 (Identity F1 Score) 作为评价指标^[13]。MOTA 表示跟踪准确度, 反映了跟踪过程中出现目标误检、漏检及 ID 切换的情况。MOTP 表示跟踪精度, 反映了跟踪过程中目标预测框与数据集标注结果之间的匹配程度。IDF1 表示正确识别的检测数与平均真实数和计算检测数之比, IDF1 值越大, 说明对特定目标进行准确跟踪的稳定性越强。

3.3 实验结果与分析

为了验证自适应链接优化方法在煤矿井下行人遮挡环境下跟踪的有效性和可靠性, 分别进行 3 组井下人员检测和跟踪实验。首先采用 YOLOv9 针对井下多目标人员进行训练, 确保在进行跟踪前提供可靠的目标信息和检测输出; 然后采用 COLinker 模块进行训练; 最后将 COLinker 模块的输出添加到 DeepSORT 的匹配级联模块, 在 MPDD 数据集上进行多目标人员跟踪实验, 并与 OLinker 模块和 AFLink 模块进行对比。

3.3.1 COLinker 模块实验分析

为验证添加运动相似度和 CPCA 注意力机制对优化轨迹链接的有效性, 本文收集每个 epoch 的概率分布数据, 实验结果以小提琴图进行可视化展示^[18]。COLinker, OLinker 和 AFLink 模块的预测概率分布情况如图 7 所示, 可看出 COLinker 模块的预测概率分布较其他模型更为集中且精度更高, 表明优化后的模型在不同场景下具有更好的稳定性和精度。

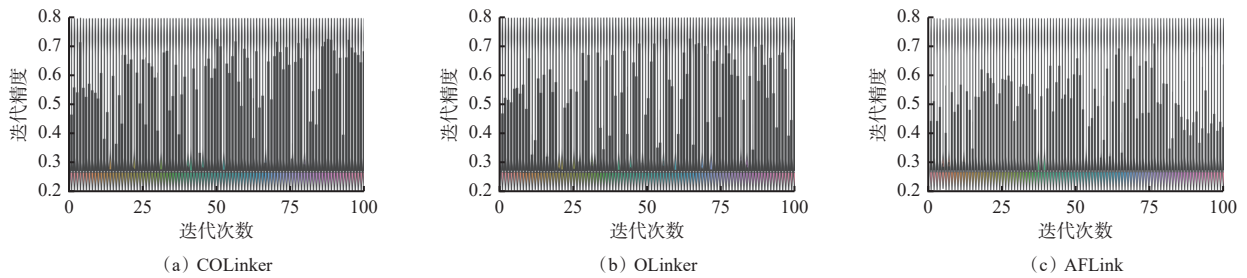


图 7 COLinker, OLinker 和 AFLink 模块的预测概率分布

Fig. 7 Predicted probability distribution chart of COLinker, OLinker, and AFLink modules

为进一步验证 COLinker 模块的有效性, 采用精确率、召回率和 F_1 分数来综合评估模型性能。COLinker, OLinker 和 AFLink 模块的综合评估指标如图 8 所示。可看出 COLinker 的精确率、召回率和

F_1 分数较 OLinker 和 AFLink 模型更高, 表明 COLinker 模块在处理验证数据集时具有更好的一致性和可靠性, 能够在更复杂的数据场景下维持较高的性能水平。

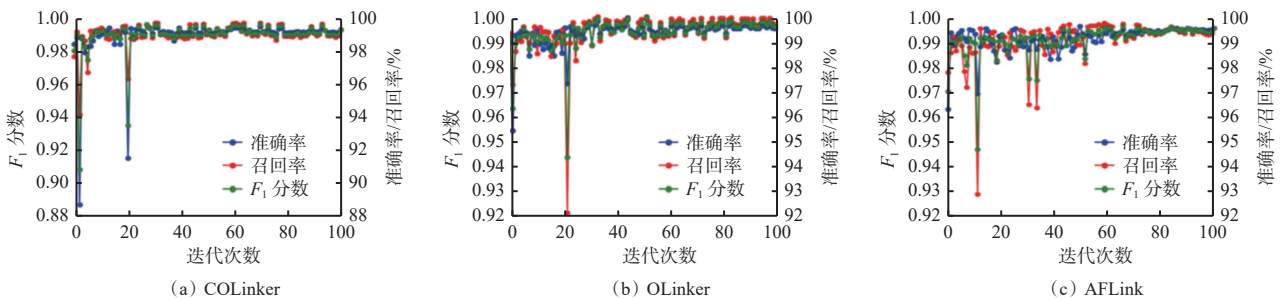


图 8 COLinker, OLinker 和 AFLink 模块的综合评估指标

Fig. 8 Comprehensive evaluation indicators of COLinker, OLinker, and AFLink modules

为直观观察每个模块的学习过程, 评估模型训练的效率和稳定性, 记录 COLinker, OLinker 和 AFLink 模块在训练过程中损失的变化情况, 结果如图 9 所

示。可看出 COLinker 模块的训练损失显著减少且验证损失趋于稳定, 说明其在捕捉轨迹之间相似性时具有较高的稳定性和泛化能力。

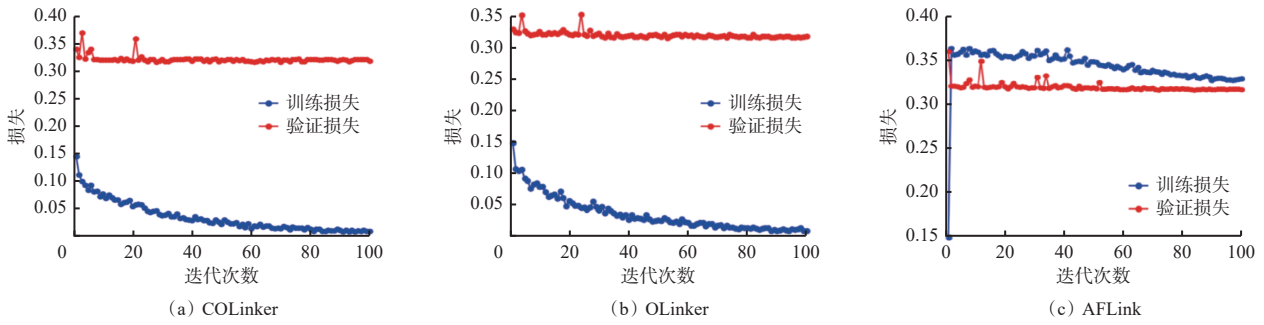


图 9 COLinker, OLinker 和 AFLink 模块的训练结果

Fig. 9 Training results of COLinker, OLinker, and AFLink modules

为进一步验证 COLinker 模块在遮挡条件下的性能,实验设计了不同遮挡程度的评估任务,选取一段总长为 2 760 帧且遮挡程度分布具有良好再现性的视频序列,视频序列中各遮挡程度的分布情况见表 2。

表 2 一段视频序列中各遮挡程度的分布情况

Table 2 Distribution of occlusion degrees in a video sequence

遮挡程度	无遮挡 (0)	轻度遮挡 (1%~10%)	部分遮挡 (10%~35%)	严重遮挡 (35%~80%)	完全遮挡 (≥80%)
帧数比/%	4.9	9.1	18.9	53.8	13.3
帧数	135	251	522	1 485	367

由表 2 可看出,严重遮挡(35%~80%)的帧数占比最高,为 53.8%,部分遮挡(10%~35%)的占比为 18.9%。这一分布反映了煤矿井下视频中遮挡的复杂性和多样性,尤其是部分遮挡到严重遮挡程度占绝大多数,能很好地用于评估本文方法在处理复杂遮挡场景下的鲁棒性和适应性。

不同遮挡程度下各模块的轨迹断裂恢复率对比如图 10 所示。可看出在不同遮挡程度下, COLinker 模块显著优于其他模块,尤其是在部分遮挡和严重遮挡条件下,其轨迹断裂恢复率较 OLinker 和 AFLink 模块分别高出 1.6% 和 2.3%,较卡尔曼滤波器提升 11.2%。表明 COLinker 模块在遮挡目标的轨迹匹配和轨迹精度上具备更强的鲁棒性和稳定性。

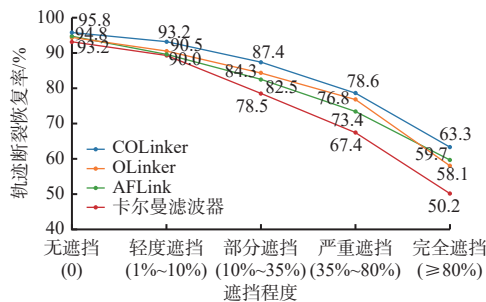


图 10 不同遮挡程度下各模块的轨迹断裂恢复率对比

Fig. 10 Comparison of trajectory fragmentation recovery rates of each module under different occlusion degrees

3.3.2 基于自适应 RB 因子的反馈机制实验分析

以每 25 帧为一时刻记录完整视频帧中的 ID 切换率 R_k 和跟踪断裂率 B_k ,经自适应 RB 因子调整后的匹配阈值变化率折线图如图 11 所示。可看出当视频序列时刻内 R_k 和 B_k 上升,即遮挡概率较大时,匹配阈值降低以适应环境变化;当时刻内 R_k 和 B_k 下降时,匹配阈值上升,使模块能更关注遮挡区域。

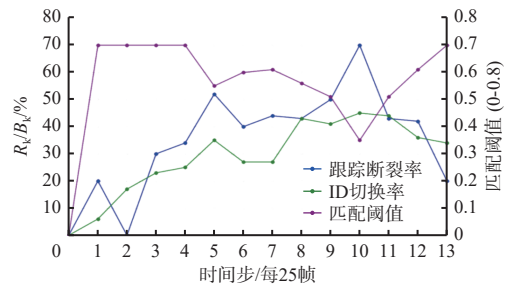


图 11 R_k , B_k 和匹配阈值变化率折线图

Fig. 11 Line chart of variation rates of R_k , B_k and matching threshold

3.3.3 自适应链接优化方法实验分析

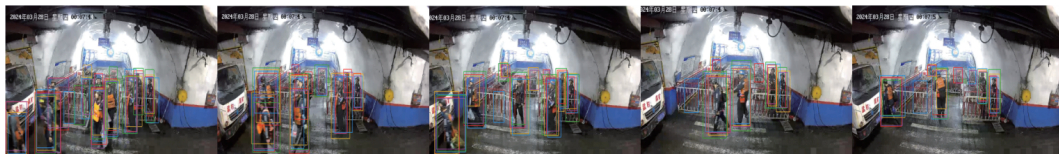
为了更加直观说明本文方法在煤矿井下行人目标遮挡下的跟踪效果,采用 DeepSORT, YOLOv7-SAM^[13], OSNet^[10], FuCoLoT^[22]和本文方法对煤矿井下视频序列进行定性实验。选取来自 MPDD 的 3 段视频序列(Video1 视频序列展示的是井口闸机区域行人监控视频; Video2 视频序列展示的是典型的井底硐室区域行人候车视频; Video3 视频序列展示的是井下作业面区域行人作业视频),实验结果如图 12 所示。可看出在 Video1 视频序列中,DeepSORT 和 FuCoLoT 在多目标遮挡情况下难以准确跟踪; YOLOv7-SAM 虽采用改进检测器方式能够更准确地检测目标,但在目标遮挡后 ID 发生变换,未能重新匹配; OSNet 优化了视频序列中多尺度的影响,但近距离下运动模糊带来的影响使得模型不能很好地跟踪目标;本文方法能够跟随目标运动,融合后的输入特征提升了目标的特征表达,并且 COLinker 模块能够更好地匹配切换后的 ID。在 Video2 视频序列

狭远处的目标并没有发生移动, YOLOv7-SAM 算法展现更好的检测能力, 但是在近景处, 本文方法能够更好地对快速移动产生的运动模糊情况下的目标进

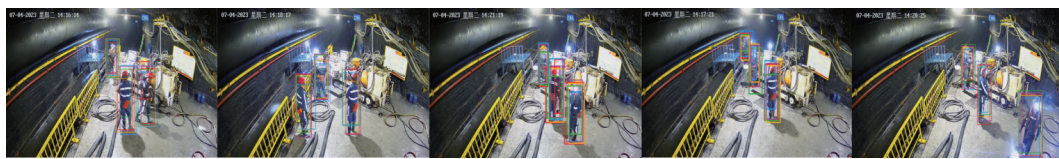
行跟踪。在 Video3 视频序列少目标情况下, 由于背景信息的干扰, 检测目标的置信度较低, 但本文方法能够自适应降低匹配阈值, 关注目标运动信息。



(a) Video1 视频序列跟踪结果



(b) Video2 视频序列跟踪结果



(c) Video3 视频序列跟踪结果

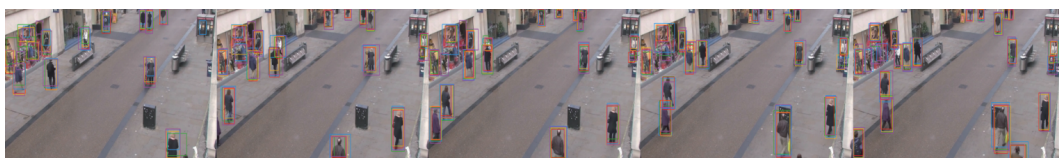
— 本文方法 — YOLOv7-SAM — OSNet — FuCoLoT — DeepSORT

图 12 5 种方法对 MPDD 数据集的目标跟踪结果

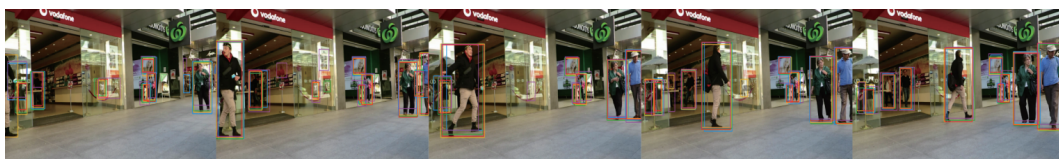
Fig. 12 Target tracking results of 5 methods on MPDD dataset

为了验证本文方法的泛化能力, 对 MOT17 数据集中的视频序列 MOT17-09-SDP 系列^[23]和 MOT15 数据集中的视频序列 AVG-TownCentre^[24]进行跟踪

实验, 结果如图 13 所示。可看出在对行人目标跟踪时, 虽然目标发生遮挡情况, 但是当目标再次出现时, 本文方法能更准确匹配跟踪目标。



(a) MOT17-09-SDP 视频序列跟踪结果



(b) AVG-TownCentre 视频序列跟踪结果

图 13 5 种方法对 MOT17 和 MOT15 数据集跟踪实验结果

Fig. 13 Tracking experiment results of 5 methods on MOT17 and MOT15 datasets

针对 COLinker 模块进行定量分析, 对比实验结果见表 3。可看出采用 COLinker 模块对行人目标进行轨迹关联与 ID 匹配时, 其精确率、召回率和 F_1 分数较 AFLink 都有很好的提升。

选取 MPDD 数据集中具有典型井下特征的 3 段视频序列、MOT17 数据集中的视频序列 MOT17-09-SDP 系列和 MOT15 数据集中的视频序列 AVG-TownCentre 进行测试, 并与 DeepSORT, FuCoLoT, YOLOv7-SAM, OSNet 进行对比, 结果见表 4、表 5。

由表 4 可看出, 在 Video1 中, 本文方法在 MOTA 上较 DeepSORT 提高了 15.6%, 本文方法在 MOTA 和 IDF1 上较 FuCoLoT 分别提高了 2.0% 和 1.8%, 较 YOLOv7-SAM 分别提高了 1.4% 和 2.7%, 较 OSNet 分别提高了 2.6% 和 1.8%。在 Video2 中, 本文方法在 MOTA 和 IDF1 上较 DeepSORT 分别提高了 16.1% 和 18.3%, 较 FuCoLoT 分别提高了 4.1% 和 1.4%, 较 YOLOv7-SAM 分别提高了 2.6% 和 2.3%, 较 OSNet 分别提高了 4.5% 和 1.6%。在 Video3 中, 本文方法

表 3 COLinker, OLinker 和 AFLink 模块的评估指标均值
Table 3 Mean values of evaluation indicators of COLinker, OLinker and AFLink modules

模型	精确率/%	召回率/%	F_1 分数
AFLink	92	89	0.90
OLinker	93	91	0.92
COLinker	96	94	0.95

表 4 3 段井下视频序列下 5 种方法指标对比

Table 4 Comparison of indicators of 5 methods in 3 underground video sequences

视频序列	方法	MOTA	MOTP	IDF1
Video1	DeepSORT	61.3	81.9	69.7
	FuCoLoT	74.9	82.1	74.1
	YOLOv7-SAM	75.5	82.3	73.2
	OSNet	74.3	82.3	74.1
	本文方法	76.9	82.4	75.9
Video2	DeepSORT	53.2	81.5	51.9
	FuCoLoT	65.2	84.7	68.8
	YOLOv7-SAM	66.7	84.5	67.9
	OSNet	64.8	84.3	68.6
	本文方法	69.3	85.3	70.2
Video3	DeepSORT	69.3	83.5	70.3
	FuCoLoT	80.9	83.7	76.6
	YOLOv7-SAM	81.6	84.4	76.5
	OSNet	80.9	83.5	78.2
	本文方法	82.3	84.7	78.6

表 5 公共视频序列下 5 种方法指标对比

Table 5 Comparison of indicators of 5 methods in public video sequence

视频序列	方法	MOTA	MOTP	IDF1
MOT17-09-SDP	DeepSORT	67.1	82.3	70.7
	FuCoLoT	76.4	82.8	74.5
	YOLOv7-SAM	78.6	83.2	75.9
	OSNet	76.4	82.9	7.7
	本文方法	78.9	83.5	75.6
AVG-TownCentre	DeepSORT	69.8	80.1	71.6
	FuCoLoT	78.5	83.1	74.7
	YOLOv7-SAM	79.5	84.3	76.8
	OSNet	80.2	83.9	74.8
	本文方法	81.1	84.8	76.4

在 MOTA 和 IDF1 上较 DeepSORT 分别提高了 1.0% 和 8.3%，较 FuCoLoT 分别提高了 1.4% 和 2.0%，较 YOLOv7-SAM 分别提高了 0.7% 和 2.1%，较 OSNet 分别提高了 1.4% 和 0.4%。本文方法的 MOTA, MOTP

和 IDF1 平均为 76.17%, 84.13%, 74.90%，较 DeepSORT 平均提升了 14.90%, 1.83% 和 10.93%，较 YOLOv7-SAM 平均提升了 1.57%, 0.40% 和 0.37%，较 OSNet 平均提升了 2.83%, 0.77% 和 1.27%，较 FuCoLoT 平均提升了 2.50%, 0.08% 和 1.80%。说明在 3 个具有井下典型特征的视频序列上本文方法的 MOTA, MOTP 和 IDF1 指标均表现最佳, 验证了其在不同场景下的鲁棒性和有效性。

由表 5 可看出, 在 MOT17-09-SDP 序列中, 本文方法的 MOTA 为 78.9%, 较 DeepSORT 提高了 11.8%, 较 YOLOv7-SAM 提高了 0.3%。在 MOTP 和 IDF1 指标上, 本文方法也分别达到了 83.5% 和 75.6%，较 DeepSORT 有显著提升。本文方法在 MOTA 上较 FuCoLoT 和 OSNet 分别提高了 2.5% 和 2.5%，在 MOTP 上分别提高了 0.7% 和 0.6%。在 AVG-TownCentre 序列中, 本文方法的 MOTA 为 81.1%，较 DeepSORT 提高了 11.3%，较 YOLOv7-SAM 提高了 1.6%。在 MOTP 和 IDF1 指标上, 本文方法分别达到了 84.8% 和 76.4%，较 DeepSORT 有显著提升。本文方法在 MOTA 上较 FuCoLoT 和 OSNet 分别提高了 2.6% 和 0.9%，在 MOTP 上分别提高了 1.7% 和 0.9%。结果表明, 本文方法不仅在特定场景下表现优异, 在公开数据集 MOT17 和 MOT15 上也能保持稳定的性能, 展现了良好的泛化能力。

4 结论

1) 针对煤矿井下环境中因遮挡频繁导致的轨迹匹配不准确问题, 提出了一种基于自适应链接优化的井下行人抗遮挡跟踪方法。在遮挡判定模块中, 通过结合目标置信度变化率和 IoU, 对目标的遮挡状态进行精准判定; 在 COLinker 模块中, 通过丰富轨迹对输入向量和添加 CPCA 注意力机制, 来扩展总成本函数, 有效缓解轨迹匹配过程中因遮挡导致的错误匹配问题; 在反馈机制中, 通过计算跟踪断裂率和 ID 切换率的变化量, 引入自适应 RB 因子, 以适应因长时间遮挡导致的轨迹断裂问题。

2) 在煤矿井下人员多目标跟踪数据集上, 对 5 种方法进行跟踪实验。结果表明所提方法能够更好地解决遮挡情况下的行人跟踪问题。

3) 所提方法在公开数据集 MOT17 和 MOT15 上也能保持稳定的性能, 表明该方法具有良好的泛化能力。

4) 虽然所提方法在井下行人遮挡跟踪中表现出优势, 但仍存在需要改进的方面: 一方面, 随着实际应用需求的增加, 可以考虑构建更大规模、更具挑战

性的煤矿井下行人遮挡数据集,以验证方法在更多复杂场景中的泛化能力;另一方面,探讨不同检测器输入对跟踪结果的影响,进一步优化检测器和跟踪模块之间的协同工作,从而进一步增强井下行人跟踪的可靠性和精度。

参考文献(References):

- [1] 张旭辉,杨文娟,薛旭升,等. 煤矿远程智能掘进面临的挑战与研究进展[J]. 煤炭学报, 2022, 47(1): 579-597.
ZHANG Xuhui, YANG Wenjuan, XUE Xusheng, et al. Challenges and developing of the intelligent remote control on roadheaders in coal mine[J]. Journal of China Coal Society, 2022, 47(1): 579-597.
- [2] 李翠芸, 衡博文, 谢金池. 模糊推理优化的抗遮挡PMBM跟踪算法[J]. 西安电子科技大学学报, 2023, 50(5): 54-64.
LI Cuiyun, HENG Bowen, XIE Jinchi. Anti-occlusion PMBM tracking algorithm optimized by fuzzy inference[J]. Journal of Xidian University, 2023, 50(5): 54-64.
- [3] LU Xiankai, MA Chao, NI Bingbing, et al. Adaptive region proposal with channel regularization for robust object tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 31(4) : 1268-1282.
- [4] 牟琦, 韩嘉嘉, 张寒, 等. 基于云边协同的煤矿井下尺度自适应目标跟踪方法[J]. 工矿自动化, 2023, 49(4): 50-61.
MU Qi, HAN Jiajia, ZHANG Han, et al. A scale-adaptive target tracking method for coal mine underground based on cloud-edge collaboration[J]. Journal of Mine Automation, 2023, 49(4): 50-61.
- [5] 冯磊, 李斌, 何勇. 基于注意力机制的煤矿工作面遮挡行人检测[J]. 价值工程, 2022, 41(14): 148-150.
FENG Lei, LI Bin, HE Yong. Pedestrian detection of coal mine working face occlusion based on attention mechanism[J]. Value Engineering, 2022, 41(14) : 148-150.
- [6] 程健, 陈亮, 王凯, 等. 一种多特征融合的复杂场景动态目标跟踪算法[J]. 中国矿业大学学报, 2021, 50(5): 1002-1010.
CHENG Jian, CHEN Liang, WANG Kai, et al. Multi-feature fusion dynamic target tracking algorithm for complex scenes[J]. Journal of China University of Mining & Technology, 2021, 50(5): 1002-1010.
- [7] COMANICIU D, MEER P. Mean shift analysis and applications[C]. 7th IEEE International Conference on Computer Vision, Kerkyra, 1999: 1197-1203.
- [8] 陈伟, 任鹏, 田子建, 等. 基于注意力机制的无监督矿井人员跟踪[J]. 煤炭学报, 2021, 46(增刊 1): 601-608.
CHEN Wei, REN Peng, TIAN Zijian, et al. Unsupervised mine personnel tracking based on attention mechanism[J]. Journal of China Coal Society, 2021, 46(S1): 601-608.
- [9] LI Xiaoyu, WANG Shuai, CHEN Wei, et al. An intelligent vision-based tracking method for underground human using infrared videos[J]. Symmetry, 2022, 14(8). DOI: 10.3390/SYM14081750.
- [10] 陈太光, 鲍新平, 汪涛, 等. 基于深度学习的矿井巷道人员计数技术[J]. 煤矿安全, 2023, 54(2): 234-238.
CHEN Taiguang, BAO Xinping, WANG Tao, et al. Mine roadway personnel counting technology based on deep learning[J]. Safety in Coal Mines, 2023, 54(2) : 234-238.
- [11] 张玉涛, 张梦凡, 史学强, 等. 基于深度学习的井下运动目标跟踪算法研究[J]. 煤炭工程, 2022, 54(10): 151-155.
ZHANG Yutao, ZHANG Mengfan, SHI Xueqiang, et al. Object tracking algorithm of moving objects in underground mine based on deep learning[J]. Coal Engineering, 2022, 54(10): 151-155.
- [12] 邵小强, 李鑫, 杨涛, 等. 改进YOLOv5s和DeepSORT的井下人员检测及跟踪算法[J]. 煤炭科学技术, 2023, 51(10): 291-301.
SHAO Xiaoqiang, LI Xin, YANG Tao, et al. Underground personnel detection and tracking based on improved YOLOv5s and DeepSORT[J]. Coal Science and Technology, 2023, 51(10): 291-301.
- [13] 范伟强, 王雪瑾, 张颖慧, 等. 改进YOLOv7和DeepSORT的井下人员检测与跟踪算法[J/OL]. 煤炭科学技术: 1-13[2024-10-07]. <http://kns.cnki.net/kcms/detail/11.2402.TD.20240119.1449.006.html>.
FAN Weiqiang, WANG Xuejin, ZHANG Yinghui, et al. Underground personnel detection and tracking using improved YOLOv7 and DeepSORT[J/OL]. Coal Science and Technology: 1-13[2024-10-07]. <http://kns.cnki.net/kcms/detail/11.2402.TD.20240119.1449.006.html>.
- [14] HUANG Hejun, CHEN Zuguo, ZOU Ying, et al. Channel prior convolutional attention for medical image segmentation[J]. Computers in Biology and Medicine, 2024, 178. DOI:10.1016/J.COMPBIOMED.2024.108784.
- [15] ZHOU Chunlun, YUAN Junsong. Multi-label learning of part detectors for heavily occluded pedestrian detection[C]. IEEE International Conference on Computer Vision, Venice, 2017: 3486-3495.

- Technology*, 2021, 35(3): 861-873.
- [14] 阙红波, 高扬, 吴成攀, 等. 基于扩展卡尔曼滤波和希尔伯特-黄变换瞬时频率的齿轮啮合刚度辨识算法[J]. *科学技术与工程*, 2021, 21(8): 3104-3111.
QUE Hongbo, GAO Yang, WU Chengpan, et al. Gear meshing stiffness identification algorithm based on extended Kalman filter and Hilbert-Huang transform instantaneous frequency[J]. *Science Technology and Engineering*, 2021, 21(8): 3104-3111.
- [15] MO Zhenling, ZHANG Heng, SHEN Yong, et al. Conditional empirical wavelet transform with modified ratio of cyclic content for bearing fault diagnosis[J]. *ISA Transactions*, 2023, 133: 597-611.
- [16] 祝文颖, 冯志鹏. 基于改进经验小波变换的行星齿轮箱故障诊断[J]. *仪器仪表学报*, 2016, 37(10): 2193-2201.
ZHU Wenyong, FENG Zhipeng. Fault diagnosis of planetary gearbox based on improved empirical wavelet transform[J]. *Chinese Journal of Scientific Instrument*, 2016, 37(10): 2193-2201.
- [17] 敬人可, 李建增, 周海林. 基于小波包变换和自适应滤波的超声信号去噪[J]. *中国测试*, 2014, 40(4): 115-118.
JING Renke, LI Jianzeng, ZHOU Hailin. Ultrasonic signal denoising based on wavelet packet transform and adaptive filtering[J]. *China Measurement & Test*, 2014, 40(4): 115-118.
- [18] 张西宁, 李霖, 刘书语, 等. 基于能量峰定位的经验小波变换及在轴承微弱故障诊断中的应用[J]. *西安交通大学学报*, 2021, 55(8): 1-8.
ZHANG Xining, LI Lin, LIU Shuyu, et al. Empirical wavelet transform based on energy peak location with applications to bearing weak fault diagnosis[J]. *Journal of Xi'an Jiaotong University*, 2021, 55(8): 1-8.
- [19] 郭燕飞, 陈高华, 王清华. 基于广义变分模式分解的滚动轴承故障微弱特征提取[J]. *机械传动*, 2023, 47(5): 150-157.
GUO Yanfei, CHEN Gaohua, WANG Qinghua. Weak feature extraction of rolling bearing fault based on generalized variational mode decomposition[J]. *Journal of Mechanical Transmission*, 2023, 47(5): 150-157.
- [20] 郭燕飞, 王清华, 陈高华. 广义变分模式分解降噪在齿轮早期故障诊断中的应用[J]. *科学技术与工程*, 2022, 22(23): 10065-10072.
GUO Yanfei, WANG Qinghua, CHEN Gaohua. Application of generalized variational mode decomposition de-noising in early gear fault diagnosis[J]. *Science Technology and Engineering*, 2022, 22(23): 10065-10072.
- [21] 何正嘉, 袁静, 瞿艳阳, 等. 机械故障诊断内积变换原理与验证[J]. *振动. 测试与诊断*, 2012, 32(2): 175-185.
HE Zhengjia, YUAN Jing, ZI Yanyang, et al. Inner product principle of mechanical fault diagnosis and validation[J]. *Journal of Vibration, Measurement & Diagnosis*, 2012, 32(2): 175-185.
- 综述[J]. *中国图象图形学报*, 2019, 24(12): 2057-2080.
LI Xi, ZHA Yufei, ZHANG Tianzhu, et al. Survey of visual object tracking algorithms based on deep learning[J]. *Journal of Image and Graphics*, 2019, 24(12): 2057-2080.
- [21] DU Yunhao, ZHAO Zhicheng, SONG Yang, et al. StrongSORT: make DeepSORT great again[J]. *IEEE Transactions on Multimedia*, 2023, 25: 8725-8737.
- [22] LUKEŽIČ A, ZAJC L Č, VOJÍŘ T, et al. FuCoLoT-A fully-correlational long-term tracker[C]. *Asian Conference on Computer Vision*, Cham, 2019: 595-611.
- [23] NEWBURGER E, CORRELL M, ELMQVIST N. Fitting bell curves to data distributions using visualization[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2023, 29(12): 5372-5383.
- [24] KEUPER M, TANG Siyu, ANDRES B, et al. Motion segmentation & multiple object tracking by correlation co-clustering[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(1): 140-153.

(上接第 75 页)

- [16] JAHANTIGHY A, TORABI H, MOHANNA F. Multiple targets video tracking based on extended Kalman filter in combination with particle swarm optimization for intelligent applications[J]. *SN Applied Sciences*, 2023, 5(3). DOI: 10.1007/S42452-023-05297-9.
- [17] 李凡长, 刘洋, 吴鹏翔, 等. 元学习研究综述[J]. *计算机学报*, 2021, 44(2): 422-446.
LI Fanzhang, LIU Yang, WU Pengxiang, et al. A survey on recent advances in meta-learning[J]. *Chinese Journal of Computers*, 2021, 44(2): 422-446.
- [18] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]. *IEEE International Conference on Image Processing*, Beijing, 2017: 3645-3649.
- [19] WANG C Y, YEH I H, MARK LIAO H Y. YOLOv9: learning what you want to Learn using programmable gradient information[C]. *European Conference on Computer Vision*, Cham, 2024: 1-21.
- [20] 李玺, 查宇飞, 张天柱, 等. 深度学习的目标跟踪算法