

基于改进ByteTrack算法的红外地面多目标跟踪方法

王雒, 李飏, 傅瑞罡

引用本文

王雒, 李飏, 傅瑞罡. 基于改进ByteTrack算法的红外地面多目标跟踪方法[J]. 计算机科学, 2023, 50(9): 176-183.

WANG Luo, LI Biao, FU Ruigang. [Infrared Ground Multi-object Tracking Method Based on Improved ByteTrack Algorithm](#) [J]. Computer Science, 2023, 50(9): 176-183.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于多模态特征融合的时间序列异常检测](#)

Anomaly Detection of Time-series Based on Multi-modal Feature Fusion

计算机科学, 2023, 50(6A): 220700094-7. <https://doi.org/10.11896/jsjcx.220700094>

[联合人体姿态估计和多目标跟踪的跨数据集学习](#)

Cross-dataset Learning Combining Multi-object Tracking and Human Pose Estimation

计算机科学, 2023, 50(6A): 220400199-7. <https://doi.org/10.11896/jsjcx.220400199>

[基于多级多尺度特征提取的CNN-BiLSTM模型的中文情感分析](#)

Chinese Sentiment Analysis Based on CNN-BiLSTM Model of Multi-level and Multi-scale Feature Extraction

计算机科学, 2023, 50(5): 248-254. <https://doi.org/10.11896/jsjcx.220400069>

[基于深度学习的视觉多目标跟踪研究综述](#)

Deep Learning-based Visual Multiple Object Tracking:A Review

计算机科学, 2023, 50(4): 77-87. <https://doi.org/10.11896/jsjcx.220300173>

[基于互相关注意力的链式帧处理多目标跟踪算法](#)

Multi-object Tracking Based on Cross-correlation Attention and Chained Frames

计算机科学, 2023, 50(1): 131-137. <https://doi.org/10.11896/jsjcx.211100097>

基于改进 ByteTrack 算法的红外地面多目标跟踪方法

王 雒 李 彪 傅瑞罡

国防科技大学电子科学学院 长沙 410000

(sagdsgjksad123@qq.com)

摘要 红外目标智能检测跟踪技术研究一直是同领域中的热点问题,尤其是在精确制导、海面监视和天空预警等方面。针对红外地面多目标跟踪场景中,由地面杂波干扰、多目标遮挡干扰、平台晃动等复杂场景造成的跟踪精度降低等问题,提出了一种基于改进 ByteTrack 算法的红外地面多目标跟踪方法。首先引用一种自适应调制噪声尺度的卡尔曼滤波器,缓解低质量检测对 vanilla 卡尔曼滤波器的影响;其次引入增强相关系数最大化算法对帧间图像进行配准,来补偿平台晃动产生的影响;然后增加了基于长短期记忆网络的运动模型,减小了卡尔曼滤波在非线性运动状态中产生的预测误差;最后引入连接模型和高斯平滑算法这两种轻量级离线算法来完善跟踪结果。在红外地面多目标数据集上进行了实验,结果表明,与 Sort 和 Deepsort 算法相比,改进算法的 MOTA 值分别提升了 8.3% 和 10.2%,IDF1 值分别提升了 6.5% 和 5.6%。与同类算法相比,改进算法表现出了更好的有效性,在红外目标智能检测跟踪场景中会有较大应用。

关键词: 多目标跟踪;红外目标;ByteTrack;卡尔曼滤波;长短期记忆网络

中图法分类号 TP391.41

Infrared Ground Multi-object Tracking Method Based on Improved ByteTrack Algorithm

WANG Luo, LI Biao and FU Ruigang

College of Electronic Science and Technology, National University of Defense Technology, Changsha 410000, China

Abstract The research of infrared object intelligent detection and tracking technology is always a hot topic in the same field, especially in precision guidance, sea surface surveillance and sky warning. Aiming at the problems that tracking accuracy is reduced due to ground miscellaneous interference, multi-object block interference, platform shaking and other complex scenes, an infrared ground multi-object tracking method based on improved ByteTrack algorithm is proposed. First of all, a modified Kalman filter which could adaptively modulate the noise scale is introduced to alleviate the impact of low-quality detection on vanilla Kalman filter. Secondly, the introduction of enhanced correlation coefficient maximization is used to settle the inter-frame images to compensate the platform shaking impact. Then ByteTrack increases the motion model based on long short-term memory network to solve the prediction error caused by Kalman filter in the non-linear motion state. Finally, two lightweight offline algorithms of link model and Gaussian-smoothed interpolation are introduced to refine the tracking results. Experiment is performed on the infrared ground multi-object dataset and the results show that compared with Sort and DeepSort, the MOTA of the improved algorithm increases by 8.3% and 10.2%, IDF1 increases by 6.5% and 5.6%, respectively. The improved algorithm shows better effectiveness and will be used in the infrared object intelligent detection and tracking scenes.

Keywords Multi-object tracking, Infrared object, ByteTrack, Kalman filter, Long short-term memory network

1 引言

红外精确制导技术是当今世界各军事强国争相发展的重要技术,红外热成像系统已经被广泛应用,用于对目标进行捕捉,然后处理目标的红外图像信息以获得目标的坐标信息,从而控制导引头对目标进行跟踪锁定。同时,具有全天候工作、被动预测、隐蔽性好和抗电子干扰能力强等特点的机载红外

预警系统已经被广泛应用于陆海空预警方面,用于实现对地面、海面 and 空中目标的搜索跟踪,使其能够为近程武器防御系统提供目标信息。近年来,远距离的红外小目标检测与跟踪技术在战斗机火控制导系统和防空预警系统中的作用越来越大,在国防建设和现代化武器装备改进升级中扮演着十分重要的角色。

随着经济和科技的飞速发展,红外热成像系统的体积和

到稿日期:2022-09-01 返修日期:2023-04-14

基金项目:国家自然科学基金(62001482);湖南省自然科学基金(2021JJ40676)

This work was supported by the National Natural Science Foundation of China(62001482) and Natural Science Foundation of Hunan Province, China(2021JJ40676).

通信作者:傅瑞罡(furuigang08@nudt.edu.cn)

成本都有所下降,逐渐被民用化。无人汽车驾驶、无人飞机巡检和红外全景监测等多个民用领域也开始广泛采用红外目标检测与跟踪技术。无人驾驶车辆在行车途中,除了捕捉静态信息之外,同时需要联合其他传感器,如红外传感器,根据目标的历史轨迹进行合理的轨迹预测,用于制定避障策略^[1]。红外热成像仪搭载在多旋翼无人机上被用于跟踪输电线路。采用红外热成像仪对线路特征较清晰的红外图像进行分析处理后,调整控制信号,以保证持续跟踪输电线路目标。油田监控预警也采用了全景红外成像,来达到准确可靠、搜跟一体、全天候和全方位的安防要求。

因此,红外目标智能检测跟踪技术的发展具有非常重要的理论意义和应用价值,有助于丰富军事侦察手段,提高军事侦察能力,具有较为广阔的应用前景。

红外目标的检测跟踪技术中,已有传统算法很难适用于所有情况,检测跟踪能力提升遭遇瓶颈。相比可见光检测跟踪,红外目标检测跟踪的检测跟踪距离、抗干扰能力以及工作范围和时长都具有可比性,因此将可见光检测跟踪的一些深度学习方法进行借鉴和发展,并将其运用到红外地面目标检测跟踪技术上。应用深度学习理论优化红外地面目标智能检测跟踪技术是一个值得研究的方向,也是一项十分紧迫的任务,这些问题对于红外目标智能检测跟踪技术在未来军事以及民用设备上的发展都至关重要。

针对红外地面多目标跟踪场景^[2]中存在的由地面杂波干扰、多目标遮挡干扰、平台晃动等复杂场景所造成的跟踪精度降低等问题,本文提出了一种基于改进 ByteTrack 算法^[3]的红外地面多目标跟踪方法。本文的主要工作如下:首先,引用 NSA(Noise Scale Adaptively)卡尔曼滤波器^[4],采用自适应计算噪声协方差来缓解低质量检测对 vanilla 卡尔曼滤波器的影响;其次,引入增强相关系数最大化算法 ECC^[5](Enhanced Correlation Coefficient)对帧间图像进行配准,来补偿平台晃动产生的影响;然后,增加了基于 LSTM(Long Short-term Memory Network)的运动模型,利用长短期记忆网络 LSTM 进行轨迹预测,减小了卡尔曼滤波在非线性运动状态中产生的预测误差;最后,引入两个轻量级、即插即用和与外观无关的算法来完善跟踪结果,第一个是与外观无关,只利用时空信息来预测两个轨迹是否属于同一 ID 的连接模型 AFLink(Appearance-free Link),第二个是用于弥补漏检情况的高斯平滑插值算法 GSI^[6](Gaussian-smoothed Interpolation),两种离线算法进一步提高了跟踪过程的准确性和鲁棒性。

2 相关工作

随着人工智能的不断发展与计算机硬件性能的逐渐提升,以卷积神经网络为基础的深度学习在计算机视觉的各个领域(如目标检测和多目标跟踪等)取得了相当大的成功。多目标跟踪在计算机视觉领域中扮演着重要角色,其任务是在给定视频中同时对多个感兴趣目标进行检测定位,并且保持目标 ID 稳定不变,最后持续跟踪目标并记录轨迹,尤其是针对车辆的跟踪,应用非常广泛,在视频监控、智能交通和军事制导等领域具有十分重要的作用。

深度学习蓬勃发展之前,目标跟踪领域大部分属于单目标跟踪,在多目标跟踪方面的研究不多。传统的多目标跟踪算法主要有联合概率数据关联滤波、多假设跟踪算法、条件随机场等方法。联合概率数据关联滤波 JPDA^[7]考虑观测值与每个目标之间的关联概率,使其在每步中进行最佳赋值。JPDA 假设所有的检测都有可能来自每个特定目标,只是来自不同目标的概率存在差异。JPDA 不依赖于跟踪目标的先验信息,但存在较多数量目标时会出现组合爆炸问题。多假设跟踪算法 MHT^[8]保留了对真实目标的所有假设,基于延迟逻辑,在连续几步中考虑多种可能的关联,基于后续接收的结果来处理当前扫描周期的相对关联歧义。MHT 是解决数据关联的最优算法,能检测出目标的出现和消失,需要依赖目标和杂波的先验信息。但由于其复杂性,应用起来通常受到限制。RMOT^[9]多目标跟踪算法通过对目标的上下文运动建模来描述其相对运动,以排除相机运动的影响,同时与贝叶斯滤波器、匈牙利匹配结合。

近年来,基于卷积神经网络的目标检测器的准确性和效率不断提高,可见光多目标跟踪取得了重大进展。Bewley 等^[10]提出的简单在线实时跟踪算法 SORT,是最早利用卷积神经网络检测目标的 MOT 算法之一,该算法重点关注帧间的预测和关联,通过线性速度模型卡尔曼滤波器预测物体运动,然后度量检测框和跟踪框的交并比,最后利用匈牙利算法将检测结果关联在一起;SORT 算法并没有过多地关注长时间跟踪期间遮挡造成的目标 ID 频繁变换问题。随后,Wojke 等^[11]提出了 DeepSort 算法,DeepSort 在 SORT 算法的基础上进行改进,在数据关联部分引入了重识别特征来提取深度表观特征,使得部分被遮挡的物体能够被重新识别,同时提出了级联匹配策略,提高了目标匹配的准确度;Chen 等^[12]提出了 MOTDT 算法,设计了一种新的多目标跟踪框架,融合了目标的深度外观信息与空间信息,并提出了一种分层数据关联策略,将预测目标与多个检测目标层级关联,根据软分配策略的轨迹评分机制来实现目标的关联,充分利用重识别特征和空间信息提升跟踪性能。Wang 等^[13]提出了 JDE 算法,从模型推理运行速度方面考虑,融合一阶段目标检测与行人重识别网络,使得网络模型同时输出检测信息和重识别信息,加快了推理运行的速度。Zhang 等^[14]提出了 FairMOT 算法,检测分支采用无锚框的目标检测算法 CenterNet,与此同时增加一个平行分支直接输出 ReID 特征用来区分不同目标,很好地将目标检测和重识别网络统一起来。Zhou 等^[15]提出了 CenterTrack 算法,实现了对检测和目标深度特征的同时提取,通过对所有目标进行并行推理来及时恢复更新因遮挡而造成的目标漏检,提高了算法对重叠遮挡问题的鲁棒性。

近年来,一些深度学习方法已经被运用到红外多目标跟踪算法中。Wang 等^[16]将可见光目标检测 YOLOv3 引入红外小目标检测中,使用 YOLOv3-tiny 网络进行检测,利用卡尔曼滤波和匈牙利算法进行多目标跟踪,虽然取得了不错的效果,但是目标 ID 发生了较多的切换。Song 等^[17]提出了一种改进的红外多类别多目标跟踪网络,它结合 RepVGG 骨干网络设计了快速金字塔结构和无锚框网络结构,进一步减少了网络参数量并缩短了推理时间,但是跟踪精度已经不能

满足现在的要求。Yang 等^[18]提出了一种基于分层数据关联的空中红外多目标跟踪方法,它构建了表观、运动、尺度 3 种特征相似性,数据关联效果不错,但是对目标的检测精度还不够好。Xie 等^[19]提出了一种融合轻量级 YOLOv4 与 KCF 算法的红外舰船目标识别方法,适合复杂背景条件下的舰船红外目标识别,实现了对舰船目标稳定持续的跟踪,具有较强的鲁棒性和实时性,但是跟踪精度依然还有提升的空间。Yang 等^[20]提出了一种基于深度学习的防空武器红外目标识别流程,目标检测阶段采用 YOLO 网络模型对全图多目标进行识别定位,目标跟踪阶段采用超分辨率重建算法提升目标局部图像分辨率,最后采用深度残差网络模型对跟踪目标进行识别分类,但数据关联效果还不够好。

基于上述问题,本文将多目标跟踪算法 ByteTrack 与 NSA 卡尔曼滤波算法、增强相关系数最大化算法 ECC、长短期记忆网络 LSTM、连接模型 AFlink 和高斯平滑插值算法 GSI 相结合,实验结果表明,改进算法表现出了更好的有效性,多目标跟踪精度得到了提升。

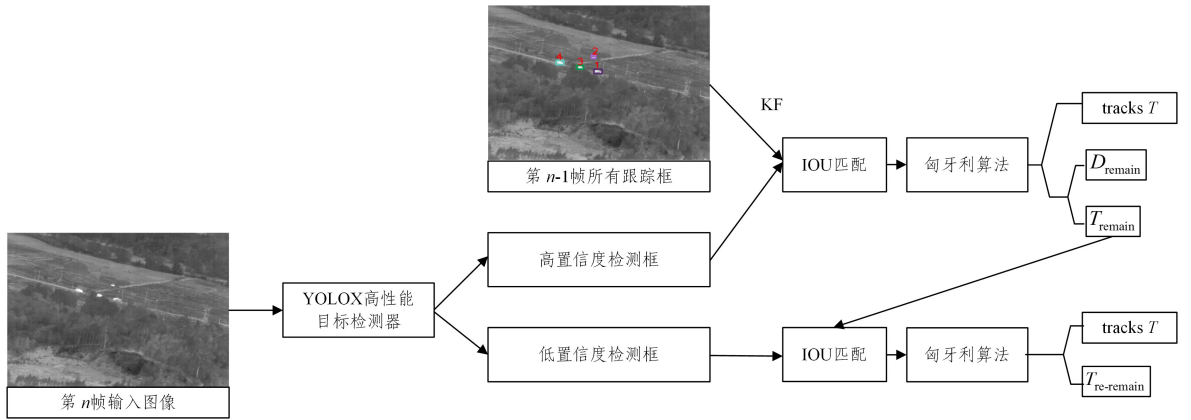


图 1 ByteTrack 算法的流程图

Fig. 1 Flow diagram of ByteTrack algorithm

3.1 高性能目标检测器 YOLOX

YOLOX 将 YOLO 系列检测器切换到无锚模式,并采用一些先进的检测技术,包括解耦头、强力数据增强(如 Mosaic 和 Mixup)以及有效的标签分配策略 SimOTA,来实现最先进的目标检测性能。骨干网络和 YOLOv3 相同,采用 CSP-Darknet 以及附加的 PAN 头。在骨干网络之后有 3 个解耦头:(1)用于分类;(2)用于回归;(3)在回归头中加入了一个额外的 IOU 注意分支。

3.2 关联方法 BYTE

BYTE 的输入有视频片段 V 、检测器 Det 和卡尔曼滤波器 KF ,同时设定 3 个阈值 τ_{high} , τ_{low} , ϵ 。前两个为检测分数阈值,后一个为跟踪分数阈值。BYTE 的输出为视频的轨迹 T ,每个轨迹包含目标的检测框和身份。

对于视频中的每一帧图像,使用检测器 Det 预测其检测框和分数。对于得分高于阈值 τ_{high} 的检测框,将它们放入高分检测框 D_{high} 中;对于得分高于阈值 τ_{low} 且低于阈值 τ_{high} 的检测框,将其放入低分检测框 D_{low} 中。将低分检测框和高分检测框分离后,对轨迹集合 T 中的所有轨迹采用卡尔曼滤波器 KF 预测,预测出 T 中每个轨迹在当前帧中的新位置。

3 ByteTrack 的基本框架

多目标跟踪的大多数方法采用检测分数高于某一设定阈值的边界框来进行关联,并且同时丢弃由遮挡和运动模糊等问题产生的低分数检测框,从而造成了不可忽略的目标丢失和轨迹碎片化。检测决定了跟踪任务的上界,而关联部分则决定是否能够更好地达到检测决定的上界,因此数据关联是多目标跟踪任务的核心。为了解决这个问题,ByteTrack 考虑关联每个检测框而不仅仅只关联高分检测框,利用多次匹配来提高跟踪精度。优先匹配高分检测框和轨迹后,再将低分检测框和剩余的轨迹进行匹配,利用它们与之前的轨迹进行相似度对比,来恢复真实目标和过滤背景。ByteTrack 算法基于高性能目标检测器 YOLOX^[21] 和关联方法 BYTE。ByteTrack 是“按检测跟踪”范式的跟踪方法,没有使用 ReID 特征计算外观相似度,它仅仅使用了卡尔曼滤波预测边界框,然后使用匈牙利算法匹配目标和轨迹,完整的计算流程如图 1 所示。

首先对高分检测框 D_{high} 和所有轨迹 T (包括丢失的轨迹 T_{lost}) 进行第一次关联,相似度由检测网络预测的检测框与卡尔曼滤波预测的跟踪框之间的 IOU 计算获得,然后采用匈牙利算法进行关联匹配。拒绝匹配相似度小于 0.2 的检测框,将未匹配的检测框存放于 D_{remain} 中;将未匹配的轨迹存放于 T_{remain} 中,第一次关联过程大致如图 2 所示。

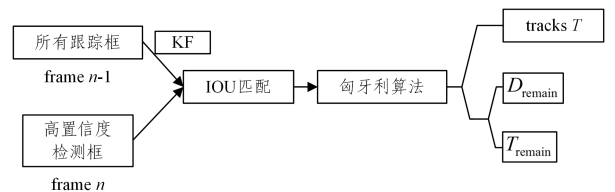


图 2 第一次关联

Fig. 2 First association

在第一次关联之后,在低分检测框 D_{low} 和剩余轨迹 T_{remain} 之间进行第二次关联,单独使用 IOU 作为相似度,因为低分检测框通常包含严重的运动模糊或遮挡,并且此时外观特征不可靠。将未匹配的轨迹存放于 $T_{re-remain}$ 中,并删除所有未匹配的低分检测框,第二次关联过程大致如图 3 所示。

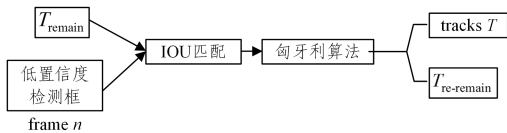


图3 第二次关联

Fig. 3 Second association

对于第二次关联后未匹配的轨迹 $T_{re-remain}$, 认为是暂时丢失了目标, 将它们放入 T_{lost} 中。对于 T_{lost} 中的每个轨迹, 只有当它超过一定数量(通常设定为 30 帧)的帧时, 才会将其从轨迹 T 中删除; 否则, 保留 T_{lost} 在轨迹 T 中。

最后, 对于在第一次关联后从未匹配过的高分检测框 D_{remain} , 如果检测得分高于 ϵ 且存活超过两帧, 则初始化为新的轨迹。每一帧的输出是当前帧中轨迹 T 的边界框和身份, 但并不输出 T_{lost} 的边界框和身份。实验中, τ_{high} 设置为 0.4, τ_{low} 设置为 0.1, ϵ 设置为 0.5, 保留下由遮挡和运动模糊等问题产生的低分数检测框, 将低分数检测框和轨迹进行匹配, 使得算法能够达到更好的跟踪效果。

4 基于 ByteTrack 的改进

4.1 NSA 卡尔曼滤波算法

vanilla 卡尔曼滤波虽然得到了广泛的运用, 但是不稳定, 它为所有目标简单地设置了一个统一的测量噪声尺度, 忽略了检测噪声尺度上的影响; 同时也没有考虑到每个目标的检测质量, 使其很容易受到低质量检测的影响。为了获得更精确的运动状态, 引用 NSA 卡尔曼滤波, 即去除噪声的卡尔曼滤波, 采用的自适应计算噪声协方差 \tilde{R}_k 的计算式为:

$$\tilde{R}_k = (1 - c_k) R_k \quad (1)$$

其中, R_k 是预先设定的常数测量噪声协方差, c_k 是状态 k 下的检测置信度。

4.2 增强相关系数最大化算法 ECC

多目标跟踪场景中通常存在着比较复杂的运动模式, 可以概括为目标的刚性和非刚性运动, 同时还包含了相机的刚性运动, 例如 MOT Challenge 数据集, 就存在着大量相机运动的场景, 有些场景甚至超过了静态相机场景参数。同时, “按检测跟踪”范式的跟踪方法也严重依赖于轨迹预测的跟踪框和检测边界框之间的 IOU。综上, 相机运动对于多目标跟踪的影响不小, 尤其是一些仅依赖于运动特征信息的算法, 相机的运动会严重干扰到运动模型。因为缺乏关于相机运动的额外参数和相机内部参数, 所以两个相邻帧之间的图像配准就是相机刚性运动在图像平面上良好的投影近似。

增强相关系数最大化算法 ECC 首先采用全局运动估计技术(GMC)来表示背景运动, 然后对图像平面上的关键点进行提取, 最后采用稀疏光流的方式, 进行基于平移的局部异常点抑制的特征跟踪。使用 RANSAC 计算获得两个相邻帧之间的仿射变换矩阵, 然后使用仿射变换矩阵将预测边界框从 $k-1$ 帧的坐标系变换到下一帧 k 的坐标。实验中, ECC 算法选择采用运动欧几里得的翘曲模式。应用 ECC 这种方法,

跟踪器对相机的运动具有较好的鲁棒性。

4.3 基于 LSTM 的运动模型

4.3.1 长短期记忆网络(LSTM)

ByteTrack 算法中使用的是基于卡尔曼滤波的运动模型, 是预测目标轨迹的一种线性运动模型, 它通过输入目标的历史运动特征进行预测, 从而进行目标匹配。但是在运动过程中出现多目标互相干扰和障碍物遮挡时, 运动状态变为非线性, 使用卡尔曼滤波并不能进行准确的轨迹预测。长短期记忆网络(LSTM)具有记忆功能, 拥有强大的序列建模能力来解决顺序和时间相关约束, 尤其是在行人轨迹预测、船舶轨迹预测和自动驾驶轨迹预测等复杂背景下的预测中展现了强大的优势, 因此 ByteTrack 算法增加了基于 LSTM 的运动模型, 来提取物体的运动特征。当出现同一目标因为遮挡或者仅仅依靠外观特征无法正确区分目标而导致目标 ID 频繁变换时, 运动特征可以帮助多目标跟踪算法正确分配目标轨迹, 实现逐帧的目标检测跟踪, LSTM 单元结构如图 4 所示。

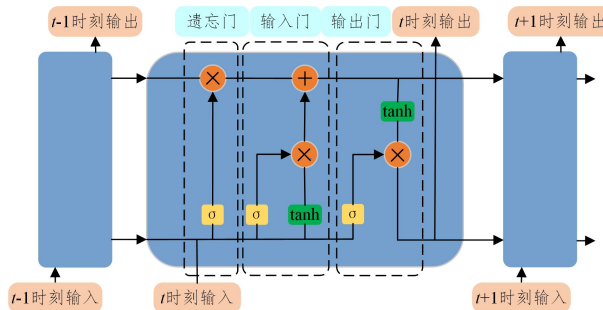


图4 LSTM 单元结构图

Fig. 4 Structure diagram of LSTM unit

4.3.2 基于 LSTM 的轨迹预测

由于地面目标的运动模式比较简单, 向循环神经网络 LSTM 中输入目标的数帧历史运动特征, 预测网络就能够准确地预测出目标未来的运动特征。目标轨迹预测网络的训练数据是红外地面多目标跟踪数据集的标注文件。经过实验后得出, 向循环神经网络 LSTM 中输入的运动特征有目标框的中心坐标 (c_x, c_y) 、宽度 w 和高度 h ; 在输入层将 $t-6$ 到 t 时刻目标的运动特征作为输入, 预测 $t+1$ 时刻目标的运动特征时, 均方差损失函数达到最小, 也就是当目标被连续跟踪 6 帧时, 使用基于 LSTM 的运动模型预测下一帧的运动特征准确率达到最高, 即经过实验发现, 当输入维度为 4, 时间步长为 6, 且循环神经网络 LSTM 的超参数设置如表 1 所列时, 均方差损失函数最小, 在非线性运动状态中产生的预测误差最小。

表1 LSTM 网络参数

Table 1 LSTM network parameters

参数	设置
迭代次数	500
神经元个数	32
隐藏层层数	1
学习率	0.005
Batchsize	30
优化算法	Adam
损失函数	均方误差

4.4 无外观连接模型 AFLink

为了将全局信息更好地利用起来,一些方法通常采用全局连接模型将短轨迹碎片关联到轨迹上,使用离线的方式将准确但不完整的轨迹与全局信息关联。虽然这些方法更好地完善了跟踪结果,但它们往往都依赖于密集型计算模型,尤其是采用外观嵌入的方法,耗时较长,实时性较差。因此,引入了一种只利用时空信息来预测两个输入轨迹是否属于同一ID的AFLink模型。

AFLink模型采用两个轨迹 T_i 和 T_j 作为网络模型的输入,其中 $T_* = \{f_k, x_k, y_k\}_{k=1}^N$ 由最近 $N=30$ 帧的帧 f_k 和位置 (x_k, y_k) 组成,零填充用于短于30帧的图像。特征提取采用时间模块,时间模块由4个卷积层组成,具有 7×1 个内核和 $\{32, 64, 128, 256\}$ 个输出通道;特征提取沿时间维度与 7×1 个卷积核进行卷积,每个卷积之后是BN层和RELU激活层。融合模块包括 1×3 卷积、BN层和RELU,它不会改变通道的数量;融合模块随后对来自不同特征维数 f, x 和 y 的信息进行 1×3 卷积,把得到的两个特征映射先后进行合并和压缩操作,从而获得包含丰富时空信息的特征向量,接着进行拼接操作。最后使用具有两个全连接层和插入其间的RELU层的MLP来判断关联的置信度分数。在关联过程中,过滤具有时空约束的不合理轨迹对,将全局连接理解为具有预测连通性分数的线性分配任务。图5给出了AFLink模型的框架。

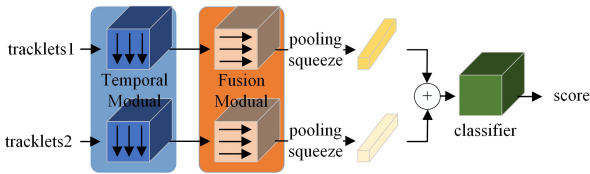


图5 AFLink模型的结构图

Fig. 5 Framework of AFLink model

将带注释的轨迹切割成具有随机时空噪声的轨迹,并且通过1:3的正负样本比例来生成AFLink模型的训练数据。在实验中,使用Adam作为优化算法,采用交叉熵作为损失函数,同时用余弦退火学习率调度进行20次迭代训练,时间距离阈值设置为30帧,空间距离阈值设置为75像素,用于过滤不合理的关联对。如果关联对的预测得分大于0.95,则考虑输入轨迹的关联性。

4.5 高斯平滑插值算法 GSI

线性插值被广泛用于弥补漏检情况,尽管线性插值简单有效,但它忽略了运动信息,限制了插值位置的精度。因此,高斯平滑插值算法GSI在高斯过程回归算法(GPR)的基础上对非线性运动进行了建模,没有增加额外的耗时组件,引入的GSI算法在精度和效率之间取得了很好的折衷。

高斯平滑插值(GSI)和线性插值(LI)之间的区别如图6所示,原始跟踪结果用橙色实线表示,它通常包含了噪声抖动;线性插值LI用蓝色实线表示,它忽略了运动信息;高斯平滑插值GSI用红色实线表示,它使用自适应平滑度因子对整个轨迹进行平滑,同时解决了噪声抖动和忽略运动信息这两个问题。对于GSI,在实验中允许插值的

最大间隙是20帧。

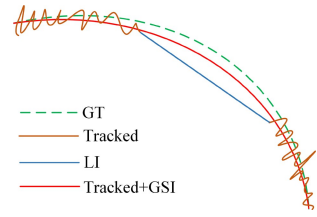


图6 线性插值和高斯平滑插值的区别(电子版为彩图)

Fig. 6 Difference between LI and GSI

5 实验与结果分析

5.1 实验环境与数据集

本实验基于Linux平台通过Python语言实现,采用Ubuntu 20.04的操作系统,深度学习框架为Pytorch 1.7,采用CUDA 11.0和cuDNN 8.0加速训练;所运行的硬件平台如下:处理器为Intel Core i9-11900K,显卡为NVIDIA RTX 3090。

实验选用红外地面多目标跟踪数据集,红外地面多目标跟踪数据集部分示例如图7所示,数据集由87条序列组成,选取其中18条序列作为测试数据集。



图7 红外地面多目标跟踪数据集部分示例

Fig. 7 Some examples of infrared ground multi-object tracking dataset

5.2 评价指标

采用MOT Challenge评价标准来评估本文提出的多目标跟踪算法,评价指标如表2所列,评价指标中(\uparrow)表示数值越高跟踪效果越好,(\downarrow)则相反。

表2 多目标跟踪评估指标

Table 2 Multi-object tracking evaluation indicators

评价指标	定义
MOTA(\uparrow)	跟踪准确度
IDF1(\uparrow)	正确识别的检测与平均真实数和计算检测数之比
IDs(\uparrow)	ID切换总数
FP(\downarrow)	误检的目标总数
FN(\downarrow)	漏检的目标总数
FM(\downarrow)	轨迹碎片化的总次数(即在跟踪过程中中断)

5.3 实验结果与分析

5.3.1 组件消融实验

为了验证改进算法各模块的有效性,使用红外地面多目标数据集对改进算法进行了一些组件消融实验,组件消融实验的结果如表3所列。

表3 组件消融的实验结果

Table 3 Component ablation results

NSA	ECC	LSTM	AFLink	GSI	MOTA/%	IDF1/%
—	—	—	—	—	76.9	85.7
✓	—	—	—	—	77.1	85.8
✓	✓	—	—	—	77.9	86.1
✓	✓	✓	—	—	80.1	87.1
✓	✓	✓	✓	—	80.4	87.5
✓	✓	✓	✓	✓	80.8	87.7

引入增强相关系数最大化算法 ECC 后, MOTA 提升了 0.8%, IDF1 提升了 0.3%, 跟踪性能有一定提升, 这表明引入的 ECC 算法对存在大量平台晃动的测试序列的跟踪效果有所改善; 增加基于 LSTM 的运动模型后, MOTA 提升了 2.2%, IDF1 提升了 1.0%, 表现出了较好的效果, 这表明增加的运动模型能够较好地处理测试序列中存在的多目标严重遮挡干扰问题; 采用无外观连接模型 AFLink 和高斯平滑插值算法 GSI 后, MOTA 提升到 80.8%, IDF1 提升到 87.7%, 跟踪性能继续有所提升, 这表明两种离线处理算法能够有效地完善跟踪结果。即使改进算法增加了 3 个模块, 但由于输入的处理图像为单通道灰度图像, 因此改进算法依然能够实时工作。多目标跟踪算法改进前后的效果如图 8 所示, 改进前, 目标 ID 发生变更; 改进后, 目标 ID 保持稳定不变。

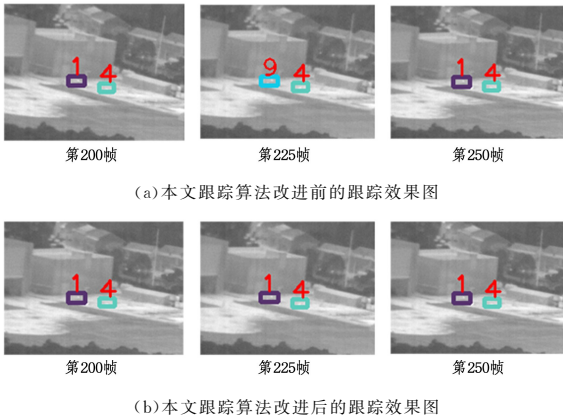


图8 组件消融实验的跟踪效果图

Fig. 8 Tracking effect of component ablation experiments

5.3.2 多目标跟踪算法的对比实验

改进算法选择红外地面多目标数据集进行实验, 与其他两种多目标跟踪算法结果进行对比, 结果如表 4 所列。

表4 不同跟踪算法的结果对比

Table 4 Results comparison of different tracking algorithms

Method	MOTA/%	IDF1/%	IDs	FP	FN	FM
SORT	72.5	81.2	18	438	4259	268
DeepSort	70.6	82.1	15	1357	3660	608
ours	80.8	87.7	11	753	2530	305

综合分析各项指标可以看出, 改进算法与其他两种多目标跟踪算法相比具有相对优势。改进算法着重考虑地面杂波干扰、多目标遮挡干扰、平台晃动等复杂场景跟踪效果不佳的问题, 因此改进算法在各项指标上都有所提升。相比 Sort, DeepSort 虽然加入了表观特征的度量方式, 但是由于红外

地面目标通常为小目标, 同时一些目标的形状结构不清晰, 不容易区分, 因此 DeepSort 算法的跟踪精度反而下降了 1.9%, 同时轨迹碎片增加到 608 个, 但是表观特征的加入依然使得 DeepSort 算法的 IDF1 提升了 0.9%, IDs 下降了 3 次, 这表明在红外地面多目标场景中, 表观特征的加入对跟踪算法性能的提升没有达到预期; 相比 Sort, 改进算法对低分框进行了二次分配并且在运动特征方面进行了改进, 因此跟踪性能有较大改善, MOTA 提升了 8.3%, IDF1 提升了 6.5%, 这表明从运动特征方面进行改进, 会给算法的跟踪性能带来较大的提升。综上, 可以验证改进算法的有效性和优势。

5.3.3 不同场景下的实验效果

测试序列 1 为白天场景, 3 种跟踪算法在测试序列 1 上的跟踪效果如图 9 所示。

测试序列 1 中存在大量平台晃动, 由于为白天场景, 因此目标与背景对比度小。3 种算法在第 60 帧时都能正确检测并跟踪 4 个目标。由于 2 号目标为小目标, 并且伴随平台晃动, 因此在第 150 帧和第 240 帧时依然能检测并跟踪 4 个目标, 但是 Sort 和 DeepSort 对 2 号目标的跟踪都发生了 ID 变换, 而改进算法却能正确检测并且稳定跟踪 4 个目标, 这说明改进算法针对运动特征方面的改进是有效的。

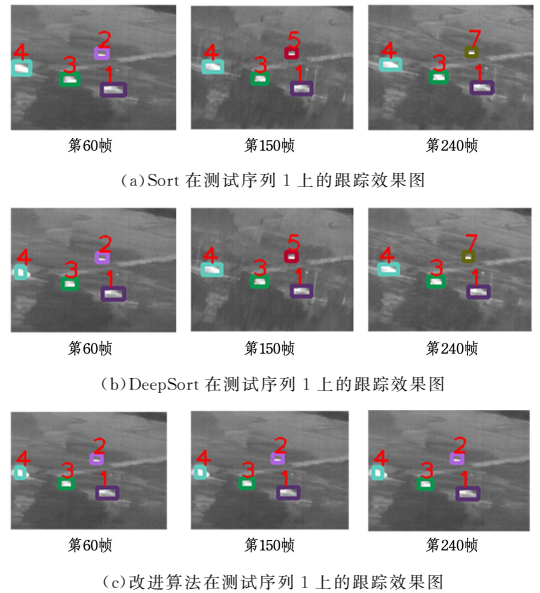


图9 3种跟踪算法在测试序列1上的跟踪效果图

Fig. 9 Tracking effect of three tracking algorithms on test sequence 1

测试序列 2 中存在平台晃动。3 种算法在第 135 帧时都能正确检测并跟踪 2 个目标; 在第 190 帧时依然能检测并跟踪 2 个目标, 但是 Sort 对 4 号目标的跟踪发生了 ID 变换, 而 DeepSort 由于加入了表观特征, 可以稳定跟踪目标, 保持目标 ID 不变, 改进算法由于做了针对运动特征方面的改进, 因此也能够持续稳定跟踪目标; 在第 245 帧时, 4 号目标变得较小, 检测分数较低, 因此被 Sort 和 DeepSort 过滤掉而造成漏检, 改进算法并没有直接过滤掉低分检测框, 对低分检测框进行了二次分配, 因此依然可以跟踪 4 号目标。

测试序列 2 为白天场景, 3 种跟踪算法在测试序列 2 上的跟踪效果如图 10 所示。

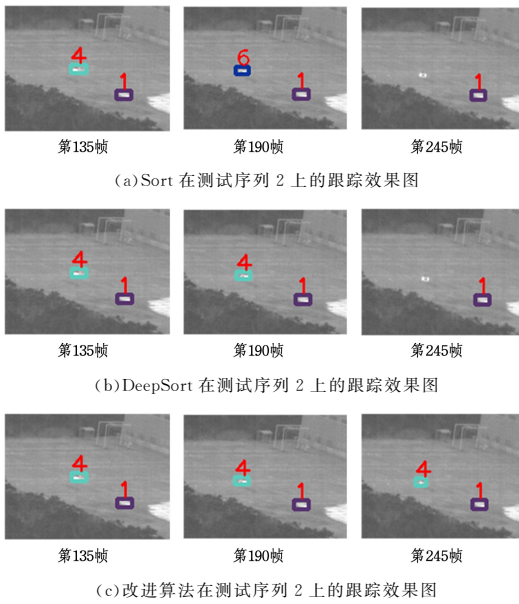


图 10 3 种跟踪算法在测试序列 2 上的跟踪效果图

Fig. 10 Tracking effect of three tracking algorithms on test sequence 2

测试序列 3 为夜间场景, 3 种跟踪算法在测试序列 3 上的跟踪效果如图 11 所示。

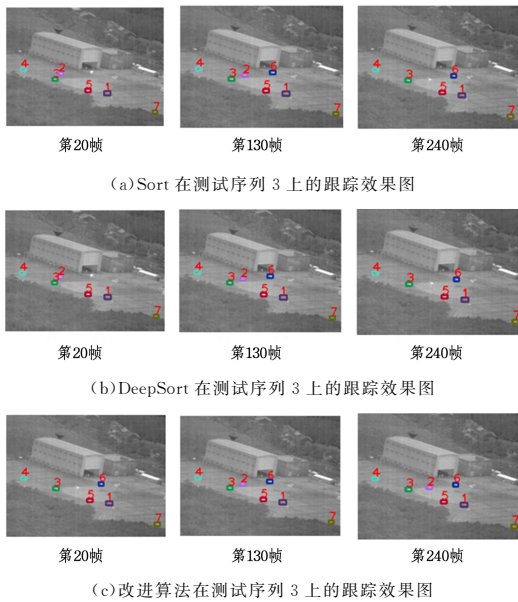


图 11 3 种跟踪算法在测试序列 3 上的跟踪效果图

Fig. 11 Tracking effect of three tracking algorithms on test sequence 3

测试序列 3 中存在平台晃动, 虽然为夜晚场景, 但由于距离较远, 因此地面目标较小且低分数检测框占大多数。第 20 帧时, 6 号目标检测分数较低, 因此被 Sort 和 DeepSort 过滤掉而造成漏检, 改进 ByteTrack 算法并没有直接过滤掉低分检测框, 对低分检测框进行了二次分配, 因此依然可以跟踪 6 号目标。3 种算法在第 130 帧和第 240 帧时均能正确检测并跟踪 7 个目标。

测试序列 4 为夜间场景, 3 种跟踪算法在测试序列 4 上的跟踪效果如图 12 所示。

测试序列 4 为夜晚场景, 目标与背景的对比较大。3 种算法均能对此测试序列上的 2 个目标进行比较稳定的跟踪。

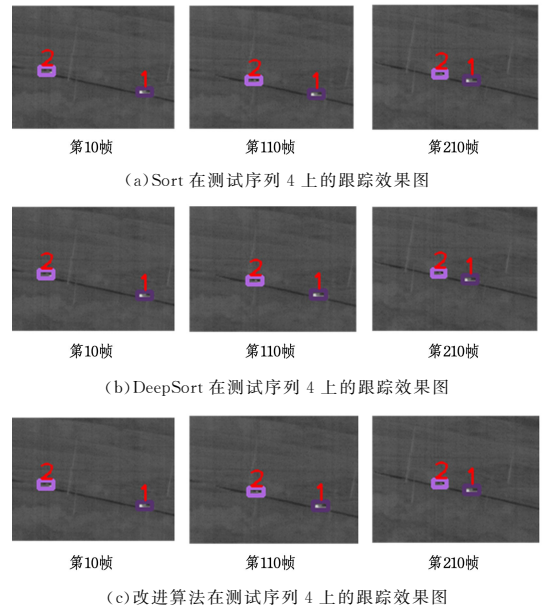


图 12 3 种跟踪算法在部分序列上的跟踪效果图

Fig. 12 Tracking effect of three tracking algorithms on test sequence 4

结束语 本文针对红外地面多目标跟踪场景中存在的地面杂波干扰、多目标遮挡干扰、平台晃动等复杂场景所造成的跟踪精度降低等问题, 提出了一种基于改进 ByteTrack 算法的红外地面多目标跟踪方法。实验结果表明, 改进算法有效地提高了多目标跟踪的跟踪精度, 有针对性地提高了整体跟踪算法的准确性和可靠性, 实现了复杂场景下持续稳定的多目标跟踪。但同时红外地面目标并没有清晰轮廓且成像不规则, 当目标与背景对比不明显并且很难与背景区分开时, 红外地面目标的检测是有困难的, 跟踪效果并不能得到保证, 因此未来的研究方向是如何更好地从与目标相似的背景中区分出红外地面目标, 来保证后续红外地面多目标的持续稳定跟踪。

参考文献

- [1] LIU C. Multi-object Fault-tolerant Tracking and Trajectory Prediction for Unmanned Vehicles [D]. Hangzhou: Zhejiang University, 2019.
- [2] FU R G, FAN H Q, ZHU Y F, et al. A dataset for infrared time-sensitive target detection and tracking for air-ground application [J]. China Scientific Data, 2022, 7(2): 203-218.
- [3] ZHANG Y F, SUN P Z, JIANG Y, et al. Byte-Track: Multi-object Tracking by Associating Every Detection Box [EB/OL]. <https://arxiv.org/pdf/2110.06864.pdf>, 2022.
- [4] DU Y H, WAN J F, ZHAO Y Y, et al. A Comprehensive Framework for MCMOT With Global Information and Optimizing Strategies in VisDrone 2021[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2809-2819.
- [5] ESS A, LEIBE B, SCHINDLER K, et al. A mobile vision system

- for robust multi-person tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008: 1-8.
- [6] DU Y H, SONG Y, YANG B, et al. StrongSORT: Make DeepSORT Great Again [EB/OL]. <https://arxiv.org/pdf/2202.13514.pdf>, 2022.
- [7] KIM C, LI F, CIPTADI A, et al. Multiple hypothesis tracking revisited[C]//IEEE International Conference on Computer Vision. 2015:4696-4704.
- [8] REID D B. An algorithm for tracking multiple targets[J]. IEEE Transactions on Automatic Control, 1979, 24(6):843-854.
- [9] YOON J H, YANG M H, LIM J, et al. Bayesian multi-object tracking using motion context from multiple objects[C]//2015 IEEE Winter Conference on Applications of Computer Vision. 2015:33-40.
- [10] BEWLEY A, GE Z Y, OTT L, et al. Simple Online and Realtime Tracking[C]//2016 IEEE International Conference on Image Processing(ICIP). 2016:3464-3468.
- [11] WOJKE N, BEWLEY A, PAULUS D, et al. Simple online and real-time tracking with a deep association metric[C]//2017 IEEE International Conference on Image Processing(ICIP). 2017:3645-3449.
- [12] CHEN L, AI H Z, ZHUANG Z J, et al. Real-Time Multiple People Tracking with Deeply Learned Candidate Selection and Person Re-Identification [EB/OL]. <https://arxiv.org/pdf/1809.04427.pdf>, 2018.
- [13] WANG Z D, ZHENG L, LIU Y X, et al. Towards real-time multi-object tracking [EB/OL]. <https://arxiv.org/pdf/1909.12605.pdf>, 2018.
- [14] ZHANG Y F, WANG C Y, WANG X G, et al. FairMOT: On the fairness of detection and re-identification in multiple object tracking[EB/OL]. <https://arxiv.org/pdf/2004.01888.pdf>.
- [15] ZHOU X Y, KOLTUN V, KRAHENBUHL P, et al. Tracking Objects as Points [EB/OL]. <https://arxiv.org/pdf/2004.01177.pdf>, 2020.
- [16] WANG L X, ZENG D, ZHU X Q. Space Based Infrared Small Target Detection and Tracking System [J]. Industry Control Computer, 2022, 35(5):39-41.
- [17] SONG Z Z, YANG J W, ZHANG D F, et al. Real-time infrared multi-class multi-target anchor-free tracking network[J]. Systems Engineering and Electronics, 2022, 44(2):401-409.
- [18] YANG B, LIN S Z, LU X F, et al. Multiple aerial infrared target tracking method based on multi-feature fusion and hierarchical data association [J]. Journal of Computer Applications, 2022, 40(10):3075-3080.
- [19] XIE X F, LIU H J, ZHANG L J, et al. Infrared Ship Target Recognition Fused with Lightweight YOLOv4 and KCF Algorithm[J]. Journal of Ordnance Equipment Engineering, 2021, 42(6):175-182.
- [20] YANG W, YAO T, GENG X T, et al. Research on application of infrared target recognition technology of air defense weapon based on deep learning[J]. Journal of Ordnance Equipment Engineering, 2022, 43(1):125-129.
- [21] ZHENG G, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO Series in 2021 [EB/OL]. <https://arxiv.org/pdf/2107.08430.pdf>, 2021.



WANG Luo, born in 1995, postgraduate. His main research interests include image processing, computer vision and automation target recognition.



FU Ruigang, born in 1991, Ph. D. His main research interests include image processing, computer vision and automation target recognition.

(责任编辑:喻黎)