

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-14

论文引用格式: Song Jian, Fu Qiang, Yang Haojie, Long Yucheng, Li Zhanwu, Kou Renke. A New Benchmark for UAV Multi-Object Tracking in Complex Scenes: MIATrack[J/OL]. Journal of Image and Graphics, XXXX: 1-14. DOI: 10.11834/jig.260317. (宋健, 付强, 杨皓杰, 龙昱承, 李战武, 寇人可. 面向复杂场景的无人机多目标跟踪新基准: MIATrack[J/OL]. 中国图象图形学报, XXXX: 1-14. DOI: 10.11834/jig.260317.) [DOI: 10.11834/jig.260317]

面向复杂场景的无人机多目标跟踪新基准: MIA-Track

宋健¹, 付强¹, 杨皓杰¹, 龙昱承¹, 李战武², 寇人可^{2,3*}

1. 陆军工程大学, 石家庄 050000; 2. 空军工程大学, 西安 710000; 3. 中国人民解放军95084部队, 佛山 528000

摘要: 目的 反无人机场景中无人机目标强机动特性对红外多目标跟踪算法的轨迹关联能力提出极高要求。现有反无人机红外数据集存在评测偏向性, 普遍关注低信杂比场景的目标检测能力。同时, 其无人机运动模式普遍较为平缓且类型单一, 既难以还原真实场景下无人机非线性、非平稳的强机动特性, 也难以对复杂动态环境中跟踪算法的鲁棒性开展定量评测。针对上述问题, 本文构建了面向高机动场景的合成反无人机红外多目标跟踪数据集 (Maneuver Infrared Anti-UAV Tracking Dataset, MIATrack)。方法 该数据集从公开红外数据源提取纯净背景与无人机目标素材, 通过物理仿真建模生成高度非线性、非平稳复杂运动轨迹, 复现无人机多样机动行为; 并通过生成参数调控使目标保持较高信杂比, 以降低检测不确定性, 将算法的评估重心从目标检测能力转向复杂动态下的跟踪鲁棒性, 为高机动红外无人机多目标跟踪提供专用评测基准。结果 在 MIATrack 数据集上对主流跟踪算法开展基准测试, 结果表明, 相比现有公开数据集, 检测性能仍保持在较高水平, 但各类跟踪算法在本数据集上的轨迹关联性能均出现明显下滑。结论 这一对比证明, 当前红外无人机跟踪算法的性能瓶颈已由目标检测转向复杂非线性机动条件下的稳健轨迹关联。数据集与项目代码已开源发布于: <https://github.com/SongJgit/TBDTracker>。

关键词: 红外弱小目标检测与跟踪; 强机动目标; 多目标跟踪; 反无人机红外数据集; 序列图像

A New Benchmark for UAV Multi-Object Tracking in Complex Scenes: MIA-Track

Song Jian¹, Fu Qiang¹, Yang Haojie¹, Long Yucheng¹, Li Zhanwu², Kou Renke^{2,3*}

1. Army Engineering University, Shijiazhuang 050000, China; 2. Air Engineering University, Xi'an 710000, China; 3. Chinese People's Liberation Army Unit 95084, Foshan 528000, China

Abstract: Objective With the rapid development and extensive application of unmanned aerial vehicles (UAVs) in both civilian and military fields, all-weather, robust multi-object tracking (MOT) technology has become a core technical requirement for low-altitude security and defense fields. Infrared thermal imaging is capable of stable perception under low-light and all-day conditions, making it the preferred sensing method for UAV monitoring. However, infrared UAV targets are characterized by small pixel scales, low signal-to-clutter ratio (SCR), high flight speed and strong nonlinear non-stationary maneuvers, which pose great challenges to existing MOT algorithms. Existing open-source infrared anti-UAV datasets generally exhibit obvious evaluation bias. They place excessive emphasis on object detection performance in low

收稿日期: 2026-06-02; 修回日期: 2026-09-02

* 通信作者: 寇人可 krkoptics@163.com

基金项目: 国家自然科学基金项目 (62605485)

Supported by: National Natural Science Foundation of China (62605485)

SCR scenarios and UAV motion patterns are predominantly stable, with mild, regular trajectories. Such datasets fail to reflect the high-speed variable motion, sharp turning, hovering and complex tactical maneuver behaviors of real UAVs and cannot quantitatively evaluate the robustness of tracking algorithms in highly dynamic and complex environments. In addition, the low SCR of existing datasets mixes detection errors with data association errors, making it difficult to locate the real causes of tracking performance degradation. To address the limitations of existing benchmarks, we construct MIATrack (Maneuver Infrared Anti-UAV Tracking Dataset), a synthetic infrared anti-UAV MOT dataset for highly maneuverable UAVs. It further reveals the real performance bottleneck of tracking algorithms under complex nonlinear maneuvers.

Method We first build a library of clean infrared backgrounds covering typical scenes, including skies, sea surfaces, deserts and urban areas, as well as a library of targets featuring various types of UAVs. Based on raw infrared images and mask annotations from public infrared anti-UAV datasets, we remove UAV targets using mask expansion and K-nearest-neighbor inpainting to produce a purified background library. Corresponding UAV instances are also extracted to establish a high-quality infrared UAV target library. We then adopt a physics-driven modeling framework to synthesize realistic infrared image sequences. Multiple kinematic models are employed to create complex flight trajectories, including constant-velocity motion, constant-acceleration motion, circular motion, serpentine maneuvers and Z-shaped zigzag motion. Manual trajectory customization is also supported. We also simulate realistic infrared target characteristics, including distance-dependent geometric scaling, atmospheric radiometric attenuation, attitude-induced coordinate rotation and point-spread-function-based optical blurring. By adjusting generation parameters, MIATrack maintains a high SCR to mitigate detection uncertainty, thereby shifting the focus from detection accuracy to tracking robustness. We adopt adjacent-frame IoU and the relative-position-switching frequency to quantitatively evaluate the motion complexity of MIATrack and conduct comprehensive benchmark tests on the YOLO26 detection model series and representative tracking algorithms, including SORT, ByteTrack, OC-SORT, SparseTrack and ImprAsso. **Result** Statistical results show that MIATrack achieves a considerably higher average SCR than existing mainstream datasets, including AntiUAV and 4th-AntiUAV, reflecting its intentional design to maintain cleaner target signatures and reduce detection uncertainty. In terms of motion characteristics, MIATrack presents a markedly lower mean adjacent-frame IoU and a notably higher relative position switching frequency, confirming that it contains highly nonlinear, non-stationary trajectories and complex target interactions. Detection experiments show that YOLO26 models of varying scales achieve favorable mAP on MIATrack, demonstrating that the high SCR design effectively suppresses clutter interference. In MOT tests, mainstream trackers maintain competitive MOTA scores thanks to high-quality detection results. However, they experience a sharp drop in core association metrics, including HOTA, IDF1, and AssA. This pronounced performance discrepancy between detection and association clearly indicates that complex maneuvers severely degrade the trajectory-matching and identity-preservation capabilities of existing tracking algorithms. **Conclusion** MIATrack compensates for the insufficient motion diversity of existing infrared anti-UAV benchmarks and enables comprehensive and rigorous evaluation of multi-target tracking algorithms in complex maneuver scenarios. Benchmark results indicate that the core bottleneck of infrared anti-UAV MOT has shifted from target detection to motion prediction and data association in complex maneuver scenarios. By maintaining a high signal-to-clutter ratio, MIATrack reduces detection uncertainty, enabling a more rigorous evaluation of robustness in motion modeling and data association under complex maneuvers. Accordingly, this work establishes a new baseline for future tracking algorithm development. The dataset and source code are publicly released to support the design, testing and practical deployment of infrared anti-UAV detection and multi-object tracking methods.

Key words: Infrared small target detection and tracking; Maneuver target tracking; Multi target tracking; Infrared datasets of anti-UAV; Sequence images

论文引用格式: Song J, Fu Q, Yang H J, Long Y C, Li Z W, Kou R K. 2026. MIATrack: A New Benchmark for UAV Multi-Object Tracking in Complex Scenes, XX(XX): XXXX-XXXX (宋健, 付强, 杨皓

杰, 龙昱承, 李战武, 寇人可. 2026. MIATrack: 面向复杂场景的无人机多目标跟踪新基准. 中国图象图形学报, XX(XX): XXXX-XXXX [DOI: 10.11834/jig.260317]

0 引言

随着无人机在民用、安防监控等领域的广泛应用,实现对无人机的全天时、高鲁棒性的持续跟踪已成为低空安防与反制领域的核心需求。热红外成像凭借其在弱光条件下的稳定成像能力,成为实现全天候无人机(Unmanned Aerial Vehicles, UAV)监控的理想技术途径。然而,红外图像通常分辨率较低,无人机目标尺度小、信杂比低,且目标飞行速度快、伴随非线性非平稳的剧烈机动,这给基于红外的多目标跟踪算法带来了严峻挑战。

高质量基准数据集是算法迭代与性能评测的重要基础。然而,受限于严格的空域管控及复杂的集群协同飞行条件,大规模、多场景的真实红外数据的获取十分困难。此外,红外小目标固有的低信杂比、低分辨率特性,加之其在时序上的剧烈机动性,使得轨迹级标注极具挑战且易引入显著误差。早期工作如(Huang 等, 2024; Jiang 等, 2023; Hui 等, 2020)为红外无人机单目标检测与跟踪研究提供了基础数据支撑;随后,规模更大、场景更丰富的基准(如4th-AntiUAV)相继建立,进一步推动了多无人机跟踪算法的发展。现有公开数据集的研究重心集中于复杂背景下的目标检测优化,重点解决目标有效识别的基础问题,但普遍存在运动模式单一的缺陷,数据集内的无人机运动轨迹平缓、规律性强,无法复现真实低空场景中无人机高速、非线性、非平稳的强机动特征。

当前先进的跟踪范式普遍采用“基于检测的跟踪(Tracking-by-detection, TBD)”框架。该框架通常将跟踪任务解耦为两个顺序执行的阶段,前者负责逐帧定位目标,后者利用运动预测模型和数据关联算法将相邻帧中的检测结果连接成连续的轨迹。然而,值得注意的是,现有红外无人机数据集在评测此类跟踪器时存在根本局限,难以客观评估算法应对高机动目标的鲁棒性。一方面,现有数据集中目标运动普遍平缓,导致运动模型和关联算法在处理非线性、非平稳运动时的性能瓶颈无法得到充分暴露与检验;另一方面,即便运动估计模型和关联模块本身存在缺陷,其影响也会与数据集中固有的检测噪声(如因低信杂比和小目标导致的漏检与误报)相互耦合,使性能归因变得困难。

针对上述问题,本文构建面向高机动红外无人机多目标跟踪的合成数据集 MIATrack (Maneuver Infrared Anti-UAV Tracking Dataset)。该数据集利用物理驱动的程序化生成方式,在确保目标具备较高信杂比以降低检测难度的同时,着重构造了高度非线性、非平稳的复杂运动轨迹。这种设计有效排除了前端检测误差的干扰,使得在检测条件可控的前提下,能够着重评估不同算法应对复杂机动运动的关联稳健性。基于此,本文进行了系统的基准评测,主要贡献如下:

- 1)构建面向机动红外无人机多目标跟踪的专用基准合成数据集 MIATrack,利用物理驱动方法生成复杂机动模式,弥补了现有公开数据集运动模式单一的不足;该数据集通过保持高信杂比的设计策略,有效分离了检测误差与关联误差,使得跟踪性能评估更加聚焦于算法在高机动场景下的关联鲁棒性。
- 2)依托自建数据集完成多组对照实验,建立完备的检测与跟踪基线,直观揭示了现有算法在高机动目标关联上的固有缺陷;
- 3)所构建数据集及基线指标可为后续算法研发提供可靠测试平台,便于高效评估各类机动红外目标跟踪方法的实际性能。

1 相关工作

1.1 红外无人机检测与跟踪数据集

红外无人机检测与跟踪领域一直存在公开数据集稀缺的问题。一方面,受成像设备成本与保密性的限制,真实数据的采集本身极为困难;另一方面,无人机目标在红外图像中通常仅占数十个像素,对其进行精确标注(无论是像素级掩膜分割还是跨帧ID关联)均需耗费大量专业人力。因此,早期研究多基于私有的小规模数据集,这使得算法的公平比较与性能评估难以实现。近年来,随着开源基准建设的推进,该领域呈现出由静态单帧向动态多帧、由目标稀疏向目标密集、由单一标注向多维综合标注的演进趋势。

在红外目标感知研究中,单帧红外小目标检测(Single infrared small target, SIRST)方向发展相对成熟(寇人可 等, 2024; 刘雨菡 等, 2026; 张迎梅 等, 2026),学界已陆续推出 MDFA (Wang 等, 2019)、NUAA-SIRST (Dai 等, 2021)、SIRST-v2 (Dai 等,

2023)、IRSTD-1K(Zhang 等, 2022)、NUDT-SIRST(Li 等, 2023)等通用公开数据集。然而, 此类数据集主要面向通用地空场景, 缺乏对无人机目标的针对性, 且不具备跟踪任务所必需的连续轨迹标注, 因此无法直接用于评估机动场景下的无人机跟踪性能。

为引入时序运动信息, 多帧红外小目标检测 (Multi-frame infrared small target, MIRST) 数据集逐渐得到发展。Hui 等人(2020)和Fu 等人(2022)先后发布了真实场景 MIRST 数据集, 开始纳入无人机目标; Sun 等人(2023)构建的IRDST数据集中同样包含部分无人机类别样本。但上述数据集普遍缺失目标身份标注, 或仅适配单目标跟踪场景, 难以支撑多无人机目标的跟踪任务评估。Jiang(2023)及Huang 等人(2024)依托反无人机挑战赛相继发布 AntiUAV、AntiUAV410 等单目标跟踪数据集, 完善了边界框与身份标注, 有效支撑了单目标跟踪算法的性能验证。第四届反无人机挑战赛(4th-AntiUAV)进一步面向多目标跟踪任务构建基准, 推动了红外无人机多目标跟踪领域的发展。此外, 可见光反无人机检测跟踪领域也已积累了一批代表性基准数据集。例如, DUT-AntiUAV(Zhao 等, 2022)、UAV-Anti-UAV(Zhang 等, 2025)构建了可见光模态下的无人机检测与跟踪评测基准。这些数据集在不同成像模态和任务需求下丰富了反无人机领域的评测体系, 为跨模态算法的研究与评估提供了重要支撑。

尽管如此, 现有数据集在目标运动模式的多样性与复杂性上存在明显局限。以 AntiUAV 系列为代表的主流基准数据集中, 无人机运动轨迹整体较为平缓, 运动规律可预测, 未能复现真实低空安防场景中无人机高速变速、紧急变向、悬停盘旋、非线性机动等复杂动态行为。运动场景单一、机动复杂度不足的问题, 导致现有跟踪任务评测难度偏低, 无法有效暴露算法在高机动目标跨帧关联中的性能缺陷, 难以真实验证跟踪算法的鲁棒性与泛化能力。

综上, 当前领域亟需构建全新基准数据集, 同时兼顾两方面需求。一是具备实例级边界框或像素掩码标注及完整的轨迹 ID, 支撑检测与跟踪任务的联合评估; 二是包含丰富且符合物理规律的高机动运动模式, 以有效检验算法在复杂动态场景下的泛化性能。针对现有数据集的上述不足, 本文依托物理知识约束下红外小目标图像视频生成系统(图1),

构建了一款全新的多帧红外小目标跟踪数据集 MIATrack。

1.2 基于检测的跟踪算法

随着检测和重识别技术的迅速发展, 基于检测的跟踪范式凭借其简洁的框架和强大的性能, 已成为当前在线多目标跟踪的主流方法, 如图2所示。在运动建模方面, SORT(Bewley 等, 2016)率先引

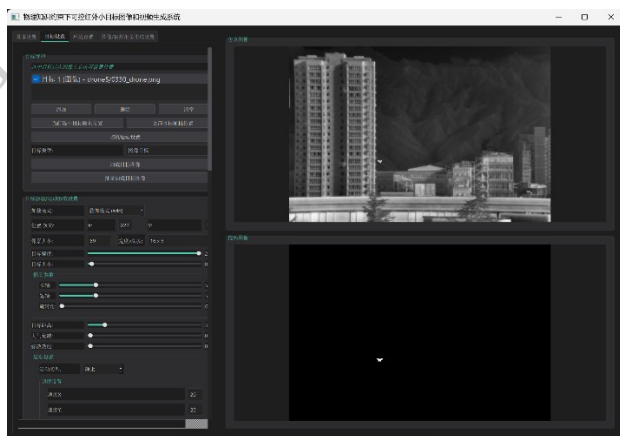


图1 物理知识约束下红外小目标图像视频生成系统

Fig. 1 Infrared Small Target Image and Video Generation System under Physical Knowledge Constraints

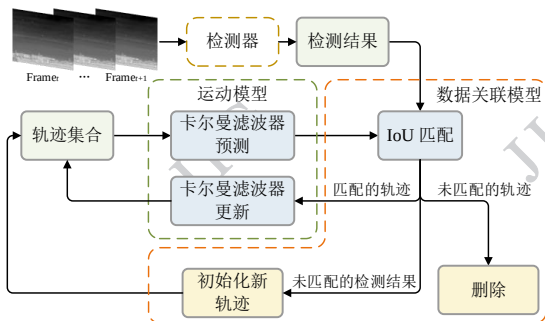


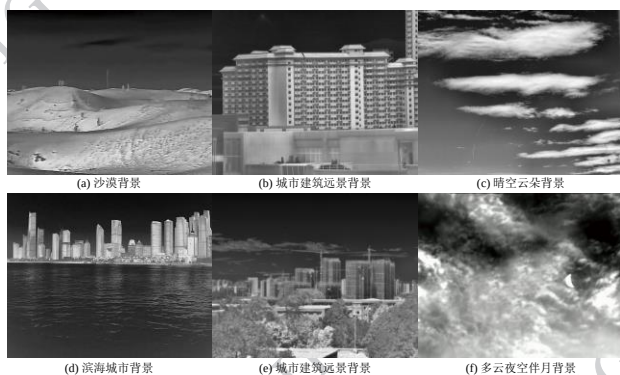
图2 基于检测的跟踪算法流程图

Fig. 2 Flowchart of a Tracking by Detection Algorithm

入卡尔曼滤波器作为运动估计器, 并利用匈牙利算法(Kuhn, 1955)完成轨迹关联, 奠定了基于运动的在线跟踪基础。

DeepSORT(Wojke 等, 2017)则在此基础上融合深度外观特征, 增强了遮挡下的身份维持能力。然而, 上述方法通常基于线性运动假设, 在面对目标尺度剧烈变化或非线性的复杂场景时表现受限。为此, OC-SORT通过改进状态更新策略, 降低了对线性运动假设的依赖, 提升了在复杂机动下的跟踪鲁棒性。在关联策略方面, ByteTrack提出的级联关联策略有效利用低置信度检测缓解了漏检导致的轨迹

断裂问题。ImprAsso (Stadler and Beyerer, 2023) 与 SparseTrack (Liu 等, 2025) 等工作从关联结构的稀疏性建模与代价函数优化入手, 提升了密集场景下的判别能力。这些进展共同推动了多目标跟踪性能的提升。尽管上述通用 MOT 方法取得了显著进展, 但它们在红外无人机跟踪场景中仍面临严峻挑战。一方面, 对于依赖外观特征的跟踪器, 红外图像仅反映热辐射强度, 缺乏颜色与纹理信息, 导致外观特征区分度极低, 难以支撑稳定的重识别。另一方面, 无人机目标具有高速、非平稳及非线性的运动特性, 这与行人或车辆等常规目标截然不同, 其运动规律难以用传统的线性动力学模型进行有效拟合。因此, 即便是那些主要依赖运动学模型的先进跟踪器, 其内置的线性运动假设或基于统计的关联策略在处理此类高机动目标时也存在固有局限。



((a) desert background; (b) cityscape in the background (long shot); (c) cloud background; (d) coastal city background; (e) cityscape in the background (close-up); (f) cloudy sky with moon background)

图3 多类型背景数据库

Fig. 3 Typical infrared background images

2 数据集构建

2.1 数据库构建

2.1.1 背景数据库构建

为保障合成数据集背景场景的多样性与真实性, 本文构建了一个包含 4875 张分辨率为 640×512 的红外图像背景库, 部分背景样本如图 3 所示。该背景库融合了来自多个权威公开数据集的红外样本, 并补充了自主采集的图像素材, 使其能够全面覆盖沙漠、城市建筑、天空、城市-海天交界、多云夜空

等多类典型红外场景。上述场景在红外成像特性上差异显著, 沙漠与城市背景纹理丰富、热辐射梯度大, 对目标检测形成较强的背景杂波干扰; 纯天空与多云夜空背景相对均匀, 但易与弱小目标在灰度上混淆, 对目标-背景区分能力要求更高。多类型、多纹理、多辐射特性的背景库为合成数据集的场景泛化能力奠定了数据基础。

上述背景样本主要来源于两类途径: 一是自主采集的不含目标的纯净红外图像, 可直接用作背景素材; 二是来自公开数据集、本身携带原始无人机目标的红外图像, 无法直接使用。对于后者, 本文结合其像素级标注完成背景净化处理, 彻底消除原有目标干扰。具体而言, 首先对标注掩码图执行膨胀操作以扩大目标区域, 随后在原图中移除该区域内的像素; 最后, 利用 K 近邻 (K-Nearest Neighbor, KNN) 算法对空缺区域进行插值修复, 从而在彻底剔除目标信息的同时, 有效保留并还原背景纹理细节, 为后续红外无人机目标仿真合成提供纯净、高质量的背景样本。上述操作中, 膨胀核尺寸统一设置为 2, 约为典型小目标短轴长度的 $1/3$, 既能完整覆盖目标边缘的过渡混合像素、彻底去除目标辐射残留, 又不致过度扩大修复区域、加重插值负担; 同时, 针对 K 近邻插值在较大空缺区域上可能引入平滑伪纹理或边缘伪影的问题, 本文通过对修复结果逐幅人工抽检、并在后续目标嵌入时优先选择远离修复区域的背景位置加以规避, 从源头上避免插值伪影与合成目标产生空间耦合。



((a) looking at portable drones head-on; (b) looking up at portable drones; (c) looking at drone models with landing gear head-on)

图4 部分目标图像

Fig. 4 Selected UAV target instances

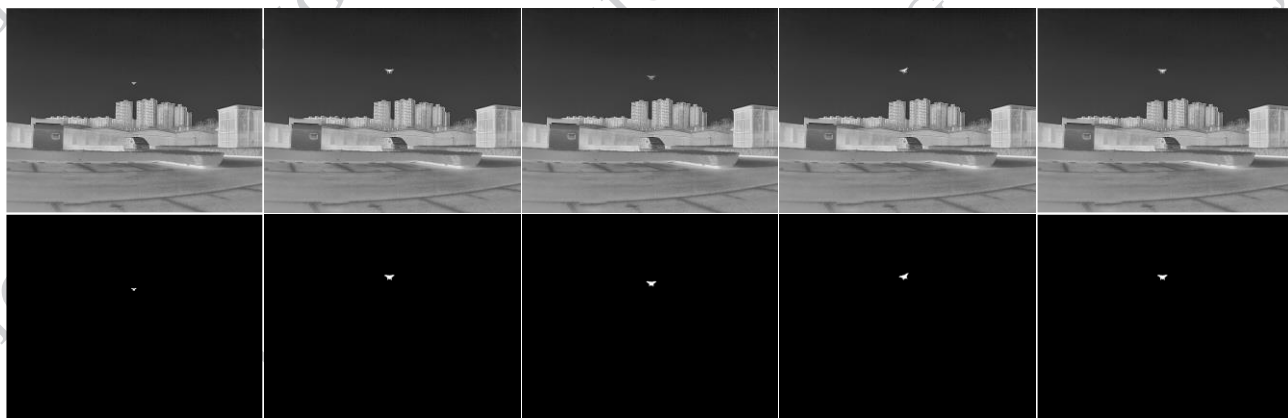
2.1.2 目标数据库构建

本文构建了包含 1036 个实例的红外无人机目标库, 涵盖多种型号、不同飞行姿态的无人机目标, 部分样本如图 4 所示。

本数据集中所有无人机目标并非随机合成, 而
© 中国图象图形学报版权所有

是全部根据公开红外反无人机数据集的像素级人工标注从原始图像中进行提取,这些标注由领域专家通过多轮校验完成,从源头上确保了目标的真实性与可靠性。具体而言,我们首先从公开数据集中筛选出标注质量较高的原始图像及其掩码;随后,利用这些高质量的像素级标注从原始图像中精确分离并提取目标区域,得到候选目标实例集。为确保目标库的纯净度与可靠性,我们对候选实例执行了三级过滤:(1)尺寸过滤,剔除面积<9像素或宽/高<3像

素的过小实例,避免噪点或伪目标误入;(2)信杂比过滤,剔除信杂比(signal-to-clutter ratio, SCR)<0.5的低质实例,确保目标具备清晰的辐射特征与成像对比度;(3)人工校验,对自动提取后的目标切片逐个人工目视复核,剔除标注模糊、目标残缺或存在贴图伪影的异常样本。经统计,共提取候选目标实例2283个,经尺寸过滤保留1537个、经信杂比过滤保留1396个,最终经人工校验后形成包含1036个高纯净度实例的红外无人机目标库。



(a) 原始目标 (b) 尺度缩放 (c) 强度衰减 (d) 姿态旋转变换 (e) 光学弥散效果

((a) original target; (b) size scaling; (c) intensity attenuation; (d) pose rotation transformation; (e) optical blurring)

图5 物理驱动建模下无人机目标的成像特征变化

Fig. 5 Imaging characteristic variations of UAV targets under physics-driven modeling

2.2 可控数据集生成

2.2.1 物理驱动的静态单目标建模

为精准刻画真实场景下无人机红外目标的成像特征,本文采用物理驱动建模方法,通过尺寸缩放、目标强度衰减、坐标系旋转变换与光学弥散校正四类核心模型,还原目标在几何、辐射、姿态与光学维度的成像规律,如图5所示。

在几何尺寸建模方面,本文采用基于距离-像素映射关系的尺寸缩放模型,刻画成像尺寸随观测距离的变化规律,公式如下:

$$\lambda_s = \frac{d_0}{d} \cdot \frac{s}{s_0} \quad (1)$$

$$\begin{cases} a = \max(0.1, a_0 \cdot \lambda_s) \\ b = \max(0.1, b_0 \cdot \lambda_s) \end{cases}$$

式中, λ_s 为缩放因子, s_0 为基大小, s 为原始目标大小, a_0 和 b_0 分别为目标的初始半长轴和半短轴, a 和 b 分别为尺度变换后的半长轴和半短轴。该模型能够反映不同探测距离下无人机小目标的像素尺寸变

化规律,贴合远距离红外小目标的几何成像特征。

对于辐射强度建模,本文采用基于距离平方反比定律与大气衰减定律的辐射强度衰减模型,表征红外辐射亮度随观测距离的变化规律,公式为:

$$I = I_0 \cdot \frac{1}{1 + \left(\frac{d}{d_0}\right)^2} \cdot \exp\left(-k \cdot \frac{d}{d_0}\right) \quad (2)$$

式中, I_0 为原始目标强度, d_0 为基距离, d 为距离观测点的距离, k 为大气衰减系数。该模型能够反映观测距离对红外能量的削弱效应。此外,由于从公开数据集提取的目标携带原始场景的热辐射残留,直接嵌入可能导致目标与新背景的亮度对比出现异常。为此,在目标嵌入时引入背景辐射匹配,根据目标在背景中的嵌入位置,对目标强度进行局部背景自适应修正— $I_{\text{adj}} = I_{\text{orig}} \times \gamma + I_{\text{bg}_{\text{mean}}} \times (1 - \gamma)$, 式中 γ 为保留系数(默认0.7-0.9), $I_{\text{bg}_{\text{mean}}}$ 为嵌入位置局部背景的平均灰度。这一修正使得目标的热辐射特征与所在背景的辐射水平协调一致,能够有效抑制贴纸

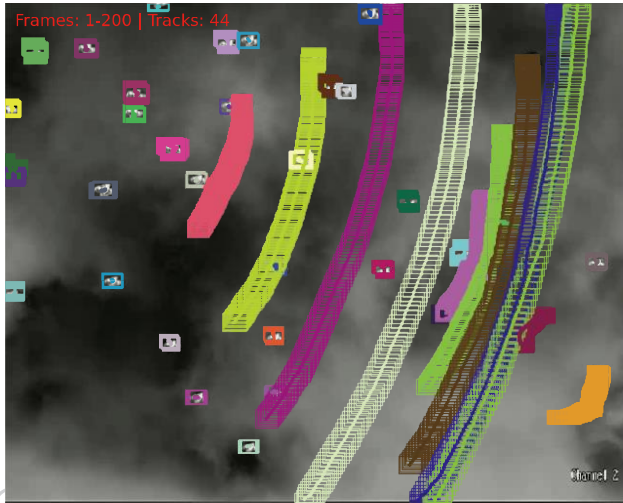
效应。

为了描述无人机姿态偏转对成像的影响,本文采用坐标系旋转变换模型描述无人机姿态偏转带来的成像变化,通过坐标变换修正目标位置,复现无人

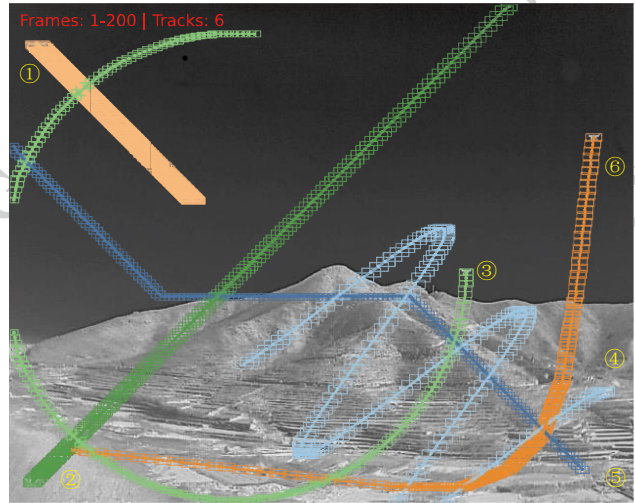
机的转向、偏转等姿态特征,其公式为:

$$(x_r, y_r) = (x - x_c, y - y_c) \quad (3)$$

$$\begin{bmatrix} x_{rot} \\ y_{rot} \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} x_r \\ y_r \end{bmatrix} \quad (4)$$



(a) 4th-AntiUAV



(b) MIATrack

((a) 4th-AntiUAV 数据集 200 帧内 44 条轨迹,多数目标近乎静止;(b) MIATrack 数据集 200 帧内 6 条典型轨迹,①匀速运动、②匀加速运动、③曲线运动、④蛇形运动、⑤ Z 字型运动、⑥ 手绘轨迹)。本图以第一帧图像为参考背景,将第 1~200 帧中同一目标的所有边界框叠加绘制在同一画面,(b) 中数字标号位置为轨迹起点。其中,①~⑤ 由物理运动学方程驱动生成,速度和加速度变化较为规律;⑥ 为人工手绘,轨迹更为多变且不规则)

((a) 44 trajectories within 200 frames in the 4th-AntiUAV dataset, where most targets remain nearly stationary; (b) six typical trajectories within 200 frames in the MIATrack dataset: ① constant-velocity motion, ② constant-acceleration motion, ③ circular motion, ④ serpentine maneuver, ⑤ Z-shaped zigzag motion, ⑥ hand-drawn trajectory. Each trajectory is visualized by overlaying the bounding boxes of the same target from frames 1 to 200 onto the first frame as the reference background. The numbered markers in (b) indicate the starting position of each trajectory. Trajectories ①~⑤ are generated by kinematic equations, exhibiting regular velocity and acceleration variations; trajectory ⑥ is hand-drawn, showing more diverse and irregular path characteristics.)

图6 4th-AntiUAV 与 MIATrack 部分轨迹可视化对比

Fig. 6 Selected trajectory visualization comparison between 4th-AntiUAV and MIATrack

式中, (x_c, y_c) 为目标的中心坐标, (x_r, y_r) 为目标局部相对坐标, (x_{rot}, y_{rot}) 为旋转变换后的局部坐标, θ 为旋转角度。

考虑到实际红外光学系统的离焦与衍射效应,本文引入点扩散函数 (Point Spread Function, PSF) 模拟实际红外光学系统的退化效应,公式如下:

$$h(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (5)$$

式中, (x, y) 为目标局部像素坐标, σ 为弥散尺度参数,用于控制目标区域模糊程度。该模型能够有效还原真实红外成像中的能量扩散与边缘柔化特征。此外,该光学弥散过程对目标边缘产生的低通滤波效应,在目标嵌入背景时起到了天然的边界柔

化作用,有助于抑制目标与背景之间的边缘突变。

通过上述多维度物理建模,可精准刻画无人机红外目标的多维成像特性,形成真实且多样的目标样本,为后续跟踪算法的训练与评测提供可靠的数据先验支撑。

2.2.2 轨迹多样性与运动学建模

本文采用多种运动学模型生成飞行轨迹,覆盖无人机在复杂空域下的各类飞行行为。首先,构建匀速直线与匀加速两类基础运动模型,模拟常规飞行状态。

匀速运动模型公式为:

$$\begin{cases} x(t) = x_0 + v_x \cdot t \\ y(t) = y_0 + v_y \cdot t \end{cases} \quad (6)$$

式中, (x_0, y_0) 为目标初始位置, (v_x, v_y) 为速度分量 (像素/秒), t 为时间。该模型可以模拟无人机平稳巡航的飞行过程。

匀加速运动模型公式为:

$$\begin{cases} x(t) = x_0 + v_x \cdot t + \frac{1}{2} a_x \cdot t^2 \\ y(t) = y_0 + v_y \cdot t + \frac{1}{2} a_y \cdot t^2 \\ v_x(t) = v_{x0} + a_x \cdot t \\ v_y(t) = v_{y0} + a_y \cdot t \end{cases} \quad (7)$$

式中, (a_x, a_y) 为加速度分量 (像素/秒²), (v_{x0}, v_{y0}) 是初始速度。该模型可以模拟无人机的变速飞行过程。

为进一步复现复杂战术动作, 本文进一步实现蛇形运动、曲线运动及Z字型机动等复杂机动模型。

曲线运动模型用于描述目标围绕中心点做圆周运动的飞行轨迹, 其运动方程为:

$$\begin{cases} x(t) = c_x + r \cdot \cos(\omega \cdot t + \phi) \\ y(t) = c_y + r \cdot \sin(\omega \cdot t + \phi) \end{cases} \quad (8)$$

式中, (c_x, c_y) 为曲线中心坐标, r 为运动半径, ω 为角速度, ϕ 为初始相位角。

蛇形运动模型在直线运动基础上叠加周期性振荡, 用以模拟目标规避行为, 公式如下:

$$\begin{cases} x(t) = x_0 + v_x \cdot t + A \cdot \sin(2\pi \cdot f \cdot t) \cdot \sin(\theta + \pi/2) \\ y(t) = y_0 + v_y \cdot t + A \cdot \sin(2\pi \cdot f \cdot t) \cdot \cos(\theta + \pi/2) \end{cases} \quad (9)$$

式中, A 为振荡振幅, f 为振荡频率, $\theta = \arctan(v_x/v_y)$ 是运动方向。

Z字型运动模型通过周期性切换运动方向实现折线飞行:

$$\begin{cases} a_k = a_{k-1} + (-1)^k \cdot \Delta a \\ x(t) = x_k + v_x \cdot \cos(a_k) \cdot (t - t_k), t \in [t_k, t_k + T] \\ y(t) = y_k + v_y \cdot \sin(a_k) \cdot (t - t_k) \end{cases} \quad (10)$$

式中, T 为单段直线运动时长, Δa 为转弯角度, k 为当前运动段索引。该模型可生成无人机连续折线转向的机动轨迹。

通过上述匀速、匀加速、曲线、蛇形机动和Z字型机动等运动学模型的组合, 可以生成如图6(b)所示的多类高机动轨迹。与现有数据集中运动平缓的轨迹 (如图6(a)所示) 相比, 本文方法生成的轨迹在相同帧数内具有显著更高的空间覆盖范围和方向变化频率。

多类运动模型的组合能够有效拓宽生成轨迹在

速度、加速度、转向率等运动学状态空间的分布范围, 全面复现无人机在复杂空域中的潜在飞行模式。同时, 支持交互式自定义轨迹输入, 可将手动绘制路径转化为平滑参数曲线。这种预设模型与自定义轨迹相结合的混合方案, 能显著提升数据生成的灵活性。

通过将高纯净度目标样本按生成轨迹嵌入背景序列, 并施加距离相关的几何缩放、辐射衰减及光学退化效应, 系统生成了兼具物理真实性与场景多样性的红外合成序列。基于上述物理模型与轨迹生成方法, 本文共生成50个红外视频序列, 每个视频包含5至12个目标, 平均长度为500帧, 帧率为25FPS, 可直接用于后续跟踪算法的训练与评测。

2.3 数据集属性与统计分析

2.3.1 基本信息

如表1所示, 为全面展示当前反无人机跟踪数据集的整体概况, 表中除红外数据集外, 同时列出DUT-AntiUAV (Zhao 等, 2022) 和 UAV-Anti-UAV (Zhang 等, 2025) 等可见光反无人机数据集作为跨模态参考。与现有数据集相比, 所提数据集在数据量上相对精炼, 但具备一个突出特征: 平均SCR显著更高, 其公式为:

$$SCR = \frac{\mu_T - \mu_B}{\sigma_B} \quad (11)$$

式中, μ_T 表示目标区域灰度均值, μ_B 和 σ_B 分别表示背景区域灰度均值和标准差。SCR正值表示目标比背景亮, 负值反之, |SCR|越大目标与背景区分越明显。这源于合成数据在生成时能对目标与背景的红外辐射特性进行更清晰的定义与控制。更高的SCR意味着, 在本数据集中, 红外目标在图像层面与背景的区分更为清晰。这一统计特性直接关联到本数据集的设计初衷: 我们旨在构建一个检测难度适中, 但跟踪挑战更为突出的基准。更高的SCR可为检测器提供更干净的输入, 有效降低检测误差对跟踪性能评估的干扰, 进而更有针对性地对比不同算法在非线性、非平稳目标运动场景下的跟踪性能。

2.3.2 运动模式分析

为深入分析所提数据集中目标的运动特性, 我们采用Sun等人(2022)定义的相邻帧间的交并比 (Adjacent-frame Intersection over-Union, AIoU) 与相对位置切换频率 (Frequency of Relative

表1 本文提出的MIATrack与现有无人机跟踪数据集比较

Table 1 Statistical information of different UAV target tracking datasets

数据集	标注类型	图像类型	视频数量	图像数量	轨迹数量	平均SCR	mAIoU	mFRPS	多目标跟踪
DUT-AntiUAV	BBox	RGB	20	24804	20	-1.86	0.717	0	
UAV-Anti-UAV	BBox	RGB	1820	1057394	1820	0.13	0.753	0	
AntiUAV	BBox	Infrared	318	296901	318	0.81	0.778	0	
AntiUAV-410	BBox	Infrared	410	438397	410	0.73	0.729	0	
4th-AntiUAV	BBox	Infrared	200	151831	5262	0.61	0.758	0.003	✓
Ours	Mask&Bbox	Infrared	50	25000	486	1.42	0.297	0.011	✓

注:SCR正值表示目标比背景亮,负值表示目标比背景暗,其绝对值越大表示目标与背景区分越明显、有利于降低检测不确定性;mAIoU越低表示帧间运动幅度越大;mFRPS越高表示轨迹交互复杂度越高。

Position Switch, FRPS)作为量化指标评估数据集的机动特性和多目标交互特性。我们将MIATrack数据集与当前主流的红外无人机跟踪数据集进行对比,并引入以目标不规则运动著称的通用多目标跟踪数据集(如DanceTrack)作为跨域参照。

相邻帧间的交并比:该指标通过计算目标在相邻两帧中边界框(bounding-box, bbox)的重叠程度(即IoU)来衡量其运动幅度。通常情况下,较低的IoU值意味着目标的相对位移更大或运动更剧烈。对于包含 N 个目标,长度为 T 帧的视频序列,其平均相邻帧IoU(mean AIoU, mAIoU)的定义如下:

$$\text{mAIoU} = \frac{1}{N(T-1)} \sum_{i=1}^N \sum_{t=1}^{T-1} \text{IoU}(B_i^t, B_i^{t+1}) \quad (12)$$

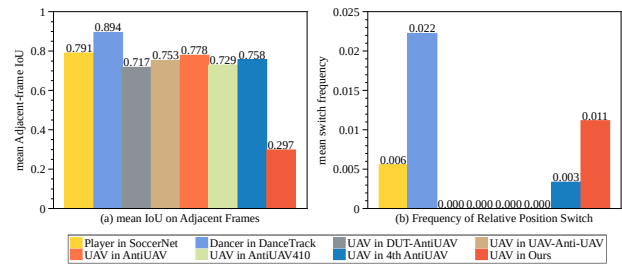
式中, B_i^t 表示第 i 个目标在第 t 帧上的边界框。

相对位置切换频率:该指标通过统计目标对在相邻帧发生左右或上下位置互换的频率,来衡量轨迹交错的复杂度。通常情况下,若目标群呈现匀速运动特征,其相对位置倾向于稳定,切换发生的概率较低。对于给定的视频序列,其平均相对位置切换频率(mean FRPS, mFRPS)定义如下:

$$\text{mFRPS} = \frac{\sum_{i=1}^N \sum_{j \neq i}^N \sum_{t=1}^{T-1} \text{sw}(B_i^t, B_j^t, B_i^{t+1}, B_j^{t+1})}{2N(T-1)(N-1)} \quad (13)$$

式中, $\text{sw}(\cdot)$ 是一个指示函数,若两个目标在相邻帧间的相对位置发生了互换,则 $\text{sw}(\cdot) = 1$,否则 $\text{sw}(\cdot) = 0$ 。该函数通过比较目标边界框中心点的相对位置来判断是否发生切换。此外,考虑到仅当目标存在重叠时,此类切换才会带来潜在的跟踪困难,因此计算时仅纳入有重叠的目标对。

如表1和图7所示,在机动性方面,以无人机



((a)相邻帧的IoU;(b)相对位置切换频率))

((a) mean IoU on Adjacent Frames; (b) frequency of relative position switch)

图7 运动信息统计

Fig.7 Motion statistics

为主要目标的红外数据集,其平均相邻帧mAIoU普遍低于以不规则运动著称的可见光数据集(如DanceTrack)。这主要源于红外无人机目标尺度与IoU之间的固有几何关系。对于行人跟踪,目标通常占据 100×300 像素,帧间移动5像素时IoU仍可达约0.87。但对于仅 10×20 像素的红外无人机小目标,同样移动5像素,IoU便骤降至0.29左右。具体而言,设目标边界框宽为 w 、高为 h ,帧间位移为 (dx, dy) 且目标仅发生平移而无形变,则相邻帧边界框的交并比可表示为:

$$\text{IoU} \approx \frac{(w - |dx|)(h - |dy|)}{2wh - (w - |dx|)(h - |dy|)} \quad (14)$$

以 8×15 像素目标为例,当帧间位移 $dx=3$ 像素、 $dy=4$ 像素时,计算可得 $\text{IoU} \approx 0.297$ 。这意味着本数据集mAIoU=0.297对应的平均帧间位移约为目标

尺寸的 37% - 60%, 属于物理上合理的连续运动, 而非轨迹跳变或视觉不连续。相比之下, 4th-AntiUAV 数据集 mAIoU=0.758, 其目标平均帧间位移仅为目标尺寸的 10% - 15%, 反映出其运动模式整体较为平缓。MIATrack 数据集 mAIoU 在所有对比数据集中最低, 印证了其包含大量高度非线性、非平稳的机动模式, 例如急转和频繁启停。其次, 在交互复杂度方面, “AntiUAV”、“AntiUAV410”、“DUT-AntiUAV”和“UAV-Anti-UAV”作为单目标跟踪数据集, 其 mFRPS 为 0; 而 MIATrack 的切换频率明显高于“4th-AntiUAV”这一最新红外多目标跟踪基准。这种高频的轨迹交错, 直接源于数据集中无人机运动模式的高度非线性所引发的复杂轨迹交互。MIATrack 数据集所呈现的丰富运动多样性, 对跟踪算法的鲁棒性提出了更严峻的挑战, 尤其是在处理不规则、非线性运动时, 算法必须具备更强大的运动建模与轨迹预测能力。

上述运动指标的定量差异与图 6 中的轨迹可视化结果高度一致, 4th-AntiUAV 数据集的大多数轨迹位移较小, 与其较高的 mAIoU (0.758) 和较低的 mFRPS (0.003) 相吻合; 而 MIATrack 数据集的轨迹覆盖范围广、方向变化多样, 则与其较低的 mAIoU (0.297) 和较高的 mFRPS (0.011) 形成对应。这一视觉与指标的双重对比, 共同验证了 MIATrack 在运动多样性和机动复杂度上的显著优势。

3 实验

3.1 实验设置

数据集设置: 本实验选取以无人机为目标且支持多目标跟踪任务的 4th-AntiUAV 数据集作为主要对比基准。由于该数据集仅公开了包含 200 个视频序列的训练集, 我们遵循 CenterTrack (Zhou 等, 2020) 的划分策略, 将每个视频序列均分为前后两段, 前段用于模型训练, 后段用于测试。对于 MIATrack 数据集, 我们依据 6:2:2 比例将 50 个视频划分为训练集 (30 个)、验证集 (10 个) 和测试集 (10 个)。

模型设置: 在检测任务中, 选用 YOLO26 (Jocher 等, 2025) 作为基础检测模型。为全面评估, 我们训练了其 n, s, m, l, x 五种规模的变体。所有模型均基于官方预训练权重, 使用默认配置 (如图像尺寸 640, 批次大小 64) 进行训练。

在多目标跟踪任务中, 我们选取检测性能最优的模型作为前端, 并在多种代表性跟踪器 (Bewley 等, 2016; Cao 等, 2023; Liu 等, 2025; Stadler and Beyrer, 2023; Zhang Y 等, 2022) 上进行基准测试。为保证评测的公平性与可比性, 所有对比跟踪器均输入完全一致的 YOLO26 检测结果, 即在同一数据集上使用同一检测模型、同一权重、同一置信度阈值与后处理参数生成的目标边界框序列, 从而消除检测性能差异对跟踪评估的干扰, 确保各跟踪器间的性能差异完全源于其运动建模与数据关联策略。

评估指标: 对于目标检测模型, 我们采用平均精度均值 (mean Average Precision, mAP) 作为衡量检测性能的主要指标, 特别是 mAP50:95, 该指标反映了模型在 IoU 阈值从 0.5 到 0.95 变化时的平均性

能。对于多目标跟踪任务, 我们采用高阶跟踪精度 (Higher Order Tracking Accuracy, HOTA) 作为主要评估指标, 以综合衡量检测与关联性能。我们还使

用关联准确率 (Association Accuracy, AssA) 和身份 F1 分数 (Identity F1, IDF1) 来专门评估关联的质量, 并使用检测精度 (Detection Accuracy, DetA) 和多目标跟踪精度 (Multiple Object Tracking Accuracy, MOTA) 来衡量跟踪器输出的检测质量。有关这些指标的详细定义, 请参阅文献 (Bernardin and Stiefelwagen, 2008; Luiten 等, 2021)。

实验设备: 实验在配备 2×NVIDIA GeForce RTX 4090 GPU、Intel Xeon Platinum 8352V CPU 及 240GB 内存的工作站上进行。

3.2 基准测试结果

3.2.1 目标检测

检测性能是 TBD 范式多目标跟踪的基础, 其可靠性直接影响后续运动估计与数据关联的有效性。为量化前文 (第 2.3 节) 所述数据集差异对检测环节的具体影响, 并确保后续跟踪评估建立在可靠的检测基线之上, 我们首先对检测模型在两个数据集上

的表现进行了对比分析。如表 2 所示, 所有 YOLO26 模型在 MIATrack 数据集上的性能均显著高于其在 4th-AntiUAV 数据集上的表现。尤其在衡量检测精度的 mAP50:95 指标上, 优势极为明显。

这一性能优势主要源于数据集的固有差异。如前文 (2.3 节) 所述, 尽管 4th-AntiUAV 数据集在规模上大于 MIATrack 数据集, 但 MIATrack 数据集凭借

表2 MIATrack数据集与4th-AntiUAV基准上的检测性能对比

Table 2 Detection performance comparison on the proposed MIATrack dataset and the 4th-AntiUAV benchmark

模型	4th-AntiUAV		MIATrack	
	mAP50	mAP50:95	mAP50	mAP50:95
YOLO26n	0.892	0.423	0.941	0.555
YOLO26s	0.895	0.426	0.949	0.560
YOLO26m	0.901	0.430	0.956	0.630
YOLO26l	0.902	0.432	0.948	0.557
YOLO26x	0.893	0.426	0.950	0.676
平均	0.897	0.427	0.949	0.596

注:加粗表示该数据集中的最佳性能。MIATrack在所有模型上的mAP50:95显著优于4th-AntiUAV。

更高的平均信杂比和精确标注,为模型提供了更优的监督信号,从而取得了显著更高的检测性能。在

面向机动目标的跟踪研究中,过于模糊或不可靠的检测会严重干扰对跟踪算法本身的评估。因此,我们通过技术手段确保了较高的检测基线,旨在将后续跟踪实验的评估焦点,从前端的“检测”环节,有效地转移至更核心的挑战,即在目标较为清晰的条件下,跟踪器能否应对复杂非线性运动下的稳健关联。

3.2.2 多目标跟踪

尽管得益于优质的检测输入,各跟踪器在 MIATrack 数据集上的 MOTA 指标(高度依赖检测精度)与 4th-AntiUAV 基准相当或更高,但在评估综合跟踪性能的 HOTA 以及核心关联能力的指标(IDF1, AssA)上,所有跟踪器的性能均出现普遍且显著的下滑,如表3所示。这一“检测强、关联弱”

的鲜明对比清晰地表明,在 MIATrack 数据集上,获得良好跟踪性能的瓶颈已非检测精度,而在于运动估计和关联环节本身。该结果印证了本数据集的核心设计,即通过引入高度不确定、非线性的运动

表3 主流跟踪器在MIATrack数据集与4th-AntiUAV基准上的性能比较

Table 3 Benchmarking of modern trackers on the proposed MIATrack dataset versus 4th-AntiUAV

跟踪器	4th-AntiUAV					MIATrack				
	HOTA	DetA	AssA	MOTA	IDF1	HOTA	DetA	AssA	MOTA	IDF1
SORT	38.821	37.524	42.439	37.053	46.801	33.546	35.849	32.223	46.501	39.788
ByteTrack	45.648	44.602	48.440	40.379	57.191	40.921	43.220	39.688	58.557	49.901
OC-SORT	41.015	43.386	41.023	41.289	49.304	22.573	32.157	18.042	35.318	23.666
SparseTrack	43.131	38.860	49.702	38.368	54.617	31.912	38.559	26.684	46.455	37.236
ImprAsso	49.293	43.441	58.231	43.020	64.216	44.535	55.747	36.439	68.856	50.732
HybridSORT	46.314	45.429	49.553	44.011	58.023	42.040	48.493	38.810	65.929	49.629
平均指标	44.037	42.207	48.231	40.687	55.025	35.921	42.338	31.981	53.603	41.825

注:加粗表示该指标的最优值。相比4th-AntiUAV,各跟踪器在MIATrack上的HOTA、AssA、IDF1均明显下降,而更偏向检测的MOTA指标基本持平甚至更高,反映出高机动场景下关联性能是主要瓶颈。

表4 各跟踪器在不同运动模式下的性能比较

Table 4 Performance comparison of trackers across different motion patterns

跟踪模型	SORT		ByteTrack		OC-SORT		SparseTrack		ImprAsso		HybridSORT	
	HOTA	AssA	HOTA	AssA	HOTA	AssA	HOTA	AssA	HOTA	AssA	HOTA	AssA
匀速运动	64.13	75.44	61.33	71.30	36.84	29.46	60.96	70.05	63.39	74.16	65.50	75.06
匀加速运动	63.82	74.88	59.44	68.22	43.10	34.44	59.81	67.80	60.96	70.05	61.10	72.09
蛇形机动	4.00	1.24	6.68	0.85	3.70	1.09	3.48	0.85	5.09	1.48	11.47	4.71
Z字型机动	54.97	60.38	58.34	66.22	50.75	50.11	55.97	57.63	59.82	67.80	63.67	73.02
曲线运动	57.99	63.59	58.17	68.99	48.57	43.93	60.78	66.05	61.77	69.34	61.26	69.21

注:加粗表示该运动模式下该指标的最优值。

模式,以模拟真实无人机飞行中未被现有数据集充分刻画动态复杂性。

在传统数据集中,平缓、可预测的目标运动使得基于简单线性假设的运动和关联模型能够有效工作。然而,当面对 MIATrack 数据集中不规则、高机动的目标时,这些模型的基础假设被打破,导致关联性能出现系统性衰减。基准测试结果表明,在 MIA-Track 数据集所呈现的高机动、非线性运动场景下,运动估计和关联已成为限制跟踪性能提升的主要瓶颈。这一结果揭示了当前跟踪算法在面对复杂目标动态时的普遍局限性。

3.2.3 不同运动模式下的性能分析

为进一步分析各跟踪器在不同机动场景下的性能差异,本小节从 MIATrack 测试集中为匀速、匀加速、曲线、Z 字型和蛇形机动五种运动模式各挑选了 30 条轨迹进行分组评测。表 4 列出了各跟踪器在不同运动模式下的 HOTA 和 AssA 结果。从表 4 可以看出各跟踪器性能在不同运动模式下的明显分化。HybridSORT 在四种模式中取得 HOTA 最优或接近最优,综合适应性最佳。SORT 和 ImprAsso 算法在基础运动上表现稳健($HOTA \approx 61-64$, $AssA \approx 70-74$),曲线和 Z 字型下略有下降。ByteTrack 和 SparseTrack 在各类非蛇形运动下性能稳定,与前列差距较小。Z 字型对各跟踪器的挑战有限,因其本质上由匀速直线段加间歇转弯构成,大部分时间内目标仍处于线性运动状态。OC-SORT 在所有运动模式下均表现最弱,且其简单运动性能反而低于 Z 字型等复杂运动,因其观测中心策略专为非线性运动修正设计,在线性运动下过度校正了本已稳定的卡尔曼预测,反而不利于关联。

蛇形机动对所有跟踪器构成断崖式挑战,HOTA 全部低于 12%。其根本原因在于蛇形运动的方向每帧连续变化,运动方程中持续的周期振荡项叠加在直线运动上,导致速度和加速度每帧都在变,严重破坏了卡尔曼滤波器的线性运动假设。整体结果进一步验证了本文核心结论:复杂非线性机动条件下,数据关联是当前红外无人机多目标跟踪的主要性能瓶颈。

4 结论与展望

本文提出了一个新的面向高机动场景的合成反

无人机多目标跟踪红外数据集 MIATrack。MIA-Track 数据集中的目标在复杂背景下呈现多样的非线性运动模式,同时确保高信杂比,从而将评估重点从检测能力转向复杂动态下的鲁棒关联。现有反无人机数据集存在明显倾向性,普遍侧重低信杂比检测场景,而忽略了对无人机多样化机动行为的刻画。复杂运动建模对反无人机跟踪系统研发尤为关键,MIATrack 可为该领域研究提供专用基准数据集。我们在 MIATrack 上评估了当前具有代表性的多目标跟踪器。结果表明在统一检测模型下,检测性能稳定优异,但关联精度显著下降,验证了本数据集可有效凸显机动无人机跟踪中运动关联建模是核心性能瓶颈。此外,作为未来工作,我们将从以下两个方向推进域泛化研究:(1)将真实红外数据集(如 4th-AntiUAV)训练的检测与跟踪模型直接部署至 MIA-Track 测试集,定量评估真实数据向仿真高机动场景迁移时的性能衰减规律;(2)利用 MIATrack 的高机动轨迹对现有真实数据集进行运动增强,探索以仿真数据扩充真实数据分布的域适应方案。MIATrack 所包含的丰富非线性机动轨迹,有望为红外反无人机的跨域泛化研究提供有价值的测试平台。

参考文献(References)

- Bernardin K and Stiefelhagen R. 2008. Evaluating multiple object tracking performance: the CLEAR MOT metrics. EURASIP Journal on Image and Video Processing, 2008 (1): 246309 [DOI: 10.1155/2008/246309].
- Bewley A, Ge Z, Ott L, Ramos F and Upcroft B. 2016. Simple online and realtime tracking//Proceedings of 2016 IEEE International Conference on Image Processing (ICIP). IEEE: 3464-3468 [DOI: 10.1109/ICIP.2016.7533003].
- Cao J, Pang J, Weng X, Khiroudkar R and Kitani K. 2023. Observation-Centric SORT: Rethinking SORT for Robust Multi-Object Tracking//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE: 9686-9696 [DOI: 10.1109/CVPR52729.2023.00934].
- Dai Y, Li X, Zhou F, Qian Y, Chen Y and Yang J. 2023. One-Stage Cascade Refinement Networks for Infrared Small Target Detection. IEEE Transactions on Geoscience and Remote Sensing, 61: 1-17 [DOI: 10.1109/TGRS.2023.3243062].
- Dai Y, Wu Y, Zhou F and Barnard K. 2021. Asymmetric contextual modulation for infrared small target detection//Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa, HI, USA: IEEE: 949-958 [DOI: 10.1109/

- WACV48630.2021.00099].
- Fu R G, Fan H Q, Zhu Y F, Hui B W, Zhang Z L, Zhong P, et al. 2022. A dataset for infrared time-sensitive target detection and tracking for air-ground application. *China Scientific Data*(2): 7 (傅瑞罡, 范红旗, 朱永锋, 回丙伟, 张志龙, 钟平, 等. 2022. 面向空地应用的红外时敏目标检测跟踪数据集. *中国科学数据: 中英文网络版*(2): 7 [DOI: 10.11922/11-6035.csd.2021.0085.zh])
- Huang B, Li J, Chen J, Wang G, Zhao J and Xu T. 2024. Anti-UAV410: A Thermal Infrared Benchmark and Customized Scheme for Tracking Drones in the Wild. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(5): 2852-2865 [DOI: 10.1109/TPAMI.2023.3335338].
- Hui B W, Song Z Y, Fan H Q, Zhong P, Hu W D, Zhang X F, et al. 2020. A dataset for infrared detection and tracking of dim-small aircraft targets under ground / air background. *China Scientific Data*, 5(3): 291-302 (回丙伟, 宋志勇, 范红旗, 钟平, 胡卫东, 张晓峰, 等. 2020. 地/空背景下红外图像弱小飞机目标检测跟踪数据集. *中国科学数据*, 5(3): 291-302) [DOI: 10.11922/csdata.2019.0074.zh]
- Jiang N, Wang K, Peng X, Yu X, Wang Q, Xing J, et al. 2023. Anti-UAV: a large-scale benchmark for vision-based UAV tracking. *IEEE Transactions on Multimedia*, 25: 486-500 [DOI: 10.1109/TMM.2021.3128047].
- Jocher G, Qiu J and Chaurasia A. 2025. Ultralytics YOLO. Zenodo [DOI: 10.5281/zenodo.15164670].
- Kou R K, Wang C P, Luo Y, Zhang Y, Xu Z L, Peng Z M, et al. 2024. Multiscale small-target detection techniques in single-frame infrared images: a review. *Journal of Image and Graphics*, 29(9): 2625-2649 (寇人可, 王春平, 罗迎, 张勇, 徐泽龙, 彭真明, 等. 2024. 单帧红外图像多尺度小目标检测技术综述. *中国图象图形学报*, 29(9): 2625-2649) [DOI: 10.11834/jig.230788]
- Kuhn H W. 1955. The hungarian method for the assignment problem. *Nav. Res. Logist. Q.*, 2(1-2): 83-97 [DOI: 10.1002/nav.3800020109].
- Li B, Xiao C, Wang L, Wang Y, Lin Z, Li M, et al. 2023. Dense nested attention network for infrared small target detection. *IEEE Transactions on Image Processing*, 32: 1745-1758 [DOI: 10.1109/TIP.2022.3199107].
- Liu Y, Ma Y P, Yang J W, Wang Z M, Aliha Aersi, Cao H Z, et al. 2026. Infrared dim and small moving target detection in recent years: an overview. *Journal of Image and Graphics*, 31(6): 1847-1874 (刘雨茜, 马亚朋, 杨建伟, 王子铭, 阿尔斯·阿力哈, 曹华钊, 等. 2026. 红外空中弱小运动目标检测方法研究综述. *中国图象图形学报*, 31(6): 1847-1874) [DOI: 10.11834/jig.260109]
- Liu Z, Wang X, Wang C, Liu W and Bai X. 2025. SparseTrack: Multi-Object Tracking by Performing Scene Decomposition Based on Pseudo-Depth. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(5): 4870-4882 [DOI: 10.1109/TCSVT.2024.3524670].
- Luiten J, Ošep A, Dendorfer P, Torr P, Geiger A, Leal-Taixé L, et al. 2021. HOTA: A Higher Order Metric for Evaluating Multi-object Tracking. *International Journal of Computer Vision*, 129(2): 548-578 [DOI: 10.1007/s11263-020-01375-2].
- Stadler D and Beyerer J. 2023. An Improved Association Pipeline for Multi-Person Tracking//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Vancouver, BC, Canada: IEEE: 3170-3179 [DOI: 10.1109/CVPRW59228.2023.00319].
- Sun H, Bai J, Yang F and Bai X. 2023. Receptive-Field and Direction Induced Attention Network for Infrared Dim Small Target Detection With a Large-Scale Dataset IRDST. *IEEE IEEE Transactions on Geoscience and Remote Sensing*, 61.
- Sun P, Cao J, Jiang Y, Yuan Z, Bai S, Kitani K, et al. 2022. Dance-Track: Multi-Object Tracking in Uniform Appearance and Diverse Motion//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE: 20961-20970 [DOI: 10.1109/CVPR52688.2022.02032].
- Wang H, Zhou L and Wang L. 2019. Miss Detection vs. False Alarm: Adversarial Learning for Small Object Segmentation in Infrared Images//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). 8508-8517 [DOI: 10.1109/ICCV.2019.00860].
- Wojke N, Bewley A and Paulus D. 2017. Simple online and realtime tracking with a deep association metric//2017 IEEE International Conference on Image Processing (ICIP). IEEE: 3645-3649 [DOI: 10.1109/ICIP.2017.8296962].
- Zhang C, Liu L, Zhang Z, Wang Yong, Wen H, Zhou X, et al. 2025. How Far are Modern Trackers from UAV-Anti-UAV? A Million-Scale Benchmark and New Baseline. *arXiv* [DOI: 10.48550/arXiv.2512.07385].
- Zhang M, Zhang R, Yang Y, Bai H, Zhang J and Guo J. 2022. ISNet: shape matters for infrared small target detection//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE: 867-876 [DOI: 10.1109/CVPR52688.2022.00095].
- Zhang Y M, Bao W T, Xiao Q, Yang Y, Wan W G, Luo Y T, et al. 2026. Selective attention-based for infrared small target detection. *Journal of Image and Graphics*, 31(3): 797-810 (张迎梅, 鲍王涛, 肖沁, 杨勇, 万伟国, 罗亦韬, 等. 2026. 基于选择性注意力的红外小目标检测. *中国图象图形学报*, 31(3): 797-810) [DOI: 10.11834/jig.250313]
- Zhang Y, Sun P, Jiang Y, Yu D, Weng F, Yuan Z, et al. 2022. ByteTrack: Multi-object Tracking by Associating Every Detection Box//Proceedings of European Conference on Computer Vision (ECCV). Springer Nature Switzerland: 1-21 [DOI: 10.1007/978-3-031-20047-2_1].
- Zhao J, Zhang J, Li D and Wang D. 2022. Vision-based anti-UAV

detection and tracking. IEEE Transactions on Intelligent Transportation Systems, 23(12): 25323-25334 [DOI: 10.1109/ITITS.2022.3177627].

Zhou X, Koltun V and Krähenbühl P. 2020. Tracking objects as points// Proceedings of Computer vision - ECCV 2020. Cham: Springer International Publishing: 474-490 [DOI: 10.1007/978-3-030-58548-8_28].

作者简介

宋健,男,博士研究生,研究方向为状态估计,多目标跟踪。
E-mail: songj9507@outlook.com

付强,男,副教授,主要研究方向为模式识别。E-mail: Fu_Qiang@aeu.edu.cn

杨皓杰,男,硕士研究生,主要研究方向为目标检测跟踪。E-mail: extantair@outlook.com

龙昱承,男,硕士研究生,主要研究方向为状态估计,3D多目标跟踪。E-mail: longyc0010@163.com

李战武,男,教授,主要研究方向为火力控制与指挥。E-mail: afeulzw@189.cn

寇人可,男,博士,通讯作者,研究方向为红外小目标检测与跟踪。E-mail: krkoptics@163.com