文章编号:1671-4598(2026)09-0282-08 DOI:10.16526/j.cnki.11-4762/tp.2026.09.035 中图分类号:TP391 文献标识码:A

面向复杂自然环境的红外目标跟踪数据集构建与评测

王 刚,董雨函,王晟贤
(中国电子科技集团公司 光电研究院,天津 300308)

摘要: 红外目标跟踪是态势感知、无人系统、安防监控等任务中的重要基础技术,高质量红外目标跟踪数据集是支撑算法训练、性能评测与技术迭代的关键基础,其数据规模、场景多样性和标注质量直接影响跟踪算法的性能上限;然而,现有红外目标检测数据集大多面向通用场景,存在特定目标类别覆盖不足、背景环境单一以及复杂干扰因素缺失等问题,难以满足复杂动态场景下算法鲁棒性评估的需求;为此,构建了面向复杂自然环境的红外目标跟踪数据集 WAR-TIR;该数据集采用实地采集、公开数据筛选与物理建模仿真相结合的方式,包含 80 条视频序列,涵盖车辆、无人机等 8 类典型对象,覆盖田野、荒漠、山地及开放场地等 5 类背景,并重点引入了重度雾霾、强光干扰、动态模糊及热辐射伪影等 6 类典型干扰;系统分析了数据集的构建流程、标注规范及统计特性,并选取现有红外目标跟踪算法建立了基准测试,测试结果表明,上述算法在 WAR-TIR 上的最高成功率和精确度为 0.642 和 0.695,证明 WAR-TIR 数据集具有较高挑战性,能够有效评估跟踪算法在复杂环境下的抗干扰能力。

关键词: 复杂环境;红外目标跟踪;数据集构建;目标跟踪评测

Construction and Evaluation of Infrared Target Tracking Dataset for Complex Natural Environments

WANG Gang, DONG Yuhan, WANG Shengxian
(Academy of Opto-Electronics, China Electronics Technology Group Corporation (AOE CETC), Tianjin 300308, China)

Abstract: Infrared target tracking is an important fundamental technology for situational awareness, unmanned systems, security surveillance, and other tasks. High-quality infrared target tracking datasets provide an essential foundation for supporting algorithm training, performance evaluation, and technological advancement, while the scale of data, diversity of scenarios, and annotation quality directly affect the performance ceiling of tracking algorithms. However, most existing infrared target tracking datasets are designed for general scenarios, exhibiting issues such as insufficient coverage of specific target categories, monotonous background environments, and a lack of complex interference factors, making it difficult to meet the requirements of robustness evaluation in complex dynamic environments. To address these limitations, an infrared target tracking dataset for complex natural environments, WAR-TIR, is constructed. The dataset is developed by combining field data acquisition, public data screening with physical modeling simulation. It contains 80 video sequences, covering eight types of typical objects such as vehicles and unmanned aerial vehicles, and spans five types of backgrounds, including fields, deserts, mountainous areas, and open areas. Additionally, it specifically incorporates six typical interferences, including dense haze, strong-light interference, motion blur, and thermal radiation artifacts. Finally, the dataset' construction process, annotation protocol, and statistical characteristics are systematically analyzed, selecting existing infrared target tracking algorithms to establish a benchmark. Experimental results show that the highest success rate and precision of the above algorithms on the WAR-TIR are only 0.642 and 0.695, respectively, demonstrating a significant challenge of the proposed dataset, and it can effectively evaluate the anti-interference capabilities of infrared target tracking algorithms in complex environments.

Keywords: complex environments; infrared target tracking; dataset construction; target tracking performance evaluation

收稿日期:2026-07-13; **修回日期:**2026-08-14。

作者简介:王 刚(1982-),男,硕士,高级工程师。

引用格式:王 刚,董雨函,王晟贤.面向复杂自然环境的红外目标跟踪数据集构建与评测[J].计算机测量与控制,2026,34(9):282-289.

0　引言

红外目标跟踪在态势感知、无人系统、安防监控等任务中扮演重要角色。相比可见光成像，红外成像利用热辐射差异，具备全天时、全天候和隐蔽性优势，支持夜间及复杂气象条件下对目标的稳定探测与跟踪。然而，目前公开可用的红外目标跟踪数据集（如 OSU、ASL-TID、LTIR、VOT-TIR、PTB-TIR、LSOTB-TIR 等）大多面向行人、车辆等通用目标，目标类别覆盖有限；虽然近年出现了针对无人机的专用数据集，如 CST Anti-UAV[1]和 Anti-UAV410[2]，以及多模态融合跟踪基准（如 LasHeR[3]和 VTUAV[4]），但它们的背景环境和干扰因素仍相对简单，缺乏雾霾、火光、爆炸等典型干扰类型，难以全面评估算法在复杂真实场景下的鲁棒性。针对上述不足，本文构建了 WAR-TIR 数据集，该数据集综合采用实地采集、公开数据收集与建模仿真相结合的方式，涵盖地面与空中 8 类典型目标（含车辆、无人机等），覆盖田野、沙漠、山地、靶场等多种背景，并引入重度雾霾、强光干扰、动态模糊及热辐射伪影等 6 类干扰，形成高挑战性测试平台。与现有同类数据集相比，WAR-TIR 在目标类型多样性、背景复杂度和干扰因素丰富性方面具有明显优势，能够更充分地评估跟踪算法在复杂环境下的抗干扰能力和泛化性能，为相关研究提供数据支撑和基准参考。

1　红外目标跟踪数据集构建

近年来，在红外目标跟踪技术的研究进程中，精准验证算法在实际应用场景中的适用性与鲁棒性是推动该领域发展的关键环节。现有 LSOTB-TIR[5]、PTB-TIR[6]等数据集虽然为红外目标跟踪领域提供了重要基准测试平台，但主要面向通用场景下的目标跟踪任务，缺乏对复杂自然环境和多样化目标的针对性覆盖，难以全面描述高复杂度场景。尽管近年来已有研究者针对特定场景构建了更具挑战性的基准，如面向夜间红外视频跟踪的 NV-SID[7]数据集、面向热红外行人多目标跟踪的 TP-MOT[8]基准，以及面向多无人机多模态场景的 M3OT[9]数据集，但这些数据集在目标类型覆盖和干扰因素多样性方面仍存在局限。因此，本文构建了红外目标跟踪数据集 WAR-TIR，旨在弥补现有数据集的不足，提供一个更加贴近实际国防应用场景的测试平台。WAR-TIR 数据集涵盖了多种典型地面目标，引入了雾霾、火光、光照、爆炸等干扰因素，模拟真实环境中的高难度场景。通过这一数据集，能够更加全面地评估红外目标跟踪算法在复杂环境中的性能表现，特别是在抗干扰能力及模型泛化性方面的改进效果。本节将从数据集构建思路、数据采集、数据标注、数据统计分类及数据集评估五个方面，对 WAR-TIR 进行详细分析。

1.1　数据集构建思路

现有红外目标跟踪数据集大多面向通用场景，缺乏对复杂干扰因素和多样化运动目标的覆盖，往往难以满足复杂自然环境下的实际需求。具体而言，存在以下三方面不足。

1）复杂环境的干扰因素模拟不足[10]：环境复杂多变，雾霾、光照等干扰因素频繁出现，而现有数据集往往缺乏对这些复杂干扰因素的充分模拟，导致算法在高干扰条件下的性能难以得到有效评估；

2）目标类型的单一性[11]：现有数据集主要针对行人、汽车等常见目标进行跟踪任务，缺乏对典型地面目标的覆盖，难以满足对多样化目标跟踪的需求；

3）干扰类型的单一性[12]：现有数据集在干扰类型的设计上较为单一，通常为背景噪声或类内干扰，缺乏真实对抗环境中火光、爆炸等干扰因素，难以全面评估算法在复杂自然环境中的目标检测和跟踪能力。

基于此，WAR-TIR 数据集的构建从当前实际应用需求出发，通过引入环境复杂干扰因素和多样化目标，确保算法能够在高复杂度场景下进行有效评估。具体而言，该数据集的构建将从以下几个方面展开。

1）多样化背景环境：不仅限于单一的地形或气候条件，WAR-TIR 数据集涵盖了田野、沙漠、山脉、靶场等多种典型场景，确保算法能够适应不同背景下的跟踪任务；

2）多样化目标的覆盖：WAR-TIR 数据集包含各类地面目标，还涵盖了多类空中飞行器目标（背景仍为对地场景），确保算法能够适应不同类型目标的跟踪任务；

3）复杂干扰因素的引入：WAR-TIR 数据集引入雾霾、火光、光照、爆炸等复杂干扰因素，有效模拟高复杂度场景，进一步提升数据集挑战性。

为全面评估红外目标跟踪算法性能，WAR-TIR 数据集在设计时特别考虑了样本的难度差异，确保数据集能够覆盖不同复杂度的跟踪任务。具体而言，数据集的难度梯度设计主要体现在以下三个方面。

1）目标尺寸差异：相同的目标在图像中的尺寸不同，这种差异通常由目标与红外摄像机的距离变化引起，距离较远时目标在图像中呈现为小尺寸，距离较近时则呈现为大尺寸。通过引入尺寸差异，数据集能够模拟真实场景中目标距离的动态变化，确保算法能够适应不同尺寸目标的跟踪任务。

2）目标运动速度差异：数据集中目标的运动速度差异较大，有低速移动的目标（如车辆），也有高速移动的目标（如无人机等飞行器），以确保算法能够适应

不同速度目标的跟踪任务，特别是高速目标跟踪场景。

3）干扰强度差异：数据集中干扰因素的强度差异较大，有轻度干扰场景（如少量雾霾、烟雾等），也有重度干扰场景（如爆炸、火光等），以确保算法能够在不同干扰强度下保持稳定的跟踪性能。

综上所述，WAR-TIR 数据集构建思路如图 1 所示，该数据集从目标类别、干扰因素以及背景环境的多样性出发，确保了目标尺度、目标运动速度和

干扰强度等样本难度的差异，为红外目标跟踪算法在复杂自然环境下的应用提供了一个全面且具有挑战性的评估平台。

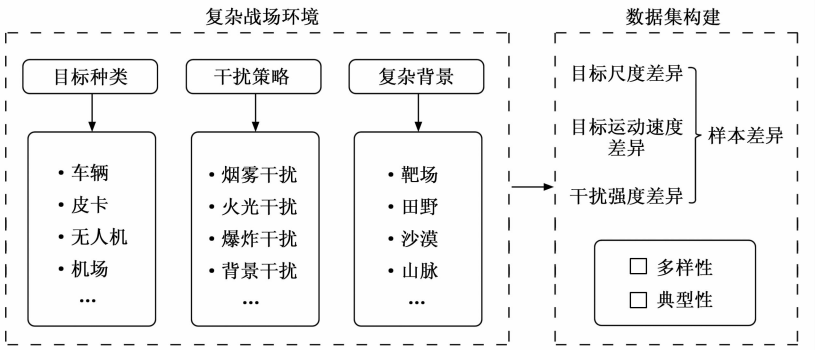

图 1 红外专用应用数据集构建框图

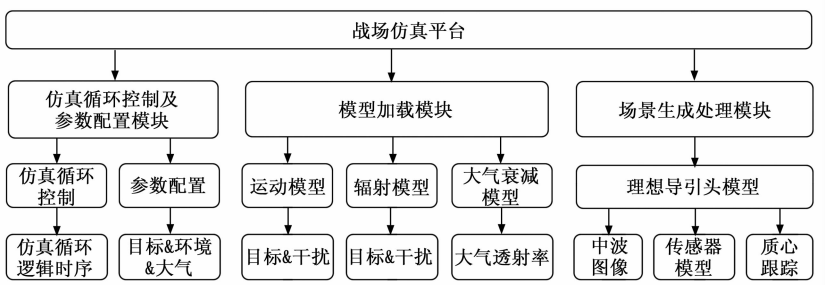

图 2 场景仿真平台组成

1.2 数据采集

由于目标跟踪数据通常处于非公开状态，获取难度较大，因此本文采用红外热像仪实地拍摄、网络搜索公开数据以及复杂干扰条件下建模仿真相结合的方式构建数据集。红外热像仪实地拍摄数据来自 2016 年在珠海航展现场拍摄的内容，共计 4 条序列；网络搜索公开数据中的跟踪目标为大型地面移动装备，共计 4 条序列；建模仿真数据由 Microsoft Visual Studio 开发的场景仿真平台生成得到。

使用的仿真平台组成如图 2 所示，基于红外理想导引头的视角进行设计，模拟红外图像的关键参数为中波红外波段 3～5 μm，图像分辨率 1 080×864 像素。该平台通过数字建模与实时红外场景渲染技术构建，完成红外目标模型、红外干扰模型和背景模型的搭建。仿真系统运行机理如下。

1）仿真循环控制及参数配置模块：进行仿真循环控制及参数配置，仿真循环控制主要是对平台的仿真循环进行控制，参数配置主要完成目标参数、环境参数与大气参数的配置；

2）模型加载模块：对仿真所需要的各个模块进行运行和调度，实时解算仿真循环所需的实时仿真数据，包括运动数据、辐射数据和大气数据；

3）场景生成处理模块：场景生成模块根据初始参数配置、模型、循环控制建构仿真方案，根据仿真数据、红外目标模型、红外场景模型和干扰模型，采用开放场景图（OSG，open scene graph）作为渲染后端，结合 OpenGL 并行渲染技术进行实时场景渲染；场景处理模块模拟红外导引头和红外跟踪系统的光电效应图像，接收其余模块的数据，包括初始态势、参数配置数据等，生成和处理实时导引头和红外跟踪序列。

完成建模仿真数据生成后，对所有数据进行汇总，最终得到 WAR-TIR 测试集的序列信息以及帧数信息，如表 1 所示。

表 1 WAR-TIR 测试集内容

序列数	最大帧数	最小帧数	平均帧数	总帧数	帧率
80	1 015	43	400	31K	30 FPS

1.3 数据标注

本文将视频序列分解为图像帧，并利用开源软件 LabelImg 对每帧图像进行标注，标注界面如图 3 所示，标注生成每帧图像对应的目标框信息并以 xml 格式存储。随后，通过自定义脚本解析 xml 文件，提取目标框信息，并按照顺序存储为统一的 txt 文件格式，该标注

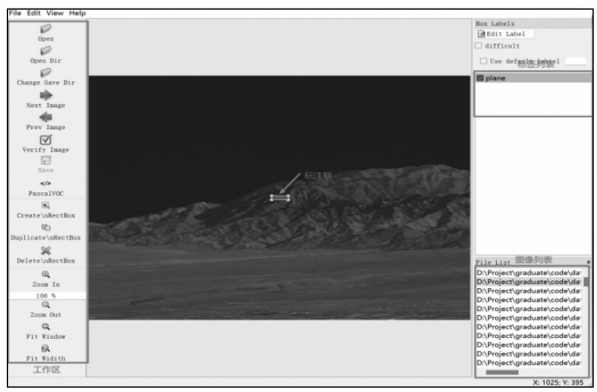

图 3 LabelImg 标注界面

格式与 LSOTB-TIR 数据集的标注格式保持一致，以确保数据格式的统一性和兼容性。在图像存储与标注格式的组织结构方面，顶层目录为数据集根目录 WAR-TIR，每个子目录对应一个独立的图像序列及其标注文件。子目录中包含一个名为 img 的文件夹用于存储图像帧，图像文件格式为 jpg。每个子目录中包含一个名为 groundtruth _ rect. txt 的标注文件，用于记录当前序列中所有帧的目标框标注信息，标注内容包括左上角坐标及宽高信息。

1.4　数据统计分类

WAR-TIR 数据集包括多种目标类型和复杂背景下干扰场景，图 4 展示了部分典型示例。具体而言，数据集涵盖了各类典型空中目标与地面目标，背景环境则涉及田野、沙漠、山脉等不同地貌。此外，仿真场景构建过程中引入了雾霾干扰、爆炸干扰、火光干扰、光照干扰等多种复杂干扰因素，力求精准模拟实际环境所具有的高难度与挑战性。这些典型示例从目标类型的多元性、背景环境的多变性以及干扰因素的复杂性等多个维度，全面展现了数据集的多样性和复杂性。这些特点为评估红外目标跟踪算法在精准性、抗干扰性以及应用等方面的性能，提供了丰富且具有挑战性的测试条件。

为进一步分析测试集的数据分布情况，对目标类别、背景类别及自然干扰类别进行统计分类，具体信息如图 5 所示。数据集包含 8 类典型目标，其中地面目标占据较大比例；背景类别包括沙漠、靶场等 5 类常见场景，其中田野、靶场和沙漠的分布比例相对较高；干扰类别涉及烟雾、火光等 6 类复杂干扰条件，爆炸、烟雾及火光干扰主要针对地面目标的干扰场景。数据集中多样化的分布设计能够有效确保算法在不同目标类型、背景环境和干扰条件下的鲁棒性评估。

为了更好地了解 WAR-TIR 数据集的特性，从位置信息和尺度信息两个方面对数据集进行了统计。

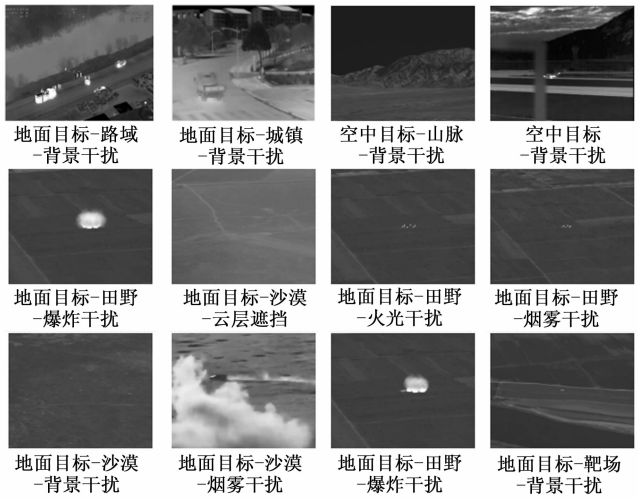

图 4　WAR-TIR 数据集各个类别典型示例

位置分布统计如图 6（a）所示，散点表示目标中心点在图像中的位置分布。结果表明，数据集中目标的位置主要集中在图像的左下角区域和中央区域，同时在水平方向和垂直方向上均表现出较大的波动性，反映了目标位置在场景中的高度多样性。这种分布特性对红外目标跟踪算法提出了更高要求，在目标位置变化较大的场景中，算法需要具备更强的目标定位能力。尺度分布分析如图 6（b）所示，其中散点的横坐标表示目标宽度与图像宽度的比值，纵坐标表示目标高度与图像高度的比值。数据集中小目标图像在整个序列中的占比较大，表明该数据集包含大量小目标场景。这种分布特性要求红外目标跟踪算法在小目标跟踪场景中具备更强的特征提取能力和抗干扰性能，以确保在低信噪比和低分辨率条件下，依然能够实现准确的目标跟踪。

1.5　数据集评估

目前在红外目标跟踪领域，有以下 6 个广泛使用的

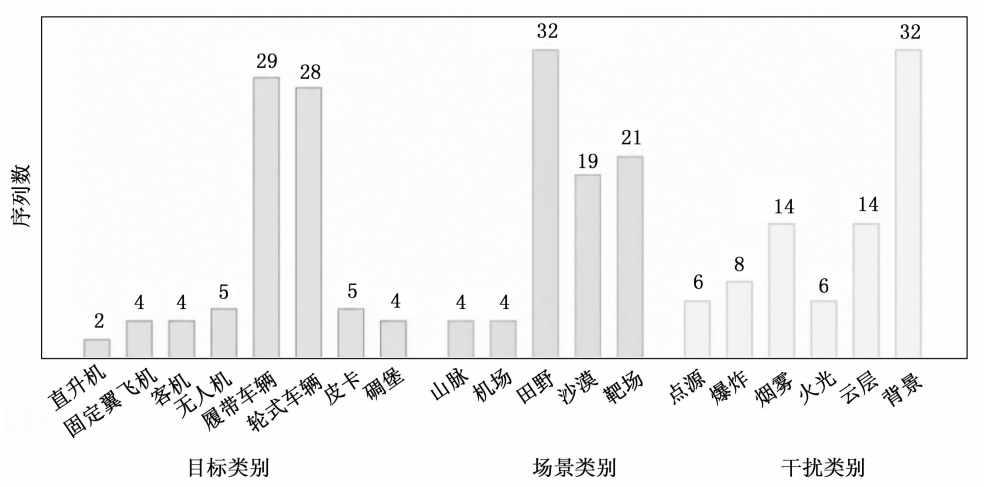

图 5　数据统计分类图

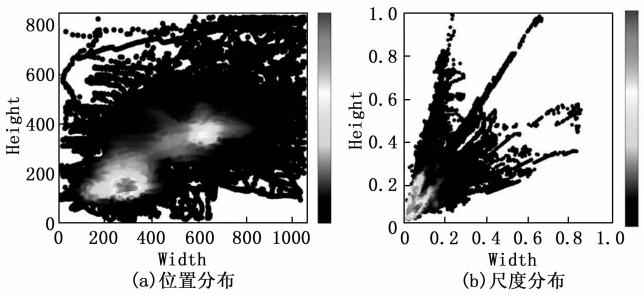

图 6 数据集中目标位置和尺度统计分布

数据集基准。

1) OSU[13]：是红外和可见光图像融合数据集，可用于红外目标跟踪。该数据集仅包含行人对象，所有六个视频都是由静态背景下的低分辨率热红外摄像机捕获的；

2) ASL-TID[14]：包含 9 个户外场景的视频序列，这些序列在不同的视角和温度条件下拍摄，具备高分辨率目标对象和低分辨率目标对象。此外，序列包含一些复杂的背景环境和遮挡；

3) LTIR[15]：第一个标准的红外对象跟踪基准，包含 20 个序列，6 个对象类和一个评估工具包，被红外目标跟踪竞赛 VOT-TIR15 所采用；

4) VOT-TIR16[16]：VOT-TIR15 [17]的扩展版本，包含 25 个序列和 8 个对象类，比 VOT-TIR15 更具挑战性，有六个挑战子集，可用于评估跟踪器的特定属性；

5) PTB-TIR[6]：侧重于红外行人跟踪，包含 60 个序列和 9 个属性挑战。数据集收集于多种设备，缺乏对不同场景属性的划分；

6) LSOTB-TIR[5]：共有 120 条测试序列，包含一个短期跟踪子集和一个长期跟踪子集，共包括 22 种目标类别及 12 种挑战属性。此外，近年来也有学者从融合跟踪技术路线出发，对红外与可见光融合目标跟踪方法进行了系统梳理与综述分析[18]。

如表 2 所示，本文提出的 WAR-TIR 数据集帧数分布均衡，最大帧数为 1 015 帧，最小帧数为 43 帧，总帧数达到 31 K。此外，WAR-TIR 针对复杂环境设计，适用于国防领域特定任务，包含直升机、无人机、履带车辆等 8 种不同运动特性的目标和点源、爆炸、烟雾等 6 种干扰种类，显著增强了测试集的挑战性。相比之下，OSU 和 PTB-TIR 测试集仅包含行人这一单一目标类型，LSOTB-TIR 仅包含 3 类空中目标与地面移动目标，在多样性和复杂性方面存在明显不足。同时，WAR-TIR 涵盖了机场、靶场等 5 种复杂场景，进一步提升了其在特定应用领域中的相关性和多样性。综上所述，WAR-TIR 测试集在序列数量、目标多样性、干扰种类和场景复杂性等方面具备优势，能够为面向国防领域的红外目标跟踪算法鲁棒性评估提供更全面的测试环境。

2 实验验证及结果分析

为验证所构建红外目标跟踪数据集的有效性、挑战性及工程应用价值，本文基于统一的训练与测试协议开展基准测试实验。选取当前主流的目标跟踪算法作为基线模型，在所构建数据集上进行训练与评估，并采用多种评价指标对其跟踪性能进行综合分析。通过比较不同算法在复杂背景、目标尺度变化、剧烈机动以及干扰条件下的跟踪表现，评估数据集对复杂红外场景的表征能力，分析现有跟踪方法面临的主要挑战，为后续高性能红外目标跟踪算法的研究提供标准测试平台与数据支撑。

2.1 基准测试算法

为全面评估所构建红外目标跟踪数据集的性能与挑战性，本文选取 6 种当前目标跟踪领域具有代表性的基准算法开展测试实验。所选算法涵盖相关滤波、孪生网络、ATOM 系列以及 Transformer 系列等不同技术路线，能够较为全面地反映现有跟踪方法在复杂红外场景下的性能表现。

1) Kernelized Correlation Filter[19] (KCF)：一种经典的相关滤波跟踪算法，将目标跟踪问题转化为岭回归问题求解，并利用循环矩阵实现密集采样，引入核函

表 2 WAR-TIR 与通用场景红外目标跟踪数据集统计信息对比

数据集	序列数	干扰序列数	最大帧数	最小帧数	总帧数	帧率	目标种类	特定应用目标类别数	干扰种类	特定应用场景数
OSU[13]	6	—	2 031	601	8 K	30 FPS	1	0	—	—
ASL-TID[14]	9	—	775	77	4.3 K	20 FPS	3	0	—	—
LTIR[15]	20	—	1 451	71	11 K	30 FPS	6	0	—	—
VOT-TIR16[16]	25		1 451	71	14 K	30 FPS	8		—	—
PTB-TIR[6]	60	—	1451	50	30 K	30 FPS	1	0	—	—
LSOTB-TIR[5]	120	70	2 110	105	82 K	30 FPS	22	3	—	—
WAR-TIR	80	80	1 015	43	31K	30 FPS	8	8	6	5

数增强目标表征能力。该方法具有计算效率高、实时性强等优点，是传统目标跟踪领域的重要代表算法之一。

此外，也有学者采用增量式子空间学习方法实现红外目标跟踪[20]，进一步丰富了传统跟踪方法的技术路线。与此同时，基于 Mean-shift 的改进算法同样被应用于运动目标跟踪系统中[21]，展现了传统跟踪方法在工程应用中的持续价值。

2）Siamese Fully Convolutional Network[22]（SiamFC）：是最具代表性的孪生网络跟踪算法之一，将目标跟踪建模为模板图像与搜索区域之间的相似度匹配问题。该方法采用共享参数的双分支网络提取特征，并通过互相关运算生成响应图，实现目标定位，具有结构简洁、运行速度快等特点。

3）Siamese Box Adaptive Network[23]（SiamBAN）：在孪生网络框架基础上提出的无锚框跟踪算法，通过统一的分类与回归分支预测目标位置和尺度信息。相比基于锚框的跟踪方法，SiamBAN 减少了超参数依赖，提高了模型的泛化能力和跟踪精度，在复杂场景下表现出较好的鲁棒性。

4）Accurate Tracking by Overlap Maximization[24]（ATOM）：是一种基于分类与回归协同优化的跟踪算法。该方法利用在线学习的目标分类器完成粗定位，并通过 IoU 预测网络对候选框进行精细优化，从而实现更加准确的目标状态估计，在尺度变化和外观变化场景下具有较好的适应能力。

5）Probabilistic Discriminative Model Prediction[25]（PrDiMP）：是在 DiMP 基础上发展而来的判别式跟踪算法，通过概率建模方式对目标状态进行预测，并进一步优化分类器学习与模型更新过程。该方法能够有效提高目标与背景之间的区分能力，在遮挡、背景干扰以及目标外观变化等复杂场景下具有较强的鲁棒性和泛化能力。

6）Transformer Tracking[26]（TransT）：是将 Transformer 结构引入目标跟踪领域的代表性方法之一。该算法利用自注意力机制和交叉注意力机制对模板区域与搜索区域进行特征交互和关联建模，实现目标特征的全局信息融合，从而提升复杂场景下的目标定位精度和跟踪稳定性。此外，针对传感器自适应跟踪问题，已有研究基于仿真系统开展了算法验证与实现[27]，为复杂环境下跟踪算法的工程化应用提供了参考。

实验在 Ubuntu 20.04.5 LTS 系统上进行，CPU 为 Intel（R）Xeon（R）Platinum 8375C、GPU 为 Nvidia GeForce RTX 3090（4 块），CUDA 版本号为 11.3，深度学习框架选用 PyTorch 1.11.0。模型训练选择 SGD 作为参数优化器，设定 Batch Size 为 32，共进行 100 个 epoch 学习轮次；搜索区域大小以及模板区域大小分别为 255×255 和 127×127。

2.2　性能评价指标

本文评价方法使用一次性评估（OPE，one pass evaluation）。OPE 指在目标跟踪评估中，跟踪器从序列的第一帧开始，通过已知的目标初始位置进行初始化，在剩余的所有帧中，持续跟踪目标，不进行任何重启或干预。整个过程中，跟踪器仅运行一遍（即“一次性”），从而评估其性能。本文使用 3 个评价指标对算法进行评估，即成功率（Success Rate）、精确度（Precision）以及帧率（FPS，frames per second）。

成功率：评估预测边界框与真实边界框的重叠率，即预测边界框与真实边界框的交集面积除以两者的并集面积，如图 7（a）及公式（1）所示。重叠率大于设定的阈值时，认为跟踪是成功的，统计成功帧所占的比例。

$$S_0 = \frac{|R_a \cap R_b|}{|R_a \cup R_b|} \tag{1}$$

其中：R_a 表示预测目标区域，R_b 表示真实目标区域。

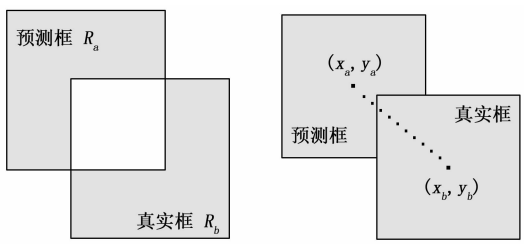

图 7　重叠率与中心距离误差图示

精确度：通过计算预测边界框中心点与真实边界框中心点的中心距离误差 E_{CL} 来衡量跟踪精度，如图 7（b）及公式（2）所示。通常以距离阈值为基准（如 20 像素），统计中心点距离小于该阈值的帧数所占的比例：

$$E_{CL} = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2} \tag{2}$$

其中：(x_a, y_a) 为预测边界框中心点的位置，(x_b, y_b) 为真实边界框中心点的位置。

帧率（FPS，frame per second）：目标跟踪器每秒处理的帧数，如公式（3）所示，反映了跟踪算法的运行效率。FPS 越高，说明该跟踪算法在单位时间内处理视频帧的速度越快。

$$FPS = \frac{\text{总帧数}}{\text{处理总时间}} \tag{3}$$

2.3　基准测试结果与分析

表 3 给出了 6 类代表性红外目标跟踪算法在 WARTIR 测试集上的性能对比结果，包括成功率、精确度以

及帧率指标。从整体结果可以看出，不同算法在 WAR-TIR 数据集上的性能差异较为明显，成功率范围为 0.313～0.642，精确度范围为 0.402～0.695。随着目标跟踪技术从传统相关滤波方法向深度学习方法不断发展，算法性能呈现出明显提升趋势，说明深度特征学习与高层语义表征对于复杂红外场景中的目标跟踪具有重要作用。

表 3 WAR-TIR 测试集上的各算法性能对比

（粗体为最优结果）

序号	跟踪算法	发表年份	成功率 ↑	精确度 ↑	帧率 ↑
1	KCF[19]	2014	0.313	0.402	275.2
2	SiamFC[22]	2016	0.494	0.603	45.2
3	SiamBAN[23]	2020	0.599	0.651	40
4	ATOM[24]	2019	0.590	0.662	33.7
5	PrDiMP[25]	2020	0.625	0.668	22.2
6	TransT[26]	2021	**0.642**	**0.695**	40.8

传统相关滤波算法 KCF 取得了 0.313 的成功率和 0.402 的精确度，虽然运行速度达到 275.2 FPS，具有极高的实时性优势，但整体跟踪性能明显落后于深度学习方法。这主要是由于 KCF 依赖人工设计特征，缺乏对目标外观变化、尺度变化以及复杂背景干扰的有效表征能力。在 WAR-TIR 数据集中，大量存在烟雾遮挡、火光干扰、爆炸干扰以及小目标场景，使得传统特征难以准确区分目标与背景，从而导致跟踪漂移甚至目标丢失。

基于孪生网络的跟踪算法表现出了明显优于传统方法的性能。SiamFC 取得了 0.494 的成功率和 0.603 的精确度，相较于 KCF 分别提升了 57.8%和 50.0%，验证了深度特征在红外目标表示中的优势。进一步地，SiamBAN 通过无锚框目标定位机制和分类回归联合建模策略，将成功率提升至 0.599，精确度提升至 0.651。相比 SiamFC，其成功率和精确度分别提高了 21.3%和 7.9%，说明更加精确的目标定位机制能够有效提升复杂场景下的跟踪性能。

ATOM 系列算法在 WAR-TIR 数据集上同样取得了较好的效果。ATOM 和 PrDiMP 的成功率分别达到 0.590 和 0.625，精确度分别达到 0.662 和 0.668。相比于 ATOM，PrDiMP 通过引入判别模型预测机制以及更加鲁棒的在线学习策略，在成功率和精确度上进一步提升了 5.9%和 0.9%。实验结果表明，在线更新机制能够较好地适应目标在运动过程中产生的尺度变化、姿态变化以及外观变化，提高了目标与复杂背景之间的区分能力。

Transformer 系列算法 TransT 取得了所有基准算法中的最佳性能，成功率达到 0.642，精确度达到 0.695，同时保持 40.8 FPS 的实时运行速度。从性能上看，TransT 相比 KCF 在成功率和精确度上分别提升了 105.1%和 72.9%，相比 SiamBAN 分别提升了 7.2%和 6.8%，相比 PrDiMP 也分别提升了 2.7%和 4.0%。这一结果表明，基于 Transformer 的全局特征建模机制能够更加有效地捕获模板区域与搜索区域之间的长距离依赖关系，在复杂背景、多干扰以及目标快速运动场景下具有更强的特征表达能力和鲁棒性。

值得注意的是，尽管 TransT 取得了当前最优性能，但其成功率和精确度仍分别仅达到 0.642 和 0.695，距离理想跟踪效果仍存在较大提升空间。这说明 WAR-TIR 数据集中的目标跟踪任务具有较高挑战性。一方面，数据集包含地面与空中多种目标，目标尺度跨度较大；另一方面，数据集中引入了烟雾、火光、爆炸和点源干扰等典型干扰因素，同时覆盖田野、沙漠、山脉和靶场等复杂背景环境，对目标跟踪算法的特征提取能力、抗干扰能力以及泛化能力提出了更高要求。因此，WAR-TIR 数据集能够有效拉开不同算法之间的性能差距，为红外目标跟踪算法的研究与评估提供具有挑战性的标准测试平台。

3 结束语

针对现有红外目标跟踪数据集存在的目标覆盖不足、背景场景描述不充分以及干扰缺失等问题，本文构建了面向复杂环境的红外目标跟踪数据集 WAR-TIR。该数据集涵盖多类典型目标，以及田野、沙漠、山脉和靶场等多种背景场景，并引入烟雾、火光、爆炸和光照等复杂自然干扰因素，形成具有较高挑战性的测试平台。通过对数据集构建流程、统计特性及挑战属性进行分析，并基于代表性跟踪算法建立统一基准测试，验证了数据集的有效性与挑战性。实验结果表明，当前先进跟踪算法在 WAR-TIR 上的性能仍存在明显提升空间，说明复杂自然环境下的红外目标跟踪仍面临挑战。

参考文献：

[1] XIE B，ZHANG C，WANG F，et al. Cst anti-uav：a thermal infrared benchmark for tiny uav tracking in complex scenes [C] //2025 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)，Honolulu，HI，USA，October 19－20，2025：6216－6225.

[2] HUANG B，LI J，CHEN J，et al. Anti-UAV410：a thermal infrared benchmark and customized scheme for tracking drones in the wild [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2024，46

(5)：2852－2865.

[3] LI C，XUE W，JIA Y，et al. LasHeR：a large-scale high-diversity benchmark for RGBT tracking [J]. IEEE Transactions on Image Processing，2022，31：392－404.

[4] ZHANG P，ZHAO J，WANG D，et al. Visible-thermal UAV tracking：a large-scale benchmark and new baseline [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition，New Orleans，USA，June 19－24，2022.

[5] LIU Q，LI X，HE Z，et al. LSOTB-TIR：a large-scale high-diversity thermal infrared object tracking benchmark [C] //ACM International Conference on Multimedia，Seattle，2020.

[6] LIU Q，HE Z，LI X，et al. PTB-TIR：a thermal infrared pedestrian tracking benchmark [J]. IEEE Transactions on Multimedia，2020，22 (3)：666－675.

[7] ALI A，PANDIT A，MANDAL S，et al. ICPR 2024 competition on VISual tracking in adverse conditions (VISTAC) [C] //International Conference on Pattern Recognition，2024.

[8] EL AHMAR W，SAPPA A，HAMMOUD R. Thermal pedestrian multiple object tracking challenge (tp-mot) [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops，Nashville，2025.

[9] NIE Z，XUE L，FANG Z，et al. M3ot：a multi-drone multi-modality dataset for multi-object tracking [J]. Scientific Data，2025，12 (1)：1－10.

[10] YUAN D，ZHANG H，SHU X，et al. Thermal infrared target tracking：a comprehensive review [J]. IEEE Transactions on Instrumentation and Measurement，2024，73：1－19.

[11] FU R G，WANG L，FAN H Q，et al. A dataset for infrared time-sensitive target detection and tracking for air-ground application [J]. China Scientific Data，2022，7 (2)：202－217.

[12] WU S，ZHANG K，NIU S，et al. Anti-interference aircraft-tracking method in infrared imagery [J]. Sensors，2019，19 (6)：1289.

[13] DAVIS J W，SHARMA V. Background-subtraction using contour-based fusion of thermal and visible imagery [J]. Computer Vision and Image Understanding，2007，106 (2－3)：162－182.

[14] PORTMANN J，LYNEN S，CHLI M，et al. People detection and tracking from aerial thermal views [C] //IEEE International Conference on Robotics and Automation，2014.

[15] BERG A，AHLBERG J，FELSBERG M. A thermal object tracking benchmark [C] //IEEE International Conference on Advanced Video and Signal Based Surveillance，2015.

[16] LEBEDA K，HADFIELD S J，BOWDEN R. The thermal infrared visual object tracking (VOT-TIR2016) challenge result [C] //European Conference on Computer Vision Workshops，Amsterdam，2016.

[17] FELSBERG M，BERG A，HAGER G，et al. The thermal infrared visual object tracking VOT-TIR2015 challenge results [C] //IEEE International Conference on Computer Vision Workshops，Santiago，Chile，December 7－13，2015.

[18] 方彦策，赵君灵，黄昭龙，等. 可见光与红外融合目标跟踪技术研究进展综述 [J]. Computer Measurement & Control，2022，30 (10)：7.

[19] HENRIQUES J F，CASEIRO R，MARTINS P，et al. High-speed tracking with kernelized correlation filters [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2015，37 (3)：583－596.

[20] 常雅男，毕笃彦，查宇飞，等. 基于增量式子空间学习的红外目标跟踪研究 [J]. 计算机测量与控制，2013，21 (6)：1668－1671.

[21] 张宝峰，李　翠，朱均超，等. 一种改进的 Mean-shift 算法在运动目标跟踪系统中的应用 [J]. 计算机测量与控制，2013，21 (2)：502－504.

[22] BERTINETTO L，VALMADRE J，HENRIQUES J F，et al. Fully-Convolutional siamese networks for object tracking [C] //European Conference on Computer Vision，2016.

[23] CHEN Z，ZHONG B，LI G，et al. Siamese box adaptive network for visual tracking [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition，Seattle，2020.

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

[25] DANELLJAN M，GOOL L V，TIMOFTE R. Probabilistic regression for visual tracking [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition，Seattle，2020.

[26] CHEN X，YAN B，ZHU J，et al. Transformer tracking [C] //IEEE/CVF Conference on Computer Vision and Pattern Recognition，2021.

[27] 贺　贺，顾　禹，徐高正，等. 传感器自适应跟踪算法研究及仿真系统实现 [J]. 计算机测量与控制，2020，28 (11)：260－264.