中图法分类号:TP392;TP18　　**文献标识码**:A　　**文章编号**:1006-8961(2026)09-3321-15

论文引用格式:Li R J, Li Z X, Chen N, Guo G W, Dou Z C, Long Z X, Luo Y H, Zeng Y Y, Sheng W D, Li B Y, Li Z J, Li M, An W, Long Y L, Li H X, Yu Z Q, Yin X Y, Zha X Y, Zeng B W, Wang W F, Zhang B, Lu Y, Chu D H, Li Z Y, Huang K W, Yang B R, Xing Y H and Zhang S Z. 2026. Infrared video satellite aerial moving target detection dataset and its evaluation. Journal of Image and Graphics, 31(9):3321-3335(李若敬, 李朝旭, 陈诺, 郭高伟, 窦泽超, 龙政兴, 罗伊杭, 曾瑶源, 盛卫东, 李博扬, 李志军, 李淼, 安玮, 龙云利, 李海鑫, 余志强, 尹小宇, 查旭阳, 曾柏文, 王吴凡, 张波, 路悦, 褚东皓, 李子怡, 黄康薇, 杨博睿, 邢颖慧, 张世周. 2026. 红外视频卫星空中动目标检测数据集及其评估. 中国图象图形学报, 31(9):3321-3335)[DOI:10.11834/jig.250536]

红外视频卫星空中动目标检测数据集及其评估

李若敬[1],李朝旭[1],陈诺[1],郭高伟[1],窦泽超[1],龙政兴[1],罗伊杭[1],曾瑶源[1*],盛卫东[1],李博扬[1],李志军[2],李淼[1],安玮[1],龙云利[1],李海鑫[3],余志强[4],尹小宇[5],查旭阳[3],曾柏文[3],王吴凡[3],张波[3],路悦[6],褚东皓[6],李子怡[6],黄康薇[6],杨博睿[6],邢颖慧[6],张世周[6]

1. 国防科技大学电子科学学院,长沙　410073; 2. 武汉大学遥感信息工程学院,武汉　430072;
3. 北京邮电大学计算机学院,北京　100876; 4. 清华大学精密仪器系,北京　100084;
5. 新疆大学计算机学院,乌鲁木齐　830046; 6. 西北工业大学计算机学院,西安　710129

摘　要:**目的**　红外视频卫星是探测空中动目标的重要手段,红外小目标检测技术是其关键基础。深度学习显著推动了单帧红外小目标检测,然而卫星红外视频中的空中动目标普遍空域显著性低、场景复杂,单帧方法难以有效检测,因此亟需发展融合时域信息的红外极弱小目标检测技术。但该领域长期缺乏视频数据集,严重制约了相关技术的发展与应用。为突破此瓶颈,该文构建了首个包含大量真实场景的红外视频卫星空中动目标检测数据集。**方法**　基于武汉一号卫星采集20 126帧真实红外视频卫星动目标数据,设计两阶段"由粗到精"的标注方法,完成29 757个空中动目标的精标注。为了丰富场景多样性,进一步融合两个真实天基背景下的仿真动目标数据,构建包含1 401个真实场景、122 265帧视频图像、454 116个目标的红外视频卫星空中动目标检测数据集。数据集提供实例级掩码标签,支持空中动目标检测与跟踪技术研究,并提出了相关评价指标。**结果**　数据分析表明,该数据集中真实目标的平均信噪比仅为3.06,超过80%的目标信噪比低于2,且实测场景中的目标与背景存在丰富的动态变化与相互干扰,呈现难以模拟的复杂性。**结论**　基于该数据集开展了首届红外视频卫星空中动目标检测比赛,充分验证了该数据集的高挑战性与实际价值,对于红外极弱小目标检测技术研究具有重要支持作用。数据集获取链接:https://github.com/TinaLRJ/DeepPro和科学数据银行ScienceDB:https://doi.org/10.57760/sciencedb.j00240.00077。

关键词:红外视频卫星;空中动目标检测;空中动目标跟踪;视频数据集;深度学习

Infrared video satellite aerial moving target detection dataset and its evaluation

Li Ruojing[1], Li Zhaoxu[1], Chen Nuo[1], Guo Gaowei[1], Dou Zechao[1], Long Zhengxing[1], Luo Yihang[1], Zeng Yaoyuan[1*], Sheng Weidong[1], Li Boyang[1], Li Zhijun[2], Li Miao[1], An Wei[1], Long Yunli[1], Li Haixin[3], Yu Zhiqiang[4], Yin Xiaoyu[5], Zha Xuyang[3], Zeng Baiwen[3], Wang Wufan[3], Zhang Bo[3], Lu Yue[6], Chu Donghao[6], Li Ziyi[6], Huang Kangwei[6], Yang Borui[6], Xing Yinghui[6], Zhang Shizhou[6]

1. College of Electronic Science and Technology, National University of Defense Technology, Changsha　410073, China;

收稿日期:2025-10-30;**修回日期**:2025-12-21;**预印本日期**:2025-12-28

*通信作者:曾瑶源　yyzeng@hotmail.com

基金项目:国家自然科学基金项目(62501618,62501609,62401591,42501589,62505362)

Supported by:National Natural Science Foundation of China(62501618, 62501609, 62401591, 42501589, 62505362)

2. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China; 3. School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China; 4. Department of Precision Instrument, Tsinghua University, Beijing 100084, China; 5. School of Computer Science and Technology, Xinjiang University, Urumqi 830046, China; 6. School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China

Abstract: Objective Infrared video satellites serve as critical tools for detecting aerial moving targets, with infrared small target detection technology forming the essential foundation for this capability. The rapid advancement of deep learning has yielded numerous single-frame detection datasets and methodologies, enabling substantial progress in identifying spatially salient targets. However, moving aerial targets captured by infrared satellites typically exhibit low spatial salience and frequently occur in complex scenarios, rendering single-frame detection methods reliant on spatial information ineffective in such challenging conditions. This urgent challenge necessitates the development of multiframe infrared small and dim target detection techniques. A major bottleneck restricting the advancement and practical application of aerial moving target detection technology has been the lack of dedicated datasets for infrared video satellite-based detection, primarily due to the difficulties and high costs associated with data collection and annotation. To address this gap and promote technological development, we construct the first infrared video satellite aerial moving target detection dataset containing many real-world scenarios, called SatVideoIRSDT dataset, and organized the inaugural detection competition based on this dataset. **Method** First, we collect 20 126 frames of real infrared video satellite data from Wuhan No. 1 satellite featuring aerial moving targets and annotate 29 757 aerial targets. During the data annotation phase, we design a two-stage annotation method to address the challenge of extremely poor spatial salience of the targets that makes them difficult to distinguish. The first stage is data preannotation, where videos are initially cropped spatially into multiple locally enlarged subvideos. By repeatedly playing these subvideos, we determine the presence of targets in each and annotate their temporal and spatial range information. The second stage involves fine-grained data annotation based on the preannotation results, following a specific workflow: confirming target positions, drawing target boundaries, and filling to generate masks. In this manner, well-annotated real data can be obtained. Then, we integrate two simulated infrared aerial moving target datasets with authentic space-based backgrounds to enhance scenario diversity. The resulting dataset comprises 1 401 real scenarios, 122 265 video frames, and 454 116 annotated targets, with mask labels distinguishing different target instances to support detection and tracking research. In the challenge based on this dataset, we establish two core tasks, i. e. , aerial moving target detection and tracking, to advance integrated detection-tracking technology. Corresponding to the two tasks, we design two special evaluation metric systems. Especially, the tracking performance metrics are proposed for the first time, including trajectory completeness rating and trajectory accuracy rating. **Result** Analysis reveals the high complexity of our dataset: The real targets are extremely small and dim, with an average size of no more than 20 pixels and an average signal-to-noise ratio (SNR) of only 3. 06. Among them, over 80% of the targets have an SNR below 2. The real videos contain rich dynamic variations of targets, including changes in motion state, grayscale distribution, and morphological size. The background also exhibits complex dynamic changes, including translation, rotation, stretching, and scaling in different directions, and local grayscale variations. Additionally, complex mutual radiation interference exists between targets and backgrounds. The diversity introduced by real-world scenarios and actual systems cannot be easily replicated by simulated data, which also causes our SatVideoIRSDT dataset more challenging than other public datasets. For the detection task, a state-of-the-art baseline algorithm achieved a recall of 0. 578 2, precision of 0. 495 8, and F1-score of 0. 533 8 on this dataset. For the tracking task, a typical baseline algorithm yielded a trajectory completeness rating of only 0. 241 9 and trajectory accuracy rating of 0. 519 1. **Conclusion** The competition fully validated the practical value and high challenge of our SatVideoIRSDT dataset, providing crucial data support for research on infrared small and dim target detection technology. It also attracted substantial attention to the field, facilitating academic exchange and technological progress. Through this event, we collected numerous outstanding aerial moving target detection and tracking solutions, which have contributed valuable new insights to drive domain development. The dataset is available at https://github.com/TinaLRJ/DeepPro (or Science Data Bank: Infrared video satellite aerial moving target detection dataset).

Key words: infrared video satellite; aerial moving target detection; aerial moving target tracking; video dataset; deep learning

0 引 言

红外视频卫星具有覆盖范围广、全天时和全天候观测的优势，因此成为发现空中动目标的重要手段（Ying等，2025a；Li等，2025）。空中动目标通常包括各类民航飞机等，其运动轨迹具有连续性，经过远距离成像后目标尺寸非常小且能量弱。红外小目标检测技术是实现空中动目标检测的重要基础。由于深度学习具有强大的数据学习能力，近年来许多领域都构建了专门的大规模高质量数据集用于算法研究，并且在充足数据集的支持下大量深度学习算法相继提出，推动不同领域取得了优秀的成绩。红外小目标检测领域也涌现出多个单帧红外小目标检测数据集（Zhang等，2022a）和大量基于空域信息的单帧红外小目标检测方法（Li等，2023a），对于空域显著目标的检测取得了显著进展。然而，基于空域信息的单帧检测方法难以实现红外视频卫星探测下的空中动目标检测，其原因主要如下：

1）空中动目标距离探测器远，在图像中通常占据更少的像素点（通常小于9×9像元），更加缺乏清晰的形状、纹理和结构等可供辨识的细节特征，极大增加了目标特征提取的难度。

2）许多目标能量弱且在传播过程中衰减显著，其成像信号非常微弱。同时，复杂的环境和传感器自身会引入背景杂波和噪声。目标信号容易淹没在杂波和噪声中，导致目标信噪比低，使得目标与背景难以区分。

3）空中动目标和某些背景干扰物在红外波段可能具有相似的热辐射特性，导致目标与某些背景干扰物之间特征相似度高，增加了误检的风险。

针对上述困难，迫切需要推动联合时域信息的红外极弱小目标检测技术的发展，利用目标与背景在时域和时空联合域上的差异性，实现对红外视频卫星空中动目标的有效检测。然而，数据是推动深度学习相关技术发展的重要基础，现有的序列红外小目标检测数据集存在各种问题，导致红外视频卫星空中动目标检测技术发展动力不足，相关成果较少且距离实际应用差距大。

具体而言，目前序列红外小目标检测数据集主要存在以下问题。首先，现有数据集包含的场景大多是近地背景且目标通常偏大，如空天杯数据集（傅瑞罡 等，2022）、Anti-UAV410数据集（Huang等，2023）、IRDST数据集（the infrared dim small target dataset）（Sun 等，2023），以及RGBT-Tiny数据集（Ying等，2025b）等，这些数据集大多通过无人机搭载相机对地拍摄的方式获取数据，目标通常包括车辆、无人机、飞机、行人和船只等，由于拍摄距离较近，目标尺寸较大且具有一定形态信息，其时空显著性远高于红外视频卫星拍摄的空中动目标，检测难度较低。其次，大部分极小目标的数据集都为仿真数据集，如SIATD数据集（small infrared aerial target detection dataset）（Sun 等，2021）、NUDT-MIRSDT数据集（multiframe infrared small and dim target dataset）（Li等，2025）、IRAir数据集（infrared aircraft dataset）（李朝旭 等，2025），以及IRSatVideo-LEO数据集（Ying等，2025a）等。然而，仿真序列难以真实模拟帧间背景动态变化，以及目标的动态变化，导致研究算法部署到实际工程应用中可能出现不适应的问题。此外，考虑到标注成本的问题，许多真实数据集采用包围框标注的形式，如空天杯数据集（傅瑞罡等，2022）、Anti-UAV410数据集（Huang等，2023）等，这限制了基于分割框架的检测算法的发展。

为了解决上述问题，本研究采集了武汉一号卫星搭载的热红外载荷拍摄的地面机场周围背景的20 126帧真实空中动目标（飞机）视频数据，并用掩码标记了29 757个空中动目标，部分场景如图1（c）（f）所示。可以看出，民航飞机在卫星载荷上呈现为点状，当飞机蒙皮温度低于背景温度时通常表现为暗目标，如图1（c）所示；当飞机蒙皮温度高于背景温度时通常表现为亮目标，如图1（f）所示。为了丰富数据集场景多样性，融合了两个红外卫星仿真数据集，一个是来自“可持续发展科学卫星1号”（SDG卫星1号）的红外图像暗弱目标检测数据集IRAir（李朝旭 等，2025），如图1（a）（d）所示；另一个是来自Landsat卫星的红外图像亮目标检测数据集IRSatVideo-LEO（Ying 等，2025a），如图1（b）（e）所示。这两个仿真数据集均采用真实红外卫星图像，通过向图像中添加观测平台抖动、运动等仿真视频图像帧间位移，同时向图像中添加噪声以及连续运动的暗、亮目标，以生成红外视频卫星仿真数据。最终形成了包含1 401个真实场景，122 265帧视频图像数据，以及454 116个目标的红外视频卫星空中动目标检测数据集SatVideoIRSDT（infrared video satel-

lite aerial moving target detection dataset)，所有目标均以掩码标注并按照不同目标轨迹区分编号。该数据集是领域内首个包含大量极具挑战性真实场景的数据集，超过80%的真实目标信噪比低于2，支持基于其进行红外极弱小目标检测与跟踪的研究。此外，本文还提出了红外小目标跟踪的评价指标，用于评估跟踪准确性和完整性。为了推动空中动目标检测技术的发展，基于该数据集开展了首届红外视频卫星空中动目标检测比赛，并设置了跟踪附加任务，吸引了国内大量优秀团体参与，实现了对SatVideoIRSDT数据集的充分评估，并取得了显著的技术进展。

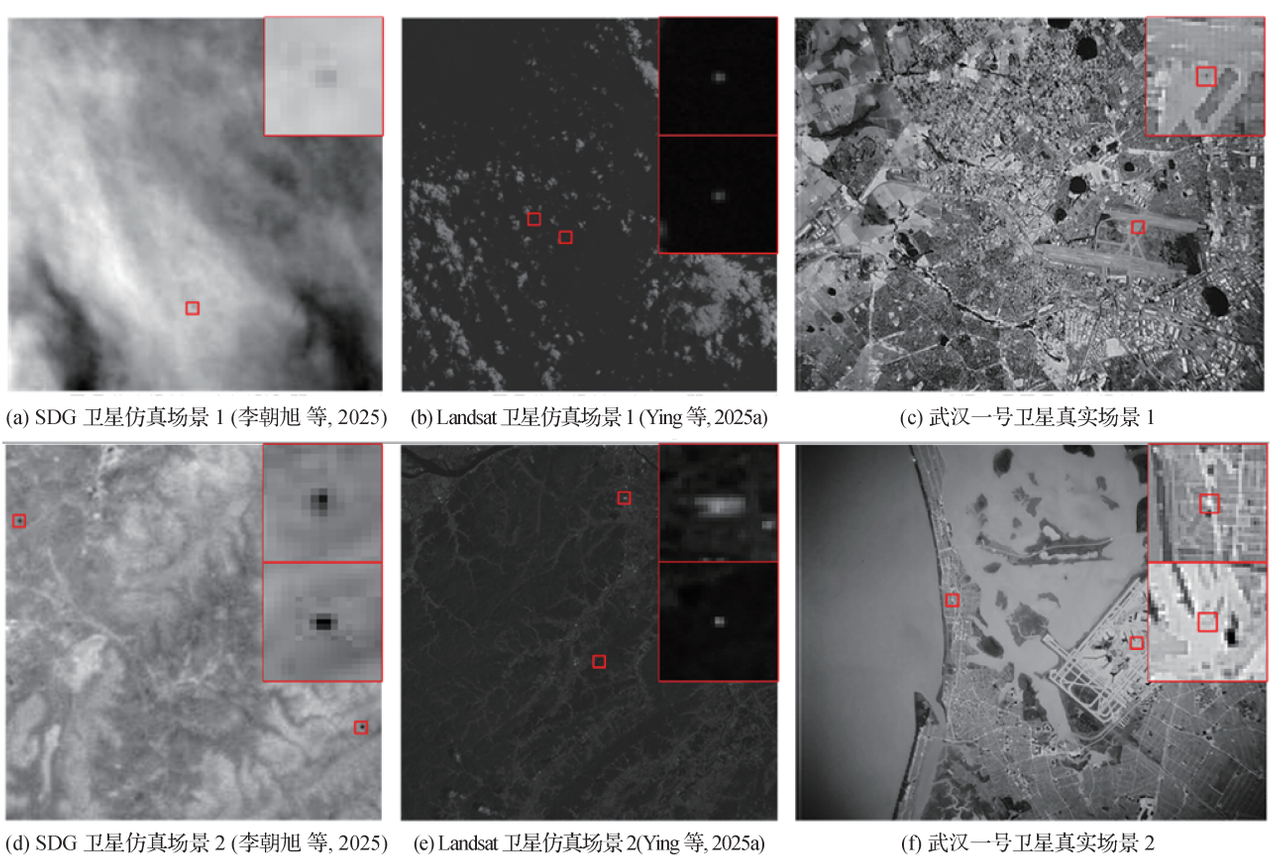
综上所述，本文的主要贡献如下：1)首次公布了大量精细标注的实测红外视频卫星空中动目标场景数据。数据包含20 126帧视频图像、29 757个空中动目标，超过80%的目标信噪比低于2，在该数据集上进行检测具有极高难度。2)构建了首个包含真实数据的大型红外视频卫星空中动目标检测数据集。为了提高场景多样性以更好地支持大规模网络训练，数据集中混合了部分红外卫星观测下的空中动目标仿真视频数据，该数据集对研究红外极弱小目标检测具有重要意义。3)基于该数据集组织了首届红外视频卫星空中动目标检测比赛，规范提出了检测与跟踪任务指标，并对该数据集进行了充分的评估，证明了该数据集的高挑战性尤其是真实数据，及其对测试相关算法能力极限的重要作用。

1 现有数据集与红外小目标检测技术

1.1 红外小目标检测数据集

1.1.1 单帧检测数据集

目前的单帧红外小目标检测数据集大部分是陆基和空基观测视角下的仿真数据，标签形式主要为掩膜标签，真实数据集较少且一般规模不大，如表1所示。Wang等人(2019)建立了首个单帧红外斑目



(a) SDG 卫星仿真场景 1 (李朝旭 等, 2025)　(b) Landsat 卫星仿真场景 1 (Ying 等, 2025a)　(c) 武汉一号卫星真实场景 1

(d) SDG 卫星仿真场景 2 (李朝旭 等, 2025)　(e) Landsat 卫星仿真场景 2(Ying 等, 2025a)　(f) 武汉一号卫星真实场景 2

图1　红外视频卫星空中动目标检测数据集部分场景与目标示意图

Fig. 1　Example scenes with targets of infrared small and dim moving target detection datasets in satellite videos (SatVideoIRSDT) ((a) simulated scenario 1 from SDG satellite (Li et al., 2025); (b) simulated scenario 1 from Landsat satellite (Ying et al., 2025a); (c) real scenario 1 from Wuhan No. 1 satellite; (d) simulated scenario 2 from SDG satellite (Li et al., 2025); (e) simulated scenario 2 from Landsat satellite (Ying et al., 2025a); (f) real scenario 2 from Wuhan No. 1 satellite)

标仿真数据集 NUST-SIRST（single-frame infrared small target dataset），通过将仿真斑目标添加到采集得到的真实红外场景图像中，从而得到大量数据。Li 等人（2023a）构建了包含更多目标类型的仿真数据集 NUDT-SIRST。随着领域的发展获得了更多的关注，逐渐出现了一些小规模的真实数据集，如 NUAA-ACMT 数据集（Dai 等，2021a）和 IRSTD-1k 数据集（Zhang 等，2022a）。随后，Wu 等人（2023b）公布了首个天基卫星观测视角下的大规模近红外实测船目标数据集 NUDT-SIRST-Sea。这些单帧检测数据集为序列检测数据集提供了大量真实背景，是扩充仿真序列检测数据集的重要基础。

表 1　单帧红外小目标检测数据集

Table 1　Single-frame infrared small target detection datasets

数据集	类型	分辨率/像素	标签类型	场景	目标类型	目标平均大小	目标数量
NUST-SIRST	仿真	128 × 128*	掩膜	陆基	斑	151	10 337
NUDT-SIRST	仿真	256 × 256	掩膜	陆基、空基	点、斑、无人机、客机、船、车	44	1 901
NUAA-SIRST	实测	303 × 221*	掩膜	陆基、空基	点、斑、无人机、船、车	40	533
IRSTD-1k	实测	512 × 512	掩膜	陆基	点、斑、无人机	85	1 001
NUDT-SIRST-Sea	实测	1 024 × 1 024	掩膜	天基	游艇、货船、军舰、游轮、油罐船	29	17 598

注：“*”代表多种分辨率的平均分辨率。

1.1.2　序列检测数据集

现有的序列红外小目标检测数据集主要以陆基和空基观测视角为主，相比单帧检测数据集，数据集规模明显更大、目标整体偏小且信噪比更低。Sun 等人（2021）建立了空基观测视角的大规模仿真数据集，有陆地和天空等多种类型场景。但该数据集仅提供了目标中心位置的标签（点标签），可能影响模型训练效果，限制了不同框架算法的研究。Sun 等人（2023）公布了一个大规模混合序列红外小目标检测数据集 IRDST，该数据集包含两部分，一部分为实测数据集，另一部分为仿真数据集，两部分目标的平均尺寸均小于 10 像素，为序列红外小目标检测技术研究提供了支撑。然而，该数据集的大部分场景是地基对空观测视角，导致大部分目标出现在背景平缓的天空中，场景多样性较小，实际检测难度较低。Li 等人（2025）仿真了首个低信噪比序列红外小目标检测数据集 NUDT-MIRSDT，其测试集按照信噪比划分为两个子集，其中一个子集目标平均信噪比低于 3；李朝旭等人（2025）建立了首个大规模天基观测视角红外暗弱目标检测数据集，分析了民航飞机在热红外卫星图像上的真实形态，首次关注了暗弱目标的检测问题。Ying 等人（2025a）分析了天基观测平台的运动特性并基于此仿真了天基观测视角下的序列数据集。

尽管已经有了许多序列红外小目标检测数据集，但实测数据非常少，仿真数据无法完全模拟真实红外卫星数据的各种挑战，包括场景的动态变化与目标的动态变化，天基观测视角下的实测数据仍然空白。

1.2　红外小目标检测技术

1.2.1　传统红外小目标检测技术

传统红外小目标检测技术经历了长期的发展，根据算法利用信息源的维度不同，分为单帧检测技术和序列检测技术。单帧红外小目标检测技术主要有 3 类方法：基于空域滤波的方法（Bai 和 Zhou，2010；张世锋 等，2016；Zhu 等，2020）、基于低秩稀疏性的方法（Kong 等，2021；Zhao 等，2022）和基于视觉注意机制的方法（Liu 等，2018；Yang 等，2023）。序列红外小目标检测技术主要有基于三维滤波的方法（Gao 等，2017；Li 等，2021）和基于低秩稀疏性的方法（Li 等，2023b；Wu 等，2023a）。

这些方法通过建模背景先验或设计手工特征提取目标来实现检测，为基于深度学习的检测技术的发展提供了重要理论基础。然而，传统方法存在两大瓶颈：1）手工特征泛化能力弱，模型鲁棒性依赖预先对场景的假设或者预设的目标匹配模板，当场景或目标偏离预设时模型容易失效；2）检测速度与精度难以兼顾，当对精度要求很高时可能需要利用更

多的手工特征或者基于更大的三维时空信息进行建模，这会导致检测速度显著下降。基于深度学习的方法能够更好地解决传统方法的瓶颈，同时带来显著的性能提升。

1.2.2　基于深度学习的红外小目标检测技术

近年来，基于深度学习的方法由于其强大的特征学习能力获得了比传统方法更好的检测性能（寇人可 等，2024）。根据算法利用信息源的维度不同，同样分为单帧智能检测技术和序列智能检测技术。单帧检测方法主要针对空域局部显著性突出的目标和背景复杂度不太高的场景（如非高噪场景），通过设计特定的模块或网络结构提取目标的局部显著性、稀疏性或背景的低秩性等特征实现检测。为了提取目标局部显著性特征，Dai 等人（2021b）将局部对比度度量融入到深度网络中，Zhang 等人（2022b）将局部对比度度量转换为深层特征空间中的非局部正交差度量。张迎梅等人（2026）基于选择注意力机制增强红外小目标的局部显著性。为了利用目标的稀疏性和背景的低秩性，Zhou 等人（2023）提出了一种深度低秩稀疏图像网络，通过将图像块分解为低秩背景分量和稀疏目标分量的叠加，将目标检测问题转化为低秩稀疏正则化的优化问题。序列检测方法具有更好的检测低空域局部显著性目标和抑制高复杂度背景的潜力，主要有基于时空显著性特征的方法和基于运动特征的方法，二者分别利用目标的时空显著性特征和运动特征。Du 等人（2022）基于帧间能量累积增强机制建立了时空特征提取的目标检测框架，利用目标时空显著性提取实现检测。Sun 等人（2021）针对目标轨迹的局部平滑性和全局连续性进行建模，并将其转化为两个约束项，以从众多候选目标中筛选出真实目标。Li 等人（2025）设计了一种基于方向编码的卷积块，能够将运动方向编码到特征中，使网络能够提取目标的运动信息，从而实现目标检测。

这些方法通过将传统先验信息与深度网络融合，取得了良好检测性能。为了进一步提升检测性能，一些方法将注意力机制（Liu 等，2023；Zhang 等，2025）融入到网络中，提取不同局部之间的相关性，使网络更好地理解语义信息；一些方法（Liu 等，2024；Li 等，2025）通过设计更适合红外小目标检测的损失函数帮助网络收敛得到更优解。基于深度学习的方法不仅取得了超越传统方法的良好检测性能，同时，得益于显卡图形加速计算能力，算法检测速度也大幅提升，突破了传统方法检测精度和速度的瓶颈。

2　数据集

2.1　数据采集与来源

2.1.1　数据采集

目前已有的红外视频卫星空中动目标检测数据集均为仿真数据集，有亮目标数据集（Ying 等，2025a）和暗弱目标数据集（李朝旭 等，2025），这些数据集向真实单帧场景中添加具有各种运动轨迹的仿真目标以获取数据，同一序列中的仿真目标通常状态维持稳定，形态等信息变化不大。

然而，真实场景中，同一目标除了运动状态（方向、速度等）会随时间发生变化，目标的灰度分布、形态大小等特征均会发生各种变化。同一真实目标不同飞行时刻可能存在亮于背景的状态，也可能存在暗于背景的状态，这与目标不同飞行状态下的表面温度、目标表面材料特性以及背景环境温度等因素相关。真实场景中目标的形态也并非一成不变，会随着不同飞行阶段（如加速、巡航阶段等）以及不同环境状况（如不同季节、不同云层高度等）等发生变化。此外，真实场景中目标与背景的辐射互相影响，并非理想的简单叠加关系（许多仿真数据直接将目标灰度矩阵加到背景中）。受目标表面与背景属性等影响，目标与背景辐射可能存在复杂的散射与反射，从而导致辐射干扰。

综上，实测场景数据的复杂性、多样性和多变性等特点突出，对于研究红外极弱小目标检测具有重要价值。然而，该领域实测视频数据始终空白，采集真实红外视频卫星空中动目标数据十分重要。为此，本研究采集了武汉一号卫星热红外相机拍摄的不同机场附近的视频数据。视频中的空中动目标主要是从机场起飞、降落以及滑行的飞机目标，目标的运动状态、灰度分布和形态大小等特性具有丰富的实时变化，而且背景变化复杂、目标状态多样。

2.1.2　数据集来源

由于真实场景数据获取难度大，数据有限，所采集标注的真实视频数据共 20 126 帧，包含空中动目标 29 757 个。为了能够更好地支持大规模网络训练，混合了部分红外视频卫星空中动目标仿真数据，

有来自SDGSAT-1卫星热红外相机影像的IRAir数据集(李朝旭 等,2025)的暗目标数据和来自Landsat8-9卫星OLI传感器红外波段影像的IRSatVideo-LEO数据集(Ying等,2025a)的仿真亮目标数据。其中,来自IRAir数据集的有1 314个场景,共65 700帧;来自IRSatVideo-LEO数据集的有82个场景,共36 439帧。3部分视频数据共同构成最终的数据集。

最终,形成了包含1 401个真实场景,122 265帧视频图像数据,以及454 116个目标的红外视频卫星空中动目标检测数据集SatVideoIRSDT。

2.2 真实场景数据标注

由于所采集真实场景数据中的目标占比非常小,且信噪比很低,因此在进行正式数据标注前,需要先检查视频数据中目标轨迹数并赋予每个目标ID号,以及确认每个目标的粗略运动范围和出现的视频帧范围(如目标1出现在约第$N1$—$N2$帧,从($x1$, $y1$)运动到($x2$, $y2$)),该过程为预标注。由此得到每个视频中目标的概略信息,并依据此信息完成数据精标注工作,最终得到mask标签。

预标注过程中,由于视场很大,目标尺寸很小,为了发现目标,首先将视频按照像空间位置裁剪成许多小尺寸的视频,然后将小尺寸视频放大重新保存成局部放大高清视频并在视频上方标注每一帧对应的原始数据的帧号。基于局部放大高清视频,通过反复播放视频的方式,人工检查每个局部视频中是否存在空中动目标(重点检查机场范围)。若发现目标,则赋予每个目标轨迹新ID号并记录目标运动范围和视频帧范围等信息,最终得到预标注结果。

基于预标注结果,开展逐个目标的精标注工作,该工作具体流程如图2所示。首先,根据预标注结果找到待标记目标的起始帧和在像平面的大致局部范围,并放大目标所在的局部区域;然后,通过来回切换当前待标记帧前后视频,确认目标详细位置,并用多边形画出目标与背景的边界线。基于此,用0灰度值标记背景区域,用目标ID号对应的数值作为灰度值填充多边形圈中的目标区域。考虑到目标边缘灰度与背景灰度相似,且部分目标能量极弱,与背景相近,边界线上的像素也被标记为目标像素,由此生成该目标的mask标签。重复上述过程,完成所有帧中该目标的连续标注和所有目标的视频标注,最终得到视频数据区分不同目标实例的分割标注结果。该标签形式支持在该数据集上开展红外视频卫星空中动目标检测与跟踪任务研究。

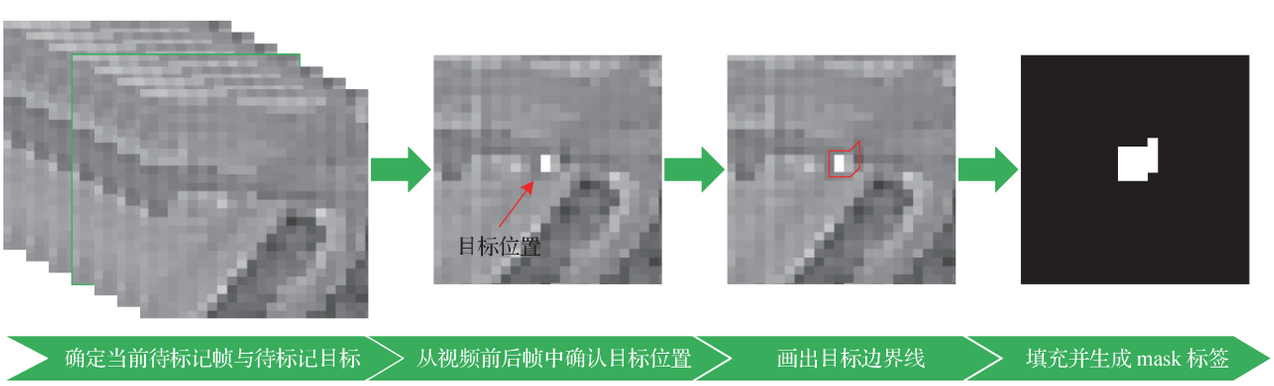


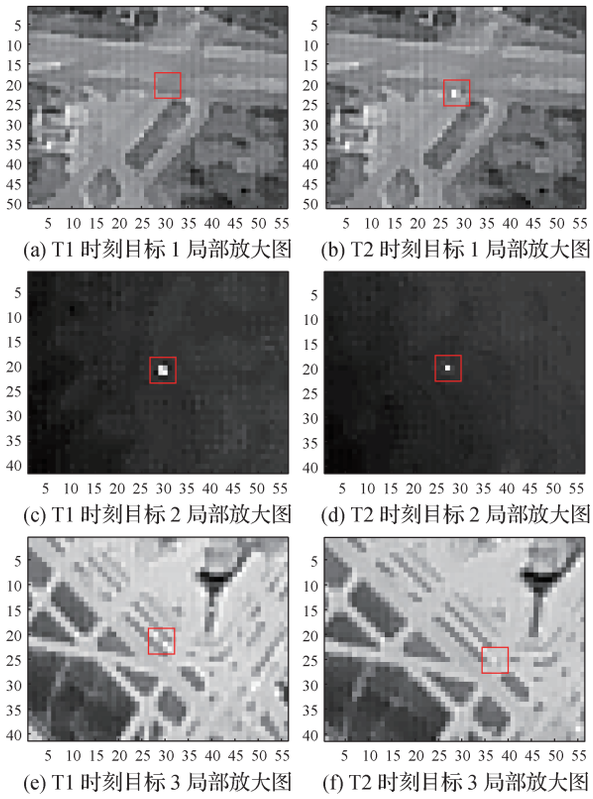
图2 目标精标注流程示意

Fig. 2 Schematic of target precision annotation process

2.3 数据特性分析

2.3.1 真实数据目标多样性分析

前面提到真实场景中目标千变万化、状态丰富,真实数据具有不可替代的重要作用。该部分详细展示分析了所采集的真实数据中目标的各种变化以及不同状态。图3展示了部分目标不同时刻的状态。从图中可以看出,所收集的真实红外视频卫星空中动目标在视频中具有丰富的变化和各种复杂状态。如图3(a)(b),目标1在某一时刻从暗目标变为亮目标,其灰度分布发生了显著变化;如图3(c)(d),目标2的尺寸从2×2变为1×1,其形态大小发生了显著变化;如图3(e)(f),目标3在T1时刻与临近背景辐射互相影响,在图像中高度耦合在一起难以划分。



(a) T1 时刻目标 1 局部放大图　(b) T2 时刻目标 1 局部放大图
(c) T1 时刻目标 2 局部放大图　(d) T2 时刻目标 2 局部放大图
(e) T1 时刻目标 3 局部放大图　(f) T2 时刻目标 3 局部放大图

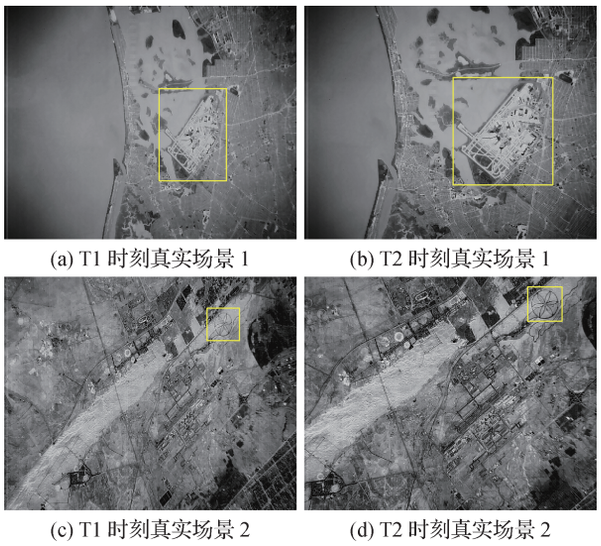
图 3　真实目标不同时刻的局部放大图

Fig. 3　Local magnification views of real targets at different times ((a) target 1 at time T1; (b) target 1 at time T2; (c) target 2 at time T1; (d) target 2 at time T2; (e) target 3 at time T1; (f) target 3 at time T2)

2.3.2　真实数据场景变化特性分析

真实情况下，除目标各种状态千变万化外，场景随时间的变化也十分复杂。对于真实卫星平台，由于轨道运动，卫星观测视角可能发生变化，这导致了场景的规律变化。在采集的视频数据中，场景变化包括平移、旋转、缩放和拉伸等各种变化，另外场景中还存在许多灰度变化(如闪烁)的背景局部。从图 4 可以看出，数据集中真实场景存在更加复杂的变化。如图 4(a)(b)，场景中的机场不仅在图像中的位置发生了变化，形状也发生了变化，从 T1 时刻到 T2 时刻，尺寸整体变大且水平方向拉伸程度更大；如图 4(c)(d)，场景右上角有一圆形图案，从 T1 时刻到 T2 时刻，该图案从椭圆形变得更加接近圆形，同时 T2 时刻该图案变得更加清晰，图像对比度有所提升。

综上，SatVideoIRSDT 数据集中包含的大量真实场景具有更加复杂的目标和背景变化以及目标状态，对红外视频卫星空中动目标检测跟踪研究具有重要意义。



(a) T1 时刻真实场景 1　(b) T2 时刻真实场景 1
(c) T1 时刻真实场景 2　(d) T2 时刻真实场景 2

图 4　不同时刻的真实场景图

Fig. 4　Real scene images at different time ((a) real scene 1 at time T1; (b) real scene 1 at time T2; (c) real scene 2 at time T1; (d) real scene 2 at time T2)

2.3.3　真实数据统计特性分析

真实数据来自武汉一号卫星热红外相机，相机波段为 3.7—4.8 μm，分辨率为 512 × 640 像素，视频数据采集的帧频为 25 Hz 和 50 Hz。对数据中目标信噪比进行统计分析，采用的信噪比计算为

$$SNR = \frac{|c_t - \mu_b|}{\sigma_b} \tag{1}$$

式中，c_t 表示目标标注中心位置的灰度值，μ_b 表示目标邻域背景的灰度均值，σ_b 表示目标邻域背景的灰度标准差。

考虑到数据中既有亮目标也有暗目标，因此采用取绝对值的方式计算信噪比。同时，由于大部分目标很微弱，与背景的边界十分模糊，如图 3(a)(e)中的目标，导致掩码标注存在误差，且存在局部亮、暗并存的目标，如图 3(b)中的目标，不适宜采用目标区域灰度均值或目标区域灰度极值代表目标灰度，因此采用目标标注中心位置的灰度值计算信噪比。对真实数据的目标信噪比统计结果如表 2 所示。从表中可以看出，有近 30% 的目标信噪比低于 1，超过 80% 的目标信噪比低于 2 dB，少量目标信噪比超过 2。可见该实测数据中的空中动目标极弱，检测跟踪难度极高，为探索和突破红外极弱小目标检测技术瓶颈提供了极佳的数据支撑。

此外，本小节还全面地分析了真实数据的统计特性，并对比了已有的序列红外小目标检测数据的

表 2　实测数据目标信噪比情况

Table 2　Signal-to-noise ratios of targets in our real data

信噪比区间	目标占比/%
[0, 1)	29.60
[1, 2)	51.70
[2, +∞)	18.70

统计特性，对比结果如表 3 所示。从表中可以看出，SatVideoIRSDT 数据集是唯一包含大量实测场景的天基序列红外小目标检测数据集，且场景与目标类型覆盖范围广，目标尺寸小且实测数据目标平均信噪比最低。SatVideoIRSDT 数据集是一个大规模、多样化和高挑战的数据集。

表 3　不同数据集统计特性对比

Table 3　Comparison of statistical properties across different datasets

数据集	类型	分辨率/像素	标签类型	场景	场景数	帧数	目标类型	目标数量	目标平均大小/像素	平均信噪比
SIATD	仿真	640 × 512	点	空基	350	150 185	亮目标	150 185	<7 × 7	6.95
IRDSTReal	实测	992 × 742*	掩膜	陆基	85	40 653	亮目标	41 819	10	6.76
IRDSTSimulate	仿真	720 × 480	掩膜	陆基	316	102 077	亮目标	102 077	6	7.54
NUDTMIRSDT	仿真	217 × 302*	掩膜	陆基	120	12 000	亮目标	11 446	34	5.04
IRAir	仿真	256 × 256	掩膜	天基	2 000	100 000	暗目标	509 448	6	4.28
IRSatVideo-LEO	仿真	1 024 × 1 024	掩膜	天基	200	91 021	亮目标	218 038	11	9.80
SatVideoIRSDT	实测 + 仿真	640 × 512*	掩膜	天基	1 401	122 265	亮目标、暗目标	454 116	7	5.03
SatVideoIRSDT-real	实测	640 × 512	掩膜	天基	> 5	20 126	亮目标、暗目标	29 757	<20	3.06

注：“*”代表多种分辨率的平均分辨率。

3　数据集评估

3.1　评估指标

由于该数据集具有实例级分割标注，因此基于该数据集进行了空中动目标检测与跟踪算法的评估。

3.1.1　检测指标

为适应不同框架网络以进行统一的公平对比，检测指标采用目标级评价指标，召回率、精确率和 F_1 分数。将模型输出结果统一转化为预测目标的质心位置，通过对比标签目标质心位置得到评估结果。

召回率，即检测概率，定义为正确检测的目标数量与目标总数量之比，评估算法发现目标的能力。其中，当检测到的目标质心与对应掩码标签质心距离不超过偏差阈值 τ 时，该检测目标被认为是正确检测目标。

精确率，定义为正确检测的目标数量与算法检测输出的目标总数量（正确检测目标数量与虚警目标数量之和）之比，评估算法抑制虚警的能力。其中，当检测到的目标质心与任意掩码标签质心距离均超过 τ 时，该检测目标被认为是虚警目标。

召回率 Re 和精确率 Pr 分别计算为

$$Re = \frac{TD}{AT} \tag{2}$$

$$Pr = \frac{TD}{TD + FD} \tag{3}$$

式中，TD（true detection）和 FD（false detection）分别为正确检测的目标数量和虚警目标数量，AT（amount of target）为目标总数量。

算法检测能力的总评估结果采用 F_1 分数，计算为

$$F_1 = \frac{2 \times Re \times Pr}{Re + Pr} \tag{4}$$

3.1.2　跟踪指标

通用多目标跟踪指标（Lu 等，2025）有多目标跟踪准确度（multiple object tracking accuracy，MOTA）、高阶跟踪精度（higher order tracking accuracy，HOTA）和多目标跟踪精度（multiple object tracking precision，MOTP）等指标，综合评估目标跟踪后的检测精度、关联精度和定位精度等。针对红外视频卫

星观测下的空中动目标，由于实测数据人工标注存在误差，目标与背景边界像素存在多种标注可能，这对小目标的质心可能带来较大扰动，这种扰动是难以避免的，因此本文不严格评估跟踪定位精度。此外，实际应用中，需要更直观反映跟踪轨迹完整程度以及轨迹点匹配准确程度的指标，因此本文提出轨迹完整度和轨迹准确度指标，对跟踪结果进行评估。

轨迹完整度，定义为经过跟踪/轨迹关联后，目标轨迹被完整关联的程度。计算为

$$T_{\text{comp}} = \frac{N_{\text{TA}}}{N} \tag{5}$$

式中，N_{TA}（number of points true associated）为目标轨迹被正确检出并关联的点数，N（number of track points）为标签中目标轨迹总点数。

轨迹准确度，定义为经过跟踪/轨迹关联后，关联轨迹中各点为真实目标点的概率。计算为

$$T_{\text{accu}} = \frac{N_{\text{TA}}}{N_A} \tag{6}$$

式中，N_{TA} 为关联轨迹中真实目标点数，N_A（number of points associated）为算法关联轨迹总点数。

3.2 方法与数据集评估

依托SatVideoIRSDT数据集，举办了首届红外视频卫星空中动目标检测挑战赛（即SatVideoIRSTD挑战赛）。该比赛以动目标检测为主要任务，附加动目标跟踪任务，在促进红外视频卫星空中动目标检测技术突破瓶颈的同时推动针对该类场景与目标的跟踪技术的探索。比赛吸引了国内各大高校、研究所与实验室以及科技公司等科研单位共73支队伍（338人）报名参赛，产生的19支获奖队伍的比赛结果为该数据集的评估提供了充分依据。

3.2.1 基准结果

将DeepPro-Plus（the plus version of deep temporal probe network）（Li等，2026）模型作为该数据集上检测任务的基准算法，输入原始红外图像序列，输出每帧图像的目标分割结果，并通过阈值分割和聚类求质心得到备选点位置信息，也即目标检测结果。同时，对比了其他序列红外小目标检测方法，如ResUnet + DTUM（residual U-Net and direction-coded temporal U-shape module）（Li等，2025）、DNANet + DTUM（dense nested attention network and direction-coded temporal U-shape module）（Li等，2025）以及ResUnet + RFR（residual U-Net and recurrent feature refinement framework）（Ying等，2025a）方法，在SatVideoIRSDT测试集上的评估结果如表4所示。从结果来看，各种算法在实测场景下表现不佳，均存在漏检多的问题，部分算法（如ResUnet + RFR）在实测场景下有极高的虚警。该数据集检测难度很大。

表4 SatVideoIRSDT数据集上检测任务基准结果

Table 4 Benchmark results of detection task on the SatVideoIRSDT dataset

检测算法	召回率	精确率	F_1分数
ResUnet + DTUM	0.364 1	0.589 4	0.454 0
DNANet + DTUM	0.464 2	**0.686 6**	**0.553 9**
ResUnet + RFR	0.515 0	0.431 3	0.469 4
DeepPro-Plus（基准算法）	**0.578 2**	0.495 6	0.533 7

注：加粗字体表示各列最优结果。

基于DeepPro-Plus检测模型，采用经典SORT（simple online and realtime tracking）（Bewley等，2016）算法作为该数据集上跟踪任务的基准算法，输入生成的序列备选点信息进行跟踪关联，输出序列中所有目标轨迹信息，也即目标跟踪结果，在SatVideoIRSDT测试集上的评估结果如表5所示（包含基准结果和比赛一等奖结果）。

数据集的特性分析结果和基准算法结果共同反映了该数据集的三大挑战：

表5 SatVideoIRSDT数据集上跟踪任务基准结果

Table 5 Benchmark results of tracking task on the SatVideoIRSDT dataset

算法	召回率	精确率	F_1分数	轨迹完整度	轨迹准确度
DeepPro-Plus + SORT（基准算法）	0.561 1	0.519 1	0.539 9	0.241 9	0.519 1
冠军方案	**0.817 0**	**0.990 1**	**0.895 3**	0.815 8	**0.990 1**
亚军方案	0.830 3	0.959 0	0.890 0	**0.829 5**	0.959 0

注：加粗字体表示各列最优结果。

1)发现目标难。由于目标极弱且多种状态动态变化,导致目标容易被漏检。

2)虚警抑制难。由于实测背景序列存在丰富的动态变化,导致许多杂波与目标高度相似,难以抑制。

3)连续跟踪难。由于目标极弱带来了检测中断问题,导致跟踪过程轨迹断续,难以关联。

3.2.2　挑战赛一等奖结果——一眼识破队

该队伍在SatVideoIRSTD挑战赛中获得了总分第1的成绩。针对检测任务,考虑了背景不稳定因素(如,相机抖动),在推理流程前设计部署了前馈相机运动补偿模块。该模块通过频域对齐来还原准静态的观测视角,补偿后,背景在时间维度趋于稳定,有助于增强运动目标在特征空间中的时空显著特性,并增大其与背景杂波在特征空间的距离,以降低漏检和误报。具体地,采用鲁棒且高效的相位相关方法(phase correlation method, PCM),以序列首帧图像为基准,计算后续每帧图像的全局平移向量$\Delta x, \Delta y$。此向量由参考帧与当前帧的归一化互功率谱峰值唯一确定。通过对每帧施加逆向平移补偿,为整个序列建立了统一且稳定的时空坐标系,同时也为后续跟踪提供可靠的运动与几何约束。该模块的处理流程如图5所示。

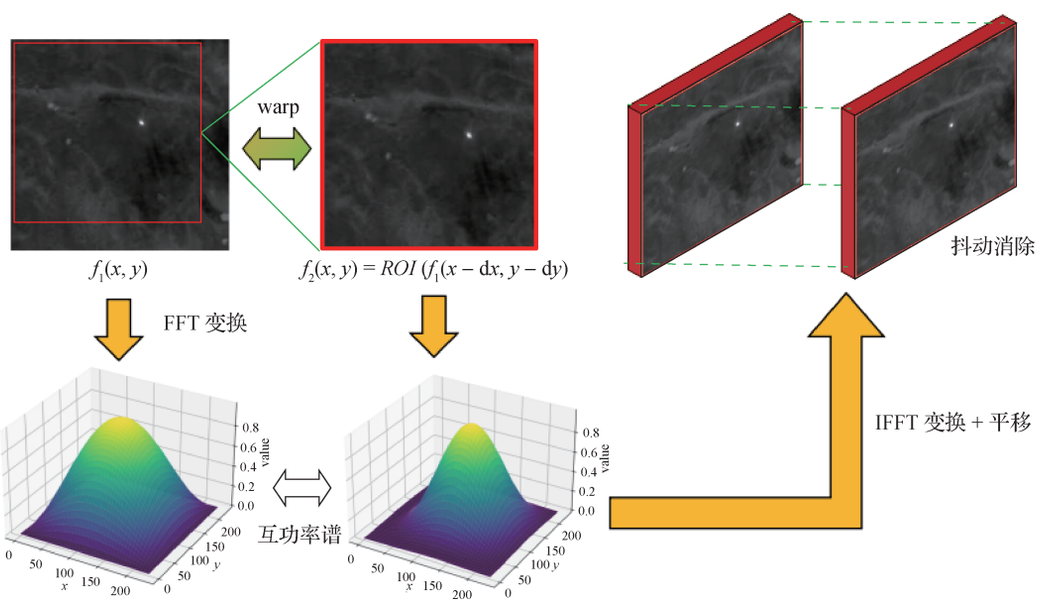


图5　运动补偿模块处理流程图

Fig. 5　Motion compensation module processing flowchart

针对DeepPro基准算法在时域信息利用过程中存在的时序边缘上下文缺失问题,提出了一种双向时序推理与融合的优化策略,如图6所示。通过双向处理,为每一帧构建充分的前向或后向参考,然后通过选择性融合,保留最佳推理结果,最终实现整体检测性能的最优化。

针对跟踪任务中轨迹碎片化与不连续问题,设计了一套多阶段、多策略的轨迹优化后处理框架,将离散的检测点转化为时空连续、语义完整的高完整度运动轨迹。首先进行在线跟踪,具体地,采用经典卡尔曼滤波对目标状态进行预测与更新,并结合匈牙利算法进行最优数据关联,从而生成初步的轨迹片段。随后,进入离线优化阶段,引入一种基于运动学与几何相似度的轨迹合并策略,将属于同一目标的多个轨迹片段进行有效拼接。最终,通过时序插值与噪声滤除,修复轨迹的微小断裂并剔除孤立的虚假轨迹。

从表5展示的评估结果来看,该方案显著降低了漏检和误报,实现了最优的检测与跟踪整体性能。此外,全面分析了该方法在真实数据上的测试性能,发现在一半以上的真实数据中目标召回率为0,即目标全部漏检。由此可见,该数据集中的真实数据具有极高挑战性和很大探索空间。

3.2.3　挑战赛一等奖结果——大红西瓜队

该队伍在SatVideoIRSTD挑战赛中获得了总分第2的成绩。SatVideoIRSDT数据集中大量实测动态背景与仿真场景存在较大差异,对深度模型泛化能力提出更高的挑战。为此,提出了一种深度学习

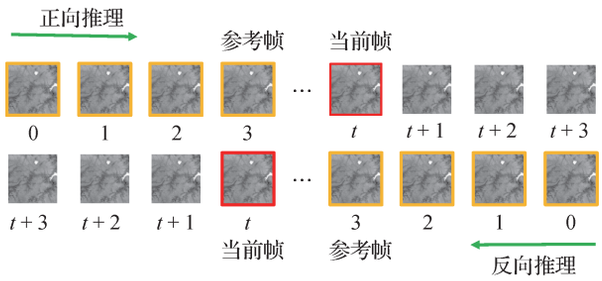


图6　双向推理示意图

Fig. 6　Schematic of bidirectional reasoning

与传统图像处理互为补充的混合策略Hybrid-MIST(hybrid moving infrared small target method):在与训练集分布较一致的子集上充分发挥深度时序模型的优势;而在分布偏离较大的子集上,则采用基于图像配准、背景建模与多帧差分的传统处理流程,并结合特征筛选机制抑制复杂背景干扰。该差异化策略使得整体方法在不同场景下均具备较强的鲁棒性,体现了深度学习与手工方法的优势互补。

具体地,深度检测模型框架如图7所示,以DTUM(Li等,2025)模型为基础,该模型通过引入运动到数据的映射与方向编码卷积块(direction-coded convolution block, DCCB),提取并编码运动方向特征,使弱小目标在时序一致性下被增强,能够有效区分目标与杂波的运动差异。将模型输入的帧数增长至9帧,在推理时用目标帧与其前8帧和后8帧分别构造两组样本,分别经过模型的配准、预测后,将两个预测置信度图做相加融合,输出深度检测结果。手工检测方法先用SIFT(scale-invariant feature transform)进行配准,再做背景建模与多帧差分,将差分图融合后使用差分高斯(difference of Gaussians, DoG)增强小尺度目标。最后,利用自适应阈值和形态学去噪,按面积筛选连通域并输出质心作为候选点得到初步结果。完成初步检测后,使用跟踪算法去除短轨迹,保留直线运动目标。

针对跟踪任务,采用正反向融合跟踪策略。以Bytetrack(Zhang等,2022c)为核心跟踪算法,在此基础上,引入中心距离匹配策略,以更好地适应低分辨率或小目标场景。为提升轨迹的完整性,在正反向两个方向分别执行跟踪,并将两者结果融合。此外,设计了一套后处理方法,用于进一步处理生成的轨迹,以提升轨迹连续性和准确性,并剔除噪声轨迹。该方法包含7个可选步骤:1)轨迹插值;2)按起终点距离剔除轨迹;3)按帧数剔除;4)补全与合并轨迹;5)直线性筛选;6)同帧近距离目标合并;7)静止目标延长。

从表5展示的评估结果来看,该方案相比基准算法显著降低了漏检和误报,实现了优越的检测与跟踪整体性能。

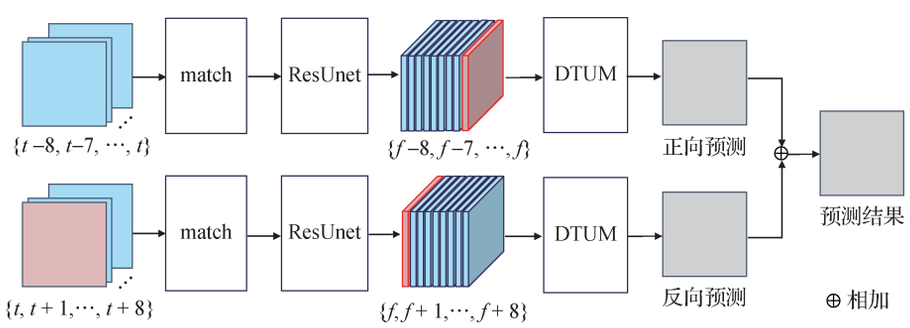


图7　基于DTUM的深度检测模型框架

Fig. 7　Framework of DTUM-based deep detection model

4　现存挑战与未来展望

基于该数据集的探索取得了初步的进展,但仍存在诸多困难挑战,需要在未来研究中重点关注。

1)极弱目标检测能力亟需提升。从表5中的结果来看,优胜方案检测能力显著强于基准算法,然而两种优胜方案均采用了轨迹插值来填补轨迹断裂,且在一半以上的实测序列中目标召回率为0,这说明并未从根本上解决极弱目标检测难的问题。

2)非稳定平台时域信息利用难。现实世界中大多数平台存在位移、旋转和抖动等情况,由此导致时域信息难以有效利用,尤其是长时域信息。对非稳定平台的目标检测依赖帧间配准,大多数方法在预

处理阶段采用传统算法进行配准，这可能导致误差传递和信息损耗，从而影响时域信息利用。

3）跟踪算法智能化水平需提升。尽管深度学习方法在各个领域取得了显著成绩，然而目前序列红外弱小目标跟踪任务仍十分依赖传统经典算法，对智能方法的探索不足。通用目标的跟踪依赖目标特征的帧间匹配，而红外弱小目标空间特征十分匮乏，难以直接适用通用智能跟踪方法，因此亟需探索针对红外弱小目标的多目标智能跟踪方法。

4）检测跟踪一体化水平需提升。目前针对红外弱小目标的跟踪任务往往分为两个阶段完成，即备选点检测和轨迹关联。然而，两阶段的方式可能带来信息的丢失，如低置信度的备选点被丢弃，以及备选点的部分特征在关联前被丢弃，这会制约目标检测跟踪性能的提升。因此，亟需探索检测跟踪一体化智能方法，减少信息丢失，充分利用弱小目标特征，实现更加鲁棒优越的红外弱小目标检测跟踪。

5 结 论

针对红外视频卫星空中动目标检测这一重要研究问题，以及其存在的发展瓶颈问题，包括单帧检测失效和视频数据集匮乏等，本文构建了首个包含大量真实场景的红外视频卫星空中动目标检测数据集SatVideoIRSDT。该数据集中的真实场景来自武汉一号卫星热红外相机，通过采集卫星相机拍摄的机场附近的场景获取所需视频数据，视频中的空中动目标主要为从机场起落的飞机。共采集实测数据20 126帧，标注了29 757个空中动目标，标签区分不同目标实例，该数据集支持开展红外极弱小目标检测与跟踪研究。为了丰富各种场景数据，更好地支持大规模网络训练，还融合了两个基于真实天基背景的仿真序列红外小目标检测数据集的一部分。考虑到从实测数据中观察到同一目标在不同时刻可能呈现暗目标和亮目标两种状态，所融合的仿真数据集一个为亮目标数据集，另一个为暗目标数据集。因此，所提出的数据集也是首个同时包含亮目标和暗目标的天基数据集。

本文对数据集进行了全面的特性分析，并对比了已有的视频数据集，证明了SatVideoIRSDT数据集具有更高的挑战性：目标尺寸小，平均不足20像元，目标强度极弱，真实目标平均信噪比仅为3.06。基于该数据集组织了首届红外视频卫星空中动目标检测挑战赛，制定了检测和跟踪任务的评价指标。比赛结果也充分证明了SatVideoIRSDT数据集的挑战性与重要价值。

数据可用性声明

支撑本研究的科学数据已在中国科学院科学数据银行（Science Data Bank）ScienceDB平台公开发布，访问地址为 https://www.doi.org/10.57760/sciencedb.j00240.00077 或 https://cstr.cn/31253.11.sciencedb.j00240.00077。

参考文献（References）

Bai X Z and Zhou F G. 2010. Analysis of new top-hat transformation and the application for infrared dim small target detection. Pattern Recognition, 43(6): 2145-2156 [DOI: 10.1016/j.patcog.2009.12.023]

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. Phoenix, USA: IEEE: 3464-3468 [DOI: 10.1109/ICIP.2016.7533003]

Dai Y M, Wu Y Q, Zhou F and Barnard K. 2021a. Asymmetric contextual modulation for infrared small target detection//Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa, USA: IEEE: 949-958 [DOI: 10.1109/WACV48630.2021.00099]

Dai Y M, Wu Y Q, Zhou F and Barnard K. 2021b. Attentional local contrast networks for infrared small target detection. IEEE Transactions on Geoscience and Remote Sensing, 59(11): 9813-9824 [DOI: 10.1109/TGRS.2020.3044958]

Du J M, Lu H Z, Zhang L P, Hu M F, Chen S, Deng Y J, et al. 2022. A spatial-temporal feature-based detection framework for infrared dim small target. IEEE Transactions on Geoscience and Remote Sensing, 60: #3000412 [DOI: 10.1109/TGRS.2021.3117131]

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, 7(2): 1-16 (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, 7(2): 1-16) [DOI: 10.11922/11-6035.csd.2021.0085.zh]

Gao J Y, Lin Z P, Guo Y L and An W. 2017. TVPCF: a spatial and temporal filter for small target detection in IR images//Proceedings of 2017 International Conference on Digital Image Computing: Techniques and Applications. Sydney, Australia: IEEE: 1-7 [DOI: 10.1109/DICTA.2017.8227499]

Huang B, Li J N, Chen J J, Wang G, Zhao J and Xu T F. 2023. 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]

Kong X, Yang C P, Cao S Y, Li C H and Peng Z M. 2021. Infrared small target detection via nonconvex tensor fibered rank approximation. IEEE Transactions on Geoscience and Remote Sensing, 60: #5000321 [DOI: 10.1109/TGRS.2021.3068465]

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]

Li B Y, Xiao C, Wang L G, Wang Y Q, Lin Z P, Li M, et al. 2023a. Dense nested attention network for infrared small target detection. IEEE Transactions on Image Processing, 32: 1745-1758 [DOI: 10.1109/TIP.2022.3199107]

Li J, Zhang P, Zhang L Y and Zhang Z Y. 2023b. Sparse regularization-based spatial-temporal twist tensor model for infrared small target detection. IEEE Transactions on Geoscience and Remote Sensing, 61: #5000417 [DOI: 10.1109/TGRS.2023.3234608]

Li R J, An W, Xiao C, Li B Y, Wang Y Q, Li M, et al. 2025. Direction-coded temporal U-shape module for multiframe infrared small target detection. IEEE Transactions on Neural Networks and Learning Systems, 36(1): 555-568 [DOI: 10.1109/TNNLS.2023.3331004]

Li R J, An W, Wang Y Q, Ying X Y, Dai Y M, Wang L G, et al. 2026. Probing deep into temporal profile makes the infrared small target detector much better. IEEE Transactions on Pattern Analysis and Machine Intelligence, 48(8): 10157-10175 [DOI: 10.1109/TPAMI.2026.3683258]

Li Y S, Li Z Z, Zhang C, Luo Z F, Zhu Y, Ding Z Q, et al. 2021. Infrared maritime dim small target detection based on spatiotemporal cues and directional morphological filtering. Infrared Physics and Technology, 115: #103657 [DOI: 10.1016/j. infrared. 2021.103657]

Li Z X, Xu Q Y, An W, He X, Guo G W, Li M, et al. 2025. A lightweight dark object detection network for infrared images. Journal of Infrared and Millimeter Waves, 44(2): 285-296 (李朝旭, 徐清宇, 安玮, 贺旭, 郭高伟, 李淼, 等. 2025. 红外图像暗弱目标轻量级检测网络. 红外与毫米波学报, 44(2): 285-296) [DOI: 10.11972/j.issn.1001-9014.2025.02.016]

Liu D P, Cao L, Li Z Z, Liu T M and Che P. 2018. Infrared small target detection based on flux density and direction diversity in gradient vector field. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(7): 2528-2554 [DOI: 10.1109/JSTARS.2018.2828317]

Liu F C, Gao C Q, Chen F, Meng D Y, Zuo W M and Gao X B. 2023. Infrared small and dim target detection with transformer under complex backgrounds. IEEE Transactions on Image Processing, 32: 5921-5932 [DOI: 10.1109/TIP.2023.3326396]

Liu P, Peng J J, Wang H L, Hong D F and Cao X Y. 2024. Infrared small target detection via joint low rankness and local smoothness prior. IEEE Transactions on Geoscience and Remote Sensing, 62: #5708315 [DOI: 10.1109/TGRS.2024.3492277]

Lu W S, Wang X Y, An W, Xiao C, Yin Q and Zhang G L. 2025. OITrack: multi-object tracking for small targets in satellite video via online trajectory completion and iterative expansion over union. Remote Sensing, 17(12): #2042 [DOI: 10.3390/rs17122042].

Sun H, Bai J X, Yang F and Bai X Z. 2023. Receptive-field and direction induced attention network for infrared dim small target detection with a large-scale dataset IRDST. IEEE Transactions on Geoscience and Remote Sensing, 61: #5000513 [DOI: 10.1109/TGRS.2023.3235150]

Sun X L, Guo L C, Zhang W L, Wang Z and Yu Q F. 2021. Small aerial target detection for airborne infrared detection systems using LightGBM and trajectory constraints. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14: 9959-9973 [DOI: 10.1109/JSTARS.2021.3115637]

Wang H, Zhou L P 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. Seoul, Korea (South): IEEE: 8508-8517 [DOI: 10.1109/ICCV.2019.00860]

Wu F Y, Yu H, Liu A R, Luo J H and Peng Z M. 2023a. Infrared small target detection using spatiotemporal 4-D tensor train and ring unfolding. IEEE Transactions on Geoscience and Remote Sensing, 61: #5002922 [DOI: 10.1109/TGRS.2023.3288024]

Wu T H, Li B Y, Luo Y H, Wang Y Q, Xiao C, Liu T, et al. 2023b. MTU-net: multilevel TransUNet for space-based infrared tiny ship detection. IEEE Transactions on Geoscience and Remote Sensing, 61: #5601015 [DOI: 10.1109/TGRS.2023.3235002]

Yang P, Dong L L and Xu W H. 2023. Small maritime target detection using gradient vector field characterization of infrared image. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16: 1827-1841 [DOI: 10.1109/JSTARS.2022.3210112]

Ying X Y, Liu L, Lin Z P, Shi Y S, Wang Y Q, Li R J, et al. 2025a. Infrared small target detection in satellite videos: a new dataset and a novel recurrent feature refinement framework. IEEE Transactions on Geoscience and Remote Sensing, 63: #5002818 [DOI: 10.1109/TGRS.2025.3542368]

Ying X Y, Xiao C, An W, Li R J, He X, Li B Y, et al. 2025b. Visible-thermal tiny object detection: a benchmark dataset and baselines. IEEE Transactions on Pattern Analysis and Machine Intelligence, 47(7): 6088-6096 [DOI: 10.1109/TPAMI.2025.3544621]

Zhang G Y, Chen S Y, Xu G S, Shen Z H, Wang H and Zhang X H. 2025. Beyond motion cues and structural sparsity: revisiting small moving target detection [EB/OL]. [2025-09-09]. https://arxiv.org/pdf/2509.07654.pdf

Zhang M J, Zhang R, Yang Y X, Bai H C, Zhang J and Guo J. 2022a. ISNet: shape matters for infrared small target detection//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 867-876 [DOI: 10.1109/CVPR52688.2022.00095]

Zhang S F, Huang X H and Wang M. 2016. Algorithm of infrared background suppression and small target detection. Journal of Image and Graphics, 21(8): 1039-1047 (张世锋, 黄心汉, 王敏. 2016. 红外背景抑制与小目标检测算法. 中国图象图形学报, 21(8): 1039-1047) [DOI: 10.11834/jig.20160808]

Zhang Y, Zhang Y, Fu R G, Shi Z G, Zhang J H, Liu D, et al. 2022b. Learning nonlocal quadrature contrast for detection and recognition of infrared rotary-wing UAV targets in complex background. IEEE Transactions on Geoscience and Remote Sensing, 60: #5629919 [DOI: 10.1109/TGRS.2022.3203785]

Zhang Y F, Sun P Z, Jiang Y, Yu D D, Weng F C, Yuan Z H, et al. 2022c. ByteTrack: multi-object tracking by associating every detection box//Proceedings of the 17th European Conference on Computer Vision. Tel Aviv, Israel: Springer: 1-21 [DOI: 10.1007/978-3-031-20047-2_1]

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]

Zhao M J, Li W, Li L, Ma P G, Cai Z Q and Tao R. 2022. Three-order tensor creation and tucker decomposition for infrared small-target detection. IEEE Transactions on Geoscience and Remote Sensing, 60: #5000216 [DOI: 10.1109/TGRS.2021.3057696]

Zhou X Y, Li P, Zhang Y, Lu X and Hu Y. 2023. Deep low-rank and sparse patch-image network for infrared dim and small target detection. IEEE Transactions on Geoscience and Remote Sensing, 61: #5003014 [DOI: 10.1109/TGRS.2023.3288574]

Zhu H, Zhang J K, Xu G X and Deng L Z. 2020. Balanced ring top-hat transformation for infrared small-target detection with guided filter kernel. IEEE Transactions on Aerospace and Electronic Systems, 56(5): 3892-3903 [DOI: 10.1109/TAES.2020.2982347]

作者简介

李若敬，女，博士研究生，主要研究方向为光学弱小目标检测和深度学习可解释性。E-mail: liruojing@nudt.edu.cn

曾瑶源，通信作者，男，副研究员，主要研究方向为空间信息处理。E-mail: yyzeng@hotmail.com

李朝旭，男，博士，主要研究方向为光学遥感小目标检测。E-mail: lizhaoxu@nudt.edu.cn

陈诺，男，博士研究生，主要研究方向为图像处理、神经形态相机和计算机视觉。E-mail: chennuo97@nudt.edu.cn

郭高伟，男，博士研究生，主要研究方向为高光谱视频目标跟踪。E-mail: guogaowei22@nudt.edu.cn

窦泽超，男，硕士研究生，主要研究方向为遥感运动小目标检测。E-mail: douzechao@nudt.edu.cn

龙政兴，男，博士研究生，主要研究方向为多类别目标的实例分割。E-mail: mikaellzx@163.com

罗伊杭，女，博士研究生，主要研究方向为多模态图像融合、小目标检测。E-mail: luoyihang@nudt.edu.cn

盛卫东，男，副研究员，主要研究方向为弱小目标检测与跟踪。E-mail: shengweidong1111@sohu.com

李博扬，男，助理研究员，主要研究方向为红外小目标检测。E-mail: liboyang20@nudt.edu.cn

李志军，男，副研究员，主要研究方向为遥感卫星载荷设计与应用。E-mail: lzj19821202@163.com

李淼，男，副研究员，主要研究方向为红外小目标检测、事件相机。E-mail: lm8866@nudt.edu.cn

安玮，女，教授，主要研究方向为空间信号处理。E-mail: anwei@nudt.edu.cn

龙云利，男，高级工程师，主要研究方向为空间信号处理。E-mail: feiyunlyi@126.com

李海鑫，男，硕士研究生，主要研究方向为红外小目标检测。E-mail: haixinli@bupt.edu.cn

余志强，男，硕士研究生，主要研究方向为红外小目标检测。E-mail: zhiqiangyu98@163.com

尹小宇，女，硕士研究生，主要研究方向为红外小目标检测。E-mail: xiaoyuyin@stu.xju.edu.cn

查旭阳，男，硕士研究生，主要研究方向为视觉SLAM。E-mail: xuyangzha@bupt.edu.cn

曾柏文，男，硕士研究生，主要研究方向为视觉SLAM。E-mail: bowing@bupt.edu.cn

王吴凡，男，特聘副研究员，主要研究方向为计算机视觉。E-mail: wufanwang@bupt.edu.cn

张波，男，副教授，主要研究方向为计算机视觉和物联网(低空)。E-mail: zbo@bupt.edu.cn

路悦，男，博士研究生，主要研究方向为计算机视觉和持续学习。E-mail: zgxd@mail.nwpu.edu.cn

褚东皓，男，硕士研究生，主要研究方向为红外小目标检测。E-mail: chudonghao@mail.nwpu.edu.cn

李子怡，女，硕士研究生，主要研究方向为红外小目标检测。E-mail: 13613227369@163.com

黄康薇，女，硕士研究生，主要研究方向为红外目标检测。E-mail: kwhuang@mail.nwpu.edu.cn

杨博睿，男，硕士研究生，主要研究方向为行人重识别。E-mail: ybr@mail.nwpu.edu.cn

邢颖慧，女，副教授，主要研究方向为多源多模态图像对齐融合与协同解译。E-mail: xyh_7491@nwpu.edu.cn

张世周，男，副教授，主要研究方向为自主进化学习和无人机视觉。E-mail: szzahng@nwpu.edu.cn